WHO ATE MY IOPS? IDENTIFYING HEAVY HITTERS USING JOBSTATS

James Beal, Principal Systems Administrator
Dave Holland, Principal Systems Administrator
V07 2022/11/30

AGENDA

INTRODUCTION

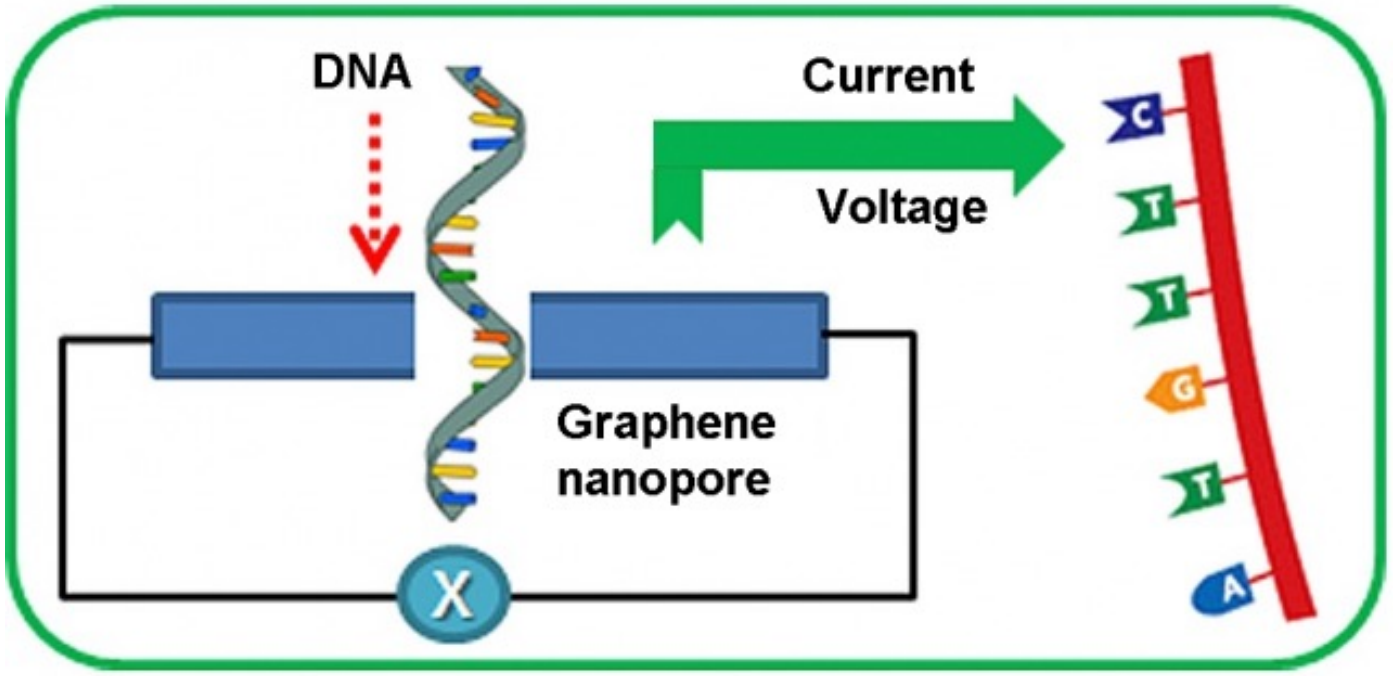
IMPLEMENTATION

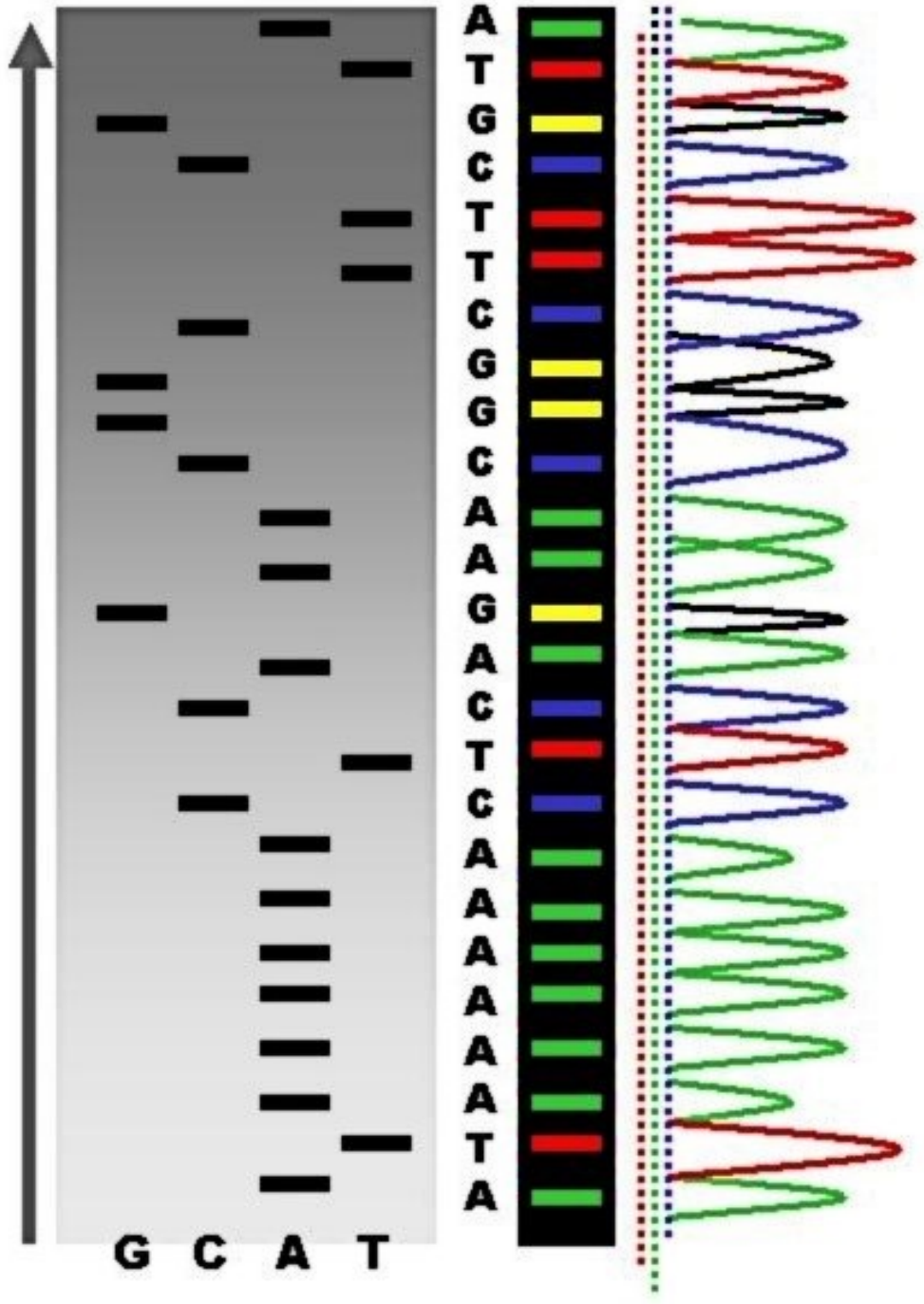
EXAMPLES

CONCLUSIONS


INTRODUCTION

DNA SEQUENCING





THE PAST

Capillary sequencer.

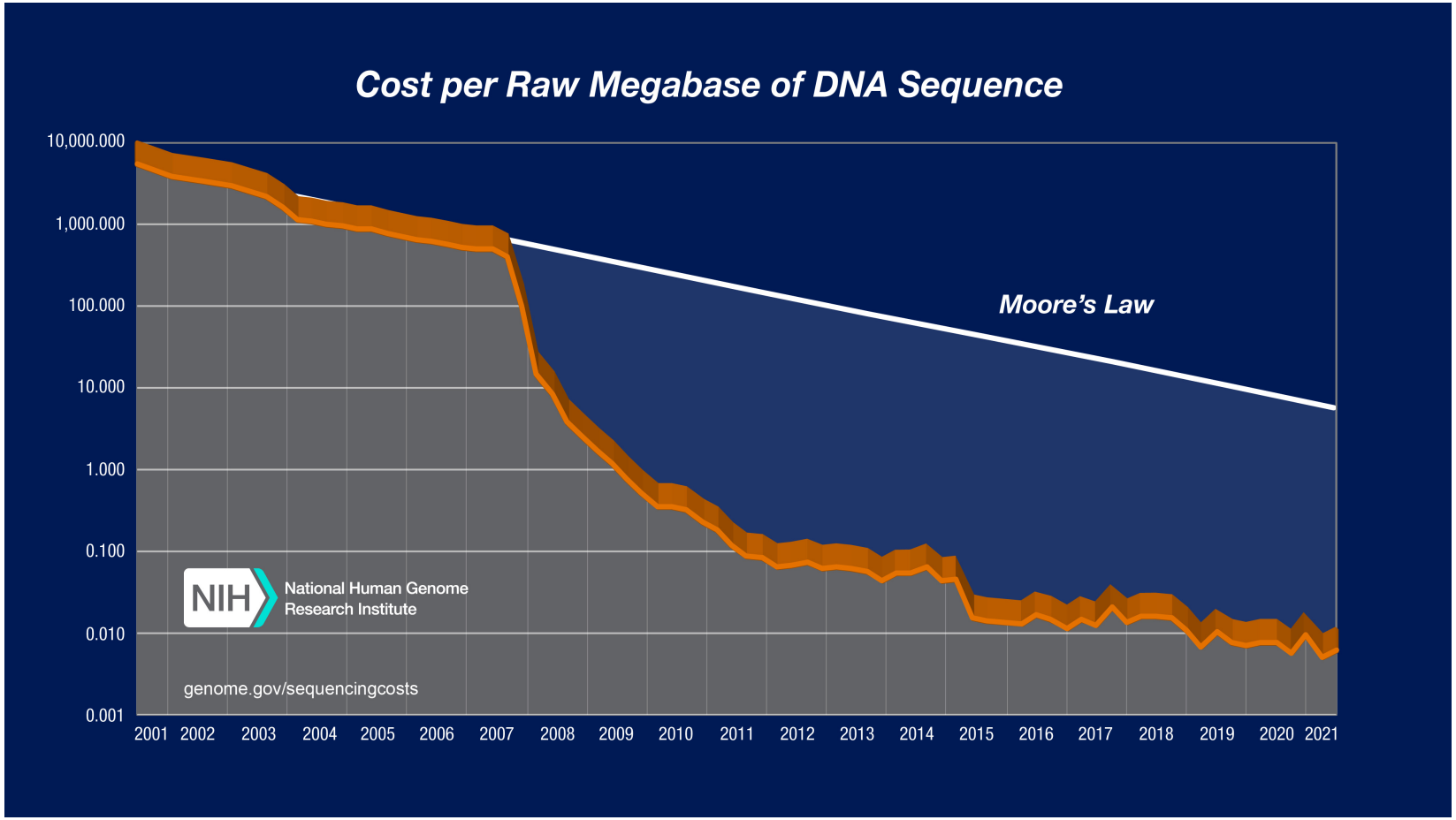
- 2 Mega ($2*10^6$) bases a day.
- Read length of about 1000 base pairs

THE PRESENT

SBS (sequencing by synthesis, Novaseq X).

- 8 Terabases ($8*10^{12}$) a day per sequencer.
- 2.5 bytes a base or 20TB per sequencer per day working space.
- 150 base pair read length
- We have 20 of the previous generation currently in production.

THE FUTURE

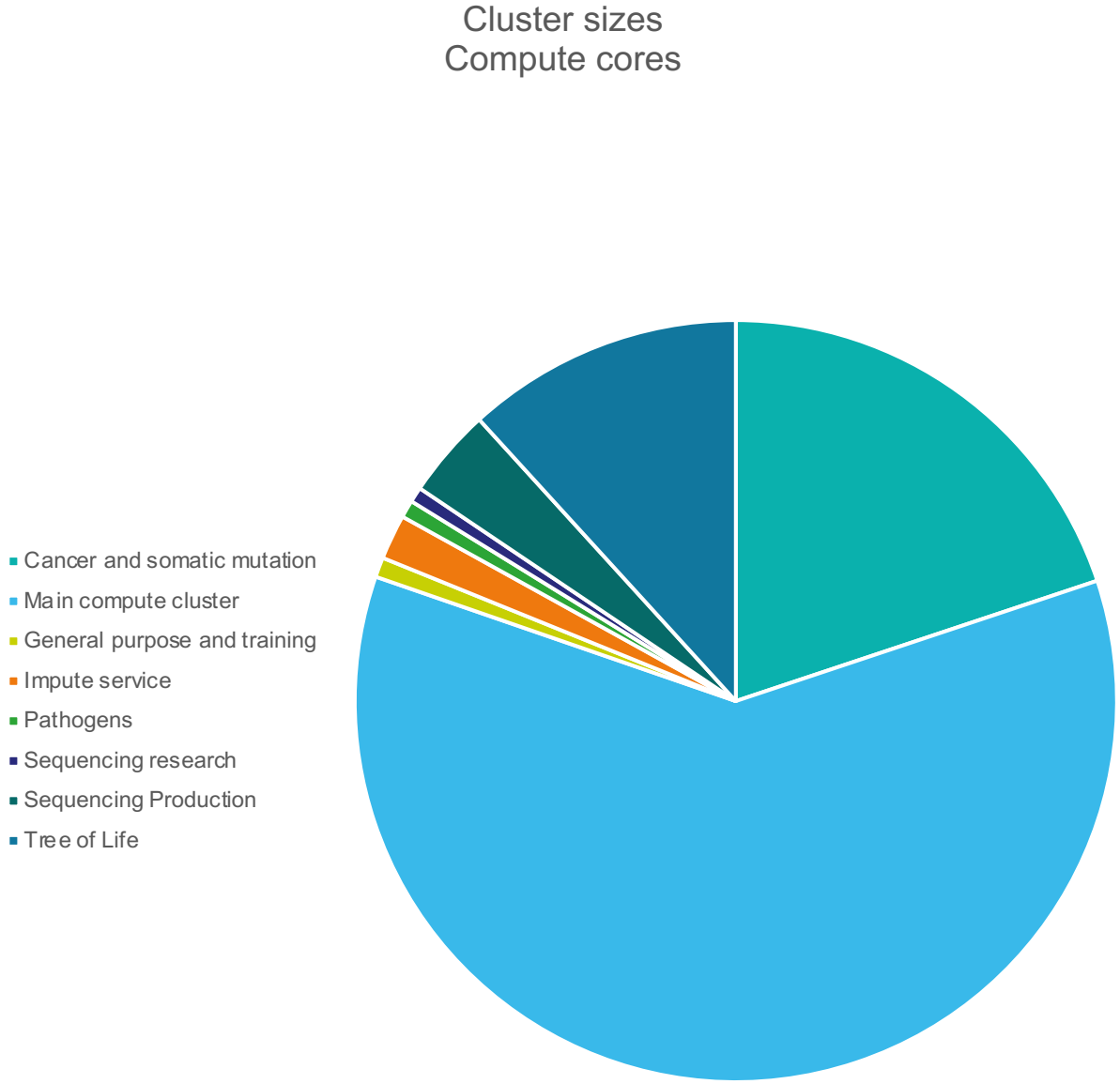


https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data

COMPUTE CLUSTERS

- A number of very similar configuration LSF clusters of different sizes. Separate clusters are used to ringfence resources.
- All* lustre filesystems are mounted everywhere.
- Lustre filesystems are not generally dedicated to a single group.

```
jb23@seqfarm3-head1:~$ df -H /lustre/*
Filesystem                                  Size  Used Avail Use% Mounted on
172.17.220.18@tcp:172.17.220.19@tcp:/lus17  3.1P  2.2P  886T  71% /lustre/scratch117
172.17.220.28@tcp:172.17.220.29@tcp:/lus18  3.1P  1.8P  1.3P  59% /lustre/scratch118
172.28.64.102@tcp:172.28.64.103@tcp:/lus19  2.5P  2.2P  254T  90% /lustre/scratch119
10.177.252.5@tcp:10.177.252.4@tcp:/lus20    3.1P  1.2P  1.9P  38% /lustre/scratch120
10.160.32.5@tcp:10.160.32.4@tcp:/lus23      5.3P  2.8P  2.5P  53% /lustre/scratch123
10.160.36.4@tcp:10.160.36.5@tcp:/lus24      5.3P  1.8P  3.5P  34% /lustre/scratch124
10.160.40.4@tcp:10.160.40.5@tcp:/lus25      5.3P  229T  5.1P   5% /lustre/scratch125
10.160.42.4@tcp:10.160.42.5@tcp:/lus26      5.3P  327T  5.0P   7% /lustre/scratch126
```



JOBSTATS

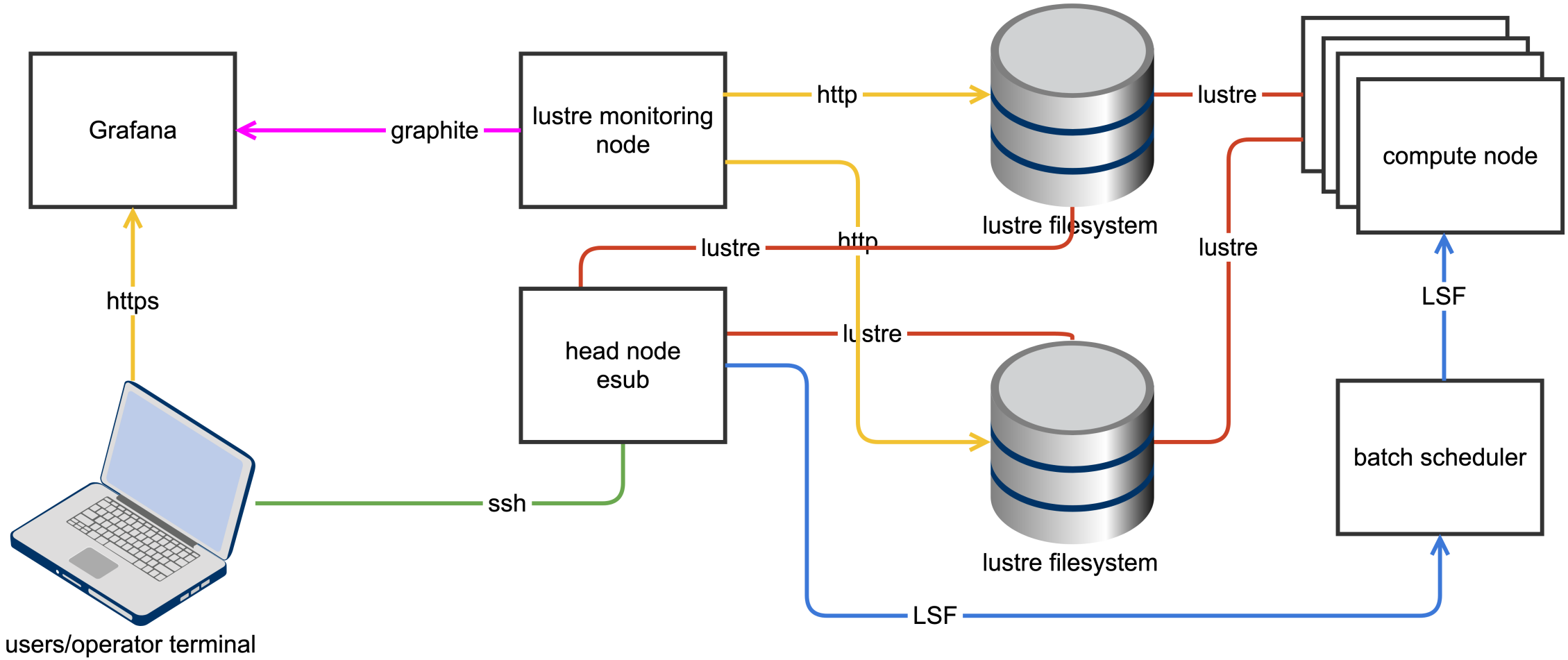
- https://doc.lustre.org/lustre_manual.xhtml#jobstats
- A method of tagging I/O with additional metadata.
- Introduced in Lustre 2.3 2011
- A 32 character tag.
- The default tag is procname_uid
  - executable_name.uid
- A filesystem or a compute node can be configured to read a environment variable as the tag.

REQUIREMENTS

- User name
- Group name
- LSF project
- Compute cluster
- A pony


IMPLEMENTATION

ARCHITECTURE



RAW DATA

- Raw data is fetched in grey with lctl get_param
- LU-16110, LU-13857 – jobs_stats not valid yaml
- LU-15870 – job_stats is sometimes corrupt
- LU-15826 – job_id not quoted

```
lctl get_param obdfilter.lus24-OST0000.job_stats
obdfilter.lus24-OST0000.job_stats=
job_stats:
- job_id:          SA1CUzMGBwYH,analysis-cgp
  snapshot_time:   1667904951
  read_bytes:      { samples:   297356696, unit: bytes, min:     4096, max: 16777216, sum: 1762446621732864, sumsq: 9915053801461514240 }
  write_bytes:     { samples:    20518685, unit: bytes, min:        1, max: 16777216, sum:   98655040843302, sumsq: 4627393946935432142 }
  read:            { samples:   297356696, unit: usecs, min:        1, max: 12945252, sum:   11870681362787, sumsq: 3234937531575968175 }
  write:           { samples:    20518685, unit: usecs, min:        1, max: 10280151, sum:      98226898749, sumsq:  50535912682848539 }
  getattr:         { samples:           0, unit: usecs, min:        0, max:        0, sum:                 0, sumsq:                 0 }
  setattr:         { samples:        5697, unit: usecs, min:        4, max:  6119237, sum:         151065923, sumsq:    327770191069603 }
  punch:           { samples:     1458531, unit: usecs, min:        5, max: 31397444, sum:       15753447595, sumsq:  36605999152526109 }
  sync:            { samples:     6517520, unit: usecs, min:        0, max:  7189473, sum:       10441362075, sumsq:  12097924287750305 }
  destroy:         { samples:           0, unit: usecs, min:        0, max:        0, sum:                 0, sumsq:                 0 }
  create:          { samples:           0, unit: usecs, min:        0, max:        0, sum:                 0, sumsq:                 0 }
  statfs:          { samples:           0, unit: usecs, min:        0, max:        0, sum:                 0, sumsq:                 0 }
  get_info:        { samples:           0, unit: usecs, min:        0, max:        0, sum:                 0, sumsq:                 0 }
  set_info:        { samples:           0, unit: usecs, min:        0, max:        0, sum:                 0, sumsq:                 0 }
  quotactl:        { samples:           0, unit: usecs, min:        0, max:        0, sum:                 0, sumsq:                 0 }
  prealloc:        { samples:           0, unit:  reqs }
```

WEB APP FOR LUSTRE SERVERS

- Minimal server
- Fix yaml errors
- JSON output for each mounted device

```
def do_GET(self):
        data={}
        self._set_headers()
        osts=map( lambda x: x.replace("/proc/fs/lustre/obdfilter/",""),glob.glob("/proc/fs/lustre/obdfilter/lus*"))
        mdts=map( lambda x: x.replace("/proc/fs/lustre/mdt/",""),glob.glob("/proc/fs/lustre/mdt/lus*"))
        for ost in osts:
          data[ost]=run_lctl('lctl get_param obdfilter.{0}.job_stats'.format(ost).split())
        for mdt in mdts:
          data[mdt]=run_lctl('lctl get_param mdt.{0}.job_stats'.format(mdt).split())
        self.wfile.write(json.dumps(data))
```

WEB APP FOR LUSTRE SERVERS

```
fix_jobname = re.compile(r'^- job_id:(\s*)(\S(?:.*\S)?)\s*$')

def process_line(line):
  # Quote the job_id
  line = fix_jobname.sub(r'- job_id:\1"\2"',line)
  # Ensure valid yaml with a space
  return line.replace("max:", "max: ").replace("min:", "min: ").replace("sum:", "sum: ").replace("samples:", "samples: ")
.replace("sumsq:", "sumsq: ")

def run_lctl(cmd):
  p = Popen(cmd, stdout=PIPE)
  # skip the first two headers lines
  # https://stackoverflow.com/questions/15571137/yaml-scanner-scannererror-while-scanning-a-directive
  results=[]
  data_stream = "\n".join(map(process_line,p.communicate()[0].split("\n")[2:]))
  return  yaml.load(data_stream)
```

SAMPLE OUTPUT

```
curl -s  http://lus23-oss1-mgmt.internal.sanger.ac.uk:8081  | jq .
{
  "lus23-OST0001": [
    {
      "write_bytes": {
        "max": 16777216,
        "sum": 1275960319678,
        "samples": 413177,
        "unit": "bytes",
        "min": 4
      },
      "job_id": "SA1TwDkwOzA7,team135",
      "setattr": {
        "samples": 7,
        "unit": "reqs"
      },
….
      "read_bytes": {
        "max": 16777216,
        "sum": 51588627873792,
        "samples": 22190157,
        "unit": "bytes",
        "min": 4096
      },
      "snapshot_time": 1667913397
    }
  ]
}
```

IS 32 BYTES ENOUGH?

SA1CUzMGBwYH,analysis-cgp

SA1	The version string, used to confirm format
C	The compute farm (CASM)
UzMGBwYH	base64 encoded (uid,pgid,gid)
,	End of fixed data
analysis-cgp	LSF project

```
echo UzMGBwYH | base64 --decode  | od -x
0000000 3353 0706 0706
jb23@farm5-head1:~$ id cgppipe
uid=13139(cgppipe) gid=1798(cancer-pipeline) groups=1798(cancer-
pipeline),15495(metanorm),1300(cancer_external),1722(musperm),1116(team113),80
8(cancer),9215(team78),1534(cgppc),1410(cgp_transfer),1554(grp14mg)
jb23@farm5-head1:~$ bc
bc 1.07.1
(3*16+3)*256+5*16+3
13139
7*256+6
1798
```

ESUB - JOB PREFLIGHT SCRIPT

```
if var_has_contents("LSB_SUB_MODIFY_ENVFILE"):
    with open(os.environ["LSB_SUB_MODIFY_ENVFILE"], "a") as modfile:
        header = "SA1"
        # Write a char which is the cluster encoded, U is unknown.
        if var_has_contents("CLUSTER"):
            cluster = encode_farm.get(os.environ["CLUSTER"].splitlines()[0]  "U")
        else:
            cluster = "U"
        uids = base64.b64encode(
                struct.pack(
                    b"hhh",
                    os.getuid(),
                    pwd.getpwuid(os.getuid()).pw_gid,
                    os.getgid(),
                )
        ).decode().rstrip("=")
        if var_has_contents("LSB_PROJECT_NAME"):
            if os.environ["LSB_PROJECT_NAME"] == "default":
                project = os.environ["LSB_DEFAULTGROUP"]
            else:
                project = os.environ["LSB_PROJECT_NAME"]
        else:
            if var_has_contents("LSB_DEFAULTGROUP"):
                project = os.environ["LSB_DEFAULTGROUP"]
            else:
                project = ""
        modfile.write('SANGER_JOB_ID="' + header + cluster + uids + ',' + project + '"\n')
        modfile.write('SINGULARITYENV_SANGER_JOB_ID="' + header + cluster + uids + ',' + project + '"\n')
```

```
if [ -n "${STAT_TYPE}" ] ; then
   lctl set_param jobid_var=${STAT_TYPE} >/dev/null
fi
```

WEB APP FOR LUSTRE SERVERS

```
def decode_jobname(name,store):
    store['project']="unknown"
    store['uid']="unknown"
    store['pgid']="unknown"
    store['gid']="unknown"
    store['farm']="unknown"
    decode_farm = {
        "F": "farm5",
        "C": "casm3",
        "G": "gen3",
        "I": "impute",
        "i": "isg-test-cluster",
        "P": "pcs6",
        "S": "seq4",
        "s": "seqfarm3",
        "T": "tol"
    }
    if name.startswith("SA1"):
      store['farm']= decode_farm.get(name[3],"unknown")
      parse = re.search("^SA1.([^,]*),(.*)$", name)
      if parse is not None:
        try:
          (uid,pgid,gid)=struct.unpack('hhh',base64.b64decode(parse.group(1)+"=="))
          store['project']=parse.group(2) or ""
          store['uid']=str(uid)
          store['pgid']=str(pgid)
          store['gid']=str(gid)
        except:
          pass
```

REAL WORLD

```
curl -s  http://lus23-oss1-mgmt.internal.sanger.ac.uk:8080  | jq .
{
  "lus23-OST0001": [
    {
      "write_bytes": {
        "max": 16777216,
        "sum": 1277816778962,
        "samples": 413738,
        "unit": "bytes",
        "min": 4
      },
      "pgid": "15152",
      "job_id": "SA1TwDkwOzA7,team135",
      "uid": "14784",
…
      "project": "team135",
      "gid": "15152",
      "farm": "tol",
      "read_bytes": {
        "max": 16777216,
        "sum": 51752874618880,
        "samples": 22381014,
        "unit": "bytes",
        "min": 4096
    }
  ]
}
```

CONFUSION

- We see corruption in the job id, LU-15870, LU-16251
- We see kworker recorded as the executable with root.
- We see thread-pool with no uid.

```
- job_id:          kworker/u518:6eamtrynka
  snapshot_time:   1668009798
  read_bytes:      { samples:           1, unit: bytes, min:  663552, max:  663552,
sum:          663552 }
  write_bytes:     { samples:           0, unit: bytes, min:       0, max:       0,
sum:               0 }

- job_id:          kworker/u130:0
  snapshot_time:   1668009796
  read_bytes:      { samples:       21459, unit: bytes, min:    4096, max: 2297856,
sum:       160104448 }
  write_bytes:     { samples:           0, unit: bytes, min:       0, max:       0,
sum:               0 }

- job_id:          thread-pool-14113912
  snapshot_time:   1668009775
  read_bytes:      { samples:           2, unit: bytes, min:   36864, max:  983040,
sum:         1019904 }
  write_bytes:     { samples:           0, unit: bytes, min:       0, max:       0,
sum:               0 }

[root@lus23-oss1 ~]# lctl get_param obdfilter.lus23-OST0000.job_stats | grep job_id
| grep thread-pool | wc -l
126
[root@lus23-oss1 ~]# lctl get_param obdfilter.lus23-OST0000.job_stats | grep job_id
| grep kworker | wc -l
155
[root@lus23-oss1 ~]# lctl get_param obdfilter.lus23-OST0000.job_stats | grep job_id
| wc -l
330
```

PUSHING METRICS TO GRAPHITE

- Fetches data from all OSS's and MDS's

```
*/3 * * * * /software/isg/scripts/jobsstats -l lus20 -s lus20-mds1-mgmt -s lus20-mds2-mgmt -s lus20-oss1-mgmt \
-s lus20-oss2-mgmt -s lus20-oss3-mgmt -s lus20-oss4-mgmt -s lus20-oss5-mgmt -s lus20-oss6-mgmt | logger -t jobstats
*/3 * * * * /software/isg/scripts/jobsstats -l lus23 -s lus23-mds1-mgmt -s lus23-mds2-mgmt -s lus23-oss1-mgmt \
-s lus23-oss2-mgmt -s lus23-oss3-mgmt -s lus23-oss4-mgmt -s lus23-oss5-mgmt -s lus23-oss6-mgmt | logger -t jobstats
```

- Adds human readable versions of UID and GUD

```
def uid_to_name(uid: str) -> str:
    try:
        return pwd.getpwuid(int(uid)).pw_name
    except (KeyError, ValueError):
        return "unknown"


def gid_to_name(gid: str) -> str:
    try:
        return grp.getgrgid(gid)[0]
    except (KeyError, ValueError):
        return "unknown"
```

- Generates multiple views
  - Per Farm
  - Per LSF project
  - Per User
  - Per Group
- Send data to Graphite

SECURE LUSTRE

- A method of delivering lustre filesystems to OpenStack projects using
  - Sub directory mounts
  - And user id mappings
- https://hpc-news.sanger.ac.uk/2021/05/20/update-on-secure-lustre-lustre-users-group-may-2020/
- Jobid is inserted on the client, the uid inserted may and often does not exist on the outside
- The nodemap the request is received on is unavailable in the jobstats data
- Our current aggregation program appears to drop this traffic.


EXAMPLES

General / Lustre per user stats

now-2d to now-4m

filesystem lus23 type Enter variable value

byfarm
bygroup
bylsf_project
byusername

NB occasional u ata "spikes" are due to a Lustre server bug and should be disregarded.

Compare with overall I/O shown ternal.sanger.ac.uk/d/000000203/lustre-stats?orgId=1

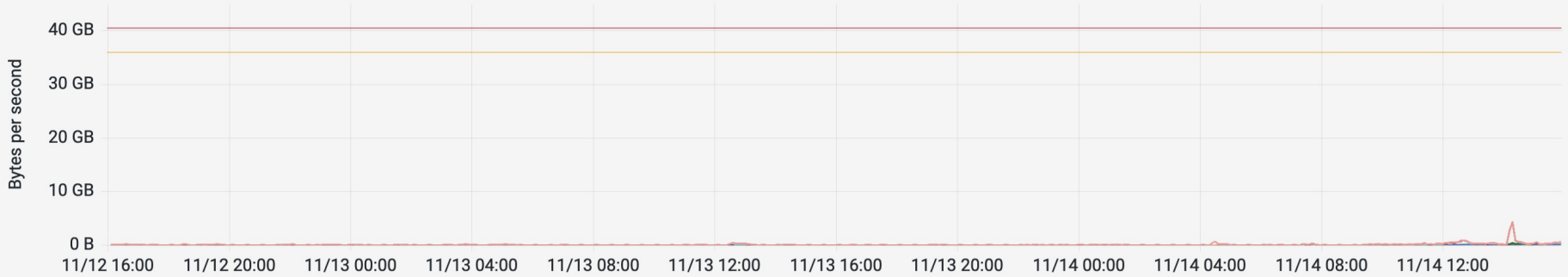
The red line indicates 90% of te ty. The yellow line indicates 85%. Around these levels, the filesystem will feel slow to interactive users.

Read/Write

data I/O (read) (stacked)

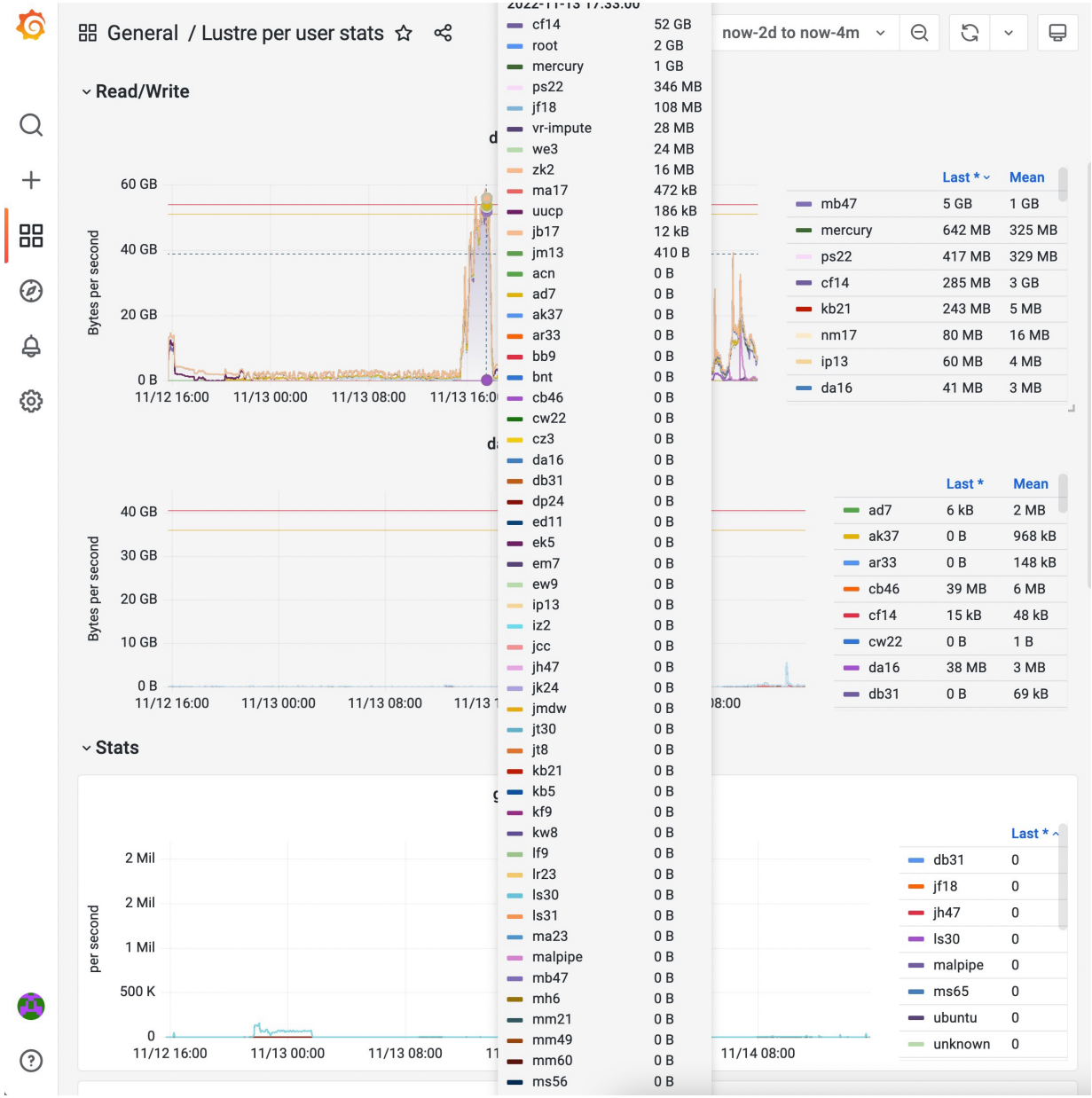
	Last *	Mean
unknown	3 GB	2 GB
team227	765 MB	3 GB
teichlab	514 MB	35 MB
team282	398 MB	7 MB
mouse_epi	189 MB	17 MB
grit-grp	133 MB	27 MB
team135	66 MB	8 MB
hgi	51 MB	17 MB

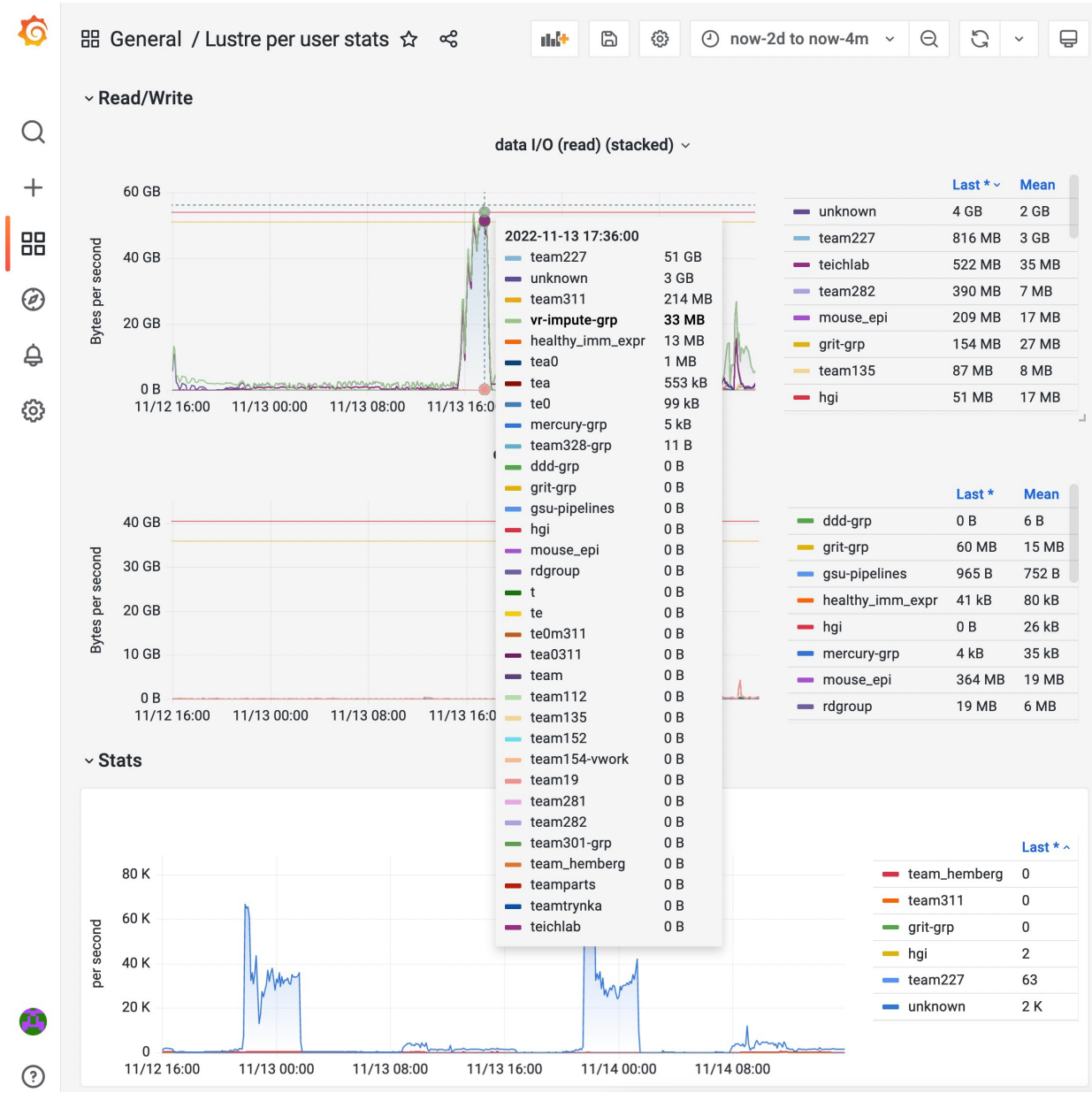
data I/O (write) (stacked)



	Last *	Mean
ddd-grp	0 B	6 B
grit-grp	78 MB	15 MB
gsu-pipelines	529 B	753 B
healthy_imm_expr	44 kB	80 kB
hgi	0 B	26 kB
mercury-grp	4 kB	35 kB
mouse_epi	337 MB	19 MB
rdgroup	24 MB	6 MB

Stats


General / Lustre per user stats
now-2d to now-4m
Read/Write
Bytes per second
60 GB
40 GB
20 GB
0 B
11/12 16:00
11/13 00:00
11/13 08:00
Last *
Mean
mb47 5 GB 1 GB
mercury 642 MB 325 MB
ps22 417 MB 329 MB
cf14 285 MB 3 GB
kb21 243 MB 5 MB
nm17 80 MB 16 MB
ip13 60 MB 4 MB
da16 41 MB 3 MB
40 GB
30 GB
20 GB
10 GB
ad7 6 kB 2 MB
ak37 0 B 968 kB
ar33 0 B 148 kB
cb46 39 MB 6 MB
cf14 15 kB 48 kB
cw22 0 B 1 B
da16 38 MB 3 MB
db31 0 B 69 kB
cf14 52 GB
root 2 GB
mercury 1 GB
ps22 346 MB
jf18 108 MB
vr-impute 28 MB
we3 24 MB
zk2 16 MB
ma17 472 kB
uucp 186 kB
jb17 12 kB
jm13 410 B
acn 0 B
ad7 0 B
ak37 0 B
ar33 0 B
bb9 0 B
bnt 0 B
cb46 0 B
cw22 0 B
cz3 0 B
da16 0 B
db31 0 B
dp24 0 B
ed11 0 B
ek5 0 B
em7 0 B
ew9 0 B
ip13 0 B
iz2 0 B
jcc 0 B
jh47 0 B
jk24 0 B
jmdw 0 B
jt30 0 B
jt8 0 B
kb21 0 B
kb5 0 B
kf9 0 B
kw8 0 B
lf9 0 B
lr23 0 B
ls30 0 B
ls31 0 B
ma23 0 B
malpipe 0 B
mb47 0 B
mh6 0 B
mm21 0 B
mm49 0 B
mm60 0 B
ms56 0 B
Stats
per second
2 Mil
1 Mil
500 K
0
11/14 08:00
db31 0
jf18 0
jh47 0
ls30 0
malpipe 0
ms65 0
ubuntu 0
unknown 0


General / Lustre per user stats
now-2d to now-4m
Read/Write
data I/O (read) (stacked)
Bytes per second
60 GB
40 GB
20 GB
0 B
11/12 16:00
11/13 00:00
11/13 08:00
11/13 16:00
2022-11-13 17:36:00
team227 51 GB
unknown 3 GB
team311 214 MB
vr-impute-grp 33 MB
healthy_imm_expr 13 MB
tea0 1 MB
tea 553 kB
te0 99 kB
mercury-grp 5 kB
team328-grp 11 B
ddd-grp 0 B
grit-grp 0 B
gsu-pipelines 0 B
hgi 0 B
mouse_epi 0 B
rdgroup 0 B
t 0 B
te 0 B
te0m311 0 B
tea0311 0 B
team 0 B
team112 0 B
team135 0 B
team152 0 B
team154-vwork 0 B
team19 0 B
team281 0 B
team282 0 B
team301-grp 0 B
team_hemberg 0 B
teamparts 0 B
teamtrynka 0 B
teichlab 0 B
Last *
Mean
unknown 4 GB 2 GB
team227 816 MB 3 GB
teichlab 522 MB 35 MB
team282 390 MB 7 MB
mouse_epi 209 MB 17 MB
grit-grp 154 MB 27 MB
team135 87 MB 8 MB
hgi 51 MB 17 MB
40 GB
30 GB
20 GB
10 GB
ddd-grp 0 B 6 B
grit-grp 60 MB 15 MB
gsu-pipelines 965 B 752 B
healthy_imm_expr 41 kB 80 kB
hgi 0 B 26 kB
mercury-grp 4 kB 35 kB
mouse_epi 364 MB 19 MB
rdgroup 19 MB 6 MB
Stats
per second
80 K
60 K
40 K
20 K
0
11/13 16:00
11/14 00:00
11/14 08:00
team_hemberg 0
team311 0
grit-grp 0
hgi 2
team227 63
unknown 2 K

General / Lustre per user stats

filesystem: lus26 | type: byusername | now-7d to now-4m

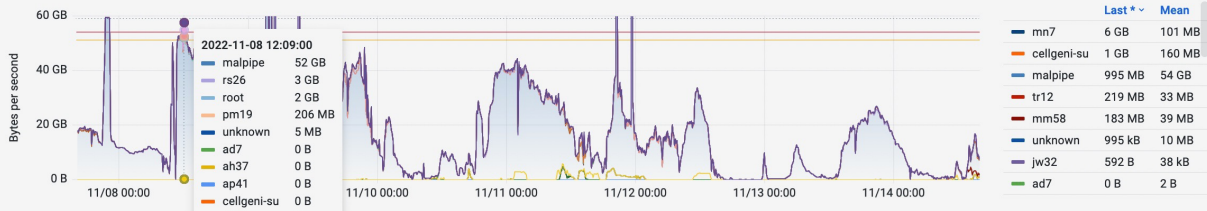
NB occasional unexpected data "spikes" are due to a Lustre server bug and should be disregarded.

Compare with overall I/O shown here: https://metrics.internal.sanger.ac.uk/d/000000203/lustre-stats?orgId=1

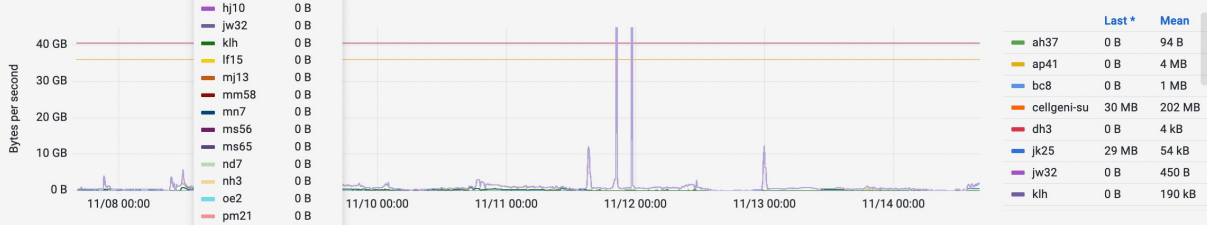
The red line indicates 90% of tested filesystem capability. The yellow line indicates 85%. Around these levels, the filesystem will feel slow to interactive users.

Read/Write

data I/O (read) (stacked)



data I/O (write) (stacked)



Stats

get stat ops per second

General / Lustre per user stats

filesystem: lus26 | type: bylsf_project | now-7d to now-4m

NB occasional unexpected data "spikes" are due to a Lustre server bug and should be disregarded.

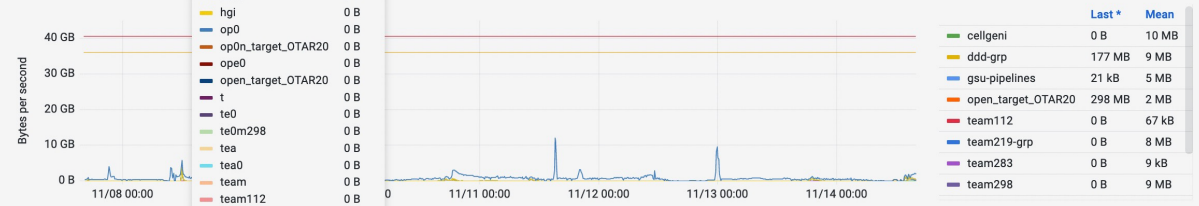
Compare with overall I/O shown here: https://metrics.internal.sanger.ac.uk/d/000000203/lustre-stats?orgId=1

The red line indicates 90% of tested filesystem capability. The yellow line indicates 85%. Around these levels, the filesystem will feel slow to interactive users.

Read/Write

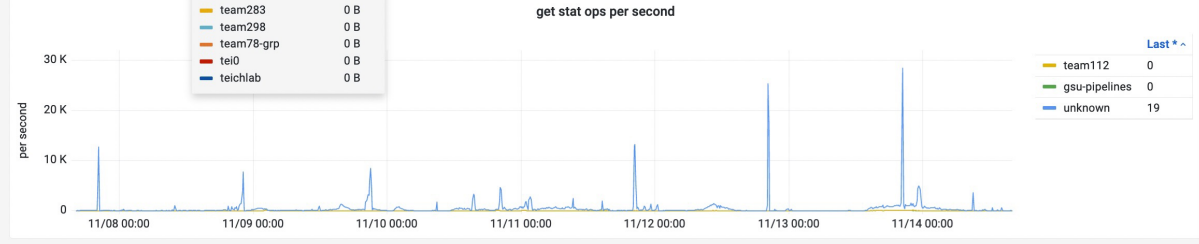
data I/O (read) (stacked)

data I/O (write) (stacked)



Stats

get stat ops per second



General / Lustre per user stats

filesystem lus26 type byusername

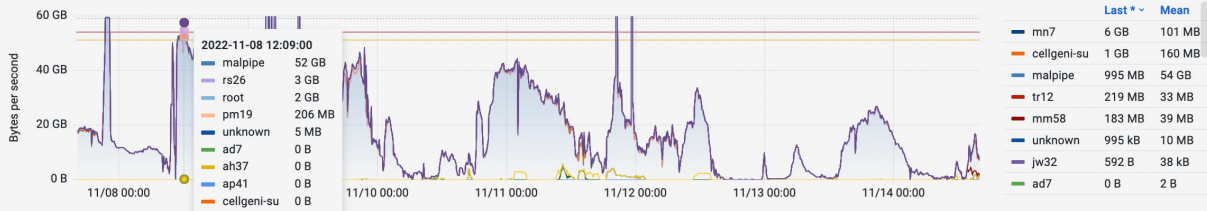
NB occasional unexpected data "spikes" are due to a Lustre server bug and should be disregarded.

Compare with overall I/O shown here: https://metrics.internal.sanger.ac.uk/d/000000203/lustre-stats?orgId=1

The red line indicates 90% of tested filesystem capability. The yellow line indicates 85%. Around these levels, the filesystem will feel slow to interactive users.

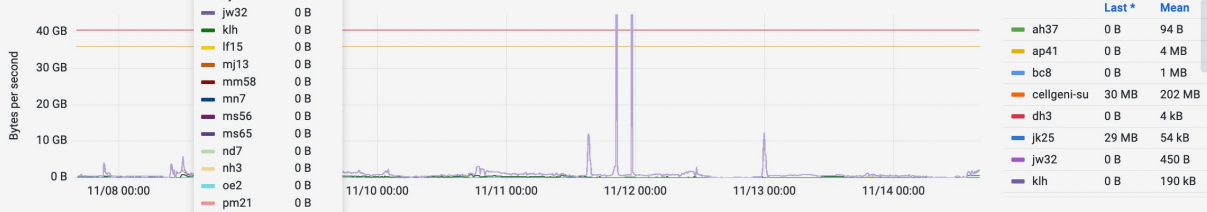
Read/Write

data I/O (read) (stacked)



	Last *	Mean
mn7	6 GB	101 MB
cellgeni-su	1 GB	160 MB
malpipe	995 MB	54 GB
tr12	219 MB	33 MB
mm58	183 MB	39 MB
unknown	995 kB	10 MB
jw32	592 B	38 kB
ad7	0 B	2 B

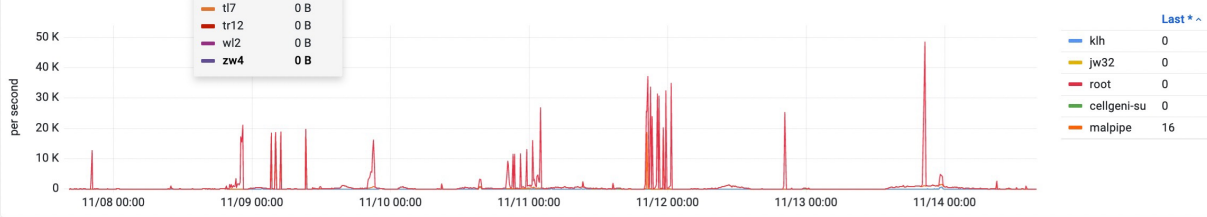
data I/O (write) (stacked)



	Last *	Mean
ah37	0 B	94 B
ap41	0 B	4 MB
bc8	0 B	1 MB
cellgeni-su	30 MB	202 MB
dh3	0 B	4 kB
jk25	29 MB	54 kB
jw32	0 B	450 B
klh	0 B	190 kB

Stats

get stat ops per second



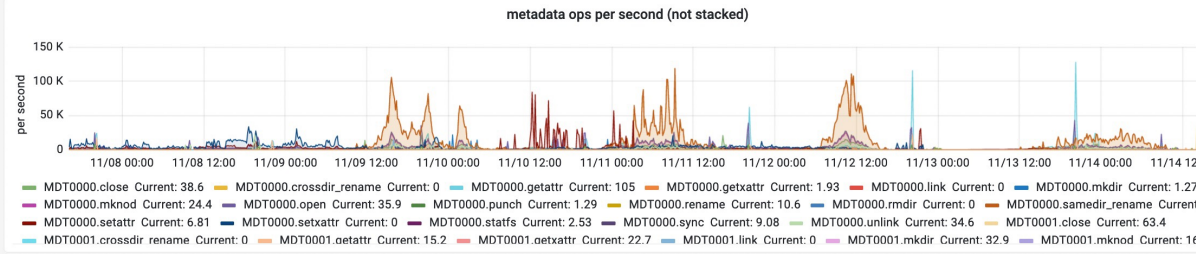
	Last *
klh	0
jw32	0
root	0
cellgeni-su	0
malpipe	16

General / Lustre stats

filesystem lus26

metadata I/O

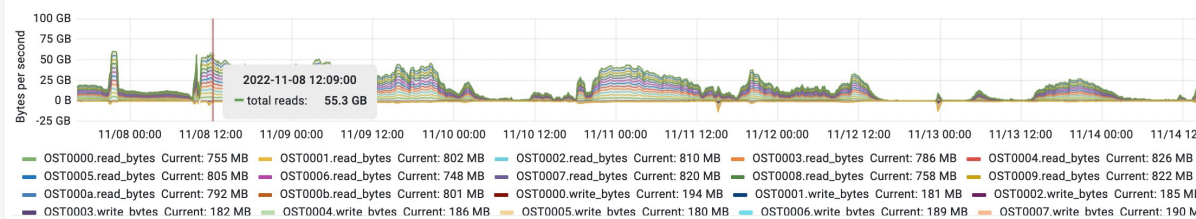
metadata ops per second (not stacked)



data-on-metadata stats (for lus24 and later) *(1 panel)*

data I/O

data I/O (read +ve, write -ve) (stacked)



CONCLUSIONS

CONCLUSIONS

- Small change on Lustre server
- Not perfect
- Still useful
  - Want fair share for I/O
    - Is NRS the answer to that question
- Released under Apache 2 license

ACKNOWLEDGMENTS

- ISG (Photo is missing Anna, Bruno, Fabio and Sai) (And includes Brett who is no longer with us).

THANK YOU

@wellcomegenomecampus

@wellcomegenome

Wellcome Genome Campus

Wellcome Genome Campus