SAS® Intelligence Platform

Overview

Second Edition

THE POWER TO KNOW®

Contents

C H A P T E R

1

Value of the SAS Intelligence Platform

What is the SAS Intelligence Platform?

The SAS Intelligence Platform is a comprehensive, end-to-end infrastructure for creating, managing, and distributing enterprise intelligence. The platform includes tools and interfaces that enable you to do the following:

- extract data from a variety of operational data sources on multiple platforms, and build a data warehouse and data marts that integrate the extracted data
- store large volumes of data efficiently and in a variety of formats
- give business users at all levels the ability to explore data from the warehouse in a Web browser, perform simple query and reporting functions, and view up-to-date results of complex analyses

□ use high-end analytic techniques to provide capabilities such as predictive and descriptive modeling, forecasting, optimization, simulation, and experimental design

□ centrally control the accuracy and consistency of enterprise data

It would be possible to build an enterprise intelligence infrastructure using applications from different vendors that specialize in specific areas. However, the implementation process would be extremely complex and time consuming. With the SAS Intelligence Platform, you can implement a fully integrated, end-to-end intelligence infrastructure by using software that is delivered, tested, and integrated by SAS. The platform draws from your existing enterprise data, and it is designed to work in even the most complex and heterogeneous information technology environments.

Using the tools provided in the SAS Intelligence Platform, you can create applications that reflect your unique business requirements and domain knowledge. The platform also serves as the foundation for SAS business solutions in areas such as customer intelligence, financial intelligence, and supply chain intelligence, as well as for SAS solutions for vertical markets such as financial services, life sciences, health care, retail, and manufacturing.

Accessibility Features in the SAS Intelligence Platform Products

For information about accessibility for any of the products mentioned in this book, see the documentation for that product. If you have questions or concerns about the accessibility of SAS products, send e-mail to **accessibility@sas.com**.

Components of the SAS Intelligence Platform

The SAS Intelligence Platform includes components in the following categories:

Data Integration and ETL
The data integration and extract, transform, and load (ETL) components enable you to consolidate and manage enterprise data from a variety of source systems, applications, and technologies. Components are provided to help you cleanse, migrate, synchronize, replicate, and promote your data. Metadata for all of your intelligence resources is stored centrally and controlled through a single management interface.

Business Intelligence
The business intelligence components enable users with various needs and skill levels to create, produce, and share their own reports and analyses. Through easy-to-use interfaces, users can obtain their own answers to business questions. Meanwhile, Information Technology retains control over the quality and consistency of the data.

Analytics
SAS offers the richest and widest portfolio of analytic products in the software industry. The portfolio includes products for statistical data analysis, data and text mining, forecasting, econometrics, quality improvement, and operations research. You can use any combination of these tools with the SAS Intelligence Platform to add extraordinary precision and insight to your reports and analyses.

Intelligence Storage
The intelligence storage options are optimized for analytical processing, enabling you to quickly retrieve and report on large

volumes of data. The options include simple relational databases, a threaded multidimensional database that supports online analytical processing (OLAP), and relational storage with a threaded multiple input/output (I/O) subsystem for intensive use by focused applications.

The following sections describe the data integration and ETL, business intelligence, analytics, and intelligence storage components in more detail.

Data Integration and ETL

The software tools in the data integration and ETL category enable you to consolidate and manage enterprise data from a variety of source systems, applications, and technologies. The data sources can include SAS data sets, database management system (DBMS) tables, and data from enterprise resource planning (ERP) systems. Metadata for all of your intelligence resources is stored centrally in the SAS Metadata Repository.

Each of the data integration and ETL tools is described briefly in the following sections.

SAS Data Integration Studio

SAS Data Integration Studio is a visual design tool that enables information architects to consolidate and manage enterprise data. The software enables you to create jobs and process flows that extract, transform, and load (ETL) data for use in data warehouses and data marts. You can also create processes that cleanse, migrate, synchronize, replicate, and promote your data.

For more information, see Chapter 6, "Clients in the SAS Intelligence Platform," on page 33.

SAS/ACCESS

SAS/ACCESS provides interfaces to a wide range of relational databases. With this product, SAS Data Integration Studio and other SAS applications can read, write, and update data regardless of which database and platform the data is stored on. SAS/ACCESS interfaces provide fast, efficient data loading and enable SAS applications to work directly from your data sources without making a copy.

SAS Data Surveyor

The SAS Data Surveyor applications enable you to build SAS Data Integration Studio jobs to help you read data in the ERP systems of vendors such as SAP and Oracle.

SAS Data Quality Server

SAS Data Quality Server analyzes, cleanses, and standardizes your data. The software uses parsing, standardization, and matching algorithms to analyze your data and help you identify and reduce inconsistencies, inaccuracies, and redundancies.

This product is often used in conjunction with products such as DataFlux Corporation's dfPower Studio, which enables you to customize the Quality Knowledge Base that the SAS Data Quality Server uses to store its data-cleansing guidelines.

Platform Process Manager

Platform Process Manager provides scheduling capabilities, including scheduling the execution of jobs, managing dependencies between jobs, and submitting jobs to Platform

LSF (Load Sharing Facility). Platform LSF manages resource requirements and provides for load balancing. The Platform Grid Management Service includes a monitoring daemon that enables administrators to monitor the load on machines running Platform LSF.

SAS Metadata Repository

All of your information assets are managed in a common metadata layer called the SAS Metadata Repository.

This repository stores logical data representations of items such as libraries, information maps, and cubes, thus ensuring central control over the quality and consistency of data definitions and business rules. The repository also stores information about system resources such as servers, the users who access data and metadata, and the rules that govern who can access what.

All of the data integration, ETL, and business intelligence tools read and use metadata from the repository and create new metadata as needed.

SAS Management Console

SAS Management Console provides an single interface through which system administrators can explore and manage metadata repositories. With this interface, administrators can efficiently set up system resources, manage user and group accounts, and administer security.

Business Intelligence

The software tools in the business intelligence category address two main functional areas: information design, and self-service reporting and analysis.

The information design tools enable business analysts and information architects to organize data in ways that are meaningful to business users, while shielding the end users from the complexities of underlying data structures. These tools include the following products:

- *SAS Information Map Studio* enables analysts and information architects to create and manage information maps that contain business metadata about your physical data.
- *SAS OLAP Cube Studio* enables information architects to create cube definitions that organize summary data along multiple business dimensions.

The self-service reporting and analysis tools enable business users to query, view, and explore centrally stored information. Users can create their own reports, graphs, and analyses in the desired format and level of detail. In addition, they can find, view, and share previously created reports and analyses. The tools feature intuitive interfaces that enable business users to perform these tasks with minimal training and without the involvement of information technology staff.

The self-service reporting and analysis tools include the following products:

- *SAS Web Report Studio* is a Web-based query and reporting tool that enables users at any skill level to create, view, and organize reports.
- *SAS Information Delivery Portal* provides a Web-based, personalized workplace to help decision makers easily find the information they need.
- *SAS Web OLAP Viewer* is a stand-alone Web-based application that uses SAS Information Maps to provide interactive and powerful navigation of OLAP data.
- *SAS Add-In for Microsoft Office* enables users to access SAS functionality from within Microsoft Office products.

□ *SAS Enterprise Guide* is a project-oriented Windows application that enables users to create processes that include complex computations, business logic, and algorithms.

As users create information maps, cubes, report definitions, portal content definitions, and stored processes, information about them is stored in the SAS Metadata Repository. Client applications and users can access these information assets on a need-to-know basis. Access is controlled through multilayered security that is enforced through the metadata.

For a description of each of the business intelligence tools, see Chapter 6, "Clients in the SAS Intelligence Platform," on page 33.

Intelligence Storage

The data storage options that are available with the SAS Intelligence Platform include SAS data tables, parallel storage, multidimensional databases, and third-party hierarchical and relational databases such as DB2 and Oracle. These storage options can be used alone or in any combination.

All metadata for your data sources is stored centrally in the SAS Metadata Repository for use by other components of the intelligence platform.

Each of the options is described briefly below.

Relational Storage: SAS Data Sets

You can use SAS data sets, the default SAS storage format, to store data of any granularity. The data values in a SAS data set are organized as a table of observations (rows) and variables (columns). A SAS data set also contains descriptor information such as the data types and lengths of the columns, as well as which SAS engine was used to create the data.

Multidimensional Storage: SAS OLAP Server

SAS OLAP Server provides dedicated storage for data that has been summarized along multiple business dimensions. The server uses a threaded, scalable, and open technology and is especially designed for fast-turnaround processing and reporting.

A simplified ETL process enables you to build consistent OLAP cubes from disparate systems. A threaded query engine and parallel storage enable data to be spread across multiple-disk systems. Support is provided for multidimensional (MOLAP) and hybrid (HOLAP) data stores, as well as for open industry standards.

Parallel Storage: SAS Scalable Performance Data Engine and SAS Scalable Performance Data Server

SAS SPD Engine and SAS SPD Server provide a high-speed data storage alternative for processing very large SAS data sets. They read and write tables that contain millions of observations, including tables that exceed the 2-GB size limit imposed by some operating systems. In addition, they provide the rapid data access that is needed to support intensive processing by SAS analytic software and procedures.

These facilities work by organizing data into a streamlined file format and then using threads to read blocks of data very rapidly and in parallel. The software tasks are performed in conjunction with an operating system that enables threads to execute on any of the CPUs that are available on a machine.

The SAS SPD Engine, which is included with Base SAS software, is a single-user data storage solution. SAS SPD Server, which is available as a separate product or as

part of the SAS Intelligence Storage bundle, is a multi-user solution that includes a comprehensive security infrastructure, backup and restore utilities, and sophisticated administrative and tuning options.

Third-party Hierarchical and Relational Databases

The SAS Intelligence Platform can access data that is stored in third-party party hierarchical and relational databases such as DB2, Oracle, SQL Server, and NCR Teradata. SAS/ACCESS interfaces provide fast, efficient data loading and enable SAS applications to work directly from your data sources without making a copy.

Several of the SAS/ACCESS engines use an I/O subsystem that enables you to read entire blocks of data instead of reading data just one record at a time. This feature, which reduces I/O bottlenecks and enables procedures to read data as fast as they can process it, is included in the SAS/ACCESS engines for Oracle, Sybase, DB2 (UNIX and PC), ODBC, SQL Server, and Teradata.

Analytics

SAS offers the richest and widest portfolio of analytic products in the software industry. The portfolio includes products for statistical data analysis, data and text mining, forecasting, econometrics, quality improvement, and operations research. You can use any combination of these tools with the SAS Intelligence Platform to add precision and insight to your reports and analyses.

One such tool is SAS Enterprise Miner, which enables analysts to create and manage data mining process flows. These flows include steps to examine, transform, and process data to create models that predict complex behaviors of economic interest. The SAS Intelligence Platform enables SAS Enterprise Miner users to centrally store and share the metadata for models and projects. In addition, SAS Data Integration Studio provides the ability to schedule data mining jobs.

SAS software provides the following types of analytical capabilities:

- statistical data analysis, to drive fact-based decisions
- data and text mining, to build descriptive and predictive models and deploy the results throughout the enterprise
- forecasting, to analyze and predict outcomes based on historical patterns
- econometrics, to apply statistical methods to economic data, problems and trends
- quality improvement, to identify, monitor, and measure quality processes over time
- operations research, to apply techniques such as optimization, scheduling, and simulation to achieve the best result

Strategic Benefits of the SAS Business Intelligence Platform

Multiple Capabilities Integrated into One Platform

The SAS Intelligence Platform combines advanced SAS analytics, high-speed processing of large amounts of data, and easy-to-use query and reporting tools. The result is accurate, reliable, and fast information with which to make decisions.

You can build data warehouses, perform data mining, enable users to query data and produce reports from a Web browser, and give users easy access to SAS processes that perform robust analytics.

The SAS Intelligence Platform provides all of this functionality in one centrally managed suite of products that are designed to work together seamlessly. This integration reduces the administration, management, and deployment costs that would be associated with providing multiple technologies to meet the needs of different users.

Consistency of Data and Business Rules

The SAS Intelligence Platform makes use of your organization's existing data assets, enabling you to integrate data from multiple database platforms and ERPs. Tools are provided to help ensure the reliability, consistency, and standardization of this data.

Users can choose from multiple tools with which to perform queries and produce reports. Since all of the tools access data through the same metadata representations, users throughout your enterprise receive consistent data. As a result, they can make decisions based on a common version of the truth.

Similarly, business logic, complex computations, and analytic algorithms can be developed once and stored centrally in SAS processes for all users to access. These processes, as well as the information in the SAS Metadata Repository, are controlled through multi-level security.

Fast and Easy Reporting and Analysis

The SAS Intelligence Platform's self-service reporting and analysis tools enable users across the enterprise to access and query data from virtually any data source. Any number of users can use wizards to create reports in the needed time frames, without waiting for support from information technology professionals. Through Web-based interfaces, users can explore large volumes of multidimensional data quickly and interactively, from multiple perspectives and at multiple levels of detail.

The reporting and analysis tools hide complex data structures, so that average business users can perform queries without having to learn new skills. The intelligence storage options are optimized for analytical processing, enabling the reporting tools to quickly retrieve large volumes of data.

As a result of these reporting and analysis capabilities, everyone spends less time looking for answers and more time driving strategic decisions.

Analytics Available to All Users

SAS is the market leader in analytics. With the SAS Intelligence Platform, you can make the full breadth of SAS analytics available to users throughout the enterprise.

SAS analytics include algorithms for functions such as predictive and descriptive modeling, forecasting, optimization, simulation, and experimental design. You can now incorporate these capabilities into self-service reports and analyses, so that decision makers throughout your enterprise can benefit from the accuracy and precision of high-end analytics.
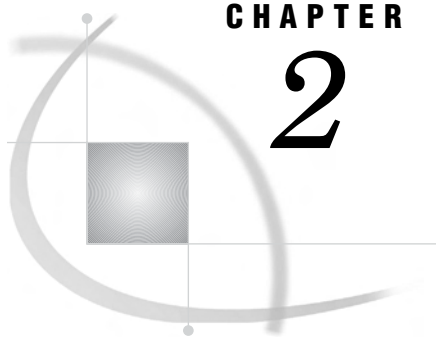
C H A P T E R

2

Architecture of the SAS Intelligence Platform

Architecture of the SAS Intelligence Platform

The SAS Intelligence Platform architecture is designed to efficiently access large amounts of data, while simultaneously providing timely intelligence to a large number of users. The platform uses an n-tier architecture that enables you to distribute functionality across computer resources, so that each type of work is performed by the resources that are best suited to the job.

You can easily scale the architecture to meet the demands of your workload. For a large company, the tiers can be installed across a multitude of machines with different operating systems. For prototyping, demonstrations, or very small enterprises, all of the tiers can be installed on a single machine.

The architecture consists of the following four tiers:

Data Sources   Data sources store your enterprise data. All of your existing data assets can be used, whether your data is stored in relational database management systems, SAS tables, or ERP system tables.

SAS Servers   SAS servers perform SAS processing on your enterprise data. Several types of SAS servers are available to handle different workload types and processing intensities. The software distributes processing loads among server resources so that multiple client requests for information can be met without delay.

Middle Tier   The middle tier enables users to access intelligence data and functionality via a Web browser. This tier provides Web-based interfaces for report creation and information distribution, while passing analysis and processing requests to the SAS servers.

Clients   The client tier provides users with desktop access to intelligence data and functionality through easy-to-use interfaces. For most information consumers, reporting and analysis tasks can be performed with just a Web browser. For more advanced design and analysis tasks, SAS client software is installed on users' desktops.

Note: The four tiers listed above represent categories of software that perform similar types of computing tasks and require similar types of resources. The tiers do not necessarily represent separate computers or groups of computers. △

The following diagram shows how the tiers interact, and the sections that follow describe each tier in more detail.

Figure 2.1   Architecture of the SAS Intelligence Platform
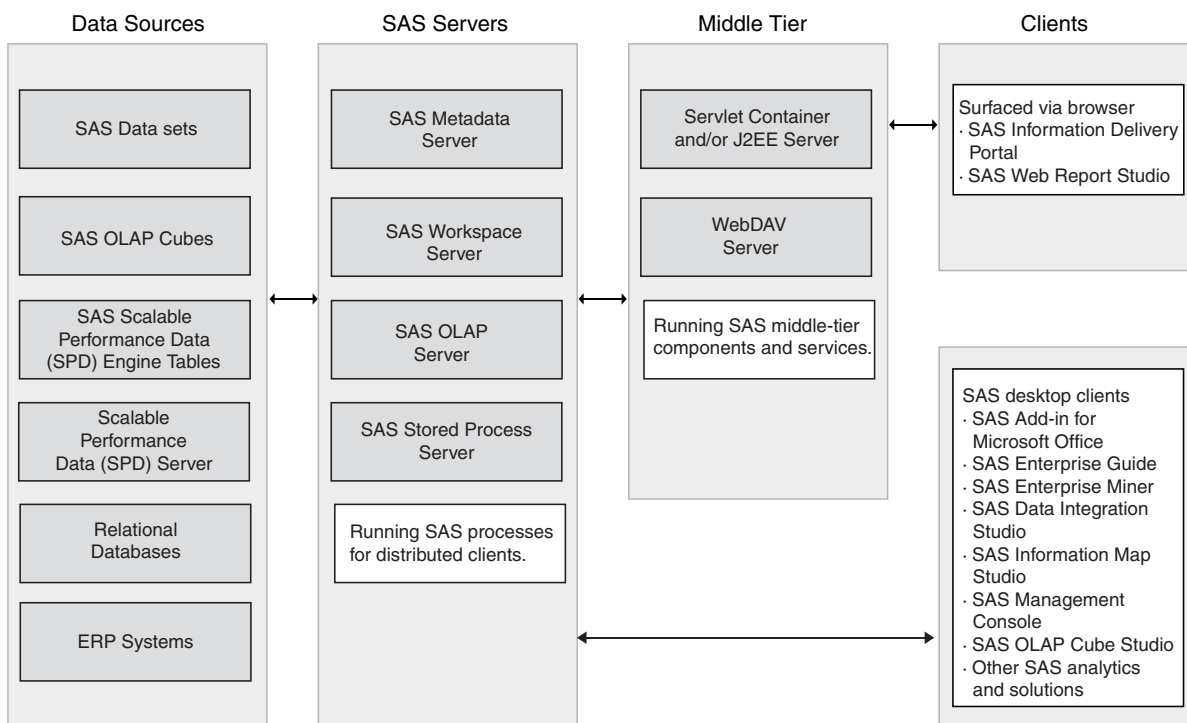
Data Sources

The SAS Intelligence Platform includes the following options for data storage:

□ SAS data sets, which are analogous to relational database tables

□ SAS SPD Engine tables, which can be read or written by multiple threads

□ SAS SPD Server, which is available as a separate product or as part of the SAS Intelligence Storage bundle

□ SAS OLAP cubes

In addition, SAS provides products that enable you to access data in your existing third-party DBMSs and ERP systems. The SAS/ACCESS interfaces provide direct access to DBMSs such as the following:

☐ Oracle

☐ DB2

☐ Teradata

☐ ODBC

☐ Microsoft SQL Server

☐ Sybase

The SAS Data Surveyor products provide direct access to ERP systems such as the following:

☐ SAP (R3 and SAP BW data types)

☐ Oracle applications

For more information about data sources, see Chapter 3, "Data in the SAS Intelligence Platform," on page 15.

SAS Servers

The SAS servers execute SAS analytical and reporting processes for distributed clients. These servers are typically accessed either by desktop clients or by Web applications that are running in the middle tier.

Note: In the SAS Intelligence Platform, the term *server* refers to a program or programs that wait for and fulfill requests from client programs for data or services.

The term *server* does not necessarily refer to a specific computer, since a single computer can host one or more servers of various types. △

The SAS servers use the SAS Integrated Object Model (IOM), which is a set of distributed object interfaces that make SAS software features available to client applications when SAS is executed on a server. Each server uses a different set of IOM interfaces and has a different purpose.

The principal servers in the SAS Intelligence Platform include the SAS Metadata Server, the SAS OLAP Server, the SAS Workspace Server, and the SAS Stored Process Server.

SAS Metadata Server

The *SAS Metadata Server* controls access to a central repository of metadata that is shared by all of the applications in the system. The SAS Metadata Server enables centralized control so that all users access consistent and accurate data. The metadata repository stores information about the following:

☐ the enterprise data sources and data structures that are accessed by SAS intelligence applications.

☐ the products that are created and used by SAS applications. These products include information maps, OLAP cubes, report definitions, stored process definitions, and portal content definitions.

☐ the SAS and third-party servers that participate in the system.

☐ the users and groups of users that use the system.

☐ the levels of access that users and groups have to resources.

The SAS Intelligence Platform provides a central management tool—SAS Management Console—that you use to manage the metadata server and the metadata repository.

SAS OLAP Server

The *SAS OLAP Server* is a multidimensional data server that delivers pre-summarized cubes of data to business intelligence applications. The data is queried using the MDX (multidimensional expression) language.

This server is designed to reduce the load on traditional back-end storage systems by quickly delivering summarized views, irrespective of the amount of data that underlies the summaries.

SAS Workspace Server

The *SAS Workspace Server* enables client applications to submit SAS code to a SAS session using an application programming interface (API). For example, when you use SAS Data Integration Studio to submit an ETL job for processing, the application generates the SAS code necessary to perform the processing and submits it to a workspace server. You can run as many instances of workspace servers as are needed to support your workload.

SAS Stored Process Server

The *SAS Stored Process Server* executes and delivers results from SAS Stored Processes in a multi-client environment. A SAS Stored Process is a SAS program that is stored centrally and that can be executed by users and client programs on demand.

You can run as many instances of stored process servers as are needed to support your workload.

For more information about the SAS servers, see Chapter 4, "Servers in the SAS Intelligence Platform," on page 21.

Middle Tier

The middle tier provides an execution environment for business intelligence Web applications such as SAS Web Report Studio and SAS Information Delivery Portal. These applications communicate with the user by sending data to and receiving data from the user's Web browser. For example, an application of this type displays its user interface by sending an HTML document to the user's browser. The user can submit input to the application by sending it an HTTP response, usually by clicking a link or submitting an HTML form.

The middle tier includes the following third-party software and SAS software elements:

☐ a servlet container or J2EE server

☐ the Java 2 Software Development Kit, Standard Edition (J2SE SDK)

☐ a WebDAV (Web-Based Distributed Authoring and Versioning) server

☐ SAS Application Services

☐ SAS Foundation Services

☐ the SAS Web Infrastructure Kit

For more information about the middle tier, see Chapter 5, "Middle-Tier Components of the SAS Intelligence Platform," on page 27.

Clients

The clients in the SAS Intelligence Platform provide Web-based and desktop user interfaces to content and applications.

SAS clients provide access to content, appropriate query and reporting interfaces, and business intelligence functionality for all of the information consumers in your enterprise, from the CEO to business analysts to customer service agents.

The software on the client tier includes Windows applications, Java applications, and a Web browser.

The following clients are Windows applications and run on Microsoft Windows systems:

- SAS Add-In for Microsoft Office
- SAS Enterprise Guide

The following clients are Java applications that are supported on Microsoft Windows systems:

- SAS Enterprise Miner
- SAS Data Integration Studio
- SAS Information Map Studio
- SAS Management Console
- SAS OLAP Cube Studio

SAS Management Console is also supported on Solaris, HP-UX Itanium, and AIX.

The following products require only a Web browser to be installed on each client machine:

- SAS Information Delivery Portal
- SAS Web OLAP Viewer
- SAS Web Report Studio

For more information about the clients, see Chapter 6, "Clients in the SAS Intelligence Platform," on page 33.

CHAPTER

3

Data in the SAS Intelligence Platform

Overview of Data Storage Options

In a SAS Intelligence Platform deployment, you can use one or more of these data storage options:

- default SAS storage in the form of SAS tables
- third-party hierarchical and relational database tables such as DB2, Oracle, SQL Server, and NCR Teradata
- parallel storage from the SAS Scalable Performance Data Engine (SPD Engine) and the SAS Scalable Performance Data Server (SPD Server)
- multidimensional databases (cubes)

All four data sources provide input to reporting applications. The first three sources are also used as input for these data structures:

- cubes, which are created with either SAS Data Integration Studio or SAS OLAP Cube Studio
- data marts and data warehouses, which are created with SAS Data Integration Studio

You can use these storage options in any combination to meet your unique business requirements. The following sections describe each storage option in more detail. Central management of data sources through the SAS metadata repository is also discussed.

Default SAS Storage

You can use SAS data sets (tables), the default SAS storage format, to store data of any granularity.

A SAS table is a file that SAS software creates and processes. Each SAS table is a member of a SAS library. A SAS library is a collection of one or more SAS files that are recognized by SAS software and that are referenced and stored as a unit.

Each SAS table contains the following:

- □ data values that are organized as a table of observations (rows) and variables (columns) that can be processed by SAS software

- □ descriptor information such as data types, column lengths, and the SAS engine that was used to create the data

Third-Party Relational Data Storage

Data can be stored in third-party hierarchical and relational databases such as DB2, Oracle, SQL Server, and NCR Teradata. SAS/ACCESS interfaces provide fast, efficient loading of data to and from these facilities. With these interfaces, SAS software can work directly from the data sources without making a copy.

Several of the SAS/ACCESS engines use an input/output (I/O) subsystem that enables applications to read entire blocks of data instead of reading just one record at a time. This feature reduces I/O bottlenecks and enables procedures to read data as fast as they can process it. The SAS/ACCESS engines for the following products support this functionality:

- □ Oracle

- □ Sybase

- □ DB2 (UNIX and PC)

- □ ODBC

- □ SQL Server

- □ Teradata

These engines, as well as the DB2 engine on z/OS, can also access database management system (DBMS) data in parallel by using multiple threads to the parallel DBMS server. Coupling the threaded SAS procedures with these SAS/ACCESS engines provides even greater gains in performance.

Parallel Storage

The SAS Scalable Performance Data Engine (SPD Engine) and the SAS Scalable Performance Data Server (SPD Server) are designed for high-performance data delivery. They enable rapid access to SAS data for intensive processing by the application.

While the Base SAS engine is sufficient for most tables that do not span volumes, SAS SPD Engine and SAS SPD Server are high-speed alternatives for processing very large tables. They read and write tables that contain millions of observations, including tables that expand beyond the 2-GB size limit imposed by some operating systems. In addition, they support SAS analytic software and procedures that require fast processing of tables.

Options for Implementing Parallel Storage

Two options are available for implementing parallel storage:

☐ The SAS SPD Engine is included with Base SAS software. It is a single-user data storage solution that shares the high-performance parallel processing and parallel I/O capabilities of SAS SPD Server, but lacks the additional complexity of a multi-user server.

   The SAS SPD Engine runs on UNIX, Windows, z/OS (on HFS and zFS file systems only), and OpenVMS for Integrity Servers (on ODS-5 file systems only) platforms.

☐ The SAS SPD Server is available as a separate product or as part of the SAS Intelligence Storage bundle. It is a multi-user parallel-processing data server with a comprehensive security infrastructure, backup and restore utilities, and sophisticated administrative and tuning options.

   The SAS SPD Server runs on Tru64 UNIX, Windows Server, HP-UX, and Sun Solaris platforms.

How Parallel Storage Works

The SAS SPD Engine and SAS SPD Server deliver data to applications rapidly by organizing large SAS data sets into a streamlined file format. The file format enables multiple CPUs and I/O channels to perform parallel input/output (I/O) functions on the data. Parallel I/O takes advantage of multiple CPUs and multiple controllers, with multiple disks per controller, to read or write data in independent threads.

One way to take advantage of the features of the SAS SPD Engine is through a hardware and software architecture known as symmetric multiprocessing (SMP). An SMP machine has multiple CPUs and an operating system that supports threads. These machines are usually configured with multiple controllers and multiple disk drives per controller.

When the SAS SPD Engine reads a data file, it launches one or more threads for each of the CPUs in the SMP machine. These threads then read data in parallel from multiple disk drives, driven by one or more controllers per CPU. The SAS SPD Engine running on an SMP machine provides the capability to read and deliver much more data to an application in a given elapsed time.

Multidimensional Databases (Cubes)

Multidimensional databases (cubes) are another storage option provided by the SAS Intelligence Platform. Cubes provide business users with multiple views of their data through drill-down capabilities.

Cubes are derived from source data such as SAS tables, SAS SPD Engine tables, and SAS/ACCESS database tables. To create cube definitions, and to build cubes based on these definitions, you can use the Cube Designer wizard, which is available from SAS Data Integration Studio and SAS OLAP Cube Studio.

Cubes are managed by the SAS OLAP Server, which is a multi-user, scalable server designed to store and access large volumes of data, while maintaining system performance.

The SAS OLAP Server uses a SAS engine that organizes data into a streamlined file format. This file format enables the engine to rapidly deliver data to client applications. The engine also reads and writes partitioned tables, which enables it to use multiple

CPUs to perform parallel I/O functions. The threaded model enables the SAS OLAP Server to create and query aggregations in parallel for fastest performance.

SAS business intelligence applications perform queries against the cubes by using the multidimensional expression (MDX) query language. Cubes can be accessed by client applications that are connected to the SAS OLAP Server with the following tools:

□ the SQL Pass-Through Facility for OLAP, which is designed to process MDX queries within the PROC SQL environment

□ open access technologies such as OLE DB for OLAP and ADO MD

How Data Sources are Managed in the Metadata

All of the data sources that are used in your implementation of the SAS Intelligence Platform are centrally controlled through metadata that is stored in the SAS Metadata Repository. In the metadata repository, you can create the following types of metadata objects to control and manage your data:

□ database servers, which provide relational database services to clients

□ database schemas, which are maps or models of the structure of a database

□ SAS Application Servers, which perform SAS processes on data

□ cubes

□ OLAP schemas, which specify which groups of cubes a given SAS OLAP Server can access

□ dimensions and measures in a cube

□ libraries, which are collections of one or more files that are recognized by SAS software and that are referenced and stored as a unit

□ the data sources (for example, SAS tables) that are contained in a library

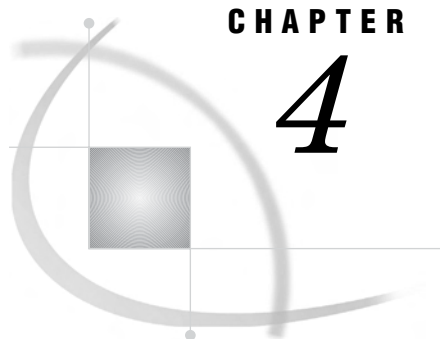□ the columns that are contained in a data source

A variety of methods are available to populate the metadata repository with these objects, including the following:

□ The data source design applications, SAS Data Integration Studio and SAS OLAP Cube Studio, automate the creation of all of the necessary metadata about your data sources. As you use these products to define warehouses, data marts, and cubes, the appropriate metadata objects are automatically created and stored in the metadata repository.

□ You can use the following features of SAS Management Console to define data source objects:

□ The New Server Wizard enables you to easily define the metadata for your database servers and SAS Application Servers.

□ The Data Library Manager enables you to define database schemas for a wide variety of schema types. You can also use this feature to define libraries if you are not using SAS Data Integration Studio to define them.

□ The Import Tables feature enables you to import table definitions from external sources if you are not using SAS Data Integration Studio to create them.

□ The Authorization Manager enables you to secure access to all of the metadata objects. Because all SAS Intelligence Platform applications use the metadata server when accessing resources, permissions that are enforced by the metadata server provide an effective level of protection.

Note: Securing the metadata object that represents the data source is not the same as securing access to the underlying data. Most SAS Intelligence Platform applications enable users to view the underlying data if the users have access to the metadata object that represents the data source and all of its parent objects. △

□ You can use the metadata LIBNAME engine, which enables a Base SAS program to read a data source and write table metadata to the metadata repository.

For detailed information about administering data sources, see the *SAS Intelligence Platform: Data Administration Guide*.

C H A P T E R

4

Servers in the SAS Intelligence Platform

Overview of Servers

The SAS Intelligence Platform provides access to SAS functionality through the following specialized servers:

☐ the SAS Metadata Server, which writes metadata objects to, and reads metadata objects from, SAS Metadata Repositories. These metadata objects contain information about all of the components of your system, such as users, groups, data libraries, servers, and user-created products such as reports, cubes, and information maps.

☐ SAS Workspace Servers, which provide access to SAS software features such as the SAS language, SAS libraries, the server file system, results content, and formatting services.

A program called the object spawner runs on a workspace server's host machine. The spawner listens for incoming client requests and launches server instances as needed.

☐ SAS Stored Process Servers, which execute SAS Stored Processes. Stored processes are SAS programs that are stored on a server and can be executed as required by requesting applications. Stored process servers have MultiBridge connections, which enable multiple processes on different ports of the same server.

A program called the object spawner runs on a stored process server's host machine. The spawner listens for incoming client requests and launches server instances as needed.

☐ SAS OLAP Servers, which provide access to cubes. Cubes are logical sets of data that are organized and structured in a hierarchical multidimensional arrangement. Cubes are queried by using the multidimensional expression (MDX) language.

□ SAS/CONNECT servers, which provide computing resources on remote machines where SAS Integration Technologies is not installed.

□ batch servers, which give you the ability to execute code in batch mode. There are three types of batch servers: DATA step batch servers, Java batch servers, and generic batch servers. The DATA step server enables you to run SAS DATA steps and procedures in batch mode. The Java server enables you to schedule the execution of Java code, such as the code that creates a SAS Marketing Automation marketing campaign. The generic server supports the execution of any other type of code.

Note: In the SAS Intelligence Platform, the term *server* refers to a program or programs that wait for and fulfill requests from client programs for data or services. The term *server* does not necessarily refer to a specific computer, since a single computer can host one or more servers of various types. △

Note: For accessing specialized data sources, the SAS Intelligence Platform can also include one or more data servers. These might include the SAS Scalable Performance Data (SPD) Server and third-party database management system (DBMS) products. The SAS OLAP Server also provides some data server functionality. For information about data servers, see Chapter 3, "Data in the SAS Intelligence Platform," on page 15. △

The following sections describe:

□ the central role of the SAS Metadata Server in the management of the SAS Intelligence Platform

□ the organizational principles that are used to manage SAS server resources, including server objects, logical servers, and application servers

□ the use of load balancing (for stored process servers and workspace servers) and workspace pooling (for workspace servers)

SAS Metadata Server

The SAS Metadata Server is a multi-user server that enables users to read metadata from and write metadata to one or more SAS Metadata Repositories. This server is a centralized resource for storing, managing, and delivering metadata for all SAS applications across the enterprise.

About the Metadata in the SAS Metadata Repository

Here are examples of the types of metadata objects that can be stored in the SAS Metadata Repository:

□ users
□ groups of users
□ data libraries
□ tables
□ jobs
□ cubes
□ documents

□ information maps

□ reports

□ stored processes

□ SAS Workspace Servers

□ SAS Stored Process Servers

□ SAS OLAP Servers

A metadata object is a set of attributes that describe a resource. Here are some examples:

□ When a user creates a report in SAS Web Report Studio, a metadata object is created to describe the new report.

□ When a user creates a data warehouse in SAS Data Integration Studio, a metadata object is created to describe each table included in the warehouse.

□ When a system administrator defines a new instance of a SAS server, a metadata object is created to describe the server.

The specific attributes that a metadata object includes vary depending on the resource that is being described. For example, a table object can includes attributes for the table's name and description, path specification, host type, and associated SAS Application Server.

The SAS Metadata Server uses the SAS Open Metadata Architecture and the Common Warehouse Metamodel (CWM), which provide common metadata services to SAS and other applications. Third parties can access metadata in the SAS Metadata Server using an application programming interface (API) that is supplied by SAS. SAS supports CWM as a standard for metadata interchange.

How the Metadata Server Controls System Access

The SAS Metadata Server plays an important role in the security of the SAS Intelligence Platform. It controls system access in the following ways:

□ The server contains a metadata object called a *metadata identity* for every user of the SAS Intelligence Platform. The object includes each user's login information, including a user ID and an encrypted password. When a user logs on to a SAS application, the application verifies the user's identity by checking it against the metadata identity. The metadata identity also includes information about the groups that each user is part of.

□ Every metadata object includes authorization information that controls which users have which permissions for accessing the metadata object (for example, reading and writing the metadata that describes a server).

In some cases, the authorization information also controls which users have which permissions for accessing the resource itself (for example, accessing a specific server).

□ Trusted peer session connections enable a SAS process (such as a SAS Workspace Server or SAS Stored Process Server) to connect to the SAS Metadata Server without explicitly providing credentials.

For more information about security in the SAS Intelligence Platform, see Chapter 7, "Security Overview," on page 39.

How Metadata is Created and Administered

Metadata can be loaded to the SAS Metadata Server in several different ways:

☐ The configuration process for the SAS Intelligence Platform automatically creates and stores metadata objects for the resources, such as servers, that are part of your initial installation.

☐ The SAS Metadata Server enables you to import metadata from a variety of sources (and to export it in a variety of formats). The server supports the Object Management Group's Common Warehouse Metamodel/XML Metadata Interchange (CWM/XMI) format, the industry standard for data warehouse metadata integration. In addition, by installing the Meta Integration Model Bridge (MIMB) third-party software, you can import metadata from market-leading design tool and repository vendors.

☐ When users create products such as reports, information maps, and data warehouses with the SAS Intelligence Platform applications, these applications create and store metadata objects describing the products.

☐ System administrators use the SAS Management Console client application to directly create metadata for system resources such as servers, users, and user groups.

System administrators also use SAS Management Console for general metadata administration tasks, including creating new repositories, promoting and replicating existing repositories, and maintaining authorization information and access rules for all types of resources.

For detailed information about the SAS Metadata Server, see the *SAS Intelligence Platform: System Administration Guide*.

Server Objects, Application Servers, and Logical Servers

In the SAS Metadata Repository, each server process that executes SAS code is represented by a server object. In the metadata, the attributes for each server object contain information such as the following:

☐ the name of the machine that is hosting the server

☐ the TCP/IP port or ports on which the server listens for requests

☐ the SAS command that is used to start the server

The intermediate level of organization is called a *logical server* object. SAS servers of a particular type can be grouped into a logical server of the corresponding type, as in the following examples:

☐ A *logical workspace server* is a group of one or more workspace servers.

☐ A *logical stored process server* is a group of one or more stored process servers.

☐ A *logical OLAP server* can contain *only one* OLAP server.

The logical servers are then grouped into a SAS Application Server. The following figure shows a sample configuration:

Figure 4.1 SAS Application Server Components: Sample Configuration
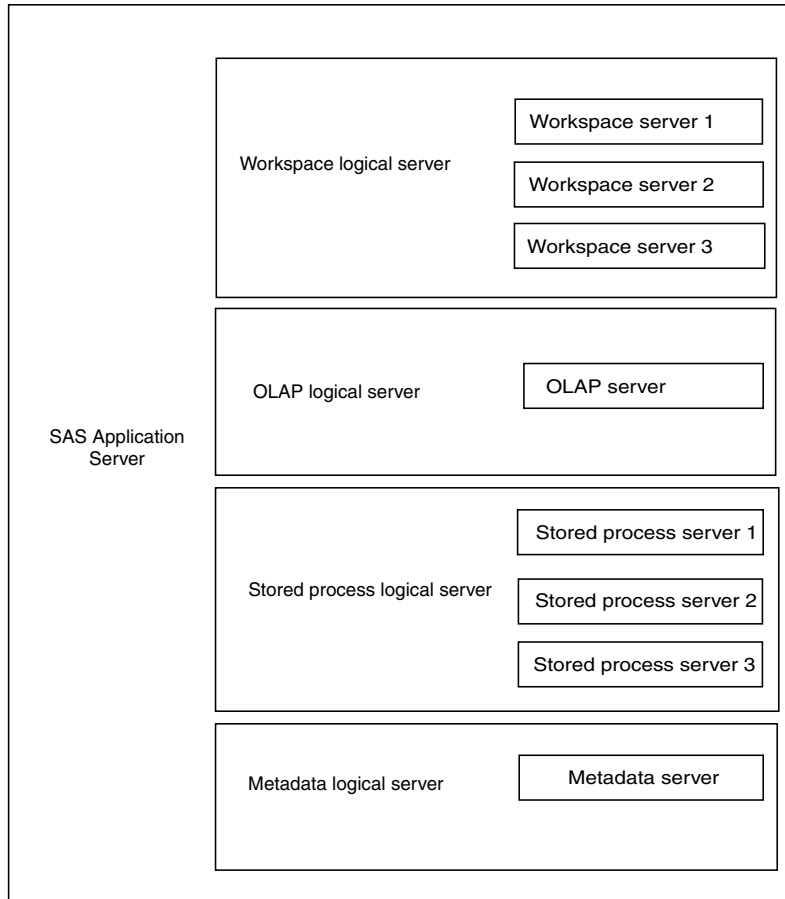
Application servers and logical servers are logical constructs that exist only in metadata. In contrast, the server objects within a logical server correspond to actual server processes that execute SAS code.

Purpose of the Application Server Grouping

Application servers, which are groupings of logical servers, provide the following functionality in the SAS Intelligence Platform:

☐ SAS applications can request the services of a group of logical servers by specifying a single application server name.

For example, a user can set up SAS Data Integration Studio to send all jobs to a specific SAS Application Server. Each time the user runs a data integration job, the SAS code that is generated is executed by a workspace server that is contained in the specified application server.

☐ System resources, such as SAS libraries or data schemas, can be assigned to a specific application server. This assignment ensures that all of the logical servers contained in the application server will be able to access these resources as needed, subject to security restrictions.

For example, if a SAS library is assigned to a specific application server, then any application that runs jobs on that server will automatically have access to the library, subject to security restrictions.

Purpose of the Logical Server Grouping

Logical servers, which are groupings of individual servers of a specific type, provide the following functionality in the SAS Intelligence Platform:

- □ Users or groups of users can be granted rights to a logical server, thereby providing access to each physical server that the logical server contains.
- □ *Load balancing* can be implemented among the servers in a logical workspace server or a logical stored process server. (Load balancing is not available for OLAP servers).
- □ *Workspace pooling* can be implemented among the servers in a logical workspace server.

Load Balancing for SAS Workspace Servers and SAS Stored Process Servers

Load balancing is a feature that distributes work equally among the server processes in a logical workspace server or a logical stored process server. Load balancing is well-suited to applications that connect to servers for long periods of time and that submit long jobs. Desktop applications often fall into this category.

The load balancer runs in the object spawner, which is a program that runs on server machines, listens for incoming client requests, and launches server instances as needed. When a logical server is set up for load balancing, and the object spawner receives a client request for a server in the logical server group, the spawner directs the request to the server in the group that has the least load.

Workspace Pooling for SAS Workspace Servers

Workspace pooling creates a set of server connections that are reused. This reuse avoids the wait times that an application incurs when it creates a new server connection for each user. You can also use pooling to distribute server connections across machines.

Pooling is most useful for applications that require frequent, but brief, connections to a SAS Workspace Server. If more than a few users will be working with Web applications such as SAS Information Delivery Portal or SAS Web Report Studio, it is essential for performance reasons that you set up workspace pooling for these applications.

For detailed information about configuring your workspace servers and stored process servers, see the *SAS Intelligence Platform: Application Server Administration Guide*.

CHAPTER

5
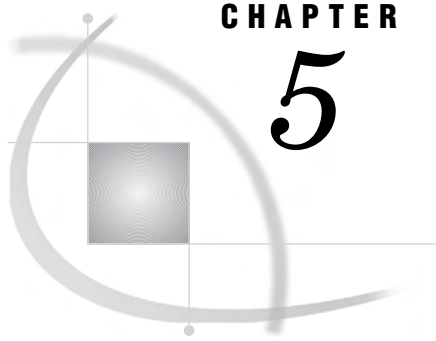
Middle-Tier Components of the SAS Intelligence Platform

## Overview of Middle-Tier Components

The middle tier of the SAS Intelligence Platform provides an environment where the business intelligence Web applications, such as SAS Web Report Studio and SAS Information Delivery Portal, can execute. These products run in a servlet container or Java 2 Enterprise Edition (J2EE) application server on the middle tier. They communicate with the user by sending data to and receiving data from the user's Web browser.

The middle-tier environment includes the following third-party software and SAS software elements:

□ a servlet container or J2EE application server

□ the Java 2 Software Development Kit, Standard Edition (J2SE SDK)

□ a WebDAV (Web-Based Distributed Authoring and Versioning) server

□ SAS Application Services

□ SAS Foundation Services

□ the SAS Web Infrastructure Kit

The following figure illustrates the components of the middle tier of the SAS Intelligence Platform:

Figure 5.1   Middle-Tier Components
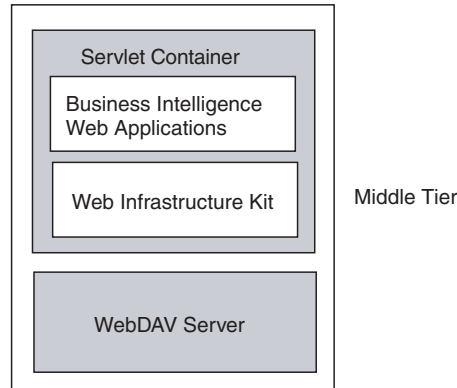

Third-Party Software Components

Servlet Container or J2EE Application Server

Servlet Container

The *servlet container* executes the Java code in SAS applications that are implemented as J2EE Web applications. A J2EE Web application is built by using Java servlet technology and is delivered as a WAR (Web archive) file (an aggregate file that contains all of the files that make up the application). For example, SAS Web Report Studio and SAS Information Delivery Portal are J2EE Web applications.

A servlet container provides a subset of the capabilities of a J2EE application server. A servlet container provides an execution environment for servlets and for JavaServer Pages (JSPs), which are translated to servlets. The following figure shows how the servlet container interacts with the client application:

Figure 5.2   Servlet Container
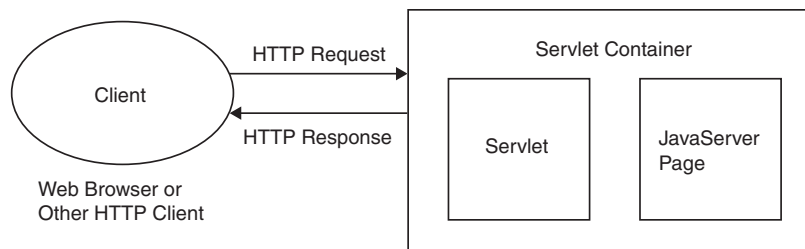
A Java Virtual Machine in the servlet container executes the Web application's Java code. It also provides additional services, such as the following:

- □ When a client sends an HTTP request to the application, the servlet container packages the contents of that request as a Java object and passes the object to the method of the servlet that will process the request.
- □ The container provides for session management. That is, the container recognizes that a sequence of HTTP requests are coming from a single client and provides for the storage of data between requests.

One servlet container that can be used in the SAS Intelligence Platform is Apache Tomcat, which is available free of charge. However, it does have a f ew limitations. You cannot run enterprise applications in this container. Also, the product might not meet the scalability and security requirements of a large enterprise. As a result, we recommend that you use Apache Tomcat only in small systems and in system prototypes. Otherwise, you should use a J2EE application server.

J2EE Application Server

The *J2EE application server* executes the Java code in SAS applications that are implemented as J2EE Web applications or J2EE enterprise applications. A J2EE enterprise application uses other technologies—in particular, Enterprise JavaBeans—in addition to Java servlet technology. The application is delivered as an EAR (enterprise archive) file. The SAS solutions are examples of J2EE enterprise applications.

A J2EE application server includes a servlet container, but also contains an Enterprise JavaBean container and a message server. It also supports a host of other technologies. The following figure shows how the J2EE application server interacts with the client application:

Figure 5.3 J2EE application server
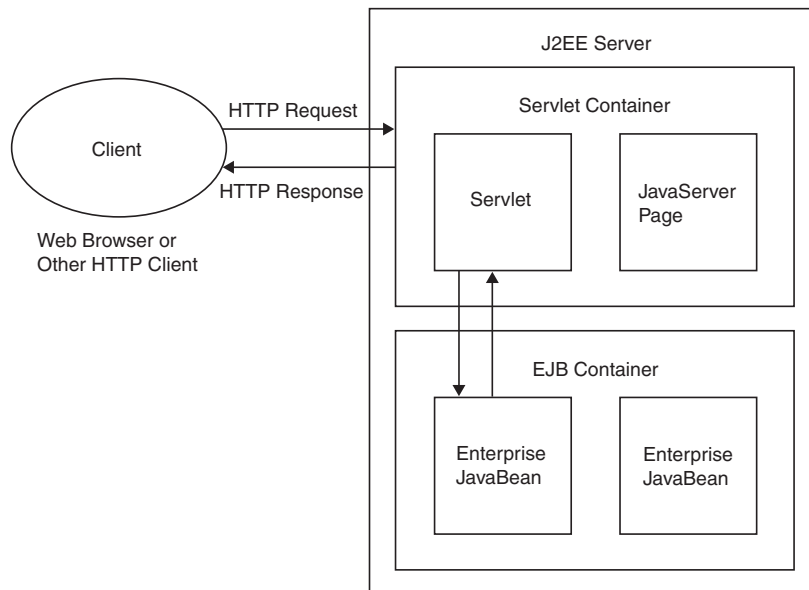
Just as the servlet container provides an execution environment for servlets, the Enterprise JavaBean container provides an execution environment for Enterprise JavaBeans (which often contain the business logic of an application). The Enterprise

JavaBean container also provides a set of services for the enterprise beans, including security and transaction management.

Most SAS business intelligence systems include such a J2EE application server. This type of server is required if you will be running any SAS solutions, and the enterprise features of the J2EE application server make it a good investment. Currently supported J2EE application servers include the BEA WebLogic Server and the IBM WebSphere Application Server.

For information about the currently supported versions of these products, see the SAS Third-Party Software Downloads page at **support.sas.com/thirdpartysupport**.

Java 2 Software Development Kit, Standard Edition (J2SE SDK)

The servlet container and the SDK are required because products such as SAS Web Report Studio, SAS Information Delivery Portal, and SAS solutions are written using JavaServer Pages and servlets. At runtime, the JavaServer Pages are translated to servlets, which must be compiled before execution. You need the SDK, which includes a Java compiler, to compile the servlets. The servlet container provides the execution environment for the compiled servlets.

WebDAV Server

The WebDAV server is an HTTP server that supports the collaborative authoring of documents that are located on the server. WebDAV enables you to manage files located on the Web server from the client desktop. The WebDAV server supports the locking of documents, so that multiple authors cannot make changes to a document at the same time. It also associates metadata with documents in order to facilitate searching. The WebDAV server acts like a network-accessible file system and stores content that users might want to access through SAS Web Report Studio or SAS Information Delivery Portal, such as documents, report definitions, and images. The SAS business intelligence applications use this type of server primarily as a report repository. WebDAV servers commonly used in the SAS Intelligence Platform include the Apache HTTP Server (with its WebDAV modules enabled) and Xythos Software's WebFile Server.

SAS Software Components

SAS Application Services

SAS Application Services, a set of services that includes SAS Java Components, SAS COM components, and SAS Query and Reporting Services, provides business-oriented query and reporting services to clients. By using a business metadata layer and a universal report definition, SAS Query and Reporting Services provides a foundation for enterprise reporting and application development. Java and COM-based interfaces to SAS Application Services make the functionality provided by SAS Query and Reporting Services available to clients. SAS Application Services can also be used by application developers to provide custom business intelligence capabilities within their solutions.

SAS Foundation Services

SAS Foundation Services is a set of core infrastructure services that enables Java programmers to write distributed applications that are integrated with the SAS platform. This suite of Java-based APIs provides core middleware infrastructure services that include the following:

- □ client connections to SAS Application Servers
- □ dynamic service discovery
- □ user authentication
- □ profile management
- □ session management
- □ activity logging
- □ metadata and content repository access
- □ connection management

Extension services for information publishing, event management, and SAS Stored Process execution are also provided.

For more information about SAS Foundation Services, see the *SAS Integration Technologies: Administrator's Guide* and the *SAS Integration Technologies: Developer's Guide*, which are available at **support.sas.com/rnd/itech/library/library9.html**.

SAS Web Infrastructure Kit

The SAS Web Infrastructure Kit, which is provided with SAS Integration Technologies, serves as the infrastructure for the SAS Information Delivery Portal Web application. The SAS Web Infrastructure Kit provides certain services to the SAS Information Delivery Portal. These services handle such tasks as the following:

- user logon and logoff
- page navigation
- searching
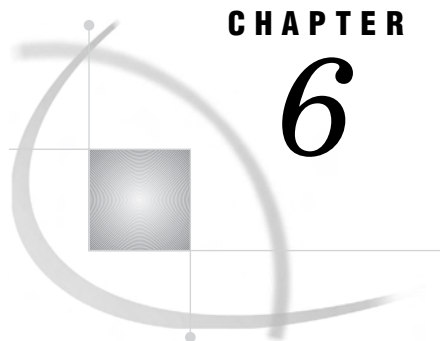- integration with SAS via stored processes
- interacting with basic content types

The SAS Web Infrastructure Kit includes the following components:

- the SAS Portal Web Application Shell, which is used by other SAS Web applications. This Web application displays content in portlets and pages. It also provides logon and logoff capabilities, metadata searching, bookmarking, and content administration features.

  Note: For full portal capabilities, you must install the SAS Information Delivery Portal. △

- the SAS Stored Process Web Application, which is a Web application that enables stored processes to be run from the Web.
- the SAS Documentation Application, which is a Web application that manages SAS documentation.
- the SAS Services Application (including Remote Foundation Services), which is a Java application that manages services that are shared by SAS applications.
- predefined portlets for content viewing and navigation.
- a portlet development kit, which includes an API and a set of best practices for developing custom portlets.
- administrative tools for deploying services, portlets, and additional Web applications.
- SAS Java components and Web infrastructure components.

For information about the SAS Web Infrastructure Kit, see the *SAS Intelligence Platform: Web Application Administration Guide*, and also the documentation that is available at **support.sas.com/rnd/itech/library/library9.html**.

Clients in the SAS Intelligence Platform

Overview of Clients

SAS Intelligence Platform clients can be Windows applications, Java applications, or Web-based applications. The following table lists the clients by type:

Table 6.1   SAS Intelligence Clients

Java Applications[1]	Web Applications[2]	Windows Applications
SAS Enterprise Miner	SAS Information Delivery Portal	SAS Add-In for Microsoft Office
SAS Data Integration Studio	SAS Web OLAP Viewer	SAS Enterprise Guide
SAS Information Map Studio	SAS Web Report Studio	
SAS Management Console		
SAS OLAP Cube Studio		

1   Require a JRE on each client machine.
2   Require a Web browser on each client machine and a servlet container or Java 2 Enterprise Edition (J2EE) application server on the middle-tier machine where the application will run.

The Windows applications and Java applications are supported only on Microsoft Windows systems. The exception is SAS Management Console, which also runs on several UNIX platforms. All of the Java applications require the Java Runtime Environment (JRE), which includes a Java Virtual Machine (JVM) that executes the application and a set of standard Java class libraries. If you have installed the SAS Foundation on a host, the JRE will already be present on that machine. Otherwise, you can install the JRE from a CD that is supplied by SAS before you install the first Java client.

The Web-based applications reside and execute on the middle tier (see Chapter 5, "Middle-Tier Components of the SAS Intelligence Platform," on page 27) , and require only a Web browser to be installed on each client machine. These products run in a servlet container or J2EE application server on the middle tier. They communicate with the user by sending data to and receiving data from the user's Web browser. For example, an application of this type displays its user interface by sending an HTML document to the user's browser. The user can submit input to the application by sending it an HTTP response—usually by clicking a link or submitting an HTML form.

SAS Add-In for Microsoft Office

The SAS Add-In for Microsoft Office is a Component Object Model (COM) add-in that extends Microsoft Office. It enables you to harness the power of SAS analytics, to access relational data sources directly from Microsoft Word and Microsoft Excel, and to create reports within Microsoft Office. You can use the SAS Add-In for Microsoft Office for the following tasks:

☐ dynamically execute stored processes and SAS tasks.

☐ embed the results of your stored processes and SAS tasks in your documents and spreadsheets.

☐ access and view SAS data sources or any data source that is available from your SAS server. There is no size limit on the SAS data sources that you can open. A SAS navigation toolbar within Microsoft Office enables you to navigate easily through data sources too large for the row application limit of Excel.

For more information about the SAS Add-In for Microsoft Office, see the SAS Add-In for Microsoft Office Online Help, which is available from within the product. For information about administrative tasks associated with the SAS Add-In for Microsoft Office, see the *SAS Intelligence Platform: Desktop Application Administration Guide*.

SAS Data Integration Studio

SAS Data Integration Studio is a Java application that enables you to consolidate and manage enterprise data from a variety of source systems, applications, and technologies. SAS Data Integration Studio is a visual design tool that provides a drag-and-drop interface, along with interactive wizards that lead you through tasks such as creating jobs and process flow diagrams. SAS Data Integration Studio enables you to create process flows that accomplish the following tasks:

☐ extract, transform, and load data for use in data warehouses and data marts

☐ cleanse, migrate, synchronize, replicate, and promote data for applications and business services

SAS Data Integration Studio enables you to integrate information from any platform that is accessible to SAS and from any format that is accessible to SAS.

For more information about SAS Data Integration Studio, see the SAS Data Integration Studio Help, which is available from within the product, and the *SAS Data Integration Studio: User's Guide*. For information about administrative tasks associated with SAS Data Integration Studio, see the *SAS Intelligence Platform: Desktop Application Administration Guide*.

SAS Enterprise Guide

SAS Enterprise Guide is a project-oriented Windows application that is designed to enable quick access to much of the analytic power of SAS software for statisticians, business analysts, and SAS programmers. SAS Enterprise Guide provides the following functionality:

- □ a point-and-click user interface to all SAS servers (V8 and later)
- □ transparent data access to both SAS and other types of data
- □ interactive task windows that lead you through dozens of analytical and reporting tasks
- □ the ability to utilize the highest quality SAS graphics
- □ the ability to export results to other Windows applications and the Web
- □ the ability to schedule your project to run at a later time

SAS Enterprise Guide also enables you to create SAS Stored Processes and to store that code in a repository that is available to a SAS Stored Process Server. (Stored processes are SAS programs that are stored on a server and are executed by client applications.) Stored processes are used for Web reporting and analytics, among other things.

For more information about SAS Enterprise Guide, see the SAS Enterprise Guide Help, which is available from within the product. For information about administrative tasks associated with SAS Enterprise Guide, see the *SAS Intelligence Platform: Desktop Application Administration Guide*.

SAS Enterprise Miner

SAS Enterprise Miner is a Java application that enables you to create and manage data mining process flows, which are sequences of steps for the examination, transformation, and processing of data to create models to predict complex behaviors of economic interest.

For more information about SAS Enterprise Miner, see the SAS Enterprise Miner Help, which is available from within the product. For information about administrative tasks associated with SAS Enterprise Miner, see the *SAS Intelligence Platform: Desktop Application Administration Guide*.

SAS Information Delivery Portal

SAS Information Delivery Portal is a J2EE Web application that enables you to aggregate data from a variety of sources and present the data in a Web browser. The Web browser content might include the output of SAS Stored Processes, links to Web addresses, documents, syndicated content from information providers, SAS information maps, SAS reports, and Web applications.

Using the portal, you can distribute different types of content and applications as appropriate to internal users, external customers, vendors, and partners. You can use the portal along with the Publishing Framework to publish content to SAS publication channels or WebDAV repositories, to subscribe to publication channels, and to view packages published to channels. The portal's personalization features enable users to organize information on their desktops in a way that makes sense to them.

For more information about SAS Information Delivery Portal, see the SAS Information Delivery Portal Help, which is available from within the product. Additional information is in the *SAS Intelligence Platform: Web Application Administration Guide* and the *SAS Web Infrastructure Kit: Developer's Guide*, which is part of the SAS Integration Technologies documentation available at **support.sas.com/rnd/itech/library/library9.html**.

SAS Information Map Studio

SAS Information Map Studio is a Java application that enables data modelers and data architects to create and manage SAS Information Maps, which are business metadata about your physical data. Information maps enable you to surface your data warehouse data in business terms that typical business users understand, while storing key information that is needed to build appropriate queries.

Information maps provide the following benefits:

- □ Users are shielded from the complexities of the physical data.
- □ Data storage is transparent to the consumers of information maps, regardless of whether the underlying data is relational or multidimensional, or whether the data is in a SAS data set or in a third-party database system.
- □ Business formulas and calculations are predefined, which makes them usable on a consistent basis.
- □ Users can easily query data for answers to business questions without having to know query languages.

For more information about SAS Information Map Studio, see the SAS Information Map Studio Help, which is available from within the SAS Information Map Studio product, and *SAS Information Map Studio: Creating Your First Information Map*. For information about administrative tasks associated with SAS Information Map Studio, see the *SAS Intelligence Platform: Desktop Application Administration Guide*.

SAS Management Console

SAS Management Console is a Java application that provides a single point of control for managing resources that are used throughout the SAS Intelligence Platform. Rather than using a separate administrative interface for each application in your computing environment, you can use the single interface in SAS Management Console to perform the administrative tasks that are required in order to create and maintain an integrated environment across multiple platforms. This client application utilizes an extensible plug-in architecture, enabling you to customize the console to support a wide range of administrative capabilities.

SAS Management Console enables you to manage the following resources:

- □ server definitions
- □ library definitions
- □ user definitions
- □ resource access controls
- □ metadata repositories
- □ SAS licenses
- □ job schedules

□ XMLMaps

SAS Management Console works by creating and maintaining metadata definitions for each computing resource or control. These metadata definitions are stored in a repository on a SAS Metadata Server, which makes them available for use by other applications.

The metadata definitions are created using plug-ins, which are application modules designed to create metadata for a specific type of resource. For example, the Server Manager plug-in creates metadata to define SAS servers and application servers. Although a set of basic plug-ins is provided with SAS Management Console, you can install other plug-ins to meet specific needs, or you can develop your own plug-in.

For more information about SAS Management Console, see the SAS Management Console Help, which is available from within the product, and the *SAS Management Console: User's Guide*. For information about administrative tasks associated with SAS Management Console, see the *SAS Intelligence Platform: Desktop Application Administration Guide*.

SAS OLAP Cube Studio

SAS OLAP Cube Studio is a Java application that enables you to register cube metadata in a SAS Metadata Repository and save physical cube data in a specified location.

SAS OLAP Cube Studio enables you to design and create online analytical processing (OLAP) cubes. Using its Cube Designer Wizard, you can specify the following:

□ the data source for a cube

□ the cube design and architecture

□ measures of the cube for future queries

□ initial aggregations for the cube

SAS OLAP Cube Studio includes plug-ins that enable you to create additional cube aggregations, to add calculated members to a cube, and to modify existing calculated members.

For more information about SAS OLAP Cube Studio, see the SAS OLAP Cube Studio Help, which is available from within the product. For information about administrative tasks associated with SAS OLAP Cube Studio, see the *SAS Intelligence Platform: Desktop Application Administration Guide*.

SAS Web OLAP Viewer

SAS Web OLAP Viewer is a Web-based application for viewing SAS OLAP data. It provides an easy-to-use interface from which you can select a data source, view the data, and customize your view with features such as sorting and filtering.

SAS Web OLAP Viewer enables you to perform the following tasks:

□ subset your data by drilling down and filtering

□ save and restore bookmarked views of your data

□ search for data values

□ calculate new data items

□ view ESRI maps

□ export your data view as a SAS report, Excel file, or PDF document

Note: You cannot use SAS Web OLAP Viewer to make changes to information maps or to physical data. △
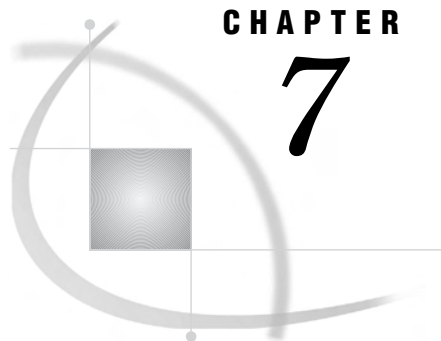
For more information about SAS Web OLAP Viewer, see the SAS Web OLAP Viewer Help, which is available from within the product. For information about administrative tasks associated with SAS Web OLAP Viewer, see the *SAS Intelligence Platform: Web Application Administration Guide*.

SAS Web Report Studio

SAS Web Report Studio is a Web-based application that enables you to create, view, and organize reports. You can use SAS Web Report Studio to accomplish the following tasks:

- □ *Creating reports.* Beginning with a simple and intuitive view of your data provided by SAS Information Maps (created in SAS Information Map Studio), you can create reports based on either relational or multidimensional data sources. You can use the Report Wizard to quickly create simple reports or the Edit Report view to create sophisticated reports that have multiple data sources, each of which can be filtered. These reports can include various combinations of list tables, crosstabulation tables, and graphs. Using the Edit Report view, you can adjust the style to globally change colors and fonts. You can also insert stored processes that take the results from a block of SAS code and embed those results directly into a report.

- □ *Viewing and working with reports.* While viewing reports using a thin client (a Web browser), you can filter, sort, and rank the data that is shown in tables, crosstabs, and graphs. With multidimensional data, you can drill down on data in crosstabs and graphs and drill through to the underlying data.

- □ *Organizing reports.* You can create folders and subfolders for organizing your reports. Information consumers can use keywords to find the reports that they need. Reports can be shared with others or kept private.

- □ *Printing and exporting reports.* You can preview a report in PDF and print the report, or save and e-mail it later. You have control over many printing options, including page orientation, page range, and size of the tables and graphs. You can also export data as a spreadsheet.

For more information about using SAS Web Report Studio, see the SAS Web Report Studio Help, which is available from within the product, and the *SAS Web Report Studio: User's Guide*. For information about administrative tasks associated with SAS Web Report Studio, see the *SAS Intelligence Platform: Web Application Administration Guide*.

CHAPTER

7

Security Overview

Introduction to the Security Overview

This chapter introduces the SAS®9 security model by describing how authentication and authorization work in the SAS Intelligence Platform. Authentication is the verification of a user's identity, and authorization is determining which permissions that user has for which resources. For a comprehensive discussion of security and detailed information about security administration activities, see the *SAS Intelligence Platform: Security Administration Guide*.

Authentication in the SAS Intelligence Platform

Introduction to Authentication

Authentication is an identity verification process that attempts to determine whether users (or other entities) are who they say they are. Authentication is a prerequisite for authorization, because a user's identity is the basis for authorization decisions about which actions the user is permitted to perform with which resources.

In the SAS Intelligence Platform, a user's identity is verified first when the user logs on to an application and again as the user requests access to other systems. For example, when a user logs on to SAS Data Integration Studio, the user authenticates to the SAS Metadata Server. When the user makes a request from SAS Data Integration Studio to run a job against an Oracle table, the user must authenticate to the SAS Workspace Server that processes the request and to the Oracle server that manages the table.

For a comprehensive discussion of this subject, see the *SAS Intelligence Platform: Security Administration Guide*.

How Identities Are Verified

In most cases, SAS servers rely on their host operating systems to verify identities. This process is called host authentication. For example, before allowing a user to run a stored process, a stored process server asks its host computer to authenticate the user. The host computer compares a provided user ID and password to a list of valid accounts in the operating system (or in a back-end authentication database that the operating system is using). If the provided user ID and password correspond to a valid account, the authentication is successful.

Note: As an alternative to relying on the host operating system, the metadata server and the OLAP server can make direct use of Lightweight Directory Access Protocol (LDAP) or Microsoft Active Directory to verify identities. However, the preferred way to use an alternative authentication provider is as a back-end user store behind host authentication, because direct use of LDAP and Active Directory Direct can significantly increase the need to store user IDs and passwords in the metadata repository. △

In some cases, SAS servers trust verification that has been performed by other components. The SAS Intelligence Platform supports the following trust relationships:

- □ The metadata server trusts the identity verification that the SAS OLAP Server performs.
- □ By default, the metadata server trusts the identity verification that a connecting SAS process performs.
- □ If SAS Web applications are configured to use Web server authentication, the metadata server trusts the identity verification that a Web server performs.

Single Sign-On

Single sign-on enables users to access a variety of computing resources without being repeatedly prompted for their user IDs and passwords. The SAS Intelligence Platform provides these single sign-on features:

- □ Most applications can cache the credentials that a user submits to log on.
- □ All applications can retrieve credentials that have been stored in the metadata repository.
- □ Web applications can share user and session contexts.

Identity Management

In addition to managing user accounts in external systems, administrators must create and maintain some user information in the metadata repository. You can minimize the amount of identity information that you need to replicate in the metadata by choosing your authentication providers carefully and making appropriate use of shared accounts.

The SAS Intelligence Platform provides the following tools for management of identity information in the metadata:

- □ Administrators can use SAS Management Console to define and manage metadata identity information.
- □ Administrators can use batch processes to extract identity information from sources such as LDAP or UNIX /etc/passwd files and create corresponding identity

information in the metadata repository. Batch processes can also be used to periodically update the identity information. Batch processes cannot be used to manage passwords.

☐ Users can use the SAS Personal Login Manager desktop application to manage their own account information.

Authorization in the SAS Intelligence Platform

Introduction to Authorization

Authorization is the process of determining which users have which permissions for which resources. For a comprehensive discussion of this subject, see the *SAS Intelligence Platform: Security Administration Guide*.

Metadata-Based Authorization

The SAS Intelligence Platform includes an authorization mechanism that consists of access controls that you define and store in a metadata repository. These metadata-based controls enable you to manage access to metadata and, in some cases, to the computing resources that the metadata represents.

The available metadata-based permissions are summarized in the following table.

Table 7.1   Metadata-Based Permissions

Permissions	Use
ReadMetadata, WriteMetadata, CheckInMetadata	Use to control user interactions with a metadata object.
Read, Write, Create, or Delete	Use to control user interactions with the underlying computing resource that is represented by a metadata object.
Administer	Use to control administrative interactions (such as starting or stopping) with the SAS server that is represented by a metadata object.

The consequences of granting or denying a metadata-based permission vary depending on factors such as these:

☐ whether you assign the permission to an individual user or to a user group. The metadata-based authorization mechanism includes an identity precedence ranking that is based on your group membership hierarchy.

☐ whether you assign the permission by applying a pre-existing permission pattern. The metadata-based authorization mechanism enables you to set controls by using either ad-hoc assignments or reusable patterns.

☐ whether you set the permission on an object from which other objects can inherit access controls. The metadata-based authorization mechanism includes a set of rules that determine which objects can be parents to which other objects.

☐ whether the application that the requesting user is using enforces the permission. In the current release, not all applications enforce the metadata-based Read, Write, Create, Delete, and Administer permissions.

Multiple Authorization Layers

Your ability to perform a particular action is determined not only by these metadata-based access controls but also by external authorization mechanisms such as operating system permissions and database controls. In order to perform a particular action, you must have the necessary permissions in all of the applicable authorization layers.

For example, regardless of the access controls that have been defined for you in the metadata repository, you cannot access a particular file if the operating system permissions do not permit the action.

A P P E N D I X

1
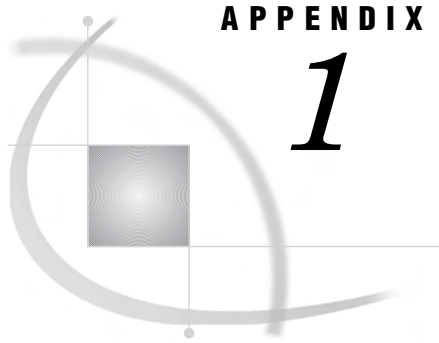
Recommended Reading

Recommended Reading

Here is the recommended reading list for this title:

SAS Intelligence Platform Administration Documentation
- □ *SAS Intelligence Platform: Application Server Administration Guide*
- □ *SAS Intelligence Platform: Data Administration Guide*
- □ *SAS Intelligence Platform: Desktop Application Administration Guide*
- □ *SAS Intelligence Platform: Installation Guide*
- □ *SAS Intelligence Platform: Security Administration Guide*
- □ *SAS Intelligence Platform: Single-User Installation Guide*
- □ *SAS Intelligence Platform: System Administration Guide*
- □ *SAS Intelligence Platform: Web Application Administration Guide*

For a complete list of administration documentation for the SAS Intelligence Platform, see **http://support.sas.com/913administration**.

Other Titles of Interest
- □ *SAS Data Integration Studio: User's Guide*
- □ *SAS Information Map Studio: Creating Your First Information Map*
- □ *SAS Management Console: User's Guide*
- □ *SAS Web Infrastructure Kit: Developer's Guide*
- □ *SAS Web Report Studio: User's Guide*

For a complete list of SAS publications, see the current *SAS Publishing Catalog*. To order the most current publications or to receive a free copy of the catalog, contact a SAS representative at

SAS Publishing Sales
SAS Campus Drive
Cary, NC 27513
Telephone: (800) 727-3228*
Fax: (919) 677-8166
E-mail: `sasbook@sas.com`
Web address: `support.sas.com/pubs`
* For other SAS Institute business, call (919) 677-8000.

Customers outside the United States should contact their local SAS office.

Index

Your Turn

If you have comments or suggestions about *SAS® Intelligence Platform: Overview, Second Edition*, please send them to us on a photocopy of this page, or send us electronic mail.

For comments about this book, please return the photocopy to

SAS Publishing
SAS Campus Drive
Cary, NC 27513
E-mail: **yourturn@sas.com**

For suggestions about the software, please return the photocopy to

SAS Institute Inc.
Technical Support Division
SAS Campus Drive
Cary, NC 27513
E-mail: **suggest@sas.com**

SAS® Publishing gives you the tools to flourish in any environment with SAS®!

Whether you are new to the workforce or an experienced professional, you need to distinguish yourself in this rapidly changing and competitive job market. SAS® Publishing provides you with a wide range of resources — including publications, online training, and software — to help you set yourself apart.

Expand Your Knowledge with Books from SAS® Publishing

SAS® Press offers user-friendly books for all skill levels, covering such topics as univariate and multivariate statistics, linear models, mixed models, fixed effects regression, and more. View our complete catalog and get free access to the latest reference documentation by visiting us online.

support.sas.com/pubs

SAS® Self-Paced e-Learning Puts Training at Your Fingertips

You are in complete control of your learning environment with SAS Self-Paced e-Learning! Gain immediate 24/7 access to SAS training directly from your desktop, using only a standard Web browser. If you do not have SAS installed, you can use SAS® Learning Edition for all Base SAS e-learning.

support.sas.com/selfpaced

Build Your SAS Skills with SAS® Learning Edition

SAS skills are in demand, and hands-on knowledge is vital. SAS users at all levels, from novice to advanced, will appreciate this inexpensive, intuitive, and easy-to-use personal learning version of SAS. With SAS Learning Edition, you have a unique opportunity to gain SAS software experience and propel your career in new and exciting directions.

support.sas.com/LE

§sas | SAS Publishing

THE POWER TO KNOW®