Diamond Light Source Site Report

James Thorne

Scientific Computing Team

diamond

Diamond Light Source

Existing Storage

- Lustre03 500 TB
- ~~Lustre04 300 TB~~
  - Decommissioning last week
- GPFS01, 900 TB, 16 GB/sec
- GPFS02, 5 PB, 40GB/sec
- Tape archive at STFC RAL of around 15 PB

diamond

Existing Compute

- Univa Grid Engine, RHEL 7.6
- 3,200 cores
- 190 nodes
- NVIDIA GPUs
  - 40 P100
  - 24 K80
  - 24 M2090
- Mix of single core jobs and multi-node GPU jobs (160 cores / 8 GPUs)
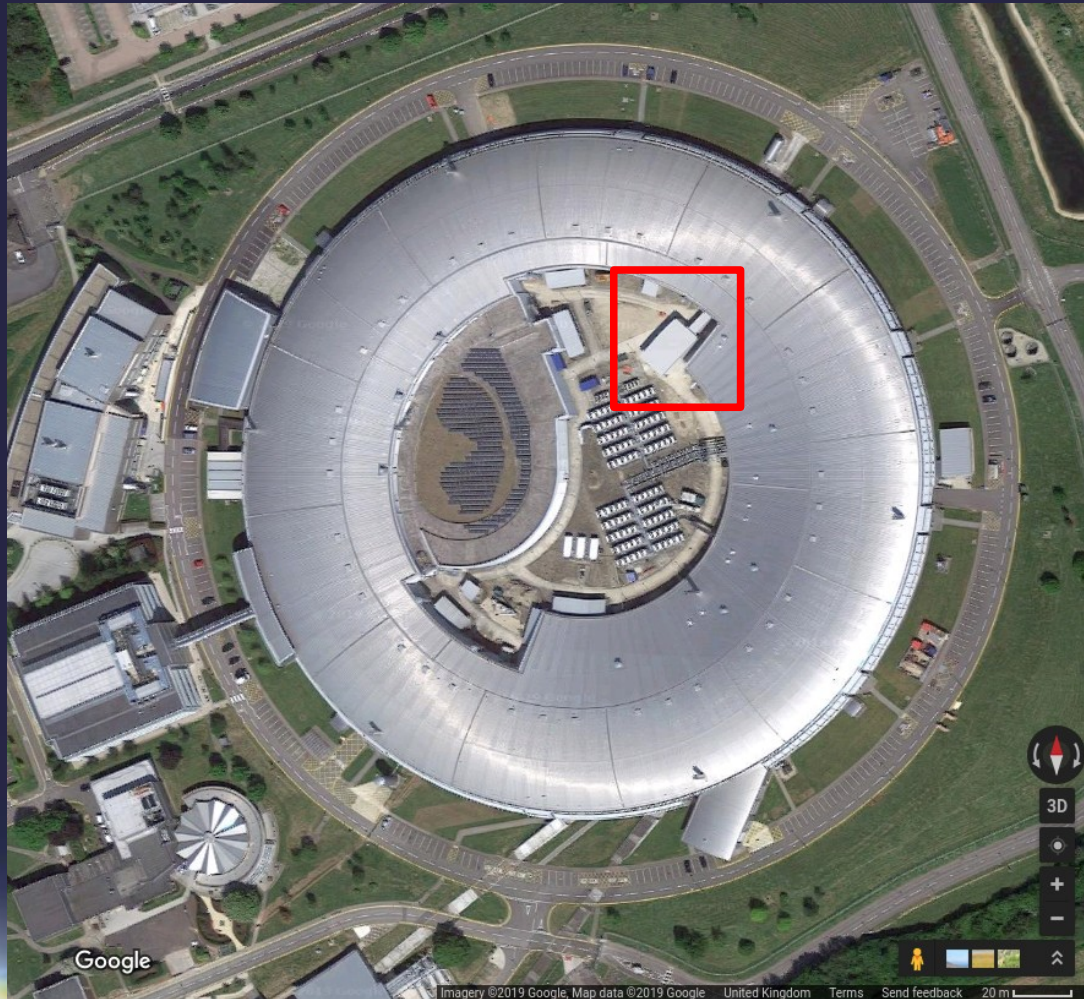- Other small, special purpose clusters

diamond

Networking

- 10 Gb/s and 40 Gb/s core
- 10 Gb/s Internet connection (via RAL)
- 4 x FDR Infiniband between compute nodes and storage

diamond

Management tools

- CFEngine policies in Gitlab
- Cobbler for yum repos
- Grafana
- UGE accounting logs into ELK stack

diamond

New data centre

HAMILTON (compute)

- 64 Lenovo nodes
  - ○ 24 x ThinkSystem SD530s
  - ○ 40 x ThinkSystem SR670s
- 2,560 cores (Xeon Gold 6148)
- 384 GB RAM per node
- 152 NVIDIA P100 GPUs
- 8 NVIDIA V100 32GB GPUs
- "Excelero" scratch space
- 4 x EDR Infiniband

diamond

HAMILTON (storage)

- Lenovo DSS-G240 systems
- 8PB GPFS 5
- 60 GB/sec
- In production use as of this week

diamond

Cloud

- Use cases
  - Post-processing
  - Bursting
  - Accelerator Physics
- Jupyter notebooks
- Kubenetes
- On-premise or STFC OpenStack and Echo
- Public Clouds
- All of the above? None of the Above?

diamond

The future

- The future is cloudy
- Higher data rate detectors
- Diamond II
- More storage
- More compute