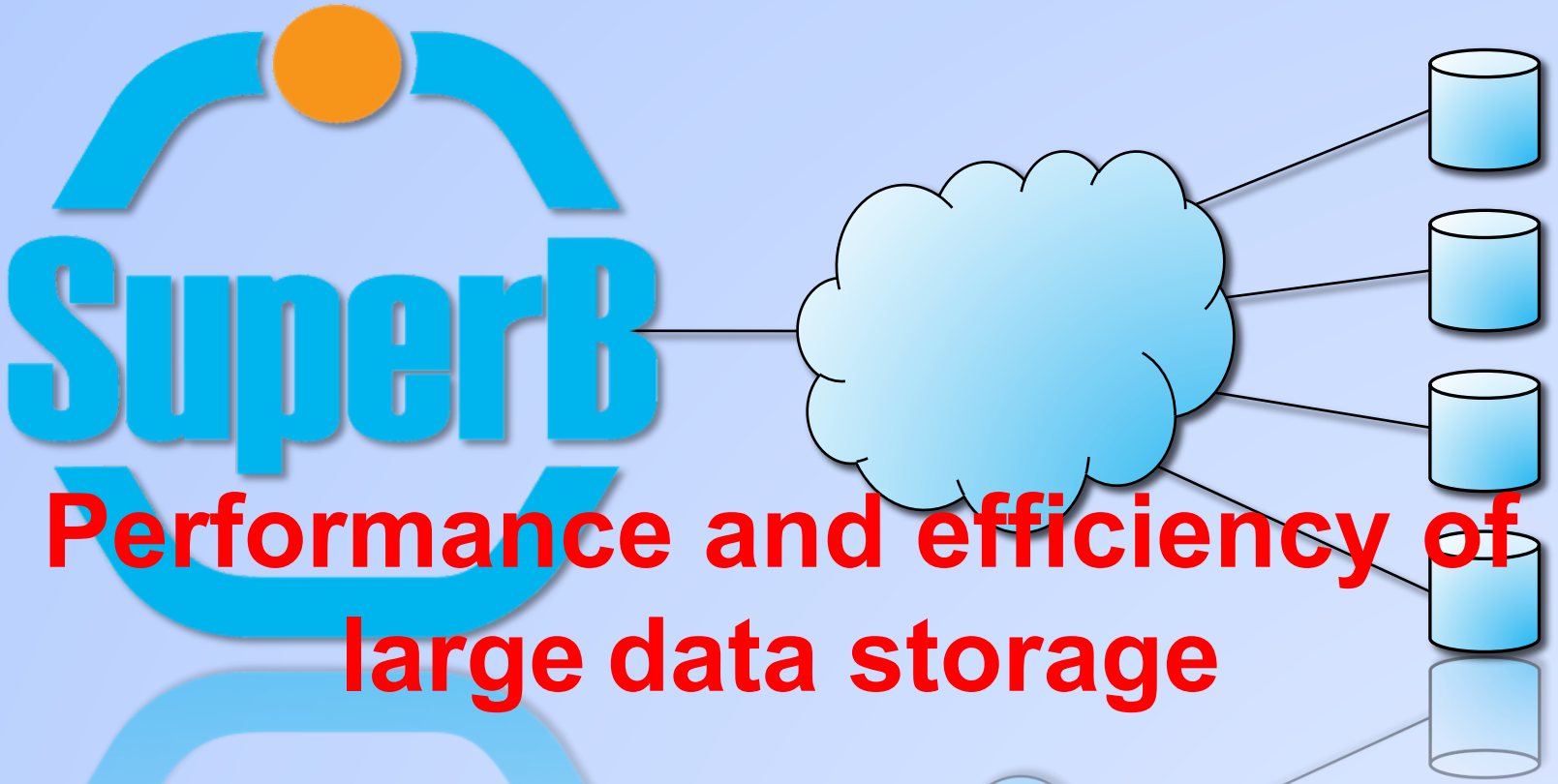
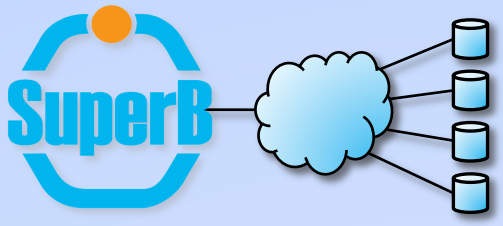


Performance and efficiency of large data storage

Vincenzo Vagnoni (INFN Bologna)
Fabrizio Furano (CERN)

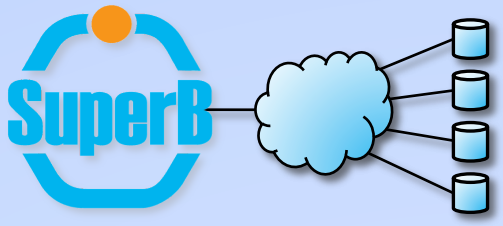
SuperB Computing R&D Workshop
Ferrara, 12th March 2010

Introduction



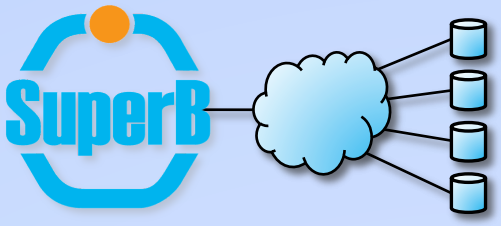
- Efficient data storage and management of large scale systems in a distributed environment are complex tasks
  - Probably the most difficult challenge and weakest point of the existing HEP computing infrastructure
- As an example, in WLCG computing reports, storage problems are always in the first bullet
  - We should learn and avoid repeating the same mistakes

Why complex? Many strongly coupled components



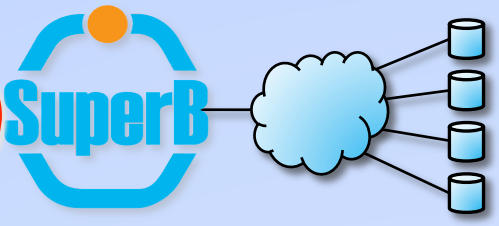
- **Fabric**
  - IB/FC/Ethernet networks
  - Disk controllers and disk arrays
  - Disk servers
  - Tape robots
- **Software/middleware**
  - Low level data format
  - Data management interfaces
  - Data access protocols
  - Tape robot management, HSM software
  - File Transfer services
  - File Catalogues
- **Failures, instabilities or inefficiencies of any component can cause the disruption of the service as a whole**

Common HEP storage models (I)



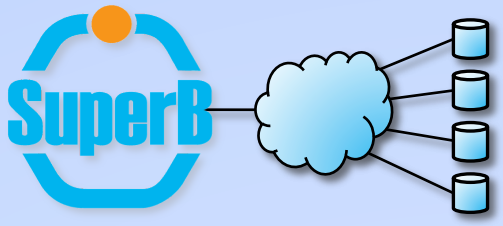
- Data custodial is realized by means of tape based MSS's
- Frequently accessed data files are preferably kept on disk
  - Full HSM approach (aka large-tape/small-disk) is not considered very reliable and efficient, e.g.
    - ATLAS tries to put minimal reliance on tape storage, having the bulk of active storage resources as disks
    - CMS is much more tape-oriented apparently, but *de facto* they work with a disk stage area of the same order of the tape space in size

Common HEP storage models (II)



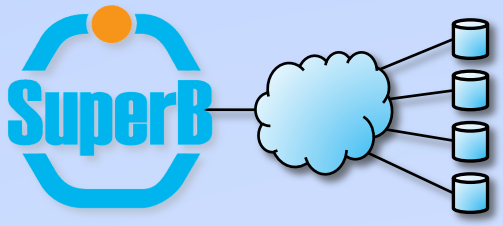
- **Two mainstream models employed for data access from disk**
  - **Disk servers with Directly-Attached-Storage and a redirection mechanism for load balancing**
    - **Products usually developed within the HEP community itself, but paradigm also expanding beyond HEP**
  - **Parallel filesystems with Storage-Area-Networks**
    - **Mainstream products from the commercial world, although some examples also coming academy/research**

The reality today (at LHC)



- Storage systems still critical in many large sites, e.g.
  - Severe instabilities of storage elements
    - Many site experience failures on a daily basis
  - Large data loss taking place too frequently
    - e.g. recently half a PB of data were lost at a Tier-1
  - Loss of sync between File Catalogues and actual contents of storage systems

This session



- **Before the coffee break**
  - Overview of data access in HEP
  - Research, Development and Scientific Application of Gfarm File System
  - Overview of new technologies and evolution of storage systems for handling large data volumes
  - Simulating storage system performance: a useful approach for SuperB?
- **After the break**
  - Discussion
    - What we have learned so far
    - Evolution of storage models
    - Most promising products already available or being developed
    - Identification of new areas where R&D efforts are needed