

# Redefining Oracle Infrastructure with Flash & Dell EMC

## Storage, Backup and Protection for Oracle Environments



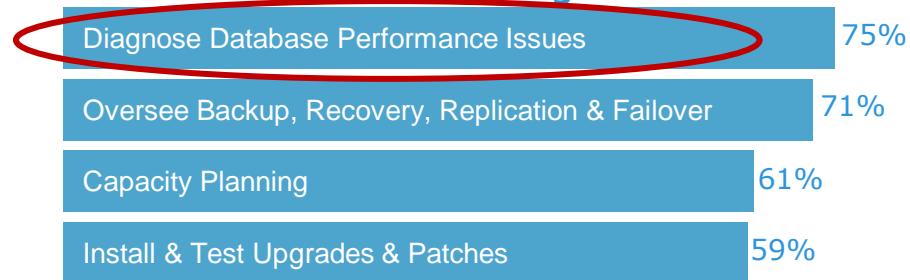
Bart Sjerps | Principal Systems Engineer | Oracle Specialist - EMEA  
[bart.sjerps@dell.com](mailto:bart.sjerps@dell.com) | +31-6-27058830  
<http://bartsjerps.wordpress.com>

**DELL**EMC

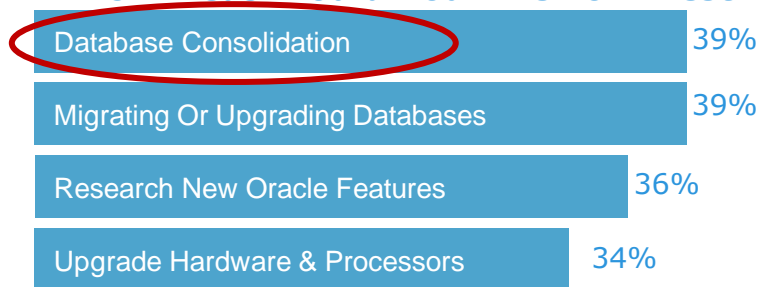
# Oracle Customer Challenges

## Too Much Time & Budget Spent Maintaining Oracle

### Which Tasks Are Consuming Most DBA Time?

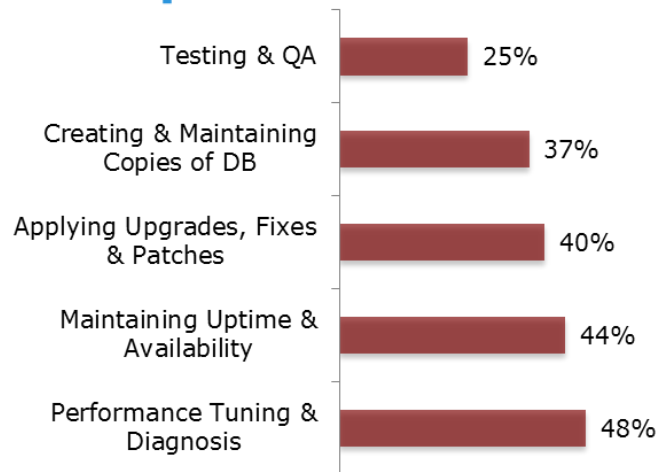


### Which Areas Would You Like To Invest More?



(Source: **IOUG** - DBA Survey)  
independent oracle users group

## Time Spent on Maintenance



### Top 5 Database Activities

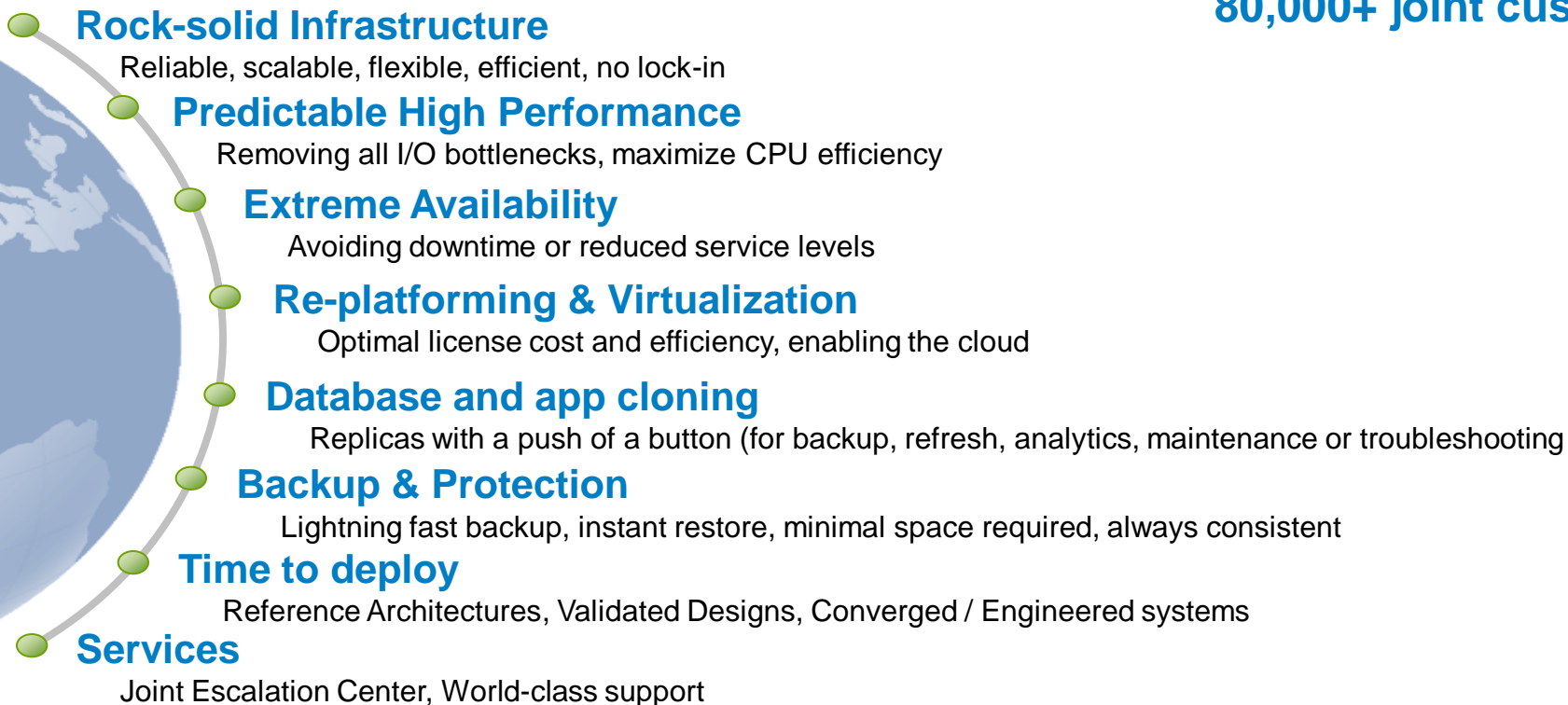
(Source: 2014 IT Resource Strategies Survey)

# Dell EMC Solutions for Oracle

**ORACLE®**

**1995–Present**

**80,000+ joint customers**



An aerial night view of a city skyline, likely New York City, showing numerous skyscrapers and buildings illuminated with lights. The word "Performance" is overlaid in large white text in the center of the image.

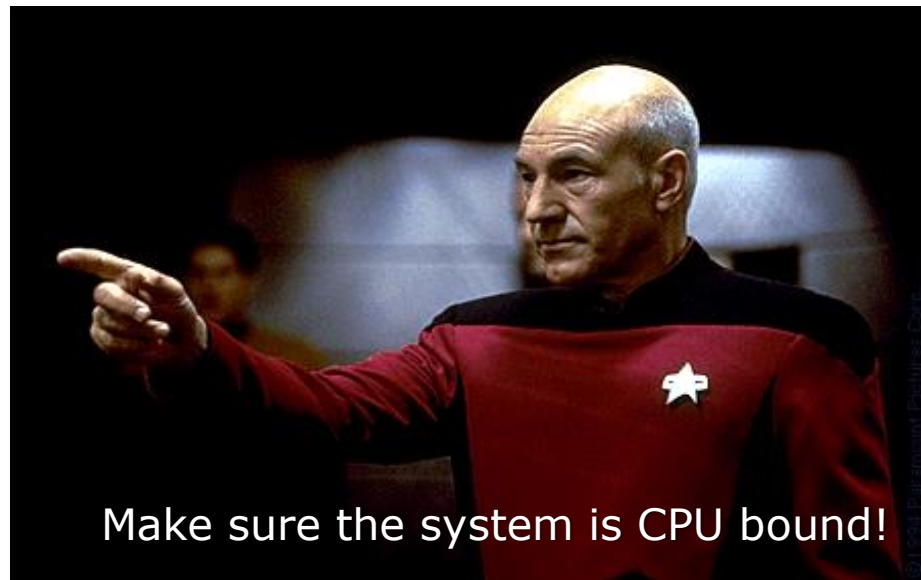
# Performance



# Databases should not have I/O Wait

## Everything should be a CPU problem

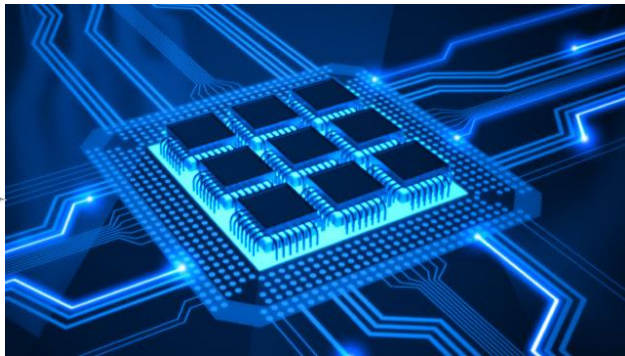
- Adding CPU does not speed up I/O bottlenecks
  - Memory does somewhat
- IOPS are relatively (!) cheap
- CPU cycles are expensive
  - Because of licenses
- Consolidation can lead to
  - Higher IO requirements
  - I/O bottlenecks
  - Bandwidth issues
  - Backup window limitations
- Flash storage can solve these limitations



**STORAGE IS NO LONGER THE BOTTLENECK**

# Spinning Disk vs Flash

Spinning disk is 60+ years old and slowly dying



- **One seek at a time**
- **Mechanical latency (7ms)**
- **Max 150 random IOPS**
- **Good at sequential IO**
- **No write penalty**

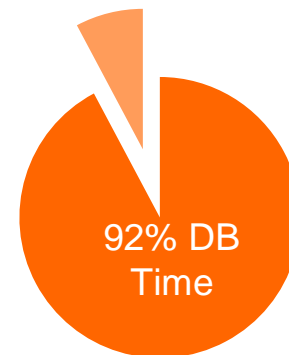
**VS.**

- **Parallel requests (0.1 ms)**
- **No mechanical latency**
- **Virtually unlimited IOPS**
- **Good at sequential IO**
- **Write penalty compensated by controller (cache, leveling)**

**"When 900 years old you reach  
Look as good you will not"  
-- Yoda (Return of the Jedi)**

# I/O bound DB server

**4 hours** time wasted by high-latency random reads



## Top 5 Timed Foreground Events

Event	Waits	Time(s)	Avg wait (ms)	% DB time	Wait Class
db file sequential read	1,420,110	14,362	10	92.7	User I/O
free buffer waits	89,072	986	11	6.4	Configurat
DB CPU		158		1.0	
library cache lock	138	51	371	.3	Concurrenc
write complete waits	9	17	1919	.1	Configurat

# After moving to All-Flash

Event	Waits	%Time -outs	Total Wait Time (s)	Avg wait (ms)	Waits /txn	% DB time
direct path read	361,425	0	930	3	7,376.0	16.4
db file parallel read	365,088	0	898	2	7,450.8	15.9
db file scattered read	64,628	0	114	2	1,318.9	2.0
db file sequential read	56,748	0	84	1	1,158.1	1.5
gc cr multi block request	137,524	0	62	0	2,806.6	1.1
direct path read temp	19,942	0	44	2	407.0	.8
read by other session	17,389	0	31	2	354.9	.6

Event	Waits	Time (s)	Avg wait (ms)	% DB time	Wait Class
DB CPU		3,427		60.5	
direct path read	361,425	930	3	16.4	User I/O
db file parallel read	365,088	898	2	15.9	User I/O
db file scattered read	64,628	114	2	2.0	User I/O
db file sequential read	56,748	84	1	1.5	User I/O

- ✓ Avg. latency went down by more than 70%
- ✓ CPU utilization doubled



# The IOPS race is over

Million IOPS? 2 million IOPS? 5 million IOPS? Who cares?



**Every flash vendor these days offers:**

Huge IOPS numbers  
Very low latency

Average small AFA: >> 200K IOPS  
Average large DB: << 50K IOPS

**Competing in the FLASH business is not about IOPS and latency any more \*)**  
**-- Bart Sjerps, 2016**

\*) Considering generic Oracle workload consolidations here - Large scale DWH/BI is a special case.

## What are differentiating features of Flash solutions for databases - if it isn't IOPS and latency?

# XtremIO Data Services

Always-on, Inline, Zero Penalty

ALWAYS ON THIN  
PROVISIONING



INLINE DATA  
DEDUPLICATION



INLINE DATA  
COMPRESSION



INLINE DATA  
ENCRYPTION



FLASH OPTIMIZED  
DATA PROTECTION



WRITABLE  
COPIES



*Up To 6:1 Database Reduction On Average  
Just For Production Databases*

# Obsolete Storage Practices

Simplify Architecture and operational management



Under-allocation  
(disk "short stroking")



Data Separation



Pre-fetching  
(attempting to  
predict the future)



Striping



Monster caches  
(preloading everything)

# More Obsolete Practices



**Classic RAID**  
(Having only one extra safety rope)



**Full data copies**  
(Avoiding I/O interference)



**Storage Layers / Tiering**  
devil's dilemma:  
performance or economics



**Direct Attached Storage**  
(Trading redundancy & sharing  
for bandwidth & cost)

# The Quick & Dirty Dedupe Analyzer

**Figure out how much efficiency savings are possible**

- Linux tool to scan disks, files and data streams for duplicate blocks
- Can scan a running system
- Safe & Fast
- Variable blocksize
- Free (Open Source)
- Bonus: Compression and thin analysis
- Based on C++ / SQLite as Key-value store
- Run your own queries against the data

[QDDA Landing page](#)



The best thing about being me... There are so many “me”s.

— Agent Smith, The Matrix Reloaded



```
[root@db11 ~](-) # qdda /dev/oracleasm/* -b0
qdda 1.5.1 - The Quick & Dirty Dedupe Analyzer
File 01, 524288 blocks, 4096 MiB processed, 0 MB/s, 257 MB/s avg, Processing took 2.75 sec
File 02, 262144 blocks, 2048 MiB processed, 0 MB/s, 393 MB/s avg, Processing took 0.60 sec
File 03, 131072 blocks, 1024 MiB processed, 0 MB/s, 282 MB/s avg, Processing took 0.62 sec
blocksize          =          8 KiB
total              =       7168.00 MiB (   917504 blocks)
free               =       2978.05 MiB (   381190 blocks)
used              =       4189.95 MiB (   536314 blocks)
unique            =       3434.32 MiB (   439593 blocks)
deduped 2x        =        177.36 MiB (    22702 blocks)
deduped 3x        =        128.15 MiB (    16403 blocks)
deduped 4x        =         0.04 MiB (         5 blocks)
deduped >4x       =         0.05 MiB (         6 blocks)
deduped total     =       3739.91 MiB (   478709 blocks)
stream compressed =        394.79 MiB (    89.44 %)
compress buckets 2k =       855.68 MiB (   438110 buckets)
compress buckets 4k =        96.67 MiB (    24748 buckets)
compress buckets 8k =       123.84 MiB (    15851 buckets)
total compressed  =       1076.19 MiB (   137752 blocks)
*** Summary ***
percentage used   =          58.45 %
percentage free   =          41.55 %
deduplication ratio =          1.12
compression ratio =          3.48
thin ratio       =          1.71
combined        =          6.66
raw capacity     =       7168.00 MiB
net capacity     =       1076.19 MiB
[root@db11 ~](-) #
```



# Still need more?

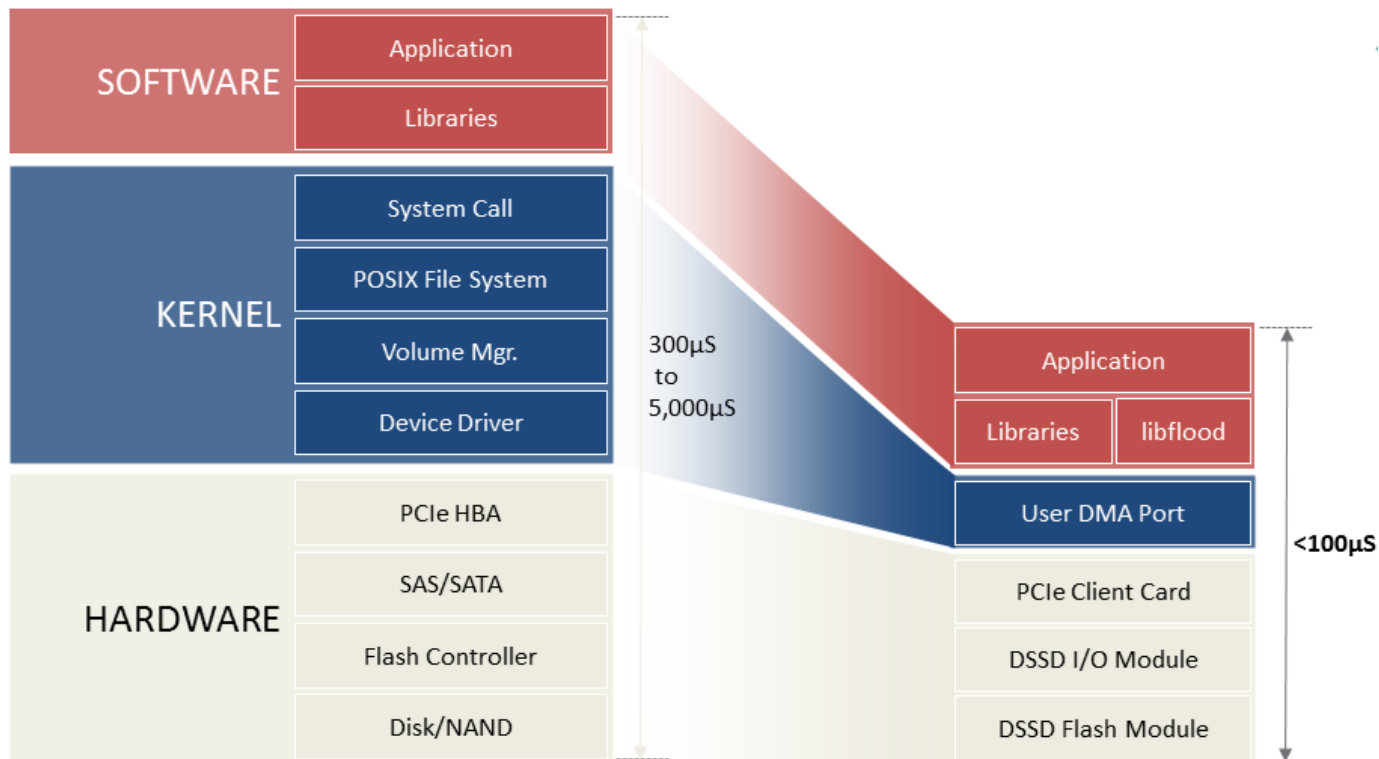


## ENGINE POWER

Lots is good  
More is better  
Too much is just enough

# Innovation: NVMe I/O protocol

## “TAFKAD”: PCIe Direct Access to SSD



- ✓ 10M IOPS
- ✓ 100 GB/s
- ✓ <0.1 ms

Enterprise  
Readiness

- ✓ Data protection
- ✓ Redundancy
- ✓ Consistent performance

Tradeoff:  
Data services  
⇒ Future

An aerial night view of a city skyline, likely New York City, with numerous skyscrapers illuminated. The word "Availability" is overlaid in large white text in the center of the image.

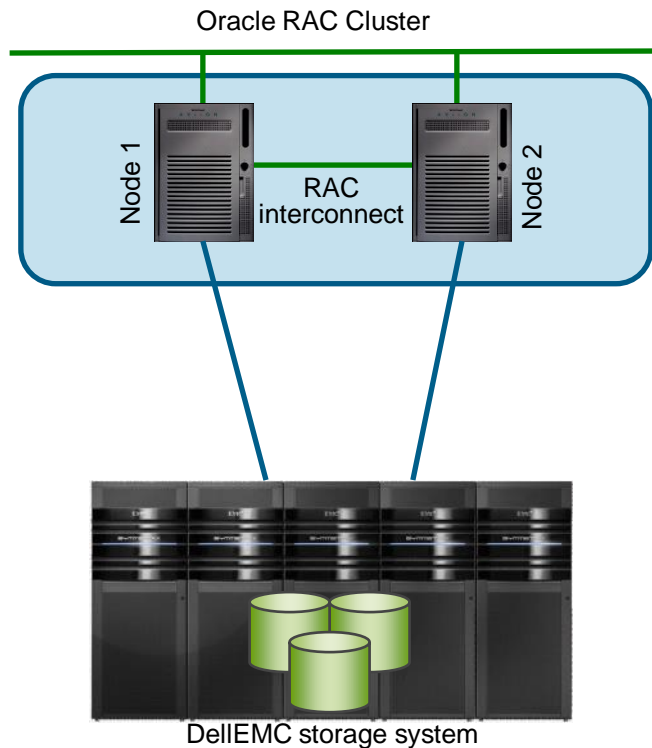
# Availability



# Oracle Fault Resilience

Where are the limitations in this picture?

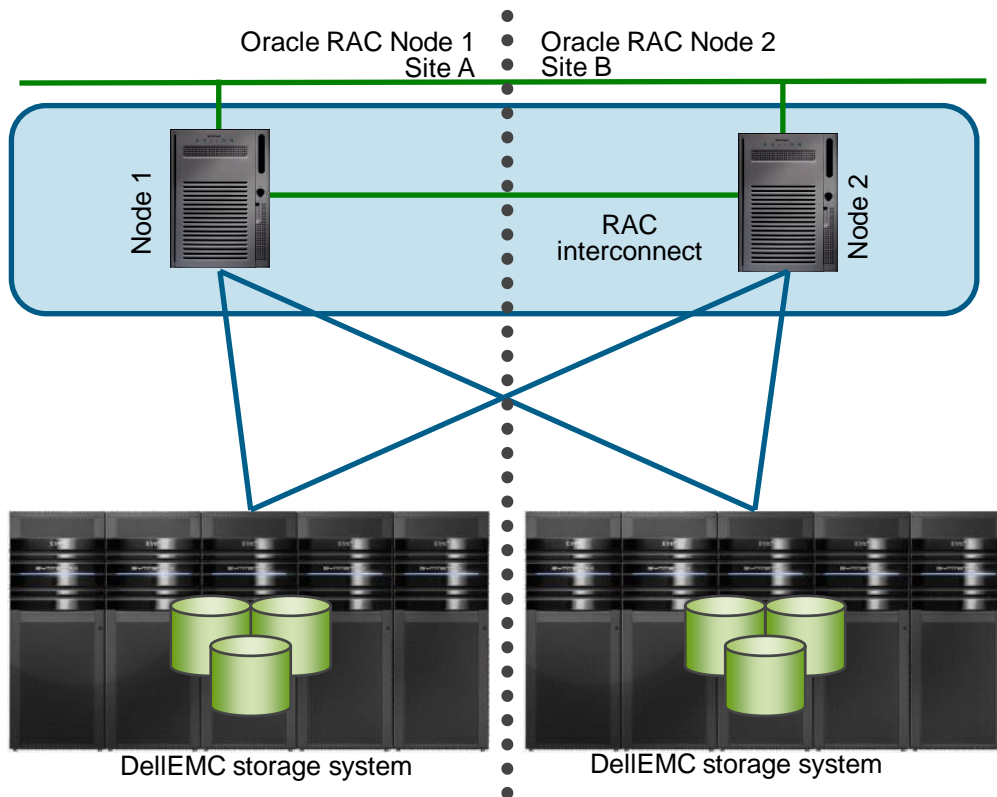
- Node failure:
- Planned maintenance:
- Storage failure:
- Site failure:



# Classic solution

## Host based mirroring

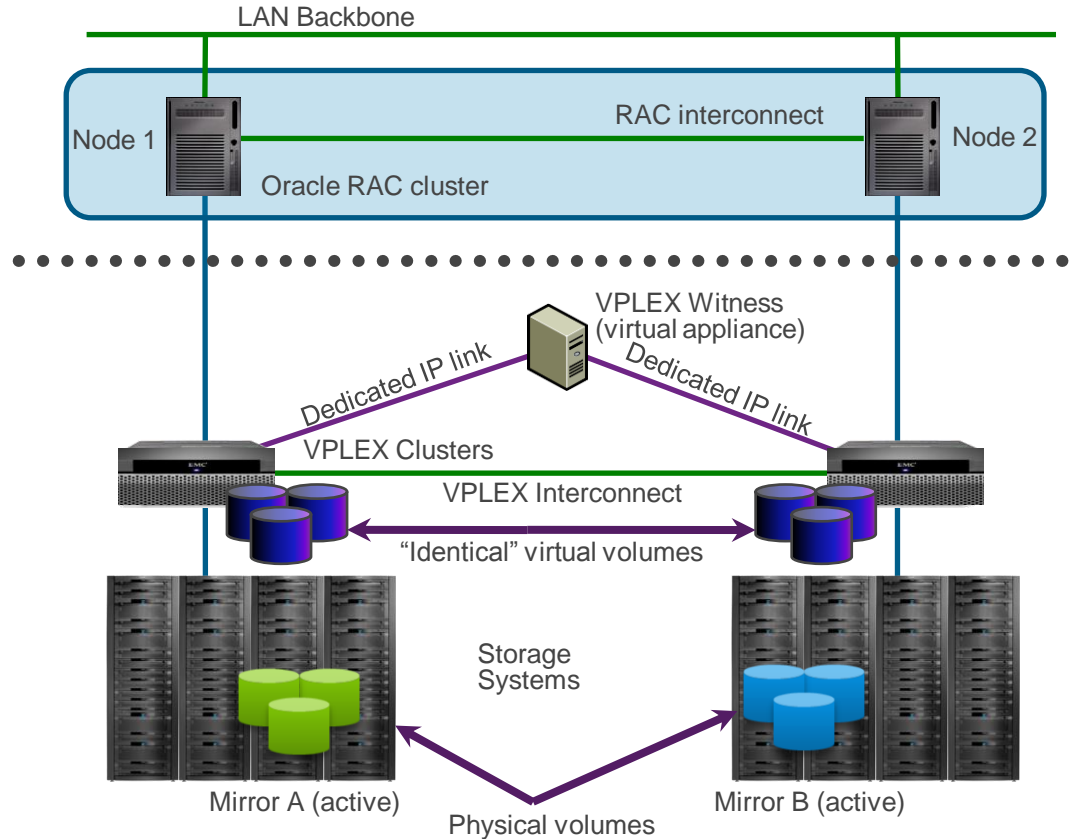
- Typically via ASM Mirroring
- Appears simple at first glance
- Devil is in the details
  - Link failure?
  - Resilvering?
  - Cloning/snapshots?
  - Data integrity?
  - I/O overhead?
  - Configuration mistakes?
  - Failure groups?
  - Split Brain issues?
  - Scaling # of nodes?
  - Application failover?



# Dell EMC Solution

## Keep it simple

- No changes on host level
  - Except adjusting timeouts
- No complex configuration
  - Reduces risk of configuration errors
- DB/Hosts are unaware of replication
  - Complexity is hidden
  - No failure groups
- Survives split brain issues
  - Due to the witness
  - No “NFS arbitration” required
- Applications “follow” Database
  - If placed on the same infrastructure
- No failover when storage unavailable
- No host overhead
- No DB performance impact for re-sync
- Plays nice with snapshots/cloning



→ Also available as native VMAX solution (SRDF Metro vs VPlex)

An aerial night view of a city skyline, likely New York City, with numerous skyscrapers illuminated. The word "Replatforming" is overlaid in large white text in the center of the image.

# Replatforming

# Database Re-platforming

## Goals

1. *Maximize use of license investment*
2. Maintain or (better even) improve performance
3. Reduce downtime / increase SLAs
4. Avoid Vendor lock-in
5. Simplify server & storage refresh cycles
6. Speed up provisioning of new databases
7. Improve security, compliance and auditing
8. Simplify management

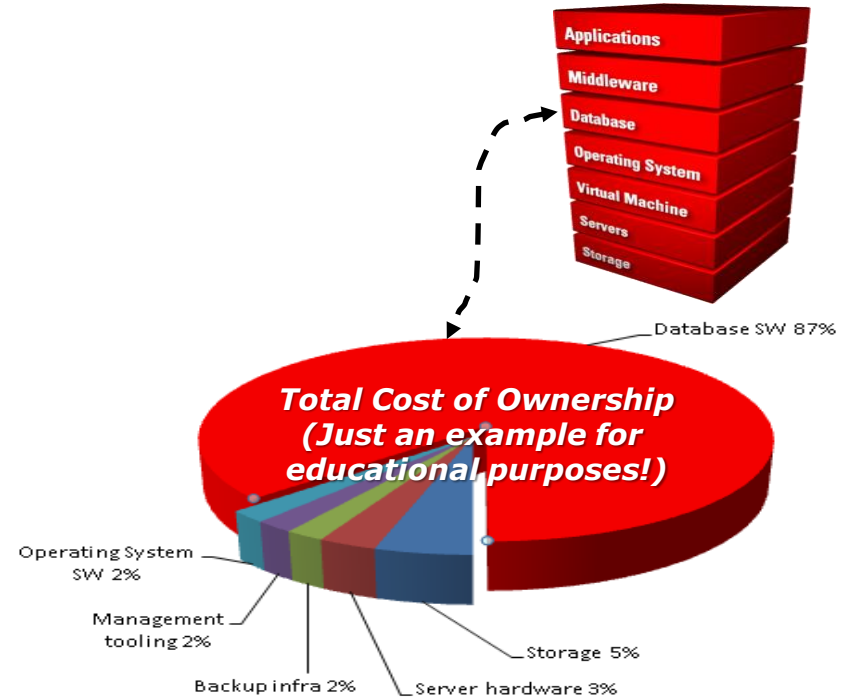




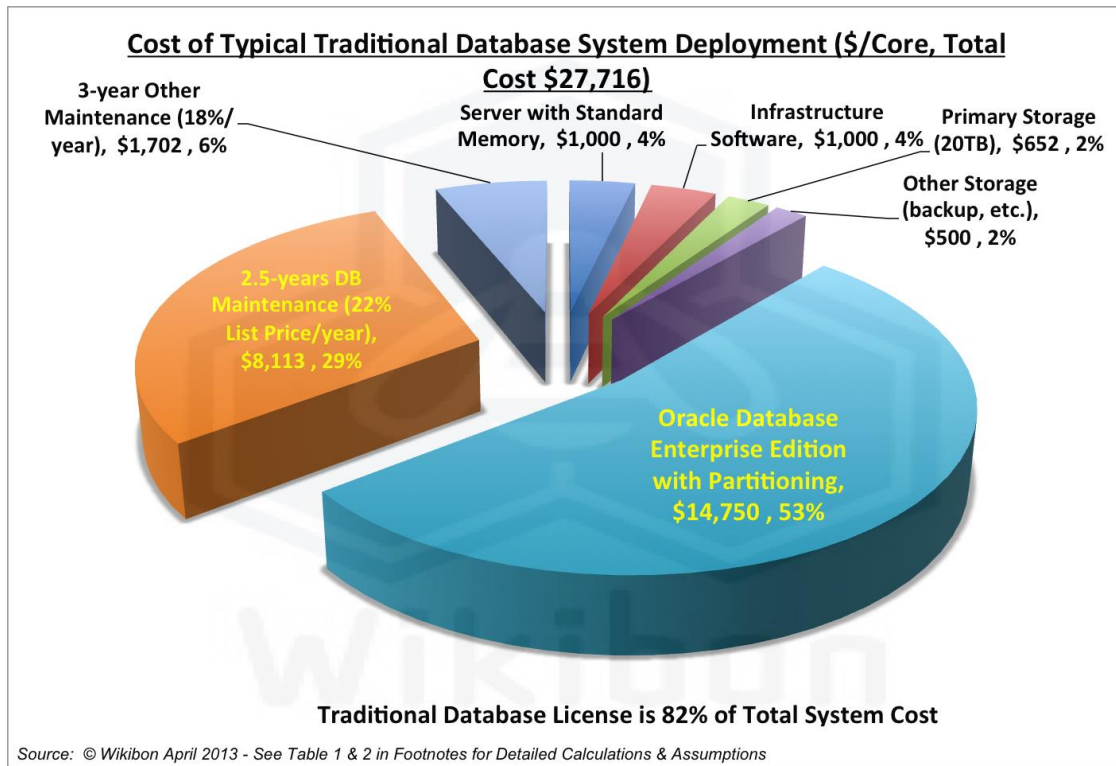
# Why look at Oracle licensing?

- Oracle DB licensing is expensive
  - Midsize server (44 cores):  
HW ~ \$ 50,000  
SW ~ \$ 913,000 @ 50% discount  
5 year maintenance ~ \$1,000,000  
(Enterprise Edition + basic options <sup>1)</sup>)
  - What if we add RAC? Active DG?  
Multitenant?
- Large part of the TCO of a database infrastructure stack
- Many DB servers are poorly utilized
  - CPU 90% idle is not unusual

<sup>1)</sup> Enterprise Edition, Partitioning, Adv. Compression, Diagnostics & Tuning Pack



# Validation: Wikibon Research



Wikibon Article: [Virtualization of Oracle Evolves to Best Practice for Production Systems](#)

# Before we start...

Beware of the license demon

Are you 100%  
"BET YOUR  
paycheck" sure  
THAT YOU'RE  
compliant?™



ORACLE®  
LICENSE MANAGEMENT  
SERVICES



If needed...  
Bring in the superheroes  
They help you with licensing  
& legal issues



[Licenseconsulting.eu](http://Licenseconsulting.eu)



[House of Brick](http://House of Brick)



[Madora Consulting UK](http://Madora Consulting UK)

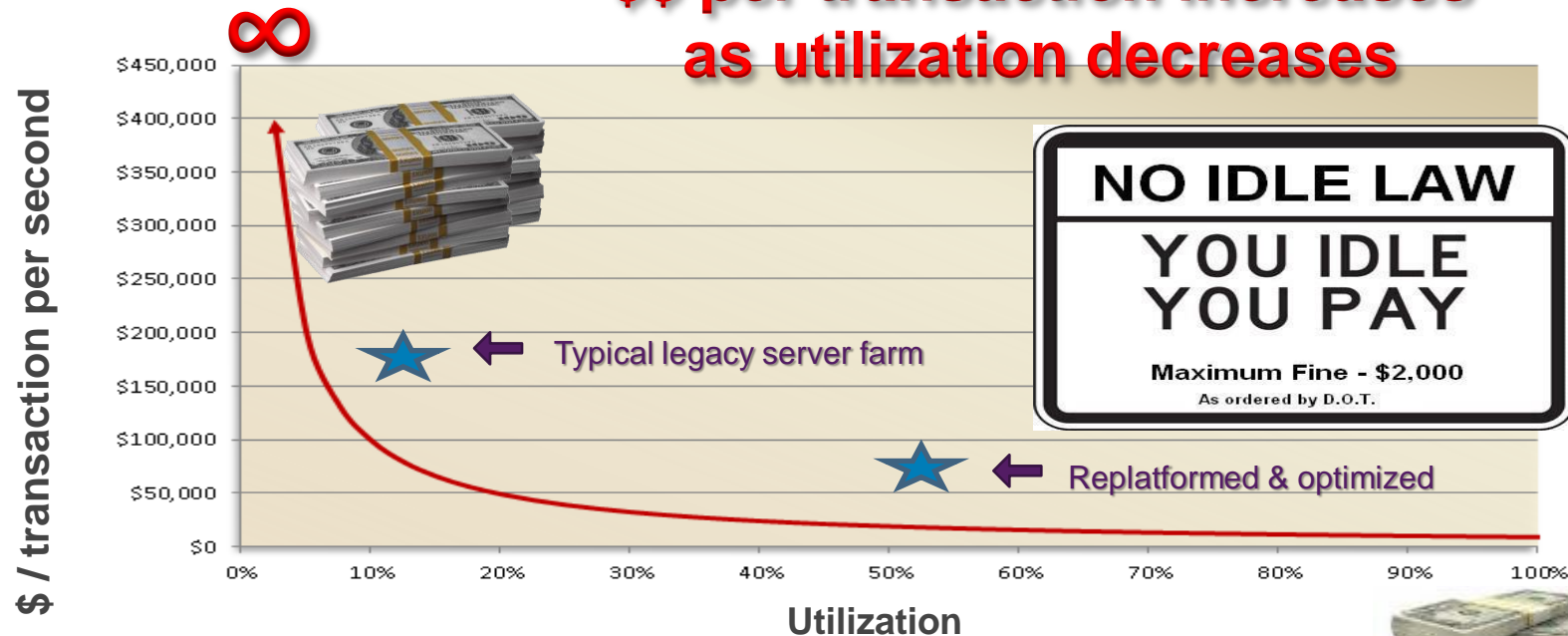
# Support & Licensing



- Oracle is FULLY supported on VMware
  - Including Oracle RAC
  - Any other claim is **FALSE**
  - Platform certification is NOT required
  - Escalation paths exist from Oracle/Dell EMC and VMware to avoid fingerprinting
  - Need to reproduce on physical is RARE but easy with Dell EMC (using clones)
- All potential licensing problems can be avoided
  - Including recent Oracle myths about Vsphere 5.5 and Vsphere 6 (cross-vcenter vmotion)
  - You only need to license servers where Oracle is installed and/or running (nothing else)
- Performance scaling & overhead is no issue
  - 1 VM: 128 vCPU, 4TB memory, 1M+ IOPS

# Transaction cost versus Utilization

**\$\$ per transaction increases  
as utilization decreases**



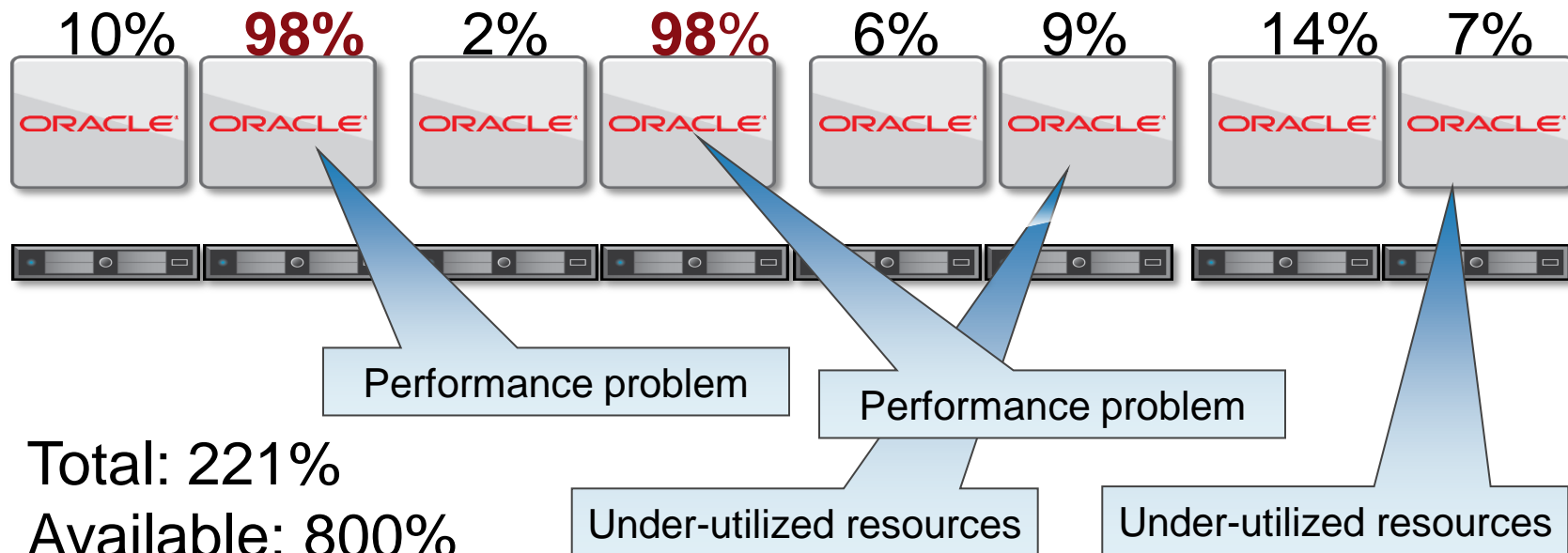
Cost per TPS for a four-node Oracle RAC 11g cluster running EE  
Software license cost: around \$2,200,000  
TPS: Around 4,000 at peak utilization





# Classic problem of Resource Management

Applied to Database processing power

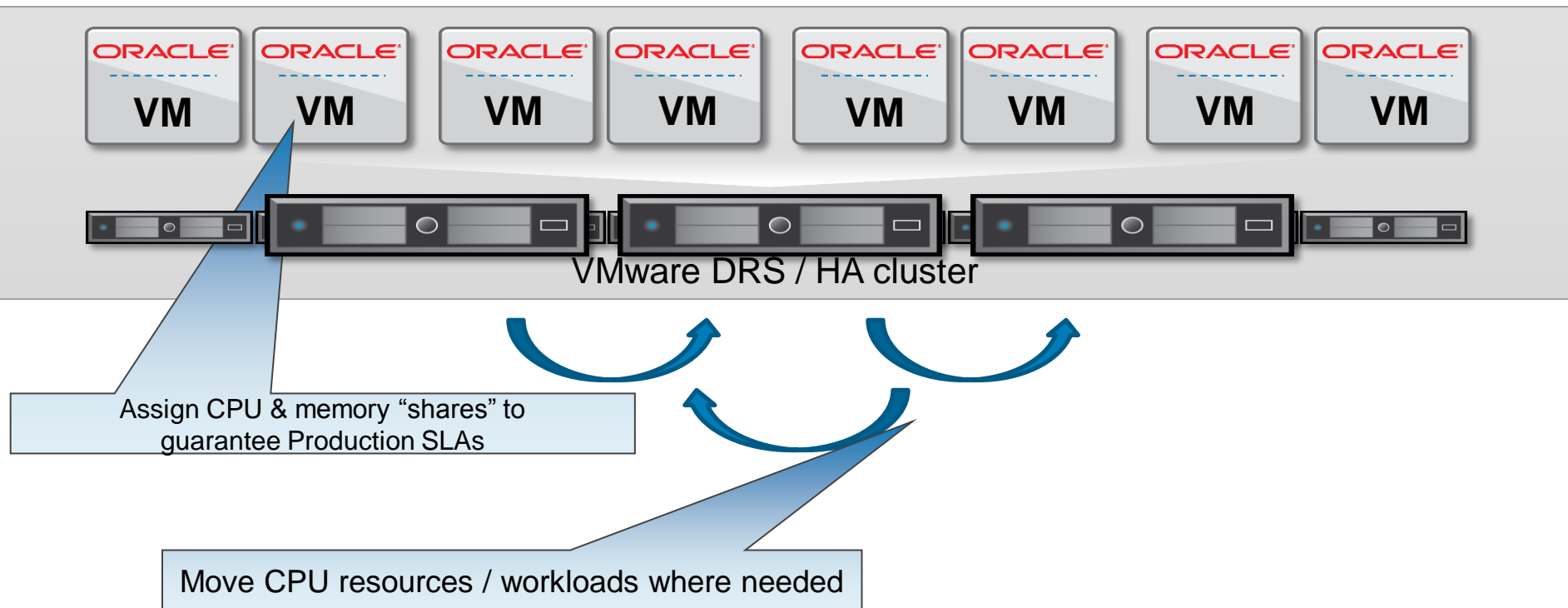


Total: 221%

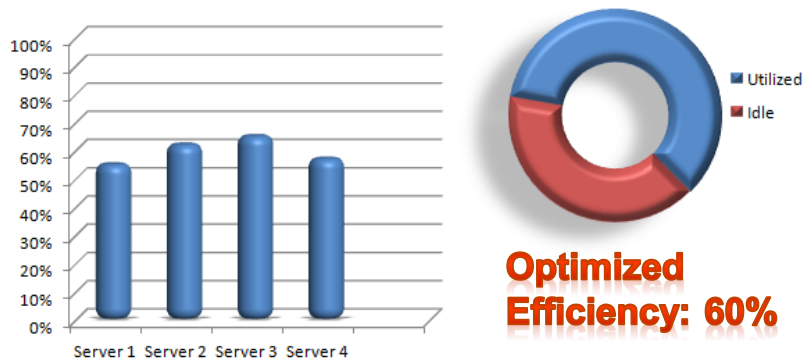
Available: 800%

Avg: 30%

# Resource Management “Mainframe style”



# Traditional vs Optimized



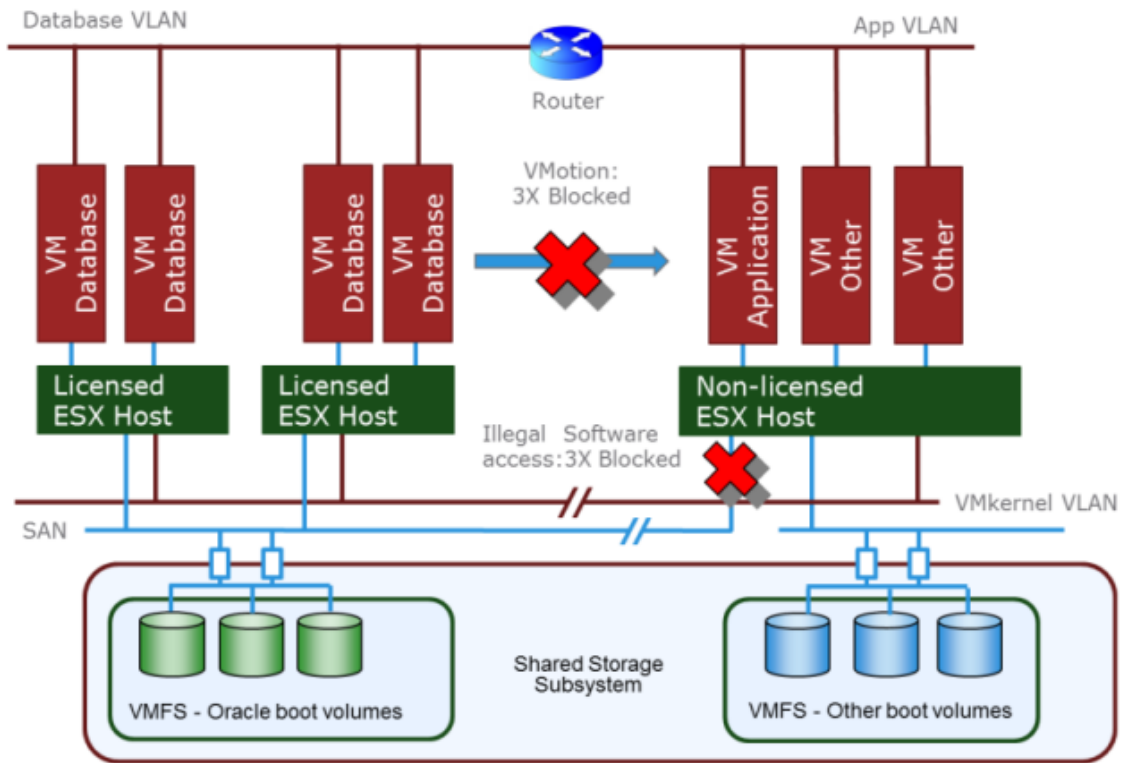
Typical legacy database server farm:

- Physically deployed
- Oversized
- Outdated platforms
- Very poor CPU utilization
- IO & CPU bottlenecks
- Servers running mix of:
  - Apps, middleware & DB
  - Tooling
  - Replication & Backup

Optimized database server farm:

- Virtualized
- Significantly less CPUs on Modern HW
- High average CPU utilization
- No I/O bottlenecks
- Sized correctly
- Servers running ONLY Oracle
- Minimal required licenses & options

# AVOIDING THE VMOTION TRAP



## Oracle on VMware: Caging the license dragon

### Do's

- Prevent “illegal” Vmotion moves by creating multiple barriers
- Keep Vmotion audit trails
- Watch the [IOUG “straight talk” video](#) on my blog
- **Hire external licensing expertise**

### Don'ts

- Believe Oracle sales reps
- Give LMS all info they ask for
- Run hypervisors that don't achieve TCO reduction

### Know

- You only have to license Oracle where it IS running (not where it might run in the future)
- Oracle FUD/Scare tactics

# 5 Steps to maximize cost efficiency

## 1. Hardware Replatform for lowest \$ / transaction

Get the best CPU type for transactional workload

Eliminate I/O problems, backup window, etc

## 2. Virtualize servers to drive up CPU utilization

Make sure you stay compliant

## 3. Remove unnecessary licensed options

Or go to different license model (i.e. Standard Edition)

## 4. Only run DB transactions on licensed CPU

Run other stuff elsewhere

## 5. Re-negotiate license contracts →

Suspend maintenance, etc

Avoid non-compliance, audits, support issues, ...

Use independent license expert services !



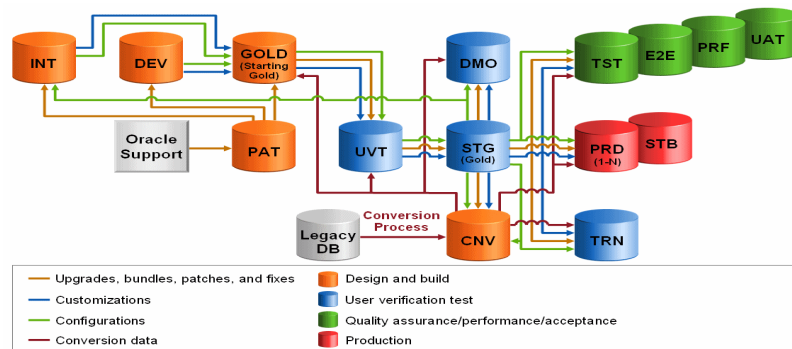


An aerial night view of a city skyline, likely New York City, with numerous skyscrapers illuminated. The word "Cloning" is overlaid in large white text in the center of the image.

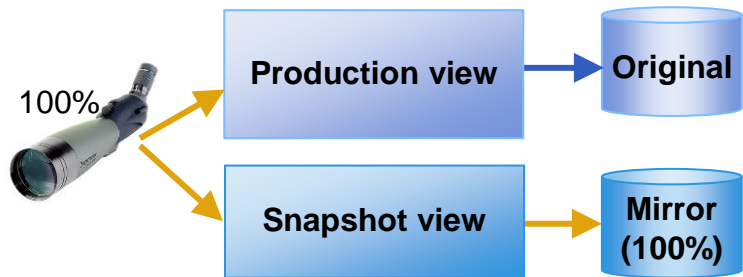
# Cloning

# Benefits of instant database copies (cloning)

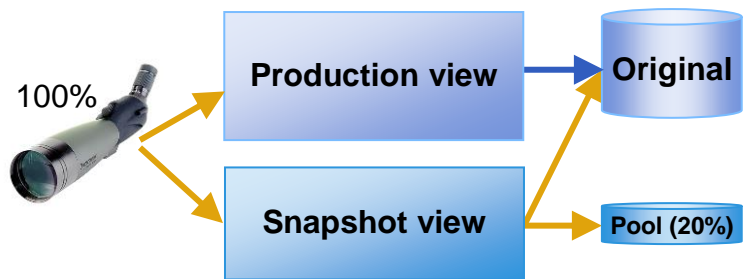
- "Serverless backups" & Quick restores
  - Zero production performance impact
  - RTO seconds to minutes
  - Protected & Instant restores
  - Out-of-order restores
  - Instant Restore from remote copies
- Firefighting
  - Creating a quick production copy to solve application problems (VMware Virtual-to-Physical)
  - Without messing with production data
- Creating Test / Dev / Acceptance copies
  - Automated, no tape restores, low people effort
- Creating copies for reporting / staging / analytics
  - Data warehouse queries can bring production performance down
  - Moving reporting workload to copy relieves production
- Application / Database Upgrades
  - Creating application "checkpoints" avoids having to fall back to starting point due to small errors
  - Easy upgrade testing



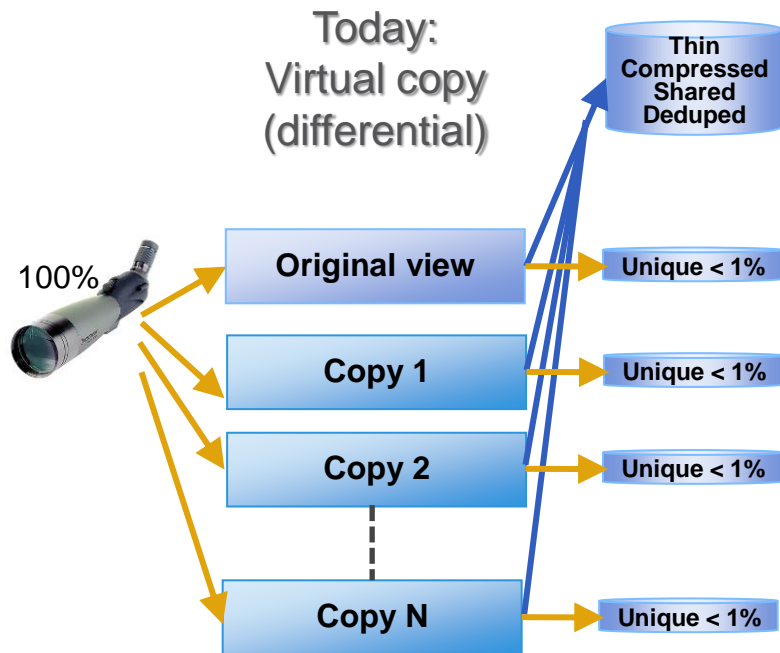
# Storage Replication - Innovations



1997: EMC Timefinder (Full copy)



2000's: Snapshots (incremental)





An aerial night view of a city skyline, likely New York City, featuring numerous illuminated skyscrapers and a dense grid of city lights. The image is dark, with the city lights providing the primary illumination. The text "Data Protection" is overlaid in the center in a large, white, sans-serif font.

# Data Protection

# Backup Optimization

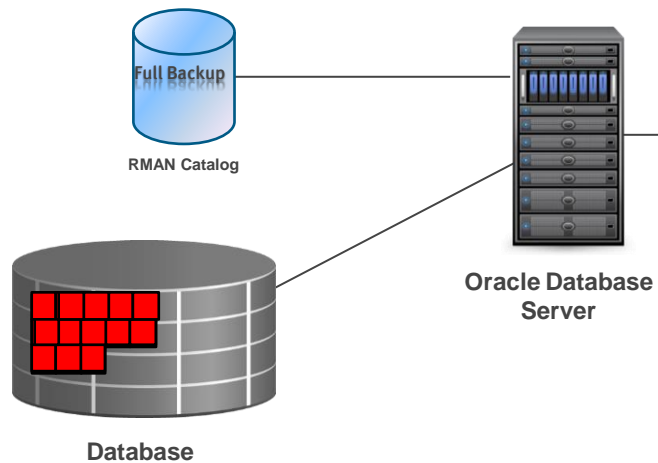
## Limitations of traditional approaches

- Real world (extreme) example:
  - Customer with several 1-2 Terabyte databases (tier 1 production)
  - SLA demands 2x full backup / day, Retention: 1 Month
  - Tape capacity required for a 2 TB database:  $2 \text{ TB} * 2 \text{ copies / day} * 31 \text{ days} = 124 \text{ TB tape}$
- Considerations
  - How much savings would be achieved when reducing DB by 20% ?
  - What if we could store only 1 full copy plus 61 delta sets?
  - What if we need 6 months retention?
  - How fast can we recover from backup using tape?
  - Is it reliable?
  - What's the performance impact on production?
  - What's the backup window?





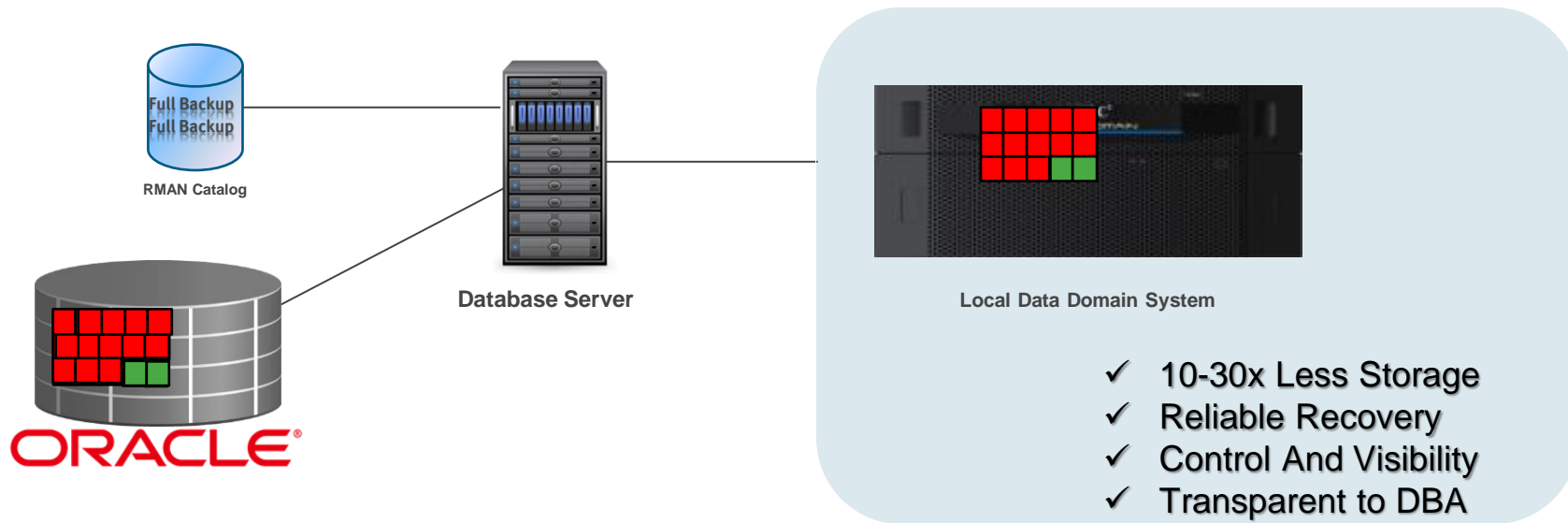
# Oracle Backup to Data Domain



Local Data Domain System

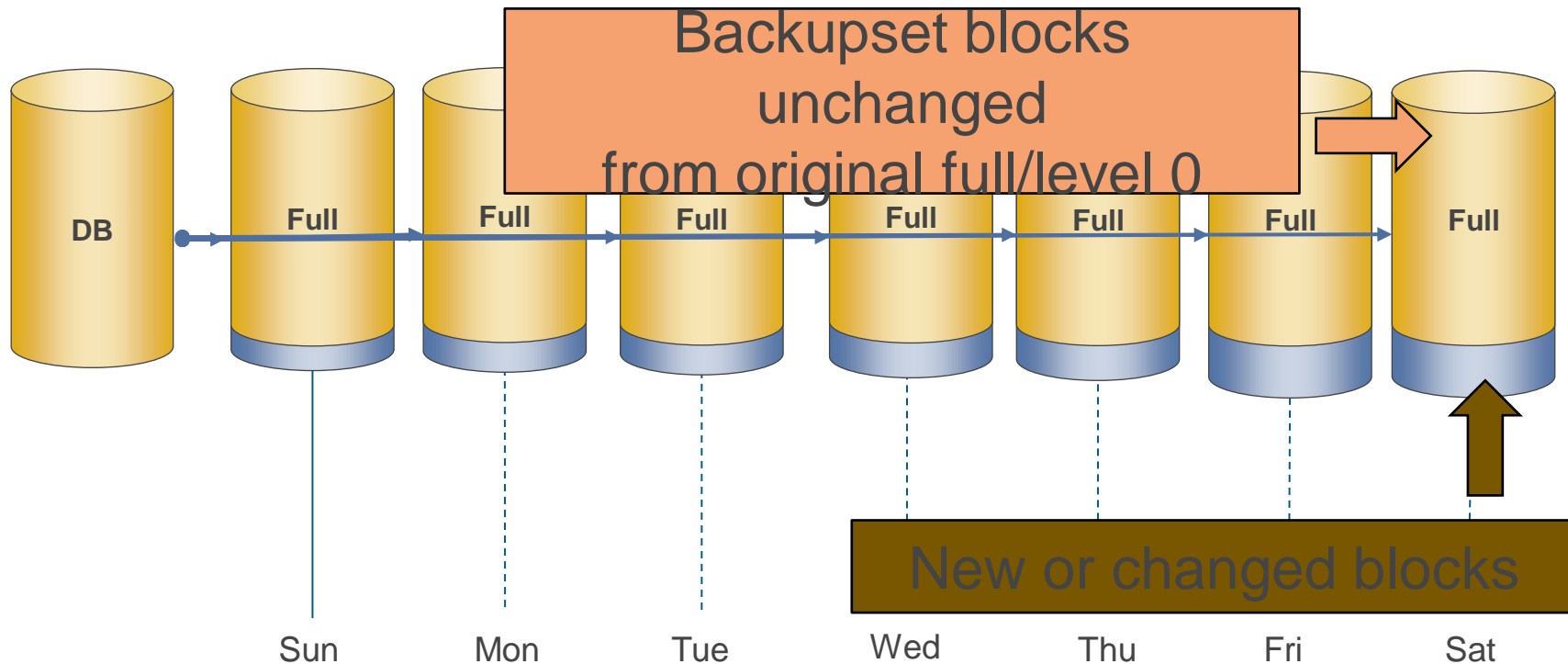
- ✓ Compression
- ✓ Encryption
- ✓ Standard (d)NFS or
- ✓ Tape emulation
- ✓ No 3<sup>rd</sup> party backup tool

# Oracle Backup to Data Domain



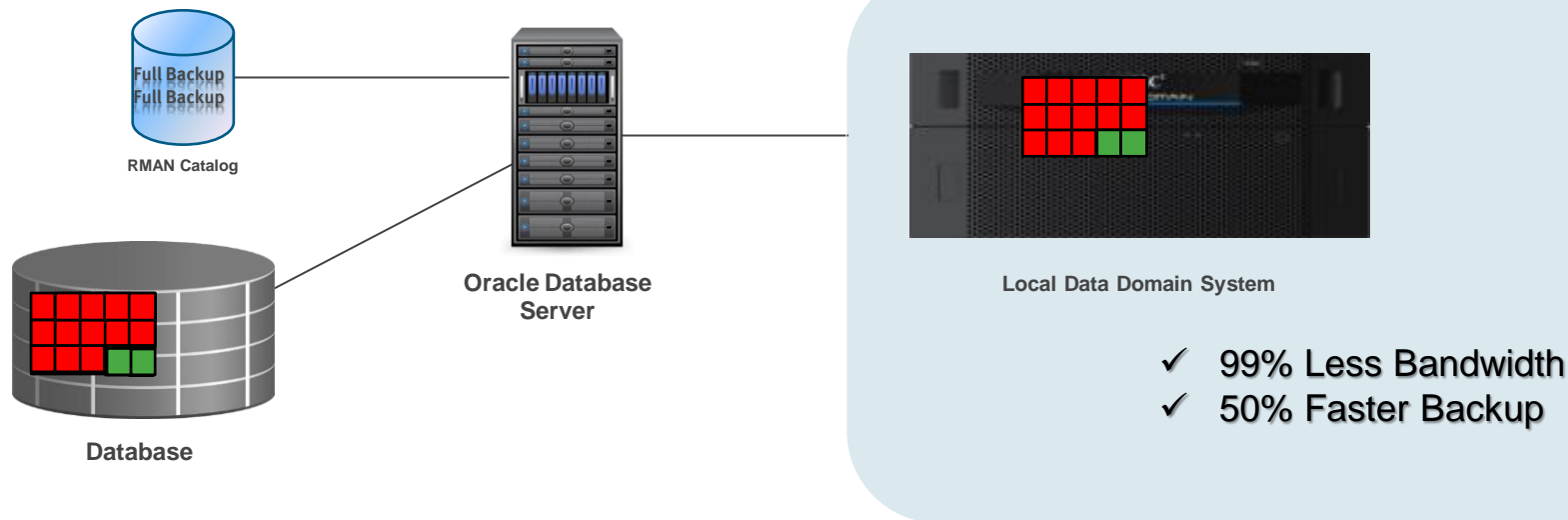
Dedupe makes backup fast and secure.  
No extra backup software or catalogs. No tape.

# Daily full backups?



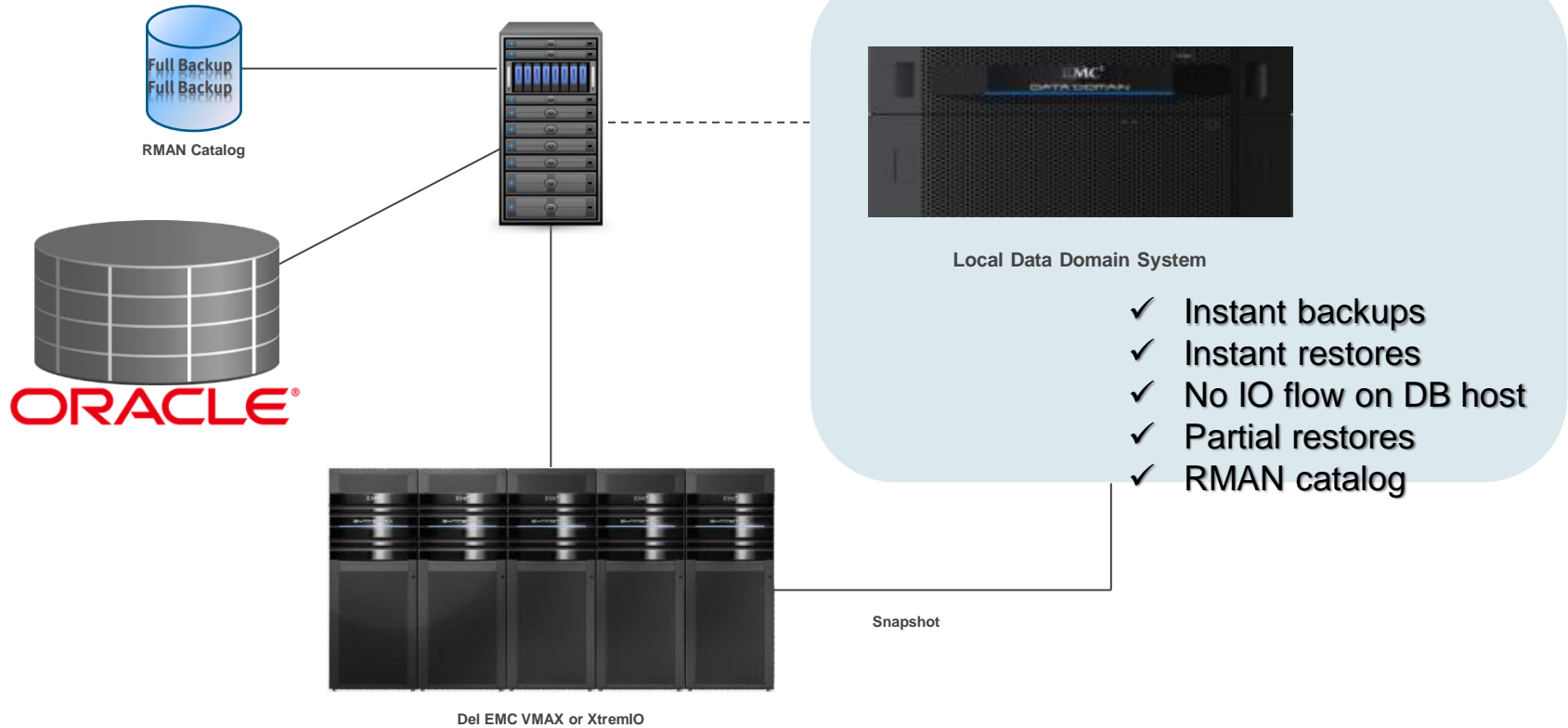
# DD Boost

Reducing backup window & network traffic



Boost de-dupes at source, dramatically accelerating RMAN backups

# Dell EMC ProtectPoint





An aerial night view of a city skyline, likely New York City, featuring numerous skyscrapers and dense urban development. The image is dark, with city lights providing the primary illumination. The text "Time to deploy" is overlaid in the center in a large, white, sans-serif font.

# Time to deploy

# Dell EMC Converged Systems

Various deployment models – pre-tested, built, validated, one-stop support



## Hyper-Converged Appliance

- ✓ Easy to Use
- ✓ Low Cost
- ✓ Start Small & Grow



## Hyper-Converged at Rack-Scale

- ✓ Scale Small to Big
- ✓ Software Defined
- ✓ Lowest TCO at Scale



## Converged Infrastructure

- ✓ Scalability
- ✓ Rich Data Services
- ✓ TCO



An aerial night view of a city skyline, likely New York City, showing numerous skyscrapers and buildings illuminated with lights. The perspective is from a high angle, looking down on the city. The lights from the buildings and streets create a vibrant, glowing pattern against the dark night sky.

# Services & Solutions

# Services & Solutions



## The EMC Oracle Joint Escalation Center

NOVEMBER 14, 2012 1 COMMENT (EDIT)



EMC and Oracle have supported each others products since 1995 and both spent millions of dollars in making them work together. EMC actually became famous in the late nineties because of our "Guilty until proven innocent" support mentality. We are known for the first company to give meaning to the concept of "Remote Support / Phone Home", and the success stories still go around that EMC field engineers sometimes surprised customers with a visit in order to repair components (mostly disk drives), often before they were broken, and if they were actually broken the customers would not even notice (needless to say that replacements were done online).

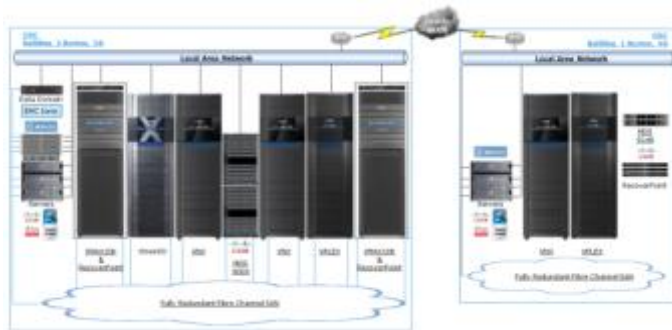


# EMC/Oracle Solutions Center



Oracle Campus, Reston, VA

- Shared services for Oracle & EMC
  - Over 500 servers
  - Over 1PB EMC storage
  - Fully Virtualized on VMware
- Provides infrastructure for
  - Oracle's Training & demos
  - EMC Demos
  - EMC POCs
- Oracle Integration Demos
  - Storage integration, cloning & replication
  - HA Stretched clusters
  - Management tooling



[Leverage EMC at Oracle Solution Centers](#)



# References

My Blog “Dirty Cache”

<http://bartsjerps.wordpress.com>

Everything Oracle @ EMC (community):

<http://emc.com/everythingoracle>

XtremIO

<http://xtremio.com/>

Outrun:

<http://outrun.nl>



## Dirty Cache

*A storage infrastructure perspective on optimizing business applications*

HOME

INDEX

RESOURCES

PRESENTATIONS

ABOUT

ORACLE

PERFORMANCE

INNOVATION

GENERAL

VPLEX

FAQ

VARIOUS

VIRTUALIZATION

← Thank you, Larry Ellison

### Stop Idling – Start Saving

OCTOBER 23, 2012 [LEAVE A COMMENT](#)



One of my missions is to help customers saving money (Dirty Cache Cash). So considering the average enterprise application environment, I frequently ask them where they spend most of their IT budget on. Is it servers? Networks? Middleware? Applications?

Turns out that if you look at the operating cost of an Oracle database application, a very big portion of the TCO is in database licenses. Note that I focus on Oracle (that's my job) but for other databases the cost ratio might be similar. Or not. But it makes sense to look at Oracle as that is the most common platform for mission-critical applications. So let's look at a database environment and forget about the application for now. Let's say that 50% of the operating cost of a database server is spent on Oracle licensing and maintenance



➤ Follow Blog via En

Enter your email address  
to receive notifications of new

Join 37 other followers

DELLEMC