Strategic Information Management Technology:

Workloads Matter in Managing Gigabytes to Petabytes

Prepared by
William McKnight

A starting point for information technology managers who need to select systems to store valuable data for their business

Contents

Preface

This white paper is intended to provide a consolidated starting point for information technology managers who need to select systems to store retrievable analytic data for their business. Due to increasing data volume and data's high utility, there has been an *explosion of capabilities* in the past few years brought into the early mainstream of computer system buyers. While stalwarts of our information, like the relational row-based enterprise data warehouse (EDW), are highly popular, it is widely acknowledged that no single solution will satisfy all enterprise data management needs.

Though storage remains historically inexpensive, costs for keeping "all data for all time" in an EDW, are still escalating into discussions by upper management due to the higher volumes of data. That is driving some heterogeneous platform deployment as well. The key to making the correct data storage selection is an understanding of your workloads – current, projected and envisioned. This white paper will explore the major categories of *information data stores* available in the market and help you make the best choices based on the workloads.

The categories of information data stores being made available are being influenced by three major current trends. The three trends are:

1. Heterogeneity of Systems

The rapid spread of systems within companies, without centralized oversight, has led to unprecedented numbers of technical system footprints in companies today, along with unprecedented levels of skills required to support the systems. The vendor consolidations in the marketplace have done little to thwart this. The days of single-vendor system environments are over. The days of data integration as the "perpetual short-term" solution to interoperability are here.

2. All Data Required

Most companies have analysts who require immediate access to all information ever collected. Archiving to inaccessible storage mediums is increasingly a relic of the past.

3. Immediate Actions Required

The one-way data flow into a downstream analytical environment where all company decisions are made must also rest in peace now. Waiting on a flow of data to a separate environment involves too much latency for an agile, competitive organization that needs to make immediate decisions and capitalize on specific opportunities. The data warehouse is still mandatory, but so is decision making and process execution at the very earliest opportunity. Operational business intelligence comprises the ideas behind supporting the immediacy of action with information. The popular Master Data Management project – master data developed and stored operationally, and potentially augmented with analytics – is one such implementation of operational business intelligence.

A summary of future information stores and their characteristics can be found in Figure 1.

	Real-Time	Small Data ok	Terabytes	Petabytes	Historical Data	Known Queries	Unknown Queries	Ad-hoc	Highly Specialized Queries	Unstructured Data	Source Data Supplier to Other Systems
Operational Systems	x	x									x
Data Warehouse			x	x	x	x	x	x			x
Data Mart[1]		x	x			x	x	x			
Multi-dimensional Mart		x					x		x		
Columnar Database			x		x	x		x			
Data Stream Processing	x	N/A	N/A	N/A	N/A	x			x		x
MapReduce			x	x					x	x	x

Figure 1. Recommended use of information stores.

1 In this paper, data mart is a subset of the data warehouse – in size and/or function.

Operational Systems functions are as diverse as it comes. And to use "ERP" as a label for some operational systems is almost equally meaningless[3] since ERP has just about become synonymous with software applications. However, from an information data store perspective, these terms still provide meaning given that almost all today are completely undergirded by row-based relational databases.

These databases are largely prebuilt data models with generic extensions ("column 56", "column 57", etc.) and, due to the need to have widespread appeal and support a superset of functions, complex to use for any given company. Despite the need for flexibility and vendor claims, many companies struggle with extending the data models to support their specific needs.

While the lack of flexibility of the data model is a primary database-related complaint for operational systems, the structure of operational systems can also be limiting when making information-based, real-time decisions and triggering information-based, real-time actions. Operational systems must extend or be supplemented by these capabilities.

Acme

Acme has many modules of an encompassing popular ERP system installed but has also numerous other systems that either predate the major ERP or provide specific functionality not found in the ERP.[4] It has systems for CRM, ordering, customer support, servicing, call center management, general ledger, payroll, product lifecycle management, quote-to-cash and numerous others that no single person knows. Like most enterprises, most of these have an associated relational database and fixed data model which present a constant challenge to Acme, as its business expands, both by volume and into new areas.

Acme 2011

Acme is our theoretical Fortune 100 telecommunications company. In 2011, it finds itself at a crossroads. Each business unit has grown an IT organization competitive with the centralized IT group, which came out of the company's first business unit about 20 years ago. There is an uncertain relationship with central IT where company executives want to move fast and execute on sponsored initiatives. Business unit executive clout and assertiveness goes a long way in influencing behavior at Acme.

Acme has a series of data warehouses, some overlapping in data and function. Some users routinely switch between warehouses throughout the day. Speaking of users, they are continually challenged by the constant system changes. They spend as much time trying to understand the systems as they do using them. Yet despite many failed attempts to update systems, 40-year-old mainframe and file-based systems remain rampant. New employees spend months getting productive – in a single system!

Interestingly, with the rush to find cost savings, cloud infrastructure is being used by Acme's customers. However, Acme, finds itself unable to deploy its own systems there. Acme's five data centers represent a menagerie of information technology circa 1976 - 2006. Only in small areas has it had the ability to take advantage of the explosion in capabilities the market provides. Acme's potential is being missed as a result. Acme's information management stores need a significant transformation.

I will delve into Acme's information infrastructure in each of the information stores, and we'll find out at the end how Acme is doing after some transformation.

2 Classifying systems as "operational" belies a certain prejudice that is rightfully, yet slowly, being eradicated. However, from a data warehouse practitioner's perspective, it refers to the systems that "run" the business and feed the data warehouse.

3 Using vendor terminology.

4 And, let's face it, another reason would be politics.

The relational theory is based on tables which are a collection of rows for a consistent set of columns. The rest of the relational database is in support of this basic structure. Row orientation describes the physical layout of the table on disk as a series of rows with all columns of the row stored.

By far, most data warehouses are stored in a relational row-oriented database. The data warehouse has been the center of the post-operational universe for some time as it is the collection point for all data that is interesting to the post-operational world. Reports, dashboards, analytics, ad-hoc access and more are either directly supported by or served from the data warehouse. Furthermore the data warehouse is not simply a copy of operational data, but frequently the data goes through transformation and data cleansing before landing in the data warehouse.

Workload

Given the capabilities of the alternative information stores, the relational row-oriented data warehouse ("data warehouse") should be considered the default data store for reports and analytics.

Given the predominance of the data warehouse today, and if your urgency factor is low, it could be an evolutionary, not revolutionary, process to move data and function off the data warehouse.

At a minimum, however, the data warehouse will be the historical information data store for an organization. Unlike other information stores that exist for solving a tactical analytical need, the data warehouse exists as a permanent[5] record. Increasingly, laws are dictating the minimum level of record-keeping that companies must keep causing data to be retained longer in the data warehouse.

The onsite data warehouse will increasingly contain solid-state components and other means for high-use data to support buffering of data and reuse of previously queried results and optimizer plans.

Since most historical data should be kept in the data warehouse, it should consequently serve as the source data to any multi-dimensional structure or other data marts. It should also continue to serve as the point in the environment where data quality is assured.

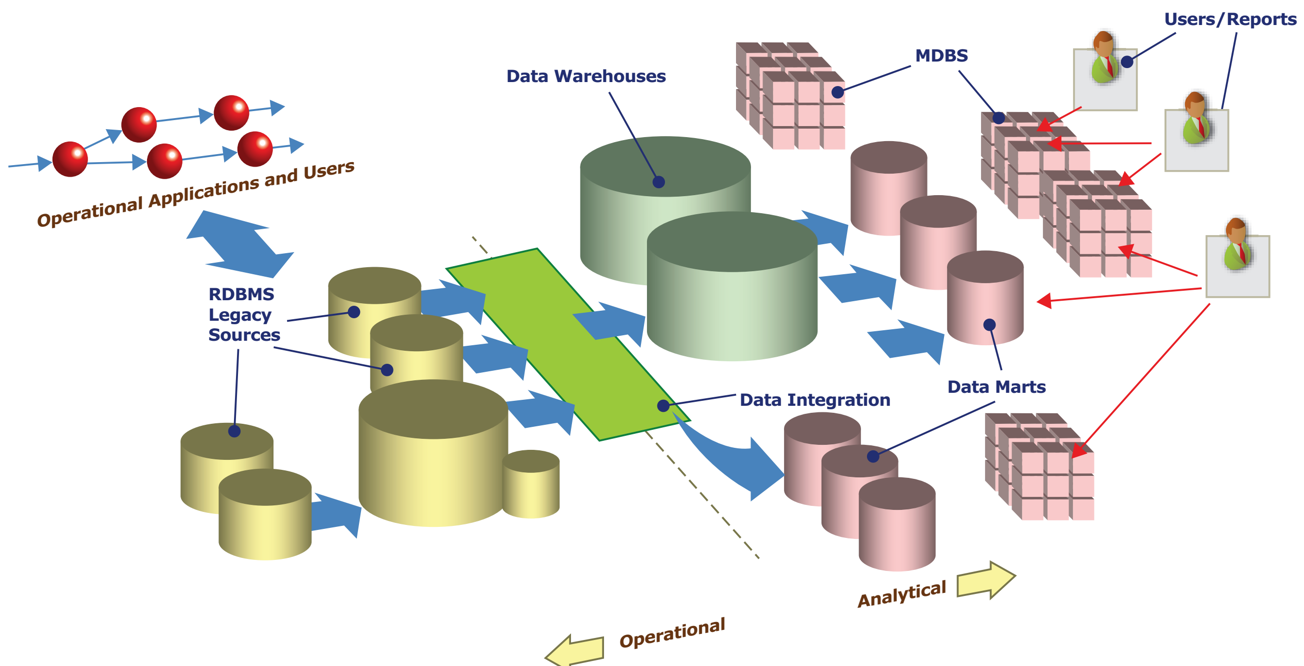

Figure 2. Information management environment before optimizing for workloads.

5 Or as long as the company wishes to keep the information.

Teradata Corporation

Teradata's offerings undoubtedly stand out for data warehouse and data mart platforms. The Active EDW line, based on the Teradata® Database, supports over 50% of large-scale data warehouses today. All database functions in Teradata Database are done in parallel all of the time using multiple server nodes and disks. These functions include: Table Scan, Index Scan, Indexed Read, Join, Sort, Row Redistribution, Parse, Catalog Read, Index Creation, Insert, Delete, Update, Load, Backup and Restore. All units of parallelism participate in each database function.

Teradata's cost-based optimizer is grounded in the knowledge that every query will be executing on a massively parallel processing system. It knows the system and leverages parallelism for all operations. It has data statistics and demographics and knows how many of a table's rows each parallel unit will be handling during each task. It uses this information to select the most efficient query plan to ensure the least resources are used to produce the fastest response to the user request.

Teradata manages contending requirements for resources through dynamic resource prioritization that is customizable by the customer. The past several years have seen tremendous investment in Teradata Database and the hardware platforms.

The server node interconnect was designed specifically for a parallel processing multi-node environment. The interconnect is a linearly scalable, high-performance, fault tolerant, self-configuring, multi-stage network. It supports guaranteed delivery, point-to-point and broadcast connections at the hardware level. It also provides direct access to high-performance messaging services with rich additional functionality for parallel and distributed processing applications.

In addition, Teradata has extended the leadership from their active EDW into their new appliance family for midmarket enterprise data warehouses as well as data marts for large companies like Acme.

The Teradata Data Warehouse Appliance supports the EDW approach to building the data warehouse and is the Teradata appliance family flagship product. It is suitable for an upper midmarket true EDW or as the platform for a focused application. With four MPP nodes per cabinet and scaling up to 11 cabinets with 12.6 terabytes each, the Teradata Data Warehouse Appliance can manage up to 140 terabytes, with the workload characteristics of a typical data warehouse – multiple, complex applications serving a wide variety of users.

The Teradata Data Mart Appliance is a more limited-capacity equivalent of the Teradata Data Warehouse Appliance and is ideal for the departmental or midmarket platform. It's a single node, single cabinet design with a total user data capacity of six terabytes. It should be noted, though, that a single node environment comes with the potential for downtime in the unlikely event that the node fails – there is no other node to cover for the failure.

The Teradata Extreme Data Appliance is also part of the Teradata appliance family, and represents affordability for managing large quantities of data. While the Active EDW tops out at 10 petabytes, the Extreme Data Appliance will scale to 50 petabytes. A system of this size would have fewer concurrent users because it supports deep history analytics, not recent data reporting. Statisticians like this kind of system with huge amounts of data to calibrate their predictive analytic models.

The Teradata Extreme Data Appliance is designed for high-volume data capture such as that found in click stream capture, call detail records (CDRs), high-end POS, scientific analysis, sensor data, and any other specialist system useful when the performance of straightforward, non-concurrent analytical queries is the overriding selection factor. It also can serve as a surrogate for near-line archival strategies that move interesting data to slow retrieval systems, and it will keep this data online.

Acme

Acme has several data warehouses[6] built on a variety of platforms. It also has specialized single purpose data marts. Some are built by Acme and others are prepackaged by various vendors. Those that are built by Acme are likely to have the same data model as the source system(s) that send it data, which may have gotten the data out of the operational system, but does nothing for data quality or modeling for accessibility.

6 Which, cynically speaking, could mean it has none, but let's go with what they say.

Multi-dimensional databases (MDBs), or cubes, are specialized structures that support 1-2 second access to summarized data. The information data store associated with multi-dimensional access is often overshadowed by the robust data access speed and financial calculation capabilities. However, it is the multi-dimensional databases that create storage and processing overhead for the organization.

MDBs are what I call "hyperdimensional." The logical model is dimensional, but the physical rendering is fully denormalized so there are no "joins." Like a data warehouse, they are built from source data. The point that must be made is that these structures grow (and build times grow) tremendously with every column added. The data access layer sometimes hides the complex underlying multi-dimensional structure.

Workload

MDBs are a "last resort" when a highly tuned relational database will not give the performance that is required. MDBs can be quickly built to support a single query or just a few queries, but there often is a high price for this tempting approach.

Since many multi-dimensional databases land in companies as a result of coming in with packaged software, it must also be said that if the overall package provides a true return on investment – versus the alternative of building the package functionality in the shop – then that is another valid reason for a multi-dimensional database.

If a query is paired well with the MDB (i.e., query asks for most columns of the MDB), the MDB will outperform the relational database, potentially by an order of magnitude. Ensure that the maintenance (cycles, storage) is well-understood for a multi-dimensional database, and MDBs can be well positioned in support of workloads that are well-defined and understood.

Teradata Corporation

Teradata supports MDB-like structures with the Teradata BI Optimizer (BIO), consisting of the Teradata Schema Workbench, Teradata Aggregate Designer and Teradata OLAP Connector components. Teradata BIO designs multi-dimensional summaries needed into Teradata's aggregate join indexes. This allows, for example, Microsoft Excel Pivot Tables to connect to the data warehouse using the MDX Query Language so you can do ad-hoc query and analysis on the data.

In addition, Teradata supports the MDBs from SAP BW and Oracle Essbase by extending their capability. Teradata BIO allows queries to pass-through the MDB query to the Teradata Database to get the detailed data that the aggregate summaries are derived from. This is a great way to utilize an MDB while managing its growth.

Acme

Multi-dimensional products came to Acme early on through the Finance department. Before the Finance department turned the MDB systems over to an unprepared IT organization, other departments copied the approach, and now overall, Acme has 259 MDBs. Most of these are no longer accessed. Acme does not know which MDBs are not used. There are so many undocumented MDBs that it is not unusual for Acme to build new cubes that are exactly the same as ones already built – and reload them nightly in long batch cycles. In the eyes of some of the user community, query has become synonymous with MDB.

Some relational database products use columnar data structures instead of row-oriented structures. In columnar databases, each physical structure contains all the values of one column of one table. This isolates each column, making the column the unit of I/O and bringing only the useful columns into a query cycle. This is a way around the all-too-common I/O bottleneck that analytical systems face today. Another way is data compression.

Workload

The columnar information store has a clear ideal workload. That workload is when the queries require a small subset (of the field length, not necessarily the number of columns) of the entire row. Columnar databases will show their benefits best with large row lengths and large data sets. Single-row columnar retrievals will underperform the row-oriented database and, since loading will be of multiple structures, loading will take longer in a columnar database. Columnar functions like average and sum will perform well in a columnar database, compared to a row-wise database.

It must be the value of performance of that workload that differentiates the columnar database for it to make sense. Interestingly, upon further analysis, many shops have several workloads that will perform better in a columnar database.

Aster Data, a Teradata Company

With the acquisition of Aster Data, Teradata has added columnar capabilities to its abilities. Aster Data gives the user a choice of column or row storage for reporting or iterative analytic queries. Sitting alongside row-based storage capabilities, queries can use the row-based information store when the query is CPU intensive and complex, or the query can select from the column-based information store if it is believed that the query meets the columnar requirements.

Acme

Some IT professionals at Acme have heard about columnar databases, but have not had the time nor the energy to educate the organization to get one into a procurement cycle. Acme does have significant long-running queries that count, summarize, group and average data that holds promise for columnar databases.

MapReduce (MR) is a parallel programming framework for large-scale data. It is not a database nor a direct competitor to the other information stores. However, if you look closely, the tasks you will do with MR will be a small subset of the tasks you might do with a relational database. It's just that the MR data is a different profile of data than what would be stored in a relational database.

A new scale of business challenge requires new solutions like MR. Most of the massive, unstructured, web-scale data is relatively unimportant, but each bit contributes to multiple aggregations that enhance profiles and otherwise contribute to operations. Each bit may also be a 'gem' that should drive a business process or be interesting to a batch process. And finally, just finding a way to actually store the data will allow for future processing on that data, should it be necessary. If any of the data is thrown away, that's value lost. In the competitive economy and where the market is providing solutions, even though it may be different from what's currently deployed, it must be considered.

MR consists of two programs the programmer builds – the map module and the reduce module. These are passed to the MR framework that runs them on the target nodes. So most of MR is focused on "how many nodes" and "which nodes" to run the map and reduce functions on. What makes MR different from relational databases is its highly flexible support for any programming language. A programmer can use almost any language to perform sophisticated functions and analysis without being limited by the confines of a relational database.

Workload

The workload for MR will be data that is massive not only from the standpoint of collecting history over time, but also from the standpoint of high volume in a single day. From a processing perspective, you would put data into MR where the functionality required was limited to batch processing with a limited set of query capabilities. Since most MR systems are flat file-based with no relational database for performance, nearly all queries run a file

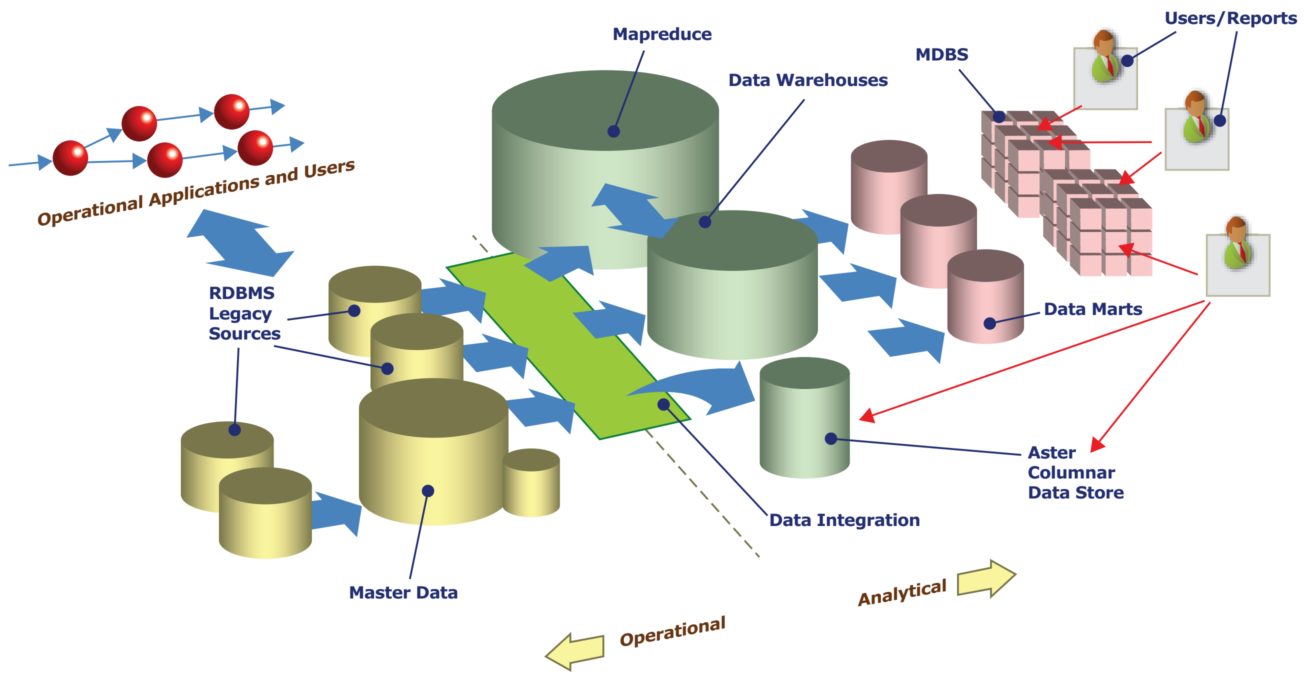

Figure 3. Information management environment after optimizing for workloads.

scan for every task, even if the answer is found in the first block of disk data. MR systems are primarily for unstructured data, for that is the data that grows enormously large quickly yet only needs batch processing and a basic set of query capabilities.

Aster SQL-MapReduce®

Aster Data has patented In-Database MapReduce and SQL-MapReduce® with nCluster, a hybrid row/column store with a MR approach. Its MPP architecture makes it work for predictable as well as ad-hoc analytic use cases. Aster blends the performance of a relational database (i.e., indexes, optimizers, etc.) with the programming flexibility of MapReduce (Java, Perl, Python, .Net, etc.). A programmer can run complex programs in parallel without the traditional limitations of other MR systems. Aster can run MR extremely fast without being limited to batch jobs only.

SQL-MapReduce is architected for optimal in-database analytics execution, and is best for custom transformations and aggregations, inter-row analysis, nested sub-queries and analysis that requires the reorganization of data into new structures.

Acme

Acme has 213 web properties. It has been storing selective click stream data into various ODS databases, one per web site in most cases. Some of the data is analyzed, and this has influenced web design. Some of the web sites have received trade press mentions for industry "best practices." However, customer identities are not tracked across properties, and the "next best offers" made to customers are based on 1990s-style market-basket analysis. And no one would say Acme has forged an online community among its millions of customers.

Some in IT have heard of MR and have even formed a working group to look into the possibilities for Acme. However, whenever the idea is floated beyond this small group to others in IT, it is quickly shut down for a different reason each time. Sometimes it's the lack of full SQL capabilities in MR. Sometimes it's because there are no skill sets at Acme in MR to support it, and other times it comes from the fact that relational databases keep expanding their capabilities, and it is thought they will one day be able to manage the Web scale workload. There seems to be a high barrier to entry for Acme for MR. Consequently, interest in the group is not consistent. It has been difficult to keep interest, and MR has almost become a meaningless "free time" activity.

After Acme executives realized that Acme's future is based on harnessing the power of its information, changes began to take place rapidly. Knowledgeable experts in information store capabilities were given the ability to make changes. The data warehouses remain a solid foundation for Acme. They are actually in more use, not less, with the additions all around it. Some consolidation of the data warehouses and mart layer has led to more calculations in the data warehouses and fewer warehouses and marts. Data mart consolidations each generated substantial cost savings in the hundreds of thousands of dollars per year in software, hardware, and labor costs.

Looking at their raw click stream data, programs were implemented, like churn management, revenue attribution modeling, analyst productivity analysis and lifecycle marketing with Aster's SQL-MapReduce.

Analysis within the data warehouse was done on CDRs to generate fraud patterns. The fraud model was fed into data stream analysis, which began picking up fraudulent calls in real time based on the patterns. Subsequently, most of the CDRs started being loaded into the data warehouse continuously. So in fewer than three minutes, reports and dashboards offered the fraud information to special investigators who could follow up further.

Simultaneously, several existing, slow-running queries and workloads were looked at, and it was decided that a column store, like Aster Data, could process the customer counts and geographical pattern analysis that was being done on a daily basis. Consequently, much of the data now resides in both a row-based store and a column-store, but the overhead is worth it.

A more active approach was taken to dropping unused MDB databases. By turning off access to MDB databases in a scheduled approach over time, it was found that most were not used. Eventually, the unnecessary data loads stopped and the cubes removed. History maintenance was removed from the MDBs and ensured to be sufficient in the data warehouse. The data warehouse now feeds the MDBs, simplifying some of the complex infrastructure. Most importantly, IT took up a standard to combine MDBs and manage a smaller set of broader-based MDBs on a needs basis (i.e., not "cubes for data access by default").

With the realization of the need to steward data to remain competitive, a heterogeneous data management future for Acme is certain. With heterogeneity comes the increased need for integration and the acceptance of a diverse and active business intelligence environment. Acme can rely on Teradata, as a full stack information management company, offering an unparalleled focus on information management, to provide the technology necessary for making information its corporate asset.

About the Author

William functions as Strategist, Lead Enterprise Information Architect, and Program Manager for complex, high-volume full life-cycle implementations worldwide utilizing the disciplines of data warehousing, master data management, business intelligence, data quality and operational business intelligence. Many of his clients have gone public with their success stories. William is a Southwest Entrepreneur of the Year Finalist, a frequent best practices judge, has authored more than 150 articles and white papers and given over 150 international keynotes and public seminars. His team's implementations from both IT and consultant positions have won Best Practices awards. William is a former IT VP of a Fortune company, a former engineer of DB2 at IBM and holds an MBA.

www.mcknightcg.com