

The benefits of enabling fallback in the active data warehouse

Deciding when the cost of system downtime overrides the cost of fallback protection. *by Jim Dietz and Lynn Hedegard*

The active data warehouse is rapidly becoming a vital element in today's service-oriented IT environment. As business processes become more real time, interactions between the enterprise and its stakeholders become more reliant on operational decision-making services offered by the active data warehouse. Indeed, it is becoming an operational system and, therefore,

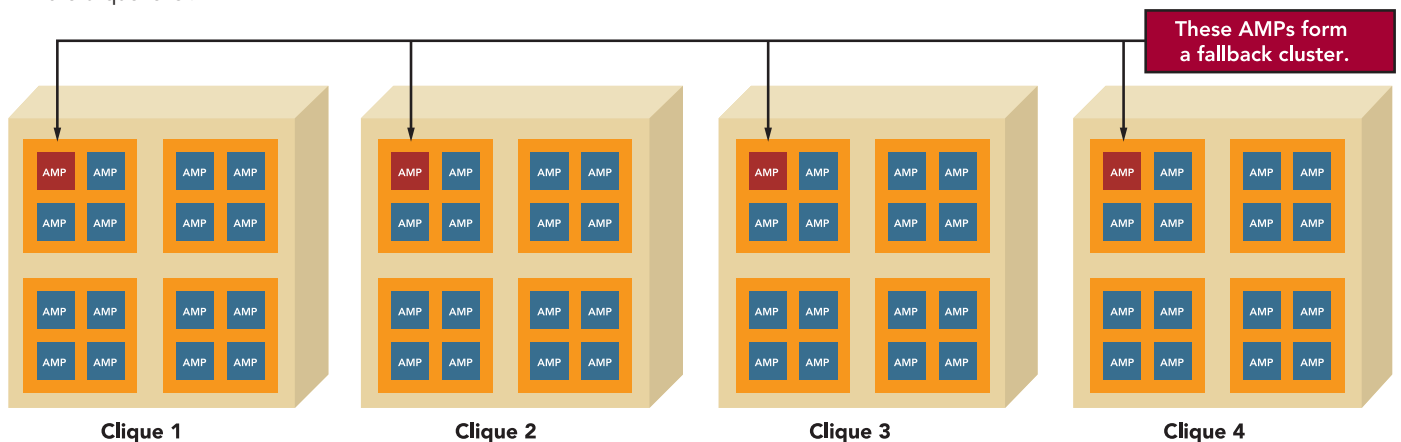
must deliver higher levels of online availability than the back-office data warehouses of the past. An active data warehouse with fallback enabled provides extreme levels of online availability required by today's operational systems.

To extend the availability of database services, Teradata engineers have developed software techniques that ensure rapid recoverability from

failure scenarios. Fallback, one of the features unique to the Teradata Database, manages redundant copies of data on alternate storage subsystems within a single database. This secondary copy of the data is referred to as the fallback copy. Thus, the Teradata Database can minimize the impact of major failure scenarios (such as the complete malfunction of a disk

Figure 1 Fallback clusters

- > A fallback cluster is a logical set of AMPs and Vdisks.
- > Each AMP in a fallback cluster has an equal proportion of its data stored redundantly on the other AMPs' Vdisks.
- > Each AMP in a fallback cluster usually exists in a different clique. This preserves data availability in the event of hardware failures at the clique level.



On a fallback-enabled system, a major system failure will not adversely affect the efficiency of the database. This diagram illustrates how data is stored evenly on AMPs within a fallback cluster. Once the system is repaired and functioning, the data stored on the fallback AMPs is updated to the primary database.

array cabinet) that would disable other relational database management systems (RDBMSs).

Fallback clusters

Fallback is a feature that provides enhanced data protection beyond that offered by redundant hardware components. When a failure occurs and fallback is enabled, Teradata can continue to perform insert, update and delete functions because redundant copies of data exist on one or more fallback AMPs. A small number of AMPs are logically grouped into a fallback cluster. A system with fallback enabled will contain many of these clusters. (See figure 1, page 59.)

Typically, each AMP in a fallback cluster operates on an independent clique. Thus, the failure of a major hardware component such as a disk array will not disable the database. Instead, the data is available via the fallback copy until the failed component is repaired. This architecture ensures that when a given AMP is down, replicated data from each AMP is evenly distributed

across the other AMPs in the fallback cluster. (See figure 2, below.)

Failure detection and recovery

If an AMP can no longer access its data, or if its data is corrupted, it is flagged as a down AMP, meaning it is deconfigured from the system. After all failed AMPs have been deconfigured, the Teradata Database will perform a restart and resume operations using the surviving components. In the Teradata Database, the remaining members of the cluster are aware there is a down AMP and thus will perform all database operations—reads and writes—against the fallback copy in the database.

While the Teradata system is running in fallback mode, each AMP in the associated fallback cluster maintains a journal of all changes made to the fallback data. This architecture ensures that updates to the database can continue without impacting

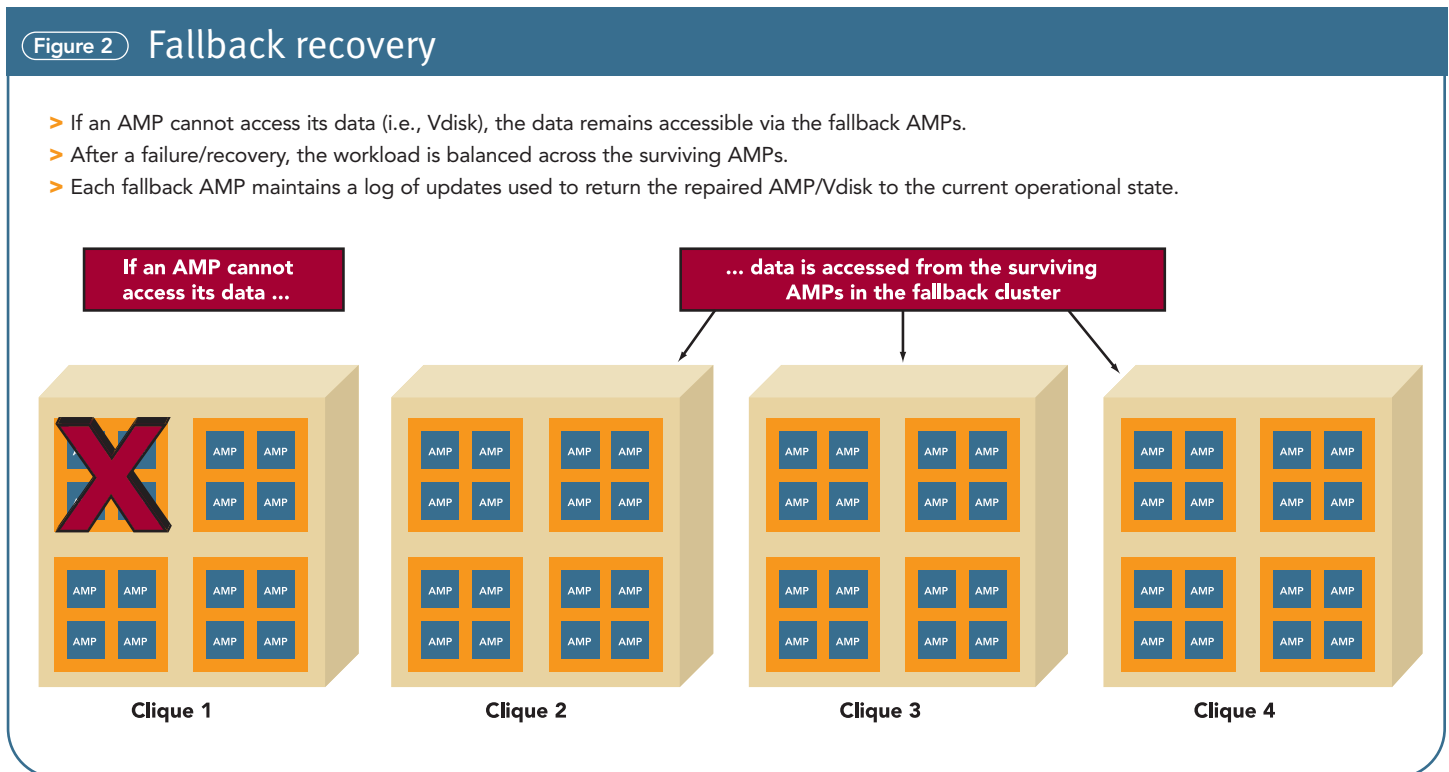
the extract, transform and load (ETL) processes. When the primary AMP is restored to normal service, its data will be updated from the fallback journals.

Furthermore, since all failed components have been deconfigured from the system, repairs to each physical component can be performed while the database system is operational. Before they can be reintegrated into the system, however, their file systems must be brought back to a current and consistent state. (See “Fallback recovery tools,” page 61, for a description of this process.)

Performance impacts of fallback

Fallback has no effect on read performance. Read operations are performed by the primary AMP only—unless it is unavailable, at which time the fallback AMP will deliver the data.

The performance of write operations, however, will be affected by fallback since every write must be performed twice—once



Data is distributed among surviving AMPs if any of the fallback AMPs are blocked. This process ensures that data is not lost during a system failure.

for the primary AMP and a second time for the fallback copy. This does not mean that a fallback-enabled system needs to be twice as large as its non-fallback counterpart. The amount of disk capacity required will increase, but the computation regarding additional disk capacity should only be considered for those tables that are fallback-enabled.

With fallback, the system consumes more resources than one without. To use the fallback option for every table in the system requires that the system be configured with 1.7 times the current disk capacity and additional processing power commensurate with the update load on the system. The necessity to add space can be ameliorated by choosing to fallback-protect only some of the tables in the system. In this case, the additional disk and processing power need to be calculated on a table-by-table basis.

Deploying fallback: A business decision

Whether an organization should enable fallback on its active data warehouse can be answered by analyzing the business—not the technology. The proper business decision is one where the cost of downtime is balanced by the cost of the additional system components necessary to achieve a desired level of availability.

Cost of downtime

When database services become unavailable, various application services also become unavailable. This, in turn, creates operational and financial impacts on the business. Some of these impacts will be seen immediately, such as lost revenue due to a reduction in sales. In other cases, the impact is more insidious, as is the case when a disgruntled customer takes his business to a competitor.

Ultimately, the firm needs to derive a function that relates downtime to cost. Extremely short periods of data unavailability will have virtually no cost to the firm—they aren't noticed by the stakeholders. At the other extreme, long periods of data unavailability will result in a dramatic cost to the firm.

Service levels for database availability have two primary dimensions: availability and time. The first dimension is availability at the macro level. For example, there may be a service-level requirement regarding the aggregate hours of downtime per year. This is the metric that is commonly used to discuss availability.

The second primary dimension for service levels involves the time to recover from a

given failure scenario. This is referred to as the recovery time objective (RTO). RTO is important because most critical business processes can tolerate a short outage—typically measured in minutes—but cannot tolerate an outage measured in hours to days. Teradata systems with the fallback feature enabled can achieve high RTO metrics for single systems because major failure scenarios

Fallback recovery tools

The Teradata Database includes a set of utilities that leverages the architecture of fallback to validate the integrity of data and/or reconstruct damaged data. These utilities minimize the work necessary to restore the Teradata Database to normal service after a failure scenario.

Rebuild

For systems not fallback-protected, a disk array failure would typically necessitate a restore from an external source while the relational database management system (RDBMS) is offline. With fallback enabled, however, the Rebuild tool can put the contents of an AMP back together in a few hours while the rest of the system is fully operational. The Rebuild tool is used to restore the entire contents of a failed AMP that are rebuilt from data on the fallback AMP.

In addition, if data corruption occurs to only one table on one AMP (usually found by the CheckTable tool), Rebuild can recover the contents of that single table from the fallback copy. This can be done with the system online rather than necessitating a lengthy restore process and requiring the applications using that table to be offline.

Failure scenarios that would bring most other RDBMS engines to a complete halt—followed by a lengthy rebuild procedure—can be managed by a couple of short database restarts when using the Teradata Database with fallback enabled.

CheckTable

CheckTable is a Teradata Database utility that exploits the redundant copy of data present in the system. It compares primary and fallback copies of data and reports any discrepancies. If primary rows are missing, the table can be rebuilt. If fallback rows are missing, they can be rebuilt from the primary copy.

Fallback options

The architecture's implementation provides flexibility as to which database objects are replicated on fallback AMPs. The database administrator (DBA) can evaluate the business value of each database object to determine whether or not the object should be protected. Fallback can be declared at the table, database and system levels.

Some tables (e.g., temporary tables for storing query results) should not be fallback-protected as their contents are easily re-created. At the other end of the spectrum, a table that is critical to operating the business and is costly or time-consuming to restore or re-create is a prime candidate for this protection.

—J.D. and L.H.



For an organization to optimize its return on investment, it should consider protection strategies that ensure the availability of database services in the face of severe component failures.

can be recovered with a system restart, which can take anywhere from one minute to four minutes to complete.

A key factor when evaluating the cost of downtime is the expected time to recover from a single failure scenario. (See figure 3, below.) In business-critical environments, the cost grows exponentially for every minute the system is unavailable. Rapid recovery from failure is, therefore, a valuable attribute in a database system.

Cost of high availability

To sustain a given workload, a fallback-enabled system requires more components than a system without fallback. The fallback-enabled system also requires additional compute nodes and disk subsystems.

Having calculated the cost of downtime and the cost to deliver a highly available database solution, the organization can make a decision regarding fallback—or any other alternative for high availability. The costs for deployment should balance the cost of downtime.

The proper business decision is to deploy a system where the system cost is balanced with the cost of downtime. Deploying an extremely reliable system, such that the cost of the system overwhelms the cost of downtime, is a poor use of business capital. On the other hand, deploying a system with less than desirable levels of service availability is also a poor business decision. A single failure scenario could reduce revenues beyond the initial cost of deploying the more reliable system.

What is it worth?

The active data warehouse is an operational system used daily during interactions with customers, suppliers, employees and business partners. It provides business-critical analytic services for tactical and strategic decision making.

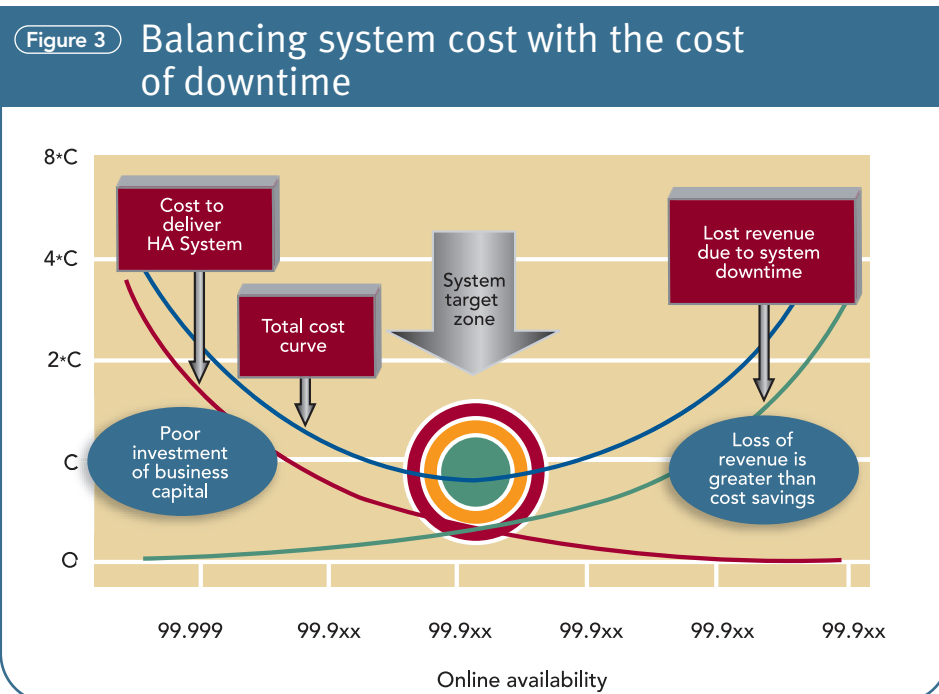
For an organization to optimize its return on investment (ROI), it should consider protection strategies that ensure the availability of database services in the face of severe component failures. The Teradata fallback feature provides the highest metrics for data availability and RTO possible for a single system. It heightens the perceived availability of the database by reducing the overall downtime and improving RTOs.

IT managers and system designers need to assess the impact to the enterprise when data suddenly becomes unavailable. The business impact due to system downtime must be balanced with the cost to deliver database services that meet a given service-level objective.

Finally, IT managers need to make an honest evaluation regarding how much risk they want to take with their active data warehouse, and how much they are willing to pay to protect it from major failures—both likely and unlikely. **T**

Jim Dietz, product marketing manager for the Teradata platform, has more than 12 years of experience developing and managing the Teradata server, storage and BYNET products.

Lynn Hedegard works in the Office of the Chief Technology Officer for Teradata's research and development division. Lynn is involved in the development of technical strategies and architectures for the active data warehouse.



Business and IT leaders must assess the overall cost of an unusable system against the cost of investing in fallback protection.