

STARBURST ENTERPRISE FOR PRESTO ON RED HAT OPENSIFT

Available on the Red Hat marketplace, the Starburst Enterprise for Presto platform is a fully supported, production-tested, enterprise-grade distribution of the open source Presto MPP SQL query engine. Starburst integrates the scalable cloud storage and computing services provided by Red Hat OpenShift with a more stable, secure, efficient, and cost-effective way to query all your enterprise data, wherever it resides.

Leading organizations across multiple industries rely on Starburst Enterprise for Presto and Red Hat.

Analytics Anywhere

The traditional data warehouse model breeds inefficiency and keeps business analysts from running fast analytics on their data, potentially delaying critical insights. Starburst Enterprise for Presto gives analysts the freedom to interrogate diverse data sets wherever they are located, without building separate data warehouses. Organizations can run multiple Presto clusters, scaling up or down dynamically and optimizing for query speed and cost as desired.

This unique functionality lets organizations access multiple software-defined data storage platforms deployed with Red Hat OpenShift Container Platform. For instance, analysts can instantly and simultaneously query data in a Red Hat Ceph Storage data lake, a SQL or NoSQL database running on Red Hat OpenShift Container Storage, and many other environments.



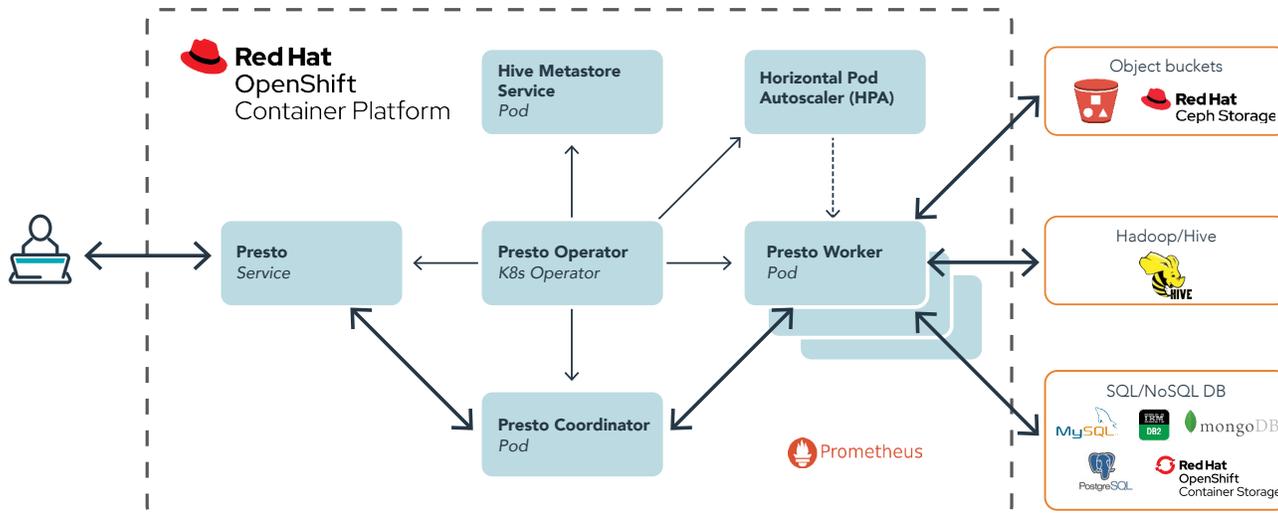
Starburst Enterprise Presto on Red Hat OpenShift

Many organizations rely on a combination of traditional and modern applications to run operations and make critical business decisions. Starburst Enterprise provides a modern solution built on the open source Presto distributed SQL query engine. It harnesses the value of open source Presto, the fastest distributed query engine available today, while adding multiple enterprise-critical features. Presto was designed and written for interactive analytic queries against data sources of all sizes, ranging from gigabytes to petabytes. Starburst Enterprise Presto adds the enterprise features and 24x7 support that organizations need for big data access at scale.

The Starburst Enterprise for Presto platform provides distributed query support for varied data sources including:

- NoSQL systems (MongoDB, Cassandra, Redis)
- SQL databases (Microsoft SQL Server, MySQL, PostgreSQL)
- Data warehouses (IBM DB2 Warehouse, Teradata, Oracle Exadata, Snowflake)
- Hive (HDFS, Cloudera, MapR)
- Data Services (Kafka, Elasticsearch, OpenShift Container Storage)
- Cloud Object Storage (AWS S3, ADLS, Azure Blob, Ceph, IBM COS)

Starburst Enterprise Presto on OpenShift



Combining Starburst Enterprise Presto with Red Hat OpenShift Container Platform offers automation, high availability, elasticity, and monitoring for Presto clusters. Together the benefits include:

Automation.

Starburst and Red Hat OpenShift operators provide auto-configuration, auto-tuning, and auto-management of Starburst Enterprise Presto clusters. The Red Hat OpenShift operator determines what to deploy, including identifying the hardware and provisioning new instances. The Starburst operator manages updates to the environment.

High availability.

Continuous operation of the Presto Coordinator is essential to delivering business value. Using liveness probes, the Red Hat OpenShift load balancer can keep services like the Presto Coordinator in an always-on state.



Elastic scalability.

Red Hat OpenShift can automatically scale the Presto worker cluster based on query load. Using the Kubernetes Horizontal Pod Autoscaler (HPA), organizations can specify thresholds for Presto worker pods. As the number of queries increases, the HPA will automatically spin up additional Presto worker pods based on specified system constraints.



Graceful scale-down and decommissioning.

With Red Hat OpenShift, reduced load doesn't mean system downtime or killed queries. The Kubernetes HPA will gracefully decommission unused Presto worker pods and free system resources for other tasks without an interruption in service.



Monitoring for all hardware and software layers.

Prometheus, OpenShift's cluster monitoring service, delivers metrics and alerts that inform Kubernetes orchestration and populate the Red Hat OpenShift dashboard. Prometheus informs Kubernetes if a pod is down, and otherwise provides metrics to the HPA to let it know whether to commission or decommission additional Presto pods.



Support for Red Hat Storage.

Starburst Enterprise Presto lets organizations make use of data storage platforms associated with Red Hat OpenShift. Application developers can make use of SQL and NoSQL databases backed by Red Hat OpenShift Container Storage. Businesses can pull data from data lakes running on Red Hat Ceph Storage archives. Parquet, ORC, and text files on Ceph are also useful for data analytics.

Common Enterprise Use Cases

Organizations are using Starburst Enterprise for Presto on Red Hat in a variety of ways, but the following use cases are the most popular.

✓ **Data Lake Query Engine**

Presto has become the de facto standard data lake query engine. A Starburst deployment enhances its functionality for the enterprise with role-based access control, autoscaling, high concurrency, ANSI SQL compatibility, and other benefits.

✓ **Data modernization**

Starburst Enterprise for Presto lets you modernize data at your own pace, even as you work with the environment you have. Organizations can update, migrate, and move data as it makes sense for the business—without forced data migrations.

✓ **ETL workloads**

Starburst Enterprise for Presto is ANSI SQL compliant for support of create table, and insert statements. It can act as the SQL engine for ETL jobs, providing a single platform for both query and migration needs. For example, archive data from an Apache Hadoop cluster could be moved to a data lake on Red Hat Ceph Storage, allowing federated

Presto queries against that data as well as data from other sources that are not ready for migration.

✓ **Interactive data investigation**

Starburst Enterprise for Presto enables rapid ad-hoc interactive queries from a range of data sources—including traditional, real-time, object stores, and so on. DBAs can query underlying sources from their SQL or business intelligence tools of choice. Data can be queried rapidly from a single source, or combined through federated joins.

✓ **Business intelligence dashboarding and reporting**

Data consumers can work with their favorite BI tool of choice, such as Tableau, Microstrategy, or Qlik for dashboarding and reporting. Because Starburst Enterprise for Presto separates compute and storage resources, it provides the interactive responsiveness that these tools require.

BUSINESS OUTCOMES

Overall, enterprises and large organizations that combine Starburst with their Red Hat OpenShift investment benefit from:

→ Shorter Time to Value

→ Increased User Adoption

→ Improved Price vs. Performance

→ Enhanced Security



Starburst Enterprise for Presto and Red Hat OpenShift Container Platform generate better, faster insights by letting organizations rapidly analyze data across multiple disparate and distributed data platforms. The combination provides critical automation, high availability, elasticity, and monitoring that meets the demands of large enterprises across industries.

Unlock the value of your data with Starburst and Red Hat today.

For more information, contact us at starburstdata.com



STARBURSTDATA.COM