**SCHEDULE F**

**ACCEPTANCE CRITERIA AND TESTING**

Contents

[1 Acceptance Overview 1](#_Toc415578403)

[2 Pre-Delivery 2](#_Toc415578404)

[2.1 Pre-Delivery Assembly 2](#_Toc415578405)

[2.2 Factory Trial 2](#_Toc415578406)

[3 Post-delivery Integration and Test 3](#_Toc415578407)

[3.1 Site Integration 3](#_Toc415578408)

[3.2 Acceptance Test 3](#_Toc415578409)

[3.2.1 Functionality Demonstration 4](#_Toc415578410)

[3.2.2 System Test 4](#_Toc415578411)

[3.2.3 System Resilience Test 5](#_Toc415578412)

[3.2.4 Performance Test 6](#_Toc415578413)

[3.2.5 Availability Test 6](#_Toc415578414)

[4 Failures and Downtime 7](#_Toc415578415)

# Acceptance Overview

Article 17 of the Subcontract provides terms and conditions of acceptance. Deliverables subject to acceptance testing are identified in the Subcontract’s Schedule E.

Acceptance of a deliverable shall be based upon successful completion of the acceptance test for that deliverable. A deliverable shall be deemed to have been accepted once it completes its acceptance criteria.

1. The acceptance test period shall begin at a mutually agreed time after the Subcontractor has completed the deliverable’s installation and end when a deliverable has met the acceptance criteria, as specified below.
2. The Subcontractor will comply with the agreed test procedures and acceptance criteria of a deliverable identified in Schedules B, D, E and H, in addition to those of any subsequently exercised options and/or Modifications to the Subcontract. UCAR assumes no contractual obligation to perform any test and acceptance for the benefit of the Subcontractor unless specifically set forth elsewhere in the Subcontract.
3. In the event a deliverable does not meet the acceptance criteria during its acceptance test period, the acceptance test period shall, upon UCAR’s sole discretion, either (1) continue on a day-by-day basis until the criteria are met, (2) be halted and restarted, or (3) be abandoned and the deliverable rejected. A deliverable must meet the acceptance test criteria within ninety (90) days after installation. UCAR reserves the right to extend the testing period beyond ninety (90) days up to an additional sixty (60) days or to terminate the Subcontract in accordance with the Subcontract’s “Termination” article without further obligation by UCAR.
4. All maintenance, support, service and parts shall be furnished by the Subcontractor without charge prior to and during an acceptance test period, unless such maintenance, support, service and parts are required as a result of fault or negligence of UCAR.
5. In the case of a rejection after testing or inspection, UCAR shall not be liable for any costs for correction. UCAR reserves the right to charge the Subcontractor any additional cost of testing or inspection when the Subcontractor is not ready at the time such test or inspection is requested, or when retest or re-inspection is necessitated by prior rejection.
6. Inspection and/or acceptance by UCAR do not relieve the Subcontractor from responsibility regarding defects or other failures to meet the Subcontract requirements, which may be discovered prior to acceptance.

# Pre-Delivery

The Subcontractor shall demonstrate that 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. UCAR recognizes that optional 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 systems (cf. Schedule E), or agreed-upon sub-configurations thereof, at the Subcontractor’s location prior to shipment. At its option, UCAR may send representatives to observe and perform 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 Subcontractor-supplied software
* Implementation of the UCAR-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

## Factory Trial

The Subcontractor shall provide UCAR on-site and/or remote access to the NWSC-2 systems (cf. Schedule E) in order to verify that each system demonstrates the ability to pass acceptance criteria.

The Factory Trial 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 |
| System management | The management subsystem performs all requisite system management functions correctly, including independent operation of the HPC and PFS systems and any partitions thereof, if applicable |
| Unix commands | All UNIX/Linux and vendor-specific commands function correctly |
| Monitoring | Monitoring software shows correct 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 |
| RAS | Demonstrate RAS capabilities and system robustness (fault injection, disconnect cables, power down of a subsystem) |
| Fail Over/Resilience | Demonstrate proper operation of all fail-over or resilience mechanisms |
| Benchmarks | The HPC and PFS systems shall demonstrate the ability to achieve the required performance level on all benchmarks, as applicable |
| 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 NWSC-2 systems (cf. Schedule E) shall be delivered, installed, fully integrated, and shall undergo Subcontractor stabilization processes. Post-delivery testing shall include replication of all of the Factory Trial 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 UCAR 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 a system (cf. Schedule E) has been delivered, physically installed, and undergone stabilization and site-specific integration (including HPC and PFS integration with the existing Yellowstone and NCAR GLADE systems) and customization. A not-to-exceed duration of the Acceptance Test Period shall be defined as part of the Subcontract.

The Subcontractor shall not be responsible for failures to meet the performance, performance reproducibility, or availability metrics set forth in this Subcontract, if such failure is the direct result of modifications made by UCAR to benchmark or Subcontractor source code, system configuration or tuning parameters. 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 UCAR. All software used for acceptance testing shall be included as part of the delivered system unless it is acquired separately by UCAR.

All tests shall be performed on a production configuration of the NWSC-2 system as defined by UCAR and as it will be deployed to the UCAR user community. UCAR 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 Test, System Resiliency Tests, Performance Tests, and an Availability Test, performed in that order and as described below.

### Functionality Demonstration

Subcontractor and UCAR 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:

* All items of the Factory Trial Test except “Benchmarks” and “72 Hour test”
* Remote monitoring, power control and boot capability
* Network connectivity
* File system functionality
* Archive storage and retrieval
* Batch system, scheduler and resource management software
* System management software
* Program building and debugging (e.g. compilers, linkers, libraries, etc.)
* Unix functions

### System Test

Subcontractor and UCAR 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 Subcontract. Demonstrations shall include, but are not limited to, the following:

* One successful PFS system shutdown and cold boot[[1]](#footnote-2) to production state in accordance with required timings and with no intervention to bring the system up. Production state is defined as the PFS-hosted file systems are available for mounting on the HPC system and Yellowstone.
* Two successful HPC system shutdowns and cold boots to production state in accordance with required timings, and with no intervention to bring the system up. Production state is defined as 98% of the HPC resource is available, the system is running all services required for production use, and it is possible to compile and run parallel jobs across the system.

### System Resilience Test

Subcontractor and UCAR 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 production 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 data or access to file systems, 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 a HPC system:

* Single HPC node power-fail/reset test: Failure or reset of a single compute node shall not cause a system-wide failure. A compute node shall reboot to production state after reset in accordance with required timings.
* Single HPC 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.

The following tests shall be performed on a PFS system:

* Single PFS node power-fail/reset test: Failure or reset of a single filesystem server node shall not cause loss of accessibility to the hosted file system(s). A PFS node shall reboot to production state after reset in accordance with required timings.
* Storage component removal: removal and/or power-fail of a redundant storage component (disk, drawer, controller, etc.) shall not cause loss of accessibility to data stored on or the ability to write to the host file system.
* 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 HPC benchmarks and PFS benchmark tests, a minimum of three times each as described in the NWSC-2 Benchmark Instructions. Benchmark answers must be numerically correct, and each benchmark performance result must meet or exceed performance commitments. Performance must be consistent from run to run; specifically, performance Coefficient of Variation shall exceed 3% from run-to-run in dedicated mode and 5% in production mode.

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 UCAR, Subcontractor shall reconfigure the resource management software to utilize only a subset of compute nodes, specified by UCAR.

### 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, UCAR 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 the Availability Test, UCAR shall have full access to the system and shall monitor the system. UCAR and users designated by UCAR 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.

Upon the commencement of the Availability Test, Subcontractor personnel will not be permitted to make any changes to the HPC or PFS systems’ hardware or software configuration without prior approval of UCAR, with the exception of those needed to effect hardware repair.

The Subcontractor shall adhere to the testing requirements as defined below:

1. The HPC system shall meet or exceed 98% System Availability.
2. The PFS shall meet or exceed 99% File System Availability.
3. System and File System availability will be measured on a node hour basis.
4. 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.
5. 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.
6. No significant problems shall be open during the last 7 days.
7. 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, UCAR 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.
8. Every test in the Functionality Test, Performance Test and UCAR-defined workload shall obtain correct numerical results in both dedicated and non-dedicated modes.
9. In dedicated mode, each benchmark in the Performance Test shall meet or exceed the performance commitment. The measured Coefficient of Variation of performance results from each performance test shall not be greater than 3%.
10. In non-dedicated mode, the mean performance of each performance test shall meet or exceed the performance commitment. The measured Coefficient of Variation of performance results from each performance test shall not be greater than 5%.

# Failures and Downtime

Node and system outages are defined below. Additional definitions of downtime are contained in Article 1 of the RFP’s NWSC-2 Sample Subcontract Terms and Conditions.

* 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.
* A file system shall be classified as down if data residing on it cannot be read or if data cannot be written to the file system, or if metadata for the filesystem is inaccessible.
* 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 cannot be met (“system-wide outage”). If the HPC and PFS systems are provided by different awardees, only the system which has experienced the outage shall be classified as down. Failures due to the UCAR network or other UCAR-supplied subsystems do not contribute to downtime.
	+ Complete a POSIX ‘stat' operation within 10 seconds 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.
	+ Successfully run any part of the performance 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.
	+ Benchmark runs produce correct numerical answers.
	+ 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, UCAR 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 UCAR 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 substituting 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 UCAR and available for production use.
	+ For software downtime, the downtime shall end when the Subcontractor supplies a fix that rectifies the problem or when UCAR reverts to a prior copy of the failing software that does not exhibit the same problem.
* A failure due to UCAR 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 UCAR, they shall be resolved prior to the end of the acceptance period.
1. In a cold boot, all elements of the system are completely powered off before the boot sequence is initiated. All components are then powered on. [↑](#footnote-ref-2)