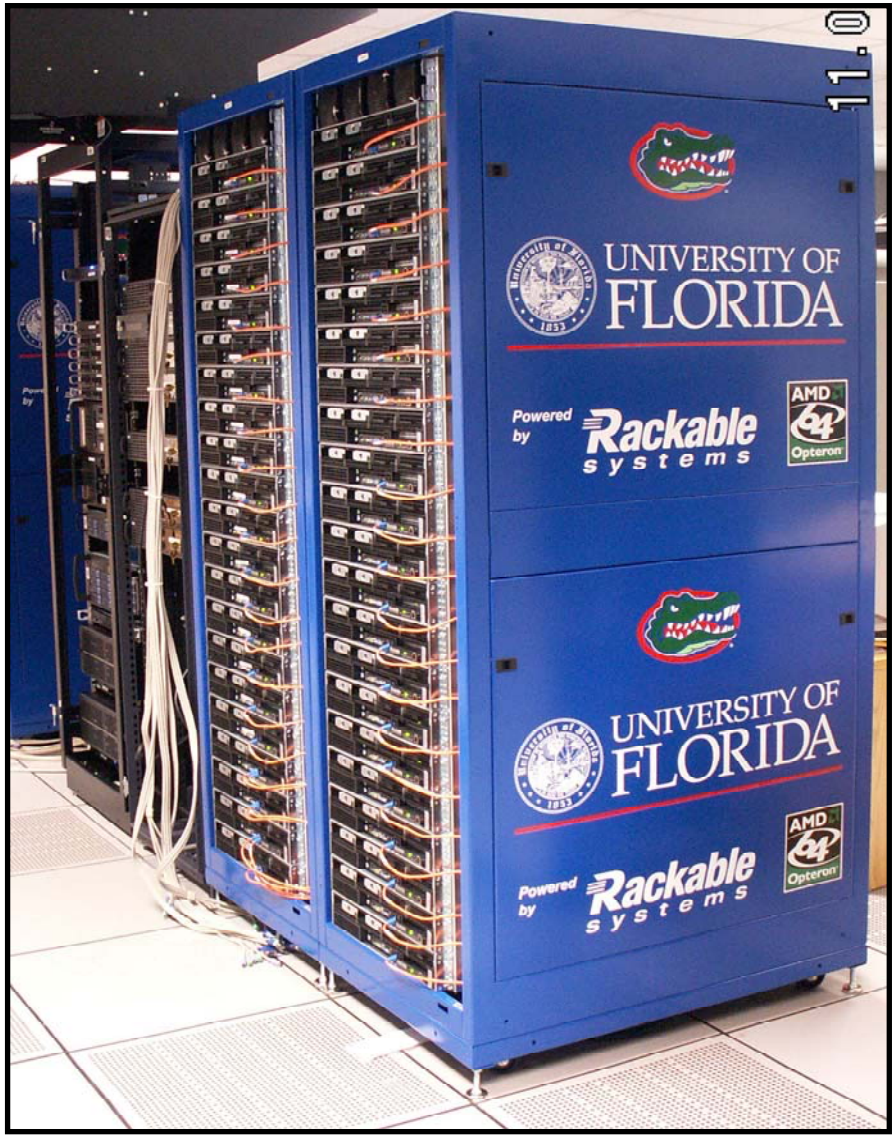
Using the Tier2 Facility at UF

Jorge L. Rodriguez
Florida International University
Southeastern CMS Physics Workshop

UFTier2 Computing Hardware



Computational Hardware

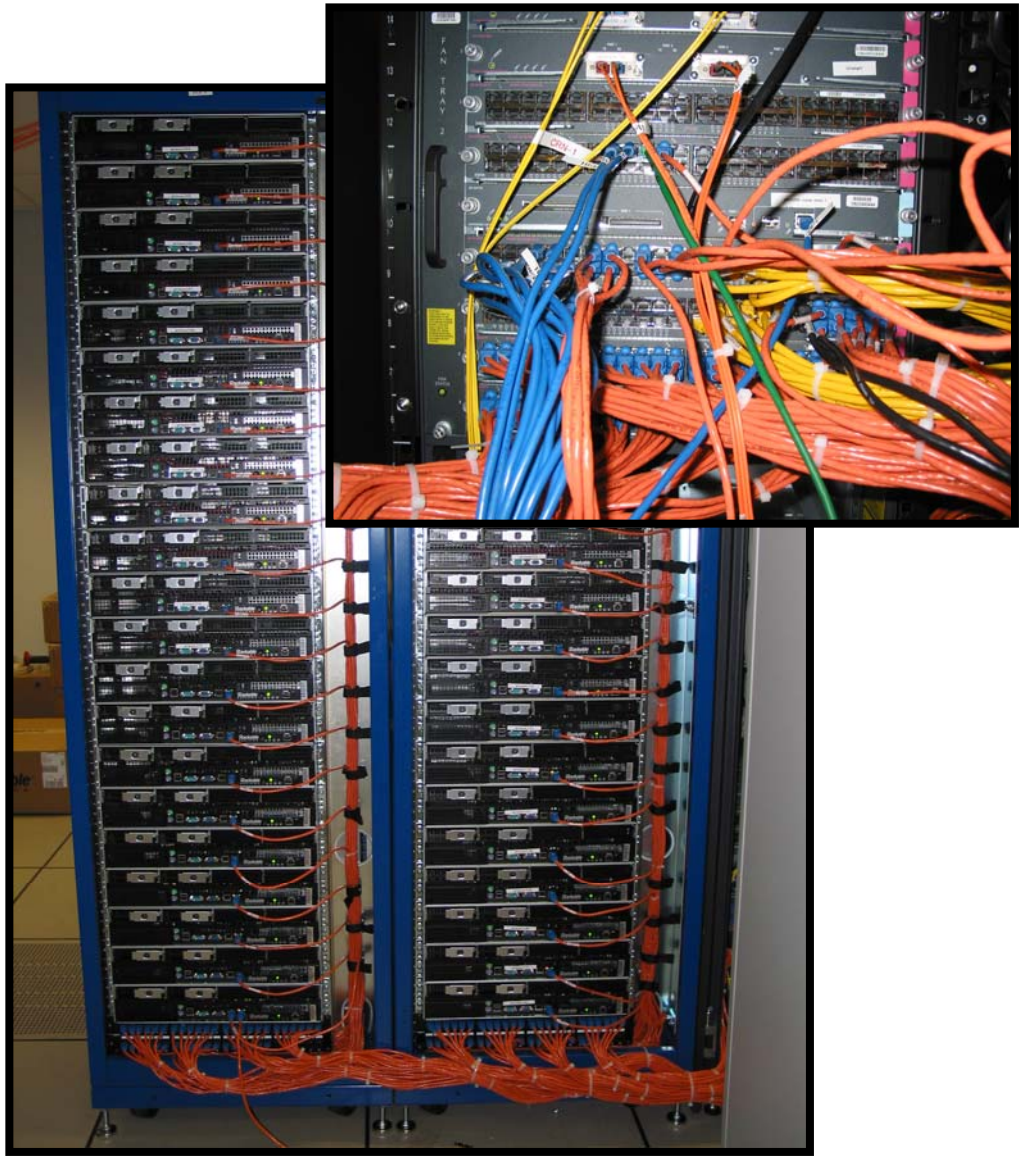
- 130 Dual dual-core (520 CPUs)
  - Opteron 275/280 (2.2/2.4 GHz), 4GB RAM, dual GigE NIC
  - 2 x (250 or 500) GB disks/node
- 740 kSpecInt2K

Storage Hardware

- 48 TB Hardware RAID
  - 30 TB in dCache RAID pools
  - Assigned to local users!
- 83 (37) TB on resilient dCache
  - One dCache pool/node (126 nodes)
  - For cmsdata, cmsprod & gusers
  - In production and available!

UFTier2 Networking

- Mostly a Cisco 6509
  - All 9 slots populated
  - 4 x 10 GigE ports (2 blades)
  - 288 x 1 GigE ports (6 blades)
  - Supervisor Engine Sup720
- Wan connectivity @ 10 Gbps
  - From out 6509 to CNS router @ 2 x 10 Gbps
    - Campus Research Network
    - Dedicated to UF research
  - From CNS router FLR over Ultralight wave @ 10 Gbps
  - All Tier2 public interfaces on Ultralight network

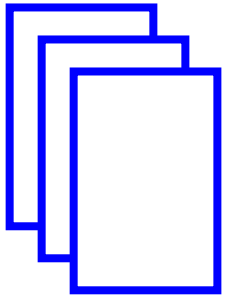


UFTier2 Hardware at t0

- Tier2 design specification
  - 1000 kSpecInt2K
    - Now at 75%, 110% if we include new HPC purchase
  - 200 TB of disk
    - Now at 60% in dCache more in local and NFS volumes
    - Soon will add more storage to resilient dCache
  - 10 Gbps connectivity to WAN
    - Campus Research Network
    - Tier2 network traffic does not compete with other campus traffic

- A sizable regional facility to be exploited

UFTier2 Configuration



Services
&
development

CMS Interactive Analysis Farm & User Interface

alachua

newberry

melrose

archer

CMS Compute Cluster

ufloridaPG

grinux143

grinux226

grid access only

dCache & Scratch Storage (NFS)

grid (SRM) and local access

grid and local condor access only

Uflorida-IHEPA

grinux101

grinux142

iHEPA Compute Cluster

The UF Tier2 Meta Cluster

CMS Analysis Farm

- alachua
- archer
- newberry
- melrose

Service Nodes

- Teir2 webser
- ultralight
- ufdev01
- ufdev02

CMS Service Nodes

- gums
- PhEDEx
- squid

ROCKS Mngmt

- frontend2
- frontend41
- frontend41-64

Production Clusters

- iHEPA
  - grinux
    - 101-142
      - grinux
- ufloridaPG
  - grinux
    - 143-226
      - grinux

dCache Headnodes

- srm1
- gridftp1
- gridftp2
- gridftp3
- dcache-admin
- dcap1
- admindoor

RAID5 dCache pools

- wpool
- wpool

nfs servers

- gatoraid
- gatoraid

Outreach Cluster

- ufloridaEO
- grinux
  - grinux

OSG Integration

- ufloridaIGT
- grinux
  - grinux

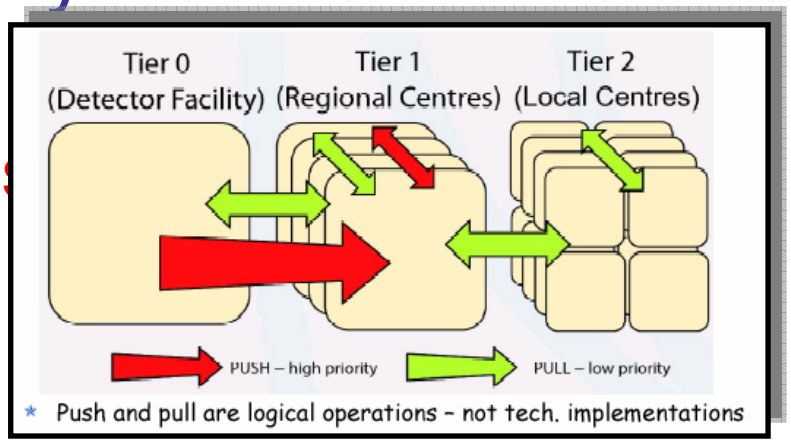
Services at the UFTier2

- Data movement, management and storage
  - dCache space: access via srm interface and local dccp
  - PhEDEX, DBS/DLS
    - Automated systems of downloading and publishing CMS data
    - Services fully tested and commissioned during CSA06
- User login and Grid Services
  - Interactive CMS Analysis Cluster: UNIX login
  - User Interfaces: grid access to OSG and LCG resources
  - Complete CMSSW installation
    - Deployed by cmssoft
    - Can be used to compile your CMS jobs
  - Webspace (~/public_html), twiki…

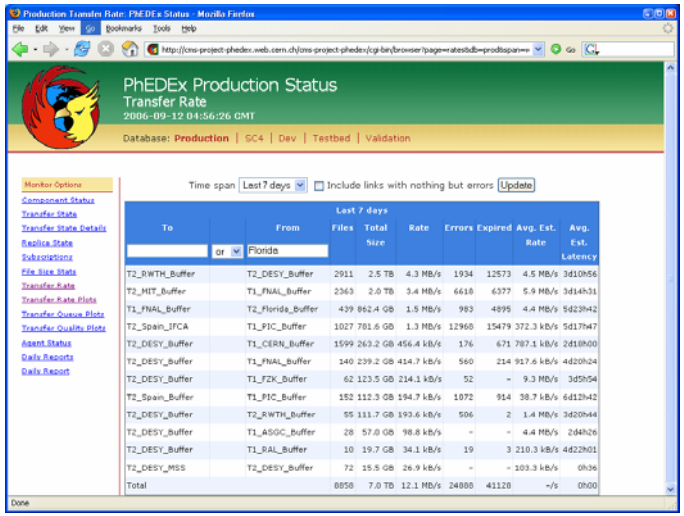
CMS Data Management System

Distributed Data Mgt & Movement System

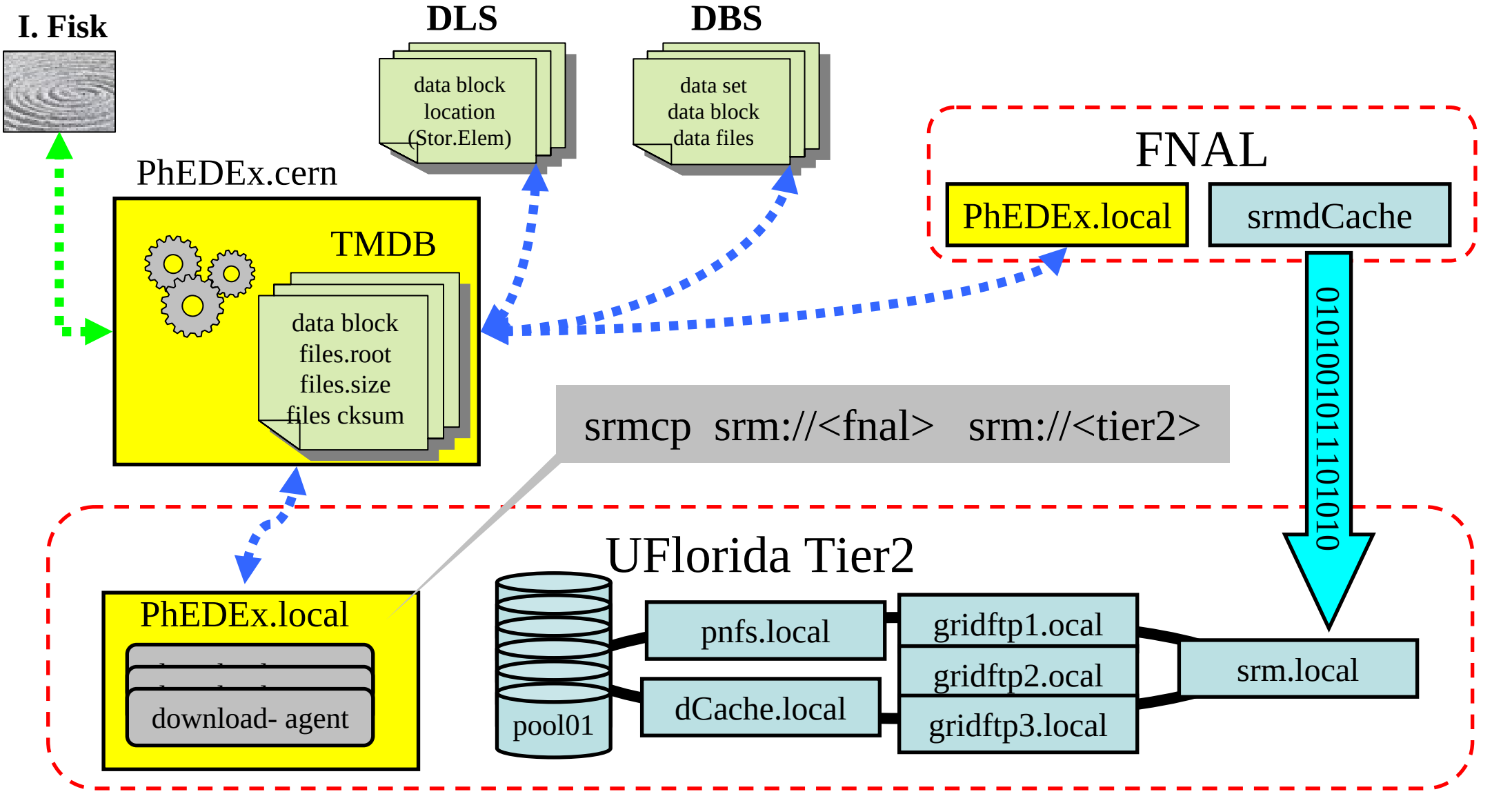
- Identification & location of data
- Worldwide movement of data replicas
  - Tier0 ⇔ Tier1 ⇔Tier2 ⇔Tier0…
    - Volume ~ PBs ≈ O(10 M) files
    - Transfer speeds ~ 5 Gb/s
- Components include
  - PhEDEx:
    - Hart of the system
  - DLS: Data location Service
  - DBS: Data Bookkeeping System



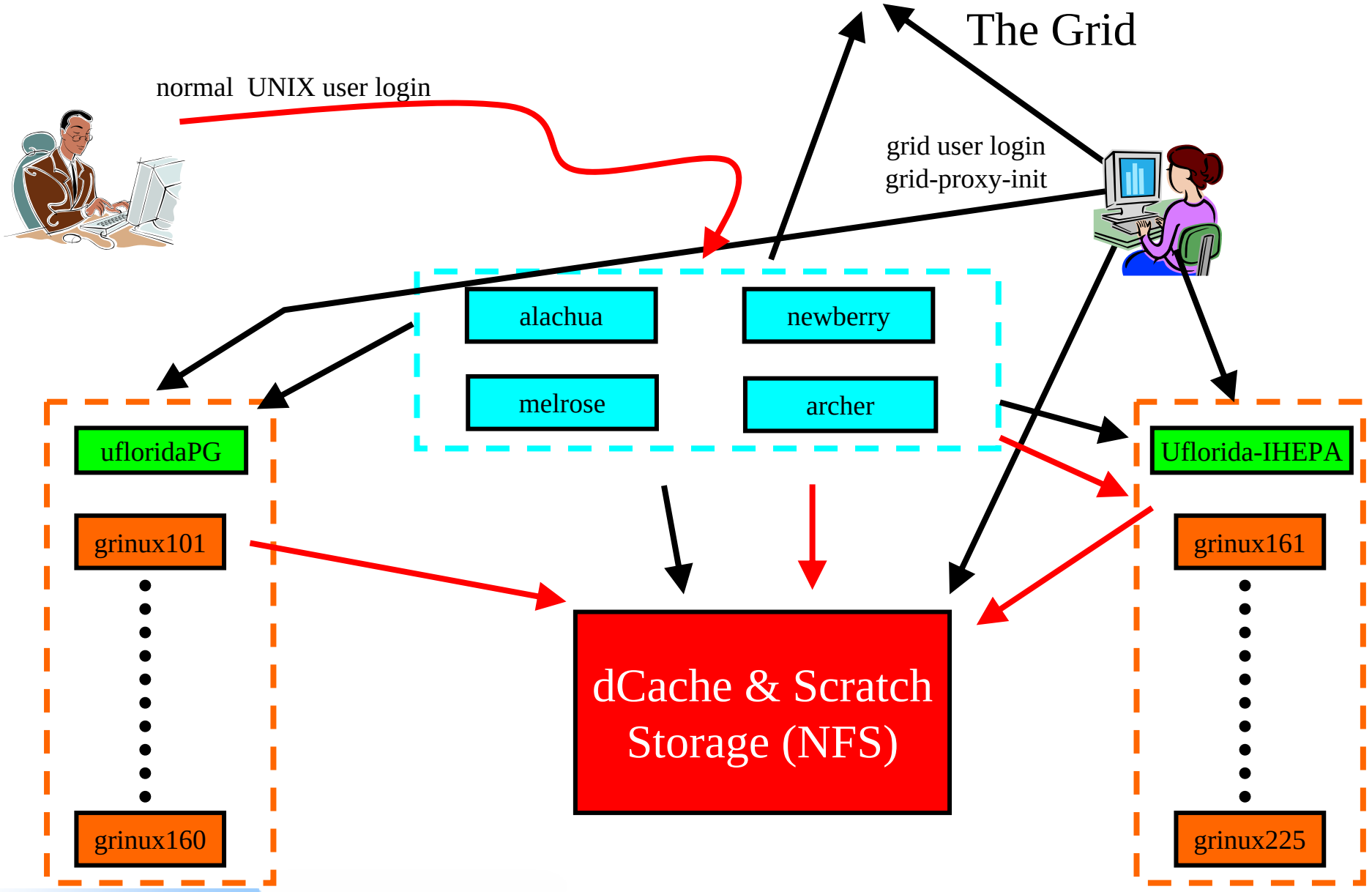
- Management at the block level (many files ~10TB)
- Monitoring



Data Mngmt & Movemnt System



Accessing the Facility at UF



Accessing the Analysis Farm

- You can request a local UNIX account
  - Send email to t2-admins@phys.ufl.edu
    - Name, Institution, phone number and email
    - We recommend that you use ssh keys in lieu of passwords
- For the workshop you can login into analysis farm
  - Use the cmstutor account provided for you

    cmstutor@alachua(archer)(newberry)(melrose).phys.ufl.edu
  - A password will be issued but we prefer that you use your ssh keys
    - In principle a more secure way to login
    - With ssh agents and forwarding is much more convenient!

Computational Resources

CMS Analysis Farm at UF

- Interactive analysis, batch computing
- The CMSSW environment is installed there for you

```
[cmstutor@alachua cmstutor]$ mkdir jorgeSpace
[cmstutor@alachua cmstutor]$ cd jorgeSpace
[cmstutor@alachua jorgeSpace]$ scramv1 project CMSSW
```

- CMS and Grid applications installed
  - Use root.exe to analyze your data
  - Use to run cmsRun interactively
  - Use to submit cmsRun to local condor batch on the iHEPA cluster
  - Can be used to launch jobs to the Grid and to move data via SRM

Storage Resources

- Home area: namespace: $HOME
  - Entire partition is < 1.0TB & is shared by all users
  - Mounted (visible) via NFS on analysis farm nodes only
  - No space management, no backup, RAID 5 redundancy
  - Source code, other work but users are encouraged to backup their own data
- Scratch storage: namspace: /scratch/<cmsanal1(2)>/<uname>
  - Each partition is on a 250 GB hard drive
  - local and cross mounted on analysis farm nodes only
  - No space management, no backup, no redundancy whatsoever!
- NFS mounted storage: Namespace = /raid/raid1(2)(3)/<uname>
  - Partitions are larger from 500GB to 2.0 TBs
  - Mounted everywhere from the analysis farm nodes & workernodes
  - No space management, no backup, RAID 5 redundancy!
  - NFS can be un reliable and perform poorly under heavy load

Storage Resources: dCache

What is dCache storage? from a user's point of view

- One giant disk (virtual filesystem)
- High performance, robust, reliable, expandable…
- A storage system designed to meet the needs of HEP experiments (CMS) in particular
- Various means of accessing the files
  - SRM: Storage Resource Manager
    - GSI authentication, VO level authorization
    - Managed storage (space reservation, request queuing…)
  - CMSSW
    - dCache is accessible from CMSSW framework
  - ROOT
    - There is a dCache xrootd implementation but its not implemented yet at UF
  - Auto mounted across all nodes via pNFS
    - POSIX like access to dCache but careful dCache is not POSIX!

Storage Resources: dCache

- dCache area: namespace: RAID & resilient
  - dCache-RAID
    - dCache pools are mapped onto large RAID arrays
    - High bandwidth & fast I/O access to local disk
    - How to access:

      from CMSSW: ”dcap1.local:22135/pnfs/phys.ufl.edu/data/raid/… “

      from shell: ls or dccp… /share/dcache/data/raid/… “

      from srm: srmcp srm://srm1.phys.ufl.edu:8443/raid/…

      This has not yet been enabled on RAID pools, also you’ll need a grid certificate and authorization to use (t2-admins@phys.ufl.edu)

      from root: The xrootd feature has not yet been enabled at UF

Storage Resources: dCache

- dCache area: namespace: RAID & resilient
  - dCache-resilient
    - dCache pools are scattered across all 126 workers
    - More efficient use of network topology for accessing files
    - Even higher bandwidth & faster I/O possible due to distributed nature of data storage
    - Access is just like previously with the RAID dCache space, other subdirectories include
      - …/data/resi (dccp accessible)
      - …/data/cmsuser (srmcp accessible)

      Note: file permissions do not currently restrict access

Storage Resources: dCache

dCache area: Summary

- UF dCache can do ~ 10-25 TB/day via SRM
- 30 TB available to local users as RAID pools
- 37 TB in resilient pools (shared, cmsprod…)
- Resource allocation: (t2-admin@phys.ufl.edu)
  - Large data samples via PHEDEX operator
  - Generate your own and dccp or write directly into RAID pools

Login to Tier2 for the Workshop

- There is an account (cmstutor) on:
  - alachua, archer, newberry, melrose .phys.ufl.edu
- password is :
- If you'd like to try ssh logins via agents let me know.
  - Its neat and secure
  - you only have to remember one password for all your accounts
- Once you are in type
  - create your space ie, jorgeSpace/
  - source /raid/raid4/analfarm/setup/setup.sh
- Follow the CMS Offline Workbook tutorials at
  https://twiki.cern.ch/twiki/bin/view/CMS/WorkBook

Using ssh with agents

A more secure and less troublesome way to login

Procedure:

1. Generate priv/pub key pair (windows: puttygen)
   - save both to a file on your PC
   - install your public key on the target server cmstutor@alachua.phys.ufl.edu
2. Launch ssh agent (windows: pageant)
   - select your private key and type passphrase
   - use putty to login, you don't need to type your password again, furthermore you daisy chain to another host

END

Regional Grid Resources

A Developing Florida Grid

- Grid Computing Resources in Florida
  - FIUPG (already on OSG prod)
  - FSU (physics site was on OSG ITB)
    - Expect to join OSG soon
    - HPC has expressed interest in joining OSG
  - FIT (resources being developed)
    - 10 node PIII cluster + analysis machines
    - Expected on OSG soon
  - UF HPC (large resource at UF, VO…)
    - Faculty lead initiative, P. Avery *et.al.*,
    - Joined OSG in 01/07 (C. Prescott)!
- The Florida Grid Operations Center fGOC
  - Not a VO support center
  - fGOC representative at OSG Operations meeting
  - OSG integration point
  - Help support fabric level infrastructure