Aster Database: "Always On" Availability

The more organizations depend on customer-facing, data-driven applications for competitive advantage, the more they demand richer, deeper data processing at ultrafast speeds on a 24x7 basis. This results in increased availability requirements on these mission-critical systems, especially when compared to traditional massively parallel (MPP) data warehouses. But these availability requirements are increasingly difficult to meet due to growing data volumes. Large data volumes result in bigger, more complex systems and processes, with a much higher number of failure points.

A traditional data warehouse and analytic system can become unavailable in one of two ways:

- Unplanned Downtime – System is completely unavailable due to hardware failures, system shutdown, etc.
- Planned Downtime – System is unavailable due to planned activities such as capacity expansion or upgrades. System may also be partially unavailable as some or all users see significant performance degradation due to routine operations such as loads, exports, backups, indexing, concurrent queries, upgrades, etc.

Aster Database delivers the first analytic platform, a massively parallel (MPP) row and column database with an integrated analytics engine. It is the first MPP data warehouse architecture that allows applications to be fully embedded within the database engine to enable ultrafast, deep analysis of massive data sets. Teradata Aster's solution is built from the ground up using award-winning Recovery-Oriented Computing technology to provide high system availability. Teradata Aster's market-leading "Always On" technology is built to address all issues that have an impact on availability, including planned tasks and unplanned failures.

Automated Recovery from Hardware and Software Failures

Aster Database has built-in fault-tolerance and self-healing capabilities to handle server hardware and software failures. Other MPP data warehouses focus on disk and processor failures, using redundant hardware (e.g. RAID controllers, battery-backed cache, hot stand-by servers). Aster Database goes above and beyond disk failures to protect data and queries from other component failures such as server motherboard, CPU, network, operating system, using Recovery-Oriented Computing software, reducing hardware and administrative costs.

Planned Downtime	
Cause of Downtime in Traditional Systems	Aster Database Solutions
Capacity Expansion	Online precision scaling
Loads and Exports	Online loads and exports
Backups	Online Backup
Upgrades	Fast, parallel upgrades (duration is cluster-size independent)

Unplanned Downtime	
Cause of Downtime in Traditional Systems	Aster Database Solutions
Server failures	• Data replication with transparent fail-over • Online replica restoration
Network failures	Fault-tolerant network aggregation
Transient failures (software/network/etc.)	Online Resync
Human errors	• Single-system view • Simplified management • Online recovery • Bare-metal installs
Concurrent workloads	Workload Manager for role-based QoS (fair-share queuing, priority scheduling)
Data corruption/media failure/site disaster/etc.	• Full and incremental backups • Cluster-wide and table-level backups • Massively parallel fast backup and recovery • Offline and online recovery • Compressed backups for fast recovery

Figure 1: Aster Database "Always On" capabilities.

Overview

Aster Database "Always On" capabilities let data-driven applications and data warehouses provide continuous support to business operations. Aster Database Recovery-Oriented Computing ensures high availability via its fault-tolerance and self-healing capabilities. Aster Database Live Admin features allow administrative operations that traditionally required planned downtime, to be done online in an Aster Database database.

Highlights

- Uninterrupted queries even in the event of hardware and software failures
- Online replica restoration for continued fault tolerance
- Online scaling by incorporating new servers into cluster without downtime
- Online loads and exports through specialpurpose servers
- Online backup and recovery for data protection

"We needed excellent uptime and real-time insights into our customers' business 24x7. Aster Database allows us to have superior scalability on commodity hardware, giving us a tremendous cost and performance advantage. Teradata Aster was the only company we could find that gave us this capability."

Shawn Farshchi, Chief Operating Officer
Coremetrics

Coremetrics
Relentless Made Easy

In-Cluster Data Replication and Transparent Fail-Over

Aster Database allows administrators to select the replication factor and keeps an appropriate number of data copies within the cluster. Replicas for various partitions are algorithmically placed across the cluster to ensure fair sharing of the load among multiple servers for performance consistency, in case of a server failure.

A server failure is handled by transparently failing-over queries to data replicas within the cluster. Due to such transparent fail-over capability, ongoing queries complete successfully even as a server fails. This ensures that business processes continue even as administrative staff repairs the failures.

Aster Database automatically manages replica consistency to ensure replicas are transactionally consistent (full ACID properties), even under heavy workloads.

Replica Restoration and Online Resync

Aster Database has a continuous self-diagnostics process that monitors the cluster for failures. This process is aided by a built-in knowledge repository to distinguish between permanent and transient failures and deal with them differently. This minimizes the impact of failures and dramatically reduces the burden on administrative staff.

- If a server is determined to have permanently failed, Aster Database can restore the replication factor by creating new replicas in the background – on existing servers or by adding a new server – without causing any downtime.
- In case of a transient failure (e.g., temporary network unavailability at a server), Aster Database temporarily fails-over ongoing queries to the servers containing data replicas. During transient failures, admins may change or add new data in the active replica. Aster Database has the ability to track incremental data changes during the failure. Once the original server recovers, Aster Database transfers only the incremental changes from the replica to the original server using its Online Resync feature. Hence Online Resync requires only a small amount of data be transported over the network to bring the original server up-to-date, resulting in rapid restoration of the replication factor. And unlike other systems that require complete and time consuming data restoration, all of this is accomplished with zero administrative intervention.

Fault-Tolerant Network Aggregation

The Aster Database network aggregation feature leverages multiple Ethernet network cables/ports in parallel to increase network speed beyond the limits of a single network interface card (NIC) or onboard network ports, as the case may be. Parallel use of the cards is programmed to deliver redundancy for higher system availability. Network hardware (NIC/cable/port) failures are automatically detected, and the remaining NICs transparently continue to handle network traffic with no disruption to system operations. This allows for graceful degradation of performance – if one out of four NICs fails, the network performance for that specific server will continue transparently at 75% of the original. This allows network administration staff to repair the failure by taking the server offline when convenient without causing system downtime. The system continues functioning with its network errors transparently masked by software algorithms.

Expand System Capacity with Ease

Aster Database online precision scaling eliminates the need for downtime traditionally required to expand capacity of the system. Adding new servers to the cluster requires a single click on the Aster Management Console (AMC) and zero preparation for the server. When adding a new machine to the cluster, the administrator only needs to input the MAC address of the first network interface and power on the bare-metal machine. The system automatically gets the software (including the operating system), formats the drives, configures the network, and rebalances the existing data. All this is done in the background and the system continues to be available to users during this process. Similarly, servers can be taken out of Aster Database and repurposed for other use with a single-click on AMC without incurring any system downtime.

"We've been very pleased with both the reliability and availability of Aster Database. We've had cases where disks and other components have failed in our cluster, but the MPP database has kept on going answering queries with the same performance as before."

Kris Wehner, Chief Technology Officer
Aggregate Knowledge

Easily Avoid and Minimize Human Errors

Human errors are a large contributor to database downtime. To minimize the occurrence and impact of such errors, Aster Database includes the following capabilities:

- A greatly simplified set of administrative tools that minimize the number of parameters administrators need to optimize. Simplified management of administrative tasks reduces points of failure and chances of errors.
- Single-system interface for administrative tasks such as software installation, upgrades, network configuration, etc. These are centrally managed by Queen servers using the single-system view capability. The administrator does not need to directly manage any other servers (Workers or Loaders), minimizing chances of errors.
- Backup and restore capabilities at both system and object levels. User errors can be rectified by restoring the appropriate backups.
- Bare-metal incorporation of new servers minimizes chances of error during capacity expansion.

Seamless Data Loads and Exports

Traditionally, data loading is done in a dedicated window when the query workload on the system is at a minimum. This is done to reduce the impact of loads on query performance. In today's business-critical data management and processing systems, large volumes of data need to be loaded frequently and made rapidly available for analysis. Under these circumstances, separating loads and query workloads to different time windows is not an option. Similarly, data may need to be frequently exported for consumption by downstream applications, which can interfere with ongoing queries. Aster Database architecture includes dedicated servers for loading and exporting data. This separation of responsibilities means loads and exports can be done concurrently with user queries, with no impact on query performance.

Fast, Online Data Backups and Restores

Aster Database provides backup using its massively parallel infrastructure. Aster Database can back up data to a cluster of servers dedicated for backup (backup cluster) or to network storage (SAN/NAS). Worker servers of Aster Database stream out data in a massively parallel way to Backup Cluster. Given the parallel use of cluster resources, backups can be done online in a much shorter time frame. Since the backup process is managed by a dedicated cluster of servers, ongoing queries experience no impact. Similarly, restores are also massively parallel and can be done in an online manner for table-level recovery. In other words, tables can be restored efficiently without requiring a full database restore. Backup Cluster can be in a geographically separated location from the Aster Database massively parallel analytic platform to protect from site-wide disasters. And backups can be restored to a secondary site running Aster Database software to enable disaster recovery.

Concurrent Administrative Operations

Database administrators routinely carry out activities that require significant data movement over the internal-server network. Here are some examples of such activities:

- Creating table copies for data transformation, summarization, and staging for more complex queries (e.g., CREATE TABLE AS SELECT operation)
- Repartitioning a table to optimize for changing query workload or for data redistribution across newly added servers

In other systems, such operations can significantly degrade the system performance or even require downtime. Aster Database includes network algorithms that minimize the impact of such operations on ongoing queries and enhance system availability.

Concurrent Workloads

Aster Database dynamic mixed workload management provides automated rolebased and policy-driven tools to predictably manage high-concurrency, multi-user workload environments and ensure quality of service to 100's of users. Queries and other jobs can be scheduled according to assigned priorities so high priority queries take precedence, and to dynamically allocate resources for incoming queries. Aster Database allows administrators to easily manage workloads ranging from interactive queries and trickle loads to large-scale reporting and batch data mining. In addition, Teradata Aster provides the first-ever workload management that works on commodity hardware.

About Teradata Aster

The Teradata Aster MapReduce Platform is the market-leading big data analytics solution. This analytic platform embeds MapReduce analytic processing for deeper insights on new data sources and multi-structured data types to deliver analytic capabilities with breakthrough performance and scalability. Teradata Aster's solution utilizes Aster's patented SQL-MapReduce® to parallelize the processing of data and applications and deliver rich analytic insights at scale.

For more information visit
http://www.asterdata.com or
call 1.888.ASTER.DATA

DS_AlwaysOnAv_0412_v1