A Flexible Data Warehouse Architecture

Building the Ideal Data Warehouse Platform

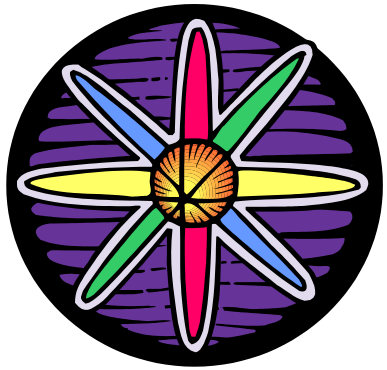
Mike Ault

Oracle Guru

Texas Memory Systems

NYOUG Dec 2010

Michael R. Ault
Oracle Guru



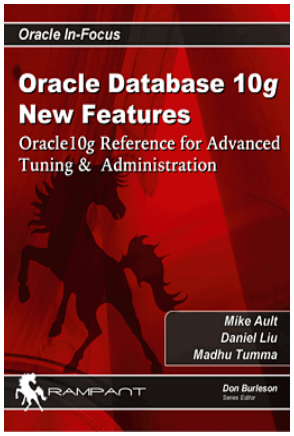
- Nuclear Navy 6 years
- Nuclear Chemist/Programmer 10 years
- Kennedy Western University Graduate
- Bachelors Degree Computer Science
- Certified in all Oracle Versions Since 6
- Oracle DBA, author, since 1990

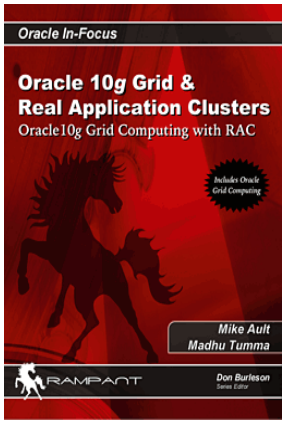
ORACLE | CERTIFIED PROFESSIONAL

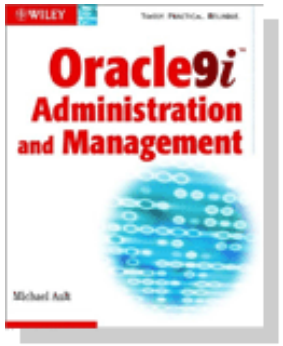
Texas Memory Systems

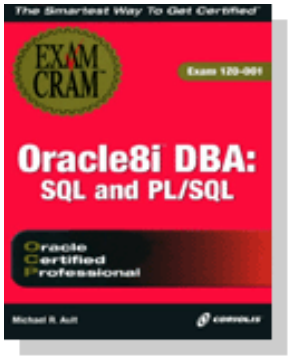
Texas Memory Systems

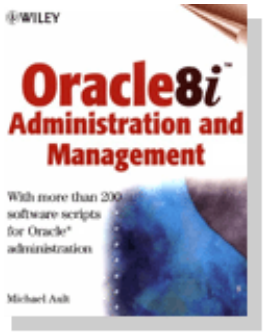
Books by Michael R. Ault

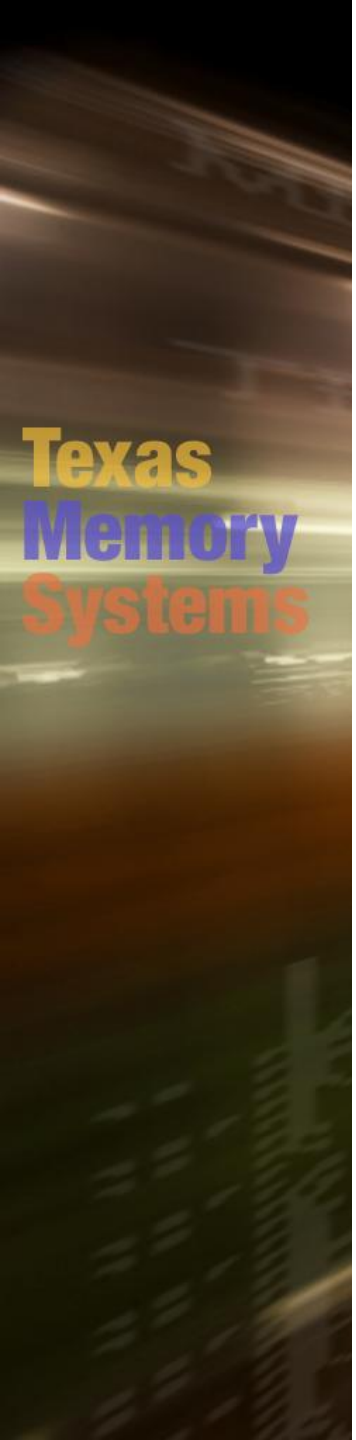












Statspackanalyzer.com

Free Statspack/AWR Analysis

Sponsored by Texas Memory Systems

-Looks for IO bottlenecks and other configuration issues.

-Straightforward tuning advice

Texas Memory Systems

General Requirements

- Flexible Architecture
  - Can use multiple OS
  - Can use multiple Databases
- Not locked to a single DBS or OS
- Expandable
- Low cost (relatively speaking)

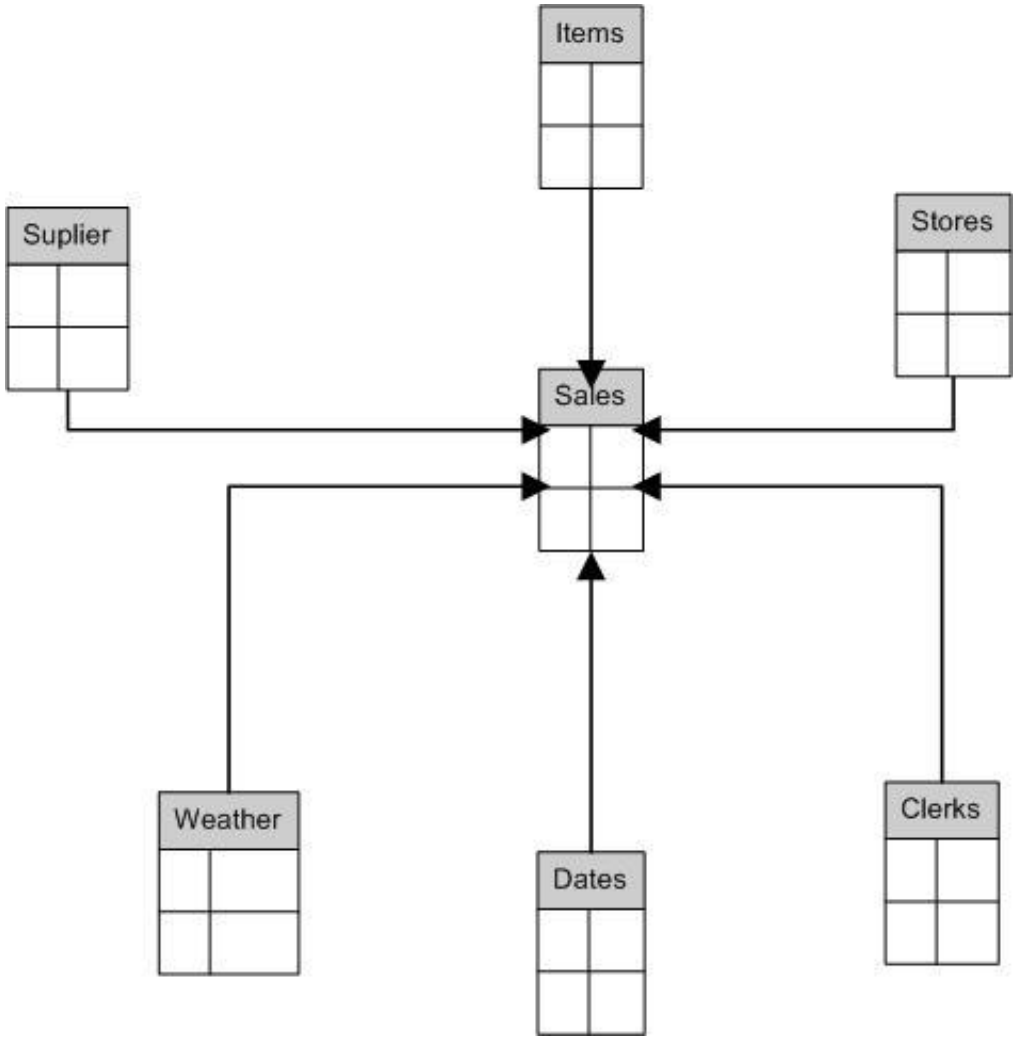
Texas Memory Systems

What is a Data Warehouse?

- It is not just a large database
- It has a specific purpose
- It has a specific design

Texas Memory Systems

Typical Data Warehouse Design



Texas Memory Systems

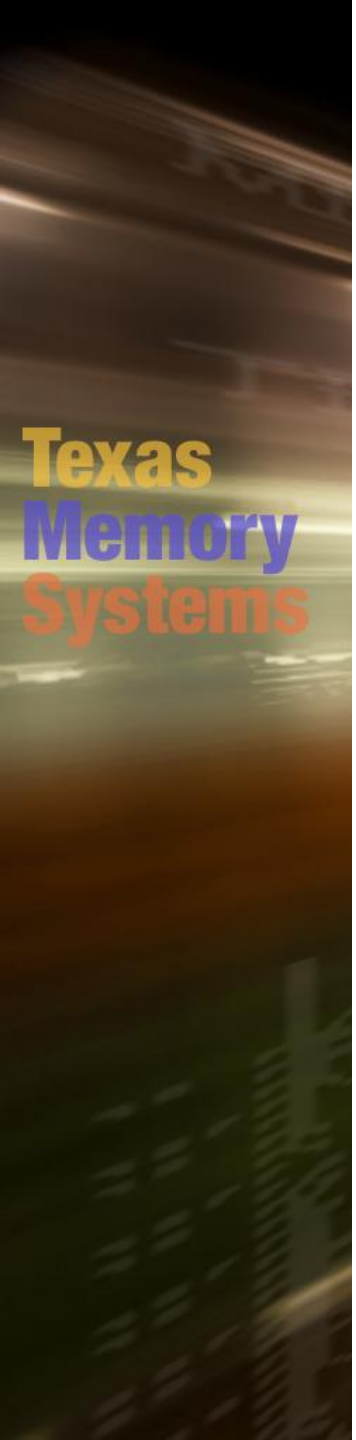
Data Warehouse Characteristics

- Usually very large (hundreds of GB to TB)
- Structured for data retrieval
- Provides for Summaries
- Requires large numbers of IOPS
- Usually only a small number of users

Texas Memory Systems

DWH IO Characteristics

- Usually access is dimensions into facts
- Bitmap index merges usually most efficient
- Called a star join
- May have full table scans
- Can exceed 200,000 IOPS for large joins with sorts



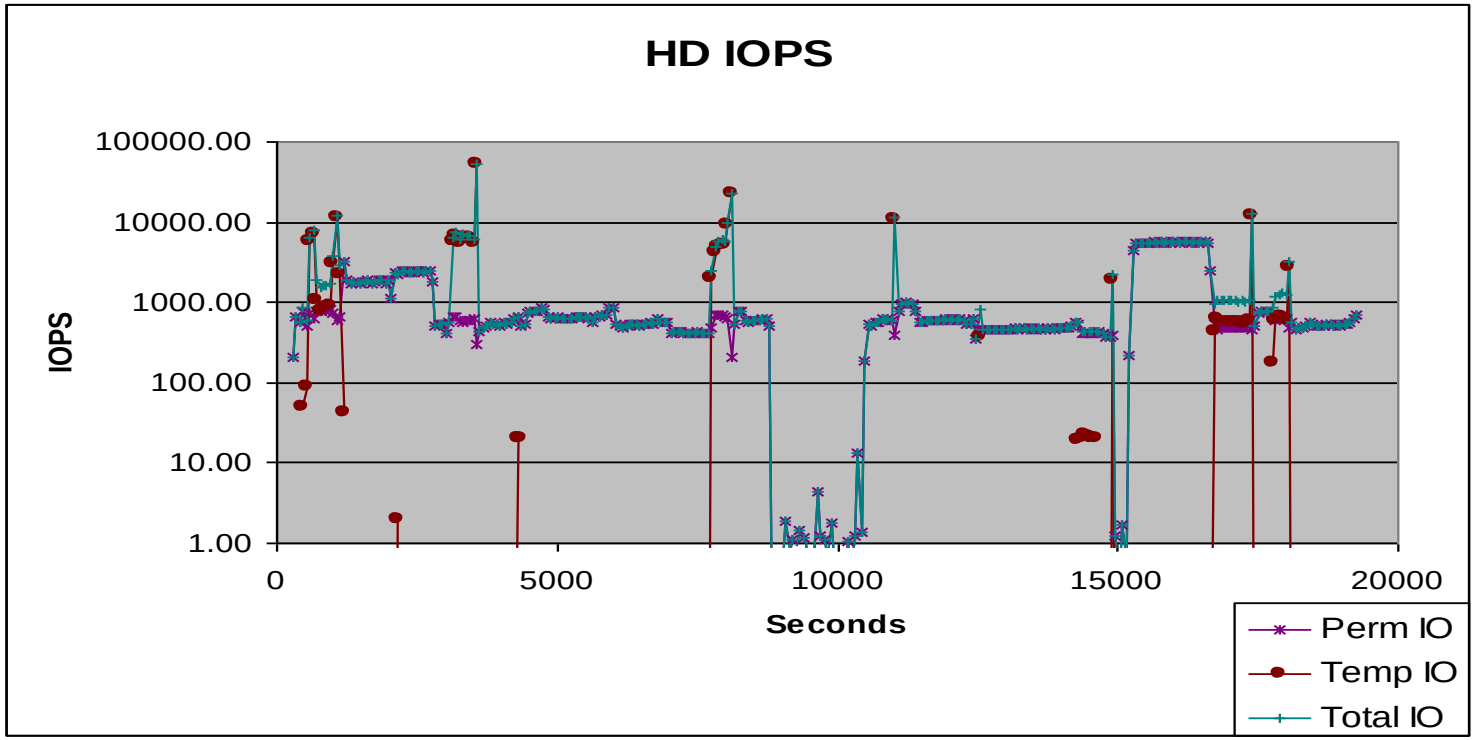
Hard Disk IO Characteristics

- Limited by physics
- Maximum of 200 random IOPS
- Latency of 2-5 milliseconds for a 15K rpm drive
- Regardless of storage capacity!

Texas Memory Systems

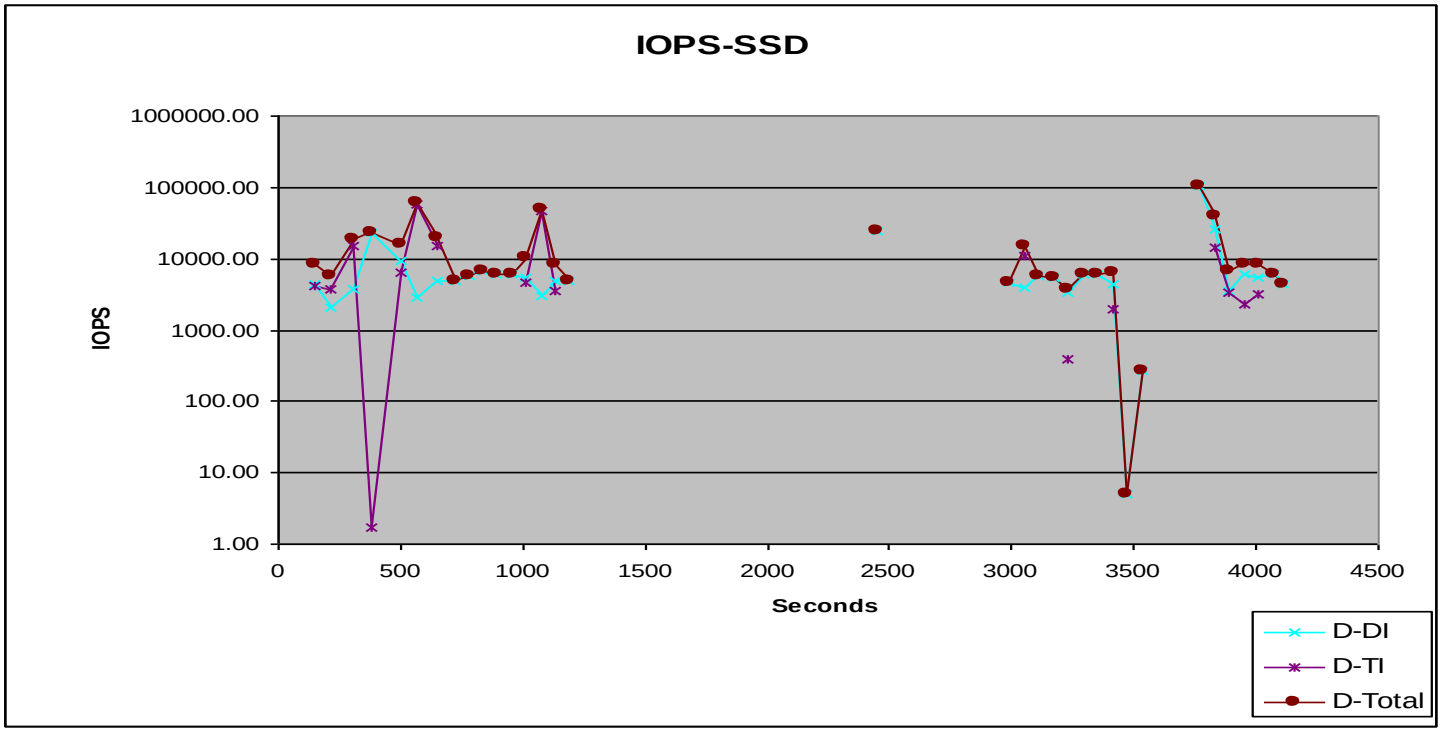
Hard Drive IOPS (28 – 15K)

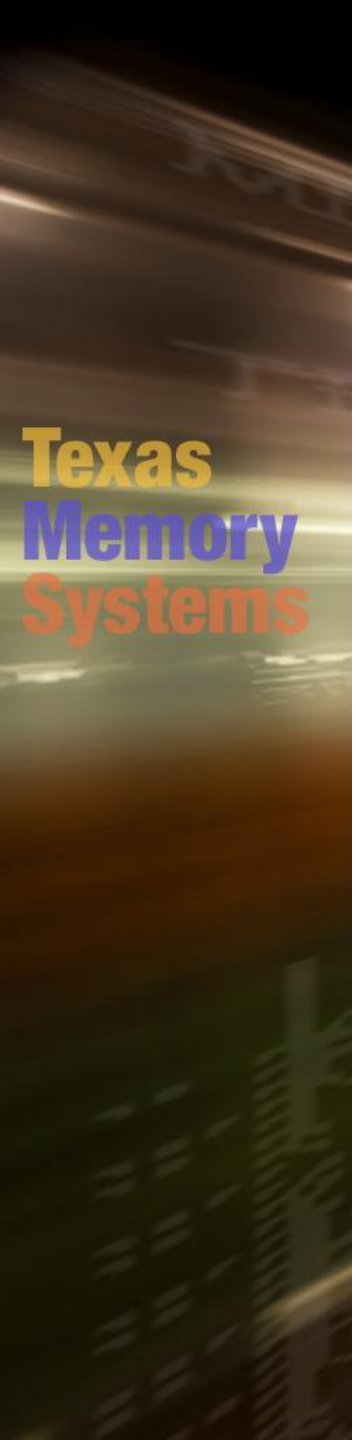
- 28 – 15K rpm 146GB drives
- 300 Gigabyte data, 600 Gigabyte total TPC-H profile



SSD IOPS Profile

- 1-1 TB Flash unit
- 2- 128 GB DDR units
- Same DB and query set



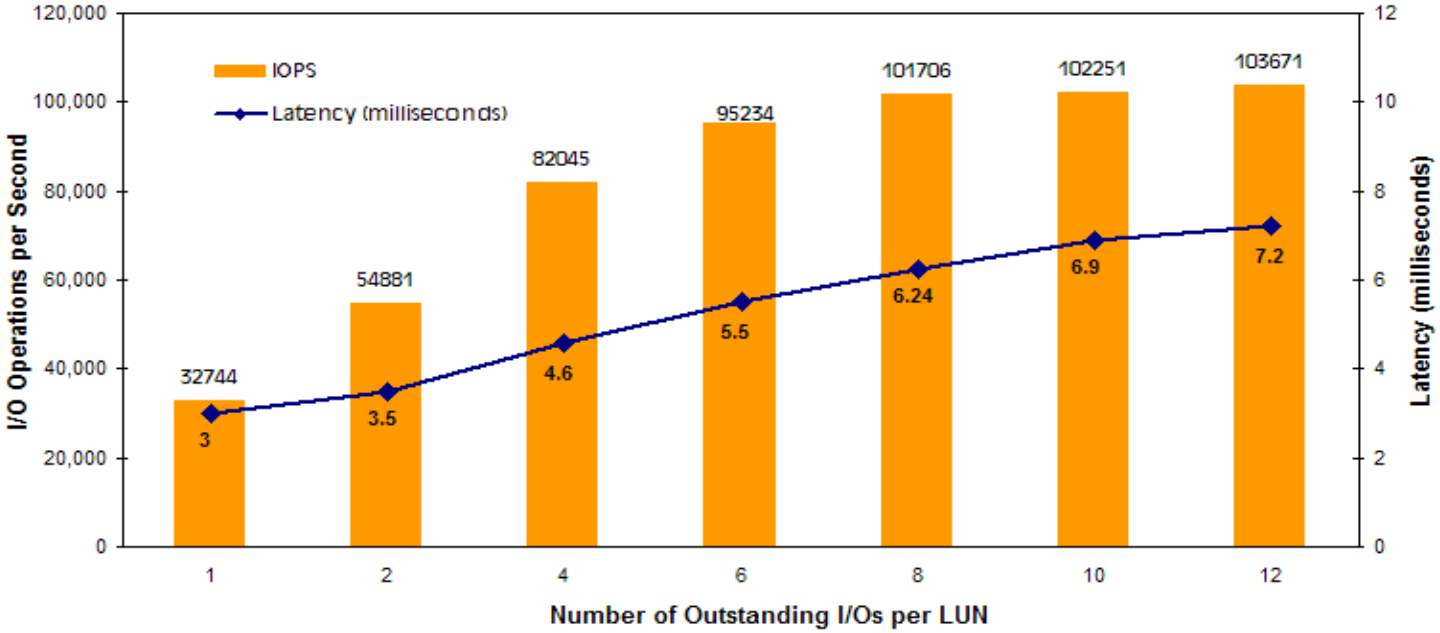


To Get to 100,000 IOPS

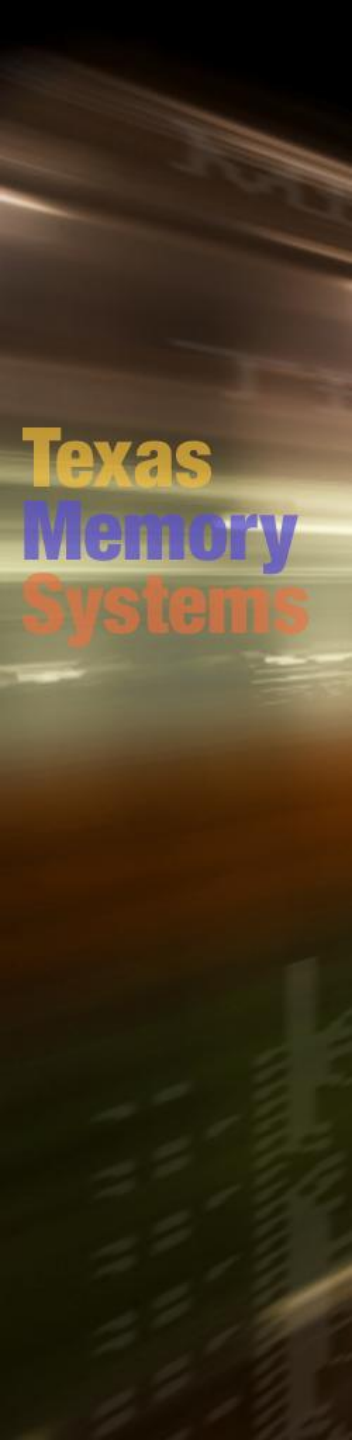
- You need about 500 hard drives
- 3-CX3 from EMC should do it…
- However…

Texas Memory Systems

Can't get away from the Latency!



EMC IOPS and Latency
(From: http://blogs.vmware.com/performance/2008/05/100000-io-opera.html)

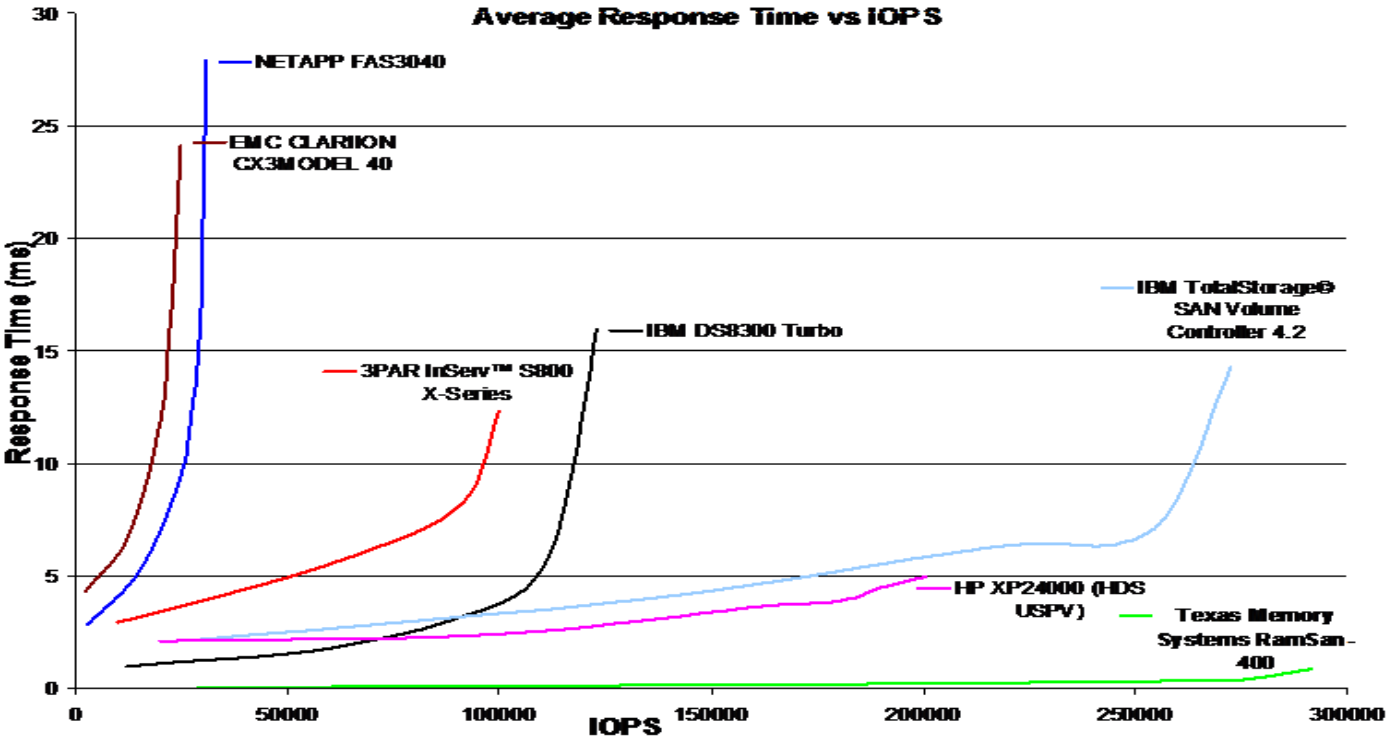


Why Am I Concerned?

- Latency will kill DWH performance
- Can’t get below 2-5ms latency with disks
- Have to get rid of physical movement to retrieve data

Texas Memory Systems

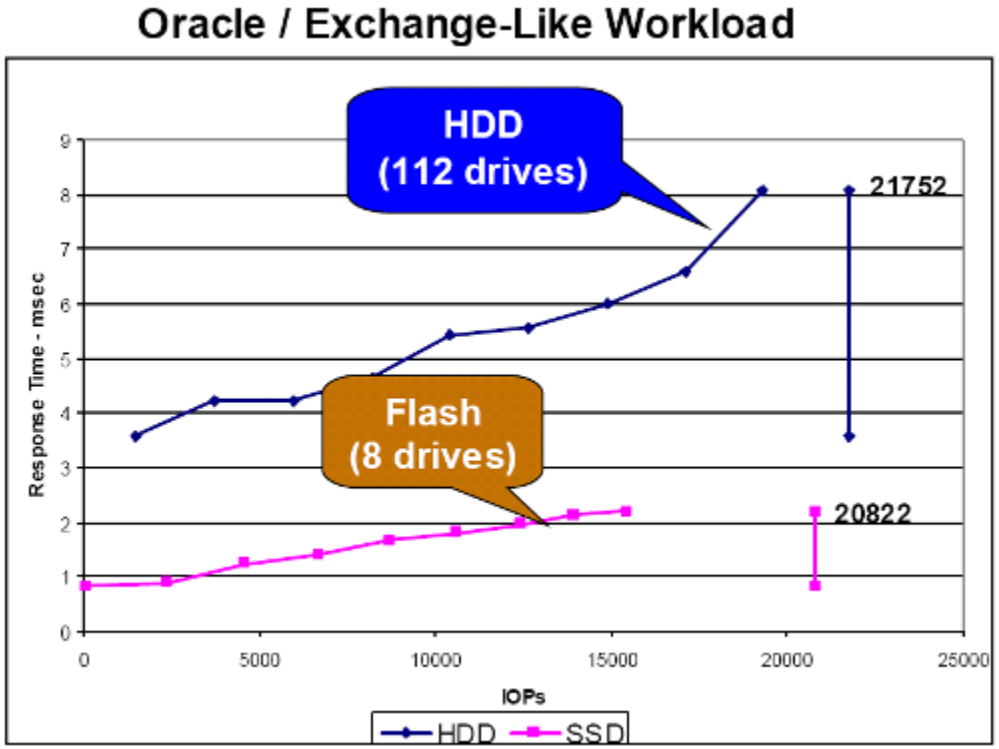
Comparative Latencies



IOPS Comparison for Various SANs
(Source: www.storageperformance.org)

New SSD Technology

- EMC putting SSD into “disk drive” size package 4 per tray maximum

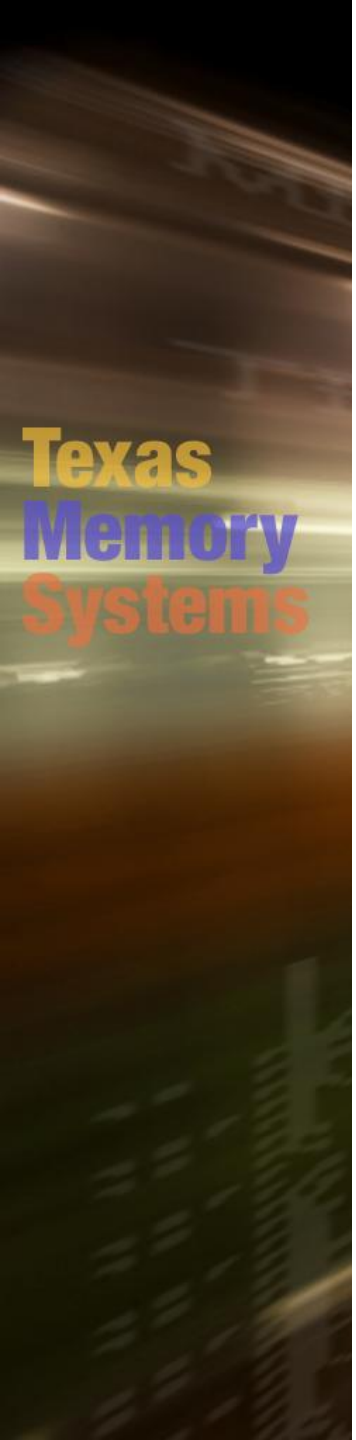


EMC SSD Response time: 1-2 MS EMC HDD Response time: 4-8 MS (Source: ”EMC Tech Talk: Enterprise Flash Drives”, Barry A. Burke, *Chief Strategy Officer, Symmetrix Product Group, EMC Storage Division, June 25, 2008*

Texas Memory Systems

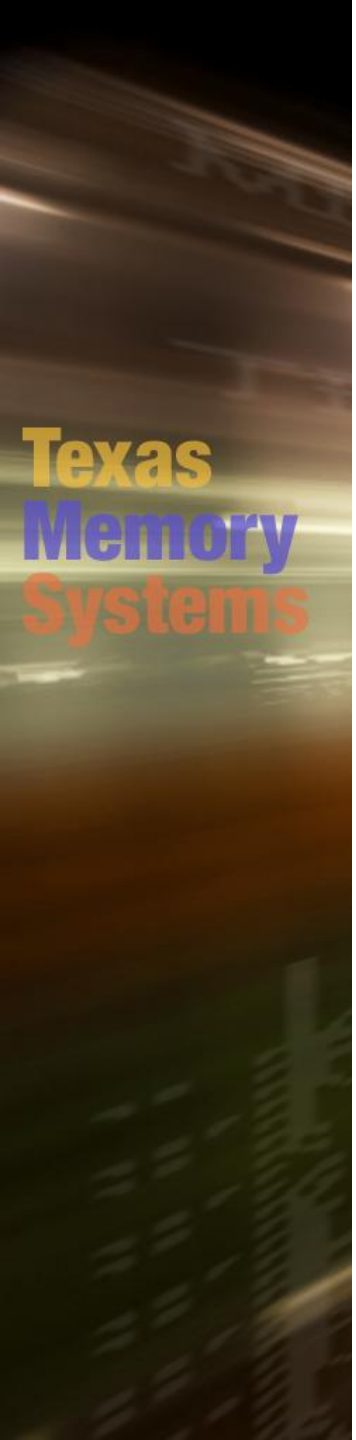
So…

- Disks too slow (4-8 ms)
- Even new SSD technology in disk format is slow (1-2 ms)
- Performance oriented SSD best choice (.015-0.2 ms)



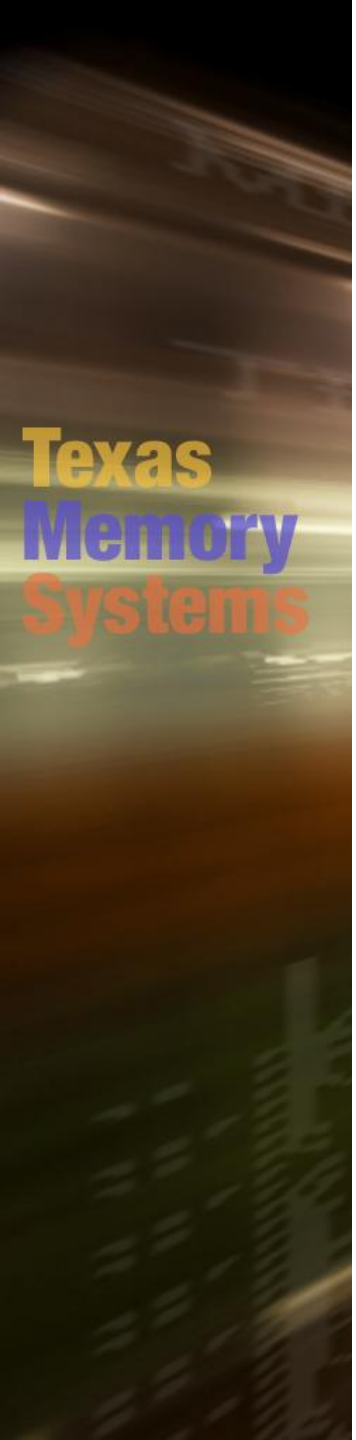
DWH Processing Characteristics

- Sorts
- Summaries
- Parallel Operations
- Need multiple CPUS
- Need expansion capabilities



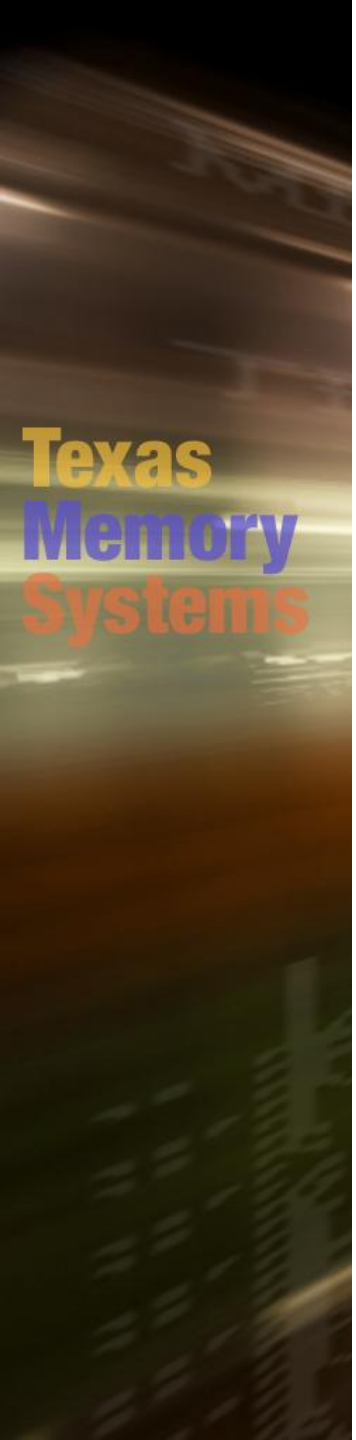
Available Server Technology

- Multi-CPU Blade Servers with RAC
- Large Multi-CPU Monolithic servers
- Multi-CPU Individual Servers with RAC



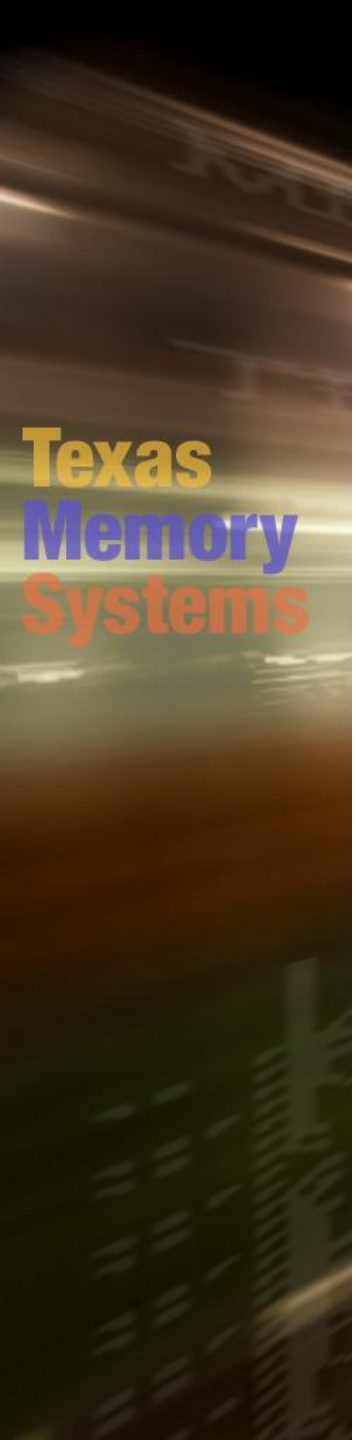
So What About Blades

- Flexible (within limits)
- Can’t switch to other manufacturer
- Limited to their upgrade/release cycle
- May have to dump entire stack if new architecture comes along
- Flexible to DB
- Flexible to OS



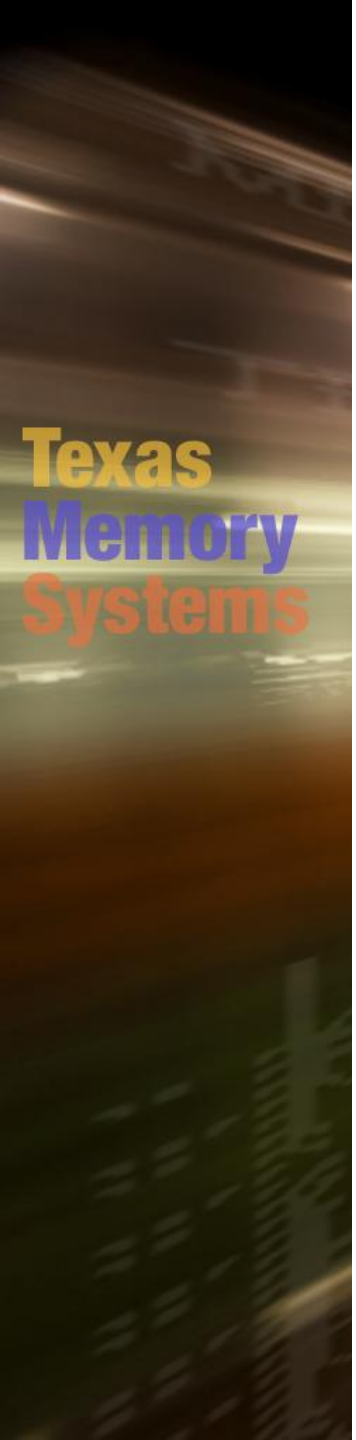
How About Monolithic Severs?

- Not flexible
- Locked to a single vendor
- Need to toss the whole thing if new architecture comes along
- Limited growth
- Flexible to DB (sometimes)
- Flexible to OS (sometimes)



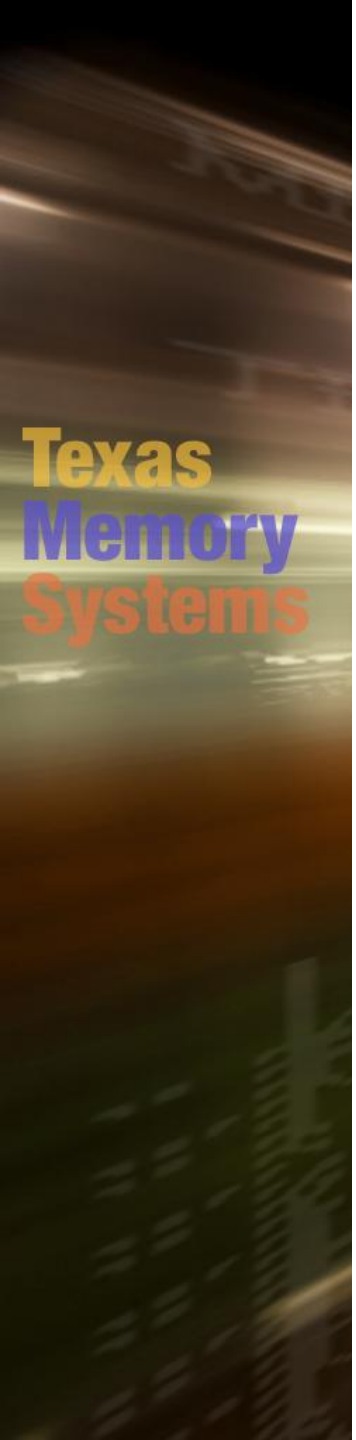
Individual Servers

- Each is self-contained
- Can be upgraded individually
- Flexible to OS
- Flexible to DB
- Expandable



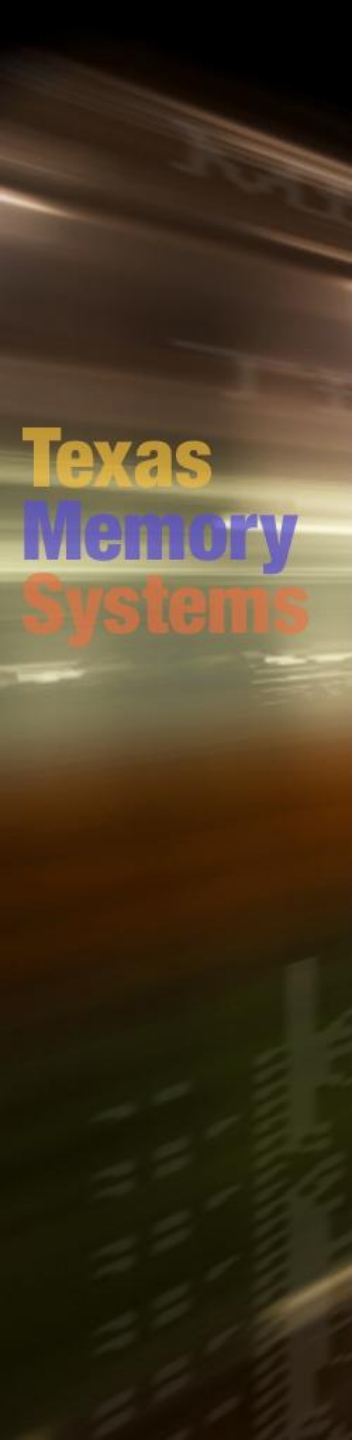
So… So Far…We Need:

- Large data storage capacity
- Able to sustain large numbers of IOPS
- Able to support high degrees of parallel processing (supports large numbers of CPUs)
- Large core memory for each server/CPU
- Easily increase data size and IOPS capacity
- Easily increase processing capability



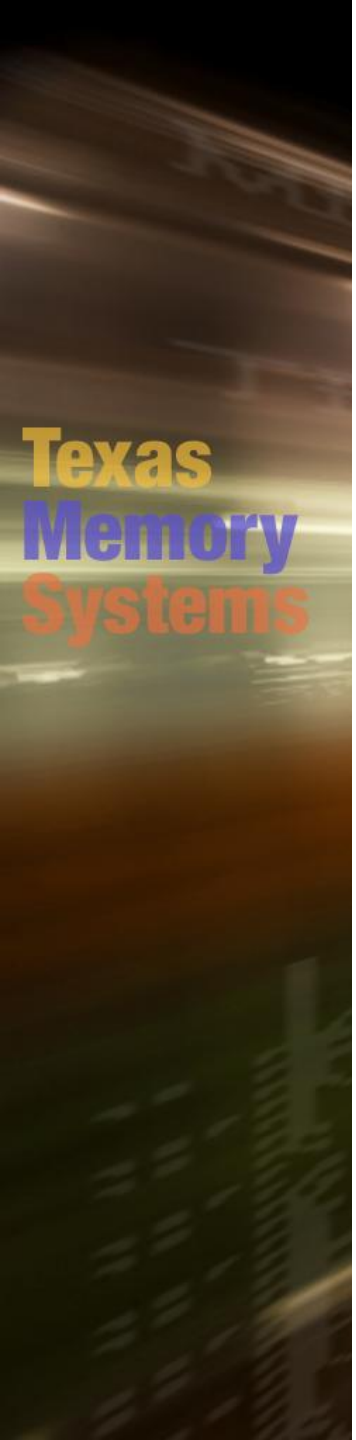
Oracle Requirements

- Real Application clusters
- Partitioning
- Parallel Query
- Maybe Analytical



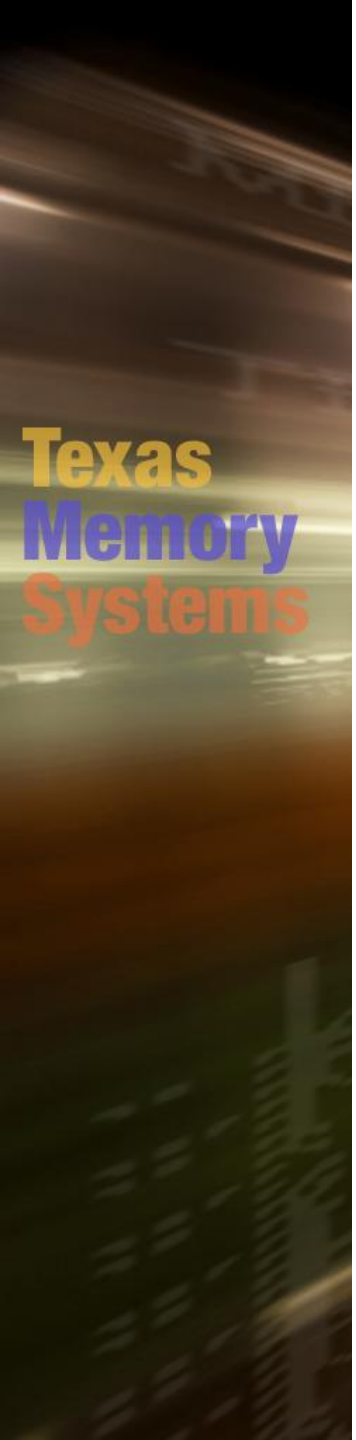
Server Requirements

- Multi-high speed CPU servers
- Multiple servers
- High speed interconnect such as infiniband between the servers
- Multiple High bandwidth connections to the IO subsystem



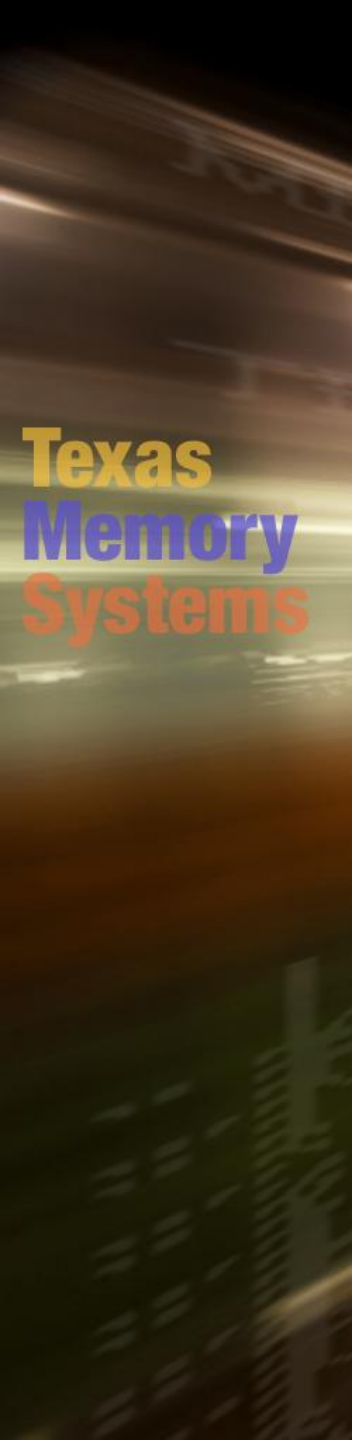
IO Subsystem Should:

- Be easily expandable
- Provide high numbers of low latency IOPS
- Provide high bandwidth



Let's build it: Servers

- DELL R905 PowerEdge .
  - 4-quadcore Opteron 8393™,3.1GHz processors,
  - dual 1 gb NIC
  - 2 dual channel 4 Gb fibre channel connections.
  - 10Ghz NIC
- 2 servers giving us 32 – 3.1 GHz processors

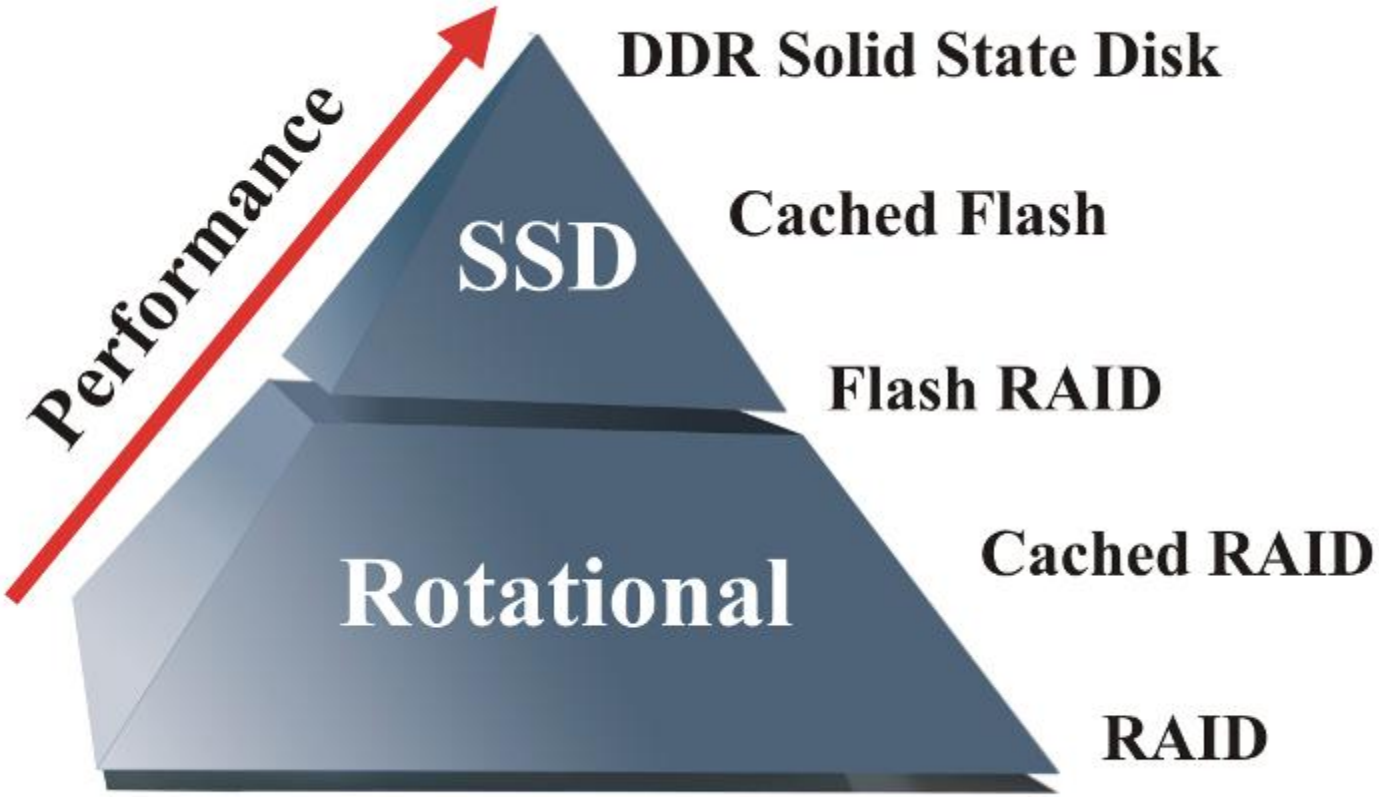


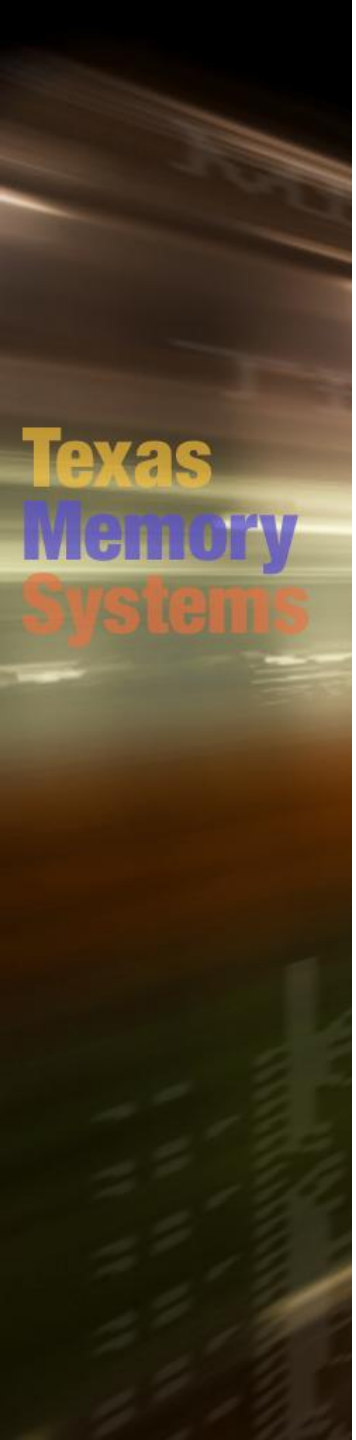
Why?

The 8393 at 3.1 GHz gives best SPEC benchmark results for the DELL 905 server

Texas Memory Systems

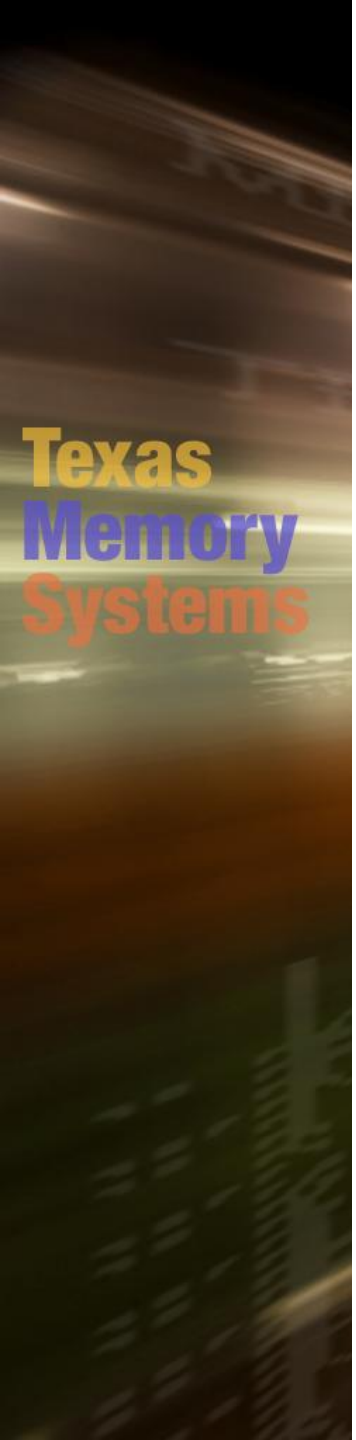
IO Subsystem





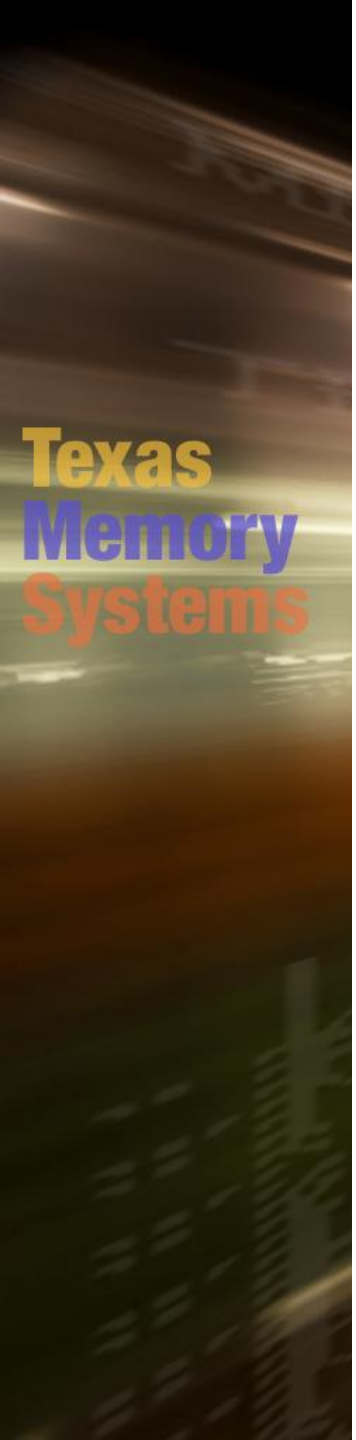
Tier Zero Storage

- Lowest Latency, smaller needed capacity
- 4 – Mirrored RamSan 440's loaded with 512 Gigabytes of DDR Ram (1 terabyte usable from mirror)
- 15 microsecond latency
- 600,000 IOPS each
- Projected cost: $668K



Tier 1 Storage

- Not as fast
- Need more capacity
- 2-mirrored RamSan-630's (8 terabytes of capacity)
- 250 microseconds response time
- 500,000 IOPS
- Projected Cost: $697K



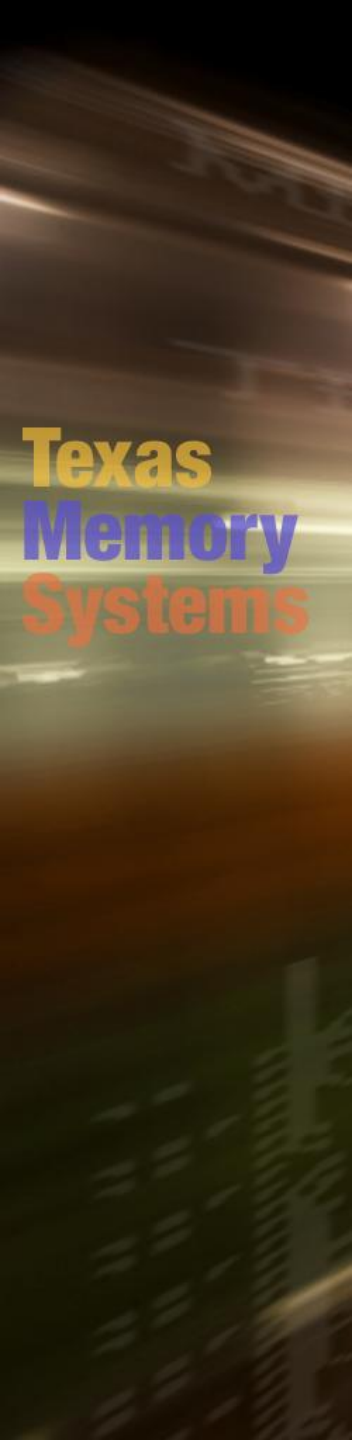
Tier 2 Storage

- Actually only used for Backup
- Should provide for de-duplication and compression
- An appliance would be best
- DataDomain DD120
- Projected cost: $12.5K

Texas Memory Systems

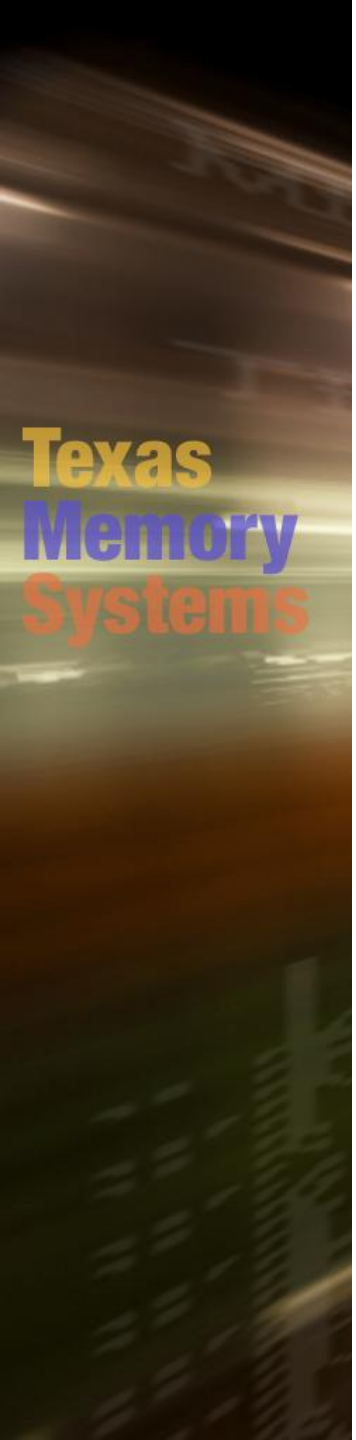
Switches

- Must provide for all connections
- Must be low latency high bandwidth
- Need 4-16 channel 4 Gb switches
- I suggest QLogic SanBox 5600Q
- Projected cost: $13.1K
- 10 Gig Ethernet router Fujitsu XG700, projected cost $6.5K
- Total switches: $19.6K



Misc.

- Cabinet
- Cables
- Etc.
- Estimate at $1.5K



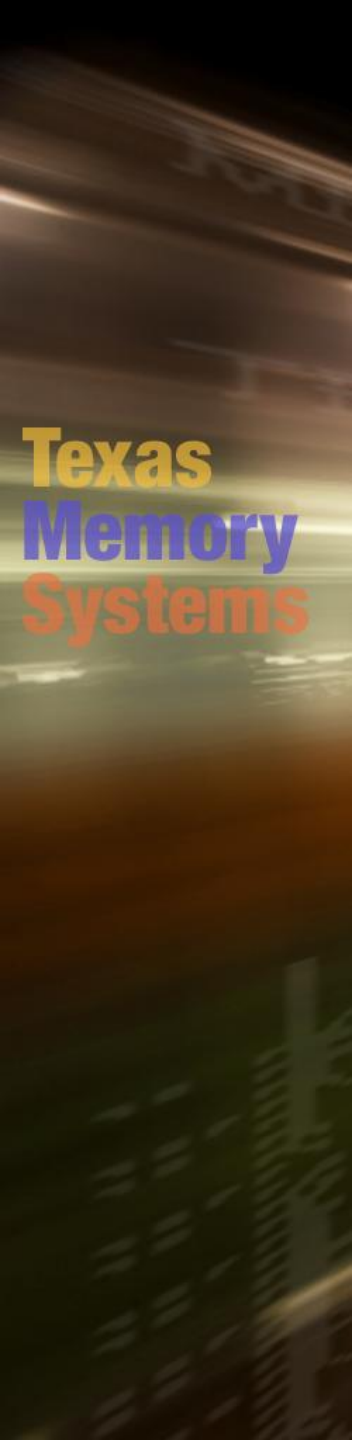
Total:

Servers:	36,484.00	
RamSan-440:	668,400.00	
RamSan-620:	697,400.00	
DataDomain:	12,500.00	
Switches:	19,600.00	
Misc.	1,500.00	(cables, rack, etc.)
Total	1,435,884.00	

Texas Memory Systems

But What About Oracle?

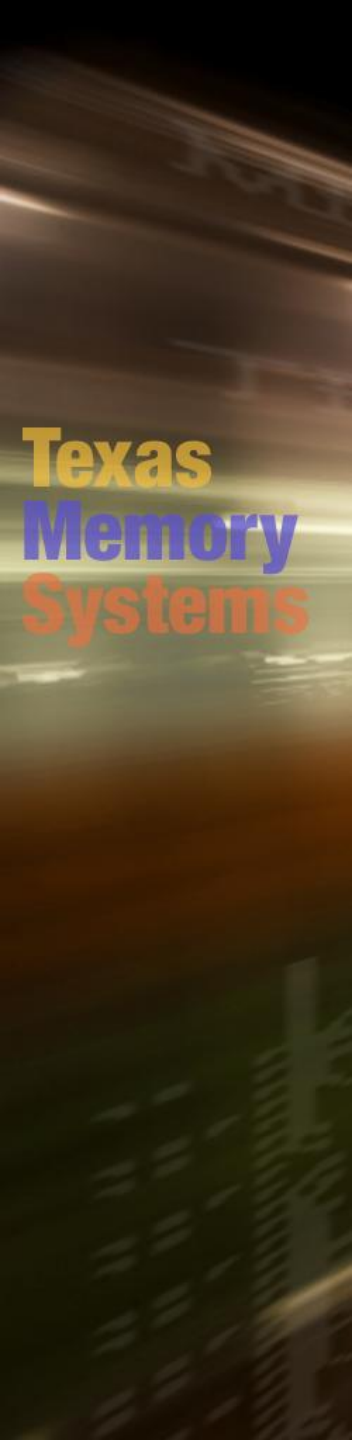
- Just can make a guess
- For a similar CPU architecture with RAC and Parallel Query
- Cost in June 2009 was $440K with three years maintenance based on a 12 named user license



Grand Total:

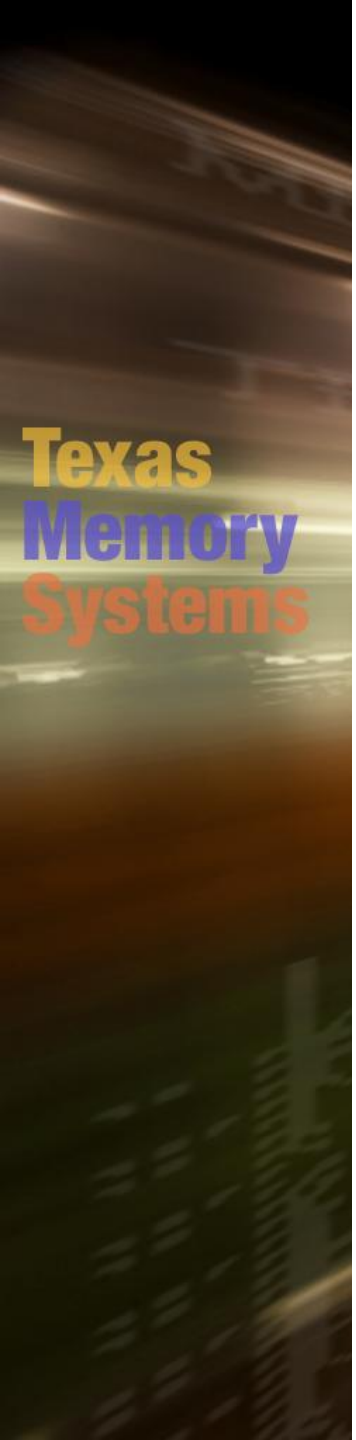
$1,875,654.00

Should I wrap that for you?



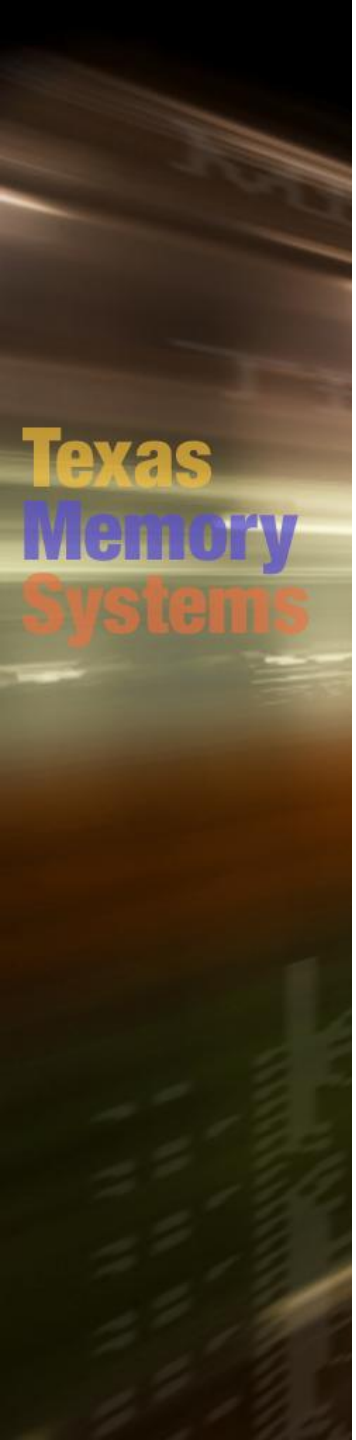
How does that Compare?

- Let's look at the new Exadata
- Won't need a full build
- We will look at a half Rack configuration
- 4 - servers
- 7 - Exadata cells
- Meets space requirements but only provides 18,900 IOPS
- Would need 741 cells to match IOPS at a cost for hardware and software in excess of $62 million



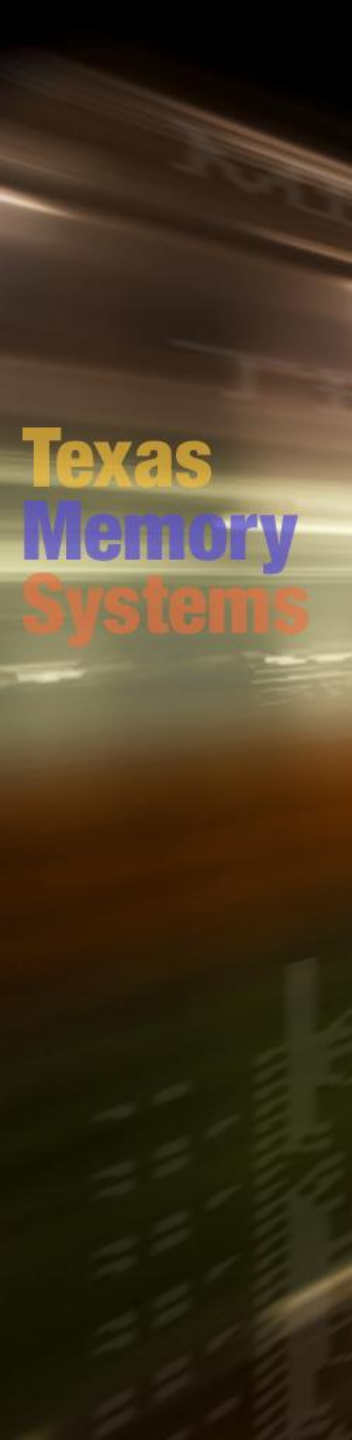
Projected Oracle DBM Cost

Half Rack:	$671,000.00
Cell Licenses:	$960,000.00
Support (3 yr):	$432,200.00
Oracle Licenses:	$440,000.00
Total	$2,503,200.00



For those in the back…

$2,503,200.00



Let's Think Green

- The energy consumption of the Exadata cell is projected to be 900 watts.
- For 7 cells about 6.3KW
- 600W for each of the RamSan440 systems
- 500W for each RamSan630
- total of 3.4KW.
- Over one year of operation the difference in energy and cooling costs could be close to half those of the Exadata using the SSD solution
- Once all of the license and long term ongoing costs are considered the ideal solution provides quite a savings.

Texas Memory Systems

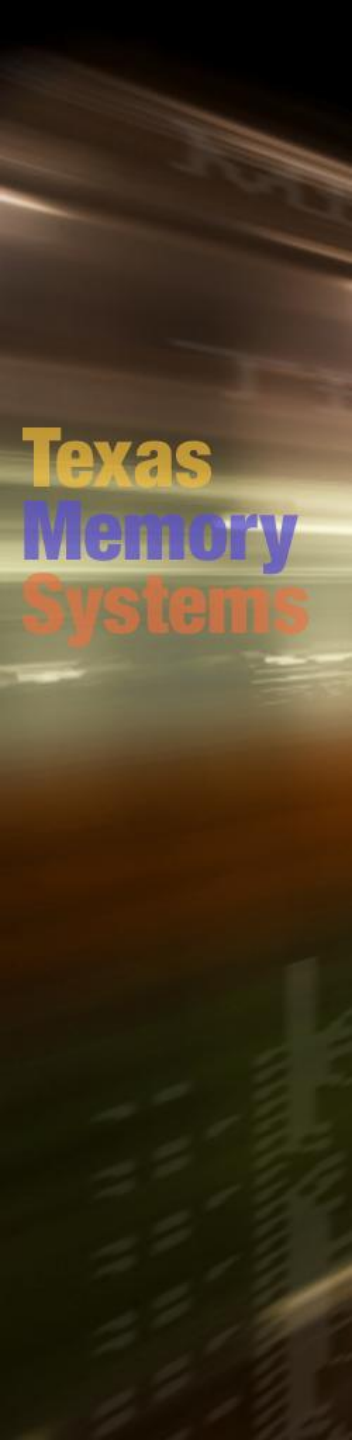
Final Scorecard

Consideration	Flexible Configuration	Oracle DWHM
OS Flexible	Yes	No
DB Flexible	Yes	No
Expandable	Yes	Yes
High IO Bandwidth	Yes	Yes – but low IOPS
Low Latency	Yes	No
High IOPS	Yes	No
Initial cost	Lower	Higher
Long term cost	Good	Poor

Texas Memory Systems

Summary

- I have shown a flexible data warehouse server architecture.
- This type of system is a moving target
- An ideal architecture allows high IO bandwidth. Low latency, capability for high degree of parallel operations and flexibility as far as future growth, database system and OS are concerned.
- Weigh all factors before selecting a solution



Questions?

- Mike.ault@texmemsys.com
- www.superssd.com
- www.ramsan.com