



GridSQL Overview - JPUG Cluster Meeting

November 19, 2009

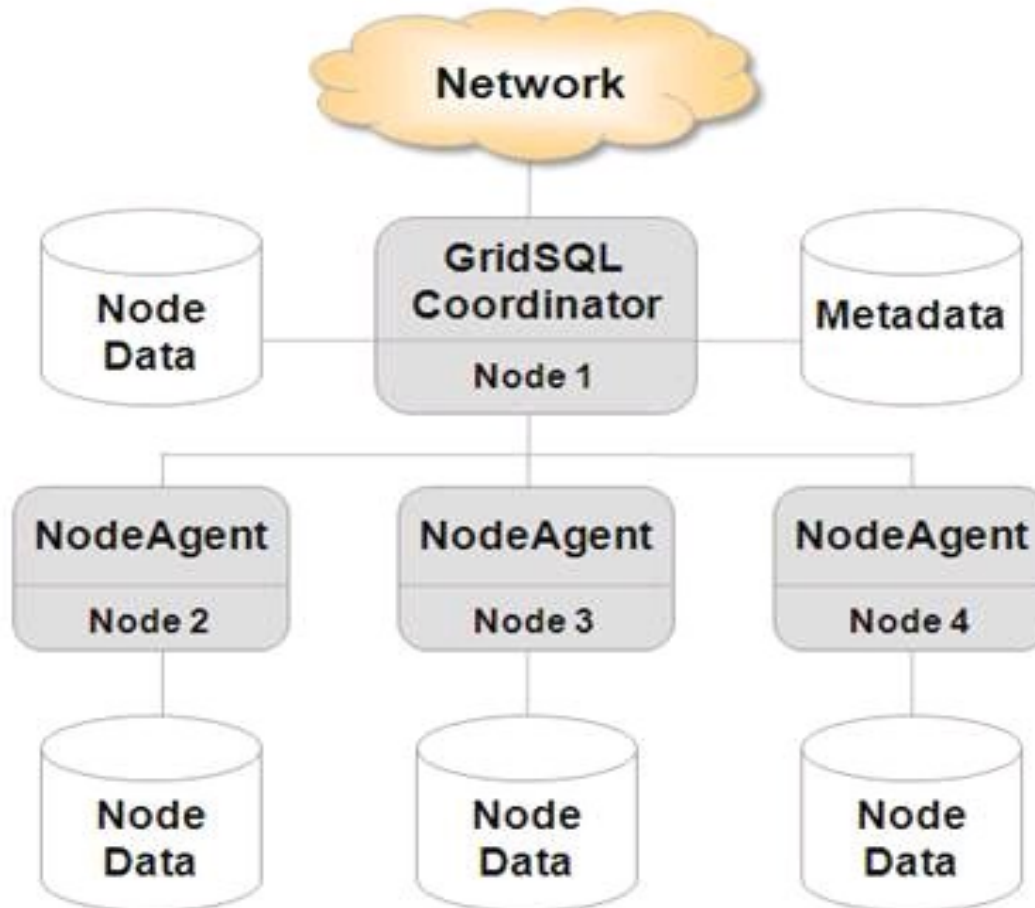
Overview

- Designed for Data Warehousing - parallel querying
- Appears as a single database to the application
- Utilizes PostgreSQL – but entirely external
- Data Loader for parallel loading
- Not just “Read-Only”, can execute UPDATE, DELETE, transactions
- Not just for star schema, can do cross-node joins
- Standard connectivity via PostgreSQL compatible connectors (supports PostgreSQL protocol): JDBC, ODBC, ADO.NET

General Design

- Outside of PostgreSQL as separate layer
- Shared-nothing architecture
- Command line tools
- End user apps should be mindful of architecture, as in any Data Warehousing solution

GridSQL



More Details

- No native HA
- Tables may be replicated, handled synchronously
- Intended to be used in same geographic location
- No direct stored procedure support, but can use external database and dblink

Scalability

- Good for reading large tables
 - Also, can set up multiple logical nodes per physical server (multi-core, multi-processor)
- Good for loading large tables
- Not good for OLTP or high concurrent reads

Status

- Being used at EnterpriseDB customers
 - Web log file analysis, billing records, data analysis
 - Several TB range
 - One customer loads 1 billion rows per day into GridSQL
- Small community



Thank you

mason.sharp@enterprisedb.com