# NCAR Globally Accessible Data Environment (GLADE)

Integration Guide

*Updated: 23 March 2020*

## Overview

The Globally Accessible Data Environment (GLADE) provides centralized file storage for HPC computational, data-analysis, visualization and science gateway resources as well as connectivity to the Campaign Storage and Tape Archive resources. The environment provides 300GB/s high-performance bandwidth accessible over EDR IB and 40/100Gb Ethernet networks. The GLADE environment also provides data transfer services to facilitate data movement both within NCAR and to external resources. These data access nodes support common data transfer protocols, Globus transfer service endpoints, Globus data sharing services and high-bandwidth access to the NCAR Campaign Storage and HPSS system.

## File Systems

NCAR HPC file systems are currently all part of the Globally Accessible Data Environment (GLADE) and can be accessed by all current HPC systems simplifying data sharing between platforms. File systems are configured for different purposes with a total of seven spaces available served by four file systems. All file systems are implemented using IBM’s Spectrum Scale TM parallel file system, formerly GPFS.

* Home (/glade/u/home): Permanent, relatively small storage for data like source code, shell scripts, etc. This file system is not tuned for high performance from parallel jobs; therefore, it is not recommended for usage during batch job runs.
* Project (/glade/p): Large, allocated, permanent, high-performance file space. Project directories are intended for sharing data within a group of researchers and are allocated as part of the annual allocations process. It is recommended that this file system be used for computational runs where the data will be accessed post run in the near future.
* Work (/glade/work): Medium, permanent, high-performance file space. Project directories intended for individual work where data needs to remain resident for an extended period of time.
* Scratch (/glade/scratch): Large, high-performance file space. Place large data files in this file system for capacity and capability computing. Data is purged as described below, so you must save important files elsewhere, like Campaign Storage, prior to the purge period.
* Flash (/glade/flash): Small, high-performance file system. Intended for high-IOPs data processing jobs with a short purge policy. Available on request.
* Share (/glade/p/datashare): Medium, purged, high-performance file space. Accessible through Globus Online services and intended to facilitate data transfers to non-NCAR users. Available on request.
* Campaign Storage (/glade/campaign): Large, allocated, high-performance file space. Campaign warm archival storage with a 5 year retention period. Available through standard allocation processes.

### Summary of File Space Policies

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **File Space** | **Peak Perf** | **Quota** | **Backups** | **Purge** |
| /glade/u/home/*user* | 8 GB/s | 25 GB | Yes | No |
| /glade/p/lab/*project* | 300 GB/s | N/A | No | 1 year |
| /glade/work/*user* | 300 GB/s | 1 TB | No | No |
| /glade/scratch/*user* | 300 GB/s | 10 TB | No | 120 days |
| /glade/flash/*user* | 150 GB/s | N/A | No | 2 weeks |
| /glade/collections | 300 GB/s | N/A | No | No |
| /glade/p/datashare | 300 GB/s | 50 TB | No | 45 days |
| /glade/campaign | 76 GB/s | N/A | No | 5 years |

### File Space Intended Use

|  |  |  |
| --- | --- | --- |
| **File System** | **Intended Use** | **File Optimization** |
| /glade/u/home | Hold source code, executables, configuration files, etc. NOT meant to hold the output from your application runs; the scratch or project file systems should be used for computational output. | Optimized for small to medium sized files. |
| /glade/p/project | Sharing data within a team or across computational platforms. Store application output files or common scripts, source code and executables. Intended for actively used data that needs to remain resident for an extended period of time. | Optimized for high-bandwidth, large-block-size access to large files. |
| /glade/work | Store application output files intended for individual use. Intended for actively used data that needs to remain resident for an extended period of time. | Optimized for high-bandwidth, large-block-size access to large files. |
| /glade/scratch | Scratch is intended for temporary uses such as storage of checkpoints or application result output. If files need to be retained longer than the purge period, the files should be copied to project space or to Campaign Store. | Optimized for high-bandwidth, large-block-size access to large files. |
| /glade/flash | The flash file system is intended for temporary uses such pre-processing/post-processing data | Optimized for high-bandwidth, small and large sized files. |
| /glade/collections | The collections space is for curated data collections served through NCAR Science Gateways | Optimized for high-bandwidth, small and large sized files. |
| /glade/p/datashare | Sharing data with non-NCAR users. Intended for transient data being delivered offsite. Only available through the Globus Online service. | Optimized for network-based transfers. |
| /glade/campaign | Warm archival space to preserve data prior to publishing as a collection. | Optimized for med-bandwidth, small and large sized files. |

### Summary of File Space Capacities

|  |  |  |
| --- | --- | --- |
| **File Space** | **Capacity** | **Block Size** |
| /glade/u/home | 50 TB | 512 KB |
| /glade/p/project | 9.277 PB | 8 MB |
| /glade/work | 1.953 PB | 8 MB |
| /glade/scratch | 14.65 PB | 8 MB |
| /glade/flash | 446 TB | 8 MB |
| /glade/collections | 10.74 PB | 8 MB |
| /glade/p/datashare | 200 TB | 8 MB |
| /glade/campaign | 25 PB | 8 MB |

## Systems Served

|  |  |  |  |
| --- | --- | --- | --- |
| **System** | **Description** | **Purpose** | **Connectivity** |
| cheyenne | SGI ICE-XA Cluster | main computational cluster | EDR IB |
| casper | Linux Cluster | GPGPU computational cluster for data analysis & visualization | 100GbE |
| data-access | Linux Cluster | Globus data transfer services, data sharing services | 100GbE |
| RDA science gateway | Linux Cluster web services | Research Data Archive Service | 10GbE |
| ESG science gateway | Linux Cluster web services | Earth Systems Grid Service | 10GbE |
| CDG science gateway | Linux Cluster web services | Climate Data Gateway Service | 10 GbE |

## Data Access Nodes

|  |  |
| --- | --- |
| **Service** | **Bandwidth** |
| Globus / GridFTP | 400 Gb/s (50 GB/s) |
| Globus+ Data Share | 400 Gb/s (50 GB/s) |
| HPSS (hsi, htar, Globus) | 80 Gb/s (10 GB/s) |

## GLADE Architecture





## Connectivity Options

The GLADE I/O Network supports both TCP/IP based and InfiniBand (IB) based connectivity and is designed to easily migrate to newer technologies supporting faster data rates. The current IBM Spectrum ScaleTM (GPFS) NSD servers provide the ability to connect to two networks, 40/100GbE and EDR IB. Bridging or routing technologies may allow better integration in the future particularly when looking at 100GbE as the primary I/O Network.

## Cluster Integration with GLADE

The current GLADE resources utilize IBM’s Spectrum ScaleTM (GPFS) parallel file system. This file system is supported natively for most Linux systems and installing the client software consists of loading an RPM and compiling a small kernel module. NCAR’s HPC Data Infrastructure Group (HDIG) can provide a quick start guide for this process and can aid if necessary.

A GLADE client is defined as a system that utilizes the shared file system for its own use. A GLADE server is defined as a system that makes data on the shared file system available to a serving application or another system. Currently NCAR can provide client licenses freely, however, server licenses will incur an additional cost based upon the capacity of the additional storage resource. NCAR runs Spectrum ScaleTM in a multi-cluster mode allowing for separate management of individual clusters. The main GLADE cluster contains file system servers and storage, while computational clusters are considered diskless.

Each computational cluster is configured as a Spectrum ScaleTM cluster. This allows clusters to be shut down without impacting the primary GLADE cluster. Each cluster will need a minimum of 3 Spectrum ScaleTM management nodes for cluster operation. I/O gateway nodes may also be needed to expose the file system into a system if the primary network isn’t connected directly to either the IB or IP central networks.



## Client Software Requirements

Both the Linux distribution and the Spectrum ScaleTM version must stay within a supported range. There may some flexibility in the Linux distribution as long as the kernel is similar enough in release level to the supported level. Testing may be required to determine if the Spectrum ScaleTM software is compatible or not. However, any software outside of the supported range will not be supported by IBM.

The following chart defines the current supported releases for the GLADE environment. The latest information from IBM is available in the [GPFS FAQ](http://www-01.ibm.com/support/knowledgecenter/SSFKCN/com.ibm.cluster.gpfs.doc/gpfs_faqs/gpfsclustersfaq.html). Note that Spectrum ScaleTM is only supported on 64-bit kernels. The Kernel Level column is the latest tested kernel version for x86\_64 systems. Spectrum ScaleTM is also supported on Power based systems.

### Spectrum ScaleTM (SS) Version: 5.0

|  |  |  |
| --- | --- | --- |
| **OS Distribution** | **Latest Kernel Level Tested** | **Min SS Level** |
| RHEL 7.6 | 3.10.0-957.el7 | 5.0.2.2 |
| RHEL 7.5 | 3.10.0-862.el7 | 5.0.1.1 |
| SLES 15 | 4.12.14-23-default | 5.0.3.0 |
| SLES 12 SP4 | 4.12.14-95.3-default | 5.0.2.3 |
| SLES 12 SP3 | 4.4.103-6.38.1 | 5.0.0.1 |

### Minimum Hardware Requirements for Server Nodes

|  |  |
| --- | --- |
| **Processor** | **Min Memory** |
| Intel EM64T | 2 GB |
| AMD Opteron | 2 GB |