WHITE PAPER

Data Resilience with Fallback Protection

teradata.

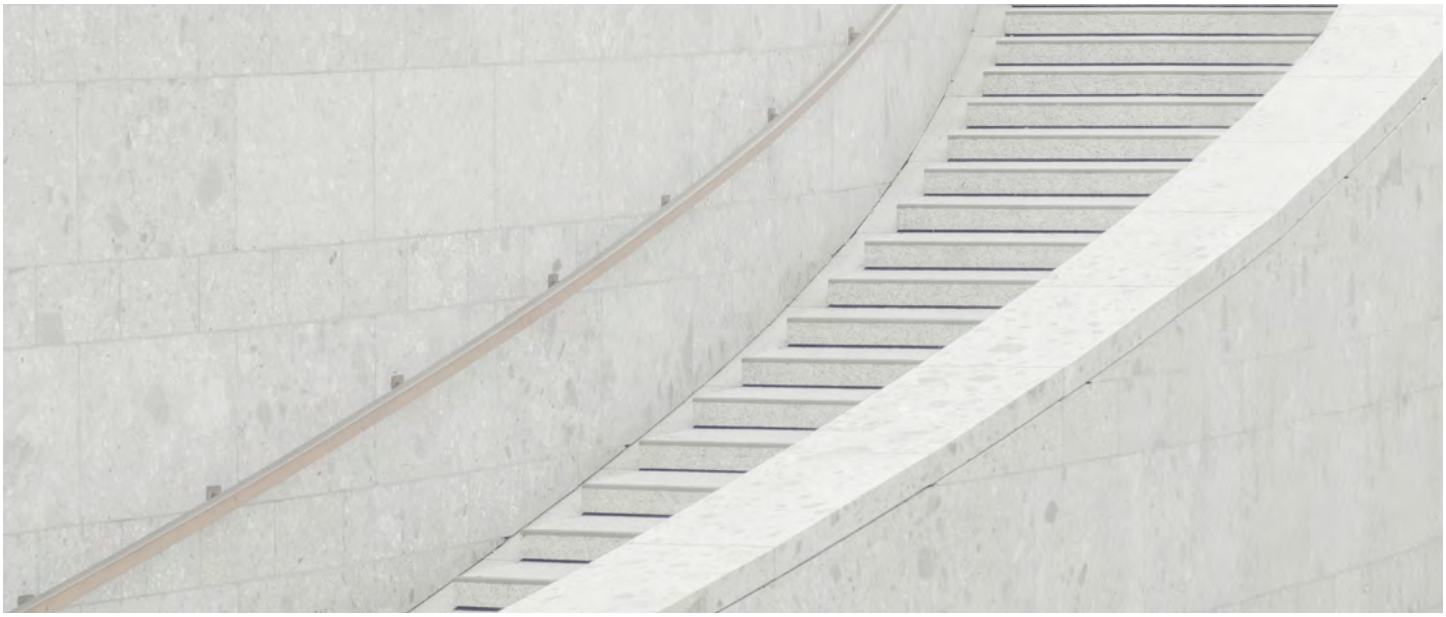
Table of Contents

Executive Overview

Historically, a data warehouse was used primarily in back office business processes. By contrast, today's analytics platform, which includes the integrated data warehouse (IDW), is an operational system used during interactions with customers, suppliers, employees, and business partners. As a result, requirements for higher levels of availability have emerged for the analytics platform. Data Resilience provides enhanced data protection for every Teradata deployment option to ensure your data is always available and your business is protected.

Teradata Vantage offers levels of data availability expected in a world-class operational system—regardless of deployment in the public cloud or private cloud using a software-only architecture, on Teradata Cloud, or on-premises with Teradata IntelliFlex™ or Teradata IntelliBase™. A Teradata system with Data Resilience—using the Fallback feature within the database—provides recovery from myriad failure scenarios, such as the complete failure of RAID devices, disconnection of I/O cables, various software errors, and more. Teradata Fallback achieves levels of single system availability unmatched by other RDBMS vendors. This paper provides a high-level discussion of the features of Fallback that enhance the availability of a single system.

teradata.



Introduction

"Build your house on a solid foundation..."
The same adage applies to building an analytics platform that includes your integrated data warehouse (IDW). In the past, the IDW was used primarily for back office report generation. However, today's IDW is a mission-critical system used during interactions with customers, suppliers, employees, and business partners. As a result, requirements for higher levels of availability and reliability are emerging for analytics platforms.

Teradata Vantage—the platform for Pervasive Data Intelligence that delivers real-time intelligent answers to users and systems across all parts of an organization—offers the levels of data availability and reliability you expect in a world-class operational system. As customers shift towards a hybrid Teradata environment, many customers load and store data in several different places. Some is on-premises in IntelliFlex or IntelliBase; some is on Teradata in the public cloud on Amazon AWS or Microsoft Azure; another portion could be on their own hardware using Teradata Vantage on VMware; and other data could be in Teradata Cloud. This means that customer data can exist in multiple places, including in systems where Teradata does not control the hardware. Yet, our customers still expect their Teradata solution to be 'always-on', because expectations of technology have evolved. Downtime of any system is no longer acceptable—therefore, automated and transparent Data Resilience for on-premises, software-only, and cloud deployments—ensures that data in Teradata is always accessible, regardless of the deployment option(s) chosen.

Critical components on Teradata hardware platforms are monitored by a sophisticated diagnostic subsystem. In the event of a component failure, the hardware subsystem(s) will identify and isolate the faulty component(s) and return the system to normal service with the surviving components. The diagnostic subsystem will generate a report of the faulty component(s), as well as status of the subsystem recovery and reconfiguration to the Teradata Global Support organization. Usually, the only people who become aware of a failure of this kind are the DBA and the field engineers.

The Teradata IntelliFlex platform offers additional levels of availability with its hot standby nodes. Each hardware subsystem—such as power, cooling, compute, interconnect, and storage—is constructed with redundant hardware components. Most single-point component failures are masked by the hardware subsystem without impact to Teradata Vantage, online applications, or user sessions.

To extend the availability of database services beyond the limits offered by traditional hardware redundancy,

Teradata engineers have developed novel software techniques to rapidly recover from failures. Fallback is a feature unique to Teradata Vantage, and is always enabled with our IntelliFlex, IntelliBase, Teradata Cloud, public cloud, and VMware deployment options. It manages redundant copies of data objects on alternate storage subsystems within a single database instance. This secondary copy of the data is referred to as "the Fallback copy of data". Thus, Teradata can minimize the impact of major failure scenarios—such as the complete failure of an entire disk array—a failure that would disable other RDBMS systems.

Fallback Overview

Fallback protection provides enhanced data protection beyond redundant hardware components by keeping a complete second copy of the data on backup AMPs (Teradata unit of parallelism). If data gets corrupted, a disk array fails, two or more drives fail, or a node or AMP goes down, Vantage continues to operate normally using the protected copy of data stored on the secondary AMPs.

Fallback Clusters

With Fallback, copies of data from a given AMP are replicated on one or more Fallback AMPs. A small number of AMPs are logically grouped into a Fallback cluster. Thus, Fallback clusters represent the method by which the redundant copy of data is spread across the AMPs. A system with Fallback will contain many Fallback clusters. See Figure 1.

Typically, each AMP in a Fallback cluster is isolated from failures that impact other AMPs in that cluster. 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 such time as the failed component is repaired. Replicated data from each AMP are typically evenly distributed across the other AMPs in the Fallback cluster. See Figure 2.

Failure Detection and Recovery

Teradata Vantage has a capability called Active Fallback that provides a layer at the database level to handle read errors. When there is a read error on

- A Fallback cluster is a logical set of AMPs and VDISKs.
- Each AMP in a Fallback cluster typically has an equal proportion of its data stored redundantly on the other AMPs' VDISKs.
- Each AMP in a Fallback cluster is isolated in a different domain from the other AMPs in that cluster. This preserves data availability in the event of hardware failures.
- If an AMP cannot access its data (i.e., VDISK), the data remains accessible via the Fallback AMPs.
- One or more Fallback AMPs maintain a log of updates, used to return the repaired AMP/VDISK to the current operational state.

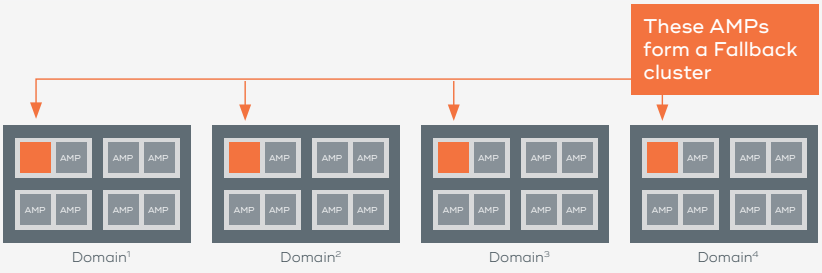


Figure 1. Fallback Clusters

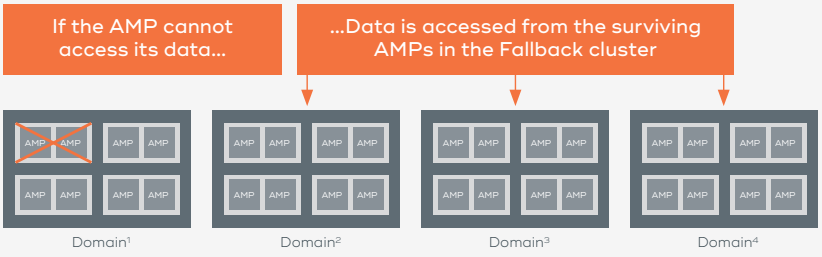


Figure 2. Fallback Recovery

teradata.

the primary data, the Fallback data is automatically retrieved and used to fix the primary data. This is all done internally and transparently to the user.

If, for some other reason, an AMP can no longer access its data or if its data is corrupted, it's flagged as a down AMP. The database system is not concerned with the cause of the failure—which may be failure in the disk, the controller, or the I/O cables. The bottom line is that the AMP can no longer access its data and is deconfigured from the system. A database restart is required to deconfigure any failed AMPs from the system. Teradata Vantage will then resume database operations using the surviving components. In Vantage, the surviving members of the cluster are aware that there is a down AMP and take responsibility for performing all database operations—reads and writes—against its Fallback copy.

Note that Vantage is not functional if there are two or more AMPs down in the same Fallback cluster.

While the Teradata system is running in Fallback mode with a primary AMP down, a journal of all changes made to the Fallback data is maintained in one or more AMPs in the associated Fallback cluster. This architecture ensures that updates to the database can continue without impacting the ETL processes. Since all failed components have been deconfigured from the system, any repair to a physical component can be performed while the database system is operational. When the primary AMP is ready to join other AMPs in the system, its data may be updated from the Fallback journals. If there was a corruption or loss of data in the primary AMP, a rebuild is needed on the primary AMP to restore the data. Whether restoring the data through the Fallback journals or the rebuild, that work can be done while the database system is up and running.

This attribute contributes to the overall uptime of database services. Eventually the physical components will be repaired or replaced and will be ready to resume service to the database. But before they can be reintegrated into the system, their file systems must be brought back to a current and consistent state. This process is described in the following sections.

Fallback Recovery Tools

Teradata Vantage 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 Vantage to normal service after a severe failure scenario.

Table Rebuild

Table Rebuild is a Vantage utility that is used to restore data for a failed AMP with a corruption or loss of data. The entire contents of the failed AMP are rebuilt from data on the Fallback AMP. Table Rebuild can put the contents of an AMP back together while the rest of the system is fully operational.

Without Fallback and Table Rebuild, such a disk failure would necessitate a restore from an external source while the RDBMS is offline. Table Rebuild can also recover a single table. If data corruption of some sort occurs to only one table on one AMP (usually found by CheckTable), Table Rebuild can recover the contents of that table from the Fallback copy with the system online rather than requiring 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 data loss and a lengthy restore procedure—can be managed by a couple of short database restarts when using Vantage with Fallback.

CheckTable

CheckTable is a Vantage 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, they can be rebuilt from its Fallback copy. If Fallback rows are missing, they can be rebuilt from the primary.

Recovery Manager

Recovery Manager is a Vantage utility that allows the user to monitor and interact with the system recovery. When a previously down AMP joins the system configuration, its data may be updated from the Fallback journals. Depending on the type of operations that took place on a table when the AMP was down, the update can involve just the affected rows in the table, as opposed to a complete rebuild.

Recovery Manager can display the status and progress of this down AMP recovery so the user can get an idea of when it is expected to complete. In addition, Recovery Manager can be used to change the priority of this down AMP recovery to speed it up and allow the AMP to be online as soon as possible. Reducing ETL operations on tables undergoing down AMP recovery also helps to speed up the process.

Performance Impacts of Fallback

With Fallback, each row that is stored on a primary AMP is also duplicated and stored on the Fallback AMP(s). Note, however, this is only for user accessible data, such as base tables and materialized indexes. Fallback does not duplicate system-generated data such as spool, temporary tables, and logs. The storage space required on a system with Fallback is basically twice as much user data space as a system without Fallback. Therefore, a system with Fallback consumes more system resources than one without Fallback to handle the additional writes associated with the redundant copy of user data.

Fallback has no effect on read performance if there is no down AMP because read operations are performed by the primary AMP only. Fallback does have some effect on read performance if there is a down AMP. In the down AMP case, the other AMPs in the Fallback cluster have to retrieve their primary data as well as the Fallback data for the down AMP.

The performance of write operations, excluding spool, is affected by Fallback even if there is no down AMP. This is because every write must be performed twice—once for the primary data and a second time for the Fallback copy. The performance impact will vary by customer and workload. As a general rule for normal system operation, the performance impact is minimal, as most data warehouse workloads are 80-90% reads and 10-20% writes. During load windows, this will shift to more write intensive.

Because Fallback protection is a built-in, always-on feature in every new Teradata system, the CPU, I/O, and capacity requirements are already factored in. It is transparent to all users (IT, administrators, application developers, users, etc.) and applications. It operates in the background, silently protecting your data, without ever requiring outside attention to do so.

The Cost of Downtime vs. the Cost of High Availability

Fallback improves the availability of database services. In order to understand the value of Fallback, you must weigh the cost of downtime versus the cost of the

Fallback is Data Insurance

When homeowners buy a house, they usually purchase insurance against disasters, such as floods, fires, and earthquakes. The odds of these extreme events occurring are low, but the costs involved in restoring one's home are extremely high should the event occur. Most people cannot tolerate the thought of losing everything should such a disaster occur—they're willing to pay insurance premiums to mitigate risk.

The "insurance premium" of Fallback is included in every Teradata system, so you will not lose your data should a failure occur.

The proper business decision is to deploy a system where the system cost is balanced with the cost of downtime. Choosing a system on either extreme of the spectrum is a poor business decision. Deploying an excessively 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.

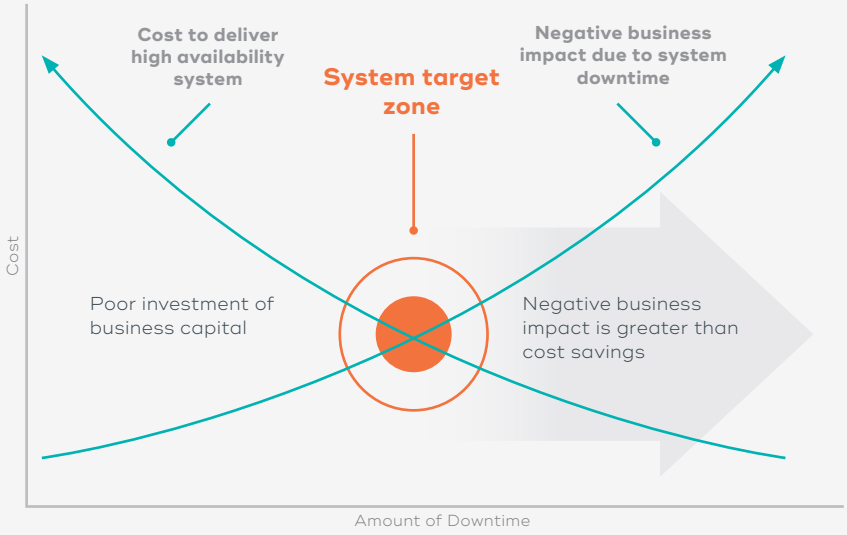


Figure 3.

additional system components necessary to achieve a desired level of availability.

Cost of Downtime

Data used during interactions with customers, suppliers, employees, and business partners are valuable assets. These end users invoke a variety of application services to perform their daily activities, and they expect that these services will always be available. By proxy, they expect the underlying database services will always be available.

When database services become unavailable, various application services become unavailable—which in turn, creates a financial impact to 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—such as the case when a disgruntled customer takes 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. 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 dimension involves the time to recover from a given failure scenario. This is referred to as the Recovery Time Objective (RTO). RTO is very 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. Fallback-enabled Teradata systems can achieve extremely favorable RTO metrics for single systems because major failure scenarios can be recovered with a system restart, which takes three to ten minutes to complete.

Analysis of historical data corruption events indicates the recovery time for non-Fallback systems can be from multiple hours to days, depending on access to back-ups; whereas, outage to customers with Fallback can be measured in minutes. The same analysis indicates a full recovery time is a few hours with some degraded performance until the Teradata systems are 100

teradata.

percent online. One proven strategy to shorten the recovery time is to minimize ETL processing.

Cost of High Availability

To sustain a given workload, a Fallback-enabled system will require additional compute nodes and disk sub-systems than a system without Fallback. However, this additional hardware is already factored into the system and its cost will balance the cost of system downtime.

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

A recent Teradata Customer Services study found that extended downtimes, those over four hours, were reduced by 87% when Fallback was enabled.

Disaster Recovery

Fallback should not be considered a disaster protection solution. Disasters brought about by tornadoes, hurricanes, floods, fires, or malicious individuals, have the potential to destroy an entire data center. If that occurs, both the primary and Fallback copy of data will most likely be destroyed.

Fallback is not guaranteed to be 100 percent effective even without a disaster. For example, it does not have a 100 percent separate path for the data. Most of the database code above the level of the file system is common, as are all of the components outside the database, such as the network and the application. Fallback does not protect from failure of these components of the total system. However, Fallback can significantly reduce the probability of data loss.

If the enterprise places a high enough value on the analytic services of the Teradata solution, then a disaster recovery plan must be considered. The first step is a robust Teradata Backup and Restore (BAR) solution for every Teradata system. This can be at the same physical location, or for even more protection, at a second geographically distinct data center with a redundant Teradata analytics platform. Another alternative is to secure the services of a data recovery center that can be used in the event of a disaster at the primary data center. Again, the RTO metric needs to be analyzed and weighed against the time required to reestablish IDW services.

Conclusion

The analytics platform, which includes the integrated data warehouse, is an operational system used daily during interactions with customers, suppliers, employees, and business partners. It provides mission-critical analytic services for tactical and strategic decision making. Data Resilience using the Fallback feature protects your data wherever it is.

To optimize your return on investment, you should consider protection strategies that ensure the availability of database services in the face of severe component failures.

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 needs to be balanced with the cost to deliver database services that meet a given service level objective.

Teradata Vantage provides a wide range of availability features, including Fallback. Fallback enables the highest metrics for data availability and RTO possible for a single system and improves the perceived availability of the database by reducing the overall downtime and improving recovery time objectives.

Teradata has raised the bar for data resilience and data availability for the analytics platform, including your mission-critical data warehouse. Teradata Vantage with Fallback ensures that the Teradata solution is always protected, whether on-premises, in a managed cloud, on customer hardware with VMware, or in the public cloud.

Case Study

A large retailer implemented Fallback on all production tables to enable continued processing (online and batch) during AMP down condition. They temporarily dropped Fallback on many large tables to overcome disk space issues for the period of six to nine months prior to their next planned upgrade.

They then experienced a data corruption event that impacted all non-Fallback tables. Tables were then recreated from data sources, dumps, and tapes.

However, critical applications were unavailable for three days, which created a business impact prolonged by the time needed to process the back log. Annual system availability decreased by .245 percent, which amounted to an additional two hours of downtime each month (not all additional downtime can be attributed to the decreased use of Fallback).

The company has since re-implemented Fallback on the upgraded system for all production tables, and annual availability measurements have returned to previous levels.

About Teradata

Teradata leverages all of the data, all of the time, so you can analyze anything, deploy anywhere, and deliver analytics that matter. By providing answers to the complexity, cost and inadequacy of today's analytics, Teradata is transforming how businesses work and people live. Get the answer at **Teradata.com**.

17095 Via Del Campo, San Diego, CA 92127 Teradata.com

teradata.