



Disaster Recovery for Tableau Server

Fatima Calcutta, Senior Product Manager, Tableau

Introduction

Today, self-service analytics are helping to bring data-driven decision-making to more people across the organization. Employees and executives are using analytics that help them discover relevant data and answer questions in real time. Tableau Server is a leader in providing mission-critical self-service analytics. It enables rapid self-service data exploration, promotes trust in content and data through robust governance, and is easy to deploy, manage, and scale across any enterprise.

As analytics are often mission-critical, Tableau Server must run reliably to support the enterprise—even during unexpected events that can compromise the platform’s availability. The [Tableau Server High Availability whitepaper](#) describes how you can configure fault tolerance and redundancy within the server to ensure Tableau Server is always available to the user. In this paper, we will explore how an enterprise can prepare and plan for disaster should Tableau Server go down.

Please note: This paper assumes you are running Tableau Server version 2018.2 or more recent (with Tableau Services Manager).

Contents

- Understanding Disaster Recovery** 3
- Disaster Recovery for Tableau Server** 4
 - Understanding Backups 4
- Disaster Recovery Planning** 6
 - Recovery Point Objective 6
 - Recovery Time Objective 6
 - Low RTO (RTO of a few minutes) 6
 - Medium RTO (RTO of a few hours) 8
 - High RTO (RTO of several hours or days) 9
- Additional Considerations** 9
- Summary** 9
- About Tableau** 10
- Additional Resources** 10

Understanding Disaster Recovery

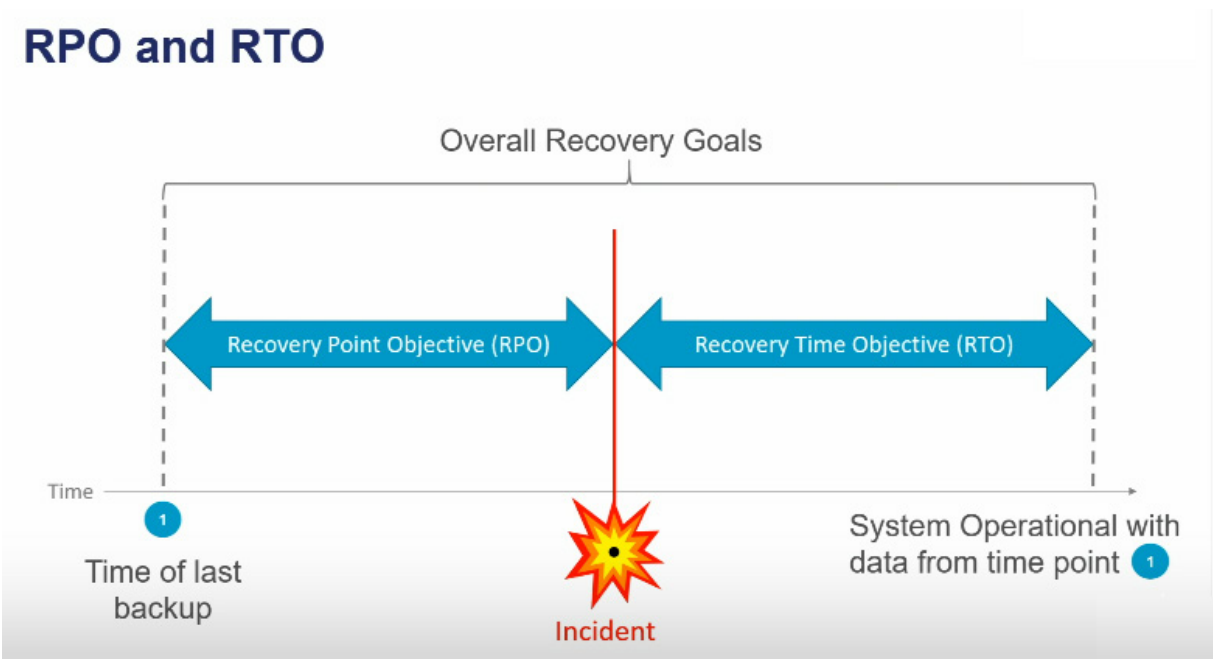
Disaster recovery (DR) means being prepared for—and recovering from—disaster. Any event such as hardware failure, software failure, natural calamity, human error, or infrastructure failure that causes disruption to business continuity can be considered a disaster.

When planning for disaster recovery, there are two main factors to consider:

- 1. Your business tolerance for downtime as part of the recovery process
- 2. Your business tolerance for lost data

These are defined in terms of:

- 1. **Recovery Time Objective (RTO)** — A measure of how much time can elapse before full recovery; that is, the maximum length of time within which a business process must be restored after a disruption.
- 2. **Recovery Point Objective (RPO)** — A measure of tolerance of data loss in terms of time, i.e. the duration of time for which data loss is acceptable.



When creating your disaster recovery plan, it's important to determine RTO and RPO thresholds. To do this, you need to understand the tradeoffs. It's costly to support a very low RTO when you consider the infrastructure and staff that is necessary to bring the system back online rapidly. If you choose to spend less on the infrastructure and responsiveness of staff, you should be able to tolerate much longer RTO (i.e. longer downtime). If Tableau Server is a business-critical application, any downtime can impact the business. Investing more in a DR solution would be beneficial to ensure minimal impact to the business.

Disaster Recovery for Tableau Server

Understanding Backups

A Tableau Server backup includes:

1. **Tableau Data** — This includes user and workbook metadata that is stored in the repository and data extract file. A backup allows you to move this data from your production server to DR server.
2. **Configuration and Topology Data** — Configuration data includes most of the server configuration information like SMTP setting, certificates, authentication settings, etc. Topology data includes how processes are configured on your deployment. A backup of configuration and topology data is used for setting up your DR server.

Tableau Data

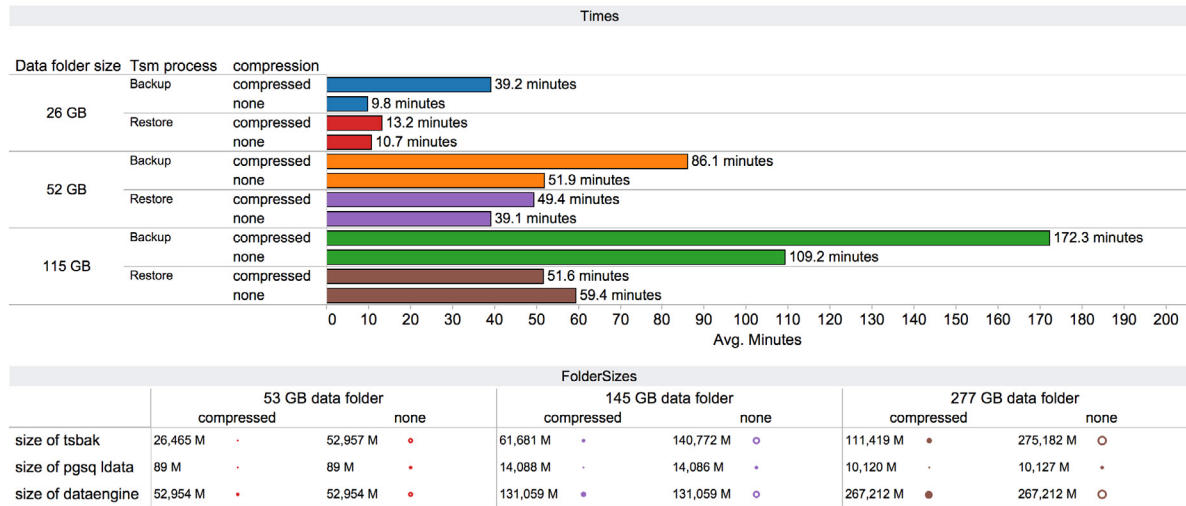
A backup of Tableau data is generated by the *tsm maintenance backup* command. The command creates a backup by saving data in a single file with .tsback extension. The backup file is generated and saved on the initial node of Tableau Server.

The size of the backup file depends on the amount of data in Tableau Server. Before generating a backup, you should always ensure there is enough disk space on the node to save the backup file. The time taken to generate your backup largely depends on the amount of data and CPU and memory resources available on the machine.

If you have a large amount of data, the backup can take a long time. Here are couple of recommendations to improve your backup speed:

1. Configure a File Store instance on the initial node where the backup command runs.
2. Run the *tsm maintenance backup* command with `--skip-compression` option. This option can reduce the amount of time a backup takes by up to 2x but results in a larger backup file because it does not use compression. See [tsm maintenance backup command help](#) for more details.

Here are some general figures for backup size and time with and without using the `--skip-compression` option.



After the backup file is generated, move the file to another data center or save it in the cloud. If the backup is stored on the Tableau Server computer, it is not protected from any disasters that impact Tableau. Backup files contain lot of sensitive data. We recommend you secure every backup file by storing it in a location which has limited access. If needed, you can encrypt the backup files for additional security.

Configuration and Topology Data

You create a backup of configuration and topology data using the `tsm settings export` command. The data is exported in a JSON file. These settings are used to configure your DR server. Some configurations are not included when settings are exported using the `tsm settings export` command—these settings need to be documented and manually restored. See online help to learn how to [backup assets that require a manual step](#).

Similar to .tsbak file, we recommend that you save the JSON file on a separate data center or cloud so you have access to them in case of a disaster. Whenever you modify or apply new configuration or topology changes, you should re-export the configuration and topology data.

You can learn more about Tableau Server [Backup and Restore](#) process in our online help documentation.

Disaster Recovery Planning

When planning for disaster recovery you need to factor in RPO and RTO that your business can tolerate. Ideally, you want to minimize both RPO and RTO metrics in case of a disaster.

Recovery Point Objective

RPO is defined by how often you take backups. You can choose to take a backup once a week, once a day, or every few hours. The more often you take a backup, the lower your RPO will be. Your minimum RPO will be limited by the length of time it takes to back up your Tableau Server installation. If a backup takes eight hours, that is your smallest RPO.

You can reduce the amount of time it takes to generate the backup and the size of the backup by running tsm cleanup command, `tsm maintenance cleanup --http-requests-table`. This will clean up old data in repository and will ensure your backup only includes the data that is needed. See [tsm maintenance cleanup command help](#).

Tableau Server often ends up with stale content, including workbooks and data sources that are no longer used. Administrators should regularly audit content for usage and clean up anything that is not used or maintained according to internal policy. Doing this helps reduce the amount of data you need to back up, and that reduces the time it takes to complete a backup. [Here](#) are some suggestions on how you can identify and clean up stale data from your server.

In most cases taking a backup at least once a day is enough. However, if Tableau is business critical, we recommend taking a backup more often.

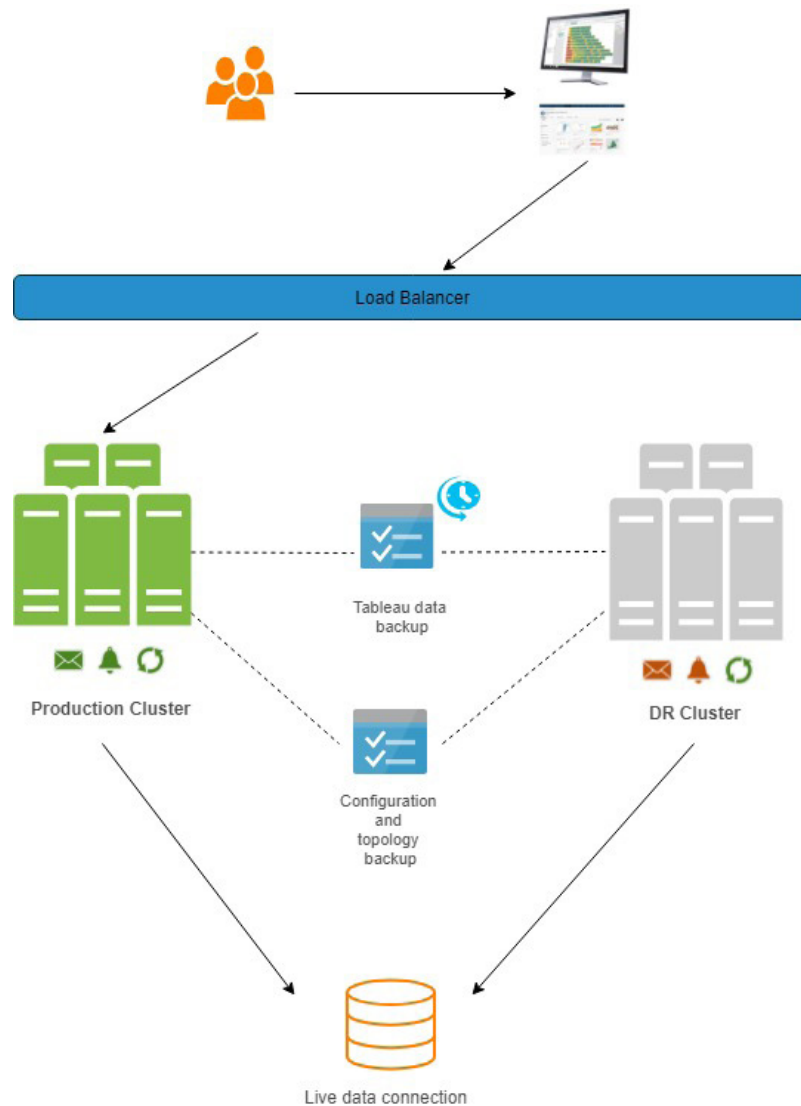
Recovery Time Objective

Based on the RTO tolerance your business has, here are some suggestions on how to plan and prepare for Disaster Recovery.



Low RTO (RTO of a few minutes)

If your business considers Tableau Server a mission-critical system which can't go down and you have a very low tolerance to Tableau Server not functioning, we recommend that you have dedicated hardware for Tableau Server provisioned in a separate DR environment. This instance of Tableau Server should run the same version and have the same configuration as your production Tableau Server cluster. The DR server should be regularly refreshed with backups from production and updates any time the configuration or topology changes in your production environment. The diagram below shows how you can set up a dedicated DR Server.



Your production cluster and your DR Tableau Server should be running in two separate data centers. Configure an external load balancer which will control where your user traffic will be redirected. On a regular day, user traffic should go to your production cluster.

Both the clusters should always be pointing to live data source(s). While your DR cluster is running and not receiving any user traffic, turn off subscriptions and alerts to ensure that the users do not get duplicate notifications.

As part of your disaster recovery plan, generate regular Tableau Data backups either nightly or as often as you can for the production cluster. Restore this backup on the DR cluster immediately. Any configuration or topology changes made to the production cluster should be applied to DR cluster to ensure it has the same settings as the production environment.

In case of a disaster, your DR plan will include the following steps:

1. Point your external load balancer to DR Tableau Server cluster.
2. Turn on subscriptions on from the Site's settings page for each site on the server cluster.
3. Turn on alerts on from the Site's settings page for each site on server cluster.

In this scenario, the duration of the outage will be a few minutes, with the data being as up-to-date as the last backup that was restored to the DR cluster.

Having a dedicated DR cluster which gets refreshed with the latest backup daily is the recommended method of having a robust disaster recovery solution for Tableau Server.

Medium RTO (RTO of a few hours)

If your business permits medium RTO, meaning it can tolerate Tableau Server being unavailable for a few hours, we recommended you have hardware provisioned for Tableau Server before you need it.

As part of the setup of your DR cluster:

1. Install the same version of Tableau Server that is running on your production cluster.
2. Import the latest configuration and topology data on the DR cluster.
3. Restore the latest Tableau Data backup on the DR cluster.
4. Power down all the nodes on DR Server.

Your DR Tableau Server will now be in a cold standby state. In case of a disaster, your disaster recovery plan will include the following steps:

1. Turn the cold standby server on.
2. Restore the latest configuration and topology data on the DR cluster.
3. Restore the latest Tableau Data backup on the DR cluster.
4. Redirect traffic to the DR Tableau Server.

In this scenario, duration of the outage will depend on the size of your Tableau Server installation and length of time it takes to restore the most recent backup. Your restored data will be as up-to-date as your most recent backup.

High RTO (RTO of several hours or days)

If your business permits high RTO, meaning it can tolerate Tableau Server being unavailable for several hours to several days, this approach may be applicable to you. In case of disaster, your recovery plan will include the following steps:

1. Acquire hardware needed to run Tableau Server.
2. Install the same version of Tableau Server that is running on production Tableau Server.
3. Restore the latest configuration and topology data on the DR cluster.
4. Restore the latest Tableau Data on the DR cluster.
5. Redirect traffic to the new Tableau Server.

In this scenario, the duration of your outage will depend on the length of time it takes to acquire the hardware, install Tableau Server, and restore your most recent backup, combined with the complexity of your production Tableau configuration. Your restored data will be as up-to-date as your most recent backup.

Additional Considerations

For many enterprises, disaster recovery policies also impose distance requirements between the data center where the primary server is running and the data center that would be the recovery target (e.g. in the United States, DR sites must be at least 250 miles apart). In these situations, Tableau can be treated as any other business-critical application being managed for disaster recovery.

For enterprises that deploy to a public cloud, you'll want to provision machines for DR Cluster in a separate region than your production environment. Provisioning your DR cluster in a separate availability zone (AZ) within the same region doesn't provide true Disaster Recovery, because if a region goes down, all of the AZs in that region will be impacted.

For enterprises that deploy Tableau Server in a public cloud, procuring machines is not a barrier. These enterprises can spin up a new machine in cloud on-demand, install Tableau Server and restore the latest backup and config quickly in the event of a disaster—and this entire process can be automated.

Summary

Tableau's built-in backup and recovery technologies can be used to meet your DR needs. This whitepaper highlights how you can optimize your Disaster Recovery plan to meet your business RPO and RTO goals. A well-defined Disaster Recovery plan will prepare you and your enterprise for a disaster. Along with a well-defined plan it is important that you test the plan on a regular basis to be able to act quickly and precisely during a disaster to have minimal impact to your business.

About Tableau

Tableau is a complete, easy-to-use, enterprise-ready visual business intelligence platform that helps people see and understand data through rapid-fire, self-service analytics at scale. Whether on-premises or in the cloud, on Windows or Linux, Tableau leverages your existing technology investments and scales with you as your data environment shifts and grows. Unleash the power of your most valuable assets: your data and your people.

Additional resources

[Online Help: Distributed and High Availability Tableau Server Installations](#)

[Whitepaper: High Availability and Disaster Recovery in Tableau Server](#)

[Sample script: Run Tableau Server Backup](#)

[Sample script: Run Tableau Server Restore](#)

