Using TSM to create a high-performance tape connection

Forschungszentrum Karlsruhe
GmbH
Institute for Scientific Computing
P.O. Box 3640
D-76021 Karlsruhe, Germany

Dr. Doris Ressmann
Dr. Silke Halstenberg
Jos van Wezel

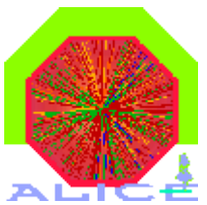
http://www.gridka.de

Introduction

- Forschungszentrum Karlsruhe and GridKa
- dCache setup
- the easiest tape connection
- improvements
- a high-performance tape connection
- conclusion

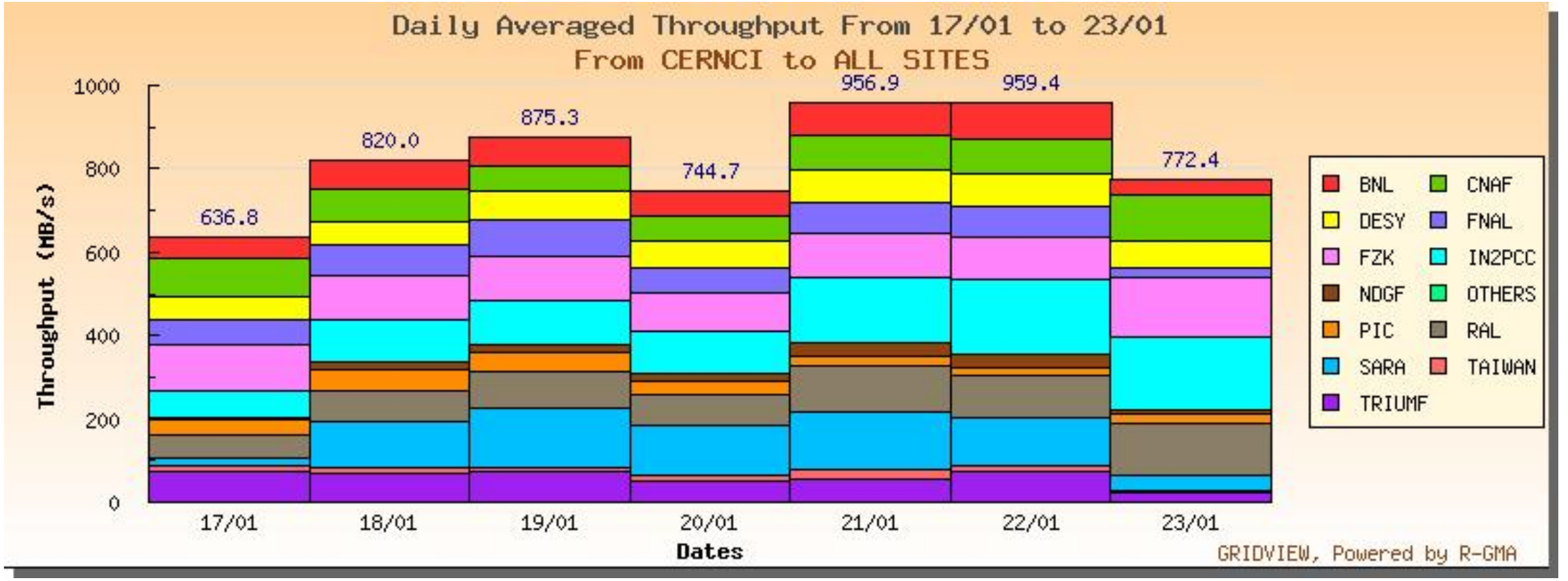
GridKa

- Forschungszentrum Karlsruhe
  - Research centre with 3800 employees
  - Program: Energy, health, nano-microsystems, earth and environment, structure of matter
  - GridKa is part of “structure of matter“
- Grid Computing Centre Karlsruhe (GridKa)
  - Tier1 centre for LHC experiments



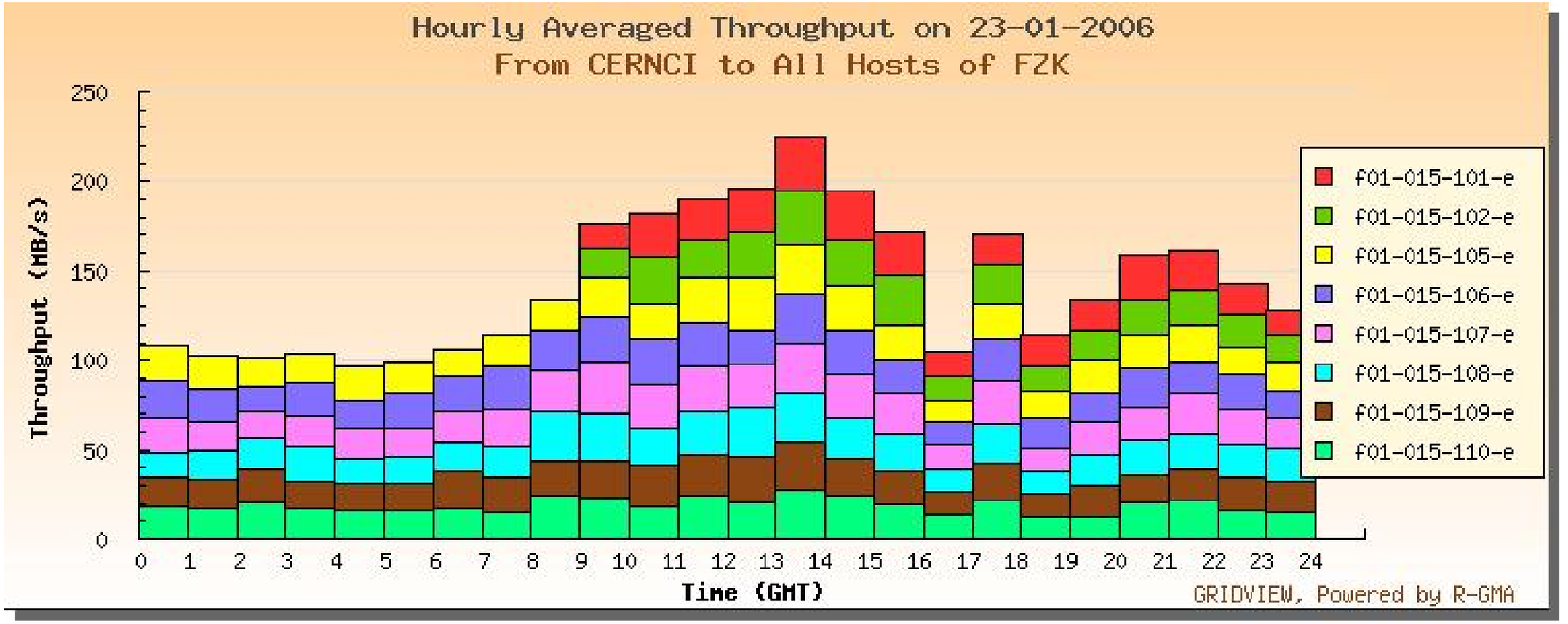
  - Production for further experiments

Service Challenge Throughput Phase

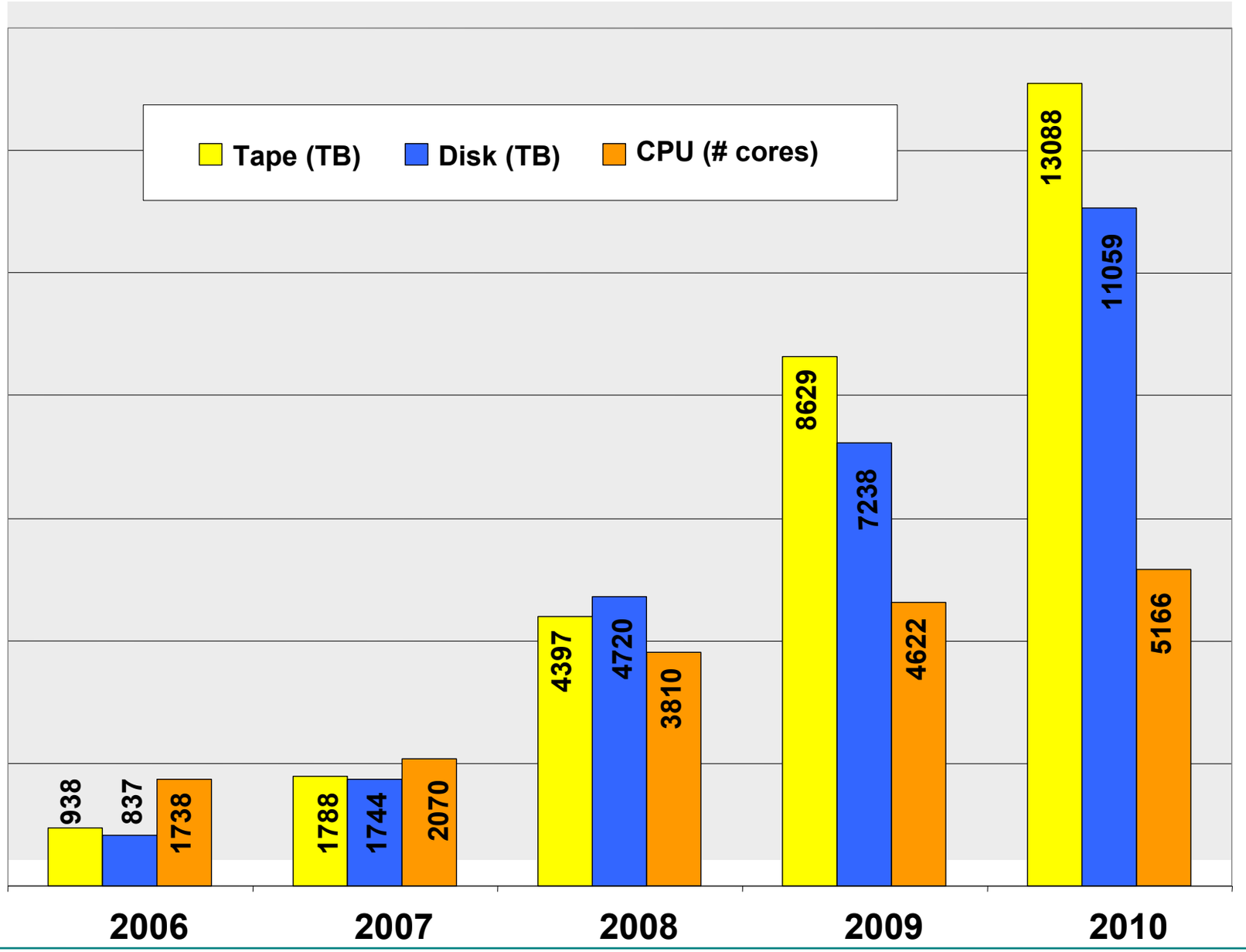


Our GridFtp doors

8 gridFtp doors but only 4 pool nodes

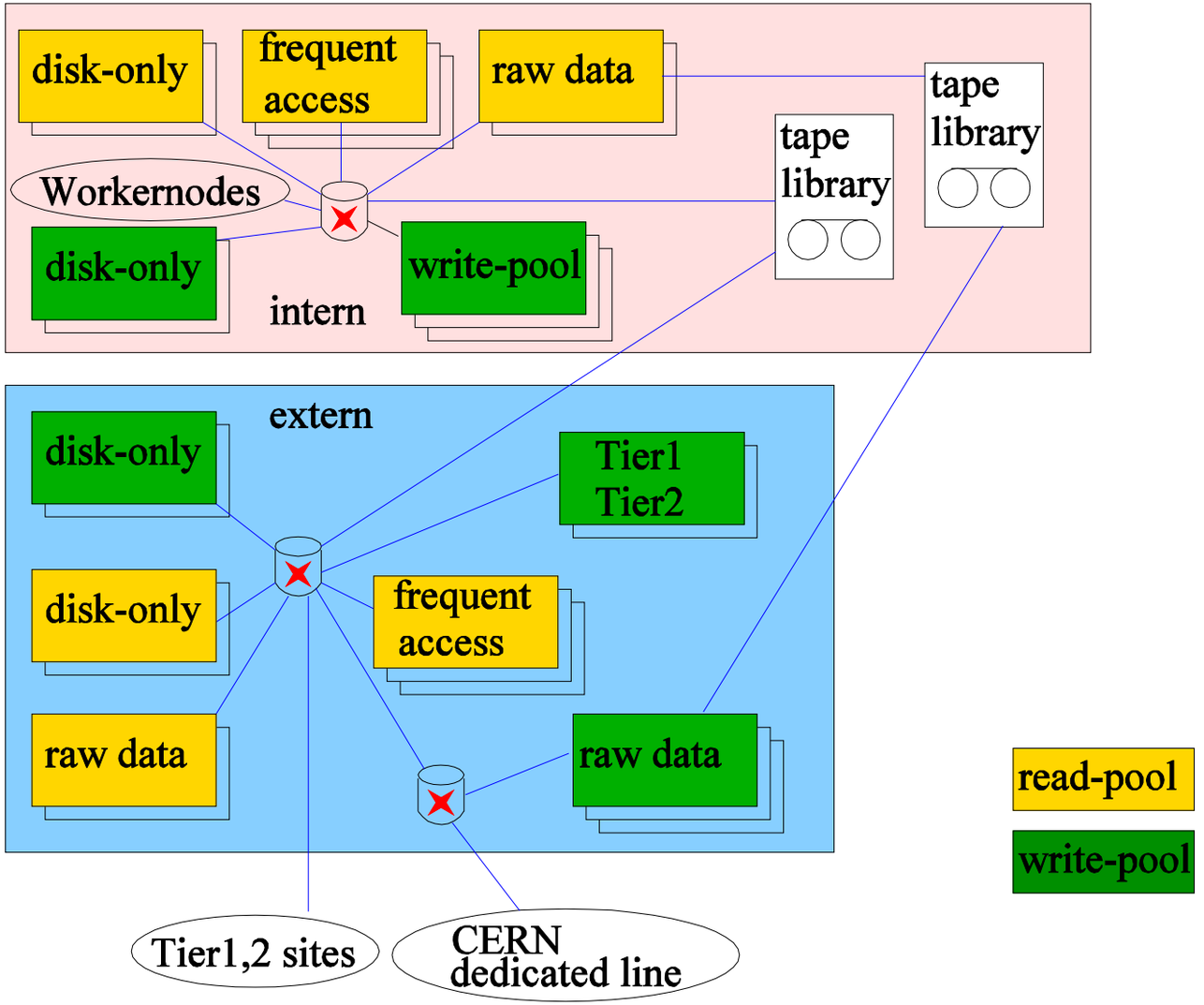


Planning numbers

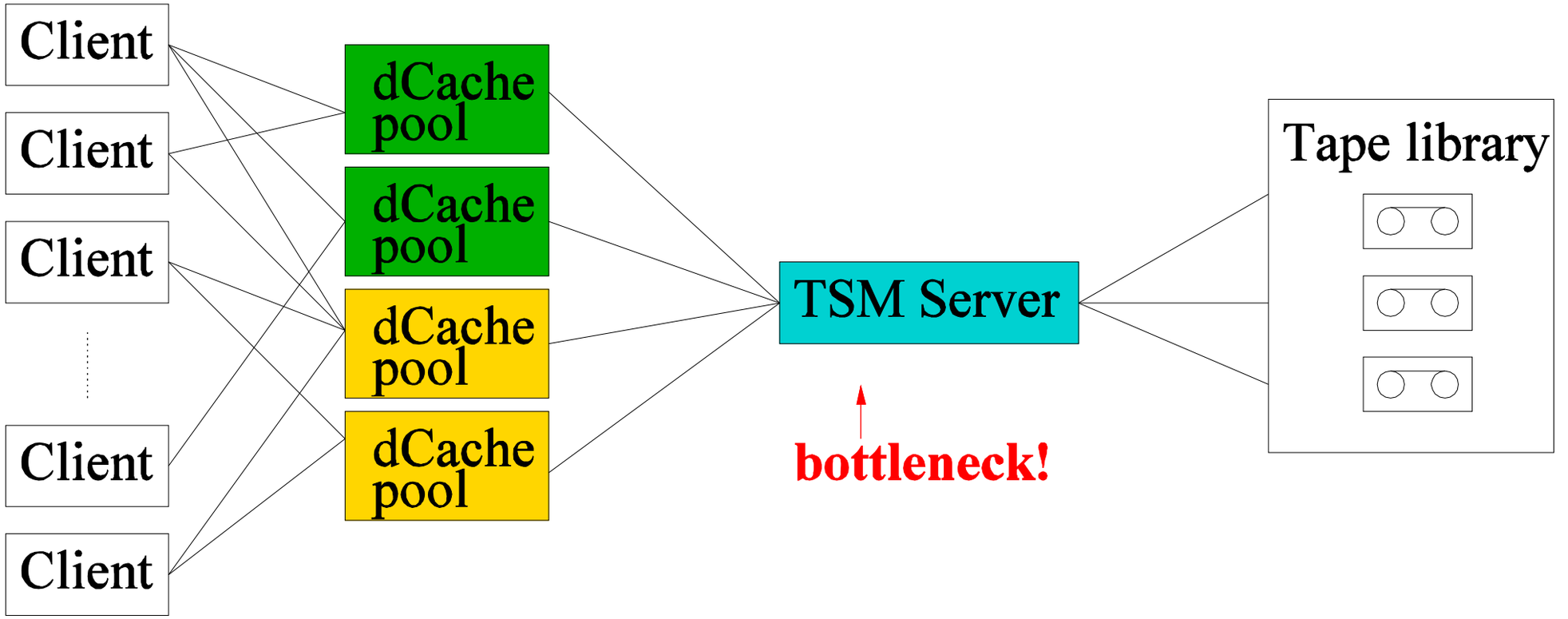


- 10% of total LHC
- Includes non-LHC
- CPU (2006) = 1300 kSI2K

dCache Setup



A Simple Tape Connection



A solution for Tier2 sites ~100MB/s

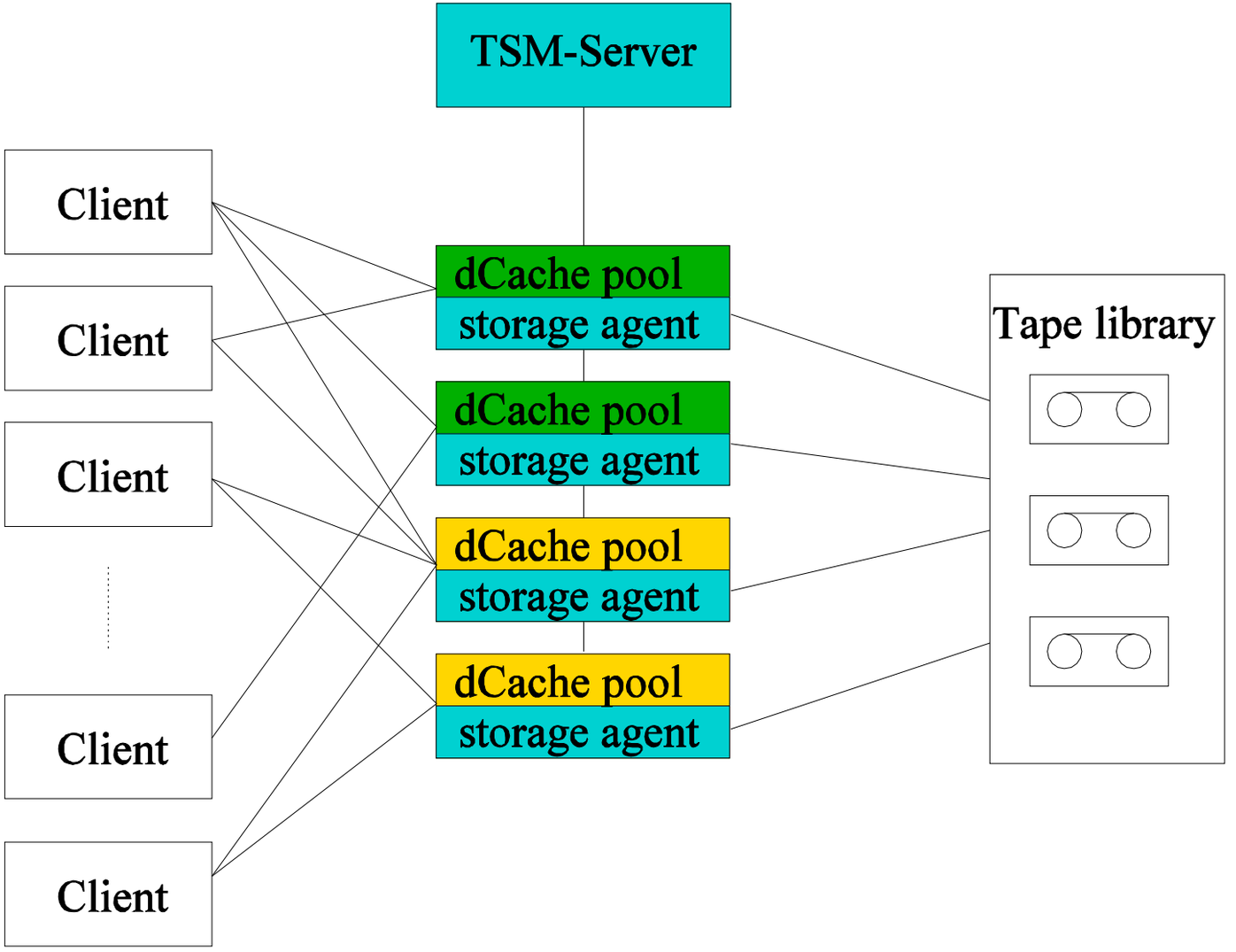
Storage Agent

- TSM Server has database about tape information
  - location
  - utilisation
- Storage Agent is a minimised TSM-Server without its own database
  - connected via Storage Area Network (SAN) to the tape library (can directly use tapes)
  - the control connection to the TSM-Server is via LAN

Proxy Node

- dCache does its own load balancing
  - files are stored from one pool and retrieved from another pool
- all nodes have the same node name within TSM
  - creates lots of info messages
  - it works only if the TSM-server does the tape connection
- same node name does not work with LAN-free storage agents
  - proxy node is common name for all nodes

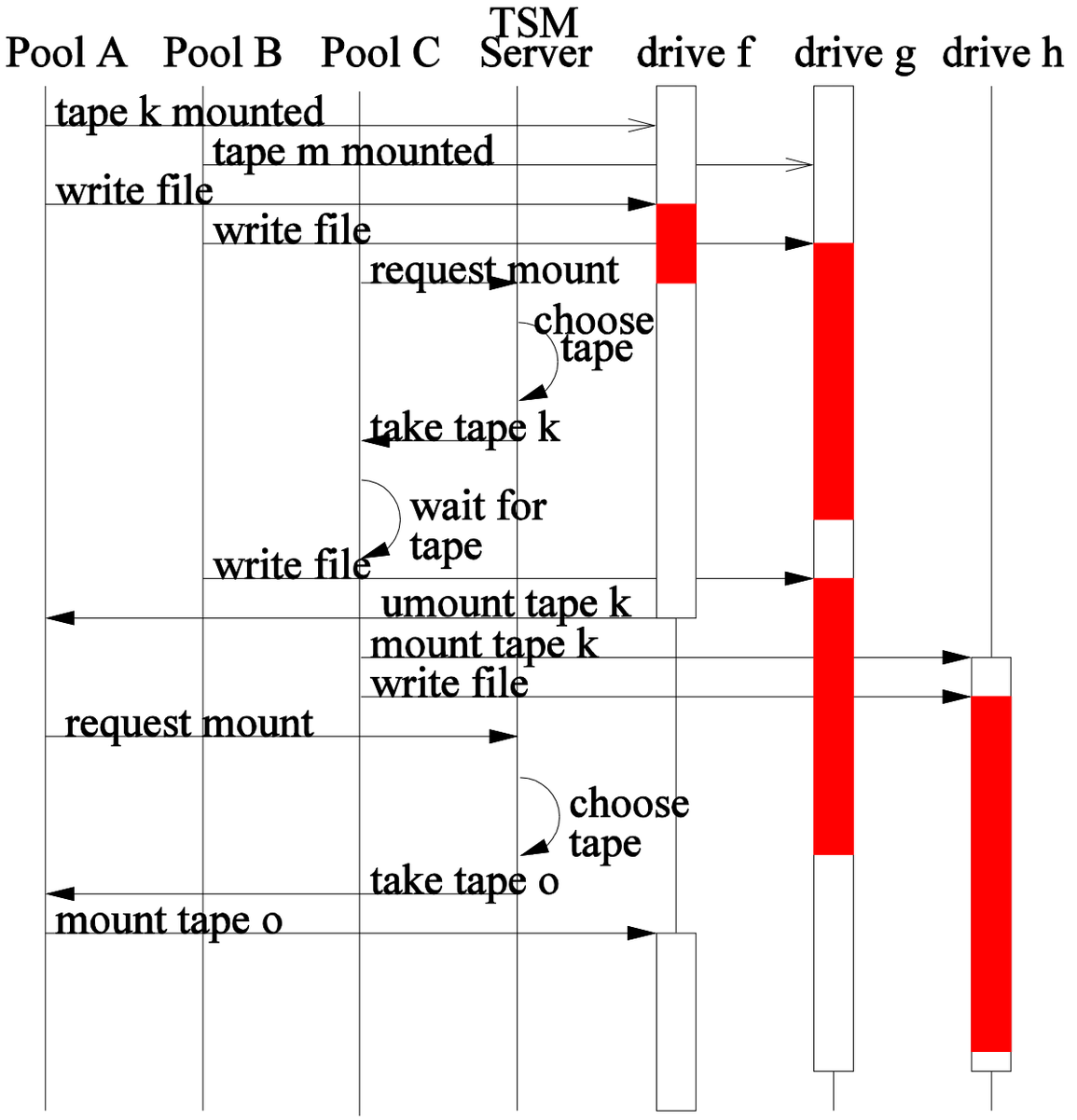
An Idea


TSM-Server
Client
Client
Client
Client
Client
dCache pool
storage agent
dCache pool
storage agent
dCache pool
storage agent
dCache pool
storage agent
Tape library

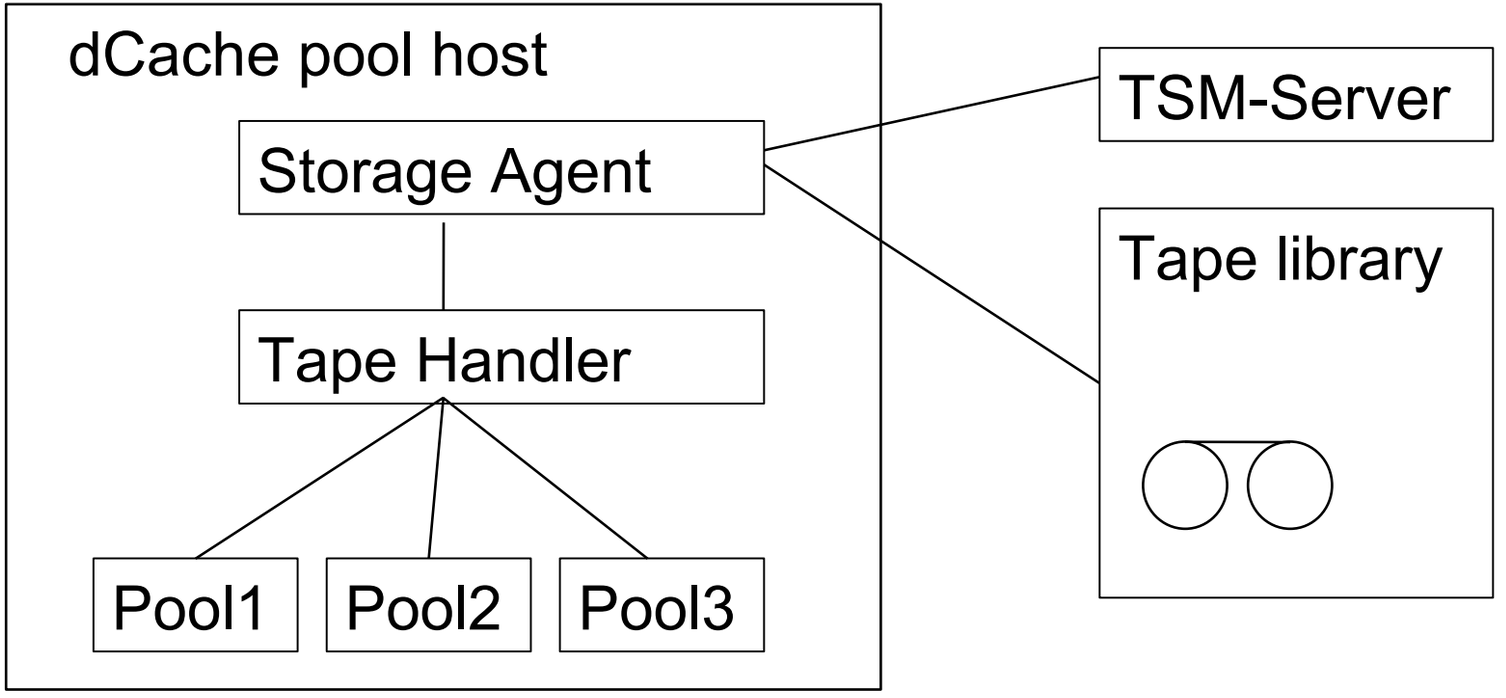
Archive Functionality

- a “normal” tsm archive
  - information lost if interruption occurs
  - no version control
- dCache creates a new session for every file
  - actual status of written files
  - TSM-server -> tape connection
    - same tape is used without dismounting
  - storage agents -> tape connection
    - frequent tape dismount

TSM With Storage Agents



A Solution



tape handler keeps connection open

- dismounting is minimized

Conclusion

- TSM Server max throughput is 100MB/s
  - 3 drives (LTO2) a 35MB/s = 105 MB/s
- Storage Agent max throughput is 100MB/s
  - without tape handler
    - >> # storage agents >> tape mount actions
    - >> # storage agents << throughput
  - tape handler
    - max 1 or 2 drives connected
    - max used is 70 MB/s
    - >> # storage agents >> throughput

Poster Session

- Dr. Andreas Heiss: Connecting WLCG Tier-2 Centers to GridKa (today)
- Dr. Sven Hermann: Operating a Tier1 centre as part of a grid environment (today)
- Bruno Hoeft: LHC-OPN Network at GridKa (Wednesday)