Accelerators

Misc

- Using AI – Access to open source models in house.
- Don't want a platform centric approach learn concepts and approaches not 'just' platform – portability (& how to do when people are mission driven?)

Running Technical Tenders

Management

- Procurement
  - ~~Running technical tenders~~
  - Getting added value from suppliers
  - Client conversations
- How to communicate what you do and offer
- Funding models

RSE Co-design sessions between Users, RSE's and RIE's

- RSE vs RIE who does what – containers?
- Support for code porting
- Code optimisation
- Efficient use of GPU's

Terminology & Management Speak

Meeting Vendors' Techies

Accelerator/Vendor Specific Training
(e,g AMD/Nvidia/Intel GPUs, TPUs, FPGAs, RISC-V, and ASICs)

- Software stack and drivers
- Scheduling – Slurm, condor, Kubernetes/Kubeflow, RunAI
- Containers
- Licencing conditions
- Recognising/debugging hardware failure
- Access to hardware sandpits

DPU's

User Training

Multiple Tenancy GPU's

- Overview of available options e.g MIG
- Security implications
- Scheduling
- Containers

GPU Slicing

RDMA

- Overview of available options e.g GPU Direct
- Security implications
- Hardware requirements (networking/storage)

Networking -> RDMA

Overview of Accelerators

- GPU vs CPU architecture
- Comparison of GPUs Nvidia/AMD/Intel
- Superchips (e.g GH200 or BH200) vs regular GPU's
- Other accelerators TPU's FPGA's ASIC's Quantum?
- Key use cases in HPC and AI and specific applications.
- Advantages, disadvantages of each
  - Technical novelty
  - Technical cost
  - Cost (upfront and TCO)
  - Sustainability
- When and how to progress from GPU's
- Horizon Scanning/Emerging Tech – what's coming next and why
- Jargon buster
- Gotcha's
  - licencing conditions
- Impact on Networking
- Impact on Scheduling

Profiling Tools -> Scheduling

Coding

Maintaining accelerated compute

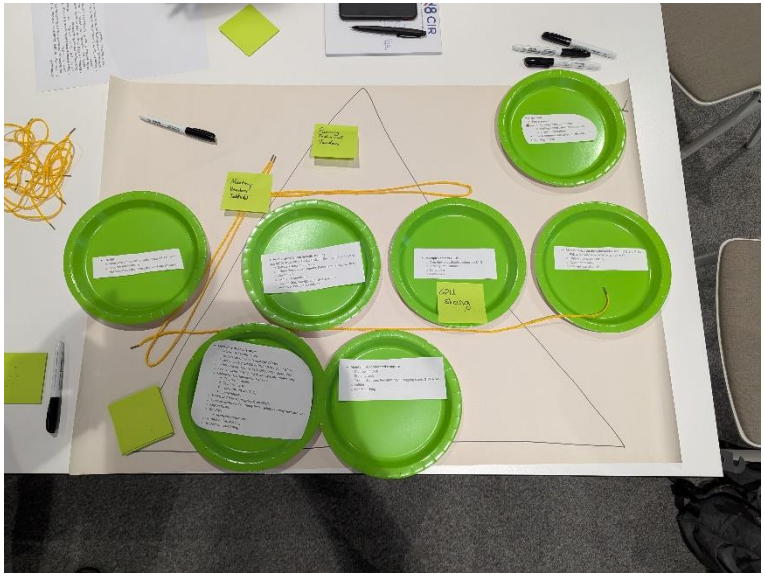
- Monitoring tools
- Profiling tools
- Troubleshooting: Recognising/debugging issues (HW & SW)
- Validating
- Benchmarking

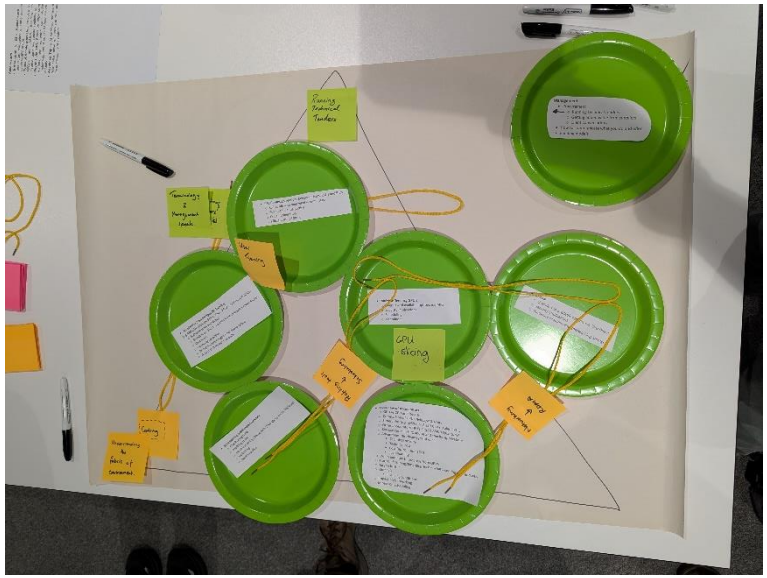
Understanding the fabric of the environment

Running Basic MPI/Slurm Job

User first information / user notes

Accelerators – Photo of progression through groups

1. 

2. 

3. 

4. 

Storage and Research Data Management (RDM)

RDE/Data steward

RIE

Management
- Data retention policy
- Data storage funding models
- Managing data corruption incidents
- Avoiding vendor lock-in
- Policy landscape – uni/funder/industry
- Applying policies retrospectively

Engagement with Key Stakeholders
- Institutional Leads
- Funding Bodies

Lifecycle replacement policy

Investment in training
Time & budget

Platform/Tool specific training
- How to configure, build, operate, maintain
- Identifying corruptions

Introduction to Data Security
- Governance
- Sensitive Data
- Data Destruction
- air gaps
- Secure Access
- TRE's
- Supporting OCI containers on networked Filesystems
- Treat Data as Sensitive
- Security of archived & live data

Data provenance and Auditing

Performant Storage
- Benchmarking Storage systems
- Identifying Bottlenecks
- Impact on networking
- RDMA

ETL
MLEng
LLM

RDE/Data Stewards/Champions co-design sessions
- Meta data training
- Institutional Processes
- Funder requirements
- Publishing, DOI, Figshare etc.
- Data formats/standardisation
- RDM best practice - FAIR
- Research data lifecycle
- Disciplinary Data Services (External Institutions)
- Users understanding
  - Sensitivity
  - Security
  - Performance vs protection
  - Data Staging
  - Notifications of storage limits and usage
- Understand Characteristics of data
- Enabling ethical reuse
- Review Processes

Mapping research data lifecycle to platforms

Difference in policy between institutions

Sensitive data on HPC
Read & Active
VMs? Container?

Disaster Recovery

Tenancy and Namespacing

Migration between local systems

Introduction to Data Protection
- Protection vs performance
- Snapshots
- Scratch deletion – timed? notifications? docs?
- Archiving Post project 10yrs+

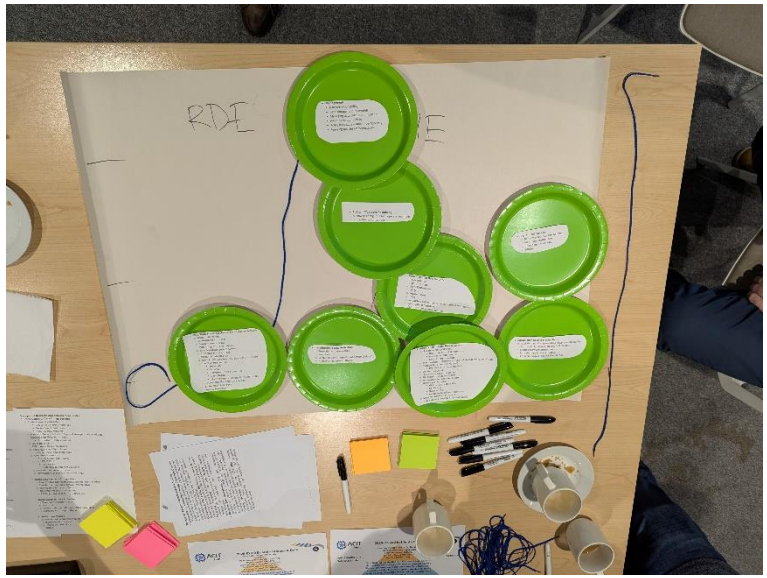
Introduction to Data Storage systems
- Understanding your data
  - Big/Large dataset differences
  - Protection requirements
  - Security requirements
- Standards/best practices of types of storage for types of data
- Storage Tiers (hot, warm, cold)
  - Promotion – auto vs manual
- On Prem vs Cloud
- HPC data vs Research Storage
- File Systems vs Object Stores
- Hardware (Disc, solid state, tape)
- Overview of options
  - NFS: Luster, ceph, GPFS
  - OS: PVC, s3
  - JBOD vs RAID
  - suitability to different use cases
- Migration between providers
- Notifications of storage limits and usage

Introduction to Data Storage Tools
- Apps for data Management /Catalogue/audit: starfish
- Transfer technologies (globus, MPI file utils)
- Sharing between Institutions
- Data sharing platforms: Globus
- API Technologies (e.g FTP, Filezilla, S3)

Sustainability of hardware

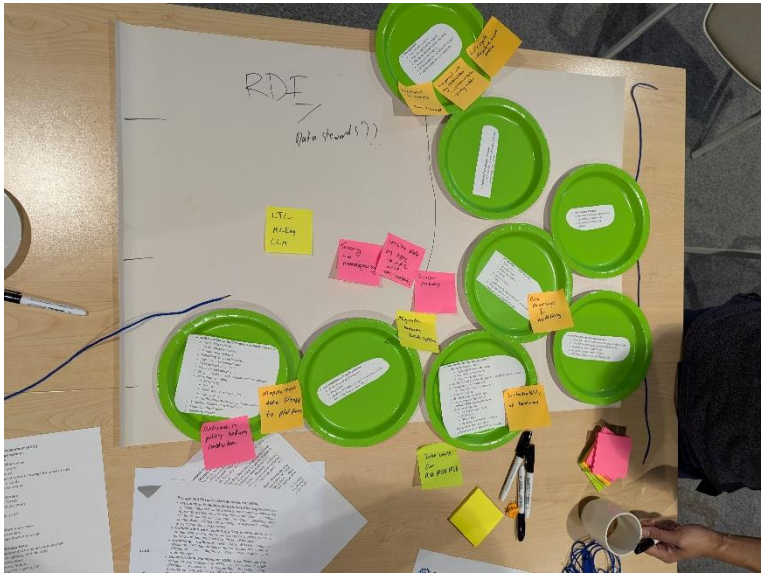
Intro course of RSE/RDE/RIE

- **Storage & RDM – Photos of progression through groups**

1. 

2. 

3. 

4. 

Networking and Datacentres

Management
- How to make the business case
- CoLo – choosing good partners
- Tenders
- Let researchers into the datacentres – political power
- Naming systems
- how to move from AC -> DLC
- how to train if it's all outsourced - S.E.P Addressing learned/taught helplessness when it's all external
- cost implications of new tech/change
- longevity planning
- E&F involvement
- Facility level buy in (Funding models)
- Funders awareness of E&F/Data centre issues

Customer Client management of outsourcing
- Learning vendor/Ops separation knowledge
- maintain involvement in system design
- understand limits of your remote fingers/ smart hands
- infrastructure and network mapping is sub-optimal at most levels (on-site, contractors)

Vendor junkets

Vendor specific accreditation for engineering on hardware

Hands on time in Datacentres for HPC
- Training HPC cluster for RIE's
- Diagnostics/troubleshooting – cooling/networking
- infrastructure and network mapping

RIE's allowed to do basic H/W support tasks

In house decisions/reasons for in house data centres

HPC Networks (vendor/product specific)
- network configuration
- network monitoring stats
- Troubleshooting N/W - "intuition checklist"
- Hands on time with hardware
  - old kit to "play" with for training

Alternative vendors for specialist n/w hardware

RSE
- Engage with RSE's to stress systems
- Signposting info for RSE's
- Hackathons with SSI – power/energy

Use of up to date software libraries 'Re-compile'

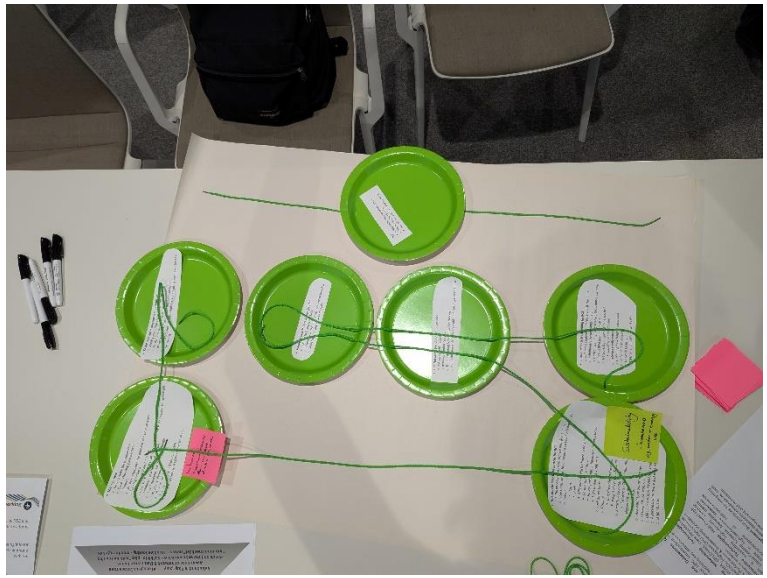
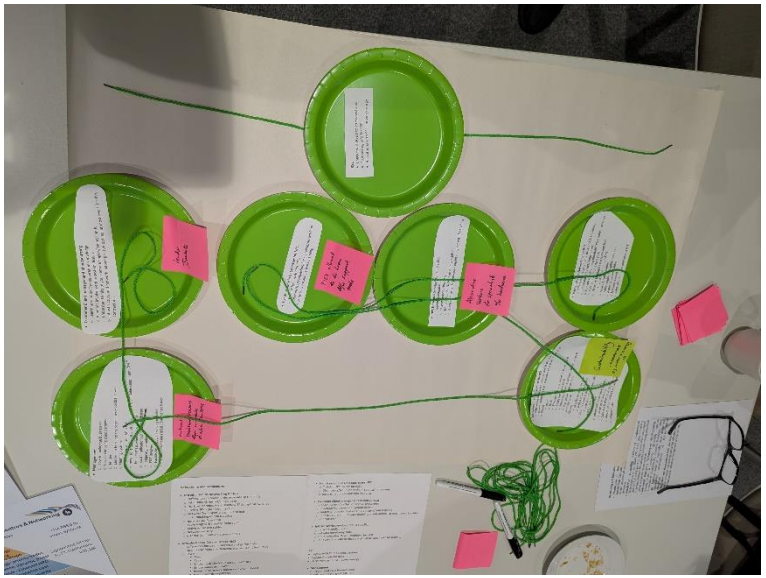
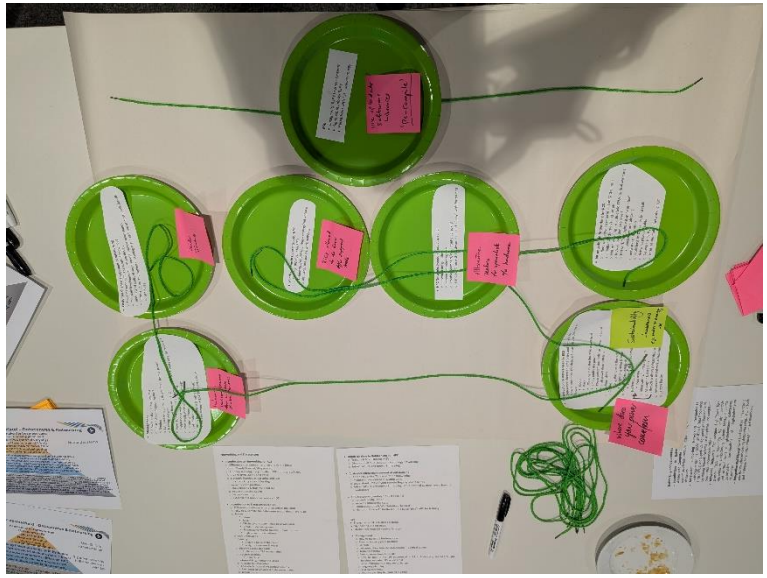
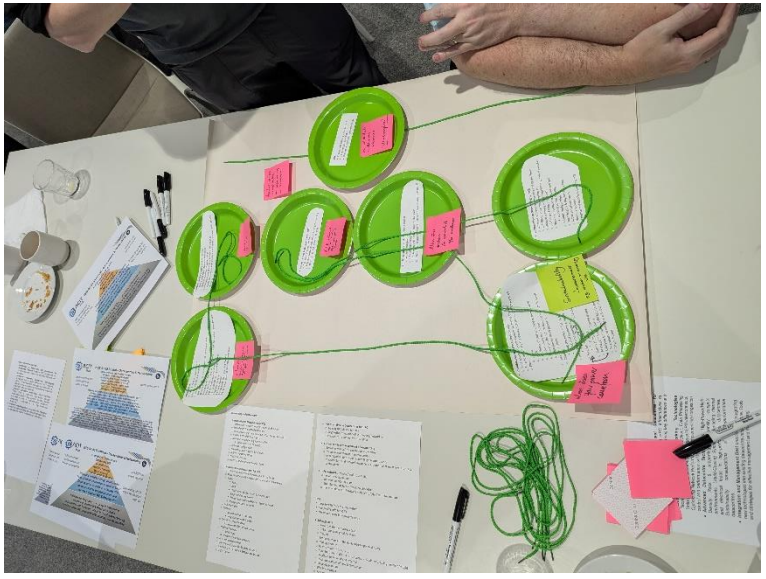
Introduction to Datacentres for RCS
- Differences between enterprise and @ the limits
- How to automate the datacentre so you don't have to go
- Power
  - Room
  - Rack
  - Efficiency of cluster linked to power used
  - Empty nodes use power
  - Power performance balance – dynamic boost
  - High power draw patterns
- rack dimensions
- Cooling
  - plug the holes in the racks
  - weight of systems & water
- Security access standards
  - Datacentre [TRE] accreditation
- NetZero Strategy
  - heat reuse
- observability/monitoring issues
- AI usage in Datacentres
- Lifecycle & disposal WEEE/data security
- Awareness of network & Datacentre teams
- Jargon buster

Where does your power come from?

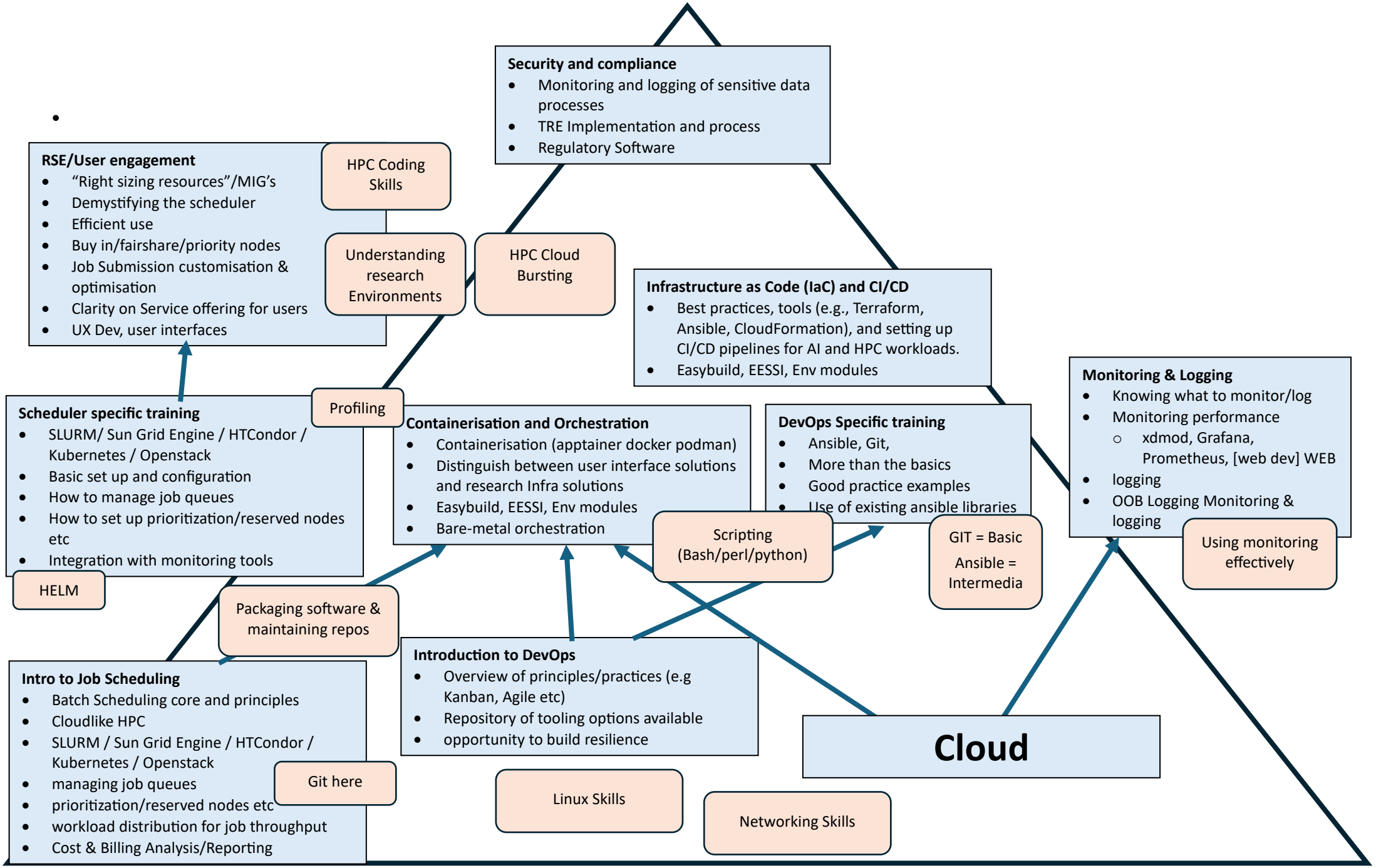
Sustainability awareness e.g Water and Energy use

Introduction to Networking for RCS
- differences between enterprise and @ the limits
- InfiniBand/Ethernet/slingshot etc
- Horizon scanning – e.g Bluefield TPUs, optical switches
- IPV4 vs IPV6 (don't do it yet!)
- Network Topologies e.g spine and leaf
  - blocking vs non blocking
- network configuration
- cable lengths & fragility, bend radii
- Active vs passive cables
- Network security
- external and mobile networks cf IOT

- **Networking and Datacentres – Photos of progression through groups**

1. 
2. 
3. 
4. 

Orchestration



Orchestration and Cloud – Photos of progression through groups

1. 

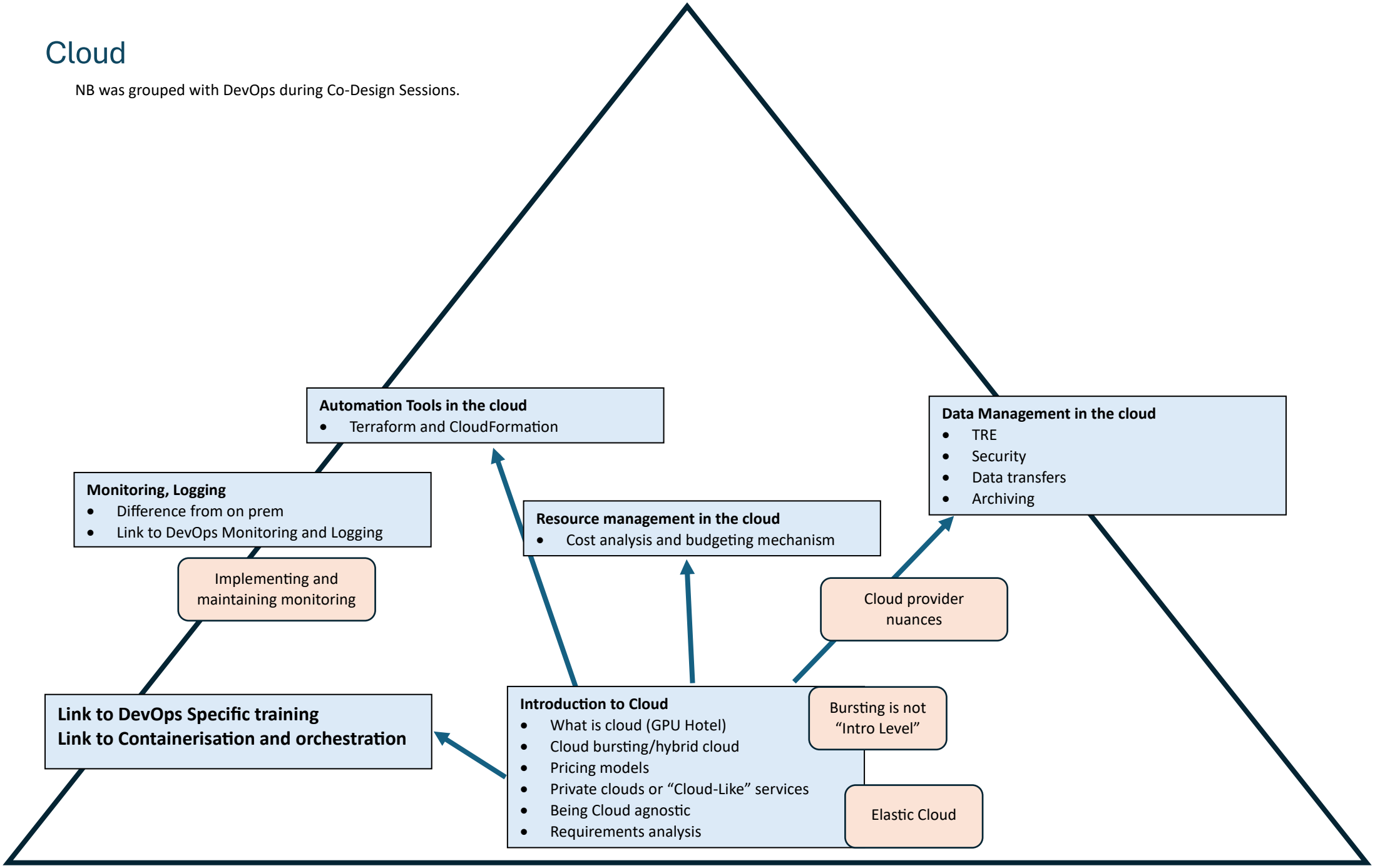
2. 

3. 

4. 

Cloud

NB was grouped with DevOps during Co-Design Sessions.



Misc

Understanding and clarity on different roles within Research IT
More appreciation for specialisations
collaborate with internal teams
Lack of team structure – better onboarding
HPC -> Lack of certs, certs pathway – structured development
Onboarding with “inclusivity” Junior feeling safe to ask questions
More sharing of best-practice
Onboarding – 1 Big tool at a time
mandatory training for ppl with no HPC b/g

- Case Studies
  - Implementation of Emerging tech
  - Novelty that enables pioneering research
  - Sustainability wins
  - Cost savings/efficiency wins