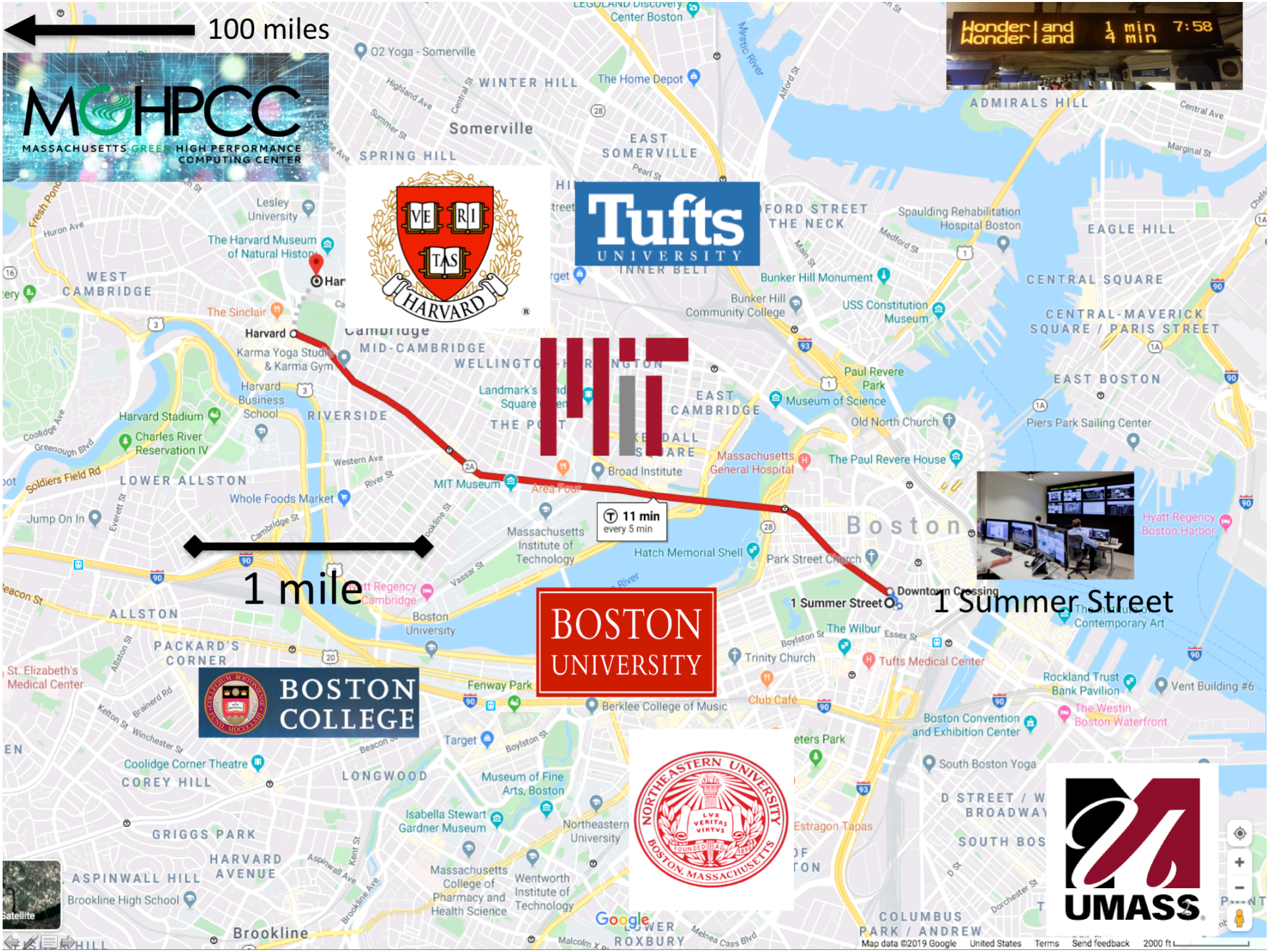
NET2 Evolution

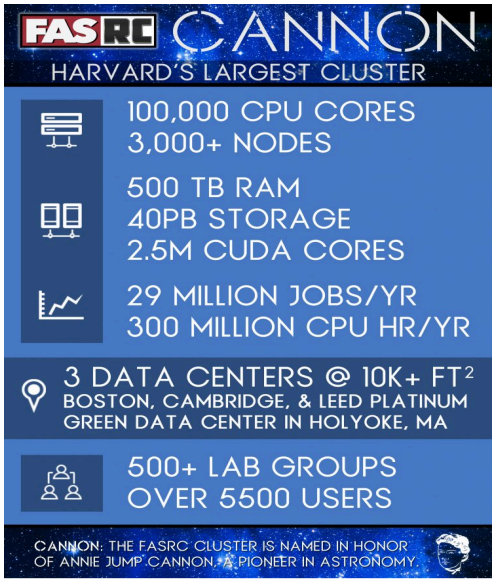
Facilities Meeting
October 30, 2019
Saul Youssef


100 miles
MGHPCC
MASSACHUSETTS GREEN HIGH PERFORMANCE COMPUTING CENTER
Wonderland 1 min 7:58
Wonderland 4 min
ADMIRALS HILL
Somerville
WINTER HILL
SPRING HILL
EAST SOMERVILLE
WEST CAMBRIDGE
MID-CAMBRIDGE
RIVERSIDE
LOWER ALLSTON
ALLSTON
PACKARD'S CORNER
COREY HILL
LONGWOOD
GRIGGS PARK
HARVARD AVENUE
ASPINWALL HILL
Brookline
EAST CAMBRIDGE
CENTRAL SQUARE
EAGLE HILL
EAST BOSTON
CENTRAL-MAVERICK SQUARE / PARIS STREET
FORD STREET THE NECK
COLUMBUS PARK / ANDREW
Boston
Harvard
MIT Museum
Broad Institute
Massachusetts General Hospital
Massachusetts Institute of Technology
Boston University
Fenway Park
Berklee College of Music
Northeastern University
Downtown Crossing
1 Summer Street
11 min every 5 min
1 mile
1 Summer Street
Tufts UNIVERSITY
HARVARD
BOSTON UNIVERSITY
BOSTON COLLEGE
UMASS
Google
Map data ©2019 Google
United States
Terms
Send feedback
2000 ft

MGHPCC

Boston University
Harvard University
MIT
Northeastern University
University of Massachusetts

- 15 megawatt $90M single purpose data center
- Near zero Carbon footprint
- Space power and cooling for 780 racks
- More than 300,000 x86 cores, millions of gpu cores
- 100Gb/s multi-fiber ring to internet2 and Esnet
- Three new top500 in the past few months
- Exascale storage plans via NESE project
- Located in Holyoke, MA
- Thousands of researchers, >100,000 student population

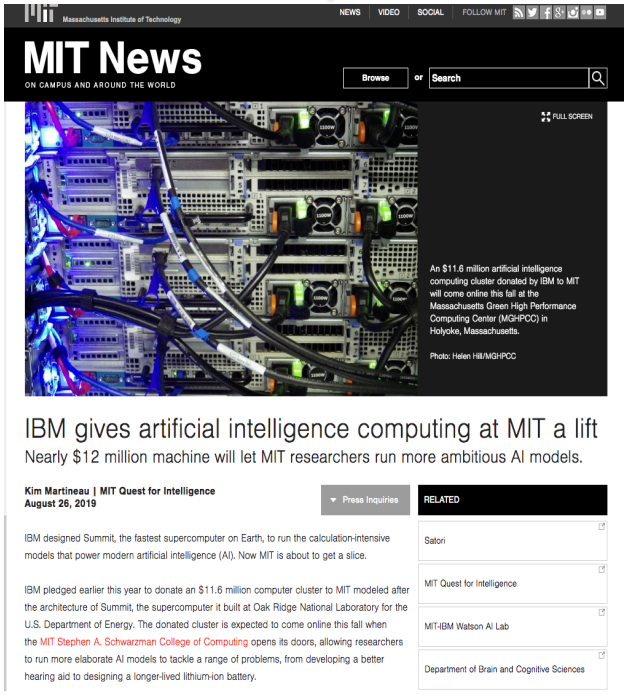


top500

M.I.T. Plans College for Artificial Intelligence, Backed by $1 Billion

The Massachusetts Institute of Technology is taking a particularly ambitious step in preparing students to develop, and consider the implications of, artificial intelligence. It is creating a new college, backed by a planned investment of $1 billion.
Cody O'Loughlin for The New York Times

Boston University launches University-wide Computing and Data Science initiative, new building.



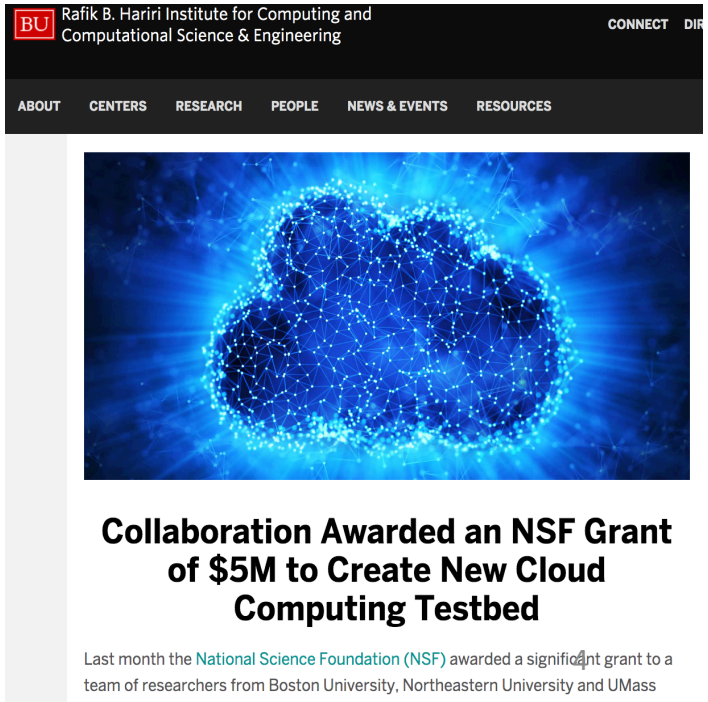
Lincoln Laboratory's new AI supercomputer is the most powerful at a university

TX-GAIA is tailor-made for crunching through deep neural network operations.

SEPTEMBER 26, 2019 | Kylie Foy | Communications & Community Outreach Office

TX-GAIA is housed inside of a new EcoPOD, manufactured by Hewlett Packard Enterprise, at the site of the Lincoln Laboratory Supercomputing Center in Holyoke, Massachusetts. Photo: Glen Cooper.

The new TX-GAIA (Green AI Accelerator) computing system at Lincoln Laboratory's Supercomputing Center (LLSC) has been ranked as the most powerful artificial intelligence (AI) supercomputer at any university in the world. The ranking comes from TOP500, which publishes a list of the top supercomputers in various categories biannually. The system, which was built by Hewlett Packard Enterprise, combines traditional high-performance computing hardware — nearly 900 Intel processors — with hardware optimized for AI applications — 900 Nvidia GPU accelerators.



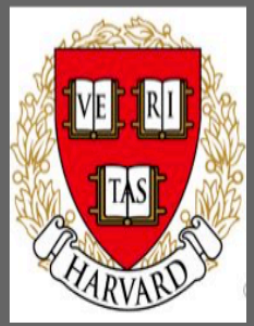
Collaboration Awarded an NSF Grant of $5M to Create New Cloud Computing Testbed

Last month the National Science Foundation (NSF) awarded a significant grant to a team of researchers from Boston University, Northeastern University and UMass Amherst. The grant will support the development of a testbed for new cloud

NESE: Northeast Storage Exchange

National Science Foundation CIF21 DIBBs award 1753840

Boston University
Harvard University
Massachusetts Institute of Technology
Northeastern University
University of Massachusetts



PHASE I: HARDWARE SPECS

Single Socket Intel Xeon 4114 Skylake 2.2 Ghz (3.0 Ghz Turbo), 10C/20T with AVX-512 (AES Encryption offloading)

Bonded 2x10G SFP+ (NESE Net)
Single 1Gbps OBM (Admin net)

Deployed at two sites with 100Gb ethernet.

12 x 10TB HDD Data Drive

4 x 480GB SSD 3x RocksDB
1x CephFS
metadata OSD

1 x 64GB NVMe Optane ~~RockDB WAL~~

2x 100GB SATADOM SW RAID1 Root Filesystem

Starter deployment is 15 PB raw already with ~50% buy-in.

Respond effectively to data growth

The Dell EMC PowerEdge R740xd2 helps you plan for future growth with large internal storage and cost-efficient drive capacities. Deliver two-socket compute performance with flash and fast networking options to meet streaming demands. Simplify management of large data sets with automated administration and front-serviceable drives. The R740xd2 lets you keep your data safely on-premise with built-in security, even as you scale capacity.

<- Major buy of DELL r740xd2 with 14TB drives, 25G nics

6PB raw for NET2
3PB raw for Northeastern University
7PB more expected soon

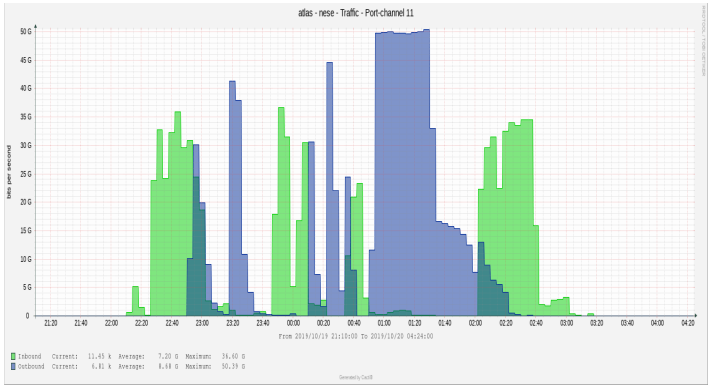
Northern Crossroads at MGHPCC

100G

100G

BU Networking

Harvard Networking



Demo saturating 2 gateways

2 x 100G dedicate for NET2

gridftp.nese.mghpcc.org -> DNS round robin

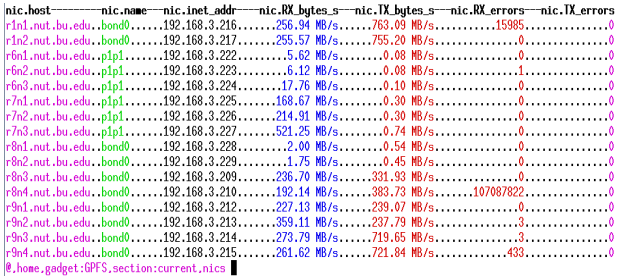
DELL S5048-ON

ARISTA 100G
ARISTA 100G
NESE gateway 25G
NESE gateway 25G
NESE gateway 25G
NESE gateway 25G
NESE gateway 25G
4 more on order

DELL S5048-ON
DELL S5048-ON
OSD
OSD
.
.
.
OSD

...

26 x 10G



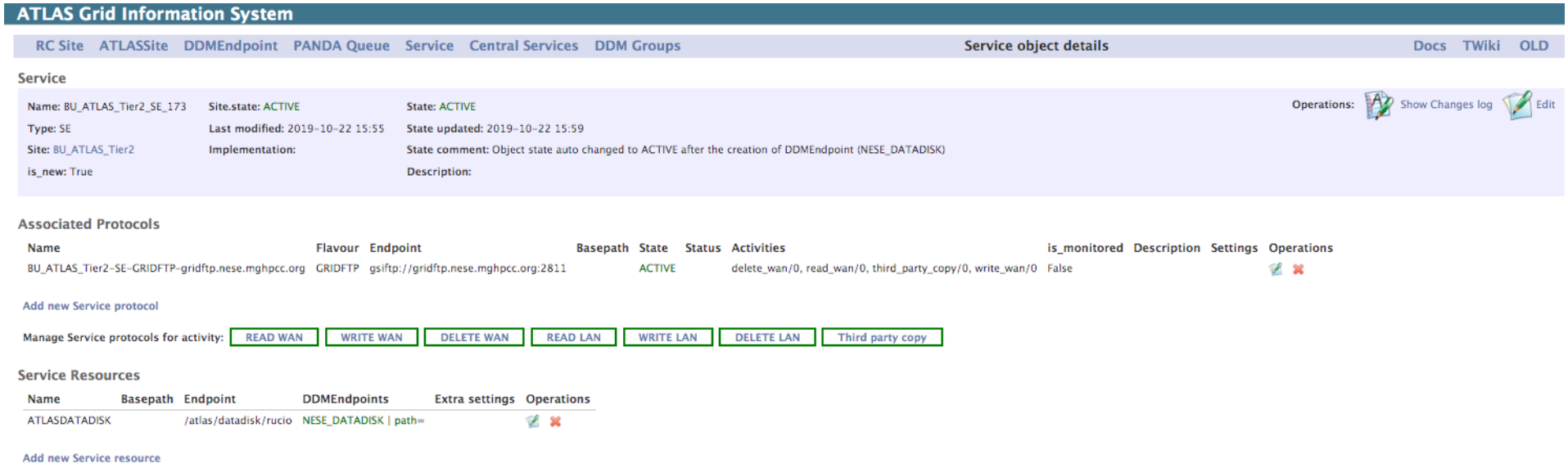
nic.host	nic.name	nic.inet_addr	nic.RX_bytes_s	nic.TX_bytes_s	nic.RX_errors	nic.TX_errors
r1n1.nut.bu.edu	bond0	192.168.3.216	256.94 MB/s	763.09 MB/s	15985	0
r1n2.nut.bu.edu	bond0	192.168.3.217	255.57 MB/s	755.20 MB/s	0	0
r6n1.nut.bu.edu	p1p1	192.168.3.222	5.62 MB/s	0.08 MB/s	0	0
r6n2.nut.bu.edu	p1p1	192.168.3.223	6.12 MB/s	0.08 MB/s	1	0
r6n3.nut.bu.edu	p1p1	192.168.3.224	17.76 MB/s	0.10 MB/s	0	0
r7n1.nut.bu.edu	p1p1	192.168.3.225	168.67 MB/s	0.30 MB/s	0	0
r7n2.nut.bu.edu	p1p1	192.168.3.226	214.91 MB/s	0.30 MB/s	0	0
r7n3.nut.bu.edu	p1p1	192.168.3.227	521.25 MB/s	0.74 MB/s	0	0
r8n1.nut.bu.edu	bond0	192.168.3.228	2.00 MB/s	0.54 MB/s	0	0
r8n2.nut.bu.edu	bond0	192.168.3.229	1.75 MB/s	0.45 MB/s	0	0
r8n3.nut.bu.edu	bond0	192.168.3.209	236.70 MB/s	331.93 MB/s	0	0
r8n4.nut.bu.edu	bond0	192.168.3.210	192.14 MB/s	383.73 MB/s	107087822	0
r9n1.nut.bu.edu	bond0	192.168.3.212	227.13 MB/s	239.07 MB/s	0	0
r9n2.nut.bu.edu	bond0	192.168.3.213	359.11 MB/s	237.79 MB/s	3	0
r9n3.nut.bu.edu	bond0	192.168.3.214	273.79 MB/s	719.65 MB/s	3	0
r9n4.nut.bu.edu	bond0	192.168.3.215	261.62 MB/s	721.84 MB/s	433	0

@.home.gadget:GPFS,section:current,nics

5PB GPFS

8.2 PB raw -> 6.0 PB useable for NET2 with 8+3 EC

1. Retire old Harvard worker nodes
2. Retire LSM software, switching to rucio-mover
3. 100G dedicated networking NET2-NESE
4. Docker containers with Gridftp & Wei's adler32 callout
5. DNS round robin from gridftp.nese.mghpcc.org to gridftp01,02,…



The plan is to use both NET2_DATADISK (read/write)
NESE_DATADISK (read only)

…for our PanDA queues

We might give Globus endpoint space to our BU/Harvard/UMass Tier 3 peeps at some point.