DATA WAREHOUSING
10.12 | EB 6296

TERADATA ANALYTICAL ECOSYSTEM

By Imad Birouty, Program Marketing Manager, Teradata Corporation

TERADATA

Not Your Father's Data Warehouse

TABLE OF CONTENTS

2 Executive Summary
3 Teradata Ecosystem: An Overview
5 Building on Pillars of Innovation
7 Design, Implementation, and Maintenance Requirements
7 Products, Features, and Services for the Analytical Ecosystem
10 Multi-system Enabling
14 Services
15 Use Cases
17 Why Use the Teradata Analytical Ecosystem?
17 About the Author

EXECUTIVE SUMMARY

This white paper focuses on the Teradata® Analytical Ecosystem, a new strategic way of looking at your data warehouse environment.

Most Teradata customers already create their data warehouse environment with multiple systems that serve specific purposes, including production, test, development, quality assurance (QA), and disaster recovery (DR).This multiplicity has traditionally been driven out of direct need for these functions; it has not been viewed as a strategic approach to data warehousing. Further, each of these systems has generally been built on Teradata's Active Enterprise Data Warehouse (EDW) platform, our flagship offering. With the introduction of the Teradata Workload-Specific Platform Family, this approach can be taken to a new level.

The Teradata Analytical Ecosystem can be thought of as the *total* data warehouse environment: everything needed to run enterprise analytics. The Teradata Analytical Ecosystem takes in all the concepts mentioned above, but expands their scope. It continues to include the Teradata systems, but with the Teradata Workload-Specific Platform Family, customers now have a choice of platform types to use in their data warehouse environment; with each platform having unique strengths, price points, and performance characteristics. Additionally, the Teradata Analytical Ecosystem includes the interaction between the systems, such as the processes that are run and the data synchronization between systems at any point in time. It also includes supporting systems, such as extract, transform, load (ETL) servers used to bring data into the data warehouse, as well as business intelligence (BI) servers, and any application servers used to deliver results to the users. In other words, the Teradata Analytical Ecosystem is a concept that is unique to Teradata in looking at the entirety of the data warehouse environment, with all the component parts, processes, and data necessary to meet the analytical requirements of the business.

In this white paper, we'll first look at the Analytical Ecosystem from an architecture perspective and look back at the business requirements that have driven the need for it. Then we will look at the strategies customers use today, and the use cases that customers typically face. Next we will look at the products and services required to deliver and maintain such an environment. We will end with sample implementations as the Analytical Ecosystem is not just one specific implementation, but a broad range of possible architectures that will meet customers' business and technical requirements.

TERADATA ECOSYSTEM: AN OVERVIEW

Below is a pictorial view of the Teradata Analytical Ecosystem. It provides the foundation to host and analyze all the business's data. Most customers won't have everything that is portrayed in this graphic, but may have subsets as appropriate for their business.

Data starts from the source or transactional systems, such as sales, inventory, and customer information. The data is then loaded into the primary production integrated data warehouse (IDW) using commercial ETL along with Teradata bundled utilities. Backup, Archive, and Restore (BAR) is used for backup/archival purposes as well as for restoring systems in case of a system loss.

This primary production IDW is the starting point for data warehouse analytics. Customers may choose to stay in this type of environment for a period of time without the need to add appliances. But over time, customers will add other appliance-class systems to complement their integrated data warehouse environment, or they may automatically inherit other data warehouses due to consolidations or acquisitions.

Data will be synchronized among the IDW and the appliances in a variety of ways. We will discuss synchronization options later.

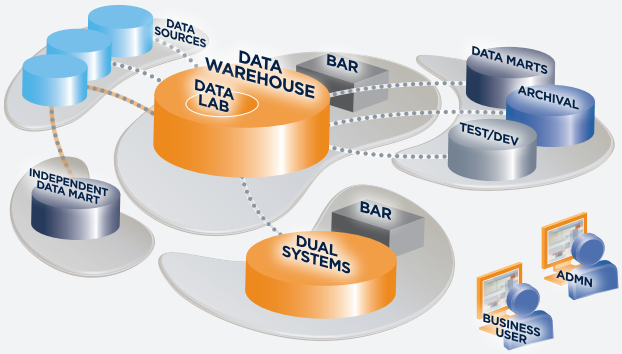


Figure 1. Visual representation of the Teradata Analytical Ecosystem.

The Ecosystem is the Warehouse

When customers employ special purpose systems for their distinct workloads, we call these architectures dual systems for workload optimization. There are two key points that stem from this term. First, what has been created is a dual-system pairing between the IDW and the appliance. Second, the intent of the pairing is to put the right workload on the best platform to achieve the price/performance/value required for that application.

As the business requirements increase demand for high availability, customers can choose to not only increase single-system availability, but may choose to employ a dual-system strategy for high availability and disaster recovery. We call these architectures dual systems for availability, as two systems are paired for availability and failover of workloads.

The most common availability architecture is Dual Active, in which two systems are configured in an Active/Active configuration. Another common availability architecture is Active/Standby, followed by an environment that only delivers disaster recovery where the primary production system is coupled with a disaster recovery system at a remote location that is only used in a true disaster.

Business users access the analytical applications, such as customer relationship management (CRM) and supply chain management (SCM), as well as BI solutions which complete the ecosystem by providing subject-based analytics with familiar user interfaces, such as dashboards and scorecards, and also direct access to the analytical data.

Finally, the database administrators (DBAs) and system administrators get a single operational view to the entire Analytical Ecosystem for systems management, alerts, and resource optimization.

The resulting Teradata Analytical Ecosystem is a flexible environment that enables enterprise analytics.

THE BEST DATABASE FOR ANALYTICS	ARCHITECTURAL FLEXIBILITY	OPTIMIZED DECISIONING	SUPERIOR OPERATIONAL EXECUTION
Leverage unmatched performance, scalability, and manageability	Meet your requirements with any architecture, price point, and data model	Discover smarter insights faster using superior in-database and analytical processes	Drive faster front-line actions with better, more relevant insights

Figure 2. Teradata leadership and innovation.

LOOKING BACK

It's helpful to look back at events in the data warehouse market that have driven the need for an Analytical Ecosystem. In the early days of computerized data collection, simple transaction processing systems collected modest amounts of data about specific subjects. Today, with the proliferation of computing systems, the number of transactional systems creating data is astounding. Additionally, data is not just created by the traditional rigid transactional system. Rather, new kinds of transactional systems have automated means of creating their own data, including inventory management systems or location information such as location-based data. Web sites, automated phone systems, ATM machines, trading systems, and many others are continuously creating data, with consumers – not corporate employees – driving the data creation for the business. Going beyond business-created data, we have social data which are also created and driven by consumers and are extremely valuable for businesses to collect and mine.

There is an abundance of data that companies need to collect, store, analyze, keep or discard, experiment with, and turn into useful information that their users leverage to make decisions. The amount of data, the speed at which they're being created, and the need for making informed business decisions continue to accelerate. To deal with the growth and management of data, customers have employed a variety of techniques that are advocated by Teradata.

But not all data has equal value. Historical data, for example, may have reduced usefulness as more time goes by. Or it may have an uneven rate of usefulness reduction. As an example, historical data may have reduced value over time, but spikes in value at the one-year mark for year-over-year comparisons before continuing its downward slope of usefulness. Other data may be considered mission critical and require continuous availability.

This is the core requirement driving the need for the Teradata Analytical Ecosystem – vast amounts of data with differing business value, usage pattern, and service level agreements (SLAs).

Teradata's multi-system architectures have been designed from the beginning for flexibility and growth, providing a perfect path for the Analytical Ecosystem. So, we designed the Teradata Analytical Ecosystem to help customers start early with a key business solution, and then grow that solution or add other components to the analytical ecosystem as the business demands grow while maintaining visibility, control, and a single view of the business across the entire data warehouse environment.

To be clear, Teradata is not deviating from the core belief that a single, integrated data warehouse is, has been, and continues to be the advocated path for delivering the lowest total cost of ownership, the highest business value, and the shortest time to market for new applications. Teradata advanced workload management features enable different workload priorities to coexist on the same system providing business value that supports company objectives and user SLAs. This is the gold standard for corporate analytics.

Teradata does, however, recognize that there are legitimate reasons why companies will need more than a single data warehouse system. Among those reasons are high availability and disaster recovery. While Teradata single-system availability is best-in-class, the only way to provide guaranteed continuous access to users, even through disaster situations, is by employing a second system that would be used to shift workloads for planned and unplanned outages, including extended outages that result from disaster situations.

HOW ARE BUSINESSES ORGANIZING AND MANAGING DATA?

- Integrated Data Warehouse
- Dependent and independent data marts for departmental or regional requirements
- Dual systems for higher availability
- Data labs or sandboxes for prototypes
- Test/development systems
- Dual systems for workload optimization

Figure 3. System architecture strategies.

BUILDING ON PILLARS OF INNOVATION

Teradata is a global leader in analytics and business intelligence. We've built that leadership position on four distinct pillars of innovation that enable our customers to build world-class data warehouse environments. Teradata Analytical Ecosystem appropriately fits in the Architecture Flexibility pillar since it comprises a nearly unlimited combination of architectures and tools to meet customer requirements.

As a best practice, customers are encouraged to pick a specific solution and deploy it on any of the Teradata platforms knowing that our analytical ecosystem will accommodate their known business requirements.

Customers can start expanding their environment immediately by adding more nodes to their existing integrated data warehouse. Or, they may choose to expand by adding a second system to support specific needs such as high availability, disaster recovery, deep history storage, performance sensitive applications, sandbox, test, development, or quality assurance, using the Teradata Workload-Specific Platform Family.

But the real Teradata value add is in not only providing the platform family to support the growth, but in our expertise and consulting experience to lead our customers to successful implementations, all while leveraging the same application code and the same developers. The growth is non-disruptive and allows our customers to protect their investments while achieving business competitive advantage.

While the use cases highlighted in Figure 4 seem similar to the architectures presented earlier, they are different in that these focus on the usage/value/business problem side of the equation that's driving the customer need, rather than the actual architectural design of the Teradata systems.

ANALYTICAL ECOSYSTEM USE CASES

USE CASE	DESCRIPTION/EXAMPLE
Integrated Data Warehouse	Single version of the data
High Availability	Dual active systems
Dual Systems for Workload Optimization	Corporate reporting + web analytics
Analytical Data Archival	Telco compliance (appliance with CDR)
Data Lab/Sandbox	Prototype, new data exploration
Query Federation	Query federation with external data sources
Backup/Disaster Recovery	Data warehouse with an appliance/BAR as backup/DR
Dependent Data Marts	Data warehouse + regional data marts
Independent Data Marts	Independent business units
Data Preprocessing	Oil and gas (preprocessing seismic data)
Data Transformation Hub	Appliance as ETL transformation/staging hub
Test/Development	Testing and application development

Figure 4. This is not an exhaustive list, but it represents some interesting use cases we've seen directly from customers.

DESIGN, IMPLEMENTATION, AND MAINTENANCE REQUIREMENTS

Now that we've seen what the Teradata Analytical Ecosystem looks like graphically, let's review the key requirements to properly design, implement, and maintain this environment. As shown in Figure 5, the technical requirements are around data synchronization, user and query routing, monitoring and control of the environment, and status reporting and exception handling. These are at a very high level. Actual implementation details are determined based on the customer's technical requirements/constraints and their business requirements.

ANALYTICAL ECOSYSTEM REQUIREMENTS

- Load data from transaction systems
- Copy/move data and objects between systems
- Synchronize data based upon business requirements
- Master data management
- Route users/queries
- Manage mixed workload
- Monitor environment
- Handle exceptions
- Report status
- Backup, archive, and restore

Figure 5. Key functional requirements of the analytical ecosystem.

PRODUCTS, FEATURES, AND SERVICES FOR THE ANALYTICAL ECOSYSTEM

For planning purposes, Figure 6 highlights the major products, features, and services that enable the Teradata Analytical Ecosystem. This is not an exhaustive list and not all elements shown must be present or used.

THE TERADATA DATABASE

At the heart of the Teradata Analytical Ecosystem is the powerful Teradata Database, which has been built from the ground up to support scalable, high-performance parallel processing for decision support. It is the gold standard in decision support databases with a long list of industry-leading features, including:

- Built from inception for data warehousing
- One hundred percent parallel design that eliminates bottlenecks
- Automatic database management for ease of use
- Dynamic query optimization
- Multi-dimensional scalability prevents technology limitations
- Built-in high availability
- Continuous innovation and thought leadership

THE TERADATA WORKLOAD-SPECIFIC PLATFORM FAMILY

All Teradata platforms are powered by the same Teradata Database, enabling application portability, human and systems resource optimization, and investment protection.

It is important to note that the Teradata Database is built from one code base. That means that "Teradata is Teradata," regardless of the platform on which it runs. And that means that customers can expect the same database features, and hence application portability, as they move from platform to platform. Not all databases can make this claim.

The Teradata Workload-Specific Platform Family is flexible; any family member can take on most any role. But there are some natural fits in which some platforms are more appropriate for a certain role than others. (See Figure 7.)

CORE PRODUCTS AND CONSULTING SERVICES

TERADATA SOFTWARE	CONSULTING SERVICES
Teradata Viewpoint	Platform Family Services
Unity Director	Systems Management Services
Unity Loader	Dual Systems/High Availability
Unity Data Mover	Workload Management
Unity Ecosystem Manager	Teradata Master Data Management
Teradata Parallel Transporter	
Teradata Master Data Management	
BACKUP, ARCHIVE, RESTORE	
TERADATA PLATFORM FAMILY	
Data Mart Appliance	
Extreme Data Appliance	
Data Warehouse Appliance	
Extreme Performance Appliance	
Active Enterprise Data Warehouse	
TERADATA DATABASE	

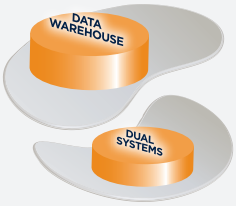
Figure 6. Core products and consulting services of the Teradata Analytical Ecosystem.

TERADATA WORKLOAD-SPECIFIC PLATFORM FAMILY

ACTIVE EDW

- Optimized for integrated data warehouse, active workloads and Dual Active solutions
- Strategic and operational intelligence
- The highest levels of concurrency, tight SLA requirements, and mission critical availability

DATA WAREHOUSE APPLIANCE



- Platform can deliver integrated data warehouse, as well as disaster recovery or Dual Active solutions
- Optimized for analytic workloads
- Foundation for data marts, test/development, or data preprocessing

EXTREME PERFORMANCE APPLIANCE

- Ideal for Dual Active or operational applications
- Optimized for hyperanalytics
- High levels of concurrency, low data volumes, and very tight SLAs

EXTREME DATA APPLIANCE

- Designed for archival analytics
- Built for analytics on large data volumes

DATA MART APPLIANCE

- Optimized for data marts and test/dev environments
- Departmental analytics

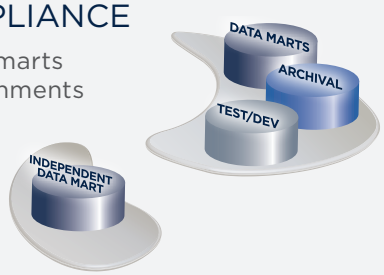


Figure 7. Suggested platform roles

TERADATA BACKUP, ARCHIVE, AND RESTORE SOLUTIONS

Backup, Archive, and Restore (BAR) is an important part of any data warehouse environment, including the Teradata Analytical Ecosystem.

Teradata BAR Solutions protect data to ensure availability during:

- System hardware failures
- Application failure or corruption
- Data corruption or loss
- User errors
- Disasters

Teradata BAR Solutions provide businesses with the peace of mind that their data is stored in an accessible, non-volatile, and known working state. They also ensure the integrity of the data for their recovery objectives.

The value of a Teradata data warehouse solution stems from its unmatched performance and a parallel architecture that distributes data across the entire Teradata system. Teradata BAR Solutions are the most effective way to leverage the performance and reliability of the Teradata data warehouse for protection of key data.

Teradata BAR Solutions provide a fully integrated set of industry-leading, matched, and certified components that store and manage data in a way that is compatible with a Teradata data warehouse, allowing the solutions to minimize the interruption of critical business processes.

TERADATA VIEWPOINT

Teradata Viewpoint is another key ingredient in the Teradata Analytical Ecosystem. It provides a centralized console for monitoring, managing, and controlling all the parts of the data warehouse environment. Its web-based delivery model is a perfect solution for making information available to a broad set of users, ranging from system administrators and DBAs to end users who can submit queries and check the status of the system and their queries.

Teradata Viewpoint grew out of a vision in which monitoring, management, and control of all Teradata management utilities could be displayed from one platform or centralized infrastructure.

Not only did Teradata migrate older management applications to this web-based platform, but we also built in never-before-seen functionality. The result of this combination is a single, centralized deployment mechanism, accessing an entire upgraded suite of management and monitoring applications. It can be combined with brand new functionality, such as the Teradata self-service portlets, to give business users access to vital system information that will help them perform their daily tasks. All of this functionality, deployed on top of an intuitive, easy-to-use web portal, allows user access on a massive scale from any browser.

Teradata Viewpoint not only displays real-time information, but has a unique capability called rewind. This capability lets the administrator go back to any specific point in time to see the state of all parts of the Teradata Analytical Ecosystem. It's similar to taking continuous snapshots and being able to go back to view them.

This is extremely powerful and useful for debugging issues or finding the root cause of a problem. It's also useful for standard performance monitoring and capacity planning. Rewind allows an administrator to actually rewind the Viewpoint page and associated portlets, providing a snapshot of the portlets and environment at a past moment in time. The amount of history and the granularity of the time slices can be controlled by the Viewpoint administrator. Note that rewind is dependent on the history data one keeps stored in the Viewpoint data collection.

So, for example, if a DBA notices that the system is down, he simply opens Viewpoint, and rewinds several portlets back to the moment in time in question and immediately gets a picture of what the environment looked like right before the system went down. In addition, the DBA can check other portlets to diagnose why the system went down and determine what happened. This is one of the most useful and intuitive system analysis and diagnostic tools that Teradata offers.

Rewind also lets an administrator:

- Pause a portlet.
- Refresh a portlet to get the most current information.
- Allow certain portlets to be rewindable.
- Keep some portlets fresh, while rewinding others.

Rewind is a tool that is primarily for administrators, but access is based on user role and portlet access permissions.

TERADATA PARALLEL TRANSPORTER

Getting data into the data warehouse is the first step in populating a data warehouse environment. Teradata offers a full range of data loading tools, each with its own characteristics and benefits. Teradata Parallel Transporter is a data loading tool that of all the previous versions of the stand-alone tools, enables bulk loading of data as well as trickle loading. The primary purpose of Teradata Parallel Transporter is to get data into the data warehouse from external transactional source systems.

MULTI-SYSTEM ENABLING

When the Analytical Ecosystem consists of more than one production system, there is a need to route queries between systems, copy and/or move data between these systems, and monitor and control the entire data warehouse environment.

Enter Teradata Unity, a portfolio of powerful products that are integrated to work together to turn a multi-system environment into an orchestrated analytical ecosystem. Within a Teradata analytical ecosystem, Teradata Unity comprises all the features and capabilities needed to simplify and synchronize systems. There are myriad benefits for both IT and business:

IT BENEFITS

- Automated management of the analytical ecosystem enables the warehouse to grow in sophistication and scale without a corresponding increase in staff
- Architectural flexibility to deliver operational and strategic intelligence meets both business and technical needs
- A comprehensive, proven solution reduces risk, saves time, and provides seamless growth in enterprise analytics

BUSINESS BENEFITS

- Transparent access to diverse analytics gets query results to users while shielding them from IT complexities
- Continuous access to analytics allows uninterrupted decision making
- A cost effective solution minimizes the IT investment needed to deliver business analytics

DATA SYNCHRONIZATION BETWEEN TERADATA SYSTEMS

Applying data and database updates to the primary production warehouse isn't enough; these updates must also be applied and synchronized between systems. Data synchronization can be on 100 percent of the data, or it may be on subsets of data. The same goes for database changes. For example, a business may determine that 30 percent of its data is mission critical and choose to keep a copy of that data on a second system for high-availability purposes.

To achieve this data and database synchronization, Teradata offers three different options that work together, each its own characteristics, benefits, and appropriate time to use.

UNITY DATA MOVER

Unity Data Mover is ideal for copying data between systems and offers a variety of flexible benefits. It is useful for synchronizing both small and large tables between systems. It can also work on full table copies or on partial table copies where just a portion of the data is new or changed and needs to be copied to another system.

Copying partial data would require identifying the changed data; a timestamp column is a popular way to achieve this. Unity Data Mover can also copy full table partitions, which may be helpful for some. And it uses the Teradata load utilities to copy data between systems and intelligently chooses which load protocol to use. Unity Data Mover is a sound choice when large amounts of data need to be copied between systems.

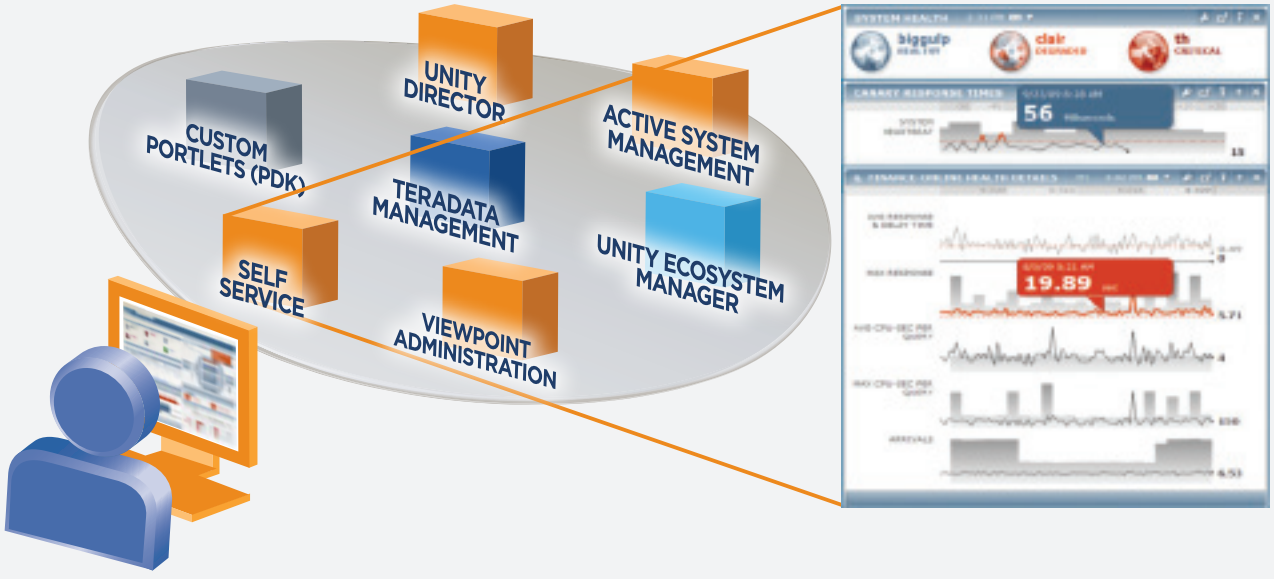


Figure 8. Teradata Systems Management Portal.

UNITY LOADER

Unity Loader offers "dual load" capability for those times when the goal is to get data from source systems loaded to more than one Teradata system.

It provides the robust capability of directing Teradata Parallel Transporter-based bulk loads to more than one Teradata system and keeping track of completion status so you don't have to. Additionally, it brings intelligence to data loading. It will analyze the incoming data loads and automatically determine which systems to send the loads to; freeing you from having to figure it out for each load job.

And because Teradata Unity is an integrated set of products, load jobs and queries that depend on the data are sequenced so that users always get the most up-to-date results.

UNITY DIRECTOR

In a Teradata Analytical Ecosystem, users and queries need to be routed to the correct system. This is a requirement for high-availability systems where the expectation is that when one system becomes non-operational, users are easily routed to the alternate system. Unity Director delivers this functionality. It allows user and query routing to be selective and transparent; no additional effort is required by the user. Unity Director is a powerful feature that allows administrators to control how to route and re-route users, and even helps load balance users between systems. It also has an auto-route feature that requires no setup or configuration. Unity Director intelligently determines which Teradata system can satisfy the query and automatically routes the query accordingly. It also routes SQL for database updates (DDL) and data updates (DML). In doing so, it keeps databases in sync across systems, applying the same changes to multiple systems.

UNITY DIRECTOR PROVIDES FULL USE OF ACTIVE/ACTIVE TERADATA SYSTEMS

- Resides between the client application and one or more Teradata systems
- Automatically routes queries based on which system can satisfy the query
- Routes sessions according to rules set by the administrator
- Helps manage workload between Teradata systems
- Provides failover from one Teradata system to another system during maintenance or downtime

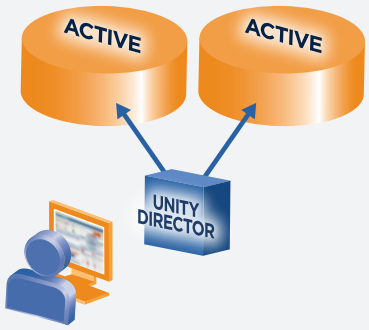


Figure 9. Unity Director software.

UNITY ECOSYSTEM MANAGER

Unity Ecosystem Manager is Teradata's product for providing monitoring and control capabilities. It pulls together all parts of the Teradata Analytical Ecosystem to help model the relationship among all parts of the environment; it monitors the hardware components, the processes, and the state of the data throughout the ecosystem. Unity Ecosystem Manager displays its user interface through Viewpoint. But while Viewpoint provides the capability to view and monitor all parts of the ecosystem, it is Unity Ecosystem Manager that ties all the parts together so that the interdependencies are defined. So for example, whenever there are users and applications that are dependent on new data being loaded into the warehouse, Unity Ecosystem Manager can model that dependency so that when the ETL server is non-operational, the downstream implications are easily identified. Unity Ecosystem Manager, in essence, models the dependencies of the Teradata Analytical Ecosystem.

Further, Unity Ecosystem Manager is action-oriented and can be programmed to take automated action based on certain events. For example, if users on a Teradata system are dependent on the latest data being loaded into the

system – and if that system has not been updated with new data, but another Teradata system has been – Unity Ecosystem Manager can be programmed to take action and re-route those users to where the latest data is located. Unity Ecosystem Manager will work with other products, such as Query Director, to accomplish this task.

Features of Unity Ecosystem Manager include:

- Provides a unified view of the data warehouse environment.
- Monitors Teradata systems, supporting servers, processes, and applications.
- Simple point-and-click controls to manage the state of system-wide resources and query routing.
- Monitors load jobs and data synchronization status.
- Enables grouping processes together and tracking dependent processes.
- Monitors thresholds and reports of discrepancies.
- Enables user routing based on application readiness (data currency/process completion).
- Supports graceful failover and failback during planned and unplanned outages.

TERADATA MASTER DATA MANAGEMENT

Although it's not a requirement, Teradata Master Data Management (MDM) can be an important asset to maintaining a consistent representation of data. It is not only used within the analytical ecosystem, but also with transactional systems that often hold the master version of the data. It is used to classify and define master data through the use of a centralized integration manager, sometimes referred to as a hub. It leverages policies and procedures for access, update, and overall management of this central resource and its coordination with other participating systems across the enterprise. Areas such as customer data integration (CDI), management of customer reference data and product information management (PIM), and management of product and supplier reference data, are domain-specific subsets of MDM.

AVAILABILITY MANAGEMENT SERVICES

Provides full use of active/active Teradata systems.

ENVIRONMENT – Physical conditions in data center

INFRASTRUCTURE – IT assets, architecture, compatibility

TECHNOLOGY – HW/SW models and versions

SUPPORT LEVEL – Coverage hours, response times, proactive processes

OPERATIONS – Daily system and database administration

DATA PROTECTION – Processes and features that minimize or eliminate data loss

RECOVERABILITY – Strategies and processes for back-up, archive, restore data or complete system recovery

Figure 10.

SERVICES

In addition to the rich set of products that Teradata offers for the Analytical Ecosystem, services are yet another key component.

TERADATA PROFESSIONAL SERVICES

Teradata offers all the necessary services to help customers select the right platform for the right applications, taking into account the performance and value characteristics of the applications and matching them with the right platforms. Systems management services also help set up Teradata Viewpoint and Unity Ecosystem Manager to provide the right monitoring and control for system administrators, database administrators, users, and even management dashboards. Teradata Dual Systems for Availability services help architect and implement an environment that delivers on the high availability and disaster recovery requirements of a company; ensuring that users' SLAs are met. Workload management services help optimize the workloads and usage of each Teradata system so that system resources are used in accordance with company expectations. In other words, that high priority work takes precedence over lower priority work which helps distribute and balance workloads between systems during normal operation and during system outages. Finally, Teradata Professional Services consultants offer Master Data Management services to help customers plan and implement Teradata MDM between their analytical and transactional systems.

TERADATA CUSTOMER SERVICES

Teradata Customer Services personnel help keep data warehouses operational through proactive maintenance and through planned and unplanned outages. They help coordinate and deliver software and hardware upgrades to keep a business running during normal business operations and during outage scenarios.

Teradata Customer Services' availability initiative is called Availability Management Services (AMS). It is a customer-centric approach for supporting greater Teradata system productivity by mitigating the risk of planned, unplanned, and degraded downtime. Although a Teradata data warehouse is designed and built with intrinsic features to deliver high data availability right from the start, risk events occur both inside and outside a Teradata system that can reduce or degrade the availability that end users experience.

TERADATA DUAL ACTIVE SOLUTION

High availability enabler for mission-critical applications.

- Delivered through two systems
- Eliminates planned and unplanned downtime
- Automatic query routing
- Products
  - Teradata Platforms + BAR
  - Unity Director
  - Unity Loader
  - Unity Data Mover
  - Unity Ecosystem Manager
  - Teradata Viewpoint
- Services
  - Platform Family Services
  - Dual Active/High Availability
  - Systems Management Services

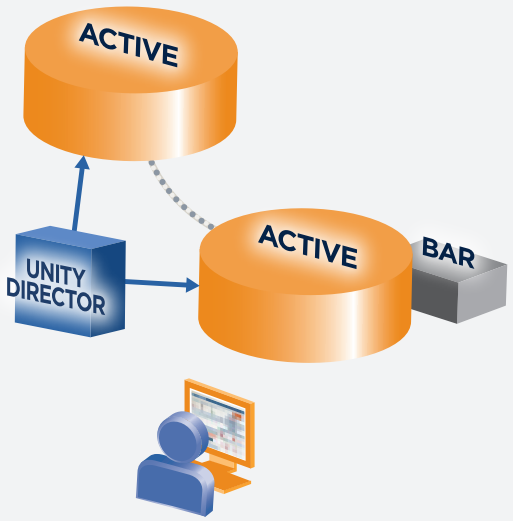


Figure 11.

AMS provides a best-practice framework for understanding and mitigating availability risk, tools for identifying specific availability management gaps, and a portfolio of products and services to match availability needs.

USE CASES

DUAL SYSTEMS FOR AVAILABILITY

Now that we've reviewed the definition of the Teradata Analytical Ecosystem and the products and services that it comprises, here are some specific use cases.

Dual systems for availability is one of the most common implementations. As companies use their Teradata data warehouse for more and more mission-critical applications to run the daily operations of their business, high availability and disaster recovery become must-have capabilities.

In this environment, two (or more) systems work together to provide high availability and disaster recovery. This can range from our high-end dual active solution where systems work together in an active/active configuration to an active/standby configuration. It can also be a single active production system coupled with a disaster recovery system.

For dual active, users and applications are running on both systems during normal operations. If one system becomes unavailable, all users or a subset of users can be transparently re-routed to the second system. Through our workload management capabilities, we can adjust the amount of system resources assigned to the workloads to ensure high value work gets the highest priority on the system. This active/active configuration delivers the greatest ROI and workload throughput because both systems are actively sharing in the production workload.

For an active/standby configuration, both systems are considered production systems, but only one is actively running a production workload at any point in time. The second standby system is kept up to date with database and data changes so that it's ready to take over production workload within the defined SLAs. During planned or unplanned outages of the active system, work can be redirected to the stand-by system to continue business operations. The time to switch users/workloads depends on company SLAs, which would be built into the design of the analytical ecosystem environment and enabled by the products and services used.

When a production/DR implementation is used, only the primary production system is utilized to handle daily workloads. The DR system is kept up to date with database changes, but it's only put into use during a true disaster. As such, the time to load current data and shift workload to this DR system could range from a minimum of hours to a typical three- to five-day recovery period. While it doesn't meet the strict SLAs of the other architectures, it takes less effort to design this solution, which works well for many companies.

~ Teradata data warehouse + Teradata appliance

~ Appliance for web analytics

~ Appliance for intensive analytics

PRODUCTS

- Teradata Platforms
- Unity Director
- Unity Loader
- Unity Data Mover
- Unity Ecosystem Manager
- Teradata Viewpoint

~ Services

- Platform Family Services
- Systems Management Services
- Workload Management Services

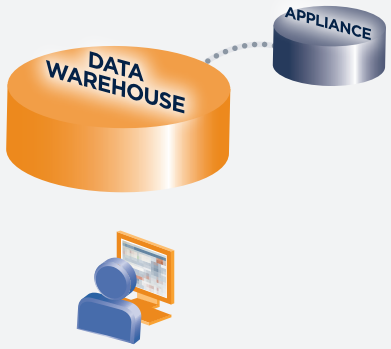


Figure 12. Platform flexibility and workload optimization.

DUAL SYSTEMS FOR WORKLOAD OPTIMIZATION

Another class of systems working together within the Analytical Ecosystem is dual systems for workload optimization. In this case, there is a matching of user/application requirements with the platform on which they're running. A mission-critical workload that requires the absolute best performance and dedicated system resources may be isolated on its own (second) system. This second system could run on any of the Teradata systems but would most likely be an appliance-class system

with the price/performance characteristics that support the application. Data would typically flow from the primary integrated data warehouse to this downstream dependent system. However, that doesn't always have to be the case; data can move in either direction.

ARCHIVAL DATA

Another class of architecture that can be found within the Teradata Analytical Ecosystem is that which supports deep historical data. This is data that was once part of the primary integrated data warehouse, but as time passed, it developed less immediate value. Companies don't want to discard this data because having it available for historical reports or reference is useful. This is especially true for industries that have regulations governing how long data needs to be available for auditing or reporting. In some cases, the regulations come from the government, other times they are industry-specific standards. Having data readily available helps organizations meet these requirements. But storing historical data on the primary data warehouse may not be a cost-effective option. Companies may choose to host this data on a dedicated system that runs the Teradata Database and is optimized for large volume storage, with performance as a secondary goal.

TERADATA DATA LAB

Yet another architecture within the Analytical Ecosystem is the data lab. Sometimes referred to as a *sandbox*, the data lab is used to load data that is temporary in nature and has not gone through the rigorous IT process of cleansing and rationalizing; nor have recurring processes been established for keeping them current. In other words, this is data that has limited initial value until later proven to have enough business value to be promoted into the primary data warehouse. While the data lab can be a physically separate system, Teradata advocates hosting the data lab in a non-production portion of the production warehouse so that the new unproven data can be analyzed along with the rest of the production data that is already in the warehouse.

The data lab is useful for experimental analytics where all the data needed does not already exist in the warehouse; sometimes referred to as the 90:10 rule where 90 percent of the data needed for the analysis is in the warehouse, and 10 percent represents new data that isn't in the warehouse. Best practices are to bring the 10 percent new data to the 90 percent existing data, and run the analysis. If the new 10 percent yields incremental business value, then the standard IT process is used for making it part of the production data in the warehouse. If the new temporary 10 percent does not bring business value, it is discarded.

FEDERATED ARCHITECTURE

The federated systems concept has been around for more than a decade and has had mixed reviews and success. However, Teradata believes in offering customers choice. We have partnered with top companies in this area to enable their query federation products to work with the Teradata Database. There are use cases that may be beneficial to customers, especially where companies have a distributed approach to managing data.

TERADATA ANALYTICAL ECOSYSTEM BENEFITS	
Investment Protection	One Analytical Infrastructure
Reusability	Flexibility
Ease of Administration	Workload Optimization

Figure 13.

WHY USE THE TERADATA ANALYTICAL ECOSYSTEM?

The Teradata Analytical Ecosystem is a unique, flexible, and powerful approach to delivering enterprise analytics. It offers investment protection, reuse of hardware, application portability, ease of centralized administration, flexibility on where to host data and applications, and workload optimization, all delivered with one analytical infrastructure to meet your business and technical SLAs. And you get all of this with a single source for support. Teradata is your premiere partner for analytics and continues to provide the vision, products, and services to meet your most demanding needs.

ABOUT THE AUTHOR

Imad Birouty is the Program Marketing Manager for Teradata High Availability Solutions, which includes features such as the Teradata Dual Active Solution and Recovery Center offerings. Imad is also responsible for the Teradata Data Mart Consolidation program, Value of Integrated Data, and Total Cost of Ownership.

Prior to this, Imad led the Product Management team responsible for Teradata Platforms including the processing nodes, Teradata BYNET® Interconnect, Server Management, and Operating Systems. He set product strategy and direction and was responsible for seven major platform releases spanning eight years.

Imad joined NCR/Teradata in 1992. He holds an MBA in Finance from San Diego State University and a BA in Human Communication.

10000 Innovation Drive Dayton, OH 45342 teradata.com

TERADATA. THE BEST DECISION POSSIBLE™