The EuroStore Project – Results and Deployment @ DESY

M. Gasthuber, P. Fuhrmann   *DESY/Hamburg*
D. Roweth   *Quadrics Limited Bristol/UK*

Overview

- The EuroStore Project

  - and Results

- The Model (@ *DESY*)

- The Implementation (@ *DESY*)

The EuroStore Project

- Esprit Project (funded by EU)

- Started March 1st 1998 – for 2 years (extended to July 2000)

- Organizational Structure

  - 4 Full Partners – CERN(S), DESY(S), QSW(I), HCSA(I)

  - 3 Associate Partners – TERA(S), HNMS(I), AMC(I)

  - Separated into *Development* and *End−user* partners

    - Development Tasks – QSW, DESY

- GOAL

  - develop a compact and scalable high performance Storage System based on existing components and long term experience.

Work Structure

- Tasks separated into
  - Development
    - Parallel Filesystem (PFS) – by QSW
    - Hierarchical Storage Manager (HSM) – by DESY
  - Requirements/Evaluation & Tests
    - Prototype housing and tests – by CERN
    - Tests – by TERA, HNMS, AMC, HCSA
  - Commercialization/Market Survey
    - by QSW, HCSA

Development Tasks

- Parallel Filesytem (PFS) – done by QSW, started with *background*

  - high performance parallel FS, running on single or multiple nodes

    - data distribution on several nodes and disks

  - UNIX System V layered FS

- HSM – done by DESY, started from scratch

  - distributed – network centric architecture (for scaling purposes)

    - service built out of distributed sub–services – implemented as discrete process or thread

  - the *black hole* (and inverse) for all type of data, managing all tertiary storage components

Results (PFS Extensions)

- File migration (*transparent move data to/from HSM*)

- Global Lock Manager

- Performance (internal structure, Data Path)

- RMS Integration

- ELAN Integration (*ELAN based Kernel RPC*)
    - proprietary network, high bandwidth/low latency

- MPI/IO

- mmap support

- HSM *aware* Backup and Restore

Results (HSM)

- JAVA based (less than 5% in C)

- internal interface for easy adoption of new robotics & drives

- Small & Compact (Philosophy)

- Secure – control and administration/monitoring connections

- Easy Admin – JAVA/Applets or interactive *ssh*

- Scalable (*Mainframe to Laptop – or many of these*)
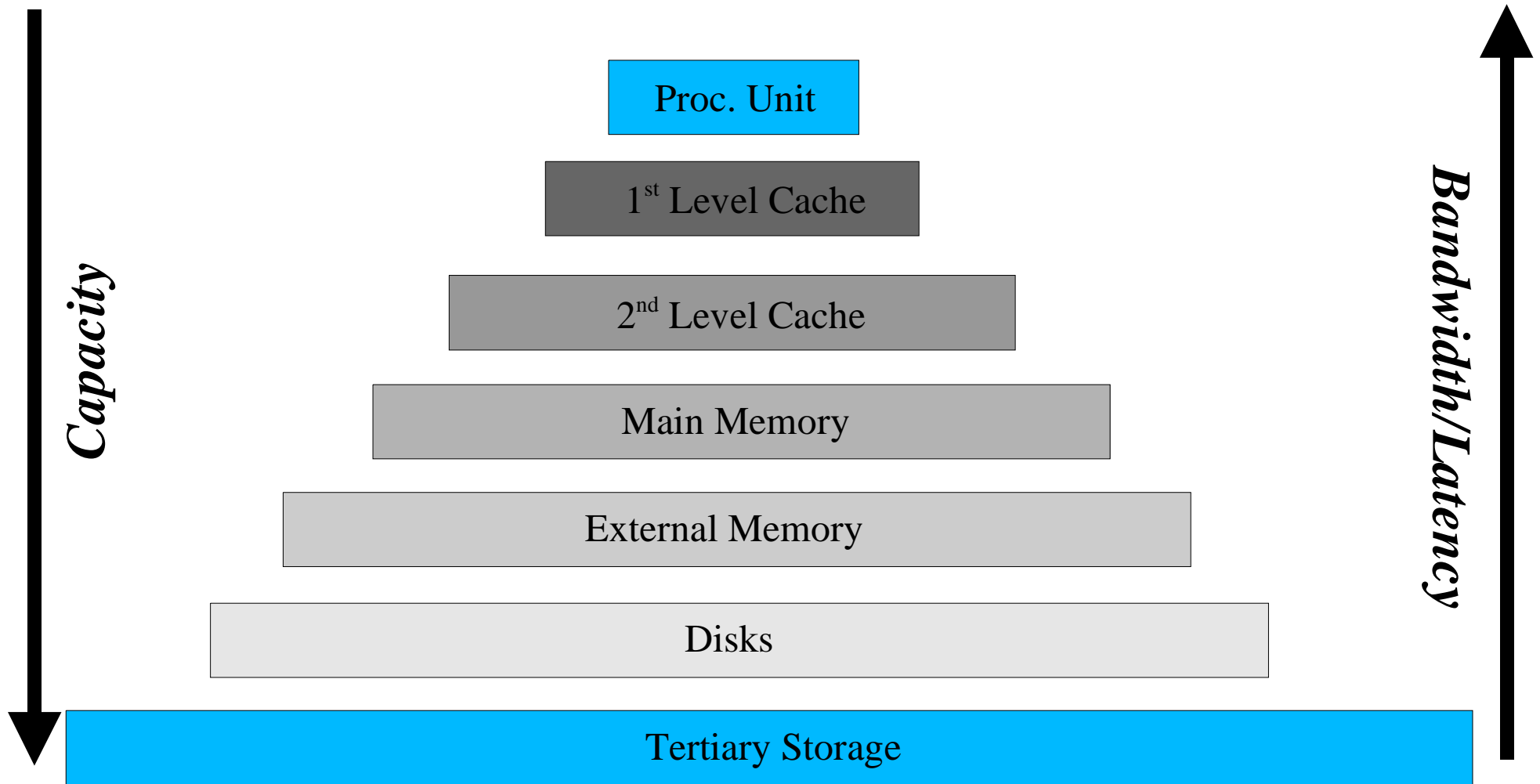
- simple *PUT/GET* C based API (*versatile*

Applicability for DESY

- Need
  - natural cut between namespace and IO operations
  - single instance namespace service
  - Arbitrary number of IO agents
  - homogenous access profile from arbitrary number of heterogeneous client systems
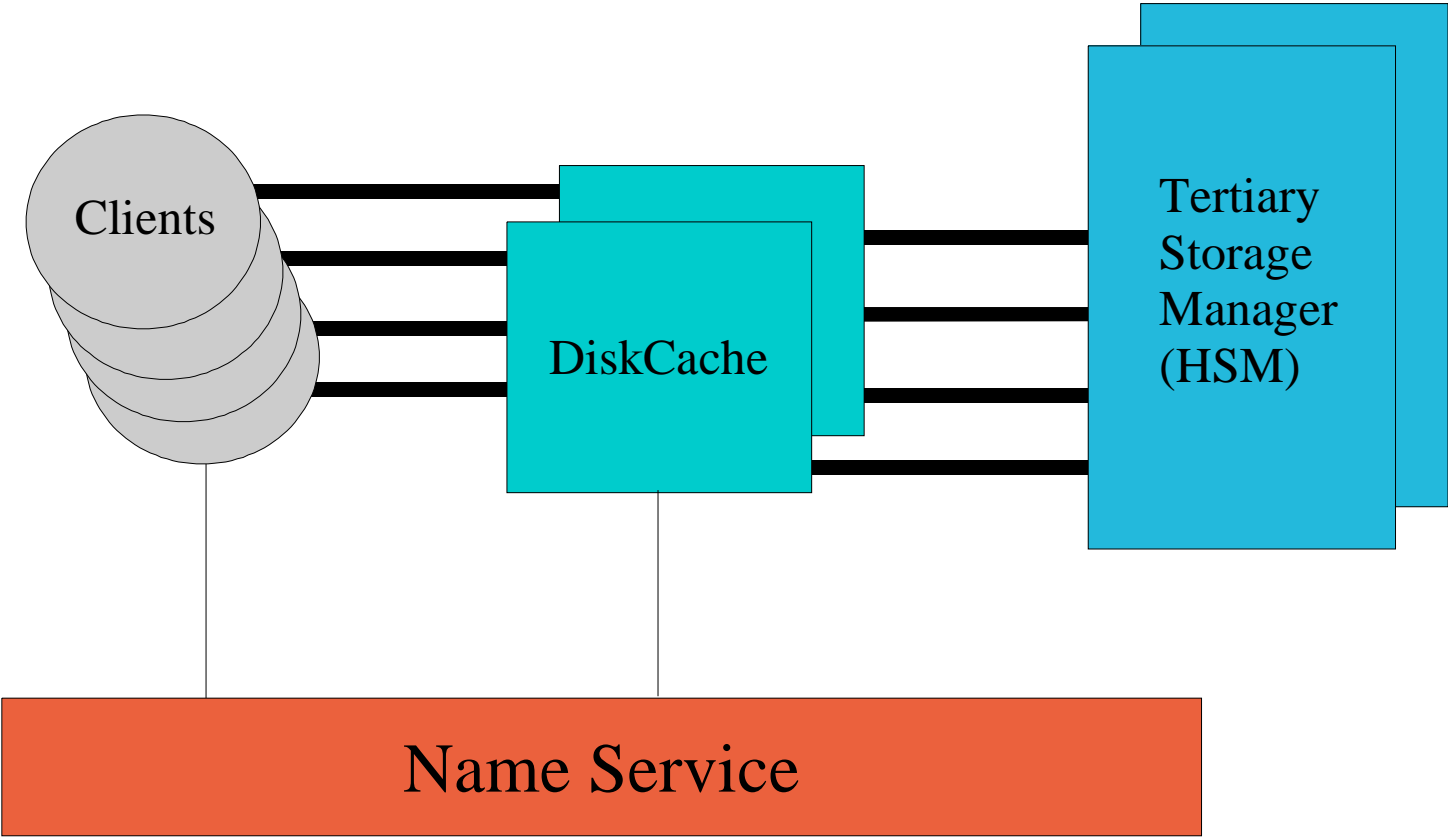
  ➡ Only HSM component of EuroStore

Base Model

Capacity

Bandwidth/Latency

Proc. Unit

1st Level Cache

2nd Level Cache

Main Memory

External Memory

Disks

Tertiary Storage

Model

Implementation – *Nameservice*

- PNFS
  - maps pathname to opaque metadata (stored in DB)
  - introduced 1996
    - See
      - CHEP'97 (Berlin) – Talk of P. Fuhrmann
      - http://watphrakeo.desy.de/pnfs
  - in production with OSM

Implementation – *DiskCache*

- Old idea – evolving design

- Increase Demand @ DESY (OSM)

- Started Collaboration with FNAL in September 99

- See

  - Poster of P. Fuhrmann (at this conference)

  - http://wizard.desy.de

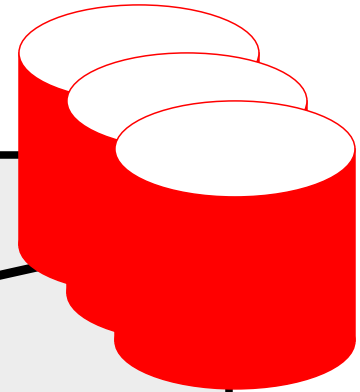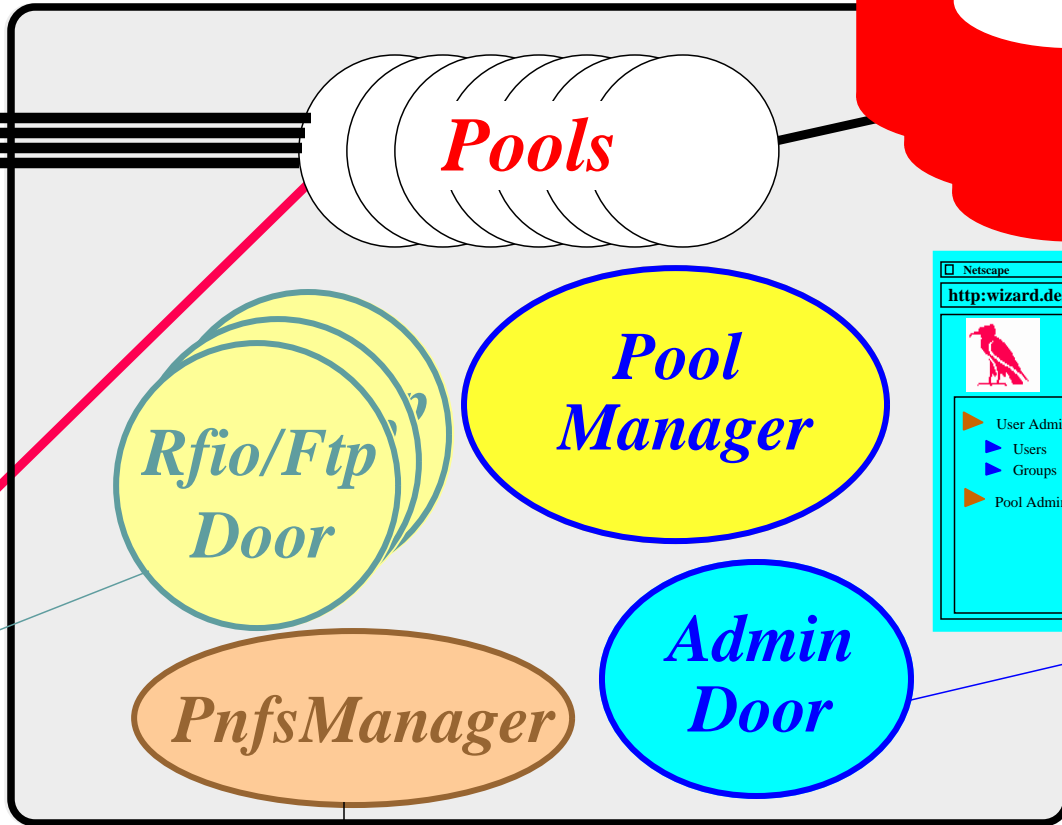# Implementation – *HSM*

- EuroStore HSM Prototype as *Base*
  - continue Development efforts in EuroStore II (*proposed follow−on Project*)
    - Hints
    - SAN Integration
    - More Platforms (Linux)
    - etc.

Tape Library

OSM
Enstore
Eurostore

Pools

Rfio/Ftp Door

Pool Manager

Rfio/Ftp Client

PnfsManager

Admin Door

Filesystem Namespace Provider

pnfs

Netscape

http:wizard.desy.de-admin

Attracting Cache Wizard
Administration

► User Admin
  ► Users
  ► Groups
► Pool Admin