

The NCAR Globally Accessible Data Environment (GLADE) Overview

Updated: 11 Mar 15

Overview

The Globally Accessible Data Environment (GLADE) provides centralized file storage for HPC computational, data-analysis, visualization and science gateway resources. The environment provides >90GB/s high-performance bandwidth currently accessible over FDR IB or 10Gb Ethernet networks. The GLADE environment also provides data transfers 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 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 four spaces available served by three file systems.

- 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 system. 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/p/work): Medium, permanent, high-performance file system. Project directories intended for individual work where data need to remain resident for an extended period of time.
- Scratch (/glade/scratch): Large, high-performance file system. 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 HPSS).
- Share (/datashare): Small, med-performance file system. Only accessible through Globus Online services and intended to facilitate data transfers to non-NCAR users. Available on request.

Summary of File Space Policies

File Space	Type	Peak Perf	Quota	Backups	Purge
/glade/u/home/user	GPFS Fileset	5GB/s	10 GB	Yes	No
/glade/p/project	GPFS Fileset	90GB/s	none	No	No
/glade/p/work/user	GPFS Fileset	90GB/s	10 TB	No	No
/glade/scratch/user	GPFS	90GB/s	10 TB	No	> 90 days
/datashare	GPFS	5GB/s	none	No	No

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	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/p/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	The scratch file system 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 HPSS.	Optimized for high-bandwidth, large-block-size access to large files.
/data/share	Sharing data with non-NCAR users. Intended for transient data being delivered offsite. Only available through the Globus Online service.	Optimized for small to medium sized files and modest network bandwidths.

Summary of File Space Capacities

File Space	Capacity	Block Size	Sub-block Size
/glade/u/home	10 TB	512 KB	16 KB
/glade/p	11 PB	4 MB	128 KB
/glade/p/work	500 TB	4 MB	128 KB
/glade/scratch	5 PB	4 MB	128 KB
/datashare	1.5 PB	512 KB	16 KB

Systems Served

System	Description	Purpose	Connectivity
yellowstone	IBM iDataPlex Cluster	main computational cluster	FDR IB
geyser	IBM xSeries Cluster, nVidia GPU	data analysis and visualization	FDR IB
caldera	IBM iDataPlex Cluster, nVidia GPU	GPGPU computational cluster	FDR IB
pronghorn	IBM iDataPlex Cluster, phi	coprocessor computational cluster	FDR IB
data-access	IBM xSeries Cluster	Globus data transfer services, data sharing services	FDR IB, 10GbE
RDA science gateway	Linux Cluster, web services	Research Data Archive Service	10GbE
ESG science gateway	Linux Cluster, web services	Earth Systems Grid Service	10GbE
CDP science gateway	Linux Cluster, web services	Community Data Portal Service	10GbE

Data Access Nodes

Service	Bandwidth
Globus / GridFTP	10 GB/s
Globus+ Data Share	5 GB/s
HPSS	10 GB/s

Current GLADE Architecture

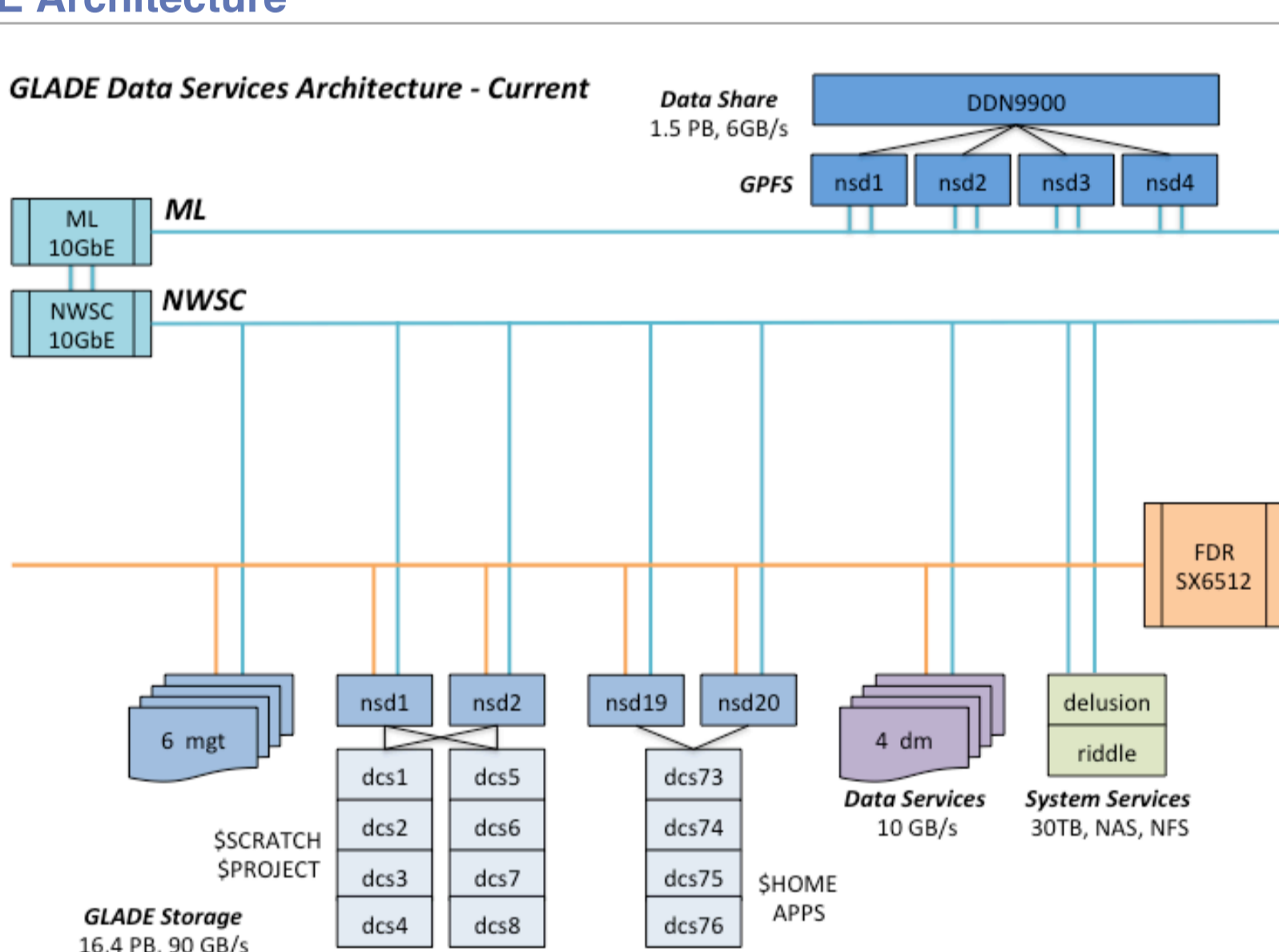


Figure 1. Current GLADE Architecture

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 GPFS NSD servers provide the ability to connect to multiple types of networks, but bridging technologies may allow better integration in the future.

The IB network provides a top level switch hosting the GLADE servers and supporting FDR IB. This switch is then connected with uplinks to yellowstone core switches and the geyser/caldera switch. Additional uplinks can be provided to additional IB switches by adding more blades within the GLADE switch.

The GLADE servers are also connected to a 10Gb Ethernet network. Currently this network is hosted by the NWSC core network, but an additional switch at 10Gb, 40Gb or 100Gb could be utilized to host high-bandwidth IP based connections if required.

Cluster Integration with GLADE

The current GLADE resources utilize IBM's GPFS parallel file system, recently renamed *IBM Spectrum Scale*. 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. DASG can provide a quick start guide for this process and can provide assistance if necessary.

A GLADE client is defined as a utilizes the shared file system for its use. A GLADE server is defined as a node that makes data on the shared file system available to a serving application or another system. NCAR can provide client licenses freely, however, server licenses will incur an additional cost. NCAR runs GPFS in a multi-cluster mode. The main GLADE cluster contains file system servers and storage, while computational cluster as considered diskless.

Each computational cluster is configured as a GPFS cluster. This allows clusters to be shutdown without impacting the primary GLADE cluster. Each cluster will need a minimum of 3 GPFS management nodes for cluster operation. These nodes are considered directly serving nodes. 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. These will also need to be licensed as servers. Compute nodes will be considered clients in all cases. NCAR's Data Analysis Services Group (DASG) can work with you to determine if this is the case and what the cost may be.

Client Software Requirements

Both the Linux distribution and the GPFS version must stay within a supported range. There may be 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 GPFS 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](#). Note that GPFS is only supported on 64-bit kernels. The Kernel Level column is the latest tested kernel version.

GPFS Version: 4.1

OS	Kernel Level	Min GPFS Level
RHEL 7.0	3.10.0-123	4.1.0.4
RHEL 6.6	2.6.32-504	4.1.0.5
RHEL 6.5	2.6.32-431	4.1
RHEL 6.4	2.5.32.358.2.1	4.1
SLES 11 SP3	3.0.76-0.9	4.1
SLES 11 SP2	3.0.101.0.21.1	4.1
Debian 7.1	3.2.0-4-amd64	4.1
Debian 6.0.8	3.2.5.1-1	4.1

Hardware Requirements for Server Nodes

Processor	Memory
Intel EM64T	2 GB
AMD Opteron	2 GB