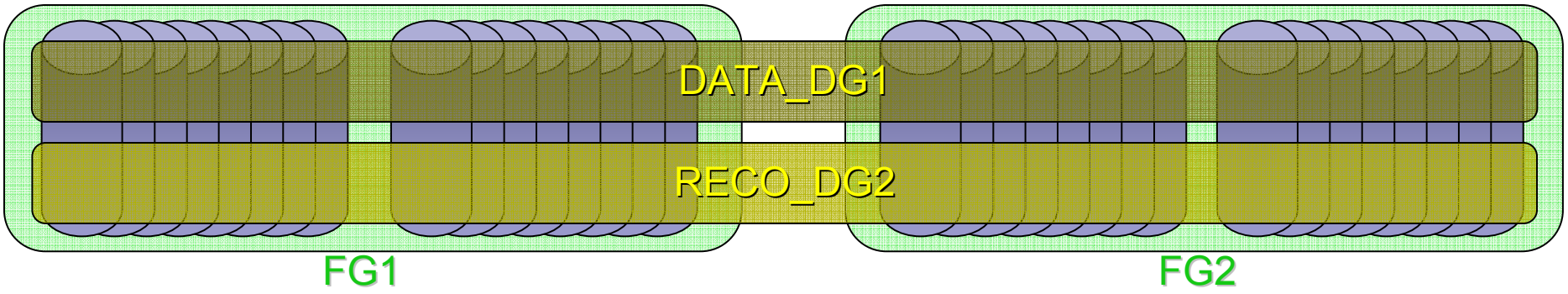
RAC/ASM update

CERN openlab II
Monthly Technical Review
27th March 2007

Dawid Wójcik

ASM on CERN RACs

- Oracle Automatic Storage Management
  - volume manager and cluster file system (online storage reconfiguration and rebalancing)
  - high performance/cost ratio
  - built-in mirroring (between failgroups) and striping (within diskgroups)
- Current production setup
  - each Fibre Channel storage has 8 SATA disks (250GB each)
  - each disk divided in 2 partitions (external, faster, used for data diskgroup, internal for recovery diskgroup)
  - normal redundancy used (two failgroups for each diskgroup)



ASM summary & issues

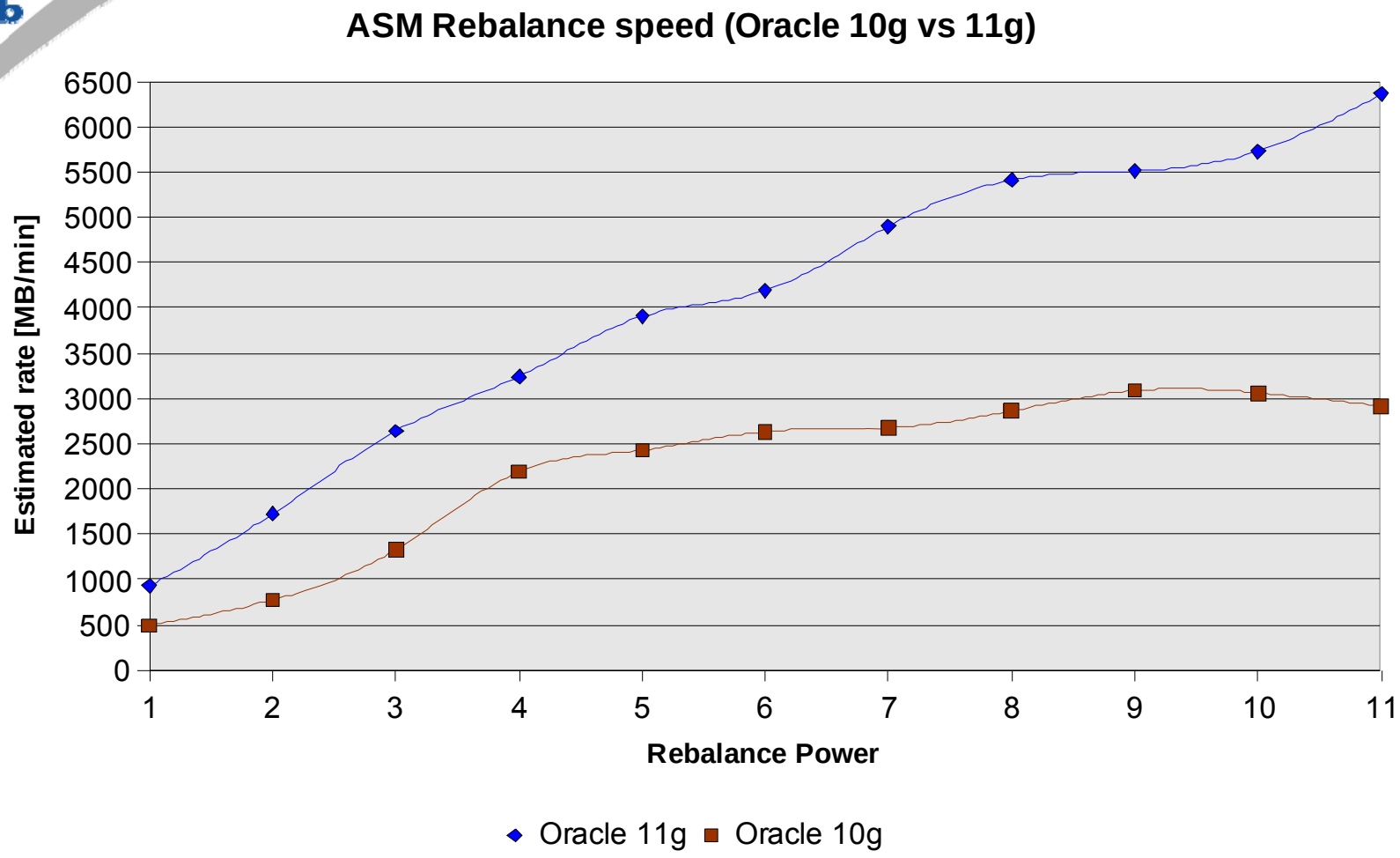
- ASM stable since 10gR2
- Recent changes in production setup
  - underlying QLogic's driver multipathing replaced with multipath and devicemapper
  - storage expansion
    - biggest setup – 6 storages ≈ 35TB (after mirroring: ≈ 9TB of data & backup space each)
    - biggest database >4TB of data
- Performance issues encountered – rebalance operations are slow due to
  - disk partners handling
  - cluster communication

ASM issues continued

- Slow rebalance – *ASM disk partners handling*
  - becomes an issue when maintaining large storages (we have max 96 LUNs)
    - disk failure – no data loss, but time consuming rebalance (potential data loss if second failure occurs)
  - exchanging remarks with oracle support
  - Bjørn confirmed that improvements are coming ...

- Slow rebalance – *cluster communication*
  - many cluster wait events (DFS), buffer busy
  - rebalance speed nonlinear with *rebalance power*
  - non CPU or I/O bound

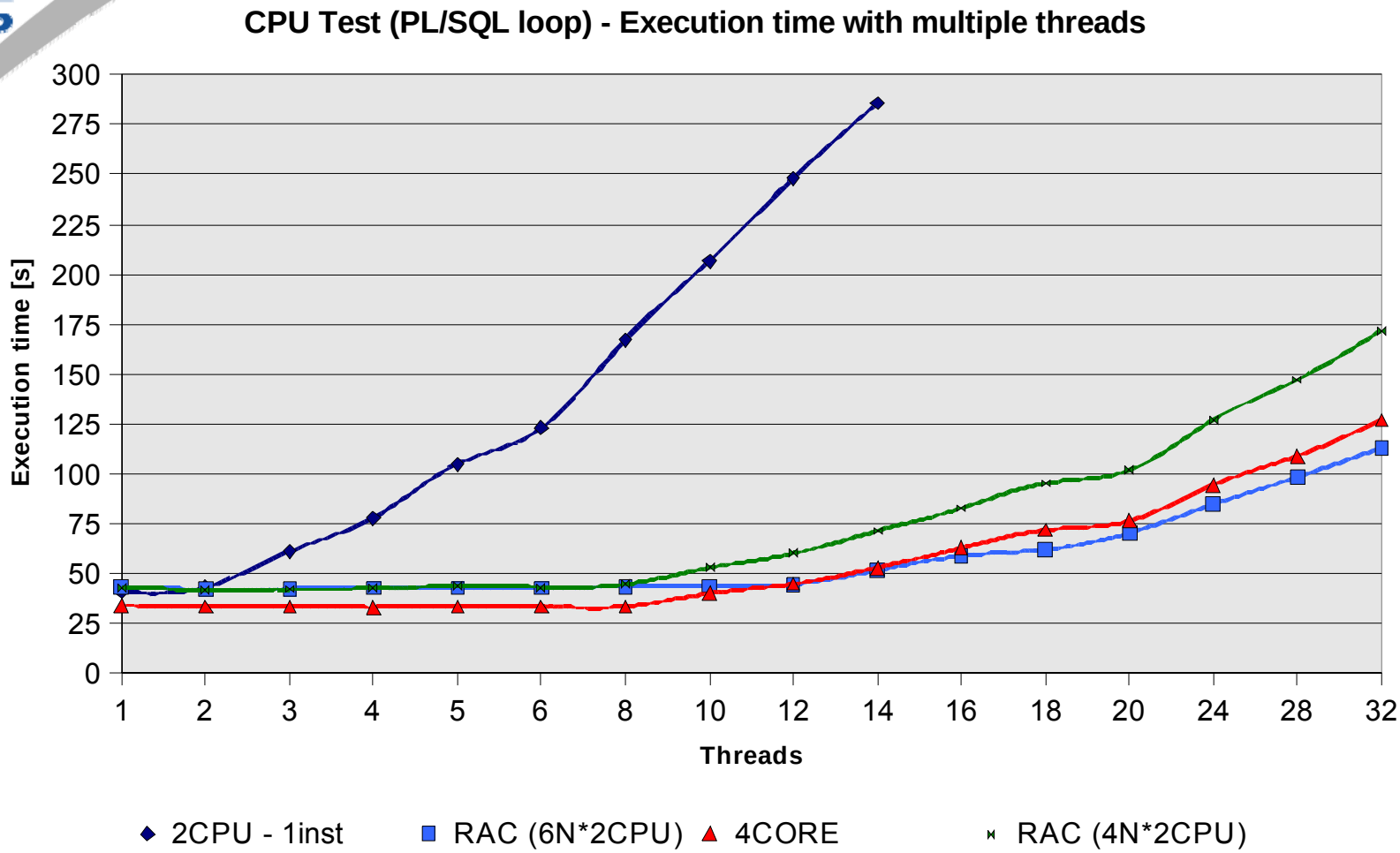
ASM rebalance test



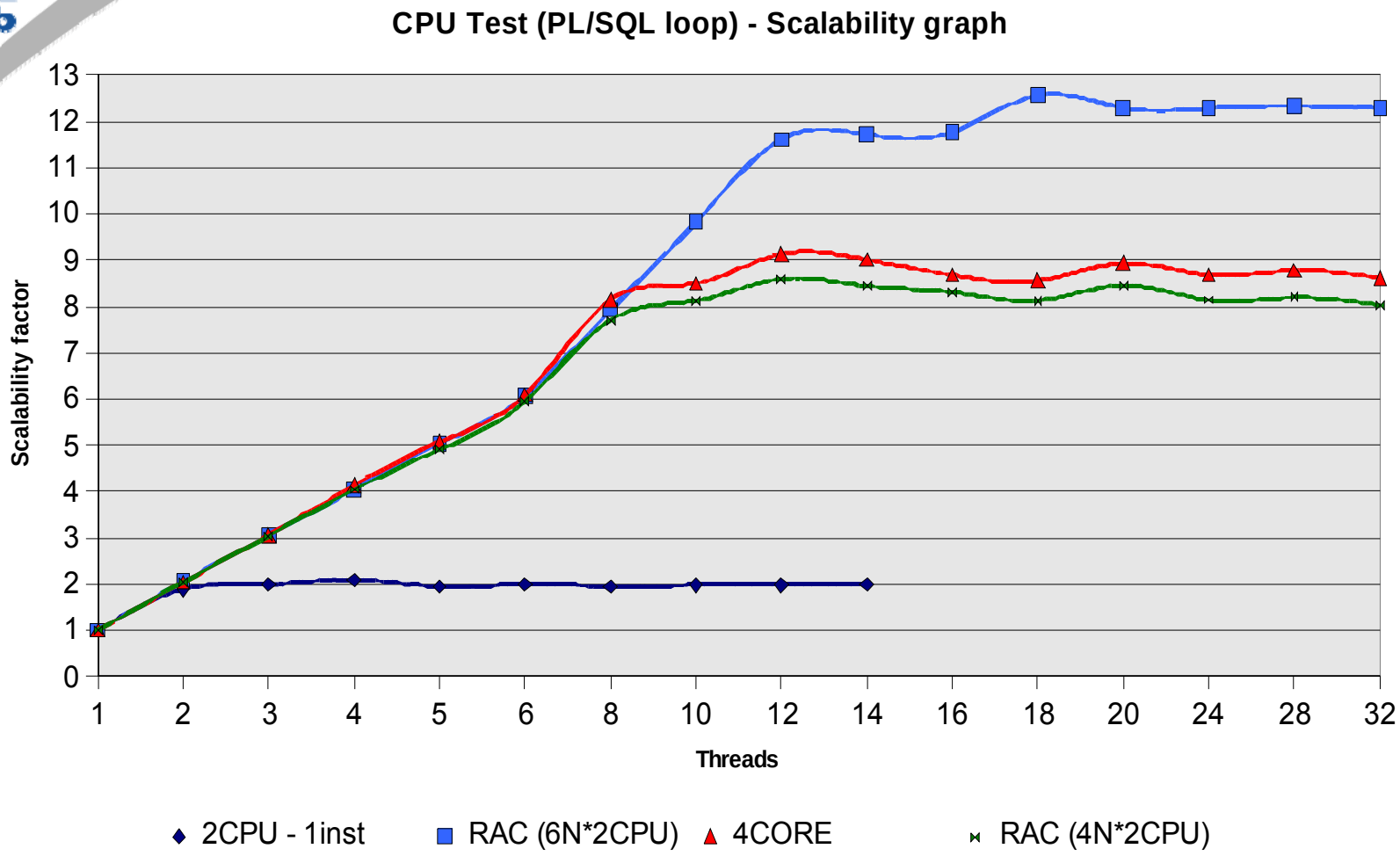
Quadcore tests - setup

- Current setup
  - 2x Pentium IV @ 3GHz, E7520 Memory Controller RAM 4GB DDR2 400MHz (2.5 ns) - memory bandwidth=6.4GB/s
  - 2 x e1000 + Qlogic HBAs 2312
  - installed with RHEL 4 U4 kernel 2.6.9-42.0.3-ELsmp i386 (32 bit)
  - Oracle 10.2.0.3 for i386 with RAC option (6-node RAC) and ASM
  - Average power consumption (per machine): 260W (loaded)
- Quadcore setup
  - 2x Intel quadcore CPUs - Xeon E5345 @ 2.33GHz - L1 cache =128kB, L2 cache=8MB, Intel_5000p Chipset Memory Controller Hub , RAM = 16GB -> 8 x 2GB FB DIMM 667MHz (1.5 ns) - memory bandwidth = 21GB/s
  - 2 x e1000 + Qlogic HBAs 2312
  - installed with RHEL 4 U4 kernel 2.6.9-42.0.8-ELsmp x86_64
  - Oracle 10.2.0.3 for x86_64 with RAC option (1-node RAC) and ASM
  - Average power consumption (per machine): 420W (loaded)
- Test
  - Multithreaded Java application connecting through JDBC

Quadcore tests – CPU results



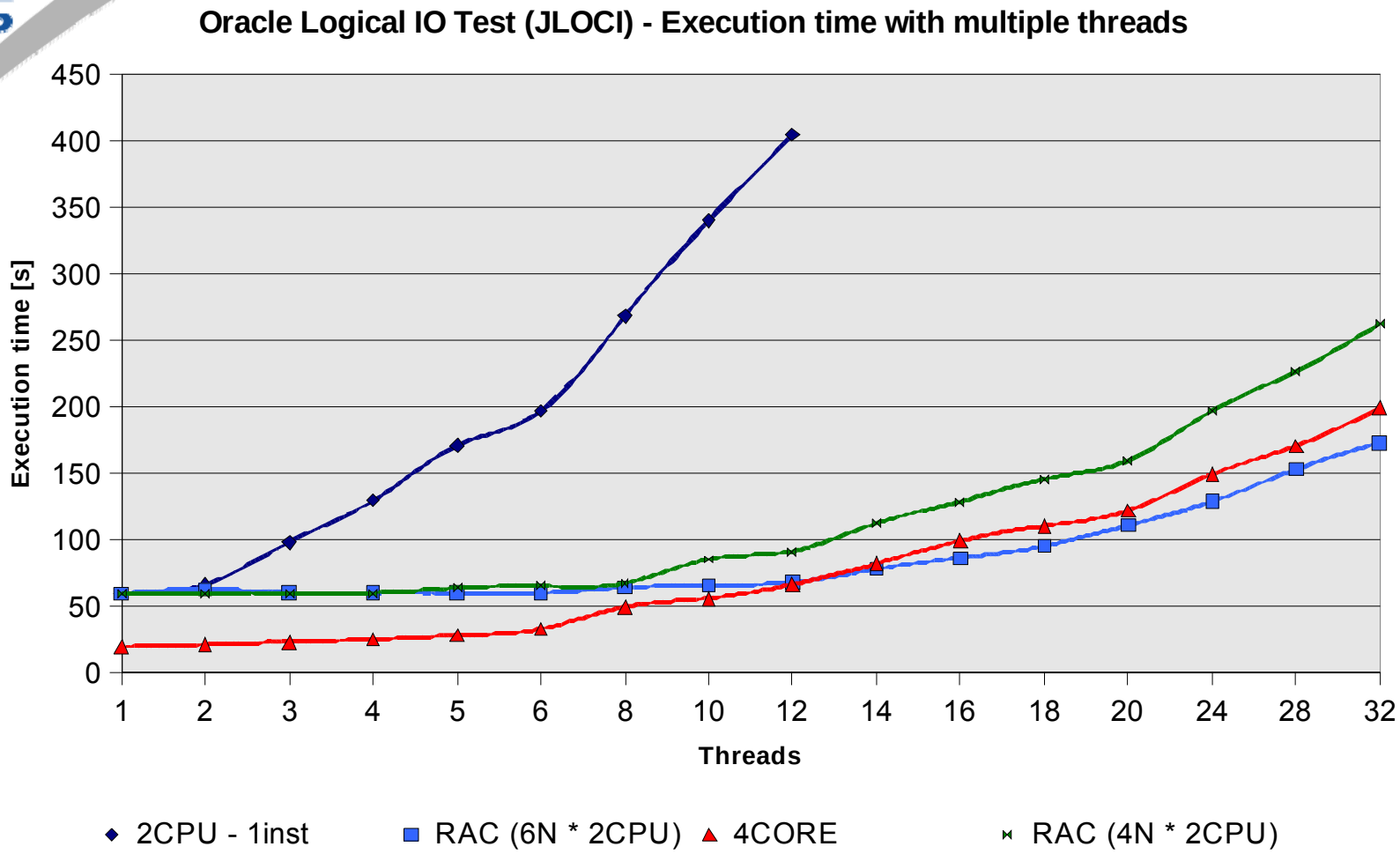
Quadcore tests – CPU results



Quadcore tests – CPU results

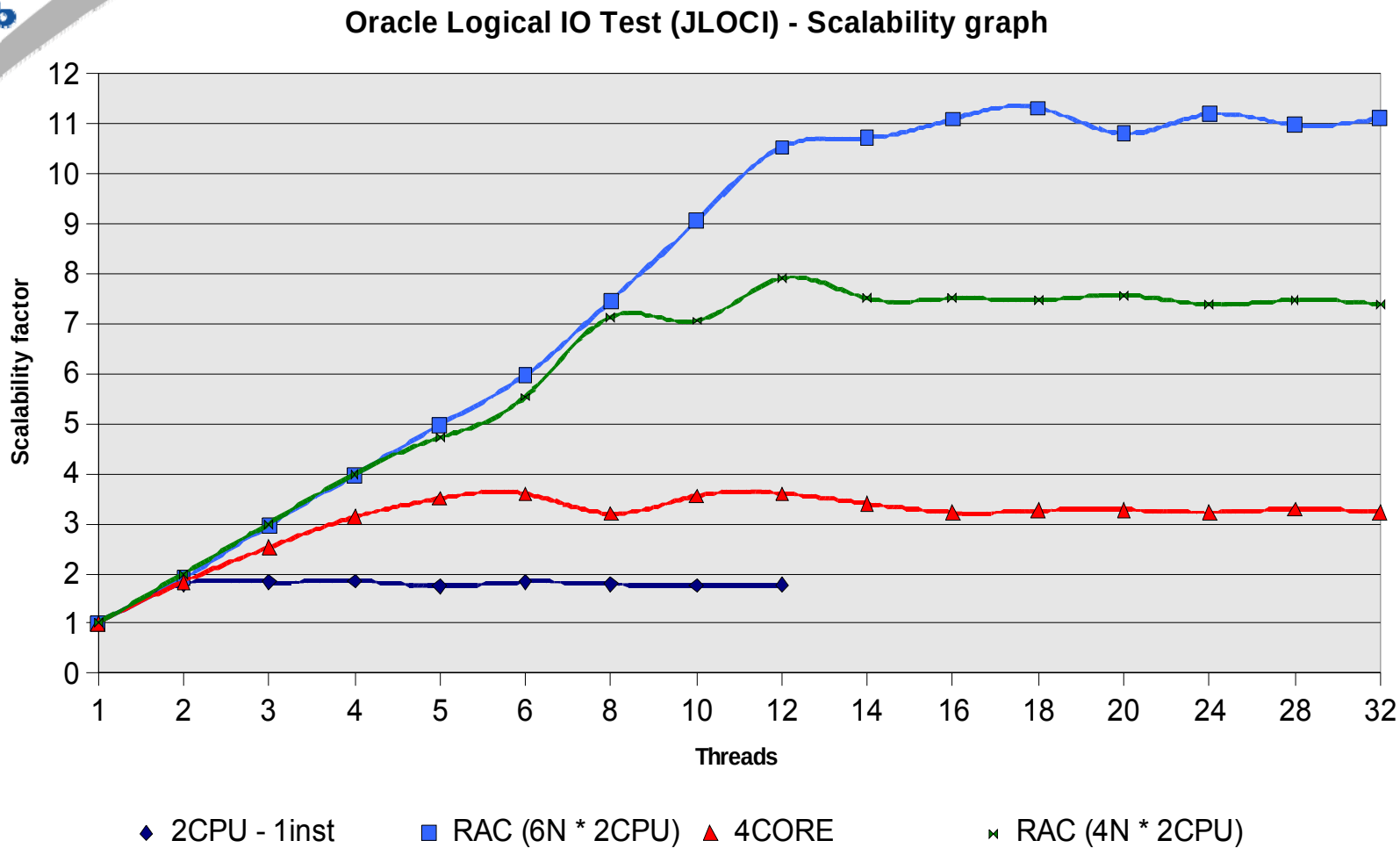
- Observations for CPU-only workload tests
  - Quadcore scales like a 4-node RAC
  - It's faster than 4-node RAC

Quadcore tests – JLOCI results



*JLOCI - Jonathan Lewis Oracle Computing Index

Quadcore tests – JLOCI results



Quadcore tests – JLOCI results

- Observations for JLOCI (logical IO) tests
  - Quadcore scales like a 2-node RAC
  - It's faster than 4-node RAC and only a bit slower than 6-node RAC (for 32 active sessions running)

Programme's Feedback

The feedback is circulated between the people involved.

Oracle EMEA	
Monica Marinucci Lopez	Management of the programmes
June Farmer	Management of the programmes
Graeme Kerr	Technical liaison
Oracle Development	
Bjørn Engsig	Primary Development Contact
CERN Openlab	
Sverre Jarp	Chief Technologist Officer
Juergen Knobloch	**IT-PSS Group leader**
Maria Girone	**IT-PSS-DP Team Leader**

Q & A