

# Storage

Wahid Bhimji

DPM Collaboration :

- Tasks.

Xrootd:

- Status;
- Using for Tier2 reading from “Tier3”;
- Server data mining

# DPM

- DPM collaboration is alive – had our first joint dev mtg.
- GridPP have some tasks on a task list. This is mainly for me and Sam but others are welcome to contribute (and we may need it...)
  - Support (beyond local) - e.g. a rota for dpm-user-forum
  - Admin-tools ; DMLITE/python DMLITE/shell (mainly Sam)
  - DMLite/Profiler / logging (“interest” – again Sam mainly)
  - Documentation and web presence
  - Configuration / puppet :

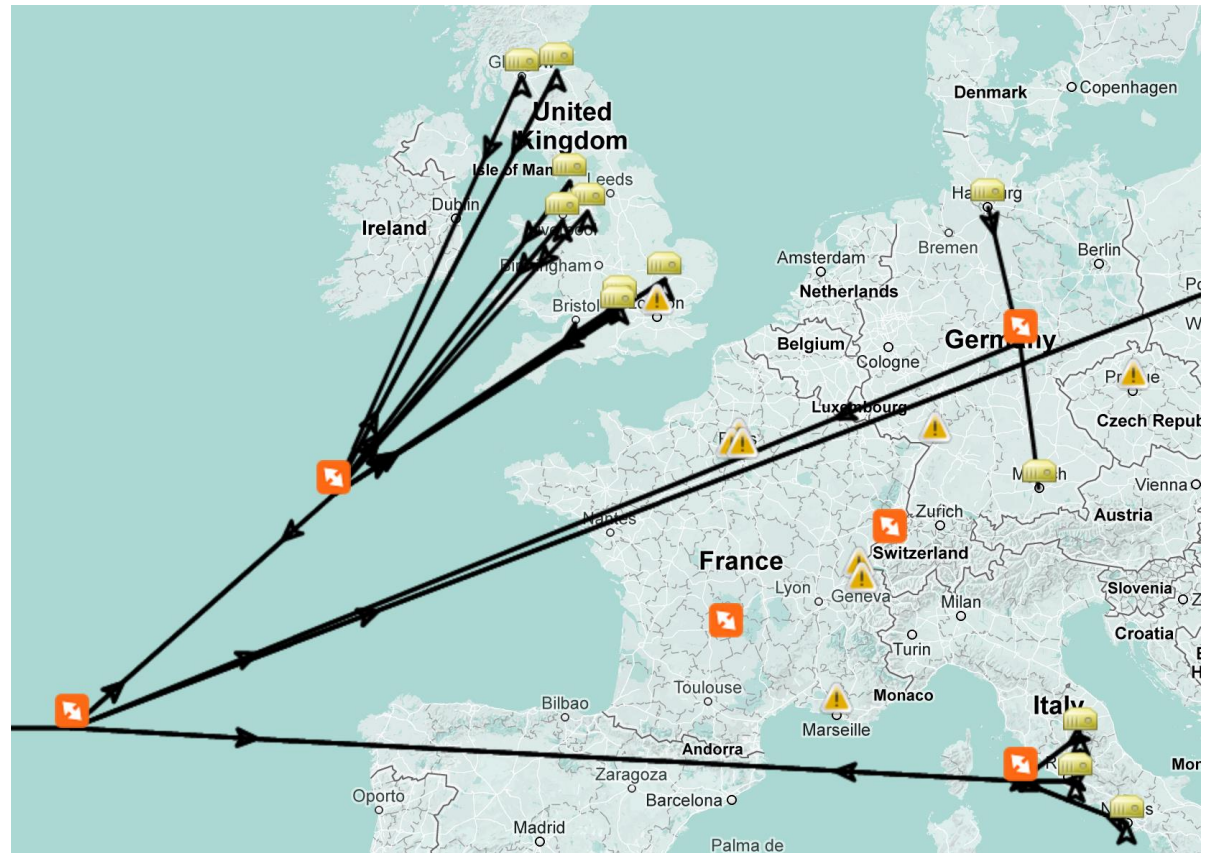
<https://svnweb.cern.ch/trac/lcgdm/wiki/Dpm/Admin/Puppet>

Possibly best to wait for release – real soon now – but then input from sites who actually use puppet would be valuable (only test setup at Ed.)

- Permanent pilot testing (other sites welcome if they have test machines)
- HTTP performance / S3 validation (I have run some tests -need to do more).

# Xrootd – Atlas Federation (FAX)

- Lots of UK sites – I encourage the last few to join.
- Still in a testing phase for ATLAS
- But just enabled fallback to FAX (if no file found on 3<sup>rd</sup> attempt) at ECDF – still no way to test properly



# Xrootd for local grid job access on DPM sites

- Glasgow, Oxford, Edinburgh been using xrscp instead of rfcop for Atlas jobs for some time now
  - **No problems** – I encourage other sites to switch
- Glasgow and Oxford have been using direct xrd access (not copy) for atlas analysis jobs for a few weeks
  - **No problems** – I think we should try at more sites.
  - For ECDF changing that exposed an issue in tarball
    - As it is a python26 one – hoping it goes away in SL6.

# Xrootd for local local access from “Tier 3”

- Being used at (at least) Edinburgh and Liverpool
- Works well -I recommend it – just with Root installed can do

TFile::Open(“root://srm.glite.ecdf.ed.ac.uk//dpm...”)

- They must use [TTreeCache](#) in their code for decent performance.
- To get a list of files for ATLAS for example (maybe easier way but this works...)

```
dq2-ls -f -p -L UKI-SCOTGRID-ECDF_LOCALGROUPDISK  
mc12_8TeV.167892.PowhegPythia8_AU2CT10_ggH125_ZZ4lep_noTau.merge.NTUP_HSG2.e1622_s1581_s1586_r3658_  
r3549_p1344/ | grep srm | sed 's/srm:\/\srm.glite.ecdf.ed.ac.uk/root:\/\srm.glite.ecdf.ed.ac.uk\/g'
```

- Can also get global names (which would use FAX so get the files remotely if not available locally – not sure I would expose users to that yet...)

```
export STORAGEPREFIX=root://YOURLOCALSRM://; dq2-list-files -p DATASET
```

# Xrootd access of T2 from T3 continued...

- Can provide a “POSIX” interface with gFal2 [gFalFS](#):

```
yum install gfal2-all gfalFS gfal2-plugin-xrootd
```

Xrootd plugin from epel-testing currently (though maybe in epel since release yesterday...)

```
usermod -a -G fuse <username>
```

```
gfalFS ~/mount_tmp root://srm.glite.ecdf.ed.ac.uk/dpm/ecdf.ed.ac.uk/home/atlas
```

Can ls etc. Test of reading files for single jobs. (100% of 7929 events from 533 MB file )

	Fuse	Xrootd directly
TTreeCache on (ReadCalls = 17)	39.47, 38.74	31.54, 33.24
TTreeCache off (ReadCalls = 115445)	16.11, 15.19	34.03, 35.57

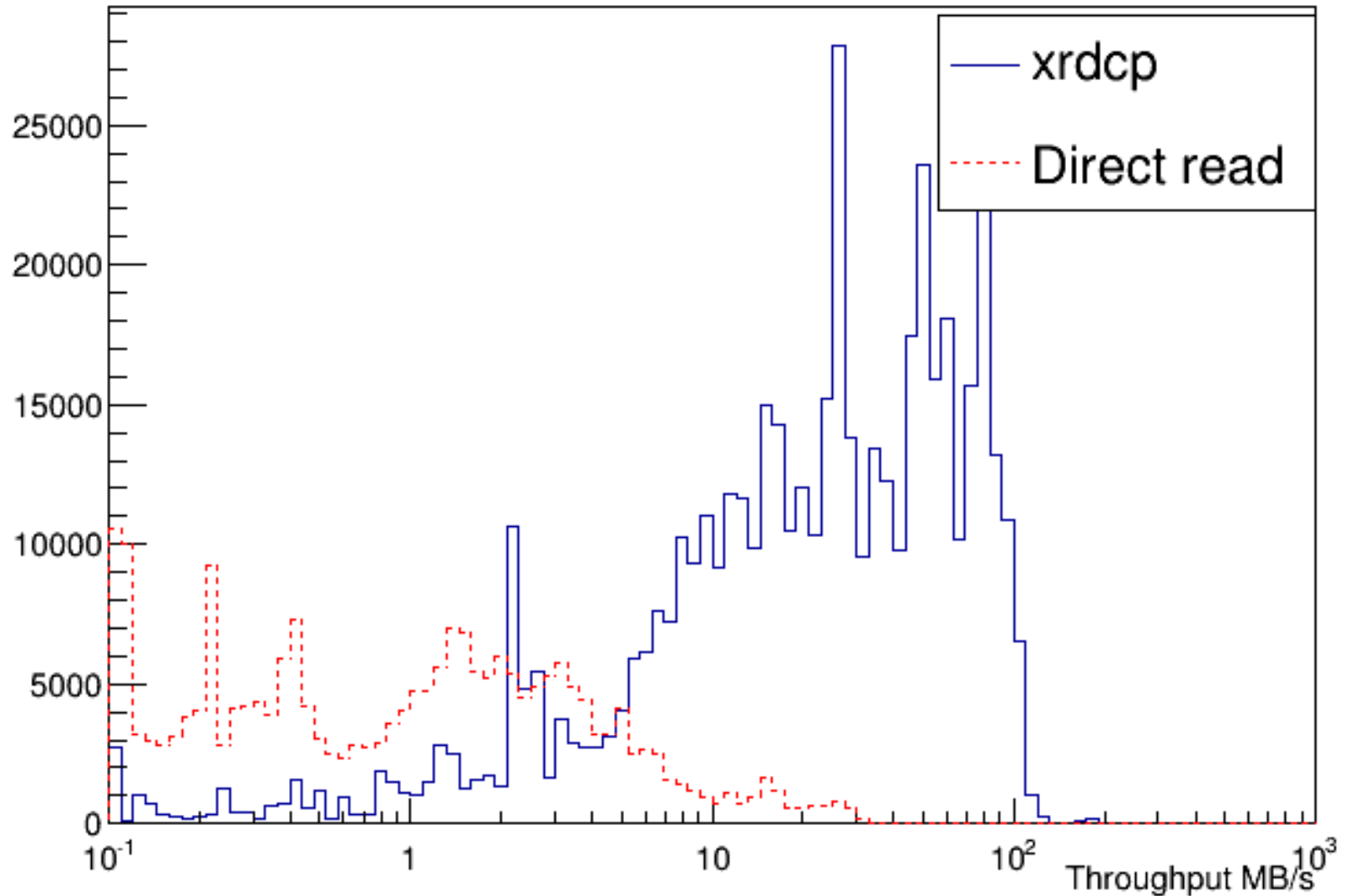
John and I have fed back a few bugs – but it basically works.

# Xrootd server data mining...

- **Source:** see, for example, this [GDB talk](#). Comes from UDP stream from every xrootd server at every site.
- [Monitoring ATLAS FAX traffic](#), and also local traffic if xrootd is used (e.g. ECDF, Glasgow and CERN)
- Lots of plots obtainable from [dashboard](#)
- Here we do some [analysis on the raw data](#)
- Tests are a significant part of current FAX data
  - So “spikes” can be these rather than user activity.
  - Can provide an extra angle on tests .
- Non-EOS non-Fax traffic has a lot of xrdcp – though some DPM sites recently moving to direct access for Analy queues

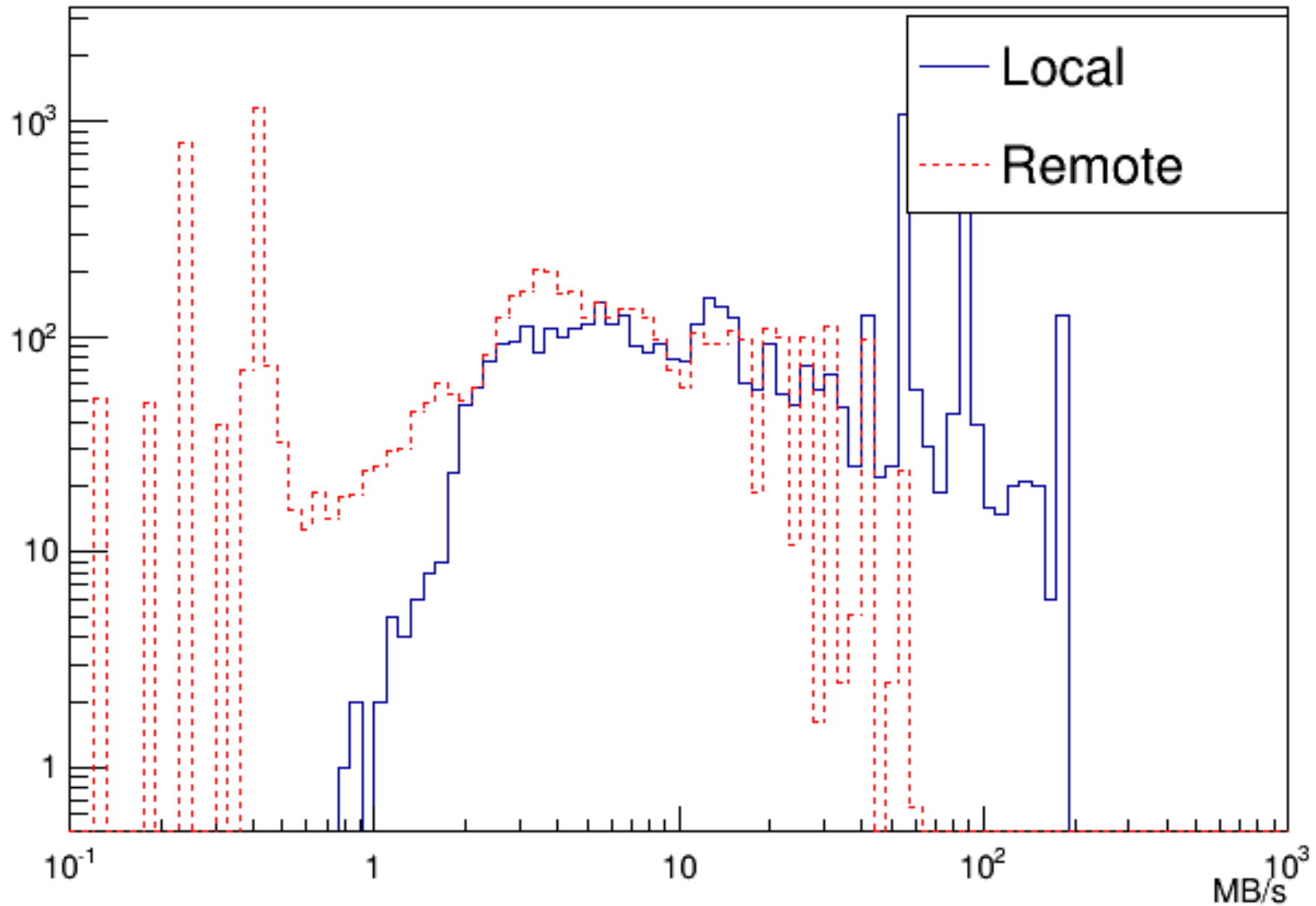
# Throughput : Edinburgh (DPM) : (Direct and copy traffic)

“xrdcp” are those where read exactly equals filesize – probably copies

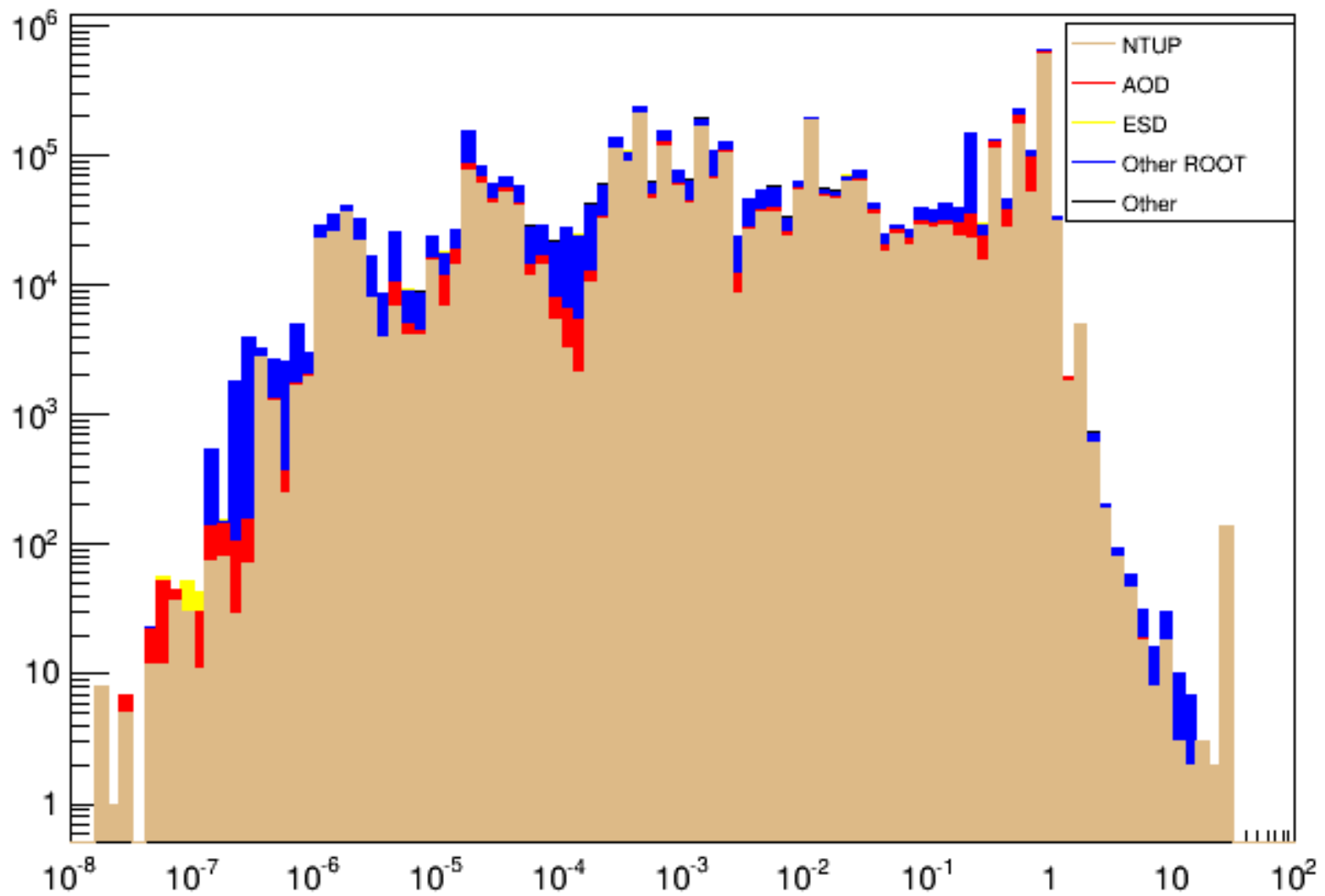




# Throughput: QMUL (Lustre and Storm): Local and Remote (xrdcp)

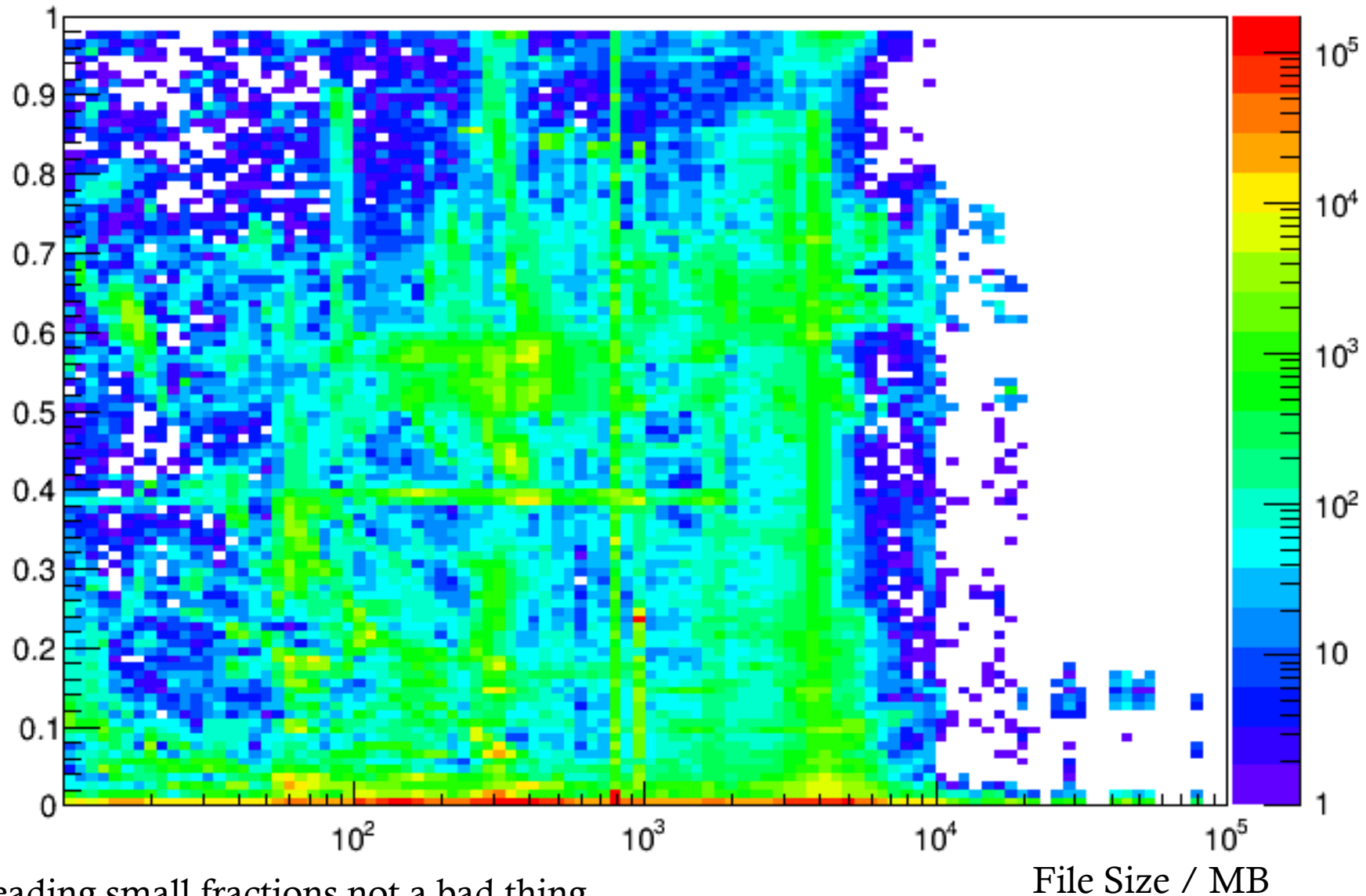


# Fraction of file read (non xrdcp) All sites except CERN



ReadRatio < 1

Read Ratio Vs Size

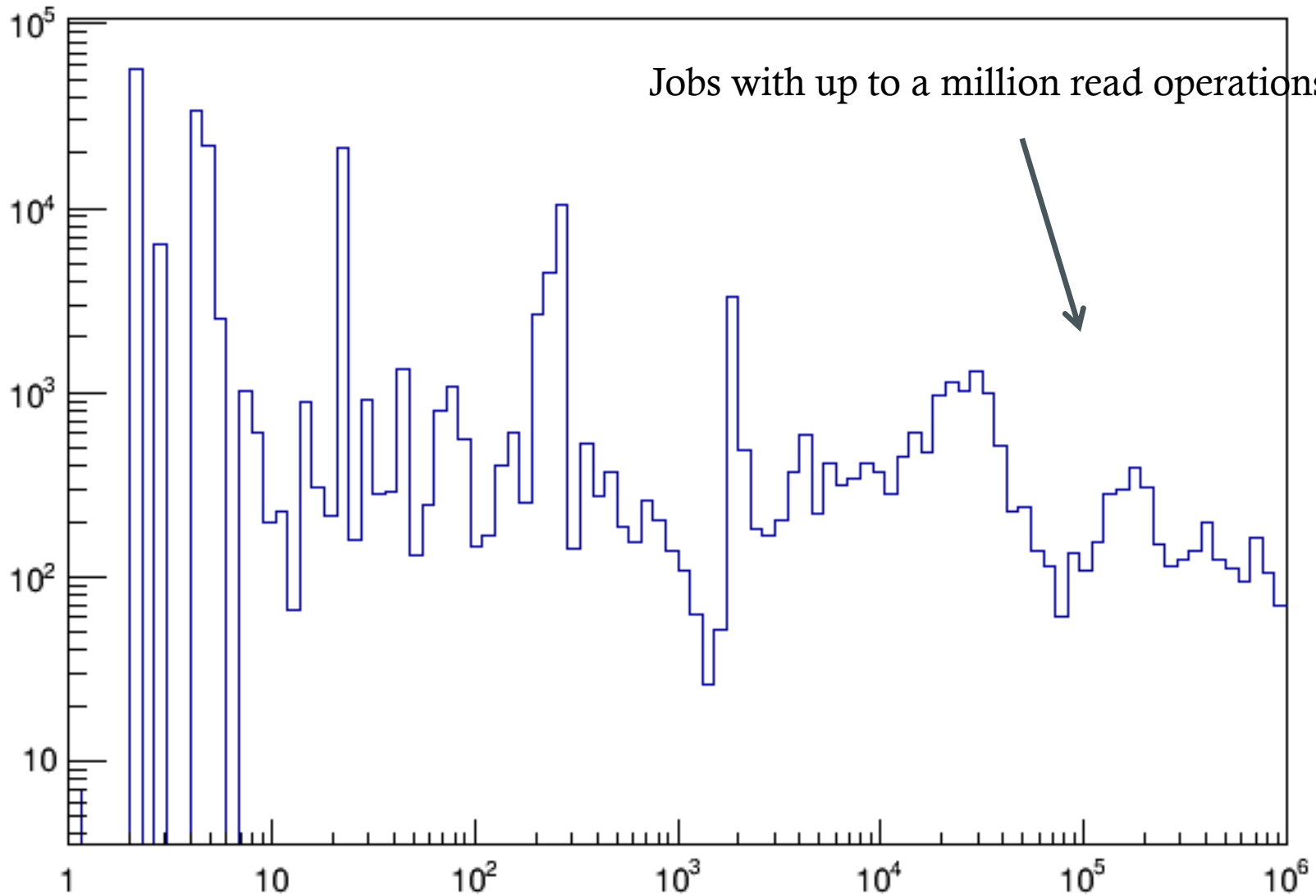


Reading small fractions not a bad thing.

Reading 10% events with TTreeCache (e.g. Ilija's tests) gives a read ratio near 100%:

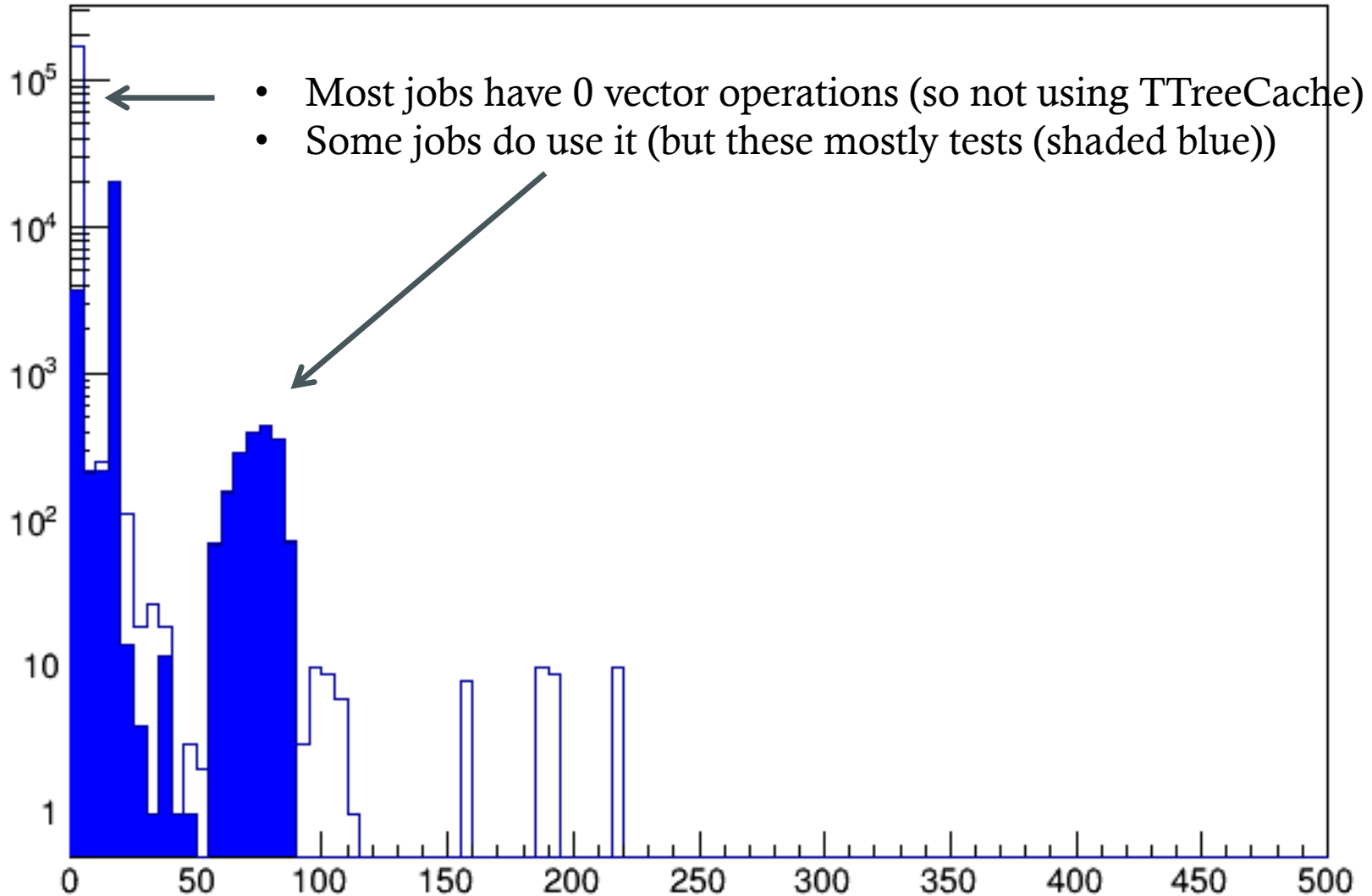
# ECDF

## Read Operations



# Vector Operations

ECDF



# Conclusions

- Performance data from xrootd is useful – we can use this to probe sites or application.
- Many sites currently configured for local access rather than remote.
- Throughput numbers appear low (may not be a problem). Current jobs are doing OK.
- Large no. of jobs with low read ratios suggest ATLAS doing more direct access would conserve bandwidth
- Jobs which read files multiple times, have high no. of read operations (and low bandwidth) are concerning:
  - If ATLAS can force bulk of traffic to well behaved applications: (e.g central service or app or wrapper ) then we will be better off.
- Encourage DPM sites to switch to xrootd for local access – copy for production and direct access for analysis ... and to join atlas FAX.
- Volunteers for DPM task list items are welcome.