

Exadata X7-2 POC with OVM

Jacques Kostic

Principal Consultant IMS Lausanne

Jacques.Kostic@Trivadis.com

Mobile: +41 79 909 72 63



Trivadis
makes IT
easier.

BASEL BERN BRUGG LAUSANNE ZUERICH DUESSELDORF FRANKFURT A.M. FREIBURG I.BR. HAMBURG KOPENHAGEN MUNICH STUTTGART VIENNA

■ About me...



ORACLE

Certified Master

Oracle Database 11g
Administrator

ORACLE

Certified Master

Oracle Database 12c
Administrator



Experience:

- Initially C/C++ developer
- In touch with Oracle since 1990 from version 4 on SCO Unix!
- High Availability and Backup & Recovery Architect
- SQL and Instance Performance & Tuning
- License Audit and Consolidation


Certifications:

- Oracle Certified Master 11g & 12c
- Oracle 11g Performance Tuning Certified Expert
- Oracle RAC 11g and Grid Infrastructure Administration
- Oracle Exadata Administrator Certified Expert
- Oracle Certified SQL Expert 11g

Teaching Courses at Trivadis:

- Oracle 11g & 12c Grid Infrastructure & RAC
- Oracle 11g & 12c Data Guard
- Oracle 11g & 12c Performance & Tuning
- Oracle 11g & 12c Administration
- SQL & PL-SQL
- OEM – 12 & 13

■ Our company.

Trivadis is a **market leader in IT consulting, system integration, solution engineering** and the provision of **IT services** focusing on **ORACLE®** and  **Microsoft** technologies in Switzerland, Germany, Austria and Denmark.

We offer our services in the following strategic business fields:



■ With over 600 specialists and IT experts in your region.



- 14 Trivadis branches and more than 600 employees
- 200 Service Level Agreements
- Over 4,000 training participants
- Research and development budget: CHF 5.0 million
- Financially self-supporting and sustainably profitable
- Experience from more than 1,900 projects per year at over 800 customers

■ AGENDA

1. Introduction
2. Current Oracle Architecture
3. Customer constraints
4. Fundamentals with Exadata
5. POC execution
6. Our Proposal with Exadata
7. Conclusion
8. Q&A



Introduction

Customer Overview

Customer Environment

The name will not be disclosed but the most relevant characteristics to the project are reported below.

- Relevant customer from banking sector
- Three DWH databases of 15 TB each
- Lack of resources
- Licensing constraints
- Consolidation opportunities with the new Exadata X7-2

Current Oracle architecture

■ Current Oracle architecture

- IBM AIX P7 PowerVM technologies, 2 LPAR PROD and QA on capped CPU POOLS
 - PROD: 9 VCPU max with 148 GB of RAM
 - QA: 6 VCPU max with 148 GB of RAM
 - SMT4 enabled
- 2 distinct sites with IBM Storage SVC replication
- Licensed 10 CPU Enterprise Edition with:
 - Partitioning
 - Diagnostic Pack
 - Tuning packs
- Poor IO performances
- System CPU saturated

Customer constraints

■ Customer constraints

- 2 more environments (5 in total)
- Have a performant and scalable architecture where running the DWH
- Have a real disaster recovery solution
- Limit the investment

**To do the best with the
minimum!**

Fundamentals with Exadata

■ Exadata X7-2: Without OVM

Pros.

- Use the entire machine capacity
- Less servers to manage
- Pay-as-you-grow approach (COD) for software licensing is another way in which Exadata helps to align costs with business growth
 - Minimum 14 cores per DB nodes (8 for Eighth Rack)
 - All additional options must follow the same allocation
 - https://docs.oracle.com/cd/E80920_01/DBMLI/exadata-capacity-on-demand.htm#DBMLI147

Cons.

- Isolation between databases and environments
- License optimization

■ Exadata X7-2: With OVM

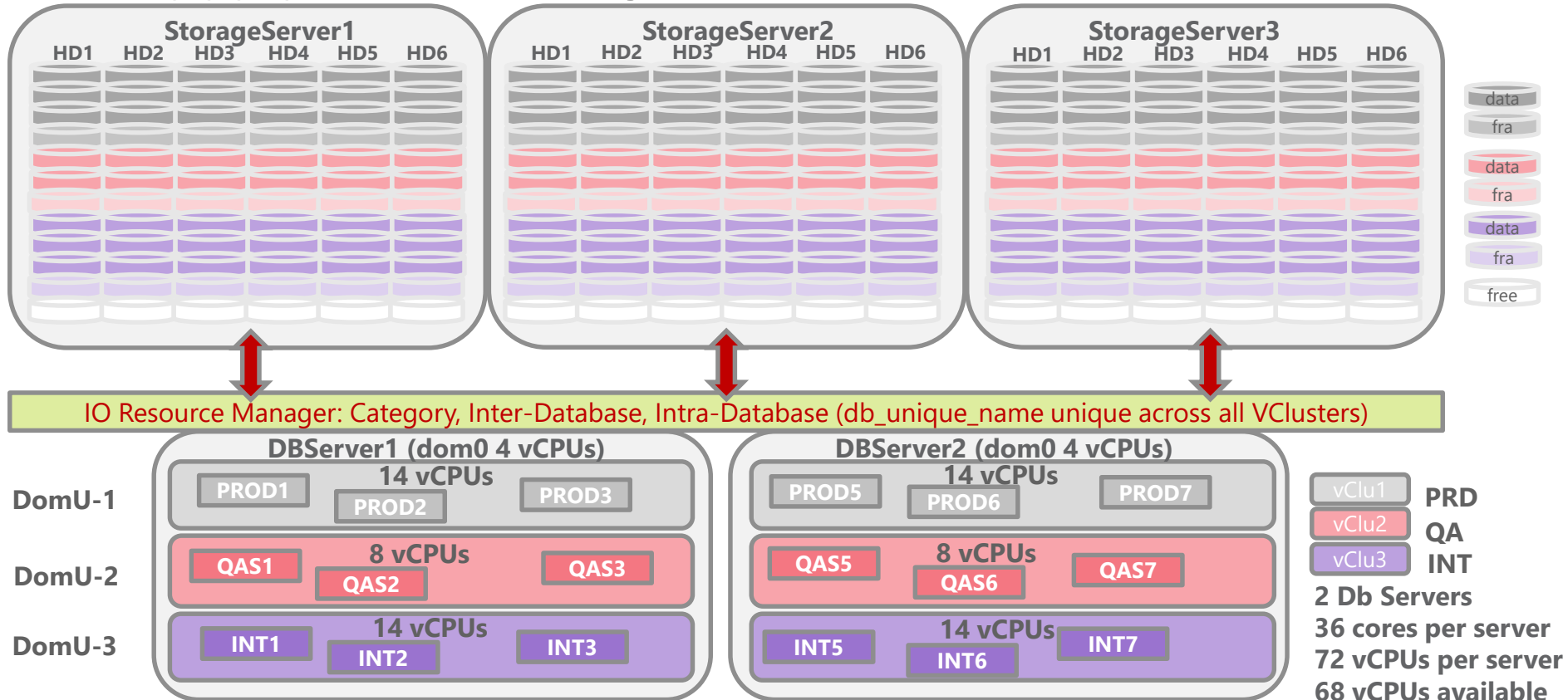
Pros.

- Environment and database isolation
- Hard partitioning facilitate licensing optimization
 - Minimum 14 cores per DB nodes (8 cores for Eighth Rack) must be licensed for Enterprise Edition
 - For other options, it's linked to CPU allocation for each VM
 - *Two cores per database node dedicated to dom0 (out of software licensing)*
- Very flexible, dynamic vCPU allocation
- Allow IO resource management between all database from all virtual machines.
 - Db_unique_name **must** be unique across the entire Exadata

Cons.

- Might appear more complex to manage

Exadata X7-2: With OVM



■ Exadata X7-2: OracleVM overview on Exadata

■ Deployment

- Create configuration (clusters) with Oracle Exadata Deployment Assistant (OEDA) Configuration tool
 - OEDA Configuration tool version May 2018
 - <http://www.oracle.com/technetwork/database/exadata/oeda-download-2076737.html>
- Prepare system
 - IP allocation, customer requirements
- Deploy configuration using OEDA Configuration tool

Exadata X7-2: Cluster deployment example

The screenshot displays the Oracle Exadata Deployment Assistant interface for configuring a cluster. The left sidebar shows a navigation tree with 'Cluster 1' selected under 'Review and Edit'. The main window is titled 'Cluster 1' and contains the following configuration sections:

- Cluster 1**
 - Cluster name: cluster-clu1 OVM Virtual Cluster
 - Virtual Guest size: Large Cores: 26 Memory: 96 Local Disk: 60GB
 - Prefix: sgexa Reset name masks
 - DNS: 172.17.15.213 172.17.15.215
 - NTP: ntp.customer.ch
 - Domain Name: customer.ch
 - Region: Europe TimeZone: Zurich
- Users and Groups**
 - Role Separated
 - User name: oracle ID: 1001 base: /u01/app/oracle
 - DBA Group name: dba ID: 1002
 - OINSTALL Group name: oinstall ID: 1001
- Software Locations**
 - Inventory Location: /u01/app/orainventory
 - Grid Infrastructure Home: 12.1.0.2 DBBP5 /u01/app/12.1.0.2/grid
 - Database Home Location: 11.2.0.3 BP25 /u01/app/oracle/product/11.2.0.3/dbhome_1
 - Software Languages: en
- Disk Group Details**
 - Diskgroup Layout: Legacy 80%:20% Legacy 40%:60% Custom
 - DBFS DiskGroup: DBFS_DG NORMAL Size: 500GB
 - DATA DiskGroup: DATA1 HIGH Size: 3TB
 - RECO DiskGroup: RECO1 NORMAL Size: 2TB
- Initial Database**
 - Database Name: TST01 Block Size: 8192 Type: OLTP DW
- Admin Network**
 - Base Adapter: Admin Domain: customer.ch
 - Start IP: 172.21.81.60 Pool size: 2
 - Name mask: sgexadm%vm01 Start id: 1
- Client Network**
 - Base Adapter: Client Domain: customer.ch
 - Start IP: 172.21.82.32 Pool size: 7
 - Gateway IP: 172.21.82.254

At the bottom, there are buttons for 'Help', 'Import...', 'Oracle_YM Defaults...', 'Save', '< Back', 'Next >', 'Finish', and 'Cancel'.

POC Execution and Result

■ POC Execution: the context

Our Competitor

- IBM P8
- Full Flash Storage
- Max 16 Cores with SMT8
 - Tests done with 10, 12,14 cores SMT8
 - Corresponding CPU licenses: 10, 12, 14
- 1 database 18 TB with 90 GB of SGA
- Oracle 12.2.0.1

■ POC Execution: the context

Our Environment

- Exadata X7-2 1/4 rack
- OVM Configuration
- Single instance mode
- Two-node cluster with various vCPUs configurations
 - 36, 28, 24, 20, 16
 - Corresponding CPU licenses: 9, 7, 6, 5, 4
- 1 database 18 TB with 90 GB of SGA
- Oracle 12.2.0.1

■ POC Execution: the context

In Summary

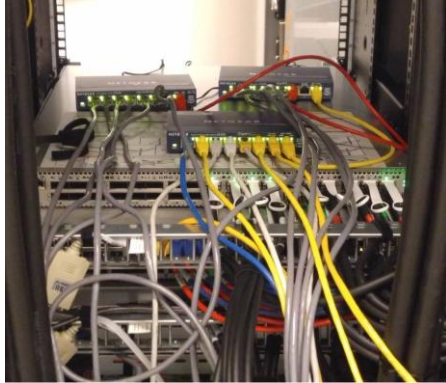
- Trivadis is proposing a complete change of architecture
- IBM is just replacing P7 by P8 and adding Full Flash Storage

■ POC Execution: the setup

- We had a lot of problems to setup the stuff
- We had to use October 2017 Image
- Thanks a lot to Arrow for the help!

■ POC Execution : the setup

We had network problems with the management switch!



In reality the step was failing because the Switch was OFF



■ POC Execution: the initial load

- Import took more than 54 hours for 3M

- It took

**We get finally
ready to start!**

- We used Multi-tenant capabilities during the POC
 - Pluggable database snapshots

■ POC Execution: the result

- IBM was able to increase the load speed by a factor of four.
- But it was achieved by:
 - using the 14 cores (SMT8) configuration
 - Setting the optimizer to 11.2.0.4 features!
 - Many ORA-00600 on stats export/imports during the load processing
 - High CPU usage during the processing
- Runs with 12 and 10 cores were CPU bound
 - But still performing around 2.5 better than the current state
- Because run using 14 cores was not CPU bond, they stopped at that level.
- But they did not tried to run using optimizer_feature='12.2.0.1'!

■ POC Execution: the result

- We start our first run with 36 vCPUs and we achieved a speed increase with a factor of two.
- But it was achieved with:
 - Leaving the optimize to the default 12.2.0.1 value.
 - Low CPU usage
 - Average IO wait time of 35 microseconds!
- Some jobs were running very badly and we discovered that the optimizer setting was not the same used by our competitor 😊
- We decided to fix the underlying queries!

■ POC Execution: the result

- Some queries were hinted to use optimizer_feature='12.1.0.2'
- Some queries were hinted to use optimizer_feature='11.2.0.4'
- Some queries were hinted to avoid view merge
- Some queries were hinted to avoid materialize of a particular factoring clause

■ POC Execution: the result

- We were finally able to achieve the same performance result obtain by IBM!
- We decided then to start downsizing the vCPU configuration to see what we can get from this beast!
- Runs with 28 and 24 did not change the performances at all!
- We got 3% less performance with 20 vCPU and around 8% less with 16 vCPU!
- CPU usage was high but acceptable with the 16 vCPU configuration

■ POC Execution : Conclusions

Following our different runs

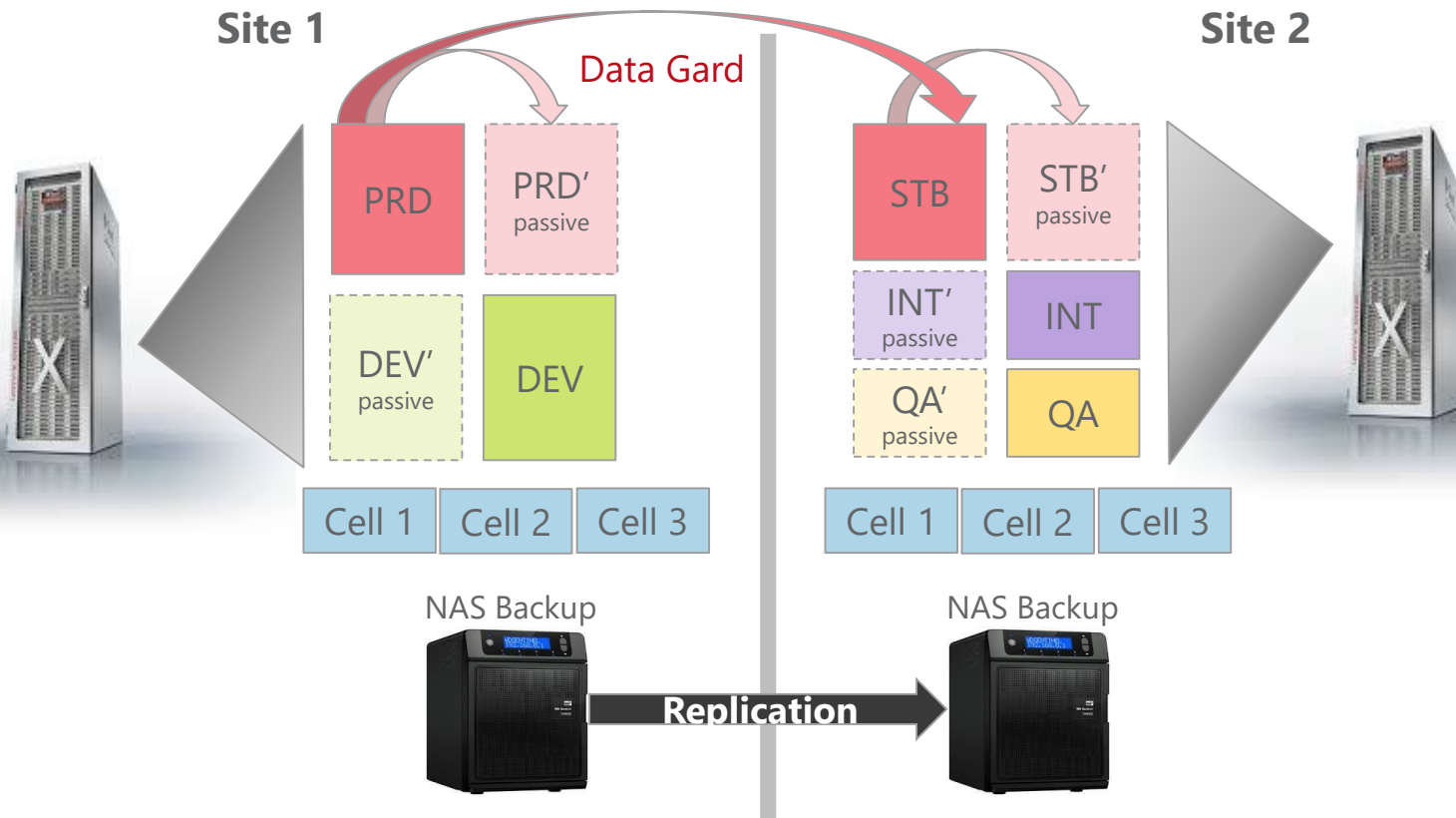
- We decide to adjust our final offer to customer with 24 vCPU
- There are still lot of optimizations to be done!

Our Proposal with Exadata

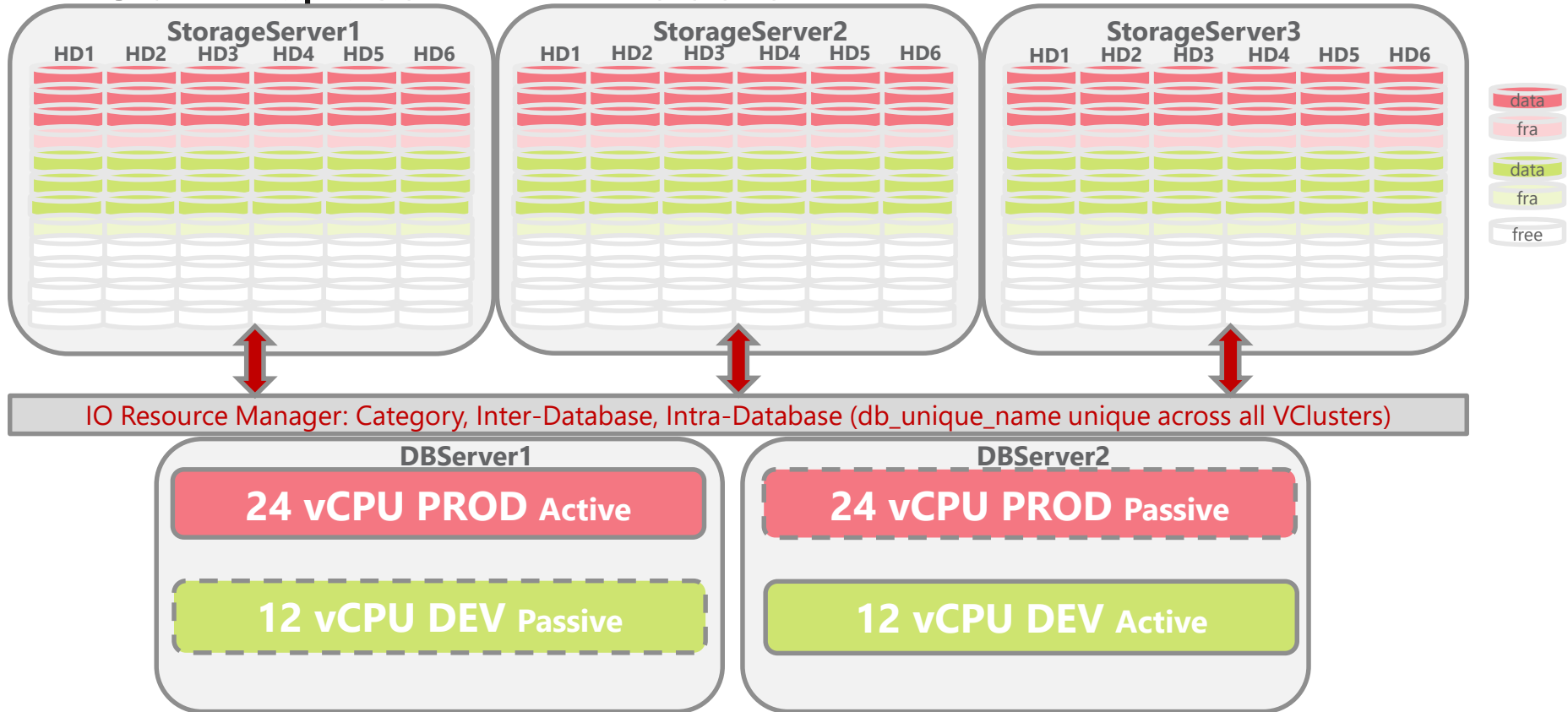
■ Our Proposal with Exadata

- Cold Failover mode
 - Oracle Active/Passive *10-days-per-year*
 - <http://www.oracle.com/us/corporate/pricing/data-recovery-licensing-070587.pdf>
- 18 CPU Licenses required
 - Enterprise Edition (minimum 14 CPU Licenses)
 - Partitioning
 - Diagnostic and Tuning Packs
- Single instance mode with Oracle 12.2.0.1

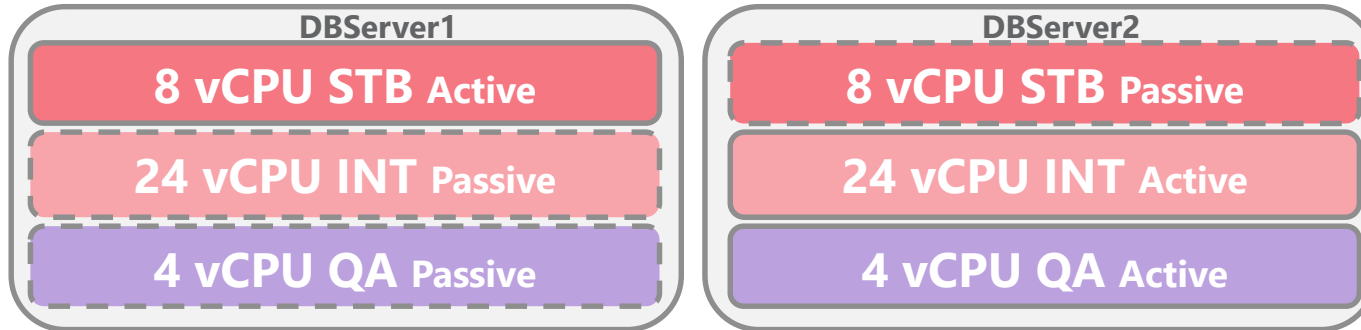
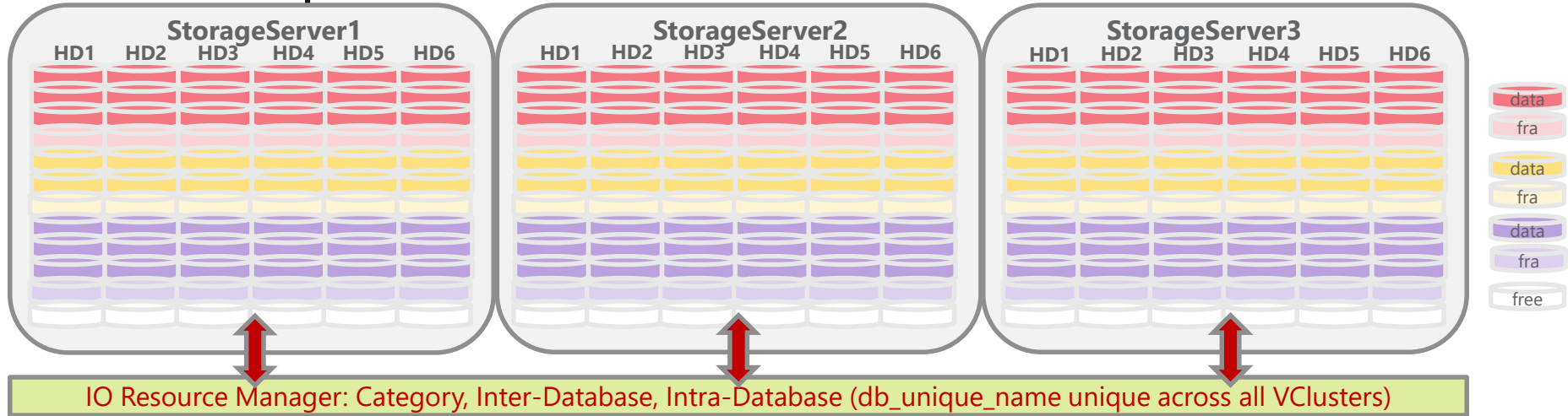
■ Our Proposal with Exadata



■ Our Proposal with Exadata



■ Our Proposal with Exadata



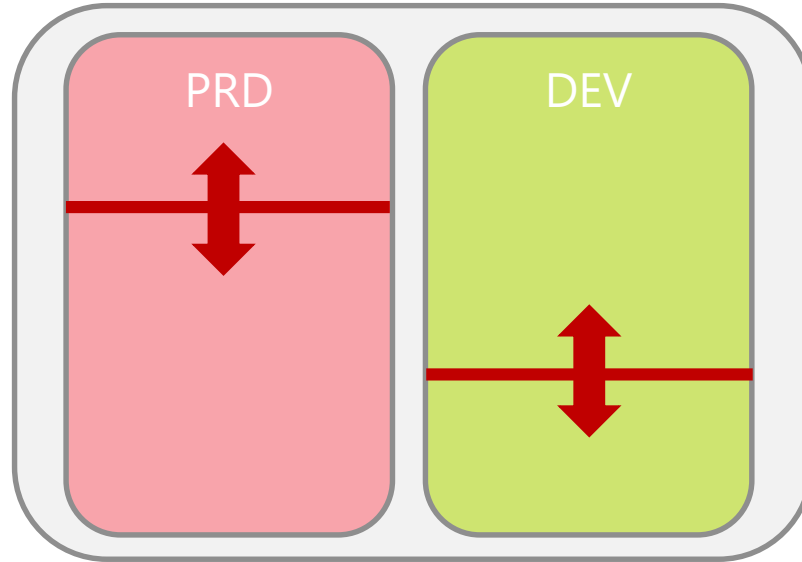
■ Our Proposal with Exadata

Adjust power on demand: MAX 34 vCPUs per VM

34 vCPUs

18 vCPUs

2 vCPUs



- Dynamic host cpu reconfiguration using: `xm vcpu-set`
- Dynamic oracle CPU_COUNT adjustment as of Oracle Oracle 12c
 - Dynamic resource management update

■ Our Proposal with Exadata

- Use ASM Sparse Disk Group for DEV environment
 - Allow fast snapshot cloning to provisions new DEV database
 - Smart Scan is supported!
- Need a standby from production to manage the clones
 - Apply Off → Clone → Apply On
- IO Performance degradation:
 - 100 time slower
 - 35 microsecond vs 3.5 millisecond
- More info's here: <https://emilianofusaglia.net/tag/asm-sparse-disk-group/>

Conclusion

■ Conclusion

- We t
- Scal
- Con
- the g



ctor

ident for

Questions...

Jacques Kostic

Principal Consultant IMS Lausanne

Jacques.Kostic@Trivadis.com

Mobile: +41 79 909 72 63



Trivadis
makes IT
easier.

BASEL BERN BRUGG LAUSANNE ZUERICH DUESSELDORF FRANKFURT A.M. FREIBURG I.BR. HAMBURG KOPENHAGEN MUNICH STUTTGART VIENNA