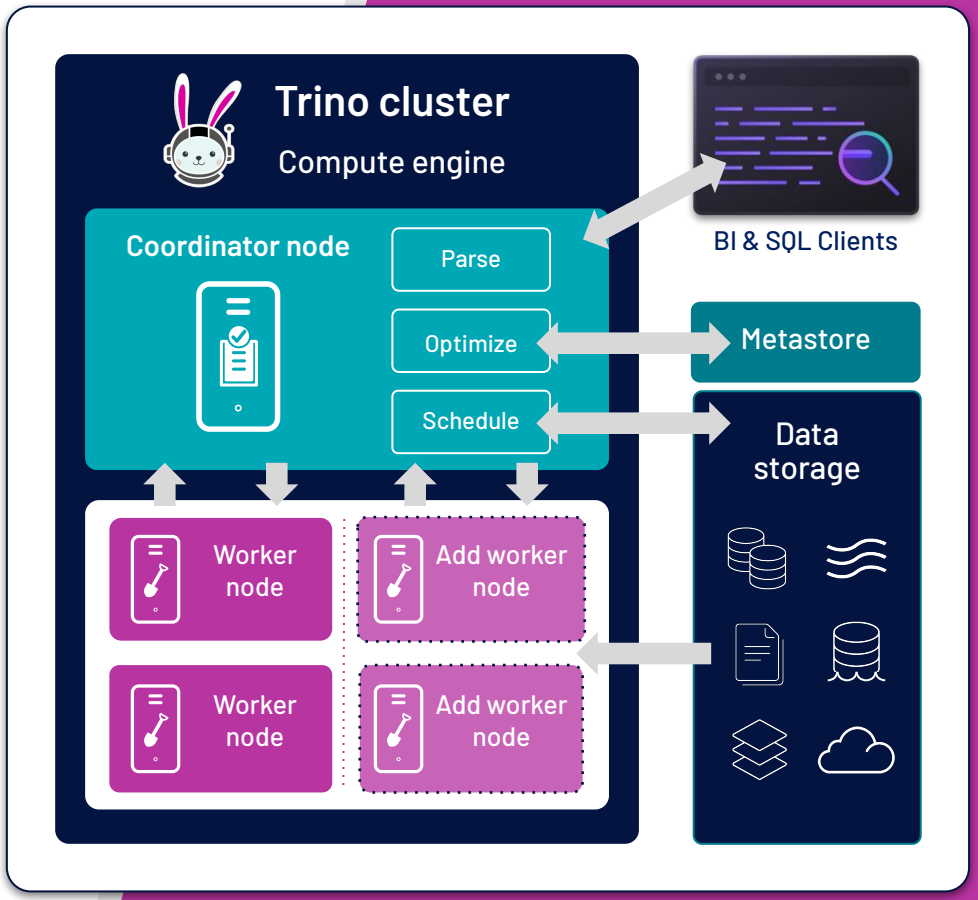
Starburst

Starburst Galaxy Architecture

Technical Overview for Starburst Galaxy

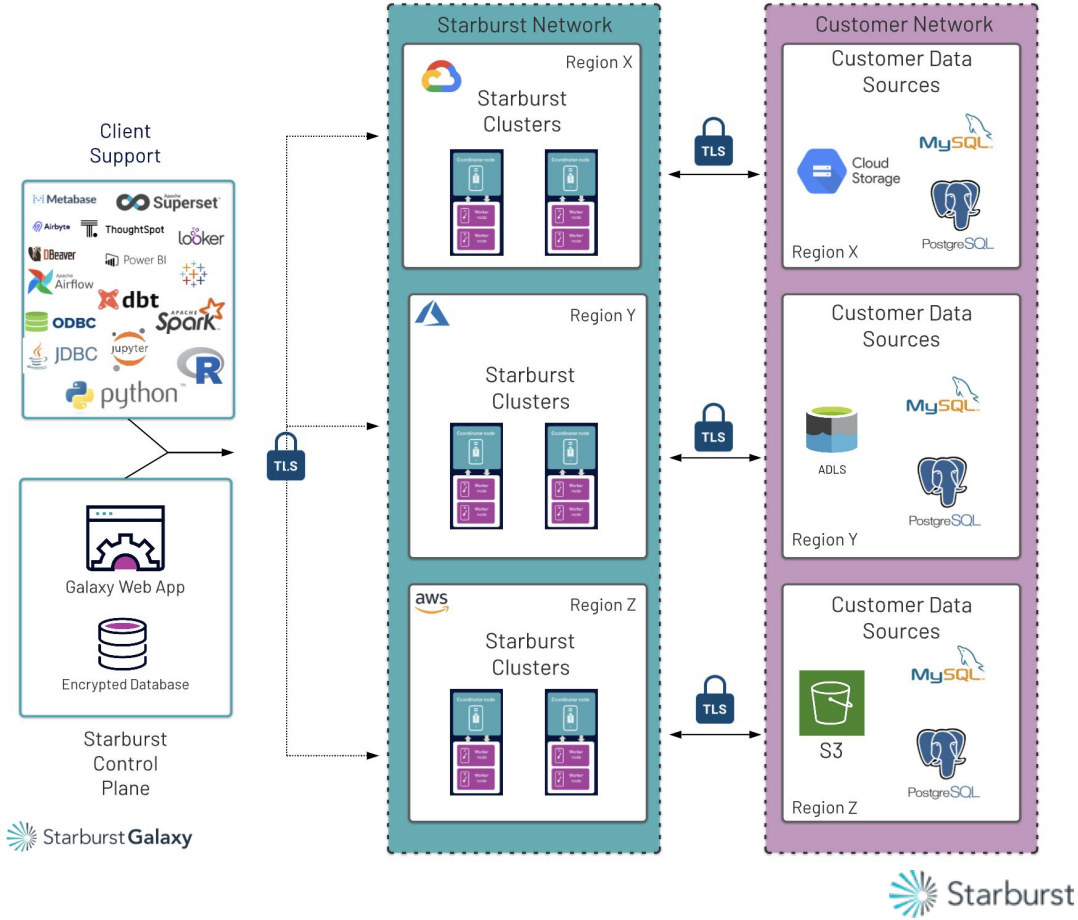
An architecture for today and tomorrow

- True separation of storage and compute
- Horizontally scales to meet demand by adding worker nodes and additional clusters
- Cost-based optimizer automatically picks the join order with the lowest computed costs
- Graceful shutdown API ensures worker nodes terminate without affecting running queries
- Cluster scheduling and autoscaling to optimize compute utilization
- Strict quality control on all Trino contributions

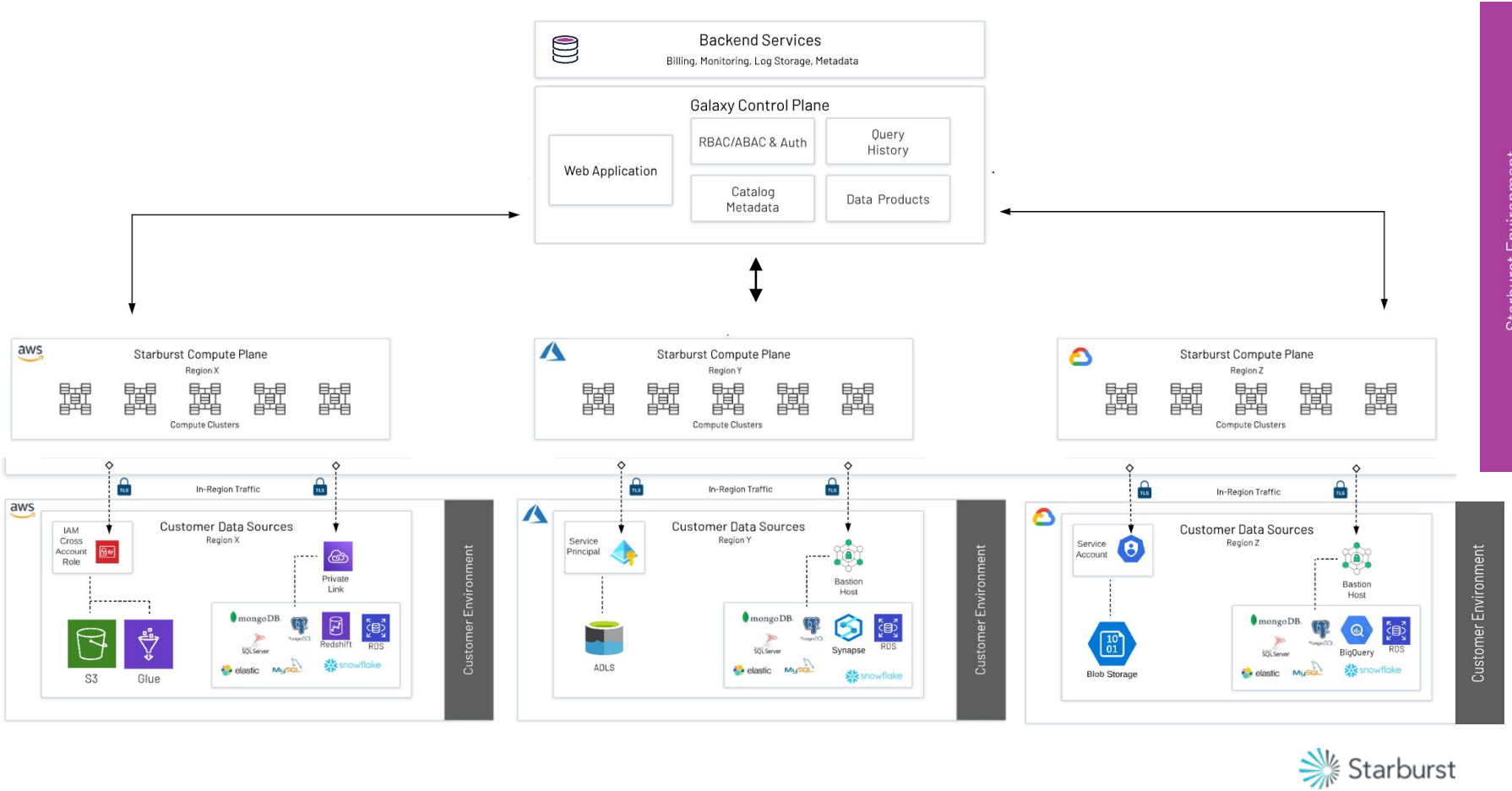


Starburst Galaxy - Architecture

- Compute in Starburst Network
- Allows for suspension and faster deployment of clusters
- Clusters deployed in region where the data lives
- Blue/Green - seamless to end users for changes/upgrades
- Deploy via control plane to all 3 clouds



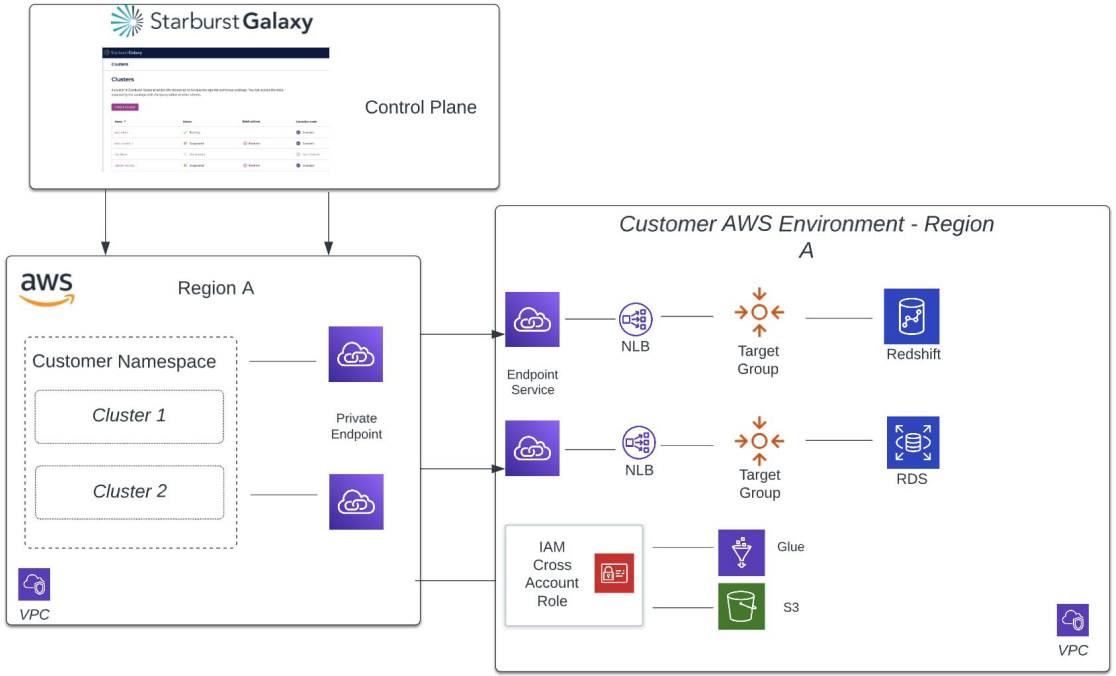
Starburst Galaxy - Architecture



Starburst

Catalog Auth Methods - Private Link

- Direct Private Endpoint connection between Galaxy Clusters and Customers Data Sources
- Ensures secure access without exposing data source to the public internet
- A more secure alternative to using a bastion host

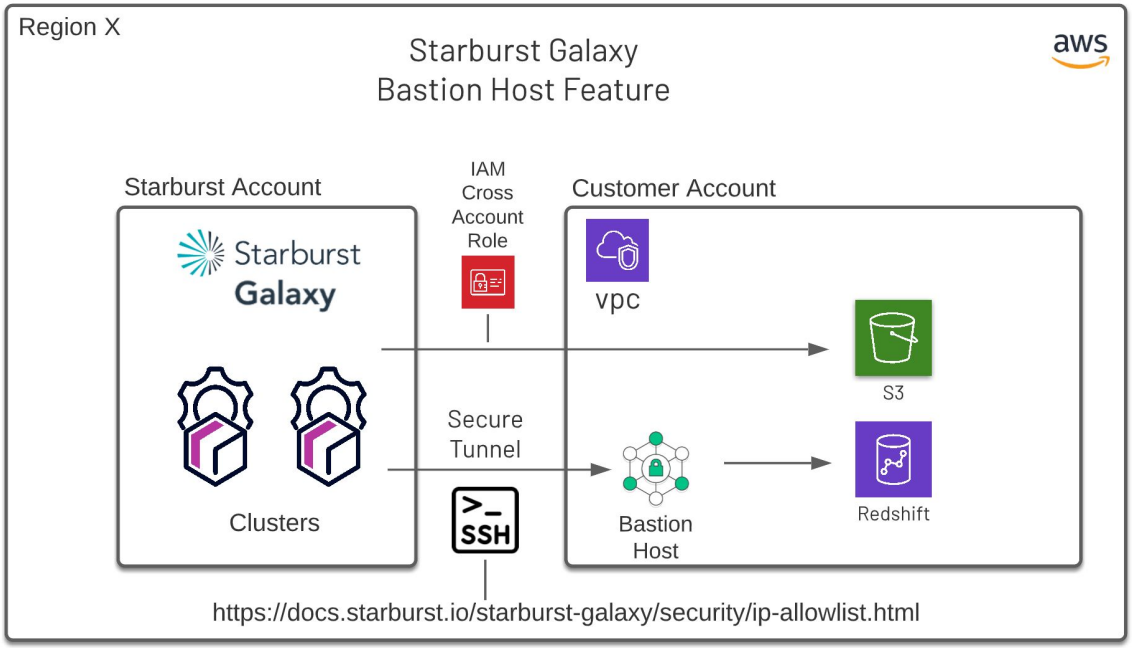


Docs: https://docs.starburst.io/starburst-galaxy/security/privatelink.html

Starburst

Catalog Auth Methods - Bastion Host

- Bastion host running Linux and ssh
- Allow list for Galaxy applied to Bastion host
- Bastion host can be used to access multiple data sources



Starburst