

高通量科学计算集群及Lustre文件系统

High Throughput Scientific Computing Clusters And Lustre Filesystem In Tsinghua University

- 清华信息科学与技术国家实验室（筹）公共平台与技术部
- 清华大学科学与工程计算实验室
- 清华大学计算机系高性能所

Outline

- Scientific Computing Cluster In Tsinghua
 - Configurations
 - Clusters & Lustre
 - Usage Statistics
 - Jobs
 - Data Characters&Statistics
- Lustre File Systems On Our Cluster(s)
 - Reliability&Recoveribility Considerations At System Installation & Deployment Phase
 - Tuning/Monitoring Parameters/Metrics
 - Problems We Have Encountered

Clusters for Scientific Computing

CPU :

xeon5670

740 nodes, 8880 cores

memory:

32G/node × 370,

48G/node × 370

Interconnect :

QDR, 2 level

37 leaf switchs+1 core switch



Storage:

总容量: ~1PB

~2TB SATA*500

IO nodes: 24

Connected to the core IB switch

Lustre

1.8.5

Clusters for Scientific Computing

CPU :

xeon E5 2670

140 nodes, 2240 cores

memory:

32G/node

Interconnect :

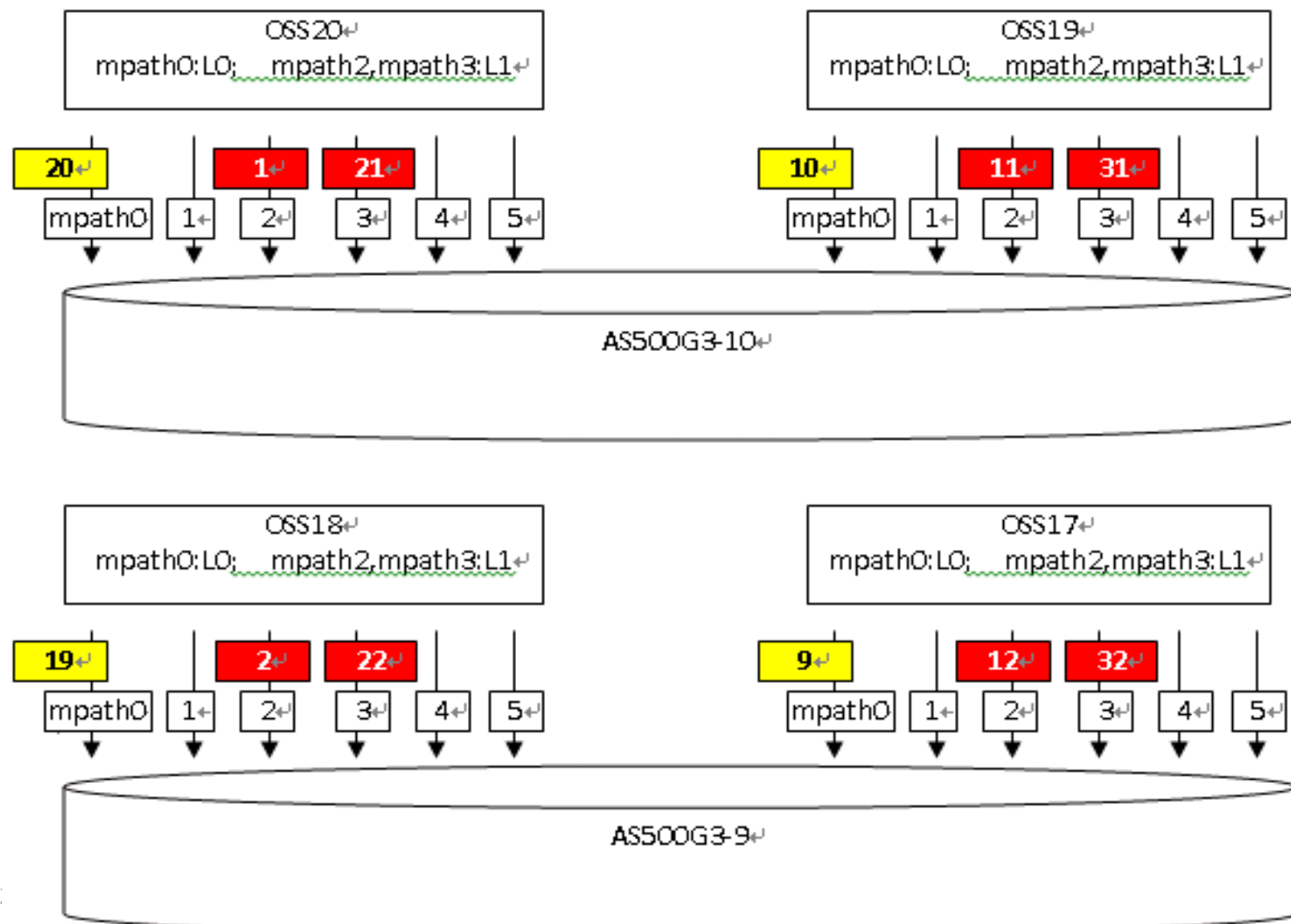
QDR, 1 core switch

Lustre

2.4

The Lustre Filesystem Configuration in Detail

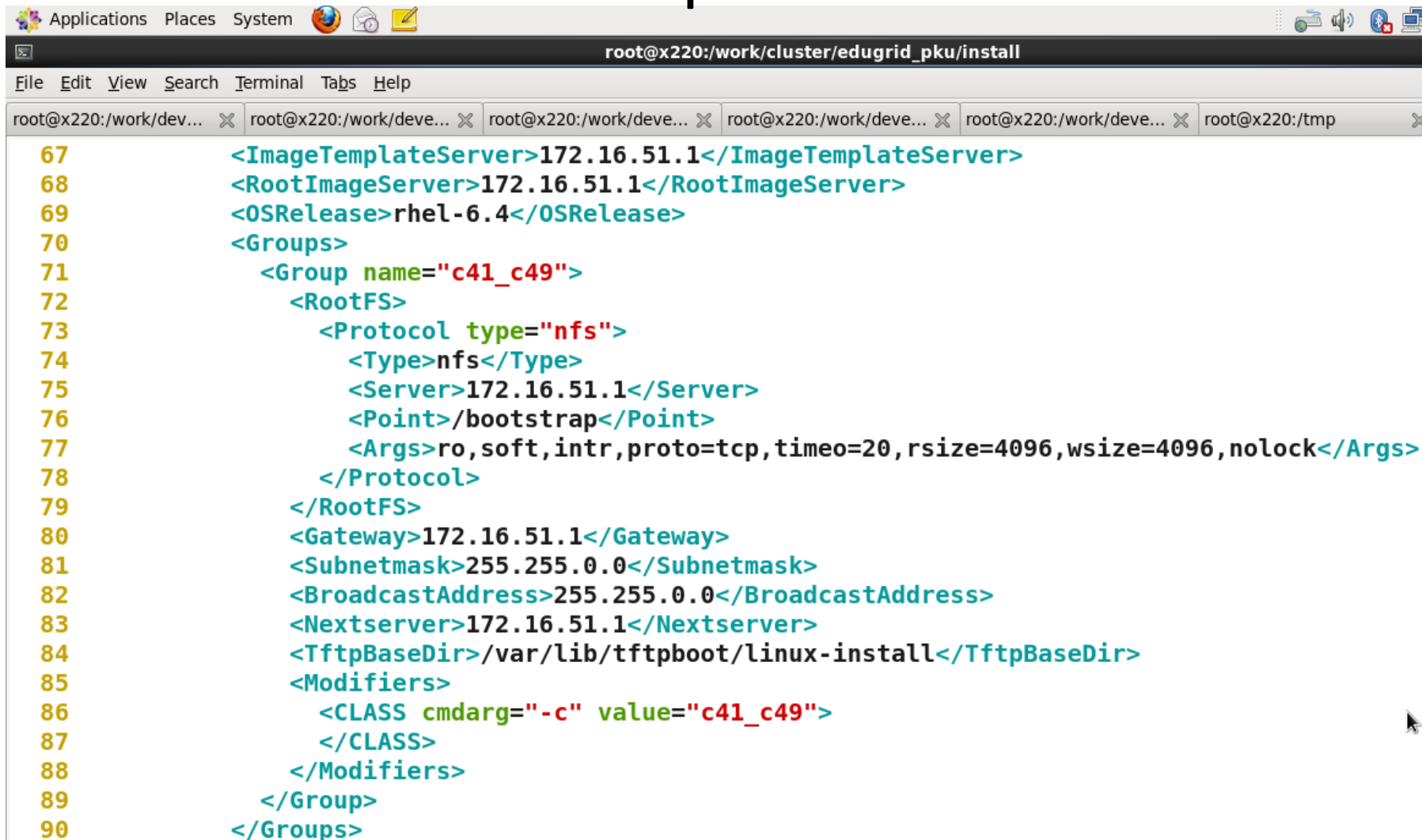
- Effective Capacity: \sim 560TB
- 22 OSS nodes
- 2 MDS nodes
- Connected to Backbone IB Switch with direct link
- 12 SAN boxes, with 8Gb fiber connection. 11 for 22 OSS nodes, and 1 for 2 MDS nodes
- Each OSS nodes provides 4 disk partitions
- Each SAN box provides 2x4 disk partitions for 2 OSS nodes



Operation Situation

- Using Customized Cluster Operating System(CoS) Developed By Ourselves (On top of RedHat)
 - Centralized Single System Image, Customized Kernel, Specializable by Machine Group
 - Online System Image Service are provided over Network with Failover and load balancing(LVS)
 - Flexible Machine Group Definition
 - All Roles, Configurations and System information are Managed by Single XML file—hierarchical, extendable, inheritable, and overridable
 - Automated running, maintenance, and recovery .
 - Smart power control.

ClusterDescriptor.default.xml



```
root@x220:/work/cluster/edugrid_pku/install
File Edit View Search Terminal Tabs Help
root@x220:/work/dev... x root@x220:/work/deve... x root@x220:/work/deve... x root@x220:/work/deve... x root@x220:/work/deve... x root@x220:/tmp
67 <ImageTemplateServer>172.16.51.1</ImageTemplateServer>
68 <RootImageServer>172.16.51.1</RootImageServer>
69 <OSRelease>rhel-6.4</OSRelease>
70 <Groups>
71   <Group name="c41_c49">
72     <RootFS>
73       <Protocol type="nfs">
74         <Type>nfs</Type>
75         <Server>172.16.51.1</Server>
76         <Point>/bootstrap</Point>
77         <Args>ro,soft,intr,proto=tcp,timeo=20,rsz=4096,wsz=4096,nolock</Args>
78       </Protocol>
79     </RootFS>
80     <Gateway>172.16.51.1</Gateway>
81     <Subnetmask>255.255.0.0</Subnetmask>
82     <BroadcastAddress>255.255.0.0</BroadcastAddress>
83     <Nextserver>172.16.51.1</Nextserver>
84     <TftpBaseDir>/var/lib/tftpboot/linux-install</TftpBaseDir>
85     <Modifiers>
86       <CLASS cmdarg="-c" value="c41_c49">
87     </CLASS>
88   </Modifiers>
89 </Group>
90 </Groups>
```


ClusterDescriptor.default.xml

```
Applications Places System root@x220:/work/cluster/edugrid_p
File Edit View Search Terminal Tabs Help
root@x220:/work/dev... root@x220:/work/deve... root@x220:/work/deve... root@x220:/work/deve...
90     </Groups>
91     <Hw>
92         <UserName>admin</UserName>
93         <Password>admin</Password>
94     </Hw>
95 </Globals>
96 <Hosts>
97     <Host>
98         <RootFS>
99             <Protocol>
100                 <Type></Type>
101                 <Server></Server>
102                 <Point></Point>
103                 <Args></Args>
104             </Protocol>
105         </RootFS>
106         <Hostname>c41b01</Hostname>
107         <NextServer></NextServer>
108         <TftpBaseDir></TftpBaseDir>
109         <TftpImageFile></TftpImageFile>
110         <Interfaces>
111             <Interface name="eth0">
112                 <Identifier>00:E0:81:E1:6A:4D</Identifier>
113                 <Ip>172.16.41.1</Ip>
114                 <Bootproto>static</Bootproto>
```

Operation Situation

- Dedicated for High throughput Computing
 - Over 600 Users, 400 of daily active
 - Over 2Million jobs a year
 - Over 90% usage rate for more than 77% days in a year
- All costs are covered by the platform itself.
 - Electric charge, Operator's salary, Room charge, equipment maintenance

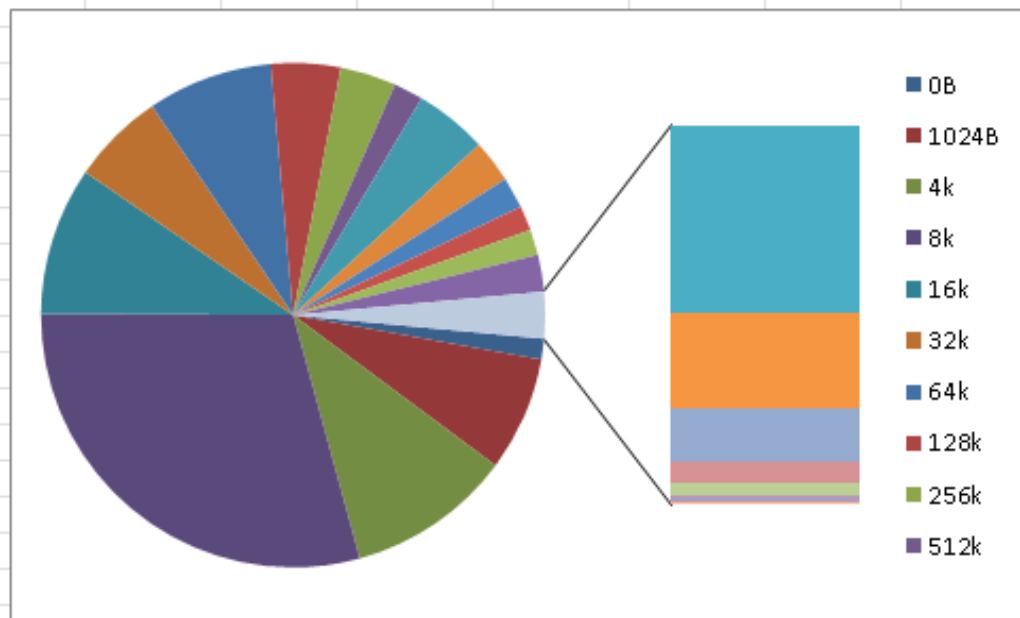
Jobs on the cluster(s)

- Mainly Numerical Computing codes
 - ~50% first principle and related———Vasp, Siesta, Dvm-Dac. Named, XMD, Gromacs, Material Science...
 - ~20% Fluid Dynamics———Earth science, Aeronautics and Astronautics, petroleum,...
 - ~20% Life science
 - ~10% Miscs
 - Supports Leading researchs(1~2 Nature and/or Science Publications a year)

Data Statistics

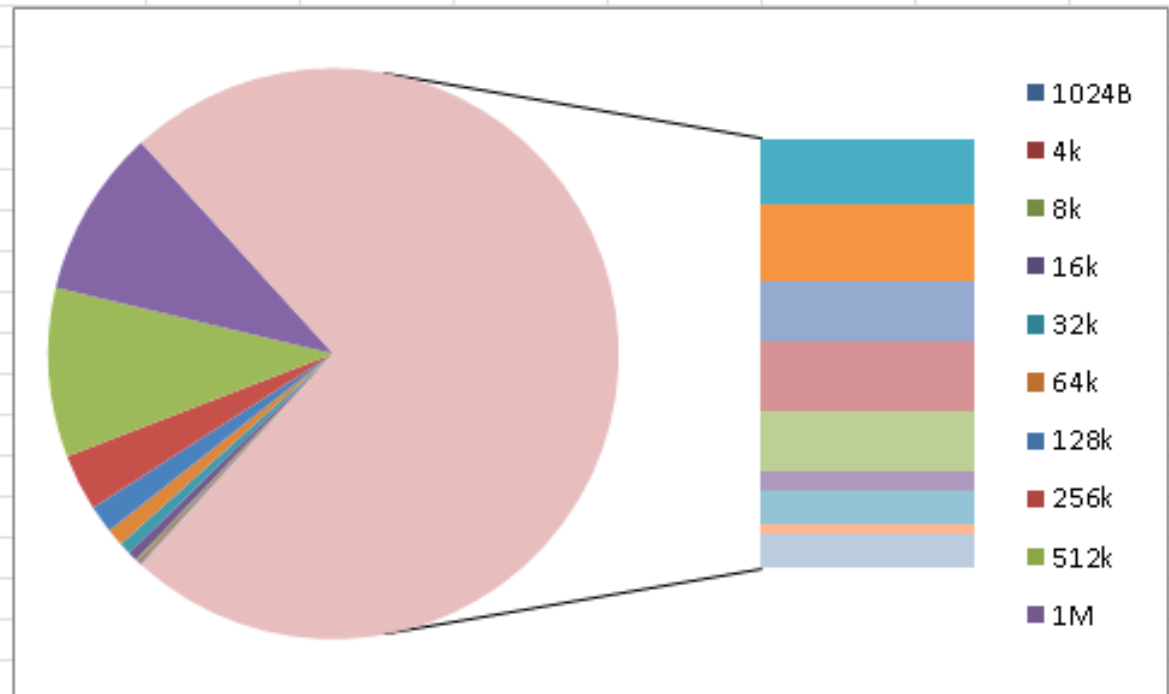
- Number of files: 30Million ~ 50Million

-1B	0	0
0B	290638	0.013240979
1024B	1610712	0.073381334
4k	2324746	0.105911525
8k	6443954	0.293575725
16k	2082072	0.094855705
32k	1305886	0.059493974
64k	1764531	0.08038907
128k	977871	0.044550161
256k	786367	0.03582556
512k	416011	0.018952763
1M	1026241	0.046753817
2M	598828	0.027281598
4M	440102	0.020050308
8M	349609	0.015927599
16M	358316	0.016324275
32M	517161	0.023560987
64M	323814	0.014752422
128M	164611	0.007499401
256M	93693	0.004268496
512M	36456	0.001660874
1G	23172	0.001055677
2G	9426	0.000429433
3G	1632	7.43512E-05
4G	2258	0.000102871
10/15G	502	2.28703E-05
6G	1278	5.82235E-05



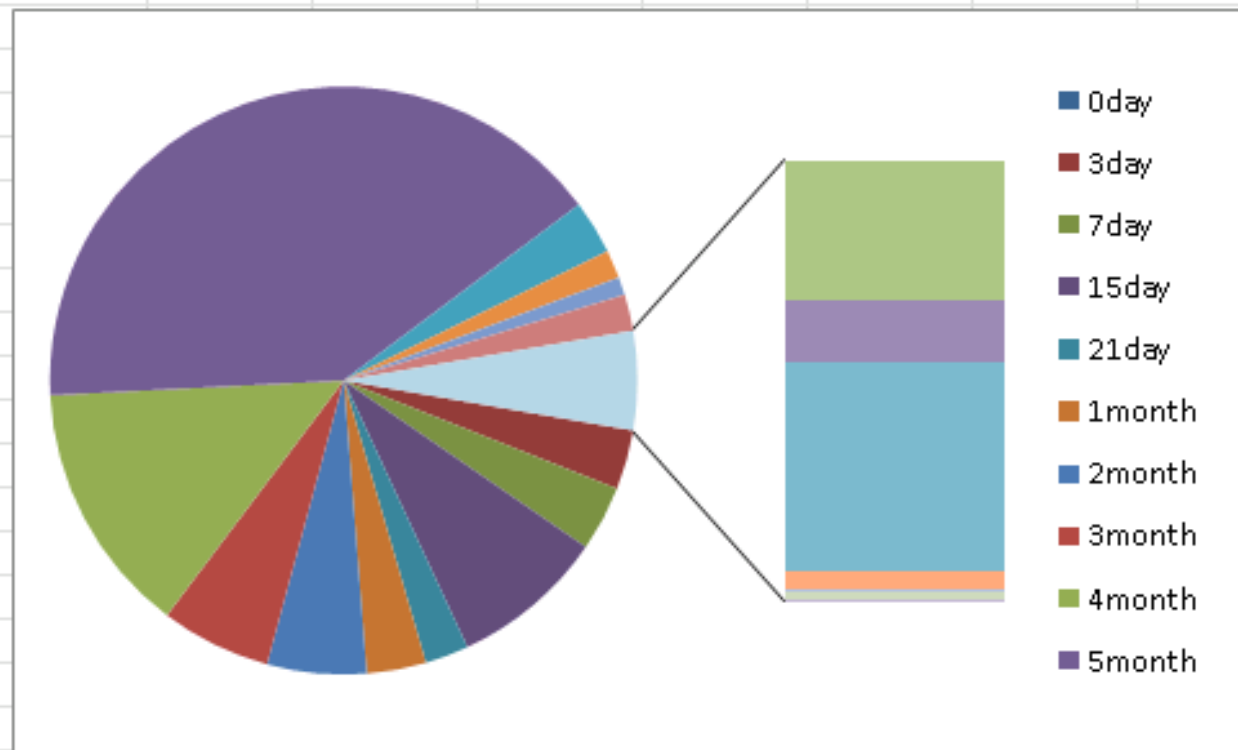
Data Statistics

1024B	525455227	3.86709E-06
4k	6430816462	4.73277E-05
8k	45509772502	0.00033493
16k	23132795456	0.000170246
32k	34739069569	0.000255663
64k	68014441891	0.000500553
128k	95388343942	0.000702012
256k	133897693782	0.000985422
512k	167138549970	0.001230058
1M	727106113395	0.005351147
2M	925408259787	0.006810554
4M	1336583761538	0.009836605
8M	2116162132099	0.015573923
16M	4374533250216	0.032194434
32M	13020633219257	0.095825518
64M	12939353138311	0.095227336
128M	15059814246405	0.110832897
256M	17993710889500	0.132424947
512M	14119693068930	0.103914063
1G	16448209990084	0.121050814
2G	13551834767976	0.099734903
3G	4553455850774	0.033511217
4G	7815057723100	0.057515018
5G	2318587152938	0.017063672
6G	8003637446005	0.058902873



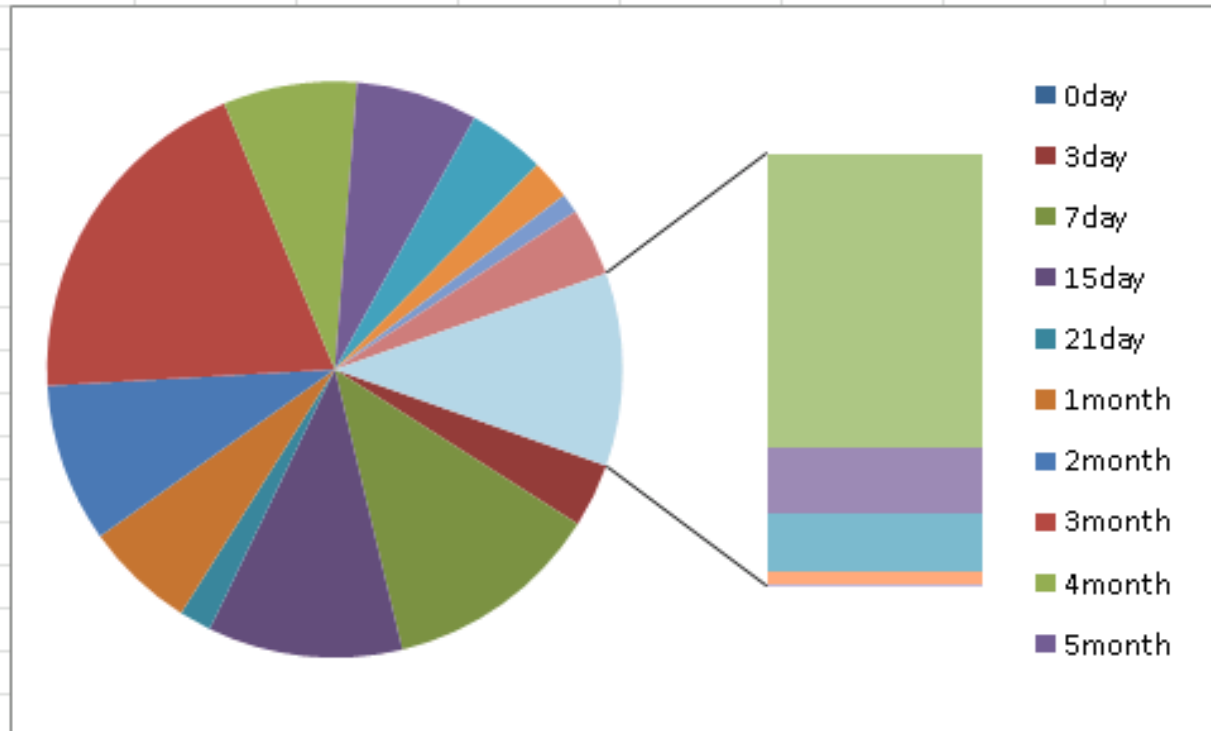
Data Statistics

0day	0	0
3day	719777	0.032788535
7day	787166	0.035858356
15day	1865295	0.084971165
21day	526756	0.023995706
1month	718468	0.032728905
2month	1198087	0.054577345
3month	1333548	0.0607481
4month	3055227	0.139177019
5month	8908069	0.405795866
6month	651407	0.029674026
7month	339796	0.015478979
8month	216086	0.009843526
9month	439499	0.020020824
10month	376969	0.017172348
11month	173889	0.007921294
12month	561418	0.02557469
15month	51225	0.00233349
18month	5014	0.000228406
21month	447	2.03625E-05
24month	23951	0.001091058



Data Statistics

3day	6444553489920	0.03596
7day	21897563464393	0.122185
15day	19605454735415	0.109395
21day	3238292936215	0.018069
1month	11121958236125	0.062059
2month	15989774639037	0.089221
3month	35169555053603	0.196241
4month	13446360264209	0.075029
5month	12273163430561	0.068482
6month	7790770958563	0.043471
7month	4012976429004	0.022392
8month	1975893613856	0.011025
9month	6845388461267	0.038196
10month	13199901536607	0.073653
11month	2996182976505	0.016718
12month	2614267209436	0.014587
15month	485296688541	0.002708
18month	20693645508	0.000115
21month	1315632507	7.34E-06
24month	86989856998	0.000485



FileSystem Metrics—Client

- /proc/fs/lustre/llite/*/read_ahead_stats
- /proc/fs/lustre/llite/*/stats

Dirty_pages_hits

Dirty_pages_misses

Readkb/Reads

Writekb/Writes

Open

Close

Seek

Fsync

Setattr

Getattr

Hits

Misses

Readpage not consecutive

Miss inside window

Failed grab_cache_page

Failed lock match

Read but discarded

Zero length file

Zero size window

Read-ahead to EOF

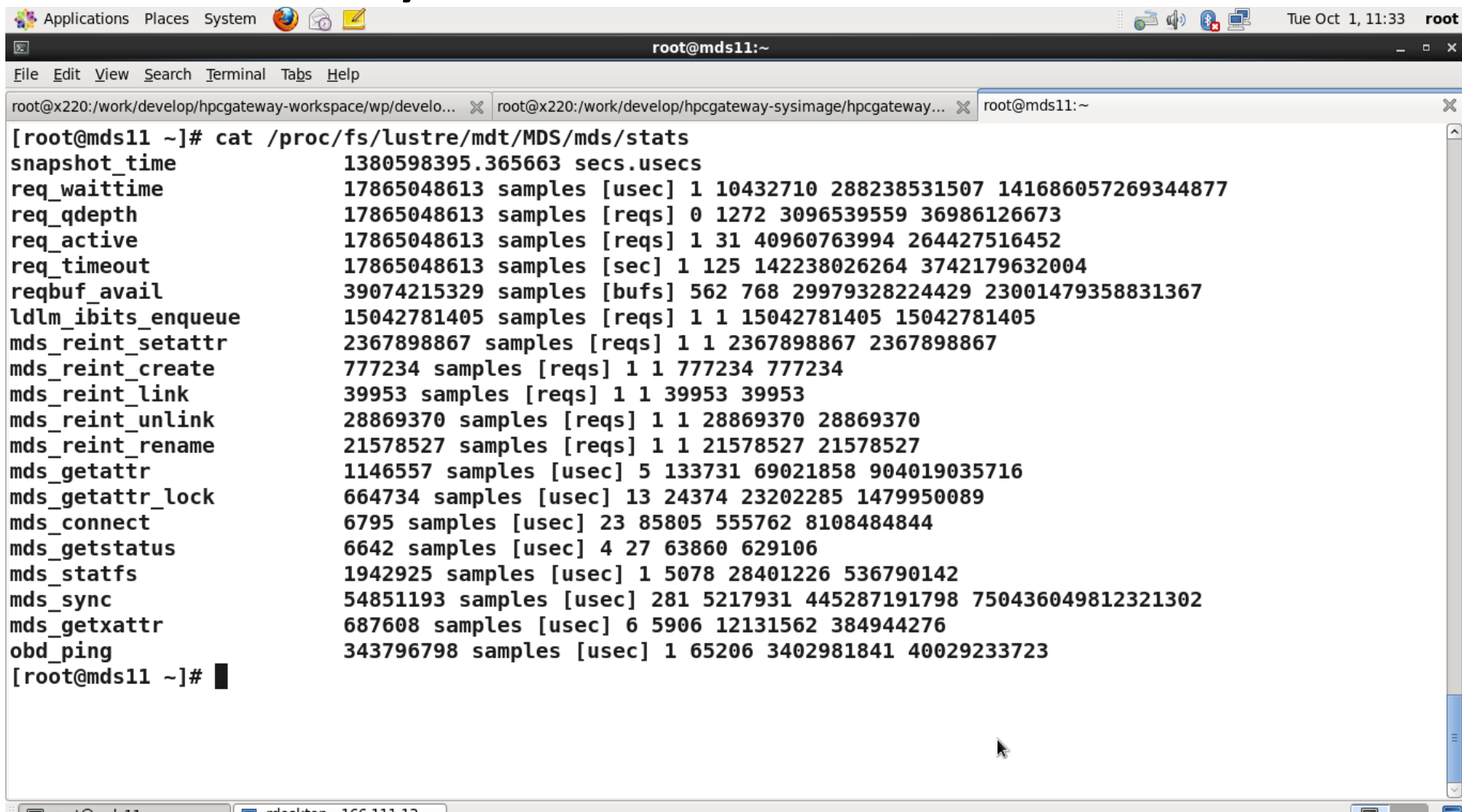
Hit max r-a issue

**Wrong page from
grab_cache_page**

FileSystem Metrics——Client

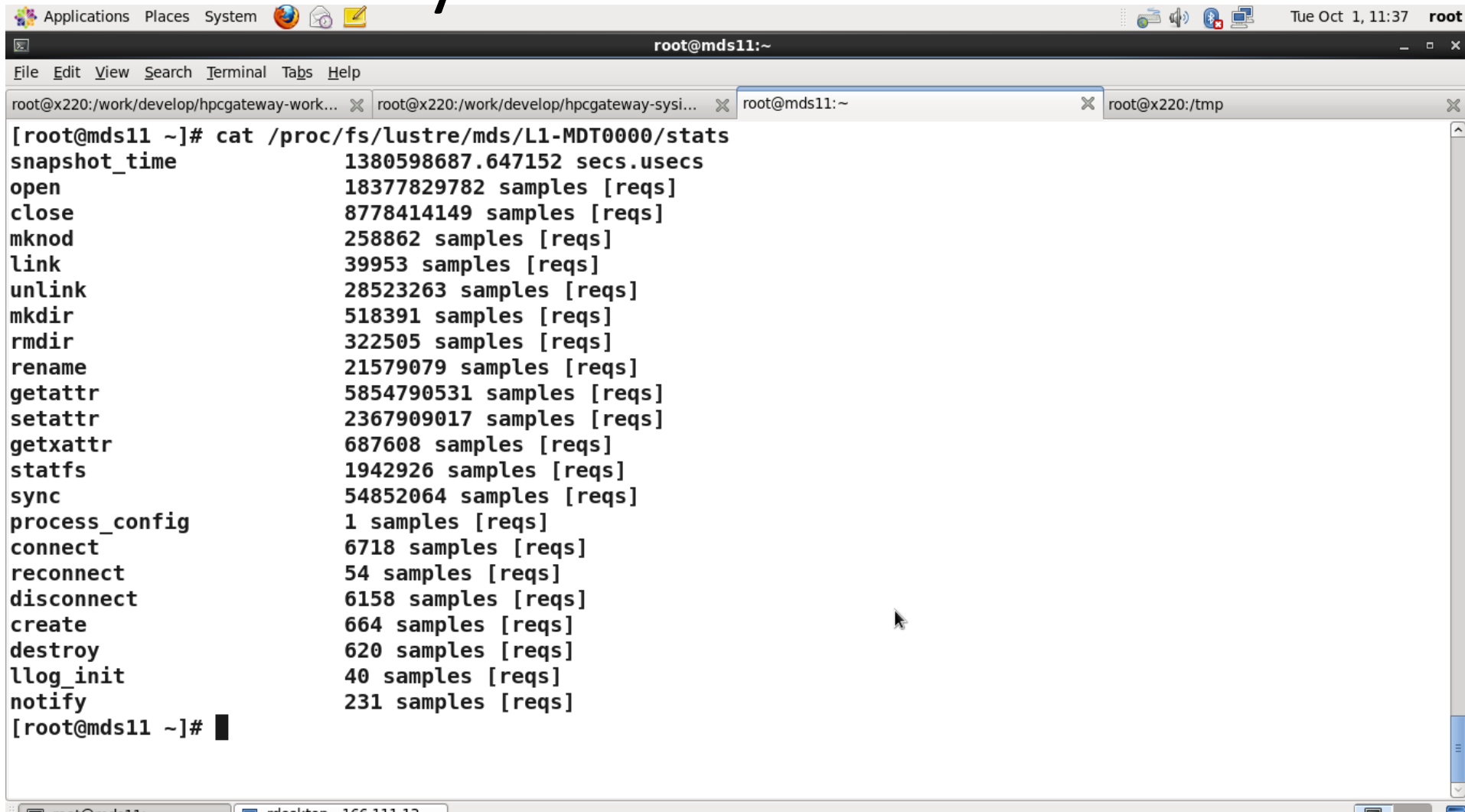
- **Client OSS Agent Statistics, RPC Buffer statistics**
- **`/proc/fs/lustre/osc/(.*)/rpc_stats`**
- **`/proc/fs/lustre/osc/(.*)/stats`**
 - 1K
 - 2K
 - 4K
 - 8K
 - 16K
 - 32K
 - 64K
 - 128K
 - 256K

FileSystem Metrics—MDS



```
root@mds11:~  
File Edit View Search Terminal Tabs Help  
root@x220:/work/develop/hpcgateway-workspace/wp/develo... x root@x220:/work/develop/hpcgateway-sysimage/hpcgateway... x root@mds11:~  
[root@mds11 ~]# cat /proc/fs/lustre/mdt/MDS/mds/stats  
snapshot_time          1380598395.365663 secs.usecs  
req_waitempty          17865048613 samples [usec] 1 10432710 288238531507 141686057269344877  
req_qdepth             17865048613 samples [reqs] 0 1272 3096539559 36986126673  
req_active             17865048613 samples [reqs] 1 31 40960763994 264427516452  
req_timeout            17865048613 samples [sec] 1 125 142238026264 3742179632004  
reqbuf_avail          39074215329 samples [bufs] 562 768 29979328224429 23001479358831367  
ldlm_ibits_enqueue    15042781405 samples [reqs] 1 1 15042781405 15042781405  
mds_reint_setattr     2367898867 samples [reqs] 1 1 2367898867 2367898867  
mds_reint_create      777234 samples [reqs] 1 1 777234 777234  
mds_reint_link        39953 samples [reqs] 1 1 39953 39953  
mds_reint_unlink      28869370 samples [reqs] 1 1 28869370 28869370  
mds_reint_rename      21578527 samples [reqs] 1 1 21578527 21578527  
mds_getattr           1146557 samples [usec] 5 133731 69021858 904019035716  
mds_getattr_lock      664734 samples [usec] 13 24374 23202285 1479950089  
mds_connect           6795 samples [usec] 23 85805 555762 8108484844  
mds_getstatus         6642 samples [usec] 4 27 63860 629106  
mds_statfs            1942925 samples [usec] 1 5078 28401226 536790142  
mds_sync              54851193 samples [usec] 281 5217931 445287191798 750436049812321302  
mds_getxattr          687608 samples [usec] 6 5906 12131562 384944276  
obd_ping              343796798 samples [usec] 1 65206 3402981841 40029233723  
[root@mds11 ~]#
```

FileSystem Metrics—MDS

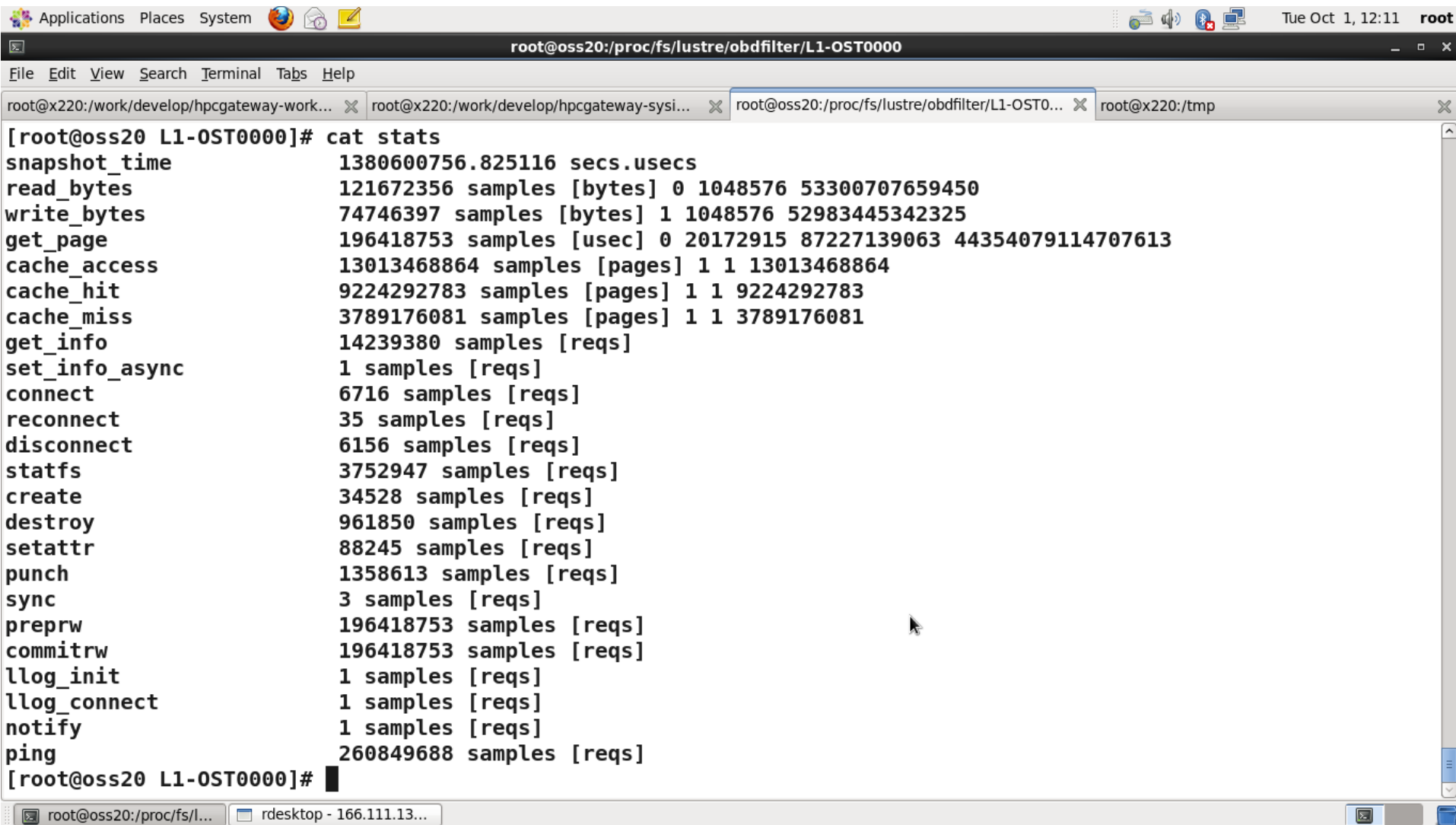


The screenshot shows a terminal window on a Linux system. The terminal title is 'root@mds11:~'. The user has executed the command 'cat /proc/fs/lustre/mds/L1-MDT0000/stats'. The output displays various metrics for the MDS process, including snapshot_time, open, close, mknod, link, unlink, mkdir, rmdir, rename, getattr, setattr, getxattr, statfs, sync, process_config, connect, reconnect, disconnect, create, destroy, llog_init, and notify. Each metric is followed by its value and unit.

```
[root@mds11 ~]# cat /proc/fs/lustre/mds/L1-MDT0000/stats
snapshot_time      1380598687.647152 secs.usecs
open               18377829782 samples [reqs]
close             8778414149 samples [reqs]
mknod             258862 samples [reqs]
link              39953 samples [reqs]
unlink            28523263 samples [reqs]
mkdir             518391 samples [reqs]
rmdir             322505 samples [reqs]
rename            21579079 samples [reqs]
getattr           5854790531 samples [reqs]
setattr           2367909017 samples [reqs]
getxattr          687608 samples [reqs]
statfs            1942926 samples [reqs]
sync              54852064 samples [reqs]
process_config    1 samples [reqs]
connect           6718 samples [reqs]
reconnect         54 samples [reqs]
disconnect        6158 samples [reqs]
create            664 samples [reqs]
destroy           620 samples [reqs]
llog_init         40 samples [reqs]
notify            231 samples [reqs]
[root@mds11 ~]#
```

FileSystem Metrics—OST

- `/proc/fs/lustre/obdfilter/(.*)/stats`
- `/proc/fs/lustre/obdfilter/(.*)/brw_stats`



The image shows a terminal window on a Linux system. The window title is `root@oss20:/proc/fs/lustre/obdfilter/L1-OST0000`. The terminal output shows the command `cat stats` and its output, which lists various file system metrics for the L1-OST0000. The metrics include snapshot_time, read_bytes, write_bytes, get_page, cache_access, cache_hit, cache_miss, get_info, set_info_async, connect, reconnect, disconnect, statfs, create, destroy, setattr, punch, sync, preprw, commitrw, llog_init, llog_connect, notify, and ping. Each metric is followed by its value and units.

```
[root@oss20 L1-OST0000]# cat stats
snapshot_time      1380600756.825116 secs.usecs
read_bytes         121672356 samples [bytes] 0 1048576 53300707659450
write_bytes        74746397 samples [bytes] 1 1048576 52983445342325
get_page           196418753 samples [usec] 0 20172915 87227139063 44354079114707613
cache_access       13013468864 samples [pages] 1 1 13013468864
cache_hit          9224292783 samples [pages] 1 1 9224292783
cache_miss         3789176081 samples [pages] 1 1 3789176081
get_info           14239380 samples [reqs]
set_info_async     1 samples [reqs]
connect            6716 samples [reqs]
reconnect          35 samples [reqs]
disconnect         6156 samples [reqs]
statfs             3752947 samples [reqs]
create             34528 samples [reqs]
destroy            961850 samples [reqs]
setattr            88245 samples [reqs]
punch              1358613 samples [reqs]
sync               3 samples [reqs]
preprw             196418753 samples [reqs]
commitrw           196418753 samples [reqs]
llog_init          1 samples [reqs]
llog_connect       1 samples [reqs]
notify             1 samples [reqs]
ping               260849688 samples [reqs]
[root@oss20 L1-OST0000]#
```

FileSystem Metrics—OST

- `/proc/fs/lustre/obdfilter/(.*)/stats`
- `/proc/fs/lustre/obdfilter/(.*)/brw_stats`

```
snapshot_time:          1380600772.322557 (secs.usecs)
                        read           |           write
pages per bulk r/w      rpcs  % cum % |  rpcs  % cum %
1:                      29438516  60  60 | 18535154  24  24
.....

                        read           |           write
discontiguous pages    rpcs  % cum % |  rpcs  % cum %
0:                      45892121  93  93 | 71007282  94  94
[root@oss20 L1-OST0000]# cat brw_stats
snapshot_time:          1380600772.322557 (secs.usecs)
                        read           |           write
pages per bulk r/w      rpcs  % cum % |  rpcs  % cum %
1:                      29438516  60  60 | 18535154  24  24
.....

                        read           |           write
discontiguous pages    rpcs  % cum % |  rpcs  % cum %
0:                      45892121  93  93 | 71007282  94  94
.....

                        read           |           write
discontiguous blocks   rpcs  % cum % |  rpcs  % cum %
0:                      45785164  93  93 | 68930663  92  92
```

FileSystem Metrics—OST

- `/proc/fs/lustre/obdfilter/(.*)/stats`
- `/proc/fs/lustre/obdfilter/(.*)/brw_stats`

```

                                read      |      write
disk fragmented I/Os      ios      % cum % |      ios      % cum %
0:                          2769489      5   5 |      0       0   0
.....

                                read      |      write
disk I/Os in flight      ios      % cum % |      ios      % cum %
1:                          21263953      31  31 |     37654247      44  44
.....

                                read      |      write
I/O time (1/1000s)      ios      % cum % |      ios      % cum %
1:                          7158700      14  14 |     19371817      25  25
                                read      |      write
.....
disk I/O size            ios      % cum % |      ios      % cum %
4K:                        37408204      56  56 |     25300879      30  30
```


FileSystem Metrics——OST

OST	ReadKB	Reads	SizeKB	WriteKB	Writes	SizeKB
L1-OST00	1038	2	519	0	0	0
L1-OST14	1022	2	511	0	0	0
L1-OST00	996	5	43	0	0	0
L1-OST14	983	4	566	0	0	0
L1-OST00	453	0	659	0	0	0
L1-OST14	894	3	784	0	0	0

Problems

- StripCount = 1 and StripCount > 1 shows totally different behavior on Cpu/Memory Usage?
 - ls, du, ... , client crash
- lfs conf_param xxx not work?
- mdt backup tools (like rsync/cpio other than dd?)
- space balances among all OSTs(StripCount=1)
- More smart read/write separation policy?
 - Codes from different type of research field show different characteristics(in usage pattern and size)
 - Based on size and/or usage patterns and/or last access time
- Small file issues
 - Compress or relocate?