CSC Storage Projects

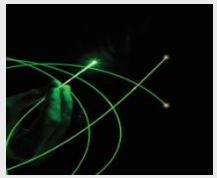
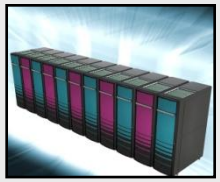
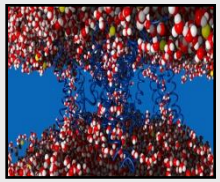
CSC = CSC - IT Center for Science

NSC'08

Ari Lukkarinen

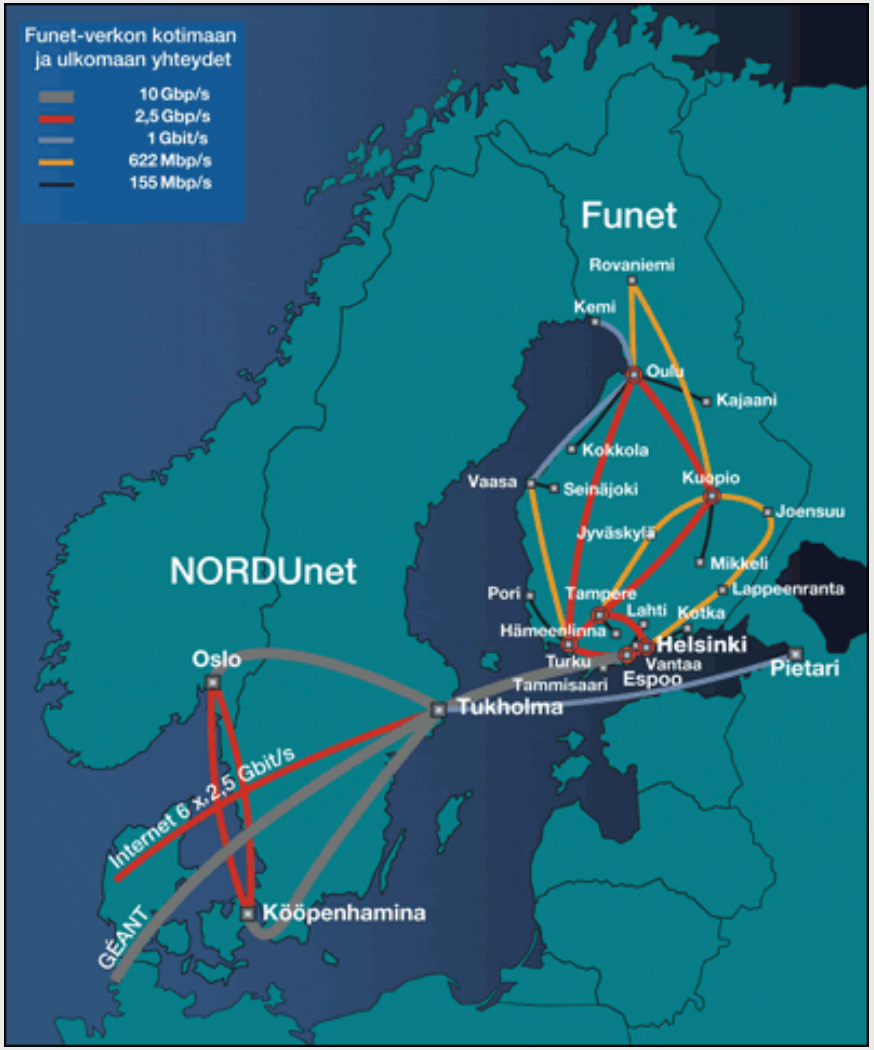
CSC

- Company owned by ministry of education
- Staff: 150 people
- Services:
  - Funet
  - Computing
  - Application
  - Information management
  - Data services for science and culture

CSC

Funet - Finnish University and research NETwork



- ✓ **High quality academic Internet services for Finland**
  - **Advanced services like IPv6 and IP multicast**
  - **Core line speed – 2,5 Gbps**
  - **Funet member access at speeds up to 1 Gbps**
  - **Lightpaths**
  - **International connectivity through NORDUnet**
- ✓ **85 member organizations:**
  - **All Finnish universities (20) and polytechnics (30)**
  - **Public research institutions and administrative organizations close to higher education (35)**
  - **Over 350 000 end users**

CSC

COMPUTING

- **CRAY XT4/XT5 alias Louhi**
  - **2356 AMD Quad Opteron 2,3 GHz CPUs**
  - **9424 cores (4048 XT4 + 5376 XT5).**
  - **Memory ~ 10,3TB**
  - **Theoretical computing power 86,7 teraflop/s**
  - **Lustre filesystem (70 TB)**

- **HP-CP4000BL ProLiant Supercluster alias Murska**
  - **554 dual-core AMD Opteron 2.6 GHz CPUs 2178 cores,**
  - **Memory ~5 TB**
  - **Theoretical computing power 11,3 Teraflops**
  - **HP SFS filesystem (100 TB)**

Power consumption ~ 1 MW

CSC

DATA

- **CSC Storage environment**
- **Projects**
  - **RTVA (Radio and TV archive)**
  - **HIP/CERN**
  - **Long Term Storage**

Operational environment

- **Universities have a lot of cheap labour available for IT administration.**
- **Storage capacity is inexpensive. Research groups/laboratories can easily have systems being able to store tens of terabytes of data.**
- **Problem: Who will take care of the data ?**

Organizations and research groups have lot of data, but they do not have the manpower, resources, or knowhow, how to make sure that their data will later be easily accessible and easily usable. They would like someone to do a long term storage solution for them.

CSC

CSC Storage environment
Storage hardware, midrange

2x HDS AMS1000
(400 TB)

Other midrange storage systems:

- EMC CX300 (10 TB)
- Fujitsu Siemens FibreCAT CX700 (50 TB)
- 2 x LSI/Engenio 6998 (70 TB total)

+ other smaller disk arrays

+ HP SFS (100 TB)

CSC

CSC Storage environment
Storage hardware, enterprise

HDS USP 100

- Heavy duty databases
- “Not for scientific customers”
- Library systems

CSC

Storage Projects: RTVA

- Radio and TV Archive, -> National audio visual archive will archive broadcasted TV and radio programs.
- CSC was chosen to be a technical partner taking care of building up the system.

- Specs:
  - Number of TV channels, about 10-20.
  - Number of radio channels, about 30.
  - Data volume, about 30 TB/year.
  - Limited number of users.

- Project started 21.1.2008. New law was needed.

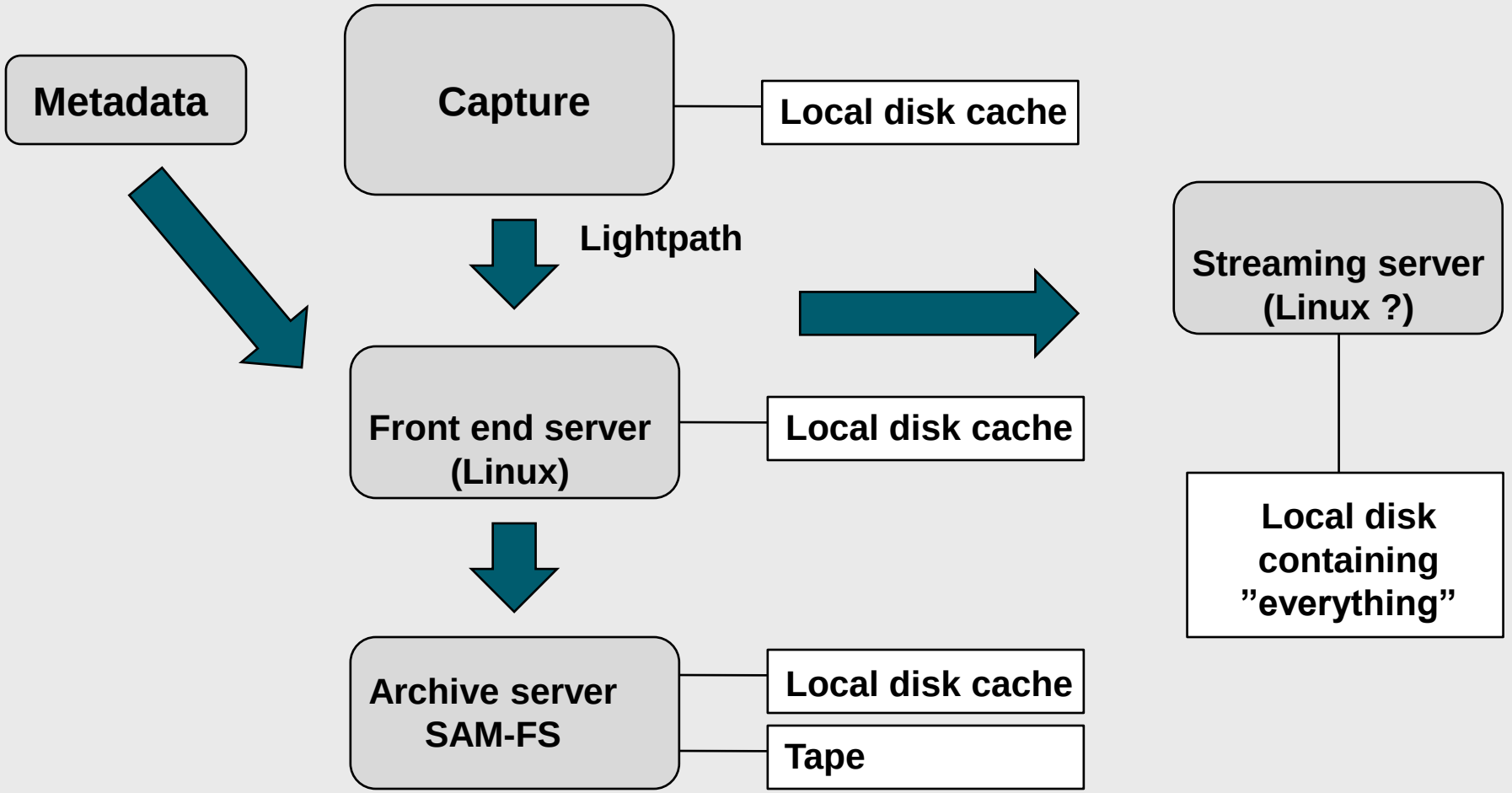
CSC

RTVA: Data capture

- Problem: Nation wide channels are relatively easy to handle. Small local TV/radio operators are problematic, because they require local precense.
- Subcontractor will be used to capture the data.
- Another subcontractor will make the metadata software.
- Schedule:
  - Data capture: Test phase starts 1.11.2008
  - Production phase will start 1.1.2009

CSC

Data storage



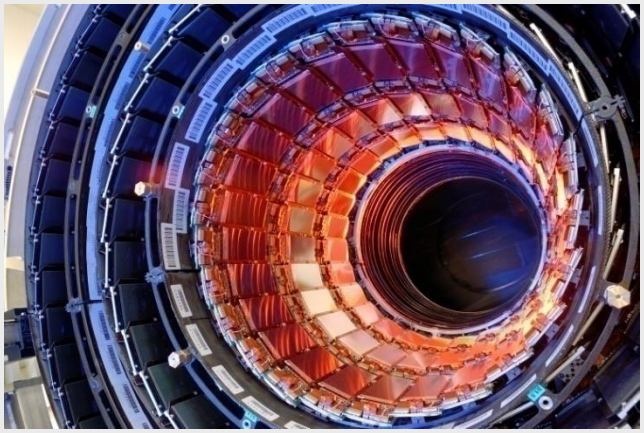
Storage challenges

- Reliable data transfer ?
  - Dedicated fiber connection (+USB disks)
- Data transfer tools
  - Rsync + other tools
- Long term storage issues (intergrity)
  - SAM-FS tuning (verify after write, checksums)
  - Additional site wide data integrity software under preparation
- Linux filesystem issues
  - Solaris and ZFS initial candidate.
  - XFS propably the best candidate (no RHEL)
  - XFS can not be resized beyond 2 TB

CSC

Storage Projects: HIP / LHC

- HIP (Helsinki Institute of Physics) decided to store date from LHC experiment to CSC.
- Finland participates three CERN experiments: CMS, Totem, ALICE
- CSC is now a part of joint nordic Tier-1 center and a Tier-2 center.
- Tier-1 centers store the data (disk and tape)
- Tier-2 centers use the data and compute new physics.

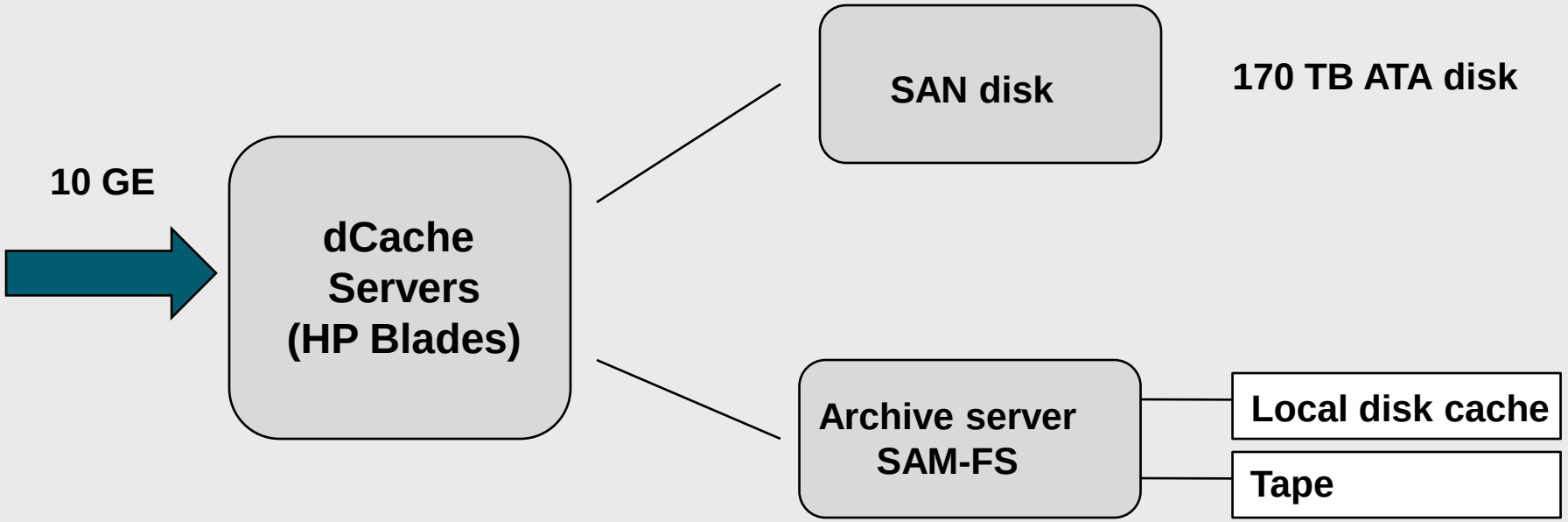


CSC

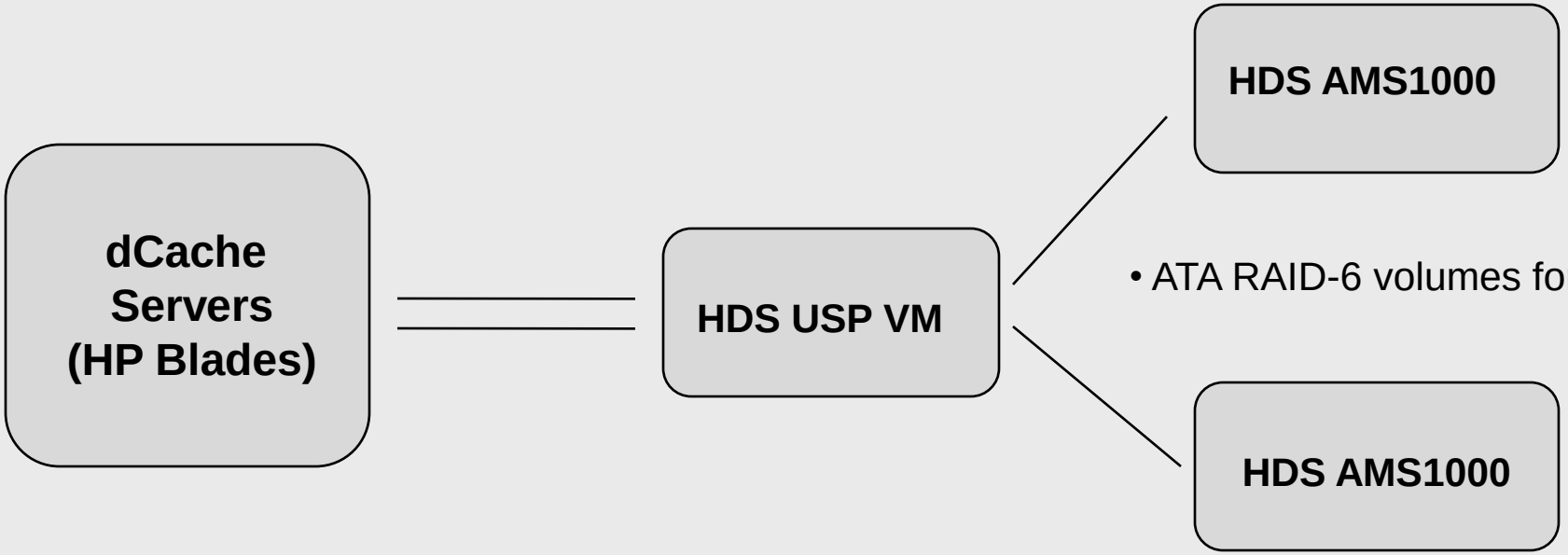
dCache

- Accelerator creates 40 PB data. 10 TB will be stored to nordic countries.
- Data managed with dcache software.
- Dcache ”virtualizes” storage devices.
- Data has an ”URL”. Physical location available from database.
- CSC provides 170 TB disk and 70 TB tape capacity.
- Capacity will be doubled 2009.

dCache Implementation



dCache SAN HW



- ATA RAID-6 volumes for data

CSC Storage environment
Storage hardware, virtualization

HDS USP VM

- Why ?
  - Large LUNs
  - Distributes load
  - Active-active controllers (hides midrange complexity)
- Cache 12GB
- No internal disks
- 200 TB virtualized storage capacity
- Old arrays will be virtualized.

CSC

dCache Challenges

2 TB limits

- **XFS**
- **Disk arrays**

Max LUN count limitation

dCache

- **Unsufficient documentation**
- **Lack of recovery tools**
- **Not a ”product”, but a prototype software.**

Virtualization tuning

CSC

Storage Project: Long Term Storage (LTS)

- **Long Term Storage (10 - ∞ years).**
- **Problems:**
  - **Disks and other HW are unreliable.**
  - **Media change interval is short.**
  - **People make mistakes.**
- **Need for a secure storage.**

CSC

Secure storage challenges

- How to garantee the data integrity ?
  - Checksums
  - Is the problem on a database or on the data ?
- How tho make sure that the data exists ?
  - External database on data, cheksums, cheksums
  - WORM media.
  - Many copies of data
    - Physical separation (more than one data center)
    - Technological separation
      - More than one archive software.
      - Different HW components

CSC

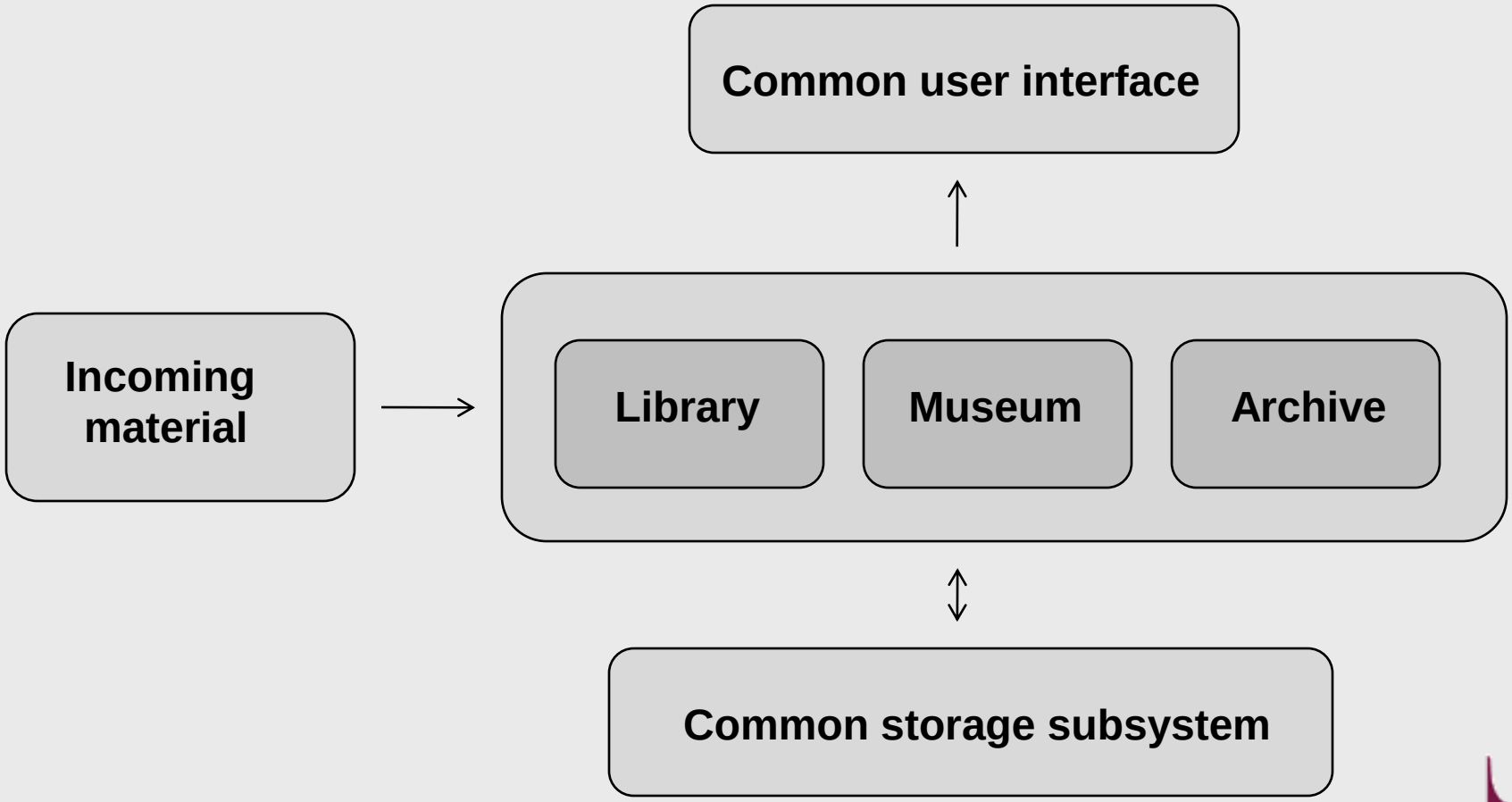
LTS, Datacenter challenges

- Datacenters have to able to communicate with each other
  - Metadata database of all data needed.
  - Files have to have common properties that can be used to identify the data.
  - There must be a way to transfer some or all data from one datacenter to another.

LTS and Finland

- Government decision:
  - A national digital library (NDL) will created.
  - Information from museums, national archive, and from libraries will be stored into a centralized system.
  - The project includes 35 organizations and 70 persons.

NDL schema


Common user interface
Incoming material
Library
Museum
Archive
Common storage subsystem

Schedule

- 2008: Planning started
- 2009: Planning continues, piloting
- 2010: Storage subsystem building
- 2011: User interface building
- **2011: System operational**

Challenges

- Standard metadata set has to be created.
- Different ”vocabulary” btw. organizations. For example the definition of ”collection” depends on organization.
- Data formats and migrations from a format to another.
- Lot of participants.
  - Technological skills of organizations vary.
  - Various needs. All requirments can propably not be met.
- Schedule

CSC

LTS and CSC

- CSC participates national digital library project.
- CSC could be one site taking care of an LTS archive.
- Metadata database program under development.
  - Collect information on filesystems (+ from dCache) to a central database.
    - Intergrity of the data
    - Reporting, file history log
    - Data migrations
- CSC environment is being evaluated by using TRAC. (Trustworthy Repositories Audit & Certification: Criteria and Checklist)

Storage Challenges

- **Lack of Object Storage Devices !!!!**
  - **Limits disk array functionality**
- **Data Life cycle management**
  - **Disk array provides disk for tens of servers.**
  - **Disk array replacement are extremely difficult.**
- **SAN networks**
  - **Is there really need to certificate the whole chain from an application disk array ?**
  - **We need more clever SAN networks (or they will become obsolete).**

CSC