

# Attachment J - CLIN 00006 – High Performance Computing (HPC) Supplemental Information

## CLIN 00006 – High Performance Computing (HPC)

- **HPC Compute and Storage Operations Support**
  - Joule
    - Current Joule 2.0 has 40 racks (avg. 850 KW) consisted of around 1,856 Compute Nodes and Storage Nodes in MDC (Modular Data Center) and Joule Backup Nodes in B-39 Data Center. This backup data is replicated to Storage Nodes in PGH B-922 Data Center.
    - The next 3<sup>rd</sup> HPC, Joule 3, is anticipated to have up to 40 racks, with around 812 CPU Compute (104K AMD EPYC Zen4 CPU cores, 384GB/768GB Memory per node) and 9 GPU Compute (72 H100-80GB GPUs, 1.5TB Memory per node), 32 storage nodes (around 20PBs), Infiniband 200Gbps network, etc. (Installation is expected to start from February 2024 and operate in May 2024, TBD)
  - Watt
    - Current Watt 1.0 is consisted of 19 racks (13 server racks and 6 vertical cooling racks) with 39 storage nodes (40PBs of raw storage), 24 GPU compute nodes (96 P100 GPUs), and 8 CPU-only pool servers.
      - 24 x 40-core Intel Skylake nodes (384 GBs of RAM and Four Nvidia P100 GPUs per node)
      - 40 PBs of raw storage - Configured as Lustre parallel file system, XFS, and ZFS shares
      - 100 Gbps OmniPath interconnect for parallel file system and high-performance file access
      - 50 TFLOPs performance across all of the CPUs
      - 323 TFLOPs performance across all of the GPUs
      - 66 GB/sec write and 122 GB/sec read performance on the parallel FS
      - OpenStack is installed and used to manage all ML compute and storage resources. Unlike traditional HPC environments where OS versions and libraries are often standardized and difficult to reconfigure, OpenStack allows users to “bring their own” OS images and environments. Either Windows or Linux, and any Linux variant can be supported. We provide baseline Windows and Linux image files to get started. Can work with researchers to build OS images tailored to the specific application requirements.
    - Next Watt 1.1 is anticipated to have existing Watt 1.0 GPU Compute and Joule 2.0 Compute servers and Storages. NETL has a plan to retrofit to Watt 1.1 once Joule 2.0 is removed from MDC.

- Next Watt 2.0 is anticipated to be procured and have up to 36 racks in the future.
    - The Contractor shall support new AI infrastructure as well such as, but not limited to, wafer-scale engine (WSE), Cerebras CS-2, to support Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) for NETL AI/ML research.
    - The Contractor shall maintain Short-term and Long-term file system using Lustre parallel distributed file system, and Network File System (NFS, not backed-up; 60-day purge policy) as well as Robinhood (or equivalent) Policy engine. Various file system and file transfer protocol should be managed, including, ZFS, XFS, EXT4, and File Transfer Protocol (FTP). Long-term file system should be housed in a separate, physically segregated, data center with backup power generation to make storage of data as safe as possible.
- **HPC Network Operations Support**
  - The Contractor shall manage Computational Network Interconnects within HPC cluster using diverse network technologies such as, but not limited to, Omnipath 100Gbps and Infiniband EDR 100Gbps / HDR 200Gbps / NDR 400Gbps or faster, low-latency network using diverse vendor products.
  - All switches must be compatible and supported on existing SciLAN and ESNET infrastructure, supported by Juniper and Infiniband.
- **HPC Virtualization, OS, Middleware, Runtime, and Application Support**
  - The Contractor shall provide HPC management, including underlying Operating System (Unix/Linux – CentOS, Ubuntu, RHEL, Rocky Linux 9, etc. and Windows – Windows 10, Windows Server, etc.), Virtualization/Hypervisor & Containerization, Middleware, Runtime, and Application Layers.
  - HPC Cluster Management Tool (Slurm or equivalent)
  - Open-Source Cloud Computing Infrastructure and Management Tool (OpenStack or equivalent) and Virtualization/Containerization technologies (Kubernetes, Docker Swarm, or equivalent)
  - HPC Simulation-Based Engineering Tools/Applications and Parallel Programming language and numeric computing environment, such as
    - Joule: MFix, CFD++, ANSYS FLUENT/Workbench, Coarsening, VASP, Comsol, Matlab, Mathworks, CUDA, Nvidia Nsight Visual Studio, OpenMPI, LAMMPS, OpenGeoSys, OpenFOAM, TurboMole, CP2K, NWChem, Octave, Materials Studio, TensorFlow, TensorBoard, Paraview, Blender, OpenGL/VirtualGL, JDTFx, Phonopy, Python, Spyder, QuickFF, QT Creator, Thermo-Calc, AmberTools, APML, ADF, BoltzTrap, XTB, etc.
    - Watt: Tensorflow, Jupyter Notebook, Anaconda, Python/Perl, PyTorch, R, Apache Spark - MLib, Apache MXNet, Matlab, etc.
  - HPC Access Client application, CAP (Cloaked Authentication Protocol), which is developed by NETL HPC Support team. The CAP client does bundle in "helper" applications that are open-source including TurboVNC (C/Java), Yubico libraries (C), WinSCP (C++), SPICE viewer (C), etc.

- The Contractor shall provide Development-Operational (DevOps) and Performance Monitoring (Saltstack, Grafana, or equivalent) support.
  - The Contractor shall run High-Performance LINPACK (HPL) Benchmark to submit NETL HPC performance to the biannual TOP500 supercomputer list, whenever running the benchmark and achieving a better result.
  - HPC system should be secured through MFA with Yubikey and encrypted SSH tunnel connection through HPC Access Client.
- **HPC Account Management Support**
    - The Contractor shall support Directory Service for HPC by managing LDAP-Samba4 (or equivalent) Active Directory (AD) Domain Controller (DC).
    - The Contractor shall program MFA RSA Token (cryptosystem for public-key encryption) Yubikey (or equivalent) and deliver to HPC users. The Contractor shall support the existing RSA Token for HPC.
- **HPC Data Center Facility Operations & Maintenance Support**
    - The Contractor shall ensure that all HPC compute nodes, storage nodes, interface and maintenance nodes, and networks are fully monitored at all times using HPC monitoring tool (Saltstack and Grafana are the primary tools) and MDC environments (louvers, fans, actuators, PLC pars, remote I/O parts, etc.) are fully monitored by MDC monitoring tool (Rockwell Automation, Factory Talk software), with automatic notifications configured to immediately alert staff of problems or failures. The Contractor shall maintain HPC/MDC monitoring tools used to monitor the health status and availability of all HPC environments. Upon detection of problems or failures, the Contractor shall perform immediate remedial actions to stabilize or restore the associated HPC services and notify Federal designee of the problems or failures.
    - **Current Joule Modular Data Center (MDC):**
      - Site/Location: Morgantown/MDC
      - Size (sq.ft): 1,391 (1,958 including vestibule)
      - Racks: 40
      - Power Capacity (KW): 1,125 (Averaged HPC Power 650KW, Total Power 750KW)
      - Alt Pwr-B/U: Small UPS only, No Generator
      - Cooling: Adiabatic (Air cooling with evaporative media)
      - PUE Actual – Planned: 1.02-1.2 (Average PUE 1.151)
      - In-Service Date: August 2012
      - Other: Adding a chilled water supplemental cooling system for use during hot/humid months in 2023 (being delayed b/c of supply chain issues). Currently we have only small 50 kva UPS for only maintenance and storage nodes.
      - MDC Maintenance List – monthly/3months/6months/12months operation/task should be performed.

Operation / Task	monthly	3 months	6 months	12 months
Verify pressure drop across filter media via the PLC Software. Schedule corrective maintenance if indicated.	X			
Visually inspect the condition of solenoid valves. Verify they are functioning properly by turning them on/off manually via the PLC UI.	X			
Check the condition and lubricate of access door hinges for easy operation and proper locking.	X			
Visually inspect and lubricate dampers and actuators. Verify they are functioning properly by opening/closing manually via the PLC UI.	X			
Visual inspect UPS system signals and warnings on control screen per UPS Manufacturer provided Instructions.	X			
Using the PLC UI, verify proper Water Flow to the evaporative media system. Physically Check Inlet Piping, Drain Flow, and Water Pressure via the unit pressure gauge.		X		
Check the fan motor running current.			X	
Verify PLC MANUAL controls and effects on unit components. Verify all fans, dampers, and actuators perform properly in their full range; verify data indicated in PLC matches actual conditions.			X	
Check fan operation/rotation.			X	
Visual inspection the protective covering of fan motor			X	
Visual inspection of fan for damage			X	
Visual inspection of fan mounting			X	
Visual inspection of entire MDC material condition, including interior, roof, doors, building envelop, and all moving mechanical parts (e.g., dampers, actuators, louvers, solenoids).			X	
Visually inspect Heat Tracing wiring. Verify fuse/breaker functioning properly. (Autumn just prior to Winter)				X
Physically verify PLC UPS Battery voltage and condition				X
Visual inspection of all electrical and controls connecting cables, terminations, and insulation				X
Visual inspection of protective earth connection				X
Replace ALL evaporative cooler media				X
Replace ALL Evaporative media HEPA filters				X
Check the filter frame for proper sealing.				X
Deep cleaning of the MDC				X
Inspect Entire UPS system based on manufacturer recommended/required best practices. Check UPS, UPS Transformer for proper voltages, amperages.				X
Replace PLC battery backup				X
Check all electrical panel breakers (below 200Amps) are operational, fuses good, switches operational				X
Visual inspection of all electrical and controls connecting wiring, isolating devices, terminations, and insulation.				X
Perform checks of electrical Low Voltage Circuit Breakers >200Amps				X

○ **Current Watt Data Center**

- Size (sq.ft): 760
- Racks: 19
- Power Capacity (KW): Approx 40 KW typical, 120KW max
- Alt Pwr-B/U: small UPS only, No Generator
- Cooling: Chiller with cooling racks
- In-Service Date: 2018
- Other: Power consumption is not monitored. Values estimated based on UPS readouts (40 KW). Currently we have small 72 kva UPS.

- **HPC Service Desk Support**

- Provide courteous, prompt, high-quality end user support services, including Yubikey, HPC RSA Token delivery to HPC users.
- HPC website:
  - Joule website (<https://hpc.netl.doe.gov/>)
  - Watt website (<https://ml.netl.doe.gov/>)
- HPC Service desk
  - SBEUC-Access@netl.doe.gov: New Access Request for HPC
  - SBEUC-Support@netl.doe.gov: All other HPC support