Draft Technical Specifications

NCAR's Next Generation HPC System - NWSC-2

NWSC-2 Computing Platforms: Draft Technical Specifications

1 Introduction

The University Corporation for Atmospheric Research (UCAR), on behalf of the Computational Information Systems Laboratory (CISL) at the National Center for Atmospheric Research (NCAR) will release a Request for Proposal (RFP) for the next-generation high-performance computing (HPC) system to be installed at the NCAR-Wyoming Supercomputer Center (NWSC). This new system, herein referred to as NWSC-2, is expected to be delivered in the second half of CY2016.

1.1 Scope of the NWSC-2 Procurement

NCAR requires a production HPC system in the 2016/2017 timeframe to support the rapidly increasing computational demands of the atmospheric science community. The system must provide a significant upgrade in computational capabilities, with a target increase of 2-5 times greater than the current sustained performance of the Yellowstone system [1].

The NWSC-2 procurement includes the following **Required** and **Optional** components:

- High-performance computing system (**Required**)
- High-performance storage system that is integrated with the HPC system (**Required**)
- Support for access to NCAR's existing GPFS-based Globally Accessible Data Environment file system (**Required**)
- Support for access to NCAR's existing HPSS mass storage system (**Required**)
- Many-core nodes which reside on the HPC interconnect (**Optional**)
- General Purpose GPU nodes which reside on the HPC interconnect **(Optional)**
- Visualization, data-analysis, and post-processing nodes which reside on the HPC interconnect (**Optional**)
- Innovative Storage and Memory Technologies **(Optional)**

The scientific and strategic drivers for Yellowstone and the NWSC data center were set forth in the NWSC Science Justification [2]. Though published in 2009, it remains highly relevant for this procurement in articulating the long-term vision and opportunities for using HPC to address scientific challenges in the atmospheric sciences.

CISL supports over 1,000 users annually and 200 applications in the atmospheric and related sciences. The primary models used by these communities include the Community Earth System Model (CESM) and its constituent components which form a fully-coupled, community, global climate model; several variants of the Weather Research and Forecasting (WRF) model which serves as a mesoscale numerical weather prediction system designed to serve both atmospheric research and

operational forecasting needs; the Model for Prediction Across Scales (MPAS), a collaborative model using unstructured Voronoi meshes to provide both quasi-uniform and locally refined grids for atmosphere, ocean and other earth-system simulation components; and a broad range of applications for the post-processing, analysis, and visualization of simulation and observational data.

An overview of NCAR's computational workload [3] summarizes key aspects of the application and job mix, and provides a quantitative assessment of how the current Yellowstone system is being used.

This discipline-specific user community results in a more narrow and focused applications base than for many other HPC centers; over half of Yellowstone's delivered core-hours are for running the CESM. The scientific goals driving the need for additional computational capability and capacity are clear. Atmospheric science already relies on large-scale simulation, and over the next several years it will move to incorporate the additional physical processes and higher resolution needed to achieve more faithful representations of real-world systems. Additionally, a large, and significant portion of the scientific discoveries are made not at the largest computational scales, but rather, performed using ensembles of individual, mutually-independent simulations.

The NWSC-2 system will be housed in the NCAR-Wyoming Supercomputer Center, in Cheyenne, WY, and is expected to run for four years, with options to extend beyond that. The system must support integration with the existing GPFS-based NCAR Globally Accessible Data Environment (GLADE) and HPSS-based NCAR archive.

1.2 High-level Schedule

The following is the tentative schedule for the procurement.

RFP Released	H1CY15
Subcontract Awarded	H2CY15
On-site System Delivery and Build Complete	H2CY16
Acceptance Complete	December 2016

1.3 Estimated Available Funding

The estimated available funding for this procurement is $30M. This is the not-to-exceed funding amount for all **Required** and **Optional** elements listed in §1.1 and includes all equipment, support, and maintenance for 4 years of operation. UCAR intends to make a single award through a competitive, Best Value procurement process. It is possible for an Offeror proposal to exceed the available funding since the total of all options will exceed this amount and not all options will be exercised.

UCAR reserves the right to award any amount of the funding—including making no award if no compelling Offeror proposal substantively addresses the requirements.

2 Mandatory Elements of Offeror Response

An Offeror shall address all mandatory requirements and its proposal shall demonstrate how it meets or exceeds each one. A proposal will be deemed non-responsive and will receive no further consideration if any of the following mandatory requirements is not met.

2.1.1 The Offeror shall provide a detailed architectural description of the NWSC-2 production and test systems. The description shall include: a high-level architectural diagram that includes all major components and subsystems; detailed descriptions of all the major architectural hardware components in the system to include: node, cabinet, rack architecture up to the total system, including the high-speed interconnect and network topology; detailed descriptions of the system software components; the storage subsystem and all I/O and file system components; electrical and cooling requirements; and a proposed floorplan.

2.1.2 The Offeror shall provide a detailed plan for delivery, installation, maintenance and support services necessary to meet the NWSC-2 target reliability through the proposed system lifetime, and a proposed delivery, installation and acceptance testing schedule for the NWSC-2 system, including the number and roles of any temporary or long-term on-site Offeror personnel.

2.1.3 The Offeror shall describe how the proposed system does or does not fit into their HPC roadmap for the period of deployment as well as a potential NWSC-3 acquisition in the 2020-2023 timeframe.

3 Target Design Requirements

This section contains detailed system design targets and performance features. It is desirable that the Offeror's design meets or exceeds all the features and performance metrics outlined in this section. Failure to meet a given Target Design Requirement will not make the proposal non-responsive. However, if a Target Design Requirement cannot be met, it is highly desirable that the Offeror provide a development and/or deployment plan and schedule to satisfy the requirement.

The Offeror shall address all Target Design Requirements and describe how the proposed system meets, exceeds, or does not meet the Target Design Requirements. The Offeror shall also propose any hardware and/or software architectural features that will provide improvements for any aspect of the system. Areas of interest include, but are not limited to: application performance, storage and memory technology, power usage, overall productivity of the system, and systems management.

3.1 Scalability

The NWSC-2 production system workload will include very large-scale jobs, up to and including full-system size; therefore, the system must scale well to ensure efficient usage. However, the anticipated job mix is likely to be dominated by jobs at smaller scales [3], and thus CISL anticipates the system is likely to reflect scalability trade-offs in order to serve the overall production workload.

3.1.1 The system shall support hundreds of concurrent users and thousands of concurrent batch jobs. The Offeror shall describe and provide details on how the system supports this requirement.

3.1.2 The system's high-speed interconnect shall support high bandwidth, low latency, high throughput, and minimal inter-job interference. The Offeror shall describe the high-speed interconnect in detail, including topology, all performance characteristics, mechanisms for adapting to heavy loads (including mixes of I/O and interprocess communication) or inoperable links, and dynamic responses to failure and repair of links and/or nodes.

3.1.3 The Offeror shall describe how the high-speed interconnect design represents an appropriate balance between supporting single full-system jobs at full-scale and supporting a workload of smaller-scale, i.e. <=64 nodes jobs.

3.1.4 The system shall support running a single application at full scale.

3.1.5 The system shall provide reproducible numerical results and consistent runtimes. The Offeror shall describe strategies available to system administrators and to user applications for minimizing runtime variability. An application's runtime (i.e., wall clock time) shall not change by more than 3% from run-to-run in dedicated mode and 5% in production mode. Variability will be measured by using the Coefficient of Variation as defined in the glossary.

3.1.6 If the system has heterogeneous node architectures, the Offeror shall describe any associated scalability limitations, impacts to the high-speed interconnect, and the scalability of applications running on each set of homogeneous nodes.

3.2 System Software and Runtime

The Offeror shall propose a well-integrated and supported system software environment that is high-performing and reliable.

3.2.1 The Offeror shall provide a system that includes a full-featured Unix-like operating system (OS) environment on all user-visible OS instances including compute nodes, login nodes, service nodes, and management servers.

3.2.2 The Offeror shall describe any system software optimizations or support for a low-jitter environment for applications and shall provide an estimate of a compute node OS's noise profile, both while idle and while running non-trivial MPI application. If core specialization is used, describe the system software activity that remains on the application cores.

3.2.3 The Offeror shall describe the security capabilities of the full-featured, Unix-like OS. The OS for the service nodes and for the system management workstations shall provide at a minimum the following security features: ssh version 2, POSIX user and group permissions, POSIX access control lists, kernel-level firewall and routing capabilities, syslogging, and auditing.

3.2.4 The compute partition OS shall provide a trusted, hardware-protected supervisory mode to implement security features. The Offeror shall describe how the supervisor/kernel provides authoritative user identification, ensures that user access controls are in place, employs the principle of least privilege, and interoperates with the same features on the service nodes and management workstation(s). Logging and auditing features supported by the compute node OS shall have the capability to be enabled, disabled and custom-configured to site preferences.

3.2.5 The Offeror shall describe the processes and procedures in place for generating and providing software updates for the proposed OS. Processes should allow expeditious updating of kernel and non-kernel packages to address issues, including security vulnerabilities in the suite of software.

3.2.6 The Offeror shall describe how the system provides support for static libraries and objects and/or dynamic loading of shared objects. The Offeror should describe how the system will support applications using these techniques at the full scale of the system.

3.2.7 The Offeror shall describe how the system provides efficient, secure, interprocess communication that allows cooperating applications running anywhere on the high-speed network to inter-communicate (e.g., the compute partition, the service partition, or both). The provided mechanism shall be as close to the underlying network stack as possible. The security model shall allow applications and users to set access controls based on authenticated or trusted values for process and user identifiers.

3.2.8 The Offeror shall describe and provide a documented and efficient application programming interface (API) for the native network layer(s) of the high-speed network software stack.

3.2.9 The Offeror shall provide a job scheduler and resource management subsystem capable of simultaneously scheduling both batch and interactive workload. The Offeror shall describe the features and capabilities available to administrators and users, including: checkpoint-restart, job migration, backfill, targeting of specified resources, advance and persistent reservations, job preemption, monitoring of running and pending jobs, job reporting and accounting, architecture-aware job placement, and Eclipse integration. The Offeror may propose multiple options for a third-party supported resource manager.

3.2.10 The job scheduler and resource management subsystem shall support an efficient mechanism to launch applications at sizes up to full scale. The Offeror shall describe the factors (such as executable size, number of jobs currently running or queued, and so on) that affect application launch time. The Offeror shall describe expected application launch times and how the factors noted increase or decrease the launch time.

3.2.11 The job scheduler and resource management subsystem shall utilize an optimized job-placement algorithm to reduce job runtime, lower variability, minimize latency, etc. The Offeror shall describe in detail how the algorithm is optimized to the system architecture.

3.2.12 The Offeror shall describe the centralized system(s) for operation, management, monitoring and administration. The capabilities should include monitoring, logging, and notification of hardware failures and hardware and software events.

3.2.13 The Offeror shall describe how the system can be operable and administrable from a remote location via lights-out management. This may be in the form of an embedded controller, or an equivalent capability, that provides the ability to manage the system remotely.

3.3 Software Tools and Programming Environment

The primary parallel programming model currently used on existing NCAR systems is hybrid Message Passing Interface (MPI) with MPI/OpenMP. To support current climate and weather models that form the large majority of NCAR's production workload, the Offeror's proposed system shall continue to support the hybrid MPI/OpenMP programming model for its primary production workload.

In general, CISL intends to evaluate and acquire third-party software tools separately from the NWSC-2 system procurement. For each of the following software categories, the Offeror should describe the Offeror's own optimized and/or integrated software tools and programming environment components. The Offeror should also describe any third-party software that CISL may have a compelling reason to make part of the NWSC-2 procurement.

3.3.1 The production system shall support the Message Passing Interface 3.0 (MPI-3) standard specification. The Offeror shall describe the MPI implementation, including version, optimizations for collective operations, support for features such as accelerated collectives, and the ability for applications to access the physical-to-logical mapping of the job's node allocation, and describe any limitations relative to the MPI-3 standard.

3.3.2 The Offeror shall describe at what parallel granularity the production system can be efficiently utilized by MPI-only applications.

3.3.3 The Offeror shall describe and provide licenses (50 seats) for at least one set of high-performance, optimizing compilers capable of creating executables for the compute partition of the system offered. These compilers shall support the latest International Standards for C, C++, and Fortran.

3.3.4 The Offeror shall describe in detail all programming APIs, languages,, compiler extensions, etc., other than MPI that will be supported (e.g., OpenMP, OpenACC, CUDA, OpenCL, and so on) on the NWSC-2 production system. Describe the advantages and disadvantages of each node-level programming API from a programming and performance perspective on the proposed architecture. In addition, describe any interoperability limitations (e.g., thread interoperability).

3.3.5 The Offeror shall describe support in the programming toolchain(s) for mixed-language programs. In particular, use of parallel features of any one of C, C++ or Fortran shall not interfere with the use of parallel features in the others.

3.3.6 The compilers identified in §3.3.3 shall describe the support for partitioned global address space (PGAS) features, including coarrays in Fortran and, possibly, Unified Parallel C (UPC). The Offeror shall describe production system hardware and programming environment software for exploiting PGAS capabilities.

3.3.7 The Offeror shall list and briefly describe all optimized mathematical and I/O libraries that will be included and supported on the production system.

3.3.8 The system shall allow for control of task and memory placement within a node for efficient performance of applications. The Offeror shall describe the controls provided and any limitations that may exist.

3.3.9 The Offeror shall describe any other optimized and tightly coupled programming environment components and software tools that will be provided as part of the programming environment and/or optimized compiler tool suite. These may include debuggers, performance analysis tools, event-tracing tools, and stack-tracing tools. Such tools should support at least 20 simultaneous users and a single job up to one-fourth (1/4) scale of the production system.

3.4 Parallel File System

3.4.1 The system shall include a parallel file system (PFS) that presents a global, consistent name space to the platform. The Offeror shall provide a detailed description of the PFS implementation including:

- File system architecture and proposed implementation
- Expected scaling characteristics
- Management, diagnostic, deployment, security and configuration tools
- Externalized error and diagnostic information

3.4.2 The parallel file system shall be designed to minimize the risk of data corruption and data loss. Reliability of the PFS will be assessed as part of the overall system reliability metrics. For the proposed PFS implementation, the Offeror shall provide a detailed description of the following:

- How the I/O solution strives to achieve zero corruption and zero data loss over the life of the system
- How silent data corruption errors (e.g., high-fly writes, short writes, misdirected I/Os) are addressed
- How the I/O solution will function when a failure is exposed beyond the RAID level
- Ability to run in a partial/degraded mode
- Estimates of failure rates of the various components comprising the PFS implementation.

3.4.3 The PFS shall achieve the target aggregate bandwidth, as determined by the FLOPS/BYTE ratio given in Table 5. FLOPS/BYTE is the ratio of sustained performance (as measured by the NCAR Benchmarking Suite, §3.6), to the maximum achieved performance of the PFS. For example, if the sustained performance is 50 TF, and the FLOPS/BYTE ratio is 250, the aggregate bandwidth of the PFS shall be 200 GB/s.

3.4.4 The PFS shall have a minimum usable capacity given in Table 5.

3.4.5 The bandwidth shall be achievable using any arbitrary collection of compute nodes starting at 10% of compute nodes up to the full scale of the system when the PFS is up to 70% full. The Offeror shall describe the extent to which performance differs with the size of transfers or the number of files being read/written.

3.4.6 The PFS shall provide a robust, interactive environment for users. The time required to insert, delete, enumerate, and retrieve file system object status within a single directory on login, compute, or file transfer service nodes shall be prompt and not be substantially impacted by unrelated applications running on the rest of the compute partition. Given a single directory with 1 million files, describe how long the following metadata operations will take on the proposed file system to:

- Insert one million objects,
- Delete one million objects,
- Enumerate and retrieve one million objects.

3.4.7 The system shall provide POSIX I/O and MPI I/O functionality that is tightly integrated with file system software to provide high-performance small- and large-block I/O. MPI I/O shared file performance shall achieve 80% of POSIX I/O performance using a single file per processor at full system bandwidth.

3.4.8 The PFS system shall support access by external clients with the same functionality as internal clients. External clients are nodes, servers, or other systems not part of the delivered production NWSC-2 system. This would include, for example, the existing Yellowstone compute system.

3.4.9 The PFS shall be externally accessible with an aggregate external bandwidth as specified in Table 4.

3.4.10 The PFS shall be able to operate independently of the production NWSC-2 system and be capable of serving data to external clients when the production NWSC-2 system is down.

3.4.11 The Offeror shall describe expected PFS maintenance procedures and their impacts on the PFS performance under normal load and other routine operations including purging, file system health monitoring, performance statistics, problem alerts, diagnosis and repair, and reconstruction after a drive replacement, including reconstruction time.

3.4.12 The Offeror shall describe how the PFS design allows for upgrades in performance, capacity, or both. Specific options for upgrading are requested in §4.3.

3.5 Integration with Existing NWSC Data Services

3.5.1 Offeror must provide a solution that allows the HPC system to interoperate with the existing GLADE resource, providing an aggregate, sustainable bandwidth specified in Table 4. This bandwidth is in addition to the bandwidth provided for access to the NWSC-2 production system's PFS, and those specified in §3.5.2, and §3.5.3. Figure A1 in Appendix A provides a high-level diagram of the network architecture of the current GPFS deployment and integration with Yellowstone.

3.5.2 The system shall interoperate with NCAR's HPSS-based data archive providing an aggregate, sustainable bandwidth in excess of that specified in Table 4. This bandwidth is in addition to the bandwidth provided for access to the NWSC-2 production system's PFS, and those specified in §3.5.1, and §3.5.3.

3.5.3 The system shall interoperate with the NWSC local area network, providing an aggregate, sustainable bandwidth in excess of that specified in Table 4. This bandwidth is in addition to the bandwidth provided for access to the NWSC-2 production system's PFS, and those specified in §3.5.1, and §3.5.2.

3.6 Application Performance Requirements

Assuring that real applications perform well on the NWSC-2 platform is key to the success of the system. Because the full applications are large, often with millions of lines of code, NCAR has put together a suite of application kernels, referred to as the NCAR Benchmarking Suite (NBS), and micro-benchmarks for RFP response evaluation, capability improvement, and system acceptance. The NBS constituent applications, which are listed in Table 1, are representative of real applications, but are smaller in terms of the compute resources required to run them. As noted in Table 1, the NBS and input cases will be available no later than 2 months prior to the formal release of the RFP.

The performance of the micro-benchmarks listed in Table 2 will be evaluated as indicated in the table. All performance tests must continue to meet acceptance criteria throughout the lifetime of the system.

In addition to the NBS, and micro-benchmarks, the NCAR Simulation Suite (NSS) defined in Table 3 will be evaluated at acceptance and must demonstrate correct numerical solutions. The NSS and input cases will be made available at the same time as the NBS.

3.6.1 The Offeror shall provide performance results (actual, predicted and/or extrapolated) for the proposed systems for the NBS and micro-benchmarks listed in Table 1 and Table 2. If predicted or extrapolated results are provided, explain the methodology used in detail, including the algorithm(s) used.

The Offeror shall report all NBS and micro-benchmark results in the "Benchmark Run Rules and Results" worksheet. This, along with the benchmarks, input data sets and run rules will be available on the NWSC-2 Procurement website [3].

3.6.2 The Offeror shall provide licenses for the delivered system for all compilers and tools used to achieve benchmark performance. Licenses maybe be provided on a temporary basis if they are not part of the delivered production system.

3.6.3 The Offeror's proposal shall state a minimum Sustained Performance as measured by the geometric mean of the NBS constituent applications. The NWSC-2 system must meet or exceed the stated minimum Sustained Performance at acceptance. The target sustained performance is a 2x-5x increase over the Yellowstone platform [2]. The baseline Yellowstone performance for the NBS will be available from the NWSC-2 procurement website [3].

3.6.4 Several of the codes in the NBS will be used to judge Capability Improvement at acceptance. Using these, the NWSC-2 system shall achieve, on average, a 2x-5x Capability Improvement over the Yellowstone platform. The specific codes required for the Capability Improvement will be identified when the NBS is released.

- Capability Improvement is defined as the product of an increase in problem size and an application-specific runtime speedup factor. For example, if the problem size is 4 times larger and the runtime speedup is 1.25, the capability improvement is 5x.
- Yellowstone results shall be collected using the NBS components and will be provided to the Offeror.
- NWSC-2 results shall be collected using at least two-thirds (2/3) of the NWSC-2 system.
- Source code and example runtime rules for components of NBS to be used will be provided.

3.6.5 Following award, and prior to delivery of the production NWSC-2 system, Offeror will be required to validate that the production HPC system generates correct solutions on the NSS. These runs will be made on the same processor and interconnect technology that will be delivered on the production HPC system, but run at smaller scale. The validation runs will be performed jointly by Offeror and NCAR staff. The Early Access System (§4.6) can be used for this testing, provided it meets the criteria above.

Table 1. NCAR Benchmark Suite (NBS)

App Kernel Name	Description	Availability	RFP Response	Acceptance
CAM/HOMME	CESM communication kernel	Code and input cases available two months prior to RFP release	X	X
LES-K	Large Eddy Simulation kernel	Code and input cases available two months prior to RFP release	X	X
Morrison-Gettelman microphysics-K	Cloud microphysics/general circulation kernel	Code and input cases available two months prior to RFP release	X	X
Pyreshaper	Application I/O kernel	Code and input cases available two months prior to RFP release	X	X
POPPerf	Ocean circulation model	Code and input cases available two months prior to RFP release	X	X
RRTMP-k	Radiation transport kernel	Code and input cases available two months prior to RFP release	X	X
WRF	Weather Research and Forecasting model	Code available from: http://www2.mmm.ucar.edu/wrf/users/downloads.html Input cases available two months prior to RFP release	X	X
HPCG	High Performance Conjugate Gradient Solver	Code available from: https://software.sandia.gov/hpcg/ Input cases available two months prior to RFP release	X	X
LSQR Seismic Tomography	GPGPU benchmark	Code and input cases available two months prior to RFP release	X	X

Table 2. Micro-benchmarks

Benchmark	Description	Availability	RFP Response	Acceptance
STREAM	Memory bandwidth	Code available from: http://www.cs.virginia.edu/stream/ref.html Standard input cases	X	X
OMB (OSU Micro-Benchmark)	Interconnect performance	Code available from: http://mvapich.cse.ohio-state.edu/benchmarks/ Standard input cases	X	X
mdtest	Metadata performance	Code available from: http://mdtest.sourceforge.net/ Input cases available two months prior to RFP release	X	X
IOR	I/O latency and bandwidth	Code available from: http://sourceforge.net/projects/ior-sio/ Input cases available two months prior to RFP release	X	X

Table 3. NCAR Simulation Suite (NSS)

Application	Description	Availability
CESM	Community climate model	Code available from: https://www2.cesm.ucar.edu/models/current Input cases available two months prior to RFP release
MPAS	Community climate model	Code available from: http://mpas-dev.github.io/ Input cases available two months prior to RFP release
WRF	Weather forecasting	Code available from: http://www2.mmm.ucar.edu/wrf/users/downloads.html Input cases available two months prior to RFP release

3.7 Reliability, Availability, and Serviceability

For each metric specified below, the Offeror must describe how they arrived at their estimates. Terms used in this section can be found in the Definitions and Glossary section of this document.

3.7.1 The NWSC-2 production system shall meet or exceed 98% System Availability.

3.7.2 The NWSC-2 production PFS shall meet or exceed 99% System Availability.

3.7.3 The NWSC-2 production system's System Mean Time Between Interrupt (SMTBI) shall exceed 14 days.

3.7.4 Failure of the system management and/or RAS (Reliability, Availability, and Serviceability) system(s) shall not cause a system or job interrupt. This requirement does not apply to a RAS system feature which automatically shuts down the system for safety reasons, such as an overheating condition.

3.7.5 An immediate resume of an interrupted job shall not require a complete resource reallocation. If a job is interrupted, there shall be a mechanism that allows resume of the application using substantially the same allocation of resource, e.g. compute nodes, that it had before the interrupt.

3.7.6 A complete system initialization shall take no more than 30 minutes to start production jobs. The Offeror shall describe the full system initialization sequence and timings. System initialization is defined to be the time to initialize 98% of the compute resource including all peripherals and 100% of any service resource to the point where a job can be successfully launched.

3.7.7 The Offeror shall discuss the RAS mechanisms and capabilities of the proposed system including, but not limited to:

- Any condition or event that can potentially cause a job interrupt
- Resiliency features to achieve the availability targets
- Single points of failure, hardware or software, and the potential effect on running applications and system availability
- How a job maintains its resource allocation and is able to relaunch an application after an interrupt
- How is efficiency and availability monitored and calculated when the system hangs

3.8 System Operations

System operation capabilities provide the ability to effectively manage system resources with high utilization and throughput under a workload with a wide range of concurrencies. System management must be an integral part of the overall system. The overall objective is to provide system administrators, security officers, and user-support personnel with productive and efficient system configuration management and an enhanced diagnostic environment.

3.8.1 The Offeror shall describe and provide scalable integrated system management and montoring capabilities that provide: human interfaces and APIs for system configuration and its ability to be automated; software management; change management; local site integration; and system configuration.

3.8.2 The Offeror shall describe and provide a means for centralized management, tracking and analyzing all software updates, software and hardware changes and failures, and hardware replacements over the lifetime of the system. Notwithstanding upgrades required for reasons such as security, it is highly desireable that all system software must have full support for the life of the machine without need for complete reinstall or major service interruptions.

3.8.3 The system management capabilities shall provide a single, scalable log analysis capability for all logs originating from any component of the proposed system.

3.8.4 The Offeror shall describe system configuration management and diagnostic capabilities, addressing the following topics:

- Detailed description of the system management support
- Any effect or overhead of software management tool components on the CPU, network, or memory available on compute nodes
- Release plan, with regression planning, testing and validation, for all system related software and security updates
- All updates must allow previous version and next version to be running on production system for updates that do not require outages
- Support for multiple simultaneous or alternative system software configurations, including estimated time and effort required to install both a major and a minor system software update
- User activity tracking, such as audit logging and process accounting

3.8.5 CISL system administrators shall have unrestricted privileged access to all hardware components delivered with the system.

3.8.6 All critical components must provide hot-swap redundancy for all services that are required for normal production operations.

3.8.7 The Offeror shall provide CISL with a complete plan for support of offline backups and bare metal recovery of all components and systems.

3.9 Buildable Source Code

3.9.1 Source code, and necessary build environment, shall be provided for all software except for firmware, compilers and third-party products. Exceptions will be granted for vendor-supplied software that is proprietary in nature and/or where copyrights do not permit distribution of source.

3.9.2 Updates of source code, and any necessary build environment, for all software shall be provided over the life of the subcontract.

3.10 Test System

The Offeror shall propose a test system. The system shall contain all the same hardware and functionality of the main system, including file systems, but scaled down to an appropriate configuration.

3.10.1 The test system shall have at least 128 compute nodes and 50 TB of storage.

3.10.2 The test system and the main system shall not share any internal or external communications paths, peripherals, storage, power supplies or subsystems required for normal operation.

3.10.3 All procedures for the production system shall be testable on the test system, including upgrades, patches, etc.

3.10.4 Depending on the specific support requirements (e.g., power and cooling), the test system may be installed at the NCAR Mesa Lab in Boulder.

3.10.5 The test system shall be delivered to and installed at the designated NCAR facility at least 30 days prior to initial production system equipment delivery.

3.11 Facilities and Site Integration

Notional diagrams of site network connectivity and use can be located in Appendix A of this document. These are intended to assist Offerors in understanding the integration requirements for NWSC-2 and related systems. These are subject to change prior to the NWSC-2 deployment. Additional integration details will be made available at the NWSC-2 procurement site [3].

3.11.1 The computational system shall use 3-phase 480V AC. Other power sources are available to support the system's infrastructure (disks, switches, consoles).

3.11.2 All equipment and power control hardware shall be Nationally Recognized Testing Laboratories (NRTL) certified. All equipment shall bear appropriate NRTL labels. All equipment shall comply with the IEEE, NEC and NFPA environmental standards and codes as referenced in Section 7, particularly for Performance Level 2 systems as defined in IEEE 1156.2-1996. All proposed equipment shall comply with the new Class 1 & 2 recommended operating environment range as specified in the 2008 ASHRAE Environmental Guidelines for Datacom Equipment

3.11.3 Every rack, network switch, interconnect switch, node, and disk enclosure shall be clearly labeled with a unique identifier visible from the front of the rack and/or the rear of the rack, as appropriate, when the rack door is open. These labels will be high quality to be usable and readable throughout the lifetime of the system. Nodes shall be labeled on the service-access side with a designator matching that assigned via administrative software and a unique serial number for inventory tracking.

3.11.4 The Offeror shall describe the features of the system related to facilities and site integration including:

- Remote environmental monitoring capabilities of the system and the proposed integration into facility monitoring.
- Detailed descriptions of power and cooling distributions throughout the system including power consumption and cooling requirements for all subsystems, at idle, observed maximum (e.g., HPL), and design limit states.
- Detailed descriptions, quantities and types of all electrical, mechanical connections made to facility infrastructure.
- OS distributions or other client requirements to support off-platform access. This would include, for example, remote management and troubleshooting of hardware and software components

3.11.5 The Offeror shall provide a description of facility and installation planning services with their proposal. The description shall include the facility preparation and planning processes to be conducted with UCAR, including shipping, receiving and staging of NWSC-2 equipment and all on-site assembly thereof, as well as all logistics information, including crated and uncrated sizes, weights and floor loading.

3.11.6 The Offeror shall provide transportation, delivery and installation of all NWSC-2 equipment as well as replacement and spare parts. The Offeror shall provide unpacking, uncrating, assembly and interconnection of the NWSC-2 system components at the NWSC facility in Cheyenne, WY, (and the NCAR Mesa Laboratory facility in Boulder, CO, if necessary). The Offeror shall remove all packing materials and trash associated with delivery and installation.

Table 4. NWSC-2 Facility Requirements

NWSC-2	
Location	NCAR-Wyoming Supercomputer Center, Cheyenne, Wyoming. The system will be housed in NWSC Module A, which will be built out to support the system.
Altitude	6,260 feet
Seismic	N/A
Water Cooling	The system must operate within ASHRAE TC 9.9 temperature ranges. Total flow requirements may not exceed 6KGPM or a differential pressure of 25PSI at the system cabinets. De-ionized water is available.
Air Cooling	The system must operate within ASHRAE TC 9.9 temperature ranges. The total required airflow must not exceed 60K CFM. No more than
Maximum Power	3 MW
Maximum Power Rate of Change	No restrictions.
Floor	10’ raised floor
Ceiling	12’ ceiling; maximum cabinet height is 9’9”
Maximum Footprint	64’x70’, or 4480 square feet (inclusive of compute, storage and service aisles)
Shipment Dimensions and Weight	For delivery, system components shall weigh less than 7,000 pounds. All doors and pathways are 6’ 0” in width and 9’ 9” in height or larger.
Floor Loading	The floor loading shall not exceed a uniform live load of 250 pounds per square foot with a deflection of not more than 0.04 inch, and a concentrated load of 2,500 pounds on

NWSC-2	
	one square inch
Cabling	All power cabling and water connections shall be below the access floor. All other cabling (e.g. system interconnect) should be above floor and integrated into the system cabinetry. All cables shall be plenum rated. All signal cables shall be labeled with a unique serial number at both ends. All communications cables, wherever installed, shall be source/destination labeled at both ends.
External network interfaces supported by the site for connectivity requirements specified below	1GbE, 10GbE, 40GbE, 100GbE
Aggregate external bandwidth on/off the system for general TCP/IP connectivity	20 GB/s
External bandwidth on/off the PFS for access by other systems (e.g., existing Yellowstone)	100 GB/s
Aggregate external bandwidth on/off the system for accessing NCAR's external, GLADE file system (GPFS)	100 GB/s
Aggregate external bandwidth on/off the system for accessing external NCAR archive (HPSS)	12 GB/s

3.12 Target System Configurations

Table 5. Target Configuration and Performance Requirements

NWSC-2	
NCAR Benchmarking Suite sustained performance increase over Yellowstone [3] system	2-5x
Minimum memory on a compute node. This metric is for main memory capacity only, e.g. DDR4. It does NOT include memory associated with caches, accelerators, scratch pads, etc.	64 GB
Disk Capacity of the system's parallel file system	20 PB
FLOPS/BYTE used to determined the	250 FLOPS/BYTE

NWSC-2	
bandwidth of the system's parallel file system	
Parallel Debugger Licenses	20 simultaneous users; A single job up to one-fourth (1/4) scale
Compiler licenses for each compiler suite proposed	50
Resource Manager and/or Scheduler Licenses	Licenses must be provided

3.13 Maintenance, Support, and Technical Services

The Offeror shall propose maintenance and support with the following minimum features.

3.13.1 Pricing and the Maintenance Period

The Offeror shall propose technical services, warranty, and Maintenance and Support for a period four (4) years subsequent to the date of Acceptance of the NWSC-2 system by UCAR. The Maintenance and Support pricing shall be for each year of the above period after the warranty expires. The warranty period begins at the date of Acceptance of the system.

3.13.2 Required Maintenance and Support

The Offeror shall, at a minimum, provide replacement hardware for all failed components and return shipping of failed components to the Offeror. The Offeror shall train and certify NCAR staff to perform hardware failure diagnosis, isolation and repair activities, provide Offeror employees or contractors to perform those activities, or a combination thereof. The Offeror shall supply hardware maintenance procedural documentation, training, and manuals. No component may fail and remain out of service longer than ten (10) calendar days unless mutually agreed to by the Offeror and UCAR.

3.13.3 On-site Parts Cache

The Offeror shall maintain a parts cache on-site at the NWSC facility. The parts cache shall be sized and provisioned sufficiently to support all normal repair actions for two weeks without the need for parts refresh. The initial sizing and provisioning of the cache shall be based on Offeror's MTBF estimates for each Field Replaceable Unit (FRU) and each rack, and scaled based on the number of FRU's and racks delivered. The parts cache configuration will be periodically reviewed for quantities needed to satisfy this requirement, and adjusted if necessary, based on observed FRU, component or node failure rates. The parts cache will be resized, at the Offeror's expense, should the on-site parts cache prove to be insufficient to sustain the actually observed FRU or node failure rates.

3.13.4 Hardware and Software Protection Plan

The Offeror shall support and maintain the hardware comprising the NWSC-2 system and the software stack delivered with the NWSC-2 system for the duration of the base subcontract plus any extensions.

3.13.5 Software and Firmware Update Service

The Offeror shall provide an ability for UCAR to update all software and firmware provided with the NWSC-2 system. This shall include new releases of software/firmware and software/firmware patches as required.

3.13.6 Problem Reporting and Resolution Service

The Offeror shall provide 24x7 telephone and web-based problem reporting and resolution services.

3.13.7 Production Transition Support

The Offeror shall provide remote and/or on-site assistance to UCAR to transition the production workload and key production applications within NCAR's production workload to the NWSC-2 system from its predecessor system. This shall include system and file system configuration and tuning, software problem isolation and resolution, and application tuning and optimization. This support shall be provided from subcontract execution through six months after system Acceptance.

4 Technical Options

This section contains options to the base system described in §3. Some options may represent collaborations between NCAR, with the Offeror providing functionality that doesn't currently exist, or is inadequate in the current marketplace, but could be delivered after acceptance. Final quantities and delivery dates for Technical Options shall be negotiated at the time of award. Table 6 provides Offerors with guidance to assist them in developing their responses to each of the specific technical options.

Table 6. Technical Options Overview

Technical Option	Guidance
§4.1 Many-core partition	Development work may be required. In proposing a delivery date for the many-core component, Offeror should consider market maturity, reliability, and programming model, among other criteria.
§4.2 General Purpose GPU	General market availability of the GPGPU is assumed. If exercised, some quantity is expected at initial system delivery. Subsequent quantities are possible.
§4.3 Parallel File System	Upgrade to performance and/or capacity of PFS. Specific quantities and delivery date will be negotiated at award. Subsequent quantities are possible.

§4.4 Visualization, Data Analysis and Post Processing	General market availability assumed. If exercised, some quantity is expected at initial system delivery. Subsequent quantities are possible.
§4.5 Software Tools and Programming Environment	General market availability assumed. Tools that are part of Technical Options that require development may be deployed with corresponding technology.
§4.6 Early Access System	Minimum configuration for NCAR Simulation Suite due 6-months prior to initial system delivery. Other configurations will be negotiated as part of specific technical options that are exercised.
§4.7 Innovative Storage and Memory Technologies	Development work may be required. Offeror should propose a delivery date consistent with general market availability.
§4.8 Maintenance, Support, and Technical Services	Enhanced services beyond the base §3.13 services. Offeror should propose incremental enhancements, if applicable, as described.
§4.9 Upgrades and Expansion to the NWSC-2 System	Based on the delivered system configurations. Options could be negotiated at any time during life of the contract.

The Offeror shall provide all relevant technical, business and price information (as defined in §2) for all options listed below. Options shall be priced separately and shall include delivery, installation, warranty, maintenance, and support for the life of the system. Pricing shall be firm fixed prices. The technical, business and price information for vendor-proposed options will be evaluated during the selection process and represent a critical element of the overall Offeror proposal.

Beyond the options described below, Offeror is encouraged to propose additional areas of collaboration that they feel provide substantial value to the NWSC-2 system and its user community.

4.1 Many-core Compute Partition

An emerging class of processors, so-called many-core, have the potential to address the power and performance challenges in moving towards exascale computing. Characteristics of these processors may include: self-hosted, i.e., unlike co-processors, they are not accessed via the PCI bus from a host processor; provide binary compatibility with traditional processors allowing developers to maintain a single application source tree and deploy on both traditional and many-core systems; and they may have an additional level of high-speed local memory that improves memory bandwidth. In the time frame of the anticipated NWSC-2 delivery, CISL does not expect that most of its production workload, which is composed primarily of climate and weather models, will be able to fully leverage these emerging architectures. However there are significant efforts underway by CISL and the model development teams to transition the models to, and evaluate model performance on, these future processors. Therefore, while the production

NWSC-2 system should support the anticipated production workload, CISL will be investing in sufficient capabilities through technical options to expand the system in support of these model transition and evaluation efforts.

4.1.1 The Offeror shall describe the node-level architecture for the many-core partition. It is highly desirable that the many-core partition be based on the same node-level architecture as the main compute nodes. If the Offeror determines that the proposed compute node architecture is not consistent with the roadmap of the many-core processor, the Offeror shall describe the alternative architecture(s).

4.1.2 The Offeror shall describe how the many-core components will be integrated into the system. It is highly desirable that these components be integrated on the same high-speed network as the main compute resources and have equal access to all other system resources, e.g., file systems and storage. If this is not possible, Offeror should explain why and what the performance implications are of having them on an independent network.

4.1.3 The Offeror shall describe how the many-core partition affects the scalability of the main system partition and impact the high-speed interconnect of the cluster. The Offeror should address the scalability of the different partitions as well as the combined scalability of the integrated system.

4.1.4 Offeror shall provide options for the following sizes (in terms of node count) of many-core partitions:

4.1.4.1 5% of the total system is comprised of many-core nodes.

4.1.4.2 10% of the total system is comprised of many-core nodes.

4.1.4.3 20% of the total system is comprised of many-core nodes.

4.1.5 All application performance requirements in §3.6 shall apply to the many-core partition. If the many-core partition is delivered at a date later than the primary NWSC-2 production system, the contribution to the sustained performance (NBS, §3.6.3), and the Capability Improvement (NSS, §3.6.4) shall be applied at the time of acceptance of the many-core partition.

4.2 General Purpose GPU Compute Partition

CISL provides support applications that make use of general purpose GPUs (GPGPU). Offeror shall provide options for a GPGPU compute partition as follows:

4.2.1 The Offeror shall describe the node-level architecture for each GPGPU partition. It is highly desirable that the partition be based on the same node-level architecture as the main compute nodes. If the Offeror determines that the proposed compute node architecture is not consistent with the roadmaps of the GPGPU, the Offeror shall describe the alternative architecture(s).

4.2.2 The Offeror shall describe how the GPGPU components will integrated into the system. It is highly desirable that these components be integrated on the same high-speed network as the main compute resources and have equal access to all other system resources, e.g., file systems and storage. If this is not possible, Offeror should explain why and what the performance implications are of having them on an independent network.

4.2.3 The Offeror shall describe how the GPGPU partition affects the scalability of the main system partition and impacts the high-speed interconnect of the cluster. The Offeror should address the scalability of the different partitions as well as the combined scalability the integrated system.

4.2.4 Offeror shall provide options for the following sizes (in terms of node count) of GPGPU partitions:

4.2.4.1 5% of the total system is comprised of GPGPU nodes

4.2.4.2 10% of the total system is comprised of GPGPU nodes

4.2.4.3 20% of the total system is comprised of GPGPU nodes

4.2.5 All of application performance requirements in §3.6 shall apply to the GPGPU partition. The application in Table 1 identified as GPGPU-specific shall be required for the GPGPU, while the others shall be required for the host processor. If the GPGPU partition is delivered at a date later than the primary NWSC-2 production system, the contribution to the sustained performance (NBS, §3.6.3), and the Capability Improvement (NSS, §3.6.4) shall be applied at the time of acceptance of the GPGPU partition.

4.3 Parallel File System Upgrade

The initial storage deployment specified in Table 5 is considered sufficient in capacity and performance to accommodate the major application use cases during the first two years of production. However, it is likely that additional storage capacity and performance will be required and thus, Offeror shall provide options for upgrading storage in accordance with the design described in §3.4.

4.3.1 Options will be described as scalable units (SU) of which NCAR may purchase in variable quantities. Any additional infrastructure required to support the SU will either be included in the SU or identified as additional requirements to support the SU.

4.3.2 Offeror shall describe and provide an option that increases the capacity of the initial storage deployment in a balanced way by 50% and 100% without necessarily increasing performance. Describe any impact on performance of adding capacity only.

4.3.3 Offeror shall describe and provide an option that increases both bandwidth performance and capacity of the initial storage deployment by 50% and 100% over the initial deployment. I.e., one option where performance and capacity increases by 50% each; and one option where performance capacity increases by 100% each.

4.4 Visualization, Data Analysis and Post-processing

The system shall be capable of supporting post-processing, visualization and analysis workloads, described as follows:

1) *Post-processing:* reading simulation outputs stored on the parallel file system(s), performing operations on the data, and writing the new, derived outputs back to the file system(s). Post-processing tasks are typically memory- and/or CPU-intensive, and almost always perform substantial I/O. Few post-processing applications support distributed-memory parallelism, but more may support shared memory. Scalability is typically achieved by simultaneously running multiple serial instances across one or more nodes.

2) *Batch and interactive 2D visualization:* reading simulation or post-processing outputs, transforming the data into graphical representations, and displaying (via X11) or storing the resulting imagery.

3) *Interactive 3D visualization:* interactively exploring data residing on the parallel file system with advanced, OpenGL-based visualization applications. Rendering, and in some cases general-purpose computing, may be performed on one or more GPUs per node. Remote display and interaction between the visualization resource and the geographically distant user is accomplished via image-based remote display technology (e.g., VirtualGL).

In summary, data analysis and visualization workloads require substantial compute, memory, and GPU resources, and especially must support good bandwidth into the parallel file system.

4.4.1 The Offeror shall describe how/confirm that the visualization resources shall support the following packages:

- VAPOR - https://www.vapor.ucar.edu/
- VisIt - https://wci.llnl.gov/simulation/computer-codes/visit/
- Kitware's ParaView (open source).
- VirtualGL
- NCL - http://www.ncl.ucar.edu/
- Matlab
- IDL

These packages use some or all of the following system software capabilities: full support of sockets, dynamic linked libraries, POSIX threads, Python scripting, MPI, OpenGL, NumPy, NetCDF, HDF5 and MPI I/O. NCAR will assume responsibility for porting of NCAR packages and separately provide any required licensing for third-party visualization applications.

4.4.2 It is highly desirable that the visualization/data analysis partition(s) have a similar node-level architecture (e.g., CPU, motherboard) as the main compute nodes. If the Offeror determines that the proposed compute node architecture is not consistent with the roadmaps of all the visualization packages listed above, the Offeror shall propose an alternative architecture that is consistent.

4.4.3 The visualization/data analysis resource shall be tightly integrated into the system, sharing the same high-speed network as the main compute resources and have equal access to all other system resources, e.g., file systems and storage.

4.4.4 Offeror shall describe and provide an option for a visualization/data analysis partition that is nominally 0.5% of the total number of compute nodes. If the Offeror proposes a GPGPU partition that is also suitable as a visualization resource, they must provide suitable application and performance information to support this.

4.4.5 The visualization/data analysis partition shall have a minimum of 256 GB of memory per node.

4.4.6 Offeror shall describe and provide an option for the visualization/data analysis partition with a minimum of 1 TB of memory per node.

4.5 Software Tools and Programming Environment

The Offeror shall provide information about software tools and programming environment software in the following product categories. The Offeror may choose to respond either by identifying compatible or available third-party offerings (for CISL to evaluate separately) or by describing and providing full technical options for software that CISL may elect to execute as part of the NWSC-2 procurement.

Descriptions of the software tools should reflect that the primary programming model used by application scientists running in production on the NWSC-2 systems will be the Message Passing Interface (MPI) with hybrid MPI/OpenMP parallelism used by key applications. However, significant development activity for NCAR's climate and weather models is expected to target the transition to programming models on the many-core partitions proposed by the Offeror in §4.1.

4.5.1 The Offeror shall identify or describe additional options for a vendor-supported job-scheduler/resource manager, consistent with the details requested in §3.2.

4.5.2 The Offeror shall identify or describe additional available or supported MPI 3.0 implementations, consistent with the details requested in §3.3.

4.5.3 It is highly desirable for the production NWSC-2 system to support multiple compilation environments. The Offeror shall identify or describe all compilation suites and languages, including version numbers and integrated components (such as debuggers or profiling tools), that are compatible with the proposed system.

4.5.4 The Offeror shall identify or describe the interactive debuggers or debugging suites (available separately from a previously described compiler or tool suite) with a graphical user interface that are compatible with the proposed system and provide a single-point of control for debugging applications using all granularities of parallelism and programming environments provided by the system.

4.5.5 The Offeror shall identify or describe the tools available for the proposed system that provide detailed performance analysis and profiling of user applications (available separately from a previously described compiler or tool suite). The tools shall support all granularities of parallelism and the programming environment of the system. The Offeror shall describe all tools supported and any limitations, e.g. limits on scalability.

4.5.6 The Offeror shall identify or describe the tools available for the proposed system (separate from a previously described compiler or tool suite) that provide event-tracing capabilities and any limitations. Events of interest include Message-Passing Event Tracing, I/O Event Tracing, Floating Point Exception Tracing, and Message-Passing Profiling.

4.5.7 The Offeror shall identify or describe the stack-tracing tools available for the proposed system (separate from a previously described compiler or tool suite). The tools identified should include a source-level stack trace back, including an API that allows a running process or thread to query its current stack trace.

4.6 Early Access System

To allow for early and/or accelerated development of applications, development of functionality required as a part of the statement of work, or for the correctness validation of the NCAR Simulation Suite results (§3.6.5), the Offeror shall propose options for early access development systems (EAS). These can be either systems delivered to NCAR or dedicated access to remote systems. The systems can be in support of the baseline requirements or any proposed options. EAS(s) systems shall be delivered or made available six months prior to the delivery of hardware or software that is planned for production.

4.6.1 The Offeror shall propose Early Access System(s). The primary purpose is to expose the application to the same programming, runtime, and system software environment as will be found on the final system, and to provide access to hardware or software that may be proposed as part of required or technical options. Except for tests described in §3.6.5, it is acceptable for the early access system to **not** use the final processor, node, or high-speed interconnect architectures. However, the programming, runtime, and operating system environment must be sufficiently similar that a port to the final system is trivial. The early access system shall contain similar functionality of the final system, including file systems, but scaled down to the appropriate configuration.

4.6.2 The specific size of the EAS shall be negotiated with Offeror and will be based on details of the specific Technical Option awarded.

4.7 Innovative Storage and Memory Technologies

NCAR is interested in new storage and memory technologies that have the potential to dramatically improve the performance of NCAR's user applications. These may include, for example: stacked memory, non-volatile memory, burst buffers, hybrid SSD/HDD storage systems, processor-in-memory, or software enhancements that leverage these new technologies. These innovative technologies may be deployed subsequent to the initial system deployment, but no later than one year following Acceptance of the NWSC-2 production system. It is understood that long-term projections of the performance of these technologies is difficult given the level of development work that may exist between now and the expected deployment of them. Therefore, NCAR will negotiate these acceptance criteria at the time a decision is made to exercise these options. The Offeror shall propose separately priced options for innovative storage and memory technologies as follows:

4.7.1 Storage and Memory Technology Description

Offeror shall fully describe the technology, including all relevant architectural details, how it will be integrated into the production HPC system, many-core, or GPGPU partitions, and the expected performance characteristics, software programming model, and reliability. Where NVM technologies are proposed, Offeror shall address endurance or any other attribute where performance is expected to degrade or substantially change with time or I/O patterns.

4.7.2 Cost/benefit of Storage and Memory Technology

Offeror shall provide a cost/benefit analysis of deploying the proposed technology. For example, if a burst buffer is proposed, Offeror shall provide the rationale for the amount, and to what extent it will reduce the amount of storage required in other parts of the system.

4.7.3 Market Maturity of Storage and Memory Technology

Offeror shall provide information that can be used to assess the state of the technology development and adoption at the expected time of contract award, and at the time it is proposed to be deployed in the NWSC-2 system. This can include, for example: current or planned installations of the technology; committed market release dates; data or other results that indicate the current state of development and the risks that remain in bringing it into production.

4.7.4 Performance of Storage and Memory Technology

Offeror shall estimate the expected performance improvements to NCAR's application workload and propose a set of benchmarks that could be used to demonstrate these performance benefits.

4.8 Maintenance, Support, and Technical Services

The Offeror shall propose separately priced maintenance and support options with the following enhanced features beyond the basic set proposed in response to §3.13:

4.8.1 Maintenance and Support Options

The Offeror shall propose the following Maintenance and Support Options. UCAR may purchase or execute one of the Options or none of the Options at its discretion. Different maintenance options may be selected for the test systems and the production NWSC-2 system. Additionally, a combination of options may be used for components of the NWSC-2 system, for example, 7x24 for components whose failure would induce a system interrupt and 5x9 for other components. Each Option shall be separately priced. The Offeror may propose other maintenance and support solutions in addition to Options 1 and 2 below.

4.8.1.1 Option 1 – 7x24

The Offeror shall price Option 1 as full on-site hardware maintenance for Offeror-provided hardware components. The principal period of maintenance (PPM) shall be for 24 hours by 7 days a week with a four hour response to any request for service.

4.8.1.2 Option 2 – 5x9

The Offeror shall price Option 2 as full on-site hardware maintenance for Offeror-provided hardware components. The principal period of maintenance (PPM) shall be on a 9 hours by 5 days a week (exclusive of holidays observed by NCAR).

4.8.2 Software and Firmware Update Service

The Offeror shall provide an update service for software and firmware delivered with the NWSC-2 system. Offeror personnel in cooperation with UCAR staff shall perform this update service. This shall include new releases of software/firmware and software/firmware patches as required.

4.8.3 Call Service

If not already included by the Offeror in the support services proposed in response to §3.13.*, the Offeror shall propose an option which would make key technical personnel with knowledge of the proposed equipment and software available to NCAR. These personnel shall be available for consultation by telephone and electronic mail with NCAR personnel. In the case of degraded system, file system or application performance, Offeror's services shall be made readily available to develop strategies for improving performance, i.e. patches, workarounds.

4.8.4 On-Site Node Cache

The Offeror shall maintain an on-site spare node inventory of at least 1% of the total nodes in all of the system. These nodes shall be periodically maintained and tested for hardware integrity, functionality and compatibility with the NWSC-2 system and its software and firmware by the Offeror.

4.8.5 Extended Services, Warranty, Maintenance, and Support

The Offeror shall propose separately priced options to extend all technical services, warranty, and Maintenance and Support for each of two additional years beyond the end of the initial 4-year maintenance period.

4.9 Upgrades and Expansions to the NWSC-2 system

It is possible that NCAR will have future requirements for system upgrades and/or additional quantities based on the configurations proposed for NWSC-2. To address these potential requirements, the Offeror shall propose separately price options for system upgrades and expansions as indicated below. The Offeror shall address any technical challenges foreseen with regard to scaling and any other production issues. Proposals may be as detailed as possible based on proffered solutions. Proposals that do not address all of the additional system options in a materially responsive manner will be downgraded.

4.9.1 Upgrade/expand the NWSC-2 configuration by the following fractions of the proposed systems as measured by the NCAR Benchmarking Suite and/or the Capability Improvement metrics:

4.9.1.1 10%

4.9.1.2 25%

4.9.1.3 50%

5 Delivery and Acceptance Requirements

Testing of the system shall proceed in three steps: pre-delivery, post-delivery and acceptance. Each step is intended to validate the system and support subsequent activities. A sample acceptance test plan is include in Appendix B.

5.1 Pre-delivery Testing

NCAR and vendor staff shall perform pre-delivery testing at the factory on the hardware to be delivered. Any limitations for performing the pre-delivery testing need to be identified including scale and licensing limitations. During pre-delivery testing, the successful Offeror (Subcontractor) shall:

- Demonstrate RAS capabilities and robustness, using simple fault injection techniques such as disconnecting cables, powering down subsystems, or installing known bad parts.
- Demonstrate functional capabilities on each segment of the system built, including the capability to build applications, schedule jobs, and run them using the customer-provided testing framework. The root cause of any application failure must be identified.
- Provide a file system sufficiently provisioned to support the suite of acceptance tests.
- Provide onsite and remote access for NCAR staff to monitor testing and analyze results.
- Instill confidence in the ability to conform to the statement of work.

5.2 Site Integration and Post-delivery Testing

NCAR and vendor staff shall perform site integration and post-delivery testing on the fully delivered system. Limitations may exist for vendor access to the onsite system.

- During post-delivery testing, the pre-delivery tests shall be run on the full system installation.
- Where applicable, tests shall be run at full scale.

5.3 Acceptance Testing

NCAR and vendor staff shall perform onsite acceptance testing on the fully installed system.

- The Subcontractor shall demonstrate that the delivered systems conform to the subcontract's Statement of Work. A sample Acceptance test plan is provided in Appendix B as a basis for responding to this RFP.

6 Risk Management and Project Management

The Offeror's proposal shall:

6.1.1 Provide a risk management strategy for the proposed system in case of technology problems or scheduling delays that affect availability or achievement of performance targets in the proposed timeframe. Describe the impact of substitute technologies on the overall architecture and performance of the system as described in §2.1.1. In particular, the Offeror shall address the three technology areas listed below.

- Processor
- Memory
- High-speed interconnect, both hardware and software
- Technical Options

6.1.2 Identify any other high-risk areas and accompanying mitigation strategies for the proposed system.

6.1.3 Provide a clear plan for effectively responding to software and hardware defects and system outages at each severity level and document how problems or defects will be escalated.

6.1.4 Provide a roadmap showing how the response to this procurement aligns with their plans for high-performance computing over the next 5-10 years.

6.1.5 Discuss additional capabilities including the Offeror's:

- Ability to produce and maintain the proposed system for the life of the platform
- Ability to achieve specific quality assurance, reliability and availability goals
- In-house testing and problem diagnosis capability, including hardware resources at appropriate scale

6.1.6 A sample project management plan is included in Appendix C.

7 Documentation and Training

The Offeror shall provide documentation and training for the proposed solution to the operators, system administrators, and users of the NWSC-2 system in order to effectively operate, configure, monitor and use the platform. NCAR may, at its option, make audio and video recordings of presentations from Offeror's speakers at public events targeted at the NCAR user community. Offeror grants NCAR use and distribution rights of vendor-provided documentation, session materials and recorded media to be shared with NCAR staff and all authorized users and support staff for NWSC-2.

7.1 Documentation

7.1.1 The Offeror shall provide documentation for each delivered system describing the configuration, interconnect topology, labeling schema, hardware layout, etc., of the system as deployed before the commencement of system acceptance testing.

7.1.2 The Offeror shall supply and support system-level documentation necessary for operation and maintenance of the system for all components of the system.

7.1.3 The Offeror shall supply user-level documentation that can be shared with all authorized users of NWSC-2 for all user-accessible software tools and programming environment components. The Offeror shall describe any limitations on the distribution of these materials.

7.1.4 The Offeror shall describe how system-level and user-level documentation will be distributed and updated electronically (e.g., as electronic documents for inclusion on the CISL web site, or as electronic documents accessed via the Offeror's support site).

7.1.5 All documentation shall be updated in a timely manner. Changes to the system shall be accompanied by relevant documentation updates. Documentation of changes and fixes may be distributed as release notes. Full reference manuals may be updated later, but should be no older than two minor releases behind the installed version.

7.2 Training

7.2.1 The Offeror shall provide the following types of training at facilities specified by NCAR:

Class Type	# Classes annually	
	1st Year	Subsequent Years
Customer-replaceable hardware support and service	3	1
System Operations and Advanced Administration	2	1
Application Programming and Runtime Optimization	4	1

7.2.2 The Offeror shall describe how all proposed training relevant to the systems will be delivered (offsite classroom training, onsite training, or online training).

8 References

[1] https://www2.cisl.ucar.edu/resources/yellowstone

[2] https://www2.cisl.ucar.edu/sites/default/files/NWSC_Science_Objectives.pdf

[3] https://www2.cisl.ucar.edu/NWSC-2

Facilities Interfaces

Figure A1: Current NCAR HPC Environment

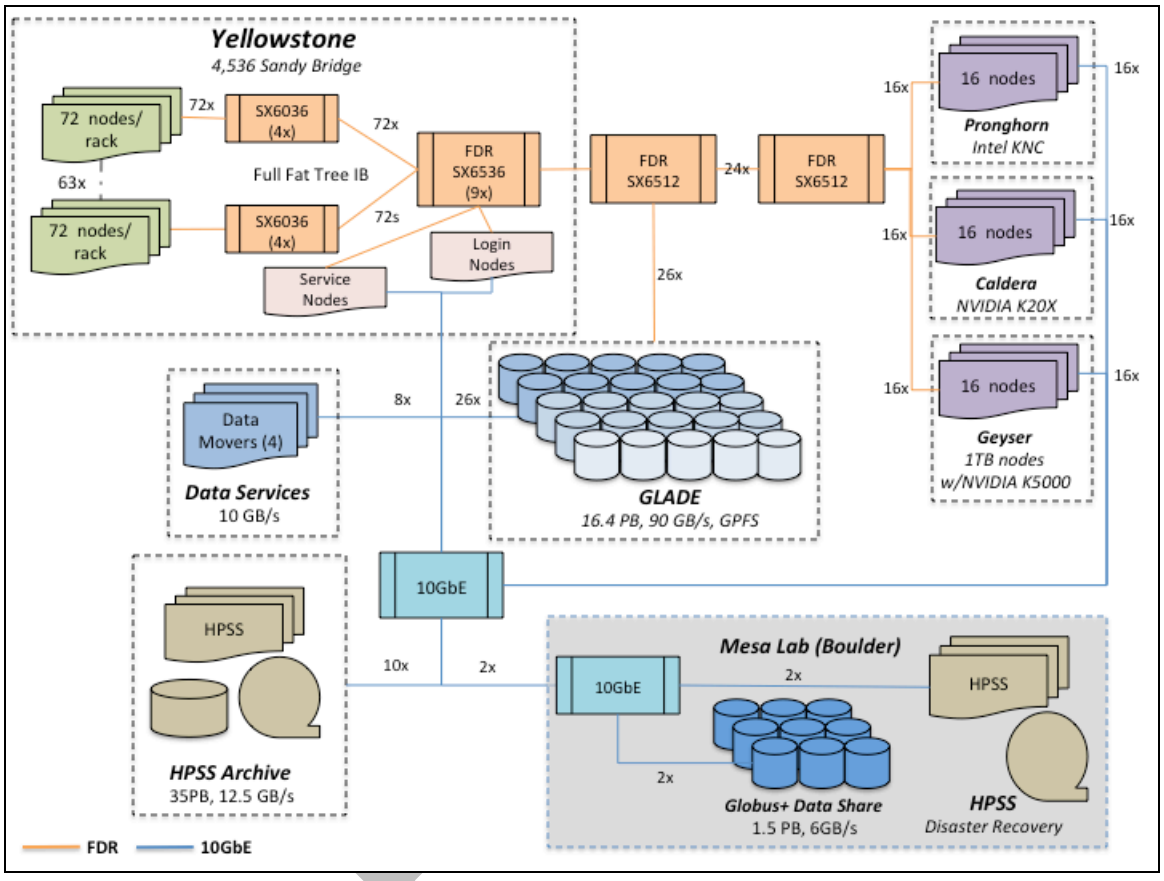
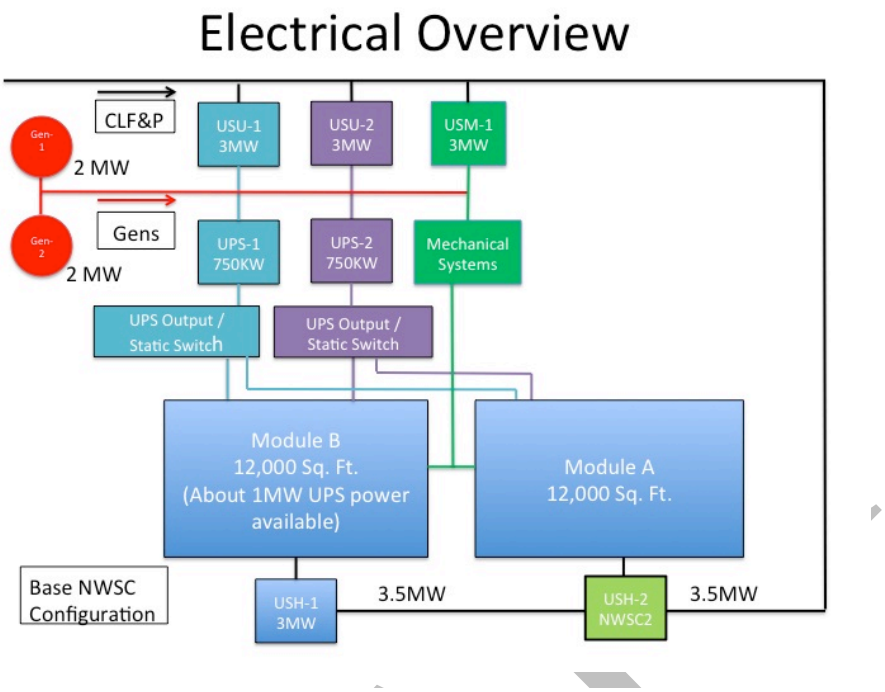
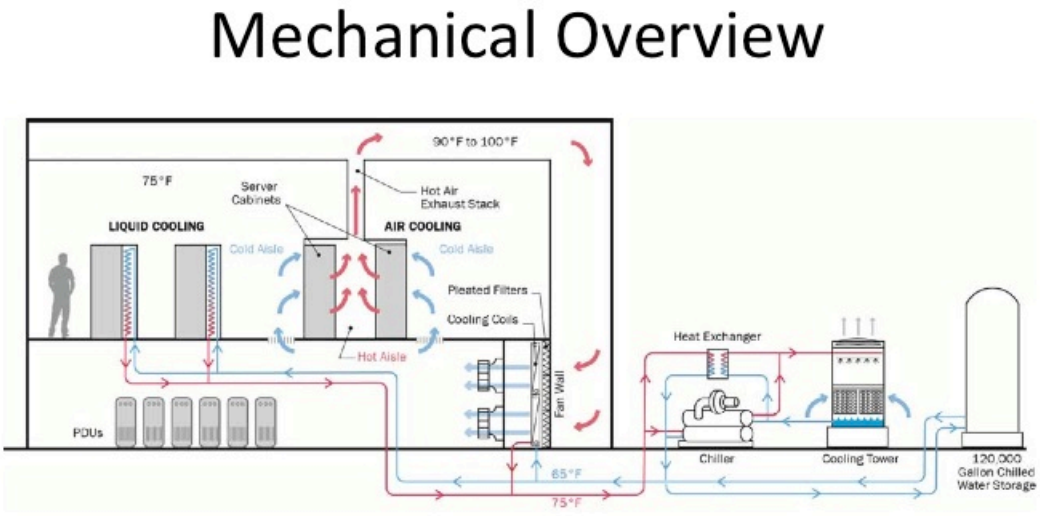


Figure A2: Current NWSC Module B Electrical and Mechanical Environment





Appendix A: Sample Acceptance Plan

Pre-delivery

The Subcontractor shall demonstrate all hardware is fully functional prior to shipping. If the system is to be delivered in separate shipments, each shipment should undergo pre-delivery testing. NCAR recognizes that Early Access Systems are not part of the pre-delivery acceptance criteria.

Pre-Delivery Assembly

The Subcontractor shall perform the pre-delivery test of the NWSC-2 system or agreed-upon sub-configurations of NWSC-2 at the Subcontractor's location prior to shipment. At its option, NCAR may send representatives to observe testing at the Subcontractor's facility. Work to be performed by the Subcontractor includes:

- All hardware installation and assembly
- Burn-in of all components
- Installation of Offeror-supplied software
- Implementation of the NCAR-specific production system-configuration and programming environment
- Perform tests and benchmarks to validate functionality, performance, reliability, and quality
- Run benchmarks and demonstrate that benchmarks meet performance commitments

Pre-Delivery Test

Subcontractor shall provide NCAR on-site access to the system in order to verify that the system demonstrates the ability to pass acceptance criteria.

The pre-delivery test shall consist of (but is not limited to) the following tests:

Name of Test	Pass Criteria
System power up	All nodes boot successfully
System power down	All nodes shut down
Unix commands	All UNIX/Linux and vendor-specific commands function correctly
Monitoring	Monitoring software shows status for all nodes
Reset	"Reset" functions on all nodes
Power On/Off	Power cycle all components of the entire system from the console and remotely
Fail Over/Resilience	Demonstrate proper operation of all fail-over or resilience mechanisms

Name of Test	Pass Criteria
Full Configuration Test	Full Configuration Test runs successfully on the system
Benchmarks	The system shall demonstrate the ability to achieve the required performance level on all benchmarks
72 Hour test	High availability of the production system for a 72 hour test period under constant throughput load

Post-delivery Integration and Test

The Subcontractor's system(s) shall be delivered, installed, fully integrated, and shall undergo Subcontractor stabilization processes. Post-delivery testing shall include replication of all of the pre-delivery testing steps, along with appropriate tests at scale, on the fully integrated platform.

Site Integration

When the Subcontractor has declared the system to be stable, the Subcontractor shall make the system available to NCAR personnel for site-specific integration and customization. Once the Subcontractor's system has undergone site-specific integration and customization, the acceptance test shall commence.

Acceptance Test

The Acceptance Test Period shall commence when the system has been delivered, physically installed, and undergone stabilization and site-specific integration and customization. A not-to-exceed duration of the Acceptance Test Period shall be defined as part of the subcontract.

All tests shall be performed on the production configuration as defined by NCAR.

The Subcontractor shall not be responsible for failures to meet the performance or availability metrics set forth in this Section, if such failure is the direct result of modifications made by NCAR to Subcontractor source code. Such suspension will be only for those requirements that fail due to the modification(s) and only for the length of time the modification(s) result(s) in the failure.

The Subcontractor shall supply source code used, compile scripts, output, and verification files for all tests. All such provided materials become the property of NCAR, unless the softward is being used exclusively for acceptancet testing and is not part of the delivered system.

All tests shall be performed on a production configuration of the NWSC-2 system, as it will be deployed to the NCAR user community. NCAR may run all or any portion of these tests at any time on the system to ensure the Subcontractor's compliance with the requirements set forth in this document.

The acceptance test shall consist of Functionality Demonstrations, System Tests, System Resiliency Tests, Performance Tests, and an Availability Test, performed in that order.

Functionality Demonstration

Subcontractor and NCAR will perform the Functionality Demonstration on a dedicated system. The Functionality Demonstration shall show that the system is configured and functions in accordance with the statement of work. Demonstrations shall include, but are not limited to, the following:

- Remote monitoring, power control and boot capability
- Network connectivity
- File system functionality
- Archive storage and retrieval
- Batch system
- System management software
- Program building and debugging (e.g. compilers, linkers, libraries, etc.)
- Unix functions

System Test

Subcontractor and NCAR will perform the System Test on a dedicated system. The System Test shall show that the system is configured and functions in accordance with the statement of work. Demonstrations shall include, but are not limited to, the following:

- Two successful system cold boots to production state in accordance with required timings, with no intervention to bring the system up. Production state is defined as 98% of the compute resource and 99% of the PFS is available and the system is running all services required for production use, and being able to compile and run parallel jobs across the system.
- In a cold boot, all elements of the system (compute, login, I/O, network) are completely powered off before the boot sequence is initiated. All components are then powered on.

System Resilience Test

Subcontractor and NCAR will perform the System Resilience Test on a dedicated system. The System Resilience Test shall show that the system is configured and functions in accordance with the statement of work.

All system resilience features of the NWSC-2 system shall be demonstrated via fault-injection tests when running test applications at scale. Fault injection operations should include both graceful and hard shutdowns of components. The metrics for resilience operations include correct operation, any loss of access or data, and time to complete the initial recovery plus any time required to restore (fail-back) a normal operating mode for the failed components.

The following tests shall be performed on the production NWSC-2 system and PFS:

- Single node power-fail/reset test: Failure or reset of a single compute node shall not cause a system-wide failure. A node shall reboot to production state after reset in accordance with required timings.
- Single cabinet power-fail/reset test: Power loss to a single compute cabinet shall not cause a system-wide failure.
- Single high-speed network switch failure test: Failure of a single tier-1 communications switch shall not cause a system-wide failure.
- Redundant power-feed test: All components provided with redundant power shall survive the loss of one power feed.

Redundant power-feed test: All components provided with redundant power shall survive the loss of one power feed.

Performance Test

The Subcontractor shall run the NWSC-2 tests and application benchmarks, Full Configuration Test, External Network Test and file system metadata test, a minimum of three times each as described in the Benchmark Run Rules section. Benchmark answers must be correct, and each benchmark result must meet or exceed performance commitments.

Benchmarks must be run using the supplied resource management and scheduling software. Except as required by the run rules, benchmarks need not be run concurrently. If requested by NCAR, Subcontractor shall reconfigure the resource management software to utilize only a subset of compute nodes, specified by NCAR. Performance must be consistent from run to run.

Availability Test

The Availability Test will commence after successful completion of the Functionality Demonstration, System Test and Performance Test. The Subcontractor shall perform the Availability Test; at this time or before, NCAR will add user accounts to the system. The Availability Test shall be 30 contiguous days in a sliding window within the Acceptance Test Period during which the system must demonstrate the required NWSC-2 system availability.

During the Availability Test, NCAR shall have full access to the system and shall monitor the system. NCAR and users designated by NCAR shall submit jobs through the NWSC-2 resource management system. These jobs shall be a mixture of benchmarks from the Performance Test and other applications.

The Subcontractor shall adhere to the System Availability and Reliability requirements as defined below:

- All hardware and software shall be fully functional at the end of the Availability Test. Any down time required to repair failed hardware or software shall be considered an outage unless it can be repaired without impacting system availability.

- Hardware and software upgrades shall not be permitted during the last 7 days of the Availability Test. The system shall be considered down for the time required to perform any upgrades, including rolling upgrades.
- No significant problems shall be open during the last 7 days.
- During the Availability Testing period, if any system software upgrade or significant hardware repairs are applied, the Subcontractor shall be required to run the Benchmark Tests and demonstrate that the changes incur no loss of performance. At its option, NCAR may also run any tests it deems necessary. Time taken to run the Benchmark and other tests shall not count as downtime, provided that all tests perform to specifications.
- Every test in the Functionality Test, Performance Test and NCAR-defined workload shall obtain a correct result in both dedicated and non-dedicated modes.
- In dedicated mode, each benchmark in the Performance Test shall meet or exceed the performance commitment and variation requirement.
- In non-dedicated mode, the mean performance of each performance test shall meet or exceed the performance commitment. The measured Coefficient of Variation of results from each performance test shall not be greater than 5%.
- Node and system availability will be measured on a node hour basis as described in the Glossary.

Node and system outages are defined in the following section.

Definition of Node and System Failures

- A node shall be defined as down if a hardware problem causes Subcontractor supplied software to crash or the node is unavailable. Failures that are transparent to Subcontractor-supplied software because of redundant hardware shall not be classified as a node being down as long as the failure does not impact node or system performance. Low severity software bugs and suggestions (e.g. wrong error message) associated with Subcontractor supplied software will not be classified as a node being down.
- A node shall be classified as down if a defect in the Subcontractor supplied software causes a node to be unavailable. Communication network failures external to the system, and user application program bugs that do not impact other users shall not constitute a node being down.
- Repeat failures within eight hours of the previous failure shall be counted as one continuous failure.
- The Subcontractor's system shall be classified as down (and all nodes shall be considered down) if any of the following requirements can not be met ("system-wide outage"):

- Complete a POSIX ‘stat' operation on any file within all Subcontractor-provided file systems and access all data blocks associated with these files.
- Complete a successful interactive login to the Subcontractor's system. Failures in the NCAR network do not constitute a system-wide failure.
- Successfully run any part of the performance test. The Performance Test consists of the NWSC-2 Benchmarks, the Full Configuration Test and the External Network Test.
- Full switch bandwidth is available. Failure of a switch adapter in a node does not constitute a system-wide failure. However, failure of a switch may constitute failure, even if alternate switch paths were available, because either full bandwidth would not be available for multiple nodes or more than 1% of the nodes are unreachable.
- User applications can be launched and/or completed via the scheduler.
- Other failures in Subcontractor supplied products and services that disrupt work on a significant portion of the nodes shall constitute a system-wide outage.

If there is a system-wide outage, NCAR shall turn over the system to the Subcontractor for service when the Subcontractor indicates they are ready to begin work on the system. All nodes are considered down during a system-wide outage.

- Downtime for any outage shall begin when NCAR notifies the Subcontractor of a problem (e.g. an official problem report is opened) and, for system outages, when the system is made available to the Subcontractor. Downtime shall end when:
  - For problems that can be addressed by bringing up a spare node or by rebooting the down node, the downtime shall end when a spare node or the down node is available for production use.
  - For problems requiring the Subcontractor to repair a failed hardware component, the downtime shall end when the failed component is returned to NCAR and available for production use.
  - For software downtime, the downtime shall end when the Subcontractor supplies a fix that rectifies the problem or when NCAR reverts to a prior copy of the failing software that does not exhibit the same problem.
  - A failure due to NCAR or to other causes out of the Subcontractor's control shall not be counted against the Subcontractor unless the failure demonstrates a defect in the system. If there are disputes as to whether a failure is the fault of the Subcontractor or NCAR, they shall be resolved prior to the end of the acceptance period.

Appendix B: Project Management Requirements

Project Management

The assembly, pre-shipment testing, installation and acceptance testing of the NWSC-2 system will require close cooperation between the Subcontractor and NCAR, and therefore a clear and comprehensive project management plan is a critical element of a successful system deployment and subsequent transition to production. Additionally, the NCAR-Wyoming Supercomputer Center is located in Cheyenne, Wyoming and there are logistics and transportation issues which must be addressed.

Project Planning Workshop

NCAR and the Subcontractor shall schedule and complete a workshop to mutually understand and agree upon project management goals, techniques, and processes. The workshop shall take place within 45 days after award.

Project Plan

Subcontractor shall provide NCAR with a detailed Project Plan – which includes a detailed Work Breakdown Structure (WBS) and milestone schedule. The Project Plan shall contain all aspects of the proposed Subcontractor's solution and associated engineering (hardware and software) and support activities. The project plan shall be submitted within 60 days of award. The Project Plan shall address or include:

- Project Management approach and process
- WBS
- Facilities Planning (e.g., floor, power & cooling, cabling)
- Transportation and delivery of equipment
- Computer Hardware Planning
- Installation & Test Planning (including pre-delivery factory tests and acceptance tests)
- Shipping, Delivery, Deployment and Integration
- System Stability Planning
- System Scalability Planning
- Software Plan
- Development
- Interdependencies
- Testing
- Risk Assessment & Risk Mitigation.
- Staffing, through the life of the subcontract

- Remote and on-site Support and Services Planning, through the life of the subcontract
- Training & Education of systems, application, and operations staff

Project Management Team

The Subcontractor shall appoint a Project Manager (PM) for the purposes of executing the Project Management Plan on behalf of the Subcontractor.

The Project Manager shall be assigned for the duration of the subcontract and be based at the NWSC-2 site in Cheyenne, Wyoming through the delivery, installation and acceptance of the System. When the PM is unavailable due to vacation, sick leave, or other absence, the Subcontractor shall provide a backup who is knowledgeable of the NWSC-2 project and has the authority to make decisions in the absence of the PM. The PM or backup shall be available for emergency situations via phone or pager on a 24x7 basis through the delivery, installation and acceptance of the delivered system.

Subcontractor Management Contacts

The following positions in the Subcontractor management chain are responsible for performance under this subcontract:

- Technical Contact
- Service Manager
- Contract Manager

Roles and Responsibilities

The **Project Manager** has responsibility for overall customer satisfaction and subcontract performance. It is anticipated that he/she shall be an experienced Subcontractor employee with working knowledge of the products and services proposed. The Subcontractor's PM can and shall:

- Delegate program authority and responsibility to Subcontractor personnel
- Establish internal schedules consistent with the subcontract schedule and respond appropriately to schedule redirection from the designated NCAR authority
- Establish team communication procedures
- Conduct regularly scheduled review meetings
- Approve subcontract deliverables for submittal to NCAR
- Obtain required resources from the extensive capabilities available from within the Subcontractor and from outside sources
- Act as conduit of information and issues between NCAR and the Subcontractor
- Provide for timely resolution of problems

- Apprise NCAR of new hardware and software releases and patches within one week of release to the general market place and provide NCAR with said software within two weeks of request

The PM shall serve as the primary interface for NCAR into the Subcontractor, managing all aspects of the Subcontractor in response to the program requirements.

The **Technical Contact** shall be responsible for:

- Developing System configurations to technical design requirements
- Updating NCAR on the Subcontractor's products and directions
- Working with the PM to review the Subcontractor's adherence to the subcontract

The **Contract Manager** is:

- The Subcontractor's primary interface for subcontract matters
- Is authorized to sign subcontract documents committing the Subcontractor
- Supports the Project Manager by submitting formal proposals and accepting subcontract modifications.

The **Service Manager** has the responsibility for:

- Compliance with the Subcontractor's hardware maintenance and software service requirements
- Determining workload requirements and assigning services personnel to support NCAR
- Managing the Subcontractor's overall service delivery to NCAR
- Meeting with NCAR personnel regularly to review whether the Subcontractor's service is filling NCAR's requirements
- Helping Subcontractor's service personnel understand NCAR's business needs and future directions

Periodic Progress Reviews

The Subcontractor's PM or designate shall communicate daily with NCAR's Project Manager, Technical Representatives or designate and appropriate NCAR staff. These daily communications shall commence shortly after subcontract award and continue until both parties agree they are no longer needed. The topics covered in this meeting include:

- Project status vs schedule
- System problems – status including escalation
- Non-system problems
- Impending deliveries
- Other topics as appropriate
- The Subcontractor's PM (or designate) is the owner of this meeting.

Target duration for this meeting is one-half hour. Both Subcontractor and NCAR may submit agenda items for this meeting.

Weekly Status Meeting

The Subcontractor's PM shall schedule this meeting. Target duration is one hour. Attendees normally include the Subcontractor's PM, Service Manager, NCAR's Project Manager, Procurement Representative, Technical Representative and System Administrator(s) as well as other invitees.

Topics covered in this meeting include:

- Review of the past seven days and the next seven days with a focus on problems, resolutions, and impending milestones
- Review of NCAR's top-10 list of problems and issues.
- System reliability
- System utilization
- System configuration changes
- Open issues (hardware/software) shall be presented by the Subcontractor's PM. Open issues that are not closed at this meeting shall have an action plan defined and agreed upon by both parties by close of this meeting
- Other topics as appropriate

Extended Status Review Meeting

Periodically, but no more than once per month and no less than once per quarter, an Extended Status Review Meeting will be conducted in lieu of the Weekly Status Meeting. The Subcontractor's PM shall schedule this meeting with the agreement of NCAR's Project Manager. Target duration is one to three hours. Attendees normally include: Subcontractor's PM, Technical Contact, Field Service Manager and Line Management, NCAR's PM, Procurement Representative, Technical Representative and Line Management as well as other invitees. Topics covered in this meeting include:

- Review of the past 30 days and the next 30 days with a focus on problems, resolutions and impending milestones (Subcontractor PM to present)
- Implementation schedule status (Subcontractor PM to present)
- High priority issues (issue owners to present)
- Facilities issues (changes in product power, cooling, and space estimates for the to be installed products)
- All topics that are normally covered in the Weekly Status Meeting
- Other topics as appropriate

Quarterly Executive Meeting

Subcontractor's PM shall schedule this meeting. Target duration is six hours. Attendees normally include: Subcontractor's PM, Subcontractor's Senior Management, NCAR's Project Manager, Procurement Representative, Technical Representative, selected Management, selected Technical Staff and other invitees. Topics covered in this meeting include:

- Program status (Subcontractor to present)
- NCAR's satisfaction (NCAR to present)
- Partnership issues and opportunities (joint discussion)
- Future hardware and software product plans and potential impacts for NCAR
- Participation by Subcontractor's suppliers as appropriate
- Other topics as appropriate
- Both Subcontractor and NCAR may submit agenda items for this meeting.

Hardware and Software Support

- Severity Classifications
    - The Subcontractor shall have documented problem severity classifications. These severity classifications shall be provided to NCAR along with descriptions defining each classification.
- Severity Response
    - The Subcontractor shall have a documented response for each severity classification. The guidelines for how the Subcontractor will respond to each severity classification shall be provided to NCAR.

Problem Search Capabilities

The Subcontractor shall provide the capability of searching the problem database via a web page interface. This capability shall be made available to all individual CISL staff members designated by NCAR.

Problem Escalation

The Subcontractor shall utilize a problem escalation system that initiates escalation based either on time or the need for more technical support. Problem escalation procedures are the same for hardware and software problems. A problem is closed when all commitments have been met, the problem is resolved and NCAR is in agreement. NCAR initiates problem notification to onsite Subcontractor personnel, or designated Subcontractor on-call staff.

Risk Management

The Subcontractor shall continuously monitor and assess risks affecting the successful completion of the NWSC-2 project, and provide NCAR with documentation to facilitate project management, and to assist NCAR in its risk management obligations to the NSF.

The Subcontractor shall provide NCAR with a Risk Management Plan (RMP) for the technological, schedule and business risks of the NWSC-2 project. The RMP describes the Subcontractor's approach to managing NWSC-2 project risks by identifying, analyzing, making contingency plans for, mitigating, tracking, and ultimately retiring project risks. The initial plan is due 30 days after award of the subcontract. Once approved by NCAR, NCAR shall review the Subcontractor's RMP annually.

The Subcontractor shall also maintain a formal Risk Register (RR) documenting all individual risk elements that may affect the successful completion of the NWSC-2 project. The RR is a data-base managed using an application and format approved by NCAR. The initial RR is due 30 days after award of the Subcontract. The RR shall be updated at least monthly, and before any Critical Decision (CD) reviews with the NSF. After acceptance, the RR shall be updated quarterly. Along with each required update to the RR, the Subcontractor shall provide a Risk Assessment Report (RAR) summarizing the status of the risks and any material changes. The initial report and subsequent updates will be reviewed and approved by NCAR's Project Manager or his/her designee.

Risk Management Plan

The purpose of this RMP, as detailed below, is to document, assess and manage Subcontract's risks affecting the NWSC-2 project:

- Document procedures and methodology for identifying and analyzing known risks to the NWSC-2 project along with tactics and strategies to mitigate those risks.
- Serve as a basis for identifying alternatives to achieving cost, schedule, and performance goals.
- Assist in making informed decisions by providing risk-related information.

The RMP shall include, but is not limited to: management, hardware, software; risk assessment, mitigation and contingency plan(s).

Risk Register

The RR shall include an assessment of each likely risk element that may impact the NWSC-2 project. For each identified risk, the report shall include:

- Root cause of identified risk
- Probability of occurrence (low, medium, or high)
- Impact to the project if the risk occurs (low, medium, or high)
- Identifies the consequence of a risk event affecting cost, schedule, performance, and/or scope.
- Risk mitigation steps to be taken to reduce likelihood of risk occurrence and/or steps to reduce impact of risk.

Execution of mitigation plans are subject to NCAR approval and may include:

- Technology substitution - subject to the condition that substituted technologies shall not have aggregate performance, capability, or capacity less than originally proposed;
- 3rd party assistance - especially in areas of critical software development;
- Performance compensation - possibility of compensating for performance shortfalls via additional deliveries.
- Dates for risk mitigation decision points.
- Contingency plans to be executed should risk occur; subject to NCAR approval
- Owner of the risk.

Risk Assessment Report

The RAR shall include the following:

- Total number of risks grouped by severity.
- Summary of newly identified risks from last reporting period.
- Summary of any risks retired since the last report.
- Identification and discussion of the status of the Top 10 (watch list) risks.

Definitions and Glossary

Coefficient of Variation: the ratio of the standard deviation to the mean.

Downtime: Any period of time during which a node or component of the system cannot be used operationally. In calculating System Availability, only downtime due to the failure of Offeror-supplied hardware or software applies.

External Network Test: A test or suite of tests, mutually agreed to by the Offeror and UCAR, which demonstrates the system's ability to communicate with other systems on, and transfer files across, the NCAR LAN and WAN.

Flops-to-bytes: A ratio defined by the sustained floating point operations of the system divided by the achieved bandwidth of the parallel file system (PFS). Units are Flops/Byte.

Full Configuration Test: A test or suite of tests, mutually agreed to by the Offeror and UCAR, which demonstrate the full use of the system.

Full scale: All of the compute nodes in the system. This may or may not include all available compute resources on a node depending on the use case.

Job Interrupt: Any system event that causes a job to unexpectedly fail or unintentionally terminate, for example, to not exit or complete as expected due to a system failure, not an application failure.

System Availability:

$$SystemAvailability = \frac{\sum_i^N (S_i - D_i)}{\sum_i^N (S_i)}$$

where:

N is the number of nodes in the system

S_i is the number of scheduled hours for node i

D_i is the number of hours of downtime for node i

Scheduled hours is the wallclock time minus any downtime scheduled by NCAR

System Interrupt: Any system event, or accumulation of system events over time, resulting in more than 1% of the compute resource being unavailable at any given time. Loss of access to any dependent subsystem, e.g. parallel file-system or service partition resource, which affects the system's ability to run the scheduled workload will also incur a system interrupt.

System Mean Time Between Interrupt (SMTBI): Average time between system interrupts over a given time interval.