Capacity Management for Oracle Database Machine Exadata v2

Dr. Boris Zibitsker, BEZ Systems

NOCOUG 2010

About Author

Dr. Boris Zibitsker, Chairman, CTO, *BEZ Systems.*

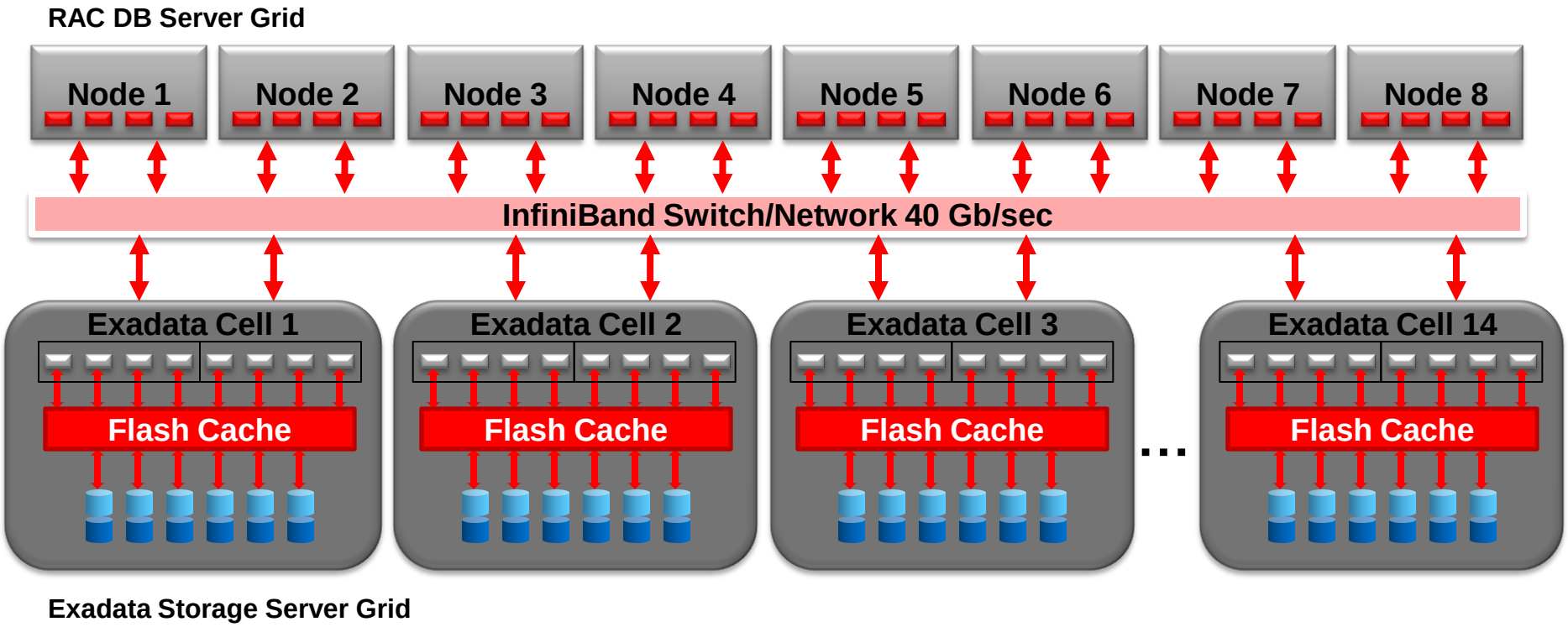
- Boris and his colleagues developed modeling technology supporting multi-tier distributed systems based on Oracle RAC, Teradata, DB2, and SQL Server. Boris consults, and speaks frequently on this topic at many conferences across the globe.

Outline

- Overview of Oracle Database Machine V2 architecture
- Capacity management challenges for DB Machine
- Role of predictive analytics
- How to justify strategic capacity planning decisions
- How to justify tactical performance management actions
- How to justify operational workload management actions
- Summary

Oracle DB Machine Exadata v2 Architecture

Massively Parallel Grid



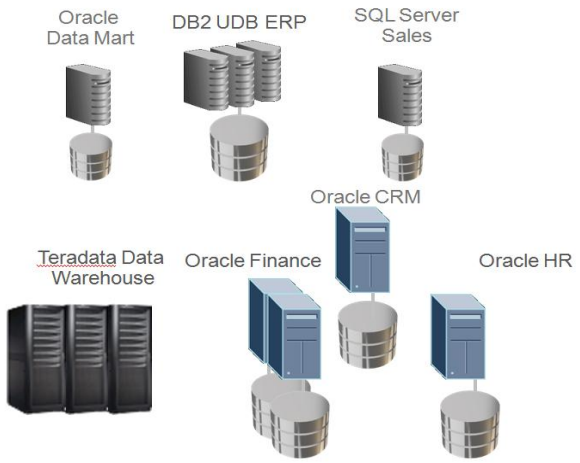
Smart Scans
Hybrid Columnar Compression
Storage Indexes
Flash Cache

Exadata Cell Disk Storage Capacity:

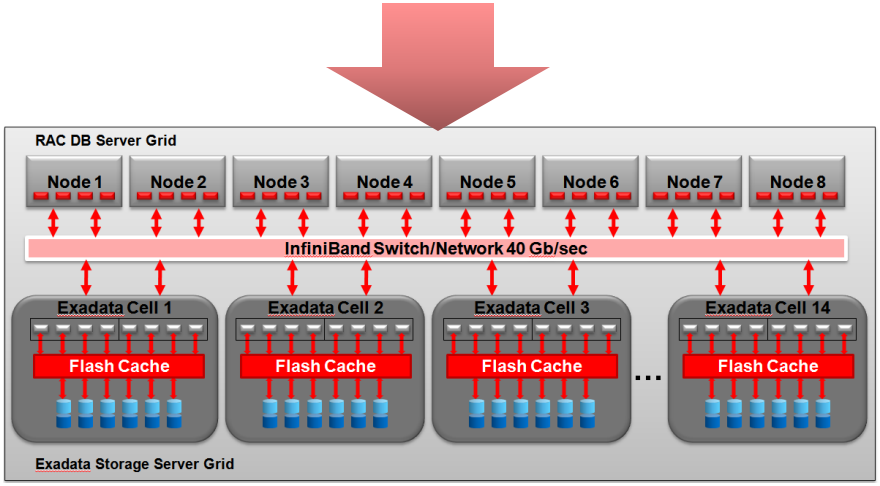
- **12 x 600 GB SAS disks (7.2 TB/cell @ 100 TB total)**
- **12 x 2TB SATA disks (24 TB/cell @ 336 TB total)**
- **4 x 96 GB Sun Flash Cards (384GB/cell @ total 5TB)**

Capacity Management Challenges for DB Machine

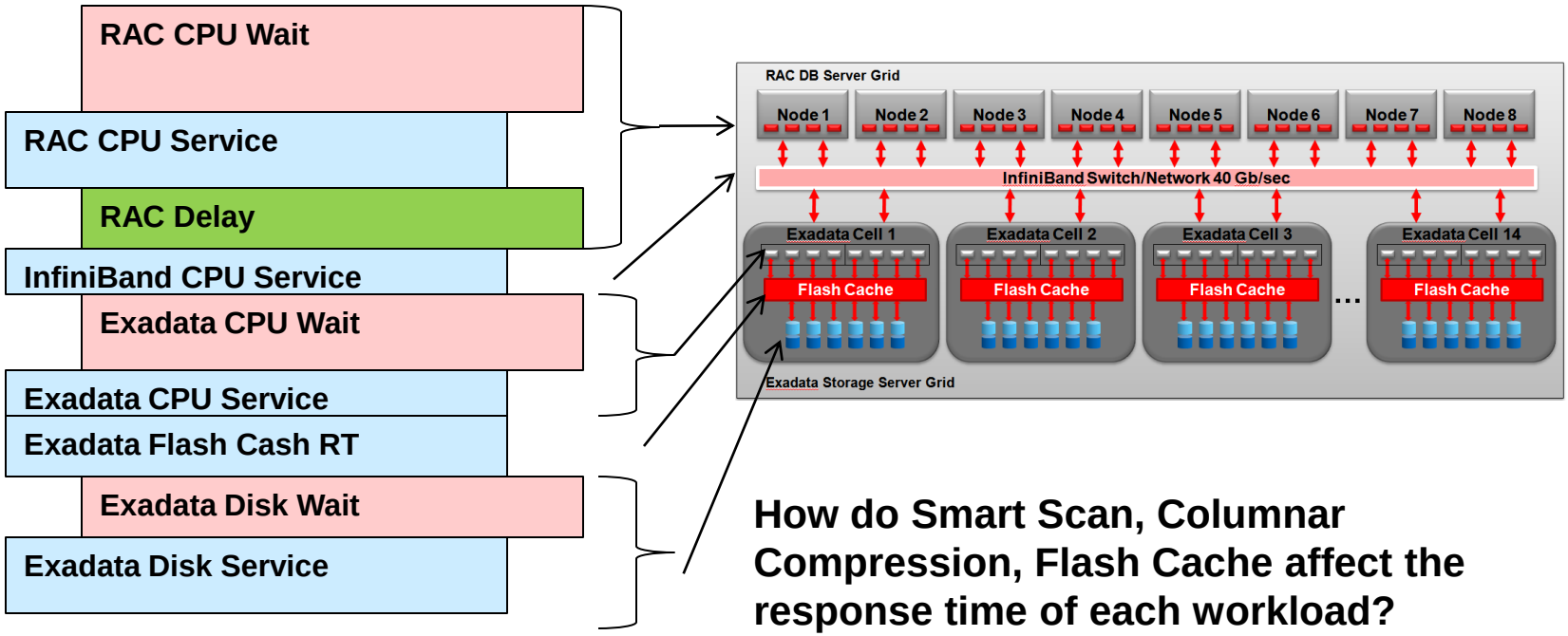
- Data collection
- Workload characterization
- How to set realistic SLO
- Performance prediction
- How to justify DB Machine
- How to justify server consolidation
- What will be the impact of new application implementation
- How to justify
  - strategic capacity planning decisions
  - tactical performance management actions
  - operational workload management actions



OLTP, DW, ETL, Arch, New Appl.



Response Time, Throughput and Cost Affect Capacity Management Decisions

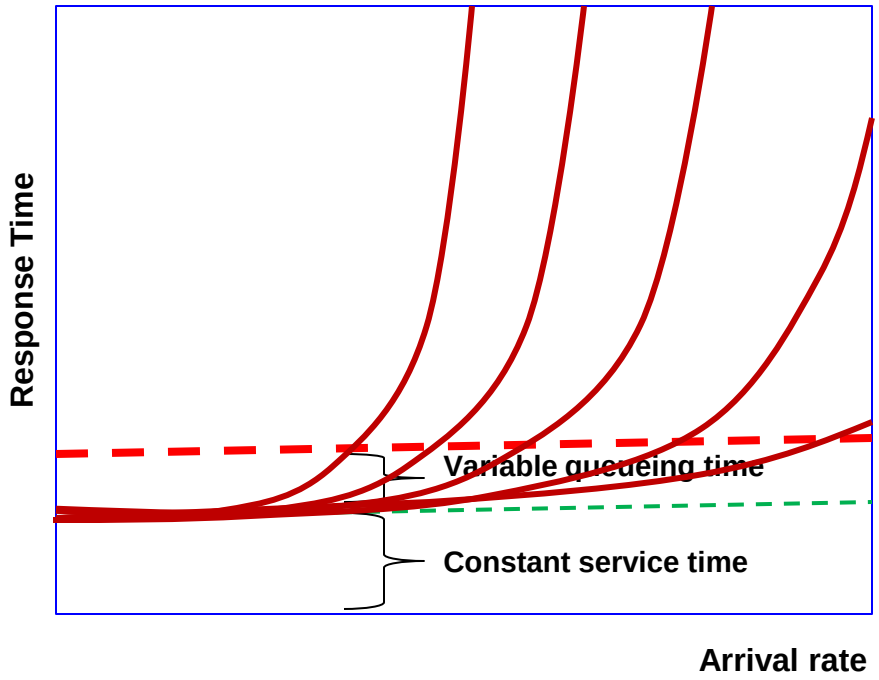


How do Smart Scan, Columnar Compression, Flash Cache affect the response time of each workload?

What should be done proactively to support acceptable response time and throughput for each workload?

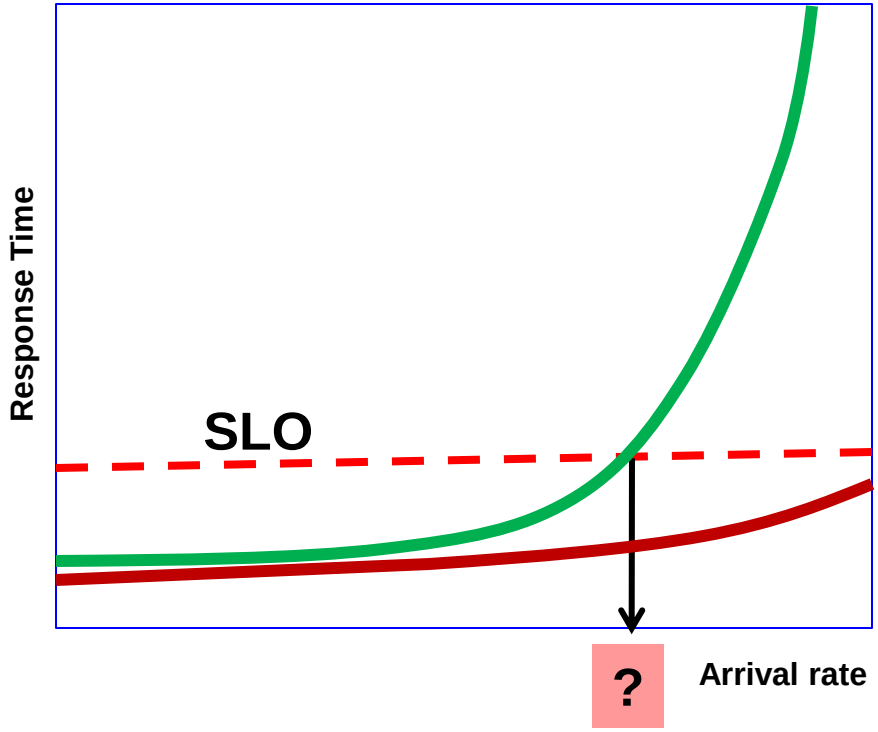
Response Time Depends on Many Factors, including Arrival Rate of Requests and Server Utilization

- Utilization of the server depends on arrival rate of requests and time required to serve an average request
- Response time of the average request depends on the service time and server utilization
- When arrival rate and server utilization are low then time waiting for service is low and the response time is equal to service time
- When workload activity is growing it increases contention for the server and significantly increases the response time
- How can you manage if you do not know what will be an impact of expected growth?

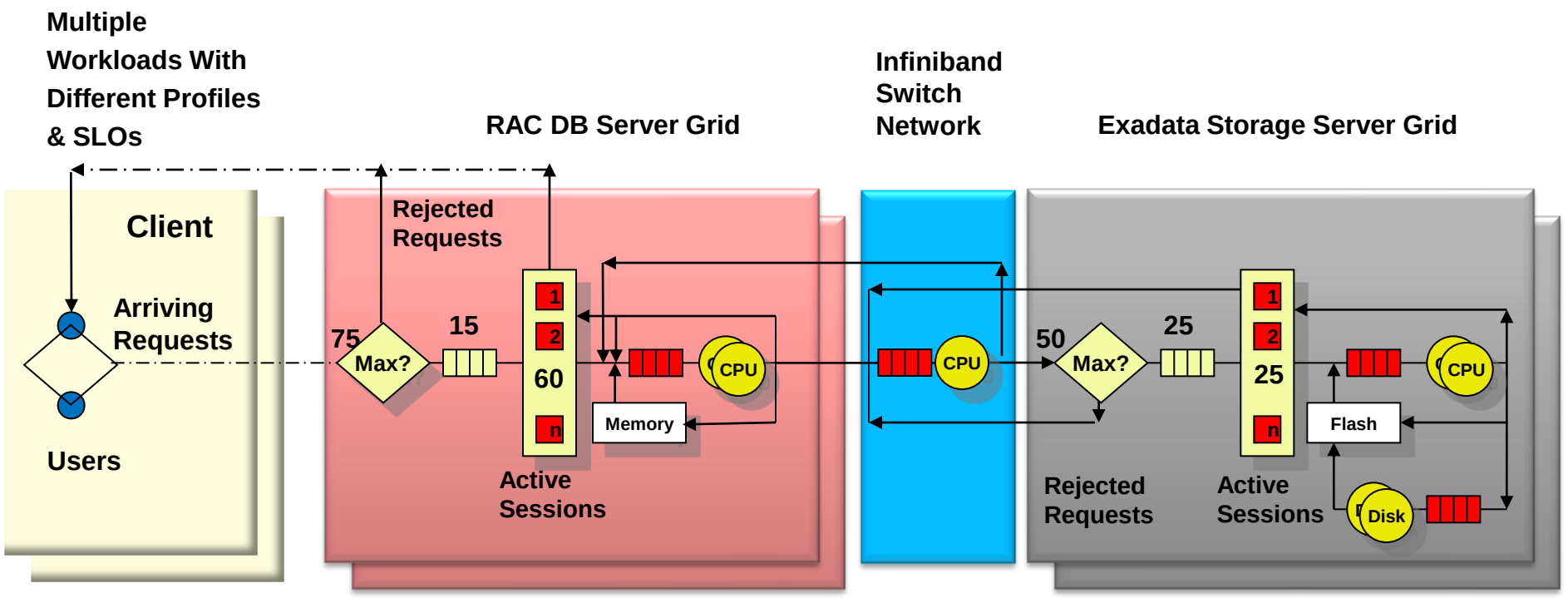


How to Decide When and What Should be changed to Support SLO for each Workload?

- Gut feelings
- Rules of thumb
- Benchmarks
- Regression analysis
- Analytical models



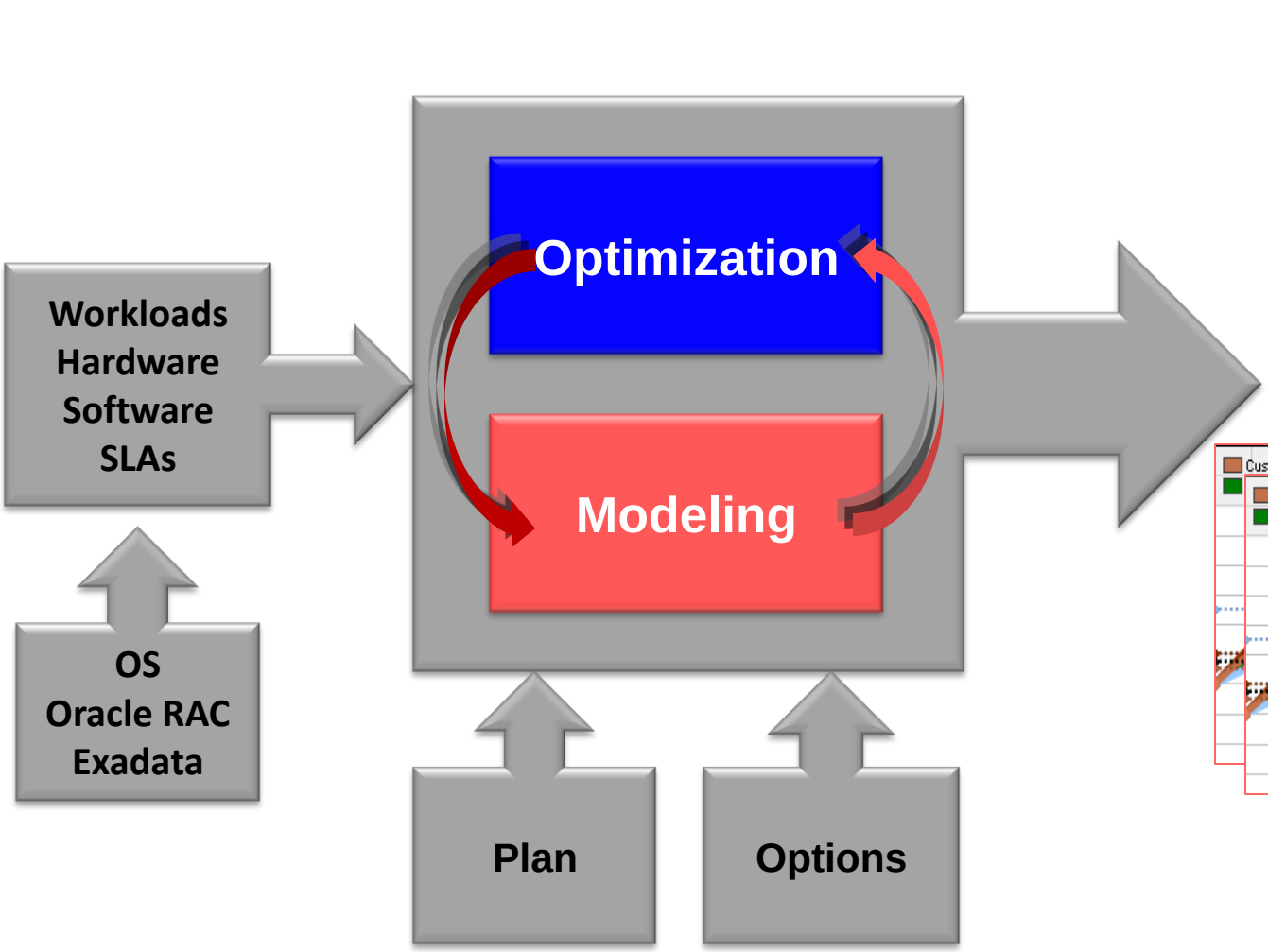
Simplified Analytical Model of the Database Machine

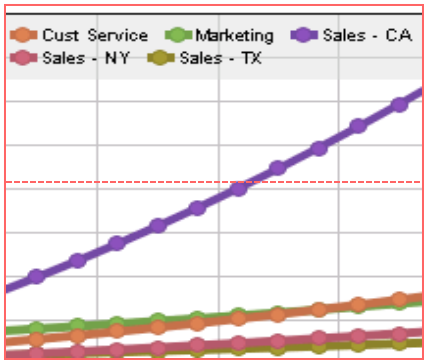


- Up to 8 servers
- 2 Intel quad-core
- Xeons each

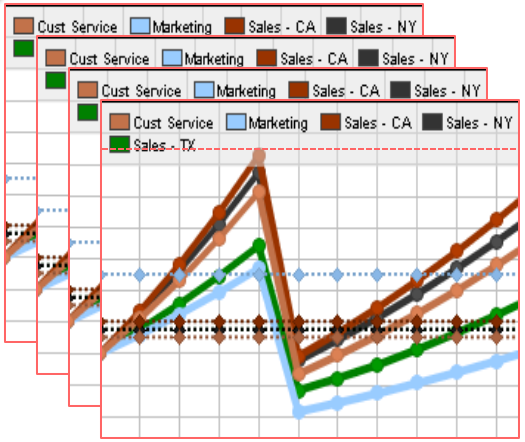
- 14 storage servers
- 100 TB raw SAS or
- 336 TB raw SATA disk
- 4 x 96GB PCI Express Flash Cards per Exadata Server
- 5TB+ flash storage

Input and Output of Modeling and Optimization





What if?

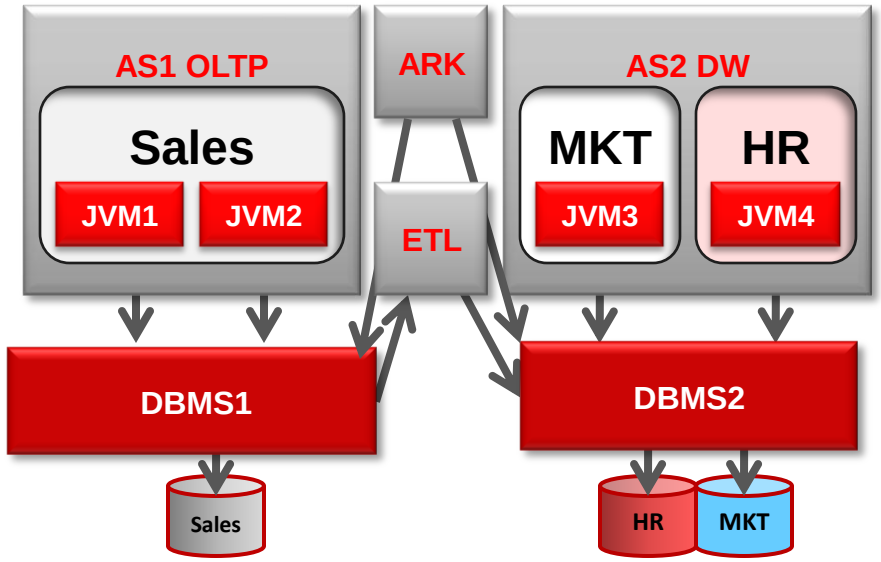


What is the best decision?

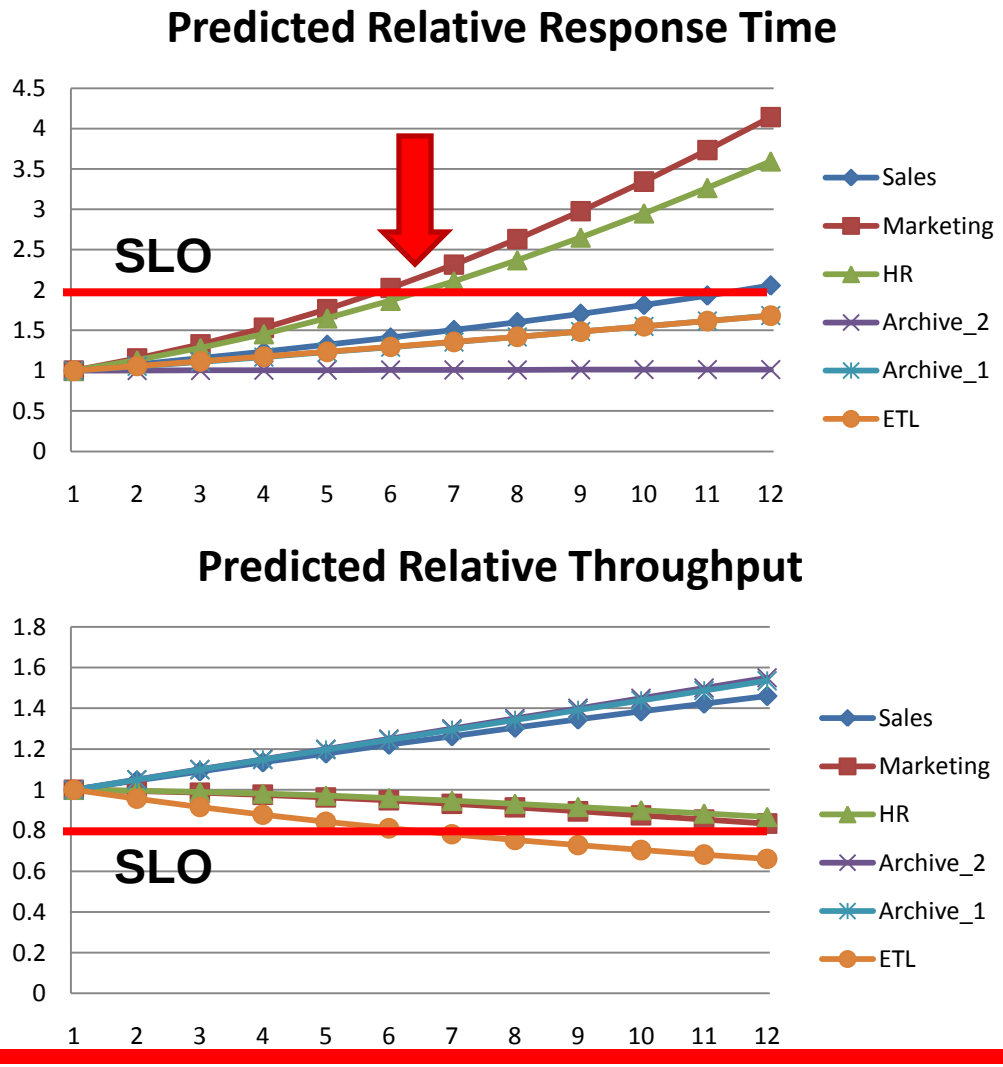
Case Study

How to justify capacity planning, performance management and workload management decisions to support workloads' SLOs effectively?

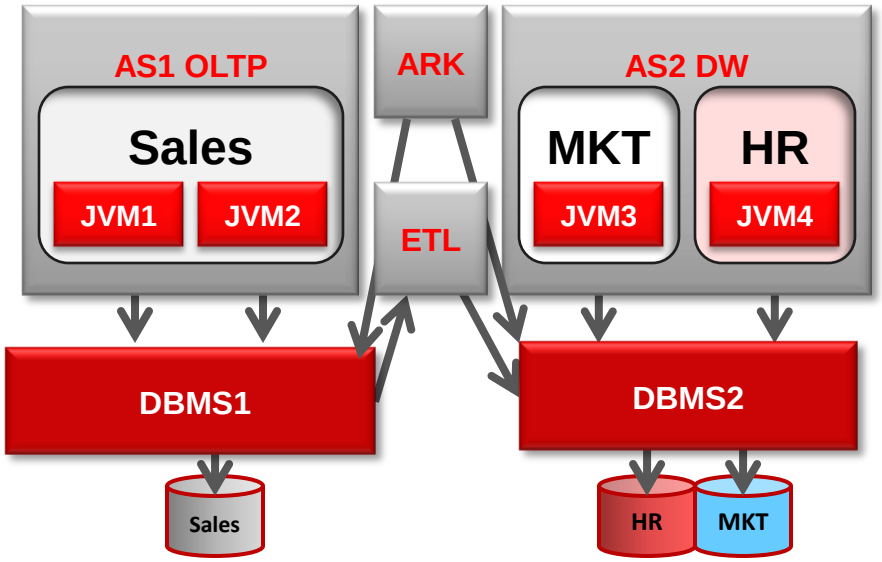
1. **What will be the impact of workload growth, and when will the current system be out of capacity?**
2. **What will be the impact of implementing Oracle DB Machine v2?**
3. **What will be the impact of hardware upgrade?**
4. **What will be the impact of performance tuning?**
5. **What will be the impact of changing workload concurrency?**
6. **What will be the impact of changing workload priority?**
7. **What will be the impact of new application?**
8. **What will be the impact of limiting CPU utilization?**



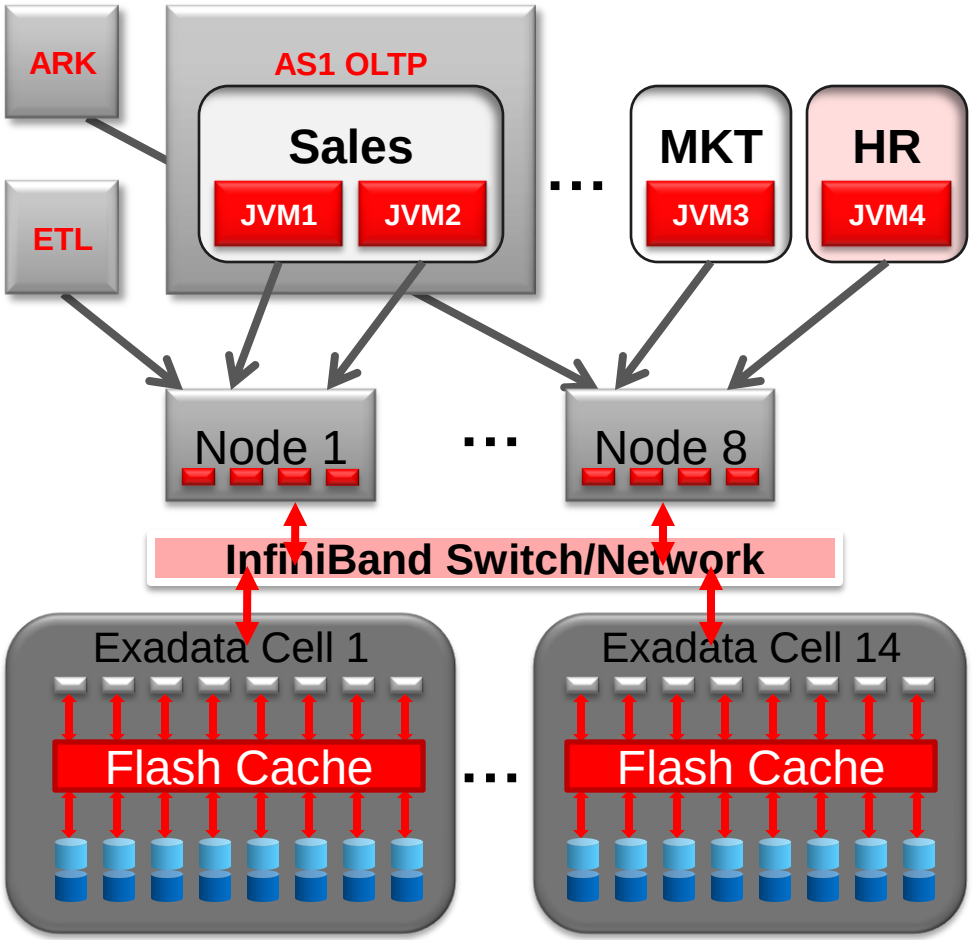
How Will Workload and Database Size Growth Affect Performance of Each Workload in Current Multi-tier Distributed Environment



How Will Server Consolidation on Oracle DB Machine Affect Each Workload's Performance?

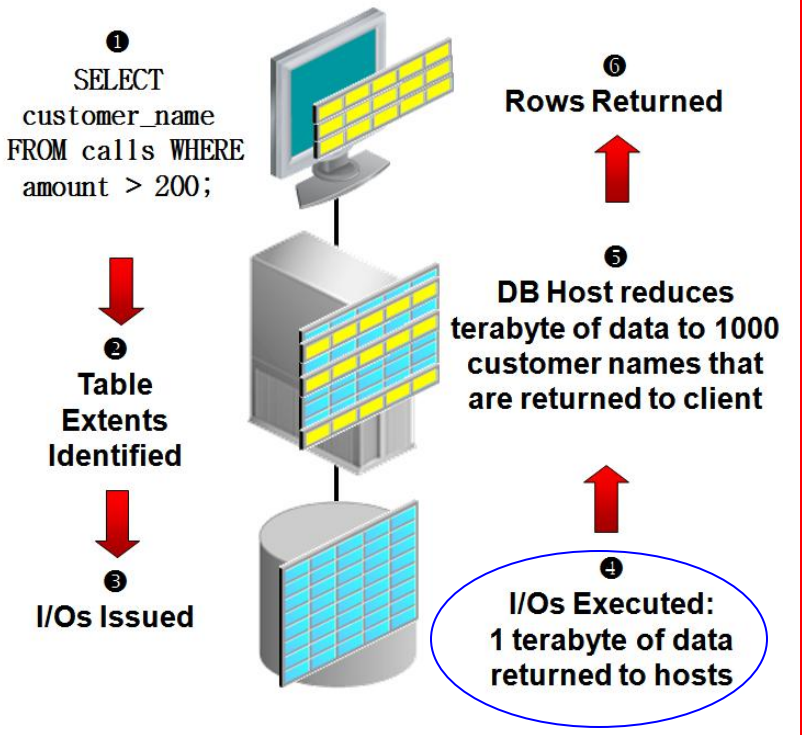


Current Multi-tier Distributed System with OLTP & DataWarehouse workloads

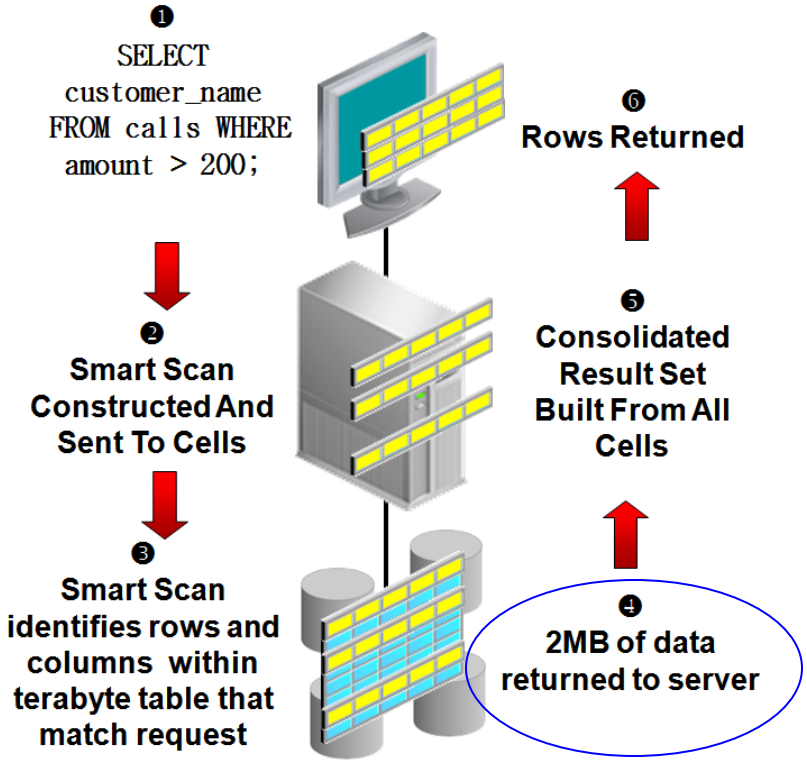


Exadata Smart Scan Reduces Volume of Data Returned to Server

Traditional Scan Processing

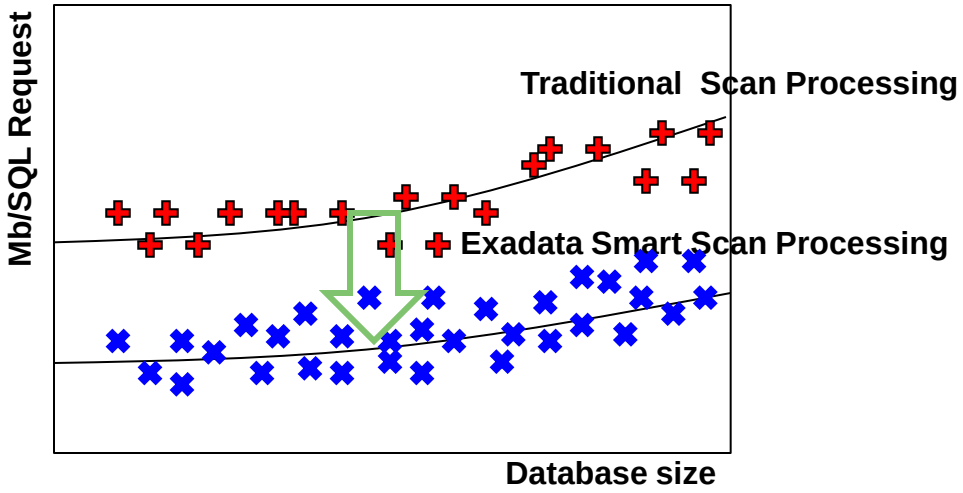


Exadata Smart Scan Processing



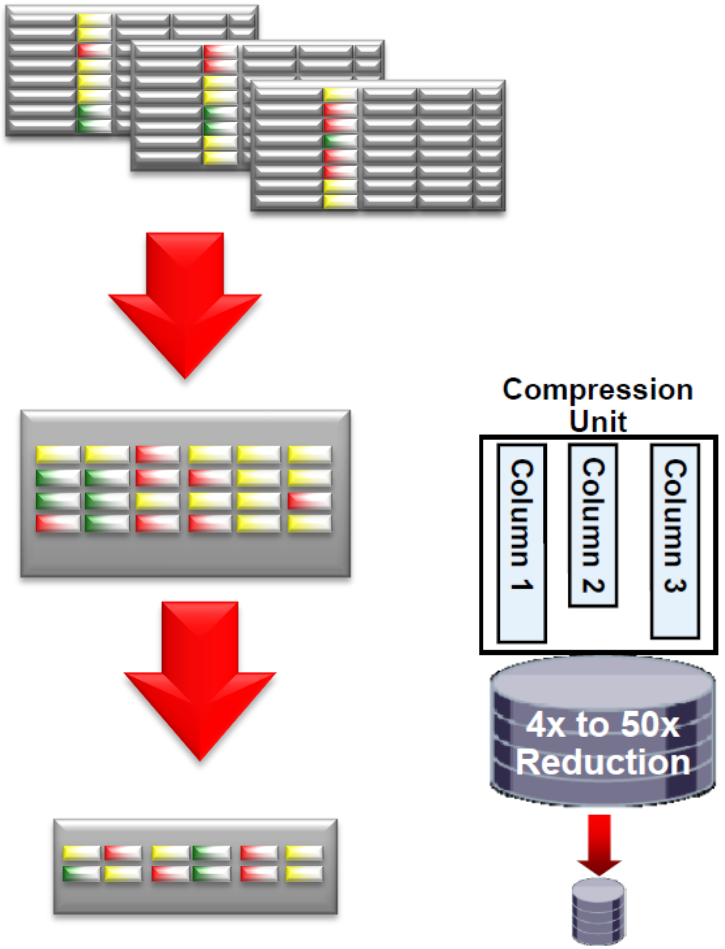
How will Smart Scan Reduce Volume of Data Returned to Each Server by each Workload?

- Smart scan will have different impact on volume of data send to RAC server
- We will use models to predict how smart scan will affect the response time and throughput for each workload



OLTP (1 – 1.5)
DW (2 – 10)
Archiving (50 – 500)

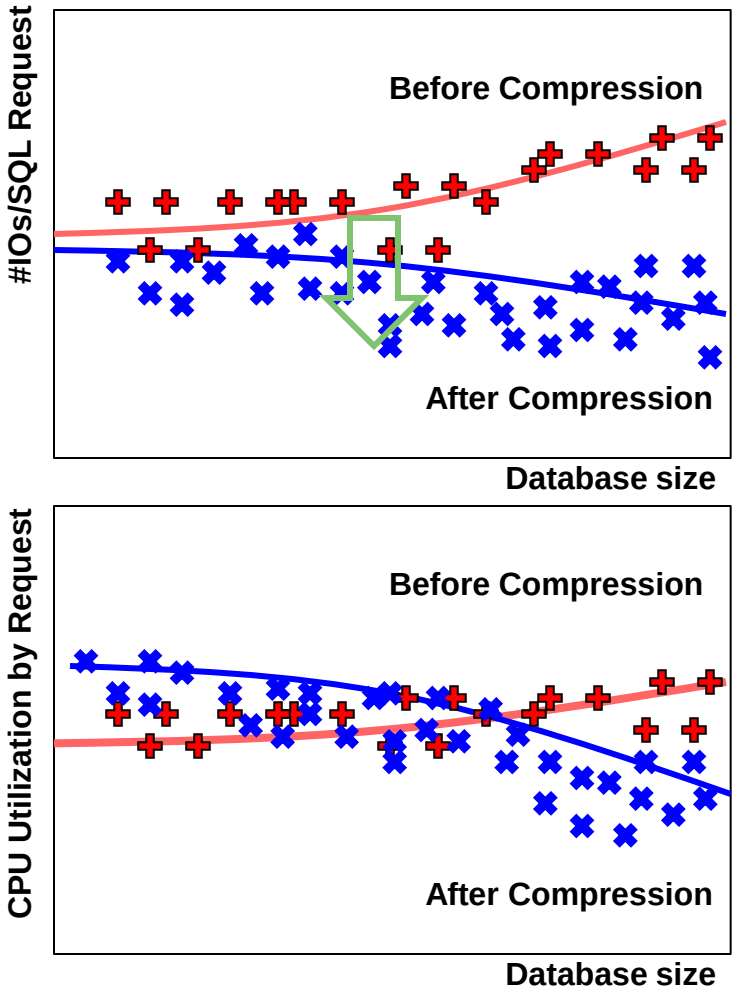
How will Exadata Hybrid Columnar Compression Affect Performance of Each Workload?



- Data is stored by column and then compressed
- Factors affecting a compression ratio:
  - Table size
  - Data cardinality
  - Read/write ratio

How Exadata Hybrid Columnar Compression Affects the # of I/Os, and CPU Utilization for Each Workload

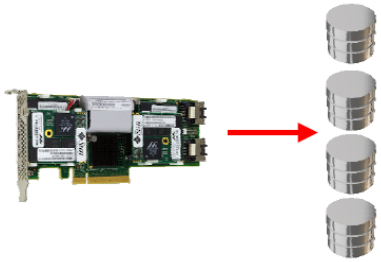
- Compression ratios depends on workload's profile:
  - OLTP (1 - - 6)
  - DSS (5 -- 30)
  - Archive (10 -- 50)
- Regression analysis shows the impact of Compression on #IO/SQL Request and CPU utilization
- We will use models to predict how compression will it affect the response time and throughput of each workload



How Will Smart Flash Cache Size Will Affect Performance of Each Workload?

Disk IO Service Time is about 3 ms and it can handle up to 300 IOPS

Flash Cache Service Time is small allowing 1mln iops and scan 50gb/sec



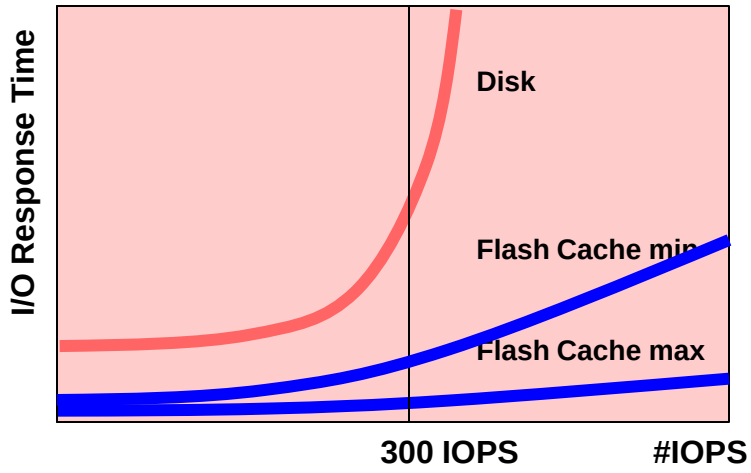
DBA control:
Alter Table Customer Storage (cell_flash_cache_keep/default/none)

- **Disk utilization as well as Flash Cache utilization depends on IO Rate and I/O Service time :**

$$Udisk = IO\ Rate * I/O\ Service\ Time$$

- **I/O Response time depends on I/O Service Time and Disk or Flash Cache Utilization" :**

$$I/O\ Response\ Time = \frac{I/O\ Service\ Time}{(1 - Udisk)}$$



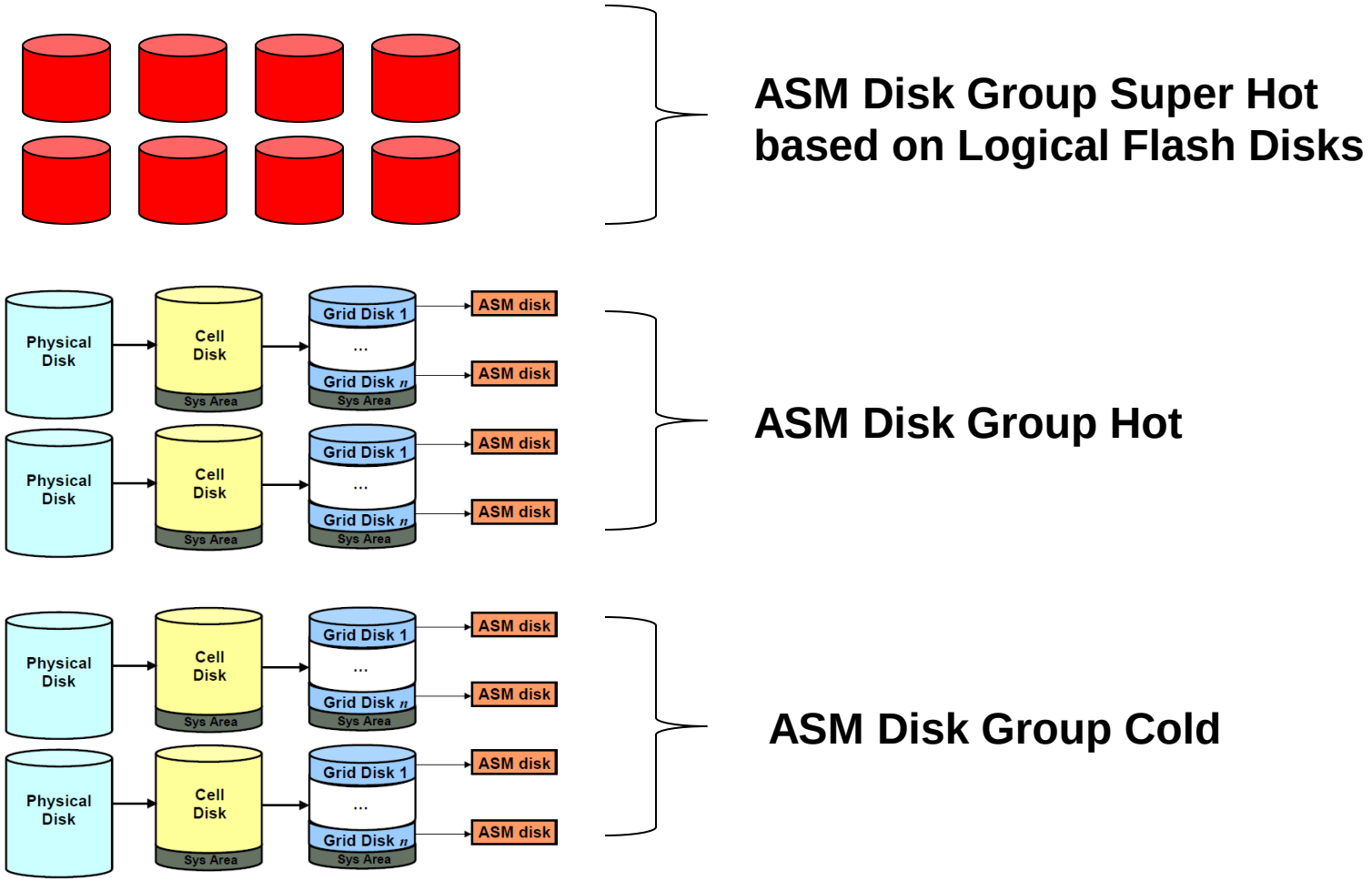
- **A lot of value for OLTP workloads**
- **Cost is a limitation**

Exadata I/O Resource Management in Multi-Database Environment

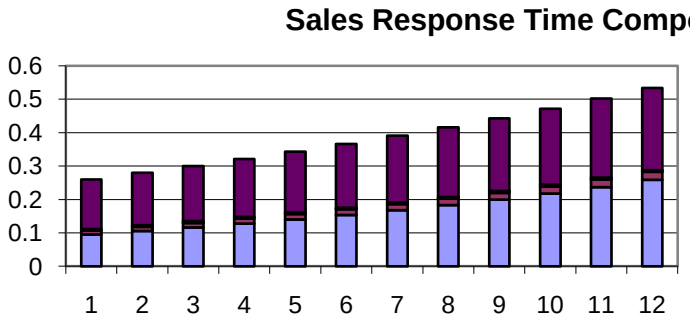
How to optimize allocation of I/O resources / bandwidth to different Databases and Workloads

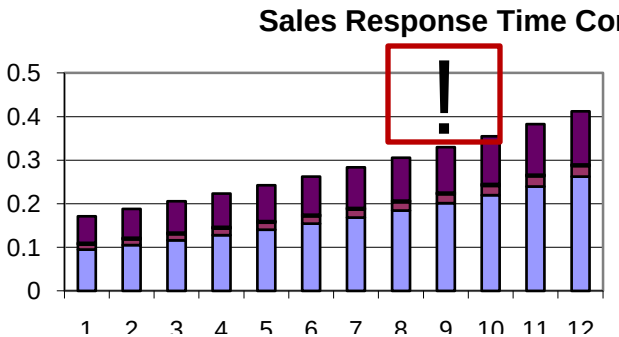
- Database Sales 40% I/O resources
  - OLTP 70% of I/O recourses
  - ETL : 30% of I/O resources
- Database Marketing: 35% of I/O resources
  - Interactive: 45% of I/O resources
  - Batch: 55% of I/O resources
- Database HR: 25% of I/O resources

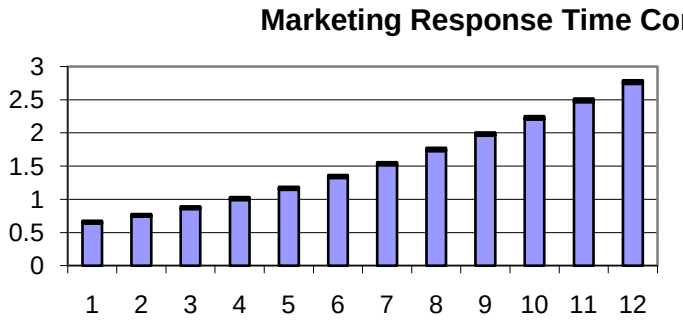
What will be an Impact on Cost/Performance as a Result of Moving Data between ASM Disk Groups: Super Hot, Hot and Cold

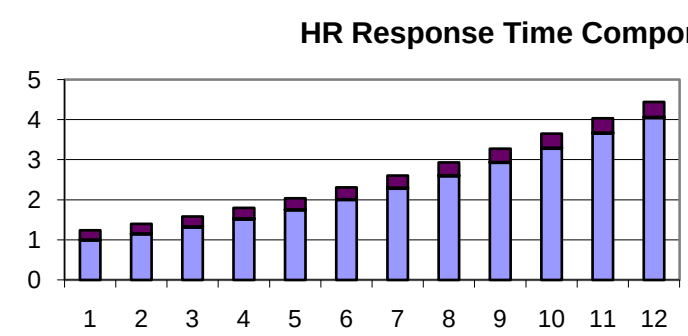


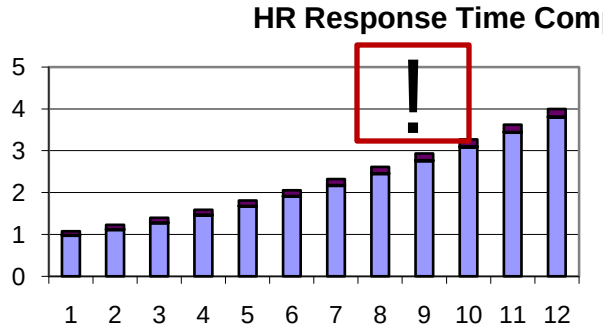
Predicted Server Consolidation Impact on DB Machine



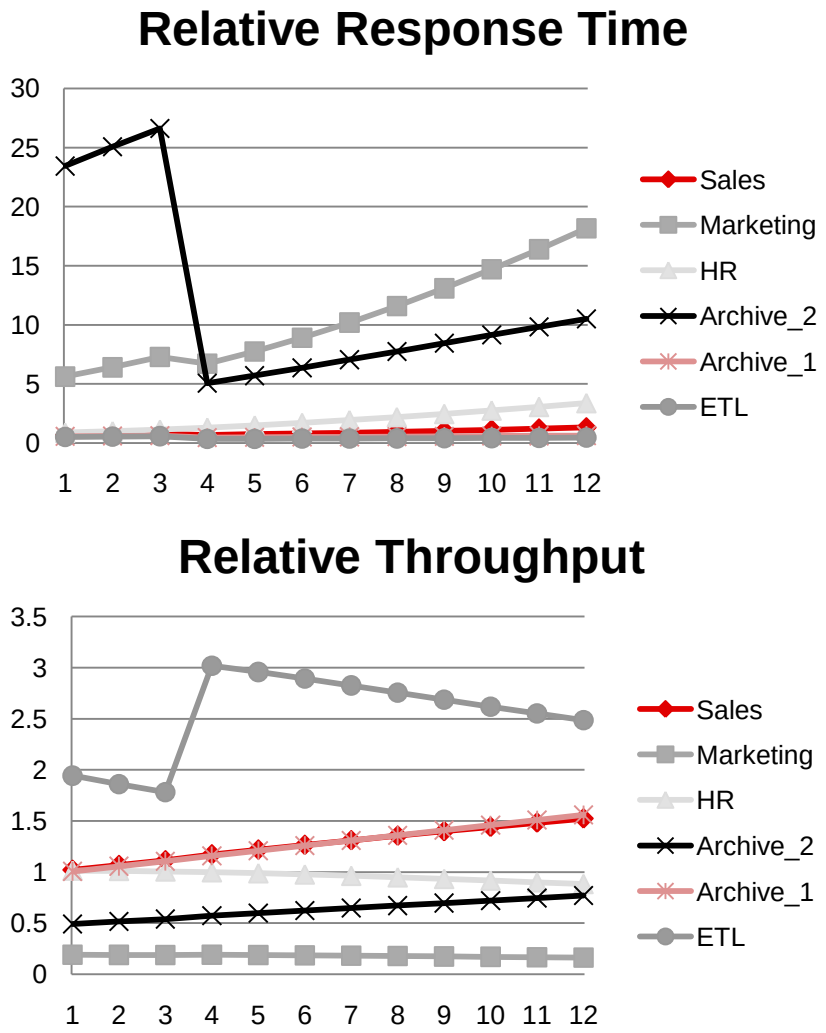




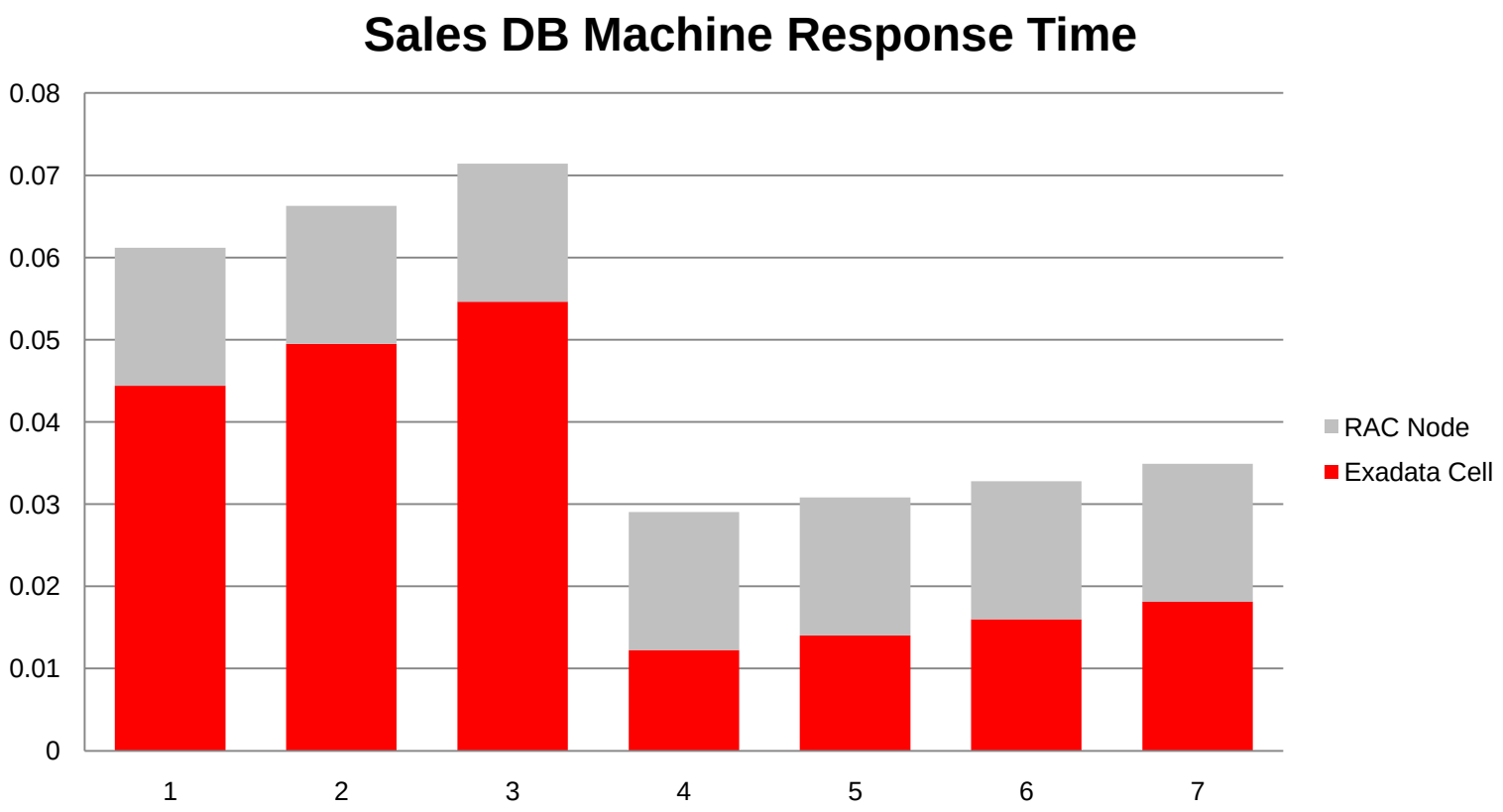




What Will be the Impact of Increasing Number of Exadata Cells?



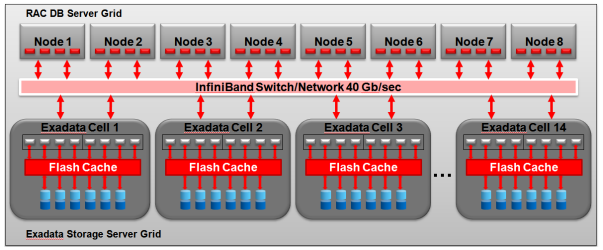
Predicted Impact of Adding Exadata Cells on Sales Response Time



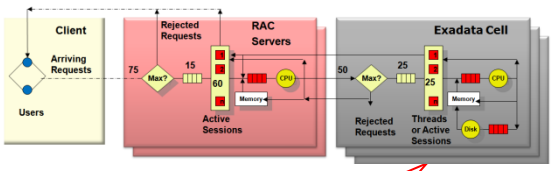
How to Predict New Application Implementation Impact

Stress Testing

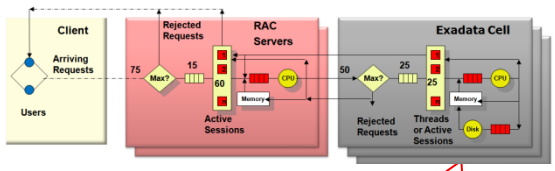
Production DB Machine



New Application



Copy New Workload



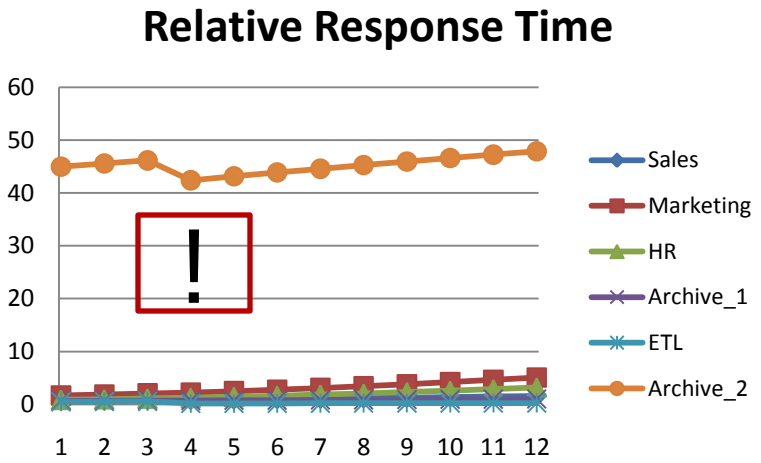
Build Model of the Test System

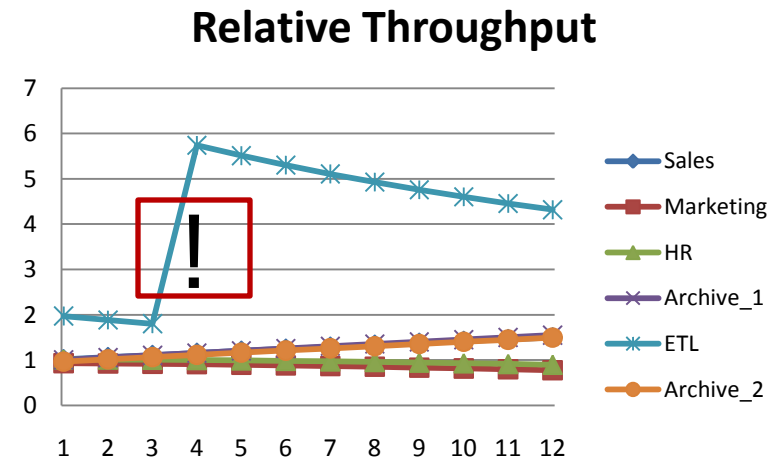
Build Model of the Production System

Predict how new application will affect performance of existing applications

Predict how new application will perform in production environment

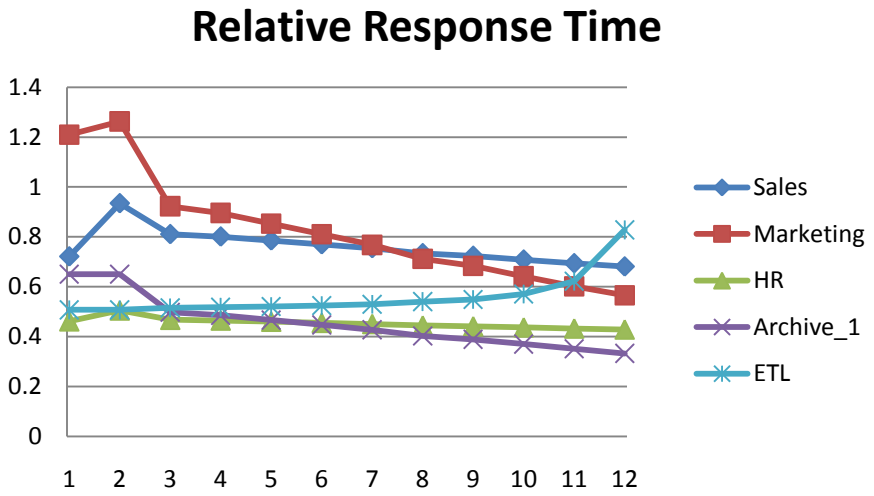
Predicted Impact of Database Tuning

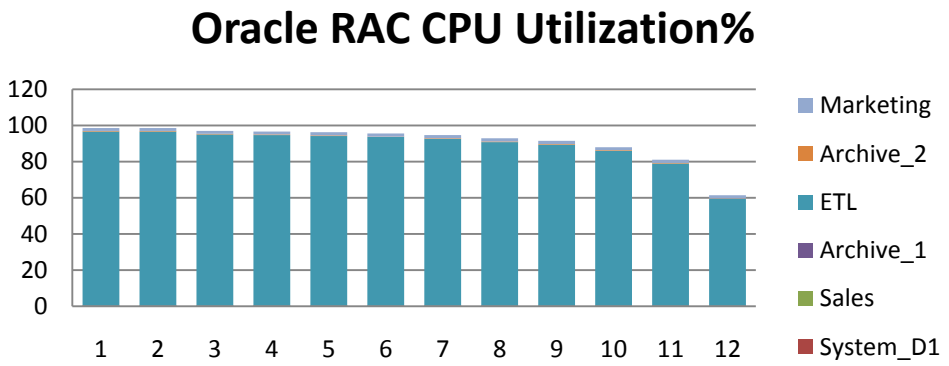


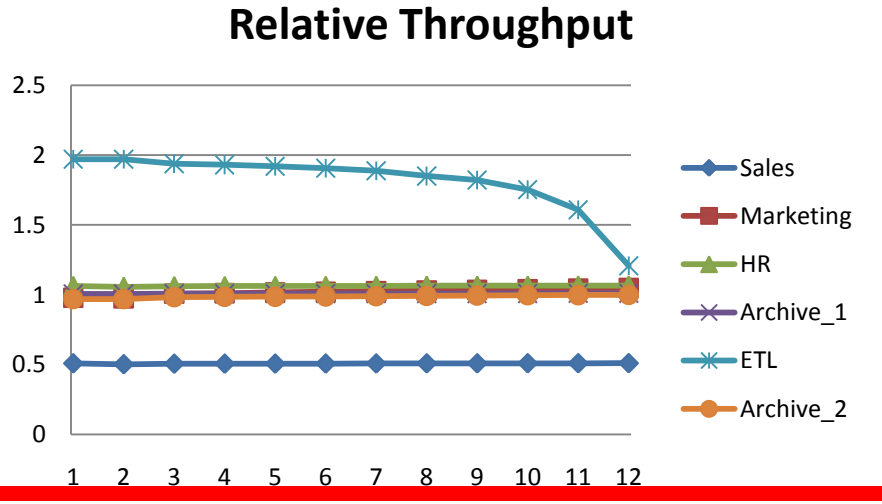


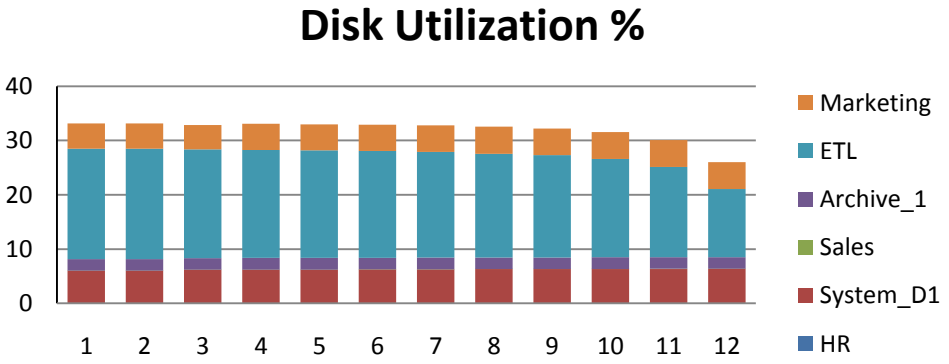
- **New Index?**
- **Materialized view?**
- **Data Compression?**

Predicted Impact of Reducing ETL Concurrency From 110 to 10 Starting P3

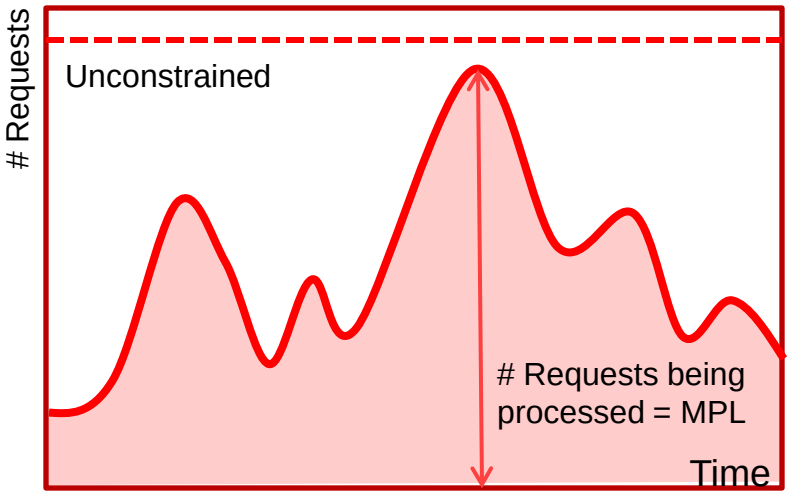


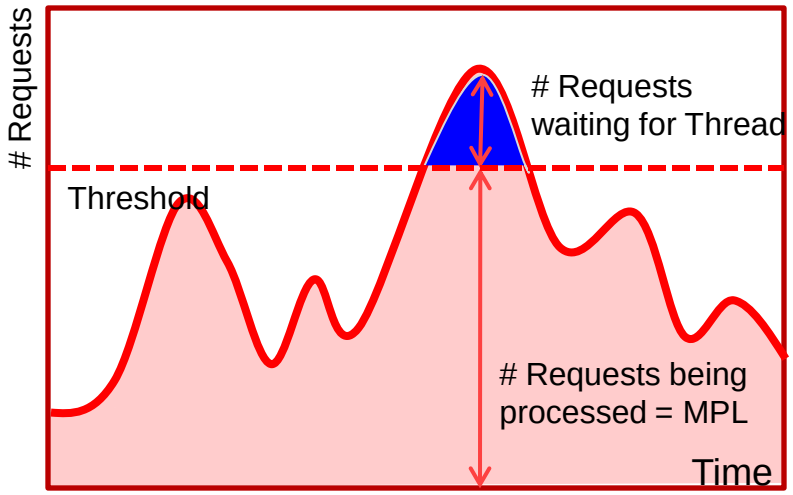


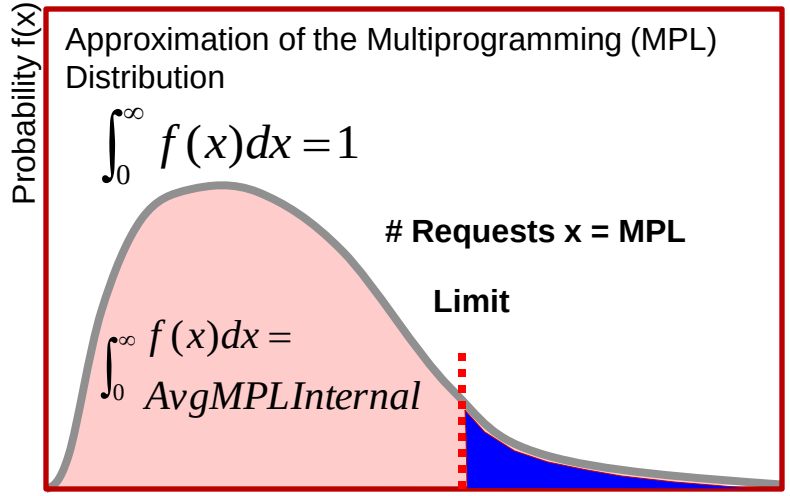


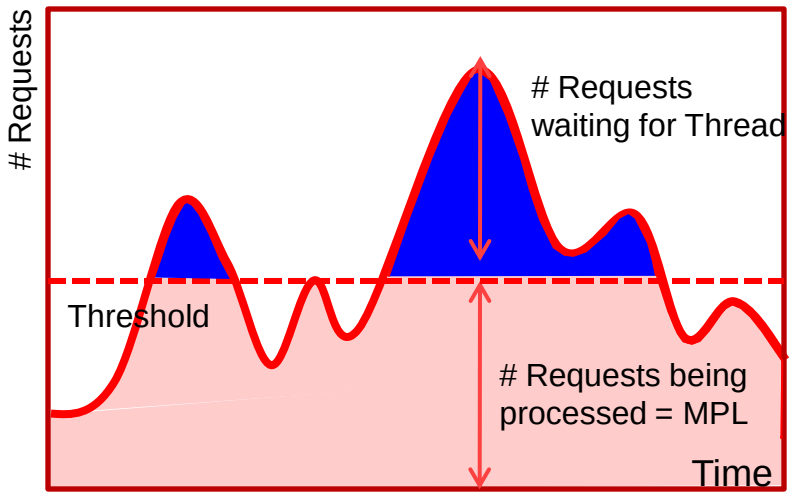


Limit Concurrency Reduces Contention but Increase # of Requests Waiting for the Thread



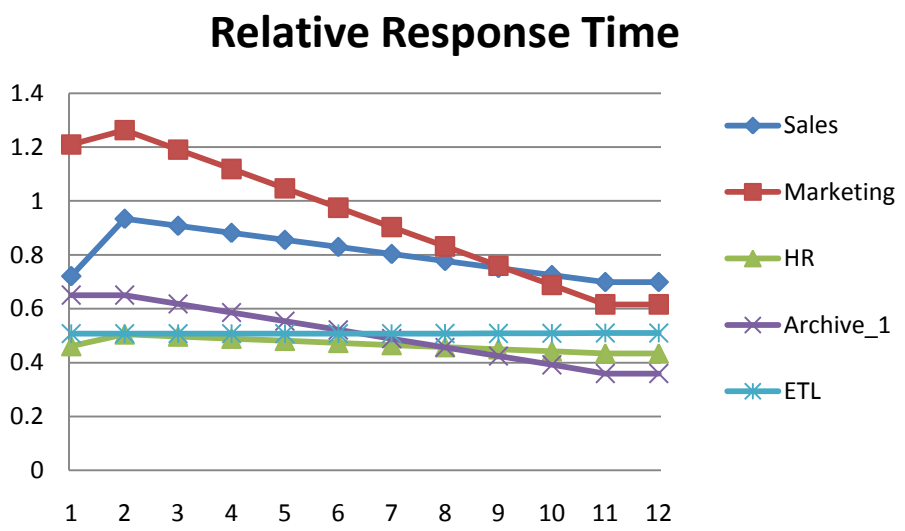


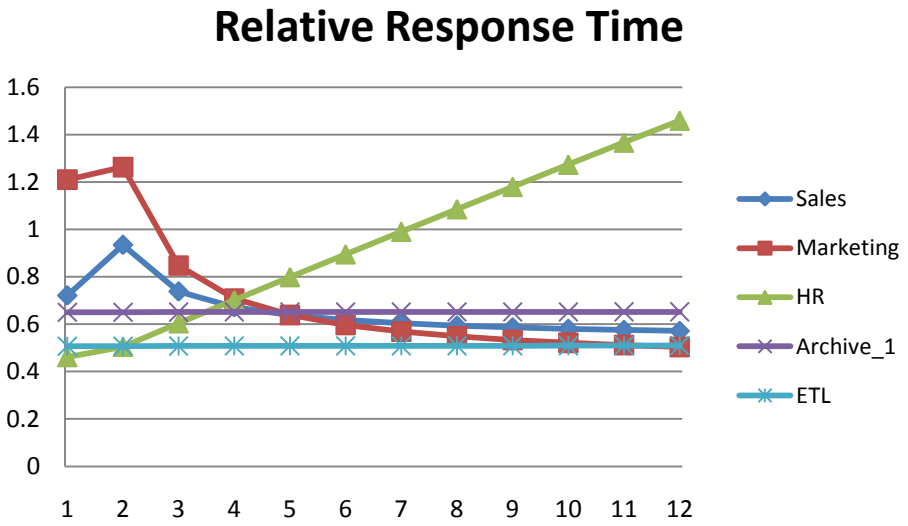




$$\int_{Limit}^\infty f(x)dx = AverageWaitngQueue$$

Predicted Impact of Changing Workloads' Priority

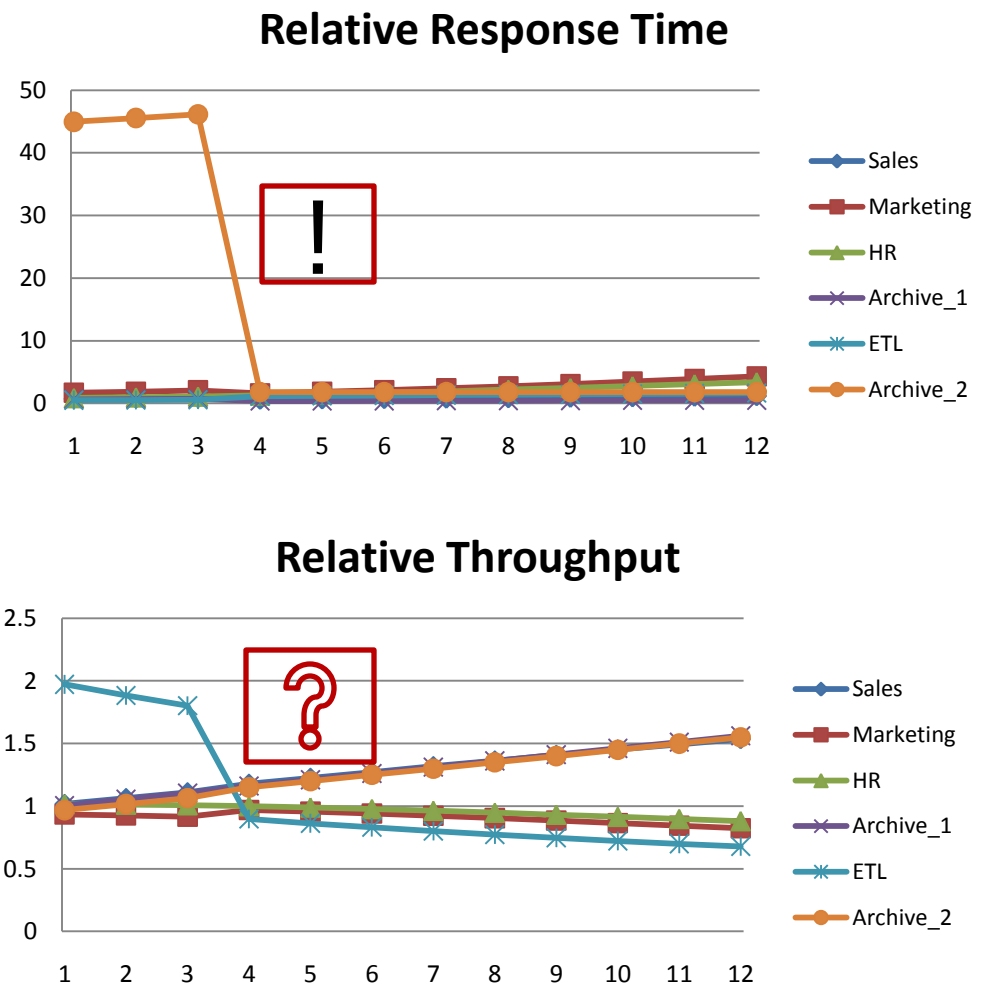




Reducing ETL Priority from 0.9 to 0.1 Starting P3

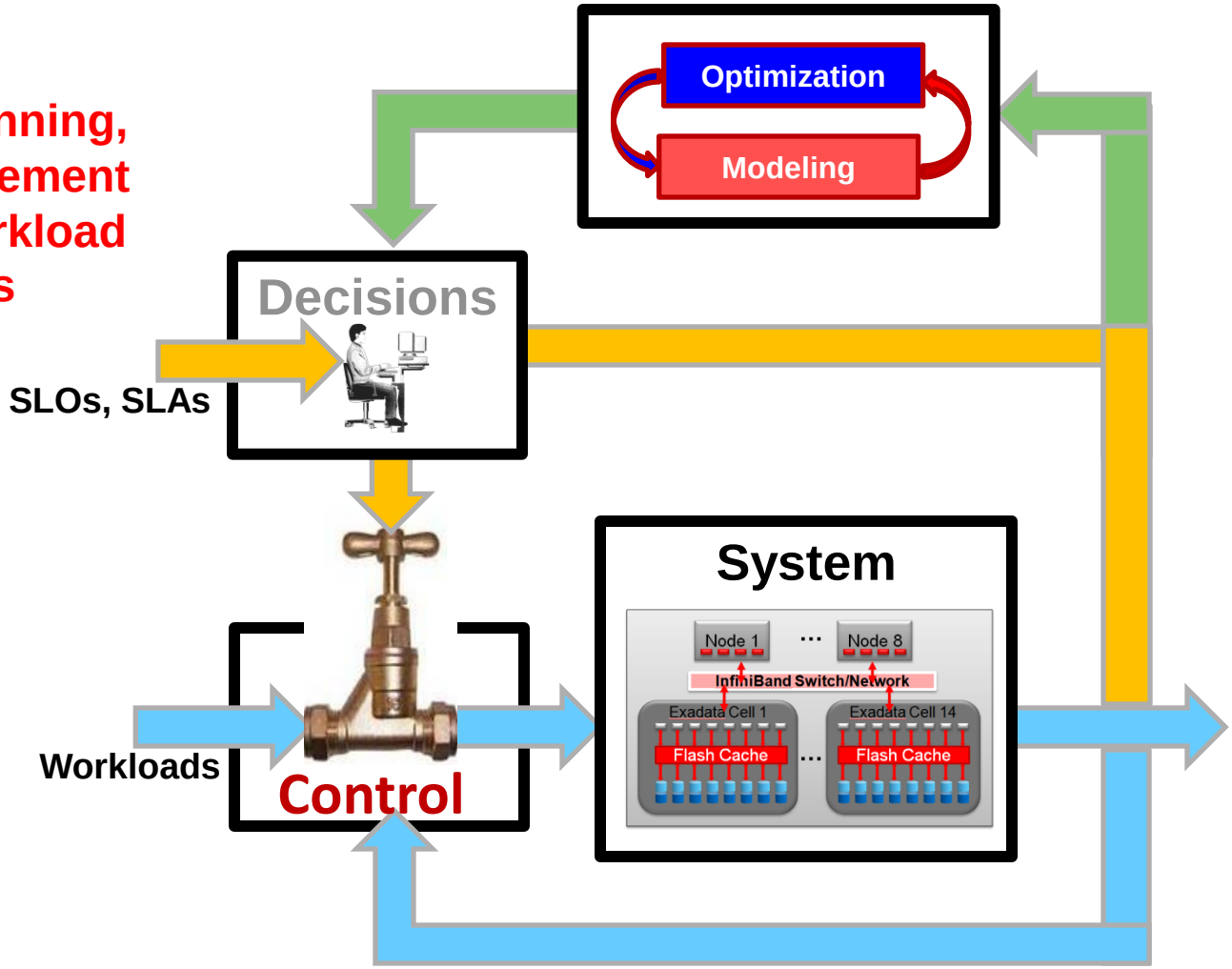
Increasing Marketing and Sales Priority from 2 to 11 Starting P3

Predicted Impact of Limiting CPU Consumption by ETL Workload to 50%?

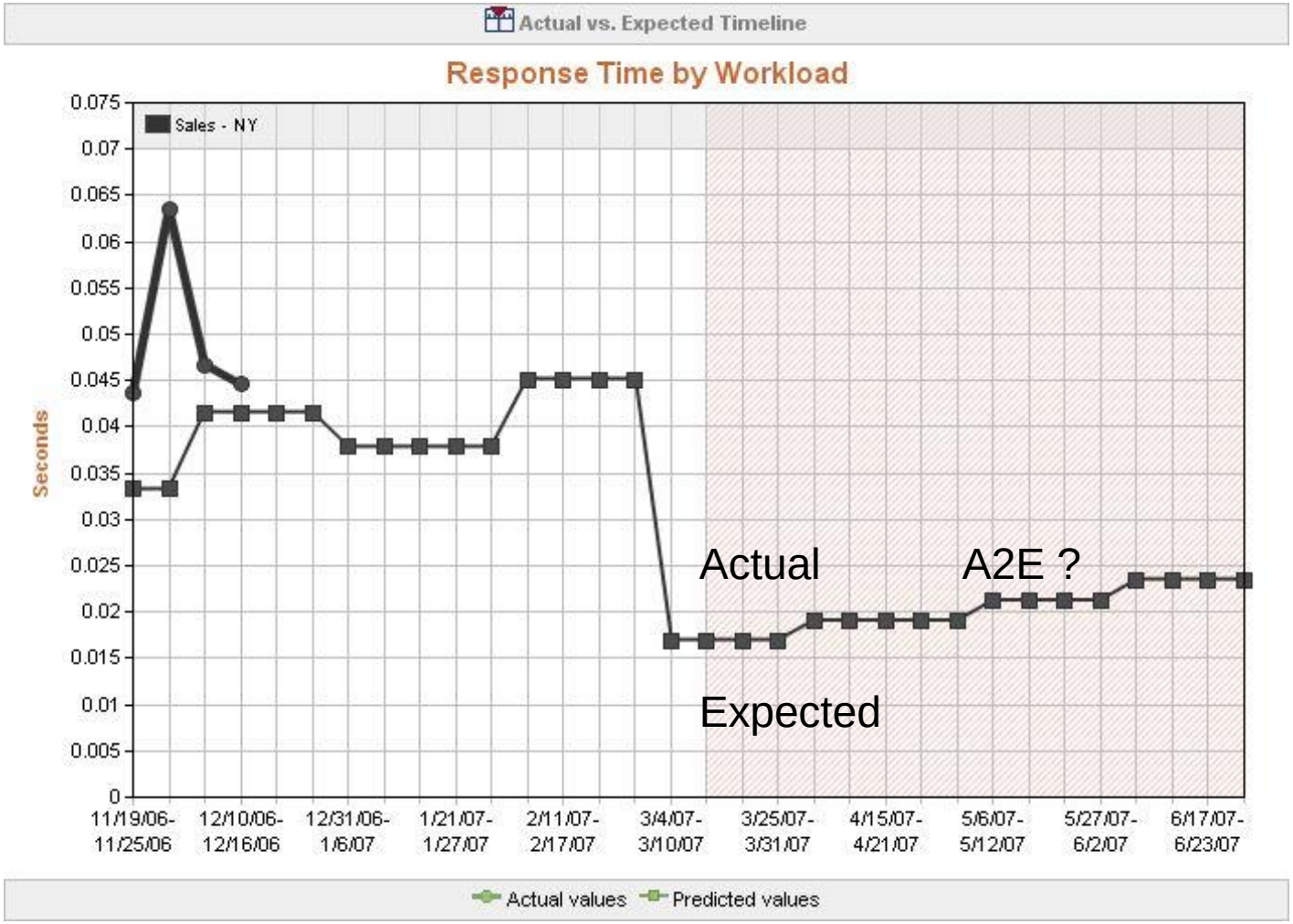


Role of Modeling and Optimization

- **Justify Capacity Planning, Performance Management and Operational Workload Management Actions**



Organizing Continuous Proactive Performance Management Process



Conclusion

1. Oracle Database Machine smart scan, columnar data compression and flash memory affect scalability of OLTP and DW workloads differently
2. Capacity management decisions based on gut feelings can be misleading
3. Rules of thumb do not take into consideration specifics of your environment
4. Benchmarks produce the most accurate results, but benchmarks are expensive, time consuming and not flexible
5. We demonstrated a value of predictive analytics in evaluating options and justification of capacity planning, performance management and workload management decisions
6. Collaborative efforts between business people and IT in workload forecasting and evaluation of results helps to understand how complex system works
7. Prediction results set realistic expectations and enable comparison of the actual results with expected
8. It reduces an uncertainty and minimizes risk of surprises

Thank you!

Contact Information

boris@bez.com

www.bez.com

bezsys.blogspot.com

References

- B. Zibitsker, A. Lupersolsky , IOUG 2009, “Modeling and Optimization in Virtualized Multi-tier Distributed Environment”
- B. Zibitsker, IOUG 2008. “Reducing Risk of Surprises in Changing Multi-tier Distributed Oracle RAC Environment”
- B. Zibitsker, DAMA 2007, “Enterprise Data Management and Optimization”
- B. Zibitsker, CMG 2008, 2009 “Hands on Workshop on Performance Prediction for Virtualized Multi-tier Distributed Environments”
- J. Buzen, B. Zibitsker, CMG 2006, “Challenges of Performance Prediction in Multi-tier Parallel Processing Environments”
- B, Zibitsker, G. Sigalov, A. Lupersolsky “Modeling and Proactive Performance Management of Multi-tier Distributed Environments”, International conference “Mathematical methods for analysis and optimization of information and telecommunication networks"
- B. Zibitsker, C. Garry, CMG 2009, "Capacity Management Challenges for the Oracle Database Machine: Exadata v2“
- Oracle Enterprise Manager Grid Control documentation library at:
  http://www.oracle.com/technology/documentation/oem.html