



High Performance Computing (HPC): Cluster Administration, Slurm and GPU Nodes



High Performance Computing (HPC): Cluster Administration, Slurm and GPU Nodes

Introduction:

High performance computing HPC clusters now carry crash and fluid-dynamics simulation, genomics pipelines and AI model training, yet many organisations run them with idle GPU nodes, jammed queues, saturated scratch storage and users who cannot tell why their MPI job stalls. This Core Concept course equips infrastructure and research computing staff to design, schedule, tune and support a shared cluster: node roles and interconnect, Slurm partitions and fairshare, MPI and OpenMP execution, GPU allocation, parallel storage, Apptainer containers and on-premises versus cloud bursting economics. Participants finish with an HPC Cluster Design and Service Plan for a case engineering and AI centre.

Course Objectives:

- Characterise simulation, data-intensive and AI training workloads by their compute, memory, communication and I/O profile to size nodes and interconnect
- Configure Slurm partitions, quality-of-service limits, multifactor priority and slurmdbd accounting so that queue wait and utilisation meet agreed targets
- Launch and diagnose hybrid MPI and OpenMP jobs, including process binding, rank placement and GPU affinity on multi-socket and accelerator nodes
- Design scratch, project and archive storage tiers on a parallel file system with quotas, purge policies and data movement rules
- Benchmark and tune cluster throughput using HPL, HPC Challenge and application profiling, and package user software reproducibly with Apptainer
- Produce an HPC Cluster Design and Service Plan covering capacity, cost model, operations runbook and user support tiers

Target Audience:

- Infrastructure engineers who build and maintain Linux compute clusters, head nodes and high-speed fabrics
 - Research computing staff who onboard scientists and engineers and keep shared queues productive
 - Engineering simulation teams who run CFD, structural and crash solvers across many nodes
 - AI platform administrators who allocate GPU nodes for distributed model training and inference
 - Data centre operations staff responsible for power, cooling and hardware lifecycle of dense compute racks
-

Course Outline:

Day 1: HPC Workloads, Supercomputing Landscape and Current Cluster Assessment

- Capability Versus Capacity Computing: Tightly Coupled Solvers, Embarrassingly Parallel Sweeps and AI Training
- TOP500 List, HPL Rankings and FLOPS, Memory Bandwidth and Latency Metrics
- Amdahl's and Gustafson's Laws for Strong and Weak Scaling Estimates
- Workload Profile Matrix: Cores, Memory per Core, Message Size and I/O Pattern
- Current Cluster Health Check: Queue Wait, Node Utilisation and Storage Fill Baseline

Day 2: Cluster Architecture: Node Roles, Interconnect, Scheduler and Storage Design

- Login, Head, Compute, GPU and Data Transfer Node Roles and Isolation
- InfiniBand and High-Speed Ethernet Fabrics: Fat-Tree and Dragonfly Topologies, RDMA and Latency
- Slurm Architecture: slurmd, slurmd, slurmd, Partitions, Jobs and Job Steps
- Parallel File System Layout: Metadata Servers, Object Storage Targets and Striping
- Cluster Provisioning: Stateless Node Images, Configuration Management and Module Environments

Day 3: Cluster Administration Lab: Slurm Scheduling, MPI and OpenMP Jobs, GPU Nodes

- Slurm Batch Scripts: sbatch, srun, Job Arrays, Dependencies and Interactive Sessions
- Fairshare, Quality-of-Service Limits, Backfill and Preemption Policy Configuration
- Open MPI Launch, Rank Placement and Hybrid MPI Plus OpenMP Process Binding
- GPU Scheduling with Generic Resources, Multi-GPU Jobs and Accelerator Affinity
- Apptainer Image Builds, Bind Mounts and MPI Inside Containers

Day 4: Performance Tuning, Storage Pressure, Failure Cases and Cost Trade-offs

- HPL and HPC Challenge Benchmark Runs and Result Interpretation
- Application Profiling: Load Imbalance, Communication Overhead and NUMA Effects
- Scratch Purge, Quotas, Small-File Storms and Data Staging to Archive
- Node Drain, Hung Jobs, Fabric Errors and Scheduler Failover Troubleshooting
- On-Premises Versus Cloud HPC Bursting: Cost per Core-Hour and GPU-Hour Model

Day 5: Lab Capstone: HPC Cluster Design and Service Plan for an Engineering and AI Centre

- Case Brief: Demand Forecast for Simulation Teams and AI Model Training Groups
 - Node Mix, Fabric and Storage Tier Sizing Worksheet
 - Scheduler Policy and Allocation Scheme for Competing Research Groups
 - User Support Tiers, Onboarding Guide and Operations Runbook Drafting
 - HPC Cluster Design and Service Plan Presentation and Peer Critique
-

Skills You Will Gain:

- HPC Workload Characterisation
- Slurm Scheduler Administration
- Parallel Job Diagnostics
- GPU Resource Allocation
- Parallel Storage Management
- Cluster Benchmarking
- HPC Container Packaging
- Compute Cost Modelling

Why Attend This Course:

- Leave with an HPC Cluster Design and Service Plan for a case engineering and AI centre, critiqued by peers
- Shorten queue waits and raise node utilisation by applying scheduler policies tested in the lab
- Resolve stalled MPI jobs, full scratch file systems and drained nodes with a practised troubleshooting routine
- Compare cluster practice with administrators from universities, engineering firms, energy and AI groups

Conclusion:

A shared compute cluster earns its investment only when workloads, scheduling, storage and support are designed together. The week moves from HPC workload profiles and scaling laws, through node roles, fabrics, Slurm and parallel storage, to hands-on scheduling, MPI and OpenMP execution, GPU allocation and Apptainer packaging, then benchmarking, failure cases and on-premises versus cloud cost. The final day turns this into an HPC Cluster Design and Service Plan that participants can take back to guide capacity, policy and user support decisions.
