ACM HPDC 2018
The 27th International Symposium on
High-Performance Parallel and Distributed Computing

Argonne NATIONAL LABORATORY

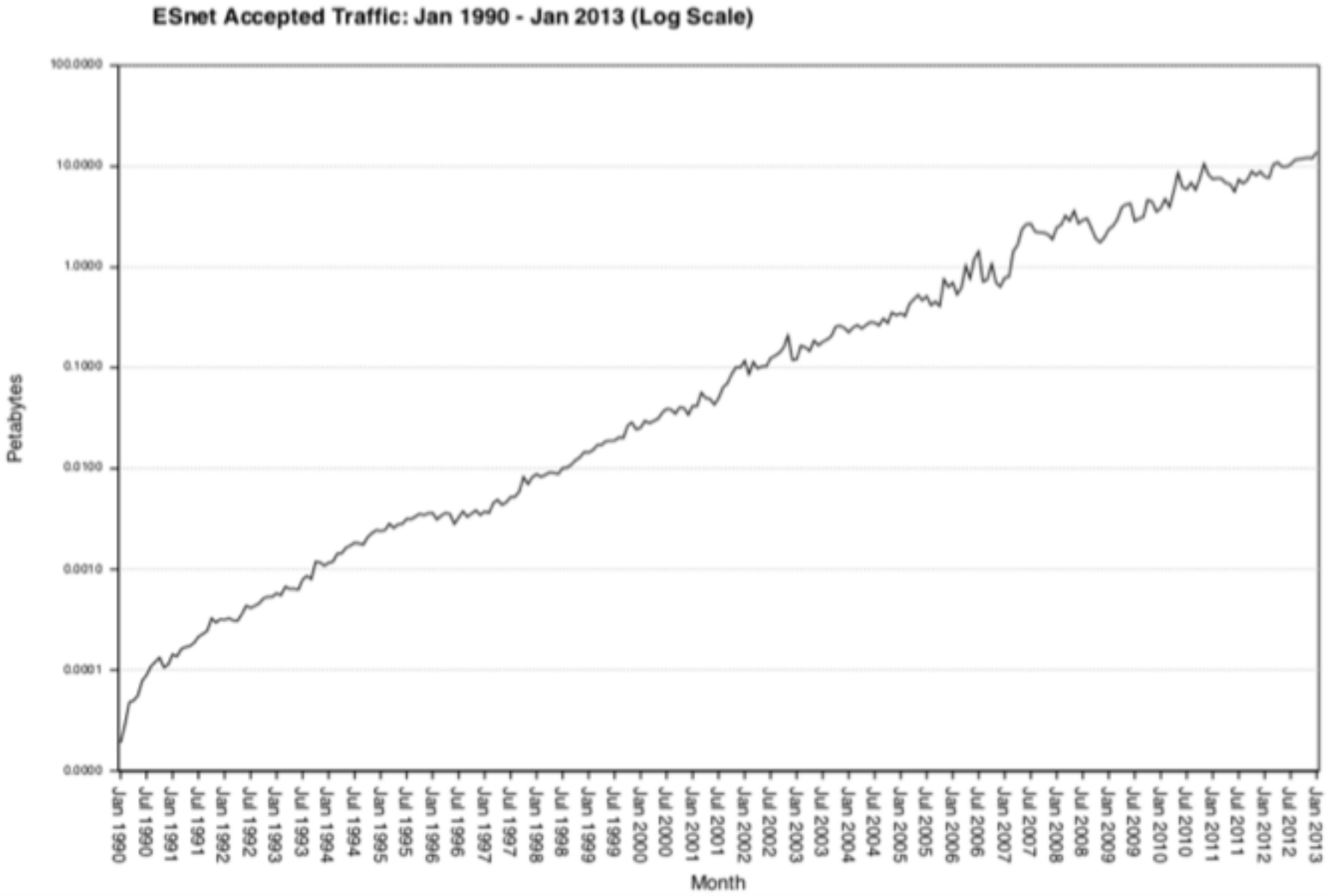
Cross-Geography Scientific Data Transferring Trends and Behavior

Zhengchun Liu, Rajkumar Kettimuthu, Ian Foster and Nageswara S.V. Rao

Presented by: Rajkumar Kettimuthu

Tempe, Arizona, USA - June 15, 2018

WIDE AREA SCIENCE DATA TRANSFERS

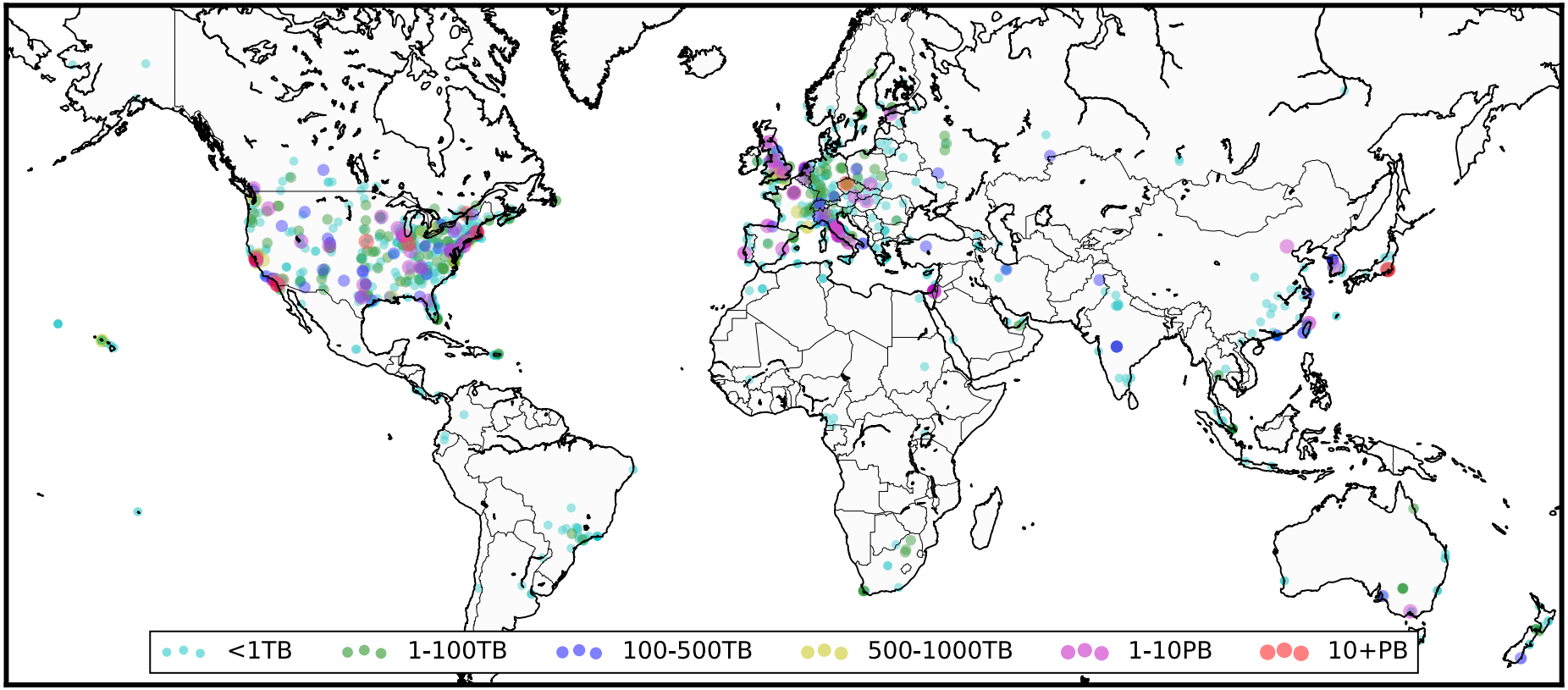


OUR STUDY

- We use file transfer application logs to characterize wide area science data transfers over a four-year period
- Resource providers optimize the resources used for data transfer
- Researchers and tool developers build new (or optimizing the existing) data transfer protocols and tools
- End users organize their datasets to maximize performance
- Funding agencies plan investments
- Results provide a number of insights
  - Utilization of DTNs, Data corruption
  - Repeat transfers, File types transferred
  - Transfer performance, User behavior

GRIDFTP (65% OF INCOMING, 42% OF OUTGOING ESNET TRAFFIC TO/FROM DOE LABS IN 2017)

- End-to-end wide area file transfers carried out by tools such as GridFTP, FTP, rsync, SCP, BBCP, FDT, XDD, Aspera, etc.
- Geographical distribution of bytes moved by GridFTP, per city in 2017



GRIDFTP

- GridFTP, an extension of the standard FTP protocol for high performance, better security, and reliability
- Standardized through the Open Grid Forum
- Multiple implementations of that standard exist
- Globus and dCache implementation are the most popular
- Many different implementations of GridFTP clients exist
- Top 5 heavily used clients

Petabytes and millions of files transferred via **GridFTP** using different tools over the past four years

Year	fts_url_copy		libglobus_ftp_client		globusonline-fxp		globus-url-copy		gfal2-util		Total	
	PBytes	MFiles	PBytes	MFiles	PBytes	MFiles	PBytes	MFiles	PBytes	MFiles	PBytes	MFiles
2014	N/A	N/A	111.23	746.59	39.81	1646.10	13.13	816.67	N/A	N/A	176.24	3431.78
2015	48.09	77.29	103.21	841.96	52.89	2424.58	19.27	947.78	0.93	6.70	267.33	4435.13
2016	244.46	295.67	105.75	998.96	88.56	3600.78	14.76	850.76	10.03	74.05	466.91	5922.83
2017	342.12	550.57	40.11	885.65	113.45	3901.27	16.89	898.14	45.93	234.65	585.01	6671.79
Total	634.67	923.53	360.3	3,473.16	294.71	11,572.73	64.05	3,513.35	56.89	315.4	1,495.49	20,461.53

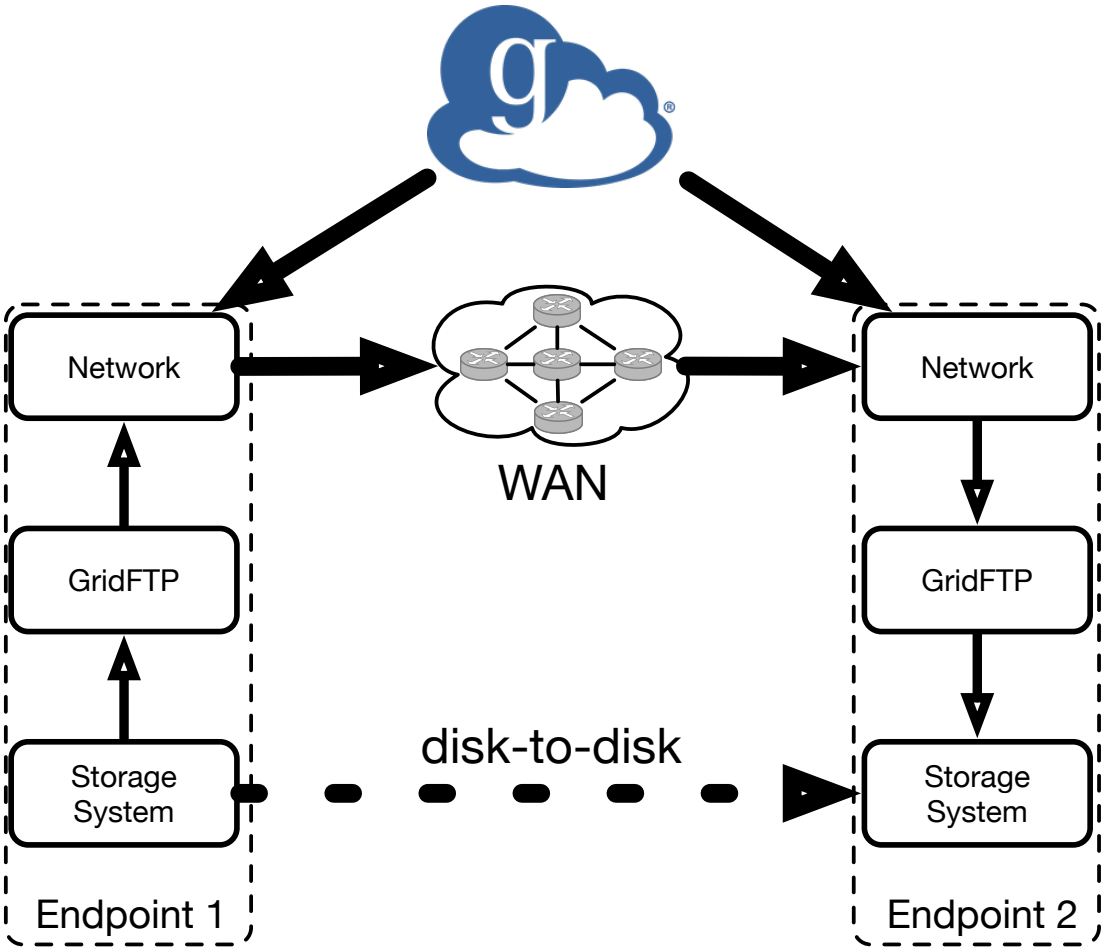
GRIDFTP

- Globus implementation reports limited usage information

Key	Value	Description
num_streams	4	Parallel TCP streams
appname	globusonline-fxp	Application name
hostname	grid-cr2.desy.de	Server hostname
start_time	1452637794.757349	POSIX time with μsec
ftp_return_code	226	RFC959 completion code
ip_address	212.189.205.173	Host IP address
num_bytes	976526	Bytes transferred (file size)
end_time	1452637794.877834	POSIX time with μsec
trans_type	STOR	FTP command (RFC959)
buffer_size	174758	TCP buffer size

- Command logs for 20B files (1.8 EB of data) transferred between any two of 63K endpoints from 2014 to 2017
- Limitations
  - No information on the other end of the transfer
  - Only file level information, no mapping of files to transfer requests

Globus transfer service (Globus online)



- Globus transfer service logs have more information
- ~5M transfers (13.1 billion files and 305.8 PB) involving ~42K unique endpoints, ~72K unique edges pairs and 26K users

Analysis overview

We group our analysis into four categories:

Dataset characteristics

dataset size, # files, average file size, directories, file type and dataset sharing behavior

Transfer characteristics

Data integrity checking, encryption, and reliability, transfer direction, performance, duration and transfer parameters

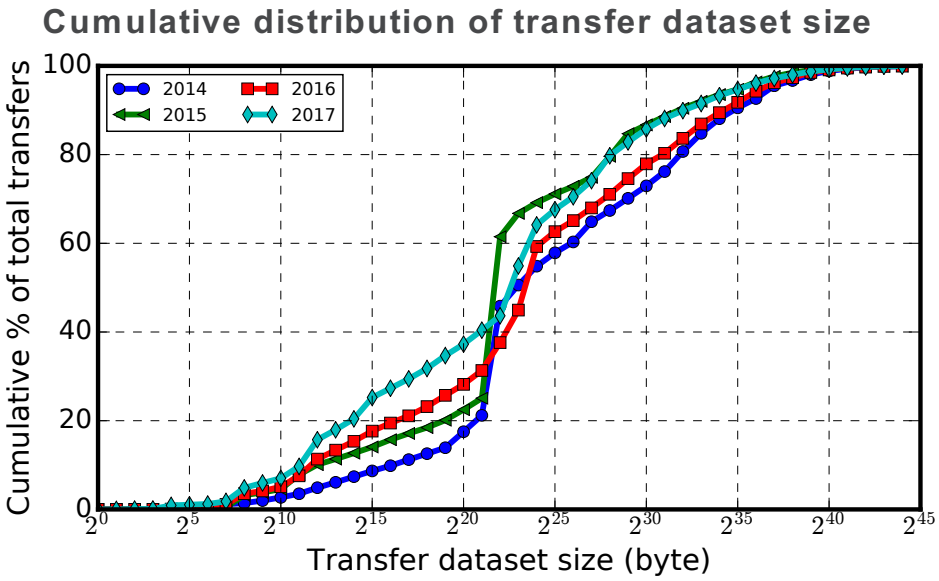
User behavior

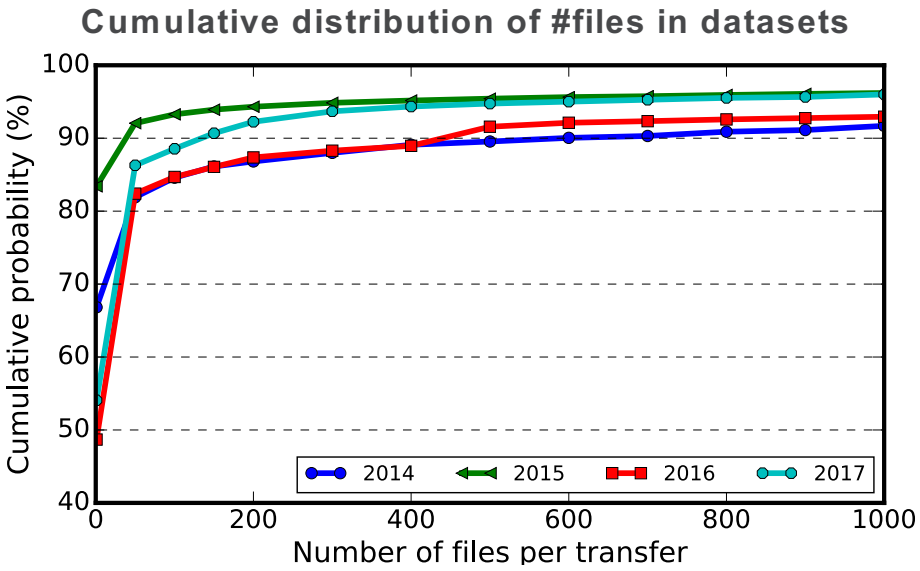
transfer frequency, transfer volume, degree of connection to endpoints and pattern of users access endpoint

Endpoint characteristics

degree of sharing to users, resource utilization (idle time percentage), source-to-destination edge

Dataset size and number of files





- Most of the datasets moved over the wide area are small
- 50th, 75th, 95th percentiles are 6.3 MB, 221.5 MB, 55.8 GB
- Dataset size has decreased year by year from 2014 to 2017
- Most of the datasets transferred by Globus have only one file
- And 17.6% of those datasets (or 11% of the total) have a file size of ≥ 100 MB

File size

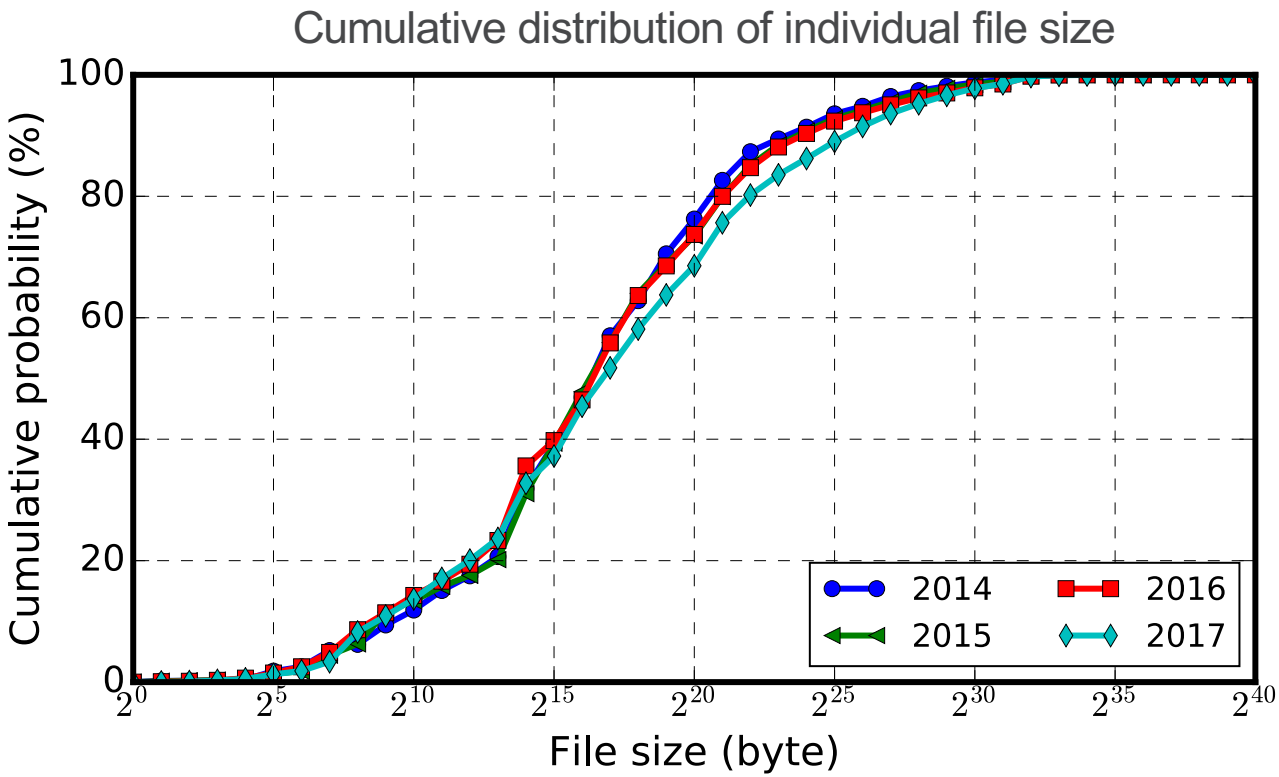


Table 4: Average file size (in MB) by application and year.

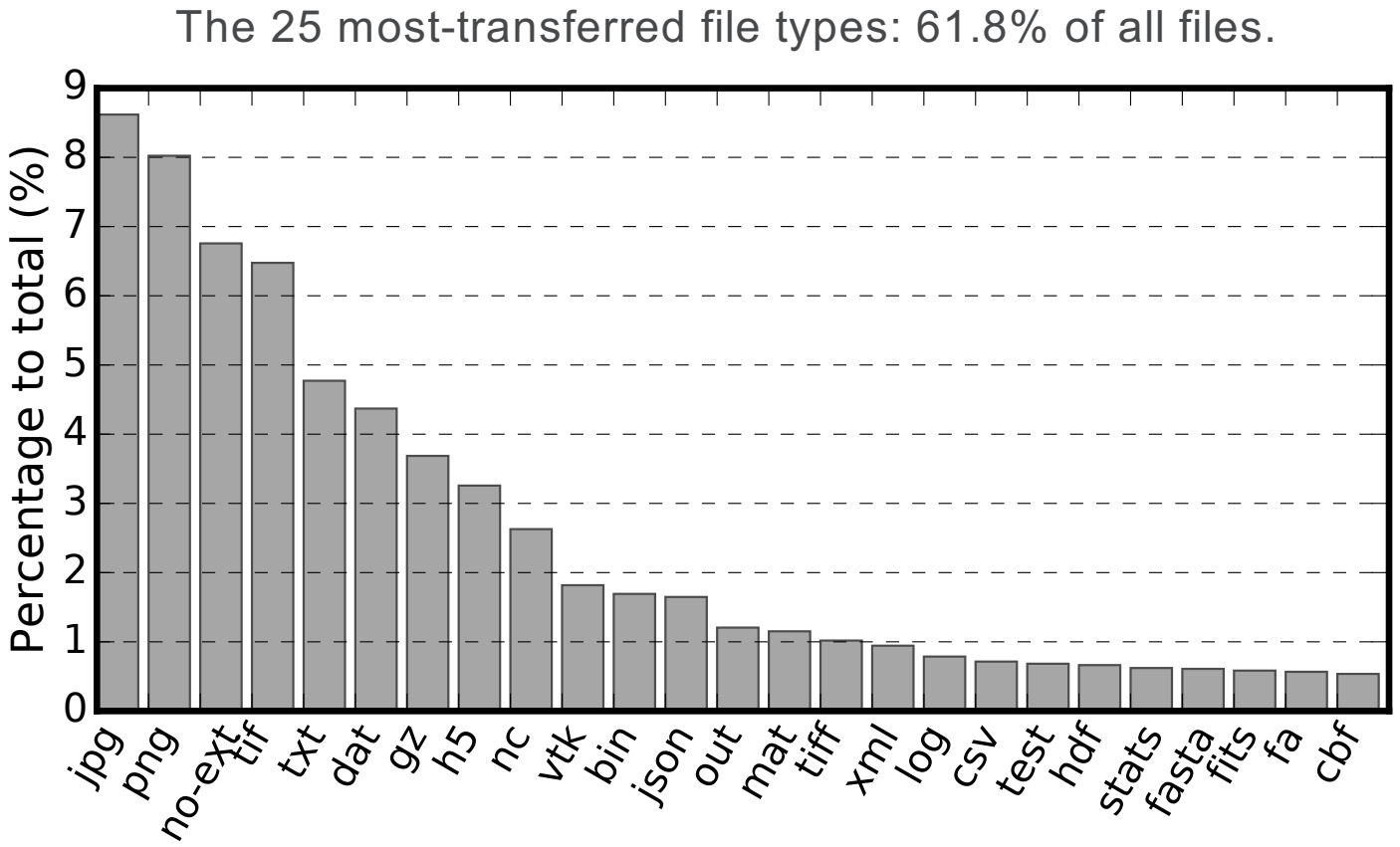
Year	fts_url_copy [5]	libglobus_ftp_client	globusonline-fxp [34]	globus-url-copy [2]	gfal2-util [6]	Overall
2014	–	142.96	27.31	8.86	–	53.89
2015	652.44	133.78	23.89	18.41	32.72	69.18
2016	856.98	193.83	26.28	45.20	252.22	105.28
2017	719.65	153.42	30.78	29.18	182.29	111.81

Repeat transfers

N_{src}	N_{dst}	N_{usr}	N_{trs}	Size
1	120	111	131	10.2GB
3	26	24	73	5.0MB
7	8	3	72	14.7GB
1	58	57	64	9.1GB
9	7	6	53	170.4MB
3	12	33	52	3.1GB
1	4	30	51	3.1GB
1	44	43	51	9.3GB
1	47	47	49	8.3GB
1	4	32	42	365.0MB
2	39	39	40	7.4GB
1	5	4	33	3.7GB
2	6	6	31	17.7GB
1	17	17	25	13.3MB
1	4	17	25	0.3MB

- Less than 7.7% of the datasets are transferred more than once
- Distributed mostly from one (or a few) endpoints to multiple destinations
- Multiple users transfer the same data to the same destination

File type



- Image files are the most common file type transferred, followed by raw text files
- Scientific formats such as .h5 (hierarchical data format) and .nc (NetCDF) are in the top 10

Analysis overview

We group our analysis into four categories:

Dataset characteristics

dataset size, # files, average file size, directories, file type and dataset sharing behavior

Transfer characteristics

Data integrity checking, encryption, and reliability, transfer direction, performance, duration and transfer parameters

User behavior

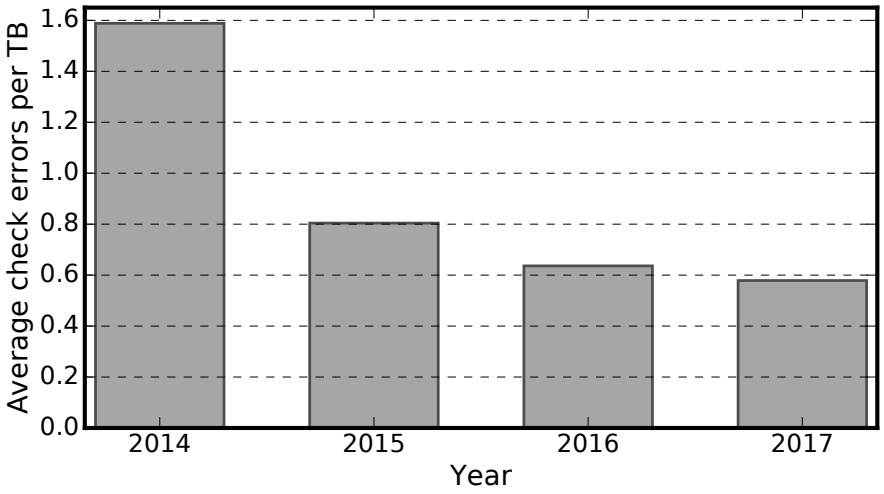
transfer frequency, transfer volume, degree of connection to endpoints and pattern of users access endpoint

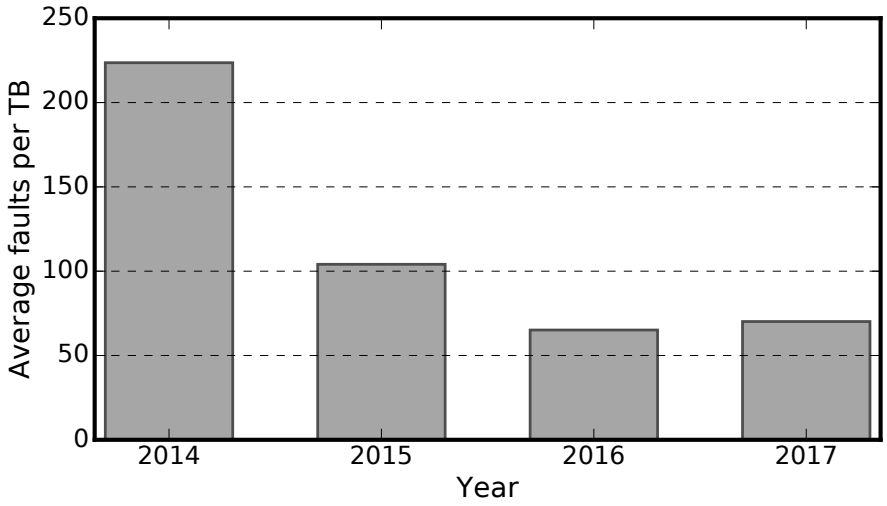
Endpoint characteristics

degree of sharing to users, resource utilization (idle time percentage), source-to-destination edge

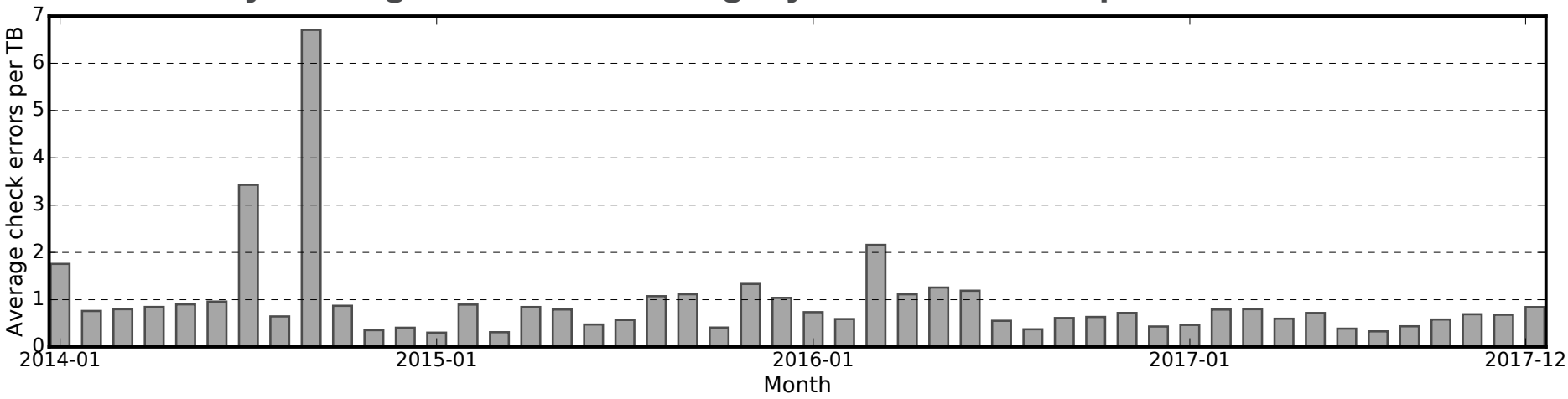
Checksum failures and other faults

Annual average number of checksum failures and faults per TB transferred

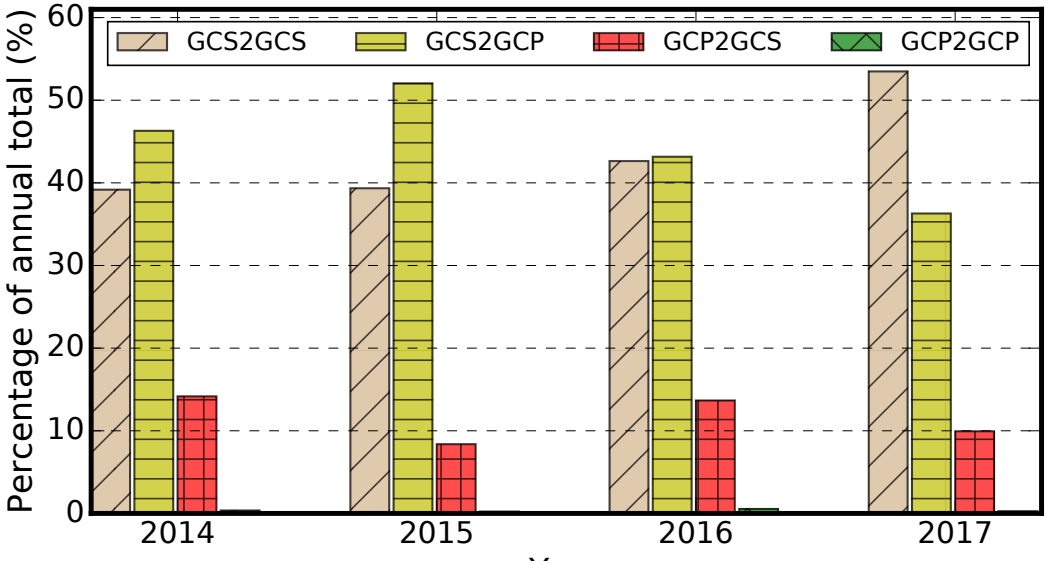




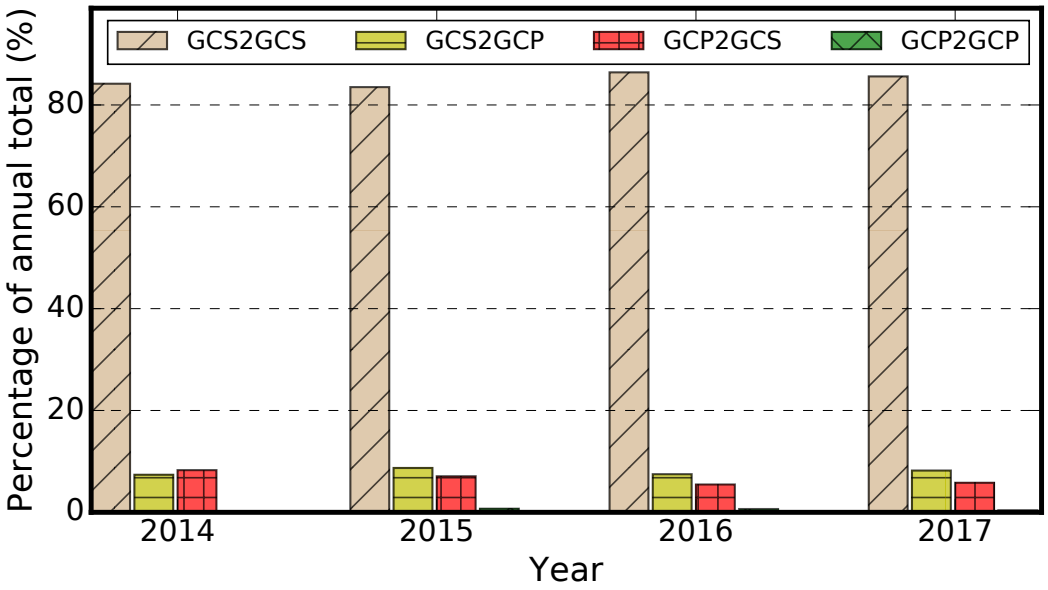
Monthly average number of integrity check failures per TB transferred



Server-to-server transfers, downloads, uploads

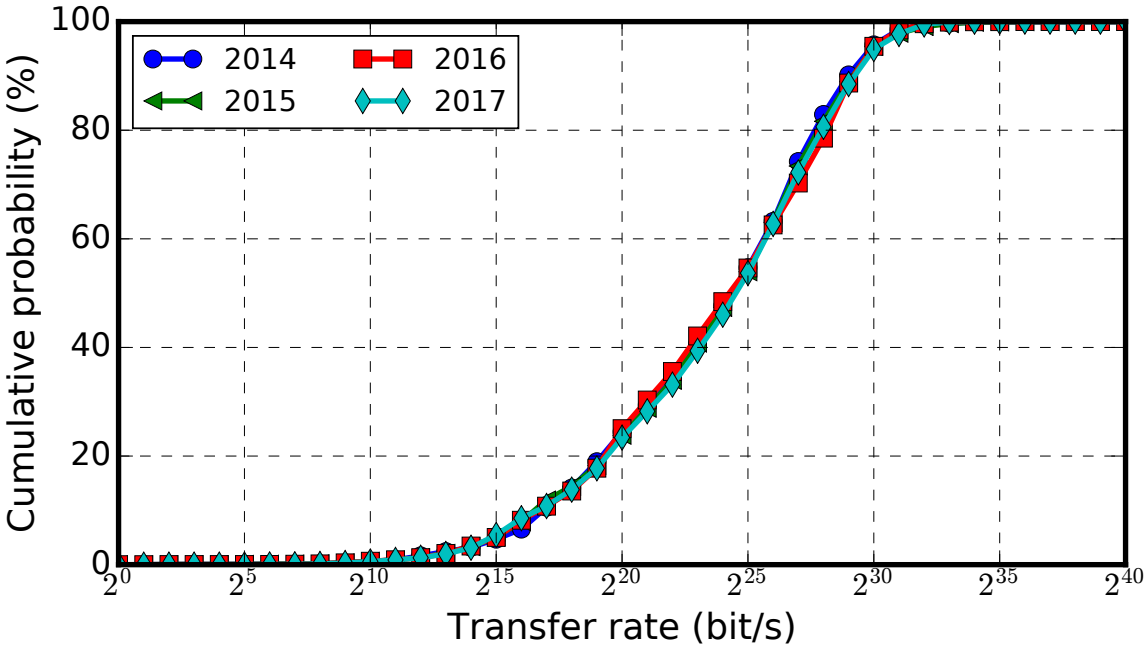


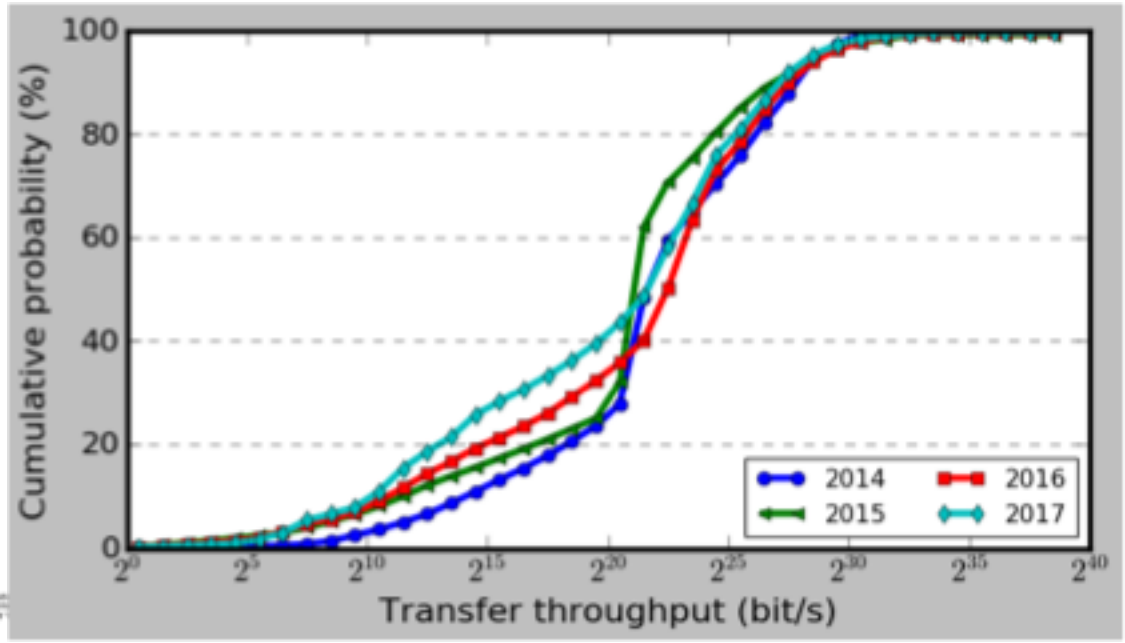
transfers



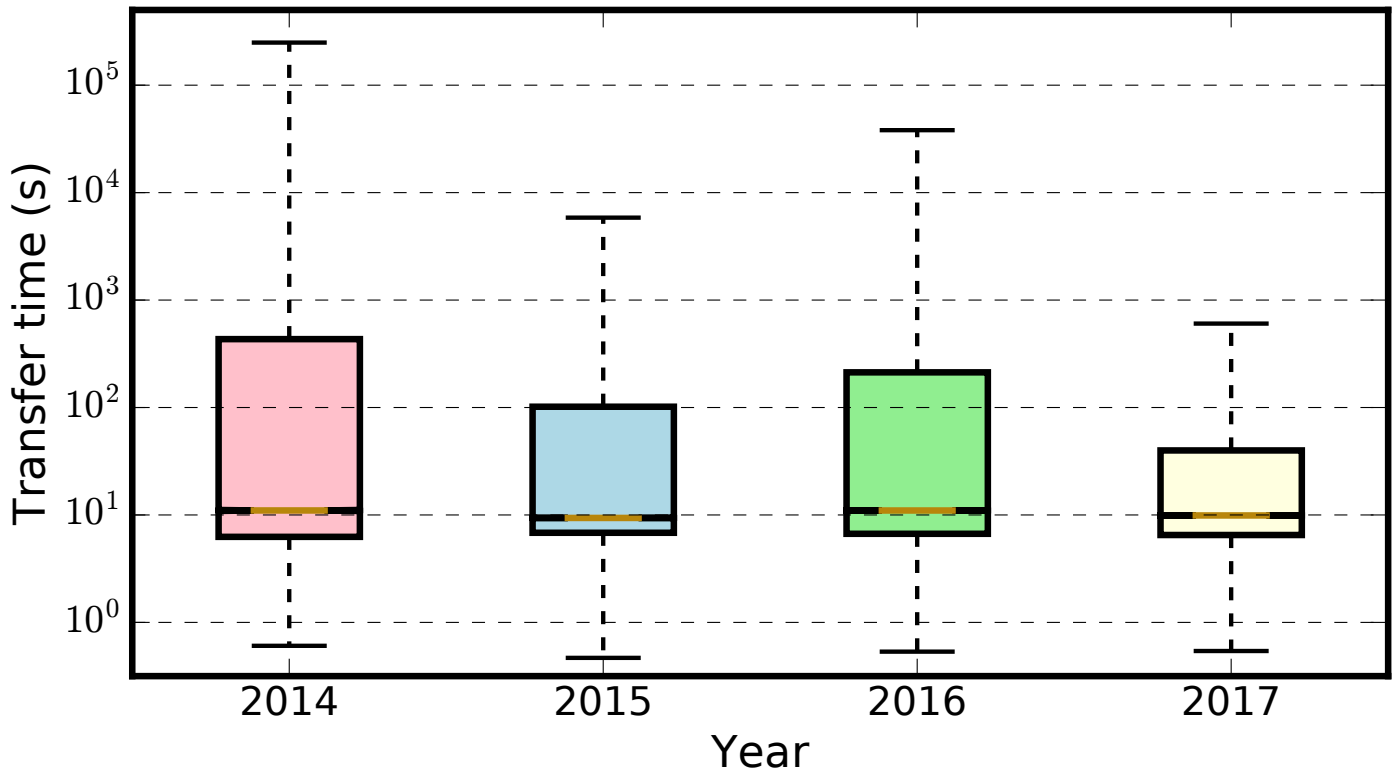
Bytes transferred

Transfer performance





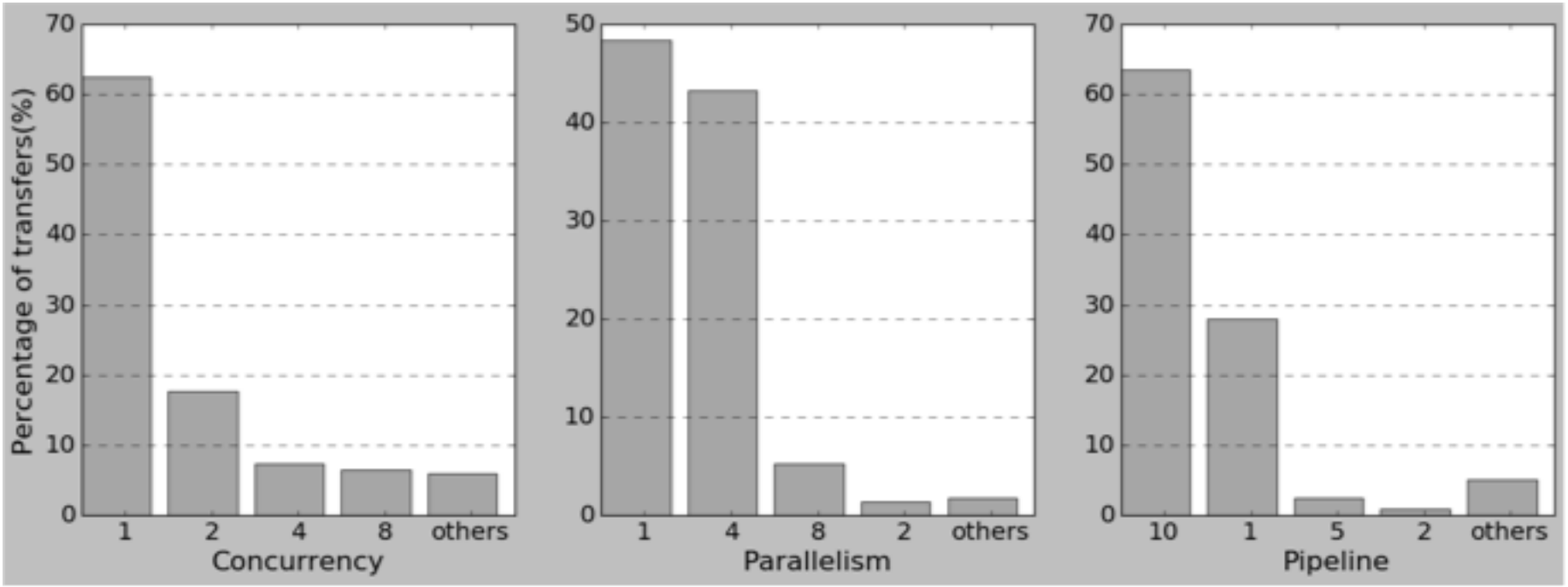
Transfer duration



- 50% of all the transfers finished in less than 10 seconds
- The longest-running transfer to date ran for six months
- 0.004% ran for more than a month, 0.09% for more than a week, 1.2% for more than a day, 8% for more than an hour

Transfer tuning parameters

- 94.6% of globus-url-copy transfers, 93.4% of fts_url_copy transfers and almost all of gfal2-util transfers use 1 TCP stream (default)



- Most users do not manually tune the transfer parameters
- Transfer tools should be smart enough to choose the best parameters to achieve maximum performance

Analysis overview

We group our analysis into four categories:

Dataset characteristics

dataset size, # files, average file size, directories, file type and dataset sharing behavior

Transfer characteristics

Data integrity checking, encryption, and reliability, transfer direction, performance, duration and transfer parameters

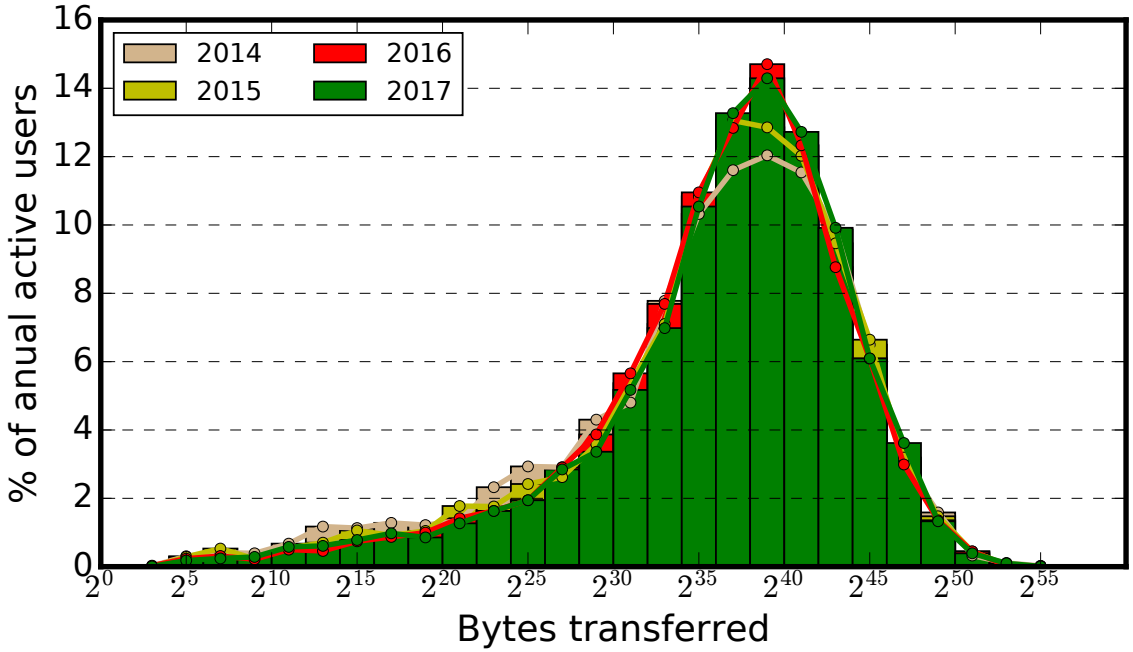
User behavior

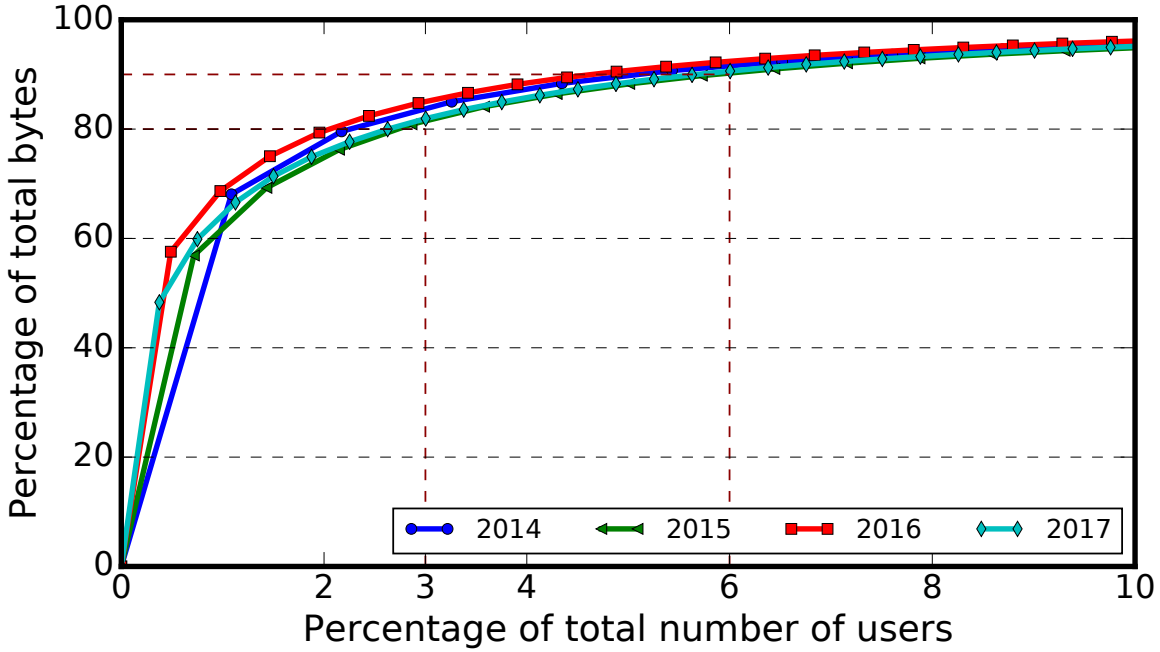
transfer frequency, transfer volume, degree of connection to endpoints and pattern of users access endpoint

Endpoint characteristics

degree of sharing to users, resource utilization (idle time percentage), source-to-destination edge

Distribution of bytes transferred by users

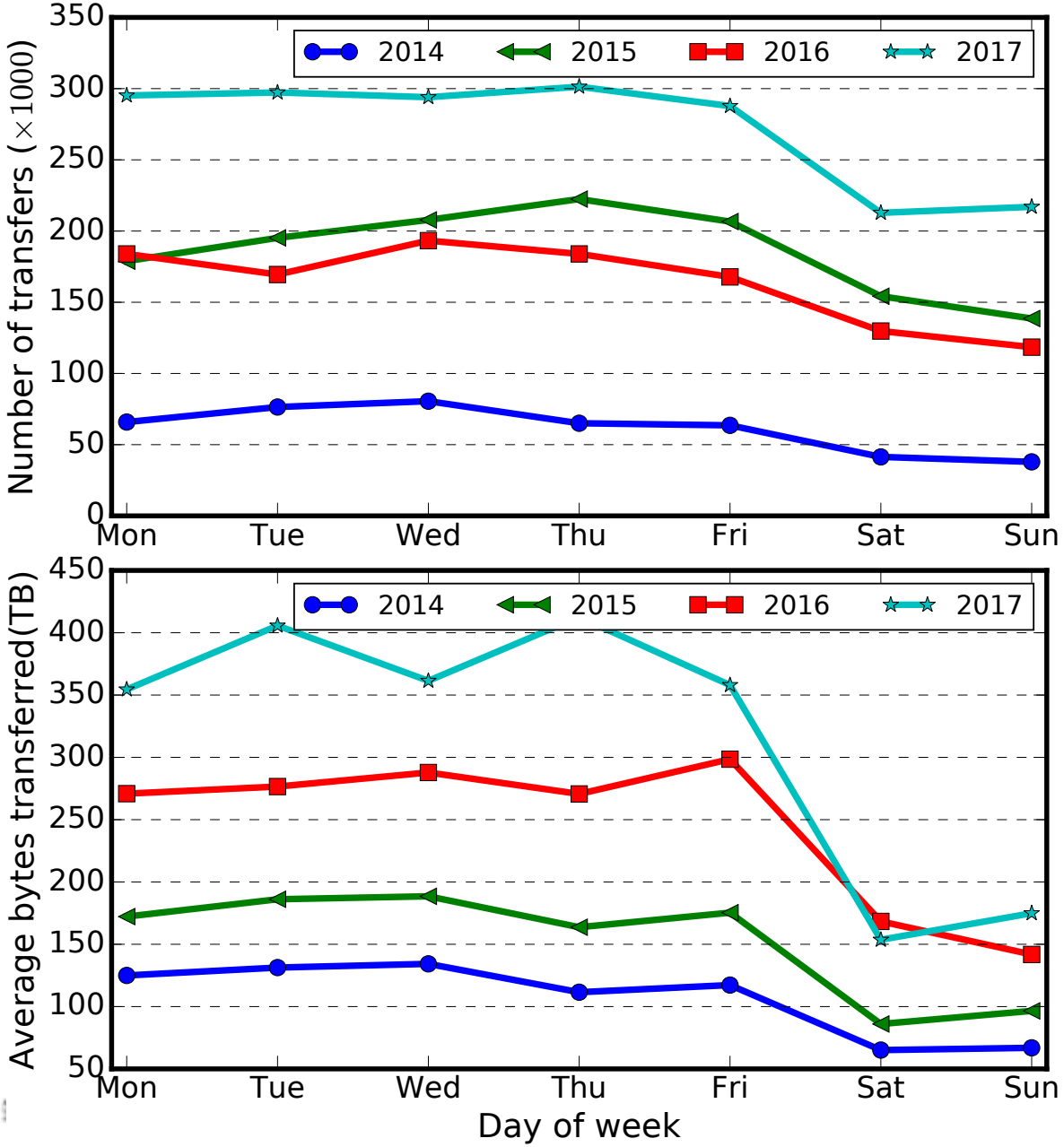




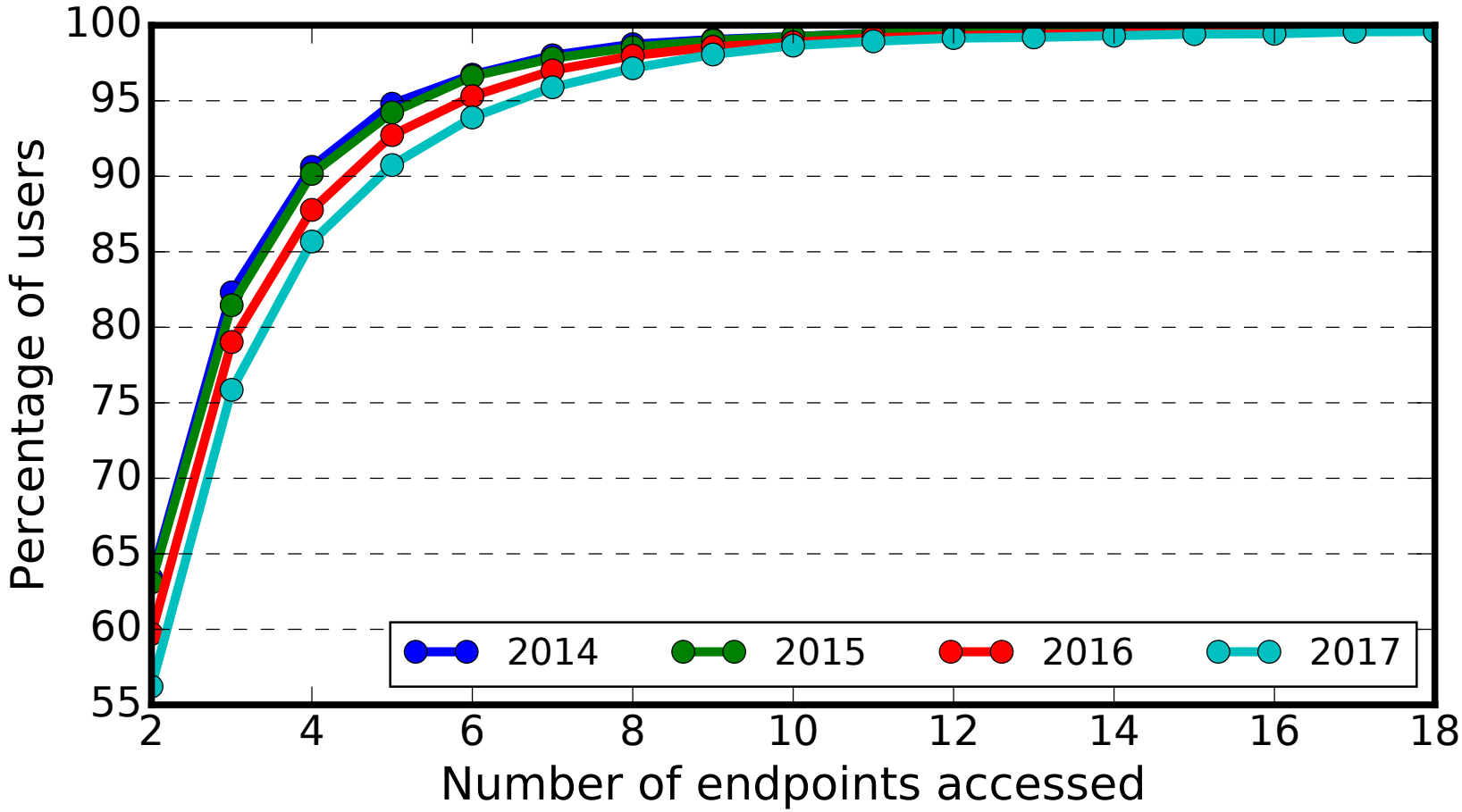
- Most users transferred dozens of gigabytes
- Few users who transferred 100s of TBs accounted for the majority of total bytes moved
- Of all the bytes transferred, 80% are by just 3% of all users; 10% of the users transferred 95% of the data

Argonne

User transfer pattern



Users access of endpoints



- >50% users accessed 2 or fewer endpoints
- 1.5% users accessed only one endpoint (copy files locally)
- Local copy – 71K transfers, 17PB, ~3K users, ~2K endpoints

Analysis overview

We group our analysis into four categories:

Dataset characteristics

dataset size, # files, average file size, directories, file type and dataset sharing behavior

Transfer characteristics

Data integrity checking, encryption, and reliability, transfer direction, performance, duration and transfer parameters

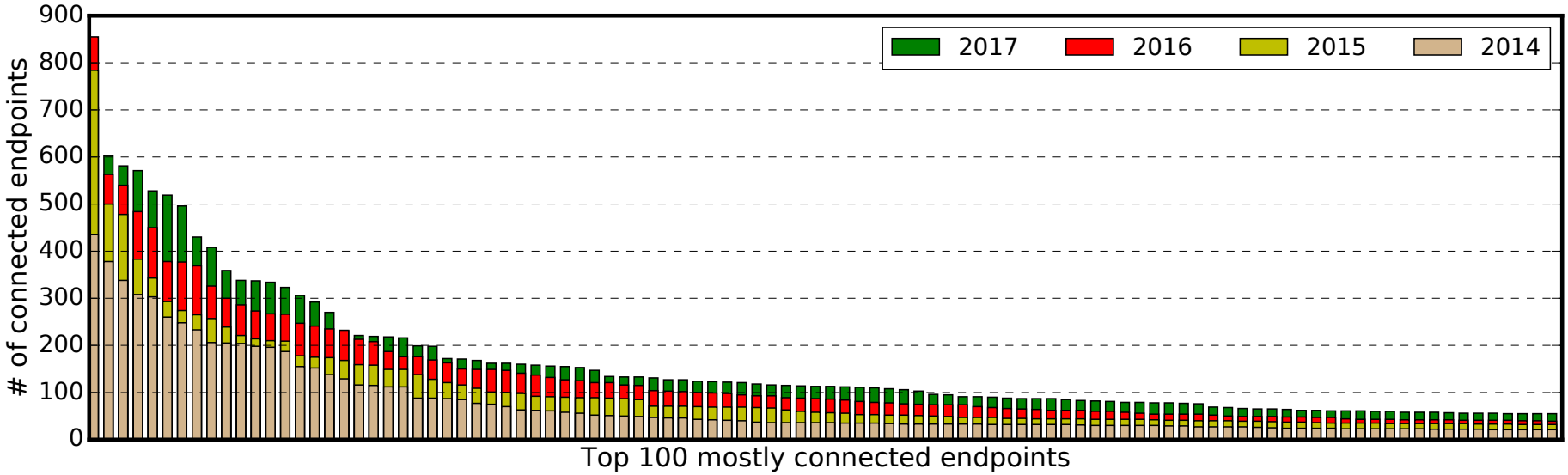
User behavior

transfer frequency, transfer volume, degree of connection to endpoints and pattern of users access endpoint

Endpoint characteristics

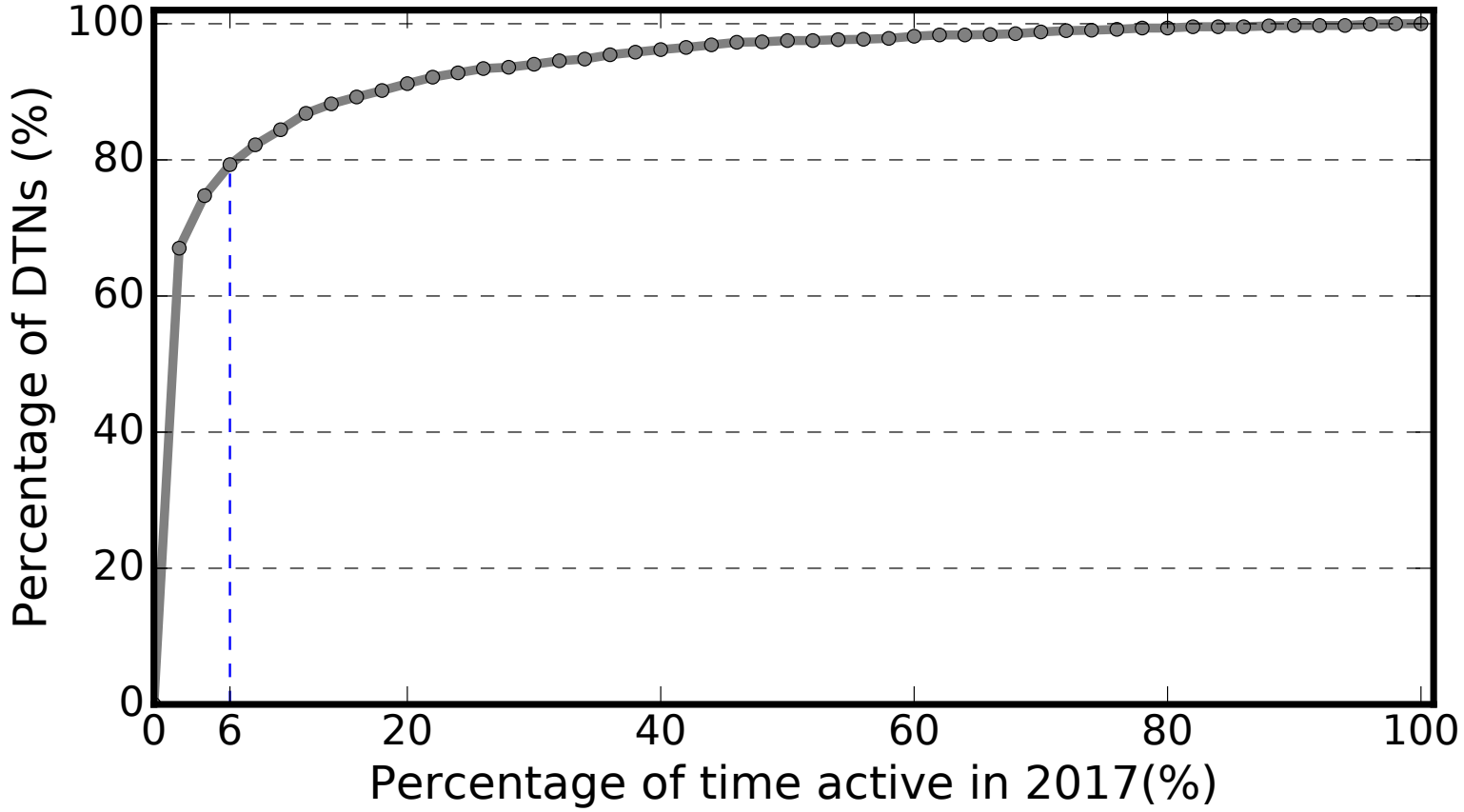
degree of sharing to users, resource utilization (idle time percentage), source-to-destination edge

Endpoint's popularity



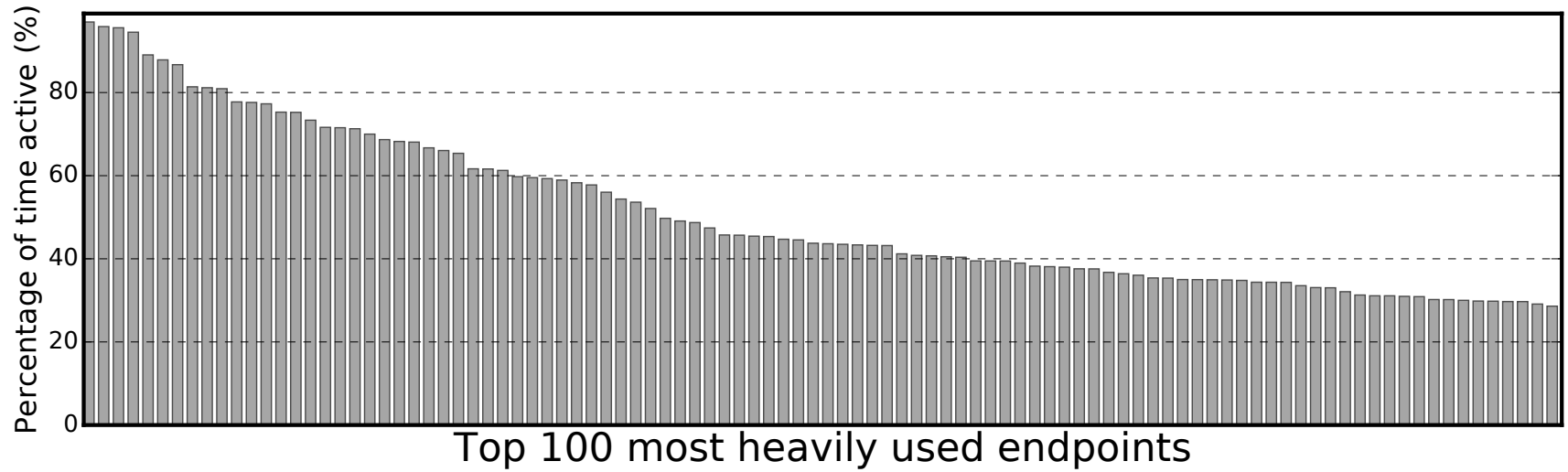
- Number of unique endpoints with which it has engaged in one or more transfers
- In 2017, 81% of the endpoints connected to only one other endpoint, 11% to two other endpoints, and only 8% to more
- Some endpoints are highly connected, and the degree of connection is increasing over time

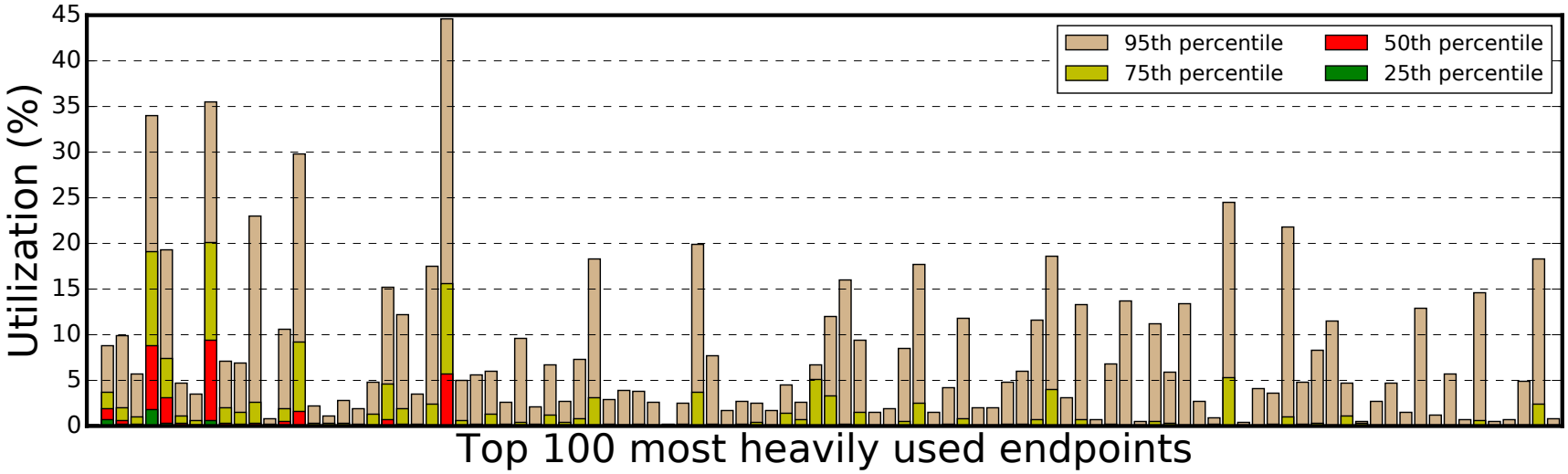
Endpoint usage



- On average, DTNs are completely idle for 94.3% of the time. 80% of the endpoints are active less than 6% of the time.

Endpoint usage





Conclusion and future work

- We analyzed 20 billion GridFTP command logs totaling 1.8 exabytes transferred, by a total of 63,166 GridFTP servers
- We supplemented our analysis with 4.8 million transfers logs collected by the Globus transfer service
- Analysis revealed a number of insights in terms of the utilization of the data transfer nodes, data corruption in wide area transfers, repeat transfers, file types transferred, transfer performance, and user behavior
- We believe our analysis can help researchers, tool developers, resource providers, end users, and funding agencies from different perspectives

THANKS FOR YOUR ATTENTION !

QUESTIONS ?