Small site approaches - Sussex

- Research Computing within ITS established 2010
  - Shared cluster with EPP
  - Lustre (27 TB / 54 TB EPP)
  - 20 nodes (11 ITS / 8 EPP / 1 Informatics)
- Shared management ITS and EPP
- Expanded to other schools / research groups

- Connectivity to JANET 2Gb
- To cluster, multiple 1Gb

- Current cluster located in new (2011) DC
    - QDR Infiniband interconnect
    - 27 x 12 core Westmere
    - 2 x 48 core AMD
    - 16 x 8 core Nehalem
    - 32 x 64 core AMD
    - 32 x 16 core 2.6GHz (Sandy Bridge onwards)
    - 4 GPU nodes (K20, K40)
    - 3 x 20 core E5-2630 v4 2.2GHz
    - 600TB Lustre, 8 OSS 19 OST
    - Univa Grid Engine

- Total physical cores 3168
- Dedicated grid cores 256 (2114 HS06), 80TB Lustre
- Connectivity to JANET 20Gb; to cluster multiple 1Gb

8 x 18kW racks
Water cooled doors

1MW generator
 online UPS

- New cluster
  - Upgrade for mid-scale MPI
    - Cosmology, CFD, Computational Chemistry, fMRI, EM
  - Update fabric with Omni-Path for new nodes
  - Replace 6 year old nodes
  - LNET router to Infiniband for Lustre
  - Bright Cluster Manager, Univa Grid Engine
  - Hadoop on Lustre, Spark
  - HPC nodes Centos 7.3, Grid nodes SL6.8

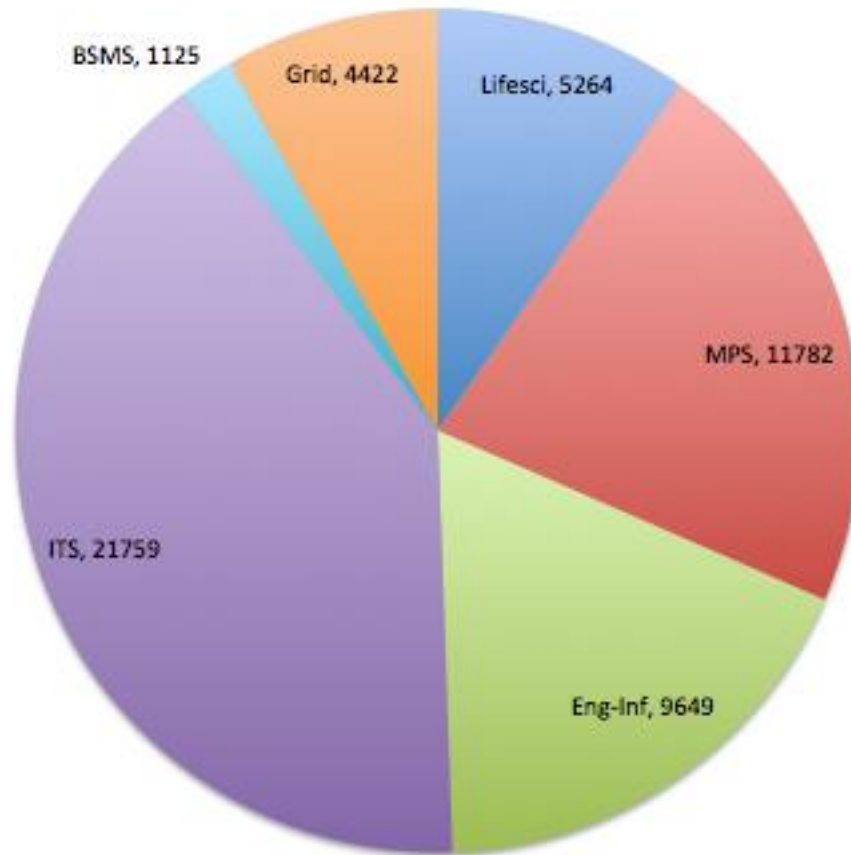 - Connectivity to JANET 20Gb; to cluster multiple 10Gb

- New cluster Centos 7.3
  - Omni-Path interconnect
  - 62 x E5-2640 v3
  - 1 x E5-2690 v4
  - physical cores 1020
- From old cluster
  - 32 x 64 core AMD
  - 32 x 16 core 2.6GHz (Sandy Bridge onwards)
  - 4 GPU nodes
  - 3 x 20 core E5-2630 v4 2.2GHz
  - physical cores 2620
  - 600TB Lustre
- Total cores 3640
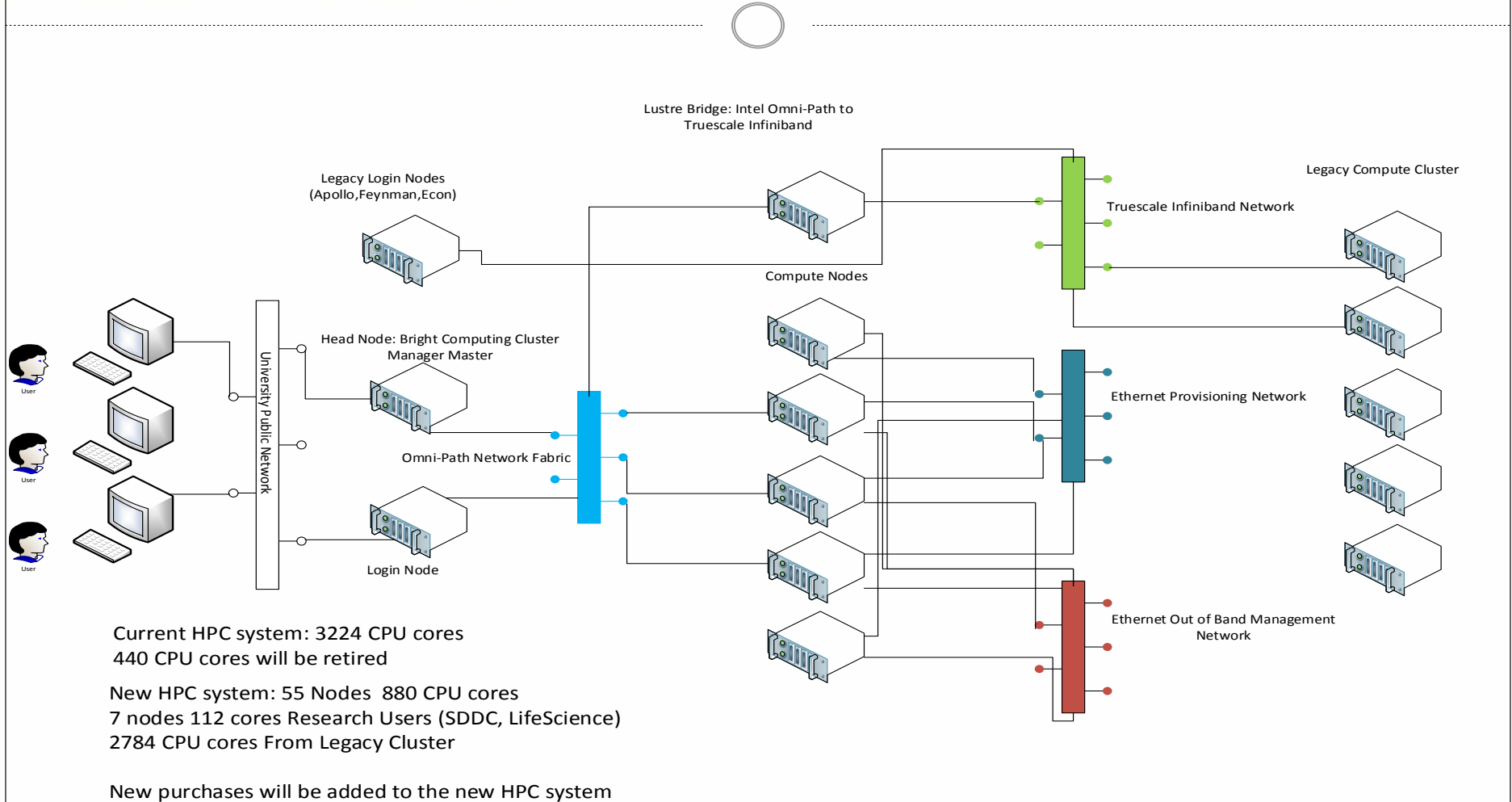- Dedicated grid cores 472 (4421 HS06) on SL6.8

54k HSO6

HS06

BSMS, 1125
Grid, 4422
Lifesci, 5264
MPS, 11782
ITS, 21759
Eng-Inf, 9649

High Performance Computing Cluster Architecture

Lustre Bridge: Intel Omni-Path to Truescale Infiniband

Legacy Login Nodes (Apollo,Feynman,Econ)

Legacy Compute Cluster

Truescale Infiniband Network

Compute Nodes

Head Node: Bright Computing Cluster Manager Master

Omni-Path Network Fabric

University Public Network

Ethernet Provisioning Network

Login Node

Ethernet Out of Band Management Network

User

Current HPC system: 3224 CPU cores
440 CPU cores will be retired

New HPC system: 55 Nodes 880 CPU cores
7 nodes 112 cores Research Users (SDDC, LifeScience)
2784 CPU cores From Legacy Cluster

New purchases will be added to the new HPC system

GridPP

- Maintain Grid Infrastructure
  - Lustre shared file system
  - Move SRM (Storm) from VM to new hardware (10Gbe)
  - Cream, Apel, Argus, Bdii …
  - XRootD
  - Webdav
- VO: ATLAS production, SNO+

- Previous EPP sysadmin (Emyr, Matt) funded by MPS school
- Last year, no EPP sysadmin.
- Grid kept running by ITS Research Computing (3 FTE)

- Now have Leo! + another sysadmin in Astronomy/Theory
- More sysadmin resource so better reliability and scope for development (Monitoring, OpenStack, Ceph etc.)

GridPP – backfill

- Opportunistic use of idle CPU time via Singularity
  https://indico.cern.ch/event/612601/
  - Current utilisation ~ 75%
  - Perhaps ~750 core, 7000 HS06, extra grid resource
  - Pre-emptables jobs best…

- Under used GPU nodes for CMS, LHCb ?

- Maths and Physics storage renewal
    - Replace 400TB obsolete Lustre storage
        - Move to Omni-Path
        - New MDS/MDT
        - 460TB (at least) new OST
        - Probably using ZFS

- Maths and Physics compute renewal
    - Replace AMD 64 core nodes (9k HS06)