

Advanced Research Computing and Globus

Todd Raeker
raeker@umich.edu
www.arc-ts.umich.edu



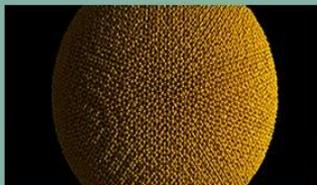
ARC·TS

ADVANCED RESEARCH COMPUTING
TECHNOLOGY SERVICES
UNIVERSITY OF MICHIGAN



Leading advances in data-intensive and computational research

COMPUTATIONAL SCIENCE



The Michigan Institute for Computational Discovery and Engineering (MICDE) focuses on the interdisciplinary development of mathematical algorithms and models on high performance computers.



DATA SCIENCE



The Michigan Institute for Data Science (MIDAS) is the focal point for the new multidisciplinary area of data science at the University of Michigan.



TECHNOLOGY SERVICES



Advanced Research Computing – Technology Services (ARC-TS) provides access to and support for the use of advanced computing resources



CONSULTING SERVICES



Consulting for Statistics, Computing and Analytics Research (CSCAR) provides support and training relating to the management, collection, and analysis of data.



Research Data Storage at UM

ARC-TS

High Performance for all of Campus*	1.85 PB
Scratch on main cluster*	1.25 PB
Special purpose cluster**	1.35 PB
Private cloud service	0.25 PB
Hadoop/Spark cluster	0.30 PB
DSI special purpose cluster	1.00 PB
OSIRIS	5.00 PB

Total

~10 PB

Research Data Storage at UM

ARC-TS		Others	
High Performance for all of Campus*	1.85 PB	Central IT*	1.5 PB
Scratch on main cluster*	1.25 PB	Michigan Health IT*@	2.5 PB
Special purpose cluster**	1.35 PB	Colleges and Schools*@	2 PB
Private cloud service	0.25 PB	Individual research groups*@	>20 PB
Hadoop/Spark cluster	0.30 PB		
DSI special purpose cluster	1.00 PB		
OSiRIS	5.00 PB		
Total	~10 PB	Total	~ 26 PB

Globus at UM

ARC-TS umich# subscription

~20 active endpoints

umich#flux

- local scratch
- mounts points for HP Storage

umich#conflux (Power8)

- local GPFS only

Other units with storage umich#

- clone admin created endpoint to umich#

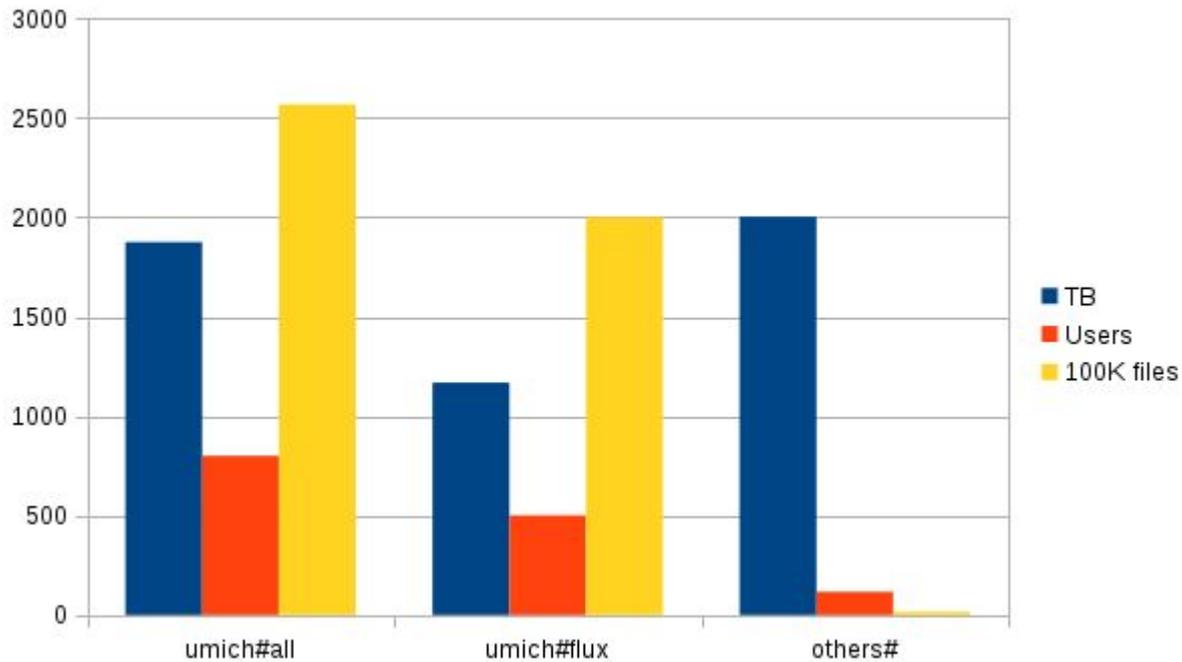
Others

~10 active endpoints

never associated with umich management

special high profile projects

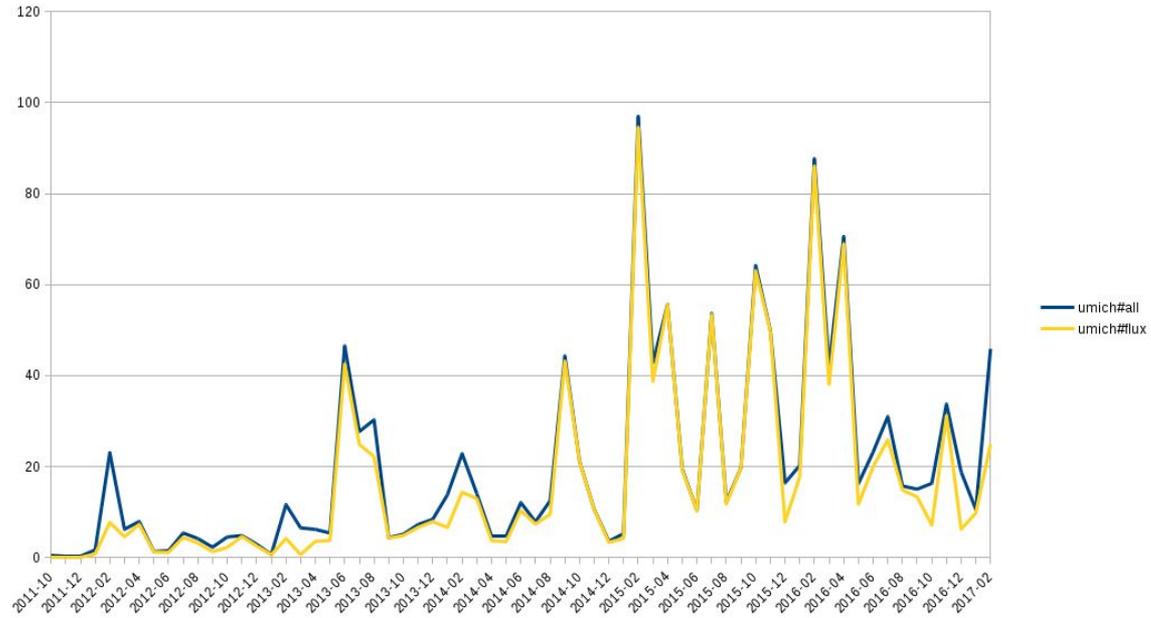
Total Globus usage for all of UM



ARC·TS

ADVANCED RESEARCH COMPUTING
TECHNOLOGY SERVICES
UNIVERSITY OF MICHIGAN

Globus usage for endpoints under managed subscription



General Observations

Storage systems widely distributed on Campus

Many groups do not see need for managed subscription (no sharing)

If sharing wanted not happy with umich# versus their own identity (Display Name is a good thing)

Future Directions

Consolidate Storage systems and endpoints where possible.

DSI will be putting heavy demands on Storage and Data Transfer.

Think about Data Management now rather than after the fact

Faculty Data Management Plans. Library taking lead, where does ARC/ARC-TS fit in?

Secure Data computing (data in and out concerns).

What would we like to see from Globus

Can we somehow associate these “other” endpoints with subscription for monitoring?

What would we like to see from Globus

Can we somehow associate these “other” endpoints with subscription for monitoring?

Support for endpoint writing to HDFS.

What would we like to see from Globus

Can we somehow associate these “other” endpoints with subscription for monitoring?

Support for endpoint writing to HDFS.

HIPAA?

What would we like to see from Globus

Can we somehow associate these “other” endpoints with subscription for monitoring?

Support for endpoint writing to HDFS.

HIPAA?

Public cloud support. Azure and Google?

What would we like to see from Globus

Can we somehow associate these “other” endpoints with subscription for monitoring?

Support for endpoint writing to HDFS?

HIPAA?

Public cloud support. Azure and Google?

Power8 and ARM?