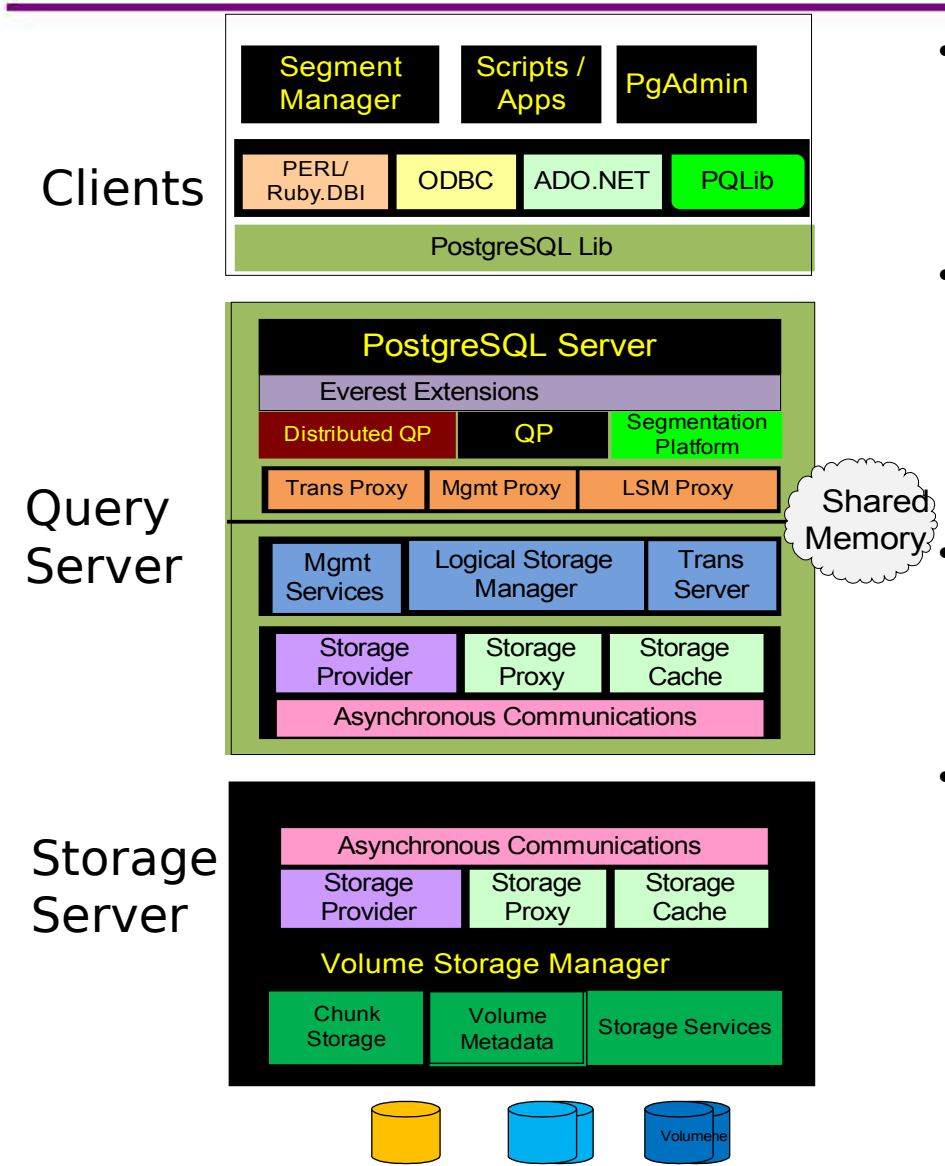
Everest

Scaling to Petabytes

Yahoo!
May 2008

Everest Architecture



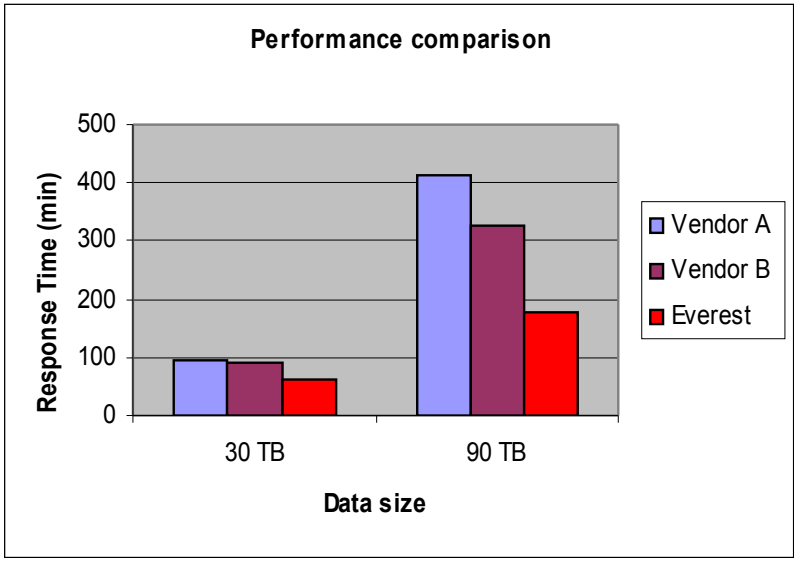
- **Massively Parallel (Tens of PB)**
  - Commodity Clusters
  - Multi-tier scalability
  - Distributed Columnar Storage
- **Smart**
  - Optimized compression
  - Parallel Vector Query Processing
  - Query and Storage optimizations
  - Query Expression and Columnar caching
- **Leverage PostgreSQL**
  - Tools and Connectivity (ODBC)
  - extensibility
  - UDF & UDAF framework
- **Inexpensive**
  - COTS

Performance and Scale

- Proven Petabytes scale in production
  - Approaching 2 PB, projected to grow > 30 PB by 2009
  - Largest table: 3.5 Trillion rows (time partitioned)

- 10x Price-Performance relative to commercial systems

Data size	Everest (min)	Vendor A (min)	Vendor B (min)
90 TB (600 B rows)	177	414	325
30 TB (200 B rows)	60	95	91
HW Cost (1 PB)	250	1200	1200



Everest Performance Advantages

- Source of Performance and Scale
  - Distributed Compressed Columnar Storage
  - Highly Parallel and Asynchronous
    - Multi threaded Query Execution as well as Storage
  - Vector Query Processing
  - Multi-level data partitioning and query partitioning
  - Cluster-level Compressed Columnar caching
  - Query expression caching
  - Yahoo! specific language extensions and UDF & UDAF