+ a b | e a u ®
SOFTWARE

Ensuring High Availability and Reducing Downtime with Tableau Server

Business Intelligence Is Becoming Mission-Critical

In the past, business intelligence (BI) was often used and managed in departmental settings, where it was not considered mission critical. As BI has become more widespread in organizations, it has also become an integral part of the business and must support required business continuity service levels.

Users and decision makers have come to depend on immediate access to data and information to understand quickly changing market and customer trends, and respond with strategic and tactical adjustments to stay on top of the competition.

Most executives understand the mission-critical importance of business intelligence at their companies. For example, a study done by Ventana Research shows that 77 percent of respondents feel business intelligence technologies are now critical to their sales teams.[i]

This reliance on data requires a high degree of availability from the supporting BI systems. However, configuring enterprise systems for high availability has traditionally been a complex process. With the growing importance of business intelligence, this capability needs to be more accessible and easily configurable by organizations.

Making High Availability Accessible

At Tableau, we know how much users have come to rely on their ability to quickly and readily access and visualize data in order to make timely business decisions. We also realize there will always be events that will threaten the availability of business intelligence systems, whether these events are hardware, software, network or even people related. Beginning with version 7.0.4, Tableau Server has the capability to automatically failover in the event of a failure. This is accomplished by replicating critical components: the Data Engine and the Repository.

Within Tableau Server, the Data Engine and Repository hold the data extracts and the metadata, including workbooks, projects, views, user settings, data source connections. They are the moving parts that are constantly


being updated as users work in Tableau. They are also the parts at most risk in the event of a failure.

Tableau Server can be easily configured to replicate the data engine and repository on a second node. Tableau Server can also automatically send an e-mail notification to the administrator in the case of an error.
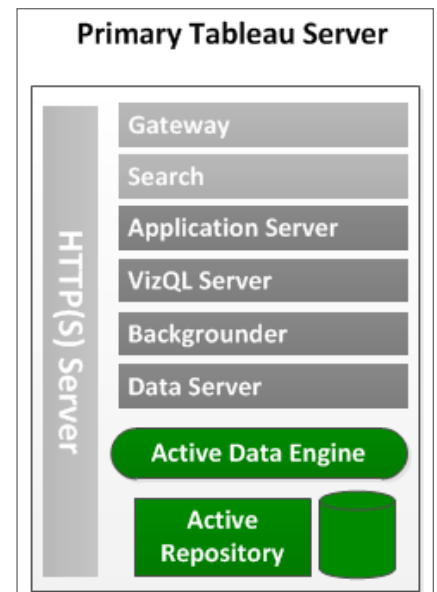
Understanding High Availability with Tableau Server

Although Tableau can run on just one server, this configuration does leave the system at risk for potential downtime in the event of a failure. To reduce the risk of downtime, Tableau Server can be configured for automatic failover, with little or no impact to the users.
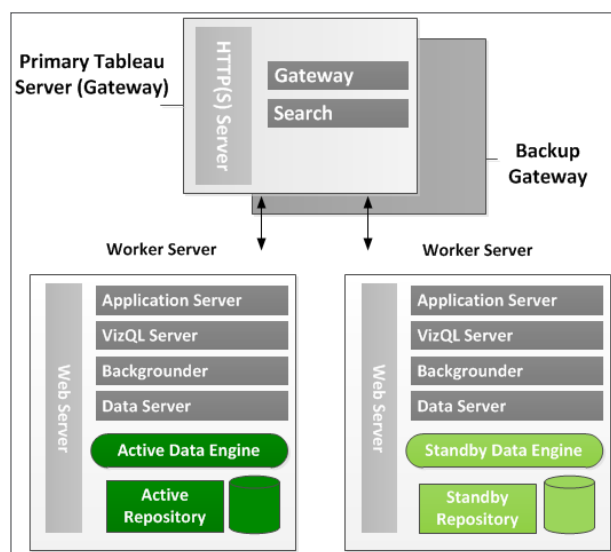
At a minimum for high availability, the Tableau Server cluster should consist of three nodes. One node will serve as the Primary Tableau Server running as the Gateway, while the other two nodes are Worker servers, hosting the core Tableau components. While Tableau Server can be configured to run on more than three nodes for additional scalability, the high availability function will only be able to replicate data from one active Data Engine and Repository to one standby Data Engine and Repository.

The Gateway server will load balance the user requests between the Worker servers. The Gateway can be installed on a small, one-core server with four gigabytes of RAM. Since only one Gateway is allowed per Tableau cluster, the Gateway should run on a dedicated server, running no other processes. This reduces the server load and reduces the risk of potential failures. Please note that currently, external load balancers are not supported with Tableau Server.

The two additional servers will run the core Tableau processes. These servers can be set up for high availability by configuring both servers to host the Data Engine and Repository processes. The Tableau Server cluster will automatically designate one of these servers as the host for the 'active' Data Engine and Repository, while the other server is designated as the host for the 'standby' Data Engine and Repository. The data will be replicated across the two servers in near real-time.

From a configuration standpoint, it is important to note that because both workers will be processing requests, the secondary server should have adequate processing power to handle the full load of the cluster itself, should the active server fail. Also, because the system automatically designates the active and standby Data Engines and Repositories, it's advisable that both servers be similar in capacity and configuration.

When the system starts, replication will begin across the nodes. The standby components will be ready after the content has been synchronized across the nodes. The time to replicate is dependent on the size of the current repository and data extracts.

The Gateway server is another potential single point of failure. If the Gateway does fail, Tableau Server will not be accessible at all. To reduce the risk of downtime due to Gateway failure, we advise installing a backup Gateway on an additional node. Currently, the failover to the secondary Gateway requires a manual process. This process is well documented and can be done easily and quickly.

Maintenance

☑ Service waiting for request ☑ Service standing by ☑ Service handling request ■ Service unlicensed ☒ Service down

Status

Machine	Repository	Data Engine	Server Web Application	VizQL Server	Data Server	Background Tasks	Web Server
123.45.67.89							Gateway
123.45.67.88	☑	☑	☑	☑	☑	☑	☑
123.45.67.87	☑	☑	☑	☑	☑	☑	☑

How Tableau's Automatic Failover Works

Failover setup in Tableau is straight-forward and easy. Once Tableau Server is installed, both workers are identified and configured. Two nodes are configured to host the Data Engine and Repository in order to ensure replication. Then, just set up the list of e-mail addresses that will receive notifications if there is a system failure. Refer to the *Real-Time Replication & Failover Support* document for detailed steps and diagrams.

When a critical incident happens that causes the active Data Engine and Repository to be unavailable, Tableau Server automatically fails-over from the active Data Engine and Repository to the standby Data Engine and Repository. At the same time, an e-mail alert goes out to the recipients identified during setup notifying them of the failure.

The failover process should take just a few minutes. There is a built-in delay to account for issues that may resolve themselves, such as momentary network issues. This time period is configurable to meet specific service-level agreements. During this short time, Tableau Server will not be able to respond to user requests until the failover process has completed. When issues with the original worker are resolved and it comes back online, it will be automatically designated as the host for the standby Data Engine and Repository. Another e-mail notification will be sent when the system is available again.

The primary Gateway can also be switched to a backup Gateway, should it fail for any reason. Although the steps are manual, they are very simple. After just a

couple command-line and Gateway file changes, the backup Gateway is ready be engaged.

System Administrators can monitor the state of the Tableau Server processes in the Administration section of Tableau Server. This provides administrators with an understanding of the active and passive components in the system as well as whether any components have issues.

Beyond High Availability: Ensuring Resiliency in Your Organization

Achieving high availability requires more than features and functionality built into BI systems. True high availability comes from understanding what users need and following best practices and processes to meet those needs. Although Tableau makes high availability easier to achieve with its failover features, it should not replace best practices which include tasks such as regular maintenance and backups.

Although many organizations create redundancy through geographically distributed deployments across different data center sites, we don't recommend this approach for Tableau Server. The resulting network latency will adversely impact users.

Tableau realizes the importance of business intelligence applications to help your business make critical business decisions and it is our intention to increase the confidence and resiliency of your Tableau system. Users can be confident that they can continue to analyze and make critical business decisions whenever they need to.

End notes

[i] http://www.ventanaresearch.com/resources/resources.aspx?id=2469

About Tableau

Tableau Software helps people see and understand data. Ranked by Gartner and IDC in 2011 as the world's fastest growing business intelligence company, Tableau helps anyone quickly and easily analyze, visualize and share information. More than 7,000 companies get rapid results with Tableau in the office and on-the-go. And tens of thousands of people use Tableau Public to share data in their blogs and websites. See how Tableau can help you by downloading the free trial at *www.tableausoftware.com/trial*.