

# Xrootd status and ongoing/future work

## (Status report)

Pavel Jakl

**S&C meeting**

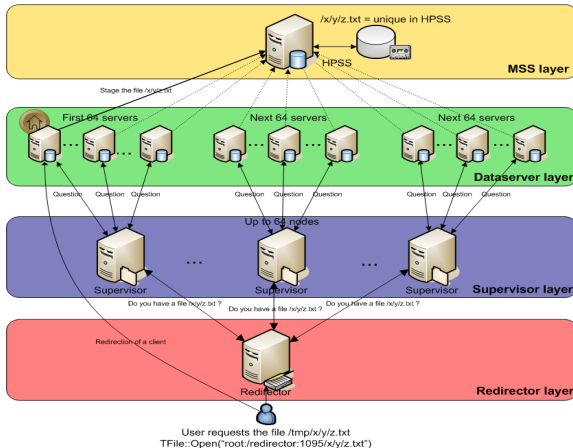
7th of June 2006

# Outline

- Enabling HPSS plugin with DataCarousel
- Understanding load balancing
- Automatic purging of used space
- Additional fault-tolerance to xrootd cluster

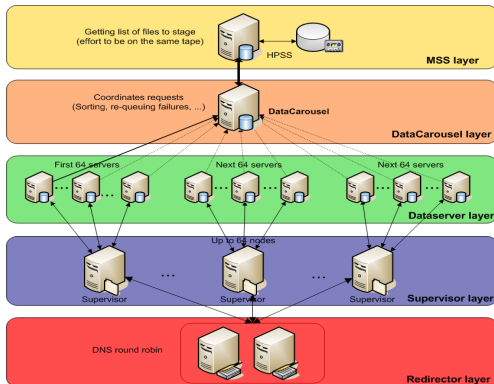
# Reminder of XROOTD architecture

## XROOTD architecture interfering with user's request



# DataCarousel integration with XROOTD

- DataCarousel resolved un-coordinated requests to HPSS
  - queuing of user requests
  - sorting algorithm for effort having many files on a same tape
  - re-queuing in case of stage failure



# Understanding load balancing

- load balancing is used to select best node for serving user request
- two possibilities:
  - 1 file is already on distributed disk and has multiple copies
  - 2 file doesn't have copy and need to be staged from HPSS
- load of server is reported by external script giving information about cpu,memory,network usage etc.
- overall load is computed by combination of these information
- we needed to understand the load on CAS machines for best set up of these thresholds
- <http://www.star.bnl.gov/~pjakl/xrootdStat>

# Automatic purging of used space

- effort to have automatic purging in absence of free space
- node selection is also based on available space (second cycle of selection)
- purging is controlled by 2 thresholds:
  - 1 node is selected when it has minimum of free space
  - 2 node is purged when **selected** and hasn't limit of free space to put a file
- both of these thresholds need to be considered when space is shared by other users

# Additional fault-tolerance to xrootd cluster

- redirector node is main point of failure in xrootd architecture
- problem could be resolved by **DNS round robin** which has 2 nodes configured as redirector
- xrootd had problem to build the tree described above
- problem reported to xrootd development team
- in meanwhile redirector node can serve just on LSF batch job to reduce a load
- possible failure of supervisor's node was resolved by one additional redundant supervisor node

# Production cluster

- xrootd cluster under bigger load of users didn't respond on specific time
  - time was extended, but it waste of CE
  - within this time xrootd redirector is waiting for responses from dataservers
  - improvement could be immediate answer to client when redirector got response from server which has requested file
- **XROOTD Maintenance**
  - **When ?**: Today 7 PM
  - **What does it mean for me ?**: all jobs will be suspended