SYBASE®

Sybase Business Intelligence and Data Warehousing Solutions for Sybase IQ™

Changing the Rules of the Game: Sybase IQ

By: Frank Teklitz
Director Strategic Programs

analytics

WHITE PAPER

Table of Contents

Traditional Relational Databases

Traditional relational databases are made up of a data space and index space. This architecture works very well for online transaction processing (OLTP) - Airline Reservation systems, Claims Processing systems, etc. where one record or row is processed at a time. This horizontal approach breaks down for ad hoc queries. Traditional relational databases are not designed to retrieve multiple rows of data or to do aggregation on the fly. This forces DBAs to pre-aggregate data, create materialized views and a lot of indexes (b-tree or bit-map).

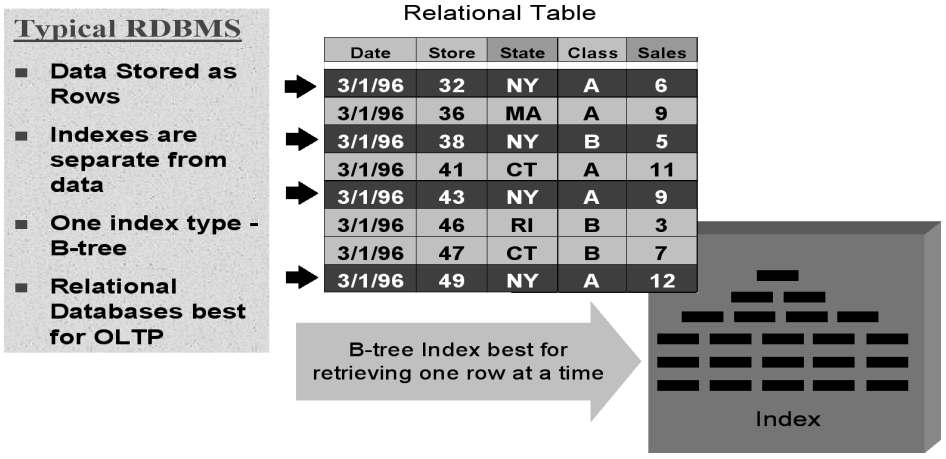

Diagram 1: Traditional Relational Databases

With conventional relational databases you must tune for every query. This costs time, money and resources. It also causes huge data and index explosion, which further increases maintenance overhead.

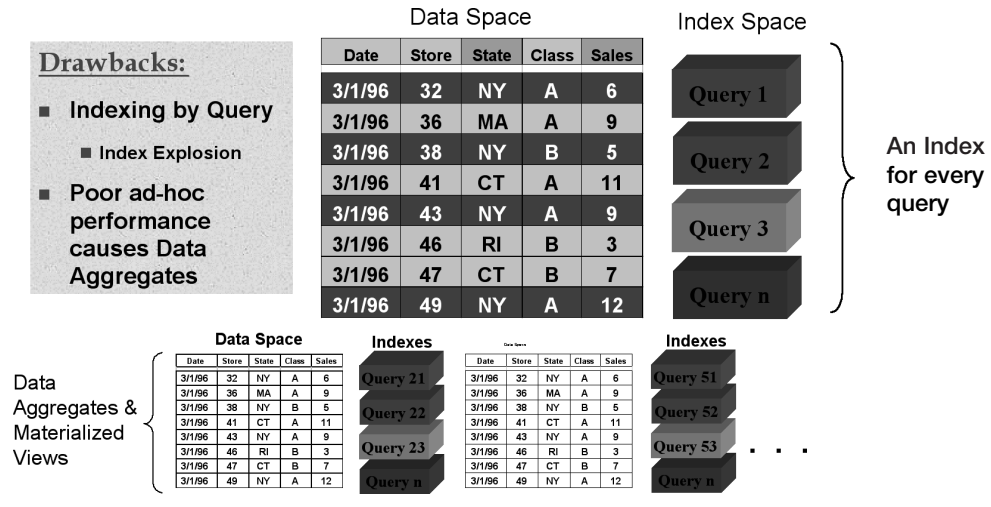

Diagram 2: Traditional Relational Databases are subject to index and data explosion to improve performance

Sybase IQ: Designed From The Ground Up For Enterprise Analytics

Sybase IQ is designed for the kinds of dynamic business analysis needed to support large numbers of users and large amounts of data. Sybase IQ delivers high-speed access to business information 10 to 100, even 1,000 times faster than conventional relational databases. At the core of Sybase IQ are our exclusive and patented Bit-Wise™ indexes. Sybase IQ stores the data by column. This is also called Vertical Partitioning. Traditional relational databases store data by row. With Sybase IQ, your analysis query simply and quickly zips through only the bits in the columns that you need, while conventional relational databases slowly chug through all the bytes in a row, clogging I/O channels, memory and disk.

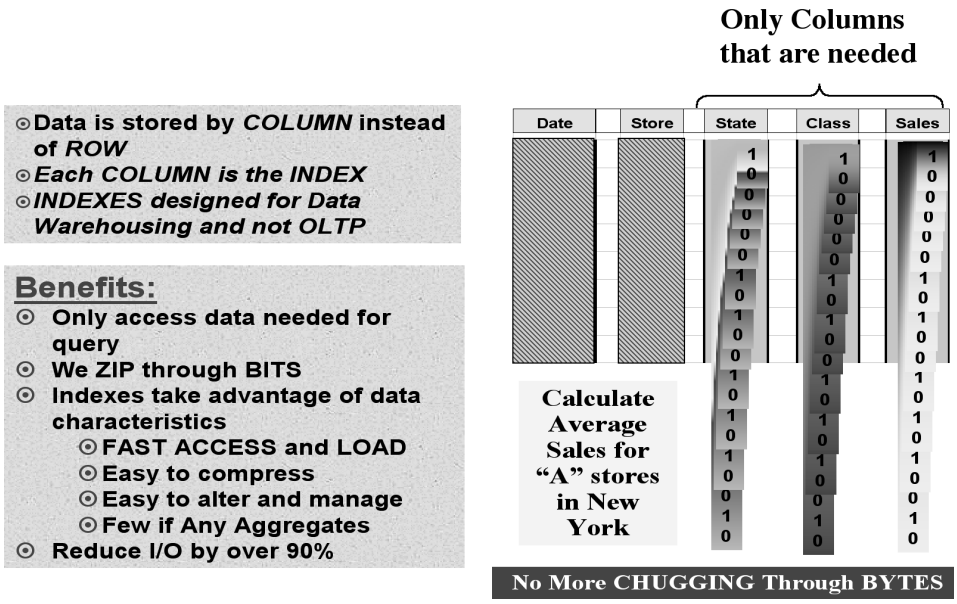

Diagram 3: Sybase IQ uses Bit-Wise indexes to zip through queries

For example, let's say you want to calculate "Average Sales for 'A' stores in New York."

With Sybase IQ, you only access data columns needed for the query. From there Sybase IQ uses its specialized Bit-Wise indexes to zip through bits for fast access and load. The Bit-Wise indexes take advantage of data characteristics, making them very efficient. Because of the vertical partitioning, it is easy to alter and manage the database. You need few if any aggregations. Finally, because of the Bit-Wise indexes and vertical partitioning, Sybase IQ reduces I/O by over 90%. This reduction in I/O is key to delivering linear user scalability and multiplexing.

For example, if two columns out of 20 are used from a table for an ad-hoc query, with Sybase IQ only those two columns are passed through from disk to the I/O buffer and into memory. With traditional relational databases (again designed for OLTP), the entire row must be processed. Also, Sybase IQ uses a 512KB buffer designed to pass large amounts of detail data into memory for fast, on the fly access.

Traditional Relational Databases Versus Sybase IQ

To Gain Performance, Traditional Relational Databases Are Inherently Inflexible

The first limitation of using traditional relational databases for business analysis is that they tend to be inflexible. With traditional relational databases you theoretically have unlimited access to information – n-way joins, entity models, 3NF. But in reality, to improve the performance and scalability of traditional relational databases, DBAs will use techniques like pre-aggregation or materialized views to gain some level of performance. Aggregates or materialized views become a bottleneck for flexibility. If the user performs a query that is outside the available aggregates or materialized view, performance for the query can be measured in hours and days. That query also takes resources from other users and applications, slowing down every user's performance. The DBA must add new indexes, aggregates or materialized views at a significant cost in data management, maintenance overhead and disk. Because traditional relational engines are inherently slow when doing calculations on the fly, much of the numeric information must be pre-calculated, aggregated and constantly tuned. This pre-calculation and aggregation is very difficult and costly, even with the best transformation tools.

Sybase IQ Delivers Performance And Flexibility

Sybase IQ is faster for ad hoc business analysis for a number of reasons. First, Sybase IQ's patent pending index structures index and optimize each column of data based on the characteristics of the column. Today there are 5 core indexes:

- **Projection Indexes** are the data stores for Sybase IQ, displaying information in the actual report or query result set.
- **Low Cardinality Indexes** are for those fields that have less than 1,500 unique values (State Description - CA, OH, etc; Gender Code; male, female; Product or Part Descriptions). These indexes deliver fast performance for data look-ups; for example, "where" clauses.
- **Bit-Wise Indexes** for high cardinality fields are used for calculations and range searches (Sum of Dollars and Units, where price is less than $50).
- **High Group Indexes** are for key fields and grouping information – Group Dollars and Units by Product (High Cardinality Group by index) and Plant location (Low Cardinality).
- **Join Indexes** deliver improved performance for join relationships between tables.

In addition, Sybase IQ delivers two specialized column indexes designed to meet new and evolving needs in the market. The ability to build specific column indexes to meet specific business needs is a huge competitive advantage for Sybase IQ customers because Sybase has the expertise to deliver powerful indexes that are easy to apply to a customer data warehouse.

- **Text Index:** With the advent of 9/11, the ability to easily and efficiently analyze text information from around the world is imperative. The Text Index enables the analysis of textual (qualitative data) and numeric data (quantitative data) together. Load email, documents from around the world, credit card transactions, transcripts and medical records into this power index and you can quickly, effectively and efficiently understand the relevance of textual information.
- **Compare Index:** Turn the interrelationships of information into real-time business analysis. The Compare Index enables businesses to turn key performance indicators (KPIs) into events that drive the business for an active, real-time data warehouse.

All of these indexes are designed to deliver query results 10 to 100, even 1,000 times faster than traditional relational databases. This performance information is based on results from the hundreds of customers who own and use Sybase IQ.

With Sybase IQ, the DBA's job is much simpler because Sybase IQ can perform aggregation on the fly (eliminating pre-aggregated data and extra indexes) and can perform index compression automatically as part of the data load process. Sybase IQ can work on any schema (3NF, Star, R-Cube) and optimize any query. Because of Sybase IQ's Bit-Wise, column oriented indexes, it is easy to add columns to a table as they are needed. There is no need to rebuild indexes, aggregates or tables as with traditional relational databases. With Sybase IQ it is easy to add and load a column to a table because aggregation is done on the fly, eliminating the need to pre-aggregate or re-index the table.

Sybase IQ dramatically reduces the transformation design and development effort. Much of the metadata developed for data transformation is used to pre-aggregate data. Because of the reduction in pre-aggregation with Sybase IQ, the transformation logic is greatly simplified.

Sybase IQ offers great design flexibility supporting 3NF, Star schema or R-Cubes. The entire dimension and fact tables model is dramatically simplified by the reduction in aggregates and materialized views.

Traditional Relational Databases Use 4 Times As Much Storage As Sybase IQ

The second limitation of using traditional relational databases for business analysis is the data explosion that results from the added indexes, aggregations and materialized views needed to improve performance and scalability. For every terabyte of input data, a traditional relational database will require at least an additional 2 terabytes for aggregations and 2 terabytes of b-tree (or bit map) indexes. It is not unusual to have a data and index explosion in the range of 4 to 10 times the original input data size using traditional relational technology.

Below is a table developed by Gartner on the amount of disk space needed to accommodate raw input data.

	Raw Data to Disk Space*	DBAs per Terabyte
IBM DB2 UDB	1 to 3-4	3
Oracle	1 to 4-5	3-4
Teradata	1 to 2.5	1-plus
Sybase IQ	1 to 0.25 -0.9	1-2 (any size)

** Does not take into account any disk mirroring database free space.*

Chart 1: Gartner Study

Hidden Costs of Administration: "Rules of Thumb," April 2002 Symposium

According to Gartner, "not all DBMS products are equal when it comes to supporting large and complex DWs. In sizing hardware configurations (disk and processor) and estimating DBA resources, enterprises make the critical mistake of assuming DBMS products are equal. We examine these issues and provide some 'rules of thumb' (see Chart 1) to determine configurations, resources and TCO. Performing this analysis can help set realistic cost estimates that reduce the likelihood of requiring unexpected funding later."

Sybase IQ Delivers Data Compression

Not only do Sybase IQ's specialized Bit-Wise indexes dramatically improve performance, they also compress the data. For example, a terabyte of input data will be compressed in Sybase IQ to 250GB-900GB (including ALL indexes). In contrast, traditional DBMSs will "explode" a terabyte of input data to 2.5TB-10TB (depending on indexing).

In a recent certification test done by Francois Raab (Auditor for TPC), 48 terabytes of input data (a world record) were loaded into Sybase IQ and used only 22TB of storage. A traditional DBMS would require 150TB-300TB.

According to Raab, "this is the largest demonstrated and verified data warehouse to date with 48.2 terabytes of input data (179 billion rows) stored in a realistic data warehouse schema using only 22 terabytes of total storage."

Traditional Relational Databases Use 10 Times As Many Disk Spindles As Sybase IQ

Today, traditional OLTP databases use page sizes in the range of 2KB to 16KB (approximately 2,000 to 16,000 total characters) – a perfect range for OLTP, where every column is used to process business transactions such as insurance claims, airline reservations, banking/financial exchanges, etc. These transactions typically fall in the range of 2KB to 16KB total characters. Every time there is a read (or write), the arm on the disk will pick up 2KB to 16KB and pass it through the I/O channel into memory to be processed by the CPU very efficiently, even for millions of transactions.

But what is good for OLTP is disastrous for business analysis. First, typically, only a few columns in a table are used for analysis. For example, if only 4 columns of a 100 column row of data are needed for analysis, using a traditional relational database (this includes IBM DB2 / Universal Server, Oracle, Teradata and Microsoft SQL Server), the other 96 columns must be dragged out of disk through the I/O channel and into memory. From there the specific columns needed must be found in the row and then passed to the CPU for processing. This explains why vendors in competitive benchmarks often put highly used columns at the front of the row to reduce the memory seek time. It is all unnecessary overhead.

Second, the amount of data needed to perform a typical analysis is often far greater than 16,000 (16KB) characters. For example, let's say you want to sum up all of the insurance claims in Ohio in January. This may total 512,000 characters (512KB). Assuming the largest page size of 16KB, it would take 32 reads to return 1,024KB of data. During this time the CPU is waiting for the information to get into memory, dramatically slowing down business intelligence queries, data loads and ad-hoc calculations.

Can You Improve On These Numbers?

If you are using an OLTP database, the best way to improve performance is to increase the number of I/O channels by spreading the data over many spindles with small disks, which many of our competitors do. As proof of this, take a look at any TPC-h benchmark and you will see that the traditional VLDB players use 8 and 16GB disks – i.e. many spindles. Given that 73GB disks are the norm today for disk size, it seems puzzling to use small (and expensive) disks. But if you are using an OLTP database for business intelligence, you have no choice. A general heuristic for spindle management with traditional databases is 10 spindles for every 1 CPU needed. Using this 10:1 ratio of spindles to CPUs, if you have 32 CPUs, you will need 320 spindles to keep your CPU, I/O and memory fully utilized. This is spindle explosion and if you need performance from a traditional database for data warehousing, it will cost you a bundle. The ratio for Sybase IQ is 1 spindle per CPU.

Sybase IQ Delivers Disk Spindle Efficiency

Sybase designed Sybase IQ from the ground up for business intelligence. Sybase IQ leverages the largest disks and delivers incredible disk spindle efficiency.

From the earlier example, if only 4 columns of a 100 column row of data are needed for analysis, using a traditional relational database, the other 96 columns must be dragged out of disk through the I/O channel and into memory. From there the specific columns needed must be found in the row and then passed to the CPU for processing.

This is all unnecessary overhead. With Sybase IQ only the 4 columns of data that are needed for analysis are taken from disk, through the I/O channel and into memory. There is no need to find the specific column of data, because Sybase IQ manages only columns, which are quickly passed directly to the CPU for processing.

In addition, the amount of data needed to perform a typical analysis is often far greater than 16,000 (16KB) characters – the maximum I/O size for a traditional database. From the previous example, let's say you want to sum up all the insurance claims in Ohio in January. This may total 512,000 characters (512KB). Assuming the largest page size of 16KB, it would take 32 reads to return 512KB of data. During this time the CPU is waiting for the information to get into memory, dramatically slowing down business intelligence queries, data loads and ad-hoc calculations.

With Sybase IQ, the maximum page size is 512KB. You only need one read to get the entire dataset.

Because of the inherent efficiency of Sybase IQ, the general heuristic for spindle management is 1 spindle for every CPU needed. Using this 1:1 ratio of spindles to CPUs, if you have 32 CPUs, you will need 32 disk spindles to keep your CPU, I/O and memory utilized (versus 320 disk spindles with a traditional relational database). Sybase IQ's spindle efficiency delivers the high performance you need, effectively and inexpensively.

Sybase IQ eliminates I/O bottlenecks by enabling page sizes that fit the specific needs of analysis. This enables maximum throughput to keep all the CPUs running and all the queries performing 10 to 100, even 1,000 times faster than traditional OLTP databases. This performance will be maintained as you add hundreds and even thousands of users. Sybase IQ's I/O efficiency can take advantage of the largest disks produced today (top end is 178GB). The table below outlines the impact of this I/O bottleneck on the number of spindles needed to keep all CPUs running at 100% efficiency.

	Sybase IQ	OLTP
1 CPU	1	10
30 CPUs	30	300
60 CPUs	60	600
90 CPUs	90	900
120 CPUs	120	1,200

Chart 2. Spindles needed per CPU: Sybase IQ vs. OLTP

I/O inefficiency has a dramatic impact on OLTP spindle counts. It is easy to see why traditional OLTP databases require small, costly disks (higher spindle counts improve CPU utilization) and why Sybase IQ can take advantage of the largest disks with no loss of scalability and CPU utilization.

Traditional Relational Databases Require Costly RAID-1

RAID (Redundant Array of Inexpensive Disks) is a method of combining multiple hard drives for data redundancy in case of a hard drive failure. RAID level 1 enables redundancy by maintaining duplicate sets of all data on separate disk drives (mirroring) for backup, recovery, high availability and fault tolerance. RAID-1 is also the only method available for shared nothing MPP environments. Why? Because shared nothing MPP does not easily or effectively take advantage of higher RAID levels or storage fabric (we will talk about RAID-5 and storage fabric in the next section).

Of the RAID levels, level 1 provides the highest data availability since two complete copies of all information are maintained. The maximum read performance that can be achieved is with disk array controllers that allow simultaneous parallel reads from both the primary and mirrored disks (maximum parallelism is two). So, what is the down side? First, RAID-1 is expensive, since twice as many disk spindles are needed. Second, the maximum parallelism is two – one from the primary disk and one from the mirrored disk.

Sybase IQ Leverages Efficient RAID-5

Sybase IQ takes advantage of RAID-5. RAID-5 stripes data across disks drives that make up one large logical volume. Since RAID-5 splits the data across multiple disks (striping), data can be accessed in parallel across all the drives. RAID-5 can read and write faster because of this advanced parallelism. RAID-1 can read across 2 drives at most (the primary and the mirror). Leveraging RAID-5 is another example of how Sybase IQ dramatically enhances the performance of queries or data loads. Also, RAID-5 disk striping is particularly effective with disk storage fabric (also known as storage area networks or network area storage). RAID-5 does not require disk mirroring like RAID-1. With disk parity, RAID-5 reduces data redundancy to the size of a single disk to withstand a single point of failure without losing data or access to data. Parity delivers good data availability and fault tolerance, which is fine for historical data (i.e. data warehouses), while at the same time saving significant disk costs.

Sybase IQ, by leveraging RAID-5, gains the full performance benefits of built-in parallelism and the additional disk reduction costs of data parity.

Traditional Relational Databases Are Static Because They Require High Levels Of Maintenance Overhead

A major limitation of traditional relational databases is the amount of development and maintenance effort needed to manage them for business analysis. This is because of the additional aggregations, materialized views and indexes needed to improve performance. The transformation process (and DBA hours) needed to aggregate detailed data and load the data warehouse can be tremendous. And once the aggregates or materialized views are built, the maintenance needed when underlying detail data changes or new columns are added is an ongoing, time consuming and extremely expensive process. In our example, if the detailed data changed (100,000 additional transactions came in for the last 6 months), to get the aggregates or materialized views synchronized with the new detail data, you would have to:

1. Drop terabytes of old aggregates, materialized views and indexes
2. Create new aggregates or materialized views based on the new detail data
3. Reload terabytes of aggregated data / materialized views into the database
4. Create terabytes of new indexes
5. Test to make sure the changes do not adversely impact current analytic applications

This is a time consuming, error prone process.

Sybase IQ Delivers Ease Of Maintenance And Dynamic Data Warehousing

Using Sybase IQ, the DBA assigns an index to each column, depending on the characteristics of the column (Gender Code - Low Cardinality, Dollars - High Cardinality, etc.). Because Sybase IQ does aggregation on the fly using the detailed information directly, few if any aggregations are needed. Generally, the DBA assigns Sybase IQ indexes per column just once. Sybase IQ automatically compresses the data as it is loaded, inserted or updated. There is little tuning needed after the initial column/index assignment.

Sybase IQ dramatically reduces load time by cutting down the number of load steps and through a parallel load/indexing facility. Typically, loading a data warehouse is a three-step process:

1. Stage the data from the transformation step
2. Load the data into the data warehouse
3. Index the data

Each of these steps can be long and involved. With Sybase IQ the Load and Index steps are combined because the data and index are a single structure. This significantly reduces processing overhead and time. In addition, this Load/Index process is done in parallel and is multi-threaded.

By delivering this kind of flexibility and ease of maintenance, Sybase IQ users are delivering new generations of applications where they add and delete columns as easily as other databases add and delete rows. Today, Sybase IQ scales to over 40,000 columns in a single table.

Traditional Relational Databases Have Poor User Scalability

As stated earlier, the current generation of relational databases is designed for transaction processing, not for ad hoc access. They are not designed for ad-hoc access by thousands of web-based users. Just imagine 10 users doing analysis when one user asks a question that does not have a supporting aggregate, materialized view or index. Your system would be brought to its knees, with queries stacking up and users waiting for results – all costing you lost manpower and lost decisions.

What would happen if you had 100s or even 1,000s of Web users, and that same user asked a question that clogs the system? What would happen if you loaded data while doing queries?

Sybase IQ Delivers Linear User Scalability

Sybase IQ uses only one-tenth the I/O bandwidth of traditional relational databases. This is one of the key reasons that as the number of users increases (which will happen quite dramatically when you extend your business analysis environment to the Web), Sybase IQ will scale with 97% linearity with the added nodes. For a traditional relational database already using 90% percent of an I/O channel, there is little room to add more users, CPUs, servers or disk without creating a huge I/O bottleneck that locks up resources.

Traditional Relational Databases Require High-Cost Backup And Restore Tools

In addition to using RAID-1 disk mirroring, extensive backup and restore technology is required to support traditional relational databases. These backup and restore applications are costly and resource intensive. They add little value to the business application running on the server. Backup and restore applications are best for OLTP applications, but are not designed for the rigors of managing 10 and 100 terabyte data warehouses.

Sybase IQ Enables Fast, Efficient And Economical Backup And Recovery

Sybase IQ provides groundbreaking Virtual Backup functionality. This capability, designed to backup and restore large tables efficiently, leverages disk array replication management from Hitachi, Sun and EMC.

Virtual Backup provides:

- Full database backups that complete in seconds or minutes, regardless of database size
- Near instantaneous application failover from primary to shadow with minimal disruption and downtime to the user during a disaster recovery scenario
- Ability to offload tasks such as maintenance process and backing up tape to the shadow devices to keep production performance at its peak level
- Power to bring terabytes of information online in minutes

By combining hardware and software technologies, the Sybase Virtual Backup functionality can bring the database backups online in seconds or minutes, rather than hours or days. The outage window to end-users is minimized while the redundancy and data protection are increased.

Traditional Relational Databases Are Poor For Real-Time, Active Data Warehousing

Traditional databases have three major deficiencies for delivering real-time, active data warehouses:

- poor user scalability
- the need for aggregations and materialized views
- the slowness of data loads and updates to the data warehouse

Real-time, active data warehousing most often comes into play to serve the real-time needs of operational users such as call center representatives, shop floor managers or distribution managers. To support operational users with real-time analysis information, you need to support a large user population. Traditional relational databases do not do this well.

Aggregations and materialized views are extract and transformation intense, causing latency between the business operations that create a transaction and the processing needed to get the data into the data warehouse for analysis. If you are looking for real-time data warehousing, latency can dramatically reduce the effectiveness of your business analysis solution. The latency of aggregation-intense extraction and transformation make real-time data warehousing difficult, if not impossible.

Today there are vendors who can "trickle" transactions into the data warehouse to enable some degree of real-time, active data warehousing. Unfortunately, organizations are exploding with transaction (atomic) and life event (sub-atomic) information. Trickling in transactions is just not good enough when millions of pieces of information are needed in real time.

As business intelligence systems become more and more complex and the need for more detailed data grows, the distance between the operational and analytical information also grows. Every transformation step diminishes the real-time value of the information. For e-business intelligence to be useful for e-business and e-commerce it must be real-time. Given the current generation of transformation technology and the limitations of conventional relational databases for e-business intelligence, this can be daunting.

Sybase IQ Delivers Real-Time Data Warehousing

Database and transformation vendors have created numerous ways to pre-aggregate data. Only Sybase has asked the question: How do you reduce the pre-aggregation process to significantly reduce integration overhead, labor costs and data explosion, while enabling real-time, active access to business analysis information by eliminating the significant time delays caused by pre-aggregation and slow data loads?

Sybase IQ is a database specifically designed for the demands of e-business intelligence. Because Sybase IQ is 10 to 100, and in many cases 1,000 times faster for data access versus traditional relational databases, Sybase IQ does not need to pre-aggregate data to gain performance. This delivers savings in operational overhead, time and expense, while enabling real-time data warehousing.

Sybase IQ delivers powerful data load capabilities by leveraging today's high-powered server technologies. Sybase IQ can load terabytes of detailed transactions per hour while having little if any impact on real-time analysis users.

Traditional Relational Databases Use MPP Ineffectively

Massive Parallel Processing (MPP) has been around for almost fifteen years. It was pioneered by Teradata and is used by IBM as the basis for their scalable solutions for very large data warehouses. For many years MPP was considered the de-facto standard for managing large amounts of data. Parallelism itself is not bad. Sybase IQ uses parallelism quite extensively to improve performance. With MPP, all operations – queries, loads, and data management – are parallelized across multiple servers, with each server owning its own disk, CPUs and memory. This is called shared nothing disk partitioning. The operation is split across the servers and each server accesses the disk it owns to process the operation with the output from each server assembled to deliver the final result. If the performance was not meeting expectations or more disks were needed, another server could be added. Fifteen years ago when memory, CPUs, disk storage and I/O channels were small, this MPP scheme made sense.

Sybase IQ Changes The Rules Of MPP

Today, there has been a revolution in hardware technology. Memory is incredibly scalable, CPUs are extremely fast, disk fabric is unlimited and I/O channels are super highways. What Sybase discovered was that the current MPP schemes are too complex, cumbersome and expensive. Why split an operation query, loads or data management across multiple servers, when with today's technology one server can handle even the most complex business analysis operation? Sybase determined that using new generation hardware technology could change the rules of the game regarding MPP. In the section "Sybase IQ and MPP," Sybase discusses the major disadvantages of MPP in more detail.

Sybase IQ and MPP

MPP has been around for almost fifteen years. For many years MPP was considered the de-facto standard for managing large amounts of data. When storage disks, CPUs, memory and I/O channels were all small, MPP made a lot of sense. Taking big business analysis problems and breaking them into small parts to be tackled by separate, independent servers was the best way to use the technology of that time. Because of the complexity and expense of this architecture, Sybase developed its own massively parallel product called Sybase MPP. Sybase realized in 1994 that one database could not do it all and that specialized databases are needed for data warehousing, OLTP and mobile applications. With the purchase of Expressway in 1995, Sybase started on the path to change the way data warehousing would be done. In the following sections we will compare and contrast current MPP technologies with Sybase IQ.

Starting With A Single Node Of MPP And Sybase IQ

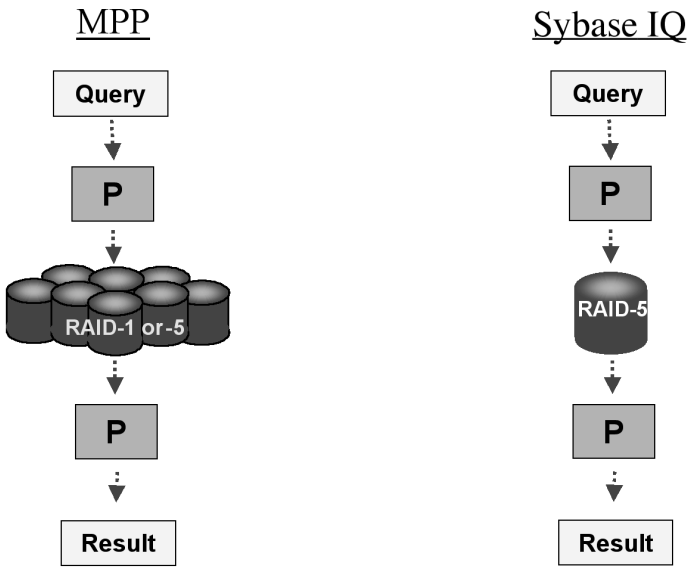

Diagram 4: A single MPP node is a standard relational database

The disadvantages of traditional relational technologies as compared with Sybase IQ have already been discussed. In the context of MPP, however, imagine starting with a single node of MPP. In actuality you would have a traditional relational database. Right from the start all the disadvantages discussed earlier that applied to traditional relational databases apply to MPP technologies as well, because MPP databases at their core leverage traditional relational technology:

MPP leverages traditional relational approaches	Sybase IQ delivers a proven and innovative approach
To gain performance traditional relational databases are inherently inflexible	Sybase IQ delivers performance and flexibility through column based indexes
Traditional relational databases need 4 times as much storage because of inefficient indexes, aggregates, and materialized views	Sybase IQ delivers compression to atomic data through efficient indexes and eliminates the need for aggregates and materialized views
Traditional relational databases use 10 times as many spindles as Sybase IQ	Sybase IQ delivers disk spindle efficiency at a rate of 1 disk spindle per CPU
Traditional relational databases require costly RAID-1	Sybase IQ leverages efficient RAID-5
Traditional relational databases have high levels of maintenance overhead because of complexity and inflexibility	Sybase IQ delivers ease of maintenance through simplicity
Traditional relational databases have poor user scalability	Sybase IQ delivers linear user scalability
Traditional relational databases require high-cost backup and restore tools	Sybase IQ enables fast, efficient and economical backup and recovery
Traditional relational databases are poor for real-time, active data warehousing because data loads drag down query performance, there is poor user (and query) scalability and high overhead from aggregates and materialized views	Sybase IQ delivers real-time data warehousing because query response is fast even while loading data. Sybase IQ lowers overhead because there is no need for materialized view or aggregates

Massive Parallel Processing – MPP

MPP works because it is massively parallel – massive numbers of CPUs working in parallel to solve problems, leveraging massive amounts of memory, which in turn leverage massive amounts of disk and I/O. With traditional MPP it is the massiveness of the parallelism that delivers scalability and performance. All queries, data loads and maintenance must be parallelized across every server processor node to gain optimal performance. In other words, every query must leverage every server to be effective and every server node must be equally fast to be effective. Each server owns and manages its own CPUs, memory, data and disks (called a shared nothing architecture) and each server in turn finds the piece of data needed for the query and passes it down to be assembled into the final result. This also means that if one node is lost, the entire MPP system is lost. MPP provides excellent performance in situations where a problem can be partitioned so that all nodes can run in parallel with little need for the nodes to share data across the server partitions.

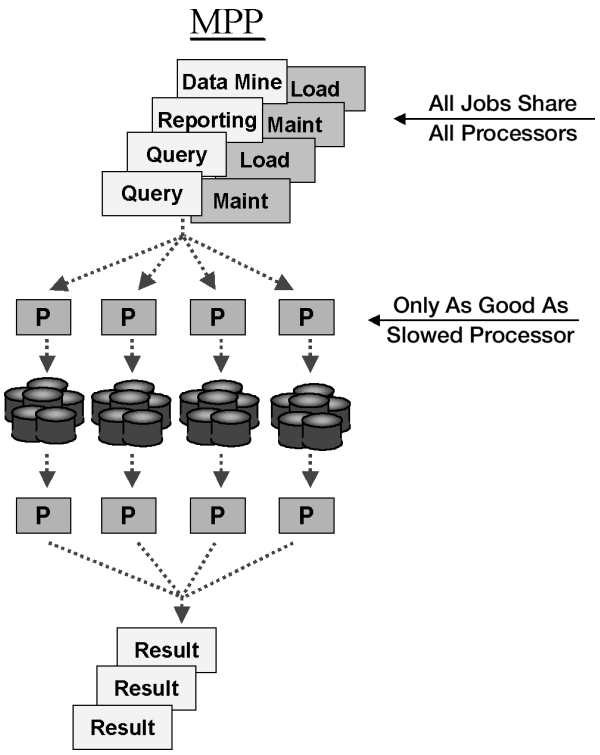

Diagram 5: MPP is about massive amounts of servers, disk, I/O and memory

Sybase IQ – A Totally Different Approach

Sybase IQ leverages a new generation of technology that delivers scalable memory, extremely fast CPUs, unlimited disk fabric storage, and I/O channels the size of super highways. Because of this new generation of technology, Sybase IQ takes advantage of individual servers performing specific functions, while sharing common disk storage. This is called multiplexing. It is very different from MPP, which attempts to leverage all the servers in parallel for an individual query, data load or maintenance operation and isolates storage, memory and CPUs per server.

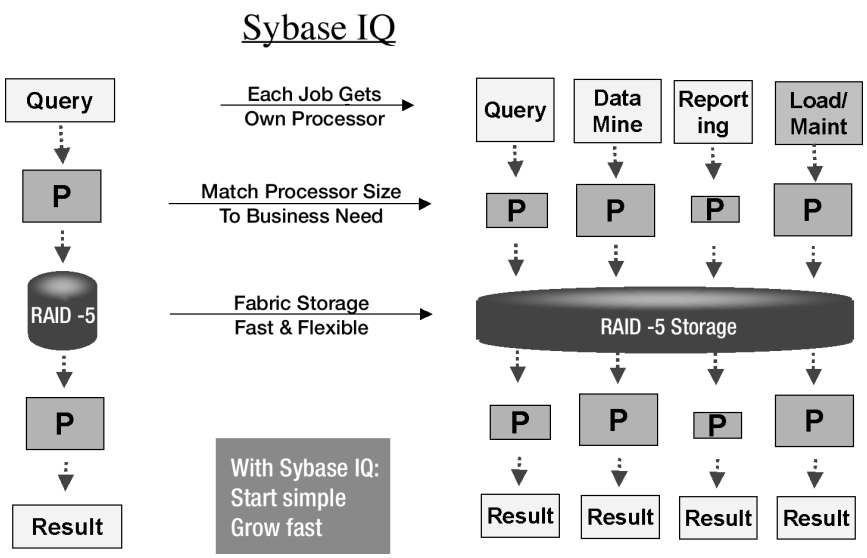

Diagram 6: Multiplexing is about smart incremental growth

Multiplexing is a very different architecture compared to MPP. Intuitively, it may seem that one query using all server resources would be faster with MPP. But this does not take into account the revolution in hardware technology. With all the advantages mentioned in the prior section, Sybase IQ enjoys considerable advantages by placing individual operations on individually multiplexed servers. For example, you could set up one server just for queries, another server for data mining and a third for data load and maintenance.

The next sections address four major issues that users of MPP systems have to face and how Sybase IQ overcomes these issues:

- **Data Skew** – What happens when business queries do effectively leverage all the servers in parallel?
- **High Data Sharing Between Nodes** – What happens when there are queries that force nodes to share high amounts of data with other nodes?
- **Runaway queries, data mining, data loads and other heavy users** – What happens when one application takes over the system?
- **Loss of a node** –What happens when a node is lost?

Distributing Data Across MPP Nodes

Before discussing data skew, you need to understand how data is distributed (also called partitioning) across MPP nodes. There are two typical ways data is distributed in an MPP environment – Range and Hash.

Range Partitioning

To illustrate range partitioning, let's start with a data warehouse with three major dimensions: Time, Markets and Products. There are also hundreds of other secondary dimensions like Customer, Brand, State, Country, Shipping Channel, etc. The goal of partitioning is to distribute the data smoothly across nodes. Let's say you pick Time for the range distribution key. Time will only work if there is no seasonality in your data, such as Christmas. Seasonality will force lumpiness in the distribution. With MPP you need smooth distribution across disks to optimize utilization. If you choose Geography for your distribution key, this will work only if the number of transactions in North America is the same as in France or China. Again, you want smooth distribution. You would have the same problem using a Product distribution key. These are all examples of what is commonly called distribution skew – the difficulty of distributing data evenly across disks to optimize the utilization of all the parallel processors.

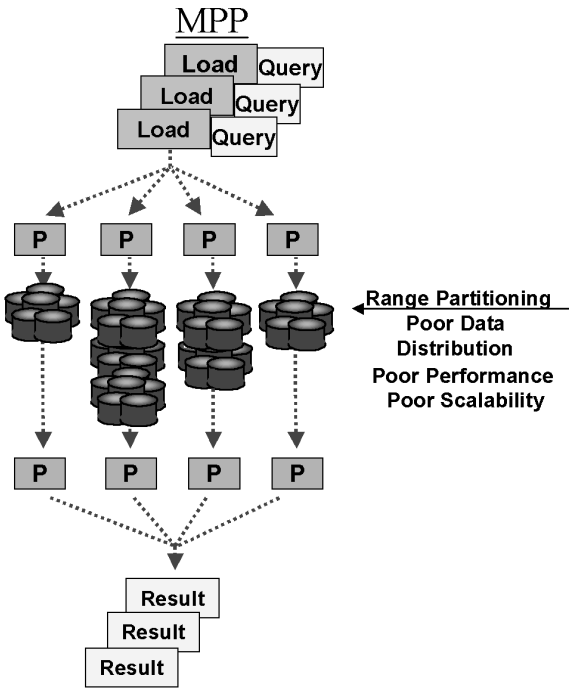

Diagram 7: Range Partitioning and Distribution Skew a constant challenge

Hash Partitioning

There is a lot of science concerning distribution skew and how to smooth distribution of data across disks. The most common way to optimize distribution is with hash partitioning. Hash partitioning uses algorithms to optimize distribution. For example, you may concatenate the Time, Market and Product dimensions together, then divide the results by some magic number, times the square root of another magic number. This would result in a hashed distribution key spreading data across nodes.

The problem is that even if you do distribute the data smoothly across MPP nodes, that does not mean the business queries will effectively leverage this distribution. We will see how this occurs in the next section.

Data Skew

The problem is that the distribution key is all about loading the data. It is less about accessing data. Even though the data is distributed evenly across disks, there can still be lumps in the distribution for those queries that use dimensions other than the ones chosen for the distribution key. In the prior example, a distribution algorithm centered on Time, Market and Product dimensions. But what if you need to do queries, data mining, reporting or advanced analysis for the other 100 dimensions? What you will get are hot spots on the disks where there are concentrations of secondary dimensions, which many want to query. There will also be cold spots for concentrations of products, market and time data where nobody needs to analyze. Count on it. For those queries using secondary dimensions that are not part of the distribution key, only a small part of the MPP resources are being used to perform analysis. Also, these nodes where there are hot spots become bottlenecks for the overall system, because parallelism works when all the nodes are being leveraged equally. This problem is called data skew. Multiply data skew by hundreds and thousands of users and queries and you can see how the problem can grow. The very nature of enterprise business intelligence to respond to unforeseen business opportunities and threats will blow huge holes in MPP architectures.

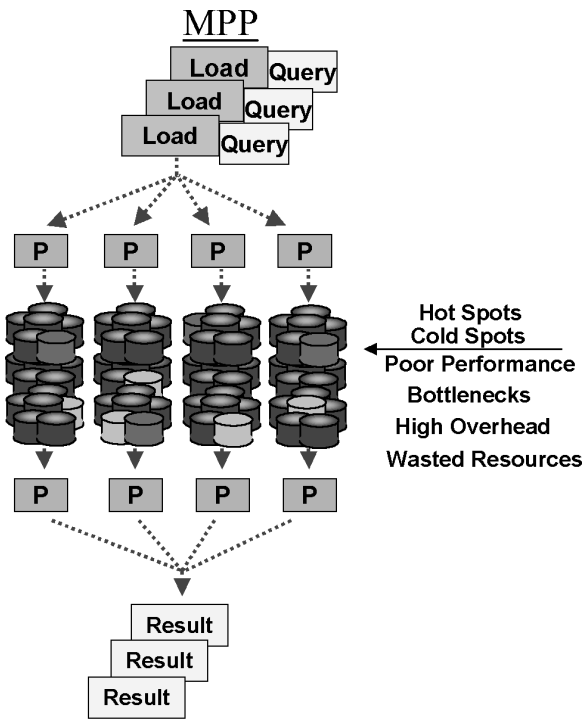

Diagram 8: Data Skew

Sybase IQ Leverages Round-Robin Disk Striping For Random Distribution

With Sybase IQ, data is distributed evenly across disks with disk striping in essentially a round-robin fashion. This round-robin method evenly distributes records across all disks, without regard to the data values being stored or distribution algorithm (like Hash). This method is simple, clean and easy. It permits even workload distribution for subsequent business queries by spreading read operations across multiple disk spindles. In addition to Sybase IQ striping, Sybase IQ leverages RAID-1 and -5 striping (Stripe on Stripe). This double layer of striping eliminates data skew.

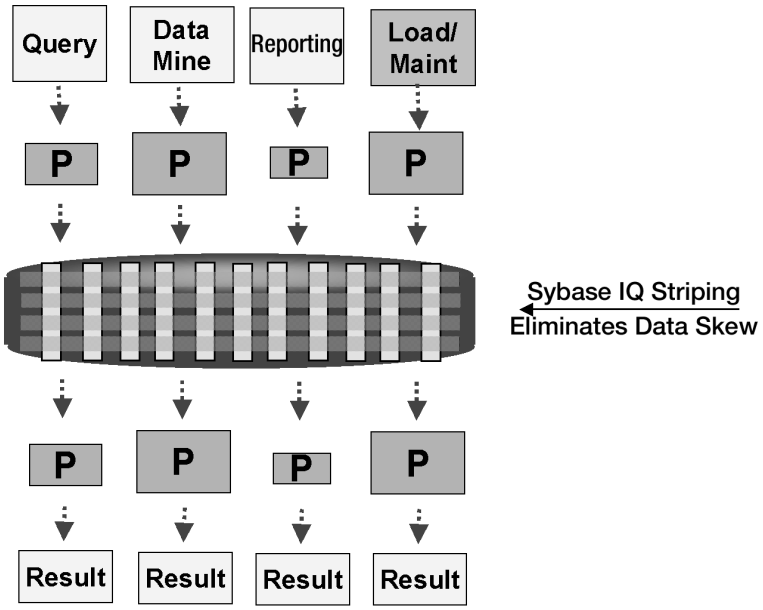

Diagram 9: Sybase IQ eliminates data skew

Heavy Data Traffic Between Nodes

Whenever there is a query that requires data from multiple tables, those tables must be joined together to determine what information is mutually shared between them. When there is a match, all the data from each table is passed to a shared processor to be combined. With MPP, a join may have very good or very poor performance depending on the relationship between the partition key and the join key. If the partition key is the same as the join key, a process on each of the MPP nodes can perform the join operation on its local data, most effectively utilizing parallel operations. If the partition key is not the join key, each record on each node has the potential to join with matching records on all of the other nodes. If there are N MPP nodes, the join operation requires each of the N nodes to transmit each record to the remaining N-1 nodes, increasing communication overhead and reducing join performance. The problem gets even more complex when real-world data having an uneven distribution is analyzed.

Unfortunately, with ad hoc queries, data mining and complex business analysis predominating in decision support systems, the case of a partition key not being the join key can be quite common. To make matters worse, MPP partitioning decisions become more complicated when joins among multiple tables are required.

Sybase IQ handles data sharing issues the same way it handles runaway queries, data mining, data loading and other heavy usage.

Runaway Queries, Data Mining, Data Loading And Other Heavy Usage

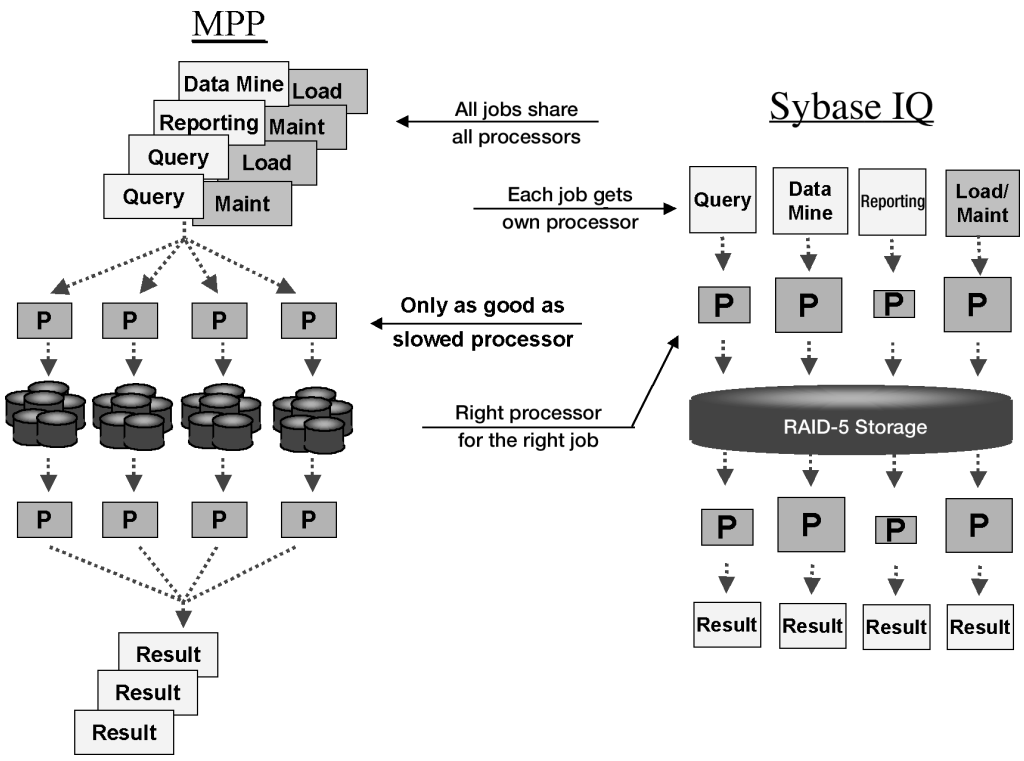

Diagram 10: MPP performance is dependent on what queries and data loads are running

With MPP, all the queries, data loads, and maintenance are processed in parallel across all of the server nodes. For optimal query performance, data must be retrieved from every server node. To perform a data load, the data must be distributed across server nodes based on distribution keys and algorithms. For maintenance, database definitions, index creation/deletions, etc., must be synchronized across servers. These operations occur in a first in/first out fashion, much like an automotive assembly line in a factory, where each assembly line is responsible for their part of the car. At the end of the process, all the parts come together at a final assembly point to produce the finished car. As long as all the assembly lines are working in parallel, the MPP factory works just fine. But what if the chassis line breaks down? What if a special order comes in? What if one of the assembly lines is upgraded with faster equipment? What if one of the lines gets overloaded? This is the problem MPP has for delivering consistent query performance, managing runaway queries, loading data, delivering reporting and performing maintenance. If one server fails or becomes overloaded, if one data mining operation hogs the system, if a data load window is missed, if there is a mass update, MPP can become problematic, volatile and fragile very quickly.

Sybase IQ Delivers Fast And Consistent Performance With No Inter-Node Traffic

Because Sybase IQ multiplexes independent server nodes for any given data warehouse process, there is total flexibility in managing query, load and maintenance performance. Because each node is independent, the data-sharing problem simply goes away. Diagram 10 illustrates an example in which query operations are on one server, data mining on a second server, reporting on a third server, and data loading on a fourth. Three of the four servers are performing read functions that will have very little impact on the individual performance of any server. The same is true for the Writer node. Because the Write node owns the write lock and because of a process called version control, the readers are unaffected by the writing process. In other words, all the read and write servers deliver fast performance. In benchmarks performed by the independent auditor, Francois Raab, Infosizing, Inc., it was determined that this scalability is nearly linear, meaning that every time memory, CPUs, disk or I/O controllers are added, they will be utilized fully and consistently with 97% scalability.

MPP Performance Limitations	Sybase IQ Solution
How do I keep a runaway query from slowing down overall performance?	With Sybase IQ a runaway query will only impact the server it is running on and will not slow down or bring down the data warehouse.
How do I keep data loads and updates from interfering with ad-hoc queries and reports?	Data loads and updates are performed on a separate server. With Sybase IQ's version, control write processing does not interfere with any reader servers.
How do I keep data mining applications from using all the resources?	With Sybase IQ read-intensive applications like data mining can be on a separate server and will not slow down or bring down the data warehouse.
How do I perform database maintenance and deliver information for 24/7 real-time worldwide operations?	Data loads and updates are performed on a separate server. With Sybase IQ's version control write processing does not interfere with any reader servers. Your data warehouse is up 24/7, even while you are doing database maintenance.
What do I do when 100 queries (or 1,000 queries) hit the server at once?	Sybase IQ delivers linear scalability. You can add servers, CPUs and memory, and know that these resources will deliver results no matter the load.
How do I keep my data mining users from running into my query or reporting users?	You can place each of these major operations on a separate server.
What do I do if I miss my data load window?	Data loads and updates are performed on a separate server. With Sybase IQ's version, control write processing does not interfere with any reader servers.

Loss Of A Node

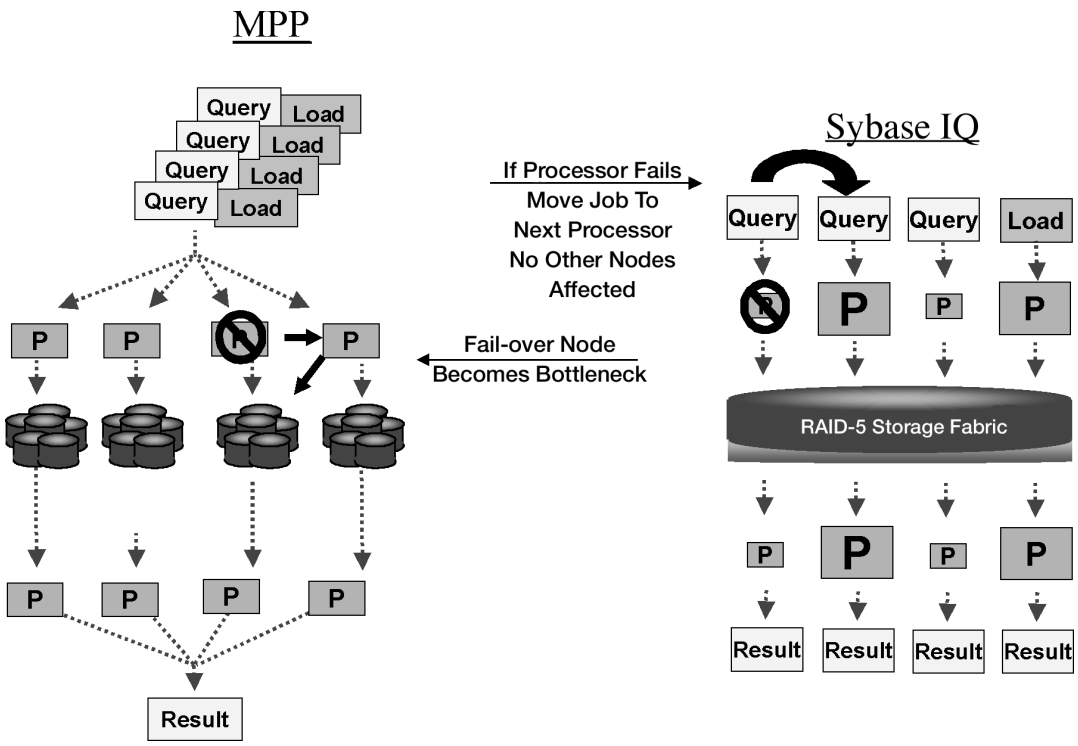

Diagram 11: Loss of an MPP node slows the entire system

With MPP the loss of a node is always costly. If you lose a node in an MPP environment, it is difficult to continue operations because when a node is lost, another node must take on the load of the lost node. This means that the recovering node has to work twice as hard and performance suffers. Because this is all happening in parallel, all the nodes are equally slowed (remember the assembly line). There is no way around this. This is especially a problem if your business intelligence environment is mission-critical. You will need high performance 24/7.

Sybase IQ Delivers Flexibility

With Sybase IQ, each node is stand-alone. If any node goes down, only the fail-over node is affected. All the other nodes run unaffected. Also, because Sybase IQ takes advantage of RAID-5, Data Parity, there is good data availability and fault tolerance, as well as time savings and reduced disk costs.

For more information, supporting case studies and success stories, please visit our website at: www.sybase.com/products/bi/asiq

International Contacts

Argentina
+5411 4313 4488

Australia
+612 9936 8800

Austria
+43 1 504 8510

Belgium
+32 2 713 15 03

Brazil
+5511 3046 7388

Bulgaria
+359 2 986 1287

Canada
+905 273 8500

Central America
+506 204 7151

Chile
+56 2 330 6700

China
+8610 6856 8488

Colombia
+57 1 218 8266

Croatia
+385 42 33 1812

Czech Republic
+420 2 24 31 08 08

Denmark
+45 3927 7913

Ecuador
+59 322 508 593

El Salvador
+503 245 1128

Finland
+358 9 7250 200

France
+33 1 41 91 96 80

Germany
+49 69 9508 6182

Greece
+30 1 98 89 300

Guatemala
+502 366 4348

Honduras
+504 239 5483

Hong Kong
+852 2506 6000

Hungary
+36 1 248 2919

India
+91 22 655 0258

Indonesia
+62 21 526 7690

Israel
+972 3 548 3555

Italy
+39 02 696 820 64

Ivory Coast
+225 22 43 73 73

Japan
+81 3 5210 6000

Kazakstan
+7 3272 64 1566

Korea
+82 2 3451 5200

Malaysia
+603 2142 4218

Mexico
+5255 5093 8500

Netherlands
+31 20 346 9290

New Zealand
+64 4473 3661

Nigeria
+234 12 62 5120

Norway
+47 231 621 45

Panama
+507 263 4349

Peru
+511 221 4190

Philippines
+632 750 2550

Poland
+48 22 844 55 55

Portugal
+351 21 424 6710

Puerto Rico
+787 289 7895

Romania
+40 1 231 08 70

Russian Federation
+7 095 797 4774

Slovak Republic
+421 26 478 2281

Slovenia
+385 42 33 1812

South Africa
+27 11 804 3740

South Korea
+82 2 3451 5200

Spain
+34 91 749 7605

Sweden
+46 8 568 512 00

Switzerland
+41 1 800 9220

Taiwan
+886 2 2715 6000

Thailand
+90 212 325 4114

Turkey
+90 212 325 4114

Ukraine
+380 44 227 3230

United Arab Emirates
+971 2 627 5911

United Kingdom
+44 870 240 2255

Venezuela
+58 212 267 5670

Asian Solutions Center
+852 2506 8700

For other Europe, Middle East, or Africa inquiries:
+33 1 41 90 41 64 (Sybase Europe)

For other Asia Pacific inquiries:
+852 2506 8700 (Hong Kong)

For other Latin America inquiries:
+925 236 6820

Sybase, Inc.
Worldwide Headquarters
One Sybase Drive
Dublin, CA 94568-7902 USA
Tel: +800 8 SYBASE
www.sybase.com/bi

Part #L02369