

# ADMIN

## Network & Security

ISSUE 72

# OpenStack

- Sensible deployments
- Small business alternatives
- What's new in OpenStack?

## Windows Subsystems for Linux and Android

## Parity Declustering RAID

Decrease resilvering times

### Siege

Website stress and  
benchmarking tool

### NetTools

Simplified AD  
troubleshooting

### Microsoft Purview

Data security in the  
hybrid working world

### Age

Simple, secure  
file encryption

LINUX NEW MEDIA  
The Pulse of Open Source



Get started with



# SysAdmin JOB HUB

Top jobs for IT professionals  
who keep the world's  
systems running

**[SysAdminJobHub.com](https://SysAdminJobHub.com)**



# Container Angst

**Just because it's new, doesn't mean it's better. How I learned to stop worrying and love the old containers.**

The IT world went container crazy a few years ago when Docker first hit the scene. Although I'm into virtualization of all types, I totally missed out on Docker-mania. I never really "got" what all the hype was about when containers had been around for about 40 years. There was never any excitement around them until the whole Docker craze struck the hearts and minds of IT folk. I like containers. They're easy to set up and easy to manage, and there are lots of advantages of a lightweight but robust virtualization solution. I'd been working with containers for years before Docker came along and spoiled it for everyone. Well, I spoiled it for everyone who knew that containers weren't new or spectacular in any way. It didn't matter how long containers had been around, these newfangled, fancier containers were somehow newer and more exciting and something we'd never seen or heard of.

To the whole ridiculous whimsical container sickness, I say, "Bah. Humbug!" I love containers. Containers aren't the problem. It's the madness surrounding them that irritates me. Other container technologies are just as exciting as Docker containers. Podman, for example, is a great container technology from Red Hat. Both technologies have their advantages and disadvantages.

I'm a fan of the old-school chroot jails, Solaris zones, and even the newer system container implementation, system-nspawn. I like the simplicity of these "standard" container technologies. Attempting to create a cross-platform, write-once-play-anywhere technology sounds great, but the problem for me is that I'm required to install so much supporting software that it takes away the *lightweight* nature of the original idea. I don't necessarily think that every technology must be cross-platform capable. Some things should be operating system-specific or at least operating system-optimized. For example, you can run an Apache web server, PHP, and MySQL on Windows (WAMP), but when Linux is free, freely available, and LAMP (Linux, etc.) stacks come prepackaged – even as container and virtual machine (VM) images, I'm not sure why one would go to the trouble to create a WAMP stack (although I've done it).

I think a mixture of containers and full-featured VMs is the best solution. The Proxmox project offers this on a single host system, which is a blending of the OpenVZ container technology and KVM/Qemu, as I recall. I'm a huge fan of the OpenVZ container implementation. I can have a fully ready-to-run container host system operable in about an hour that has the capability of hosting hundreds of container virtual machines.

Don't get me wrong. I love new technology. More than once, I've shunned currently available technology for something much improved. And it only makes sense that not every new solution will appeal to everyone. I like a new solution that is more efficient, cost-effective, or labor-saving, or so clever that I can't say anything bad about it. If a newfangled whatever isn't in some way an improvement, then I'll take a hard pass on it. I don't need to change what I've done for the past five years just because there's a new thing available. Show me how it's better, and I'll buy in; otherwise, keep moving. This is how I feel about certain container solutions mentioned above. They don't seem to improve on previous technology enough to go to the trouble of learning or using them. If you use Docker or Podman, live long and prosper. You might have the right temperament for them, but I don't. My motto is that "Just because you *can* do something, doesn't mean you should." I can create concrete furniture, but I don't see the point of doing it, so I refrain.

I'm rarely an early adopter of the latest and greatest technology, not because I'm afraid of the technology but because of its vulnerabilities and security holes. Avoiding unnecessary risk is my jam, and I'm sticking to it for better or worse. Now, where are my hammer, chisel, and stone tablets? Too bad there's not a "Find my ancient tools" app available somewhere. Now that's something I could use. Perhaps a "pinch zoom" on common household items with tiny writing is also too tall an order. People would rather invent things we don't need than those we do.

Ken Hess • ADMIN Senior Editor



# ADMIN

**Network & Security****Features**

- 10 OpenStack Interview**  
Developers Thierry Carrez and Jeremy Stanley talk about problems, innovations, and future plans.
- 14 OpenStack News**  
The unprecedented hype surrounding OpenStack 10 years ago changed to disillusionment, which has nevertheless had a positive effect: OpenStack is still evolving and is now mainly deployed where it actually makes sense to do so.
- 20 Sovereign Cloud Stack**  
Operators of OpenStack need to know whether their environment is working and quickly pinpoint problems. We look at the basics of observability in the Sovereign Cloud Stack.
- 24 Container Issues**  
As the major Linux distributors increasingly lean toward containers, many administrators have come to realize that containers are by no means a panacea for all their problems.

**Tools**

- 30 Git Versioned Backups**  
The open source Git tool from the developer world delivers backups and version control.
- 36 VPN with SoftEther**  
SoftEther is lean VPN software that outpaces the current king of the hill, OpenVPN, in terms of technology and performance.
- 40 Windows Subsystems**  
WSL runs graphical Linux applications on Windows 11, and WSA shows that Windows can be a platform for Android apps.

**Containers and Virtualization**

- 46 Siege Benchmarking Tool**  
A stress and benchmarking tool for websites controlled from the command line.
- 50 OpenStack Alternatives**  
OpenStack is considered the industry standard for building private clouds, but the solution is still far too complex and too difficult to maintain and operate for many applications. What causes OpenStack projects to fail, and what alternatives do administrators have?

**Security**

- 56 Age/Rage File Encryption**  
Age and Rage are the Go and Rust implementations of a simple, modern, and secure file encryption tool.
- 58 WSL Vulnerabilities**  
Several tactics, techniques, and procedures circulating among cybercriminals exploit Windows Subsystem for Linux as a gateway. We look at how WSL can be misused and some appropriate protections.
- 62 Zeek**  
An arsenal of scripts for monitoring popular network protocols, along with its own policy scripting language for customization.

**Management**

- 64 NetTools for AD**  
This set of utilities extracts information from Active Directory to help simplify troubleshooting and administration.
- 70 Samba AD Domain Controller**  
The open source Samba service can act as an Active Directory domain controller in a heterogeneous environment.



# 10, 14, 20 | OpenStack

## OpenStack 2022

Find out whether the much evolved OpenStack is right for your private cloud.

### Highlights

#### 50 OpenStack Alternatives

If OpenStack is overkill for your company, you should consider a solution that brings together elements of open source software.

#### 64 NetTools

This bouquet of utilities squeezes the last snippet of information out of Active Directory - and it's fun to use!

#### 76 Midmarket IAM

Identity access management is expensive and more important than ever in small and mid-sized enterprises, so make sure you know how to evaluate products and implement your solution.

### Nuts and Bolts

#### 76 Midmarket IAM

Identity and access management in midmarket organizations.

#### 80 CMMC

The US Cybersecurity Maturity Model Certification will be required by mid-2023 to handle controlled unclassified information and win federal contracts, but it can also help minimize business risk and keep information out of the hands of adversaries.

#### 84 Declustered Parity

dRAID decreases resilvering times, restoring a pool to full redundancy in a fraction of the time over the traditional RAIDz.

### Nuts and Bolts

#### 88 Analyzing Logs in HPC

Log analysis can be used to great effect in HPC systems. We present an overview of the current log analysis technologies.

#### 93 Performance Dojo

Sensor tools provide highly variable data from a variety of sources. We look at some tools to verify the temperature of components on diverse hardware.

### On the DVD

Fedora is the upstream source for Red Hat Enterprise Linux and is based on the latest server technology to provide a "stable, flexible, and universally adaptable basis for the everyday provision of digital services and information." Fedora 36 Server changes include an updated Ansible 5 and Podman 4.0, and the RPM database has been moved from `/var` to `/usr`. For more information on Fedora 36 Server, go to <https://docs.fedoraproject.org/fedora-server/>.



### Service

- 3 Welcome
- 6 News
- 97 Back Issues
- 98 Call for Papers



News for Admins

# Tech News

## OpenSSL 3.0.7 Patches Serious Vulnerabilities

OpenSSL has issued an advisory (<https://www.openssl.org/news/secadv/20221101.txt>) relating to two vulnerabilities (CVE-2022-3602 and CVE-2022-3786), which affect OpenSSL version 3.0.0. These vulnerabilities have been addressed with the release of OpenSSL 3.0.7, so users should update now.

“Users of OpenSSL 3.0.0–3.0.6 are encouraged to upgrade to 3.0.7 as soon as possible. If you obtain your copy of OpenSSL from your operating system vendor or other third party then you should seek to obtain an updated version from them as soon as possible,” the OpenSSL team says (<https://www.openssl.org/blog/blog/2022/11/01/email-address-overflows/>).

In a previous announcement (<https://mta.openssl.org/pipermail/openssl-announce/2022-October/000238.html>), these vulnerabilities were described as “critical” — possibly leading to remote code execution. However, the OpenSSL project team has since downgraded the threats to “high,” saying they “are not aware of any working exploit that could lead to remote code execution” and have no evidence of the vulnerabilities being exploited at this time.

## IBM Introduces Diamondback Tape Library

IBM recently introduced the Diamondback Tape Library, “a high-density archival storage solution that is physically air-gapped to help protect against ransomware and other cyber threats in hybrid cloud environments.”

The Diamondback Tape Library (<https://www.ibm.com/products/diamondback-tape-library>) is aimed at organizations that need to secure hundreds of petabytes of data, such as hyperscale cloud providers and global enterprises aggregating massive data sets, according to the announcement (<https://www.hpcwire.com/off-the-wire/ibm-releases-its-diamondback-tape-library/>). “It provides long-term storage and is designed to provide a significantly smaller carbon footprint compared to flash or disk storage, and with a lower total cost of ownership.”

The main benefits of IBM Diamondback, the announcement says, include:

- Sustainability
- Ransomware protection and cyber resiliency
- Data capacity and storage costs

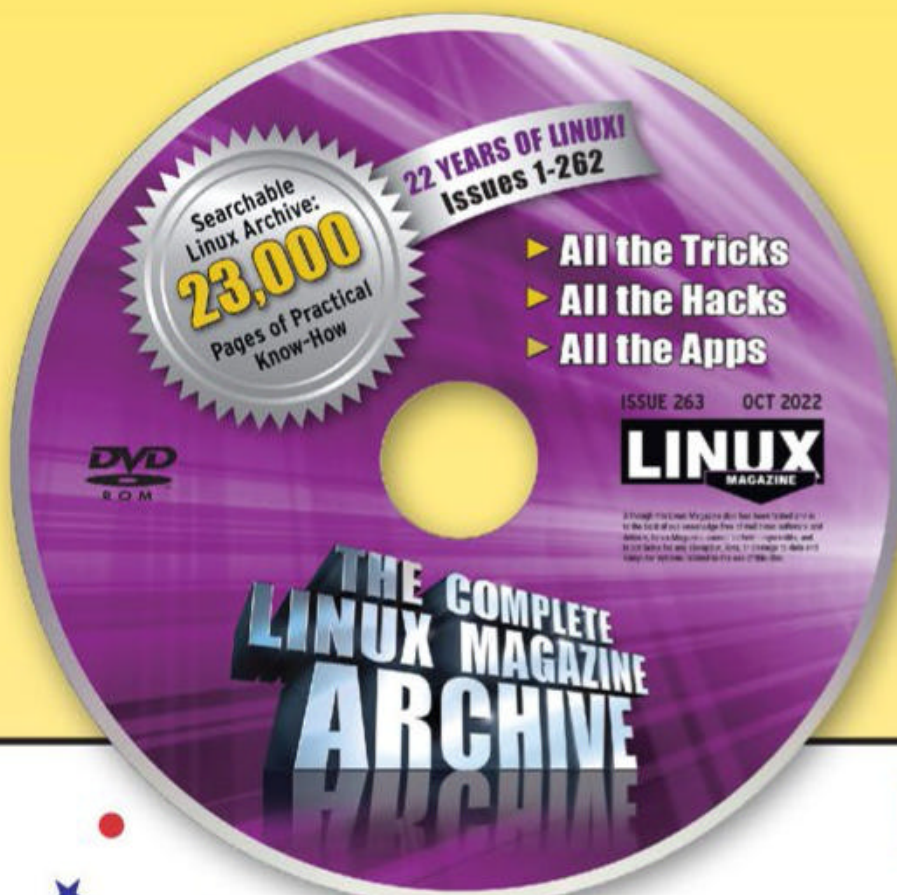
“The IBM Diamondback Tape Library provides critical protection against a variety of threats, helping minimize data center floor space requirements and organizations’ carbon footprint[s],” says Scott Baker, Vice President and Chief Marketing Officer of IBM Storage.



**Get the latest  
IT and HPC news  
in your inbox**

**Subscribe free to  
ADMIN Update  
and HPC Update**  
[bit.ly/HPC-ADMIN-Update](https://bit.ly/HPC-ADMIN-Update)





## LINUX MAGAZINE ★ ARCHIVE DVD ★

**ORDER NOW!**

<https://bit.ly/Archive-DVD>





## PostgreSQL 15 Released

The PostgreSQL Global Development Group has released version 15 of the open source PostgreSQL (<https://www.postgresql.org/>) database.

PostgreSQL 15 now also includes the SQL standard MERGE (<https://www.postgresql.org/docs/15/sql-merge.html>) command, which “lets you write conditional SQL statements that can include INSERT, UPDATE, and DELETE actions within a single statement.”

According to the announcement, PostgreSQL 15 also includes performance improvements “with noticeable gains for managing workloads in both local and distributed deployments, including improved sorting.” Specifically, PostgreSQL 15 offers improved in-memory and on-disk sorting (<https://www.postgresql.org/docs/15/queries-order.html>) algorithms, with benchmarks showing increases of 25–400 percent depending on data type.

“The PostgreSQL developer community continues to build features that simplify running high performance data workloads while improving the developer experience,” said Jonathan Katz, a PostgreSQL Core Team member.

## Hackers Weaponize Open Source Software in Targeted Phishing Attempts

The Microsoft Threat Intelligence Center (MSTIC) has recently detected a wide range of phishing attempts using weaponized open source software.

These attempts, attributed to ZINC (<https://www.microsoft.com/security/blog/2021/01/28/zinc-attacks-against-security-researchers/>), have used traditional social engineering tactics by contacting individuals with fake job offers on LinkedIn. “Upon successful connection, ZINC encouraged continued communication over WhatsApp, which acted as the means of delivery for their malicious payloads,” Microsoft says (<https://www.microsoft.com/security/blog/2022/09/29/zinc-weaponizing-open-source-software/>).

MSTIC has observed ZINC, also known as Lazarus, using weaponized versions of open source software including PuTTY, KiTTY, and TightVNC installer for these attacks, which have targeted “employees in organizations across multiple industries including media, defense and aerospace, and IT services in the US, UK, India, and Russia.”

## Linux Kernel 6.0 Announced

Linus Torvalds has released Linux kernel 6.0 (<https://lwn.net/Articles/910086/>), noting that the version number change is more a matter of practicality than reflective of any fundamental changes.

“But of course there’s a lot of various changes in 6.0,” Torvalds says, “We’ve got over 15k non-merge commits in there in total, after all, and as such 6.0 is one of the bigger releases at least in numbers of commits in a while.”

According to Jon Corbet at LWN.net, “headline features of this latest release include a number of io\_uring improvements, including support for buffered writes to XFS filesystems and zero-copy network transmission, an io\_uring-based block driver mechanism, the runtime verification subsystem, and much more.”

See change details in LWN’s merge-window summaries (part 1 [<https://lwn.net/Articles/903487/>] and part 2 [<https://lwn.net/Articles/904032/>]) and read more at LWN.net (<https://lwn.net/Articles/910086/>).

## Google Announces TensorStore for High-Performance Array Storage

Google has announced TensorStore, an open source, C++ and Python library designed for reading and writing large multi-dimensional arrays.

Many contemporary computer science applications manipulate huge, multi-dimensional datasets, says Google. “In these settings, even a single dataset may require terabytes or petabytes of data storage. Such datasets are also challenging to work with as users may read and write data at irregular intervals and varying scales, and are often interested in performing analyses using numerous machines working in parallel,” Google explains.

TensorStore is an open source software library that, according to the website (<https://google.github.io/tensorstore/#concepts>):



- Provides a uniform API for reading and writing multiple array formats
- Natively supports multiple storage drivers, including Google Cloud Storage, local and network filesystems, in-memory storage
- Automatically takes advantage of multiple cores for encoding/decoding and performs multiple concurrent I/O operations to saturate network bandwidth
- Enables high-throughput access even to high-latency remote storage

“Processing and analyzing large numerical datasets requires significant computational resources,” says Google, which “is typically achieved through parallelization across numerous CPU or accelerator cores spread across many machines.” Thus, according to Google, “a fundamental goal of TensorStore has been to enable parallel processing of individual datasets that is both safe (i.e., avoids corruption or inconsistencies arising from parallel access patterns) and high performance (i.e., reading and writing to TensorStore is not a bottleneck during computation).”

## NIST and Google Partner to Develop Open Source Chips

Google and the National Institute of Standards and Technology (NIST) have signed a research and development agreement (<https://www.commerce.gov/news/press-releases/2022/09/nist-and-google-create-new-supply-chips-researchers-and-tech-startups>) for production of open source chips.

The chips will be manufactured by SkyWater Technology in Minnesota, according to the announcement. “Google will pay the initial cost of setting up production and will subsidize the first production run. NIST, with university research partners, will design the circuitry for the chips. The circuit designs will be open source, allowing academic and small business researchers to use the chips without restriction or licensing fees.” As many as 40 different chip designs are planned, which will be optimized for different applications.

“The SkyWater foundry will produce the chips in the form of 200-mm discs of patterned silicon, called wafers,” the announcement says, and the first production run will be distributed to leading US universities. “Giving researchers access to chips in this format will allow them to prototype designs and emerging technologies that, if successful, can be integrated into production more quickly, thus speeding the transfer of technology from lab to market.”

“By creating a new and affordable domestic supply of chips for research and development, this collaboration aims to unleash the innovative potential of researchers and startups across the nation,” says Under Secretary of Commerce for Standards and Technology and NIST Director Laurie E. Locascio.

## Security and Compliance Drive Adoption of Observability Practices

New Relic’s 2022 Observability Forecast (<https://newrelic.com/observability-forecast/2022/about-this-report>) surveyed tech professionals across 14 countries to better understand their use and adoption of observability tools and approaches.

In regard to what is driving adoption of observability practices, the top response was an increased focus on security, governance, risk, and compliance (49%). Other factors included:

- Development of cloud-native application architectures (46%)
- Increased focus on customer experience management (44%)
- Migration to multi-cloud environment (42%)
- Adoption of open source technologies (39%)

The survey also asked whether practitioners viewed observability as more of an enabler for achieving business goals or more for incident response, with the following results:

- Observability was completely or generally for achieving core business goals (50%)
- Observability was equally for business goals and incident response (28%)
- Observability was more for incident response/insurance (21%)

The survey also found that “IT operations teams were most likely to be responsible for observability followed by network operations and DevOps teams.”

Questions about the present and future of OpenStack

# Maturity Processes

OpenStack has been on the market for 12 years and is generally considered one of the great open source projects. Thierry Carrez and Jeremy Stanley both work on the software and provide information about problems, innovations, and future plans. By Ulrich Bantle

**ADMIN Magazine:** The OpenStack Foundation has changed its name to the OpenInfra Foundation. Can you say something about the part OpenStack plays in this new setting?



Thierry Carrez, General Manager at the OpenInfra Foundation, was involved in founding the OpenStack project as a systems engineer and still contributes to its governance and release management. A fellow of the Python Software Foundation, he previously worked as technical lead for Ubuntu Server at Canonical, operations lead for the Gentoo Linux Security team, and IT manager at various companies.

**Thierry Carrez:** The foundation was originally created to host and promote the OpenStack project. In the process of doing so in the last 10 years, we assembled a wide community of operators, organizations, and developers interested in the concept of using open source solutions to provide infrastructure. With such an audience, it is only natural that we support and host other projects that are relevant for the same audience.

This is why we became the OpenInfra Foundation: to support other open infrastructure projects beyond OpenStack that are of interest to the group of infrastructure providers we assembled. OpenStack is still by far the largest project hosted at the Foundation, though, so it is central to all of our activities. With the change and the new project hosting offering we presented at the 2018 Summit in Berlin, we are ready to tackle the next decade of open infrastructure projects.

**AM:** OpenStack celebrates its 12th anniversary this year. Where do you see it in the coming years, and what will change?

**TC:** Today, OpenStack is the de facto open source standard for deploying

cloud infrastructure services, be it to offer private resources for a given organization or public resources for customers around the world with a credit card. Combined with the Linux kernel and Kubernetes for application orchestration, it forms a very popular open source framework, the Linux OpenStack Kubernetes Infrastructure (LOKI). Usage is growing significantly, driven by new requirements and new practices.

That said, OpenStack is functionally mature and its scope now well defined, so I expect new development to slow down and focus on maintenance going forward. In the same way Linux kernel development is driven by new capabilities in computer hardware, I expect OpenStack development also to be driven by new capabilities in data center hardware needing to be made available to the software.

**AM:** OpenStack is still a very successful project with big numbers, but some critics say the growth in installations belongs more to existing customers than to new customers. Is that right?

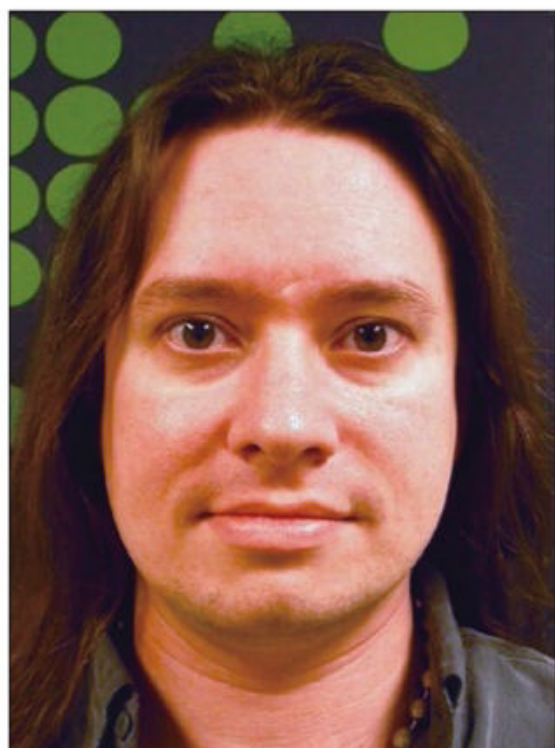
**TC:** It's really both. In June we held our first in-person event since the pandemic, the OpenInfra Summit in



Berlin [1]. We met one new user after another, many of them running at impressive scale surpassing 100,000 cores of compute on OpenStack. New requirements like digital sovereignty are driving a lot of adoption, especially in scientific use cases and the public cloud in Europe. The need to deploy new municipal and research systems rapidly for pandemic response also resulted in a lot of additional deployments worldwide. Existing implementations keep growing, as well. It's a healthy balance of energy and creativity for the community.

**AM:** When you take a look at possible competitors, why should users choose OpenStack and not AWS? What are the main differences?

**TC:** Users should always choose the solution that is the best for their specific needs. There are many reasons why someone would choose one, the other, or both. Our community tells us that they see OpenStack playing an important role to control their infrastructure costs, or to fulfill their regulatory requirements, or to answer very specific use cases for which



Jeremy Stanley has worked as a Unix and Linux sys admin for more than two decades, focusing on information security, Internet services, and data center automation. He is a core member of the OpenStack project's infrastructure team, serving on both the technical committee and the vulnerability management team.

hyperscalers can't provide a good solution. In those strategies, private OpenStack clouds are often used in combination with AWS, Azure, Google, or one of the many OpenStack-based public clouds in a multi-cloud or hybrid cloud deployment. That said, the main difference between OpenStack and proprietary clouds is the ability to participate in an open, active community and directly influence the direction of the project through personal involvement. That's something proprietary solutions will never match.

**AM:** Companies are desperately looking for developers and IT experts. What steps are planned to make OpenStack easier to use to mitigate this staff shortage?

**Jeremy Stanley:** It's funny: Every time I hear someone talk about how hard OpenStack is to deploy and manage, they inevitably have not touched it for years. The community has made operations a major point for new capabilities, and while infrastructure will always be hard, the stereotypical challenges to deploying and operating OpenStack are no longer issues.

**AM:** In detail, are there any plans regarding the development of OpenStack to make deployment less complicated?

**JS:** In recent years, the community has put a lot of energy into supporting installations for a variety of popular configuration management and orchestration solutions like Ansible, Chef, Helm, and Puppet, as well as focusing on both packaging for server distributions and container-based frameworks like Kubernetes.

**AM:** What do you think could make the installation more simple?

**JS:** Many new users don't realize that OpenStack isn't an all-or-nothing solution. It's a composable, modular suite of services, most of which are optional, and some can even be used entirely on their own. We're working to amplify this message, because a lot

of the perceived complexity is really a result of people mistakenly installing more than they actually need.

**AM:** Any plans to make the upgrading process less complex and with less downtime?

**TC:** At this point, upgrading an OpenStack cluster is a pretty established process, and it does not trigger significant downtime. At the same time, the size of deployments has grown significantly, with some now well surpassing the million-core CPU mark. That represents a lot of OpenStack clusters, so keeping them up to date with a release every six months is still significant work.

To reduce the pressure to upgrade for established deployments, starting with the OpenStack "Antelope" release in March 2023, OpenStack is offering the possibility to directly upgrade yearly instead of every six months. This should hopefully facilitate the lives of OpenStack operators.

**AM:** Can you describe the modular architecture of OpenStack?

**JS:** The fundamental design principle across OpenStack projects is that services interact with each other through REST (i.e., HTTP-based) interfaces to coordinate shared resources. Services are written in a consistent programming language and coding style, avoiding significant duplication through central libraries and consensus around dependencies. The result is a pluggable suite of service options that organizations can tailor to their particular use cases by installing just what they need, while still having the opportunity to grow their solution by adding other services as their requirements change.

**AM:** What should change in terms of the OpenStack architecture to make management easier?

**JS:** A significant undertaking, already well underway, is the implementation of an extensible role-based access control model. With this,

operators will be able to delegate permissions for some tasks to other users, reducing their own workload. Consumers of the cloud resources will likewise gain the ability to provide fine-grained control of specific actions to relevant stakeholders in their organizations. This work is still in its early stages, with the introduction of support for a read-only role to allow access to audit systems without the risk that the user might make changes; the ability to create more specific roles will follow in coming releases.

**AM:** What do you think are the key deployment challenges that organizations face with OpenStack?

**JS:** The hardest challenges, by far, are related to planning and sizing infrastructure. No two use cases are the same, and especially with added complications in obtaining hardware these days, it's more important than ever to be as accurate as possible and not overspend. Unfortunately, sizing a deployment (of anything beyond trivial complexity, not just OpenStack) is more of an art than a science. If you haven't done it often and for a wide variety of solutions, it's hard to even know where to begin. That's really one of the biggest reasons to establish a relationship with an OpenStack distribution vendor: They have more experience than anyone else at estimating how much of what sort of hardware you're

likely to need and how best to design the deployment to maximize efficiency and minimize cost.

**AM:** Maybe you could give a special view on the "Yoga" release and how it helps make things easier.

**JS:** It's hard to single out a few features, but one good example is that Ironi [for provisioning bare metal machines] continued to evolve its defaults to align with more modern server infrastructures, switching from legacy BIOS to UEFI and emphasizing local boot workflows for deployed images. Another is Kolla [production-ready containers and deployment tools], which focused on a single set of container images, rather than trying to maintain two different solutions in parallel, and simplified its design with the addition of a new Ansible collection that shared across its components. Additionally, Nova added support for offloading the network control plane to a SmartNIC, freeing up more capacity for hypervisor workloads, and can also now provide processor architecture emulation for users who want to run software built for ARM, MIPS, PowerPC, or S/390 systems, all on the same host.

**AM:** Yoga Neutron has a new feature called Local IP. Can you explain the benefits in performance?

**TC:** Local IP is primarily focused on high efficiency and performance of

the networking data plane for very large scale clouds or clouds with high network throughput demands. Local IP is a virtual IP that can be shared across multiple ports or VMs, is only reachable within the same physical server or node boundaries, and is optimized for such performance use cases.

**LM:** You gave a talk at the OpenInfra Summit to collect scaling stories. Can you give a short overview of what users reported there and what the problems in scaling could be?

**TC:** It was a standing room only session, and it was great to see operators of large deployments at UbiSoft, Bloomberg, OVHcloud, CERN, Workday, or Adevinta share their experiences and pain points. They primarily discussed two classic scaling pain points in large OpenStack clusters: RabbitMQ and Neutron. In particular, they traded tips on how best to monitor and alleviate load issues.

This discussion was driven by the operators themselves within the OpenStack Large Scale SIG and is a great example of what open communities can achieve working together – something that is only truly possible with open source software. ■

#### Info

**[1]** OpenInfra Summit Berlin: [<https://www.youtube.com/hashtag/openinfrasummit>]



# Hone your skills with special editions!

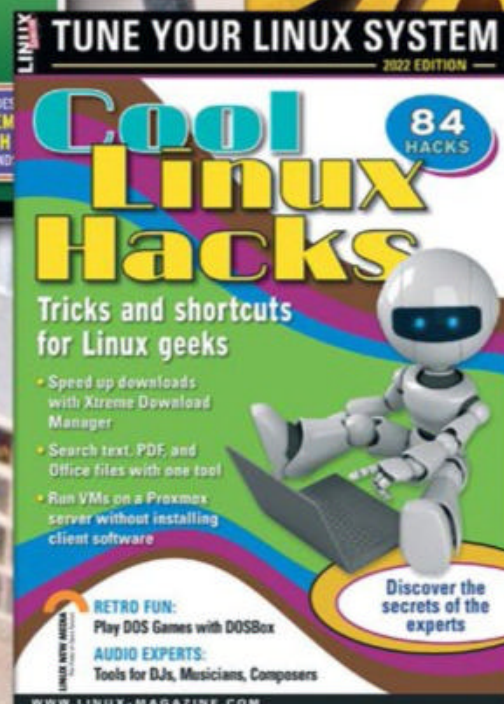
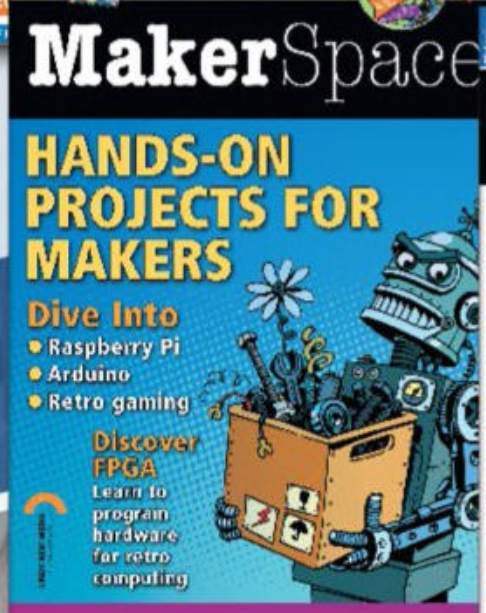
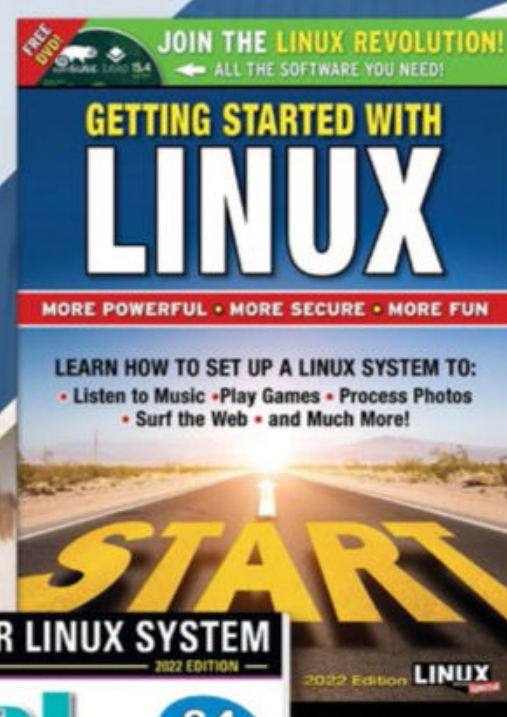
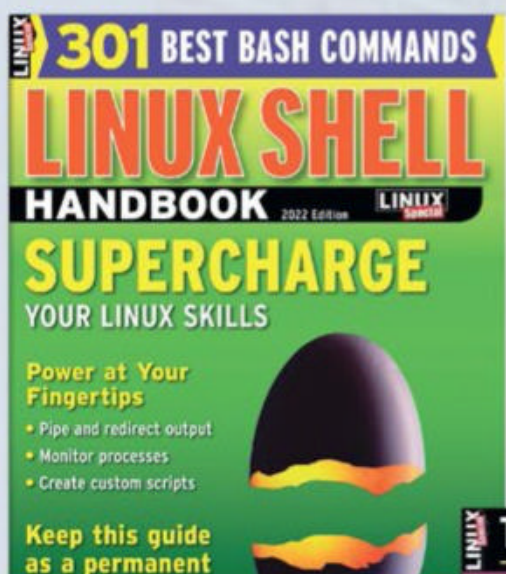
Get to know Shell, LibreOffice, Linux, and more from our Special Edition library.

The *Linux Magazine* team has created a series of single volumes that give you a deep-dive into the topics you want.

Available in print or digital format

**Check out the full library!**

[shop.linuxnewmedia.com](http://shop.linuxnewmedia.com)







The state of OpenStack in 2022

# Hangover

The unprecedented hype surrounding OpenStack 10 years ago changed to disillusionment, which has nevertheless had a positive effect: OpenStack is still evolving and is now mainly deployed where it actually makes sense to do so. By Martin Loschwitz

**In the middle** of the past decade, it seemed that the leaders of the OpenStack Foundation could hardly believe the success of their product. For a long time, OpenStack trade fairs and, above all, the OpenStack Summit were surrounded by something like a mystical aura: 8,000 participants and more regularly joined the private cloud computing environment camp to catch up on the latest technology. In 2012, an OpenStack sensation began that the industry had never experienced before, especially in the open source software context.

A big bang is inevitably followed by a big hangover at some point, and OpenStack was no exception. Some participants even left the OpenStack party while it was still in full swing. Large OpenStack projects sprang up like mushrooms, fizzling out after months or years, and leaving little behind except frustrated people. Others

turned to Kubernetes, which was about to become the next big thing. Following that, the media went quiet about OpenStack.

Some believe that OpenStack has disappeared from the scene, but – to paraphrase Mark Twain – the news of its demise is exaggerated. OpenStack is still an active project today, albeit with a much smaller community. However, it is still evolving, which is reason enough to take a fresh look at OpenStack and ask: What has changed technically, organizationally, and administratively?

## OpenInfra

The most noticeable innovation, and one that has packed the biggest punch, is the reorientation of the OpenStack Foundation. Originally, the Foundation had formed to give OpenStack a non-commercial home,

which makes sense in the US in particular for various reasons (e.g., a virtual-only project cannot hold any rights to trademarks or brands). On top of that, someone had to pay for the OpenStack party, and one of the Foundation's core tasks is to raise sponsorship money. This task is done on a corporate and individual membership basis, and the money is used to fund the (fairly expensive) OpenStack Summits ([Figure 1](#)), among other things.

At the height of the OpenStack hype cycle, the OpenStack Foundation managed to gain significant influence with major vendors like VMware, Intel, Red Hat, and others. As OpenStack's relevance waned, that influence threatened to shrink again, much to the displeasure of the folks at the Foundation. A few years ago, a course of reorientation was determined: The OpenStack Foundation

Photo by vadim kaipov on Unsplash



became the OpenInfra Foundation, and the OpenStack Summit became the OpenInfra Summit. Ever since, OpenStack has been a topic of interest, but it is no longer the only subject of interest to the Foundation. In recent months, for example, the Foundation has successively revised many tools from OpenStack development, systematically placed them under some kind of project management, and published them. The topics tackled are something like

the evergreens of the infrastructure and open source theme: Zuul, for example, the development suite that the Foundation uses to develop OpenStack, is now frequently used outside the project. As with OpenStack itself, the Foundation is responsible for ensuring that the project does not disappear without a trace. Additionally, the OpenInfra Foundation has put out feelers in quite a few other directions. Similar to the Linux Foundation, it offers an umbrella for

open source projects, although the focus is on software that is somehow used in the infrastructure sector. The log manager, Loki (Figure 2), a lightweight alternative to the classic Elasticsearch-Logstash-Kibana (ELK) stack, has now also slipped under the umbrella of the OpenInfra Foundation. Critics welcome this development because the apparent omnipotence of the Linux Foundation had gradually become a problem for many. However, it remains to be seen whether the OpenInfra Foundation will succeed in becoming a long-term and powerful counterweight.

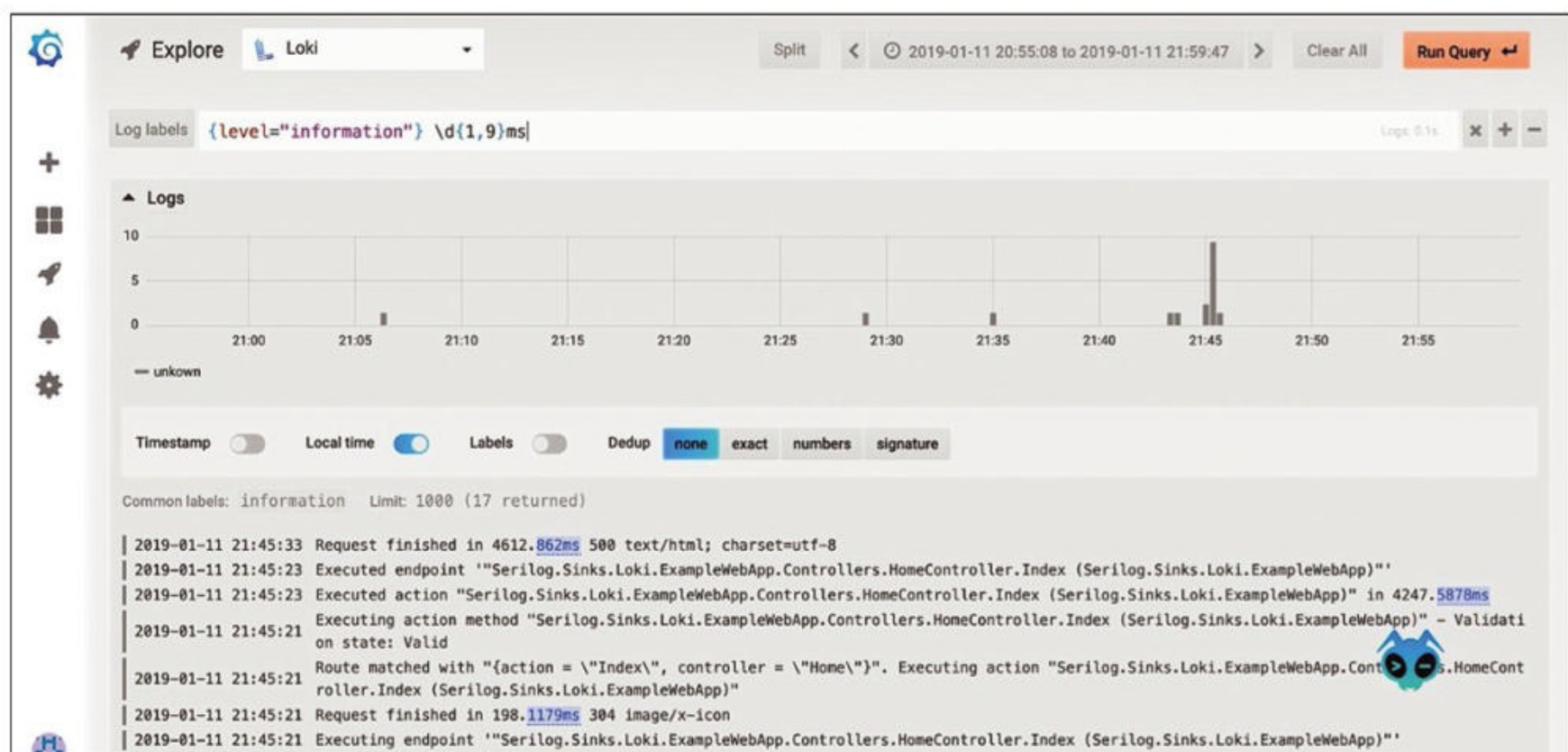
## Technical Progress

The truism “you’re always smarter in retrospect” fully applies to OpenStack. In terms of technology, many of the teething troubles of the early years are just now disappearing from the environment. Several factors play a role: Something recognizably good for the project is the waning interest of the major vendors.

Red Hat and Canonical are officially committed to OpenStack to this day, but SUSE, an OpenStack pioneer, has largely abandoned OpenStack development and pulled the plug on its own OpenStack distribution, SUSE Cloud. Accordingly, there is little coming out of those quarters



**Figure 1:** OpenStack Summit 2022 in Berlin, live again for the first time since the emergence of COVID-19, is still a reunion of sorts, albeit a much smaller one. © OpenInfra Foundation



**Figure 2:** Loki is a lightweight alternative to the ELK stack and is now one of the OpenInfra Foundation projects. © Loki



in terms of OpenStack source code at the moment. However, it is apparent that the large number of companies that wanted to influence OpenStack, especially in the initial phase of the project, was probably more of a curse than a blessing. A good example of this is provided by Cinder, the OpenStack component that manages storage. When industry leaders argue about architecture decisions on the Cinder mailing list, the only way out is always the lowest common denominator. Now that fewer major players are involved, developers have occasionally dared to make major changes and features that break with old compatibility defaults.

On top of that, the OpenInfra Foundation has now found a stable work approach to help manage the individual OpenStack components and their development in a meaningful way. To date, each component has an engineering owner elected by the community from among its members. The owner is responsible for the development content and is newly elected for each OpenStack development cycle.

## Few Massive Changes

If you look at the present OpenStack changelog, you quickly notice that the individual components are now being developed more thoughtfully.

In pioneering days, the industry had high expectations that each new OpenStack version had to light a bonfire full of features with groundbreaking innovations. Currently, OpenStack is developing at a far slower pace. Although OpenStack releases do still include thousands of commits and a very long changelog, today the release highlights (i.e., the most important and profound changes in an OpenStack release) tend to fit on a few screen pages, mostly because new OpenStack versions no longer come with new components that no one has heard of before and then disappear into a black hole again two releases later.

The last major OpenStack release (“Yoga,” numbered 25) came out at the end of March. A look at its release highlights underscores the fact that OpenStack is now more cautious than it was a few years ago. OpenStack Blazar, for example, is a relatively new OpenStack component that lets users reserve resources for themselves without using them. This ability improves reliability for applications that need guarantees in terms of resources available and the extent of these resources at a certain point in time. In OpenStack Yoga, Blazar can now store more details about individual virtual instances so that the OpenStack Nova scheduler chooses the hosts in a more targeted way.

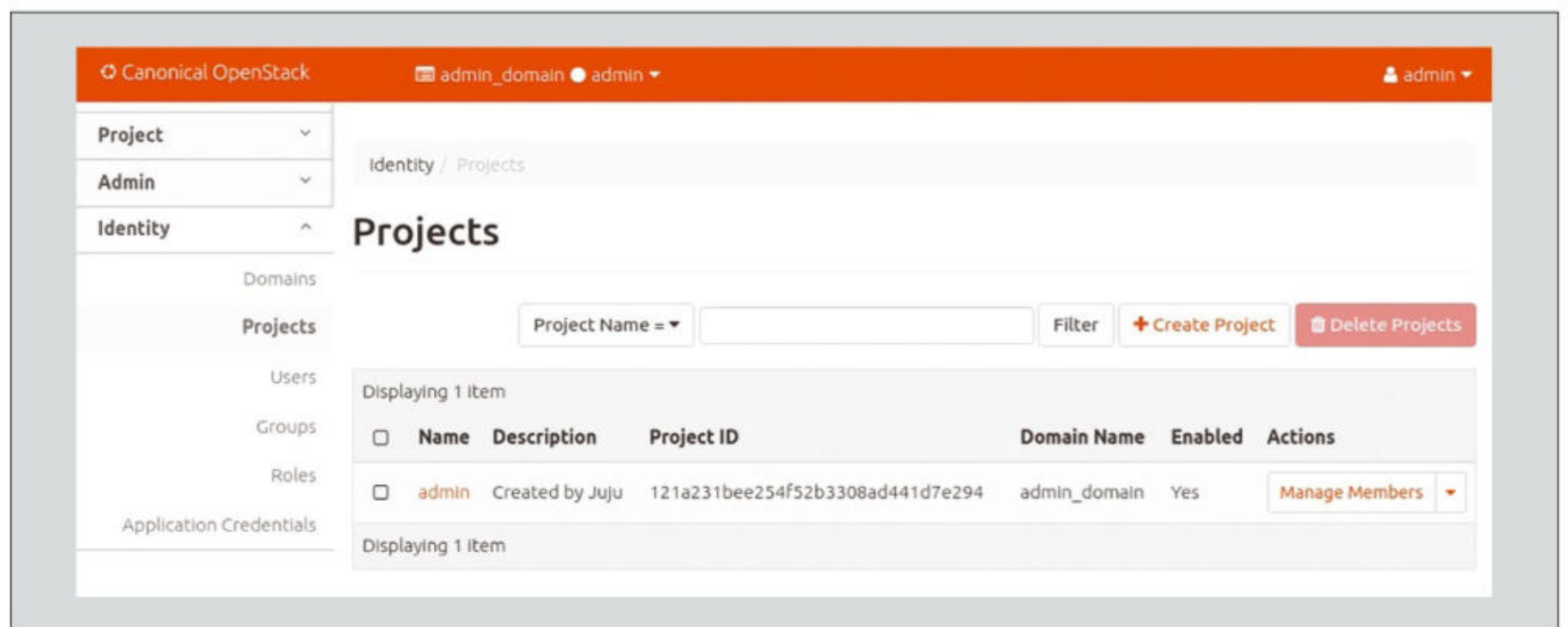
## The Big Six

Most of the changes are found in the six core platform components that together form an OpenStack instance. For example, the storage manager Cinder can now overwrite a virtual volume with the contents of a hard drive image. Until now, this was only possible during virtual machine (VM) creation. Now, however, existing VMs can be re-created without deleting and re-creating the entire VM or disk (reimaging).

Moreover, Cinder now has support for new hardware back ends. In the future, it will be able to use Fibre Channel to address its storage in the form of LightOS over NVMe by TCP, just like Netstor devices from Toyon. NEC V series storage will also support both Fibre Channel and iSCSI in the future. The Glance image service has always been one of the components with more relaxed development; consequently, it offers little new. Quota information about the stored images can now be displayed; some new API operations for editing metadata tags, for example, have been added; and there are more details about virtual images residing on RADOS block devices (RBDs) in Ceph.

## Horizon, Nova, Keystone

The Horizon graphical front end has been massively overhauled by the



**Figure 3:** The developers have successively improved the theming options for the OpenStack dashboard. Horizon on Ubuntu appears today in the typical Canonical orange. © Ubuntu

folks behind OpenStack over the past two years and now supports translation to other languages and theming in a far better way. Canonical and Red Hat placed great emphasis on this dashboard: After all, Canonical Ubuntu wants to impress in orange (Figure 3), whereas Red Hat OpenStack accents its interface with red (Figure 4). Functionally, the dashboard in Horizon now has the ability to create and delete quality of service (QoS) rules.

The most important innovation in Yoga for the Nova virtualizer is the ability to pass smart network interface cards (SmartNICs) directly through to a VM by way of network back ends. Traditionally, a software-defined network (SDN) in the OpenStack context works by creating a virtual switch on the host that takes care of packet delivery. However, this approach to processing packets has the major disadvantage that it aggressively hogs the host CPU. For this reason, practically all major network manufacturers are now offering NICs whose chips can process Open vSwitch. Therefore, control over network packet delivery is offloaded from the host to the NICs, where it is

handled more efficiently – lifting the load off the host’s shoulders.

Previously, Nova could use NICs by selecting the appropriate driver for the emulator in use. Now, however, Nova has a setting on the back end for the virtual network card, so the feature has been standardized and can theoretically be used by any network back-end driver in Nova.

You can see how established OpenStack thinks it has become by looking for Keystone in the OpenStack release highlights. The component is responsible for user and project management and is therefore an absolutely essential feature; however, the developers consider nothing in Yoga relevant enough to warrant a special changelog entry. The developers themselves would probably not call the Keystone feature complete, but the intent is to protect admins against too many changes in Keystone in the foreseeable future.

## Kubernetes on the Horizon

Kubernetes on OpenStack (actually Kubernetes on OpenStack on Kubernetes) is a fairly common deployment scenario for the free cloud

environment today. Accordingly, it is important to the developers to optimize OpenStack integration with Kubernetes to the best possible extent. More recently, a number of components have been added to OpenStack for this purpose, including Kuryr. Like OpenStack, Kubernetes is known to come with its own SDN implementation. It used to be quite common to run the virtualized Kubernetes network on top of the OpenStack virtual network – much to the chagrin of many admins, because if anything went wrong in this layer cake, it was difficult or impossible to even guess the source of the problem.

Kuryr remedies this conundrum by creating virtual network ports in OpenStack so that Kubernetes can use them natively and dispense with its own SDN. This solution implicitly enables several features that would otherwise go unused, such as the previously described outsourcing of tasks to NICs, including offloading support, or the seamless integration of as-a-service services from OpenStack in Kubernetes environments.

Octavia, the load balancer as a service, plays a prominent role in the OpenStack universe today but can

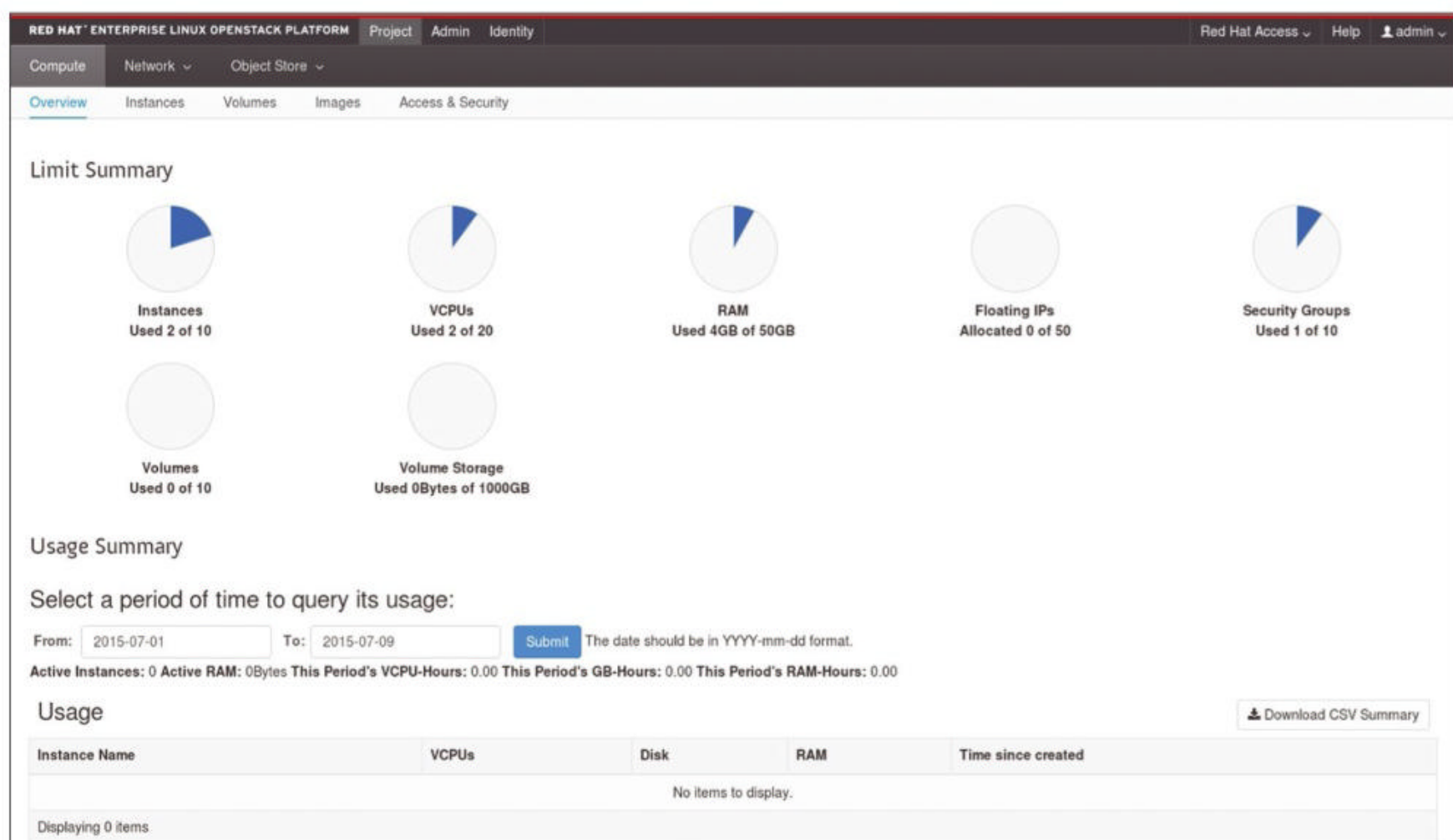


Figure 4: The Red Hat OpenStack Platform (RHOP) uses the typical Red Hat look found in other products. © Red Hat



only be meaningfully used for Kubernetes if you have Kuryr. In this case, meaningful means without stacking several balancer instances on top of each other.

## Across the Board

Many small changes to the various OpenStack services are now occurring across the board. In the past few months, for example, the developers have massively upgraded Designate. DNS as a service is important for many end users because a website hosted on OpenStack also needs to be accessible under a brand name and not just an IP address.

For a long time, however, Designate was only capable of basic functions and, depending on the network backend selected, did not work particularly well. Today, Designate is in good shape, talking to all available SDN back ends and offering a cornucopia of features. For Yoga, the developers have focused on bug and error hunting, expanding their internal testing program and eliminating various bugs. IPv4 and IPv6, distributed name server records, and special forms of DNS records (e.g., TXT lines) have all been offered by Designate for some time. All told – perhaps more than any other OpenStack component – Designate demonstrates that today the platform’s focus is on stability and reliability, not just the next big thing.

## Paradigm Shift

What OpenStack developers have done recently in terms of OpenStack deployment is almost a paradigm shift of sorts. For a long time, the topic was considered external: Tools based on Ansible, Puppet, or Chef, for example, were good enough to handle OpenStack deployment. This line of reasoning crumbled quite early – at the latest, in OpenStack Ironic. The service that establishes lifecycle management for hardware of almost any provenance in OpenStack is described as bare metal as a service. Most impressively, Ironic lets you

manage OpenStack hosts like OpenStack VMs. TripleO – OpenStack on OpenStack – is the name of the construct in this case. A server running with a base OS alone is more or less meaningless to OpenStack; you need the components for it to count as OpenStack. Although they now all come in the form of containers, you still have to roll them out and configure them.

For this reason, OpenStack-Ansible is now part of OpenStack, although a separate component. It handles tasks in the TripleO context but can be equally well used standalone to get hosts ready for use with OpenStack. That this strict separation between component and deployment from the early days is no longer practiced is good and important. All told, it adds value to the OpenStack deployment.

## What Is Not Better

Despite all the praise for the community making the right decisions in the recent past, OpenStack still has elements that don’t work reliably or at all. Some of these can be tolerated, but others are very significant.

Today, it is good form for cloud environments to offer as-a-service services to their own user bases. OpenStack has various as-a-service components (e.g., NFS as a service, alias Manila, is well known). However, not everyone who runs a web store is an IT geek with a command line background, so services such as databases need to be accessible to less savvy users. Amazon, for example, is leading the way with its database-as-a-service offering.

OpenStack doesn’t seem to be very innovative when it comes to database as a service. Although it already had a component for this service on board, in the form of OpenStack Trove, its development has been idle for more than two years, probably because the last active maintainer has since been poached by a competitor, who deleted the OpenStack topic from the new, joint portfolio shortly afterward. OpenStack users are frustrated and are forced to build their databases

manually and painstakingly – if they don’t immediately launch their own wild projects, as is the case in many places.

Another aspect of OpenStack that has shown a conspicuous absence of virtually all required features for many years is billing. The Ceilometer component writes user data and stores it in a database called Gnocchi. This component is a pure metric data meter and does not offer any possibility to create invoices with the acquired data.

Granted, OpenStack doesn’t make things easy at this point. The complexity of the subject of billing almost implicitly ensures that you need to do more than simply collect data and generate an invoice as a PDF. Instead, it is important to pay attention to various aspects. From the customer’s point of view, for example, you need to understand down to the bottom line what you are actually purchasing. If the provider claims on its invoice that a VM ran from 5:35am to 6:35am, it must be able to demonstrate this plausibly in the event of an inquiry. Moreover, various formal requirements apply to invoices, such as the obligation to number them consecutively, the correct disclosure of sales tax depending on the recipient country, and so on. OpenStack developers appear to be reluctant to deal with these details. The premise has always been that OpenStack would not handle billing and would use an external service for this purpose. For an external service to work efficiently, OpenStack at least needs some sort of generic interface for exporting arbitrary processed data. Many companies rely on proprietary accounting tools, to which OpenStack would then of course have to connect. In practice, however, OpenStack already has a problem providing data for other services. Quite a few approaches were attempted (e.g., CloudKitty, which itself accesses data from Ceilometer); however, none of the solutions has been a resounding success, and to date only proprietary products are cobbled together for specific clouds based on OpenStack.



A handful of commercial vendors have added OpenStack to their billing tools, but companies that don't use one of these tools continue to be frustrated. If you want to run OpenStack in a way that makes sense, not only technically but also commercially, you more or less have to build the entire billing system yourself. It is high time for developers to come up with a generic interface that billing service providers can dock onto.

## Conclusions

Much achieved, much to do – that is my conclusion with regard to OpenStack after a little more than 12 years of project history. What started life as a cooperation between NASA and Rackspace has become a comprehensive platform for private clouds with high standards.

However, not just OpenStack has changed, but also the industry's take on the project. It is no longer the case that the smallest service providers are trying to join the OpenStack circus, because the solution is now considered a professional tool for large environments. One concern for some is that it has also significantly reduced the number of vendors with OpenStack products in their portfolio. Apart from Canonical and Red Hat, none of the big vendors offer a serious OpenStack distribution.

Resistance is brewing. Kurt Garloff, who is very well connected in the open source and OpenStack scene, has been working for some time on his Sovereign Cloud Stack (SCS) [1], which is based on OpenStack at its core (Figure 5) and is intended to enable data sovereignty by operating

in a company's own cloud. (See the article on SCS in this issue.) Garloff has already found some helpers, and the technical development of SCS is certainly something that whets your appetite for more. Competition and diversity don't hurt, so here's hoping SCS can establish itself as an alternative to the big two and help OpenStack gain new momentum. ■

### Info

[1] Sovereign Cloud Stack:  
[<https://scs.community>]

### The Author

Freelance journalist Martin Gerhard Loschwitz focuses primarily on topics such as OpenStack, Kubernetes, and Chef.

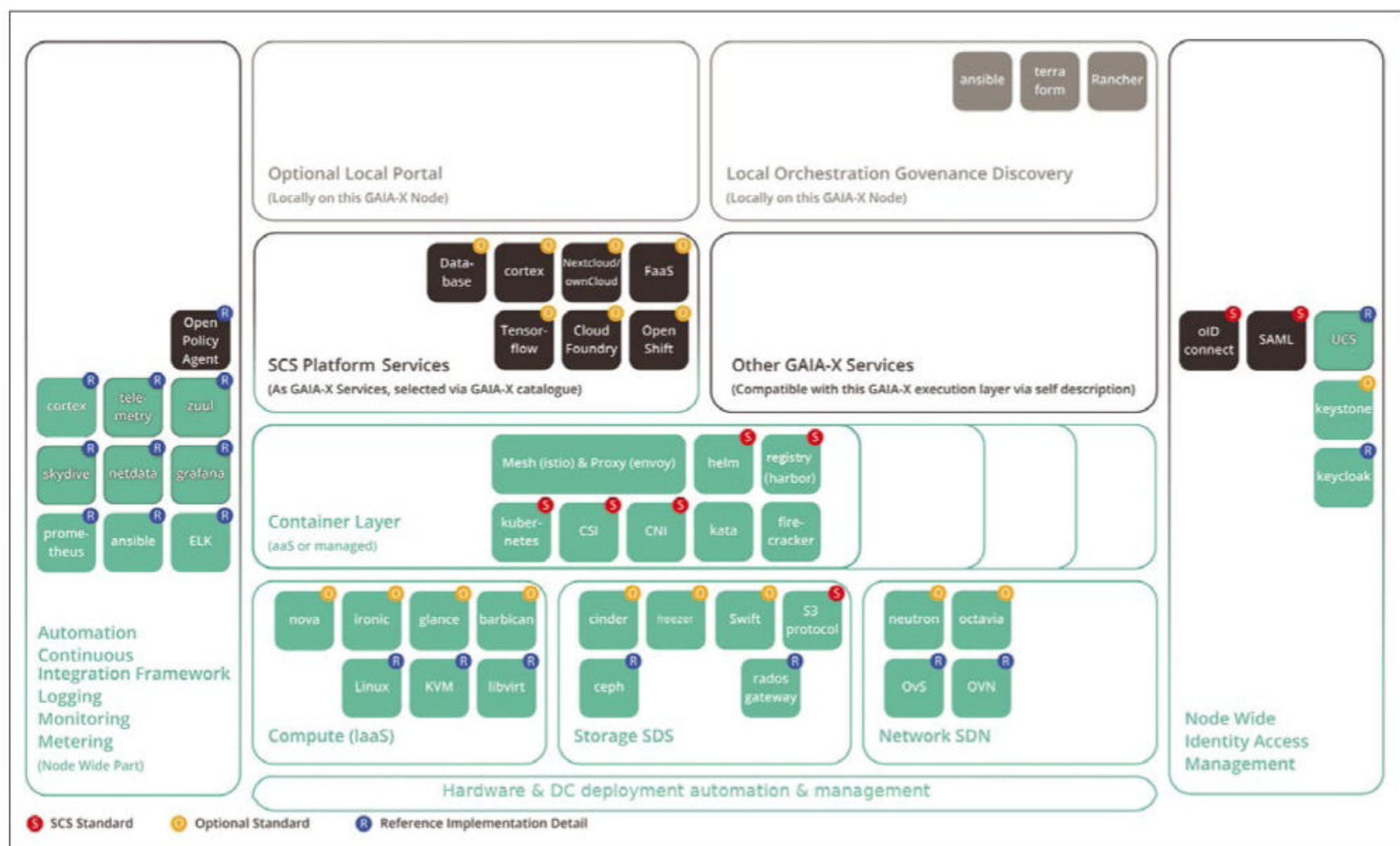


Figure 5: Sovereign Cloud Stack provides a reference architecture for clouds with OpenStack at its core. © Sovereign Cloud Stack



## OpenStack observability with Sovereign Cloud Stack

# Guard Duty



Operators of an OpenStack environment need to know whether the environment is working and quickly pinpoint problems. We look at the basics of OpenStack observability in the Sovereign Cloud Stack. By Felix Kronlage-Dammers

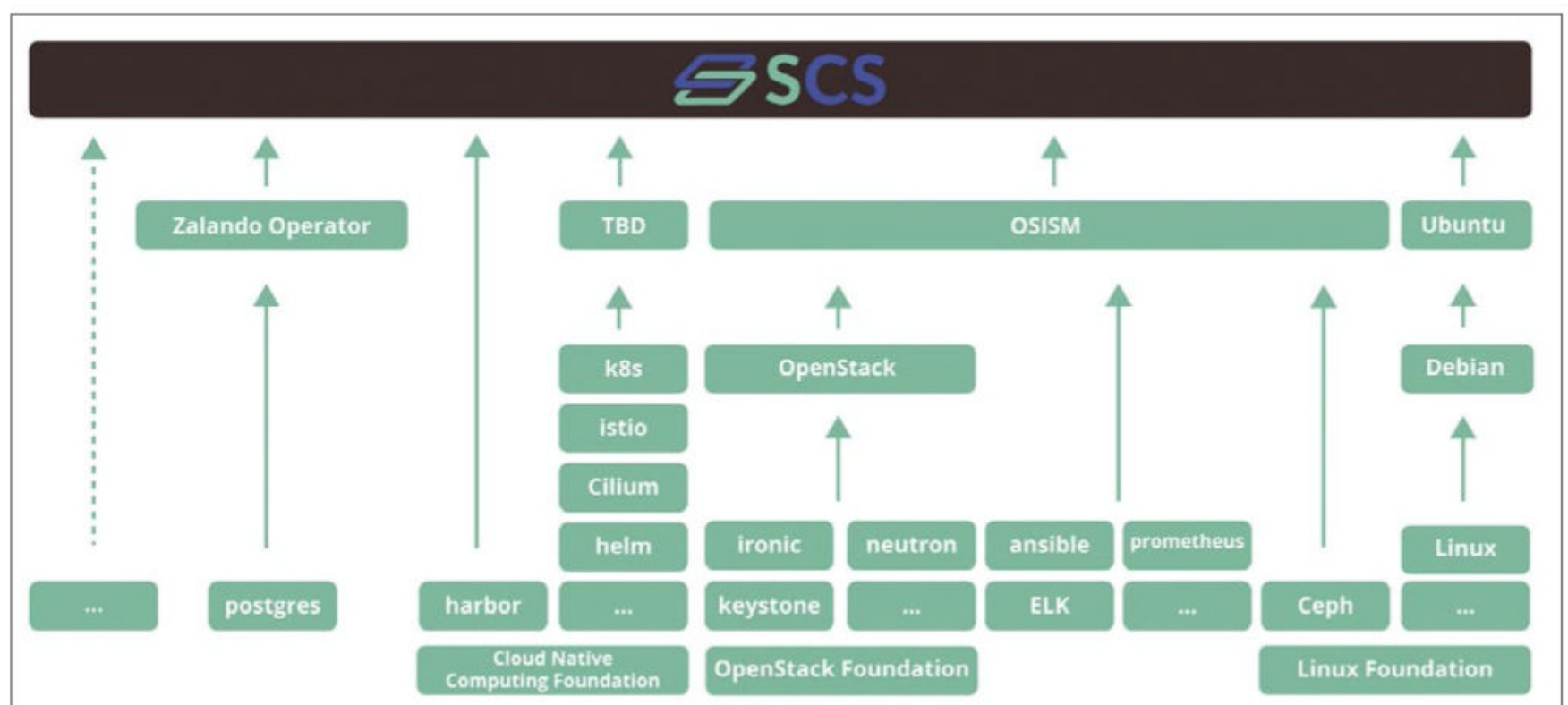
**The Open Source Business** Alliance (OSBA) Sovereign Cloud Stack (SCS) project was launched in 2019 to enable federation-capable cloud environments. In close collaboration between the community and OSBA's SCS team, the project uses an agile approach to create the appropriate

standards and a reference implementation. SCS [1] is an open, federation-capable modular cloud and container platform based on open source software. It builds on proven open source components such as OpenStack and Kubernetes in the reference implementation (**Figure 1**). As a platform,

SCS provides the foundations for creating offerings that deliver full digital sovereignty.

### SCS Community

The community comprises a wide variety of cloud service providers



**Figure 1:** Visualization of the relationships between Sovereign Cloud Stack components.

Photo by Michaela Filipcikova on Unsplash



(CSPs) and their employees, people from the OpenStack community, and companies working on deliverables that are awarded through open tenders. Collaboration between the different providers creates a level of cooperation that rarely exists elsewhere. Various teams and special interest groups (SIGs) work together on the topics, coordinating standards and requirements in terms of content. The primary goals of SCS are for teams to take an active part in the respective upstream projects, to be involved in the creation of content, and not to allow a parallel world to develop. The first such group to be formed out of the Operations and Identity and Access Management team was the Monitoring SIG, which is dedicated to the topics of monitoring and observability.

## Observability

Observability does not just mean plain vanilla state analysis (i.e.,

information on whether or not a service feature is available). Monitoring in line with contemporary standards also includes collecting, storing, and visualizing metrics and aggregating log messages. The overall system enables the operator to get a comprehensive picture of their system to locate problems, their origins, and causes quickly.

In the CSP environment, additional requirements to delete accruing data are based on appropriate specifications. Additionally, for compliance reasons, certain system messages need to be logged or billing data generated from the metrics that occur.

## Architecture

When the SIG was founded a year ago, it first collected the requirements of the participating CSPs and contrasted them with the individual components already covered in the reference implementation.

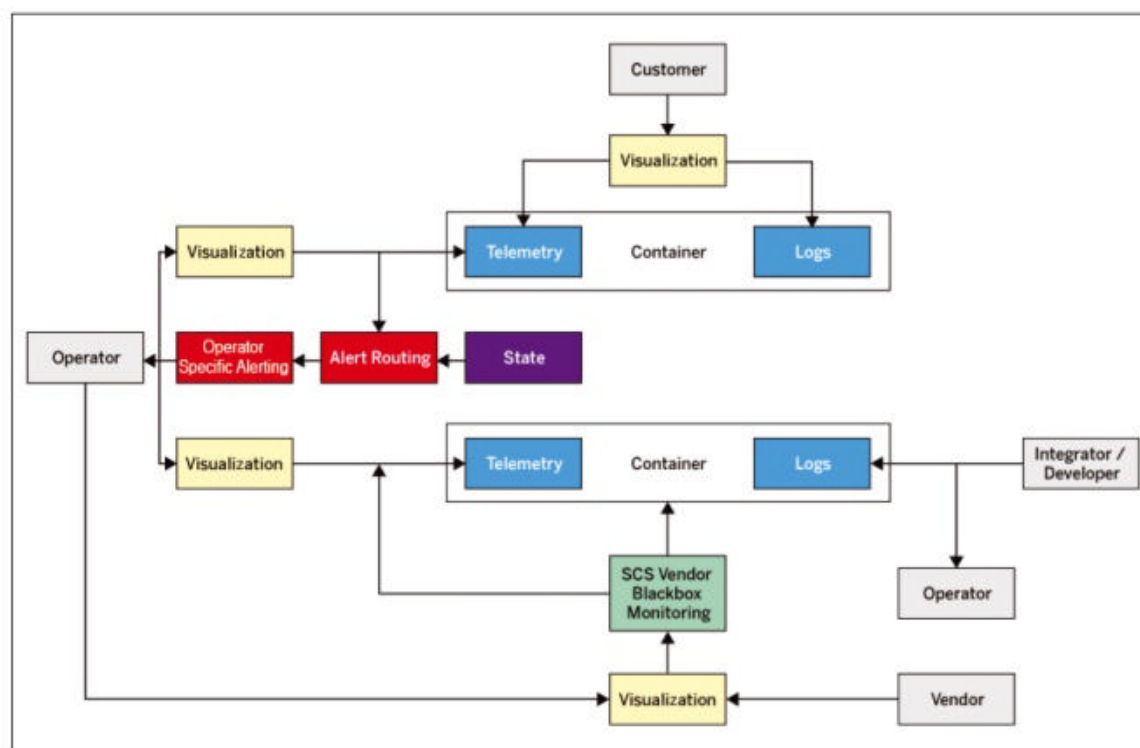
To avoid getting lost in discussions relating only to the tool right from the outset, the first step was to take a tool-agnostic look at the architecture.

**Figure 2** shows the intended architecture. First, the various roles were defined. The SCS Operator is the system operator (e.g., the CSP), whereas the SCS Integrator describes the companies that assist the operator in setting up and commissioning the system, if needed.

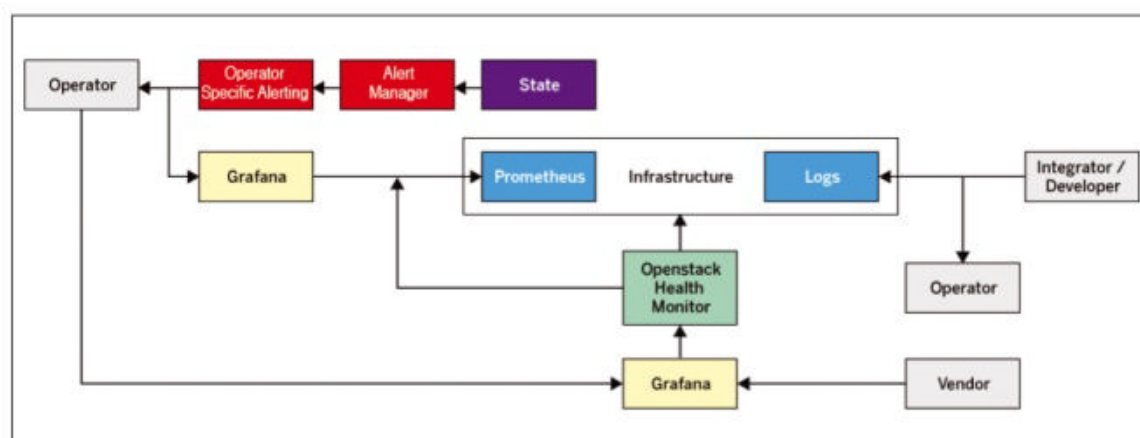
For several reasons, the developers made a deliberate decision to ignore the container layer in the first step and focus instead on the infrastructure-as-a-service (IaaS) layer. On the one hand, this approach reduced complexity and the number of decisions. On the other hand, observability was part of an award package that was not scheduled until 2022. At the IaaS level, the focus is on OpenStack with its various components (**Figure 3**).

## Components and Tools

SCS relies on the Open Source Infrastructure and Service Manager (OSISM) [2] as a tool for deployment and day 2 operations for OpenStack. OSISM itself relies heavily on Kolla Ansible [3] and on OpenStack-Ansible, a collection of Ansible playbooks for deploying OpenStack. One of the main focuses of OSISM is to simplify the operation of OpenStack-based systems and, in particular, upgrades from one OpenStack version to the next. The goal is to be able to install updates on a system at any time. Kolla Ansible comes with a Prometheus-based [4] monitoring stack out of the box, which coincided very well with the Monitoring SIG favoring an OpenMetrics-based approach from the outset. Initially, the use of traditional monitoring software, such as Zabbix or Icinga for service state monitoring, was considered. However, it became clear relatively quickly in the discussions that these scenarios could just as easily be covered by Prometheus' Blackbox exporter. With a view to



**Figure 2:** The tool-agnostic architecture of SCS.



**Figure 3:** The IaaS monitoring components of SCS.



reducing complexity, it makes sense to rely on the Blackbox exporter instead of a completely independent software solution. These changes were incorporated into OSISM; Zabbix, which had previously been included, was dropped.

As a first step, additional dashboards were provided for Grafana (Figure 4) and integrated into the Kolla Ansible project. Additionally, various exporters for Prometheus, which are currently not part of Kolla Ansible, should be included.

## Alerting

Alerting is an important component in any monitoring setup. It quickly became clear in the Monitoring SIG that every CSP that is not just starting to commission a corresponding environment already has an alerting system in operation. Ideally, the monitoring supplied with SCS would dock onto it.

Therefore, the decision was made to opt for the Prometheus Alert Manager, which is already integrated in Kolla Ansible, and to document best practices [5] for connecting to external alerting systems. The open source Alerta [6] software provides an alternative for aggregating alert occurrences at this point. Initially,

the idea of integrating it directly was considered; however, for the time being, Alert Manager was deemed sufficient.

Alert rules are an important part of Prometheus monitoring. To create a good starting point, several rule sets have been adopted from the Awesome alert rules [7] project, and they are now also making their way into Kolla Ansible.

## There's Monitoring and Then There's Monitoring

When the talk turns to monitoring, people tend to talk first about simple process monitoring. Does the *Foo* process exist, and does the *Bar* service respond on port 42? Often, instead of simply checking whether a service responds on a port, it is a good idea to use test scenarios that carry out a functional check of the service.

For example, in an environment like OpenStack, it's helpful to know whether the Horizon web front end is being delivered correctly to the browser or whether an API should be used to check that VMs can be started. However, checking a network component such as Open Virtual Network (OVN) for correct functionality can become complex.

To monitor OVN efficiently, the SIG is currently working on integrating the OVN exporter [8] upstream to provide various OVN metrics for Prometheus. Figure 5 illustrates where the exporter needs to reside to capture data from the redundant components and detect failures.

## From Health Monitor to CloudMon

OpenStack Health Monitor is an important component for the SCS project. The program relies on the OpenStack API to cover various scenarios and assess the functionality of an OpenStack cloud from a customer perspective.

The OpenStack Health Monitor periodically installs virtual machines, creates load balancers, and checks the availability of these components. It also logs runtimes so that changes such as delayed provisioning of virtual machines can be tracked. This feature is especially helpful when changes are made to the system environment, such as after an upgrade to OpenStack components, to help identify problems – ideally before end users notice them.

The OpenStack Health Monitor first saw the light of day more than four years ago in the Open Telekom Cloud

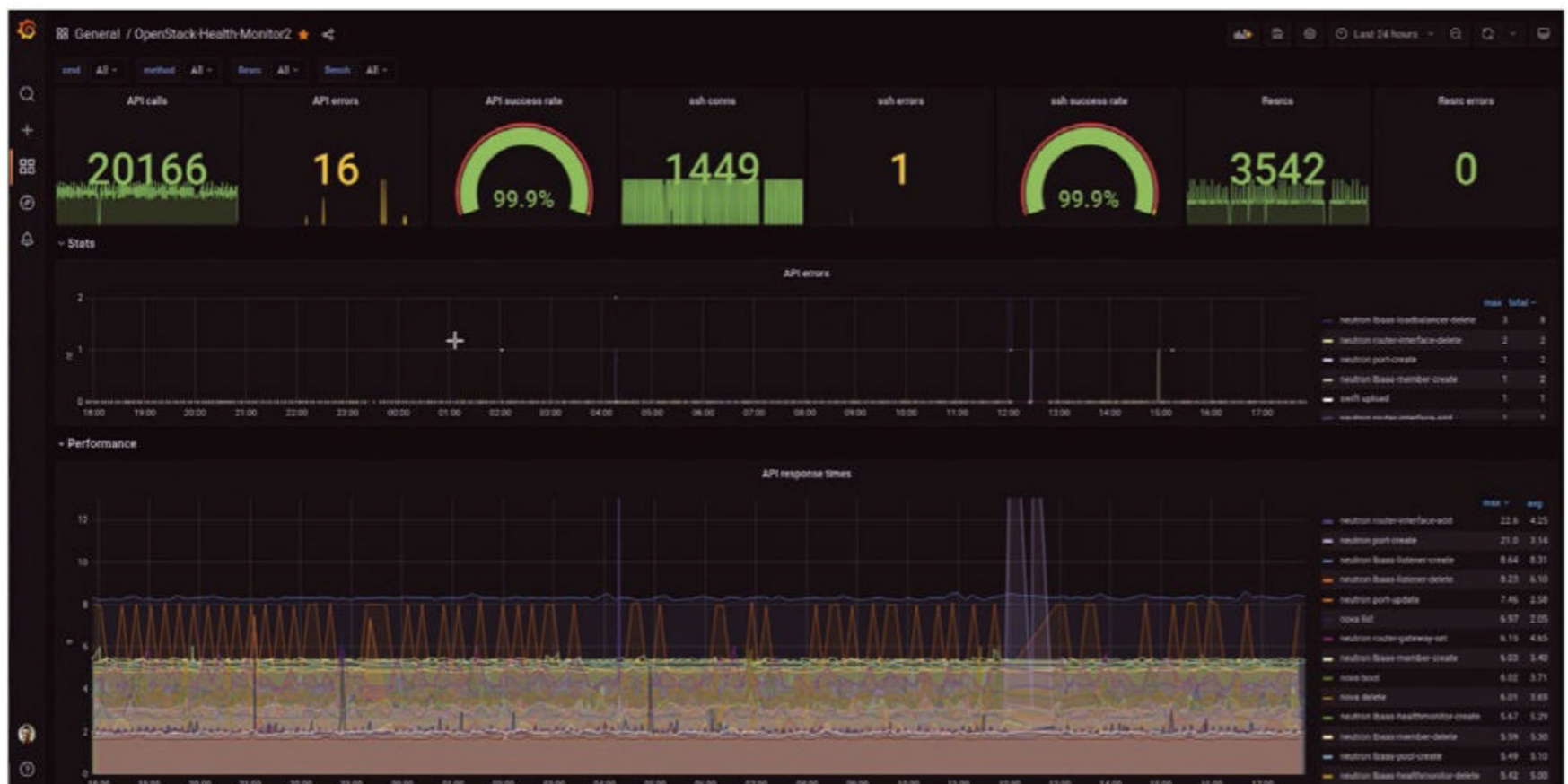


Figure 4: The Grafana dashboard of the OpenStack Health Monitor.



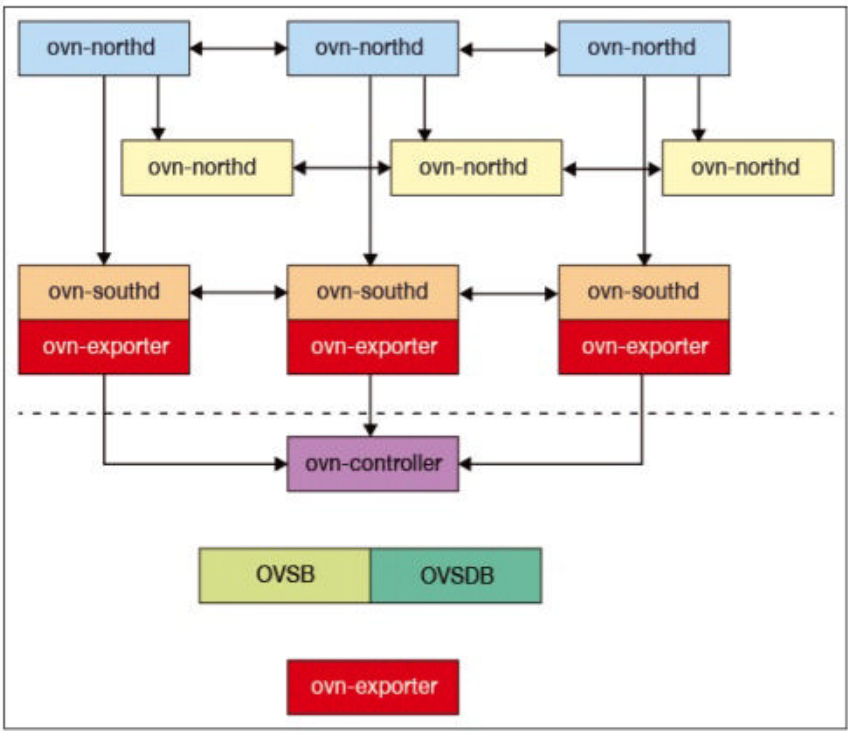


Figure 5: Open Virtual Network setup with Prometheus exporters.

(OTC). The project was initiated then by Kurt Garloff, who is the current CTO of the SCS project. Colleagues at OTC have since further developed the OpenStack Health Monitor for their environment and presented the resulting APIMon project at the OpenInfra Summit 2020 [9]. Following OpenInfra Summit 2022, APIMon developers met with members of the SCS community to consider how to collaborate on the next iteration. The intent was to develop and maintain an independent successor project, CloudMon, to serve as a reference for behavior-driven monitoring.

audit logging), an installed SCS environment would ideally already include appropriate underpinnings. It would provide SCS operators best practices for aggregating relevant log messages, such as those relating to attempted logins processed by Keycloak. The reference implementation of the Sovereign Cloud Stack can be easily put through its paces in the OSISM testbed [10], and it also provides a good introduction to the project. All meetings related to SCS can be viewed on the project website under *Contribute to SCS* [11]. SCS is currently looking for developers to join the community.

### Outlook

One of the next priorities for the Monitoring SIG will be to expand log management. Kolla Ansible currently provides a `fluentd` daemon along with Elastic-Search for log shipping. Especially in terms of compliance requirements in the CSP environment (think au-

#### Info

- [1] SCS: [https://scs.community/]
- [2] OSISM: [https://osism.tech/]
- [3] Kolla Ansible: [https://docs.openstack.org/kolla-ansible/latest/]
- [4] Prometheus: [https://prometheus.io]
- [5] Alertmanager best practices: [https://github.com/SovereignCloudStack/Docs/blob/main/Operational-Docs/integrating-alertmanager.md]
- [6] Alerta: [https://alerta.io]
- [7] Awesome alert rules: [https://awesome-prometheus-alerts.grep.to]
- [8] OVN exporter: [https://github.com/greenpau/ovn\_exporter]
- [9] Keynote on APIMon: [https://www.youtube.com/watch?v=Em8TfiUIXF4]
- [10] OSISM testbed: [https://docs.osism.tech/testbed/]
- [11] Contribute to SCS: [https://scs.community/contribute/]

#### Author

Felix Kronlage-Dammers has been building open source IT infrastructure since the late 1990s. Between then and now, he was part of various open source development communities, from DarwinPorts to OpenDarwin to OpenBSD and, nowadays, the Sovereign Cloud Stack. His interests range from monitoring and observability over infrastructure as code to building and scaling communities and companies. He has been part of the extended board of the Open Source Business Alliance (OSBA) for six years and describes himself as a Unix/open source nerd. If not working, he is usually found on a road bike.





**When are Kubernetes  
and containers not a  
sensible solution?**

# Curb Complexity

As the major Linux distributors increasingly lean toward containers, many administrators have come to realize that containers are by no means a panacea for all their problems. By Martin Loschwitz

**Containers, Kubernetes**, Rancher, OpenShift, Docker, Podman – if you have not yet worked in depth with containers and Kubernetes (aka K8s), you might sometimes get the impression that you have to learn a completely new language. Container-based orchestrated fleets with Kubernetes have clearly proven that they are not a flash in the pan, but all of this omnipresent lauding of containers is getting on the nerves of many. Whichever provider portfolio you assess, it seems that the decisive factor for success is answering a simple question: How do I get my entire portfolio cloud-ready as quickly as possible?

The problems and challenges inherent in containers and K8s are too often forgotten. I address this article to all admins who still view the container hype with a healthy pinch of skepticism. In concrete terms, the question is what are the use cases in which containers do not offer a meaningful alternative to existing or new setups. When does it make more sense to avoid the technical overhead of migration because a particular case will (hardly) benefit from migrating anyway?

## Ceph as a Negative Example

If you are in an environment far removed from a greenfield, you can face some disadvantages. Ceph, the distributed object storage solution, is one example. If you believe the vendor Red Hat, containerized deployment is now the best way to roll out Ceph. All of the Ceph components come in the form of individual containers, which `cephadm` then force-fits on systems.

This arrangement is all well and good, but if you are used to working with Ceph and then try to run a simple `ceph -w` at the command-line interface (CLI) to discover the cluster status, you are in for a nasty surprise: *Command not found* is what you will see in response. Logically, if all of the Ceph components are in containers, so are Ceph's CLI tools ([Figure 1](#)). The `cephadm shell` command does let you access a virtual environment where you can run the appropriate commands, but userland software does not benefit. Programs that need the Librados API for low-level access to the RADOS service, or even depend on `/etc/ceph/ceph.conf` existing and

containing the correct addresses for the MON servers, will not work.

In other words, Red Hat forces you to plumb the depths of containerization, whether you want to or not. Of course, from the manufacturer's point of view, this scenario makes perfect sense. Red Hat only has to maintain one version of Ceph per major release to have executable programs for Red Hat Enterprise Linux in all versions and their derivatives, and even for Ubuntu or SUSE. Whenever a runtime environment for containers is available, these containers will eventually work in an identical way. However, administrators are quite right to resist being told what's best for them.

## Containers Are Not Always Needed

Up and down the IT space, vendors promote containers as the universal panacea for all problems. Applications must be cloud-ready; if you still don't rely on containers in your environment, you are – at the very least – behind the times, if not actively blocking innovation. Anyone who



```

[root@b52-41-2-47-29 ~]# podman exec -i -t ceph-mon-$(hostname -s) ceph -s
cluster:
  id: ff63bca2-a718-11ea-9a92-52540006ff8b
  health: HEALTH_OK

services:
  mon: 3 daemons, quorum b52-41-2-47-29,b52-41-2-47-27,b52-41-2-47-25 (age 3w)
  mgr: b52-41-2-47-25(active, since 10w), standbys: b52-41-2-47-27, b52-41-2-47-29
  osd: 150 osds: 150 up (since 2w), 150 in (since 3M)
  rgw: 3 daemons active (b52-41-2-47-25.rgw0, b52-41-2-47-27.rgw0, b52-41-2-47-29.rgw0)

task status:

data:
  pools: 16 pools, 6400 pgs
  objects: 7.76M objects, 24 TiB
  usage: 76 TiB used, 484 TiB / 560 TiB avail
  pgs: 6399 active+clean
      1 active+clean+scrubbing+deep

io:
  client: 12 MiB/s rd, 46 MiB/s wr, 69 op/s rd, 214 op/s wr

[root@b52-41-2-47-29 ~]# alias ceph="podman exec -i -t ceph-mon-$(hostname -s) ceph"
[root@b52-41-2-47-29 ~]# ceph health
HEALTH_OK
[root@b52-41-2-47-29 ~]#

```

**Figure 1: A containerized Ceph presents challenges to those without experience with containers, including not being able to find the familiar tools where they normally hang out.**

hasn't learned the ideal architecture of a Kubernetes cluster by rote must have been asleep for the last 25 years. One fundamental problem with the entire container debate is that the distinction is small between technical strategies that don't necessarily belong together – even if they might complement each other. To begin, you need to distinguish clearly between containerization strategies, as such, and container orchestration. Applications in containers may well have their uses outside of Kubernetes, and not just from a provider perspective; however, if all the terms and descriptions end up in one big pot and are applied arbitrarily, at the end of the day, no one knows what the debate is actually about. Therefore, this article looks at containerization and Kubernetes separately.

## Containers Are Practical

Containers are practical in many ways. Time and time again, advocates of the principle argue that the dependency hell of classic package management almost no longer exists with containers. Ultimately, each container doesn't just come with the application itself, but also with the userland, which works autonomously. This setup also makes updating easier. If in doubt, the admin simply stops

a running container, checks out the new version of the image, starts it with the old configuration and data, and the job is done.

Containers also initially offer more in terms of security than the established competitors. Under the hood, containers are ultimately no more than a combination of various Linux kernel tools that isolate processes from the rest of the system and from each other. If a vulnerable web server is running in the container and an attacker gains unauthorized access, they will be trapped in the container and not normally be able to get out. This one additional layer of security, at least, is missing in applications without a container layer.

## The Security Myth

If you take a closer look, however, the myth of secure containers quickly and decidedly crumbles. As with packages, containers ultimately depend on who their originator is and whether it is a trusted entity.

The popular package managers offered by the major distributors have long since solved this problem. Both RPM and dpkg have mechanisms to establish a chain of trust that is documented right down to the package installed on a system. To this end, most distributors sign their package

manager's distribution lists with a digital PGP key. With a signed file, you can then check whether the installed package matches the package listed there (e.g., whether the SHA256 checksums of the individual files match). This way of tracking whether any local changes have been made to the package content is both excellent and compliant.

The situation is quite different in the case of container images. The major vendors also offer ready-made base images, but these images do not typically contain any services. You have two possible courses of action: creating your own environment for continuous integration and continuous deployment (CI/CD), or grabbing ready-made images off the web.

## Complex CI/CD Structure

If you choose the first option of creating your own environment for CI/CD, you inevitably get to use tools like Jenkins or Zuul. The idea behind these tools is that, on the basis of distribution images, CI/CD systems automatically build a new image when a new software version is released and prepare it such that you only have to confirm production deployment by pushing a button. The rollout itself is handled by the CI/CD system, so your workload is kept within narrow limits, because the degree of automation in setups of this type is very high.

The bad news is that if you don't want to buy a ready-made CI/CD system for a large amount of money, you will have to spend time building one. Jenkins and the like are complex tools that can hardly be described as self-explanatory. Much time and effort goes into putting together a CI/CD system that ultimately gives you the same conditions that RPM, dpkg, and other package management systems have provided for decades.

Don't forget updates and how they are handled. You justifiably trust that running `apt dist-upgrade` on your system will apply all the security updates offered by Canonical or Debian.



```
pablo@Shinobi:~$ sudo apt-get instal hcxdumptool
E: Invalid operation instal
pablo@Shinobi:~$ sudo apt-get install hcxdumptool
Reading package lists... Done
Building dependency tree
Reading state information... Done
hcxdumptool is already the newest version (5.1.7-1).
You might want to run 'apt --fix-broken install' to correct these.
The following packages have unmet dependencies:
 windscribe-cli:i386 : Depends: openvpn:i386 but it is not installable
                       Depends: resolvconf:i386 but it is not installable
                       Recommends: stunnel4:i386 but it is not going to be installed
E: Unmet dependencies. Try 'apt --fix-broken install' with no packages (or specify a solution).
pablo@Shinobi:~$ sudo apt-get install anbox
Reading package lists... Done
Building dependency tree
Reading state information... Done
You might want to run 'apt --fix-broken install' to correct these.
The following packages have unmet dependencies:
 anbox : Depends: lxc (>= 3.0.0) but it is not going to be installed
         Depends: libboost-log1.71.0 but it is not going to be installed
         Depends: libboost-program-options1.71.0 but it is not going to be installed
         Depends: liblxc1 but it is not going to be installed
         Depends: libstdl2-image-2.0-0 (>= 2.0.2) but it is not going to be installed
 windscribe-cli:i386 : Depends: openvpn:i386 but it is not installable
                       Depends: resolvconf:i386 but it is not installable
                       Recommends: stunnel4:i386 but it is not going to be installed
E: Unmet dependencies. Try 'apt --fix-broken install' with no packages (or specify a solution).
pablo@Shinobi:~$
```

**Figure 2:** The only systems that fall victim to the notorious dependency hell today are those that use poorly maintained third-party software. © reddit.com/Pablo

However, if you rely on CI/CD, you need to track security updates yourself and replace any affected containers individually rather than collectively.

More Overhead Detours

CI/CD systems are by no means self-perpetuating, then, and in some respects, they even cause more work than legacy package managers. Quite a few admins try to avoid this work and instead resort to ready-made images off the web. Docker Hub is a popular source for such images, but anyone can upload virtually any image there – and in any form. Although ready-made images are available directly from the manufacturers for popular services such as MariaDB or MongoDB, images of lesser known software can often only be obtained from the community. In

many cases, these images do not give you an opportunity to understand how they were built; consequently, you can only guess what is hidden inside.

In the past, several Docker Hub images for various services turned out to be Bitcoin miners in disguise. Other images are uploaded only once and are not maintained by their original authors after the event, and if a security update is due for the program, you are left out in the cold. The use of these black boxes in production operation can therefore only be strongly and categorically advised against. If you use them despite the warnings, you are turning the security advantages of containers into the complete opposite and will end up with significantly less secure systems than would be the case if you used a legacy package manager.

The Myth of Dependency Hell

The dependency hell many admins want to escape by deploying containers is actually a relic of the past (Figure 2). If you only use packages from the manufacturer, from properly maintained backport directories or (often) from application vendors on your systems, you will rarely face issues at installation and update time because Red Hat, SUSE, Debian, and others have done their homework and today even support cross-version updates (e.g., from Ubuntu 20.04 to Ubuntu 22.04 at Canonical). If you run and are familiar with a defined set of services on your systems, you will reach your objectives far faster with classic packages than with a complete CI/CD environment.

Automation is a Good Thing

Automation is necessary, especially for scenarios in which classic automation by Ansible and similar tools plays a significant role. Another advantage that container advocates like to mention is that CI/CD systems offer great automation resources, and systems can be restarted quickly if something goes catastrophically wrong. After all, so the story goes, all you need to do it is roll out the container, including the configuration, on another system and start it there again. Unfortunately, this story ignores the fact that comparable setups might not have containers and instead reside on bare metal. Physical hardware, for example, can easily be managed by

FOREMAN								
centos7-devel-lobatolan@Local SSD Monitor Content Containers Hosts Configure Infrastructure Administrator								
Hosts								
Filter Search Create Host								
	Power	Name	Operating system	Environment	Model	Host group	Last report	Actions
		carla-steel.lobatolan.home	RedHat 7.3	production	toshiba.lobatolan.home	base		Edit
		centos7-devel.lobatolan.home	CentOS 7.3	production	toshiba.lobatolan.home	ansible_test	11 days ago	Edit
		green-waito.lobatolan.home	CentOS Linux 7.2.1511	production	toshiba.lobatolan.home			Edit
		irelia-cowell.lobatolan.home	RedHat 7.3	production	toshiba.lobatolan.home	base		Edit
		isa-hurta.lobatolan.home	RedHat 7.3	production	toshiba.lobatolan.home			Edit
		ivy-kate.lobatolan.home	RedHat 7.3	production	toshiba.lobatolan.home	base	15 days ago	Edit
		mobile-greenland.lobatolan.home	CentOS 7.3.1611		toshiba.lobatolan.home	ansible_test	19 days ago	Edit
		nick-mphuber.lobatolan.home	RedHat 7.3	production	toshiba.lobatolan.home	base		Edit
		ronia-riggen.lobatolan.home	RedHat 7.3	production	toshiba.lobatolan.home	test		Edit
		toshiba.lobatolan.home	Fedora release 25 (Twenty five)		Satellite Radius P52W-8	ansible_test	19 days ago	Edit
		vmi-ehp.lobatolan.home-4		production				Edit
		win-4p4bu@msk.lobatolan.home	windows 6.1.7600	production	toshiba.lobatolan.home		about 1 month ago	Edit

**Figure 3:** Life-cycle managers like Foreman help you get systems back up and running again quickly by reinstalling after a problem. Containers are not needed. © Foreman



open source software such as Foreman (Figure 3). As the administrator, you then have access to comprehensive life-cycle management features that let you reboot remotely, reinstall, and completely wipe computers. If you can, and want to, commit specifically to one manufacturer, Red Hat (Red Hat Satellite), SUSE (SUSE Manager), and Canonical (Landscape) also have boxed products on the shelf that can be rolled out in a short time. Besides, today's administrators have a full arsenal of well-developed, perfectly tested automation tools at their disposal. Whether you go for Puppet, Chef, Salt, or Ansible, all of these tools can handle simple tasks, such as installing a service and rolling out a template-based configuration file quickly – without detouring through CI/CD. If a server suffers a hardware failure, a combination of Foreman and Ansible will cleanly reinstall it, along with its required services and their configurations, in a matter of minutes and without any involvement from Docker or Podman. Of course, you have to establish a valid automation story, but even that will be faster and easier than working with Jenkins and the like, in many cases.

### More Complicated Handling

Another factor that clearly speaks against containers is something I

briefly mentioned at the beginning, but which deserves closer consideration: Containers running on systems are unfamiliar to sys admins in their everyday work and are usually more complex to handle, precisely because the administrator is dealing with the runtime environment for containers. Access to the services is by way of this environment.

Your worries start here. If you use Red Hat Enterprise Linux 8 or one of its clones, you will only be able to access the classic Docker engine by roundabout paths and unofficial sources. Not only can this become a problem in terms of security and compliance, it's something that can really get on your nerves, because Red Hat has long since said goodbye to Docker and is instead going its own way in the form of Podman, assuring you that they are completely compatible with Docker at the command line. In everyday life, however, you will quickly discover that the purported compatibility between Podman and Docker is not great in many places. If you look at the other vendors, the situation is not much better: The Community Edition of Docker for SUSE, Debian, and Ubuntu might still be available unchanged, but Canonical would prefer its own users to rely on Snap (i.e., Canonical's own format). If you manage systems from Red Hat, Debian, and Canonical,

you will have to deal with three different runtime environments, two completely different CLIs, and various compatibility problems.

### Confused Debugging

Even if you can exclusively rely on Podman, Docker, or Snap, you will quickly realize in everyday life that containers make many things more complicated. For example, if you want a service running in the container to start automatically at system startup time, you need a systemd init file. However, this file needs to interact with the container management software instead of simply starting the respective service directly.

Speaking of container management, tales of admins desperately trying to find their output on the stdout of their running containers almost sound like urban legends. In Ceph, for example, the OSDs and MONs write their own logfiles, which you will find in /var/lib/ceph/osd/ on the host systems. However, they do not contain the messages directly from the services. If you need these, you can instead tap into the running container with `docker logs -f` (Figure 4).

In defense of containers, I have to admit, many of these difficulties can be configured, but in view of all these factors, the claim that containers work better, faster, more reliably,

```
[root@b52-41-2-47-29 ~]# podman logs ceph-mon-$(hostname -s)
2020-10-12 18:12:28 /opt/ceph-container/bin/entrypoint.sh: Existing mon, trying to rejoin cluster...
2020-10-12 18:12:35 /opt/ceph-container/bin/entrypoint.sh: SUCCESS
exec: PID 86: spawning /usr/bin/ceph-mon --cluster ceph --setuser ceph --setgroup ceph --log-to-stderr=true --err-to-stderr=true
--default-log-to-file=false --foreground --mon-cluster-log-to-stderr --log-stderr-prefix=debug --default-mon-cluster-log-to-file
=false -i b52-41-2-47-29 --mon-data /var/lib/ceph/mon/ceph-b52-41-2-47-29 --public-addr 172.23.52.10
exec: Waiting 86 to quit
debug 2020-10-12 18:12:35.151 7f398cdaf3c0 0 set uid:gid to 167:167 (ceph:ceph)
debug 2020-10-12 18:12:35.151 7f398cdaf3c0 0 ceph version 14.2.8-111.el8cp (2e6029d57bc594ecea4751373da6505028c2650) nautilus (
stable), process ceph-mon, pid 86
debug 2020-10-12 18:12:35.151 7f398cdaf3c0 0 pidfile_write: ignore empty --pid-file
debug 2020-10-12 18:12:35.182 7f398cdaf3c0 0 load: jerasure load: lrc load: isa
debug 2020-10-12 18:12:35.182 7f398cdaf3c0 0 set rocksdb option compression = kNoCompression
debug 2020-10-12 18:12:35.182 7f398cdaf3c0 0 set rocksdb option level_compaction_dynamic_level_bytes = true
debug
debug 2020-10-12 18:12:35.182 7f398cdaf3c0 0 set rocksdb option compression = kNoCompression
debug 2020-10-12 18:12:35.182 7f398cdaf3c0 0 set rocksdb option level_compaction_dynamic_level_bytes = true
debug 2020-10-12 18:12:35.182 7f398cdaf3c0 0 set rocksdb option write_buffer_size = 33554432
debug 2020-10-12 18:12:35.183 7f398cdaf3c0 1 rocksdb: do_open column families: [default]
debug 2020-10-12 18:12:35.183 7f398cdaf3c0 4 rocksdb: RocksDB version: 6.1.2
debug
debug 2020-10-12 18:12:35.183 7f398cdaf3c0 4 rocksdb: Git sha rocksdb_build_git_sha:@@
debug 2020-10-12 18:12:35.183 7f398cdaf3c0 4 rocksdb: Compile date Sep 21 2020
debug 2020-10-12 18:12:35.183 7f398cdaf3c0 4 rocksdb: DB SUMMARY
```

**Figure 4:** If you want to see the error messages generated by containerized Ceph services on stdout, you need to dock directly with the container, instead of accessing a normal logfile.



and more securely than packages out of the box is something I very much doubt.

If you also look at the way services running in containers are connected to the local network, more doubts raise their ugly heads. Because containers are designed out of the box not to be bound to the IP addresses of the host system, Docker and others use port forwarding as a ploy. The containers then open a virtual network on the host, to which the kernel forwards packets for the target IP. Conversely, then, debugging with tools like `tcpdump` is more complicated than in conventional environments.

## Interim Conclusions

No matter how you spin it, containers do have their advantages, but they require the administrator to navigate a huge learning curve, while not solving certain problems as elegantly as can a combination of classic automation and package management. If you have a clear idea of the services you want to run, and on what systems, you will fare better and enjoy more stability from a combination of Ansible, configuration templates, and classic packages from the providers. Most importantly, you will have less complexity than Docker, Podman, and the other container solutions.

## Kubernetes: Yes, but ...

In the second and significantly shorter part of this article, I really need to mention Kubernetes, the software that has undoubtedly given the container hype triggered by Docker a massive boost. Admins love modern technology, and once a solution becomes popular, it attracts admins like moths to the flame, as was the case with Kubernetes. For years, it has been the talk of the town. That said, many Kubernetes-based solutions presented today completely lose sight of what the software is all about and what problem it looks to solve: Kubernetes comes from the world of

cloud computing and makes great sense, especially against this background.

A few years ago, Docker had just made containers on Linux respectable after solutions like OpenVZ had failed to do so for decades. The primary reason Docker was successful is that it offered not only the runtime environment for containers on Linux but also created a suitable ecosystem, which undoubtedly included the aforementioned Docker Hub.

At the time, however, cloud computing had long been seen as the means of choice for companies that wanted to move from being IT heroes to major platform providers and at least play in a league similar to Amazon. Containers promised huge dividends because both orchestration and automation combined with prepackaged applications are effectively the highway to platform-as-a-service (PaaS) and software-as-a-service (SaaS) offerings. What was still missing was a solution that could manage and control containers across any number of physical hosts in a completely automated way. Kubernetes quickly established itself as a solution for precisely this task, if only because the product, as an in-house development at Google (up to that point), was explicitly designed for that purpose.

## Over-the-Top Hype

The vast majority of features that K8s has seen added in recent years, or that can be retrofitted with external tools, are for the PaaS and SaaS deployment areas already described. Kubernetes is far less well suited for classic information as a service (IaaS), for example. This niche is where classic virtual machines (VMs) based on KVM or Xen play to their strengths.

In some ways, this insufficiency is the crux of the problem, although it has never stopped many Kubernetes service providers from selling K8s as a solution to life, the universe, and everything. Local setups, so the story goes, also benefit from Kubernetes

because it becomes easier to handle applications. The fact that this only applies if the applications are suitable for operation in containers at all and, even then, that many conventional applications present their own challenges, is deliberately concealed by the advertising blurb. Most of the problems already described for containers without K8s also apply to containers in Kubernetes. Granted, some companies are using the move to containers and Kubernetes to make a clean sweep of their existing application landscape, and even rewriting some parts of it, but this is by no means possible or affordable in every case.

Kubernetes also is designed in the best possible way to scale seamlessly across the board (i.e., to provide the underpinnings of an eternally growing platform). However, you only really need this if you want to operate a platform that can handle any amount of data and services. If you have a clearly separated setup in mind with a defined target size (with some scope for fluctuation), you will be hard put to use most of Kubernetes' functionality in a meaningful way. In these cases, K8s quickly becomes a hyper-complex chunk at the end of the process, the complexity of which bears no relation to the actual needs and benefits.

## Conclusions

Kubernetes is primarily of interest to enterprises that want to run arbitrary numbers of containers for SaaS or PaaS applications across large fleets of physical systems. Kubernetes can also be used for internal purposes. For example, if you are converting your production software to as-a-service principles, Kubernetes will serve you well internally. In this respect, operating a K8s platform is by no means tied to being a major platform provider.

If you're looking for a solution that rolls out individual applications in a neatly automated and scalable way, though, you'll have an easier life with a solution that is based on classic



automation, just as you did with containers. If required, the automation can be combined with virtualization: The combination of VMware or KVM and Ansible is found in many places. Virtualization solutions such as Proxmox (Figure 5), which can be easily integrated with existing automation, are also helpful. Almost all of these

pairings are far easier to handle and maintain than K8s, and accordingly offer less functionality. For many companies embarking on the Kubernetes adventure without a clear roadmap, the functionality offered by classic automation will be completely sufficient: Containers and container automation are not

just the continuation of virtualization by other means. ■

The Author

Freelance journalist Martin Gerhard Loschwitz focuses primarily on topics such as OpenStack, Kubernetes, and Chef.

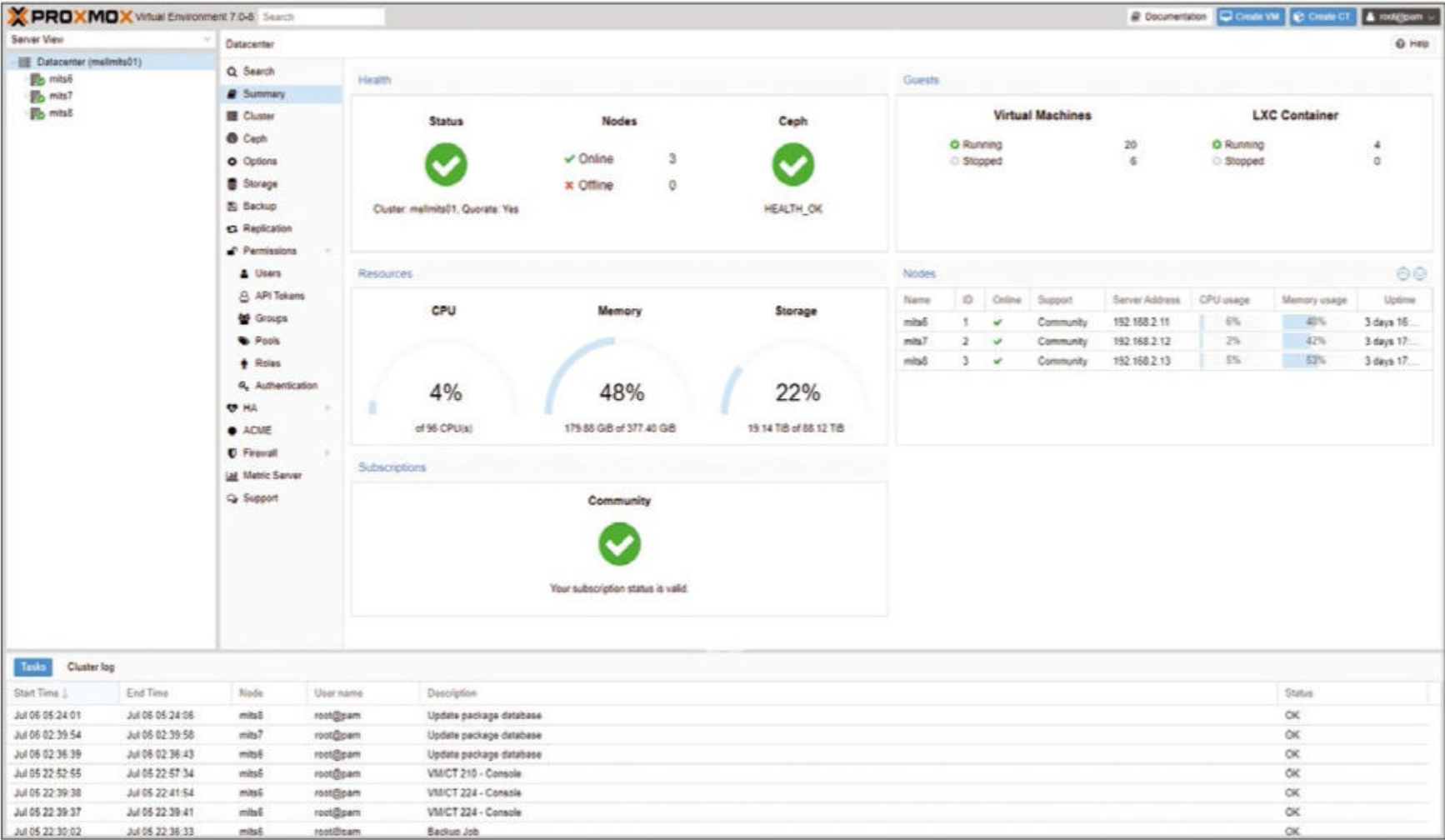



Figure 5: Like other virtualization solutions, Proxmox Virtual Environment can easily team up with many automation tools, ensuring a kind of scalability. However, the entire setup is far less complex than a full-blown K8s environment. © Proxmox

■





Versioned backups of local drives with Git

# Genealogy

Versioning is a recommended approach to back up files as protection against hardware failures and user errors. To create versioned backups, you can use established backup programs or an open source tool that originates from the developer world: Git. By Andreas Stolzenberger

**Despite cloud and file servers** on the corporate network, users still store important files on the local hard drives of their workstations or laptops. Modern solid-state drives (SSDs) lull users into a deceptive sense of security: Thanks to technologies such as self-monitoring analysis and reporting technology (S.M.A.R.T.) and wear leveling, these data storage devices can predict their demise and usually warn the user in time before a disk failure. However, valuable data is rarely lost by spontaneous disk failure. More often, the cause of data loss is the users themselves accidentally deleting or overwriting files. If you travel with your laptop, you also have to worry about losing the device or damaging it irreparably.

A device and its operating system and applications can be replaced quickly, but it's a different story for user files. Therefore, every user with important files on their local computer needs a viable backup and restore strategy that covers the following functions:

- Multilevel file versioning
- Local backup
- Remote backup over local area network (LAN)

- Optional remote backup over wide area network (WAN)

- Backup and recovery independent of the operating system

On the free market, all common operating systems have tons of backup programs – many with inexplicably confusing user interfaces. For most users, the backup chain ends at the USB drive, but if you want to back up your data, it is better to use a network share. Common cloud backup tools, on the other hand, back up directly to the connected cloud and therefore only work if you are online.

## Backup with Git

The Git [1] tool supports code versioning, shows the details of the differences between saved versions, and allows multiple developers to work together on a project. Because it works online and offline, this popular open source application is more or less a perfect backup program; moreover, it is available for all common operating systems. The backup format is openly documented – independent of the operating system

and filesystem – and not a proprietary file format.

Anyone who creates a Git repository first creates an object store with copies of the selected data. This object store keeps track of the changes in the files saved in it. As soon as the user makes a “commit” (i.e., a snapshot of the files), Git remembers the changes to the previous version. Users can roll back the entire repository or individual files to previous versions if they want and recover accidentally deleted information from a previous snapshot.

However, Git does not just back up its repository on the local system. Users can create one or more remote copies of the repository and upload the versions there. Luckily, remote repositories do not need to receive every single commit immediately. A user can write many consecutive commits to the local repository and then send an update to the remote repository. Git's delta mechanism will submit all missed commits from the local repository so that the remote repository ends up containing all intermediate steps.

Photo by Roman Kraft on Unsplash



Git is also very flexible when it comes to restoring. You can conveniently access previous versions of individual files or entire directories at the command line or in one of the many available Git user interfaces (UIs). When doing so, you do not need to overwrite the existing version but can save an earlier version under a different name. You also have a wide choice of Git servers and, here too, you can view, modify, or download individual files on the web UI, if required.

## Git on Windows

The Git homepage offers a setup for Git on Windows [2]. In addition to the plain vanilla command-line (CLI) tool for the Windows prompt and PowerShell, Git also provides the necessary OpenSSL libraries and tools and a MinGW environment, the native Windows port of the GNU Compiler Collection (GCC), that lets Windows administrators run Git from the command line or in PowerShell, as well as in Git Bash.

Windows also has a number of graphical UI (GUI) tools for Git. The popular GitHub Desktop is primarily aimed at application developers; however, this tool does not give you a very good overview, especially

for large repositories. Although it is not open source, Atlassian provides its Git client Sourcetree [3] free of charge (Figure 1). The client can be used for both development and backup repositories. Git Extensions [4] offers a somewhat older, technically overloaded look and feel (Figure 2). In return, this free client is extremely fast and includes many features.

If you use the Eclipse development environment, the JGit tool, written in Java, is included. Unlike the conventional Git client, JGit also handles the Amazon Simple Storage Service (S3) protocol. If you want to store your repository somewhere instead of, or in addition to, the regular Git server, you can use JGit to move directly into an S3-compatible object store.

## Distributing Local Git

For version control, Git creates the versioned repository in the .git subdirectory within the folder to be backed up. However, this does not work for the backup scenario here. The key to success is the --separate-git-dir parameter when creating the repository, which tells Git not to create the object store for versioning in the directory, but on a different path. In this case, .git

is not the directory with the files, but a text file with the link to the external directory (i.e., it acts independently from the operating system and filesystem).

In this example, you will be creating a Git repository inside the regular user directory for documents and using a USB drive with driver letter H: for the object store:

```
mkdir h:\git_backup
cd %HOMEPATH%\Documents
git init --separate-git-dir=h:\git_backup
git add -A
```

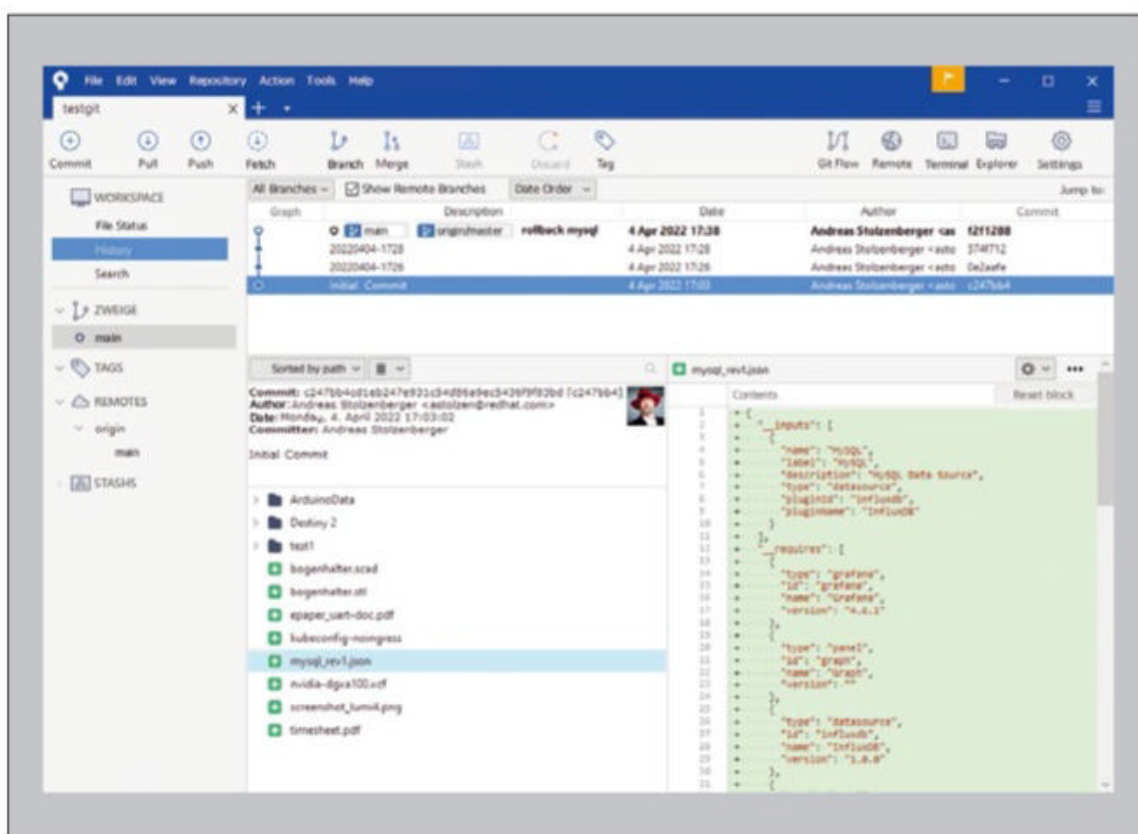
If you have not used Git on the system before, the tool will ask you for global variables such as your username and email address; then, it creates the object directory on the USB drive. Depending on the volume of data in the document directory, this step can take some time. In this setup, Git writes its data at a rate of about 1.2GB per minute. The speed also depends on whether Git needs to write small or large files. During the write process you will see a number of warnings, such as:

```
warning: LF will be replaced by CRLF
in <filename>
```

This message refers to the old dilemma that \*nix systems define the end of line (LF, line feed) in a text file differently from Windows (CRLF, carriage return-line feed). Because Git works independent of the operating system, it always saves text files within the object memory in \*nix format but leaves the original file unchanged. This automatic conversion should not bother you unless you are restoring a text file stored in Git without the Git tool. For example, if you download a text file from your repository over HTTP from a Git server with a web UI, no CRLF reconversion takes place.

Now that Git has copied all the data to the local repository, you can create a snapshot of the current state with a commit:

```
git commit -m "Repository created"
```



**Figure 1:** Atlassian's free tool Sourcetree provides insights into local and remote Git snapshots.



Now, if you make changes to files, Git tracks them and saves them for the next commit. Each commit is given a unique ID and, for clarity,

Listing 1: git restore

```
type mysql_rev1.json
Unfortunately overwritten

git log mysql_rev1.json
commit 780...
Author: Andreas ....

    2200404-1300
...

git restore --source 780... mysql_rev1.json

type mysql_rev1.json
{
  "__inputs": [
    {
      "name": "MySQL",
      "label": "MySQL",
      "description": "MySQL Data Source",
      ...
```

a name that you pass in with the `-m` parameter – in this case *Repository created*. Later, you can see in the Git history exactly which files changed in which commit, but if you add new files to the directory, Git does not automatically include them in the repository. You need to run a `git add -A` again before committing. To automate the process, create a suitable batch file:

```
set commitname=%date:~-4%-%date:~-7,2%-%date:~-10,2%-%time:~-11,2%-%time:~-8,2%
set gitdir=%HOMEPATH%\Documents
cd %gitdir%
git add -A
git commit -m "%commitname%"
```

Now you can create a shortcut on the desktop and trigger the backup at the push of a button. Alternatively, create

an automatic task that performs the backup regularly (e.g., every hour), but keep in mind that Git, like any other program, cannot include open files in the snapshot. Of course, the script can be prettified – for example, to check first whether the USB target disk is connected to the system before triggering the backup and initiating an upload to the server once a day. Now, to restore a single file to a previous state after an accidental change, use `git restore` as in [Listing 1](#).

Git Server Selection

One of the most popular self-hosted Git servers is GitLab [\[5\]](#), which of course also runs on the service at *gitlab.com*. However, the massive GitLab, written in Ruby, requires quite a bit of performance on the part of the server hardware. In return, GitLab delivers a wide range of functions, such as a wiki, a bug tracker, and an integrated container image repository. If you only

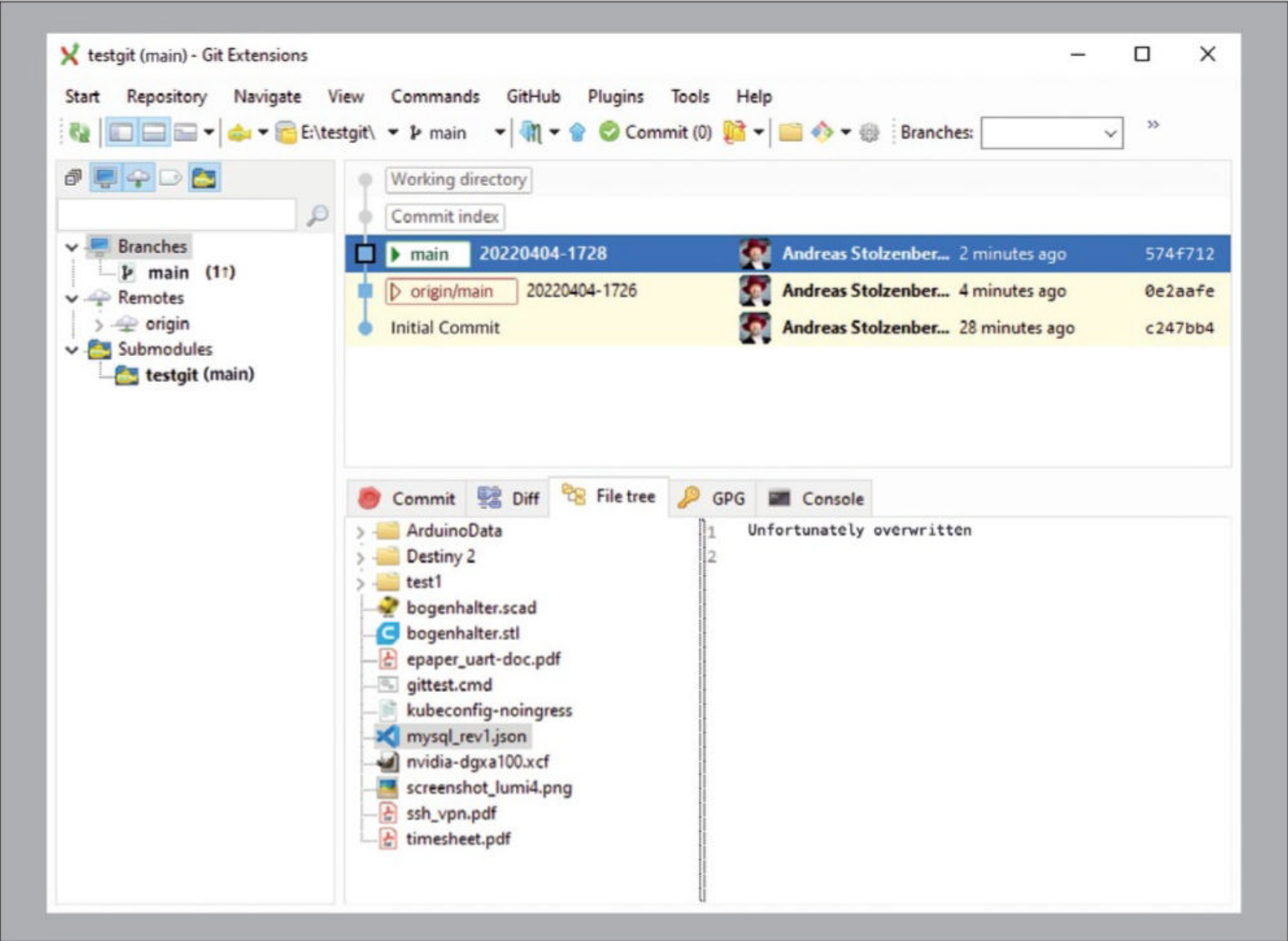


Figure 2: In the Git Extensions open source tool, you can see the complete history of commits.

use a Git server as a backup target, you don't need all of these features and are better off with a simple, but agile, Git server like Gitea [6], which handles many databases and runs in containers with low resource requirements (Figure 3). For administrators with a predominantly Windows background, the free Bonobo [7] Git Server in .NET integrates directly with Internet Information Services (IIS). To create a second, remote backup of a local repository, you need a Git server. Here, too, you have a whole range of open source offerings. Gitea is the simplest solution for newcomers. Written in Go, it derives from Google's

in-house Git server Gogs but is maintained by a somewhat more liberal developer community than the Google original. Like Git itself, Gitea is available for all major platforms, so the server service also runs smoothly on Windows. All you need is an existing Git installation and a database. Gitea supports MySQL and PostgreSQL as well as Microsoft SQL Server (MSSQL) or, for simple test setups, SQLite. Of course, Gitea can also be run in a Podman or Docker container. Once started, Gitea listens on ports 3000 (HTTP) and 22 (SSH) by default. In the simple and clear-cut web UI, you first need to create the user

accounts. To avoid the need to log in with a username and password, you will want to use SSH keys to authenticate. The key pairs usually are generated with a centralized key management system and then assigned to the users. In smaller environments, however, users on the client PCs can create key pairs themselves with ssh-keygen. Remember always to keep a copy of the SSH keys outside the client PC. With the remote server running, you first log in to the Gitea UI with your user account, where you store the public SSH key, if this has not already been done elsewhere. In your

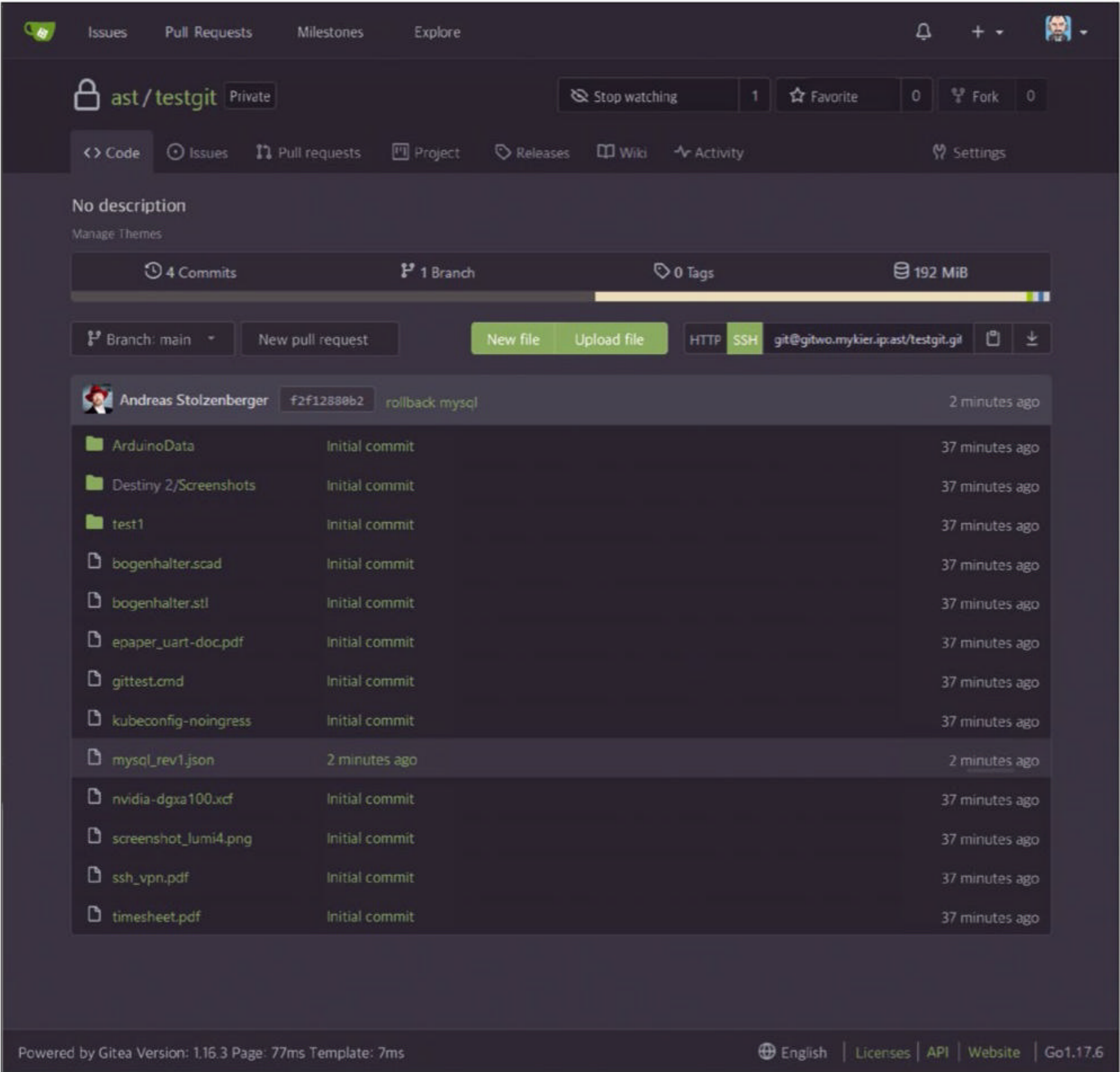


Figure 3: The simple Gitea Git server functions as a central backup server in the LAN or WAN.



account, you then create an empty repository for the backup data and mark it as a *Private repository*. Gitea returns the appropriate SSH URL. Now you can add the remote storage path *origin* to the existing local repository,

```
git remote add origin ➤
git@<Server>:<User/Repository>.git
```

and copy the local repository, including all previous commits, to the remote server:

```
git push -u origin main
```

The `-u` lets you define the remote server *origin* as the default upstream server for the repository. Future changes can then be sent to the *origin* server with the `git push` command, without any additional parameters. Git lets you specify multiple remote repositories and use

```
git push <name>
```

to send the updates to different servers. The Git server can run on the company LAN as well as on a cloud system or in the demilitarized zone (DMZ) with an Internet connection. Because communication between the client and server relies on SSH encryption anyway, the cloud backup

to the in-house Git server is secure even without a virtual private network (VPN). Either the web interface of the Gitea server on port 3000 should be protected by a reverse proxy/load balancer with HTTPS termination, or the service itself should run on HTTPS with a suitable certificate. To do this, modify the configuration of Gitea in `./gitea/conf/app.ini` and expand the following section:

```
[server]
PROTOCOL = https
ROOT_URL = https://<URL>:<Port>
HTTP_PORT = <Port>
CERT_FILE = cert.pem
KEY_FILE = key.pem
```

Hosted Git services such as GitHub or GitLab are obviously out of the question as backup targets because they limit the size of the repositories.

## Git in Git

Once you have familiarized yourself with the Git tool, you can optimize its use. For example, if you are just working on a project for a certain period of time, you can back up all related data to a separate directory and therefore to a separate Git. If the folder is inside an existing repository, Git notices and excludes the project

directory as its own repository from the underlying Git.

The user then backs up two repositories: The document directory itself and the project. Thanks to Git's group features, project data can be shared between workgroups and checked out to multiple clients. When the project is completed, the clients can delete the local project directory, and the Git server keeps all the data as an archive that is available to users at any time.

## Conclusions

Backup and version control are closely related topics, so Git is a very good choice as a backup tool for user data. Of course, Git does not back up the operating system or the installed applications or their configuration in the registry. ■

### Info

**[1]** Git: [\[https://git-scm.com\]](https://git-scm.com)

**[2]** Git for Windows: [\[https://git-scm.com/download/win\]](https://git-scm.com/download/win)

**[3]** Sourcetree: [\[https://www.sourcetreeapp.com\]](https://www.sourcetreeapp.com)

**[4]** Git Extensions: [\[http://gitextensions.github.io\]](http://gitextensions.github.io)

**[5]** Download from GitLab: [\[https://about.gitlab.com/install/\]](https://about.gitlab.com/install/)

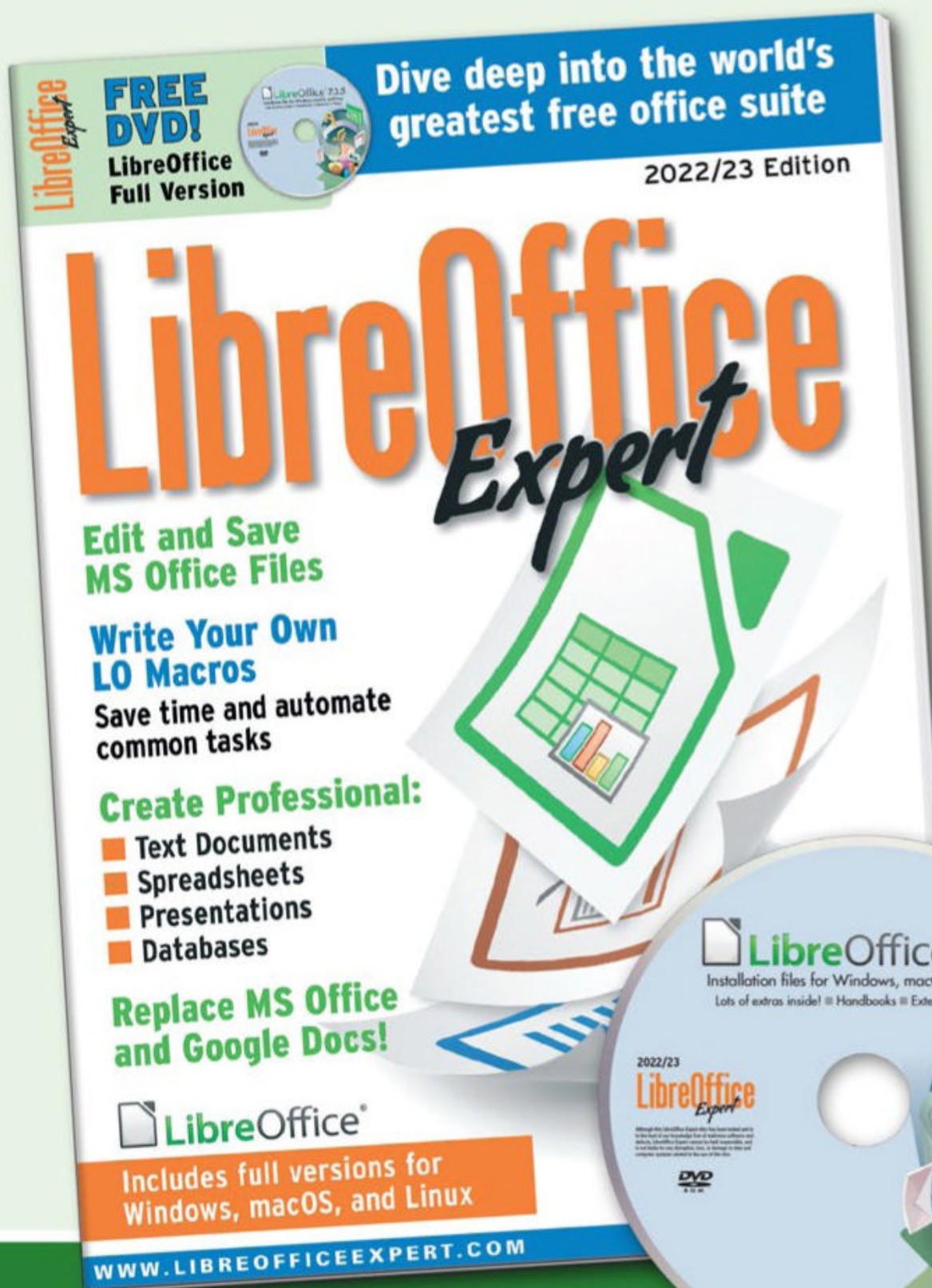
**[6]** Gitea: [\[https://gitea.io/en-us/\]](https://gitea.io/en-us/)

**[7]** Bonobo: [\[https://bonobogitserver.com\]](https://bonobogitserver.com)



Shop the Shop  
[shop.linuxnewmedia.com](http://shop.linuxnewmedia.com)

# Become a LibreOffice Expert



Explore the **FREE** office suite used by busy professionals around the world!

**Create Professional:**

- Text Documents
- Spreadsheets
- Presentations
- Databases

Whether you work on a Windows PC, a Mac, or a Linux system, you have all you need to get started with LibreOffice today. This single-volume special edition will serve as your guide!

Order online:  
[shop.linuxnewmedia.com/specials](http://shop.linuxnewmedia.com/specials)

**For Windows, macOS, and Linux users!**





SoftEther VPN software

# Speed in the Tunnel

SoftEther is lean VPN software that outpaces the current king of the hill, OpenVPN, in terms of technology and performance. By Holger Reibold

**In the age of home offices** and distributed locations, companies want security solutions that can be integrated easily into existing infrastructures, offer genuine added value, and secure communications between mobile clients and sites. OpenVPN was long considered the measure of all things VPN, but the business model was developed at the expense of the community edition and has various restrictions. For example, the classic tool only supports its own virtual private network (VPN) protocol and does not offer support for natively integrated VPN clients from Android, iOS, macOS, and Windows.

## Fast SoftEther Alternative

SoftEther [1], whose name contains elements of software and Ethernet, has something to offer to counter the aforementioned limitations. The open source VPN supports VPN protocols such as Secure Socket Layer (SSL) VPN, the Layer 2 Tunneling Protocol (L2TP)/Internet Protocol Security (IPsec), OpenVPN, and Microsoft Secure Socket Tunneling Protocol (SSTP). SoftEther supports network address translation (NAT) traversal, which means you can run the VPN server on a machine located behind

home gateways, facility routers, and firewalls. Firewalls that perform deep packet inspection do not recognize SoftEther's VPN transport packets as VPN tunnels because HTTPS is used to disguise the connection. Other highlights include:

- Site-to-site and remote access VPN connections
- Access to restricted public wireless local-area networks (WLANs) by VPN over Internet Control Message Protocol (ICMP) and VPN over DNS
- Ethernet bridging and Layer 3 over VPN
- Logging and firewalling in the VPN tunnel
- Support for relevant operating systems (Windows, Linux, macOS, Android, iOS)
- Cloning OpenVPN connections
- RADIUS/NT domain authentication of users

A table showing the direct comparison between the two VPN applications can be found online [2].

Apart from purely technological aspects, the biggest benefit is fast speed. On the basis of performance protocols, various studies show that OpenVPN has a data throughput of less than 100Mbps, which often turns out to be a bottleneck. According to the developers, SoftEther provides

speeds of more than 900Mbps. Performance is achieved through utilization of the full Ethernet frame. At the same time, the software reduces memory copying, parallel transfer, and clustering. The sum of these measures significantly reduces latency and massively boosts throughput. Another special feature of SoftEther, its modular architecture, lets you expand the basic system to include additional functions, such as VPN Gate. Thanks to versatile protocol support, you usually don't have to install the SoftEther client. However, if you make intensive use of the VPN environment, you will want to use the client for performance reasons alone.

## SoftEther in Operation

The name "SoftEther" points to the architecture of the VPN software: Virtualization creates a virtual Ethernet adapter and generates a switch that emulates a conventional Ethernet switch – in SoftEther terminology, it is referred to as a virtual hub. The VPN connection is established by the two components working together over a virtual Ethernet cable. The SoftEther version at the time of publication was 4.39. The installation packages for the supported operating systems are available from the SoftEther download center [3]. The SoftEther server is at the core of the environment. It is especially easy to install

Photo by Marc Sendra Martorell on Unsplash



on Windows, where the setup wizard guides you through the various steps. During the install, you can choose between the server and the VPN bridge. Server Manager is automatically installed and supports centralized administration of various SoftEther installations. To configure the local SoftEther infrastructure, you just connect to the VPN server with a single click and specify a server-specific password. The ease of working with SoftEther soon becomes apparent (**Figure 1**). In the *SoftEther VPN Server/Bridge Easy Setup* dialog, you can choose from the two most common installation variants: *Remote Access VPN Server* or *Site-to-Site VPN Server to VPN Bridge*. If you want to use advanced configurations such as clustering, you need to do so manually. The setup dialog supports you in the decision-making process, in that it provides a visualization and brief description of the scenario in question with information relevant to decision making. In this article, I look at the remote access scenario, where the VPN clients establish a secure connection to the VPN server, allowing access to the network behind it. Pressing *Next* in the setup wizard opens the hint dialog that informs you the server is initialized. You only have to confirm at this point and proceed to assign a name to the server configuration.

The setup wizard now reveals another special feature. The VPN server has a built-in dynamic DNS feature that assigns a permanent DNS name to the server, making it globally accessible. Pressing *Exit* completes the basic configuration and you can carry on with the next step. You can integrate the local VPN server with the Azure Cloud Service (not to be confused with Microsoft's cloud service). VPN Azure Cloud is a free cloud VPN service from the SoftEther project. In the context of this example, it is difficult to say whether or not it makes sense to use it. My recommendation is to disable the associated function by selecting *Disable VPN Azure*. Then press *OK* to close this dialog.

The next step is to set up the first SoftEther user by clicking on *Create Users* and assigning a username, real

name, password, and authentication method. If you are running a RADIUS server or Active Directory, specify the associated username. You can also use existing certificates. SoftEther is now ready for use.

## Installing the Client

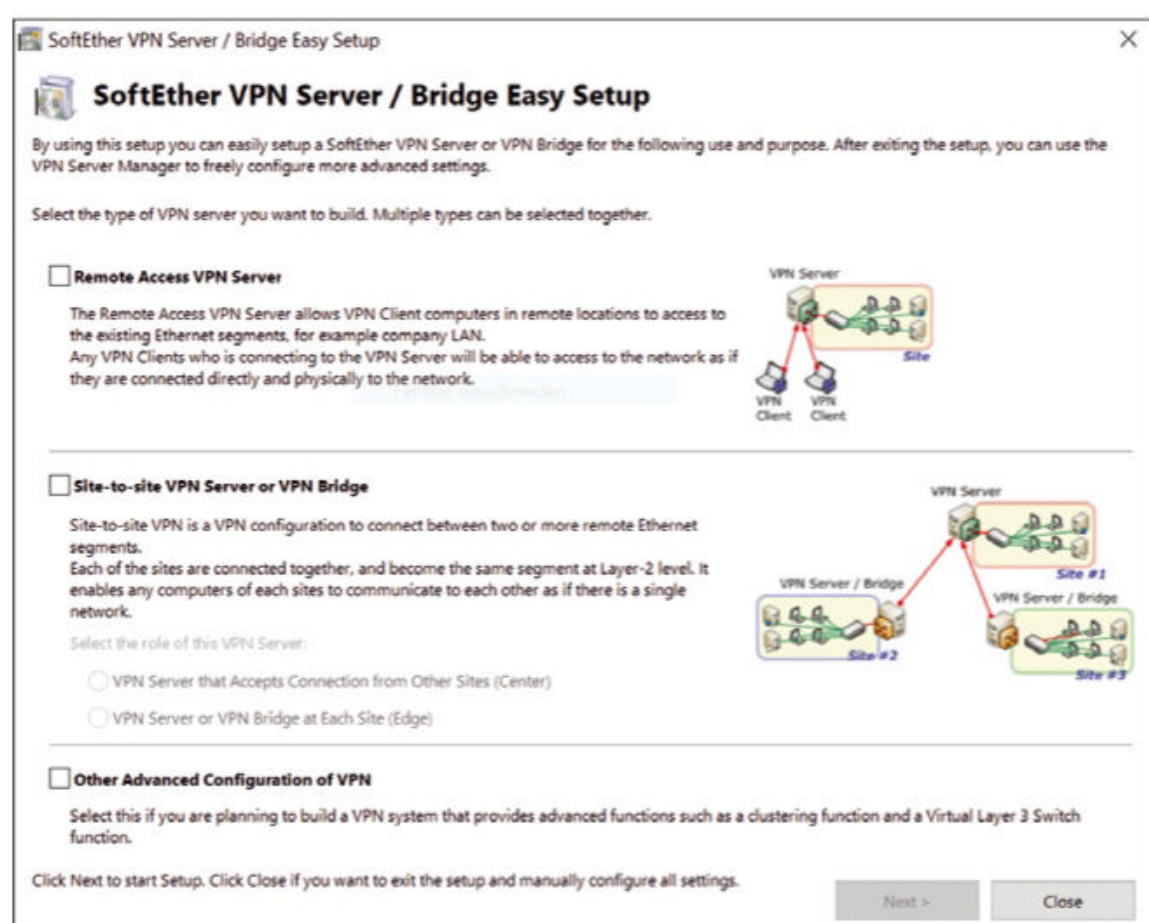
In principle, SoftEther lets you use the native VPN clients of today's popular operating systems, but the SoftEther client is recommended for regular use of the VPN environment. The installation packages are also available from the download center. The installer sets up a virtual network adapter and a background service on the client side. As with the SoftEther server, the client provides a VPN Client Manager that you use to manage various VPN connections and connection-specific settings. To connect to the SoftEther server, double-click on *Add VPN Connection* in the Client Manager and assign a name, the server data, the virtual network adapter to be used, and the access data to the connection in the associated *New VPN Connection Setting Properties* dialog. If required, you can also use a proxy server.

One highlight on the client side is the advanced connection settings, which

are hidden behind the *Advanced Settings* button. You can significantly improve performance by increasing the number of connections under *Number of TCP Connections* (**Figure 2**). For broadband connections the value can handle up to 8 parallel connections; for dial-up connections you will want to keep the default value 1.

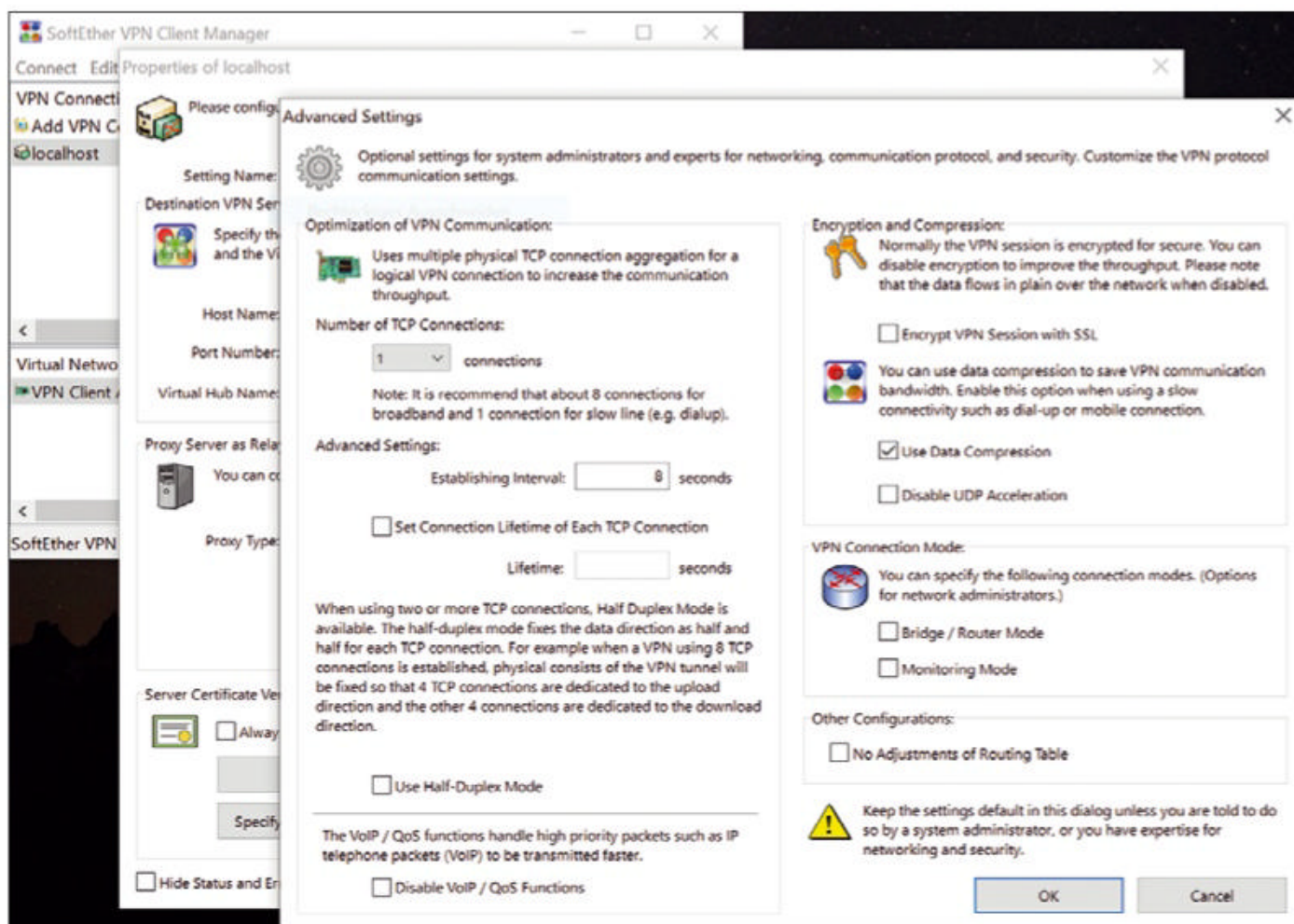
If sufficiently powerful systems are available on the client and server side, enabling data compression by checking *Use Data Compression* can provide a significant performance boost. The SoftEther VPN protocol can internally compress and transmit all Ethernet frames sent and received. The Deflate algorithm is used here. Data compression reduces the communication volume by up to 80 percent, but compression and decompression do cause a significant increase in the compute load. Compression has its limits, however: If the line speed exceeds 10Mbps, not compressing data often improves communication speed. A final *OK* accepts the connection settings, and you can open a VPN to the SoftEther server in the Client Manager just by double-clicking.

The Client Manager offers a wealth of other practical features. For example, you can export and import connection



**Figure 1:** The user-friendly SoftEther server setup wizard has information on typical deployment scenarios.





**Figure 2:** The SoftEther Server Manager is the central interface to the VPN server instances.

settings for use on third-party systems and create additional virtual adapter and smart card configurations. The Client Manager tags the existing VPN connections as *Connected* in the *Status* column. You can access the connection details by right-clicking on the connection settings and selecting the *View Status* command.

## Managing SoftEther Environments

SoftEther server administrative operations are divided into two types: Server management and administration of virtual hubs in a VPN server configuration. The developers placed great emphasis on decoupling the VPN server process from configuration tweaks when designing the environments, which ensures that VPN functionality is available without interruptions.

You only need to restart SoftEther if the operating system requires you to do so, when the VPN server program is updated, when the server process is restarted because of a hardware or software error, or when you make manual changes to the VPN server configuration file.

Changes to the clustering configuration are the only tasks that require stopping the VPN service.

For all administrative tasks, you can use the Server Manager or the `vpncmd` console tool. SoftEther supports two types of administrative permissions: management authority for the entire VPN server and for virtual hubs. The server administrator should be identical to the server computer admin. In particular, the admin is responsible for managing the certificates and ports. Furthermore, they have full access to the `vpn_server.config` SoftEther configuration file. In addition to server configuration, this file also contains the encrypted admin password and the private key of the connection setup certificate, so it requires special protection. In the Windows installation, only members of the Administrator and SYSTEM groups have read and write permissions; the same applies on Unix-based systems.

## Administrative Tasks

VPN Server Manager is the central interface for typical administrative tasks. Pressing the *About this VPN Server* button tells SoftEther to show

you a tabular overview with server information that specifically includes the server type (usually *Standalone*), the operating system, a plethora of general server data, and supported services and functions. For an overview of the current state of the server, press the *View Server Status* button; alternatively, use the `ServerStatusGet` command at the command line. The overview shows the number of active sockets, virtual hubs, and sessions. It also tells you how

many users and groups were active and what data volumes were transferred. To view the list of current connections, click the *Show List of TCP/IP Connections* button in the Server Manager. Here, you can find out which clients have opened a VPN connection to the SoftEther server and when the connection was initialized.

In practice, it is always useful to be able to take a look at the current configuration file (e.g., to verify whether any configuration adjustments you have made have been implemented in the central system configuration). To view the config file, press *Edit Config*. The dialog box comes up with the text file but does not allow editing. That said, you can restore the factory settings, save the configuration file, or load and enable an alternative file. The OpenVPN clone server feature is useful for organizations that are still using OpenVPN but want to migrate after evaluating SoftEther. You can use it to connect all OpenVPN clients, including iPhone and Android clients, to the SoftEther VPN server with minimal overhead. The settings are hidden behind the *OpenVPN/MS-SSTP Setting* button. Proceeding is very easy: Just enable the clone function and specify



the UDP ports that OpenVPN uses. SoftEther takes care of everything else. You can then disconnect the OpenVPN server from the network, but make sure that cloning works reliably with random tests beforehand.

## SoftEther as a LAN Tester

SoftEther can also be used to test and simulate network configurations. For example, SoftEther has a delay, jitter, and packet loss generator that lets you simulate a network or network segment in poor condition. To begin, define two local bridges from a virtual hub to two physical Ethernet network adapters. The two Ethernet network segments are bridged by the virtual hub, which introduces delay, jitter, and packet loss as it forwards Ethernet frames. The generator is particularly suitable for testing VoIP devices. To put such an Ethernet-based network topology through its paces,

use the VPN server, a client, and a bridge. On the SoftEther server, create multiple separate Ethernet segments; the VPN server relies on a virtual Layer 3 switch function that provides IP-based Layer 3 routing between L2 segments. An access control list (ACL) function for packet filtering is available for the virtual hub configuration. You can use these L2 and L3 functions to test the network design. The SoftEther software has a variety of other features that are of interest for enterprise use. For example, you can use the environment to implement a virtual network with connectivity to well-known cloud services. It is also possible to set up your own VPN-secured cloud.

## Conclusions

SoftEther is excellent VPN software that makes many established tools look old fashioned. In particular,

versatile support for VPN protocols, the above-average performance, and great convenience for users make SoftEther an attractive VPN alternative. Despite all the enthusiasm, certain limitations exist. Unlike OpenVPN, SoftEther cannot operate as software as a service (SaaS), and it tends to target small to medium-sized environments. It also lacks a kill switch function and policy management. Another shortcoming is that the developers do not provide commercial support to date. That said, it would be a good thing for SoftEther to step out of the shadow of OpenVPN and the like and find many new friends in the admin community. ■

### Info

- [1] SoftEther: [\[https://www.softether.org\]](https://www.softether.org)
- [2] SoftEther vs. OpenVPN: [\[https://www.vpnxd.com/softether-vs-openvpn-which-one-better/\]](https://www.vpnxd.com/softether-vs-openvpn-which-one-better/)
- [3] SoftEther download center: [\[https://www.softether-download.com/en.aspx\]](https://www.softether-download.com/en.aspx)

# IT Highlights at a Glance



Too busy to wade through press releases and chatty tech news sites? Let us deliver the most relevant news, technical articles, and tool tips – straight to your Inbox.

Linux Update • ADMIN Update • ADMIN HPC

Keep your finger on the pulse of the IT industry.

ADMIN and HPC: [bit.ly/HPC-ADMIN-Update](https://bit.ly/HPC-ADMIN-Update)

Linux Update: [bit.ly/Linux-Update](https://bit.ly/Linux-Update)





Windows Subsystem for Linux and Android in Windows 11

# Good Host

The Windows subsystem for Linux runs graphical applications on Windows 11 out of the box, and Windows Subsystem for Android, out in preview at press time, shows that Windows can be a platform for Android apps. We look at the state of the art and help you get started. By Christian Knermann

**High-ranking Microsoft employees** publicly making disparaging remarks about Linux is ancient history. Redmond has changed its course 180 degrees and impressively demonstrated that its commitment to the development of open source software is not mere lip service. For example, GitHub, the platform for collaborative version management for countless free projects, was acquired by Microsoft in 2018. However, Microsoft does not just operate the platform, it also actively contributes to many projects as one of the world's largest open source providers.

Additionally, Microsoft has developed its own Linux distribution named CBL-Mariner – where CBL stands for Common Base Linux – which you will find on GitHub, of course [1]. However, CBL-Mariner is

by no means designed to compete with established Linux distributions for general-purpose use. Microsoft mainly uses its in-house distribution itself, following an approach that is most comparable to Fedora CoreOS. CBL-Mariner, as a minimalist distribution, is primarily geared toward running containers. It forms the basis for the Azure Kubernetes Services and other services in the Azure cloud. However, CBL-Mariner is not only relevant for Microsoft's cloud activities, it now also handles tasks under the hood of the Windows Subsystem for Linux (WSL).

## Graphical Linux Apps

The second version of WSL sees Microsoft place the subsystem on completely new underpinnings. The basis

is the Windows Virtual Computer Platform feature, which also provides the foundation for Hyper-V. Unlike Hyper-V, however, WSL does not rely on full virtualization, but on containers. In concrete terms, this means that not every Linux distribution installed in WSL will run on a fully virtualized machine, but all distributions share a common kernel. The advantage is that WSL gets by with significantly fewer resources than Hyper-V. That said, during my last visit to the site, some add-ons were still missing, and I had to resort to an X server on Windows to run graphical Linux applications [2]. Native support for X11 and Wayland apps was initially only found as a preview in the Windows Insider program. After the introduction of Windows 11, Microsoft has finalized the

Photo by Jack Hunter on Unsplash



Windows Subsystem for Linux GUI (WSLg) [3]. The basis for this is the WSLg system distribution based on CBL-Mariner (Figure 1), which runs Weston, the reference implementation of the server for the Wayland remote display protocol that Microsoft has adapted for use with WSL and optimized for displaying individual applications instead of entire desktop environments over the Remote Desktop Protocol (RDP). Besides Wayland applications, WSLg can also handle classic X11 applications. The additional Pulse Audio Server takes care of the bidirectional transmission of audio signals. When you work with WSL, you will not see any of this architecture, which seems complex at first glance but offers several advantages. The applications within your user distributions, of which you can install and run several in parallel, use their familiar Linux interfaces without customization. Windows, on the other hand, accesses them by RDP without having to worry about implementing X11 or Wayland. In this way, Linux applications can also use the host's physical GPUs, if available. This setup not only benefits graphically intensive applications,

but also those that require GPUs for machine learning (ML) or artificial intelligence (AI) use cases. Microsoft has ported the DirectX 12 interface to Linux for this purpose with its own kernel driver [4].

## Upgrade from Windows 10 to 11

Microsoft has reserved WSLg exclusively for Windows 11 and has not backported it to Windows 10. The good news is that if you own a Windows 11-capable device, you can safely update. In the lab, existing Linux instances survived an upgrade installation from Windows 10 to 11 without any complications. After an upgrade, just make sure that WSL is up to date and that your Linux instances are all using version 2 of the subsystem. You can do this at the command line with admin privileges and list all existing Linux distributions by typing:

```
wsl --update
wsl --list -v
```

If you find a distribution that has not yet moved to version 2, you can change that with

```
wsl --set-version <name of distribution> 2
wsl --shutdown
```

and restart the subsystem.

## Two Approaches to a Fresh Install

If you start with a fresh installation of Windows 11, on which WSL has not run so far, two approaches can give you the desired state. Before you start, make sure that hardware virtualization is active in your system's BIOS. The WSL setup does not check this state and responds with some less than meaningful hex code as an error message if you try to launch a Linux distribution. The hardware virtualization is called *Intel VT-d/VT-x* or *AMD IOMMU*, depending on the processor manufacturer, but it can also be hidden under other terms or in a submenu in the BIOS depending on the brand of your computer. If in doubt, consult the manufacturer's documentation.

If hardware virtualization is active, you can pop up an admin command line and install WSL with just one command that will do all the work for you:

```
wsl --install
```

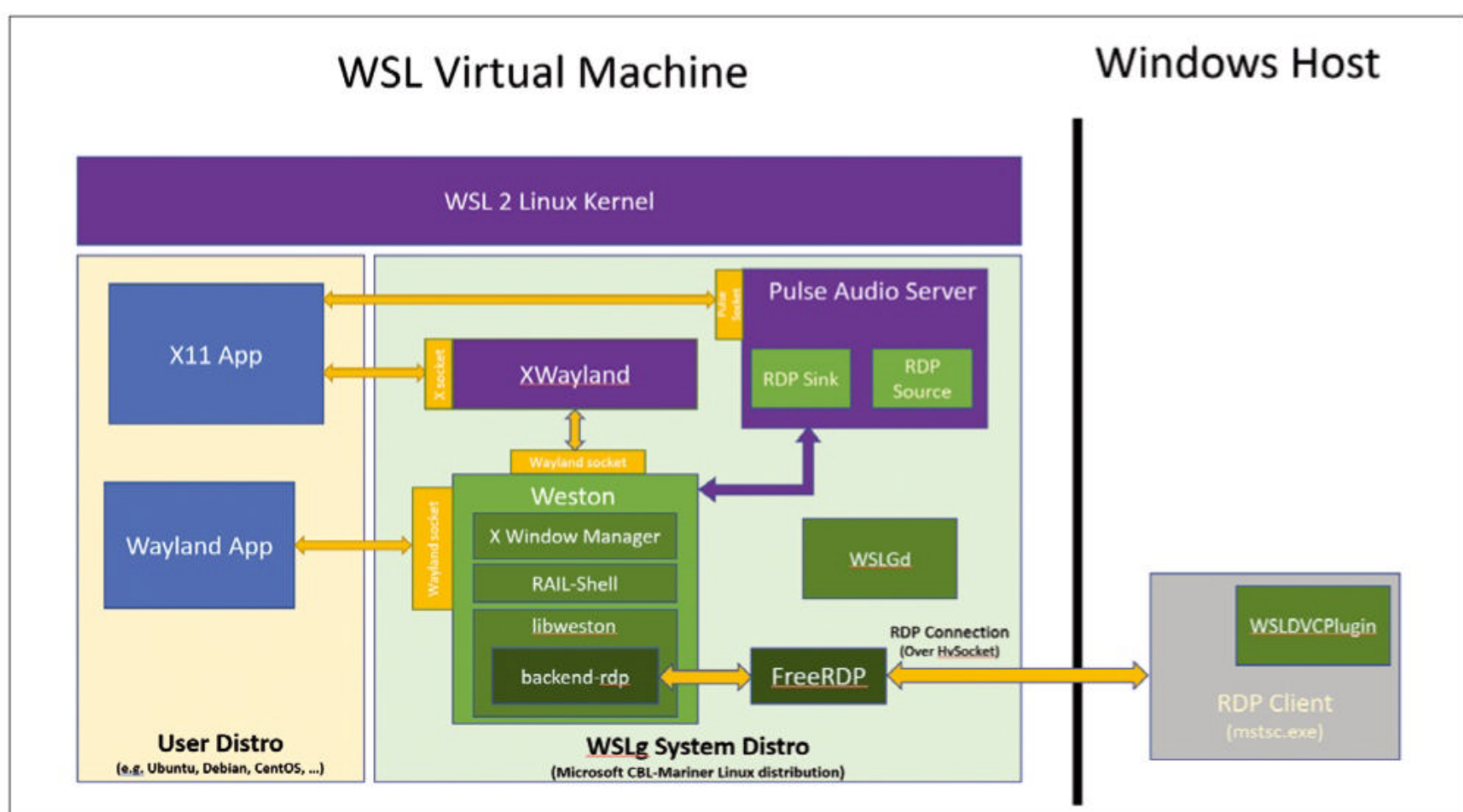


Figure 1: The CBL-Mariner system distribution converts the Wayland and X11 display protocols to RDP.



The command starts by installing the required Windows features, the Virtual Computer Platform, and the Windows Subsystem for Linux if they are not already in place. You can also see this for yourself on Windows 11 in the legacy Control Panel in *Programs and Features* | *Turn Windows features on or off*. The setup routine then loads the current long-term support (LTS) version of Ubuntu from the Microsoft Store as the default distribution. If you prefer some other distribution, such as Kali Linux, which is popular among forensic experts, try:

```
wsl --install --distribution kali
```

Although the advantage is that this command handles all the required steps in one fell swoop, this approach has a drawback.

When installing with `wsl.exe`, the subsystem itself and the Linux kernel use the Windows Update service to retrieve updates and the operating system; however, this approach is no longer preferred by Microsoft. Instead, Microsoft wants to decouple the updates for WSL from the operating system and deliver them from the Microsoft Store [5]. You can find WSL in the Microsoft Store as Windows Subsystem for Linux Preview, but don't worry that Microsoft

currently still rates the app as a preview. It gives you the identical feature set – just not as a Windows feature, but as a state-of-the-art Store app.

Unlike the `wsl` command, however, the Store app unfortunately does not take care of dependencies. Without the virtual machine platform, attempting to run a Linux distribution ends up with a hex code error and a note to the effect that a required feature is not installed. You need to add this feature to Windows first from the control panel or with

```
dism.exe /online /enable-feature /featurename:VirtualMachinePlatform /all
```

on a command line with admin privileges.

## WSLg in Action

After the obligatory reboot, you can install a Linux distribution (e.g., Ubuntu, Kali Linux) from the Microsoft Store and then find it in the Start menu. When called for the first time, WSL opens and initializes a shell for the Linux instance and prompts you to specify a user and a password. This account is independent of your

Windows user account and only exists within this Linux instance.

The usual shell commands let you update Linux and install graphical applications, such as the Gnome text editor or even the Mozilla Firefox web browser:

```
sudo apt update
sudo apt upgrade
sudo apt install gedit -y
sudo apt install firefox -y
```

When you're done, you can see for yourself that graphical applications not only launch from the shell but also show up as subfolders in the respective Linux distribution in the Start menu (Figure 2). WSL has caught up with Google ChromeOS and seamlessly integrates Linux apps into the Windows interface.

File exchange between the two worlds is bidirectional. You can access all your Windows drives in WSL as mount-points (e.g., `/mnt/c` and `/mnt/d`). On Windows, you will find the root filesystems of all your installed distributions in subfolders of the `\\wsl$` virtual network share.

If you need graphical tools that do not exist as native applications for Windows or if you develop ML/AI use cases with Linux as the target platform, WSL on Windows 11 is a useful tool.

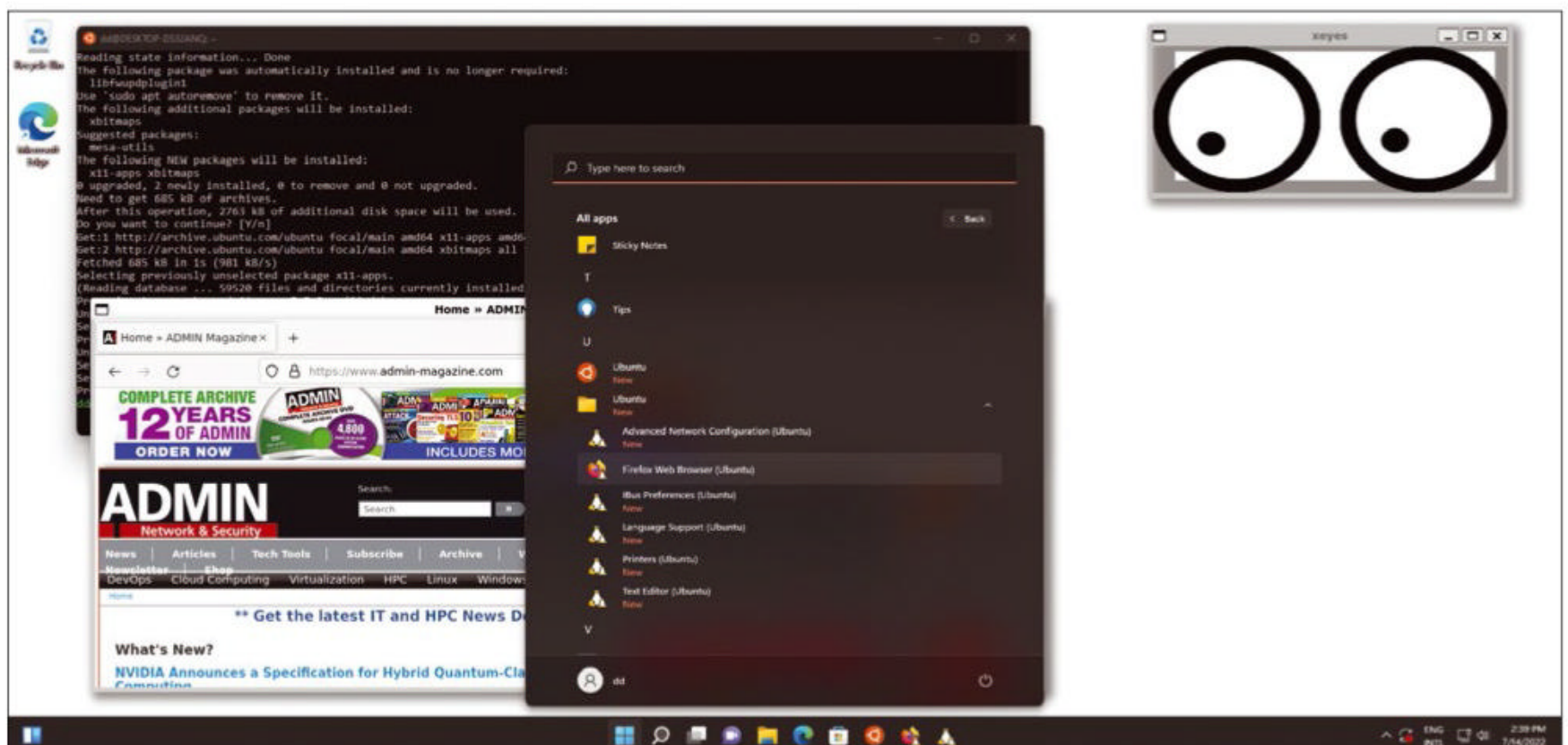


Figure 2: WSLg seamlessly integrates graphical Linux applications with Windows.



## systemd

The systemd init system has become more or less the standard in most distributions. Unfortunately, at the time of writing this article, in contrast to a full-fledged Linux VM, WSL did not have systemd and the common commands like `systemctl`. Since then, systemd for WSL has been released [6], which will allow you, for example, to start the classic trio of daemons for a web server, script interpreter, and database (e.g., Nginx, PHP, and MySQL).

## USB Support over the Network

Maybe you need a smartcard reader or you develop IoT applications and want to configure a microcontroller over USB. In this case, you will unfortunately notice that WSL cannot communicate directly with USB devices connected to Windows at the present time. An indirect approach involving the USB/IP network protocol will help you. In the past, you had to make complex adjustments to the Linux kernel to do this. But on Windows 11, current instances of WSL 2 from kernel version 5.10.60.1 upward do this by default. To begin, install the latest version of *usbipd-win* as described on the project's GitHub page [7] with a `winget` command or as an MSI package. Now launch a Linux distribution and install the appropriate tools for USB/IP there, too [8]:

```
sudo apt install \
  linux-tools-5.4.0-77-generic hwdata
sudo update-alternatives \
  --install /usr/local/bin/usbip usbip \
  /usr/lib/linux-tools/5.4.0-77-generic/\
  usbip 20
```

On the Windows side, list all the available USB devices from an admin command line and connect them to the Linux environment:

```
usbipd wsl list
usbipd wsl attach --busid <ID>
```

Another `list` command shows that the device is connected. On the

The screenshot shows a Windows command prompt window with the following commands and output:

```
C:\WINDOWS\system32>usbipd wsl list
BUSID  DEVICE                                     STATE
1-1    Microsoft Usbccid-Smartcard-Leser (WUDF)    Not attached
1-7    Intel(R) Wireless Bluetooth(R)               Not attached
1-8    Integrated Camera                             Not attached
1-15   Realtek USB 3.0 Card Reader                   Not attached

C:\WINDOWS\system32>usbipd wsl attach --busid 1-1
C:\WINDOWS\system32>usbipd wsl list
BUSID  DEVICE                                     STATE
1-1    Microsoft Usbccid-Smartcard-Leser (WUDF)    Attached - Ubuntu
1-7    Intel(R) Wireless Bluetooth(R)               Not attached
1-8    Integrated Camera                             Not attached
1-15   Realtek USB 3.0 Card Reader                   Not attached

C:\WINDOWS\system32>
```

Below the command prompt, a WSL terminal window is shown with the following output:

```
christian@LenovoX280: ~$ lsusb
Bus 002 Device 001: ID 1d6b:0003 Linux Foundation 3.0 root hub
Bus 001 Device 002: ID 08e6:3437 Gemalto (was Gemplus) GemPC Twin SmartCard Reader
Bus 001 Device 001: ID 1d6b:0002 Linux Foundation 2.0 root hub
christian@LenovoX280: ~$
```

Figure 3: WSL uses USB/IP to access USB devices connected to Windows.

opposite side on Linux, the `lsusb` command should confirm this, as the smartcard reader example demonstrates (Figure 3). After completing these steps, you can detach the device again with the command:

```
usbipd wsl attach --busid <ID>
```

In this way, the WSL can be useful for low-level development, as well.

## Outlook on the Subsystem for Android

Much like the technical underpinnings for WSL, Microsoft is also looking to run mobile apps for Android on Windows with the Windows Subsystem for Android (WSA). In contrast to WSL, WSA is still in its infancy and was not offered until the recent release of Windows 11 version 2022, aka Sun Valley 2. Unfortunately, by the editorial deadline, only a fairly vague announcement had been released on Microsoft's blog. Now, however, the pre-release version of WSA is available in the Microsoft Store [9] [10].

## Retrofitting a Graphical Package Manager

The WSA Package Manager (WSA PacMan, Figure 4) accesses the WSA developer interface and installs Android packages with a

double-click. You can download the installer for WSA PacMan from the project's GitHub page [11] and install the software. It registers as the default for the APK file type if you let it. After rebooting the system, call the package manager from the Start menu, which tells you that WSA is not yet running and offers to start the subsystem.

In the next step, download Android packages from an alternative F-Droid open source repository for Android apps [12] or Aurora OSS [13] on Windows. Now, as soon as you double-click on one of the downloads in APK format, WSA PacMan displays the permissions the respective app requests before proceeding to install the app in the subsystem. You will then find it in the Windows Start menu. Alternatively, you can call up Android's app launcher from the package manager and install additional software from the alternative app stores, such as the privacy-oriented browser by search engine coders DuckDuckGo, which is then also shown directly as a link in the Windows Start menu.

Even without the alternative app stores, various online portals offer you direct downloads of Android packages beyond the official app stores by Google and Amazon. However, be careful, because these offerings are not validated by the store providers, and you cannot rule out



the packages containing malware. Notes on which Android apps work with WSA are also available on GitHub [14].

### Conclusions

Microsoft is continuously expanding WSL's feature set. Under Windows 11, the subsystem has already reached a practical level of maturity that makes a full-fledged virtual or even physical machine unnecessary for many Linux use cases. Graphical applications integrate seamlessly, and even the systemd init system is now available. In contrast, WSA is still in its infancy, but this new subsystem demonstrates that Windows can be a candidate as a platform for Android apps. The practical benefits totally depend on the available apps. It remains to be seen whether Microsoft and Amazon will expand their offerings and open up to customers

outside the US market. Until then, the preview lets you try out WSA. ■

#### Info

- [1] CBL-Mariner: [https://github.com/microsoft/CBL-Mariner]
- [2] "Linux apps on Windows 10 and Chrome OS" by Christian Knermann, *ADMIN*, issue 66, 2021, pg. 88, [https://www.admin-magazine.com/Archive/2021/66/Linux-apps-on-Windows-10-and-Chrome-OS/]
- [3] Run Linux GUI apps on WSL: [https://learn.microsoft.com/en-us/windows/wsl/tutorials/gui-apps]
- [4] DirectX for WSL: [https://devblogs.microsoft.com/directx/directx-heart-linux/]
- [5] WSL in the Windows Store: [https://devblogs.microsoft.com/commandline/a-preview-of-wsl-in-the-microsoft-store-is-now-available/]
- [6] systemd for WSL: [https://devblogs.microsoft.com/commandline/systemd-support-is-now-available-in-wsl/]

- [7] Download usbipd-win: [https://github.com/dorssel/usbipd-win]
- [8] usbipd-win and WSL: [https://github.com/dorssel/usbipd-win/wiki/WSL-support]
- [9] WSA with Amazon Appstore: [https://apps.microsoft.com/store/detail/windows-subsystem-for-android(tm)-with-amazon-appstore/9P3395VX91NR]
- [10] Release Notes for WSA: [https://learn.microsoft.com/en-us/windows/android/wsa/release-notes]
- [11] WSA PacMan: [https://github.com/alesimula/wsa\_pacman/releases]
- [12] F-Droid: [https://f-droid.org]
- [13] AuroraOSS : [https://auroraoss.com]
- [14] WSA-compatible apps: [https://github.com/riverar/wsa-app-compatibility]

#### Author

Christian Knermann is Head of IT-Management at Fraunhofer UMSICHT, a German research institute. He's written freelance about computing technology since 2006.

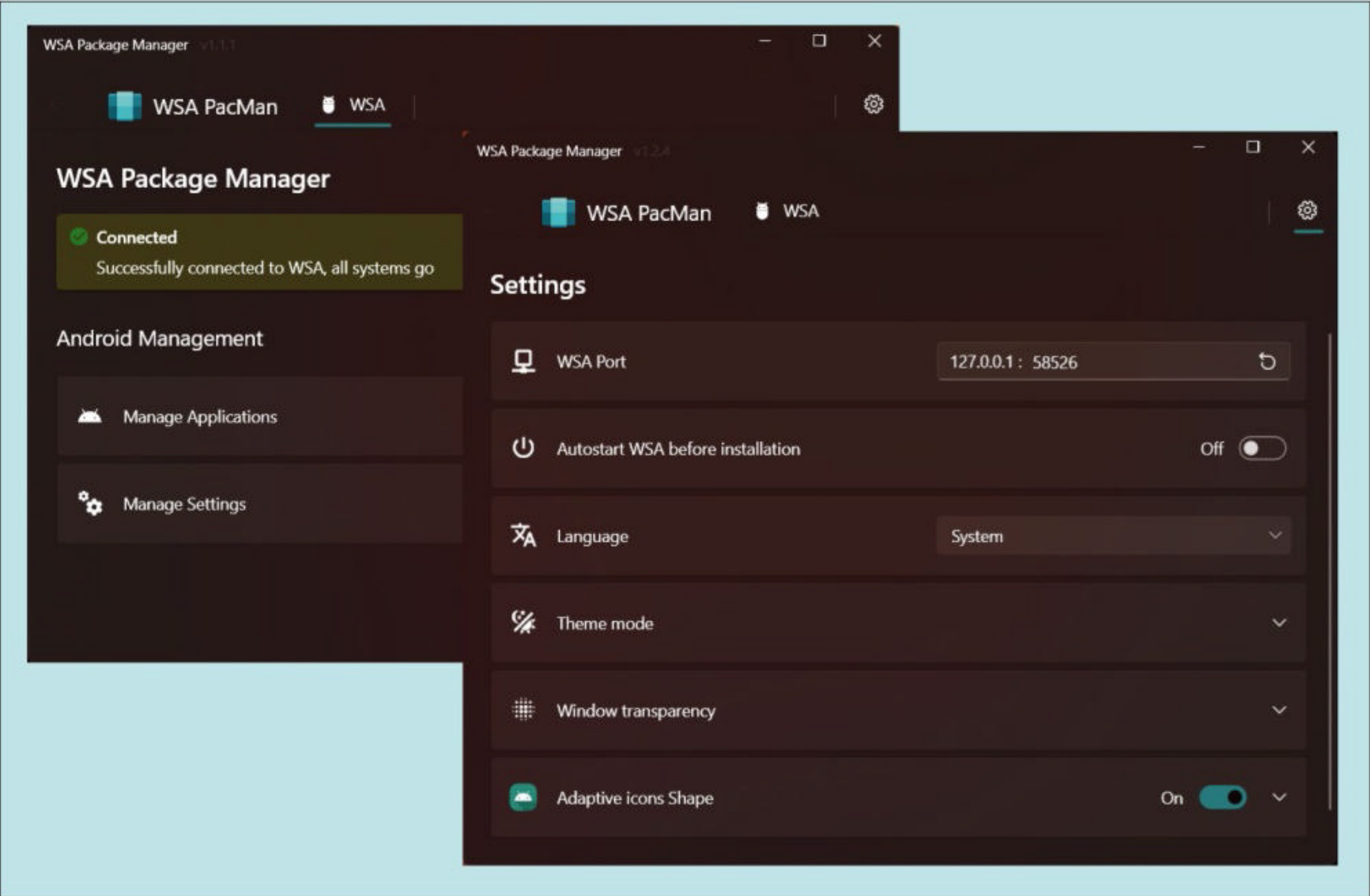


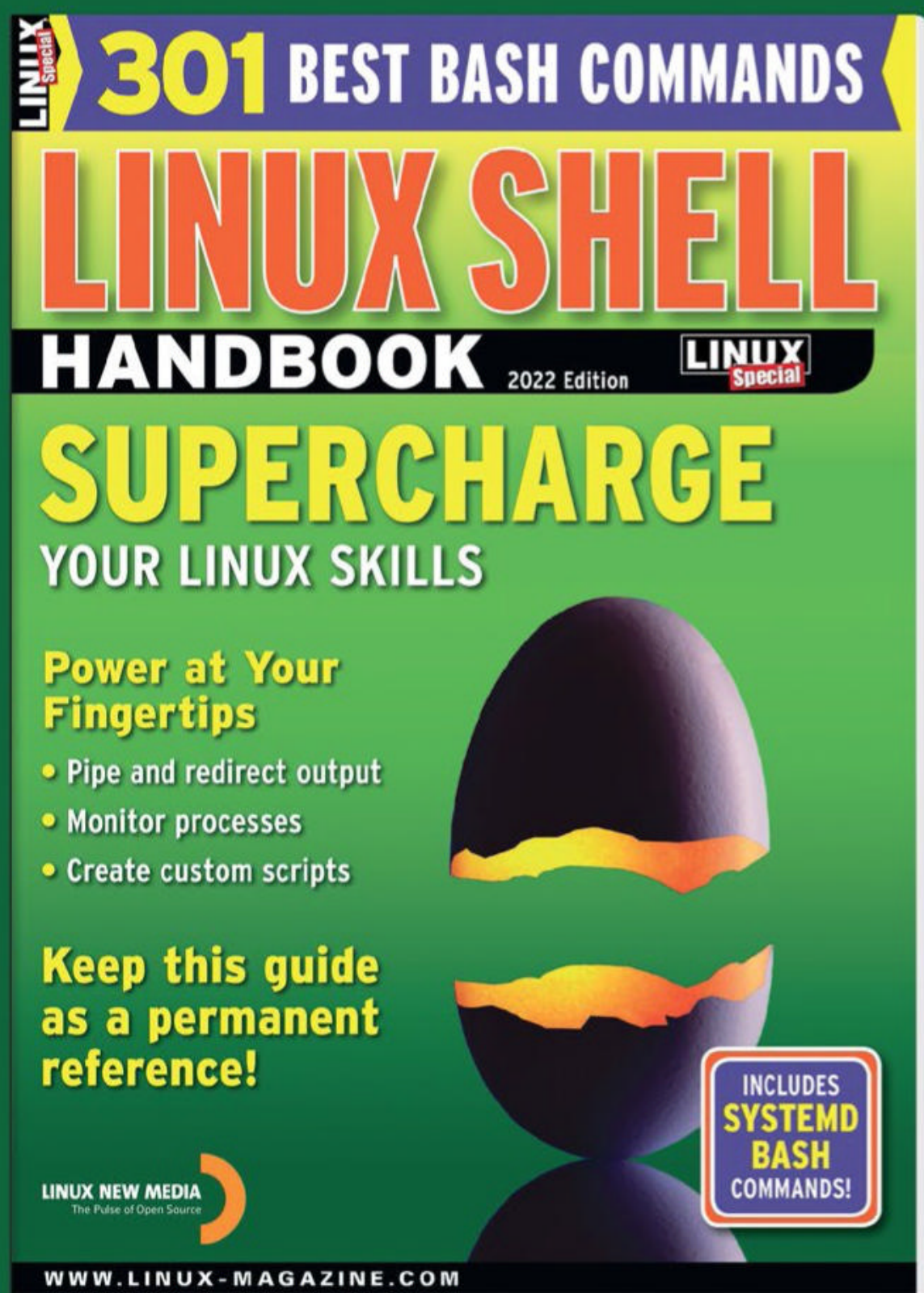
Figure 4: WSA forms the basis for Android apps on Windows, and WSA PacMan helps with the install.



# THINK LIKE THE EXPERTS

Linux Shell Handbook 2022 Edition

This new edition is packed with the most important utilities for configuring and troubleshooting systems.



Here's a look at some of what you'll find inside:

- Customizing Bash
- Regular Expressions
- Systemd
- Bash Scripting
- Networking Tools
- And much more!

ORDER ONLINE:

[shop.linuxnewmedia.com/specials](http://shop.linuxnewmedia.com/specials)





Load test your website with Siege

# Stress

A stress and benchmarking tool for websites controlled from the command line. By Ali Imran Nagori

**Monitoring the performance** of a website and fixing performance issues is a day-to-day job for sys admins. Ignoring this part of the job could result in the loss of potential customers. Siege is an open source benchmarking tool that can greatly help you in this regard. Benchmarking tools [1] perform standardized tests for accessing the relative capabilities of hardware or software and evaluating their results.

## Raison d'Stress

Siege puts your website under different stress conditions and then

### Important Note

Running Siege against third-party websites and servers without the consent of the owner is an illegal practice. Siege has a distributed denial-of-service (DDoSing) effect; as such, it can bring down a running server. This article is written only for educational purposes, and you alone are responsible for any damage caused by any misuse of this software.

evaluates the following performance parameters, among others:

- Availability of the server
- Response time
- Number of successful and failed transactions
- Throughput

Siege developers define each parameter, which I talk about later in this article. Before getting your hands dirty, though, please read the “Important Note” box.

Although running Siege does not require much expert knowledge, interpreting its result is what you should care about, and I will try to make that part simple. Running Siege requires sudo privileges, so be sure you have administrative access to your system. In this article, I provide instructions on how to use the Siege benchmarking tool to stress test a website hosted on an Ubuntu 20.04 server.

## Why Care About Siege?

Your job as a software developer doesn't end until you are confident

your code will be stable in the worst conditions. Surviving traffic spikes, system crashes, memory overuse, and so on are core milestones that software needs to achieve. Siege plays a vital role in checking the stability of your code in rush hour traffic conditions.

Siege checks the strength of your code by placing it under stress, so you will know whether your site can sail smoothly under scores of transactions or whether it will crash. In this way, you can estimate how much customer volume your website can withstand until it breaks.

## Installing Siege

Siege can be installed two ways: from the official distribution package or from source code.

To install from the official Ubuntu repositories, open a terminal and execute the command:

```
sudo apt-get install siege -y
```

That's it. Siege is installed.

To install from the source code, download the latest version of the file [2] and enter the following



```
vagrant@LHB:~$ tar -xzf siege-latest.tar.gz
vagrant@LHB:~$ ls
file1  siege-4.1.5  siege-latest.tar.gz  test
vagrant@LHB:~$
```

Figure 1: Extracting the file.

```
vagrant@LHB: ~/siege-4.1.5
vagrant@LHB:~$ cd siege-4.1.5/
vagrant@LHB:~/siege-4.1.5$ ls
acinclude.m4  AUTHORS  configure.ac  html  install-sh  README.md
aclocal.m4    ChangeLog  COPYING      include  Makefile.am  src
acspecific.m4  configure  doc          INSTALL  Makefile.in  utils
vagrant@LHB:~/siege-4.1.5$
```

Figure 2: Files inside the Siege directory.

```
vagrant@LHB:~$ siege 192.168.56.11
** SIEGE 4.0.4
** Preparing 25 concurrent users for battle.
The server is now under siege...^C
Lifting the server siege...
Transactions:          29068 hits
Availability:          100.00 %
Elapsed time:           14.44 secs
Data transferred:      89.76 MB
Response time:          0.01 secs
Transaction rate:      2013.02 trans/sec
Throughput:             6.22 MB/sec
Concurrency:           24.61
Successful transactions: 29068
Failed transactions:    0
Longest transaction:    0.16
Shortest transaction:    0.00

vagrant@LHB:~$
```

Figure 3: Running Siege against a single target.

```
vagrant@LHB:~$ siege -t1M 192.168.56.11
** SIEGE 4.0.4
** Preparing 25 concurrent users for battle.
The server is now under siege...
Lifting the server siege...
Transactions:          126815 hits
Availability:          100.00 %
Elapsed time:           59.91 secs
Data transferred:      391.60 MB
Response time:          0.01 secs
Transaction rate:      2116.76 trans/sec
Throughput:             6.54 MB/sec
Concurrency:           24.62
Successful transactions: 126815
Failed transactions:    0
Longest transaction:    0.18
Shortest transaction:    0.00
```

Figure 4: Setting the time for a Siege operation.

```
vagrant@LHB:~$ siege -c 30 192.168.56.11
** SIEGE 4.0.4
** Preparing 30 concurrent users for battle.
The server is now under siege...^C
Lifting the server siege...
Transactions:          9725 hits
Availability:          100.00 %
Elapsed time:           8.11 secs
Data transferred:      30.03 MB
```

Figure 5: Setting the number of Siege users.

commands to extract the file (Figure 1), change to the extracted directory, run the configure script (Figure 2), and complete the installation:

```
$ tar -xzf siege-latest.tar.gz
$ cd siege-<version>/
$ ./configure
$ sudo sh -c "make && make install"
```

That's it. Siege should now be installed on your system. To make sure everything is ready to go, enter:

```
$ siege -V
```

Next, you want to enable logging by opening the file `siegerc` in Nano,

```
$ sudo nano /etc/siege/siegerc
```

and searching for the line:

```
#logfile = $(HOME)/var/log/siege.log
```

Simply remove the starting hash symbol (#) to uncomment and save the file.

## Siege on the Battlefield

To understand how siege works, you first need to deploy a simple website. In this case, I use an Apache web server that comes with a default web page. In a basic test, Siege can be run with either the IP address of the target or its fully qualified domain name (FQDN). When invoked without options, Siege uses the default configuration in the `siegerc` file (i.e., the `siege.conf` file):

```
$ siege 192.168.56.11
```

As Figure 3 shows, Siege made a total of 29,068 transactions.

Runtime can be controlled with the `-t` (`--time`) flag. The format is `<NUM><m>`, where `<NUM>` denotes the proposed time and `<m>` sets the unit of time to either S, M, or H (i.e., seconds, minutes, and hours).

To set the running time for each user to one minute (Figure 4), enter:

```
$ siege -t1M http://192.168.56.11
```



The default value of the number of users in a Siege operation is 25. However, you can override the number of simultaneous users hitting a server with the `-c` flag. To set this number to 30 (Figure 5), use the command:

```
$ siege -c 30 http://192.168.56.11
```

You can also use Siege to test multiple websites at once. For this, you need to create a file containing a list of the target sites. You can create such a file by going into a text editor, adding the URLs of target sites line by line, and saving the file:

```
$ nano ~/target-sites.txt
www.host1.com
192.168.56.12
```

As you can see, you can also use the target IP addresses. Now run Siege against these sites:

```
siege -f ~/target-sites.txt
```

The results are as shown in Figure 6. So far, the Siege command line has only comprised one or two arguments. However, you can also combine multiple operations in a single command (Figure 7),

```
siege -b -c 30 -t 30s 192.168.56.11
```

where `-b` specifies the benchmark mode, `-c` sets the concurrency of

connections (in this case, 30 users), and `-t` sets the runtime.

## Interpreting the Results

Now that you have seen Siege in action, I'll explore some of the critical metrics reported in Figures 6 and 7. *Transactions* represents the number of times Siege users hit a server. As mentioned earlier, the default is 25 users. If every user hits the server 10 times, the total transactions would include 250 hits. However, this metric also includes meta redirects and responses from multiple elements of a web page, which could result in a larger transaction rate value from that calculated. *Availability* is the percentage of successful socket connections handled by the server. This percentage is derived from the ratio of the number of socket failures to total connection requests. *Response time* is another important parameter that shows the average time a Siege user takes to respond to requests. *Transaction rate* is basically the proportion of all the transactions to the total duration of the test. *Throughput* is the mean traffic (in bytes per second) sent by the server to all simulated Siege users.

## Conclusion

In this article, I looked at how Siege can be used to determine the potential of your website under stress. For

you to fine-tune your website, you need to know how it's performing under various traffic conditions. Siege is a good choice for web developers who want to check the stability of their code.

Siege puts a lot of stress on the server, so you might experience system lag while trying to log in or while using the system. If you are new to or just learning about stress testing, Siege is a good tool to play around with; however, be cautious while running Siege in a production environment.

The Siege man page [3] contains a lot of useful information, so you should consult these pages to discover more about the possibilities of Siege. ■

---

### Info

- [1] "Benchmarking," ITC 250 W14 3210 - Web App Programming 2 (retrieved October 20, 2022), [[https://canvas.seattlecentral.edu/courses/937693/pages/17-benchmarking?module\\_item\\_id=8157211](https://canvas.seattlecentral.edu/courses/937693/pages/17-benchmarking?module_item_id=8157211)]
- [2] Siege FTP page: [<https://download.joedog.org/siege/>]
- [3] Siege man page: [<https://linux.die.net/man/1/siege>]
- [4] The author on LinkedIn: [<https://www.linkedin.com/in/ali-imran-nagori/>]

---

### Author

Ali Imran Nagori is a technical writer and Linux enthusiast who loves to write about Linux systems administration and related technologies. He blogs at [[tecofers.com](https://www.tecofers.com)]. Connect with him on LinkedIn [4].

```
vagrant@LHB:~$ siege -f ~/target-sites.txt
** SIEGE 4.0.4
** Preparing 25 concurrent users for battle.
The server is now under siege...^C
Lifting the server siege...
Transactions:          7131 hits
Availability:          100.00 %
Elapsed time:           5.25 secs
Data transferred:      22.02 MB
Response time:          0.02 secs
Transaction rate:      1358.29 trans/sec
Throughput:             4.19 MB/sec
Concurrency:           24.09
Successful transactions: 7131
Failed transactions:    0
Longest transaction:    0.15
Shortest transaction:   0.00

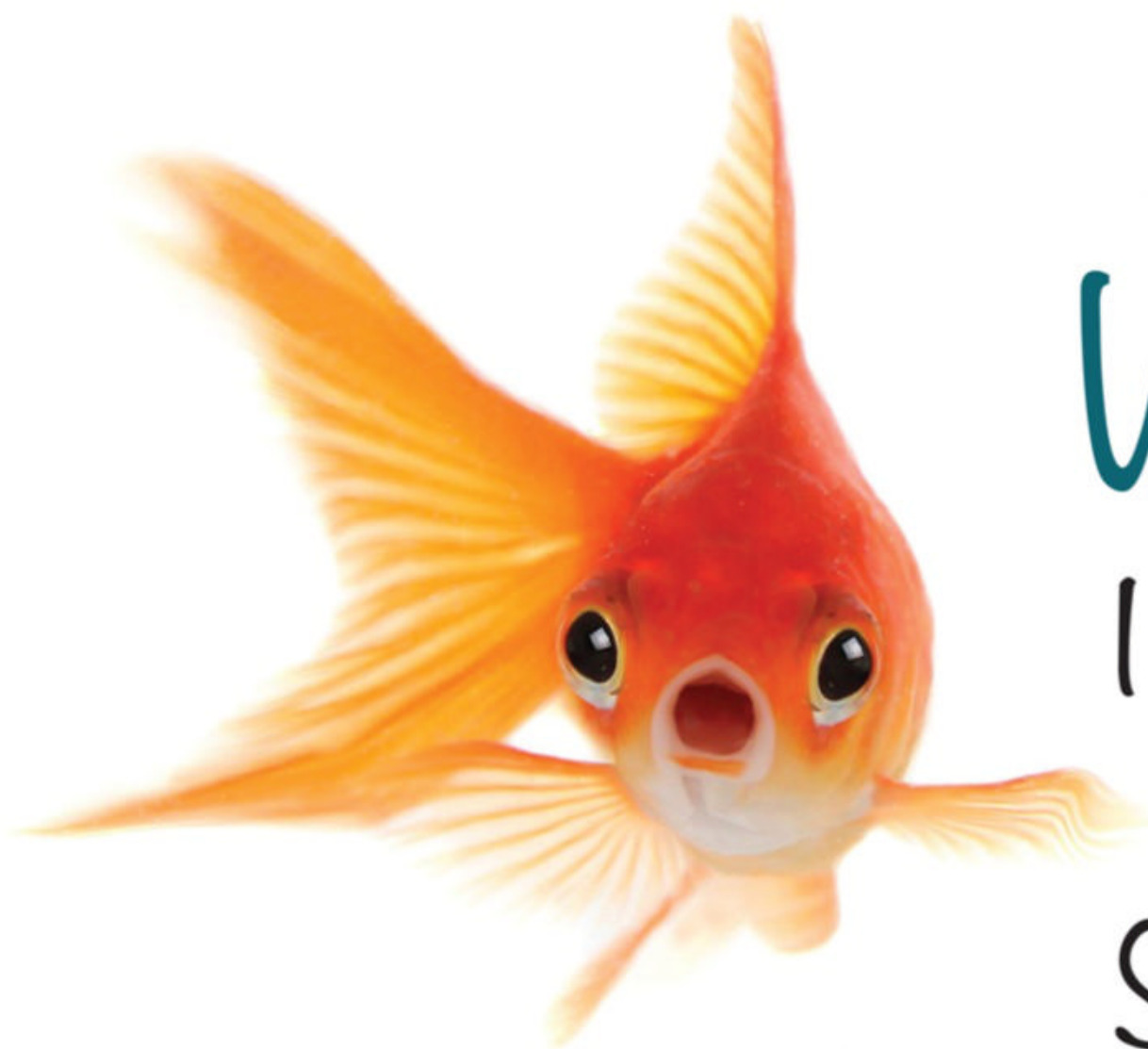
vagrant@LHB:~$
```

Figure 6: Targeting multiple sites.

```
vagrant@LHB:~$ siege -b -c 30 -t 30s 192.168.56.11
** SIEGE 4.0.4
** Preparing 30 concurrent users for battle.
The server is now under siege...
Lifting the server siege...
Transactions:          38161 hits
Availability:          100.00 %
Elapsed time:          29.17 secs
Data transferred:     117.84 MB
Response time:         0.02 secs
Transaction rate:     1308.23 trans/sec
Throughput:            4.04 MB/sec
Concurrency:          29.37
Successful transactions: 38161
Failed transactions:    0
Longest transaction:    0.30
Shortest transaction:   0.00
```

Figure 7: Combining multiple Siege operations.





# What?!

I can get my  
issues  
SOONER?



Available anywhere, anytime!

Sign up for a digital subscription to improve our admin skills with practical articles on network security, cloud computing, DevOps, HPC, storage and more!

Subscribe to the PDF edition: <https://bit.ly/digital-ADMIN>

Now available on ZINIO: [bit.ly/ADMIN-ZINIO](https://bit.ly/ADMIN-ZINIO)



Alternative virtualization solutions  
when OpenStack is too much

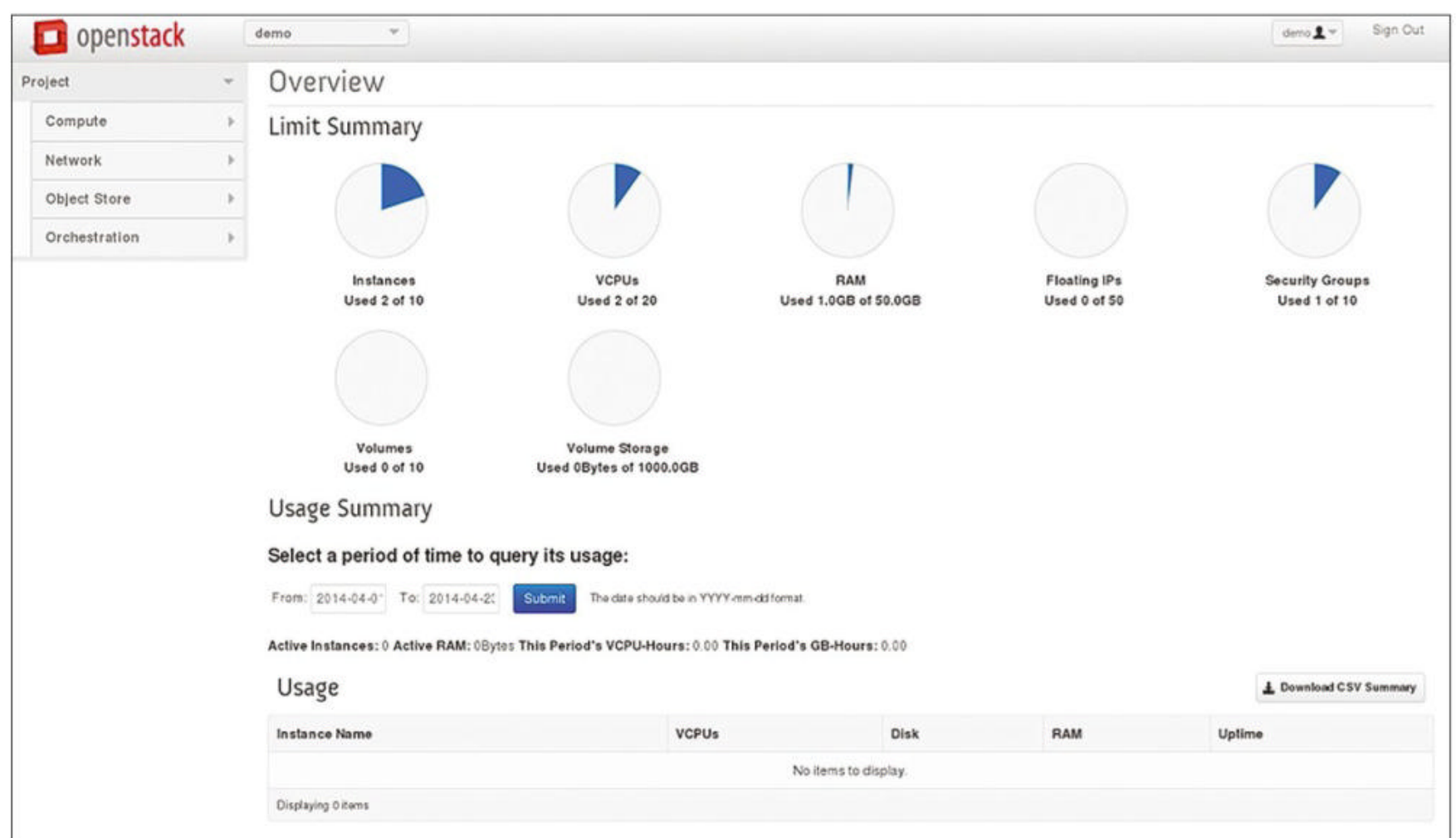
# Plan B

OpenStack is considered the industry standard for building private clouds, but the solution is still far too complex and too difficult to maintain and operate for many applications. What causes OpenStack projects to fail, and what alternatives do administrators have? By Martin Loschwitz

**Ten years ago**, reports of the decline of OpenStack would have been completely unthinkable. For a while at least, many people thought that OpenStack was a miraculous solution

for most of their IT problems. From the perspective of 2022, many, if not most, of the hopes once associated with OpenStack (Figure 1) have not been met for the majority of users.

In the Kubernetes era, OpenStack has shrunk massively as a project, with nothing like the number of developers actively involved as there were the past. Similarly, many



**Figure 1:** OpenStack is considered the ultimate in private cloud software, but the complexity of the environment often exceeds what enterprises can realistically achieve. © OpenStack Foundation



formerly prominent and almost militant OpenStack proponents have scaled back or ended their OpenStack involvement.

This development also has its good side. In the course of the past decade, several large-scale OpenStack projects failed because expectations and requirements were more or less undefined up front. Additionally, it was not clear in some cases what OpenStack is, what it can do, and what it cannot do. Often enough, OpenStack was chosen for projects that could not be implemented at all with this tool, or with massive restrictions.

That many manufacturers have dropped out is not automatically considered a great misfortune. For example, participation of all storage vendors (e.g., HP, Dell EMC, SanDisk) meant great hardware support for numerous devices, but it also meant all the major vendors dispatched as many employees as possible into the fray to determine the fate of the project. OpenStack has now largely freed itself from this stranglehold.

## Where OpenStack Fails

Administrators are still skeptical when it comes to storing their data with AWS and the like and face a dilemma: A public cloud doesn't work because people are worried about their data, but when it comes to a private cloud, many administrators automatically think of OpenStack and get cold feet. It makes sense to address the reasons that cause OpenStack projects to fail, because this is the only way to empower admins to evaluate realistically whether OpenStack is suitable for their intended use case. Where the answer to the suitability question is "no," I look at potential alternatives.

The saying "the cloud is just somebody else's computer" stopped being true a long time ago. Today, "cloud" also means infrastructure automation, Infrastructure as Code (IaC), immutable underlay, micro-architecture, and so on. Although OpenStack development has slowed,

it still remains a complex construct and today comprises more than 30 individual components; admittedly, not every admin really needs every one of them. If you reduce the number of tools to the absolute minimum, however, you are still left with at least six components. You need to know what they are and clearly understand the basic functionality of tools such as Nova, Cinder, or Glance.

Many large OpenStack projects have failed to make this simple hurdle. In many cases, problems of understanding have led to companies seeing OpenStack as a continuation of VMware by other means. In terms of comfort and ease of administration, however, OpenStack cannot and never wanted to compete with VMware. Moreover, classic VMware setups are often based on completely different premises; for example, software-defined networking (SDN) is missing as a core component. Many companies that adopted OpenStack in the hope of cost savings suddenly found themselves faced with uncontrollable complexity and eventually gave up in total frustration. Often the reason was that the goal of operating some virtual machines (VMs) was out of proportion to the overhead.

## OpenStack Is Complex

To this day, OpenStack Summits regularly feature presentations of companies proudly talking about their automated OpenStack, putting the number of servers per admin at 2,000 units and more. In fact, OpenStack can be operated in this way, but before automation and orchestration are set up, adapted, and in operation, a great deal of work must be done – and by personnel who know their stuff.

It is virtually impossible for a company with no OpenStack experience to turn to a boxed solution like Red Hat OpenStack Platform (RHOP) and build a giant cloud with full automation in just a few weeks. After too long a development period

without concrete results, quite a few companies are either running out of patience, money, or both to continue working on OpenStack. It is precisely these OpenStack behemoth projects that have done massive damage to the software's reputation over the past decade.

In summary, the reasons for the failure of OpenStack in many organizations are: too little research in advance, unrealistic assessment of OpenStack's complexity, wrong choice of OpenStack in the assumption that all of its features are needed, and excessive overhead for operating the product itself in an automated and therefore efficient way.

## Possible Alternatives

The beauty of analysis is that it also shows a way out of the predicament by setting out the parameters that ought to play a role when choosing an alternative to OpenStack.

Many companies, for example, need virtualization with a feature such as high availability, but often multi-tenancy is not a factor. However, this already accounts for a considerable part of OpenStack's complexity, because it often drags in other features, such as SDN.

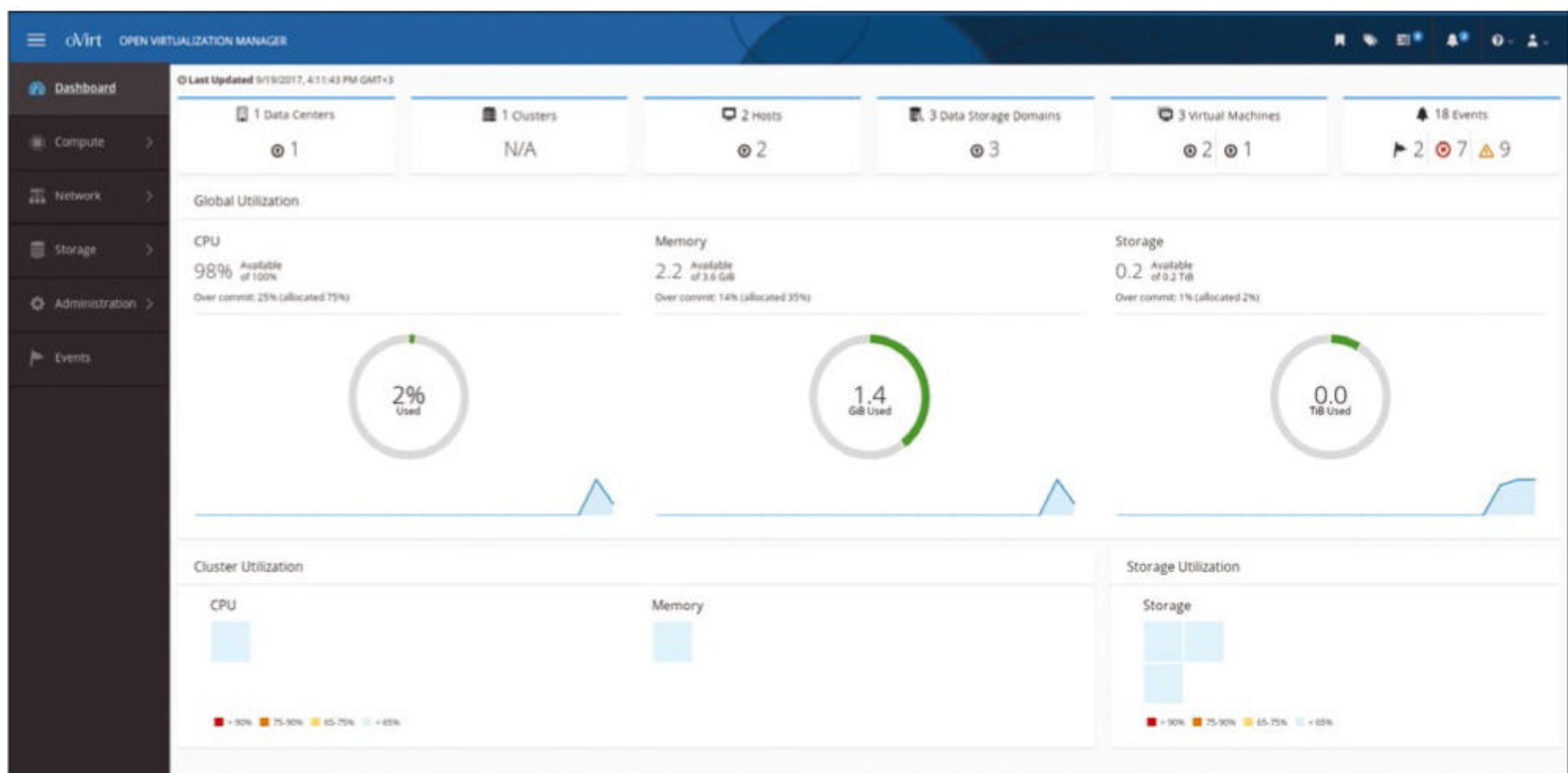
Still other companies might need a sophisticated virtual network, but not OpenStack's self-service capabilities. These capabilities frequently play a significant role in the project, because on-demand use is a central characteristic of clouds.

Classical virtualization with minor VM fluctuation, on the other hand, can often do without colorful web interfaces, which in turn contributes significantly to reducing complexity. This pruning to essentials leaves a small but select group of less complex OpenStack alternatives, which I discuss in detail.

## oVirt

Some administrators might be more familiar with oVirt ([Figure 2](#)) under the name of its commercial variant, because oVirt is the open source





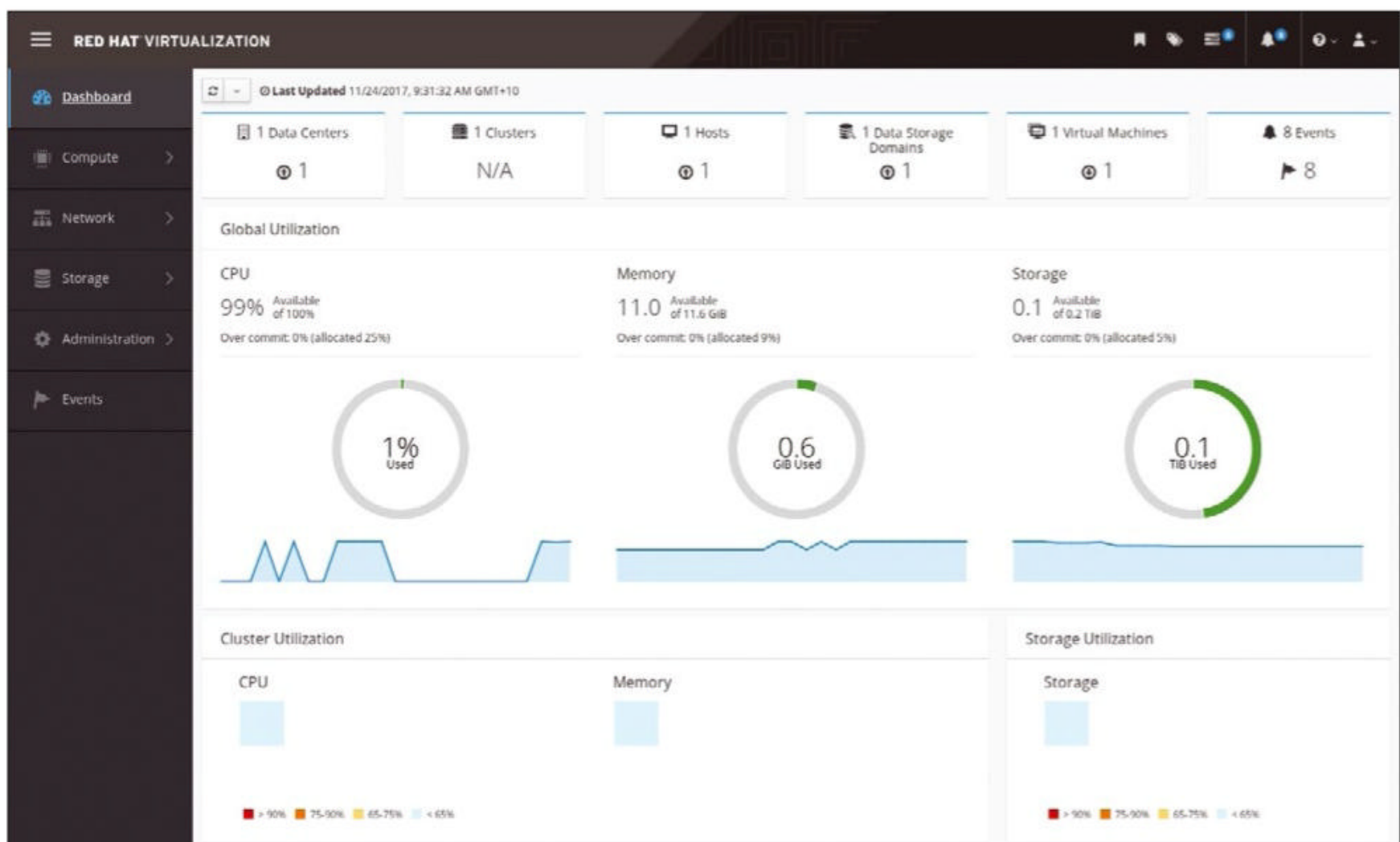
**Figure 2:** oVirt is Red Hat's virtualization manager, offering many features that can also be found in OpenStack. © Red Hat

foundation of Red Hat Enterprise Virtualization (RHEV; [Figure 3](#)). Red Hat continues to be one of the most active OpenStack developers and thus one of the few companies that still have a commercial OpenStack distribution in its portfolio – Red Hat OpenStack Platform. However, RHOP is hardly likely

to compete with its sibling RHEV internally, which means that oVirt is a plain vanilla virtualization solution for small environments, although the feature set is quite impressive.

The linchpin of the product is a controller node that runs all oVirt basic services. Red Hat fans have

a clear advantage because oVirt is only available in packages for RHEL 8 and CentOS Stream 8. However, it will run on AlmaLinux 8 and Rocky Linux 8. After installing the basic components, you first need to assign disk space to the new oVirt instance – with a cornucopia of options from which to choose, including NFS,



**Figure 3:** Those who need an option with support rely on oVirt's commercial distribution RHEV. © Red Hat



GlusterFS, Fibre Channel, and other POSIX-compatible filesystems.

Ceph is interestingly missing from the list of storage drivers for oVirt, which is surprising because Ceph has long established itself as a rock-solid solution for scalable storage, even outside of OpenStack. Unlike OpenStack, Ceph can be set up and ready for operation in a fully automated manner in a fairly short time. Unlike OpenStack, Ceph causes very little overhead in everyday administration and takes care of most of its problems itself to boot.

If you dig deeper into the documentation, you will find out how Red Hat imagines the Ceph connection in oVirt, and you will end up with – surprise – an OpenStack component. As you know, the service responsible for storage in OpenStack is named Cinder. It can run quite well without the rest of the OpenStack components, essentially communicating with the storage back ends in the background to create volumes. Cinder then outputs their paths to requesting instances such as oVirt.

If you combine oVirt and Ceph, you also get a little piece of OpenStack, but the overhead involved in using the *cinderlib* driver in oVirt doesn't even begin to compare with the complexity of the connection in OpenStack itself.

## Many Features

In addition to the controller, the administrator of an oVirt installation also provisions any number of virtualization hosts (i.e., the actual compute nodes). The nodes communicate with their controller through a controller agent process and thus know what they have to do. The reward for this effort is a small but fine virtualization solution that can even be extended to include a rudimentary kind of self-service over a Lightweight Directory Access Protocol (LDAP) connection. However, oVirt does not envisage complex virtual network setups. On the one hand, this facilitates troubleshooting; on the other hand, it drastically

reduces the number of possible sources of error.

oVirt also integrates a feature that many admins miss in OpenStack right out of the box. The data warehouse writes metrics data with thousands of values in the background and makes them directly available in oVirt through interfaces, which makes it far easier to connect to monitoring or trending systems. An oVirt exporter can even export data from oVirt's data warehouse in a format suitable for Prometheus [1].

In terms of everyday functionality, oVirt leaves little to be desired. Ready-made operating system images can be defined in this way, and it is possible to create snapshots of existing VMs, making backup and restore operations far easier. All told, oVirt is a very powerful OpenStack alternative, especially for small setups. Admins who just need virtualization should keep that in mind.

## DIY Cluster: Linux-HA

It may be down to my nostalgic feelings that I've even mentioned the virtualization cluster archetype; however, it could also be because a setup of this type is the easiest and fastest way to set up a few highly available VMs as quickly as possible.

Considering the complexity of solutions like OpenStack and even oVirt, the Linux-HA (high availability) stack is definitely not the most complex product under the sun. Additionally, Red Hat has put a great deal of work into getting the former horrors of Heartbeat 2 and its successor Pacemaker under control in recent years. Today, hacking XML files is just as superfluous as memorizing cryptic shell commands to feed the cluster with the resources you want it to manage.

At the same time, the Linux-HA stack and all its components are more or less available wherever you look. Red Hat (RHEL) and SUSE Linux Enterprise (SLES) distribute it in the form of an add-on, and matching packages are also included with Ubuntu and Debian. In other words: If you have

three systems running a popular Linux distribution, you can quickly get a Pacemaker cluster up and running.

Countless how-tos online now explain how this works in great detail.

Pacemaker has probably not been used for any task as regularly over the past 15 years as it has been for running virtual instances. The VirtualDomain resource agent has long since proved its value and is now generally thought of as being reliable and stable. Connecting external storage services such as Ceph is almost easier in a setup of this type than with oVirt. You just need to enter the path to your RADOS block device (RBD) in the *libvirt* definition of the respective instance, and you are ready to go.

On top of that, Pacemaker handles at least basic balancing of resources easily. For example, a three-node cluster can be made to distribute three VMs evenly across the three servers. You can even have a Pacemaker instance shut down should a server fail. Although this method is not a comprehensive placement strategy, like those that exist in OpenStack or oVirt, it is often good enough for small environments. In this scenario, however, administrators do have to do without many goodies, such as self-service over a web user interface (UI) or virtual networking by SDN. Colorful wizards for setting up new instances are also missing; in fact, the package does not even include image management. If you need any of these features, you are undoubtedly better off with oVirt or Proxmox, but if you really only want to virtualize a few VMs, Pacemaker will get you there quickly. Today, Pacemaker even has automation modules (e.g., Ansible) that let you handle tasks completely automatically.

## The Veteran: Proxmox

If you think Pacemaker is too much of a hassle, you'll probably be happier with Proxmox. From today's perspective, Proxmox can almost be considered a kind of veteran in the virtualization business, and what



the project has achieved over the past few years certainly speaks for itself. In many respects, Proxmox is even on a par with its competitor VMware – but let's take things one step at a time.

At the heart of Proxmox was the desire to create a simple way to manage virtual instances and their storage needs. When the first versions of Proxmox saw the light of day, neither oVirt nor OpenStack were as advanced as they are today, and a kind of race for users' favor emerged between Proxmox and oVirt in particular. As things stand, Proxmox might even have won the honors because the product is now seen as the logical VMware replacement and successor to oVirt in many places. In terms of functionality, Proxmox (**Figure 4**) really doesn't need to hide its light under a bushel. The key to the solution is a dedicated management plan that can be rolled out to one or dozens of systems, with the systems handling the coordination tasks themselves. One major strength of the solution is its web interface, which gives even less experienced

users access to quite a few functions in the background. Thanks to appropriate rights management and role-based access control (RBAC), the Proxmox UI can also be used for basic self-servicing.

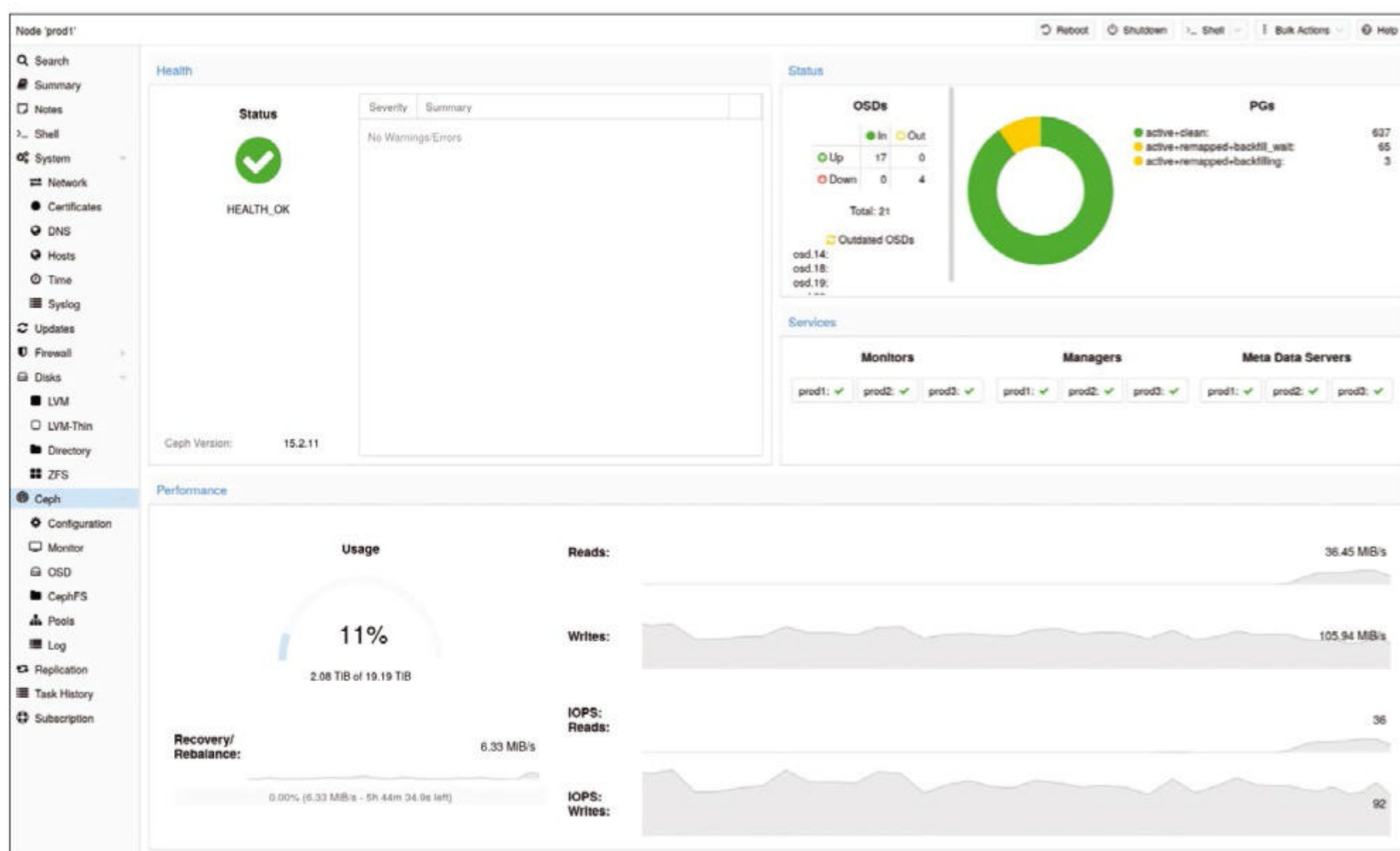
Users who found tinkering with Pacemaker too much can enjoy complete Pacemaker handling as part of the Proxmox package. Compute nodes can be defined as HA nodes, and Proxmox takes care of rolling out and configuring the required HA services on the affected systems in the background.

Moreover, the integration of various storage back ends is legendary. For HA, for example, support for the distributed replicated block device (DRBD) replication solution was offered as early as a decade ago. Today, people tend to rely on Ceph, which Proxmox can even automate when called on to do so. Of course, this only works if you have the right hardware in place. The same applies to your own storage replication stack based on ZFS if no hardware is available for Ceph or you do not want to use the product.

Beyond that, Proxmox leaves little to be desired in terms of functionality. For storing templates for virtual instances, Proxmox supports KVM virtualization, as well as container virtualization that is based on Linux containers (LXC). Proxmox handles containers and VMs as equivalent virtual instances. Its own setup is also manageable. The solution is up and running quickly, not least thanks to a comprehensive installation guide on the project page.

The provider's distribution policy causes many an administrator to frown – although this attitude is not totally fair. Proxmox virtual environment (Proxmox VE) is available under a free license and can be obtained as open source software from the provider's GitHub directories completely free of charge. However, building packages is then up to the user, and it is this step that many companies want to avoid.

Those who are only keen on the packages and do not want any support from the manufacturer are asked to pay a EUR95 (~\$93) subscription fee per year and CPU



**Figure 4:** Proxmox is a genuine veteran when it comes to virtualization, but today it coexists well with other solutions (e.g., Ceph) and even rolls them out on its own in some cases. © Proxmox



socket, so a setup with four sockets would cost EUR380 (~\$372). Given Proxmox's impressive feature set and ease of installation, this does not seem excessive.

## OpenNebula

OpenNebula belongs on any list of OpenStack alternatives. The product has basically acted as a kind of anti-OpenStack from the very beginning. Far less complexity, simpler UIs, easier operation, and more automation out of the box are just some of the promises vendor OpenNebula uses to attract customers – and OpenNebula actually delivers on most of its promises.

Automated installation of OpenNebula (Figure 5) with all the required components is an easy and quick process – at least if you get your planning right up front, including the web interface, which is so important for genuine cloud computing, as well as any software that might be required in the background, such as Ceph. Basic SDN functionality with OpenDaylight or Open Virtual Network (OVN) is now available, as is comprehensive handling of storage volumes. The provider garnishes the whole thing with neat administration of VM templates and an intuitive web

interface, including multitenant capability. All told, OpenNebula is a genuine alternative, especially for smaller cloud setups.

Some people might wonder why OpenNebula shouldn't replace OpenStack completely. As usual, the devil is in the details. In some ways, OpenNebula is what many enterprises once hoped OpenStack would be – a virtualization solution that is not overly complicated and offers basic cloud functionality, automation, and self-service.

OpenStack has never been able to live up to this claim, if only because many large telcos and other companies had a say in the fate of the project and sometimes unrestrainedly asserted their own interests. The result was the kind of complexity that exceeds the capabilities and possibilities of smaller setups a priori. For most small environments, OpenStack can simply do too much. OpenNebula offers an attractive alternative, offering sufficient numbers of useful features for small environments without being overly complex.

## Conclusions

As this article shows, you rarely will have to resort to OpenStack. OpenStack has its sweet spot in

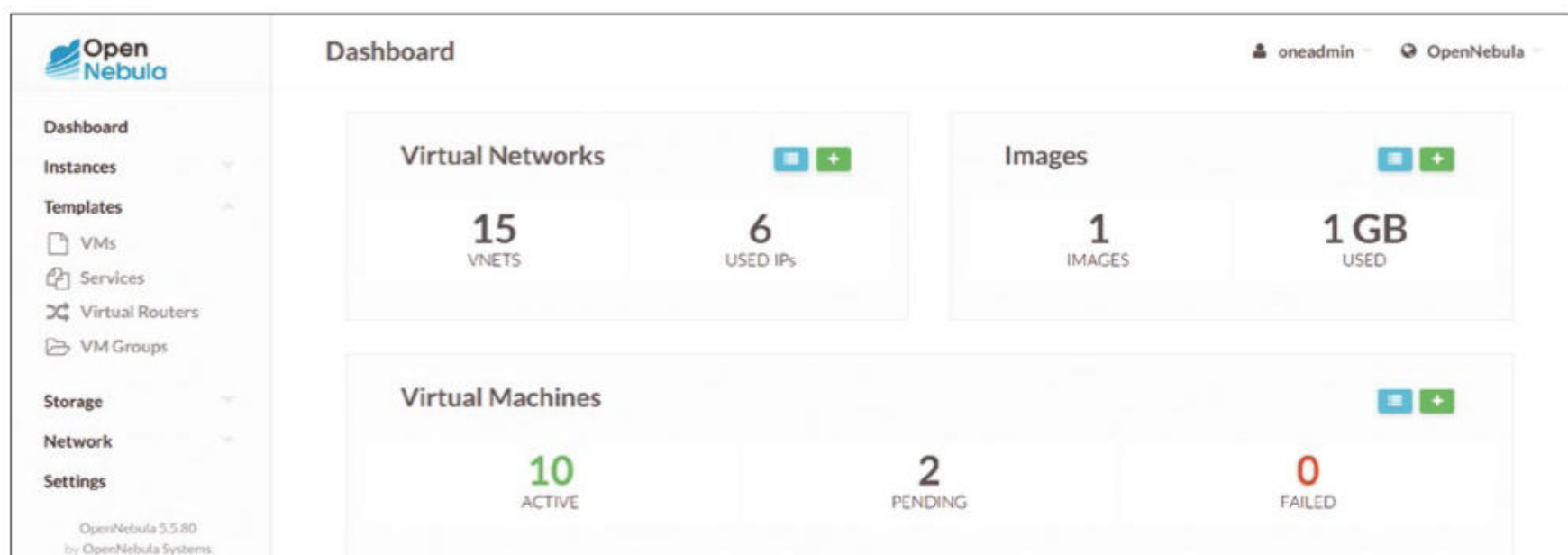
large corporations that are becoming platform providers and need to offer huge virtualization environments in the process. However, this is not the case for most smaller companies, so check before you commit. The first step in introducing a new virtualization environment should always be to document your needs accurately. If you don't need most classic cloud functionality, a DIY solution with Pacemaker, Libvirt, and Qemu is a valid option, because it is easily manageable. If you are looking for a successor to VMware, Proxmox and oVirt are logical first choices. If you really need to create a kind of cloud, but without the overburdening complexity of OpenStack, OpenNebula is likely to be the environment of choice. This example once again demonstrates the strength of the F/LOSS community, which has a ready-made solution for problems of any size. ■

### Info

[1] oVirt exporter for Prometheus: [https://github.com/czerwonk/ovirt\\_exporter](https://github.com/czerwonk/ovirt_exporter)

### The Author

Freelance journalist Martin Gerhard Loschwitz focuses primarily on topics such as OpenStack, Kubernetes, and Chef.



**Figure 5:** OpenNebula has been successfully marketing itself as a sort of anti-OpenStack for years. If you are fed up with the excessive OpenStack complexity, you will most likely find what you are looking for here. © OpenNebula



Encrypt and decrypt files with Age or Rage

# Keep It Simple

Age and Rage are the Go and Rust implementations of a simple, modern, and secure file encryption tool. By Matthias Wübbeling

**Encrypting files ensures the IT** security protection goal of confidentiality. Depending on which method you use, integrity and accountability can be ensured, as well. Asymmetric encryption is easier with Age than with GnuPG. In this article, I look at how to use Age and how you can use it in practice.

## The Role of Encryption

IT security protection goals define requirements for data or the contents of files during storage or transmission. File encryption is useful and important in many enterprise scenarios. Encryption makes sense, and not just when you need to send data over insecure channels such as the Internet, but also for data that is no longer needed in everyday life or that is already backed up and no longer needs to be kept available in

the clear. Cryptographic techniques can be used to store such data confidentially and verify its integrity on recovery.

Encrypted backups or routinely encrypted older files in an archive primarily provide protection against a potential attacker copying large volumes of information and subsequently publishing or selling this information to other market players. Of course, this does not protect you against ransomware infestation. Also, the assigned private key or password used for decryption should not simply be stored on the hard drive of your computer or server.

## Not Always GnuPG

Many distributors use GnuPG to sign their packages and distribute the public key accordingly. Therefore, most distributions are capable of

encrypting or signing files out of the box. At the command line, you can easily sign a file with a private key stored in the keychain:

```
gpg --detach-sign -o sig.gpg secret.txt
```

After doing so, the file `sig.gpg` with the signature can be forwarded to the recipient. With your public key, it is easy to check whether the file has been modified since it was signed. This process ensures integrity. Encrypting a file with GnuPG works in the same way, but now you need the recipient's public key or, in the case of multiple recipients, all of the corresponding public keys. You would need to run the following command to encrypt a file named `secret.txt`:

```
gpg --encrypt --armor ?  
-r recipient@admin-magazine.com ?  
secret.txt
```

Because GnuPG offers many other possibilities besides plain vanilla



encryption, it is not always the best choice for straightforward use. Although the software has sensibly chosen default values, they are not always transparent, and the many options certainly give users scope for errors. If you want to exchange files easily and securely on the basis of asymmetric keys, nobody is forcing you to use GnuPG.

## Easy and Quick with Age

Age is the acronym for “Actually good encryption,” and that is what the developer promises. Age is a small tool implemented in Go [1] that supports all the classic operating systems. The first stable version 1.0 was released just over a year ago. Alternatively, you can use the implementation in Rust, named Rage [2]. The Go version is well ahead in terms of performance, though, especially with large volumes of data.

Age has a clearly defined feature set with only a few configuration options, which makes both the program and the API very easy to use, virtually eliminating configuration errors by the user. In the remainder of this article, I look exclusively at how to use the program on Linux.

Age supports different file recipients, who are selected by their public keys, which you specify in each case with the `-R` argument during encryption. After the install, which is very easy because the packages exist for popular Linux distributions, you need to create a new private key with the command:

```
age-keygen > private.key
```

You can copy the public key from the output or take it from the commented lines in the `private.key` file; just ignore the hint about the potential risk posed by the file’s insecure access rights in this example. In production, you would use `umask` up front to adjust the permissions of the new file. If you only want to output the public key at the command line for the private key in the file, use the command:

```
age-keygen -y private.key
```

The keys generated are in the form of Base32-formatted X25519 identities (i.e., they are based on elliptic curves). In addition to Age’s own format keys in version 1, Age also accepts SSH keys for encryption, which can be for the same 25519 elliptic curve (`ssh-ed25519`) or RSA keys (`ssh-rsa`).

The use of SSH keys may seem unusual at first glance, but it is perfectly suitable and very practical because these keys are already distributed in many places. Especially in the developer environment, you will find many public SSH keys in profiles on GitHub or GitLab. Unlike GnuPG, however, when using Age, you do not have a web of trust to verify the keys or to establish trust in the keys.

If you want to make your data available to several recipients on a regular basis, you can store the recipients in a recipient file. The different key types of the recipients can be combined in a single file. After adding the public keys of the recipients to the `recipient.txt` file – one key per line in the normal way – encrypt your data:

```
echo Admin Magazine | 2
age -R recipient.txt -a
```

The `-a` sends the output to an ASCII wrapper, with which you are probably already familiar from GnuPG. With the command

```
echo Admin Magazine | 2
age -R recipient.txt -a | 2
age -d -i private.key
```

you can test decryption directly.

## Authenticated Data

Even though Age can be used for file encryption in a very simple way, it is particularly useful in scenarios where information is encrypted or decrypted as a data stream. By default, Age processes data from standard input and returns the results on standard output. Age’s natural habitat is

therefore the command line, scripts, or cronjobs.

Age not only ensures the confidentiality of the data, but also its authenticity and integrity. During decryption, the tool immediately checks the integrated Message Authentication Code (MAC). The principle known as “authenticated encryption with associated data” (AEAD) checks for possible changes to the ciphertext for each block, preventing various attacks on the encryption or the integrity of the data in the process. Unlike GnuPG, however, the files cannot be cryptographically signed, and Age does not support attribution to an author through a signature.

## Conclusions

Age is a simple alternative to GnuPG that lets you encrypt and decrypt data asymmetrically, easily, and reliably. The clear design and the deliberate omission of options for configuring the encryption method help ensure secure use for everyday tasks. Thanks to support for different key types, you can also use the widespread SSH keys of your recipients. ■

---

### Info

[1] Age Go implementation:

[<https://github.com/FiloSottile/age>]

[2] Rage Rust variant:

[<https://github.com/str4d/rage>]

---

### The Author

Dr. Matthias Wübbeling is an IT security enthusiast, scientist, author, consultant, and speaker. As a Lecturer at the University of Bonn in Germany and Researcher at Fraunhofer FKIE, he works on projects in network security, IT security awareness, and protection against account takeover and identity theft. He is the CEO of the university spin-off Identeco, which keeps a leaked identity database to protect employee and customer accounts against identity fraud. As a practitioner, he supports the German Informatics Society (GI), administering computer systems and service back ends. He has published more than 100 articles on IT security and administration.



## Attackers, defenders, and Windows Subsystem for Linux

# Open House

Several tactics, techniques, and procedures circulating among cybercriminals exploit Windows Subsystem for Linux as a gateway. We look at how WSL can be misused and some appropriate protections. By Akshat Pradhan

**As a compatibility layer**, the Windows Subsystem for Linux (WSL) allows Linux binaries to run directly on Windows without any modifications. Users can call processes in Linux from Windows and vice versa with WSL, accessing files on both operating systems, sharing environment variables, and linking different commands.

Two WSL versions [1] have significantly different architectures: WSL 1 makes use of a translation layer that implements Linux system calls on top of the Windows kernel and can be achieved on minimal Pico processes and providers (1xss.sys and 1xcore.sys) managed by a kernel mode driver. On its WSL blog, Microsoft provides more details on the role and history of the Pico processes [2]. In WSL 2, on the other hand, the source code of the Linux kernel is executed in a virtual machine, sized dynamically by Windows depending on the utilization level [3].

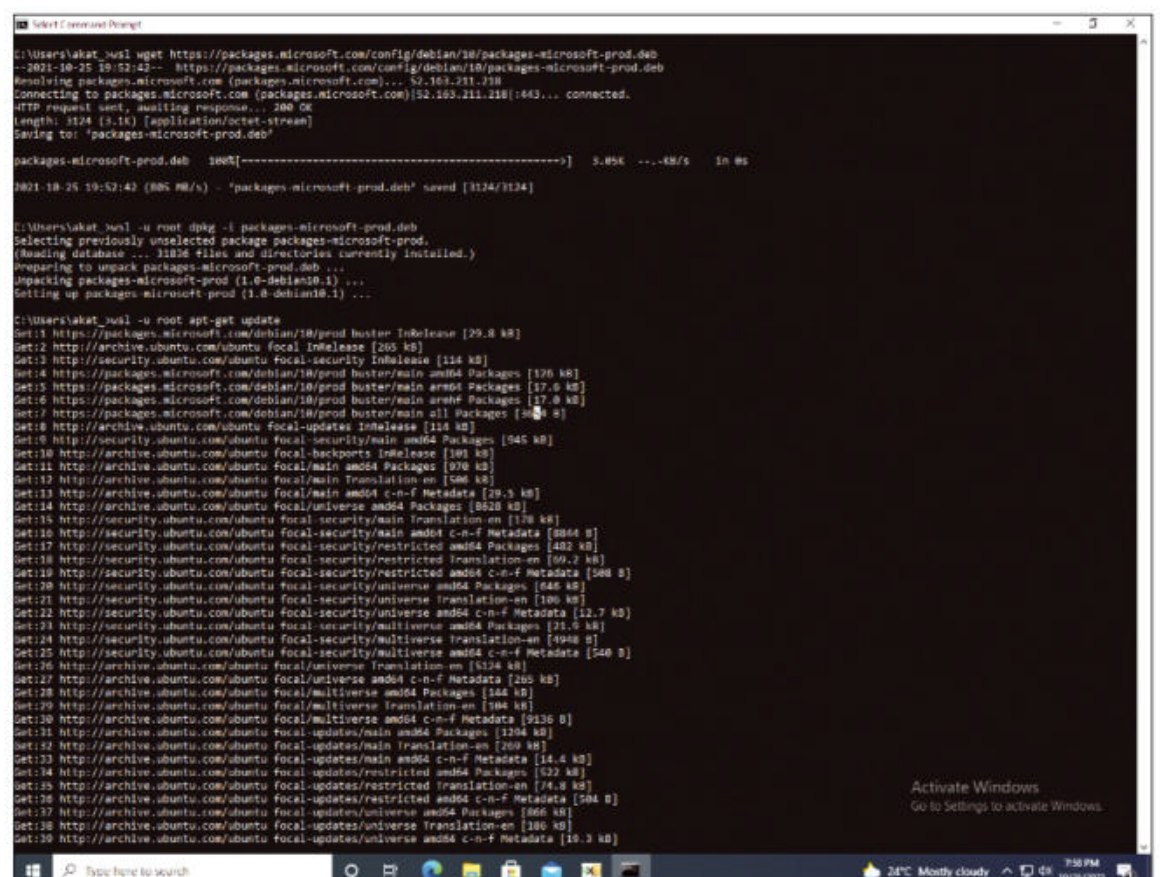
WSL is still in its early stages, but Microsoft is actively developing the project and adding additional features, such as GUI support for a fully integrated desktop experience [4]. The stated goal of WSL is to enable users to use their favorite Linux tools

on Windows. However, WSL can also be misused for attacks. To do so, cybercriminals resort to various tactics, techniques, and procedures (TTPs).

## TTP 1: Tools

Attackers bypass the requirement to enter a sudo password by passing the `-u root` argument to `ws1.exe`, making it far easier to download and deploy

arbitrary tools to run or create payloads. Cybercriminals also can add repositories of hacking distributions to deploy tools with a package installer. In a simple example of this technique, I use PowerShell, which interestingly can also be installed on Linux ([Figure 1](#)). What's more, you can easily download the tool from Microsoft's own repository. The obvious problem with this technology is that `pwsh` runs on



**Figure 1: PowerShell runs on Linux, albeit with missing cmdlets, among other things.**



Linux. Accordingly, some cmdlets are not available – for example, those for the Common Information Model (CIM) and Windows Management Instrumentation (WMI). However, attackers can execute remote commands from WSL with PowerShell or use PowerShell over SSH to access Windows-specific cmdlets, allowing them to bypass PowerShell script logging.

However, the true strength of this attack technique is shown elsewhere: Hackers can develop modular loaders that are located entirely in WSL and then exploit the interoperability to run the Windows modules they need to achieve their goals. In fact, the loader completely disappears off the radar of most security solutions. For several years now, cyberattacks have been popping up time and again that take advantage of precisely this fact [5]: yet another tool in the attackers' arsenal that cyber defenders need to consider.

Containing this TTP is often difficult because installing utilities is basically a legitimate use case. Nevertheless, potentially malicious activity can be identified, for example, by deploying perimeter solutions with an anomaly detection function that includes identifying access to hacking repositories and verifying that

the host in question has a legitimate use case for it.

## TTP 2: Injecting an Attack Distro

Attackers have two options for installing their own distributions in WSL: They can either import a tarball or install the distribution directly from the Microsoft Store. Basically, any Linux variant will do; it does not have to come from the official Microsoft Store [6]. Alternatively, the attacker can create their own Linux derivative for WSL with known threats and tools.

The command

```
PS> wsl --import <Distro>\<Install location> <File>
```

binds a distribution into WSL. You can currently obtain a whole range of Linux games from the Microsoft Store (see the box “Distributions in the Microsoft Store”). Noticeably, the popular penetration testing distribution Kali Linux is also in the store, which means that any Windows computer with WSL in its IT environment can potentially be turned into a Kali computer. Thanks to Win-KeX, Kali offers a full desktop interface

on WSL 2 [7].

Additionally, Microsoft has released a preview of WSL 2 support for running Linux GUI applications [8].

Mitigating such an attack would rely largely on system policies being enabled that prevent the features required to install WSL. If users are allowed to use WSL, you can rely on command-line identifiers to track down potential installation activity.

## TTP 3: Redirecting Execution

The Windows doskey.exe [9] utility calls the command-line history, edits the command line, and creates macros. What is exciting here is that macros take precedence over the PATH system variable search, which is why cybercriminals use it to disrupt the execution process.

Macros are instance-bound, however. To make them persist across all CMD instances, you need to create a macro definition file and add it to the AutoRun value under one of two keys, depending on the version of Windows you are using:

```
HKEY_CURRENT_USER\Software\Microsoft\
  Command Processor
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\
  Command Processor
```

Alternatively, macros can be exported with the command:

```
PS> doskey.exe /macros > <file>
```

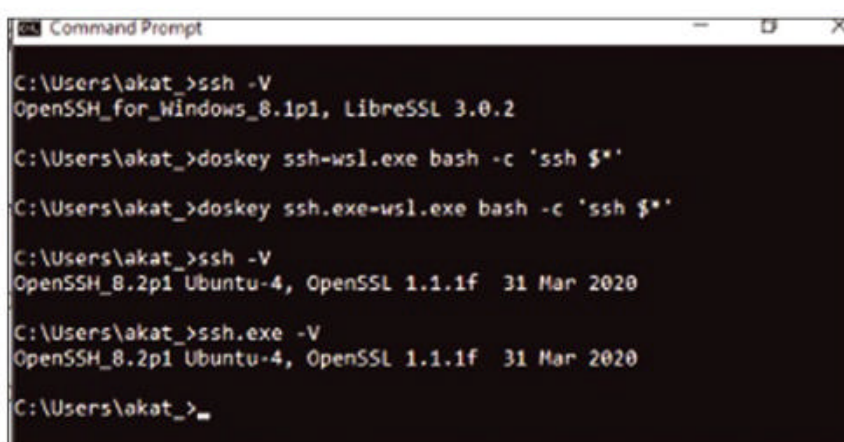
Redirecting execution to Linux can lead to persistence over patched binaries.

In a scenario of this type with SSH redirection, Windows uses OpenSSH. The required client is preinstalled in Windows 10 [10]. Even in the absence of SSH, the attacker can create an SSH macro and redirect it to WSL, as shown in Figure 2. It then becomes persistent by exporting the macros and making the necessary registry changes.

Linux has several SSH backdoors [11]. Popular backdoors include the universal SSH backdoor, which patches SSH and its associated libraries to allow SSH logins with a password set by the attacker. The patched SSH binary also records the usernames, passwords, and IP addresses of all incoming and outgoing SSH requests as plain text in logfiles. As soon as a Windows user logs in to a remote computer by SSH, their credentials are exposed and logged in WSL. Then, attackers use those credentials for lateral movements.

### Distributions in the Microsoft Store

- Debian GNU/Linux
- Fedora Remix for WSL
- Kali Linux
- Ubuntu 22.04 LTS, 20.04, 20.04 ARM, 18.04, 18.04 ARM, 16.04
- SUSE Linux Enterprise Server 15 SP3, 15 SP2, 12
- openSUSE Tumbleweed, Leap 15.3, Leap 15.2
- Oracle Linux 8.5, 7.9



**Figure 2:** Even without SSH, macros can be created and redirected to WSL.

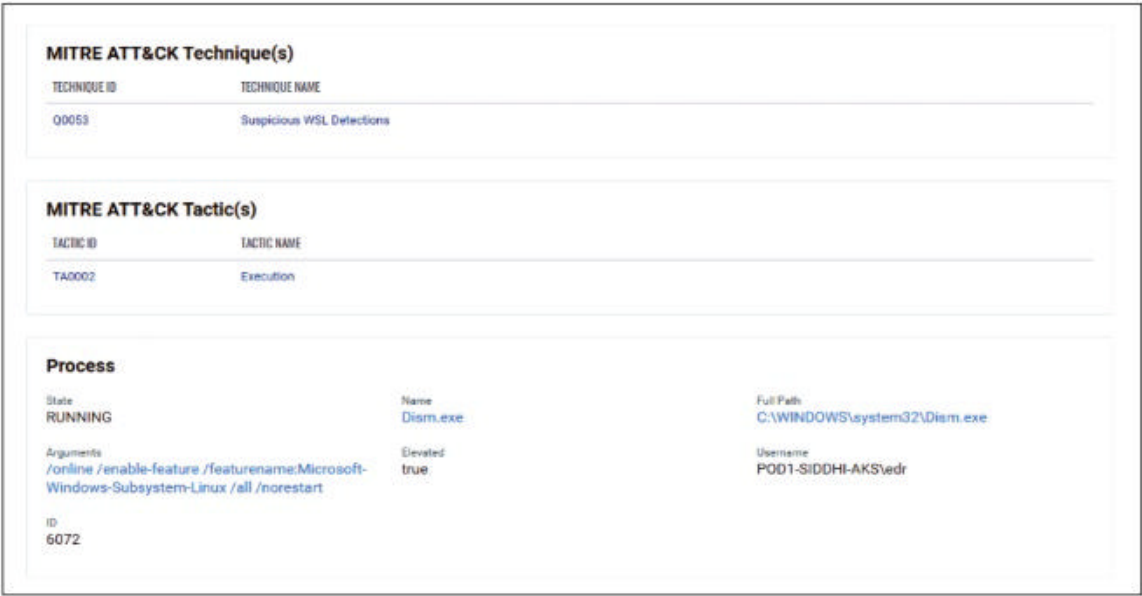


Figure 3: Endpoint Detection and Response (EDR) can help mitigate the threat of WSL vulnerabilities.

Existing security solutions should be designed to respond to the use of doskey.exe and the corresponding changes in the registry. However, a patched binary is not so easy to detect right away and would require the defender to run a check of the installed components (rpm -Va) within the WSL distributions. Another method is to use popular community tools like Rootkit Hunter to identify potentially malicious code.

Threat Hunting in WSL

The use of WSL in an environment generally looks suspicious if it is not one of the usual developer tools. You will want to monitor the environment for command lines containing wsl.exe and bash.exe. Additionally, DISM or PowerShell used to enable WSL or virtualization features can indicate unfriendly behavior. You can mitigate WSL threats with optional feature policies, and you

might want to disable virtualization and WSL with the PowerShell cmdlet

```
Disable-WindowsOptionalFeature
```

or the DISM utility. You also can hide the enable or disable Windows Features task by setting the NoWindowsFeatures value in the registry paths to 1:

```
HKEY_CURRENT_USER\Software\Microsoft\Windows\CurrentVersion\Policies\Programs
HKEY_LOCAL_MACHINE\Software\Microsoft\Windows\CurrentVersion\Policies\Programs
NoWindowsFeatures 1
```

Finally, logging and auditing processes with tools such as Endpoint Detection and Response (Figure 3) also reduces potential risks.

Conclusions

The Windows Subsystem for Linux is a useful technology designed to improve productivity by integrating applications

and utilities from various distributions into the Windows environment. However, such a large-scale project inevitably affects safety. Because attackers are focusing on WSL in their search for new TTPs, defenders need to establish appropriate protections.

Info

- [1] WSL architectures: [https://docs.microsoft.com/en-us/windows/wsl/compare-versions#whats-new-in-wsl-2]
- [2] Pico processes: [https://docs.microsoft.com/en-in/archive/blogs/wsl/pico-process-overview]
- [3] WSL 2 source code of the Linux kernel: [https://github.com/microsoft/WSL2-Linux-Kernel]
- [4] Linux GUI apps in WSL: [https://docs.microsoft.com/en-us/windows/wsl/tutorials/gui-apps]
- [5] Bashware attack on WSL: [https://www.trendmicro.com/vinfo/us/security/news/cybercrime-and-digital-threats/bashware-attack-targets-windows-system-for-linux-wsl]
- [6] Custom Linux distributions for WSL: [https://docs.microsoft.com/en-us/windows/wsl/build-custom-distro]
- [7] Kali in WSL 2: [https://www.kali.org/docs/wsl/win-kex/]
- [8] Preview of Linux GUI applications: [https://docs.microsoft.com/en-us/windows/wsl/tutorials/gui-apps]
- [9] Doskey documentation: [https://docs.microsoft.com/en-us/windows-server/administration/windows-commands/doskey]
- [10] OpenSSH: [https://docs.microsoft.com/en-us/windows/terminal/tutorials/ssh]
- [11] SSH backdoors: [https://blogs.juniper.net/en-us/threat-research/linux-servers-hijacked-to-implant-ssh-backdoor]





**2021**  
Archives  
Available  
Now!

# **CLEAR OFF YOUR BOOKSHELF WITH DIGITAL ARCHIVES**

Complete your collection of *Linux Magazine* and *ADMIN Network & Security* with our Digital Archive Bundles.

You get a full year of issues in PDF format to access at any time from any device.

<https://bit.ly/archive-bundle>





Network monitoring with Zeek

# Light into Darkness

Zeek offers an arsenal of scripts for monitoring popular network protocols and comes with its own policy scripting language for customization. By Matthias Wübbeling

If you want to know what is happening on your network, the only way is to look at the connections between devices and to endpoints on the Internet. Popular tools such as tcpdump and Wireshark are useful for occasional analysis, but for permanent network monitoring and as an alternative to intrusion detection systems, Zeek is a very interesting tool.

## Network Monitoring

Keeping track of device activity on the network is a routine task for IT administrators. Network monitoring is a large market comprising various tools, and vendors outdo each other with feature set claims, especially in the area of event processing and manual analysis.

Zeek, the first version of which was released back in 1999 (known as Bro

at that time) [1], is a kind of hidden champion. The declared objective was to develop a tool for monitoring large volumes of data with a simple option for analyzing network traffic with self-programmed scripts, known as policy scripts. The name change to Zeek (think “seek”) didn’t happen for another 20 years or so.

Zeek offers an extensive arsenal of scripts for monitoring the popular network protocols and writing the monitoring results to various logfiles on your hard drive. From there, you can integrate the files into your existing log management or your installed security information and event management (SIEM) solution. The log data is compressed and archived at regular intervals, which is an effective way to save disc space, especially on busy networks.

## Installation

Normally I use Docker when I try out software. Unfortunately, the Zeek developers do not provide their own Docker images, so my options were to find an alternative provider or create an image myself from the Dockerfiles they provided. You will need to allow some time for the Zeek build process.

The OpenSUSE Build Service is a faster approach. Developers offer ready-made packages there for the classic Linux distributions [2]. With your distribution’s package manager, you can mount the required repository and install Zeek in the usual way. Of course, you can run your distribution in its own Docker container, too. To do this, start an Ubuntu container with the command:



```
docker run -ti --net=host 2
--name=zeek ubuntu:latest 2
/bin/bash
```

The `--net=host` argument gives the container direct access to the host system's network interfaces, which it will need to read all the traffic on the interface. The `--name=zeek` argument explicitly names the container and makes connecting with the container easier later on. Before you follow the instructions for installing Zeek [2], first use the command

```
apt update && apt install 2
-y sudo curl gnupg vim net-tools
```

to install the dependencies in the running Ubuntu container.

## Configuration and Start-Up

Fortunately, you do not need to configure Zeek extensively before using it for the first time: Just specify the name of the network interface on which you want to monitor the traffic. Of course, running the server with Zeek on a mirror port of your router or switch is a good idea. All data packets that are transmitted are also transferred to the mirror port. Assuming the interface on this port is named *eth1*, you can configure the interface and the host name in the respective line of the `/opt/zeek/etc/node.cfg` file:

```
[zeek]
type=standalone
host=localhost
interface=eth1
```

Now run the management tool `zeekctl`, install the supplied policies, and launch Zeek with:

```
/opt/zeek/bin/zeekctl
[ZeekControl] > deploy
```

The output should show the command workflow. If you see an error message, you can use the Zeek `diag` command to discover why. Check the name of the network interface and make sure a process is not

already listening in the background on TCP port 47760, which is used by Zeek.

Now type `exit` to quit the management tool and begin to generate network traffic on the host. You can watch new logfiles appearing in the `/opt/zeek/logs/current` folder. The first file created here is `conn.log`, which simply lists all network connections. The list grows quickly, and files `dns.log`, `dhcp.log`, and `ntp.log` give you an initial idea of which logs Zeek pre-filters for you in the background.

If you run a DHCP server on your network and after some time use `cat` to display the content of the `dhcp.log` file, you will see requests from different network devices and the assigned IP addresses. If you open a web page, for example, with the

```
curl https://www.admin-magazine.com
```

command, you will then see an entry in both the `dns.log` and the `http.log` files. However, because Zeek cannot resolve TLS connections, you will find most calls to web pages in the `ssl.log` file.

If you let Zeek run for a little while longer during your day, you will notice that it keeps archiving the files. Each day will have a folder with the matching date in `/opt/zeek/logs`. Look for the `weird.log` file in the archive subfolders, also with the corresponding dates and compressed by `gzip`. If you unzip these files and take a look at the content, you will see that these logfiles contain unusual events, such as the reuse of existing connections or errors in UDP and TCP checksums.

## Converting Logs to JSON

To make the log data easier to handle, you can change the tab-delimited logging to a more modern JSON format. To adapt the configuration, add the following two lines to your `/opt/zeek/share/zeek/site/local.zeek` file:

```
# Output in JSON format
@load policy/tuning/json-logs.zeek
```

Now, with the `deploy` command used earlier, restart the Zeek process and check the format in the logfiles. For further analysis of your log data, you can also connect tools that expect input in JSON format.

## Conclusions

For administrators, reliable insights into network traffic are a must-have. They not only help you identify and analyze problems, but detect possible attackers. Zeek can already look back on more than 20 years of development, delivering a classic approach to monitoring network activity. The tool comes with its own policy scripting language [3] for customization. With its help, you can flexibly adapt your monitoring setup to suit your needs or expand the analysis options to include more network protocols, if required. ■

### Info

- [1] Paxson, V. Bro: A System for Detecting Network Intruders in Real-Time. *Computer Networks*, 1999;31(23-24):2435-2463, [<https://www.icir.org/vern/papers/bro-CN99.pdf>]
- [2] Zeek packages: [<https://software.opensuse.org/download.html?project=security:zeek&package=zeek>]
- [3] Policy scripts: [<https://docs.zeek.org/en/master/scripting/basics.html>]

### The Author

Dr. Matthias Wübbeling is an IT security enthusiast, scientist, author, consultant, and speaker. As a Lecturer at the University of Bonn in Germany and Researcher at Fraunhofer FKIE, he works on projects in network security, IT security awareness, and protection against account takeover and identity theft. He is the CEO of the university spin-off Identeco, which keeps a leaked identity database to protect employee and customer accounts against identity fraud. As a practitioner, he supports the German Informatics Society (GI), administrating computer systems and service back ends. He has published more than 100 articles on IT security and administration.



Active Directory management with NetTools

# Health Check

NetTools utilities extract information from Active Directory to help admins simplify troubleshooting and administration. By Klaus Bierschenk

**Active Directory administrators** have a massive choice of tools, starting with the integrated administration tools on Windows Server and including a variety of free and commercial programs. This toolbox can be nicely rounded off with the free NetTools: in total, more than 90 utilities that simplify troubleshooting and administration.

NetTools [1] notches some initial brownie points directly after you download its single EXE file. You don't need to install a tool palette, with no dependencies on frameworks or DLLs. However, this doesn't translate to a hodgepodge of command-line tools; instead, everything is available in a central management interface without the need for context changes. The only add-on is an INI file with work files for the current configuration and user-specific information. If `nettools.ini` does not exist, it is created at program launch time. Software can be that simple and makes installation on a domain controller less critical because it avoids any risk of trouble with components installed at the same time.

On the website you can look forward to some very good and, above all, up-to-date documentation – not necessarily a matter of course in the world of freeware. The FAQ section is helpful, especially for newcomers,

and helps you find your way around. A list of all functions and a carefully maintained blog round off the information collection. The author offers readers tips and tricks for the utilities, accompanied by plenty of examples. NetTools starts exactly where the on-board tools leave off. Therefore, it is not intended to be an alternative to existing tools, but deliberately closes functional gaps.

## Targeted Profiles

By default, you work with the toolbox in the context in which you have logged on. For more flexibility, you can create profiles that include both connection data for a specific domain controller, or even a different domain, along with the user account. A profile can be specified in the course of a specific action, which means you do not need to be interactively logged in as a domain admin to act in the context of the domain admin's authorizations. All profile information is stored in the INI file – except for the passwords. The passwords are not stored anywhere but need to be retyped each time.

Even at first glance, the tools' graphical user interface (GUI) is neat and tidy and the buttons are self-explanatory. For newcomers, it is still advisable to consult the help on the website, which

simplifies getting started with the GUI and includes instructions on how to handle the profiles.

## Integrated GUI Functions

The NetTools feature set is not just the individual tools lined up ready for use on the left side of a GUI. Useful elements are also embedded directly in the GUI (e.g., in the context menus of the objects). Taking a user object as an example, you can easily search for it by typing the name or part of it, without wildcards, in the search bar at the top of the screen. The results appear in the main window, and from this list you can select an element by right-clicking. The treasure trove of information in the context menu is a revelation for any admin.

*Last Logon*, for example, shows detailed information about the respective object: When and how often did the user log in recently? Which domain controller processed the login? Was the password entered incorrectly? When was it last changed? The *Use With* menu has even more to offer: The *Group Changes* subitem gives you information on the history of group assignments in addition to group memberships. You can easily see when a user has been removed from or added to a group, for example. The various options in



the context menu invite admins to browse and try them out.

## Finding Differences

Comparing two objects is a fairly common scenario. Staying with the user object, I will look at an example related to permissions in Active Directory. A comparison in NetTools always involves two steps: First, select an element from the list of users (or other objects), again using the context menu, this time with the menu item *Select left SD to compare* (where SD stands for security descriptor, the place where permissions are stored for a user object).

After selecting the option, the GUI remembers the object. In this example, assume you have selected the *Christa* user account. Now you can display the second object with another search. If you open the context menu again, you will see the *Compare to 'Christa' SD* item. A new window then shows the differences between the two user objects in tabular form. This view contains a column header at the top. Worth noting is the column with the asterisk (\*) header (Figure 1) that shows the results of comparisons between values in the columns as special symbols, allowing you to identify whether permissions are identical, partially identical, or not available for comparison for one of the two objects.

The developer chose this symbolism to illustrate the several possibilities. Clicking on the \* shows hints about the different symbols; what's more, you have the option to select a filtering character for the display, which means you can reduce the list to elements of the objects that are identical or precisely not identical.

## Other Options for Comparisons

Comparisons often help during troubleshooting group memberships, as well. How do two user accounts differ in terms of their group memberships? To find out, the procedure is similar to what you just saw. This time, choose *Use With | Group Compare*. The memberships are listed, and you can see directly what is going on in relation to two objects.

These examples are just a few that show how useful functions are integrated into the view and the context menu. A lot more is waiting, though. For example, you can see whether a conflict exists with accounts in another domain from the email address of a user account (e.g., as a check criterion before migrating the account). A look at not only user objects but also computer accounts in the context menu of an object is definitely very helpful if you want to discover the full potential.

## Connection Check Between DCs

Domain controllers (DCs) maintain active communication with each other. For this reason, it can be challenging when firewall rules do too much of a good thing – especially between remote sites – and important ports are closed. The resulting error patterns are difficult to interpret and usually show up in completely different places. A function integrated into NetTools tests the connectivity between domain controllers and, in the event of individual errors, displays them at the port level. Admittedly, the *Test-NetConnection* cmdlet does the same job, but it is not as nicely integrated into a GUI. All of the ports important for AD communication are already included. You also have the option of specifying individual ports that will be included in the test. Staying with the network, in terms of site topology, correct mapping of subnets, the site, and the domain controller is immensely important. Among other things, mapping ensures that computers can find their DC. If parts of an IP address range are assigned to multiple subnets, overlapping subnets occur, which can lead to issues that are difficult to track. Although this situation is not an immediate threat, unnecessary network traffic can try users' patience; after all, you

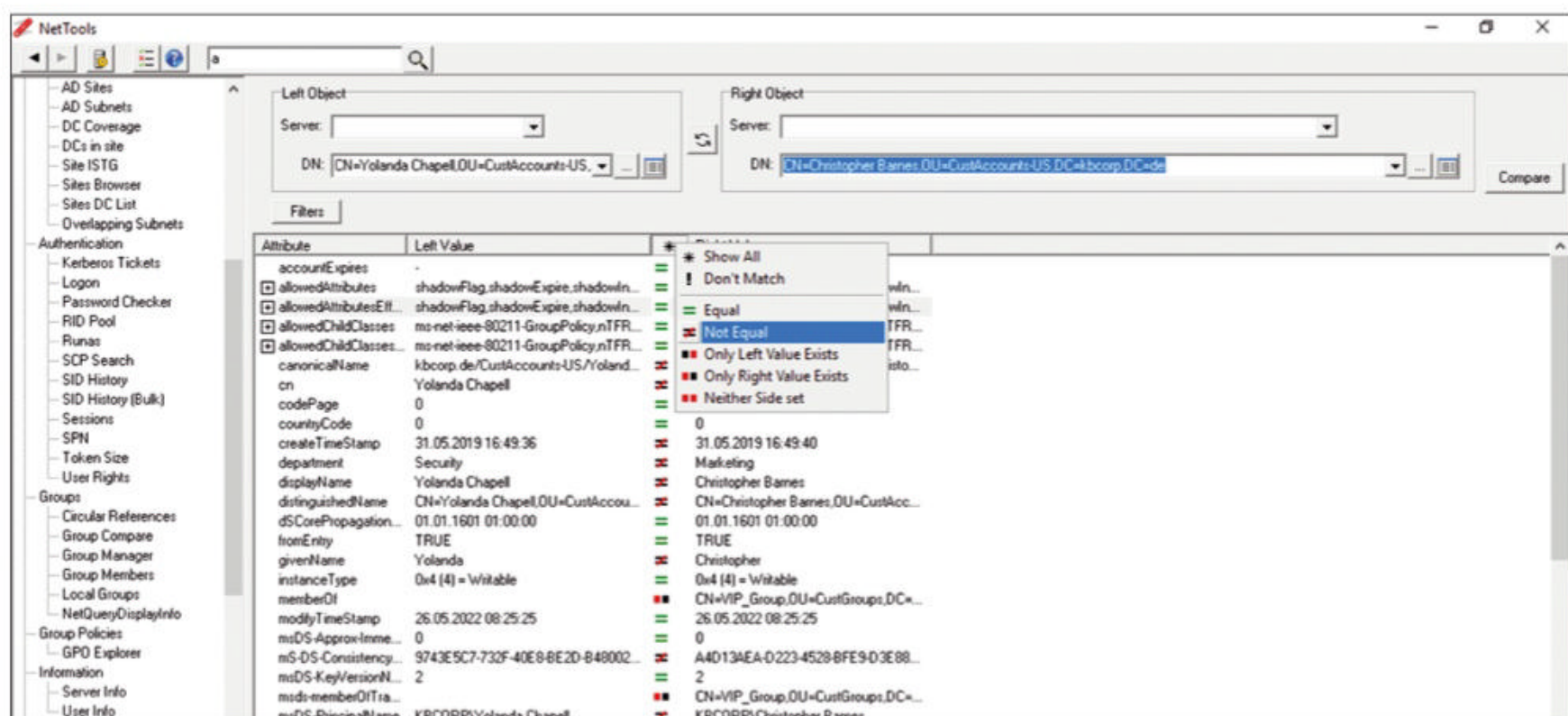


Figure 1: Two objects compared: Smart symbols help you identify differences.

probably want each computer to communicate with the closest DC. The *Overlapping Subnets* function, which can also be found in the sidebar, is where the subnets and the respective masks are calculated and overlaps are displayed. You do not have to specify a network range, because the function accesses the subnet information of the topology and calculates possible overlaps.

## Replacement for GPOTool

Some of you might remember the Windows Server 2003 Resource Kit, which included a tool named GPOTool.exe that inspected and tested group policies. However, Microsoft did not provide a replacement for this feature after the resource kit was discontinued. Luckily, NetTools has a similar function named GPO Explorer.

After launching the tool, you are taken to an overview of the state of group policy objects (GPOs), which initially appears to be similar to the Group Policy Management Console (gpmc.msc). A closer look reveals small but subtle differences. With just a few clicks, you can view assigned GPOs, output their contents by scrolling and switching between policies, and more. A *Test* button lets you test a single GPO. Alternatively, you can select the node with the domain name up front to check the entire environment by

including one or more DCs, or just a certain number of GPOs, in the test.

## Monitoring Replication

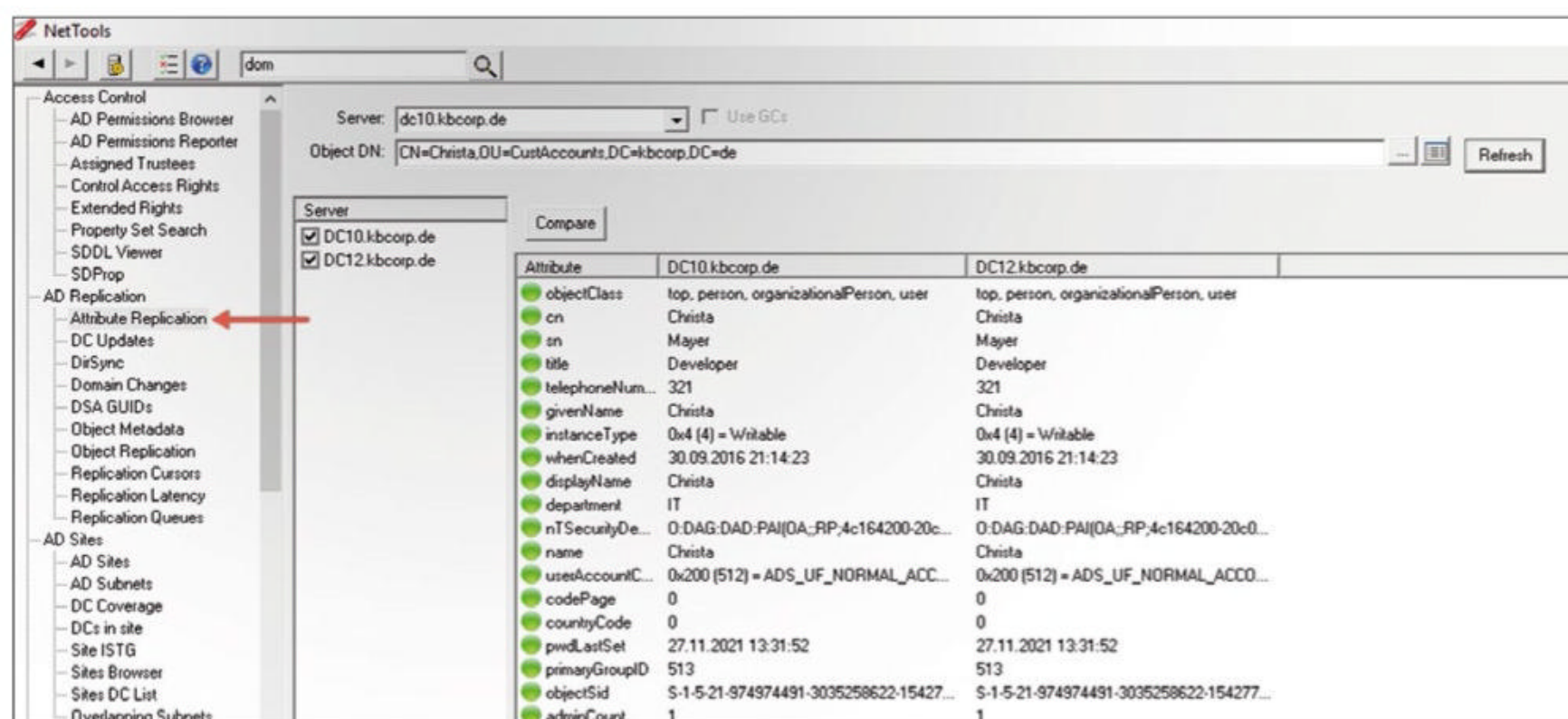
Replication, and therefore the distribution of changes in Active Directory, is one of the directory service's most important functions and is especially true for infrastructures whose sites and domain controllers provide services, although they are physically remote from each other. Unfortunately, inconsistencies in synchronization can cause very unpleasant glitches, with a wide variety of causes. The network, domain name system (DNS), or even policies can sometimes interfere. NetTools comes with a number of features in the *AD Replication* section that can help shed a light on this problem.

The monitoring functions range from health checks for a site to analyzing replication at the attribute level. The *Attribute Replication* function ([Figure 2](#)) lets you look into the details. To do this, you need to specify the distinguished name (DN) of the object to be examined. A detail window then shows the attributes of the object, along with a list of domain controllers in the left-hand section. After you decide which of the DCs need to take part in the comparison, pressing *Compare* displays the attribute content for each DC.

The *AD Replication | Domain Changes* option is somewhat simpler, but just as helpful. After specifying a domain controller, replication information is returned if a change has occurred. If you want to know how long it takes for changes in a partition to replicate, *Replication Latency* is the right place to look. To begin, you need to specify the DN of an object; it is then created and deleted again. Both write processes can be monitored in terms of time. In this way, comparisons can be made and any weak points in regional network connections can be located. Depending on the application scenario, you can use the tool that best suits each case. The current 1.31 version of the toolbox contains 10 functions that offer versatile views of replication, which is all the more valuable because the on-board tools have little to offer in this regard.

## LDAP Directory Without the Frills

NetTools enables versatile access to LDAP directories with various LDAP functions. The focus is on Active Directory, but you are not restricted to AD as long as the directory you want to access follows the LDAP API standard. An LDAP browser delivers directory information in the form of raw data, unlike the Active Directory admin tools that prettify



**Figure 2:** Help with troubleshooting: *Attribute Replication* checks whether identical object information exists on several DCs.



data here and there and validate user input before making changes. For example, the *LDAP Search* function integrated into NetTools supports SSL-based access, and even write access is possible without leaving the client.

An admin does not need any knowledge of LDAP syntax. The query criteria are created in the GUI with the use of drop-down lists (Figure 3). More than 280 predefined LDAP queries are found under *Favorites*, which is very useful. Besides illustrative material, you'll find a number of useful tools for practical admin work.

Most administrators will be really excited to see the catalog of LDAP queries. Do you need a list of groups without members? Do you want to know which GPOs were modified in the last 10 days? Predefined queries bring this information and far more

to light. New queries are quite easy to create by editing existing queries and then adding the modified versions to the catalog. The ability to import and export LDAP statements via the clipboard rounds off the feature set, leaving virtually no wishes unfulfilled.

The option to write the output to files means that you are not restricted to the GUI in the tool but can process the info downstream (e.g., in Excel).

## Advanced View 2.0

Like the standard Users and Computers console (*dsa.msc*), you can also use the various NetTools functions to display object properties. You are probably aware of the advanced display variant in *dsa.msc* if you enable the *Advanced Features* in the View menu. The Properties dialog for a user account still provides the basic info in this case, but

you can also display other content such as the object attributes.

Much more awaits you in NetTools. The information is integrated in the Properties dialog and varies as a function of the type of object you are viewing. For a user account, for example, you also have the option of displaying the last password change date, the Fine Grain Password Policy if applicable to the user, or which domain controller the user logged in with and how often they did so. This information can help narrow down individual issues.

## Checking Groups for Changes

Previously you looked at the group membership history starting with a user object. This view also works in the other direction – that is, starting

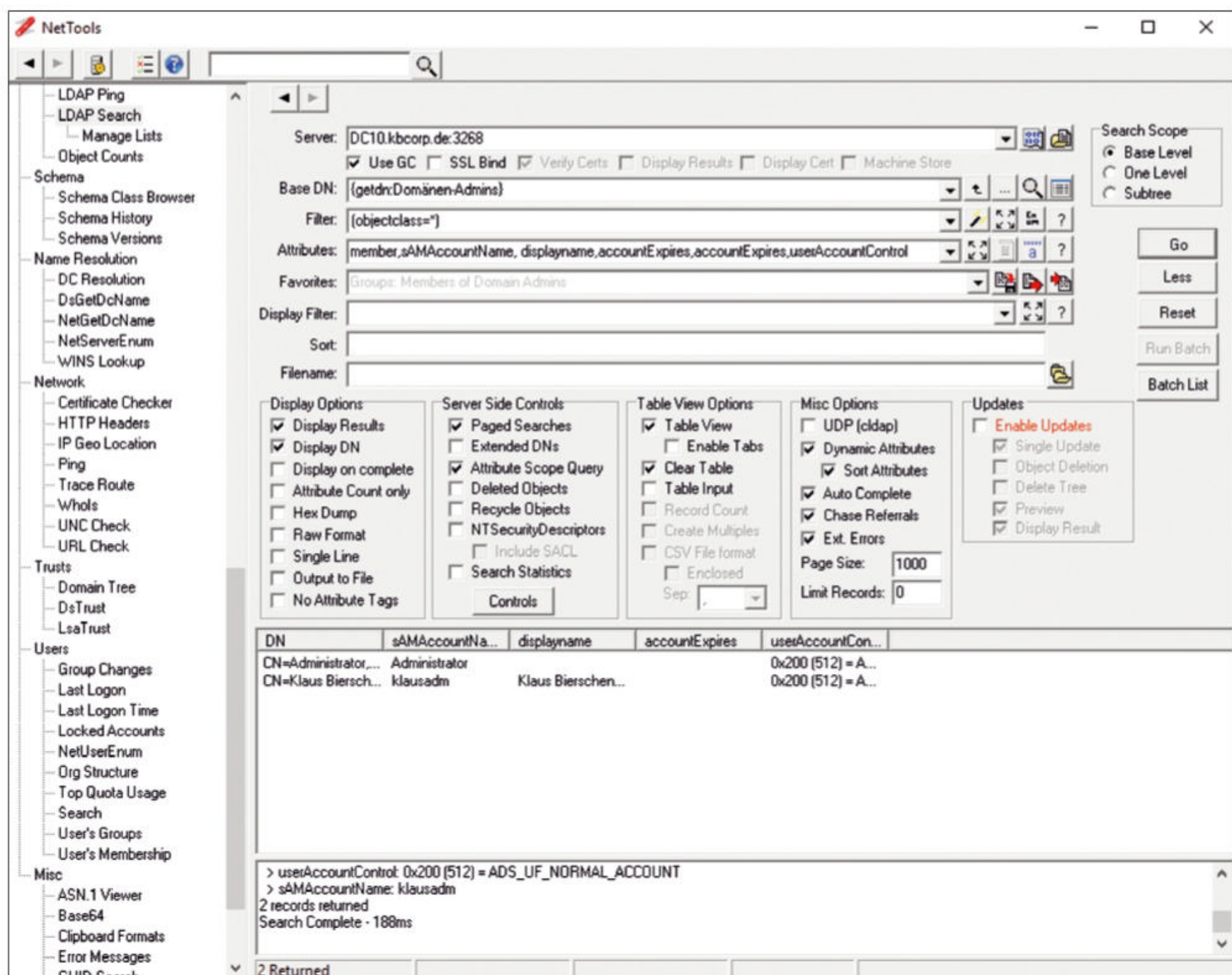


Figure 3: Like the cockpit of an airplane, the LDAP client comes with a plethora of settings.



with a group. To find out what changes have been made, for example, in the Domain Admins group, you need to view the group in one of the NetTools views. The easiest way to do this is to use the search box at the top of the toolbar. Uncheck *Return Users Only*; otherwise, the search will fail. You also need to pay attention to the spelling of the group names. In this case, one AD domain controller was a German server; accordingly, the group name was *Domänen-Admins* (Figure 4), but this is just a side note.

To find out what activities have taken place in this group, navigate to the context menu by right-clicking and then selecting *AD Properties*. In the Properties, focus on the *Members* tab to view the current members of the group. Unlike in the Users and Computers console, the *Removals* field here already displays the latest accounts to have been removed from the group. You can go into even more detail by pressing the *Changes* button. Another window provides data on when accounts were added and removed.

In this context, I need to mention once again the LDAP catalog, with its diverse query options. It also includes queries relating to group administration, especially for administrative groups. Note also that the group names in the queries might need to be adjusted to match the language of the domain controller. The LDAP queries are all based on the built-in group names in English.

## Schema - Under the Hood

Administrators are pretty sure to be familiar with the history and updates of their schemas; after all, this is “their” Active Directory and the info is essential. What happens, though, if you need to investigate the schema, versions, and update history of someone else’s AD? The functions in the *Schema* category can be helpful. The *Schema Versions* item helps you find out what changes have been made to the schema, for both the Active Directory and Exchange schemas. *Schema History* takes this one step further and shows you the schema history (i.e., when which update was made to the schema).

In addition to the quite extensive NetTools functions, a number of minor functions can brighten up your everyday life as an administrator. For example, the unique security identifier (SID) of an Active Directory object or the name of a SID is displayed by the SID converter. You have probably viewed the event log of a domain controller that revealed a problem with a user that had a SID of *123 et cetera*, but who is this user? Of course, you have PowerShell cmdlets or the `wmic useraccount` command with the `get sid` or `get name` parameters, but it’s good to know that NetTools has a converter, especially if you are working in the NetTools GUI and with identities anyway.

Speaking of converters: The *Time Converter* lets you convert a time to

some other format, including 64-bit. This function certainly is not used on a daily basis, but if you ever need it, it is reassuring to know it can be found in the toolbox.

## Conclusions

When you work with NetTools, it becomes clear what a wealth of information Active Directory offers. Experienced administrators who have managed without NetTools so far will have loads of fun with it. You will be surprised to see the bouquet of utilities that squeeze the last snippet of information out of Active Directory.

Great attention has been paid to detail in the functions themselves, which definitely helps in daily operations. Despite all the praise, people are bound to look for things that could be improved, and for the many admins who like to avoid mouse pushing, a call for a command-line interface might be justified, considering NetTools sees itself as a purely GUI-based tool. ■

### Info

[1] NetTools: [\[https://nettools.net\]](https://nettools.net)

### Author

Klaus Bierschenk is an Executive Consultant at CGI Germany, a speaker at conferences and community events, and technical author of various publications. You can find him on his blog <http://nothingbutcloud.net/>.

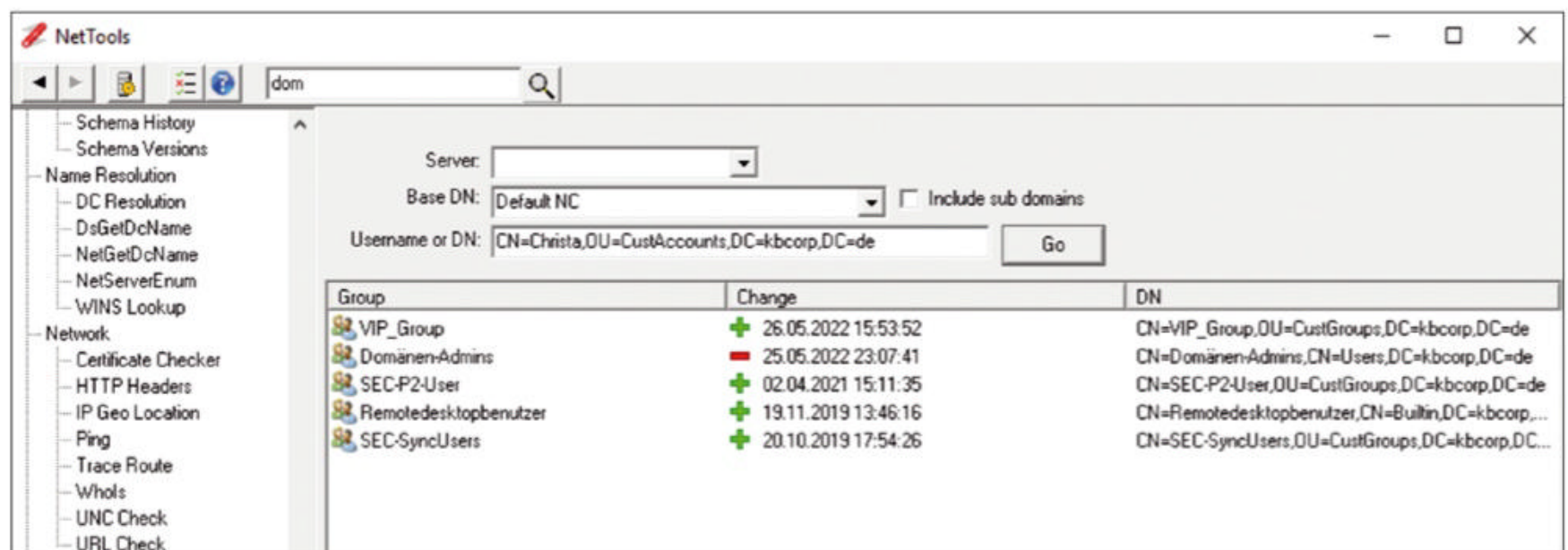


Figure 4: Group Changes lets you track the latest group changes for a user account.



# REAL SOLUTIONS *for* REAL NETWORKS



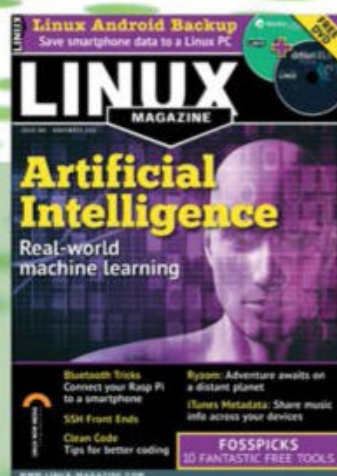
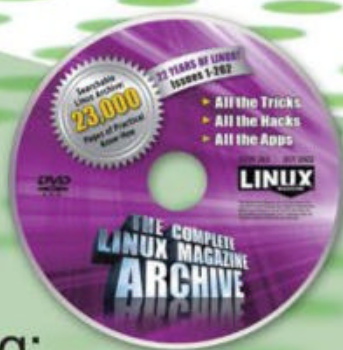
ADMIN is your source  
for technical solutions  
to real-world problems.

Improve your admin skills  
with practical articles on:

- Security
- Cloud computing
- DevOps
- HPC
- Storage and more!

**SUBSCRIBE NOW!**  
[shop.linuxnewmedia.com](http://shop.linuxnewmedia.com)

Check out  
our full catalog:  
[shop.linuxnewmedia.com](http://shop.linuxnewmedia.com)







## Samba domain controller in a heterogeneous environment

# Shake a Leg

The open source Samba service can act as an Active Directory domain controller in a heterogeneous environment. By Andreas Stolzenberger

**An Active Directory (AD) domain controller (DC)** serves as a central logon server in heterogeneous networks with Windows, Linux, and macOS clients. This task does not necessarily have to be handled by a Windows server. The open source Samba service can also act as a DC. Heterogeneous networks with servers and clients running both Linux and Windows need a centralized management server for the user directory and a standardized protocol for network shares. Windows systems naturally prefer Active Directory for this purpose, but technologies such as Kerberos and Lightweight Directory Access Protocol (LDAP) for securing user and access rights are open source. The obvious choice would seem to be the open source FreeIPA directory server. However, FreeIPA mainly targets Linux systems and user and group management. FreeIPA lacks some features needed to act as a DC that a Windows system provides over the Server Message Block (SMB) and Common

Internet File System (CIFS) protocols. Version 4 of the well-known open source Samba file service, on the other hand, provides a complete DC implementation.

The Samba project has been around for some 30 years now. It started life as a free Unix client for DEC Pathworks, which was partly based on the technology of the IBM OS/2 LAN Server and Microsoft LAN Manager. In the 1990s and early 2000s, the open source project initially fell foul of Microsoft, with repeated disputes. When Microsoft revamped its “Linux is cancer” (Steve Balmer) stance to “Microsoft loves Linux” (Satya Nadella), the software giant’s relationship to the open source project changed. Microsoft employees have been part of the Samba development team since 2011. Additionally, Microsoft has now openly documented the SMB protocol, which helps Samba developers. Since Windows Server 2003, a Samba server can become a member of an existing AD forest. However, this setup always required a Windows

server as the domain controller. In version 4, Samba itself could assume the domain controller role. The implementation also supports mixed operation of Windows and Linux servers as DCs. Of course, users with an existing Windows server infrastructure will not want to swap their systems for Samba servers. Rather, the solution is recommended for environments that use Windows, macOS, and Linux on the client side but run their server services on Linux systems. In this scenario, a Samba server serves as a central directory service for Windows, macOS, and Linux systems, as well as a file server for all.

## Installing Samba 4

A directory server hosts a whole range of services and protocols such as domain name system (DNS), Kerberos, and LDAP. Samba integrates all services to ensure they work together optimally in the Active Directory Service (ADS) DC. Other products such as FreeIPA, for example, are made up of various components such as MIT Kerberos, OpenLDAP, and Bind9. However,

Photo by Zachary Nelson on Unsplash



complete integration of the libraries into the Samba ADS package causes problems for various distributions. The Fedora and Enterprise Linux (EL) distributions rely on MIT Kerberos and ship with its packages in place. (“EL” includes all Linux distributions that are clones of Red Hat Enterprise Linux, such as Alma or Rocky Linux and CentOS Stream.) Unfortunately, this approach still does not work correctly in ADS mode. A TechPreview build of the Samba server that includes MIT Kerberos is not currently considered stable. Samba prefers the “Heimdal” implementation of Kerberos for the AD service. A Samba ADS can only be set up on Fedora and EL distributions if you use workarounds that require either third-party repositories or a Samba build from source code. For this reason, I used two Debian variants in the test setup: the current Debian 11 on a virtual machine and an ARM CubieTruck single-board computer with Armbian 5.9, which is based on Debian 10. In smaller environments, a single-board computer of this class (ARM V7 dual core, 2GB of RAM) is fine as an ADS. In preparation, the machines need a standard

Debian (Armbian) minimal setup with a static IP address and the correct configuration of `/etc/hostname` and `/etc/hosts`. The first computer with ADS also assumes the role of the DNS server. If you already have a Dynamic Host Configuration Protocol (DHCP) service running on the local area network (LAN), it must point to the IP address of the ADS as the DNS (DHCP option 6) and transmit the domain name (DHCP option 15). During the install, the primary ADS node must be able to access your existing DNS server, but you can change the DNS setup after the initial setup (Figure 1). As with all other services that use encryption, the correct system time is essential. On an ADS network, the domain controller also acts as a time source for its clients. Before the install, first make sure that the server’s system time and time zone match and that a service such as Network Time Protocol (NTP) or Chrony ensures automatic time synchronization with the Internet. This detail is especially important in a setup with an ARM single-board computer because systems like a Raspberry Pi or CubieTruck do not have a hardware clock.

If you want your Samba server to manage advanced access authorizations (access control lists, ACLs), the server’s filesystem must allow extended attributes, which are always enabled on modern Linux installations with XFS or Btrfs. This setup is now also the standard for ext4 filesystems. If you are using ext4, you can check `/boot/config-<current kernel>` to make sure the required settings are in place before you install:

```
CONFIG_EXT4_FS_POSIX_ACL=y
CONFIG_EXT4_FS_SECURITY=y
```

Now set up the required packages on your Debian server:

```
apt install samba samba-vfs-modules 2
samba-dsdb-modules
smbclient winbind libpam-winbind 2
libnss-winbind
krb5-kdc libpam-krb5 -y
```

The basic Samba installation comes with a configuration file as a template. However, the ADS setup creates a new one, so you need to remove the old one by typing

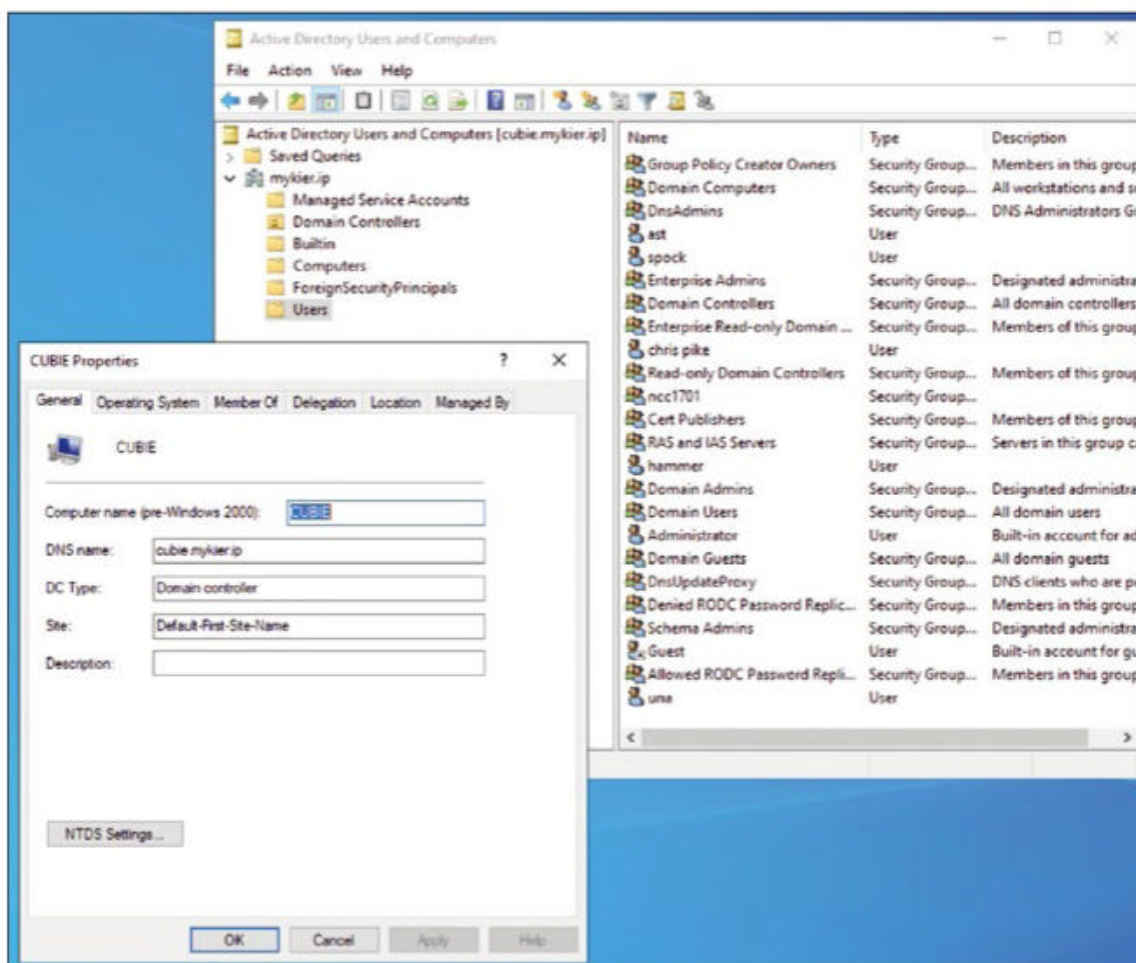
```
rm /etc/samba/smb.conf
```

before you proceed with the setup. The Samba tool then guides you through the AD setup with:

```
samba-tool domain-provision --interactive
```

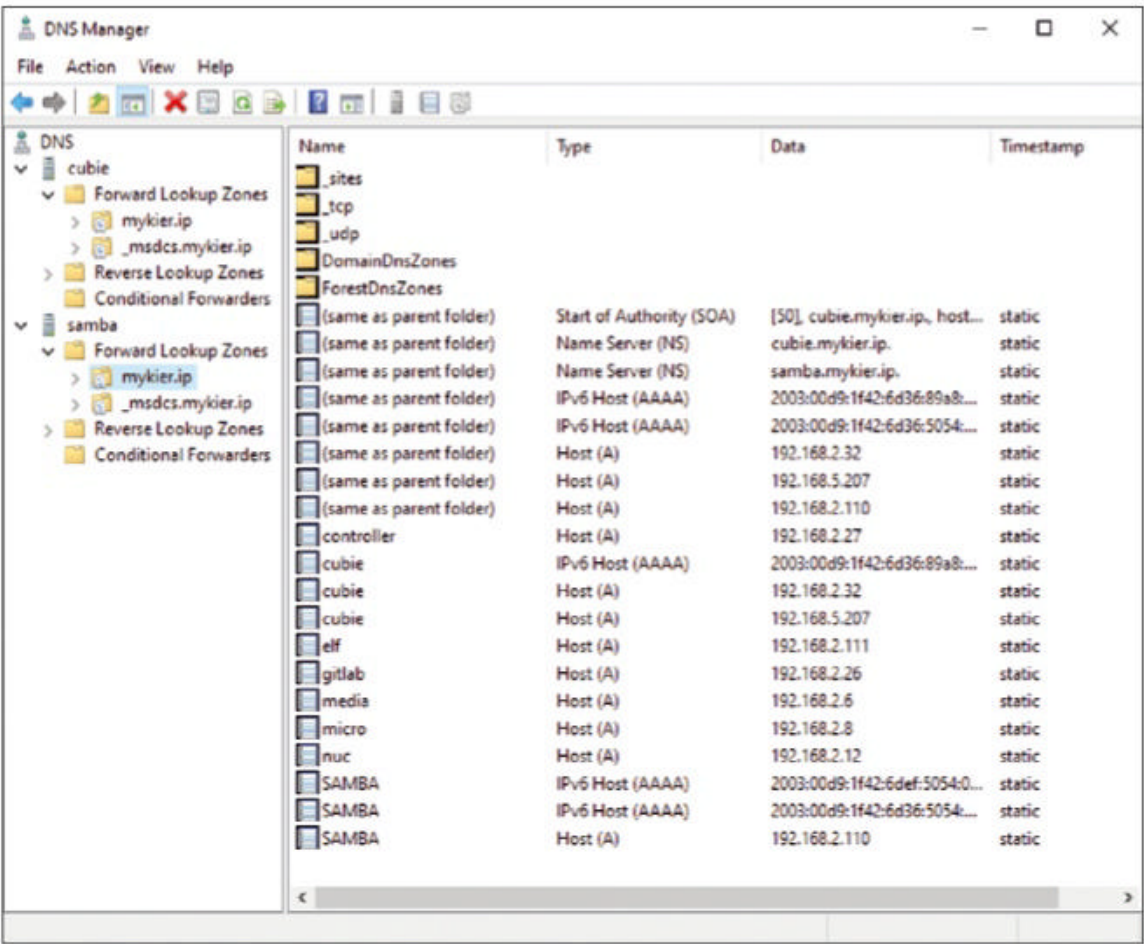
Without any further parameters, the Samba tool relies on Samba’s internal services for all required services (DNS, Kerberos). When prompted, enter the domain and realm name information. Another important item to enter is your existing DNS server as the forwarder (Figure 2). Finally, the setup creates the appropriate configuration files and a registry. The `smb.conf` file immediately creates the setup in the correct directory. The Kerberos configuration, on the other hand, needs to be copied manually to the correct directory by typing

```
cp /var/lib/samba/private/krb5.conf /etc/
```



**Figure 1:** Use the standard AD tools on Windows to customize the domain settings of the Linux DC.





**Figure 2:** The Windows DNS Manager manages the DNS zones of the forest. The Linux DC servers synchronize their settings.

Samba provides a separate binary for AD operation. To prevent the regular Samba services from running or being started accidentally, enter:

```
systemctl stop smbd nmbd winbind
systemctl disable smbd nmbd winbind
systemctl mask smbd nmbd winbind
```

and, instead, use the AD server service:

```
systemctl unmask samba-ad-dc
systemctl start samba-ad-dc
systemctl enable samba-ad-dc
```

which integrates all services.

Checking the DNS Service

Once the Samba AD server is running, it assumes the DNS role for the network. On the DC system itself, you now need to configure local DNS resolution. The Samba tool has added your existing DNS server to the configuration in `/etc/samba/smb.conf`:

```
[global]
    dns forwarder = <DNS IP address>
...
```

The local DNS resolver therefore also needs to point to the local system. Modern distributions use the `systemd-resolved` service, which detects network changes and dynamically adjusts the DNS configuration by overwriting the `/etc/resolv.conf` file as needed. This service is needed by users on clients that frequently switch LANs or establish VPN connections. The domain controller in this example, on the other hand, requires a static DNS configuration, which you can achieve by turning off the `systemd-resolved` service and unlinking the `resolv.conf` file:

```
systemctl stop systemd-resolved
systemctl disable systemd-resolved
cd /etc
unlink resolv.conf
```

In an editor of your choice, create a new `/etc/resolv.conf` with only two entries:

```
nameserver <IP address of the DC>
search <domain name>
```

Next, check that both a local and forwarded DNS request work. The command

```
host -t SRV _ldap._tcp.<Domain Name>
```

must respond with *has SRV record 0 0 389 <FQDN of the DC>*, whereas a regular DNS request to the Internet is answered by the forwarder with, for example, *www.admin-magazine.com has address <IP-Address>*.

Errors in the DNS are some of the most common causes of directory setup problems. Whatever else happens, this service must work correctly in your environment.

ADS Management at the Command Line

The Samba tool used previously can do more than just guide you interactively through the AD setup. It also acts as a command line interface (CLI) for AD management that manages users and group memberships easily. Before you can work with the Samba tool, you first need to log in as an administrator and manage the directory:

```
init Administrator
samba-tool user create hammer IamAgenius
samba-tool group addmembers ncc1701 hammer
```

The practical thing about the CLI method is that standard tasks can be handled by script. A simple Bash script could read the DNS configuration of an existing `/etc/hosts` file and transfer these hosts to the AD-integrated DNS with `samba-tool dns add`.

Adding Clients

Before you add a client to AD, you need to verify that its DNS configuration points to the DC server. This example deploys Windows 10 Enterprise and Windows Server 2022. Windows 10 Enterprise version 21H2 does not have the *Join a domain* button that previous Windows versions had in system Settings. You need to work your way through the following menus in Windows: *Start | Settings | System | About | Rename this PC (advanced)*. In the System Properties dialog, press *Change* and enter the name of your domain in



the *Member of Domain* box. Now Windows asks for a user with the domain join permission. Simply enter *Administrator* with the appropriate password and the system will become part of the domain. After restarting, you can log in to the Windows client as one of the previously created users.

If you want to use the GUI tools for AD administration on this Windows client, you can download them from Microsoft [1]. Windows Server 2022 provides the appropriate tools. The DS-Client setup works in a similarly simple way on Linux systems. Different Kerberos versions for the clients have no compatibility issues as long as they use AD for login verification and do not run a local Samba service. EL and Fedora systems can join the domain just like Debian systems. On a Fedora 36 workstation, the first step is to install the necessary packages:

```
dnf install realmd sssd oddjob 2
oddjob-mkhomedir adcli 2
samba-common-tools -y
```

If needed, adjust the DNS configuration in `systemd-resolved` as described previously. Make sure the configuration is correct and the Fedora client can contact the DC:

```
realm discover <Domain-Name>
```

If you do not get an error message here, add the Linux client to the domain:

```
realm join <Domain-Name> -v
```

The `-v` switch prints all the details of the operation on the CLI. At the end of the process, *Successfully enrolled machine in realm* confirms joining the domain. Now you can log in to the Linux GUI or CLI as with a domain user.

## Adding a Second DC

Of course, one DC should not manage a network alone, so you need to add at least one more DC. It is also considered best practice for DCs not to provide file shares but devote themselves to their tasks as domain controllers. Regular domain member servers without DC functions then host the filesystem shares – but more on this later. Setting up additional AD servers follows the same pattern as for the AD: Set up the packages, delete `smb.conf`, check the DNS configuration, and adjust if necessary. On another DC server, run the Samba tool:

```
samba-tool domain join <domain-name> DC 2
-U"Administrator@<domain-name>"
```

The tool will then prompt you for the admin password and perform the AD setup. Afterward, disable and

Shop the Shop → [shop.linuxnewmedia.com](https://shop.linuxnewmedia.com)

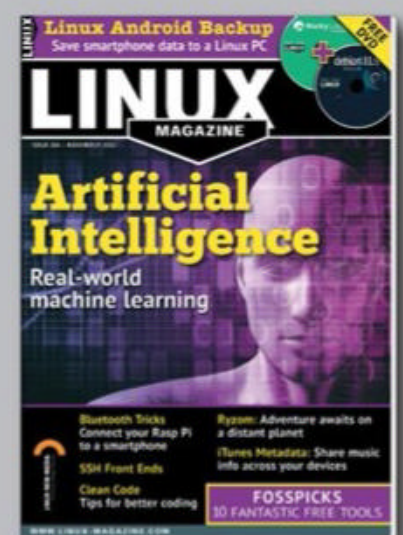
Discover the past and invest in a new year of IT solutions at Linux New Media's online store.

Want to subscribe?

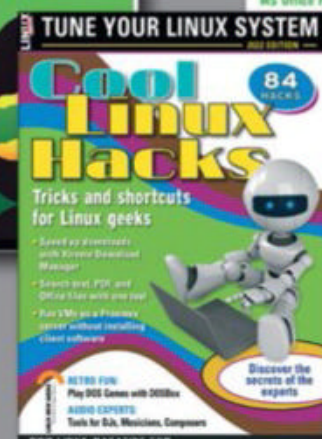
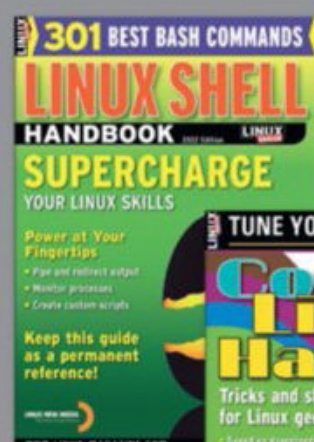
Searching for that back issue you really wish you'd picked up at the newsstand?

➤ [shop.linuxnewmedia.com](https://shop.linuxnewmedia.com)

DIGITAL & PRINT  
SUBSCRIPTIONS



SPECIAL EDITIONS





comment out the regular Samba services as on the first DC, enable the *samba-ad-dc* service instead, and apply the Kerberos configuration.

Becoming a Domain Member

If you specify

```
samba-tool domain join 2
<Domain Name> MEMBER
```

instead of DC, the system only joins the domain as a member and not as a domain controller. However, this does not always work flawlessly or with all distributions. The somewhat more complex way, which also works with other distributions such as Fedora or EL, uses a manual configuration. The setup of a member server differs from the client setup described above. For a client, you just need the System Security Services Daemon (SSSD) with Kerberos. The member on the other side connects to the domain with Samba’s Winbind service without needing SSSD. On the EL8 system (e.g., RHEL 8, Alma or Rocky Linux, CentOS Stream 8), which I use as a domain member, the necessary services need to be set up first:

```
dnf install samba samba-client 2
samba-winbind 2
samba-winbind-clients 2
oddjob oddjob-mkhomedir
```

Adjust the *smb.conf* configuration for the member server ([Listing 1](#)).

Listing 1: smb.conf for Member Server

```
[global]
realm = <ADS Realm>
security = ADS
template shell = /bin/bash
winbind enum groups = Yes
winbind enum users = Yes
winbind offline logon = Yes
winbind use default domain = Yes
workgroup = <AD Workgroup Name>
idmap config * : rangesize = 1000000
idmap config * : range = 100000-19999999
idmap config * : backend = autorid
...
```

Doing so specifies that Winbind uses ADS to query the users. To ensure that the local system also uses Winbind to determine users and their IDs, you also need to modify the */etc/nsswitch.conf* file. The entries for *passwd* and *group* should look something like:

```
passwd      compat winbind
group       compat winbind
```

The Linux system will then first search the local *passwd* and *group* files for users and groups, along with their IDs. If it doesn’t find anything there, it fetches the information from the domain via Winbind. In an AD client setup, for example, *sssd* is displayed here instead of *winbind*. Now you can register the member server as a member of the domain with

```
net ads join -U Administrator
```

and start the *winbind* and *smb* services. You might get an error message stating that DNS registration did not work, but this response is not a big deal. Just enter the server manually in the ADS-integrated DNS. After that, you can share folders on the member server and make them available as shares in the ADS forest.

Text or Registry

A Samba server usually stores its configuration in the */etc/samba/smb.conf* text file, making it difficult for a Windows administrator to configure Samba services. Alternatively, the service can save all or part of its configuration in a registry. This function is often used in Samba clusters. For example, in the domain setup described here, this means that an administrator can customize the Samba configuration of a Linux server with the registry editor while working on a Windows workstation. To put the complete configuration of a Samba server into the registry, you need an */etc/samba/smb.conf* with exactly two lines:

```
[global]
config backend = registry
```

In this case, Samba ignores the complete content of the *smb.conf* file and reads the information from the registry only. In many cases, this method is too strict, which is why you have two other options:

```
include = registry
registry shares = yes
```

The first option evaluates the complete *smb.conf* file plus the registry entries, and the second option only reads information about shares from the registry. Therefore, you can only change the shares, but not the basic server configuration in the registry. Samba does not evaluate all the share information at startup time. Only when a client wants to access a share does Samba look for the associated information in the registry. Samba saves all its keys in *HKLM/Software/Samba/smbconf*. Given sufficient rights, you can edit the registry with *regedit* over the LAN or with the Linux *samba-regedit* CLI tool in a legacy GTK design. If you want to work with a registry, you can transfer an existing *smb.conf* to the registry with `net conf import`.

Conclusions

On small and medium-sized networks with heterogeneous systems, a Samba domain controller performs well and ensures a smooth exchange of files between the various client systems. In this article, I looked at as many solutions as possible for different distributions. Of course, the whole thing works on a Mac, on Windows, and on Linux, even though I did not go into the details of setting up an AD connection for a macOS client here.

Info  
[1] Microsoft GUI Tools (RSAT):  
[\[https://www.microsoft.com/en-us/download/details.aspx?id=45520\]](https://www.microsoft.com/en-us/download/details.aspx?id=45520)





**Linux Magazine** is your guide to the world of Linux. Look inside for advanced technical information you won't find anywhere else!

### Expand your Linux skills with:

- In-depth articles on trending topics, including Bitcoin, ransomware, cloud computing, and more!
- How-tos and tutorials on useful tools that will save you time and protect your data
- Troubleshooting and optimization tips
- Insightful news on crucial developments in the world of open source
- Cool projects for Raspberry Pi, Arduino, and other maker-board systems

If you want to go farther and do more with Linux, subscribe today and never miss another issue!

## Subscribe now!

[shop.linuxnewmedia.com/subs](http://shop.linuxnewmedia.com/subs)

**GET IT NOW!**

FAST DELIVERY  
WITH OUR PDF  
EDITION





IAM for midmarket companies

# Spoiled for Choice

We look at the roll of identity and access management in midmarket organizations. By Martin Kuppinger

**Up to now, identity and access management (IAM)** has mainly been the domain of larger organizations, but it is important for organizations of all sizes to manage digital identities (not only of their employees) and their access authorizations in an efficient and effective way. However, a chronic shortage of personnel in the midmarket makes it crucial to define their own IAM requirements precisely and select the right providers.

IAM has the reputation of being complex; this opinion is sometimes justified, but by no means always true. This rumored, presumed, or perceived complexity, together with what are typically small IT teams in the midmarket, often make organizations reluctant to venture into the discussion. However, IAM is important for many reasons: security; regulatory compliance; more efficient processes for employees, business partners, and customers; simple yet secure access to systems

and applications; and, last but not least, administrative efficiency. The popular definition of the midmarket includes medium-sized companies with 51 to 1,000 employees and larger medium-sized companies with 1,001 to 10,000 employees. In the larger midmarket, genuine IAM infrastructures are very often already in place for managing users and access authorizations (IGA, identity governance and administration), for access control (access management with authentication and identity federation), and in some cases, for monitoring and controlling access by highly privileged users (PAM, privileged access management).

Smaller companies, on the other hand, often have only technical administration tools, for example, the Quest Active Roles server for Microsoft Active Directory (AD). IT managers then tend to manage permissions in applications interlocked with Active Directory in AD groups while managing other applications manually. Sometimes

specialized solutions for business applications show up, like SAP (e.g., SAP Access Control).

## New Requirements Make IAM Essential

However, these products are not up to the task in most cases, not least because requirements are growing and IT environments are changing. To begin with, every organization is a potential target for cyberattacks. Attackers are always looking to gain control of user accounts to steal data, distribute malware, or carry out attacks. However, that is only one aspect, because IT landscapes are changing in small and mid-sized enterprises (SMEs) because of the increasing use of cloud services with a tendency to use more services overall – often for specialized tasks. For each of these services, managing users and permissions securely is important.

Additionally, the role of Active Directory, used in most midmarket

Photo by Robert Anasch on Unsplash



companies, is changing. After the introduction of Microsoft 365, companies started using Azure Active Directory (AAD) and Microsoft Entra [1], which meant that, for what is often a long transition period, two central services are combined and managed, increasing the complexity that IAM can help reduce.

One issue that no midmarket company in any sector can afford to underestimate, but especially in the manufacturing industry, is the demand from important customers for certification in line with the ISO 2700x standards. These standards also include working IAM. Even if the requirement can be covered in a basic way with manual processes, certification can be achieved more easily if organizations use suitable IAM applications.

This situation is even more true for companies in the critical infrastructure sector, where the requirements for IT security management, and therefore also for user and authorization management, have been significantly tightened. Following the German KRITIS (critical infrastructure) revisions, the focus has also increasingly shifted to medium-sized companies, where working IAM has become practically mandatory.

The topic of Industry 4.0 (convergence of information and operational technology systems for seamless generation, analysis, and communication of data) is also directly related to IAM – on the one hand for access control in the manufacturing sector and on the other hand to secure business systems that interface with and reduce the risk of attacks on production systems.

Most importantly, it’s not just about employees, but also access by (and applications for) business partners and, especially, customers and consumers. Virtually all organizations are facing growing demand to provide more digital services, which are increasingly at the core of business models. The digital identities of customers and consumers need to be managed, as does employee access to these services.

### The Right Amount of IAM for the Midmarket

How much IAM is feasible for the midmarket and which aspects of IAM are genuinely needed? IAM is very diverse, as a look at reference architectures shows (Figure 1), but what parts of it do SMEs really need? IGA includes tools for user lifecycle

management for technical identity provisioning, with the creation, modification, and deletion of user accounts in target systems, complemented by support for access governance (e.g., the recertification of authorizations). These are core areas of IAM. Within IGA, however, it is primarily about having standardized processes for the essential tasks (e.g., creating and changing departments and deleting users), supporting a simple authorization request, and checking and being able to connect to critical target systems. For midsize companies in particular, connecting can turn into a challenge because many vendors support common business applications for large enterprises but not midsize solutions. Providers specializing in SMEs (e.g., Tenfold) can be an alternative. Access management for centralized control of authentication; support for multifactor authentication and, if possible, login without passwords; and interaction with target applications is also a must. This integration requires not only support for common identity federation standards such as OAuth, OpenID Connect [2], and security assertion markup language (SAML), but also web access management for legacy applications.

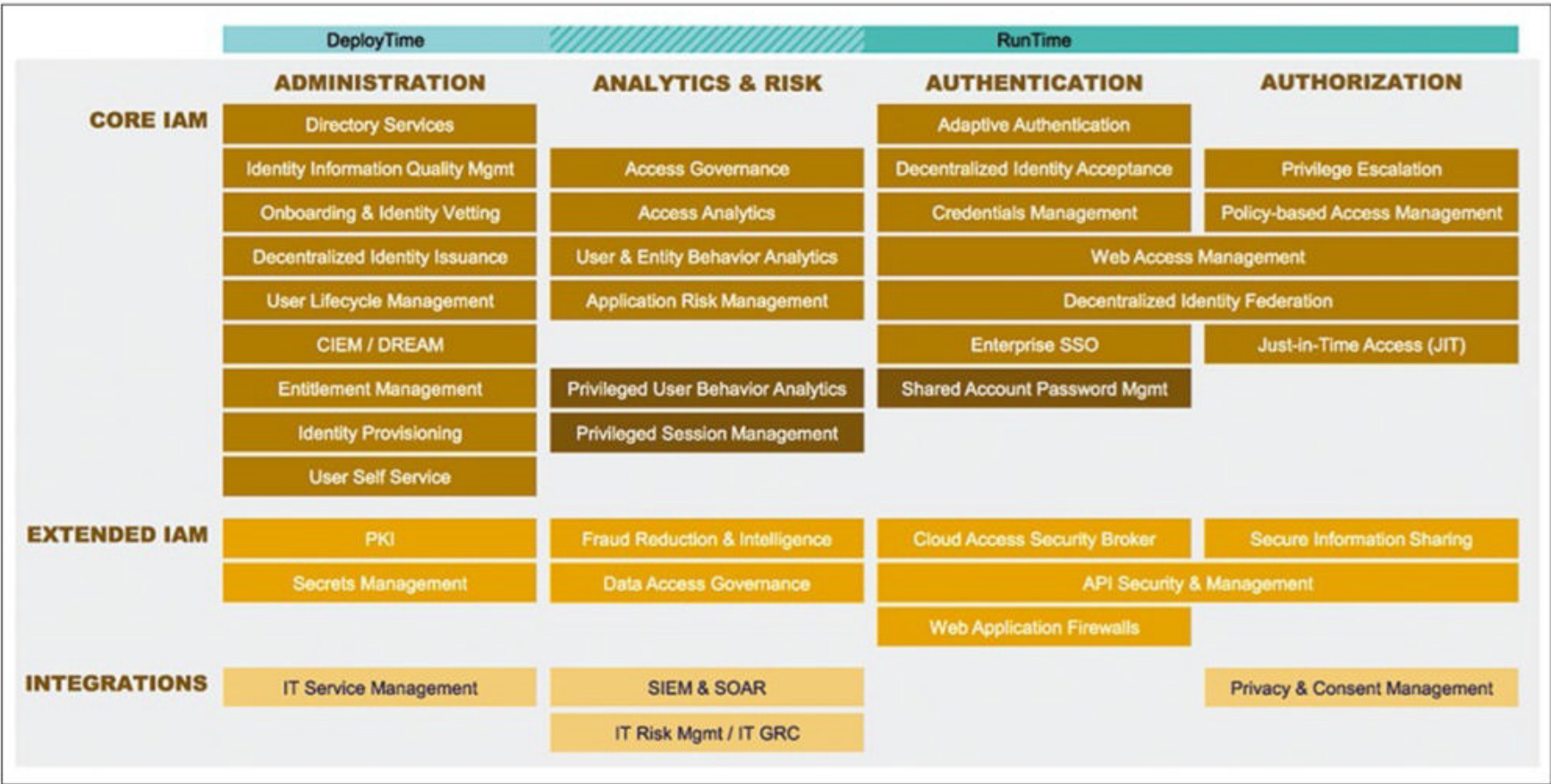


Figure 1: The IAM reference architecture shows the variety of related topics. SMEs need to think carefully about what they want to implement and what they can implement.



In addition to these building blocks, admins also need to think about a suitable customer IAM (CIAM) application, especially where companies are developing many of their own digital services. Which products are best suited here depends heavily on the application scenarios. Where many custom applications are created, developer-focused tools that provide APIs for identity functions such as user registration and authentication make the most sense (e.g., Okta [3] and Auth0 [4]).

The use of PAM in SMEs is also important, although nowhere near as widespread. However, some of today's IAM offerings offer integrated PAM features, at least for basic functions. Some PAM providers are in the market, such as Delinea, but also smaller specialists who have comparatively lean and easy-to-implement products in their portfolio. Some of the vendors in the market, such as Microsoft or EmpowerID [5], offer several of these functional areas from a single source in an integrated platform.

## Planning IAM Carefully

Each of the topics mentioned above is a project in its own right, which makes it important to plan the introduction or modernization of IAM carefully and not take on too much. Of course, the only way to ensure that

a sensible overall solution is created is to make a roadmap – a precise definition of where the company wants to go and which projects are part of the IAM program.

The first step for IT managers is to consider whether individual solutions or a suite of products are more suitable to cover as many areas as possible from a single source. Microsoft and Okta, but also EmpowerID or N8 Identity, are some of the vendors that cover multiple functional areas in a comparatively lean application.

On the other hand, specialists in each of the sub-segments are highly suitable for SMEs. When it comes to access management in particular, many cloud services beyond Microsoft and Okta can be easily implemented. In addition to One Identity and OneLogin or Ping Identity are European providers such as Nevis, Ergon, United Security Providers, or Ilex. A number of IGA manufacturers with many references in SMEs include European providers such as Omada or Beta Systems, on top of SME specialists such as Tenfold or OGITiX.

In terms of sequence, the typical starting points are either access management to enable secure user access by multifactor authentication and identity federation or IGA for basic lifecycle and authorization management functions. Both

are fine as long as the IT manager has a clear plan for the step-by-step implementation of the other functions, too.

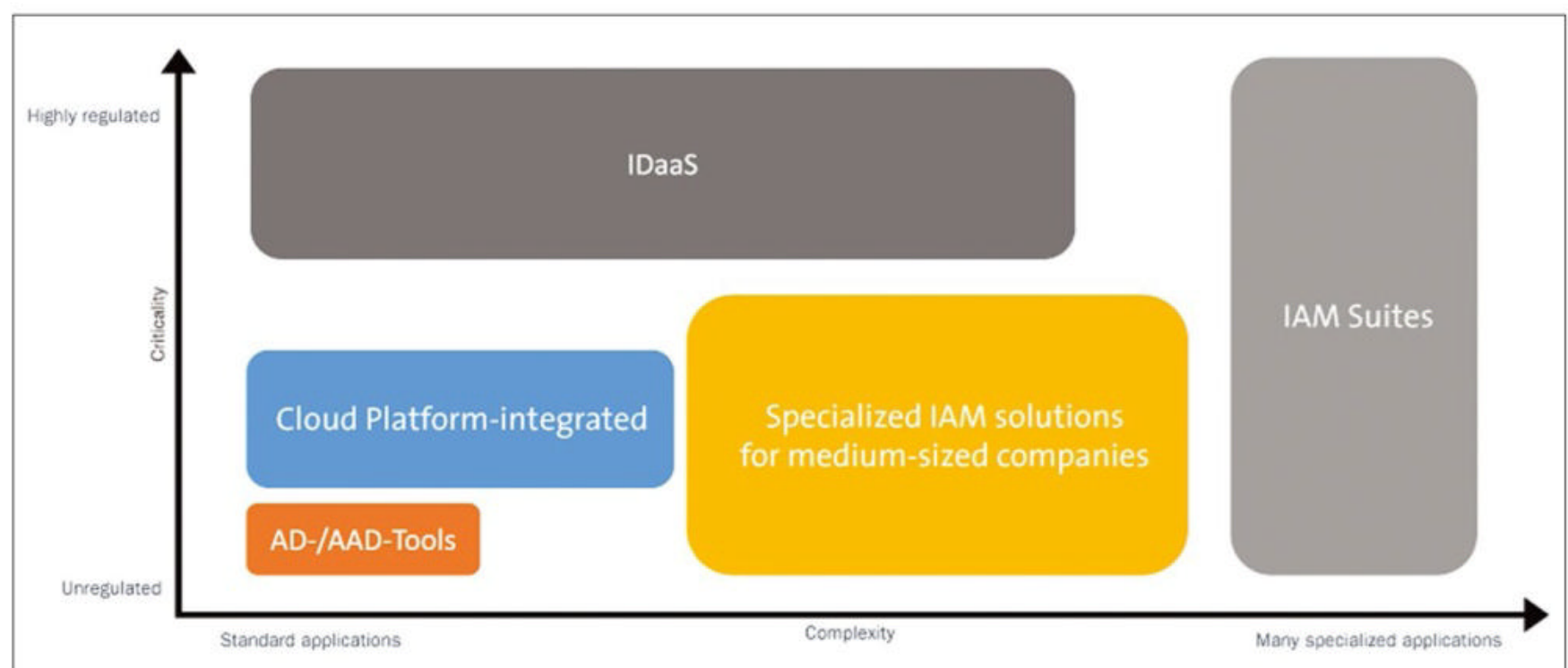
## Choosing the Right Provider

As is already clear from the list of vendors in the examples above, many companies supply platforms suitable for SMEs. Depending on your specific requirements, though, you might need to turn to large-scale, leading IAM providers. Where companies develop digital services of high complexity, ForgeRock [6] is an interesting option, especially in terms of CIAM requirements. For complex regulatory requirements, market-leading IGA vendors such as SailPoint [7], One Identity [8], or IBM are relevant options, even for the smaller midmarket.

In addition to the target image is the matter of defining your own requirements, both in terms of the required functionality and the target systems to be connected. This dataset can then be used to compare providers. Taking a closer look at the products and having them demonstrated is also important to understand whether they meet the requirements and what the internal IT team can do to help.

In essence, the providers can be divided into five categories (Figure 2):

- Cloud-integrated solutions from the major cloud providers: Both



**Figure 2:** The focal areas of use of different types of IAM tools, looking at criticality and complexity of requirements.



Microsoft and Google offer IAM functionality in tight integration with their respective Office platforms. These platforms can also provide the foundation for organization-wide IAM, especially in the case of Microsoft with its coverage of all major IAM functional areas.

- Plain vanilla identity as a service (IDaaS) (i.e., IAM products from the cloud). These solutions are now offered by almost all IAM vendors and have the advantage that the overhead required for setup, customization, and operation is significantly lower than for traditional local tools. This solution also makes applications that were previously considered too complex for the midmarket as on-premises variants an option in this market segment.
- IAM systems in the various sub-segments that can be used locally and that bundle multiple functions or are generally comparatively simple and lean in use.
- Complementary solutions for managing Azure AD and Entra and on-premises AD, where the vast majority of target applications interact with these systems anyway.
- Specialized IAM with a midmarket focus, often characterized by the fact that it also has many interfaces to typical business applications in the midmarket.

The IAM market is very large, so I can only list the vendors by way of example at this point. In any case, I

recommend doing detailed research, whether on the Internet or by calling in an analyst to identify suitable providers.

Different applications are required as a function of the complexity and criticality (i.e., the regulatory requirements in particular). Where regulatory pressure is high and environments are complex, (e.g., in KRITIS-relevant companies), classic IAM or IDaaS from established providers is more in demand; otherwise, cloud-integrated or specialist products for SMEs, for example, might be the better choice.

The two most important criteria for selection are defining your own requirements and taking the time to ensure a good overview of the market. You can then select the products, check them against your requirements, view demos, and, if necessary, perform a proof of concept.

## Conclusions

Because IAM projects are costly, IT managers need to invest both time and money to identify the right product. One thing is certain: Mistakes in terms of product selection are far more time-consuming and expensive than well-planned and -executed product selection. When defining requirements, it is important to have realistic goals. What is really needed and what is manageable? Simple authorization models and a simple

recertification instead of sophisticated functions are usually the better choice. Also important is for manufacturers to include standard features, such as predefined processes and reports. References in the SME sector are also an important criterion. Additionally, those responsible for IT should be supported, both in the selection of products and in project implementation, by suitable partners, who – depending on the phase – offer market knowledge or implementation expertise, especially in SMEs. If you get it right and the setup is carefully considered, IAM is feasible and controllable even for medium-sized businesses, where IAM is more important than ever. ■

---

### Info

- [1] Microsoft Entra: <https://www.microsoft.com/en-us/security/blog/2022/05/31/secure-access-for-a-connected-worldmeet-microsoft-entra/>
- [2] OpenID Connect: <https://openid.net/connect/>
- [3] Okta: <https://www.okta.com>
- [4] Auth0: <https://auth0.com>
- [5] EmpowerID: <https://www.empowerid.com>
- [6] ForgeRock: <https://www.forgerock.com>
- [7] SailPoint: <https://www.sailpoint.com>
- [8] One Identity: <https://www.oneidentity.com>

---

### Author

**Martin Kuppinger** is the founder of and Principal Analyst at KuppingerCole Analysts AG.

---





## Understanding Cybersecurity Maturity Model Certification

# Ready, Steady, ...

United States Cybersecurity Maturity Model Certification will be required by mid-2023 to handle controlled unclassified information and win federal contracts, but it can also help minimize business risk and keep information out of the hands of adversaries. By Christopher J. Cowen

**The US Department of Defense** (DoD or the Department) created the Cybersecurity Maturity Model Certification (CMMC) program to add a comprehensive and scalable certification process to verify the implementation of industry practices that achieve a cybersecurity maturity level. CMMC is designed to provide assurance to departments and agencies that the defense industrial base (DIB) contractor can adequately protect sensitive unclassified information such as federal contract information and controlled unclassified information (CUI). The US government is concerned with ensuring that the data and information their contractors receive is stored and used safely. This government-furnished information is more commonly known as GFI. Of great concern is the potential that government-furnished information will escape into the wild. The US government wants to protect itself and its citizens against the theft of

intellectual property and the sensitive information of US industrial sectors from malicious cyber activity. The aggregate loss of intellectual property and certain unclassified information from the DoD supply chain can threaten US economic and national security by undercutting technical advantages and innovation and significantly increasing risk to national security.

To address these concerns, the DoD issued an interim rule [\[1\]](#) September 2020 intended to create a DoD assessment methodology and CMMC framework to assess a contractor's cybersecurity posture. By issuing the interim rule, the US government is seeking to understand what requirements and business practices contractors incorporate to protect their unclassified information systems and the data that is housed on those systems from a threat actor.

The US currently requires DoD contractors to include Defense Federal

Acquisition Regulation Supplement (DFARS) clause 252.204-7012 [\[2\]](#) in subcontracts for which subcontract performance will involve covered defense information (DoD CUI). DFARS provides acquisition regulations that are specific to the DoD and outlines regulations to which DoD government acquisition officials, contractors, and subcontractors must adhere when doing business with the DoD. However, this DFARS clause does not provide the department or agency contracting this work sufficient insights with respect to the cybersecurity posture of DIB companies throughout the multitier supply chain for any given program or technology development effort. Given the size and scale of the DIB sector, the DoD cannot scale its organic cybersecurity assessment capability to conduct on-site assessments of approximately 220,000 DoD contractors every three years. As a result, the Department's organic

Photo by Braden Collum on Unsplash



assessment capability is best suited for conducting targeted assessments for a subset of DoD contractors that support prioritized programs, technology development efforts, or both. CMMC addresses the challenges of contractor assessment capabilities by partnering with an independent organization that will accredit and oversee third-party assessment and conduct on-site assessments of DoD contractors throughout their multitier supply chain contract. The cost of these CMMC assessments will be driven by multiple factors, including market forces, the size and complexity of the network or enclaves under assessment, and the CMMC level. Later I will talk about the plans to enforce CMMC.

## What Material Falls Within CMMC?

In a simplified definition, only unclassified data and information that falls below the classified category falls within the purview of CMMC. Classified data and information have distinct processes and guidelines that contractors must adhere to, to possess these types of materials. Classified data not covered by CMMC is secret (S), top secret (TS), or top secret sensitive compartmented information (TS-SCI).

Within the unclassified data and information category of material are distinct classifications of data that require CMMC protection:

- CUI Assets process, store, or transmit CUI.
- Security Protection Assets provide functions or capabilities to include people, technology, and facilities.
- Contractor Risk Managed Assets are capable of, but not intended to, process, store, or transmit CUI because of the security policy, procedures, and practices in place.
- Specialized Assets are government property, industrial Internet of Things, supervisory control and data acquisition (SCADA) systems, and restricted information systems or test equipment that may handle CUI.

## CMMC Framework

The framework has three main key features:

*Tiered Model.* CMMC requires that companies entrusted with national security information implement cybersecurity standards at progressively advanced levels, depending on the type and sensitivity of the information. The program also sets forth the process for information flow down to subcontractors.

*Assessment Requirement.* CMMC assessments allow the Department to verify the implementation of clear cybersecurity standards.

*Implementation Through Contracts.* Once CMMC is fully implemented, certain DoD contractors that handle sensitive unclassified DoD information will be required to achieve a particular CMMC level as a condition of contract award.

Building on the National Institute of Standards and Technology special publication (NIST SP) 800-171 (DoD Assessment Methodology) [3], the CMMC framework adds a comprehensive and scalable certification element to verify the implementation of processes and practices associated with the achievement of a cybersecurity maturity level. CMMC is designed to provide increased assurance to the Department that a DIB contractor can adequately protect sensitive unclassified information (i.e., CUI) at a level commensurate with the risk, accounting for information flow down to its subcontractors in a multitier supply chain.

Implementation of the CMMC framework is intended to solve the following problems:

- Verification of a contractor's cybersecurity posture. DIB companies self-attest that they will implement the requirements in NIST SP 800-171 on submission of their contract offer.
- DoD contractors that inconsistently implement mandated system security requirements for safeguarding CUI.
- Verification of a DIB contractor's cybersecurity posture. The

company must achieve the CMMC level certification required as a condition of contract award.

## CMMC in Depth

CMMC is currently on its second iteration, CMMC 2.0. The prior iteration of CMMC was CMMC 1.0, which built a framework of four elements: security domains, capabilities, practices, and processes. When combined, they built best practices for the protection of an organization and associated federal contract information and CUI. CMMC 1.0 had five cybersecurity maturity levels (1-5) that composed the CMMC framework, with level 1 being the least mature and level 5 the most mature.

The CMMC 1.0 framework consisted of 17 cybersecurity domains. A domain is a distinct group of security practices that have similar attributes and are key to the protection of federal contract information and CUI, either individually or in combination. Each domain comprises several capabilities an organization is expected to achieve to ensure that cybersecurity and the protection of federal contract information and CUI is sustainable. Capabilities are a combination of practices, processes, skills, knowledge, tools, and behaviors, which when working together enable an organization to protect federal contract information and CUI. In total (at level 5) the CMMC framework identifies 171 practices associated with the 17 security domains and mapped across the five maturity levels.

Practices applied at maturity levels 1 and 2 have been referenced from Federal Acquisition Regulation (FAR) 52.204-21 [4] for the basic safeguarding of covered contractor information systems applied to the protection of federal contract information. Practices applied at levels 3, 4, and 5 are referenced from DFARS 252.204-7012 [2] for the safeguarding of covered defense information and cyber-incident reporting. Most of the above framework and guidelines carried over to CMMC 2.0 with the one big exception that CMMC 2.0 only has



three levels, as opposed to the five levels of CMMC 1.0.

### CMMC 2.0

In March 2021, the DoD initiated an internal review of CMMC’s implementation, informed by more than 850 public comments in response to the interim DFARS rule. This comprehensive, programmatic assessment engaged cybersecurity and acquisition leaders within DoD to refine policy and program implementation, and in November 2021, the DoD announced CMMC 2.0 [5]. The view is that version 2.0 would achieve the primary goals that came out of this internal review:

- Safeguard sensitive information to enable and protect the warfighter.
- Dynamically enhance DIB cybersecurity to meet evolving threats.
- Ensure accountability while minimizing barriers to compliance with DoD requirements.
- Contribute toward instilling a collaborative culture of cybersecurity and cyber resilience.
- Maintain public trust through high professional and ethical standards.

The enhanced CMMC 2.0 program maintains the program’s original goal

of safeguarding sensitive information while trying to simplify the CMMC standard and providing additional clarity on cybersecurity regulatory, policy, and contracting requirements. The US government believes this is achieved by focusing the most advanced cybersecurity standards and third-party assessment requirements on companies supporting the highest priority programs and increasing Department oversight of professional and ethical standards in the assessment ecosystem.

The program is meant to ensure accountability for companies to implement cybersecurity standards while minimizing barriers to compliance with DoD requirements, as well as instill a collaborative culture of cybersecurity and cyber resilience and enhance public trust in the CMMC ecosystem, while increasing overall ease of execution. Jesse Salazar, Deputy Assistant Secretary of Defense for Industrial Policy, one of the leaders of the CMMC effort, said [6]:

*CMMC 2.0 will dramatically strengthen the cybersecurity of the defense industrial base. By establishing a more collaborative relationship with industry, these updates will support businesses in adopting the practices*

*they need to thwart cyber threats while minimizing barriers to compliance with DoD requirements.*

Figure 1 shows the difference between CMMC versions 1.0 and 2.0. With the implementation of CMMC 2.0, the Department is introducing several key changes that build on CMMC 1.0 and refine the original program requirements:

- A *streamlined model* that focuses on the most critical requirements and reduces the model from five to three compliance levels. This change aligns CMMC 2.0 with widely accepted standards and uses NIST cybersecurity standards.
- *Reliable assessments* reduce assessment costs and allow all companies at level 1 (Foundational) and a subset of companies at level 2 (Advanced) to demonstrate compliance through self-assessments. Higher accountability increases oversight of professional and ethical standards of third-party assessors.
- *Flexible implementation* encourages a spirit of collaboration that allows companies, under certain limited circumstances, to make plans of action and milestones

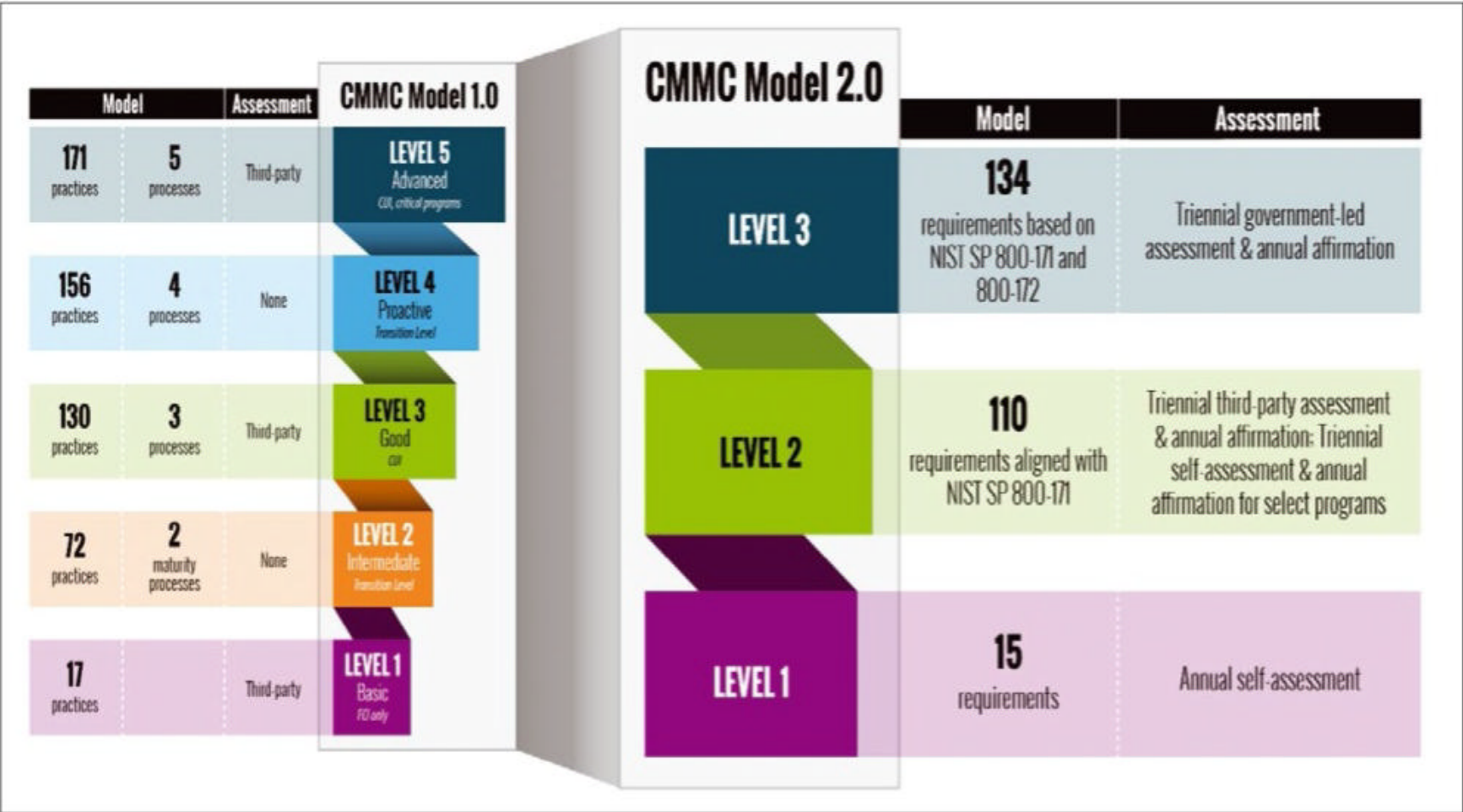


Figure 1: Key features of CMMC 2.0 compared with version 1.0.



(POA&Ms) to achieve certification. The added flexibility and speed allow waivers to CMMC requirements under certain limited circumstances.

## CMMC Is Coming

Whether you like it or not, the DoD is planning to release an interim rule on the CMMC framework by May 2023, according to Stacy Bostjanick, director of the CMMC program for the DoD [7]. CMMC will be enacted on the day the interim rule is published, and CMMC requirements will start to appear in DoD contracts by July 2023, 60 days after publication of the interim rules.

Doing nothing is not an option. As of this writing, defense contractors have little more than a year to become CMMC compliant. If your company is still at the very beginning of this timeline, the time for action is now, although you might be able to take consolation on two fronts:

First, the time and effort needed to achieve compliance will be different for every defense contractor. Variables include your baseline cybersecurity maturity level and the resources and prioritization you can assign to achieving compliance. Your organization might need less time to achieve and demonstrate compliance, or you might be able to commit more resources and energy than a typical contractor would and so make up for lost time.

Second, under CMMC 2.0, it will be permissible to sign DoD contracts with a POA&M in place, which means your organization does not need to achieve the highest assessment score possible by May 2023. However, in case that lessens your sense of urgency, consider that CMMC 2.0 will bring critical changes to the reprieve that POA&Ms have historically offered. POA&Ms will most likely have time constraints. At this point, it seems that contractors will have 180 days to remediate security gaps identified in their POA&Ms,

but that is subject to change during the CMMC 2.0 interim rule-making process.

Bostjanick also said recently [7] that the DoD is expected to permit POA&Ms only for the lowest-risk security controls (i.e., those that are worth just one point in the DoD's assessment scoring methodology). Fifty of the 110 security controls are worth one point. That leaves 60 controls that are worth either three or five points and must be met before a contractor's CMMC level 2 assessment. These 60 controls are both the most important for securing CUI and some of the most difficult to meet.

In short, although your organization doesn't have to achieve the highest possible assessment score by May 2023, it should be on the cusp of doing so by then. Your business risk is too high to be far behind the timeline of your organization's compliance journey.

## Conclusion

A company's goal should not just be to achieve eligibility to win defense contracts, but also to minimize business risk and keep CUI out of the hands of adversaries. By getting started on your organization's compliance journey, you can achieve these objectives and ensure your company is ready for ramped-up federal enforcement of cybersecurity regulations. ■

### Info

- [1] Defense federal acquisition regulation supplement: assessing contractor implementation of cybersecurity requirements (DFARS case 2019-D041). 85 *Federal Register* 61505 (Sep 2020), 18 p., [<https://www.federalregister.gov/documents/2020/09/29/2020-21123/defense-federal-acquisition-regulation-supplement-assessing-contractor-implementation-of>]
- [2] Safeguarding covered defense information and cyber incident reporting (Dec

2019). DFARS 252.204-7012 (Sep 2022), [<https://www.acquisition.gov/dfars/252.204-7012-safeguarding-covered-defense-information-and-cyber-incident-reporting.>]

- [3] "What Is the NIST SP 800-171 and Who Needs to Follow It?" by Traci Spencer. *NIST*, Oct 8, 2019, [<https://www.nist.gov/blogs/manufacturing-innovation-blog/what-nist-sp-800-171-and-who-needs-follow-it-0>]
- [4] Basic safeguarding of covered contractor information systems. FAR 52.204-21 (Oct 2022), [<https://www.acquisition.gov/far/52.204-21>]
- [5] "Strategic Direction for Cybersecurity Maturity Model Certification (CMMC) Program." US DoD release. Nov 4, 2021, [<https://www.defense.gov/News/Releases/Release/Article/2833006/strategic-direction-for-cybersecurity-maturity-model-certification-cmmc-program/>]
- [6] "DOD launches CMMC 2.0 program; Jesse Salazar quoted" by Jane Edwards. *GovConWire*, Nov 5, 2021, [<https://www.govconwire.com/2021/11/dod-launches-cmmc-2-0-program/>]
- [7] "Stacy Bostjanick shares updated DoD CMMC rollout schedule" by Summer Myatt. *GovConWire*, May 20, 2022, [<https://www.govconwire.com/2022/05/dods-cmmc-director-stacy-bostjanick-shares-updated-rollout-schedule/>]

### Author

Christopher J. Cowen is currently a Senior Cyber Security Analyst with the US Department of Defense (contractor). He has worked within informa-



tion technology for more than 20 years within both the corporate and government spaces. He is currently focused on information security for nuclear facilities and critical infrastructure security. Cowen is a Certified Information Systems Security Professional (CISSP), a Certified Ethical Hacker (CEH), and a Certified Information Security Manager (CISM). He has been a featured speaker all over the world and has written articles for *Cyber Defense Magazine* and *Cyber Security: A Peer-Reviewed Journal*, among others. You can reach Chris at [[chrisjcowen@yahoo.com](mailto:chrisjcowen@yahoo.com)].





## Introducing parity declustering RAID

# Silverware

Declustered RAID decreases resilvering times, restoring a pool to full redundancy in a fraction of the time over the traditional RAIDz. We look at OpenZFS, the first freely distributed open source solution to offer a parity declustered RAID feature. By Petros Koutoupis

**Fault tolerance has been at the forefront of data protection** since the dawn of computing. To this day, admins continue to struggle with efficient and reliable methods to maintain the consistency of stored data, either locally or remotely on a server (or cloud storage pool) and keep searching for the best way to recover from a failure, regardless of how disastrous that failure might be. Some of the methods still being used today are considered ancient by today's standards. Why replace something that continues to work? One such technology is called RAID. Initially, the acronym stood for redundant array of inexpensive disks, but it was later reinvented to describe a redundant array of independent disks.

The idea of RAID was first conceived in 1987. The primary goal was to scale multiple drives into a single volume and present it to the host as a single pool of storage. Depending on how the drives were structured, you also saw an added performance or redundancy benefit. (See the box titled "RAID Recap.")

RAID technology isn't perfect. As drive capacities increase and storage technologies move further away from movable components and closer to persistent memory, RAID is starting to show its age and limitations, which is why its algorithms continue to be improved. Papers on declustering the parity of RAID 5 or 6 arrays date back to the early 1990s. The design aims to enhance the recovery performance of the array by shuffling data among all drives within the same array, including its assigned spare drives, which tend to sit idle until failure occurs. **Figure 2** showcases a very simplified logical layout of each stripe across all participating volumes. In practice, though, the logical layout is arranged randomly to distribute data chunks more evenly over the drives (and based on their logical offsets). It is done in such a way that, regardless of which drive fails, the workload to recover from the failure is distributed uniformly across the remaining drives. **Figure 3** highlights a simplified example of how this is exercised with the OpenZFS project, as discussed later in this article.

## Why Parity Declustering?

When a traditional redundant array fails and data needs to be rebuilt to a spare drive, two major things occur:

- Data access performance drops significantly. In parallel with user read and write requests (which involve rebuilding the original data stripe from the existing calculated parities, consuming more processing power), a background process is initiated to rebuild the original (and newly updated) data to the designated spare drive(s). This process creates a bottleneck to the spare drive.
- As a result of both activities occurring simultaneously, the recovery process will take much longer to complete. Depending on the user workload and capacity sizes, we are talking about days, weeks, even months. In the interim, however, you have the risk of a secondary (or tertiary) drive failing; depending on the RAID type, it can put your array into an extremely vulnerable state – if not



cripple the entire array and render your data unrecoverable.

By leveraging parity declustering, all drives within the array, including the spares, participate in the data layout with specific regions of a stripe dedicated as the spare chunks. When the time comes to regenerate and recover lost data because of a drive failure, all drives participate in the recovery process and, in turn, do not bottleneck a single drive (i.e., the spare). This method means a

reduced rebuild time, and when the existing array is under a heavier user workload, this matters.

## OpenZFS dRAID

How does one take advantage of this approach to data management? Fortunately, and fairly recently, in a joint vendor and community effort, the OpenZFS [1] project recently introduced parity declustering. The implementation, called distributed RAID

(dRAID), was released in OpenZFS version 2.1 [2].

As an example, to configure a single-parity dRAID volume with a single spare volume on five drives, you would use the `zpool` command-line utility:

```
# zpool create -f myvol2 2
draid:3d:1s:5c /dev/sd[b-f]
```

In this example, if you want dual or triple parity, you would instead

### RAID Recap

RAID allows you to pool multiple drives together to represent a single volume. Depending on how the drives are organized, you can unlock certain features or advantages. For instance, depending on the RAID type, performance can dramatically improve, especially as you stripe and balance the data across multiple drives, thus removing the bottleneck of using a single disk for all write and read operations. Again, depending on the RAID type, you can grow the storage capacity.

Most of the RAID algorithms do offer some form of redundancy that can withstand a finite amount of drive failures. In such a situation, when a drive fails, the array will continue to operate in a degraded mode until you recover it by rebuilding the failed data to a spare drive.

Most traditional RAID implementations use some form of striping or mirroring (Figure 1). With a mirror, one drive is essentially a byte-by-byte image of the other(s). Striping, on the other hand, retrieves data and writes or reads it across all drives of the array in a stripe. In this method, data is written to the array in chunks, and collectively, all chunks within a stripe define the array's stripe size.

Common RAID types include:

**RAID 0 - Disk striping.** Data is written in chunks across all drives in a stripe, typically organized in a round-robin fashion. Both read and write operations access the data in the same way, and because data is constantly being transferred to or from multiple drives, bottlenecks associated with reading and writing to a single drive are alleviated, and data access performance is dramatically improved. However, RAID 0 does not offer redundancy. If a single drive or drive sector fails, the entire array fails or is immediately invalidated.

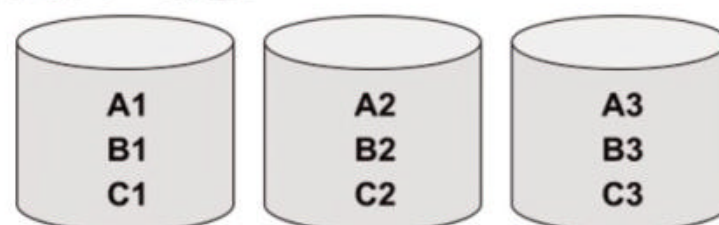
**RAID 1 - Disk mirroring.** In a RAID 1 array, one drive stores an exact replica of its companion drive, so if one drive fails, the second immediately steps in to resume where the first left off. Write performance tends to be half of a single drive because you are literally writing two copies of the same dataset (one to each drive). If designed properly, however, read performance can double through a mechanism called *read balancing*. Read requests can be split across both drives in the mirrored set so that each drive does half the work to retrieve data.

**RAID 5/6 - Redundancy.** Here is where things get a bit more complicated. RAID levels 5 and 6 are similar to RAID 0, except they offer a form of redundancy. Across each stripe of chunks exists a chunk dedicated to an XOR-calculated parity of all the other chunks within that same stripe. This special chunk is then balanced across all drives in the array, so that not one single drive will bear the burden of continuously writing updates to the same drive(s) every time a stripe is updated. These parity calculations make it possible to restore the original data content when a drive fails or becomes unavailable. A RAID 5 volume con-

tains a single parity chunk per stripe. A RAID 6 volume is designed with two parity chunks per stripe, allowing it to sustain two drive failures. Although not the focus of this article, hybrid RAID types are worth a mention. Typically hybrid types involve nesting one RAID type within another. For instance, striping mirrored sets is considered RAID 10, but if the striping includes a single parity (e.g., RAID 5), it is referred to as RAID 50.

RAID systems can be either hardware or software based. Once upon a time, processing power was limited, and it was advantageous to rely on external hardware for creating and managing RAID arrays. As time progressed, this approach became less significant. Server hardware grew more powerful, and implementing RAID solutions became much easier and less expensive through software with commodity storage drives attached to the server. However, hardware-based, vendor-specific storage arrays still exist, and they offer some additional fault tolerance and performance features that are advantageous in some settings.

#### RAID 0 - Stripe



#### RAID 1 - Mirror



#### RAID 5 - Stripe with Parity

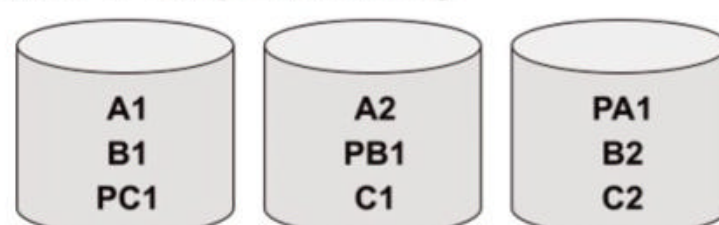


Figure 1: A general outline of RAID levels 0, 1, and 5.



substitute `draid` with `draid2` or `draid3`. The next colon-delimited field defines the number of data drives (in this case, three), then the number of spare drives, and finally the total number of children. If executed without error, `zpool` status would output something like [Listing 1](#). For reference purposes, with a traditional single-parity RAIDz

pool and a single spare, the output would look something like [Listing 2](#). Notice that in the dRAID example, all drives, including the spare, are active in the pool, whereas in the RAIDz example, a single spare drive remains idle until a failure occurs. To exercise this feature, you will need to “fail” a drive; the quickest way to

do so is to take it offline by the storage subsystem in `sysfs` (in Linux):

```
# echo offline >2
/sys/block/sdf/device/state
```

When the drive failure is detected on the next I/O operation to the pool, the distributed spare space will step in to resume where the failed drive left off ([Listing 3](#)).

The next step is to take the failed drive from the ZFS pool offline to simulate a drive replacement:

```
# zpool offline myvol2 sdf
# echo running >2
/sys/block/sdf/device/state
# zpool online myvol2 sdf
```

A resilver (resyncing, or rebuilding, a degraded array) is initiated across all drives within the pool ([Listing 4](#)).

### Conclusion

The OpenZFS project is the first freely distributed open source storage management solution to offer a parity declustered RAID feature. As stated earlier, the biggest advantage to using dRAID is that resilvering times are greatly reduced over the traditional RAIDz, and restoring the

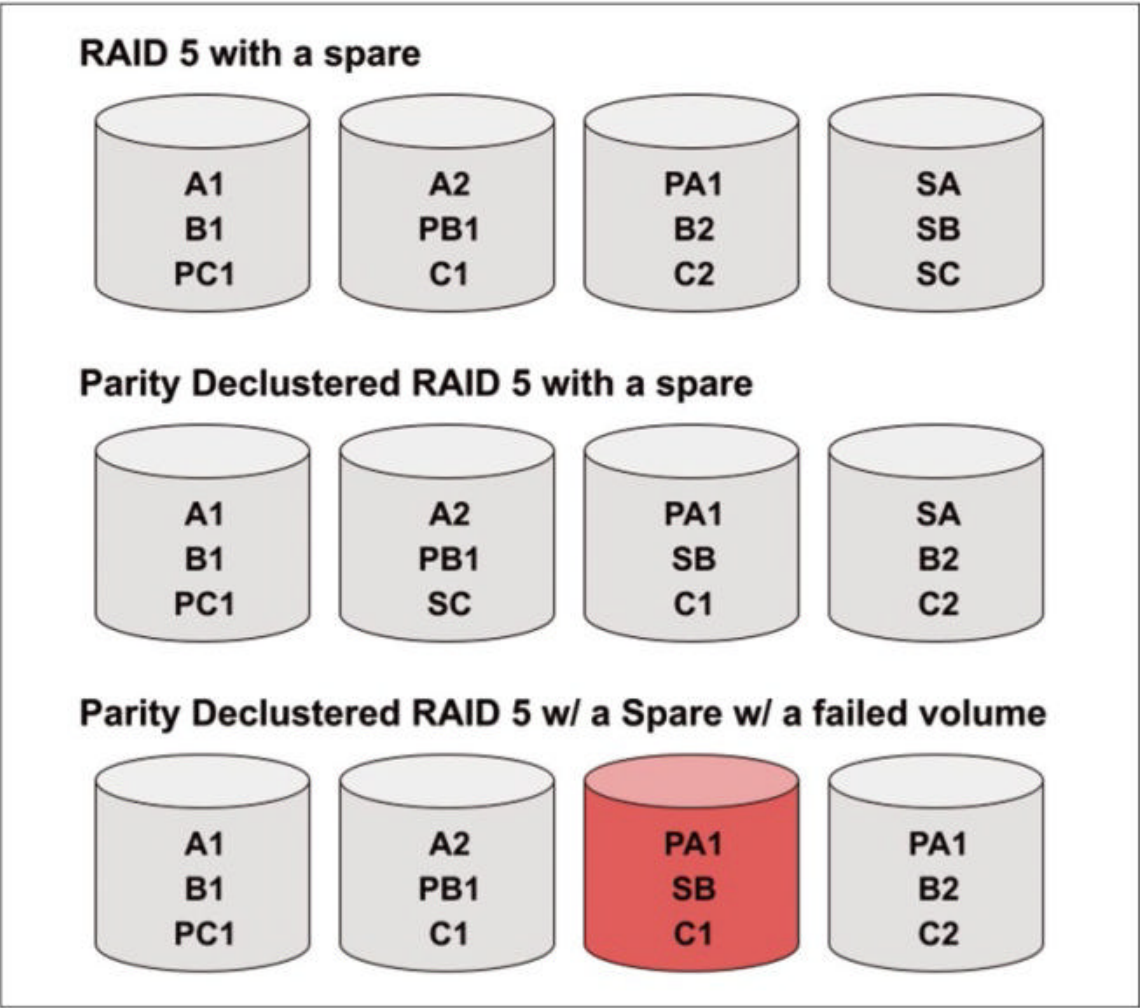


Figure 2: A simplified RAID 5 with a spare layout compared with a parity declustered layout.

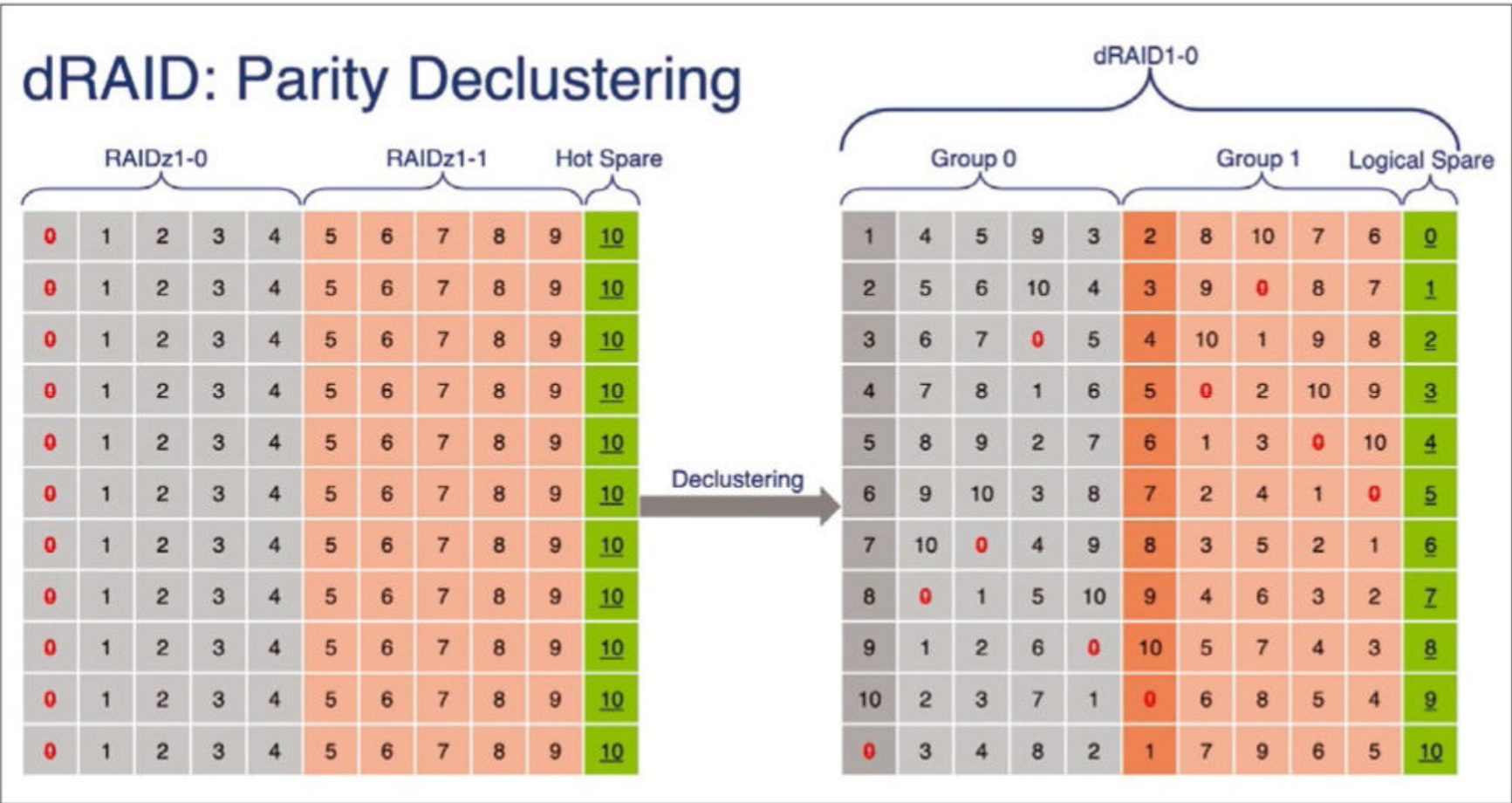


Figure 3: A simplified dRAID layout compared with RAIDz. © OpenZFS project



pool back to full redundancy can be accomplished in a fraction of the time. As drive capacities continue to increase, this feature alone will continue to show its value. ■

### Info

[1] OpenZFS Project page: [\[https://www.openzfs.org\]](https://www.openzfs.org)

[2] OpenZFS dRAID documentation: [\[https://openzfs.github.io/openzfs-docs/Basic%20Concepts/dRAID%20Howto.html\]](https://openzfs.github.io/openzfs-docs/Basic%20Concepts/dRAID%20Howto.html)

### Author

Petros Koutoupis is a senior performance software engineer at Cray (now HPE) for its Lustre High Performance File System division. He is also the creator and maintainer of the RapidDisk Project ([www.rapiddisk.org](http://www.rapiddisk.org)). Petros has worked in the data storage industry for well over a decade and has helped to pioneer the many technologies unleashed in the wild today.

**Listing 1: dRAID zpool status**

```
# sudo zpool status
pool: myvol2
state: ONLINE
config:

    NAME                STATE        READ WRITE CKSUM
    myvol2               ONLINE       0     0     0
      draid1:3d:5c:1s-0  ONLINE       0     0     0
        sdb              ONLINE       0     0     0
        sdc              ONLINE       0     0     0
        sdd              ONLINE       0     0     0
        sde              ONLINE       0     0     0
        sdf              ONLINE       0     0     0
    spares
      draid1-0-0         AVAIL

errors: No known data errors
```

**Listing 2: RAIDz zpool status**

```
# zpool status
pool: myvol1
state: ONLINE
config:

    NAME                STATE        READ WRITE CKSUM
    myvol1               ONLINE       0     0     0
      raidz1-0           ONLINE       0     0     0
        sdb              ONLINE       0     0     0
        sdc              ONLINE       0     0     0
        sdd              ONLINE       0     0     0
        sde              ONLINE       0     0     0
    spares
      sdf                AVAIL

errors: No known data errors
```

**Listing 3: Spare Drive Stepping In**

```
# zpool status
pool: myvol2
state: DEGRADED
status: One or more devices could not be used because the label is missing or
invalid. Sufficient replicas exist for the pool to continue
functioning in a degraded state.
action: Replace the device using 'zpool replace'.
see: https://openzfs.github.io/openzfs-docs/msg/ZFS-8000-4J
scan: scrub in progress since Mon Oct 24 17:11:22 2022
      80.7M scanned at 80.7M/s, 133M issued at 133M/s, 81.6M total
      0B repaired, 163.06% done, no estimated completion time
scan: resilvered (draid1:3d:5c:1s-0) 20.2M in 00:00:00 with 0 errors on Mon Oct 24
17:11:22 2022
config:

    NAME                STATE        READ WRITE CKSUM
    myvol2               DEGRADED     0     0     0
      draid1:3d:5c:1s-0  DEGRADED     0     0     0
        sdb              ONLINE       0     0     0
        sdc              ONLINE       0     0     0
        sdd              ONLINE       0     0     0
        sde              ONLINE       0     0     0
        spare-4          DEGRADED     0     0     0
          sdf            UNAVAIL       3    397     0
            draid1-0-0    ONLINE       0     0     0
    spares
      draid1-0-0         INUSE        currently in use

errors: No known data errors
```

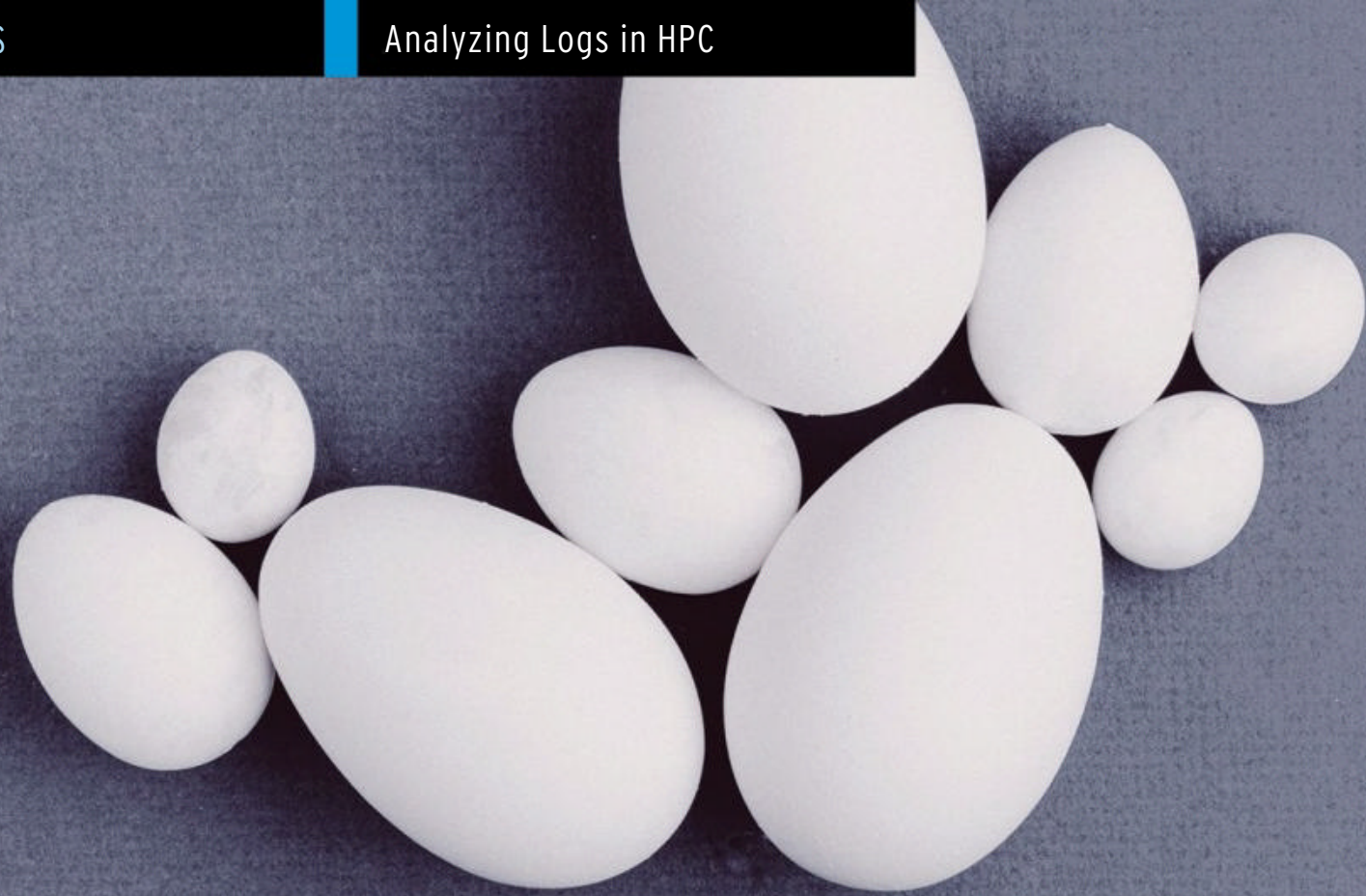
**Listing 4: Resilvering**

```
# zpool status
pool: myvol2
state: ONLINE
status: One or more devices is currently being resilvered. The pool will
continue to function, possibly in a degraded state.
action: Wait for the resilver to complete.
scan: resilver in progress since Mon Oct 24 17:13:44 2022
      15.0G scanned at 3.00G/s, 5.74G issued at 1.15G/s, 15.0G total
      709M resilvered, 38.29% done, 00:00:08 to go
config:

    NAME                STATE        READ WRITE CKSUM
    myvol2               ONLINE       0     0     0
      draid1:3d:5c:1s-0  ONLINE       0     0     0
        sdb              ONLINE       0     0     0 (resilvering)
        sdc              ONLINE       0     0     0 (resilvering)
        sdd              ONLINE       0     0     0 (resilvering)
        sde              ONLINE       0     0     0 (resilvering)
        spare-4          ONLINE       0     0     0
          sdf            ONLINE       3    397     0 (resilvering)
            draid1-0-0    ONLINE       0     0     0 (resilvering)
    spares
      draid1-0-0         INUSE        currently in use

errors: No known data errors
```





Log analysis in high-performance computing

# State of the Cluster

Log analysis can be used to great effect in HPC systems. We present an overview of the current log analysis technologies. By Jeff Layton

**Gathering logs from distributed systems** for manual searching is a typical task performed in high-performance computing (HPC) [1].

Log analysis is important for cybersecurity, understanding HPC cluster behavior, and event and trend analysis. In this article, I address the state of the art in log analysis and how it can be applied to HPC.

## Origins

Log analysis can produce information through a variety of functions and technologies, including:

- ingestion
- centralization
- normalization
- classification and logging
- pattern recognition
- correlation analysis
- monitoring and alerts
- artificial ignorance
- reporting

Logs are great for checking the health of a set of systems and can be used to locate obvious problems, such as kernel modules not loading. They can also be used to find attempts to break into systems through various means, including shared credentials. However, these examples do not really

take advantage of all the information contained in logs: Log analysis can be used to improve system administration skills.

When analyzing or just watching logs over a period of time, you can get a feel for the rhythm of your systems; for example: When do people log in and out of the system? What kernel modules are loaded? What, if any, errors occur (and when)? The answers to these questions allow you to recognize when things don't seem quite right with the systems (events) that "normal" log analysis might miss. A great question is: Why does user X have a new version of an application? Normal log analysis would not care about this query, but perhaps the user needed a new version and could indicate that others might also need the newer version, prompting you to build and make it available to all.

Developing an intuition of how a system or, in HPC, systems behave can take a long time and might be impossible to achieve, but it can also be accomplished by watching logs. If you happen to leave or change jobs, a new admin would have to start from scratch to develop this systems intuition. Perhaps you have a better way

with the use of log analysis on your HPC systems. Before going there, I'll look at the list of technologies presented at the beginning of this article.

## Ingestion and Centralization

The ingestion and centralization step is important for HPC systems because of their distributed nature. Larger systems would use methods that ingest the logs to a dedicated centralized server, whereas smaller systems or smaller logs could use a virtual machine (VM). Of course, these logs need to be tagged with the originating system so they can be differentiated. When you get to the point of log analysis, you really aren't just talking about system logs. Log ingestion also means logs from network devices, storage systems, and even applications that don't always write to /var/log.

A key factor that can be easily neglected in log collection from disparate systems and devices is the time correlation of these logs. If something happens on a system at a particular time, other systems might have been affected, as well. The exact time of the event is needed to search across logs from other systems and devices to correlate all logs. Therefore, enabling the Network Time Protocol (NTP) [2] on all systems is critical.

Photo by Kier... in Sight on Unsplash



## Normalization

Normalization is just what it sounds like: the process of converting all or parts of the log content to the same format – most importantly for items such as the time stamp of the log entry and the IP address.

The tricky bit of normalization is automation. Some logging tools understand formats from various sources, but not all of them. You might have to write a bit of code to convert the log entries into a format suitable for the analysis. The good news is that you only have to do this once for each source of log entries, of which there aren't many.

## Classification and Ordering

You can mix all of your logs into a single file or database, but this can make analysis difficult because the logs contain different types of information. For example, a log entry from a user logging onto a system has information that is different from a network storage device logging data throughput over the last five seconds. You can classify the log messages into different classes, add tags to messages with keywords, or use both techniques for better log organization.

## Pattern Recognition

Pattern recognition tools are typically used to filter incoming messages (log entries) according to a set of known patterns. This method allows common events to be separated out or handled differently from events that do not fit the pattern.

Although it might sound like you wouldn't have to collect and store so much log information, (1) you have to define the benign patterns, and (2) you can lose quite a bit of information (more on that later in the article).

## Correlation Analysis

Correlation analysis is a way to find all of the entries associated with an event, even if they are in different logs or have different tags. This method can be more important than

you might think: If a user's application doesn't run as well today as it did yesterday, you have to determine whether any new events occurred before the latest run. More specifically: What happens when a user's application crashes? Do any events in the logs across various devices explain what could have caused this problem?

## Monitoring and Alerts

Log analysis tools usually include the ability to notify you about events that might require human intervention. These events can be tied to alerts in various forms, such as email or dashboards, so you are promptly notified. A good simple example is the loss of a stream of events to a system log. This usually requires a human to find out why the stream stopped.

A second example is if environmental properties in a data center go beyond their normal levels. Early in my career a small group of engineers wanted to have a meeting in a data center, so they turned off all of the air conditioning units because they were too loud. As the ambient temperature went above a critical temperature, I got several email messages and a beeper page. (That shows you how long ago it was.)

## Artificial Ignorance

Because hardware problems are fairly rare overall, log analysis tools have implemented what they call "artificial ignorance," which uses a machine learning (ML) approach to discard log entries that have been defined as "uninteresting." Commonly, these entries are typical of a system that is operating normally and that don't provide any useful information (however "useful" is defined). The idea is to save log storage by ignoring and even deleting this uninteresting information.

My opinion, with which you can agree or disagree, is that artificial ignorance is not something I would enable for a long time. The uninteresting logs can provide information about how the system typically runs. For example, when do people log in to the system? When do users typically run jobs? A

lot of day-to-day activities are important to know but are lost when artificial ignorance is used.

Although keeping or watching everyday activities might seem like *Groundhog Day*, I feel it is important for understanding the system. Dumping this data before you have a chance to develop this understanding is premature, in my opinion.

Ignoring the uninteresting data could also hinder understanding an event.

In the previous section on Correlation Analysis, when an event occurs, you want to gather as much information around that event as possible.

Artificial ignorance might ignore log entries that are related to the event but appear to be uninteresting. They could even be deleted, handicapping your understanding of the event.

Lastly, this data can be important for other tasks or techniques in log analysis, as a later section will illustrate.

## Reporting

Log analysis tools can create notifications and alerts, but many (most) of the tools can create a report of their analysis, as well. The report is customized to your system and probably your management because you will want to see all system problems – or at the very least, a summary view of the events. Reports can also be your answer to compliance requests, with all the information needed to prove compliance (or show non-compliance). This feature is critical in Europe, where the General Data Protection Regulation (GDPR) topics are very important. If a request to remove specific data has been made, you must be able to prove that it was done. A log analysis tool should be able to confirm this in a report.

## Examples of Log Analysis Stacks and Tools

All of the log analysis tools and stacks are different, flexible, and run in a specific manner. Some are created from various components, usually open source, and others are monolithic tools, so it would be too difficult to examine them all. Rather,



I'm going to focus on a log analysis stack concept that illustrates the tools that fulfill the various tasks described at the beginning of the article. In this way, I hope to orient you in the correct direction for deploying a log analysis capability.

## Splunk

Although I'm a big supporter of open source, Splunk [3], a commercial product, is probably the gold standard for a combination log collection and log analysis tool. It is the template security incident and event management (SIEM) tool that all others use for comparison. Honestly, though, Splunk is pretty costly for an HPC system and may be overkill.

Splunk came out in 2003 and was very quickly a big success. Arguably, it was the first enterprise-grade log collection and analysis tool that could monitor logs from many different devices and applications and locate patterns in real time. It also uses machine learning in its analysis, which was very innovative at the time and set the bar for similar tools. Splunk has a great number of features, is easy to install, uses data indices and events, and can ingest data from many sources. For an enterprise-grade tool, all of these features made it unique and powerful.

Sites started using Splunk, and its popularity grew and grew. However, as I mentioned, it's fairly expensive, particularly for HPC systems, which has led to log analysis stacks and tools developed to compete with Splunk but still be open source, less expensive at the very least, or both.

## Splunk Alternatives

Given the preeminence of Splunk and the associated cost, people naturally will look for Splunk alternatives [4]. An article about Splunk alternatives presents a few options but is not necessarily comprehensive. From that article, I gleaned a few options:

- LogDNA (commercial with open source agent; software as a service (SaaS))

- Elastic Stack (aka the ELK stack), which includes Elasticsearch (search and analytics engine), Logstash (log collection transformation), Kibana (visualization), and Beats (data shippers, not really part of ELK, but added to Elastic Stack)
- Fluentd (open source)
- Sumo Logic (commercial)
- Loggly (uses open source tools, including Elasticsearch, Apache Lucene, and Apache Kafka)
- Graylog (open and commercial versions)

Most of the open source tool stacks have a company behind them that offers support, proprietary plugins, additional capability, or a combination of features for a price.

Rather than dig into the tools and stacks, I'm going to discuss the ELK stack and its components briefly. This stack was the original open source log analysis stack designed to replace Splunk but has morphed into the Elastic Stack while adding an additional tool, Beats, a collection of data shipper tools.

## Elastic Stack (ELK)

ELK stands for Elasticsearch, Logstash, and Kibana and was put together to provide a complete log management and analysis stack that is all open source and competitive with Splunk. The steps or tenets for log collection, management, and analysis and the tools that fulfill these steps in the ELK stack are:

- log collection (Logstash)
- log/data conversion/formatter (Logstash)
- log search/analysis (Elasticsearch)
- visualization (Kibana)

A fourth tool, Beats, a collection of lightweight data shippers, was later added to the stack, which was renamed Elastic Stack.

The ELK stack was a big hit because it was totally open source and provided a good portion of the capability found in Splunk. Its popularity grew quickly and even Amazon Web Services (AWS) offered the ELK stack components as managed services. These components can be used

together and with other AWS services. The Elastic company [5] develops the tools in the ELK and Elastic stacks, offers support, and develops commercial plugins.

## Gathering the Data

Logstash serves the purpose of gathering and transforming logs from the various servers (classification and ordering) by ingesting the logs or data from the specified sources and normalizing, classifying, and ordering it before sending it to the search engine, which is Elasticsearch in the ELK or in Elastic stack.

Installing and configuring Logstash is covered in an article from Elastic [6]. Each client system runs a small tool named Filebeat that collects data from files on that server and sends it to the log server that is running Logstash. This tool allows you to specify system logs or the output of any scripts you create or any applications.

Filebeat takes the place of log gathering tools such as *syslog-ng* or *rsyslog* in general but isn't strictly necessary. If you have a server already logging to a central log, you can install Filebeat on that server and the logs will be transmitted to the Logstash server as JSON (JavaScript Object Notation), which can be the same server, easing the upgrade from just log collection to log analysis.

## Searching the Data

Logstash gathers the logs and transforms them into a form the search engine, Elasticsearch [7], can then consume (JSON). Elasticsearch is an open source tool that is based on the Apache Lucene library [8]. You can configure Elasticsearch to gather whatever information you need or want, but the defaults [9] are a good place to start. Besides being open source, Elasticsearch has some attractive features. One aspect that HPC users will understand and appreciate is that it is distributed. If you have lots of data and want to improve performance, you can shard data across several servers. Elasticsearch also has near real-time performance, perhaps close



to Spunk's performance, that gives you quick insight into problems and perhaps overcomes them as quickly. While doing the searching, it creates an index for the data that is very useful if you want to see all of the data related to an event, look back at the system logs and data, or both.

The core of Elasticsearch is in Java, but it has clients in various languages, including, naturally, Java, but also .NET (C#), PHP, Apache Groovy, Ruby, JavaScript, Go, Perl, Rust, and Python. These choices provide a great deal of flexibility of Elasticsearch.

In addition to developing all the tools in the stack – Logstash, Elasticsearch, Kibana (more on that in the next section), and Beats – Elastic also created the Elastic Cloud service.

## Visualizing the Data

Humans are great pattern recognition engines. We don't do well with text zooming on the screen, but we can pick out patterns in visual data. To help, the tool Kibana [10] ties into Elastic Stack to provide visualization. At a high level, it can create various charts of pretty much whatever you want or need.

Installing Kibana (Figure 1) is easiest from your package manager. If you read through the configuration document, you will see lots of options. I recommend starting with the defaults [11] before changing any of them. The documentation starts with alert and action settings followed by lots of other configuration options.

## ELK Stack Wrap Up

Other components plug into the ELK stack. Fortunately, they have been tested and developed somewhat together so they should “just work.”

Most importantly, however, they fulfill the components of a log analysis system mentioned earlier: log collection, log data conversion and formatter, log search and analysis, and visualization.

## AI

In the previous discussion about the technologies used in log analysis, machine learning was discussed – in particular, artificial ignorance, which uses machine learning to ignore and possibly discard log entries before searching the data. Pattern recognition also uses machine learning. As I discussed, although I don't know the details of how the machine learning models make decisions, I am not a big fan of discarding data just because it looks normal. Such data might be very useful in training deep learning models.

A deep learning model uses neural networks to process input to create some sort of output. These networks are trained with sample data for the inputs and have the matching expected output(s). To train such a model adequately, you need a very large amount of data that spans a wide range of conditions.

The data should be balanced according to the desired outputs for non-language models. For example, if you want to identify an image and have defined 10 possible image classes, then the input data should be fairly evenly distributed across each of the

10 classes. You can see the effect of a bad distribution if you run test images through the model and it has difficulty classifying images into a specific class. It might identify a cat as a dog if you don't have enough data, enough data in each class, or a broad enough data set.

If you are interested in special images or events that happen only very rarely, developing an appropriate dataset is very difficult and, as a result, makes it difficult to create an adequately trained model. An example of this is fraud detection. The model is supposed to identify when a fraudulent transaction happens. However, these are very rare events despite what certain news agencies say.

If you take data from transaction processing, you will have virtually the entire dataset filled with non-fraudulent data. Perhaps only a single-digit number of fraudulent transactions are in the dataset, so you have millions of non-fraudulent transactions and maybe three or four fraudulent transactions – most definitely a very unbalanced dataset.

For these types of situations, you invert the problem by creating a model to find the non-fraudulent transactions. Now you have a data set that is useful in creating the model. You are throwing millions of transactions at the model for it to learn with basically one output: Is the transaction fraudulent? A fraudulent transaction will then be relatively



Figure 1: Sample Kibana screenshot (CC BY-SA 4.0).



easy for the model to detect. Of course, other approaches will work, but this is a somewhat common approach of training models to detect rare events. Therefore, you shouldn't discard the non-interesting data, because it can be used to train a neural network model that is very useful.

In the case of HPC systems, the trained deep learning model can be used to augment your observations as to the "normal" behavior of a system. If something doesn't seem typical, then the model can quickly notify you. In HPC systems, non-typical events can be infrequent. For example, a user starts logging into the system at a later time than usual or they start running different applications. Therefore, the data set for HPC systems could have a fair number of events, and you don't have to "invert" the model to look for non-events.

## NVIDIA Morpheus

I want to mention a new tool from NVIDIA designed for cybersecurity. In full disclosure, I work for NVIDIA as my day job, but I'm not endorsing the product, nor do I have any inside knowledge of the product, so I will use publicly available links. That said, NVIDIA has a new software development kit (SDK) that addresses cybersecurity with neural network models. The SDK, called Morpheus [12], is "... an open application framework that enables cybersecurity developers to create optimized AI pipelines for filtering, processing, and classifying large volumes of real-time data." It provides real-time inferencing (not training) on cybersecurity data.

Log analysis looks for cybersecurity-like events and includes actual cybersecurity events. The product web page lists some possible use cases:

- digital fingerprinting
- sensitive information detection
- crypto mining malware detection
- phishing detection
- fraudulent transaction and identity detection
- ransomware

The digital fingerprinting use case "Uniquely fingerprint(s) every user,

service, account, and machine across the enterprise data center – employing unsupervised learning to flag when user and machine activity patterns shift," which is great to watch for break-ins on HPC systems but also to watch for shifting patterns. Instead of developing a fingerprint of user behavior, it could be used to create a fingerprint of application versions.

For example, a user might start out using a specific version of an application, but as time goes on, perhaps they use a newer version. This event signals the administrator to install and support the new version and consider deprecating any old versions. This scenario is especially likely in deep learning applications because frameworks develop quickly, new frameworks are introduced, and other frameworks stop being developed. Another example might be job queues that are becoming longer than usual, perhaps indicating that more resources are needed. A model that detects this event and creates information as to why would be extremely useful.

The framework could also be used to watch data storage trends. Certainly storage space will increase, but which users are consuming the most space or have the most files can be identified and watched in case it is something unusual (like downloading too many KC and The Sunshine Band mp4s to the HPC). There really are no limits to how Morpheus could be used in HPC, especially for log analysis in general.

## Summary

Log analysis is a very useful tool for many administration tasks. You can use it for cybersecurity, understanding how an HPC cluster normally behaves, identifying events and trends within the cluster, the need for more resources, or anything you want to learn about the cluster. The use of log analysis in your HPC systems is up to you because it can mean adding several servers and a fair amount of storage to your

system. However, don't think of log analysis as just a cybersecurity tool. You can use it for many HPC-specific tasks, greatly adding to the administration of the system. Plus it can make pretty reports, which management always loves.

The future of log analysis will probably morph into an AI-based tool that takes the place of several of the technologies in current log analysis. Instead of a single trained AI, a federated set of trained networks will probably be used. Other models will likely go back and review past logs, either for training or to create a behavior description of how the cluster operates. This area of HPC technology has lots of opportunities. ■

### Info

- [1] "Log Management" by Jeff Layton: [\[https://www.admin-magazine.com/HPC/Articles/Log-Management\]](https://www.admin-magazine.com/HPC/Articles/Log-Management)
- [2] NTP: [\[https://en.wikipedia.org/wiki/Network\\_Time\\_Protocol\]](https://en.wikipedia.org/wiki/Network_Time_Protocol)
- [3] Splunk: [\[https://www.splunk.com/en\\_us/blog/learn/log-management.html\]](https://www.splunk.com/en_us/blog/learn/log-management.html)
- [4] Splunk alternatives: [\[https://www.mezmo.com/blog/5-splunk-alternatives-for-logging-their-benefits-shortcomings-and-which-one-to-choose\]](https://www.mezmo.com/blog/5-splunk-alternatives-for-logging-their-benefits-shortcomings-and-which-one-to-choose)
- [5] Elastic: [\[https://www.elastic.co/\]](https://www.elastic.co/)
- [6] Logstash: [\[https://www.elastic.co/guide/en/logstash/current/getting-started-with-logstash.html\]](https://www.elastic.co/guide/en/logstash/current/getting-started-with-logstash.html)
- [7] Elasticsearch: [\[https://en.wikipedia.org/wiki/Elasticsearch\]](https://en.wikipedia.org/wiki/Elasticsearch)
- [8] Lucene: [\[https://lucene.apache.org/\]](https://lucene.apache.org/)
- [9] Elasticsearch defaults: [\[https://www.elastic.co/guide/en/elasticsearch/reference/current/settings.html\]](https://www.elastic.co/guide/en/elasticsearch/reference/current/settings.html)
- [10] Kibana: [\[https://en.wikipedia.org/wiki/Kibana\]](https://en.wikipedia.org/wiki/Kibana)
- [11] Kibana defaults: [\[https://www.elastic.co/guide/en/kibana/current/settings.html\]](https://www.elastic.co/guide/en/kibana/current/settings.html)
- [12] Morpheus: [\[https://developer.nvidia.com/morpheus-cybersecurity\]](https://developer.nvidia.com/morpheus-cybersecurity)

### The Author

Jeff Layton has been in the HPC business for over 30 years (starting when he was 4 years old). When he's not grappling with a stubborn systemd script, he's looking for deals for his home cluster. His twitter handle is @JeffdotLayton.







System temperature as a  
dimension of performance

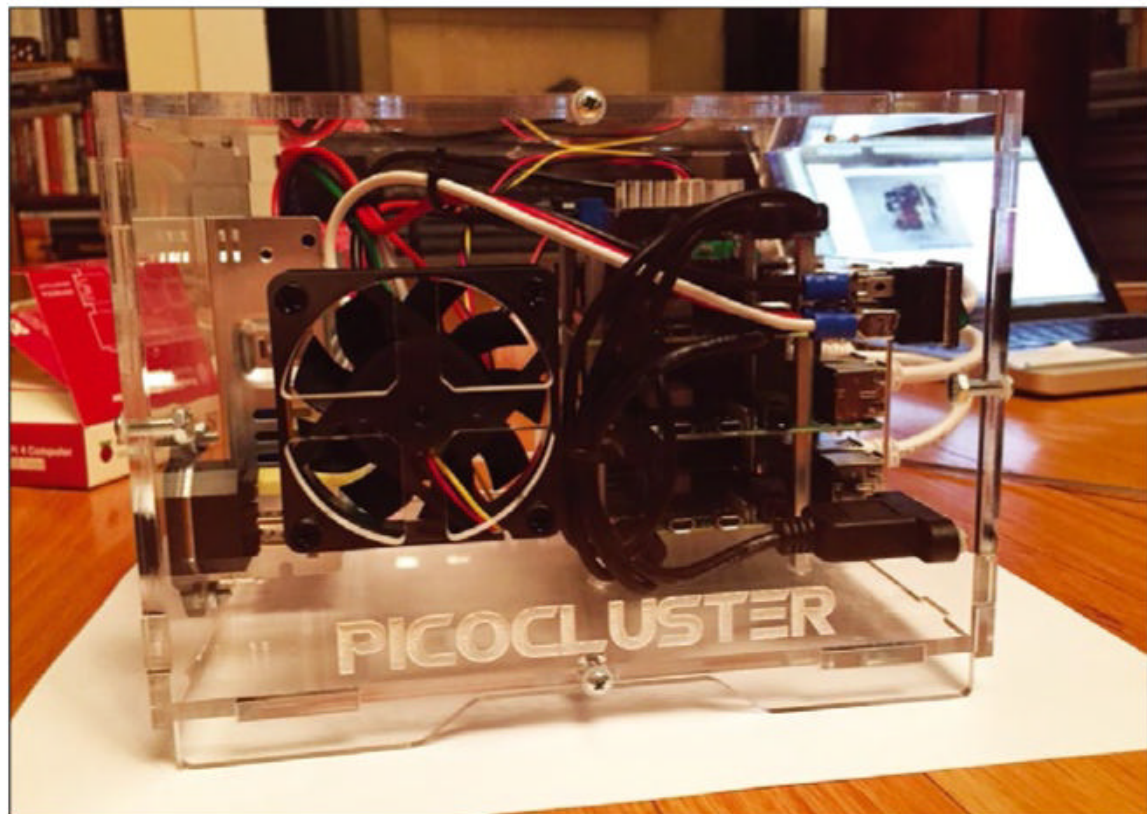
# Heat Seeker

Sensor tools can provide highly variable data from CPUs, GPUs, and a variety of sources. We look at some tools to verify the temperature of components on diverse hardware. By Federico Lucifredi

**Excessive or just elevated temperature** is sometimes the source of unexpected behavior in computing components – I faced it firsthand once, with runaway CPU heat burning through an older Slot I Pentium 3 processor. The cause was a detached heat sink, and some older hardware did not have built-in protection circuitry back then. I have also had fun experiencing intriguing boot failures from an overheating hard drive in a system equipped with one too many peripherals. In recent times I have taken a more preemptive stance, monitoring the heat build-up in a Raspberry Pi cluster as the case fan was replaced (**Figure 1**) [1]. To silence a desktop cluster, I replaced the built-in fan with a slower, silent fan made by specialty vendor Noctua [2]. Because the newer fan used fewer revolutions per minute to get the job done, I had to verify that the temperature inside the case remained roughly the same after the change:

```
vcgencmd measure_temp
```

The output quickly demonstrated that the temperature remained within the same range it had with the older fan. The `vcgencmd` utility is a tool made by Broadcom to access the state of the VideoCore GPU found on all Raspberry Pi boards [3]. It can also provide insight into voltage levels for cores and SDRAM, among other things, which can come handy



**Figure 1:** Raspberry Pi 4 Picocluster in testing.

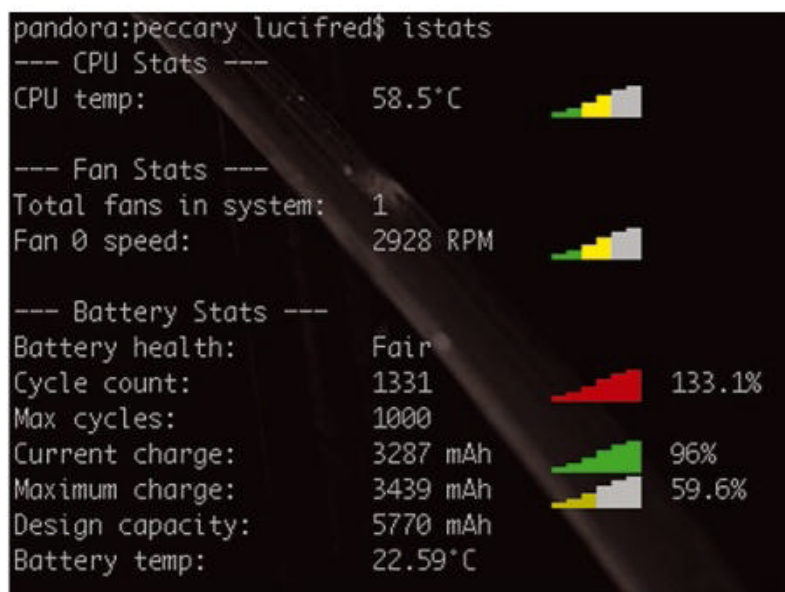
```

federico@pc0: ~
File Edit Tabs Help
federico@pc0:~ $ vcgencmd measure_temp
temp=45.0'C
federico@pc0:~ $ parallel-ssh -h nodes -i "vcgencmd measure_temp"
[1] 13:24:53 [SUCCESS] pc0
temp=45.0'C
[2] 13:24:53 [SUCCESS] pc2
temp=41.0'C
[3] 13:24:53 [SUCCESS] pc1
temp=42.0'C
federico@pc0:~ $ parallel-ssh -h nodes -i "vcgencmd measure_volts core"
[1] 13:25:24 [SUCCESS] pc0
volt=0.8683V
[2] 13:25:24 [SUCCESS] pc1
volt=0.8455V
[3] 13:25:24 [SUCCESS] pc2
volt=0.8595V
federico@pc0:~ $

```

**Figure 2:** `vcgencmd` extracting temperature and voltage information from a Raspberry Pi GPU. Note the use of parallel SSH syntax to access all nodes in a cluster.





**Figure 3:** iStats at work on a Mac. The chart of the number of battery cycles of this old battery is actually blinking in red!

while troubleshooting heat problems (Figure 2).

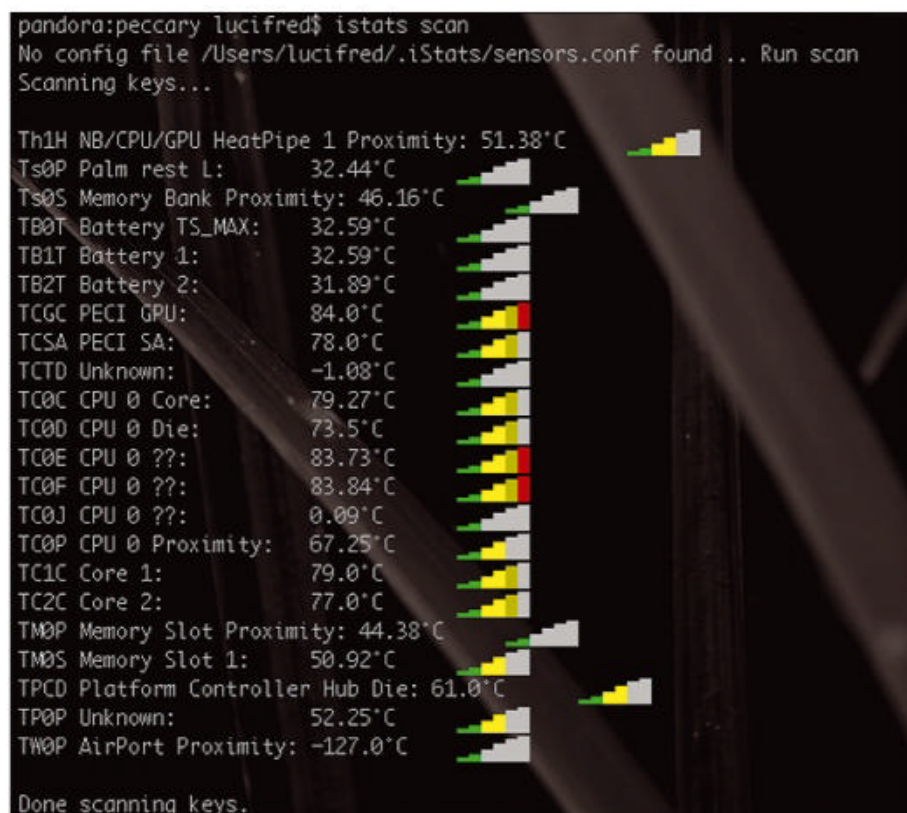
## Terminal UI Redux

Returning to a common topic on these pages, an interesting example of a terminal user interface (TUI) [4] for the Mac is the `istats` tool (install from Ruby Gems with `gem install iStats`) [5]. iStats provides convenient access to temperature, battery, and fan speed data on macOS systems (Figure 3). Its simple and easy-to-use interface belies a hidden complexity. The

`istats scan` command will look for additional “keys” of

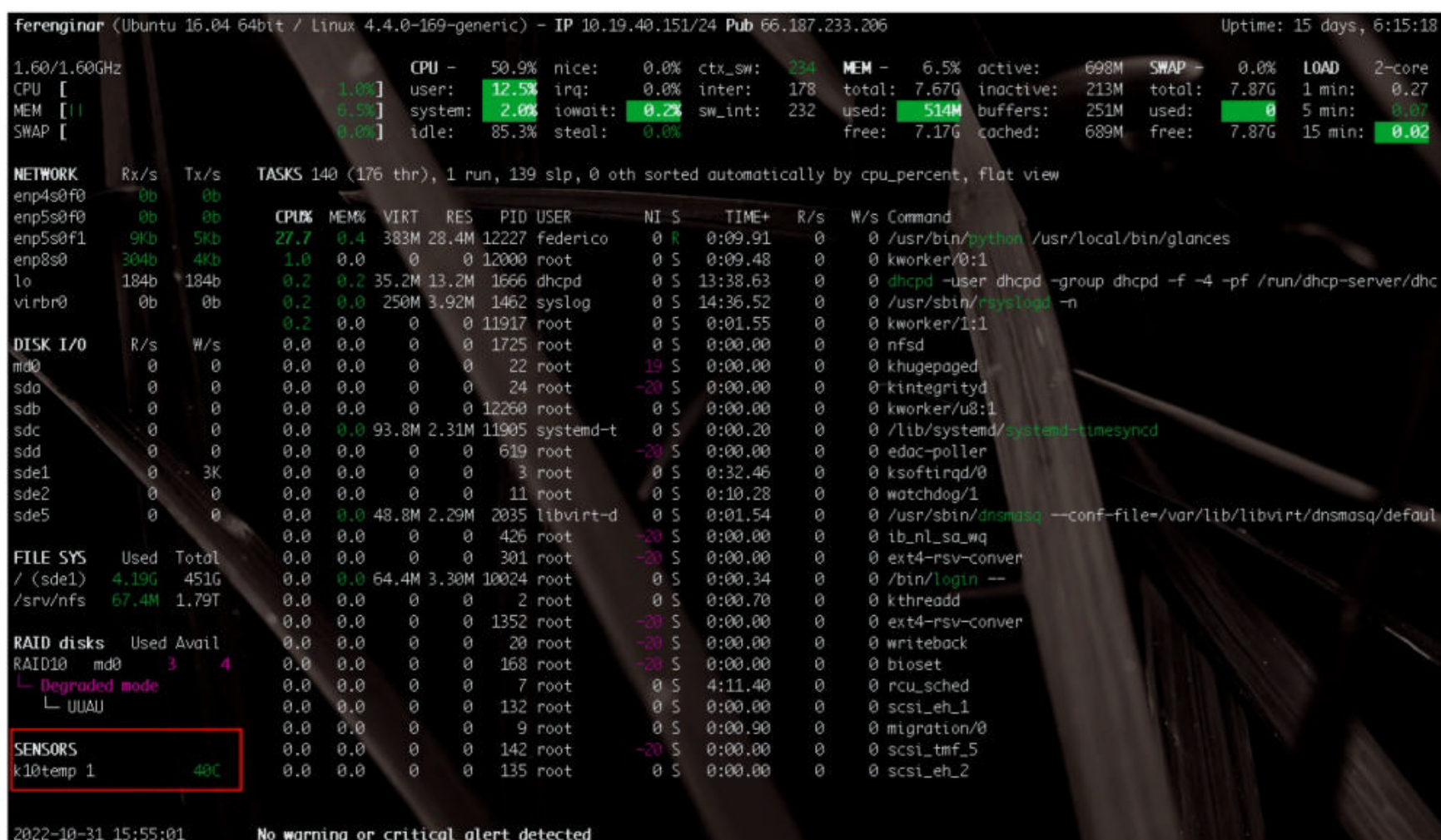
sensor data available on a given system and offer to enable them at the user’s discretion (Figure 4).

On Linux, my default choice for in-terminal monitoring TUI is `glances` [6], which I have examined previously for its many capabilities [7]. Yet, I did not look at its temperature module, an oversight Figure 5 now aims to remedy. Data is sourced from the



**Figure 4:** iStats scanning sensors available on an older MacBook Pro. Not all values discovered are temperatures.

`sensors` [8] command (`apt install lm-sensors`), which is a configurable, consistent command-line interface gathering the multitude sensor data exposed by a disparate array of Linux kernel modules (Listing 1; from the same system on which `glances` was run). Indeed, if this tour of sensor tools and interfaces has shown one thing, it is that the variety of



**Figure 5:** `glances` presents temperature data on its left-hand panel whenever it is available.



```

federico@voronoi: ~$ i7z
Cpu speed from cpufreq 2399.00MHz
cpufreq might be wrong if cpufreq is enabled. To guess correctly try estimating via tsc
Linux's inbuilt cpu_khz code emulated now
True Frequency (without accounting Turbo) 2399 MHz
CPU Multiplier 24x || Bus clock frequency (BCLK) 99.96 MHz

Socket [0] - [physical cores=2, logical cores=4, max online cores ever=2]
TURBO ENABLED on 2 Cores, Hyper Threading ON
Max Frequency without considering Turbo 2498.96 MHz (99.96 x [25])
Max TURBO Multiplier (if Enabled) with 1/2/3/4 Cores is 38x/37x/37x/37x
Real Current Frequency 398.97 MHz [99.96 x 3.99] (Max of below)
┌ Core [core-id] : Actual Freq (Mult.)    C0%  Halt(C1)%  C3 %  C6 %  Temp  VCore
└ Core 1 [0]:      398.89 (3.99x)        2.55  97.6      1    1    34    0.6284
  Core 2 [1]:      398.97 (3.99x)        1.61  98.7      0    1    35    0.6288

C0 = Processor running without halting
C1 = Processor running with halts (States >C0 are power saver modes with cores idling)
C3 = Cores running with PLL turned off and core cache turned off
C6, C7 = Everything in C3 + core state saved to last level cache, C7 is deeper than C6
Above values in table are in percentage over the last 1 sec

```

Figure 6: i7z at work, constantly refreshing CPU state data, including temperatures.

sensor data available is highly variable. Ranging from a Raspberry Pi, to a vintage MacBook Pro, to an HP

Microserver, every system has provided different datasets. Often the biggest challenge is correctly identifying

the source of the data, which can originate in a CPU, a GPU, a management card, and a variety of alternative sources.

## Let There Be Dell

Switching to a fourth system, this one running Linux on an Intel processor, I can make use of i7z [9], a tool explicitly written to visualize the state of Intel processors – including their temperature.

```

federico@voronoi: ~$ sensors
ucsi_source_psy_USBC000:001-isa-0000
Adapter: ISA adapter
in0:      5.00 V (min = +5.00 V, max = +5.00 V)
curr1:    0.00 A (max = +0.00 A)

ath10k_hmon-pci-3a00
Adapter: PCI adapter
temp1:    +32.0°C

coretemp-isa-0000
Adapter: ISA adapter
Package id 0: +44.0°C (high = +100.0°C, crit = +100.0°C)
Core 0:      +43.0°C (high = +100.0°C, crit = +100.0°C)
Core 1:      +44.0°C (high = +100.0°C, crit = +100.0°C)

BAT0-acpi-0
Adapter: ACPI interface
in0:      8.12 V
curr1:    574.00 mA

dell_smm-isa-0000
Adapter: ISA adapter
Processor Fan: 0 RPM
CPU:      +44.0°C
Ambient:  +36.0°C
Other:    +32.0°C
Other:    +39.0°C
SODIMM:   +39.0°C

pch_skylake-virtual-0
Adapter: Virtual device
temp1:    +40.0°C

nvme-pci-3c00
Adapter: PCI adapter
Composite: +42.9°C (low = -20.1°C, high = +84.8°C)
           (crit = +81.8°C)
Sensor 1:  +42.9°C (low = -20.1°C, high = +84.8°C)

acpiitz-acpi-0
Adapter: ACPI interface
temp1:    +25.0°C (crit = +107.0°C)

```

Figure 7: The complete range of sensors available on a Dell XPS 13 9360 "Sputnik."

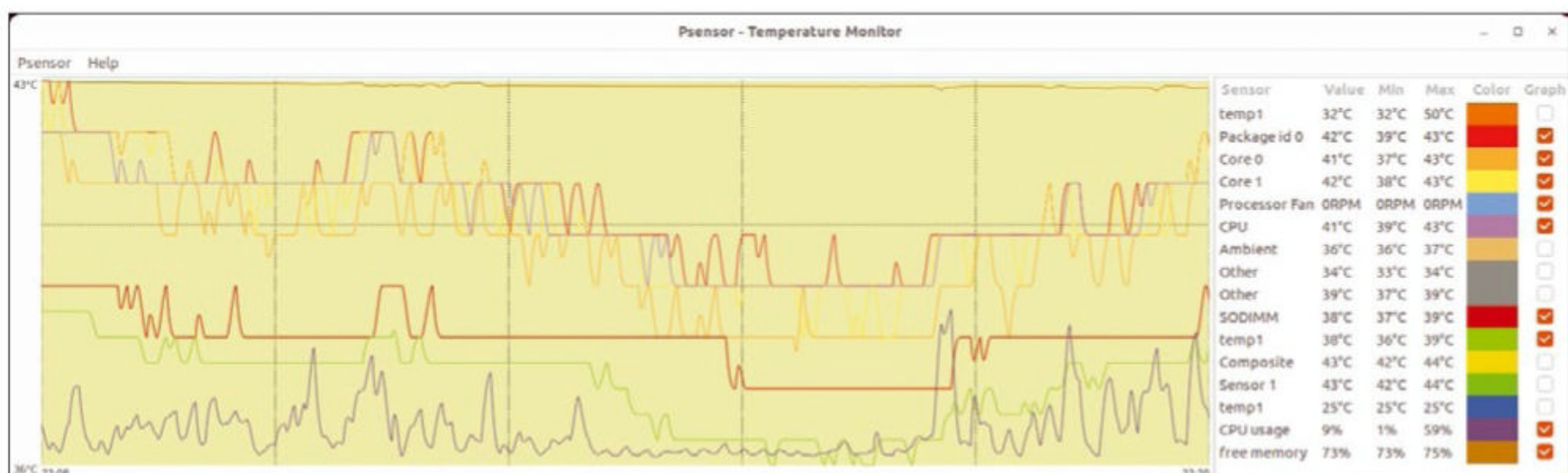


Figure 8: psensor charting some of the sensors listed in Figure 7.

### Listing 1: sensors Output

```

federico@ferenginar:~$ sensors
k10temp-pci-00c3
Adapter: PCI adapter
temp1:      +42.5°C (high = +70.0°C)
              (crit = +100.0°C, hyst = +95.0°C)

federico@ferenginar:~$

```

### Listing 2: Extracting Temperature Data

```

federico@voronoi:~$ sudo smartctl -A /dev/nvme0 | grep
Temperature
[sudo] password for federico:
Temperature:                               42 Celsius
Warning Comp. Temperature Time:           16
Critical Comp. Temperature Time:          28
Temperature Sensor 1:                      42 Celsius
federico@voronoi:~$

```

Figure 6 shows the results on an Intel Core-equipped laptop. The more general purpose solution is to configure lm-sensors to include kernel modules for all sensors available, which is done through the sensors-detect command in a way akin to what was previously described for iStats. With lm-sensors, you can actually configure the kernel to load the correct modules at boot time, as recently described by our own Bruce Byfield in sister publication *Linux Magazine* [10]. Once this configuration step is successfully accomplished, all data for a specific system becomes available in the usual filesystem locations (usually in /sys/), ready to be consumed by higher level programs (Figure 7). One such example is psensor [11], a GUI application charting select metrics and sensors to facilitate time series analysis. Not perfectly polished in its windowing toolkit interactions,



at least in Ubuntu 22.04, it nonetheless remains the best option currently available (Figure 8).

## Disk Sensors

The all-purpose `smartctl` utility offers storage device temperature data (`apt install smartmontools`) [12] equally well for spinning media and solid-state NVMe storage (Listing 2; from a persistent storage SMART source). It is generally advisable to maintain storage media under 60°C (140°F), with the actual drive's specification having the final say. ■

### Info

- [1] Sound-proofing a Picocluster: [\[https://twitter.com/0xF2/status/1244422315011645444\]](https://twitter.com/0xF2/status/1244422315011645444)

- [2] Noctua NF-A6x25 PWM, Premium Quiet Fan, 4-Pin (60mm): [\[https://www.amazon.com/gp/product/B00VXTANZ4/\]](https://www.amazon.com/gp/product/B00VXTANZ4/)
- [3] `vcgencmd` on Embedded Linux: [\[https://elinux.org/RPI\\_vcgencmd\\_usage\]](https://elinux.org/RPI_vcgencmd_usage)
- [4] "Network Performance In-Terminal Graphics Tools" by Federico Lucifredi, *ADMIN*, issue 51, 2019, pg 94, [\[https://www.admin-magazine.com/Archive/2019/51/Network-performance-in-terminal-graphics-tools\]](https://www.admin-magazine.com/Archive/2019/51/Network-performance-in-terminal-graphics-tools)
- [5] Christophe Naud-Dulude, *iStats*: [\[https://github.com/Chris911/iStats\]](https://github.com/Chris911/iStats)
- [6] Nicolas Hennion, *glances*: [\[https://github.com/nicolargo/glances\]](https://github.com/nicolargo/glances)
- [7] "Next-Generation Terminal UI Tools" by Federico Lucifredi, *ADMIN*, issue 64, 2021, pg. 93, [\[https://www.admin-magazine.com/Archive/2021/64/Next-generation-terminal-UI-tools\]](https://www.admin-magazine.com/Archive/2021/64/Next-generation-terminal-UI-tools)
- [8] `lm-sensors` repository: [\[https://github.com/lm-sensors/lm-sensors\]](https://github.com/lm-sensors/lm-sensors)

- [9] Abhishek Jaientilal, *i7z*:

[\[https://github.com/ajaintilal/i7z\]](https://github.com/ajaintilal/i7z)

- [10] "Taking Your Hardware's Temperature" by Bruce Byfield, *Linux Magazine*, issue 261, August 2022, pg. 44, [\[https://www.linuxpromagazine.com/Issues/2022/261/Beat-the-Heat\]](https://www.linuxpromagazine.com/Issues/2022/261/Beat-the-Heat)

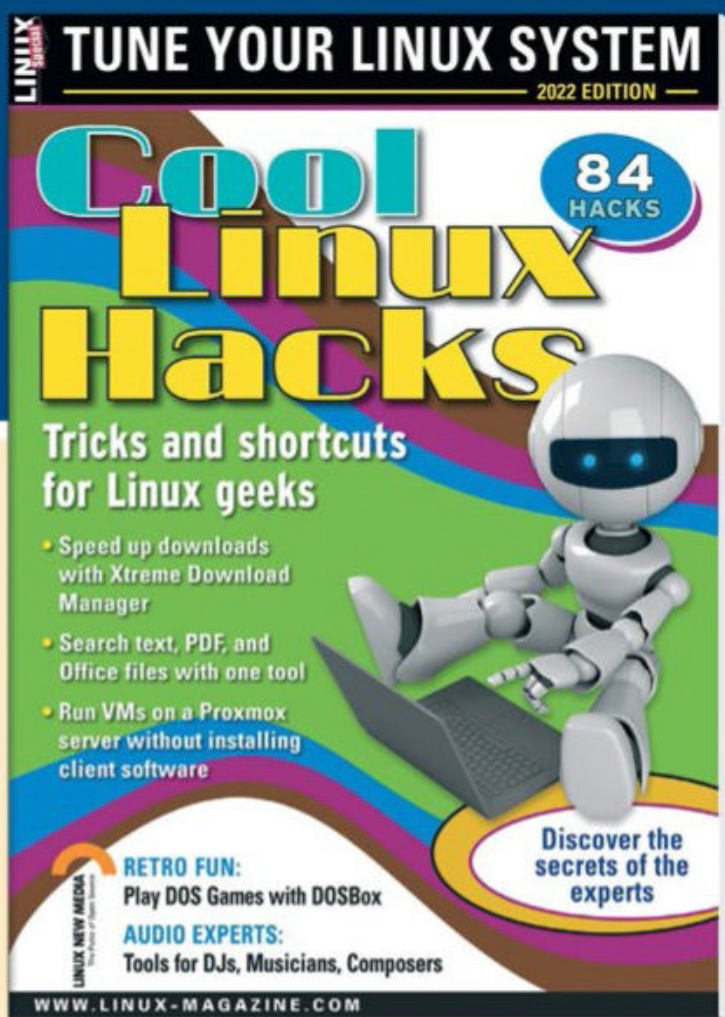
- [11] Francis Chin, *psensor*:

[\[https://github.com/chinf/psensor\]](https://github.com/chinf/psensor)

- [12] `smartmontools`: [\[https://github.com/smartmontools/smartmontools/\]](https://github.com/smartmontools/smartmontools/)

### Author

Federico Lucifredi (@0xf2) is the Product Management Director for Ceph Storage at Red Hat and formerly the Ubuntu Server Product Manager at Canonical and the Linux "Systems Management Czar" at SUSE. He enjoys arcane hardware issues and shell-scripting mysteries and takes his McFlurry shaken, not stirred. You can read more from him in the new O'Reilly title *AWS System Administration*.



SHOP THE SHOP  
[shop.linuxnewmedia.com](http://shop.linuxnewmedia.com)

GET PRODUCTIVE WITH  
COOL LINUX HACKS

Improve your Linux skills with this cool collection of inspirational tricks and shortcuts for Linux geeks.

- Google on the Command Line
- OpenSnitch Application Firewall
- Parse the `systemd` journal
- Control Git with `lazygit`
- Run Old DOS Games with DOSBox
- And more!



ORDER ONLINE: [shop.linuxnewmedia.com/specials](http://shop.linuxnewmedia.com/specials)



# ADMIN

## Network & Security

# NEWSSTAND

Order online:  
[bit.ly/ADMIN-Newsstand](https://bit.ly/ADMIN-Newsstand)

*ADMIN* is your source for technical solutions to real-world problems. Every issue is packed with practical articles on the topics you need, such as: security, cloud computing, DevOps, HPC, storage, and more! Explore our full catalog of back issues for specific topics or to complete your collection.

### #71 - September/October 2022

#### Kubernetes

We show you how to get started with Kubernetes, and users share their insights into the container manager.

On the DVD: SystemRescue 9.04



### #70 - July/August 2022

#### Defense by Design

Nothing is so true in IT as “Prevention is better than the cure.” We look at three ways to prepare for battle.

On the DVD: Rocky Linux 9 (x86\_64)



### #69 - May/June 2022

#### Terraform

After nearly 10 years of work on Terraform, the HashiCorp team delivers the 1.0 version of the cloud automation tool.

On the DVD: Ubuntu 22.04 “Jammy Jellyfish” LTS server Edition



### #68 - March/April 2022

#### Automation in the Enterprise

Automation in the enterprise extends to remote maintenance, cloud orchestration, and network hardware

On the DVD: AlmaLinux 8.5 (minimal)



### #67 - January/February 2022

#### systemd Security

This issue, we look at how to secure systemd services and its associated components.

On the DVD: Fedora 35 Server (Install)



### #66 - November/December 2021

#### Incident Analysis

We look at updating, patching, and log monitoring container apps and explore The Hive + Cortex optimization.

On the DVD: Ubuntu 21.10 “Impish Indri” Server Edition





# WRITE FOR US

*Admin: Network and Security* is looking for good, practical articles on system administration topics. We love to hear from IT professionals who have discovered innovative tools or techniques for solving real-world problems.

Tell us about your favorite:

- interoperability solutions
- practical tools for cloud environments
- security problems and how you solved them
- ingenious custom scripts

- unheralded open source utilities
- Windows networking techniques that aren't explained (or aren't explained well) in the standard documentation.

We need concrete, fully developed solutions: installation steps, configuration files, examples – we are looking for a complete discussion, not just a “hot tip” that leaves the details to the reader.

If you have an idea for an article, send a 1-2 paragraph proposal describing your topic to: [edit@admin-magazine.com](mailto:edit@admin-magazine.com).



Authors	
Amber Ankerholz	6
Ulrich Bantle	10
Klaus Bierschenk	64
Chris Cowen	80
Ken Hess	3
Christian Knermann	40
Petros Koutoupis	84
Felix Kronlage-Dammers	20
Martin Kuppinger	76
Jeff Layton	88
Martin Gerhard Loschwitz	14, 24, 50
Federico Lucifredi	93
Ali Imran Nagori	46
Akshat Pradhan	58
Dr. Holger Reibold	36
Andreas Stolzenberger	30, 70
Matthias Wübbeling	56, 62

Contact Info

**Editor in Chief**  
Joe Casad, [jcasad@linuxnewmedia.com](mailto:jcasad@linuxnewmedia.com)

**Managing Editors**  
Rita L Sooby, [rsooby@linuxnewmedia.com](mailto:rsooby@linuxnewmedia.com)  
Lori White, [lwhite@linuxnewmedia.com](mailto:lwhite@linuxnewmedia.com)

**Senior Editor**  
Ken Hess

**Localization & Translation**  
Ian Travis

**News Editor**  
Amber Ankerholz

**Copy Editors**  
Amy Pettie, Aubrey Vaughn

**Layout**  
Dena Friesen, Lori White

**Cover Design**  
Dena Friesen, Illustration based on graphics by Jakarin Niamklang, [123RF.com](http://123RF.com)

**Advertising**  
Brian Osborn, [bosborn@linuxnewmedia.com](mailto:bosborn@linuxnewmedia.com)  
phone +49 8093 7679420

**Publisher**  
Brian Osborn

**Marketing Communications**  
Gwen Clark, [gclark@linuxnewmedia.com](mailto:gclark@linuxnewmedia.com)  
Linux New Media USA, LLC  
4840 Bob Billings Parkway, Ste 104  
Lawrence, KS 66049 USA

**Customer Service / Subscription**  
For USA and Canada:  
Email: [cs@linuxnewmedia.com](mailto:cs@linuxnewmedia.com)  
Phone: 1-866-247-2802  
(Toll Free from the US and Canada)

For all other countries:  
Email: [subs@linuxnewmedia.com](mailto:subs@linuxnewmedia.com)  
[www.admin-magazine.com](http://www.admin-magazine.com)

While every care has been taken in the content of the magazine, the publishers cannot be held responsible for the accuracy of the information contained within it or any consequences arising from the use of it. The use of the DVD provided with the magazine or any material provided on it is at your own risk.

Copyright and Trademarks © 2022 Linux New Media USA, LLC.

No material may be reproduced in any form whatsoever in whole or in part without the written permission of the publishers. It is assumed that all correspondence sent, for example, letters, email, faxes, photographs, articles, drawings, are supplied for publication or license to third parties on a non-exclusive worldwide basis by Linux New Media unless otherwise stated in writing.

All brand or product names are trademarks of their respective owners. Contact us if we haven't credited your copyright; we will always correct any oversight.

Printed in Nuremberg, Germany by Zeitfracht GmbH.

Distributed by Seymour Distribution Ltd, United Kingdom

ADMIN is published bimonthly by Linux New Media USA, LLC, 4840 Bob Billings Parkway, Ste 104, Lawrence, KS 66049, USA (Print ISSN: 2045-0702, Online ISSN: 2831-9583). November/December 2022. Periodicals Postage paid at Lawrence, KS. Ride-Along Enclosed. POSTMASTER: Please send address changes to ADMIN, 4840 Bob Billings Parkway, Ste 104, Lawrence, KS 66049, USA.

Represented in Europe and other territories by: Sparkhaus Media GmbH, Bialasstr. 1a, 85625 Glonn, Germany.



# CLOUDFEST

March 21-23, 2023  
Europa-Park, Germany

**6,000+ Participants**  
**250+ Speakers**

**150+ Partners**  
**65 Countries**

The world's largest cloud industry event is ready to once again take over a spectacular European amusement park to facilitate new partnerships, deep knowledge sharing, and the best parties the industry has ever seen.



**Start your CloudFest Journey**  
**AND SAVE €399!**

With your FREE Code: **CF23ADMIN**

scan me!



[reg.cloudfest.com](https://reg.cloudfest.com)



## Workstation Edition



Iris Xe Graphics  
GeForce RTX 3050 Ti



90 Hz refresh rate



DDR4-3200 MHz



HDMI 2.0 (4K@60 Hz)

## Max Performance Edition



GeForce RTX 3060  
GeForce RTX 3070 Ti



240 Hz refresh rate



DDR5-4800 MHz



HDMI 2.1 (4K@120 Hz)

# Ultra slim workstation goes BIG!

## TUXEDO InfinityBook Pro 16 - Gen7



100%  
Linux

5

Year  
Warranty



Lifetime  
Support



Built in  
Germany



German  
Privacy



Local  
Support

# TUXEDO 18<sup>th</sup> COMPUTERS ANNIVERSARY

[tuxedocomputers.com](https://tuxedocomputers.com)