Large storage infrastructures for high-throughput computing and clouds

Michele Dibenedetto
INFN-CNAF

Roma, 30 Novembre 2012

Outline

- INFN-CNAF Computing Center
- Storage solution
- StoRM SE
- StoRM WebDAV
- From a Cloud perspective
- Summary

INFN-CNAF Computing Center

- CNAF hosts the Italian LHC Tier-1 for the ATLAS, CMS, ALICE and LHCb experiments
- Furthermore, many other experiments uses our resources
  - BaBar and CDF
  - Astro and Space physics
    - VIRGO (Italy), ARGO (Tibet), AMS (Satellite), PAMELA (Satellite) and MAGIC (Canary Islands)
  - And more (e.g. Icarus, Borexino, Gerda etc…) for a total of about 20 experiments

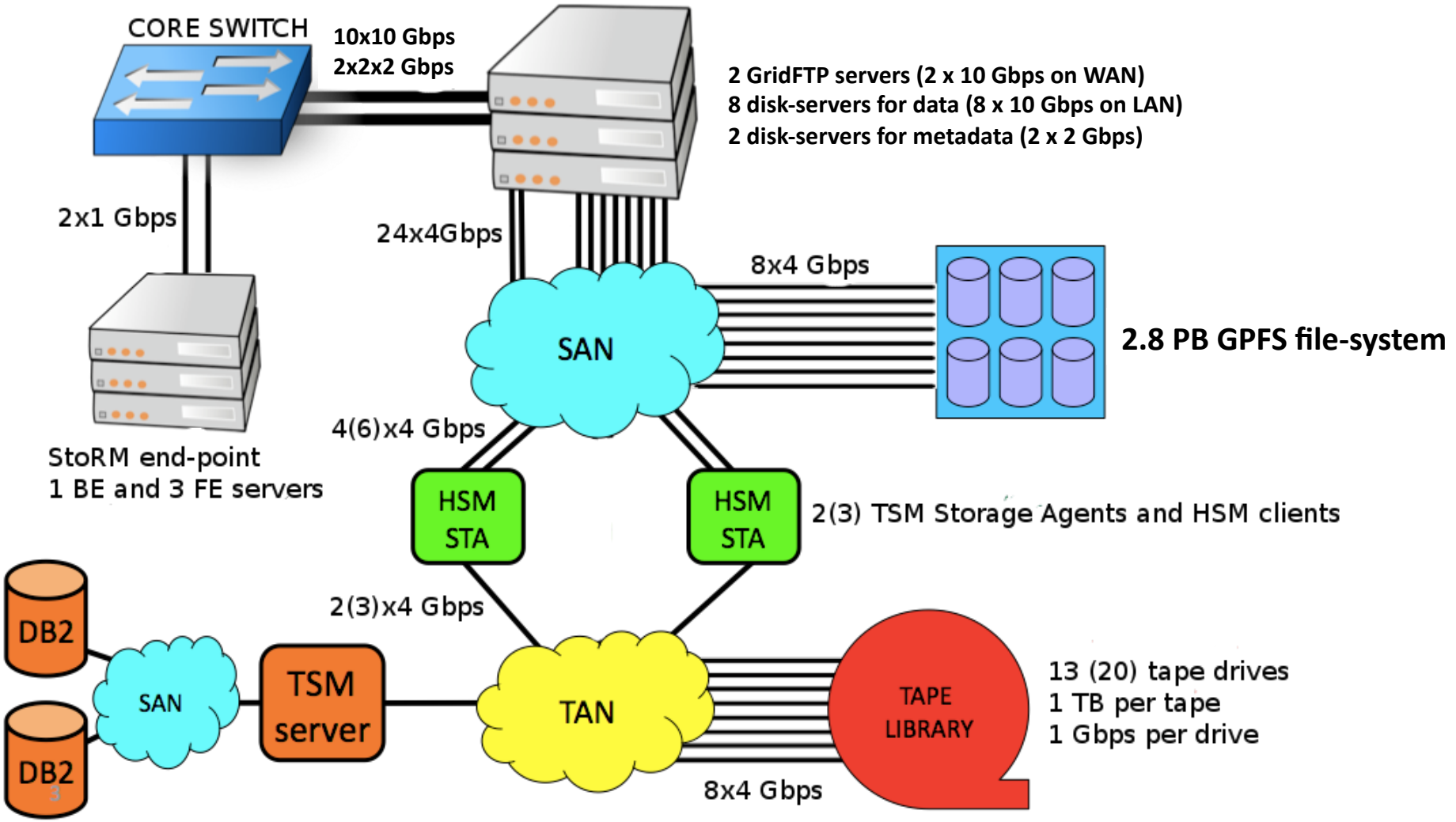
INFN-CNAF Computing Center

- **1000 m^2** with more than 120 racks
  - Redundant facility to provide 24x7 availability
- 1300 servers with about 13000 cores available
  - up to 100000 jobs/day
- **11 PBytes of disk space**
  - Raid 5/6 systems
  - ~160 servers interconnected in a SAN (4Gb/s or 8Gb/s FC links)
  - Aggregate bandwidth to storage: ~ 50 GB/s
- Tape library **15 PB**
  - 8800 x 1 TB tape capacity, ~100 MB/s of bandwidth for each drive (20 tape drives)
  - 1200 x 5 TB tape capacity, ~200 MB/s of bandwidth for each drive (10 tape drives)

Mass storage solution

- All disk space is partitioned into several GPFS clusters
  - One cluster per each main experiment
  - GPFS deployed on the SAN implements a full HA system
  - System scalable to tens of PBs and able to serve thousands of concurrent processes
- Tape library managed by TSM
- Integrating GPFS with TSM we have a complete HSM solution
- Access to storage granted through standard interfaces (SRM and WebDAV)

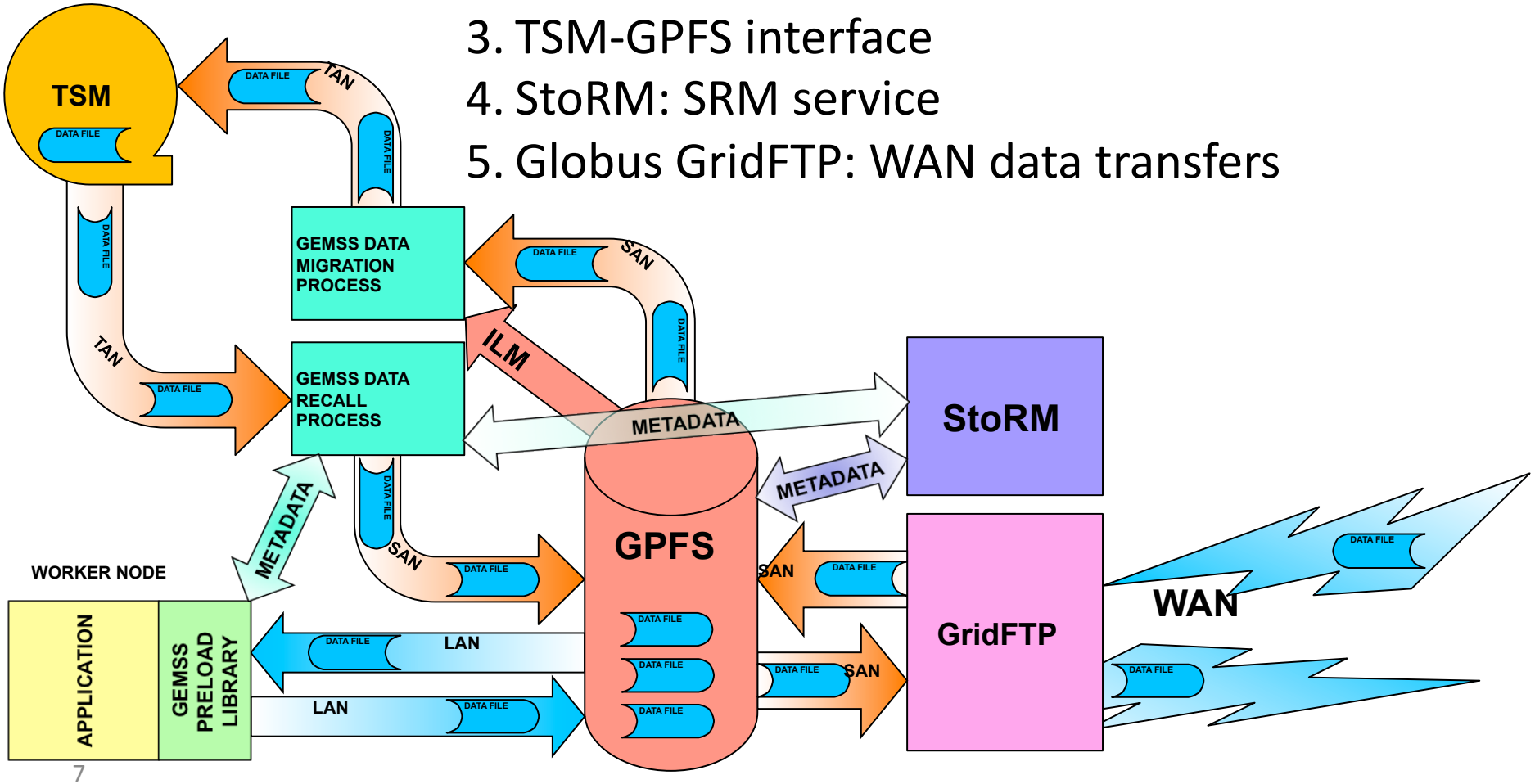
Typical storage layout for CNAF experiments



Building blocks of the storage system

Disk-centric system with five building blocks

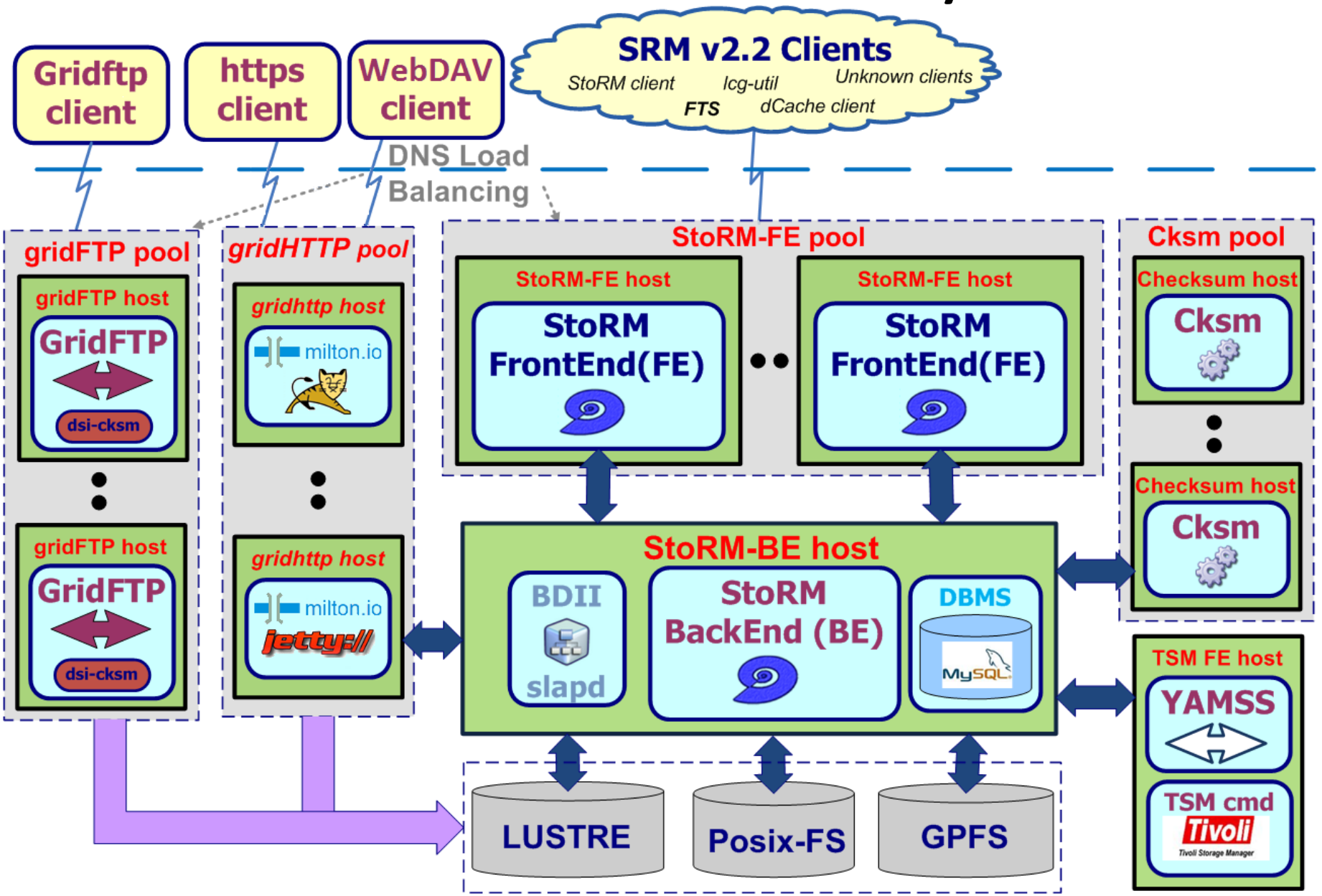
1. GPFS: disk-storage software infrastructure
2. TSM: tape management system
3. TSM-GPFS interface
4. StoRM: SRM service
5. Globus GridFTP: WAN data transfers



The StoRM service

- StoRM is a middleware component
  - Provides a commodity interface to manage storage space and user files
  - Hides the specific implementation of the storage system
- Exposes a standard SRM 2.2 interface
- Provides fine grained authorization mechanisms
- Scales from small single host installations to complex distributed scenarios
- In production since 2007 at CNAF
- Currently deployed to more than 50 data centers world-wide

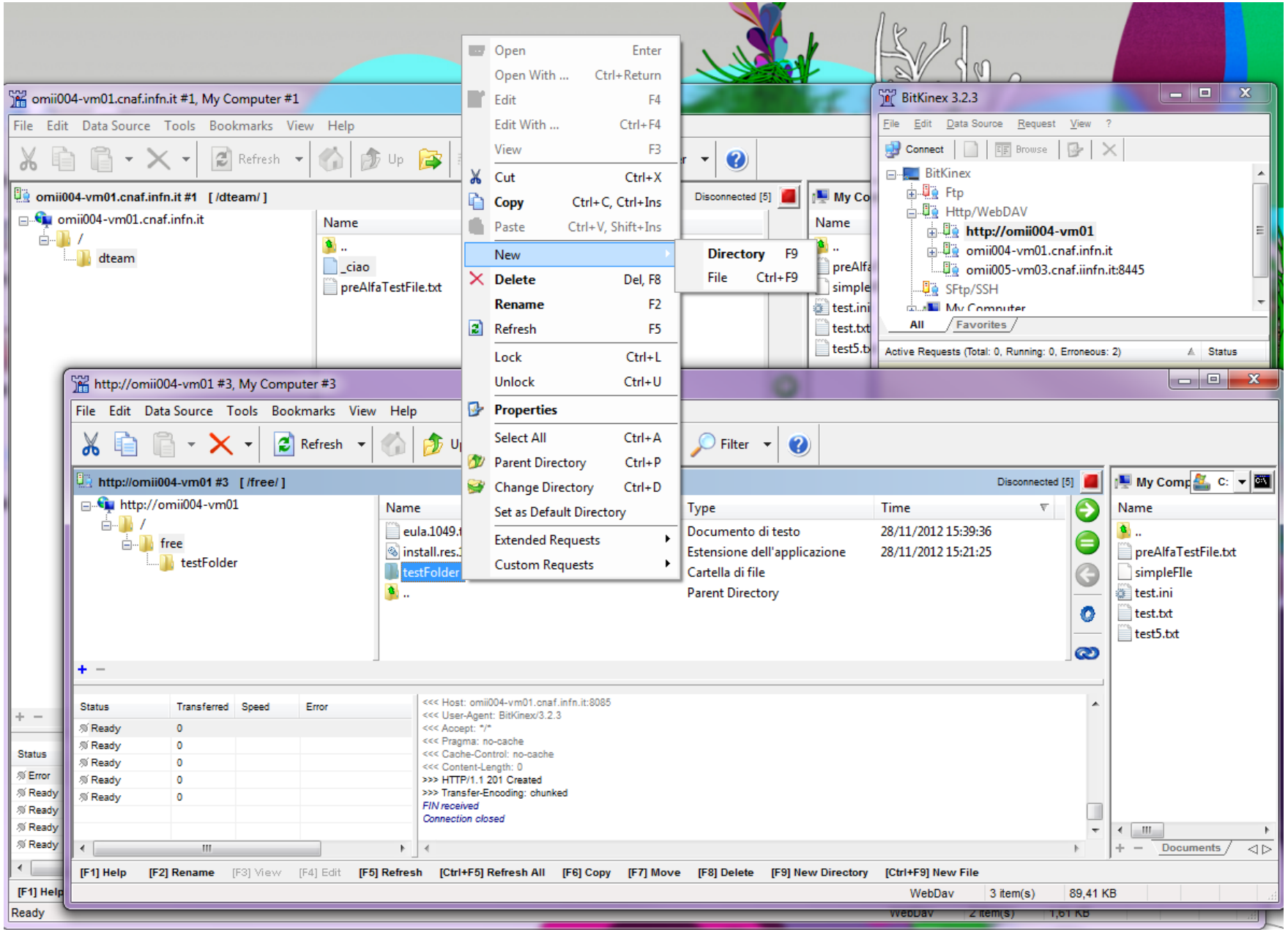
StoRM installation layout



StoRM WebDAV Interface

- New development driven in opening toward new user communities and their needs
- WebDAV
  - Standard world-wide adopted file management protocol
  - Extension of the standard HTTP protocol
  - Gives users the chance to virtually mount remote storage
  - Many user-friendly clients available
  - Even accessible by a web browser

WebDAV Client



StoRM WebDAV Interface: our implementation

- Hides SRM complexity
- Allows user to identify themselves with
  - Nothing (anonymous access)
  - x509 certificate (corporate access)
  - x509 proxy certificate (corporate enhanced access)
  - x509 VOMS proxy certificate (GRID access)
- Represents a single entry point to the storage data (both file management and transfer)

StoRM WebDAV Interface: federating storage

- The grid community is interested in federating its storage systems
- A Grid WebDAV federation solution is ready
  - A single unified files namespace is exposed
  - File replicas can be spread within the federation
  - Users asking for a file will be always redirected to the most convenient replica
  - The federation is dynamic and fault tolerant
  - Any WebDAV server can join the federation
    - Grid software providers
    - Commercial solutions

A Cloud Storage perspective

- The present CNAF infrastructure can provide large scale cloud storage resources by using StoRM WebDAV as entry point
- The main features are:
  - High network bandwidth
  - Low file system latency
  - A huge amount of available space
  - The reliability of tape systems and RAID 6 solutions
  - The simplicity of WebDAV clients

A Cloud Storage perspective

- Open to new communities
- Each community has its own file logical namespace
    - Can be private or open
- Full authentication/authorization policy specification capability
- Limited administration efforts
- Space monitoring
- User blacklisting capability

Summary

- CNAF has developed a highly available, fault tolerant, high throughput mass storage solution
  - Primarily targeted to fulfill LHC experiments needs
- Can be easily extended to support other communities
  - Native WebDAV interface as cloud storage

Questions?