Oracle for administrative, technical and Tier-0 mass storage services openlab Projects Status Review

Eric Grancher
IT-DES

Work from

Lucia Moreno Lopez, Maria Leitner, Andre Dechert, Anton Topurov, Chris Lambert and Eric Grancher

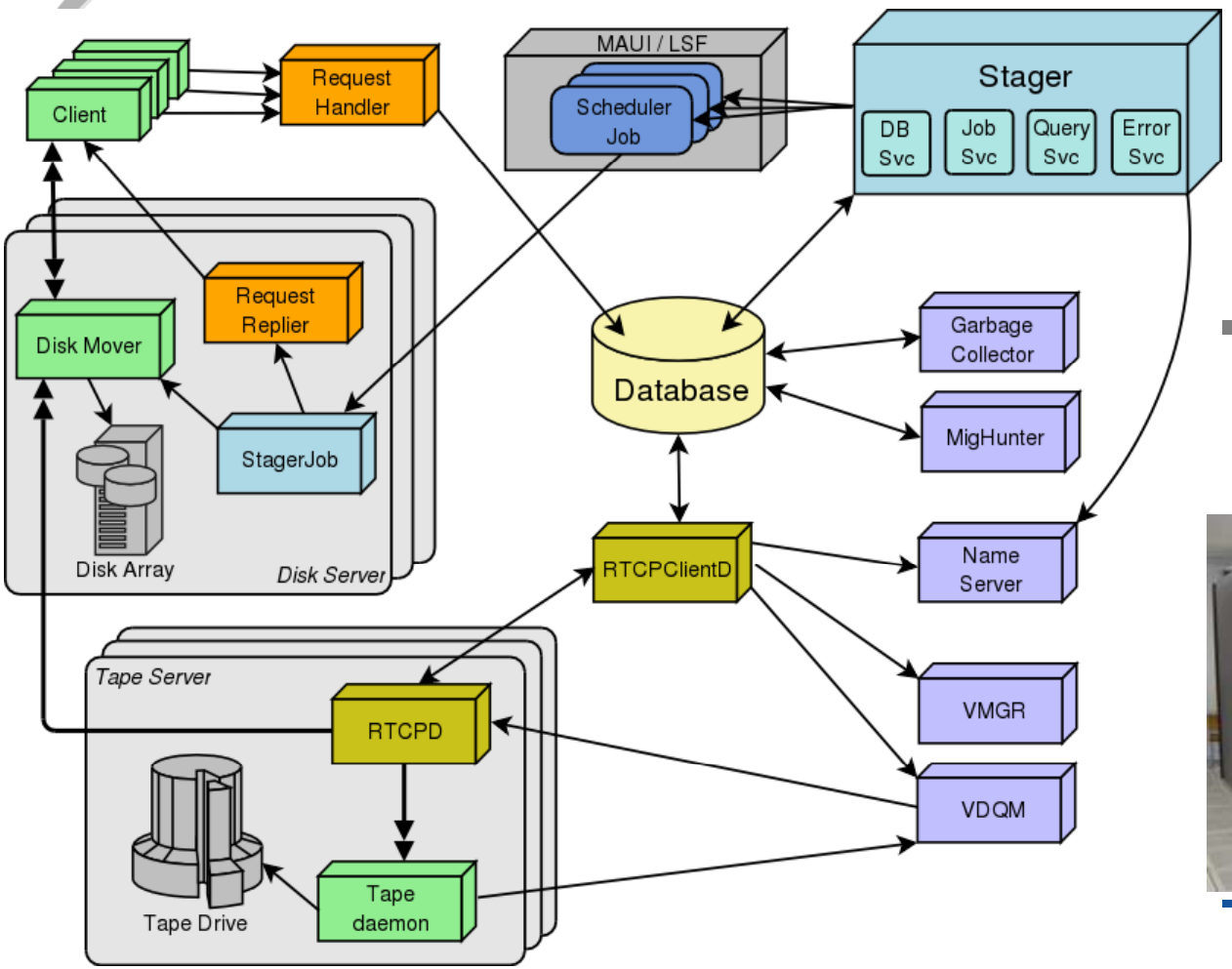
Contents

- [JST] Joint Software Testing Programme
- [OPT] Oracle Performance Testing
- [RST] RAC Scalability Tuning
- [RAC] Oracle RAC virtualisation
- [OEM] Enterprise Manager (monitoring)

[JST] – 11g RPMs

- OS packages help to be more efficient, less error-prone work, integrated with Quattor
  (already deployed on 15 clusters RAC 10.2 x86_64)
- Knowledge obtained during JST 11g tests helped for creation of 11g CRS and RDBMS OS packages (RPM)
- Test and debugging of 11g RPMs x86_64 done
- Successful deployment of “CRS” (openlab) RAC, 11.1.0.6 x86_64

[RST] Castor tests



- CERN Advanced STORage manager, hierarchical storage management (HSM)
- Store physics production files and user files

CASTOR implementation (DB)

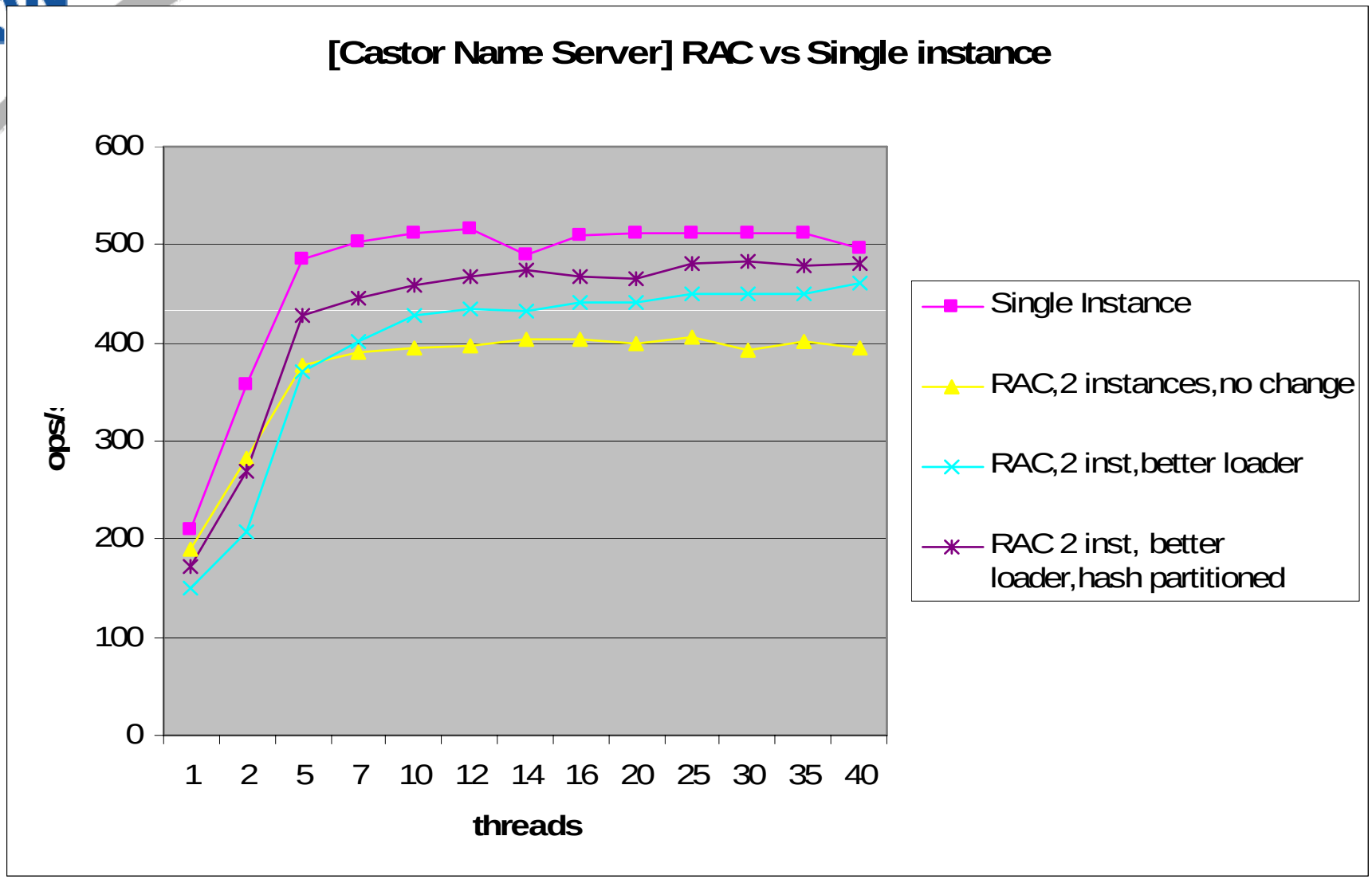
- 7 RAC databases for now (HP/Intel/NAS)
- CASTOR Name Server (CNS)
  - Implements hierarchical view of CASTOR name space. Stores file access permissions, file size, access times, file location on tertiary storage if the file has been migrated from the disk pool in order to make space for more current files
  - Multithreaded, throttled software
  - Scalability is crucial
  - Uses Oracle Database for storing files metadata

[RST] CNS Stress test application

- Multithreaded
- Used with up to 40 threads
- Each thread loops 5000 times on
  - Creating a file
  - Checking it's parameters
  - Changing size of the file

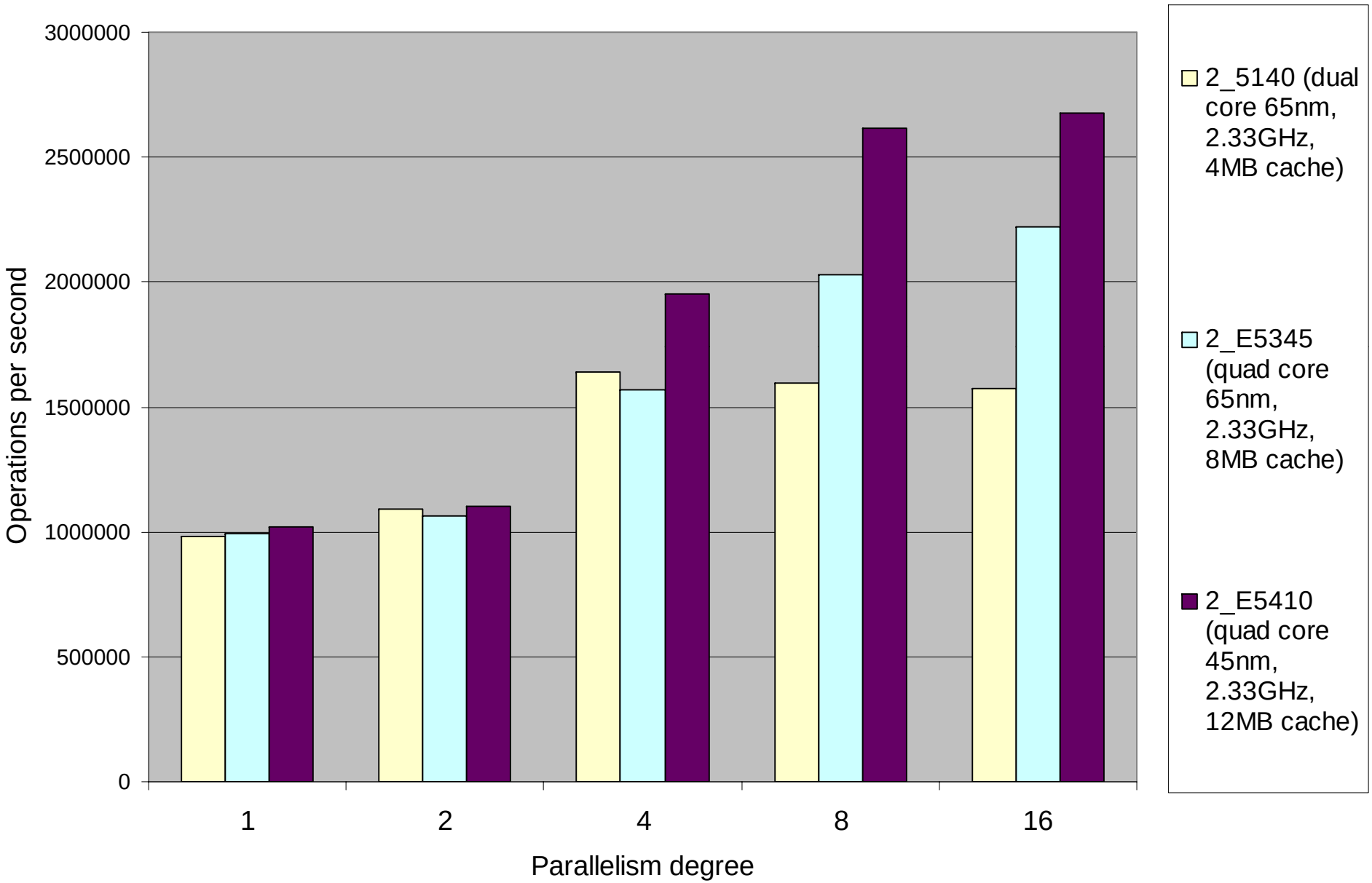
- Test Made:
  - Single instance vs. 2 nodes RAC
  - (almost) No changes in schema and application code

[RST] Results


[Castor Name Server] RAC vs Single instance
600
500
400
300
200
100
0
1
2
5
7
10
12
14
16
20
25
30
35
40
threads
Single Instance
RAC,2 instances,no change
RAC,2 inst,better loader
RAC 2 inst, better loader,hash partitioned

[OPT] Oracle Performance Tests

- Test and validate performance for newer platform
- Oracle RDBMS performance comparison between (all dual CPUs platforms):
  - 5140 (65nm, dual-core, 2.33Ghz, 4MB cache, 1333MHz FSB, “Woodcrest”), *current deployment platform for IT-DES Linux RACs*
  - E5345 (65nm , quad-core, 2.33Ghz, 8MB cache, 1333MHz FSB, “Clovertown”)
  - E5410 (45nm , quad-core, 2.33Ghz, 12MB cache, 1333MHz FSB, “Harpertown”)
- Logical IO capacity measurements (LIOCM), Oracle database 10.2 x86_64


Number of logical iops, row length 2000 bytes
Operations per second
3000000
2500000
2000000
1500000
1000000
500000
0
1
2
4
8
16
Parallelism degree
2_5140 (dual core 65nm, 2.33GHz, 4MB cache)
2_E5345 (quad core 65nm, 2.33GHz, 8MB cache)
2_E5410 (quad core 45nm, 2.33GHz, 12MB cache)

[OPT] Results and Conclusions

- Xeon E5410 versus Xeon 5140
  = +70% for parallel degree 16
- Decision to go for E5410
- New servers are in testing phase

- Waiting for 1600MHz Front Side Bus and Intel upcoming CSI
- As expected, low parallel degree applications do not profit from the new platform … good design, virtualisation

[RAC] Oracle Real Application Clusters Virtualization

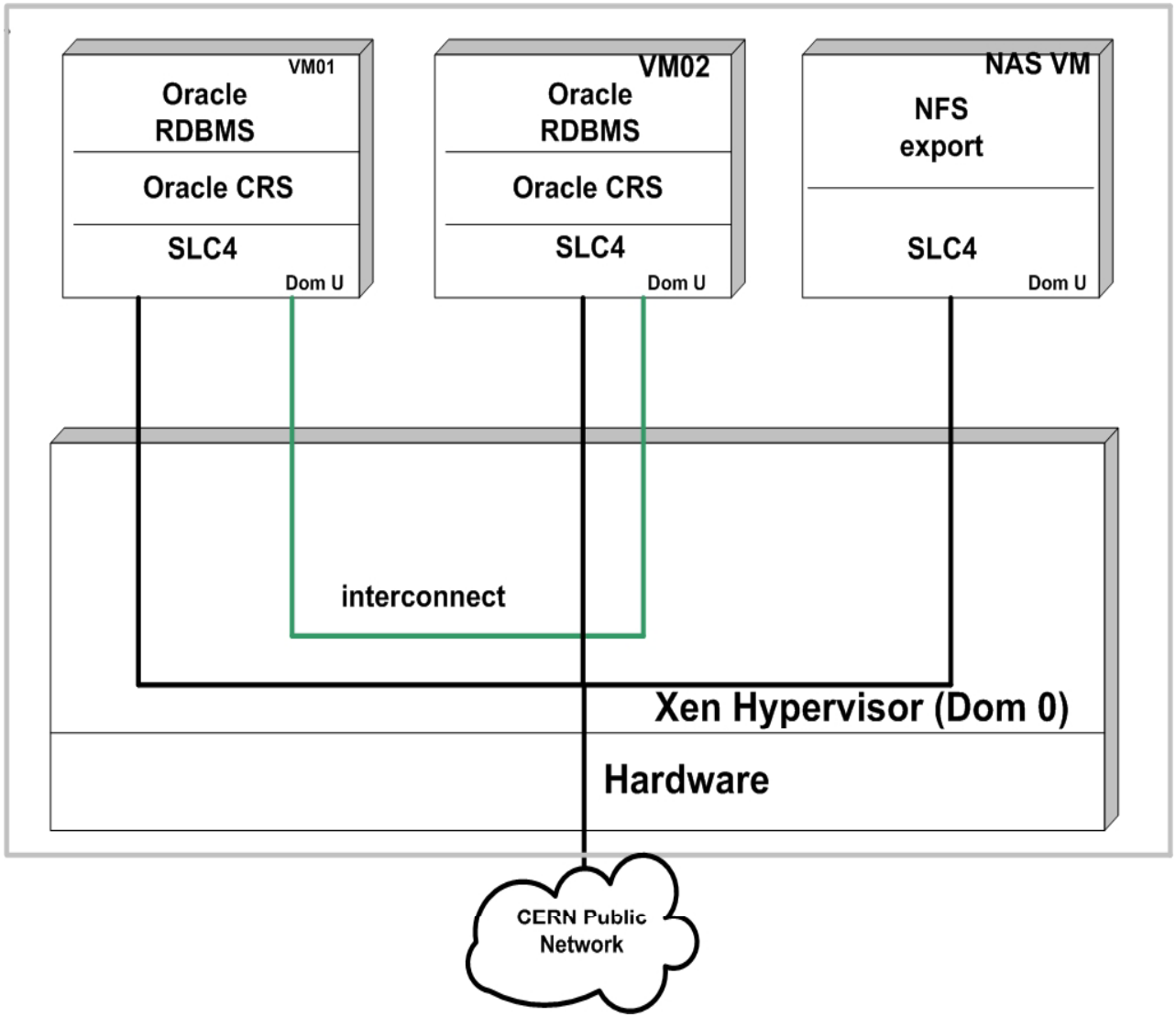
- openlab summer student project made by Maria Leitner
- Tests and conclusions by Anton Topurov

- Main goals:
  - Implementing Oracle RAC in virtualized environment
  - Identify advantages and drawback of the setup
  - Make virtual environment closest possible to our NAS-based RACs
  - Document the implementation process

[RAC] Oracle RAC virtualization

Setup 1:

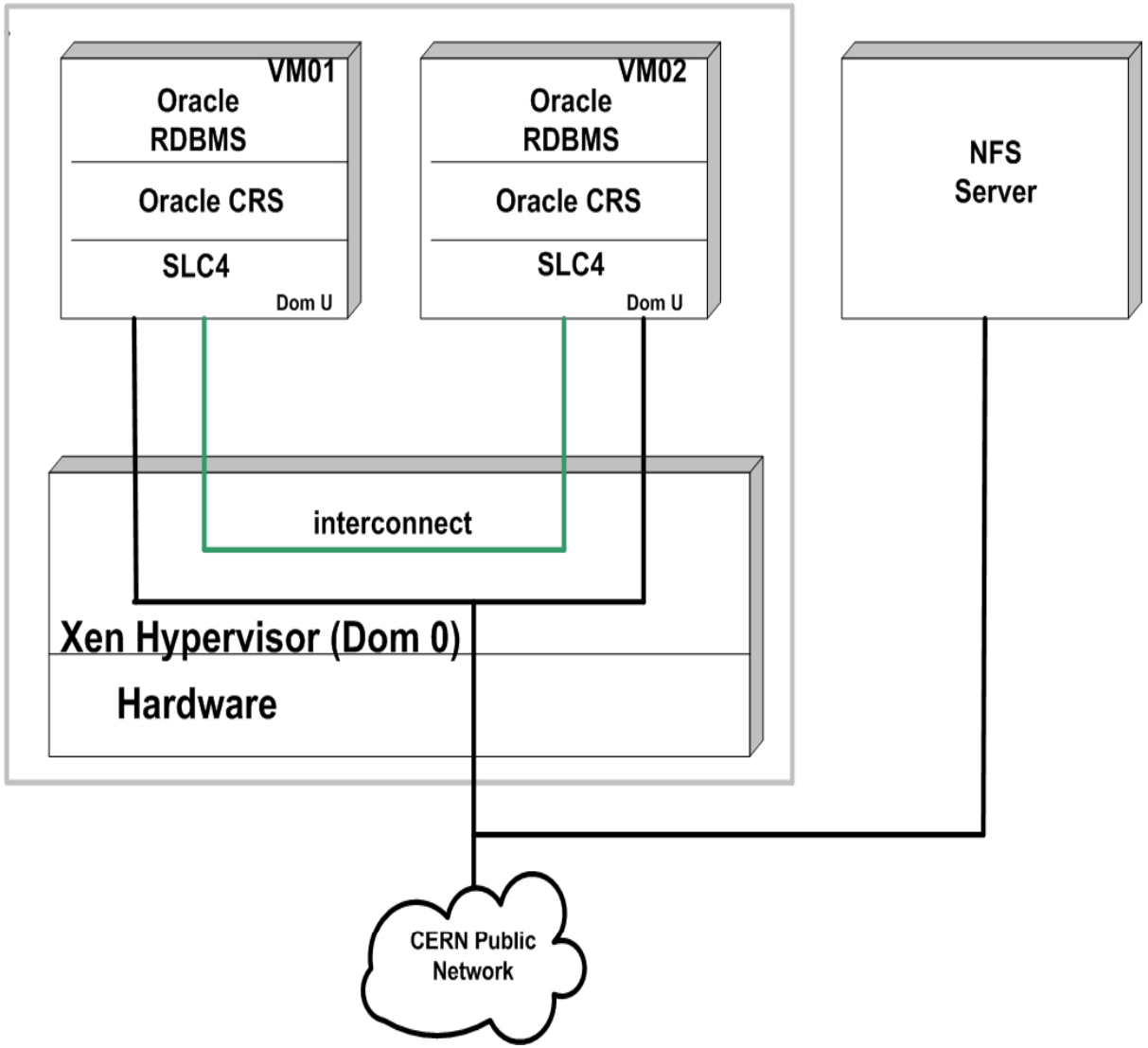
- SLC4 x86_64
- Xen 3.6
- Oracle 10.2.0.3
- 2-nodes RAC
- NFS export from 3rd VM



[RAC] Oracle RAC virtualization

Setup 2:

- Same as setup 1
- NFS export from real machine



[RAC] Progress

- Progress done:
  - Setup 1 tests:
    - done with Swingbench 2.2
    - compared with single instance on the identical node
    - Single instance in better performing than VM RAC
    - Comparison is not 1 to 1, due to different I/O subsystems
  - Setup 2 tests ongoing:
    - Both single instance and VM RAC are using external NFS exports for the database files
    - Partial test results show gain in performance for VM RAC

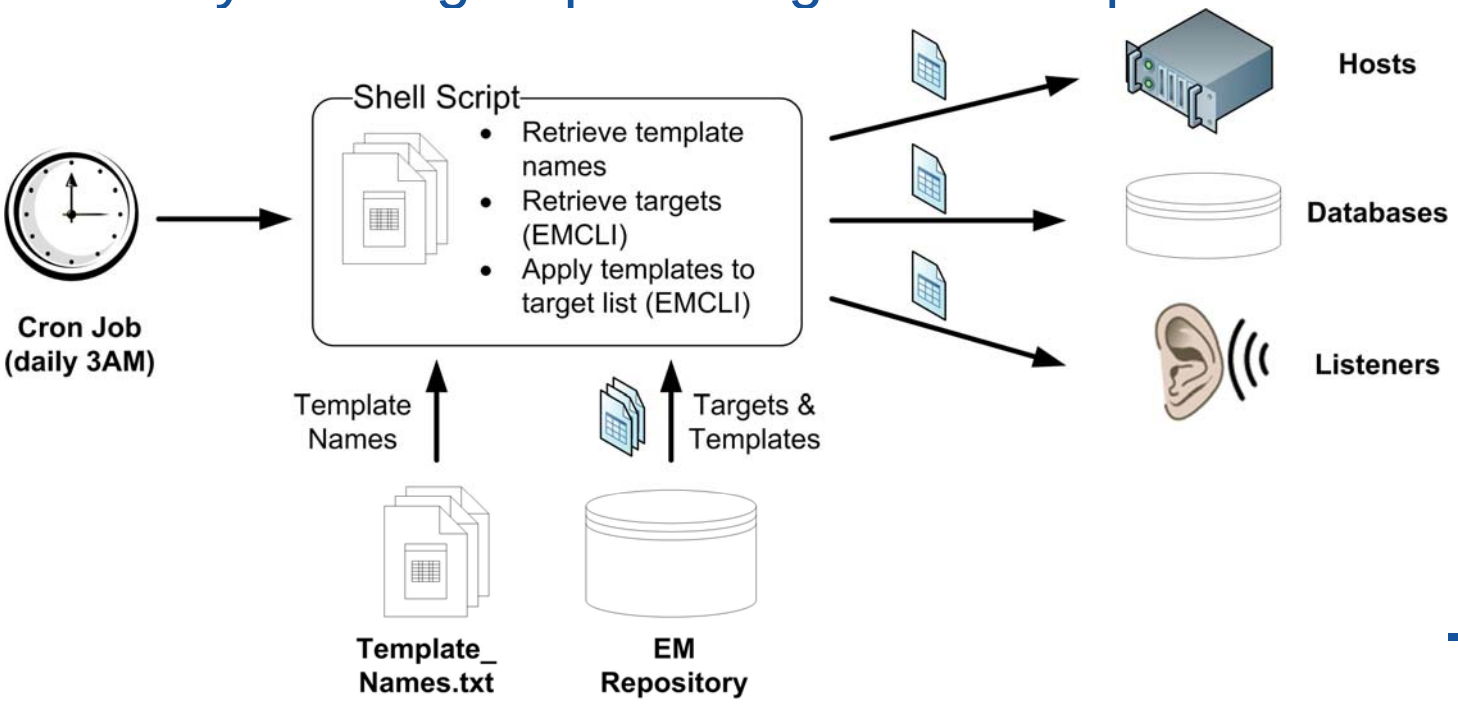
[OEM] Achievements

- **Upgrade EM to 10.2.0.4**
  - test environment has already been refreshed
  - Target for production by end March (between CPU releases)
  - Developer “more-friendly” interface
- **Automated Agent Installation**
  - Nearly all machines have 10.2.0.3 agents
  - Using RPM’s to automate installation
- **Infrastructure databases**
  - Agents installed on machines (manually/rpm)
  - Databases discovered
  - Added to groups for security/backup monitoring etc.

- **We have signed-up for EM11g Beta programme**

Automated Application of Monitoring Templates

- Automatically apply security/backup policies to targets every night
- Avoid targets falling through the net
- dynamic groups in 11g should replace this work-around

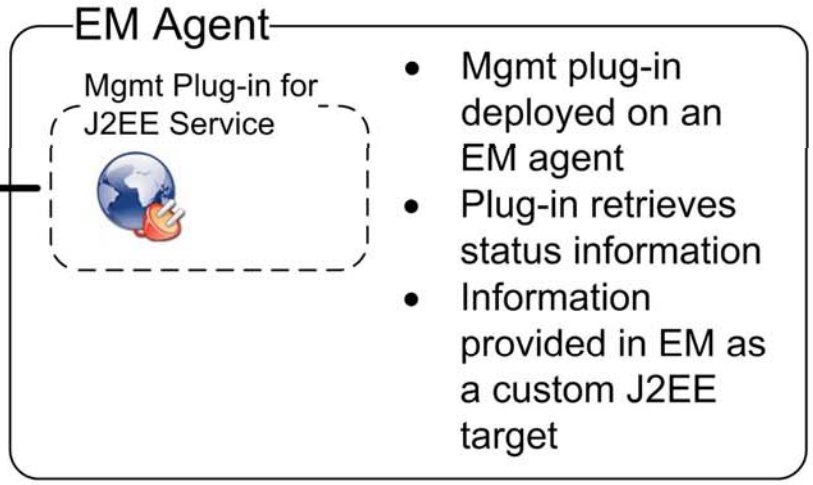


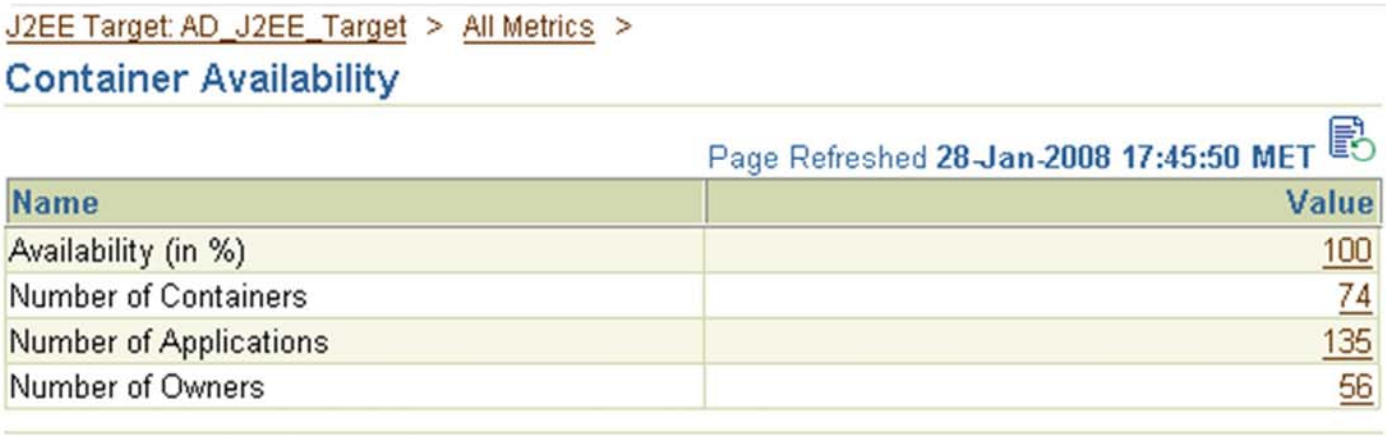
[OEM] Achievements 2/2

- **Metrics / User-defined Metrics**
  - technical aspects already in place (CPU, Memory,...)
  - moving to the service level / start to monitor business aspects (tracking the batch processes, get alerts when they run longer, get trends, how to monitor unavailability) – goal for 2008
- **Management Plug-ins for CVS & J2EE**
  - started to use the Management Plug-in functionality to integrate custom targets (early stage)

```
- <serviceupdate xsi:schemaLocation="http://sls.cern.ch/SLS/XML
    <id>JPS</id>
    <timestamp>2008-01-28T17:32:01</timestamp>
    <availability>100</availability>
    <refreshperiod>PT5M</refreshperiod>
    <notes>All applications are available.</notes>
  - <data>
      <numericvalue name="containers">74</numericvalue>
      <numericvalue name="applications">135</numericvalue>
      <numericvalue name="owners">56</numericvalue>
    </data>
</serviceupdate>
```

Status Information
(published by J2EE Service / https://jpsmanager.cern.ch/jps/central/showSLS.do)





Name	Value
Availability (in %)	100
Number of Containers	74
Number of Applications	135
Number of Owners	56

Metrics Page of the J2EE Custom Target
(Enterprise Manager 10g)

Conferences exposure: **UKOUG 2007**

Two presentations by our group

- **Eric Grancher and Anton Topurov**

  "Real life experiences of RAC scalability with a 6 node cluster"
- **Philippe Defert**
  - "Installing and managing hundreds of RAC servers"
- + 3 others from CERN

Thank you!