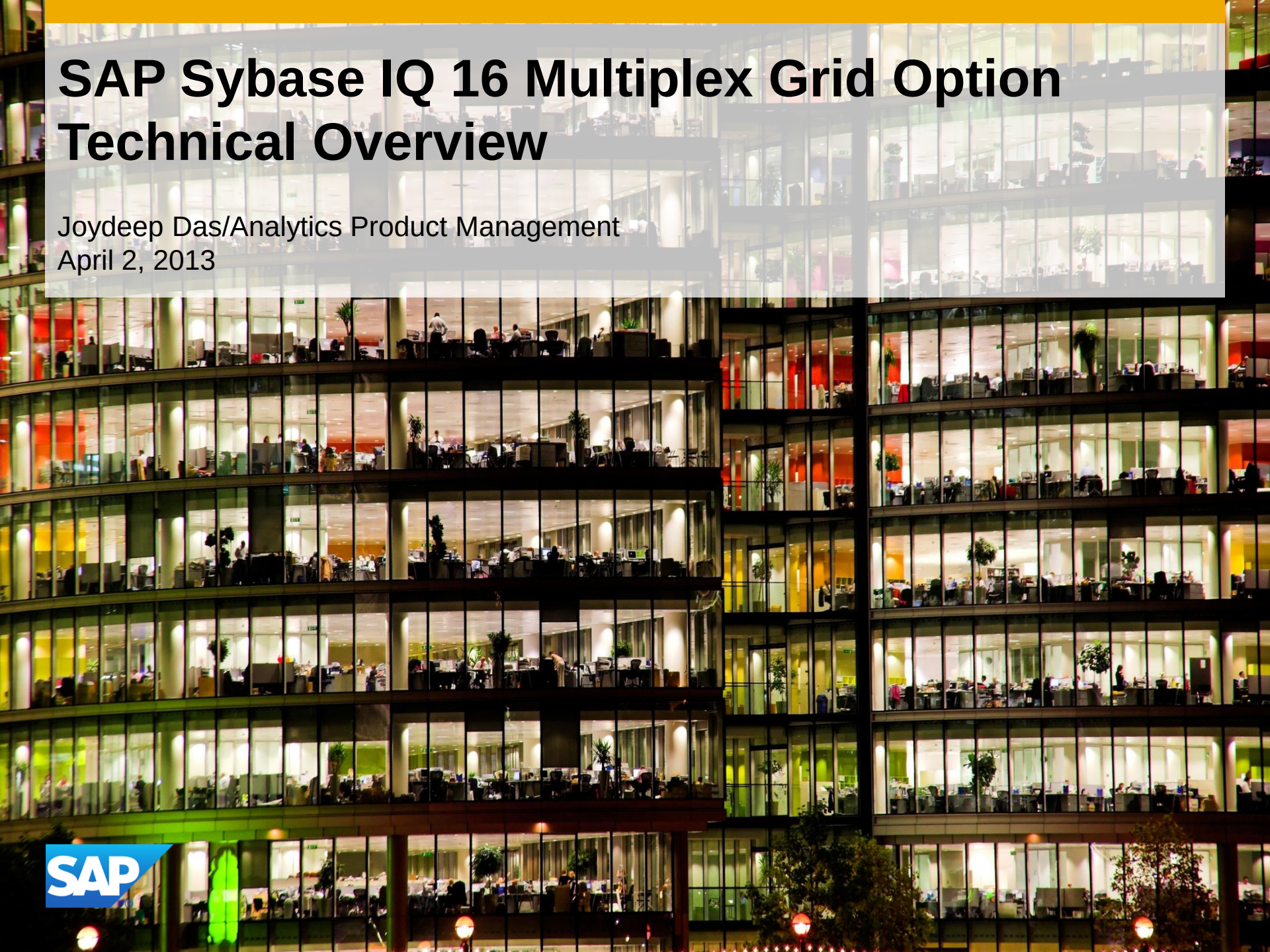
SAP Sybase IQ 16 Multiplex Grid Option Technical Overview

Joydeep Das/Analytics Product Management
April 2, 2013



Agenda

What's Happening in the Marketplace

Product Success

SAP Sybase IQ 16

Enterprise Data Warehousing Considerations

Multiplex Grid Option Details

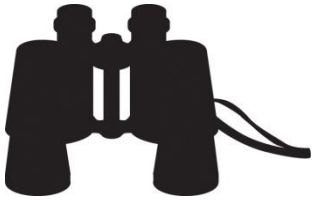
Summary

Marketplace Today

What's Happening in the Marketplace…

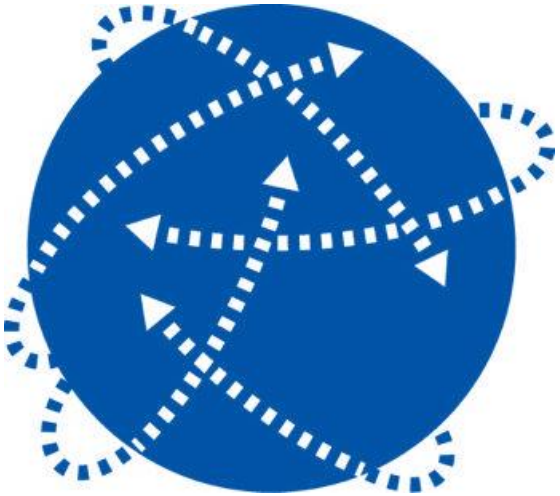
Exploding Data Volumes | **The Need for Speed** | **Rising IT Cost and Complexity**

Challenges customer face today?



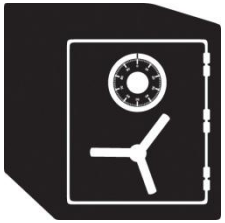
Lost revenues due to lack of insight

High Costs & Complexities



Slow Performance

Data Management Challenges

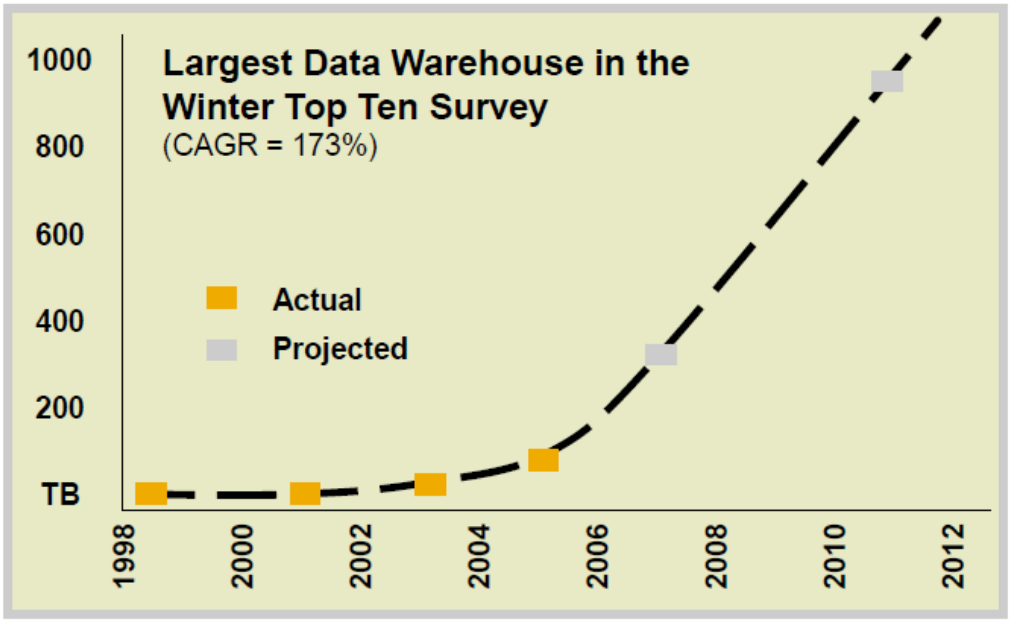


Security

SAP Sybase IQ 16 Motivators

"Petabyte is the new Terabyte" - Forbes

The data explosion continues: Data volumes in analytics environments are growing exponentially…



Source: *WinterCorp*

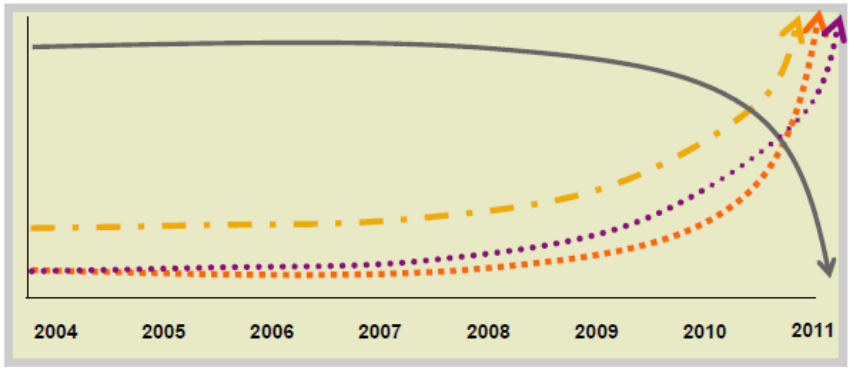
Meanwhile, what trends are you seeing…

Number of business decisions supported by data?

Number of users accessing analytics environments?

Required response time?

Resources required to maintain performance?



Product Success

SAP Sybase IQ: Market Leader for Extreme-Scale EDW and Analytics

High performance analytics server

- Columnar RDBMS (stores data in columns- versus rows)
- Optimized for managing and accessing massive amounts of data for analytics (versus transactions)

Accelerates analytics and reporting

- Up to 1000-times faster than traditional transactional databases
- Handles structured and unstructured data
- High compression and low TCO
- Highly scalable grid architecture

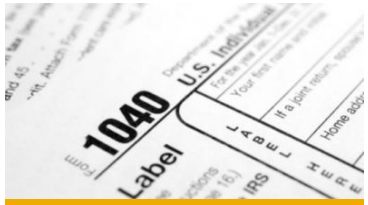
SAP Sybase IQ Facts

- 2200+ customers with over 4500+ installations worldwide
- Used by twice as many companies as the next leading provider
- Patented data compression dramatically reduces data storage requirement; cuts TCO
- Only column-based solution to support full text search, in-database analytics, and federated analytics
- 96%+ customer satisfaction rates
- Leader, 2013 Gartner Magic Quadrant for Data Warehouse DBMS

SAP Sybase IQ big data analytics

Pervasive across data intensive industries worldwide

Manage and analyze statistical measures for the entire nation of Canada

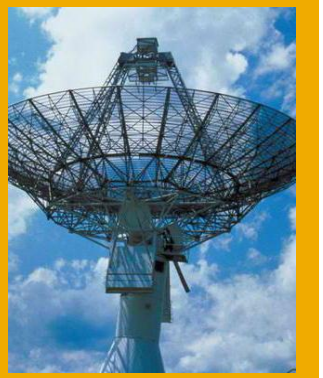


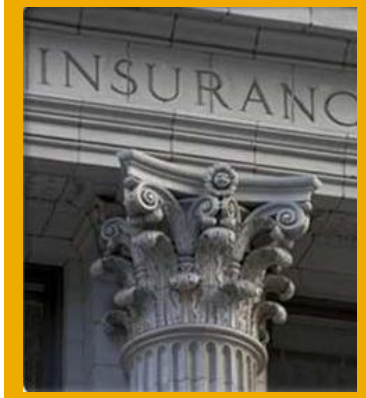
Analyze ALL Federal tax returns in the US

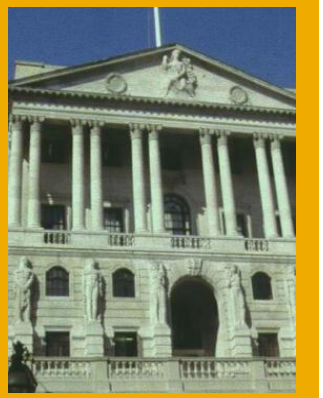
Analyze complex models in more than 200 financial institutions worldwide

Store and analyze massive amounts of industry segment data in 30 of the largest information providers in the world, including Transunion, Nielsen and Axiom

Stands out as the leading enterprise data warehouse among the largest banks, insurance agencies, and telecom operators worldwide







SAP Sybase IQ 16

Solution Overview – SAP Sybase IQ 16

SAP Sybase IQ transforms the way companies compete and win through actionable intelligence delivered at the speed of business to more people and processes.

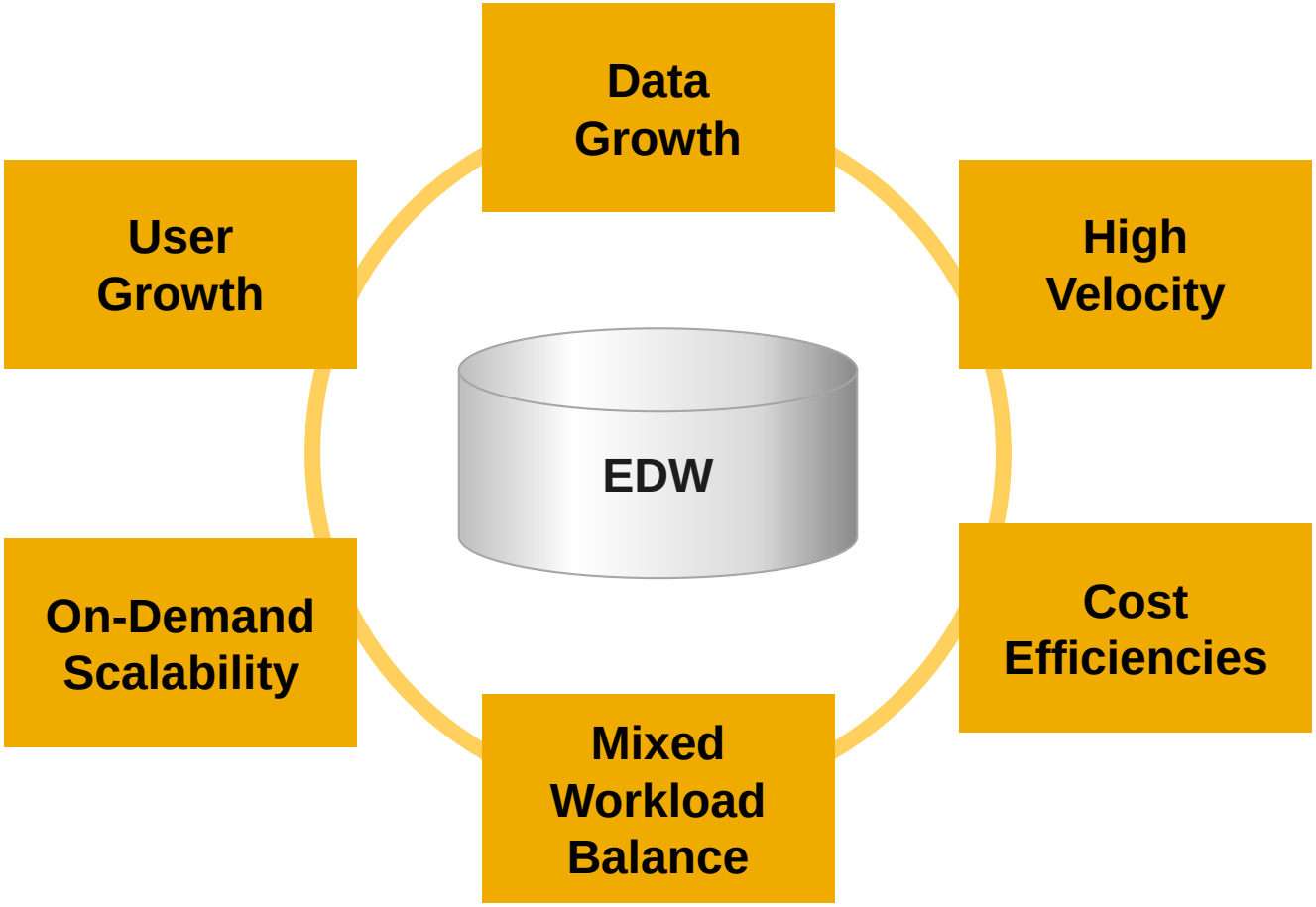
Value of SAP Sybase IQ 16

1. **Exploits the value of Big Data**
2. **Transforms businesses through deeper insights**
3. **Extends the power of analytics across the entire enterprise**

Enterprise Data Warehousing Considerations

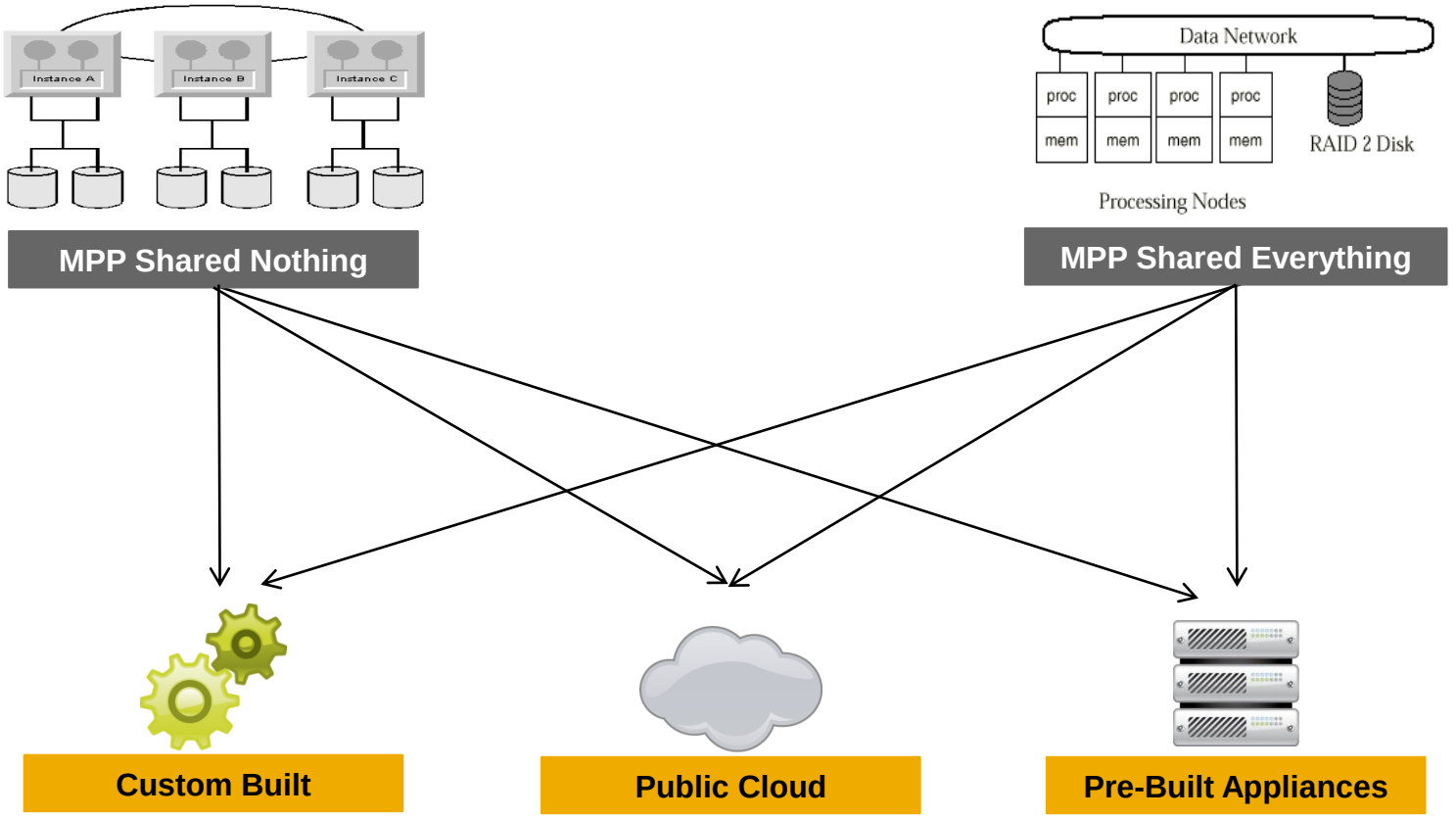
Enterprise data warehousing

Infrastructure challenges



EDW architectures

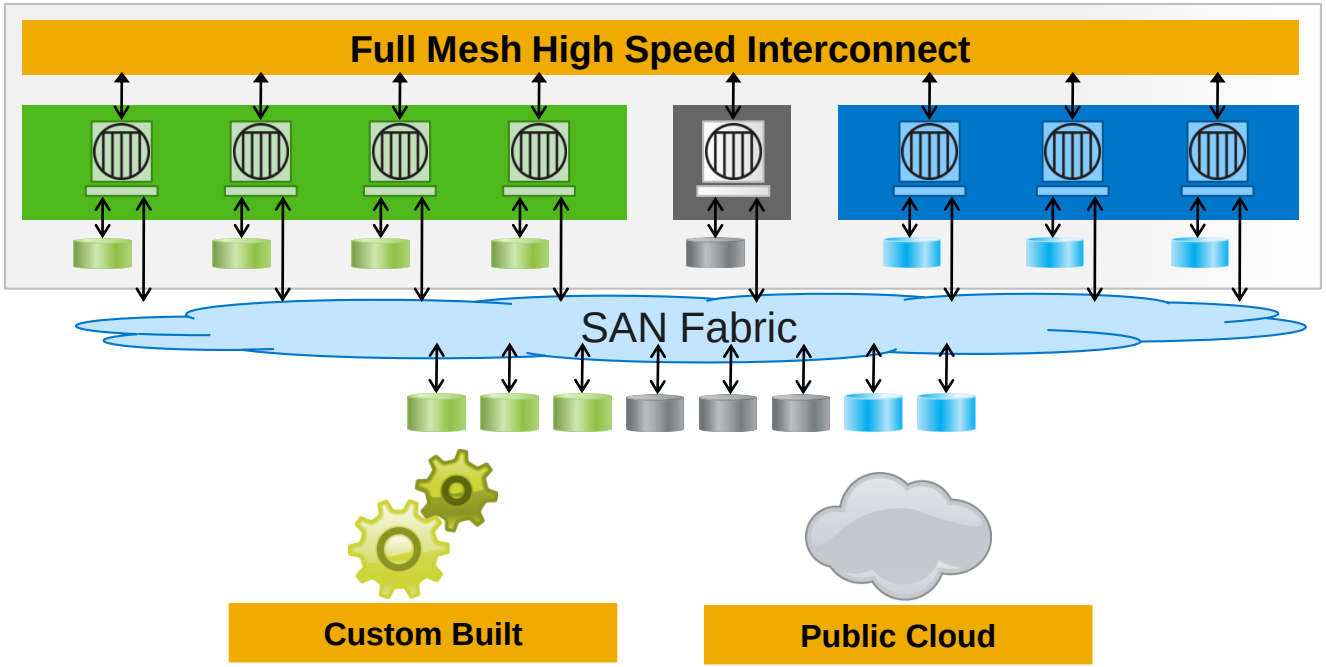
Wide array of choices



Multiplex Grid Option

PlexQ™ framework

MPP shared everything deployments



PlexQ™ is an architectural framework in SAP Sybase IQ that enables an elastic, massively parallel, high performance enterprise data warehousing infrastructure based a combination of a powerful Shared Everything MPP and a rich application services layer

SAP Sybase IQ Multiplex Grid option provides the shared everything MPP layer that unleashes the full power of PlexQ™ by enabling incremental scale-out of concurrent query jobs, load jobs, or both in an elastic, cost effective manner for high performance EDWs

Key characteristics

Foundational element of SAP Sybase IQ PlexQ™ framework

- It is a shared everything MPP architecture
  - Compute units share storage capacity
  - Compute unit capacities can be heterogeneous
    - Each unit (server machine) can have varying number of CPU cores and memory capacity

Enables incremental workload (loads, queries, users) scale out

- Simply add compute nodes and/or storage capacity
  - Compute and storage capacity can be added independent of each other

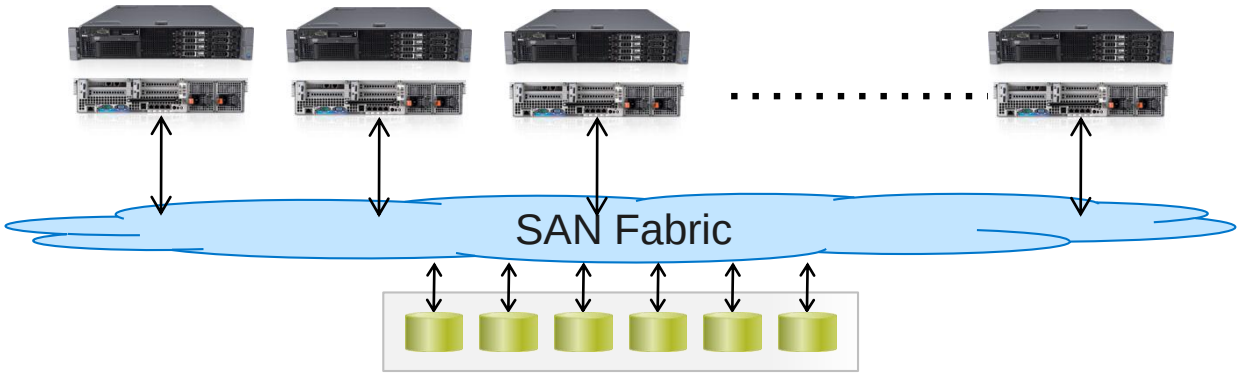
Scalability is elastic

- Add/shrink capacity on-demand based on workload increase/decrease
- Group capacity and dedicate to tasks, user groups for isolation

Deployed use case

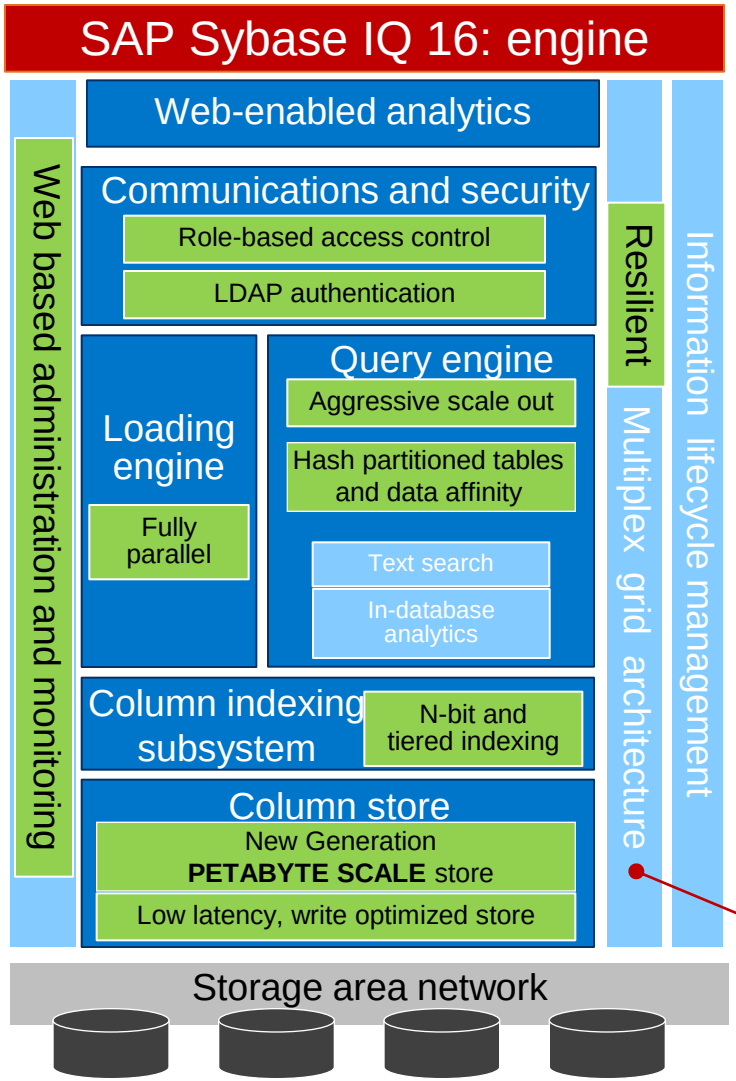
Information service provider measures the digital world

- An IQ customer provides solutions for online audience measurement, e-commerce, advertising, search, video and mobile to analysts with digital marketing and vertical-specific industry expertise
- Large SAP Sybase IQ Multiplex Grid with tens of servers and hundreds of CPU cores
- Manages more than 150TB of data with trillions of rows and tens of thousands of tables
- More than 200+ concurrent users with highly parallel and distributed workload
- Incrementally scalable on commodity hardware



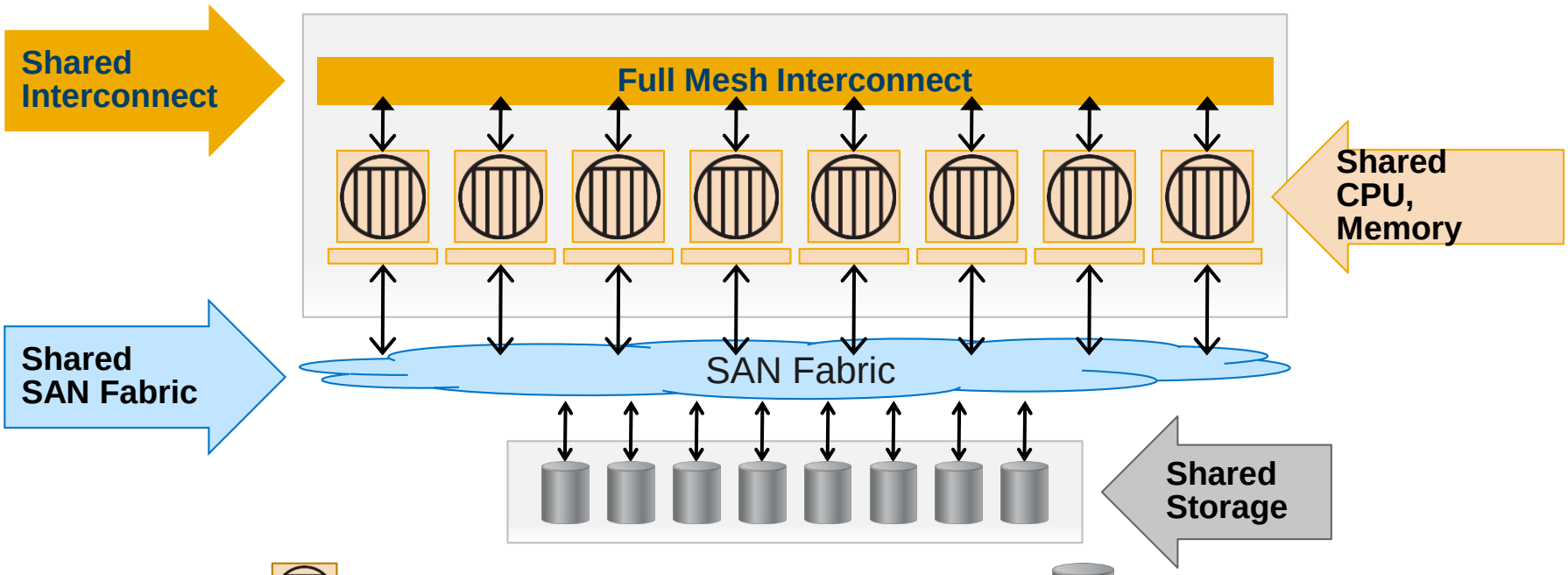
Multiplex grid option

Critical component of the SAP Sybase IQ 16 product family



Industry-leading , scalable Multiplex grid for concurrent and massively parallel processing (MPP) analytics

Architecture overview



Compute nodes

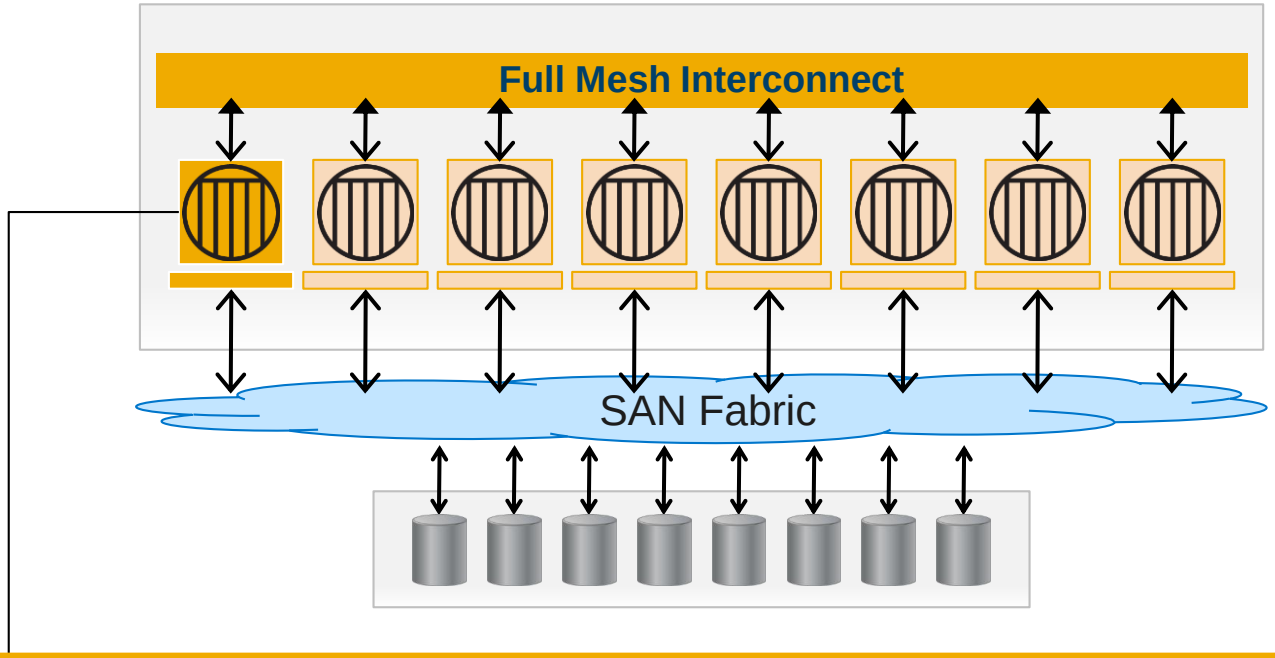
- Can be 86/BM Power/HP Itanium/Solaris Sparc servers
- Each node can have a different CPU core count, memory size
- One node is designated as the coordinator for meta data management
- Any node can be a write node
- Any node can be a read node

Shared storage

- Can be SSD, FC, SATA or a combination of all
- Shared storage can be physical SAN with FC and/or iSCSI
- Shared storage can be virtual SAN over DAS
- Storage can be logically grouped with compute
- Storage can be tiered for data aging

Architecture component

Compute node as a coordinator

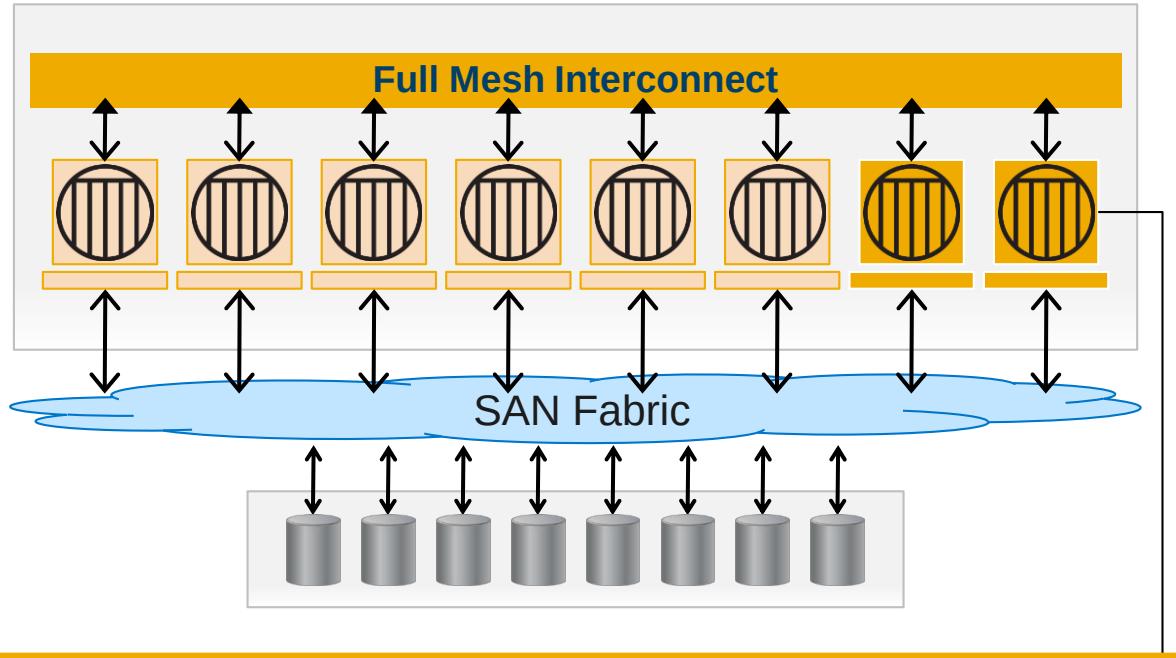


Coordinator node

- Each multiplex configuration requires a coordinator node that can also serve as a read and/or write
- Manipulates local data in system and user tables
- Coordinates all read-write operations on shared objects
  - Locking, versioning, storage mgmt, global transactions, catalog synchronizations

Architecture component

Compute node as a writer

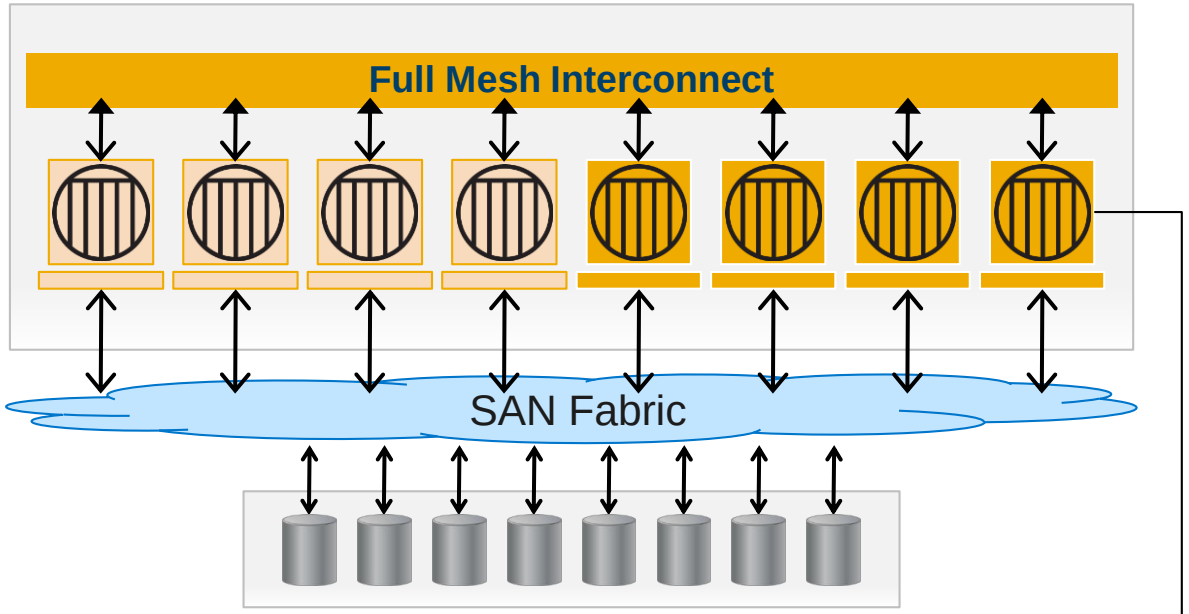


Writer node(s)

- One or more nodes can be a writer node(s)
- Writer node(s) can load, insert, update, delete, read data to/from shared store
- Multiple nodes can be designated as writer nodes
- Multiple writer nodes can be grouped together logically into a writer “farm”
- A writer node may be designated as a failover node for the coordinator node

Architecture component

Compute node as a reader

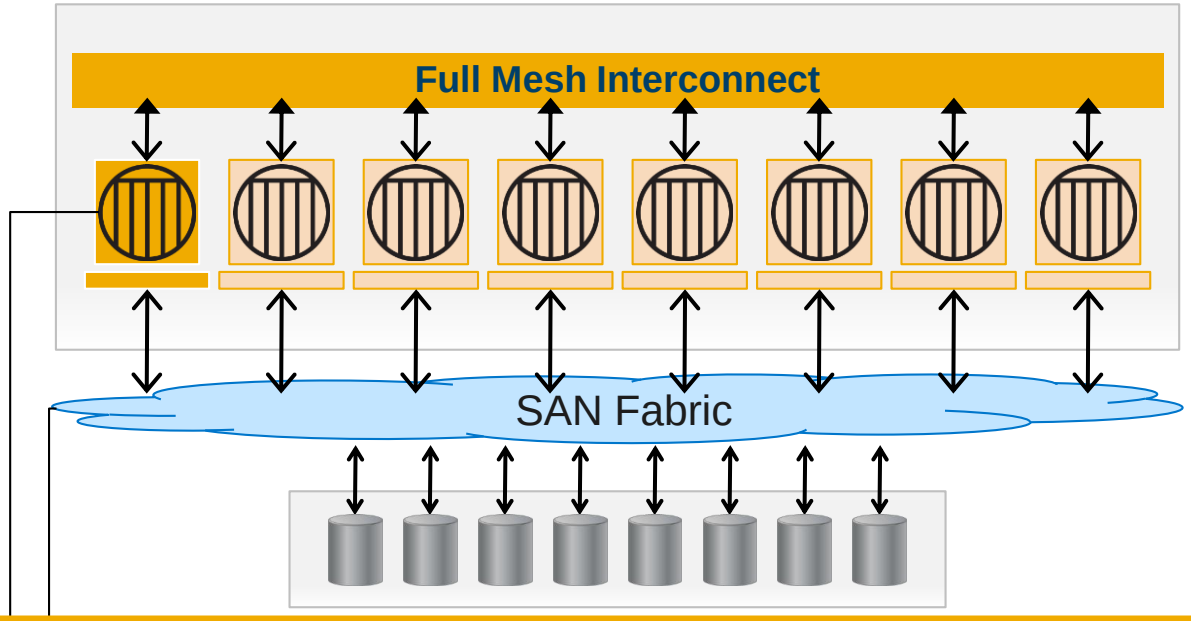


Reader node(s)

- One or more nodes can be a reader node(s)
- Reader node(s) can only read data from shared store
- Writer nodes can also serve as reader nodes
  - Writer nodes can be converted to a read only nodes dynamically
  - Reader nodes can be converted to writer nodes dynamically
- Multiple reader nodes can be grouped together logically into a reader “farm”

Architecture component

Interconnects

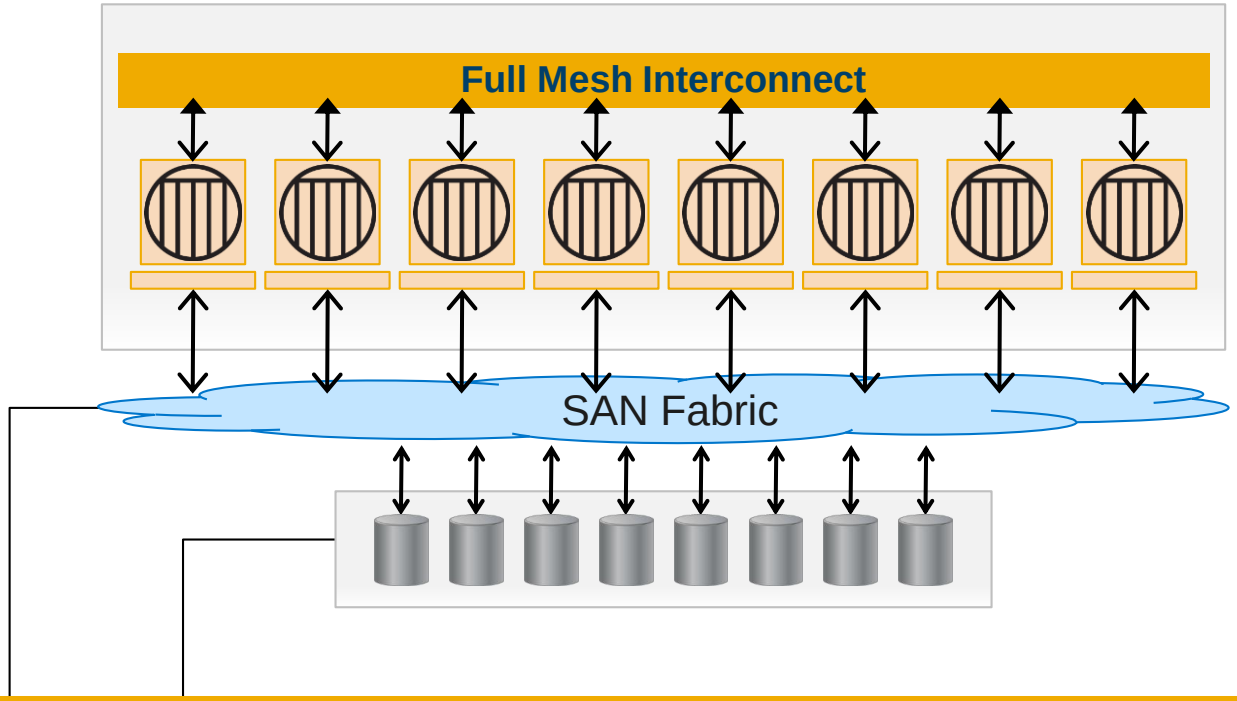


Interconnects

- Full mesh interconnect
  - 10g ethernet based
  - Peer-to-peer communication – dedicated and/or shared
  - Meta data exchange only – no user data payload
- SAN fabric interconnect
  - Fiber channel or iSCSI based
  - I/O channel isolation and/or capacity expansion via SAN switches
  - User data payload interchange only

Architecture component

Shared storage

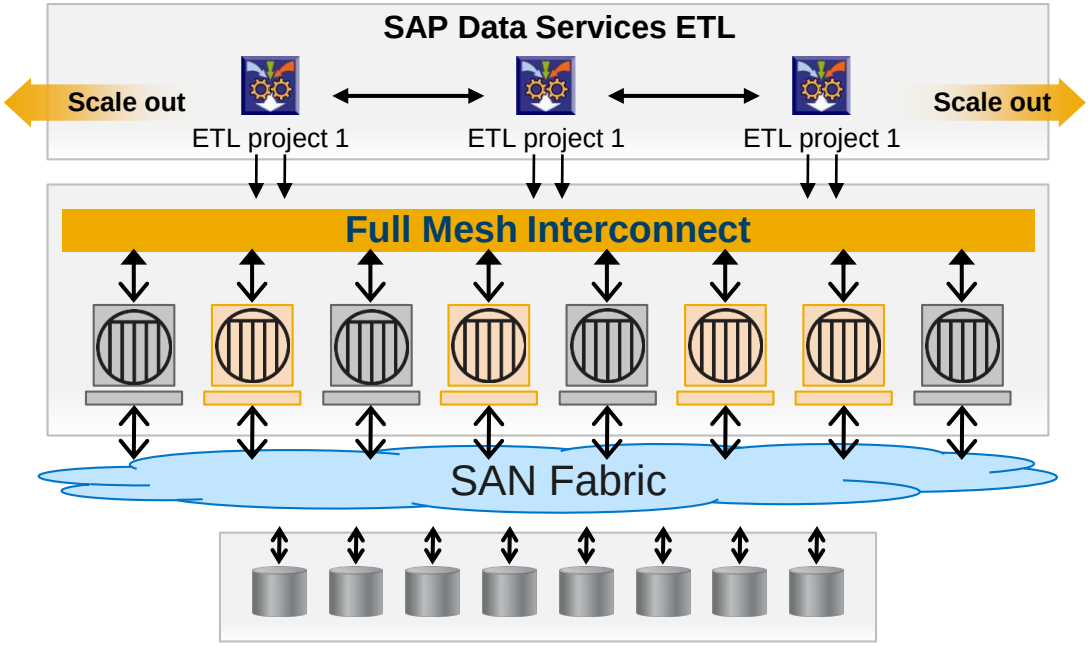


Shared storage

- Exposed to compute nodes via FC or iSCSI SAN API
  - For physical SANs
    - Disks may be a combination of SSD, FC, SATA
  - For virtual SANs
    - DAS disks (SSD, SATA) exposed via storage virtualization software

Architecture usage

Parallel, distributed writes

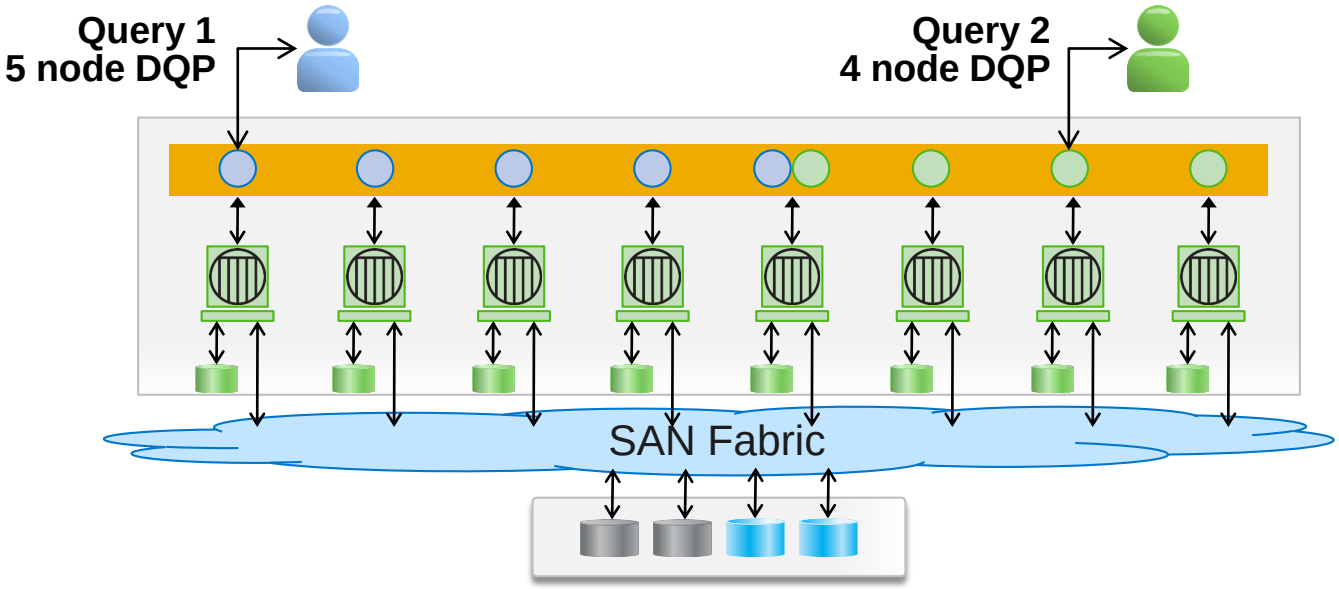


Multiple writers

- One or more writer node(s) may load/insert/update/delete at any given time
  - Writes can proceed in parallel uninterrupted via multiple nodes to disparate tables
  - Writes into same table from multiple nodes are serialized via locking
  - Writes and reads on a same table do not block – snapshot isolation is employed
- Multiple writers can be fed in parallel by ETL or other utilities

Architecture usage

Distributed query processing

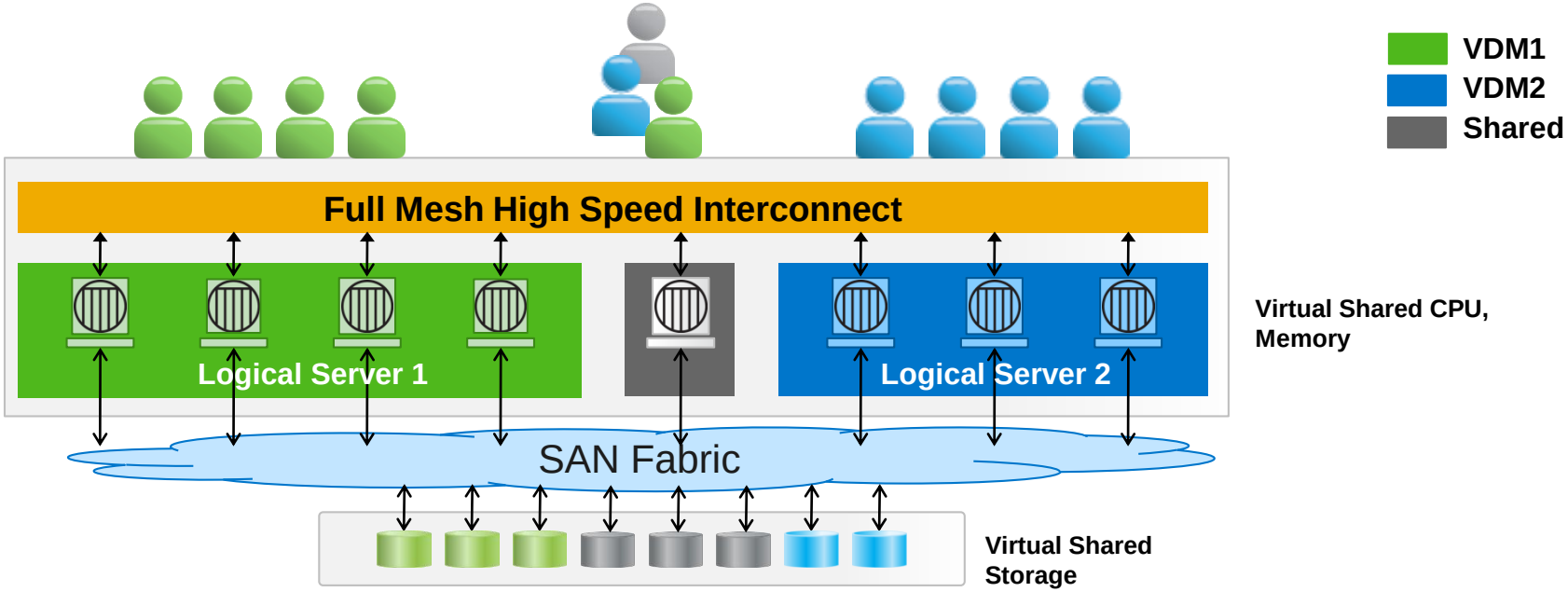


Massively parallel processing

- **Leader node:** Receives and initiates queries, including UDFs
  - Any node can be a leader, one leader per query, many concurrent leaders possible
  - Leader node may satisfy query within itself
- **Worker node:** Nodes picking up work units from leader
  - Many worker nodes per query, same worker node can serve multiple queries
  - Worker nodes are enlisted only if leader cannot satisfy query on its own

Architecture usage

Elastic virtual data marts

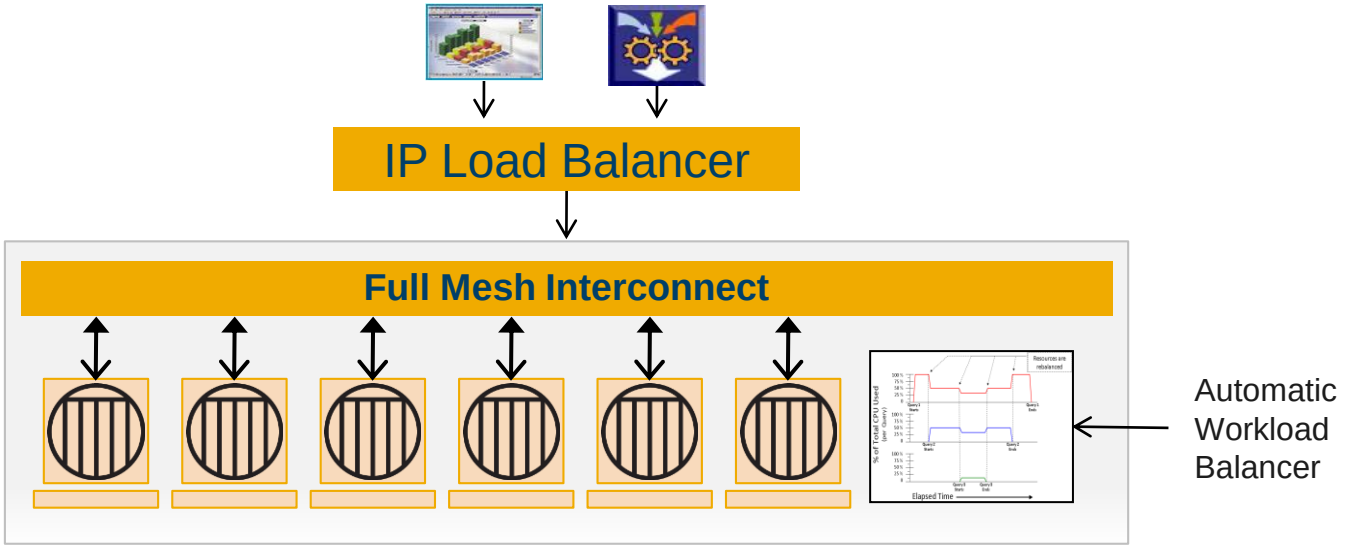


Virtual data marts

- VDM is logical binding of mutually exclusive nodes, memory, storage
  - Logical Server (LS) is a mutually exclusive logical binding of nodes, memory
  - Logical Server (LS) is a subset of VDM
  - Bindings are elastic i.e. they can dynamically grow/shrink
- VDM and works via login permission control
- VDM can isolate applications, workload, user communities
- DQP within VDM boundaries only
- Login redirection

Architecture usage

Workload balancing

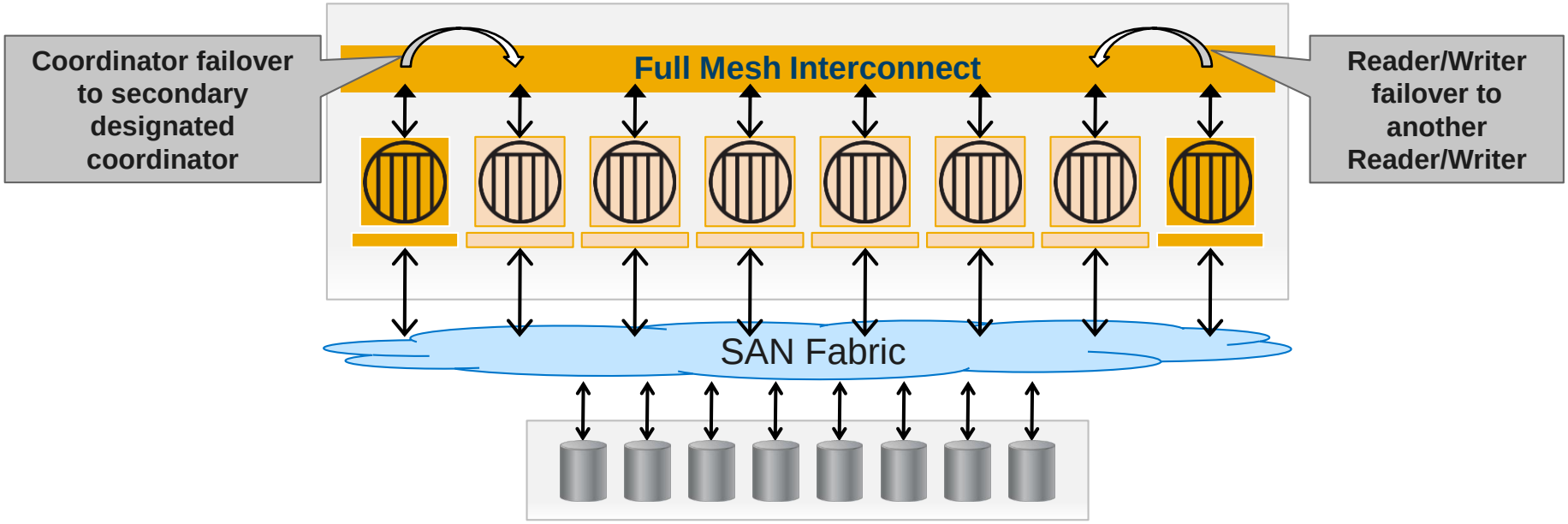


Highly efficient mixed workload balance

- Built in workload balancer
  - All work units participate in a democratic fashion i.e. they initiate participation when free
    - Makes task executions self-throttling
  - If work units fail, leader picks up and completes work
  - Workloads can be isolated by logical groupings via logical servers
- External IP workload balancer
  - Client side job submission may be done via external IP load balancers
    - External load balancers initiate based on node utilization or on a round robin basis

Architecture usage

High availability with fail over

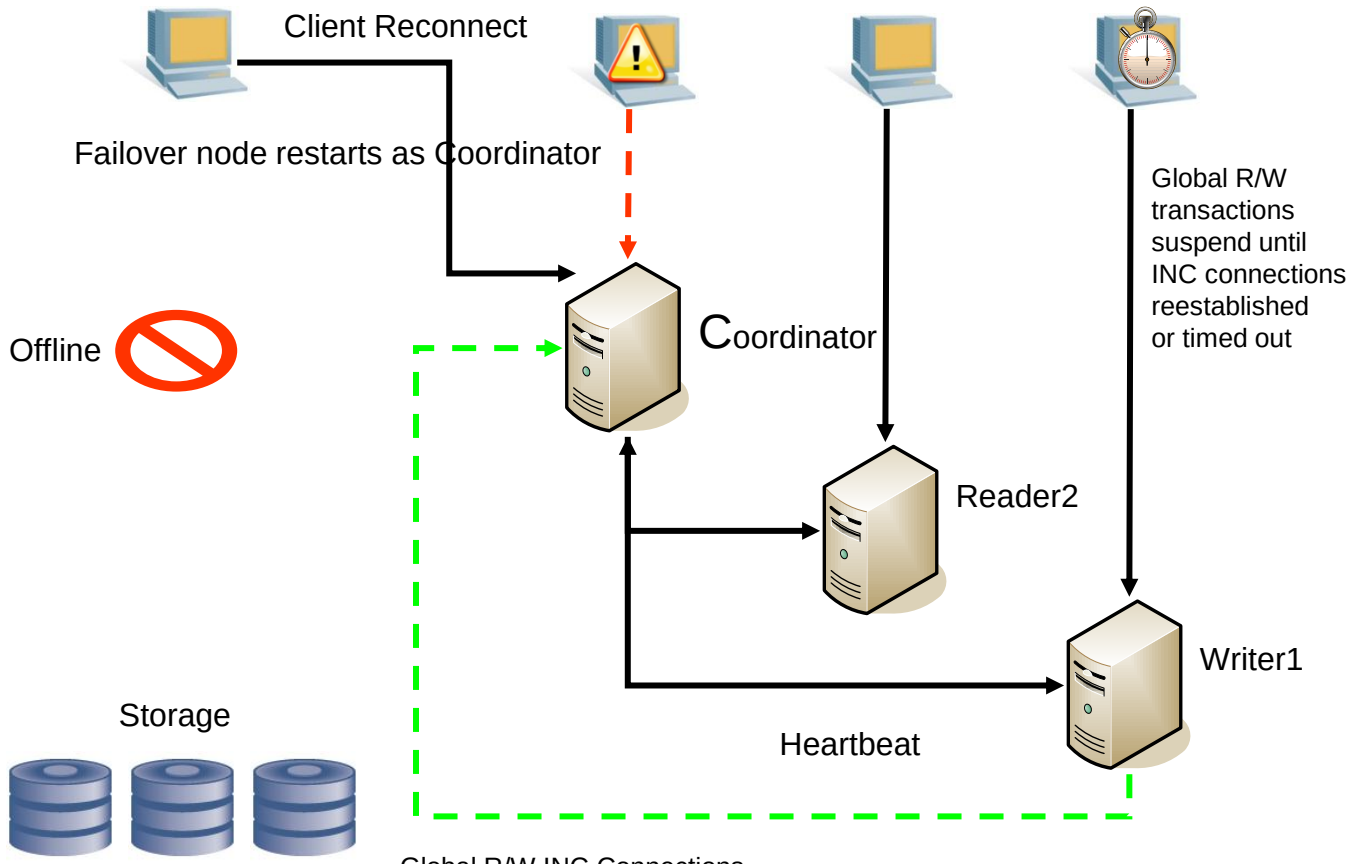


Robust failover for high availability

- Robust full mesh interconnect maintains high resiliency via continuous heartbeat monitoring
- In event of coordinator node failure, a simple process fails over primary node to designated secondary
  - Reader failover during massively parallel query processing is automatic
  - Readers can proceed without a coordinator. Writers require a coordinator to handle global transaction management.
  - Coordinator failover can be automated via scripts or partner products e.g. Symantec VCS

Multiplex grid

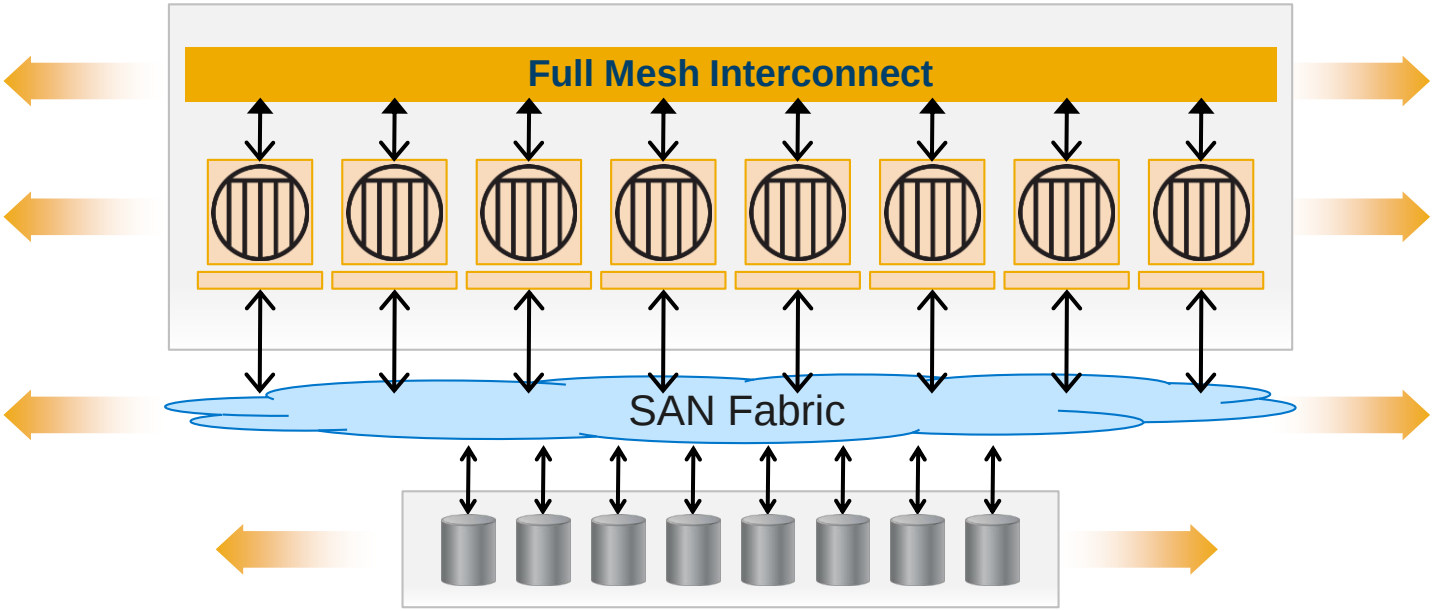
Global transaction resiliency



Global R/W INC Connections
automatically reconnect and resume transaction
after a coordinator failover.
In most cases long-running loads will transparently resume.

Architecture usage

System scale out



Multi-dimensional scale out

- Multiple resources can scale out independently
  - Storage, server (CPU, memory), SAN switches, interconnect can scale on their own
- Scale out is incremental and linear
  - No need to add large units of monolithic CPU/storage pairs

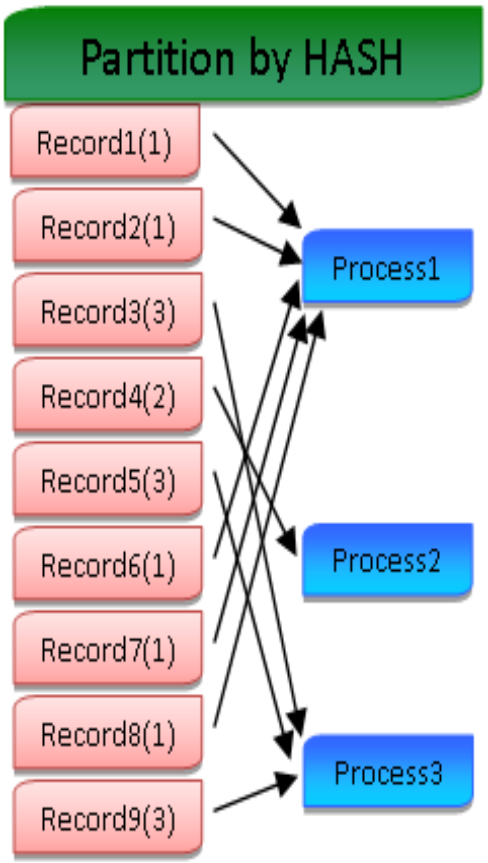
Query engine scalability

Hash partitioned tables for query scale up and scale out

Hash partitioned tables support scalable query performance

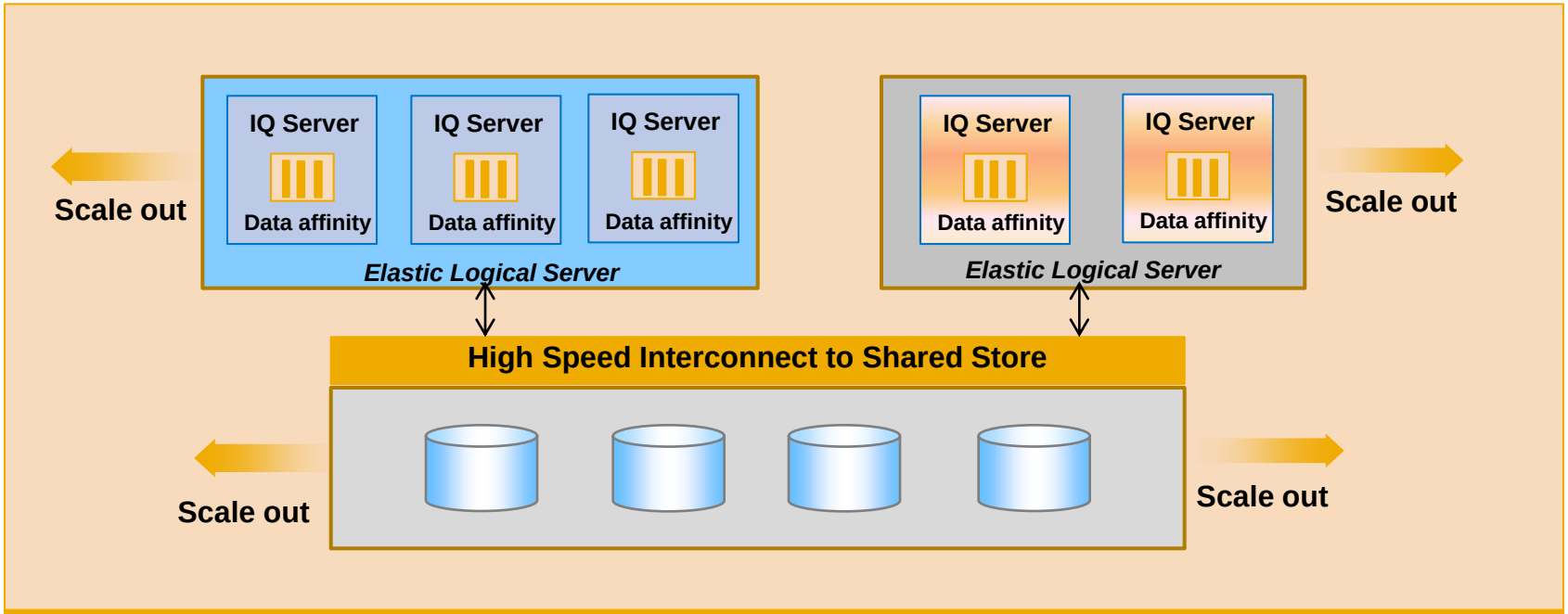
- Data is automatically partitioned during loading with built-in hash algorithms
- Data is divided into persistent subsets
  - Reduces results sharing
  - More efficient CPU usage
  - Reduces instantaneous temp usage

Optimizer will use hash partitions for join and group by when available



Query engine scale out

Data affinity for “shared nothing” performance



Data affinity

- **Query optimizer tracks which data is in which server cache**
- **Distributed query work assignments are intelligently assigned based on which server has which data set in cache**
- **Data must be hash partitioned**
- **Each partition is automatically assigned to a specific node as queries execute**
- **Data affinity is self managed – no user interaction required**
- **Caches stay “hot” and I/O is reduced**

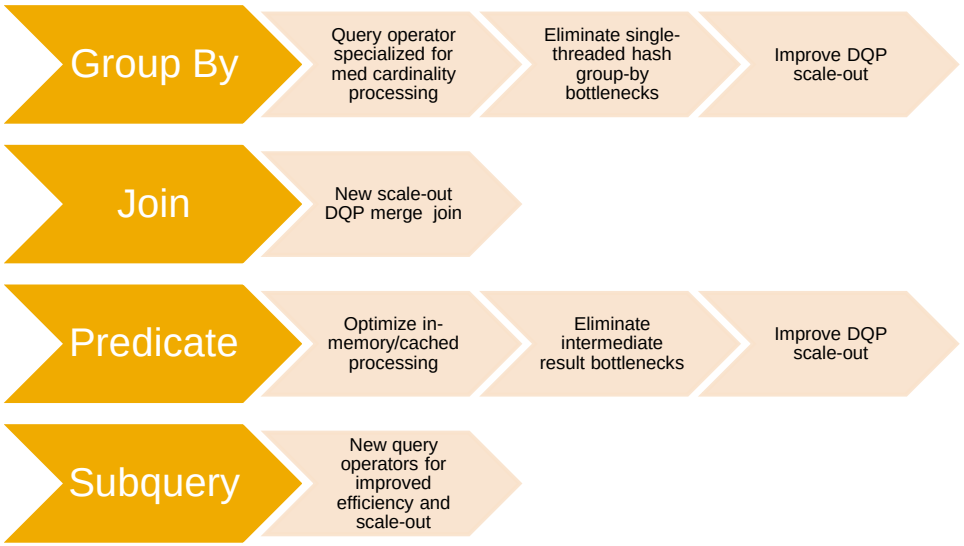
QUERY SCALE OUT – Query Runtime and DQP Optimization

Value proposition

- Eliminates SMP (single node) and DQP (distributed query processing) bottlenecks
- Leverages large memory and scale out
- Lowers shared temp and interconnect bandwidth

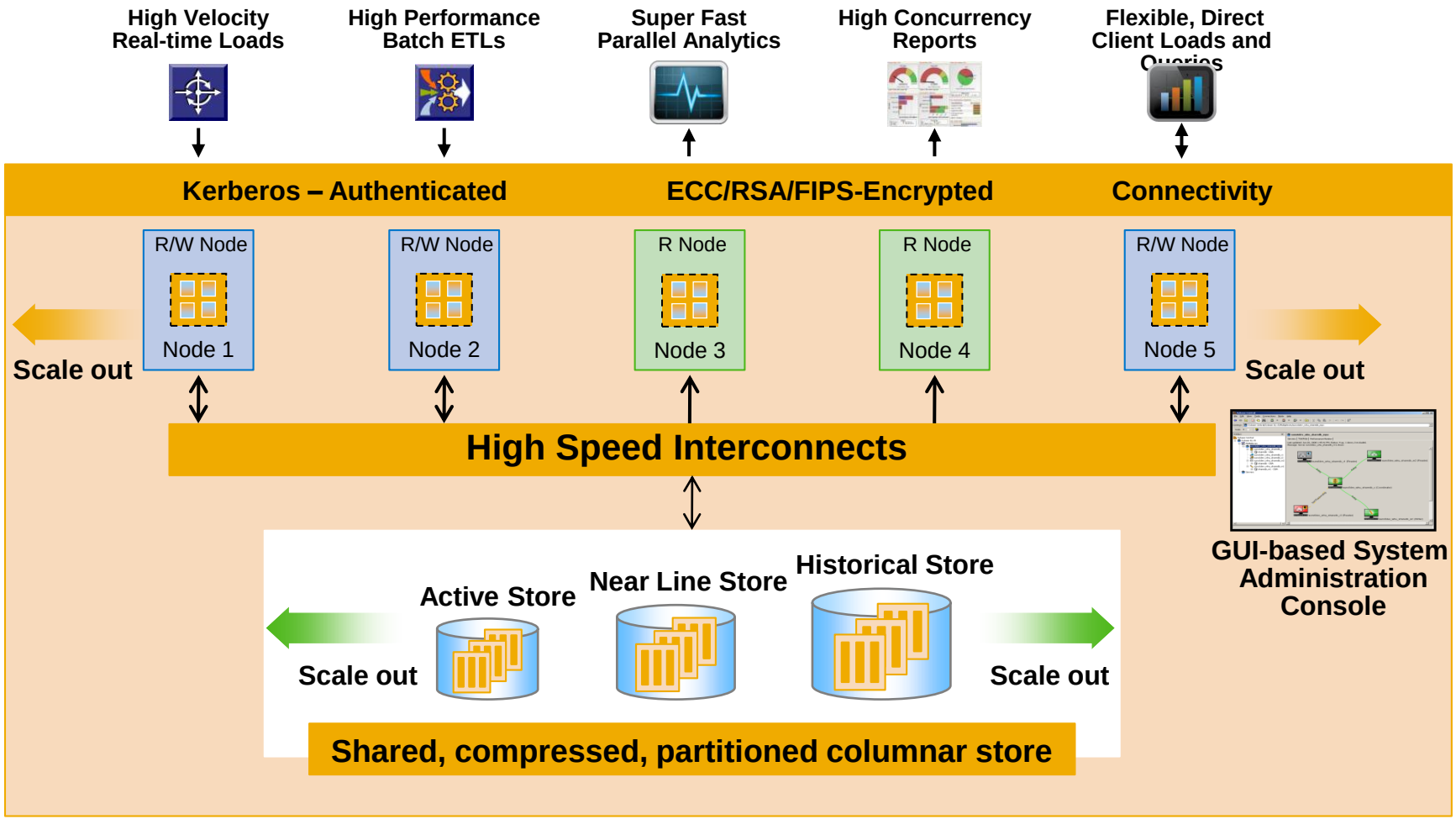
Architectural considerations

- Takes place automatically as optimizer selects best plan based on cost
- For non-partitioned data, new Join and Group algorithms reduce the amount of intermediate results exchanged



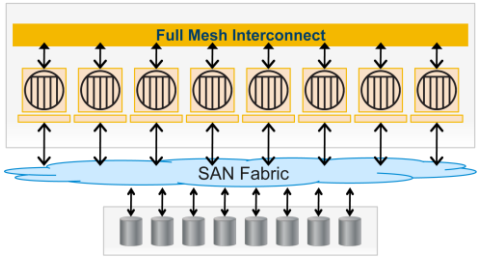
Architecture usage

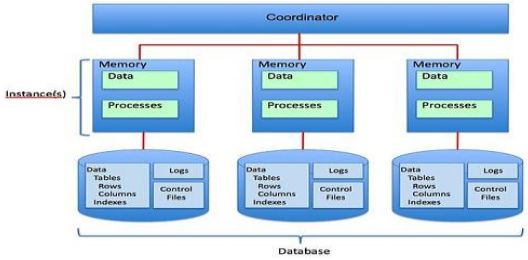
Ideal for large mixed workload EDWs



Architecture

Comparing against alternatives





PlexQ: MPP Shared Everything

- ☑ Independent, linear scale out
  - ☑ Compute OR memory OR storage
- ☑ Tremendous resource optimization
  - ☑ Dynamic resource provisioning
- ☑ No forcible data partitioning
- ☑ Well behaved, full spectrum support
  - ☑ Batch, ad-hoc, concurrent
- ☑ Mature DR support
- ☑ Less data duplication
- ☑ No single “choke-point” leader node
- ☑ SAN affordable and reliable
  - ☑ Virtualized SAN over DAS possible

PlexQ: MPP Shared Nothing Contenders

- ! Homogeneous unit scale out
  - ✖ Compute + memory + storage in chunks
- ! Sub-optimal resource utilization
  - ✖ Share Nothing!
- ! Forced data (re)partitioning
- ! Poorly behaved full spectrum support
  - ☑ Good for complex queries (planned)
  - ✖ Struggle with ad-hoc, concurrent queries
- ! Complicated DR support but better fault tolerance
- ! Significant data duplication across nodes
- ! Master “coordinator” a “choke-point”
- ☑ DAS very affordable but more redundancy needed

Multiplex grid option

Benefits

Option	Multiple Nodes	Distributed and Parallel Processing	Failover Nodes	Compute, User Scalability
Multiplex Grid Option	Yes	LOADS: Distributed & Parallel QUERIES: Distributed & Parallel UDFs: Distributed & Parallel	Yes	Yes
SAP Sybase IQ Base Product	No	LOADS: Parallel Only QUERIES: Parallel Only UDFs: Parallel Only	No	Partial

SAP Sybase IQ multiplex grid option

Summary

Powerful MPP Shared Everything based architecture that meets modern day EDW demands

Learn more:

- Visit: http://www.sap.com/iq
- Call: 1-877-727-1127 FREE ext. 11001

Thank you

© 2013 SAP AG. All rights reserved.