

TopBraid Data Platform

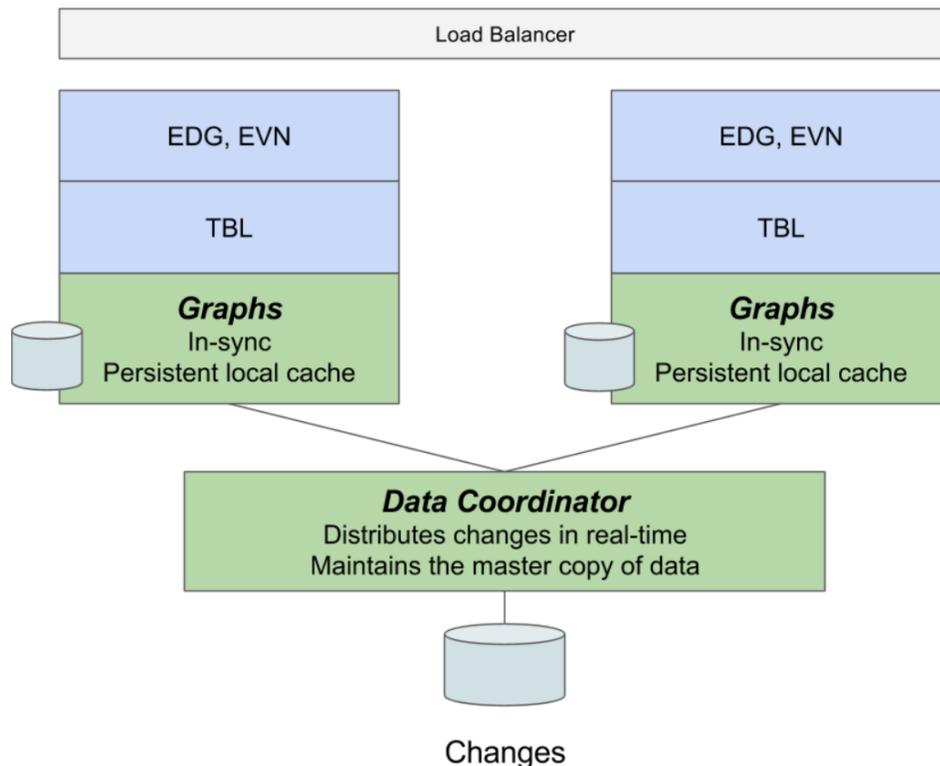
Page Contents

- 1 TopBraid Data Platform
 - 1.1 Configuration
 - 1.2 Creating Shared (DP-managed) Graphs
 - 1.3 Data Coordinator Outages
 - 1.4 Data Coordinator Requirements

TopBraid Data Platform

The TopBraid Data Platform is a high availability solution for TopBraid EDG servers. It enables continuous operation of business functions by replicating data across a cluster of EDG servers, keeping them all updated so that any one of them can respond to a client/application request. Together with a load balancer to direct requests, user and application access to the EDG data remains available even if some of the servers are offline. Note that the Data Platform feature is experimental in TBS v5.4.0.

This availability is achieved by having a cluster of EDG servers communicating with a *data coordinator* (DC) server that receives changes from any clustered server whenever it has updates. The DC server propagates the changes to all other EDG servers, keeping them all up-to-date. Additionally, whenever a new EDG server joins the cluster, it is first brought up-to-date before it starts servicing user/application requests. Each EDG server keeps a local, persistent cache of application data so that start-up only involves applying changes missed by that server.



Configuration

For each clustered EDG server, its Data Platform feature is enabled via [Server Administration > EDG Configuration Parameters > Data Platform Configuration Parameters](#).

Creating Shared (DP-managed) Graphs

When the DP feature is enabled for an EDG server and the data coordinator server is running, creation of new asset collections will show a Data Platform checkbox option, whose selection will cause the new collection to be shared among all DC-connected EDG servers.

Create New Asset Collection

This creates a new Asset Collection with yourself as the manager.

Label	<input type="text"/>
Default namespace	<input type="text" value="http://example.org/namespace-1234567890/new#"/>
Description	<input type="text"/>
Data Platform	<input type="checkbox"/> <i>Server-installed option to share this graph across all platform servers.</i>

There may be a few-second delay in propagating collection changes. The DP checkbox option only appears when the DC server is contactable. Asset collections created without the DP option will exist only on the original server.

Data Coordinator Outages

If the DC is not running then:

- shared collections can be read (viewed, queried), but
- updates to shared collections are not possible, and
- new, shared collections cannot be created.

Restarting the DC re-enables the update and sharing features.

Data Coordinator Requirements

The DC server is a Java webserver process that coordinates changes across a cluster of EDG servers. For production use, it should be run on a machine separated from all of the clustered EDG machines. For development and experimentation, it can be run on a shared machine.

The DC server stores patches on-disk. The disk storage must be backed-up, which can be done by file backup on a live server. The robustness of the system is determined by the robustness of the file storage so choosing local disk in the DC-server is limiting. It is primarily a patch server for patch files and is not computationally demanding.