



CASTOR

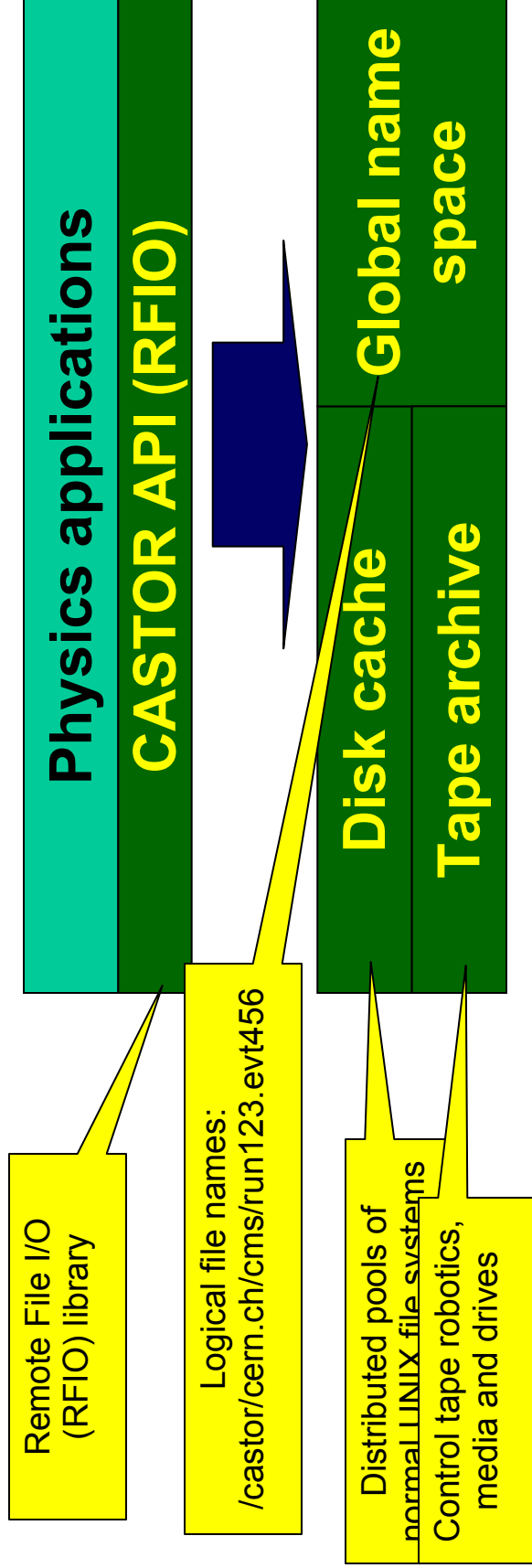
Storage and data management technologies

28-30/6/2004
Olof Barring, CERN



Background

- **CASTOR**
 - stands for “Cern Advanced STORage mgmt system”
 - Hierarchical Storage Management (HSM) system written in C (~260kSLOCs)
 - In production since 2001. Manages all physics data at CERN
 - **Currently: 20 million files, 2.4PB data**
 - Other CASTOR sites: CNAF (Bologna), PIC (Barcelona) and several Russian institutes. Being evaluated by RAL





Challenges for LHC



- Tier-0 (physics production)
 - 4GB/sec peak rate
 - 10PB/year
- Tier-1 (physics analysis)
 - 4PB disk cache
 - 1000s of concurrent clients (2-3 jobs per CPU worker node)
 - High request rates (request = file open), peaks of several 100 requests/sec

• New castor design applies established concepts from CPU cluster mgmt to disk resource mgmt

- (disk) storage resource sharing
- request scheduling
- Policy based resource allocation



CASTOR project



- Expertise in low level tape technology (drives, media, robotics)
- Currently collaborating with
 - Platform computing Inc (Canada): adapt LSF scheduler for disk resource management
 - Maui project (US): adapt their open source scheduler for disk resource management
 - GGF GSM (Grid Storage Management): Grid storage interfaces
- Operation support structure established:
 - CASTOR customers provide CERN resident development efforts in exchange for operational support
 - PIC (Barcelona)
 - IHEP (Russia)
 - Soon also CNAF (Bologna)



Recap



- Why CASTOR?
 - Commercial alternative exists and have been (and continue to be) evaluated
 - No solution that address the full problem
 - Difficult to integrate and may require special infrastructure
 - Relatively expensive
 - Open source alternative exists for parts of the system
 - Relatively new and untested at the scale of CERN Tier-0/1 requirements
 - Operation support? data management is mission critical for CERN and cannot rely on best effort support
- Collaboration with industry
 - So far relatively modest but there is room for more, e.g.
 - Scheduling systems: storage request scheduling and resource sharing
 - Expert systems: policy engines for resource allocation
 - Licensing: CASTOR is currently distributed under GPL