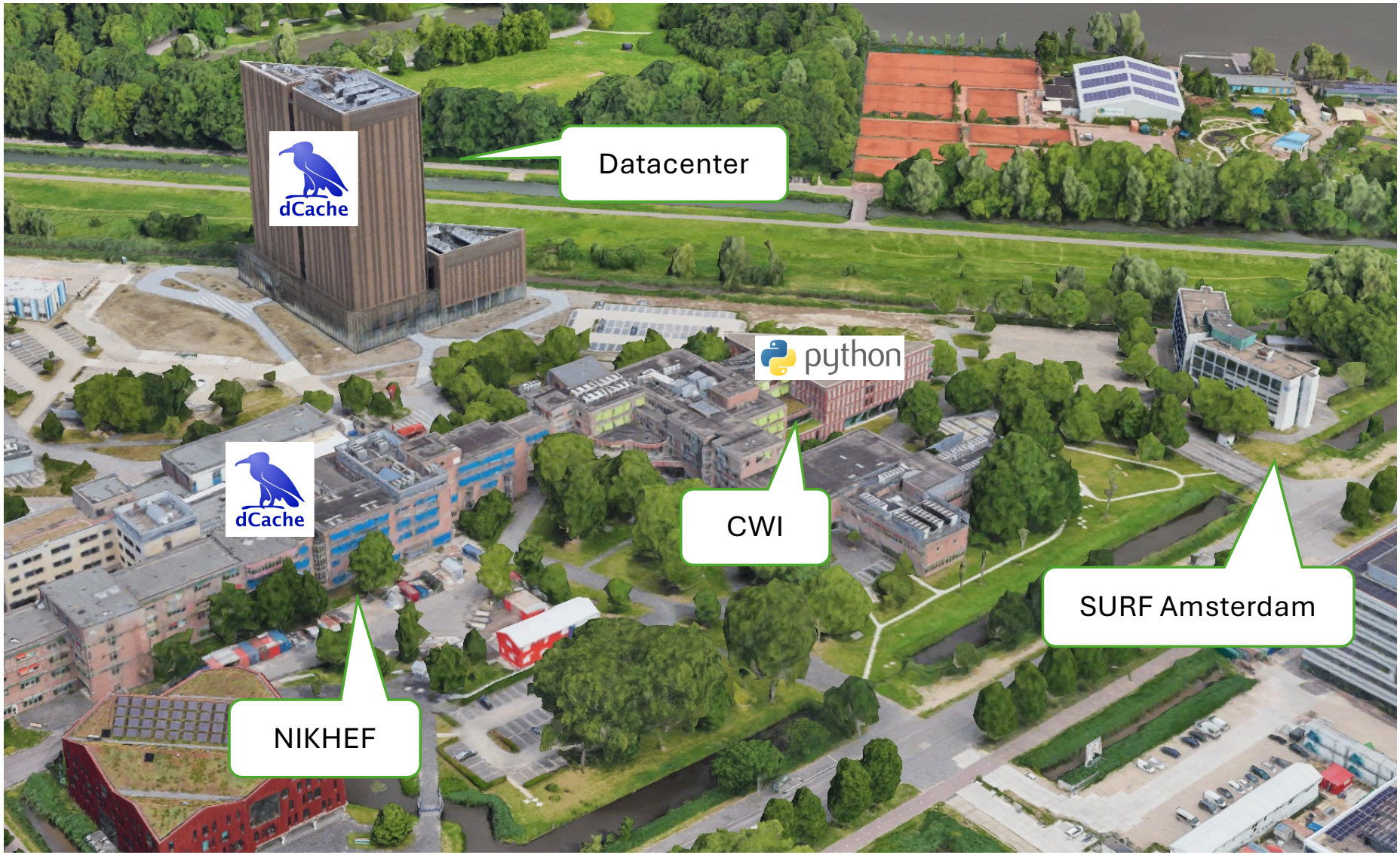
dCache @ SURF / Ada news

Haili Hu, Onno Zweers, Natalie Danezi
dCache workshop, 2026-05-06, v1

In this talk

- dCache at SURF
- Ada news


dCache
Datacenter
python
dCache
CWI
SURF Amsterdam
NIKHEF

dCache @ SURF

- New dCache team member: Mali
- Existing: Merijn & Onno

dCache @ SURF

- 94 poolgroups
- 1255 pools
- 160 nodes
- 95 PB of (unique) data, most of it on tape
- 21 PB in pool space
- 108 million files
- dCache 11.2.3
- Postgres 18 with WAL archiving
- SRM decommissioned

High Availability

- Matthias at last year's workshop (thanks!)
- Failover script
- Failover failures

```
[root@lion1 ~]# dcache-failover --status
List of replicable services and the domains they are currently running in.
L = leader, f = follower

SERVICE                DOMAIN1                          DOMAIN2
bulk                   bulk-dcache-bulk-requestsDomain
info                   lion1-centralDomain              lion2-centralDomain
PinManager             lion1-centralDomain (L)          lion2-centralDomain (f)
PoolManager            lion1-centralDomain              lion2-centralDomain
RemoteTransferManager  lion1-centralDomain              lion2-centralDomain
topo                   lion1-centralDomain              lion2-centralDomain
cleaner-disk           lion1-cleanerDomain (L)          lion2-cleanerDomain (f)
cleaner-hsm            lion1-cleanerDomain (L)          lion2-cleanerDomain (f)
gPlazma                lion1-gPlazmaDomain              lion2-gPlazmaDomain
PnfsManager            lion1-namespaceDomain (L)        lion2-namespaceDomain (f)
SpaceManager           lion1-spacemanagerDomain         lion2-spacemanagerDomain
SrmManager             srmmanagerDomain
billing                utilityDomain
```

IPv6

- IPv6 was unstable
- IPv4 default protocol
- EVPN network scalability issue
- dCache now has its own EVPN network
- Stable
- IPv6 default protocol

NFS problem

- NFS clients hanging
  - NFS door forwards clients to unreachable internal addresses
- Solution: declare the IP address(es) to be used for NFS

```
[root@otter10 ~]# grep pool.mover.nfs.multipath /etc/dcache/dcache.conf
pool.mover.nfs.multipath = 145.100.34.51,2001:610:108:203a::2:51
```

Domain memory usage

```
[root@lion1 ~]# while true ; do dcache-admin-command System@lion1-centralDomain info | grep -o 'Memory.*' | awk -F'[=; ]+' '{printf "%.2f%%\n", ($8/$4)*100}' ; sleep 1s ; done | uniq
85.36%
85.99%
86.60%
42.24%
28.49%
84.10%
84.57%
88.93%
45.19%
47.34%
48.06%
```

If it remains close to 90%, give it more heap

Tape

- Migrated from DMF to Scout (from Versity) in May 2025
- "Backwards" compatible tape format
- Except for "sparse blocks"
  - Some tape files still can't be accessed
  - Fix expected in June
- Scout can't reserve tape drives per project
  - Solution pending

Github repos: Admin space & Scripts

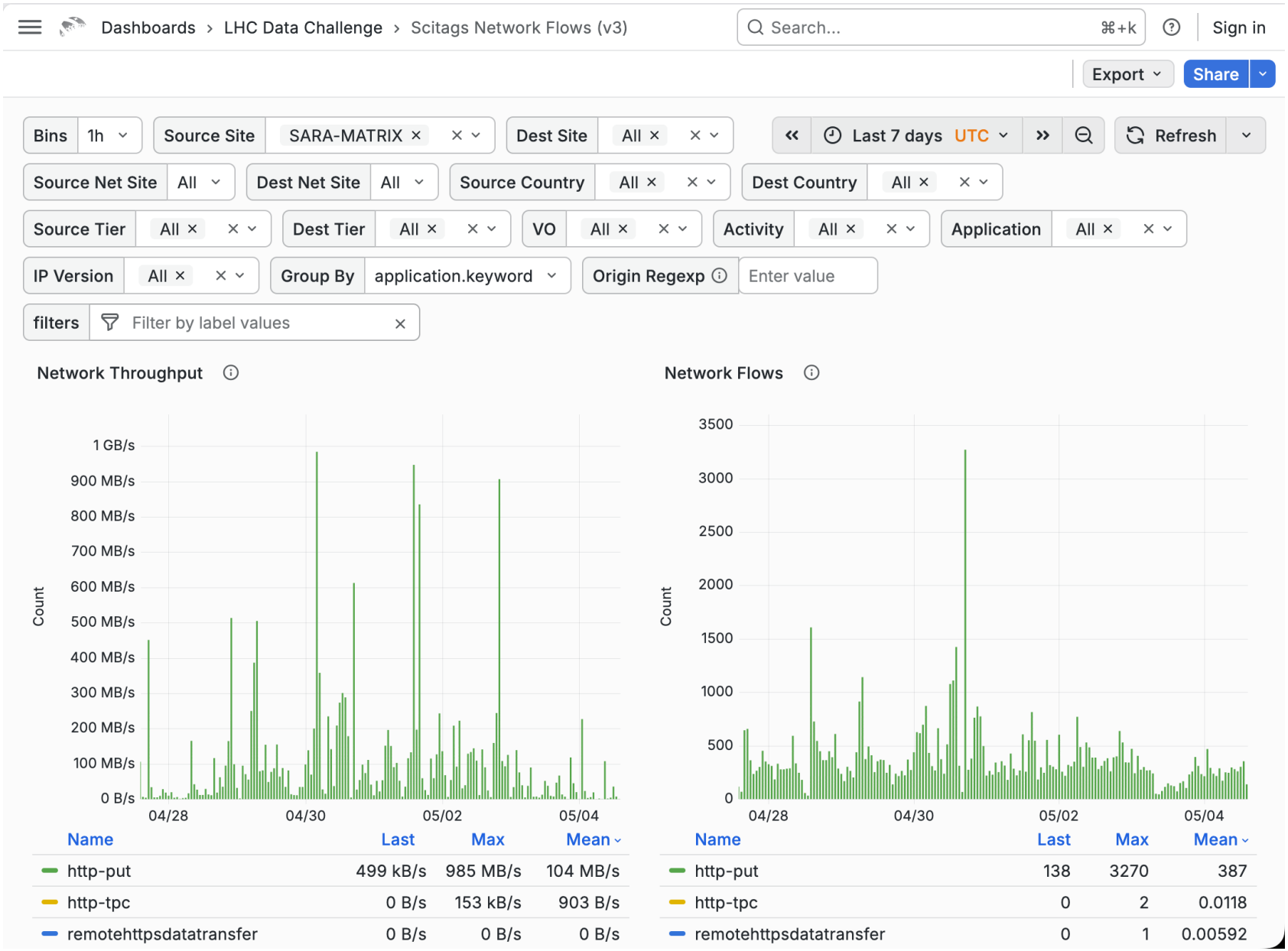
- Admin scripts
  - https://github.com/dCache/scripts
  - chimera_find.sh improved
  - Share your scripts & tools!
- Admin space
  - https://github.com/dCache/admin-space
  - Share your knowledge!
  - Make a Q&A or FAQ?

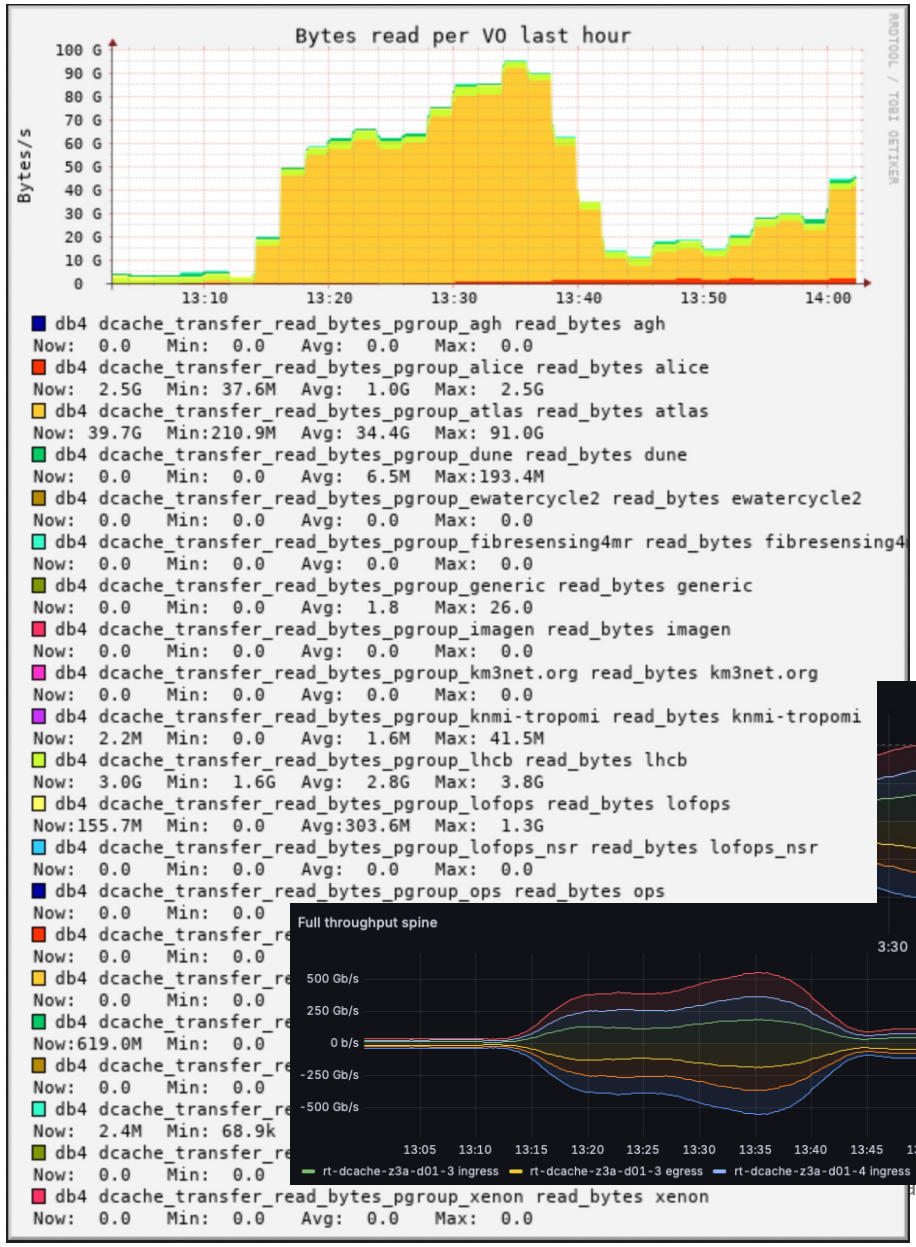
dCache 11.2.3

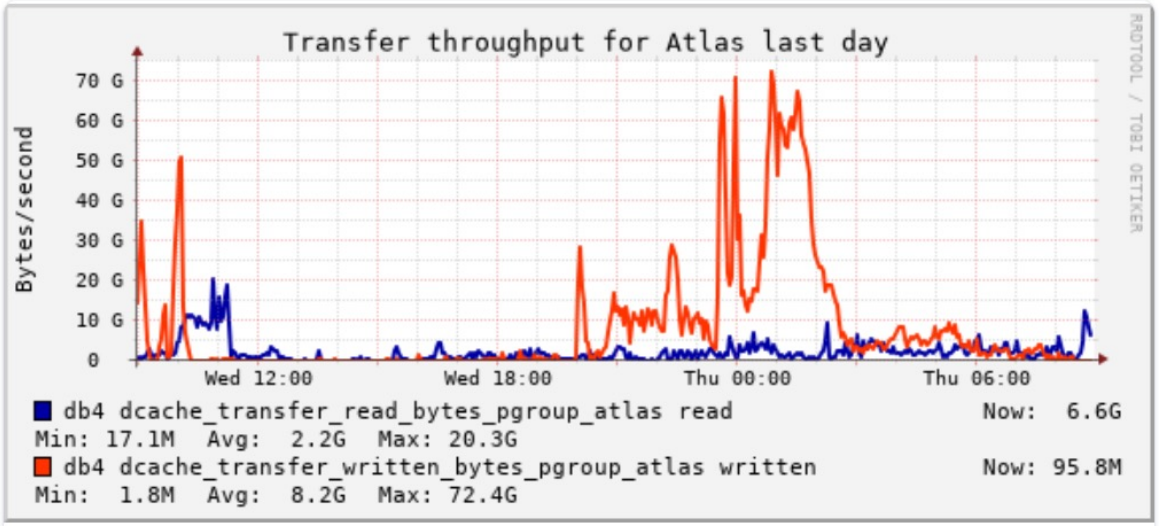
- Mostly running fine
- Proxying WebDAV doors sometimes hang
  - Log: no threads available 1<500<500
  - Only a hand full of transfers listed
  - Still investigating

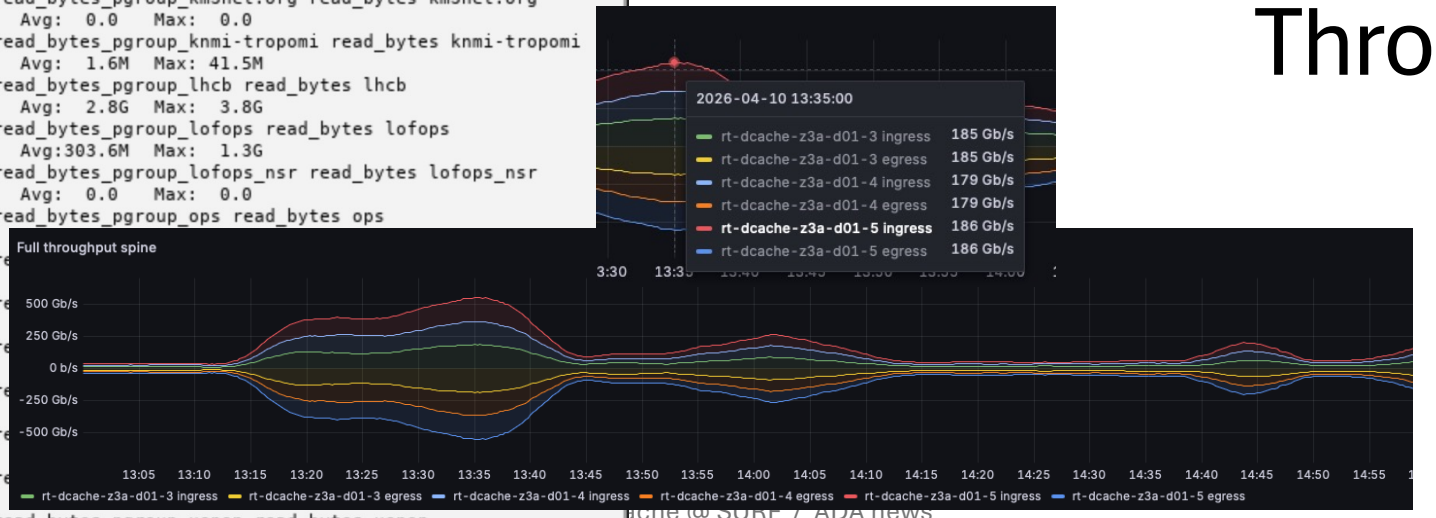
Firefly / SciTags

- Enbled
  - pool.enable.firefly=true
- Monitoring
  - https://dashboard.stardust.es.net/d/b6d7dc1e-dc56-4f39-abe0-a9e7199021f8/scitags-network-flows-v3?var-agg_bin=1h&orgId=2&from=now-7d&to=now&timezone=utc&var-src_site=SARA-MATRIX&var-dst_site=$__all&var-src_net_site=$__all&var-dst_net_site=$__all&var-src_country=$__all&var-dst_country=$__all&var-src_tier=$__all&var-dst_tier=$__all&var-vo=$__all&var-activity=$__all&var-app=$__all&var-ip=$__all&var-group_by=application.keyword&var-origin=&var-Filters=


Dashboards › LHC Data Challenge › Scitags Network Flows (v3)
Search... ⌘+k
Sign in
Export
Share
Bins 1h
Source Site SARA-MATRIX
Dest Site All
Last 7 days UTC
Refresh
Source Net Site All
Dest Net Site All
Source Country All
Dest Country All
Source Tier All
Dest Tier All
VO All
Activity All
Application All
IP Version All
Group By application.keyword
Origin Regexp
Enter value
filters
Filter by label values
Network Throughput
1 GB/s
900 MB/s
800 MB/s
700 MB/s
600 MB/s
500 MB/s
400 MB/s
300 MB/s
200 MB/s
100 MB/s
0 B/s
Count
04/28
04/30
05/02
05/04
Name Last Max Mean
http-put 499 kB/s 985 MB/s 104 MB/s
http-tpc 0 B/s 153 kB/s 903 B/s
remotehttpsdatatransfer 0 B/s 0 B/s 0 B/s
Network Flows
3500
3000
2500
2000
1500
1000
500
0
Count
04/28
04/30
05/02
05/04
Name Last Max Mean
http-put 138 3270 387
http-tpc 0 2 0.0118
remotehttpsdatatransfer 0 1 0.00592


Bytes read per VO last hour
Bytes/s
100 G
90 G
80 G
70 G
60 G
50 G
40 G
30 G
20 G
10 G
0
13:10
13:20
13:30
13:40
13:50
14:00
RRDTOOL / TOBI OETIKER
db4 dcache_transfer_read_bytes_pgroup_agh read_bytes agh
Now: 0.0 Min: 0.0 Avg: 0.0 Max: 0.0
db4 dcache_transfer_read_bytes_pgroup_alice read_bytes alice
Now: 2.5G Min: 37.6M Avg: 1.0G Max: 2.5G
db4 dcache_transfer_read_bytes_pgroup_atlas read_bytes atlas
Now: 39.7G Min:210.9M Avg: 34.4G Max: 91.0G
db4 dcache_transfer_read_bytes_pgroup_dune read_bytes dune
Now: 0.0 Min: 0.0 Avg: 6.5M Max:193.4M
db4 dcache_transfer_read_bytes_pgroup_ewatercycle2 read_bytes ewatercycle2
Now: 0.0 Min: 0.0 Avg: 0.0 Max: 0.0
db4 dcache_transfer_read_bytes_pgroup_fibresensing4mr read_bytes fibresensing4
Now: 0.0 Min: 0.0 Avg: 0.0 Max: 0.0
db4 dcache_transfer_read_bytes_pgroup_generic read_bytes generic
Now: 0.0 Min: 0.0 Avg: 1.8 Max: 26.0
db4 dcache_transfer_read_bytes_pgroup_imagen read_bytes imagen
Now: 0.0 Min: 0.0 Avg: 0.0 Max: 0.0
db4 dcache_transfer_read_bytes_pgroup_km3net.org read_bytes km3net.org
Now: 0.0 Min: 0.0 Avg: 0.0 Max: 0.0
db4 dcache_transfer_read_bytes_pgroup_knmi-tropomi read_bytes knmi-tropomi
Now: 2.2M Min: 0.0 Avg: 1.6M Max: 41.5M
db4 dcache_transfer_read_bytes_pgroup_lhcb read_bytes lhcb
Now: 3.0G Min: 1.6G Avg: 2.8G Max: 3.8G
db4 dcache_transfer_read_bytes_pgroup_lofops read_bytes lofops
Now:155.7M Min: 0.0 Avg:303.6M Max: 1.3G
db4 dcache_transfer_read_bytes_pgroup_lofops_nsr read_bytes lofops_nsr
Now: 0.0 Min: 0.0 Avg: 0.0 Max: 0.0
db4 dcache_transfer_read_bytes_pgroup_ops read_bytes ops
Now: 0.0 Min: 0.0
Now: 0.0 Min: 0.0
Now: 0.0 Min: 0.0
Now:619.0M Min: 0.0
Now: 0.0 Min: 0.0
Now: 2.4M Min: 68.9k
Now: 0.0 Min: 0.0
db4 dcache_transfer_read_bytes_pgroup_xenon read_bytes xenon
Now: 0.0 Min: 0.0 Avg: 0.0 Max: 0.0


Transfer throughput for Atlas last day
Bytes/second
70 G
60 G
50 G
40 G
30 G
20 G
10 G
0
Wed 12:00
Wed 18:00
Thu 00:00
Thu 06:00
RRDTOOL / TOBI OETIKER
db4 dcache_transfer_read_bytes_pgroup_atlas read Now: 6.6G
Min: 17.1M Avg: 2.2G Max: 20.3G
db4 dcache_transfer_written_bytes_pgroup_atlas written Now: 95.8M
Min: 1.8M Avg: 8.2G Max: 72.4G


2026-04-10 13:35:00
rt-dcache-z3a-d01-3 ingress 185 Gb/s
rt-dcache-z3a-d01-3 egress 185 Gb/s
rt-dcache-z3a-d01-4 ingress 179 Gb/s
rt-dcache-z3a-d01-4 egress 179 Gb/s
rt-dcache-z3a-d01-5 ingress 186 Gb/s
rt-dcache-z3a-d01-5 egress 186 Gb/s
Full throughput spine
500 Gb/s
250 Gb/s
0 b/s
-250 Gb/s
-500 Gb/s
13:05
13:10
13:15
13:20
13:25
13:30
13:35
13:40
13:45
13:50
13:55
14:00
14:05
14:10
14:15
14:20
14:25
14:30
14:35
14:40
14:45
14:50
14:55
rt-dcache-z3a-d01-3 ingress
rt-dcache-z3a-d01-3 egress
rt-dcache-z3a-d01-4 ingress
rt-dcache-z3a-d01-4 egress
rt-dcache-z3a-d01-5 ingress
rt-dcache-z3a-d01-5 egress

Throughput

Wish list

Every dCache workshop, I try to come up with 3 things I would like in dCache.

- Have PoolManager look at the system load of pool nodes
  - Don't select pools on nodes with high load
- Get rid of space tokens

Space Reservations 🙀

- We (still) have them for Atlas and LHCb
- They are a PAIN
  - Extra layer of bookkeeping
  - Errors very difficult to correct
- They are unnecessary
  - Poolgroup info would be more accurate
- Thinking of ways to replace space token info with poolgroup info.
  - One problem though: during pool migrations poolgroups may seem larger, but that extra space is not for users but for the migration.
  - Solution 1: exclude pools in "draining" mode from available space?
  - Solution 2: set hard limit on the available poolgroup space?

Roadmap

- Redesign interface with tape
  - Bottleneck: polling over SSH
- More Highly Availability
  - HS Proxy
- DTAP (development, test, acceptance, production)
  - Set up in Ansible
- Reorganize doors
- Test with Ceph as underlying storage
- Connect dCache with SRAM authentication platform (OIDC based)
- QoS data management

Ada news

Advanced dCache API client

Haili Hu, Onno Zweers, Natalie Danezi
dCache workshop, 2026-05-06/07

About Ada (recap)

- Command line client for dCache API
- Written in Bash
- User friendly experience
- Authentication: X509, username/password, tokens (OIDC & macaroons)
- Configurable (ada.conf)
- Unit & integration tests
- Open source https://github.com/sara-nl/SpiderScripts/

What's new

- Upload & Download
  - Prevent redirects from HTTPS to HTTP
  - Clear error messages
  - 3 ways to discover WebDAV doors
  - Upload:
    - MD5 checksum verification
  - Download:
    - Locality check:
      - Fail if file is NEARLINE
      - Explain how to stage
- But the biggest change…

Bash → Python

- Bash has limits
  - Max command line length
- Users prefer Python
  - Especially scientists
  - They want a library to include in their workflow
  - Many community attempts to write own Python API client

Reached out to our users

- Netherlands e-Science Center NLeSc (RS-DAT/dCacheFS developers)
- Astron/LOFAR (team Rainbow, 7 developers)

Explored other Python API clients

- dCacheFS Python file-system interface for dCache (builds on fsspec) by NLeSc
- "Minimal interface to dCache based on PycURL" developed by SRON/Jigsaw for Tango (ESA Scout satellite mission)
- LOFAR per-tool integration developed by Astron based on various clients/protocols: SRM (deprecated), GSIFTP, gfal, dCACHE REST API, WebDAV, FTS
- "Client library for the DCACHE api" by ASTRON/Mattia Mancini for EOSC (last commit 5 years ago)
- gfal2-python: a Python wrapper of the gfal2 C++ library developed by CERN for file operations in grids and cloud environments
- dcacheclient by Vincent Garonne (last commit 6 years ago)
- Python wrapper around ADA and rclone: https://github.com/holstegelab/dcache_cp for copying (inc checksums and staging), moving and listing files on dcache.

Decided to start from Bash ADA, using AI

- Tried several AI systems, to convert Bash to Python
- Ollama with various LLM models
  - Slow without GPU+VRAM
  - Context window too small (Bash Ada is 3500 lines, ~35k tokens)
  - Hallucinating models
  - Needed coding agent
- SURF AI service
  - Context window too small
- Devstral & vibe
  - Converted a few functions into Python, seemed to work, but free tier finished after just one query
- Claude Code
  - A colleague had a subscription and a lot of experience
  - In 20 minutes, he made us a good code foundation to build upon

Speed test: directory crawl

HTTP lib	Dir entries per second
aiohttp	251
aiohttp async 10	2875
http.client	126
http.client reuse	252
httpx	250
httpx async 10	3369
pycurl	147
pycurl reuse conn	201
requests	224

Reusing the connection: significantly faster

Asynchronous (parallel) : amazing performance increase

WINNER: httpx (performance & coding elegance)

Phase 1: Minimum Viable Product

- Importable as Python library
- Executable as command line tool
- Authentication via tokenfile
- Read configuration from config file
- Error handling
- Docstrings
- Unit tests and integration test (with pytest)

Include commands:

- help
- whoami
- list
- longlist
- mkdir
- mv
- delete
- checksum
- stage
- unstage

Phase 2 (next steps)

- Add authentication methods: netrc, proxy
- Override configuration with environmental variables
- Override override configuration with commandline arguments

- version
- viewtoken
- gettoken
- upload / download
- stat
- setlabel / rmlabel / lslabel / findlabel
- setxattr / rmxattr / lsxattr / findxattr
- stat-request
- delete-request
- events / report-staged / channels / delete-channels
- space
- quota

Phase 3: nice to have

- Get metadata from sites (whic doors are avialable, which urls?)
- Get all types of url (webdav, gridftp, etc) for a file
- Checksum validation on download data stream without having to write and read files
- Any suggestions?

Work in progress

https://github.com/sara-nl/dcache-pyclient

https://pypi.org/project/dcache-pyclient/

You're welcome to test and leave issues on the Github repo!

Thanks!

Questions?
Suggestions?

Haili and Onno at math center CWI in Amsterdam (where Python was created) (photo Ron Trompert)