

The Revolution in Experimental and Observational Science: The Convergence of Data-Intensive and Compute-Intensive Infrastructure

Tony Hey

Chief Data Scientist

STFC

tony.hey@stfc.ac.uk

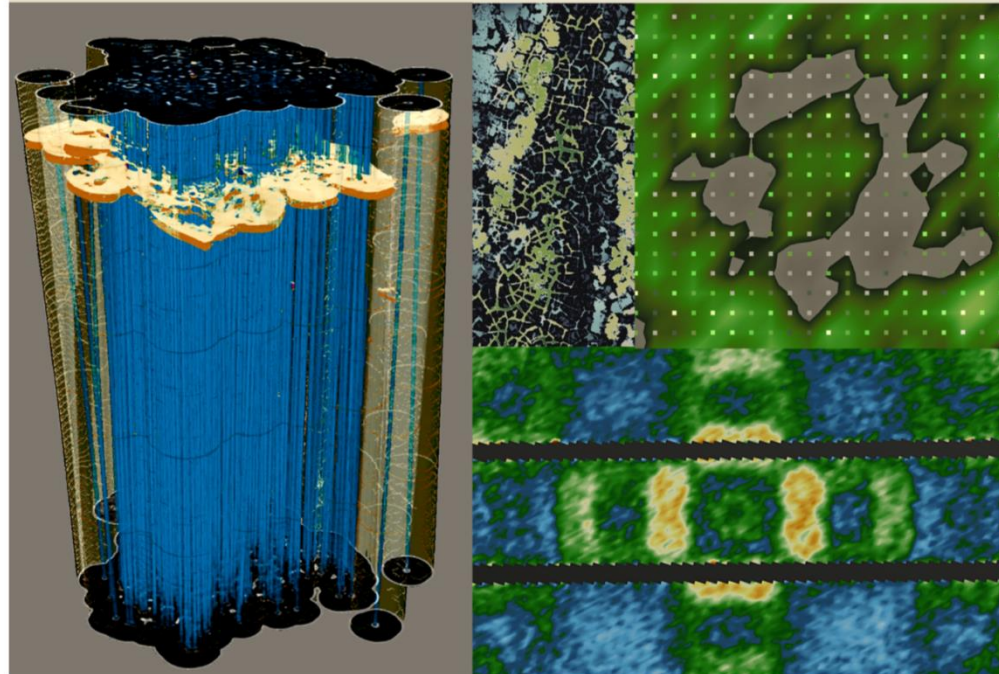
Future Directions for
**NSF ADVANCED
COMPUTING
INFRASTRUCTURE**

to Support U.S. Science
and Engineering
in 2017–2020

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Report of the
DOE Workshop on

**Management,
Analysis, and Visualization of
Experimental and Observational Data**
The Convergence of Data and Computing



U.S. DEPARTMENT OF
ENERGY

Office of
Science

September 29th - October 1, 2015
Bethesda, MD

UK Science and Technology Facilities Council (STFC)

UK Astronomy Technology Centre
Edinburgh, Scotland



Polaris House
Swindon, Wiltshire



Chilbolton Observatory
Stockbridge, Hampshire



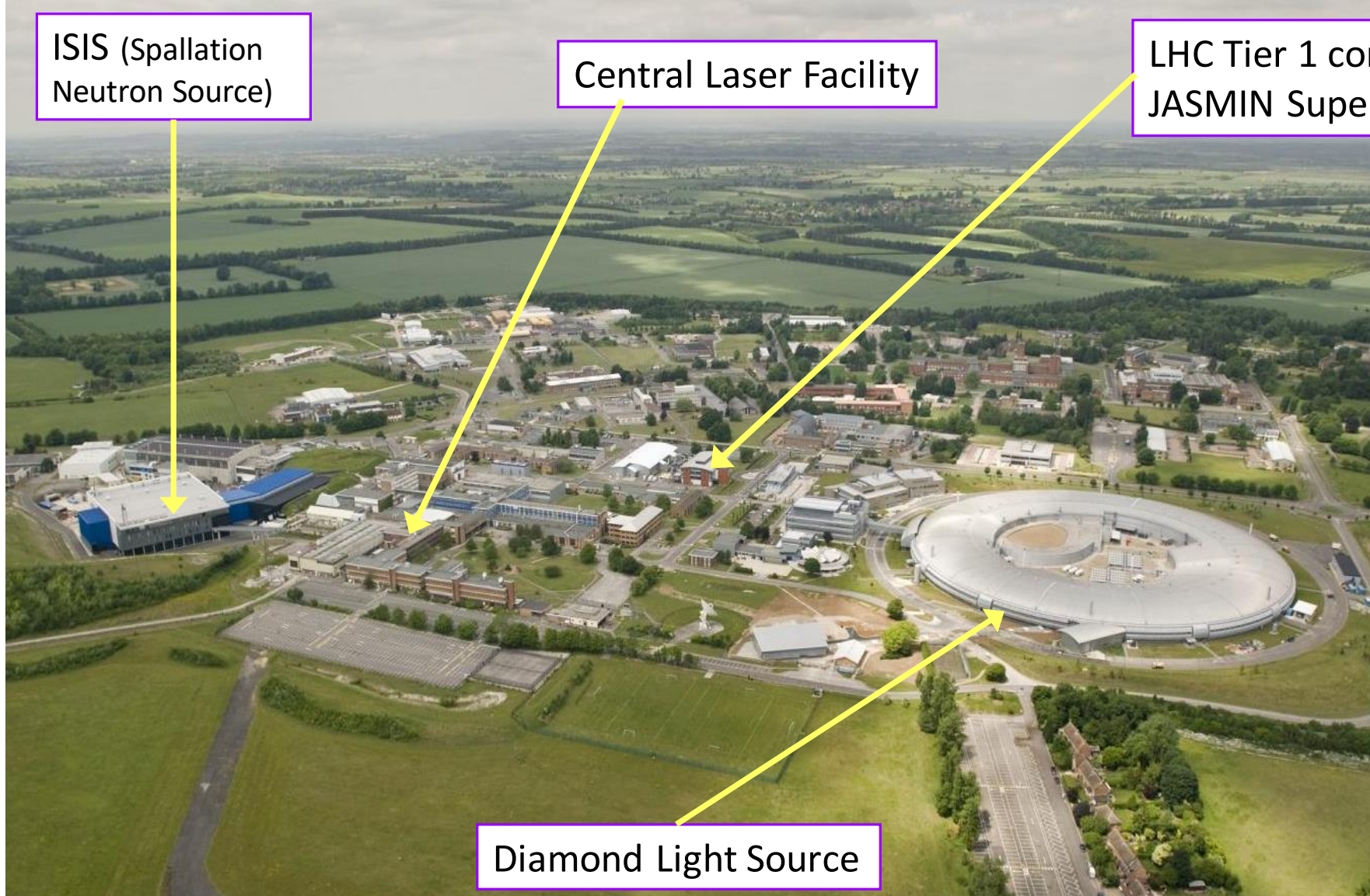
Daresbury Laboratory
Sci-Tech Daresbury Campus
Warrington, Cheshire



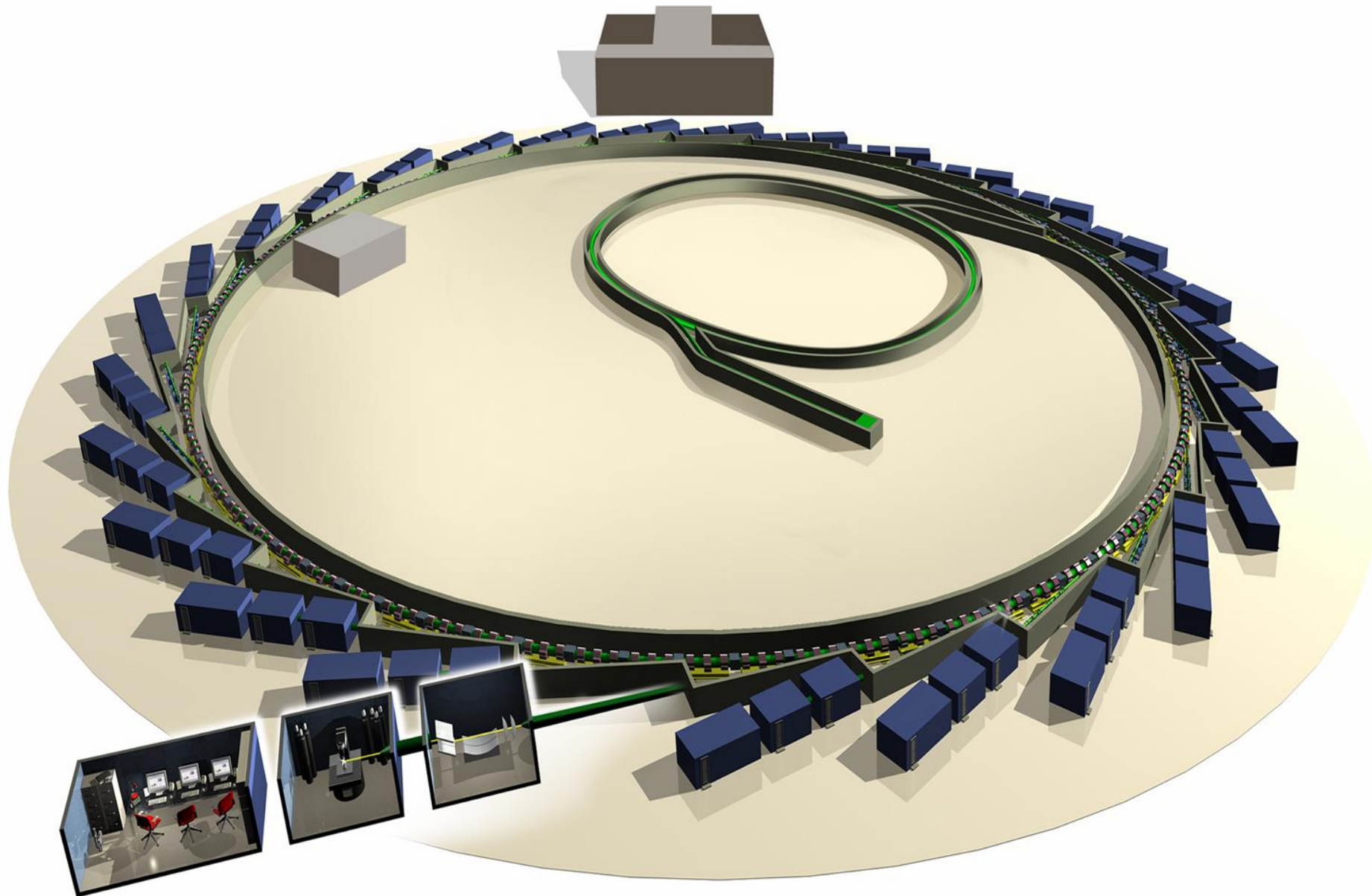
Rutherford Appleton Laboratory
Harwell Science and Innovation Campus
Didcot, Oxfordshire



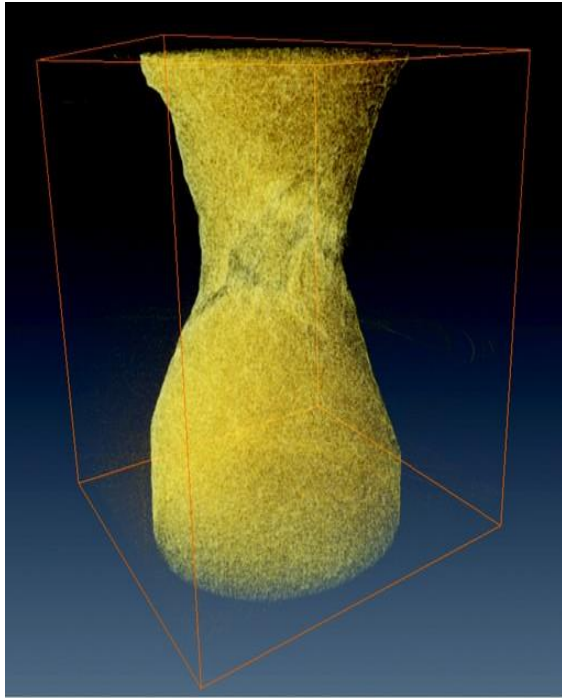
Rutherford Appleton Lab and the Harwell Campus



Diamond Light Source



Science Examples



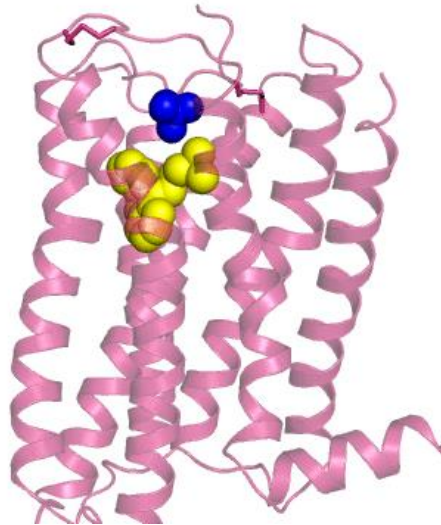
Casting aluminium



**Pharmaceutical
manufacture & processing**

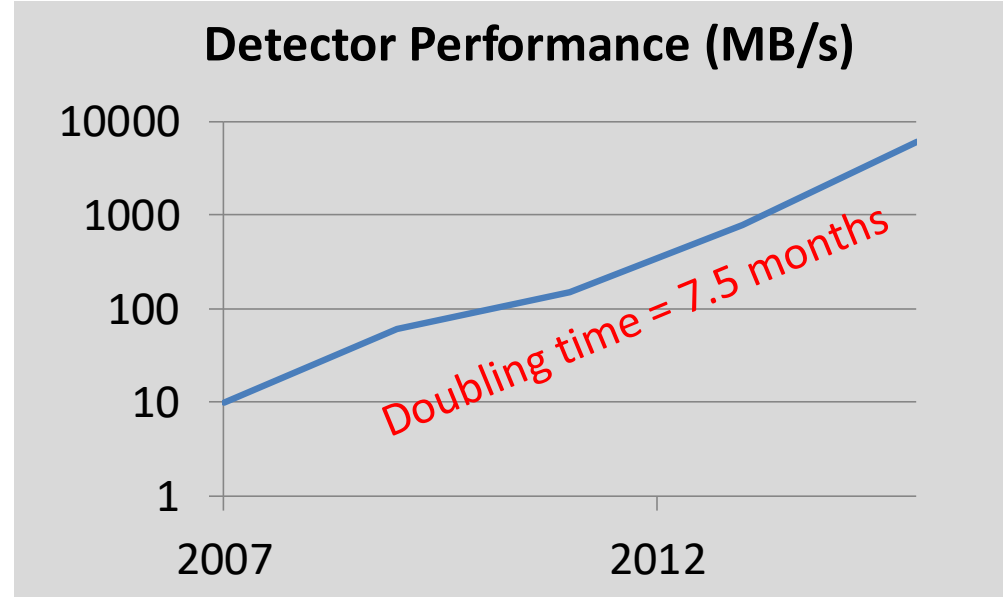


**Non-destructive imaging of
fossils**



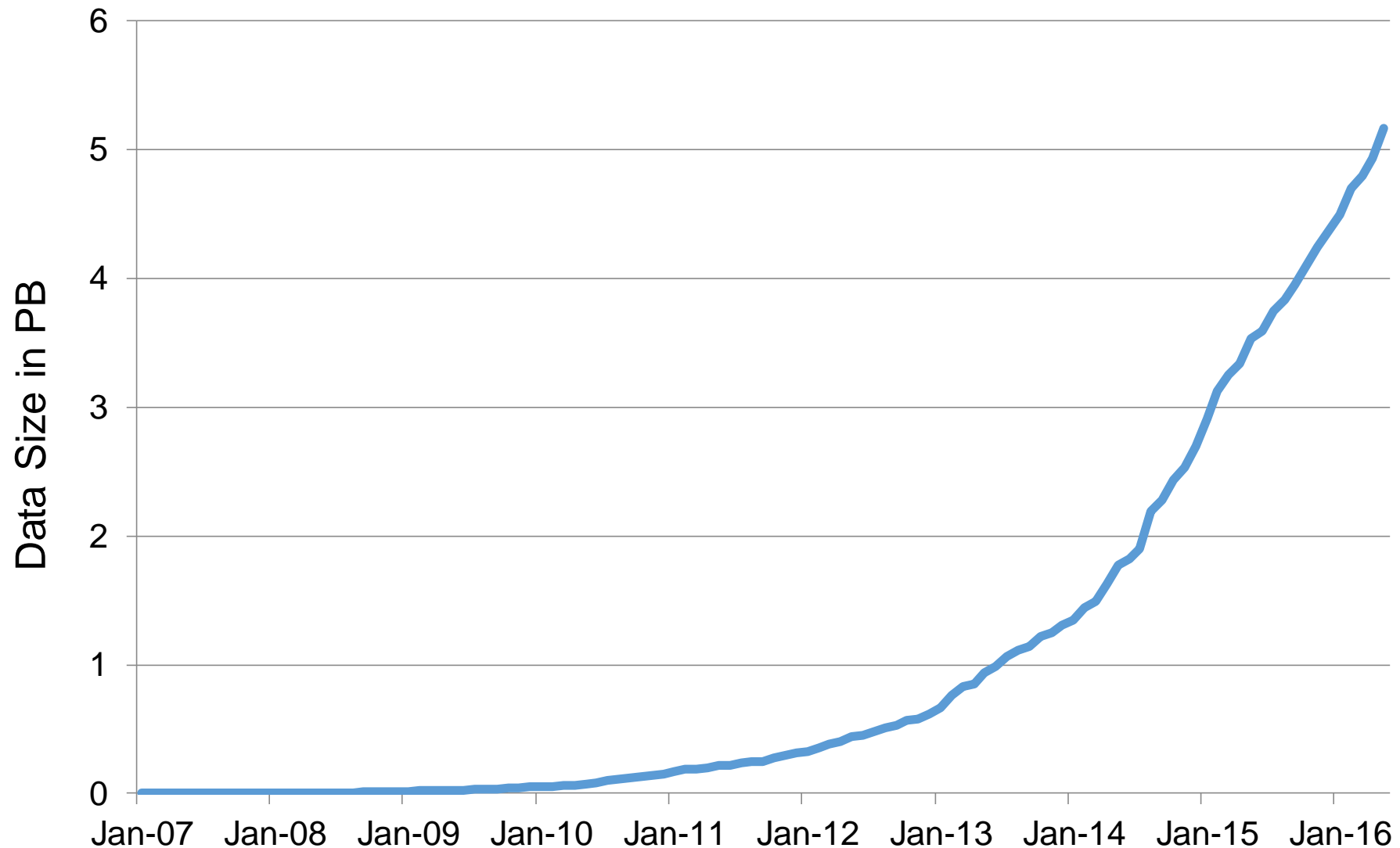
**Structure of the Histamine
H1 receptor**

Data Rates

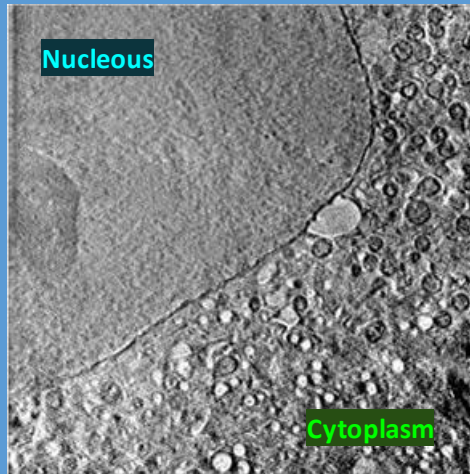


- 2007 No detector faster than ~10 MB/sec
- 2009 Pilatus 6M system 60 MB/s
- 2011 25Hz Pilatus 6M 150 MB/s
- 2013 100Hz Pilatus 6M 600 MB/sec
- 2013 ~10 beamlines with 10 GbE detectors (mainly Pilatus and PCO Edge)
- 2016 Percival detector 6GB/sec

Cumulative Amount of Data Generated By Diamond



Cryo-SXT Data



Neuronal-like mammalian cell line; single slice

Challenges:

- Noisy data, missing wedge artifacts, missing boundaries
- Tens to hundreds of organelles per dataset
- Tedious to manually annotate
- Cell types can look different
- Few previous annotations available
- Automated techniques usually fail

scientificsoftware@diamond.ac.uk

Segmentation of Cryo-soft X-ray Tomography (Cryo-SXT) data

Data

- **B24:** Cryo Transmission X-ray Microscopy beamline at DLS
- Data Collection: Tilt series from $\pm 65^\circ$ with 0.5° step size
- Reconstructed volumes up to $1000 \times 1000 \times 600$ voxels
- Voxel resolution: $\sim 40\text{nm}$ currently
- Total depth: up to $10\mu\text{m}$
- **GOAL:** Study structure and morphological changes of whole cells



B24 beamline
Data Analysis Software Group

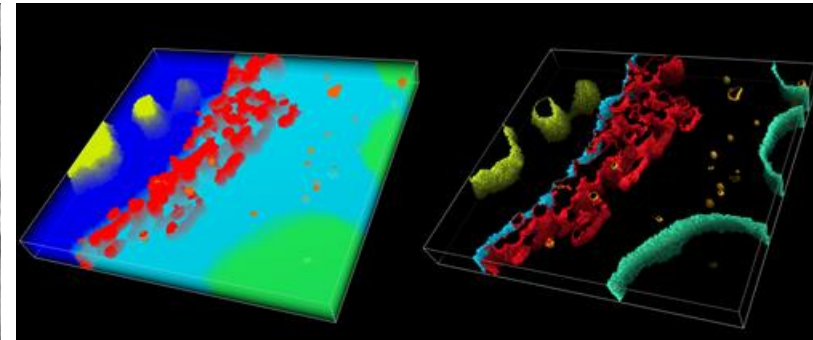
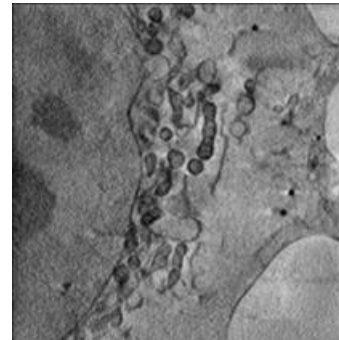


The University of
Nottingham

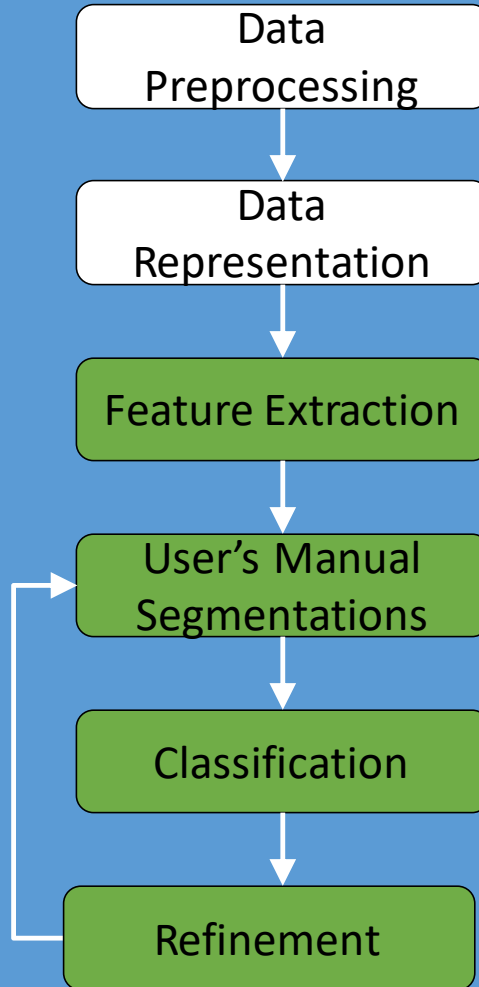
Computer Vision
Laboratory

3D Volume Data

Segmentation

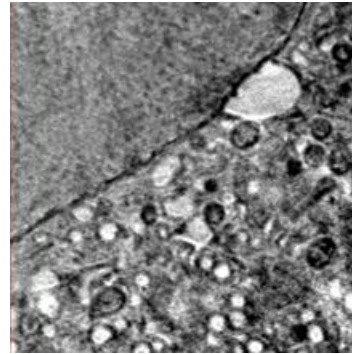


Workflow

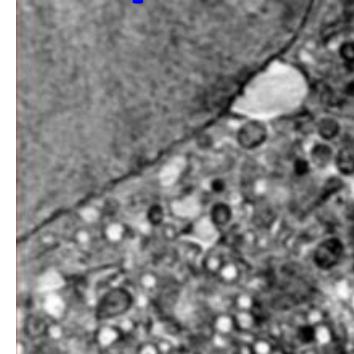


scientificsoftware@diamond.ac.uk

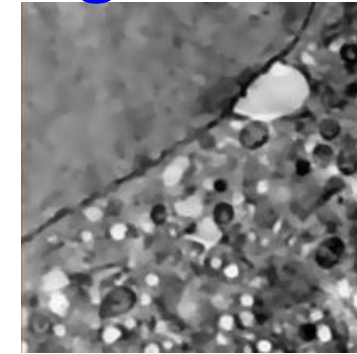
Data Preprocessing



Raw Slice

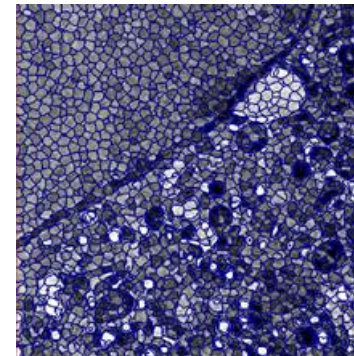
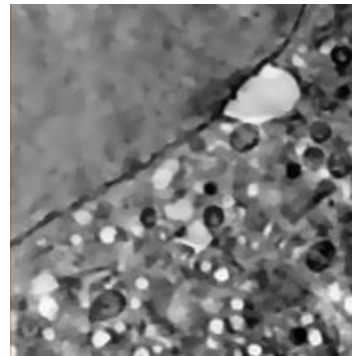


Gaussian Filter

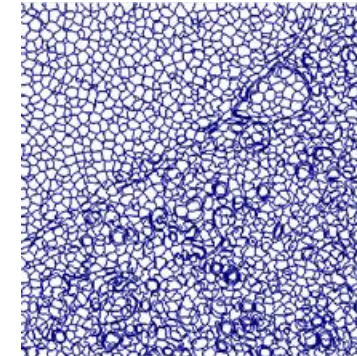


Total Variation

Data Representation



SuperVoxels (SV)

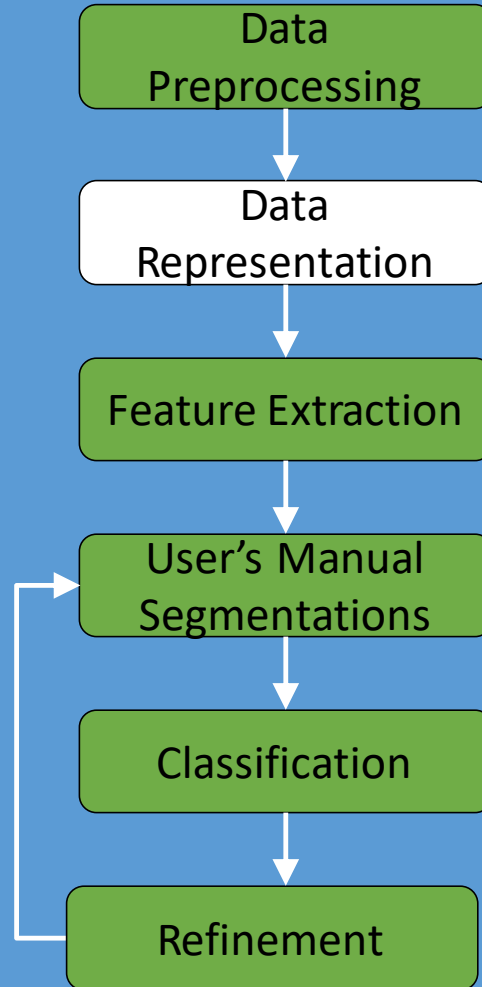


SV Boundaries

SuperVoxels:

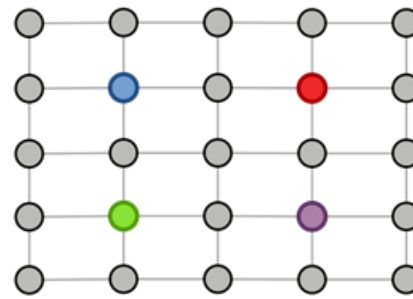
- Groups of similar and adjacent voxels in 3D
- Preserve volume boundaries
- Reduce noise when representing data
- Reduce problem complexity several orders of magnitude
- Use Local clustering in $\{xyz + \lambda * intensity\}$ space

Workflow

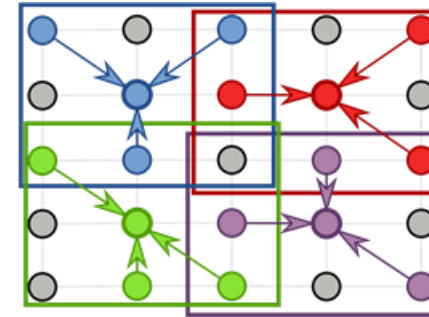


scientificsoftware@diamond.ac.uk

Data Representation

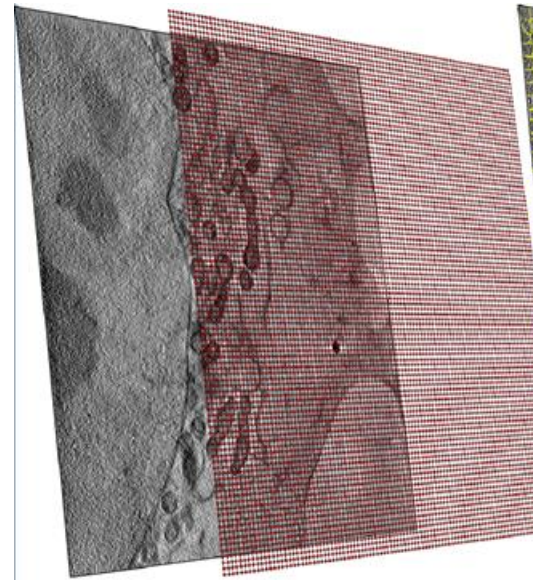


Initial Grid with uniformly sampled seeds



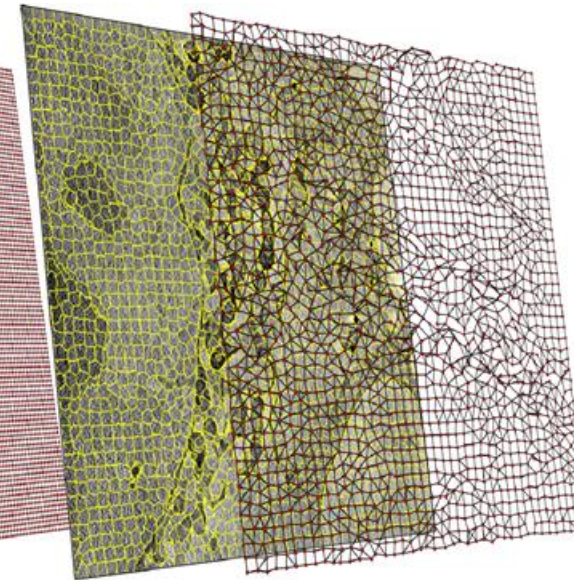
Local k -means in a small window around seeds

Voxel Grid



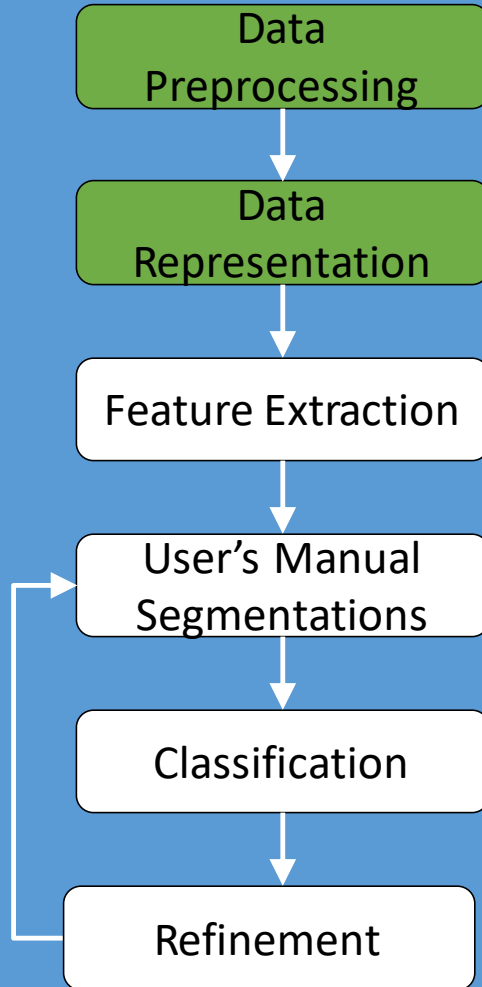
$946 \times 946 \times 200 = 180\text{M}$ voxels

Supervoxel



$180\text{M} / (10 \times 10 \times 10) = 180\text{K}$ supervoxels

Workflow



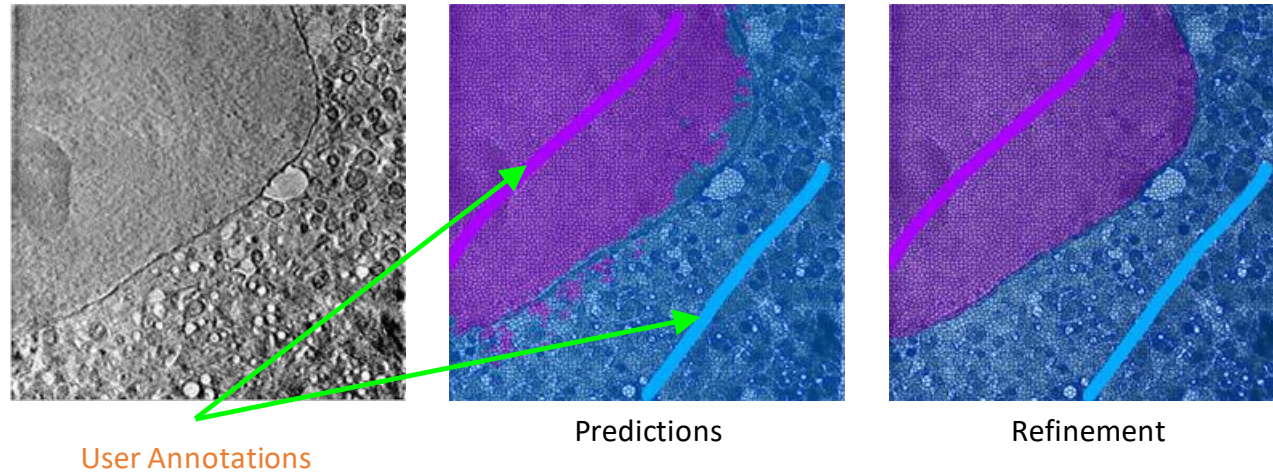
scientificsoftware@diamond.ac.uk

Feature Extraction

Features are extracted from voxels to represent their appearance:

- Intensity-based filters (Gaussian Convolutions)
- Textural filters (eigenvalues of Hessian and Structure Tensor)

User Annotation + Machine Learning

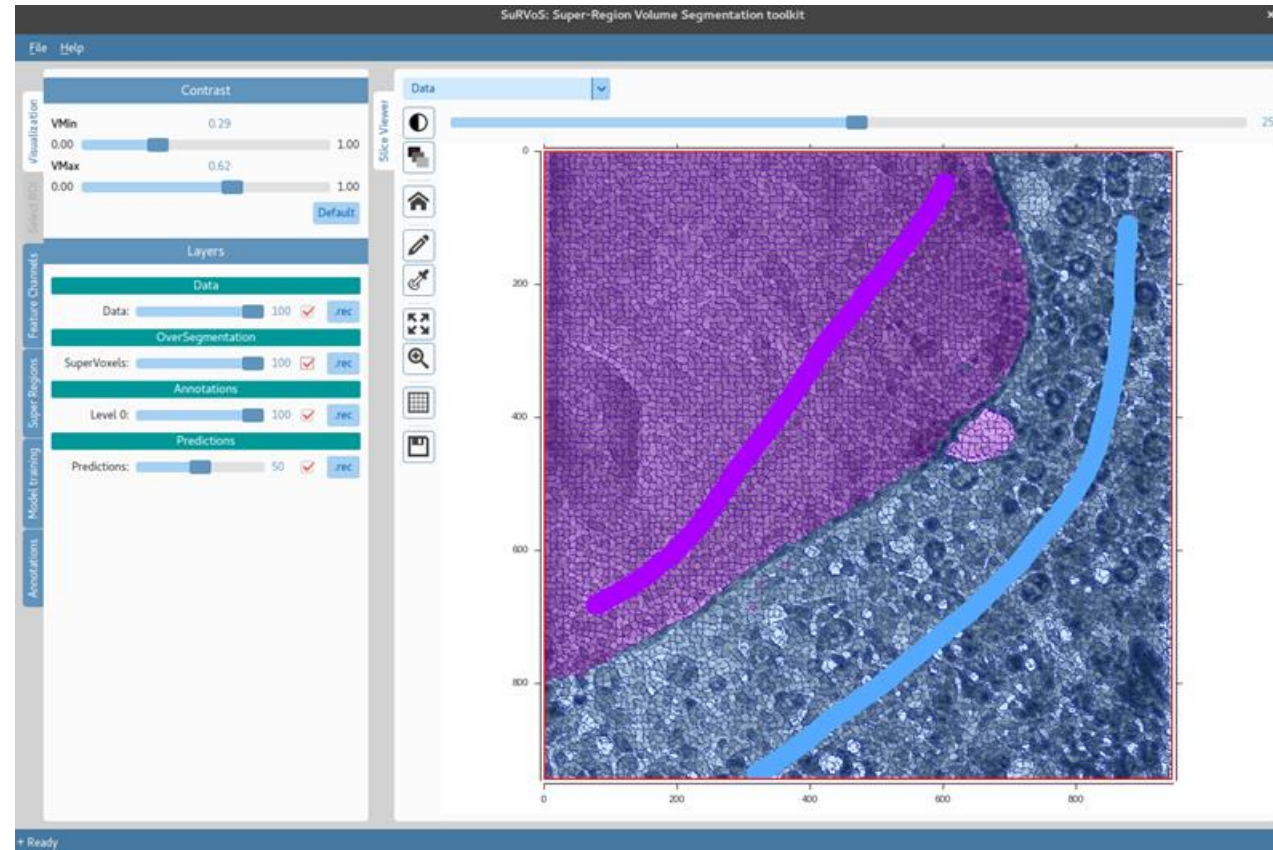


Using a few user annotations along the volume as an input:

- A machine learning classifier (i.e. Random Forest) is trained to discriminate between different classes (i.e. Nucleus and Cytoplasm) and predict the class of each SuperVoxel in the volume.
- A Markov Random Field (MRF) is then used to refine the predictions.

SuRVoS Workbench

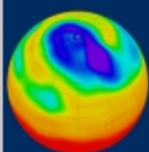
(Su)per-(R)egion (Vo)lume (S)egmentation



Coming soon: <https://github.com/DiamondLightSource/SuRVoS>

scientificsoftware@diamond.ac.uk

Imanol Luengo <imanol.luengo@nottingham.ac.uk>, Michele C. Darrow, Matthew C. Spink, Ying Sun, Wei Dai, Cynthia Y. He, Wah Chiu, Elizabeth Duke, Mark Basham, Andrew P. French, Alun W. Ashton



Centre for Environmental Data Archival

SCIENCE AND TECHNOLOGY FACILITIES COUNCIL
NATURAL ENVIRONMENT RESEARCH COUNCIL



[Contact Us](#) [CEDA News](#)

[► Home](#) [► Data Centres](#)

Search

[About CEDA](#)

[Data Centres](#)

[Services](#)

[Projects](#)

[For Academics](#)

[For Business](#)

[Training](#)

[Help](#)

[Contact Us](#)

Data Centres

The Centre for Environmental Data Archival is responsible for the running of the following data centres:



British Atmospheric Data Centre
NATIONAL CENTRE FOR ATMOSPHERIC SCIENCE
NATURAL ENVIRONMENT RESEARCH COUNCIL

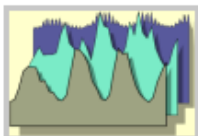
The British Atmospheric Data Centre

The British Atmospheric Data Centre (BADC), NERC's designated data centre for the UK atmospheric science community, covering climate, composition, observations and NWP data.



NERC Earth Observation Data Centre

The NEODC is NERC's designated data centre for Earth Observation data and is part of NERC's National Centre for Earth Observation.



The UK Solar System Data Centre

The UK Solar System Data Centre, co-funded by STFC and NERC, curates and provides access to archives of data from the upper atmosphere, ionosphere and Earth's solar environment.



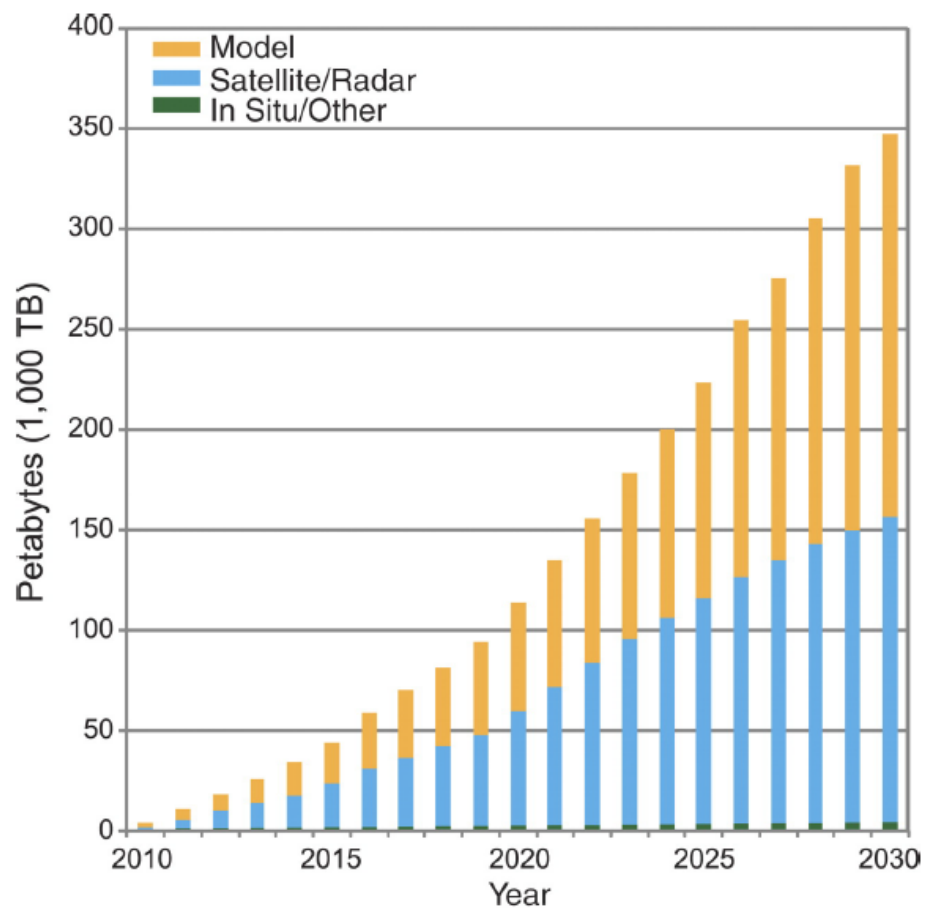
IPCC Data Distribution Centre

The **Intergovernmental Panel on Climate Change (IPCC)** DDC provides climate, socio-economic and environmental data, both from the past and also in scenarios projected into the future. Technical guidelines on the selection and use of different types of data and scenarios in research and assessment are also provided.

More Data

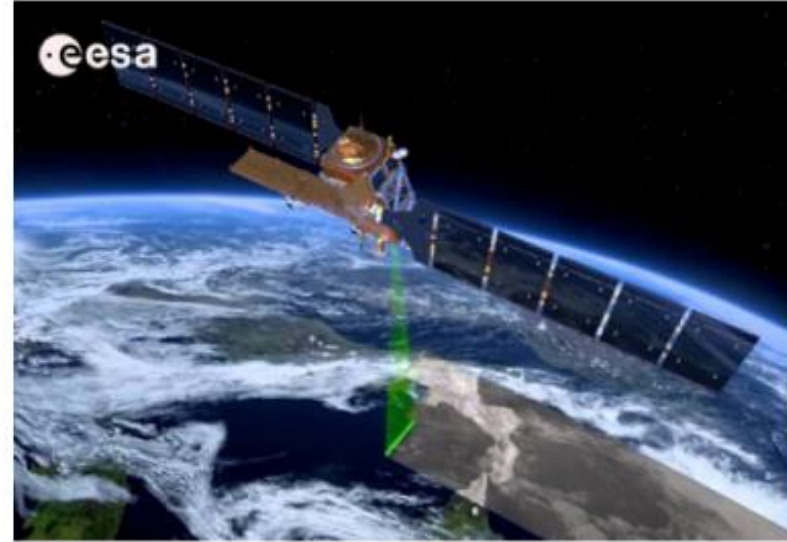
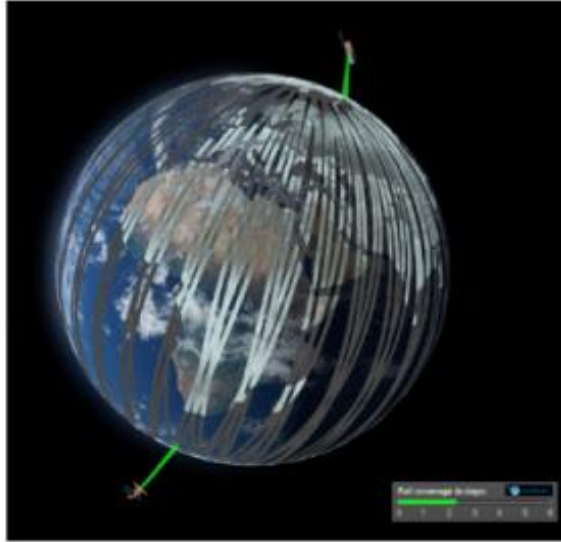
Fig. 2 The volume of worldwide climate data is expanding rapidly, creating challenges for both physical archiving and sharing, as well as for ease of access and finding what's needed, particularly if you're not a climate scientist.

(BNL: Even if you are?)



J T Overpeck et al. Science 2011;331:700-702

Large data sets: satellite observations

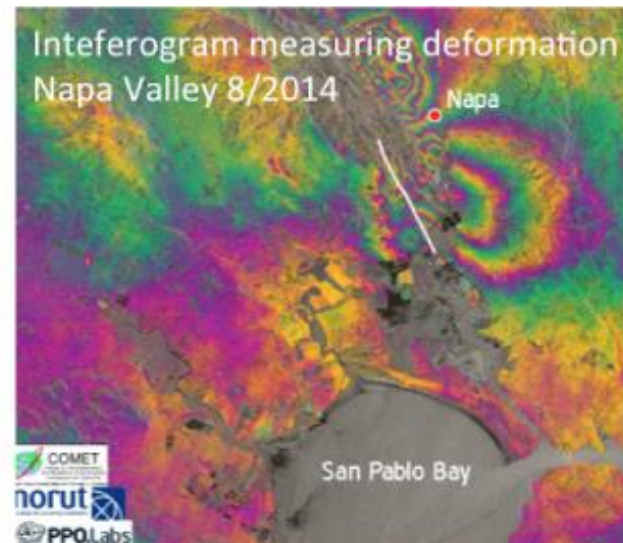


Sentinel 1A: Launched 2014 (1B due 2016)

- Key instrument: Synthetic Aperture Radar
- Data rate (two satellites: raw 1.8 TB/day, archive products ~ 2 PB/year)



COMET: Centre for Observation and Modelling of
Earthquakes, Volcanoes, and Tectonics



(Picture credits: ESA, Arianespace.com, PPO.labs-Norut-COMET-SEOM Insarap study, ewf.nerc.ac.uk/2014/09/02/new-satellite-maps-out-napa-valley-earthquake/)

Core Science Requirements



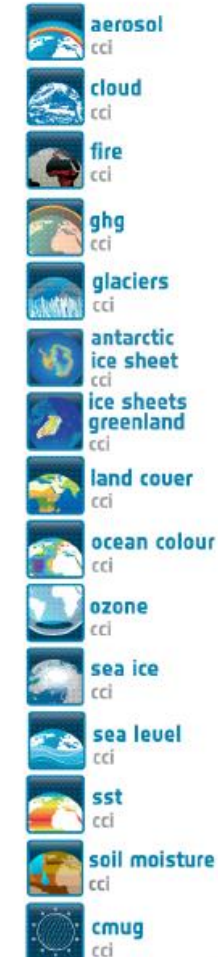
Today:	Observations	Models
Volume	20 million = 2×10^7	5 million grid points 100 levels 10 prognostic variables = 5×10^9
Type	98% from 60 different satellite instruments	physical parameters of atmosphere, waves, ocean
Soon:	Observations	Models
Volume	200 million = 2×10^8	500 million grid points 200 levels 100 prognostic variables = 1×10^{13}
Type	98% from 80 different satellite instruments	physical and chemical parameters of atmosphere, waves, ocean, ice, vegetation

→ Factor 10 per day

→ Factor 2000 per time step

→ but many more time steps needed

Big International Drivers:



Why JASMIN?

- Urgency to provide better environmental predictions
- Need for higher-resolution models
- HPC to perform the computation
- Huge increase in observational capability/capacity

But...

- Massive storage requirement: observational data transfer, storage, processing
- Massive raw data output from prediction models
- Huge requirement to process raw model output into usable predictions (post-processing)

Hence JASMIN...



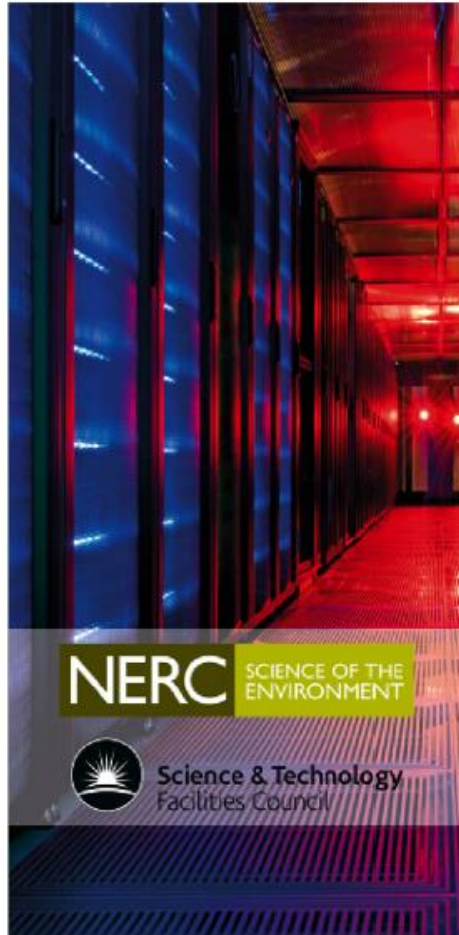
ARCHER supercomputer (EPSRC/NERC)



JASMIN (STFC/Stephen Kill)

JASMIN infrastructure

Part data store, part HPC cluster, part private cloud...



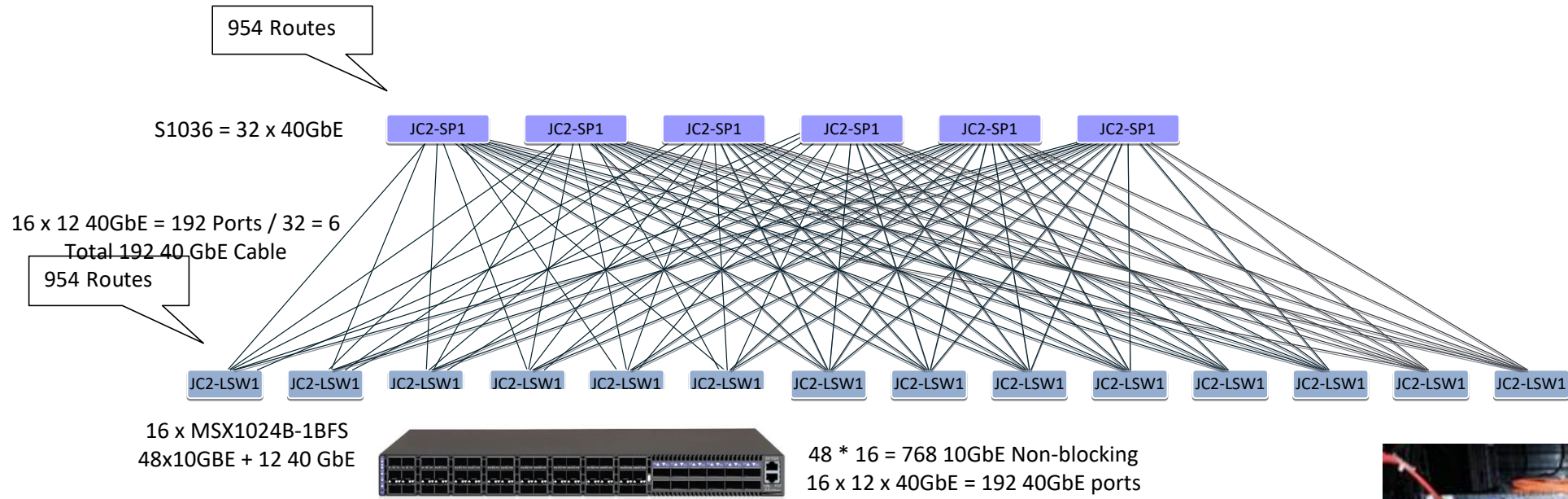
- ▶ 16 PB Fast Storage
(Panasas, many Tbit/s bandwidth)
- ▶ 1 PB Bulk Storage
- ▶ Elastic Tape
- ▶ 4000 cores: half deployed as hypervisors, half as the "Lotus" batch cluster.
- ▶ Some high memory nodes, a range, bottom heavy.



Some JASMIN Statistics

- 16 PetaBytes useable high performance spinning disc
- Two largest Panasas 'realms' in the world (109 and 125 shelves).
- 900TB useable (1.44PB raw) NetApp iSCSI/NFS for virtualisation + Dell Equallogic PS6210XS for high IOPS low latency iSCSI
- 5,500 CPU cores split dynamically between batch cluster and cloud/virtualisation (VMware vCloud Director and vCenter/vSphere)
- 40 Racks
- >3 Tera bits per second bandwidth. IO Capability of ~250GBytes/sec
- "hyper" converged network infrastructure - 10GbE + MPI low latency (~8uS) + iSCSI over same network fabric. (No separate SAN or Infiniband)

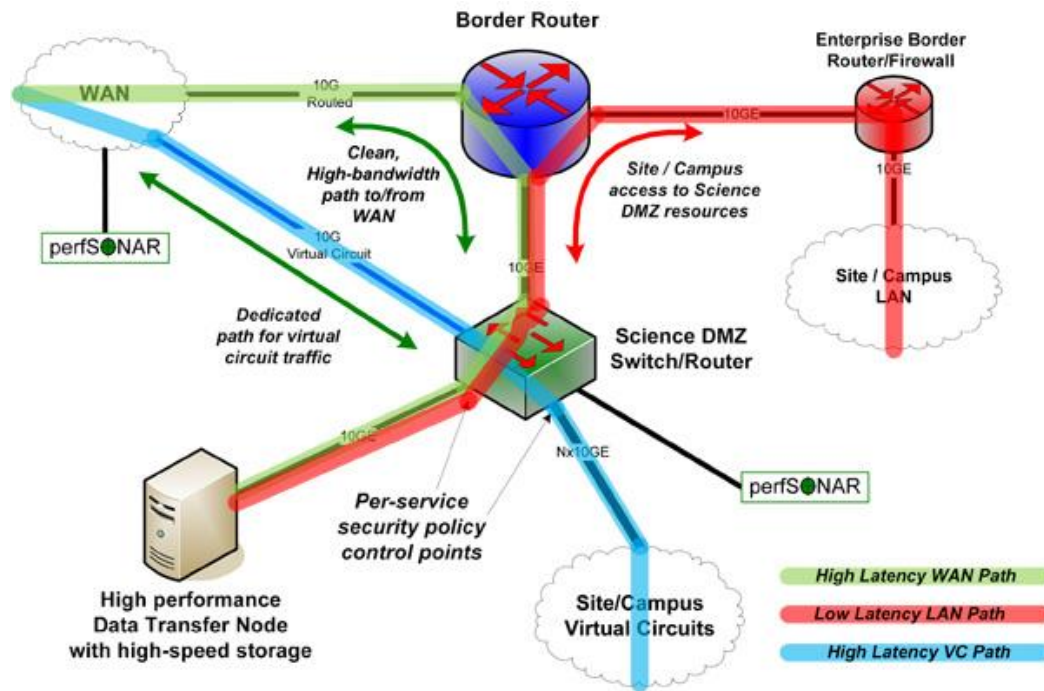
Non-blocking, low latency, CLOS Tree Network



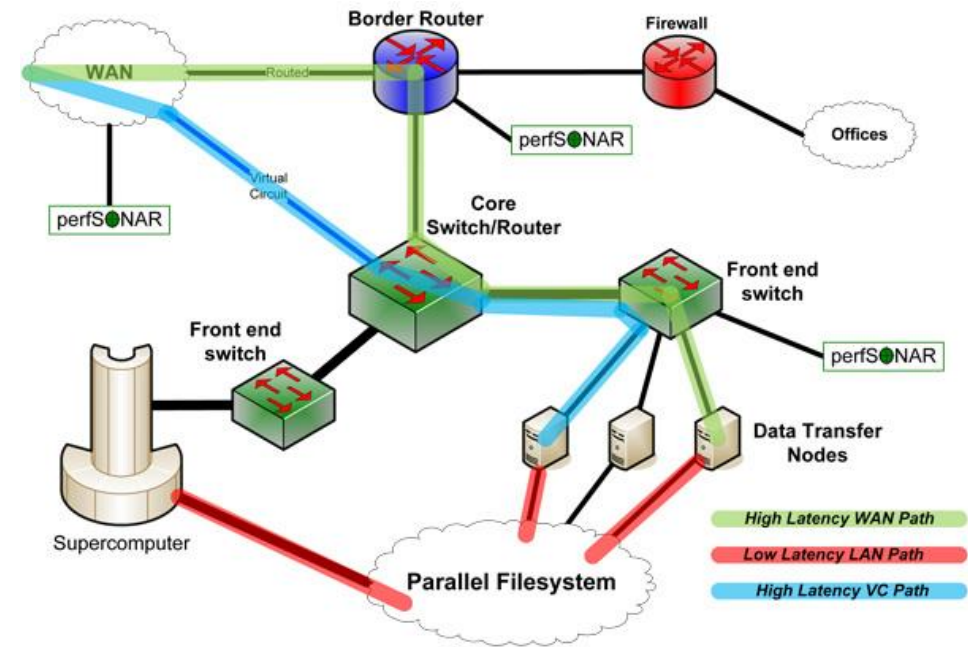
1,104 x 10GbE Ports CLOS L3 ECMP OSPF

- ~1,200 Ports expansion
- Max 36 leaf switches :1,728 Ports @ 10GbE
- Non-Blocking, Zero Contention (48x10Gb = 12x 40Gb uplinks)
- Low Latency (250nS L3 / per switch/router) 7-10uS MPI

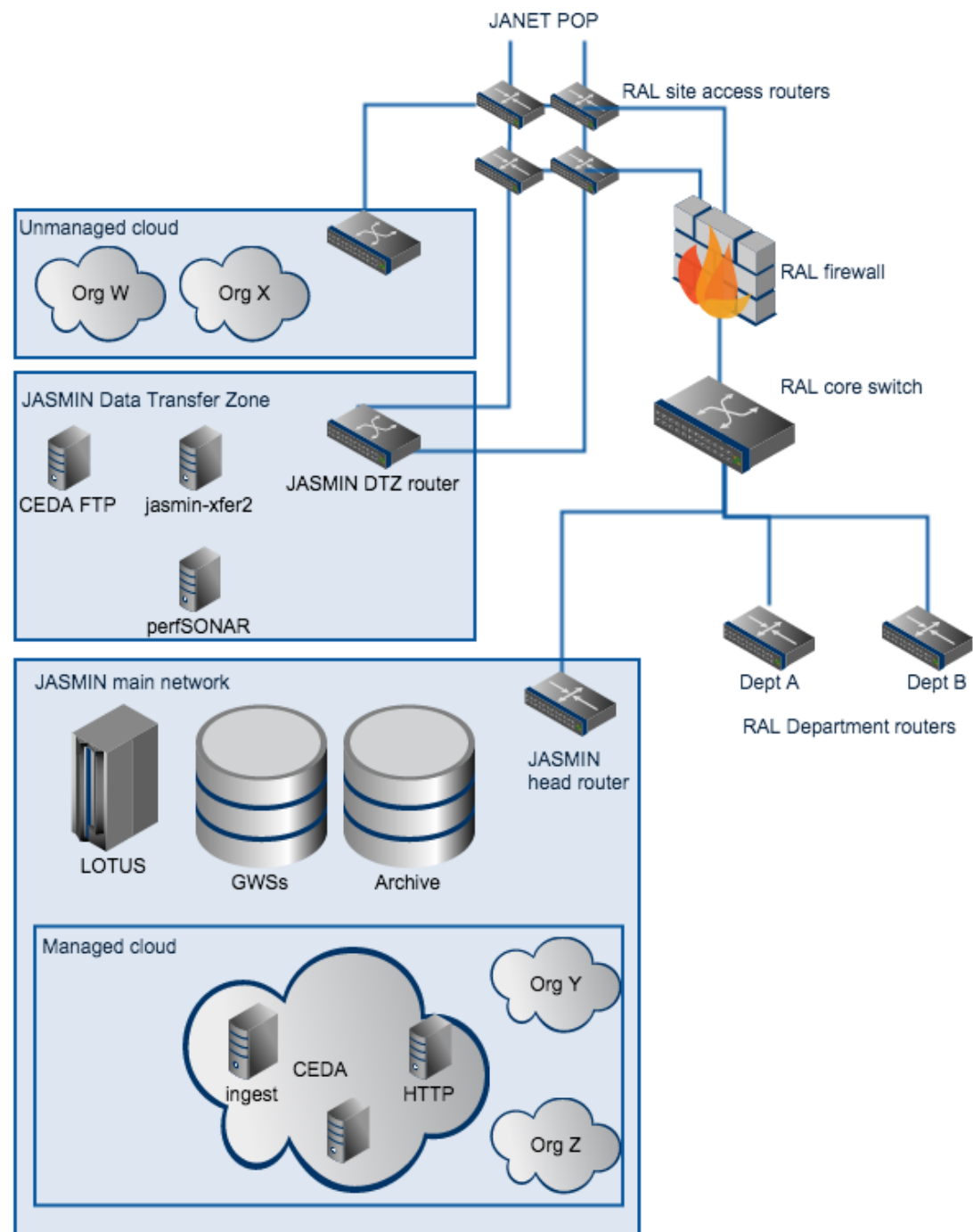
JASMIN “Science DMZ” Architecture



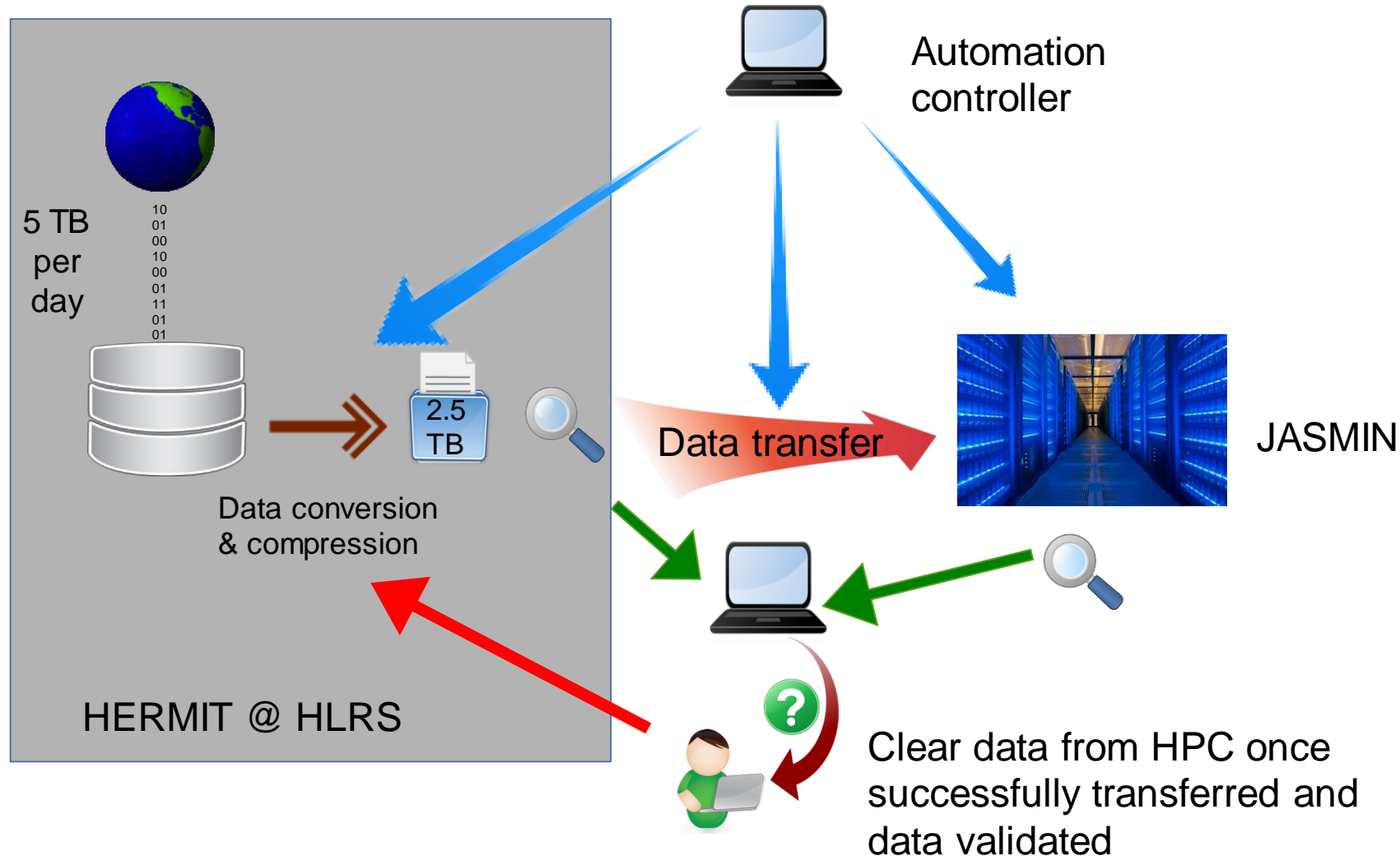
Simple Science DMZ



Supercomputer Center

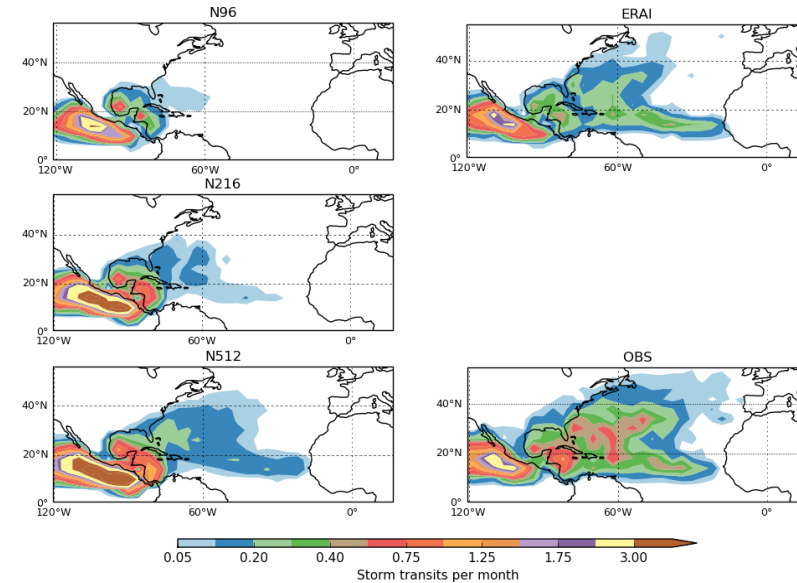


The UK Met Office UPSCALE campaign



Example Data Analysis

- Tropical cyclone tracking has become routine; 50 years of N512 data can be processed in 50 jobs in one day
- Eddy vectors; analysis we would not attempt on a server/workstation (total of 3 months of processor time and ~40 GB memory needed) completed in 24 hours in 1,600 batch jobs
- JASMIN/LOTUS combination has clearly demonstrated the value of cluster computing to data processing and analysis.



The Experimental Data Challenge

- Data rates are increasing, facilities science more data intensive
 - Handling and processing data has become a bottleneck to produce science
 - Need to compare with complex models and simulations to interpret the data
- Computing provision at home-institution highly variable
 - Consistent access to HTC/HPC to process and interpret experimental data
 - Computational algorithms more specialised
 - More users without the facilities science background
- Need access to data, compute and software services
 - Allow more timely processing of data
 - Use of HPC routine not “tour de force”
 - Generate more and better science
- Need to provide within the facilities infrastructure
 - Remote access to common provision
 - Higher level of support within the centre
 - Core expertise in the computational science
 - More efficient than distributing computing resources to individual facilities and research groups

The Experimental Data Challenge

- Data rates are increasing, facilities science more data intensive
 - Handling and processing data has become a bottleneck to produce science
 - Need to compare with complex models and simulations to interpret the data
- Computing provision at home-institution highly variable
 - Consistent access to HTC/HPC to process and interpret experimental data
 - Computational algorithms more specialised
 - More users without the facilities science background
- Need access to data, compute and software services
 - Allow more timely processing of data
 - Use of HPC routine not “tour de force”
 - Generate more and better science
- Need to provide within the facilities infrastructure
 - Remote access to common provision
 - Higher level of support within the centre
 - Core expertise in the computational science
 - More efficient than distributing computing resources to individual facilities and research groups



The ALC will significantly enhance our capability to support the Facilities' science programme:

- *Theme 1: Capacity in advanced software development for data analysis and interpretation*
- *Theme 2: A new generation of data experts and software developers, and science domain experts*
- *Theme 3: Compute infrastructure, for managing, analysing and simulating the data generated by the facilities and for designing next generation Big-Science experiments*

➤ Focused on the science drivers and computational needs of Facilities