Accelerators

- **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

- **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

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

- **Multiple Tenancy GPU's**
  - Overview of available options e.g MIG
  - Security implications
  - Scheduling
  - Containers

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

- **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

- **Beyond the GPU** Focus on edge computing, alternative markets, and real-world case studies of successful implementations of non-GPU Accelerators in academia and industry.

Management

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

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?)

Storage and Research Data Management (RDM)

- **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)

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

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

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

Networking and Datacentres

- **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

- **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

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

- **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)

- **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

RSE

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

- **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

Orchestration

- **Introduction to DevOps**
  - Overview of principles/practices (e.g Kanban, Agile etc)
  - Repository of tooling options available
  - opportunity to build resilience
- **Infrastructure as Code (IaC) and CI/CD**
  - Best practices, tools (e.g., Terraform, Ansible, CloudFormation), and setting up CI/CD pipelines for AI and HPC workloads.
  - Easybuild, EESSI, Env modules
- **Containerisation and Orchestration**
  - Containerisation (apptainer docker podman)
  - Distinguish between user interface solutions and research Infra solutions
  - Easybuild, EESSI, Env modules
  - Bare-metal orchestration
- **Monitoring & Logging**
  - Knowing what to monitor/log
  - Monitoring performance
    - xdmod, Grafana, Prometheus, [web dev] WEB
  - logging
  - OOB Logging Monitoring & logging
- **Security and compliance**
  - Monitoring and logging of sensitive data processes
  - TRE Implementation and process
  - Regulatory Software

- **Intro to Job Scheduling**
  - Batch Scheduling core and principles
  - Cloudlike HPC
  - SLURM / Sun Grid Engine / HTCondor / Kubernetes / Openstack
  - managing job queues
  - prioritization/reserved nodes etc
  - workload distribution for job throughput
  - Cost & Billing Analysis/Reporting

- **DevOps Specific training**
  - Ansible, Git,
  - More than the basics
  - Good practice examples
  - Use of existing ansible libraries

- **Scheduler specific training**
  - SLURM/ Sun Grid Engine / HTCondor / Kubernetes / Openstack
  - Basic set up and configuration
  - How to manage job queues
  - How to set up prioritization/reserved nodes etc
  - Integration with monitoring tools

- **RSE/User engagement**
  - “Right sizing resources”/MIG’s
  - Demystifying the scheduler
  - Efficient use
  - Buy in/fairshare/priority nodes
  - Job Submission customisation & optimisation
  - Clarity on Service offering for users
  - UX Dev, user interfaces

Cloud

- **Introduction to Cloud**
  - What is cloud (GPU Hotel)
  - Cloud bursting/hybrid cloud
  - Pricing models
  - Private clouds or “Cloud-Like” services
  - Being Cloud agnostic
  - Requirements analysis
- **Data Management in the cloud**
  - TRE
  - Security
  - Data transfers
  - Archiving
- **Resource management in the cloud**
  - Cost analysis and budgeting mechanism
- **Monitoring, Logging**
  - Difference from on prem
- **Automation Tools in the cloud**
  - Terraform and CloudFormation

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