ARCHER/RDF Overview

How do they fit together?

Andy Turner, EPCC

a.turner@epcc.ed.ac.uk

www.epcc.ed.ac.uk
www.archer.ac.uk

Outline

- ARCHER/RDF
  - Layout
- Available file systems
- Compute resources
  - ARCHER Compute Nodes
  - ARCHER Pre/Post-Processing (PP) Nodes
  - RDF Data Analytic Cluster (DAC)
- Data transfer resources
  - ARCHER Login Nodes
  - ARCHER PP Nodes
 - RDF Data Transfer Nodes (DTNs)

ARCHER and RDF

ARCHER

- UK National Supercomputer
- Large parallel compute resource
    - Cray XC30 system
    - 118,080 Intel Xeon cores
    - High performance interconnect
- Designed for large parallel calculations
- Two file systems
    - /home – Store source code, key project data, etc.
    - /work – Input and output from calculations, not long-term storage

RDF

- Large scale data storage (~20 PiB)
  - For data under active use, i.e. not an archive
  - Multiple file systems available depending on project
- Modest data analysis compute resource
  - Standard Linux cluster
  - High-bandwidth connection to disks
- Data transfer resources
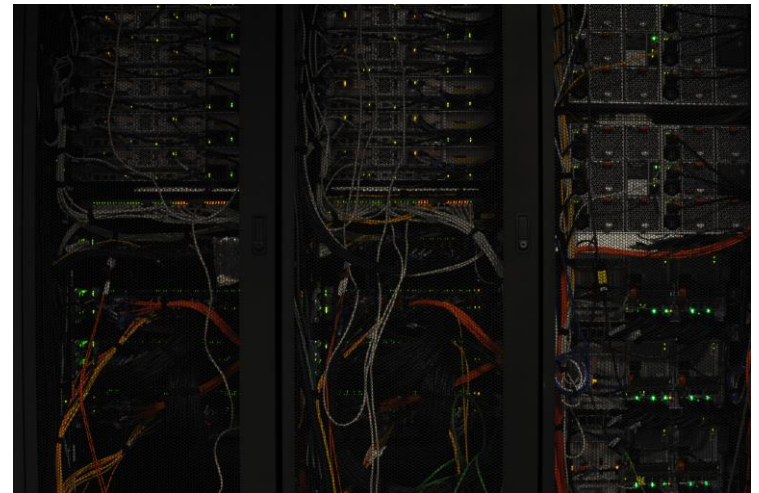
Terminology

- ARCHER
  - Login – Login nodes
  - PP – Serial Pre-/Post-processing nodes
  - MOM – PBS job launcher nodes
  - /home – Standard NFS file system
  - /work – Lustre parallel file system
    - ARCHER installation is a Sonexion Lustre file system
- RDF
  - DAC – Data Analytic Cluster
  - DTN – Data Transfer Node
  - GPFS – General Parallel File System
    - RDF parallel file system technology from IBM
    - Multiple file systems available on RDF GPFs

Overview
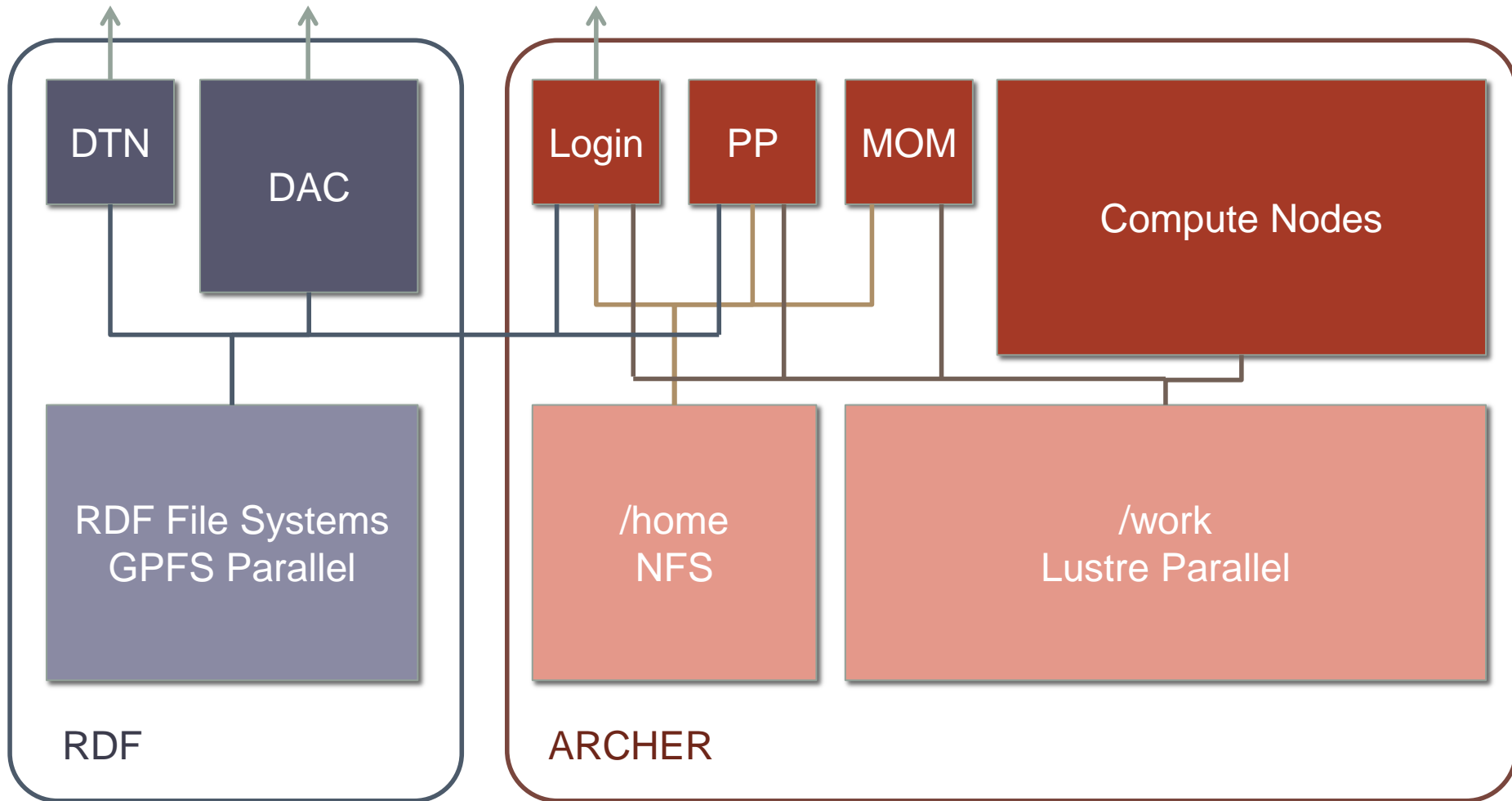
Available File Systems

ARCHER

- /home
  - Standard NFS file system
  - Backed up daily
  - Low-performance, limited space
  - Mounted on: Login, PP, MOM (not Compute Nodes)
- /work
  - Parallel Lustre file system
  - No backup
  - High performance read/write (not open/stat), large space (>4 PiB)
  - Mounted on: Login, PP, MOM, Compute Nodes

RDF

- /epsrc, /nerc, /general
  - Parallel GPFS file system
  - Backed up for disaster recovery
  - High performance (read/write/open/stat), v. large space (>20 PiB)
  - Mounted on: DTN, DAC, Login, PP

Compute Resources

ARCHER

- Compute Nodes:
  - 4920 nodes with 24 cores each (118,080 cores total)
  - 64/128 GB memory per node
  - Designed for parallel jobs (serial not well supported)
  - /work file system only
  - Accessed by batch system only
- PP Nodes
  - 2 nodes with 64 cores each (256 hyperthreads in total)
  - 1 TB memory per node
  - Designed for serial/shared-memory jobs
  - RDF file systems available
  - Access directly or via batch system

RDF

- Data Analytic Cluster
  - 12 standard compute nodes: 40 HyperThreads, 128 GB Memory
  - 2 large compute nodes: 64 HyperThreads, 2 TB Memory
  - Direct Infiniband connections to RDF file systems
  - Access via batch system
  - Designed for data-intensive workloads in parallel or serial

Data Transfer Resources

- ARCHER to/from RDF
  - Primary resource is PP nodes
    - Mounts ARCHER and RDF file systems
  - Interactive data transfer can use ARCHER Login nodes
    - Mounts ARCHER and RDF file systems
    - Small amounts of data only

- To outside world
  - RDF Data Transfer Nodes (DTNs) for large files
  - ARCHER Login Nodes for small amounts of data only

Summary

ARCHER/RDF

- ARCHER and the RDF are separate systems
- Some RDF file systems are mounted on ARCHER login and PP nodes
  - To enable easy data transfer (e.g. for analysis or transfer off site)
- A variety of file systems are available
  - Each has its own use case
  - Data management plan should consider which is best suited at each stage in data lifecycle
- Variety of compute resources available