



# Storage Review Panel

[Luca.Canali@cern.ch](mailto:Luca.Canali@cern.ch)

~ ~ ~

Distributed Database Operations Workshop  
April 20-21 2009, Barcelona

# Overview

- Motivation for such reviews
- Summary of storage config from previous workshops
- Recent developments at CERN and storage testing
- Round table discussion
  - Tier1 representative input
  - Experiments' input

# Motivation

- Share experience across Tier-0 and Tier-1 on storage technology for Oracle databases
  - Production stories
  - Discuss possible changes before implementation
  - Forum for optimizations and troubleshooting
  - New technology and test activities

# Storage Configuration for 3D Databases

- Oracle RAC and ASM
- SAN based storage
- ASM, and storage redundancy
  - CERN uses ASM for mirroring too (normal redundancy)
  - Most T1 use RAID protection
- CERN on-disk backup
  - Data and recovery areas kept separate
  - On-disk copy kept on recovery area and incrementally updated
- Related to that
  - For some critical databases standby (dataguard) protection is also available

# Storage Performance Testing

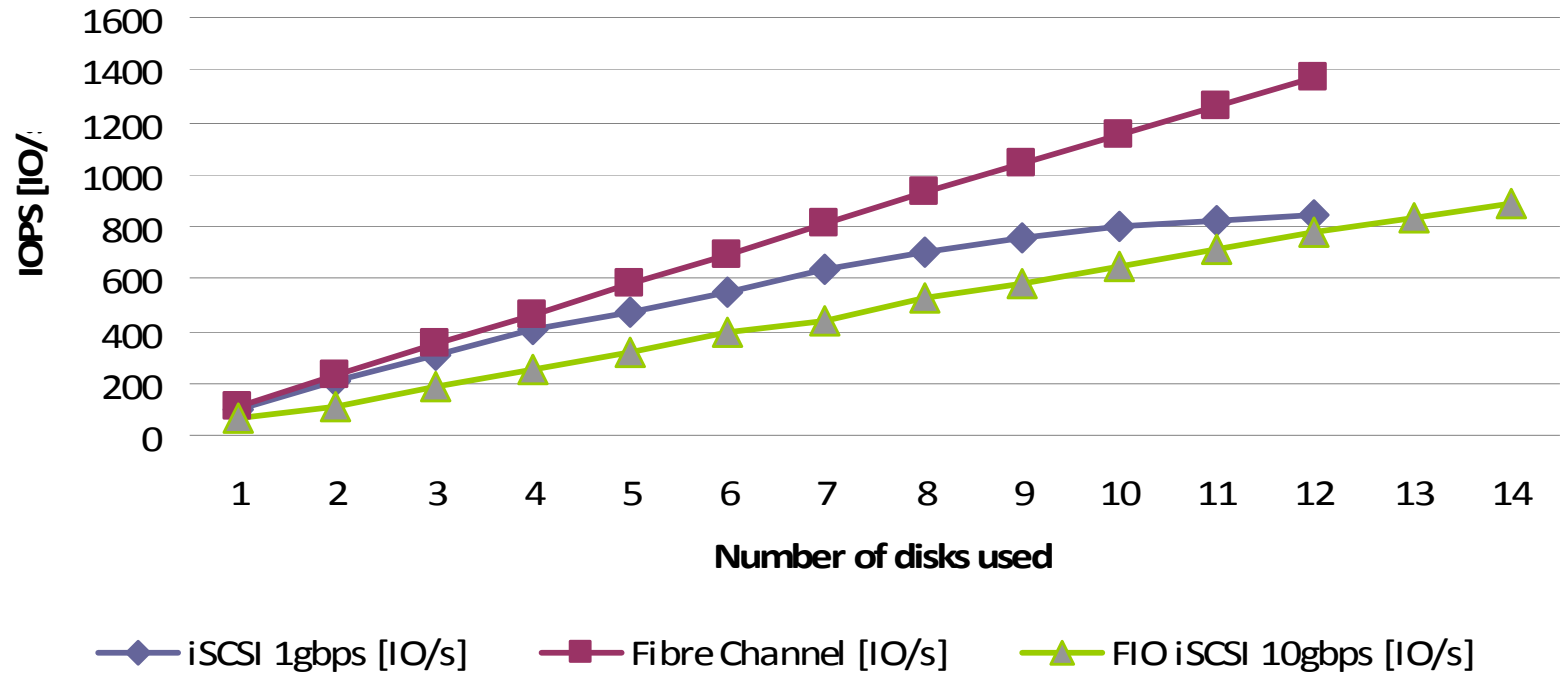
- Oracle Orion is simple but at the same time accurate
  - Used in several tests at CERN and T1s
  - Compared with Oracle workload and demonstrated accurate
  - Easy to run, although ideally should be run before putting HW in prod
  - Wiki page describes how to run Orion tests
- Some of the most important results at CERN:
  - Small random read IO (8k) -> scales linearly with ASM, tested up to 400 disks!
    - 100 IOPS per SATA disk
    - Raptor disks -> 150 IOPS per disk
    - SAS -> 190 IOPS per disk
    - Largest test ~45000 IOPS (300 raptor disks)
  - Sequential read IO, limited by HBA (4+4 Gbps per server)
    - Largest test -> 6GB/s on a cluster (parallel query)

# Recent Tests at CERN

- New features
  - ASM manageability
  - ASM performance
  - Cluster filesystem
- iSCSI vs. Fiber Channel
  - 1Gbps iSCSI
  - 10Gbps iSCSI
  - first results:
    - High CPU utilization
    - Bottlenecks in IOPS and also throughput
    - Client CPU is key
    - More tests are needed to make definitive conclusions

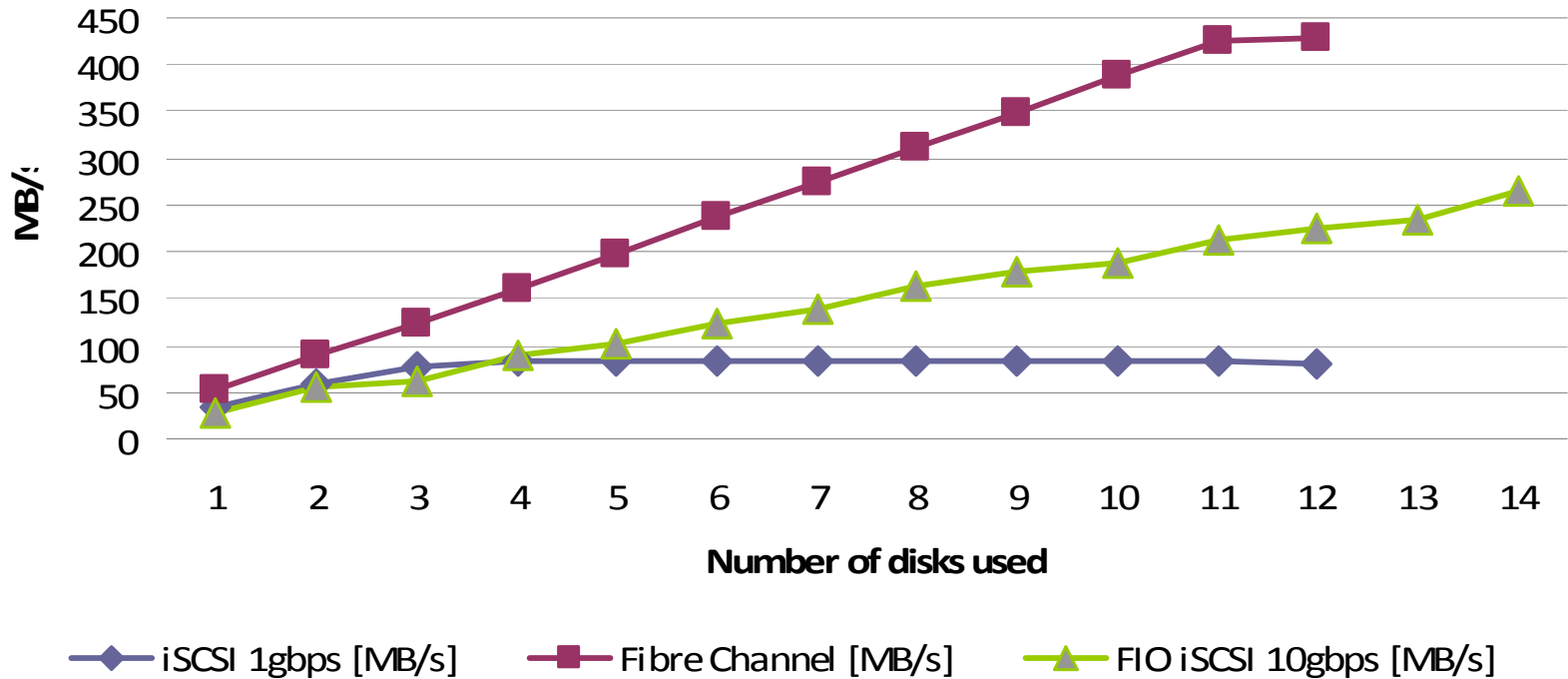
# First Tests on iSCSI

## Small (8k) random IO



# First Tests on iSCSI

## Large (1MB) sequential IO





# Round Table Discussion

- Storage in prod
  - Type, number, ASM setup
- Production stories
- Plans for new HW acquisitions
- Plans for test activities