

This document is an overview of the HPC and BD survey.

Survey content:

The survey covers six major topics: High Performance Computing (HPC) hardware resources, HPC storage, Big data (BD), Software environment, Internet bandwidth and Science DMZ, Consultation with Academic and Research Computing Systems (ARCS). Some of these are divided further into subtopics: HPC hardware resources has separate sections for parallel and serial computations; BD has separate sections for Layer 1 and Layer 2.

How this document differs from the survey:

Complete survey content for all these topics appears in this document. However, when taking the actual survey, each participant will, by his/her choice, address either all these topics and subtopics, or some customized subset of them. Each participant's experience also varies with particular responses to individual questions. For example, a participant's negative rating of a particular HPC resource will lead to further questions eliciting suggestions for improvement. This contingency structure is described throughout this document.

KEY to this document

Text in this font and color indicates a text display that appears in the survey and does not elicit an answer response. It usually introduces a topic or provides explanatory information for subsequent questions.

Text in this color and font indicates the wording of a question within the survey. Answer options are included, but the means of answering—e.g., checkboxes, dropdown lists, etc., are not explicitly indicated in this document.

SECTION TITLES in **BOLD** are for organizational clarity within this document, and do not necessarily appear in the survey.

Text marked by a blue vertical bar appear in this document to explain survey structure, e.g., conditionals and branching.

Red font calls attention to questions whose appearance is contingent on previous answer choices.

————— A thin blue line indicates separation between questions.

————— A thick blue line indicates a separation between sections.

Survey content begins on the following page.

WELCOME MESSAGE:

The purpose of this survey is to gather input from researchers on their usage of NJIT's high performance computing (HPC) and big data (BD) infrastructure. The information gathered from this survey will be used for tuning the HPC and BD environment, evaluating hardware needs, and planning the direction for HPC and BD support.

INTRODUCTION—for all participants

This survey covers the topics described below.

You may address any combination of topics.

Q: High Performance Computing (HPC) hardware resources

Clusters: Kong (general-access (GA)) and Stheno (Dept Mathematical Sciences (DMS)). Both parallel and serial computations can be done on these clusters.

GPU (graphical processing unit) nodes: These nodes contain both CPUs and GPUs. Both parallel and serial computations can be done on the CPUs. The GPUs are suitable only for parallel computations. Both Kong and Stheno contain GPU nodes.

Shared memory machines: Kong "smp" queue (GA), Cnrdp (Center for Natural Resources and Protection), Gorgon (DMS), Phi (GA). Both parallel and serial computations can be done on shared memory machines.

HTCondor (GA) Uses idle cycles in the GITC 2315C and GITC 2400 Linux classrooms. Currently only serial computations can be done on HTCondor.

Do you wish to address this topic?

Answer choices (single):

Yes; No

Q: HPC storage

AFS distributed filesystem: General computational use; accessible from any AFS client computer, including all cluster compute node

NFS distributed filesystem: General computational use; accessible only from the cluster on which it is mounted

Scratch space: Temporary files used during computations; local to a compute node

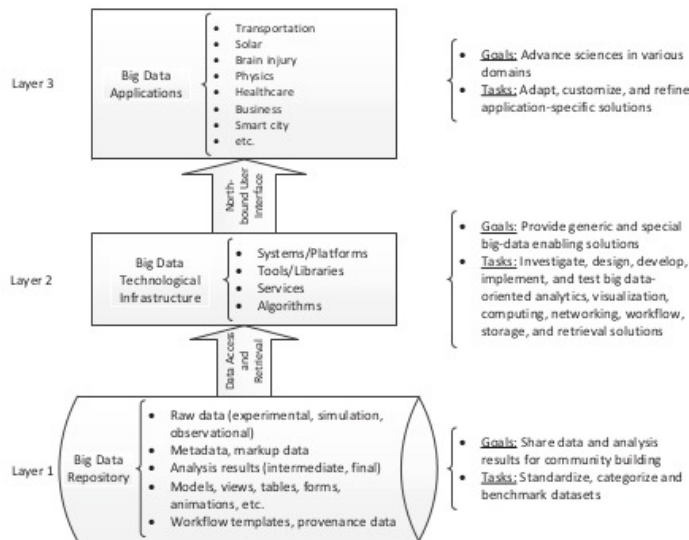
Do you wish to address this topic?

Answer choices (single):

Yes; No

Q: Big data (BD) Layer 1: Repository and/or BD Layer 2: Technological infrastructure

BDC Three-Layer Structure figure



Do you wish to address either of these topics?

Answer choices (first and/or second, or *Neither*):

Layer 1; Layer 2; Neither

Q: Software environment

Open source and commercial applications; libraries; utilities; modules (used for setting the user's environment for a specific application)

Do you wish to address this topic?

Answer choices (single):

Yes; No

Q: Internet bandwidth and/or Science DMZ

“Internet bandwidth”, including Internet 2, is the capacity of NJIT's connection to and from the Internet.

"Science DMZ" refers to a computer subnetwork that is structured to be secure, but without the performance limits that would otherwise result from passing data through a stateful firewall. The Science DMZ is designed to handle high volume data transfers, typical with scientific and high-performance computing, by creating a special DMZ to accommodate those transfers. Science DMZ is typically deployed at or near the local network perimeter, and is optimized for a moderate number of high-speed flows, rather than for general-purpose business systems or enterprise computing.

Do you wish to address either of these topics?

Answer choices (first and/or second, or *Neither*):

Internet bandwidth; Science DMZ; Neither

Q: Consultation with Academic and Research Computing Systems (ARCS)

Consultation with Academic and Research Computing Systems (ARCS) staff: installation of compilers, applications, libraries, and utilities; customized scripts to aid users in their use of HPC resources; assistance in debugging and optimizing code; assistance in getting applications to run; Assistance in running parallel code; assistance in working with and managing big data

Do you wish to address this topic?

Answer choices (single):

Yes; No

Participants who answered *No* to all topic choices are instructed to either exit the survey or continue the survey by changing any previous topic choice to *Yes*.

PARTICIPANT INFORMATION—for all participants

Q: What is your status?

Answer choices (single):

faculty/staff; student

Q: For approximately how long have you been using IST-managed HPC and/or BD resources?

Answer choices (single):

Never used; less than 6 months; 6 to 12 months; 1+ to 2 years; 2+ to 5 years; 5+ years; Don't know

Q: What is the general classification of computations for which you use IST-managed HPC (e.g., computational fluid dynamics statistical analysis, bioinformatics)?

write-in

Q: What is the specific description(s) of the computations for which you use IST-managed HPC?

write-in

HIGH PERFORMANCE COMPUTING (HPC) HARDWARE RESOURCES—only for participants who chose this topic

Terminology: parallel and serial computations

The definitions of *parallel* and *serial* computing are technical, and the distinctions between the two are often unclear. For the purposes of this survey, one of whose goals is to determine how the existing infrastructure is used, refer to the following guidelines.

Definitions of parallel and serial computing

- Parallel computations
 - Working definition: The application uses a set of **independent** cores that **can** work cooperatively on tasks at the **same time** in solving a problem
 - Common implementation platforms
 - Distributed memory clusters - e.g., Kong, Stheno
 - Shared memory machines - e.g., Kong "smp" queue, Gorgon, Cnrdp, Phi
 - Graphical processing units (GPU) - present on Kong and Stheno only
 - The CPUs that are part of the GPU nodes
- Serial computations
 - Working definition: The application can use **only one core** at a time, and processes tasks in **sequentially**
 - Common implementation platforms
 - Distributed memory clusters - e.g., Kong, Stheno
 - Shared memory machines - e.g., Kong "smp" queue, Gorgon, Cnrdp, Phi
 - The CPUs that are part of the GPU nodes
 - HTCondor (Uses idle cycles in the GITC 2315C and GITC 2400 Linux classrooms)

Note that **both** parallel and serial computations can be done on **both** clusters and shared memory machines, including the CPUs that are part of the GPU nodes. The GPUs themselves are not suitable for serial processing.

Determining whether you are doing parallel or serial computation

- You are doing **parallel** computation if your applications are using any of the following:
 - Message passing interface (MPI)
 - MP, an application programming interface (API) for multithreaded, shared memory parallelism
 - GPUs
- You are doing **serial** computation if you are not doing parallel computation. If in doubt, you are probably doing serial computation.

Q: For which type(s) of computations do you use IST-managed HPC resources?

Answer choices (multiple):

Parallel; Serial

PARALLEL—only for participants who chose Parallel in the previous HIGH PERFORMANCE COMPUTING (HPC) HARDWARE RESOURCES section

The following table provides information on HPC hardware resources:

HPC Specifications Summary

Name	Type	Access	Num nodes	Num CPU cores	CPU RAM total, GB	Num GPU cores	GPU RAM total, GB	Interconnect	OS	Cores:queue	AFS client
Kong	Cluster	General	339	3008	25472	46592	184	GigE/10GigE [1]	Scientific Linux 6.8	1400:short 800:medium 300:long	Yes
Stheno	Cluster	Restricted (DMS [2])	32	392	3840	15320	32	InfiniBand	Scientific Linux 5.5	40:ib-short 20:ib-medium 10:ib-long	Yes
Gorgon	SMP [3]	Restricted (DMS)	1	32	64	0	0	N/A	Scientific Linux 6.8	N/A	Yes
Phi	SMP	General	1	16	64	0	0	N/A	RHEL 7.4	N/A	Yes

Notes

- GigE/10GigE = Gigabit Ethernet/10Gigabit Ethernet
- DMS = Dept. Mathematical Sciences
- SMP = Symmetric Multiprocessor

Q: Please select the resources on which you run parallel computations.

Answer choices (multiple):

Kong, including “smp” queue; Stheno; Gorgon; Cnrdp; Phi; GPU; Not doing parallel computation

Participants see a separate table (below) for each resource (Kong, Stheno, etc.) selected in the previous question. Tables are identical except for that the **Node interconnect speed** row does not appear in tables for Gorgon, Cnrdp, Phi, and GPU.

Q: Please indicate the adequacy of parallel resources for [Kong/Stheno/Gorgon/Cnrdp/Phi/GPU].

Answer choices (one choice per row):

	Adequate	Moderate increase needed	Large increase needed	Don't know
Number of cores	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Number of nodes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
CPU speed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Max RAM per node	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Node interconnect speed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Participants see a separate ranking list (below), preceded by a definition box (below), for each previously selected resource (Kong, Stheno, etc.) for which they selected either **Moderate increase needed** or **Large increase needed** in the previous table. (Ranking lists and definitions for any particular resource are not presented when **Adequate** or **Don't know** are selected all table rows.)

Please refer to these definitions when answering the question below:

Terminology:

- *On-premise shared* The resource is **located** at NJIT, is **provided** by NJIT, and is **shared** amongst its users
- *On-premise dedicated* The resource is **located** at NJIT, and is **dedicated** to the purchaser

- *Off-premise shared* The resource is **not located** at NJIT, is **provided** by NJIT, and is **shared** amongst its users
- *Off-premise dedicated* The resource is **not located** at NJIT, and is **dedicated** to the researcher that purchased the resource
- *External* The resource is publically-available (e.g., at a national supercomputing center); successful proposal by researcher is required

Note: Off-premise providers include Amazon Web Services, Azure, Google Cloud Platform, IBM Bluemix, Oracle Cloud, Penguin Computing

Q: You have indicated that at least one HPC parallel computational resource on [Kong/Stheno/etc.] is inadequate. Please rank the possible ways listed below of addressing any inadequacy(ies).

Answer choices (ranking):

Increase on-premise shared - *resource use restrictions apply*

Increase on-premise dedicated - *researcher purchases resources*

Increase off-premise shared - *resource use restrictions apply*

Increase off-premise dedicated - *researcher purchases resources*

Use external grid computing, e.g., *Xsede*, which provides HPC and associated resources; requires proposal

Q: Please provide any comments on parallel computation resources.

write-in

SERIAL—only for participants who chose Serial in the previous HIGH PERFORMANCE COMPUTING (HPC) HARDWARE RESOURCES section

Q: Please select the resources on which you run serial computations.

Answer choices (multiple):

Kong, including “smp” queue; Stheno; Gorgon; Cnrdp; Phi; HTCondor; Not doing parallel computation

The Serial section has the same structure as the Parallel section: The HPC Specifications Summary Table is displayed, and participants are queried only about the resources that they select, via adequacy tables and solution rankings, preceded by text-box definitions. Adequacy tables for all six resources have the same five rows, i.e., Number of cores, Number of nodes, CPU speed, Max RAM per node, and Node interconnect speed. Participants may write in comments at the end of the section.

Note that the serial resource choices are the same as the parallel resource choices except for the inclusion of HTCondor and exclusion of GPU.

Both SERIAL and PARALLEL—only for participants who chose **High Performance Computing (HPC) hardware resources** in the **Introduction** section

Q: Please indicate the adequacy of HPC computational resources documentation at the HPC and BD Wiki.

Answer choices (single):

Adequate; Somewhat more needed; Much more needed; Don't know

Q: Please provide any comments on HPC computational resources.

write-in

HPC storage—only for participants who chose **this topic**

The base storage allocations for researchers are shown in the table below.

Base resources are available to all NJIT researchers. If resources beyond the base allocation are needed, the researchers must arrange for the purchase of such resources. Resources that are purchased are dedicated to the purchaser.

Resource type	Resource name	Default allocation	Cost of dedicated resource	Notes
Disk	AFS	500GB max [1]	DiskAndBackupCost	Accessible from all AFS clients - Linux, MacOSX, Windows
Disk	NFS	500GB max [2]	DiskAndBackupCost	Accessible only from the HPC cluster to which it is attached

[1]. Independent of NFS allocation.

[2]. Independent of AFS allocation.

Q: Please indicate the adequacy of base allocations.

AFS distributed filesystem: General computational use; accessible from any AFS client computer, including all cluster compute nodes

NFS distributed filesystem: General computational use; accessible only from the cluster on which it is mounted

Answer choices (single choice for each of two topics):

AFS storage: Adequate; Somewhat more needed; Much more needed; Don't know

NFS storage: Adequate; Somewhat more needed; Much more needed; Don't know

Text box (below) and subsequent ranking question **seen only by participants** who chose **Somewhat more needed** or **Much more needed** for either AFS or NSF storage in previous question

Please refer to these definitions when answering the question below:

Terminology:

- *On-premise shared* The resource is **located** at NJIT, is **provided** by NJIT, and is **shared** amongst its users
- *On-premise dedicated* The resource is **located** at NJIT, and is **dedicated** to the purchaser
- *Off-premise shared* The resource is **not located** at NJIT, is **provided** by NJIT, and is **shared** amongst its users
- *Off-premise dedicated* The resource is **not located** at NJIT, and is **dedicated** to the researcher that purchased the resource

Q: You have indicated that at least one HPC storage resource is inadequate. Please rank the possible ways listed below of addressing any inadequacy(ies).

Answer choices (ranking):

Increase on-premise shared - *resource use restrictions apply*

Increase on-premise dedicated - *researcher purchases resources*

Increase off-premise shared - *resource use restrictions apply*

Increase off-premise dedicated - *researcher purchases resources*

Q: Please indicate the importance of platform-independent access to HPC storage

"Platform-independent access" means that file paths and authorization are independent of which platform - Linux, MacOSX, Windows - is being used to access files.

Answer choices (single):

Very important; Moderately important; Not important; Don't know

Q: Please indicate the adequacy of HPC storage documentation at the HPC and BD wiki

Answer choices (single):

Adequate; Somewhat more needed; Much more needed; Don't know

Q: Please provide any comments on HPC storage.

write-in

BIG DATA LAYER 1: Repository— only for participants who chose this topic

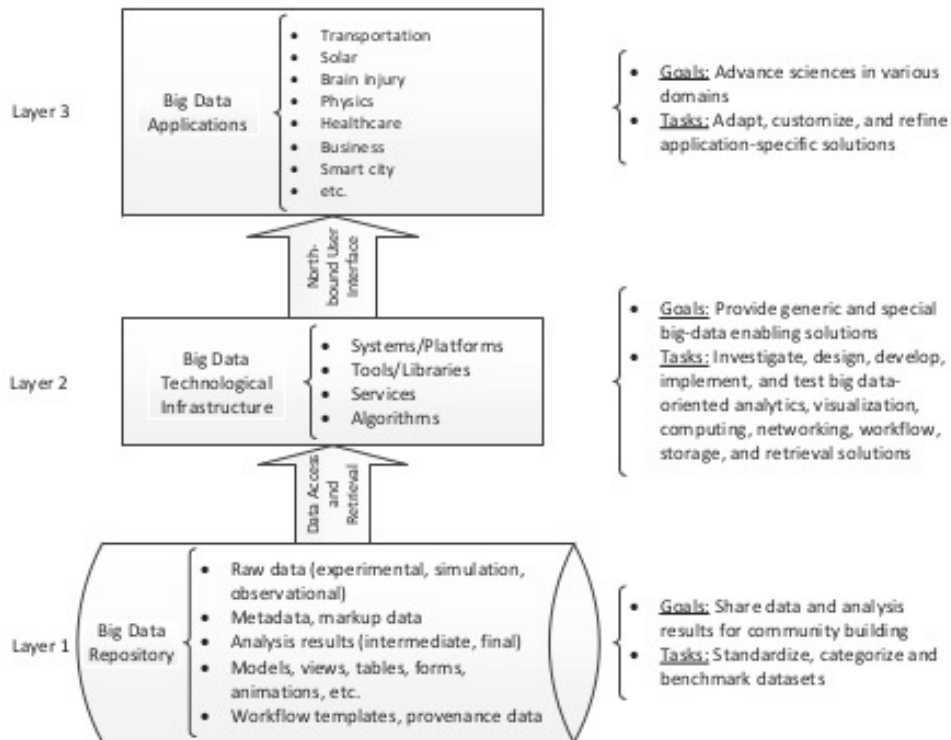
The base storage allocations for researchers are shown in the table below.

Resource type	Resource name	Default allocation	Cost of dedicated resource	Notes
Disk	AFS	500GB max [1]	DiskAndBackupCost	Accessible from all AFS clients – Linux, MacOSX, Windows
Disk	NFS	500GB max [2]	DiskAndBackupCost	Accessible only from the HPC cluster to which it is attached

[1]. Independent of NFS allocation.

[2]. Independent of AFS allocation.

Use this figure when answering the next question.



Q: Referring to the items listed in Layer 1 of the figure, please indicate the adequacy of the base allocations

Answer choices (one choice per row):

	Adequate	Some more needed	Much more needed	Don't know
Raw data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Metadata	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Markup data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Analysis results	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Models	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Views	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tables	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Forms	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Animations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Workflow templates	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Provenance data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Q: You have indicated that at least one BD Layer 1 storage resource is inadequate. Please provide any relevant information regarding the inadequacy(cies).

write-in

Q: Please indicate the importance of platform-independent access to HPC storage.

"Platform-independent access" means that file paths and authorization are independent of which platform - Linux, MacOSX, Windows - is being used to access files.

Answer choices (single):

Very important; Moderately important; Not important; Don't know

Q: Please provide any comments on big data layer 1.

write-in

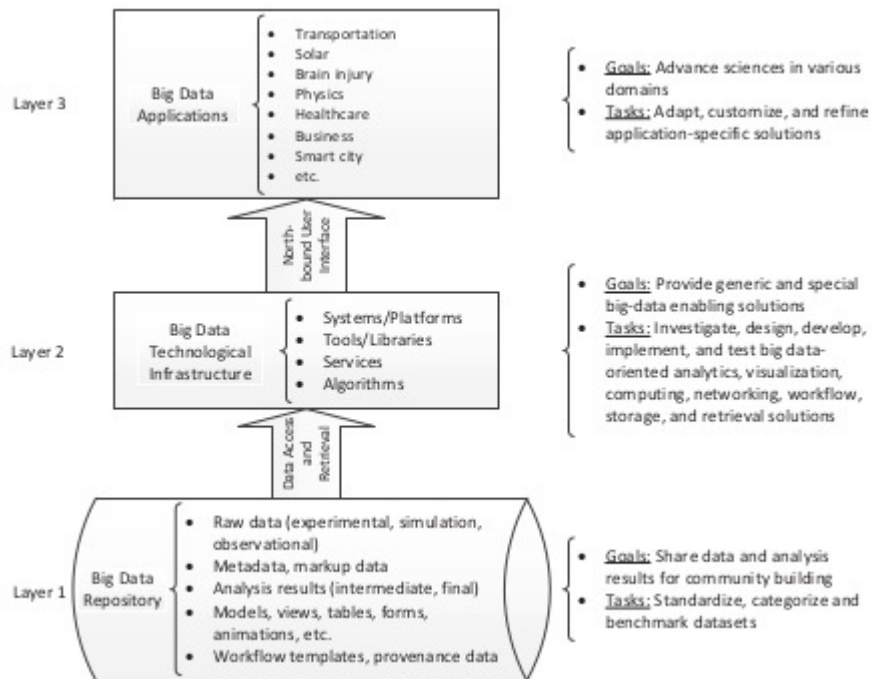
BIG DATA LAYER 2: Repository—only for participants who chose this topic

The Hadoop infrastructure is a virtual environment based on VMware Bid Data Extensions (BDE). VMware introduced Big Data Extensions, or BDE, as a commercially supported version of Project Serengeti designed for enterprises seeking VMware support. BDE enables customers to run clustered, scale-out Hadoop applications on the vSphere platform, delivering all the benefits of virtualization to Hadoop users. BDE delivers operational simplicity with an easy-to-use interface, improved utilization through compute elasticity, and a scalable and flexible Big Data platform to satisfy changing business requirements. VMware has built BDE to support all major Hadoop distributions and associated Apache Hadoop projects such as Pig, Hive, and HBase.

The hardware associated with BDE is as follows:

- 2 x IBM iDataPlex dx360 M4 nodes, each with:
 - 2 x Intel Xeon CPU E5-2680 (8 Core)
 - 16 CPU CORES @ 2.70GHz
 - 32 Logical Processors with Hyperthreading
 - 128G RAM
- 4TB HDFS (Hadoop Distributed File System) disk

Use this figure when answering the next question.



Q: Referring to the items listed in Layer 2 of the figure, please indicate the adequacy of the shared infrastructure.

Answer choices (one choice per row):

	Adequate	Inadequate	Don't know
Systems/Platforms	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tools/Libraries	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Services	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Algorithms	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Text entry question(s) (below) **seen only by participants** who selected **Inadequate** in a row in the preceding table. A separate text entry question is presented for each such row. (No text entry question appears for any individual shared infrastructure topic for which **Adequate** or **Don't know** were selected.) Thus, participants see either no follow-up text entry questions, or as many as four.

Q: You have indicated that [Systems/Platforms / Tools/Libraries / Services / Algorithms] are inadequate. Please suggest means to address this inadequacy

write-in

Q: Please provide any comments on big data layer 2.

write-in

Software environment—**only for participants** who chose this topic

The software environment is the combination of applications - open source and commercial, libraries, utilities, and modules. Modules are used for setting the user's environment for specific software. Almost all HPC and BD software has an associated module.

The following table lists the modules currently available.

AME	LAST MODIFIED	NAME	LAST MODIFIED	NAME	LAST MODIFIED
/etc/modulefiles:	Directory	fluent/16.2.0	2016/08/23	neuron/7.4	2015/11/05
compat-openmpi-psm-x86_64	2016/06/02	fluent/17.1.0	2016/08/23	node.js/6.0.0	2016/12/06
compat-openmpi-x86_64	2016/06/02	fluent/18.1.0	2017/06/22	node.js/6.10.3	2017/05/20

openmpi-1.4-psm-x86_64	2016/06/02	freefem++-sl6/3.34-2	2015/03/04	nwchem/6.6	2016/02/04
openmpi-1.4-x86_64	2016/06/02	freefem++-sl6/3.37-1	2015/05/28	ompi/2.0.0	2016/08/02
openmpi-1.5.3-psm-x86_64	2016/06/02	freesurfer/5.3.0	2016/09/28	openblas/0.2.14/gnu4/Nehalem	2015/05/28
openmpi-1.5.3-x86_64	2016/06/02	fsl/5.0.9	2016/08/09	opencv/2.4.9	2016/04/02
/afs/cad.njit.edu/ucs/modulefiles:	Directory	gcc/4.8.1	2017/02/22	opencv/3.1.0	2016/04/06
+hpc	2012/10/05	gcc/4.8.2	2017/02/22	oraclient	2014/04/11
+sysadm	2012/10/05	gcc/4.9.2	2017/02/22	oraclient-12c	2014/04/16
+utils	2015/06/01	gcc/5.3.0	2017/02/22	ovito/2.8.2	2017/02/25
+versions	2013/11/18	gcc/5.4.0	2017/02/22	paraview/5.4.1	2017/10/23
Atlas	2015/05/05	gcc/6.1.0	2017/02/22	perl/5.18.4.1805	2016/11/17
C.eclipse/4.7	2017/09/11	gflags/2.2.0	2016/04/18	perl/5.22.2.2202	2016/11/17
MPI/mpich2-intel	2014/04/30	glog/0.3.4	2016/04/18	perl/5.24.0.2400	2016/11/17
NeMo/0.7.2	2016/02/29	go/1.7.1	2016/10/14	petsc/3.5.4	2015/05/28
OpenFOAM	2014/09/19	gromacs/4.5.5	2013/11/17	pharlap/4.1.3	2017/05/31
R-Project/2.15.1	2015/04/17	gromacs/4.6.7	2016/10/13	php/current	2016/04/28
R-Project/3.1.2	2015/04/17	gromacs/5.1.1	2016/02/03	protobuf/2.5.0	2017/04/13
R-Project/3.2.4	2016/09/16	gromacs/5.1.2	2017/08/14	protobuf/2.6.1	2016/04/19
RStudio/0.98	2017/10/17	gromacs_gpu/2016.3	2017/05/03	pyadolc/20170926	2017/09/26
abaqus/2013_6.13	2016/12/09	gromacs_smp/5.1.1	2017/03/28	python2	2017/02/17
abaqus/2016	2017/10/02	gsl/1.16	2016/08/09	python3	2017/03/01
abaqus/2017	2017/10/02	hadoop/horton	2017/01/04	qe/6.0	2016/12/02
afni/14.03	2015/05/05	hdf5/1.10.0	2016/04/18	redis/3.0.7	2016/12/06
afni/15.12	2015/12/18	hdf5_parallel/1.10.0-patch1	2017/06/25	ruby/1.8.7	2015/11/17
anaconda	2017/02/17	hdf5_parallel/1.8.18	2017/06/25	ruby/2.0.0-p0	2015/11/17
android-studio/2.1.3	2016/09/08	hdf5_parallel/1.8.19	2017/06/26	ruby/2.2.3	2015/11/17
ansys/17.1	2016/08/23	intel/advisor/2016.1.0.423501	2015/10/19	saga	2013/05/23
ansys/18.1	2017/06/22	intel/advisor/2017.1.2.501009	2017/03/30	sambamba/0.6.6	2017/06/28
atomeye/20121115	2017/06/26	intel/compiler/2013_smp1.2.144	2016/09/26	samtools/0.1.18	2015/05/15
blast/2.2.31+	2015/11/24	intel/compiler/2016.0.109	2016/09/26	samtools/1.2	2015/05/15
boost/1.60-sl6	2016/04/18	intel/compiler/2016.4.258	2017/05/03	sas/9.3	2016/09/16
bowtie/0.12.9	2013/11/18	intel/compiler/2017.2.	2017/04/03	sas/9.4	2016/09/16

		174			
bowtie/2.0.5	2013/11/18	intel/inspector/2016.1.0.423441	2015/10/19	siesta/4.1-b1	2016/11/02
bowtie/2.2.5	2015/05/22	intel/inspector/2017.1.2.497970	2017/03/30	silos/4.10.2	2017/04/18
bwa/0.7.12	2015/05/15	intel/vtune_amplifier/2016.1.0.424694	2015/10/19	slatec	2014/04/21
caffe/local	2016/04/19	intel/vtune_amplifier/2017.2.0.499904	2017/03/30	spss/22	2016/09/01
casa/4.5.1	2016/02/11	ipopt/3.11.7	2015/06/15	spss/24	2016/09/01
cghub	2014/10/02	java8/1.8.144	2017/09/19	synopsys/customexplorer/L-2016.06-SP1	2016/09/21
chaste/3_3	2015/07/27	julia/0.4.6	2016/09/28	synopsys/fpga/L-2016.09	2016/09/21
cmake/3.1.0	2015/01/20	julia/0.5.0	2016/09/26	synopsys/hspice/L-2016.06-SP1	2016/09/21
cmake/3.9.4	2017/10/06	julia/0.5.0b	2016/09/28	synopsys/icc/L-2016.03-SP4	2016/09/21
comsol/5.1	2015/10/28	jython/2.7.0	2016/10/10	synopsys/identify/L-2016.09	2016/09/21
comsol/5.2a	2016/08/04	kentUtils/302	2015/10/08	synopsys/primetime/L-2016.06-SP2	2016/09/21
comsol/5.3	2017/07/31	lammps/cpu/30Jul16	2016/09/27	synopsys/tcad/sentaurus/L_2016.03-SP2	2017/09/28
condor	2014/07/28	lammps/gpu/30Jul16	2016/09/27	synopsys/tcad/taurus_medici/L_2016.03	2016/09/21
cryptopp/5.6.2	2015/08/06	lammps/serial/17Nov16	2017/01/18	synopsys/tcad/taurus_tsuprem4/L_2016.03	2016/09/21
cuda/6.5.14	2017/05/03	libreoffice/5.4.0	2017/08/29	tecplot/360ex/2016R2	2016/07/28
cuda/8.0.61	2017/05/03	maftools/01	2015/08/06	tecplot/chorus/2016R1	2016/07/28
cuda/7.5.5.1	2017/03/10	mathematica/10.2.0	2016/08/23	tecplot/chorus/2016R2	2016/09/01
cuda/8.0.5.1	2017/03/10	mathematica/10.3.1	2016/08/23	tecplot/focus/2016R2	2016/07/28
cufflinks/2.0.2	2015/05/22	mathematica/11.0.0	2016/08/23	tecplot/rs/2015R1	2016/07/28
cufflinks/2.2.1	2015/05/22	mathematica/11.1.1	2017/07/31	tmux/2.4	2017/05/20
eclipse/4.6	2016/09/16	matlab/2010b	2016/06/16	tophat	2013/12/02
eclipse/4.7	2017/09/11	matlab/2011b	2016/06/16	varanneal/20170926	2017/09/26
exonerate/2.2.0	2016/03/02	matlab/2013a	2016/06/16	visit/2.10.2	2017/04/18
f90	2013/03/21	matlab/2014a	2016/06/16	visit/2.12.2	2017/04/17
f90-64	2013/03/22	matlab/2015a	2016/06/16	vlfeat	2014/10/02

fftw3-intel	2014/04/30	matlab/2015b	2016/06/16	wien2k/13	2016/12/02
figtree/1.4.3	2017/09/07	matlab/2016a	2016/06/16	xpp/8.0	2017/08/15
flex	2012/10/05	mothur/1.39.5	2017/09/07	/afs/cad/ucs/module files/.hpc:	Directory
fluent/14.5.7	2016/08/23	mybin	2012/10/04	hpcbin	2013/07/10
fluent/15.0.0	2016/08/23	mysql/5.7	2017/02/02	hpcsbin	2013/08/06

Q: Please rate the suitability of software environment for your work.

Answer choices (single):

Excellent; Good; Fair; Poor; Don't know

Q: Is there software you want to use that is not currently available? (Check the above list for currently available software.)

You can provide up to five applications.

Answer choices (single):

Yes; No

The series of questions below are **seen only by participants** who indicate that they would like some specific currently unavailable software to be made available. The questions appear up to five times, depending on how many requests the participant wishes to provide. Thus, participants see either no question series, or as many as five.

Q: Provide the name of an application.

write-in

Q: Does this application have associated costs?

Answer choices (single):

Yes; No

Q: To what extent would you use this application for research and/or teaching?

Answer choices (single choice for each of two topics):

Research: High use; Medium use; Low use; No use

Teaching: High use; Medium use; Low use; No use

Q: Please provide any comments on the software environment.

write-in

Internet bandwidth, Science DMZ— *only for participants* who chose either of these topics

The text box below and two subsequent questions are *seen only by participants* who selected Internet bandwidth.

Internet bandwidth is the capacity of NJIT's connection to and from the Internet.

Q: Please rate the suitability of Internet bandwidth, including Internet 2 if applicable, for your work

Answer choices (single):

Excellent; Good; Fair; Poor; Don't know

Q: Please provide any comments on Internet bandwidth.

Write-in

The text box below and two subsequent questions are *seen only by participants* who selected Science DMZ.

"Science DMZ" refers to a computer subnetwork that is structured to be secure, but without the performance limits that would otherwise result from passing data through a stateful firewall. The Science DMZ is designed to handle high volume data transfers, typical with scientific and high-performance computing, by creating a special DMZ to accommodate those transfers.

Science DMZ is typically deployed at or near the local network perimeter, and is optimized for a moderate number of high-speed flows, rather than for general-purpose business systems or enterprise computing.

Q: Please rate the desirability of implementing a Science DMZ at NJIT as it relates to your work.

Answer choices (single):

High; Moderate; Low; Don't know

Q: Please provide any comments on Science DMZ.

write-in

Consultation: only for participants who chose this topic

Consultation is communication with Academic and Research Computing Systems (ARCS) staff in areas such as getting started in HPC and BD, problems encountered when running jobs, optimizing throughput, running parallel jobs, and managing disk space.

Q: Please indicate the effectiveness of consultation in your work.

Answer choices (single):

Excellent; Good; Fair; Poor; Don't know

Q: Please provide any comments on consultation.

write-in

End—for all participants

Q: What is your estimate of your research group's satisfaction with their use of IST-managed HPC and/or BD resources?

Answer choices (single):

High; Medium; Low; Not able to estimate at this time

Q: Please provide any comments on your research group's use of IST-managed HPC and/or BD resources.

write-in