

Enabling the Agile Enterprise with Active Data Warehousing

Dan Graham,
Active Data Warehouse
Marketing Manager,
Teradata Corporation

Enabling the Agile Enterprise with Active Data Warehousing

Table of Contents

<i>Executive Summary</i>	2
<i>Teradata Active Enterprise Intelligence™ Strategy</i>	3
<i>Adding Insights to Operational Tasks</i>	4
<i>Decisions per Minute from the Active Data Warehouse</i>	5
<i>Transforming A Data Warehouse into an Active Data Warehouse</i>	6
<i>Pulling it All Together</i>	14
<i>The Teradata Difference – Focus</i>	15
<i>About the Author</i>	15

Executive Summary

In the 1990s, data warehouses were used for reporting and ad-hoc query analysis. But like everything in the information technology (IT) industry, innovation raced ahead, redefining the role of the data warehouse. Driven by Teradata customers seeking new revenue and cost savings advantages, active data warehouses started being built in 2001 on top of existing data warehouses, thereby extending the power of the data warehouse to a new community of operational employees, partners, and customers. The results have been profound, giving rise to *pervasive BI*, *operational BI*, and numerous other ways to describe the technology and benefits.

Teradata can help transform an existing data warehouse into an active data warehouse (ADW) by adding some or all of these active extensions: Active Access, Active Events, Active Load, Active Enterprise Integration, Active Workload Management, and Active Availability. These design patterns and best practices are not products, so there is no additional cost except the time spent to learn how to use them. They are insights gleaned from the early pioneering firms that have long enjoyed the competitive advantage of an ADW. This paper describes Teradata's Active Enterprise Intelligence™ strategy, its value, and next steps for those who want to move beyond reporting.

Enabling the Agile Enterprise with Active Data Warehousing

Teradata Active Enterprise Intelligence™ Strategy

Since the early 1980s, Teradata has accelerated corporate performance by helping clients convert information into insights and insights into action. It started when Teradata built the world's first parallel data warehouses and data marts in 1984 for strategic decision makers. As early as 1987, Teradata espoused a single version of the truth as the best way for a company to integrate and share data across business units. The single version of the truth represents an ongoing journey that helps companies unify, simplify, and improve corporate performance.

The primary users of data warehouses from Teradata were strategic planners, budget planners, forecasters, and researchers. Sifting through oceans of data, these strategic knowledge workers produced staggering returns on their companies' data warehouse investments. This foundation of knowledge is called *Strategic Intelligence*.

With strategic knowledge workers already empowered by the data warehouse, the next evolutionary step is to extend the power of the data warehouse to a new community of operational employees, partners, and customers. This is a profound change in the use of analytic

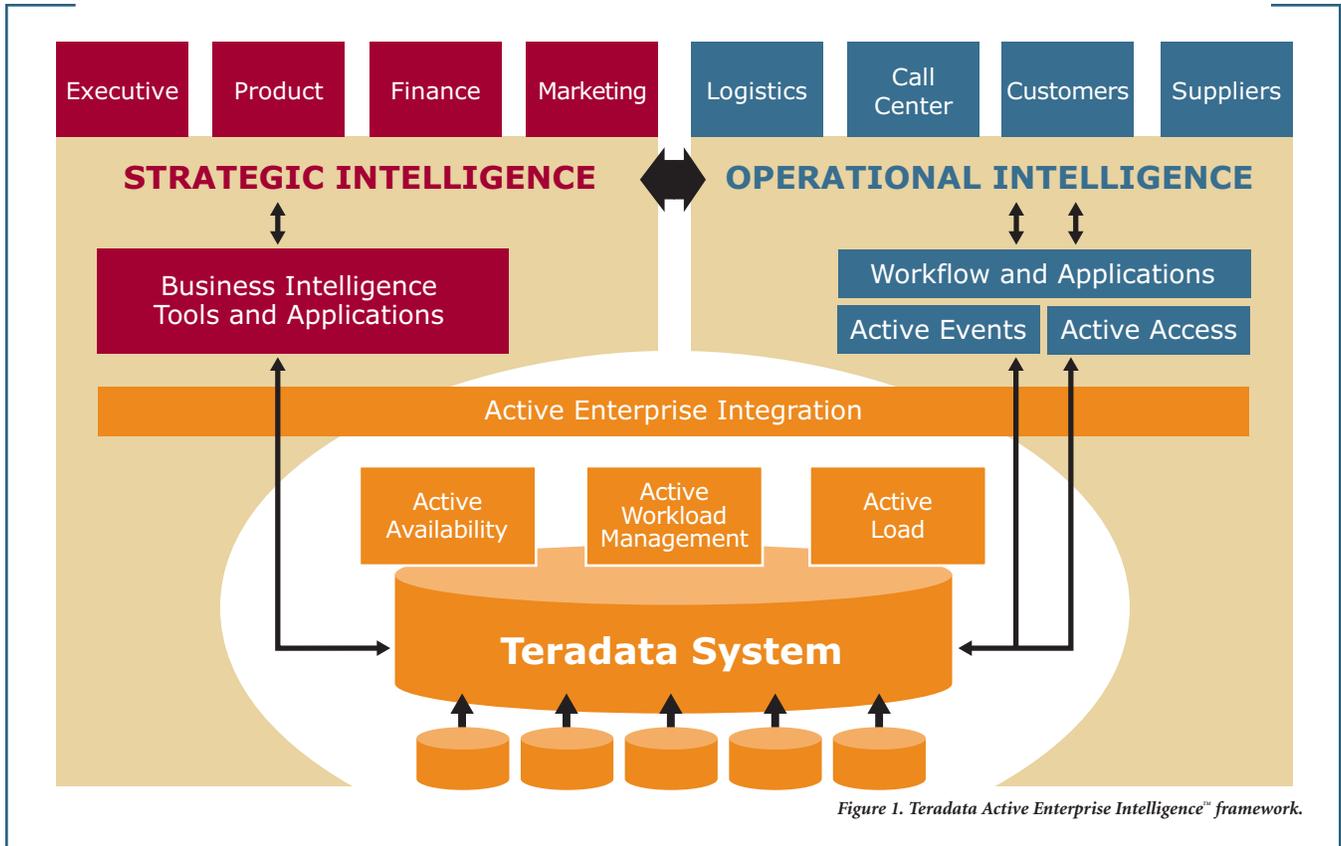
information. Corporate effectiveness can be stepped up considerably by adding insights to everyday tasks, thousands of times every day. From the loading dock, to the ATM, to the point of sale, to the sales agent, operational users need help making decisions intelligently and quickly. Today, Teradata data warehouses deliver insights every minute to front-line users to improve operational execution. Teradata makes this competitive advantage possible by extending an existing data warehouse with ADW technologies. When data warehouse insights are delivered to the front lines, it is called *Operational Intelligence*.

When strategic back-office users and operational front-line users share the same ADW, the enterprise has achieved an Active Enterprise Intelligence™ environment. With this single view of the business, companies can align back-office strategy with front-line operational execution through consistent metrics, vocabulary, priorities, and insights. Through shared metrics and insights, companies ensure employees and business processes coordinate to achieve customer profitability, higher product margins, lower costs, and many other business priorities. This alignment and acceleration is the goal of today's real-time enterprise.

“We’re using the data warehouse to turn ideas into actions that are measurable and that have measurable ROI. While real-time operation is an essential component of any data warehouse, currently, the term active warehousing puts too much focus on the real-time movement of data. Instead, the focus should be on the real-time utilization of data, regardless of where it exists. In general, successful Internet businesses will ultimately have a seamless integration of intelligence/analytics in their operational business processes. This is where the focus should be.”

– Tracy Abdo, Vice President of Enterprise Data Services, Network Solutions

Enabling the Agile Enterprise with Active Data Warehousing



Adding Insights to Operational Tasks

An ADW provides many uses for Teradata Active Enterprise Intelligence™ capabilities (See Figure 2) and still supports the traditional data analysis needs of back-office planners and managers. The change comes by extending the ADW to front-line users, making hundreds of new applications possible by reusing the same data in new ways. For example:

- > Shipping and receiving workers can monitor daily fill rates and delivery cycle times, changing carrier methods for fast-moving products or high-priority customers.
- > Sales agents, bank tellers, websites, and call center representatives can use every customer-facing event to solve customer problems, improve satisfaction, and propose relevant offers with a higher probability of acceptance.
- > Manufacturing yield drops caused by revisions or build-to-order variances can be tracked hour by hour and compared to historical norms to detect production problems before they cause excess inventory or delay committed orders.
- > Banks can deliver *one voice* to the consumer, providing a consistent treatment and next-best-offer across all communications channels, such as ATMs, tellers, the Web, and call center.
- > Transportation companies can reroute passengers or containers when weather or vehicle failures disable a shipping route, balancing passenger priority and margin optimization.

Enabling the Agile Enterprise with Active Data Warehousing

- > Insurers can automate price quotes and provide self-service claims management to consumers, reducing call center costs and increasing customer satisfaction.
- > eCommerce web sites can personalize product offers to increase take-rates and, at the same time, dynamically adjust offers to match inventories.
- > Fraudulent warranty and in-store product returns can be detected and refused as they are happening – before money is lost.

The list is endless since there are so many opportunities to add insights to front-line operational processes. This is why Gartner calls these capabilities *pervasive BI* – the analytic insights can be injected into thousands of business processes and workflows to optimize business execution.

Decisions per Minute from the Active Data Warehouse

Operational business processes have been waiting a long time for daily tasks to be enhanced with decision-support services. Consider how many operational decisions are made every minute in your enterprise that are not based on facts or corporate performance goals. There are literally thousands of them. But technology limitations prevented delivery of insights to front-line operational users from the data warehouse. Front-line users need up-to-the-minute facts, yet most data warehouses are refreshed only once per day. Front-line users need insights

Examples of Active Data Warehouse Uses	
Banking and finance	Multi-channel cross sell, automated loan authorizations, risk tracking, trading monitoring, funds transfer reconciliation, equities pricing
Fraud detection	POS shrinkage, warranty returns, insurance claims
Government	Critical supplies and labor monitoring, complaints tracking, permitting process, supplier collaboration
Healthcare	Claims fraud, billing disputes, critical supplies and labor optimization, outbreak detection
Insurance	Call center productivity and cross sell, IVR routing, claims history portal for hospitals and pharmacies, self-service consumer claim portals
Manufacturing	Defect-monitoring alerts, product consumption visibility, demand-driven replenishment, supplier collaboration, continuous demand chain optimization, service-labor optimization
Retail and eCommerce	Out-of-stock alerts, personalized Web offers, service delivery scheduling, monitor sales surges, in-store personal offers, billing disputes, custom couponing
Telecom	Provisioning scheduling, billing disputes, wireless toll fraud, complaints tracking, call center productivity, IVR routing
Transportation	Track and trace partner/consumer portals, passenger check-in optimization, container packing and routing, freight refusals/surges, hazardous goods monitoring, fuel purchase optimization

Figure 2.

delivered in seconds through familiar, easy-to-use applications. In contrast, traditional data warehouses are accessed through business intelligence tools that produce mostly reports. Perhaps most important, these complex reports often consume substantial data warehouse computing resources, causing the one-second operational inquiries to become ten-minute inquiries. For an operational employee facing customers, opportunities are lost if customers must wait minutes instead of a

few seconds. Beginning in 2001, Teradata systematically removed technology limitations to allow both strategic and operational intelligence requests to be served from a single data warehouse. Surprisingly, an ADW also allows CIOs to actually do more with less, the classic business imperative. An ADW based on Teradata technology eliminates the need to duplicate portions of the data warehouse in operational data stores, reducing costs and complexity at the same time.

Enabling the Agile Enterprise with Active Data Warehousing

Let's compare and contrast operational intelligence requirements with the older strategic intelligence requirements. Figure 3 shows the contrast between traditional strategic data warehouses and operational ADWs.

With the Teradata Active Enterprise Intelligence™ strategy, Teradata is introducing a new concept called decisions per minute. Decisions per minute focuses on delivering fast, analytical information to front-line users to support thousands of decisions every day. To support the decisions-per-minute workload, an ADW manages the mixed workload of large, intensive strategic queries side-by-side with short, tactical operational queries. In contrast, OLTP systems are aimed at update- and inquiry-intense tasks, measured in transactions per second. But OLTP systems only make the business accounting run faster. Teradata's decisions per minute helps make the business process smarter.

Transforming A Data Warehouse into an Active Data Warehouse

An ADW from Teradata is built upon a solid enterprise data warehouse foundation of strategic intelligence. Building upon this, there are six active elements to be considered when creating an ADW:

- > **Active access** – high-speed inquiries, analysis, or alerts retrieved from the ADW and delivered to operational users, devices, or systems.
- > **Active events** – operational events that need to be continuously monitored, filtered, and alerts sent based on business rules.
- > **Active load** – high-frequency data loading throughout the business day to ensure data are fresh enough to support active access and active events.
- > **Active enterprise integration** – links the ADW to existing applications, portals,

Web services, service-oriented architectures, and the enterprise service bus.

- > **Active workload management** – dynamic management of operational and strategic workloads in the same database, ensuring response times and maximum throughput.
- > **Active availability** – increasing the data warehouse availability from business critical to mission critical.

The beauty of this approach is you don't have to adopt all these capabilities at once. They can be implemented in phases as the need for more data, more users, and faster access grows. Perhaps most surprising is the implementation of the six active elements is primarily a labor cost for the IT organization. Only a few of the more than 100 existing ADW sites have purchased new hardware or software to transform their data warehouse to an ADW.

The Foundation – Teradata's High-Performance Data Warehouse

The Teradata® Database offers exceptional performance and low support costs, features that are mutually exclusive with other vendors. Exceptional performance comes from more than 30 years of continuous investment in the Teradata cost-based optimizer – ten years longer than our nearest competitor. Teradata's optimizer performance improvements come from hard lessons in solving real customer performance problems for large and small companies. At the same time, Teradata developers designed in ease of use, a

	Strategic Intelligence	Operational Intelligence
Users	Planners, analysts, managers, Finance, Marketing	Consumers, call centers, sales agents, suppliers, shipping dock workers, devices, and systems
Concurrent users	10s to 100s	10s to 10,000s
Response time	Seconds to hours	1 to 5 seconds
Update frequency	Nightly batch updates	24 hours per day
Access	BI workbench tools, Excel	Dashboards, portals, applications, alerts
Availability	Business Critical – low tolerance for downtime	Mission Critical – disaster recovery plan and no tolerance for downtime

Figure 3. Traditional strategic data warehouse capabilities vs. operational ADWs.

Enabling the Agile Enterprise with Active Data Warehousing

necessity in massively parallel systems. Thus, with Teradata, the DBA staff won't have to perform tedious file management, data reorganizations, or struggle with complex partitioning schemes. Teradata Database does it automatically. All of this enables highly normalized designs that can respond to business changes faster and with flexibility. Given executive demands, mergers, and the pace of competition, a powerful optimizer and a mostly normalized design is the only sensible CIO survival plan for the long term. Consequently, industry analysts consistently score Teradata Database as the top-rated data warehouse solution for query performance, concurrent queries, administration, data management, and platform suitability.

Strong evidence of Teradata's high-performance data warehouse quality is, of course, successful customers. Teradata customers can be seen at the annual Teradata PARTNERS user conference. With more than 3,000 attendees each year, no other data warehousing conference comes close to the size and scope of Teradata PARTNERS User Group Conference. Loyal customers exuberant about the business benefits they achieve is the best evidence of a vendor's quality and value.

Activating the foundation of your data warehouse from Teradata requires adding one or more of these elements: active access, active events, or active loads. Usually, the presence of one of these necessitates one other, so they're often deployed two at a time. Once the data warehouse is active, the three remaining

“We have been able to use our application design in combination with the Teradata [solution] to create an active data warehousing environment that conventional database platforms would not support very well. Teradata has given us the ability to utilize actual transactional data in powerful queries without requiring data extraction into a separate data warehouse, allowing us to bypass issues of data versioning and conversion. Combining Teradata services, active data warehousing, and scalable software and hardware gives us the tools we need to keep PING on top of the leader board.”

– Kent Crossland, Information Services Director, PING, Inc.

infrastructure elements need to be increased to varying degrees. These are active enterprise integration, active workload management, and active availability.

Active Access – Fast Insights for Front-line Users

Active access occurs when a front-line user or operational application retrieves information from the ADW. The simplest active access uses fast, tactical queries to retrieve historical, summarized, or fresh data. Typically, results are delivered at Web speed, often just one or two seconds. Occasionally, when the business process needs extensive data analysis, an active access can take a few minutes. But usually, when a customer is talking to a sales agent or call center representative, there are only a few seconds to decide on the right offer or problem

resolution. Similarly, a worker in shipping and receiving may only have one or two minutes to select urgent priority packages for loading before the truck leaves. And, of course, the best time to stop fraud is when it happens, not later when a report shows money has been lost. Fast access speed is not optional. Intelligence is not optional.

Fast, tactical queries depend on a number of features built into Teradata Database, including join indexes, parameterized queries, and macros. Teradata Database's join index helps front-line users retrieve frequently used data without joining tables in real time. They simply grab the precomputed answer quickly. Parameterized queries allow the Teradata optimizer to cache SQL it has seen before and reuse

Enabling the Agile Enterprise with Active Data Warehousing

the execution plan the next time it sees the same SQL. Skipping the parsing and optimization steps can save half a second or more in many cases, a crucial amount of time when trying to deliver results in a single second. Macros allow the programmer to make only one hop across the network and back, while executing multiple SQL statements. All of these combine to produce fast response times for tactical queries, ensuring that front-line users can rapidly make informed decisions and take action, instead of waiting for the system. Typically, a few million active accesses per day consume one percent of system capacity, so it's unusual to need additional hardware capacity.

Active Events – Event-Driven Operations

Real-time enterprises are event driven, meaning they're capable of responding to opportunities and problems as they happen. Business events can be as simple as receiving a purchase order or as complex as an airplane flight cancellation. Business events can be detected internally within the ADW or externally in production applications. Once detected, events are filtered to select exceptions and anomalies for immediate processing. Event filtering and processing can occur inside an application, inside the ADW, or a combination of both. An active event only occurs when event processing begins inside the ADW or when the ADW is used to augment an event alert with context and insights. The typical output of event processing is an alert to a front-line user or an update message to an operational system.

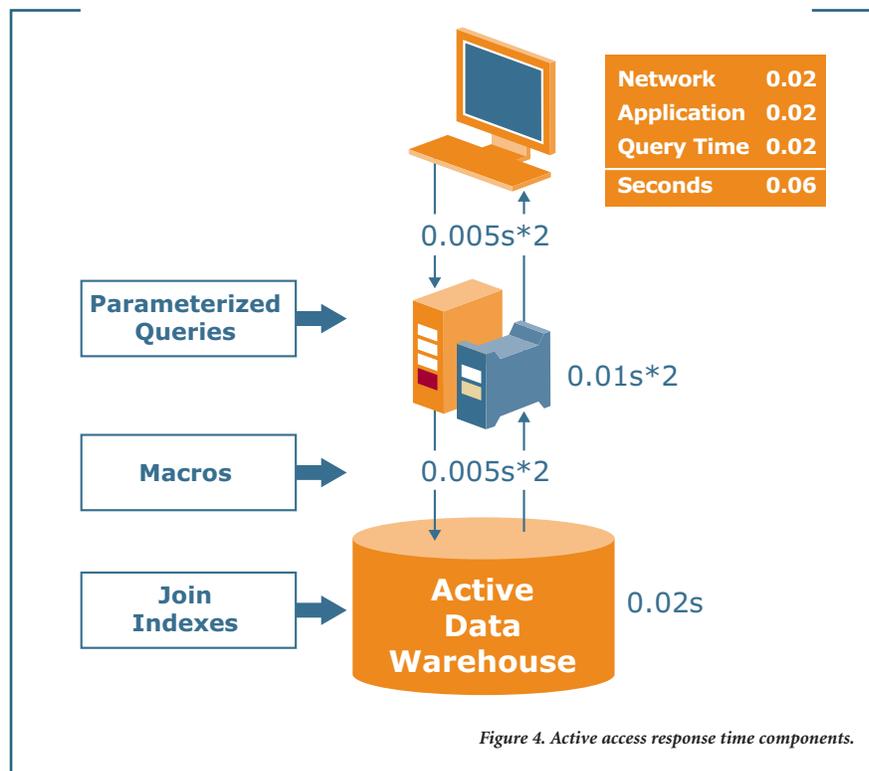


Figure 4. Active access response time components.

Teradata Database supports active events with parallel database triggers, internal and external stored procedures, and persistent queue tables. For example, consider a retail or manufacturing company concerned with out-of-stock situations. Not having goods to sell or to feed the manufacturing line is disruptive and often leads to a financial loss. Using tools called Business Activity Monitoring or Complex Event Processing, inventory managers can get alerts sent to their dashboard whenever an anomaly occurs, indicating an important commodity has fallen below prescribed thresholds. The inventory manager then drills down into the ADW for facts about the situation, using analytic insights to decide whether

or not to take action and what actions to take. At the same time, continuous data loading into the ADW inserts incoming facts from suppliers and the shipping/receiving dock. Occasionally, database triggers fire as data are loaded when they detect a threshold has been breached causing a record to be inserted into a queue table. Periodically, let's say every 15 minutes, a dashboard function checks the queue table for significant events, filters out those that aren't important, and finds one serious shipper delay that will affect profitability. Alerts are sent, and downstream applications are invoked to help the inventory manager respond to the unusual out-of-stock condition. Again, the

Enabling the Agile Enterprise with Active Data Warehousing

actions taken in response need analytic and historical information drill down in order to decide the best course of action.

All this is possible with Teradata triggers, stored procedures, and queue tables that are easy to program and link to production applications. Teradata enhances Complex Event Processing by determining what is relevant and what to do about it. Merely responding quickly to business events is not enough. Without historical

insights and contextual analysis, front-line users cannot consistently make informed decisions that improve corporate performance. This is where an ADW from Teradata excels – analyzing the situational context (what is happening over time) and supplying the best business decision recommendation (insights and optimum choices) for each active event. Active events are like active access, consuming trivial amounts of system capacity even for large numbers of alerts per day.

Active Load – A Continuous Stream of Fresh Data

Competition and customer service drive the compression of data latencies on the operational side of most businesses. For agile enterprises, this means combining what is happening (the event) with what has happened (historical context) to support smarter front-line decisions. For example, responding to the arrival of damaged critical supplies, providing in-store personal offers, or rerouting

Harrah's Entertainment

"What we've been able to do is capture all of the transactional data from each individual slot machine – down to the individual coin-in, the denomination, who plays it, when they play it – and merge that against the customer. Slot visualization then allows us to run an application over the top of the Teradata platform, which let's you graphically see, in either real time or after the fact, what is occurring at each individual slot machine, bank of slot machines, or on the casino floor, in a very graphical, compelling format. So, to be very specific, what you can do is drill down on an individual slot machine, see what type of customer has played it, perhaps their demographic, their age, the type of denomination they played, and can track a bank of slot machines compared to a next bank of slot machines and see why some customers play one versus the other. And ultimately what we're able to do is make very, very scientific decisions about what kinds of products we would put where to best service our best types of customers.

"So what happens at every individual transaction on a slot machine, it's going through a series of communications at the slot level up to what's called a slot data system, SDS. That level of transaction information is

real time, can be fed right into the data warehouse, and actually observed real time. You can see why we're so intrigued by the combination of the TIBCO platform and the Teradata environment. To be able to actually trigger and create events on those types of activities so that humans don't actually have to watch something but it can be very automated is very compelling."

– Tim Stanley, Vice President of Information Technology, Harrah's Entertainment

"Another big project for us now is VIP business. So this is a small group of customers that contribute a significant part of the revenue. These folks have individual casino hosts that help manage the relationship for them. So the active data warehouse will give these hosts real-time information about these customers. When they're on the floor, are they in the hotel, what level are they playing at? So they can understand exactly how to manage that relationship, both when a customer is on the floor, as well as when they're preparing for their next trip."

– David Norton, Vice President of Marketing, Harrah's Entertainment

Enabling the Agile Enterprise with Active Data Warehousing

passengers on delayed flights all require current data (the event) plus historical data (context) for decision making. This drives the need for near real-time data loading.

For most enterprises, data loads are done each night because most data load utilities lock up the database tables until finished. These locks can freeze incoming user requests, sometimes for hours, waiting for the load to finish. Active load on a Teradata system bypasses these limitations with a collection of products and services that continuously loads data without locking up front-line users. While some enterprises need true real-time data loading, Teradata experts prescribe “fresh enough” or “right-time data loading” to control costs while meeting business objectives.

To apply fresh operational data to the data warehouse continuously, Teradata provides several approaches for fresh enough data loading. Foremost is Teradata Parallel Transporter, a utility that continuously moves streams of data into the ADW with minimal locking. Not only can queries access a table that is also being updated by Teradata Parallel Transporter, but several load jobs can run against the same table at the same time. This gives the DBA numerous options and flexibility to manage workloads. For example, a busy Web site sends account update messages to queuing middleware, such as WebSphere MQ or Java Message Service. Next, a Teradata Parallel Transporter job reads the message queue and directly updates the ADW. Within seconds, a customer event is stored in the database and is available for both

“Eighty percent of our workload is real time or near real time. We load data from the operational system into the data warehouse in near real time; it’s never more than a minute or two behind. Marketing wants real-time information, and we give it to them. The business side wants information every hour or every other hour; finance pulls some data on a daily basis. We’re providing data for lots of different parts of the company. The important thing is that wherever a number is used, whoever is looking at it, whatever the view, it’s the same number.”

– Jack Garzella, Vice President of Data Warehouse Reporting and Analysis, Overstock.com

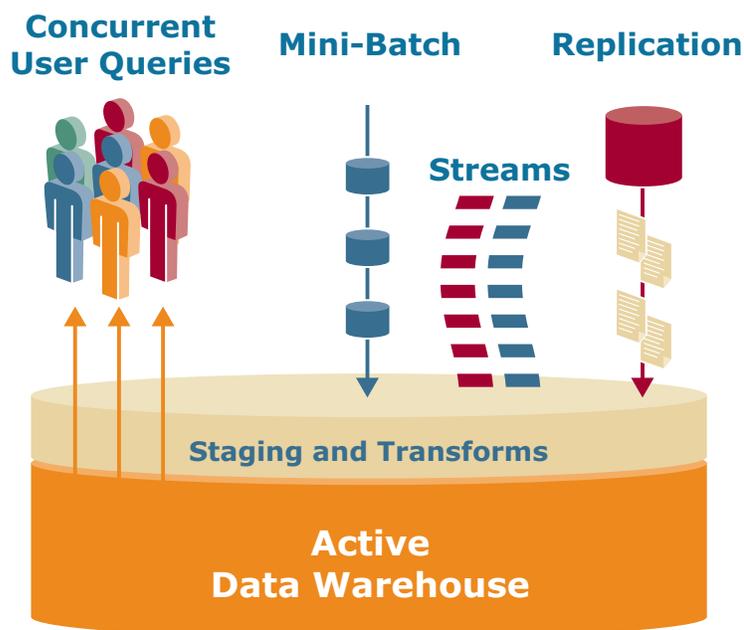


Figure 5. Active load.

Enabling the Agile Enterprise with Active Data Warehousing

operational and strategic decision making. Teradata was the first RDBMS vendor to provide easy-to-use database load utilities to continuously load streaming data. Some competitors still don't offer this capability.

Teradata Parallel Transporter also supports loading of small mini-batches at controlled intervals to load large volumes of data quickly while minimizing database locks. Mini-batch is the preferred starting point for organizations since it is relatively easy to implement and highly efficient in the use of the Teradata system hardware. A third option is Teradata Replication Services that performs change-data-capture from operational databases, such as Oracle or Microsoft SQL Server, and propagates the update to the ADW. The benefit of active load solutions is that data are fresh enough for front-line user tasks, while keeping server and IT costs as low as possible. Typically, these active load techniques consume from two to eight percent of a Teradata system, depending on the amount of data arriving and the technique used. Mini-batch tends towards the lowest system use and real-time streams tend towards higher use.

Active Enterprise Integration – Fitting into Existing IT Infrastructure

Active enterprise integration requires connecting the data warehouse with contemporary applications, workflows, Web services, and existing IT architectures via open standards. Typically, ADW deployments use browser-based portals, Web applications, Business Process Management

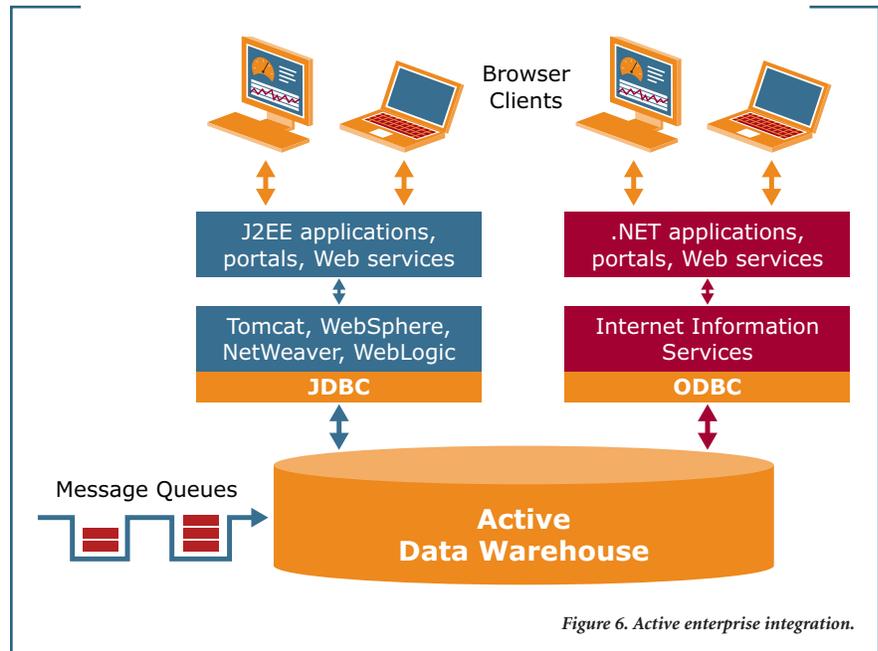


Figure 6. Active enterprise integration.

(BPM) workflows, and service-oriented architecture (SOA) components. To achieve this integration, Teradata Labs invests heavily in two major areas: fitting into the existing environment and components for accelerating deployments.

Teradata product integrations exploit open, industry standards, such as XML, SOAP, JMS, J2EE, .NET, and others.

Teradata products can also be used with Web services open standards, such as BPEL, UDDI, WSDL, and WS-Security.

With this foundation, Teradata products are integrated and tested with popular business partner products, including:

- > **Application servers:** Oracle WebLogic, SAP NetWeaver® Web Application Server, IBM WebSphere®, JBoss®, Apache Tomcat, and Microsoft® IIS.

- > **Enterprise Service Bus:** TIBCO BusinessWorks, Oracle Aqualogic™, IBM WebSphere ESB and WebSphereMQ, webMethods Fabric™, and any Java Messaging Service (JMS)-compliant message queuing software.

Connectivity to the various JMS and ESB products is achieved using the Teradata JMS Access Module or Teradata WebSphereMQ Access Module. These free software modules read and write to the message queue using the industry-standard API. Typically, these modules are used to fetch messages from the queues and pass them to Teradata Parallel Transporter for near real-time data loading. Dozens of Teradata customers load fresh data this way from point-of-sale devices, manufacturing production lines, airline

Enabling the Agile Enterprise with Active Data Warehousing

reservation systems, and more. Using the same interfaces, a programmer can also write the output from BTEQ or Teradata Parallel Transporter export to a JMS compatible queue.

Connectivity to Web application servers is primarily done via the open standard JDBC Type 4 and ODBC drivers. These interfaces manage logins, connection pooling, and the invocation of object-oriented programs. SQL is commonly either embedded directly in the object-oriented program or isolated in a data access object. Data access objects are a best practice that simplifies programming and isolates function. For Java implementations, Teradata supports Hibernate and iBATIS, popular open source that performs object-relational mapping. For .NET implementations, the object-relational mapping is built into the ADO.NET Entity Framework. Using these interfaces, a Java or .NET programmer can build Web services modules, servlets, Java beans, or plain old java objects.

Teradata Labs has also simplified the work of building object-oriented applications with the Teradata Eclipse plug-in. Eclipse is a highly productive, integrated development environment used by many Java programmers. With the free Teradata Eclipse plug-in, the programmer can write a Java program and browse the Teradata Database objects using drag and drop mouse clicks. This simplifies the coding of SQL with access to schemas, tables, views, stored procedures, macros, and even support for building data access objects.

“We have gone from a very small warehouse to a point where Teradata is tackling almost any and every aspect of the business; most of it directly and some of it indirectly. For our retail business, it is our main information source for supply chain. In addition to that, it is our single support system for the business intelligence side of things. We are in the process of actually rolling out about five or six high-end online tactical applications in the next year. So that is kind of our metamorphosis towards an active warehouse environment. That’s where we are with Teradata at this time. It’s a mission-critical system.”

– Dan Raju, Manager, DBA, Charming Shoppes, Inc.

Active Workload Management – Controlling Mixed Workloads

The crux of the ADW is that it enables front-line users and back-office users to share the same data warehouse, providing them with a single view of the business. But in the past, this caused response time for the front-line users to be painfully slow. The now obsolete method for handling this was to have separate databases, called Operational Data Stores (ODSs), for each operational group, thus duplicating costs and multiplying complexity. So, a single ADW is enormously preferable. Unfortunately, concurrent workload management has been a weak spot in relational databases for decades. Most IT operations groups have struggled

with procedures and tools to control the *query from hell* that steals the entire server away from other tasks running at the same time. Even a few complex queries can stall short tactical queries, turning them into multi-minute frustrations.

The common response of database vendors has been to offer *query police* software that queues queries according to their computer resource needs and releases them when resources are available. Unfortunately, the in-flight tasks often fool the query police into releasing queries prematurely. This happens often when in-flight tasks are in a resource usage lull just before starting a large resource-consuming step.

Enabling the Agile Enterprise with Active Data Warehousing

The query police are actually responsible for producing classic Los Angeles-style traffic jams because they're helpless to control traffic after tasks are in flight.

Teradata Active System Management tools solve the root cause of this problem. This starts with Teradata Workload Analyzer, which analyzes Teradata Database user logs and system tables to profile actual usage behavior. The easy-to-use reports and graphs describe what is really happening in the system over time. Teradata Workload Analyzer then recommends workload groups and parameters for use with Teradata Dynamic Workload Manager. The parameters allow the DBA to establish and refine the workgroup categories, throttles, and control settings. Queries and data loads can be prioritized by time (hour, day, week, or month), user groups, and type of workload. At run time, Dynamic Workload Manager analyzes each database query before and during execution based on business rules set by the DBA. Like the query police products, Dynamic Workload Manager releases queries when it knows resources are available. But unlike the query police, Dynamic Workload Manager constantly monitors the active tasks in the system, using DBA-specified rules, limits, and workload groups to allocate computer resources to in-flight tasks. By favoring the tactical queries over database loading, for example, the DBA can favor operational user requests, yet still keep the data loads running. Using cost thresholds, Dynamic Workload Manager adjusts the priority of executing tasks *continuously*. This is what

makes Dynamic Workload Manager different from simplistic query police tools. If the users go on lunch break, or there is a lull for any reason, Dynamic Workload Manager reallocates resources to get the maximum amount of work done. Runaway queries can be downgraded in priority automatically before front-line user performance suffers. Fast queries stay fast, and big workloads slow down, but they don't starve. Best of all, Dynamic Workload Manager performs with very little system monitoring effort. Since most traffic jams are avoided altogether, only serious exceptions require DBA attention.

To complete the process, Teradata Viewpoint provides an easy-to-use graphical dashboard and trend analysis of real-time and historical workload performance. It provides ad-hoc and standard reports on workload trends, such as high use tables, high resource consuming user groups, and workload history. Teradata Viewpoint allows the DBA to easily track workload performance against service level goals to plan ahead and make changes proactively.

The most obvious benefit of active workload management is that system administrators can control response times for operational and strategic users. Another benefit is the deferral of server upgrades to improve performance until absolutely necessary. With Teradata Dynamic Workload Manager, consistent response times are ensured by controlling priorities and categories, not by throwing hardware at the problem. And Teradata Dynamic Workload Manager contributes

to the consolidation of data marts and ODSs into a single data warehouse because mixed workloads can share a single database. This reduces complexity, as well as server and license costs. Finally, active workload management is what enables operational intelligence and strategic intelligence work to share a single view of the business in the ADW.

Active Availability – Mission-Critical Service Levels

ADWs typically support from hundreds to thousands of front-line users who need the same high availability as transaction systems. Consequently, existing data warehouses must shift from being business critical to mission critical. To achieve this, *all* Teradata systems provide a tightly interlocked, highly available foundation. Redundant components throughout the server platform are included in the base configuration and are fully utilized during normal operation to maximize system performance. Teradata systems mask failures with dual power supplies, uninterruptible power supplies, dual I/O controllers, heavy duty cycle disk drives, RAID controllers, and dual interconnections to the Teradata BYNET®. Best of all, Teradata software automatically manages component failover, reducing the risk of human error during the critical recovery process. Furthermore, redundant components can be replaced while the system is running, minimizing a repair's impact on system availability.

Additionally, Teradata's exclusive clique architecture minimizes the impact of

Enabling the Agile Enterprise with Active Data Warehousing

complex failure scenarios and provides rapid recovery from a wide class of failures. If any server node in a clique fails, Teradata Database redistributes the workload across the surviving nodes in the clique by migrating software processes to other nodes. However, moving workloads can degrade front-line user response times. When a four-node clique experiences a single node failure, there is only 75 percent capacity to serve 100 percent of the work. Response times may degrade as much as 33 percent. When performance is critical, Teradata customers add a hot standby node per clique to maintain consistent query performance. Large, tightly interlocked cliques plus innovative recovery software features handle the failover smoothly, minimizing human intervention and recovery complexity.

When Teradata customers need the highest levels of mission-critical performance, they choose the Teradata Dual Active solution. The dual active system provides business continuity via a remotely placed secondary server. But the secondary server is not an idle insurance policy against disaster. Teradata Query Director continuously directs work requests to both of the Teradata systems based on DBA-established rules, providing maximum use of hardware investments. If a major planned or unplanned outage occurs, dual active monitoring software automatically makes the necessary changes so current work continues, minimizing the risk of human error at a critical time. Best of all, the secondary servers need not be duplicated in size of the primary servers. They can be configured to serve users of only the most

mission-critical data while the primary server is offline. Typically, because less than half of the data is characterized as life-support critical to the enterprise, the potential cost savings over a traditional fully redundant system is enormous.

Teradata Business Continuity Solutions experts can deliver an availability assessment service that maps out the lowest cost configuration plan for meeting enterprise service level goals. This could range from large cliques, hot standby nodes, and fallback, to Teradata Recovery Centers and the Teradata Dual Active Solution. Teradata Professional Services consultants can help plan backup and recovery, disaster recovery, availability procedures, and development systems.

Pulling it All Together

For enterprises with a Teradata data warehouse in place, the good news is that an ADW is within reach. The next step is to identify an operational process, preferably a small project, which can benefit from tactical insights. It's best to start small and grow so the IT organization can strengthen its mixed workload skills before tackling a large project. Working with line-of-business users, start by mapping the end-to-end business process showing the before and after processes. Always clearly identify the business value of the project and set measurable goals for the project results. Teradata industry consultants can help business managers in this step. Then, and only then, should technical designs be started. With a clear project goal, it's easier

Foundation	Teradata system	<ul style="list-style-type: none"> > Teradata active data warehouse systems > Teradata Warehouse 8.1 or higher
Activating the data warehouse	Active events	> Teradata stored procedures, triggers, queue tables, external table functions; BAM/BPM tools, rules engines
	Active access	> Aggregate Join Index, macros; BAM/BPM dashboards, rules engines, federated query tools
	Active load	> Teradata Parallel Transporter 8.1 or higher; ETL/ELT tools
Critical infrastructure	Active enterprise integration	> Teradata Eclipse plug-in; Enterprise Service Bus, Web Application Servers, Message Oriented Middleware, EAI
	Active workload management	<ul style="list-style-type: none"> > Teradata Active System Management 8.1 or higher > Teradata Dynamic Query Manager 8.1 or higher
	Active availability	> Cliques, hot standby nodes, fallback, Backup-Archive-Recovery, Teradata Dual Active Solution

Figure 7. Teradata active data warehouse technology.

Enabling the Agile Enterprise with Active Data Warehousing

to determine which active elements are most important to the new application. Figure 7 is a high-level summary of technologies that are frequently deployed with an ADW. The exact combination depends on the goals.

The Teradata Difference – Focus

An active data warehouse from Teradata is a tightly interwoven, purpose-built system of products and experts, not a loose federation (See Figure 8). Every component has been developed to enhance and reinforce every other. Teradata delivers products that

are ready to work together – not a do-it-yourself bag of parts. A Teradata solution is harmonious, repeatable, and predictable. All components run in parallel, support high availability, and work together to make administration easier. The servers, storage, and software all behave as one solution, having been repeatedly tested and challenged by discriminating Teradata customers. Teradata Professional Services consultants are the top specialists in the world for providing end-to-end data warehouse expertise. They practice continuously on one solution set, honing their data warehouse skills as a cohesive team. Teradata salespeople are trained and *certified* to

identify the business value of an ADW. Why? Because your success defines our success.

Teradata has a single, passionate focus – not 10 different agendas. Teradata specialists are focused on driving business value – profit and cost savings – out of every data warehouse they deliver. Our passion has always been to exploit integrated data, converting information into insights and insights into actions. The active data warehouse is what truly provides the single view of your business, aligning operational execution with strategic goals.

About the Author

Daniel T. Graham

In the 1970s and 1980s, Dan was an application developer and database administrator at Univac and California Federal Savings and Loan. He joined Teradata Corporation in 1989, where he was a senior product manager for the Teradata DBC/1012 parallel database servers and software. Dan joined IBM in the 1990s where he was responsible for product planning on the RS/6000 SP parallel server. He later became Strategy and Operations manager for IBM's Global Business Intelligence Solutions division. During this time, he managed the IBM Teraplex Centers, competitive teams, and analyst relations. Dan returned to Teradata and is currently a Program Director for Active Data Warehouse Marketing.



Figure 8. Teradata provides tightly coupled solutions that offer seamless integration of all of these strengths.

Raising Intelligence is a trademark and BYNET, Teradata, and the Teradata logo are trademarks of Teradata Corporation and/or its affiliates in the U.S. and worldwide. TIBCO BusinessWorks is a trademark and TIBCO is a registered trademark of TIBCO Software, Inc., in the United States and/or other countries. IBM is a registered trademark and WebSphere is a trademark of IBM in the United States. WebLogic is a trademark of Oracle Corporation. SAP NetWeaver is a registered trademark of SAP AG in Germany and in several other countries. JBoss is a registered trademark of Red Hat, Inc. Microsoft and BizTalk are registered trademarks of Microsoft Corporation. Teradata continually improves products as new technologies and components become available. Teradata, therefore, reserves the right to change specifications without prior notice. All features, functions, and operations described herein may not be marketed in all parts of the world. Consult your Teradata representative or Teradata.com for more information.

Copyright © 2007-2010 by Teradata Corporation All Rights Reserved. Produced in U.S.A.