Working with Problematic Nodes

Jesse Hanley
Sarp Oral
Neena Imam

January 2017
Overview

• Problematic clients within a Lustre environment
• How to identify the client
• Responding to these issues
• Resolving the issue (long-term)
Noticing the problem

• Monitoring alerts
• Performance metrics
• User reports
• More data == More potential understanding of the issue
Starting from the Servers

- Request history
- Adjustable buffer
- Can view recent requests

```
[root@mds ~]# lctl get_param mds.MDS.mdt.req_history | tail -n 15000 \
> | awk -F':' '{print $3, $8}' | sort | uniq -c | sort -rn
14861 12345-192.168.1.8@o2ib 0s(-317s) opc 101
  56 12345-192.168.1.8@o2ib -1s(-318s) opc 101
   21 12345-192.168.1.95@o2ib 0s(-317s) opc 1102
   20 12345-192.168.1.90@o2ib 0s(-317s) opc 1102
   20 12345-192.168.1.112@o2ib 0s(-317s) opc 1102
   20 12345-192.168.1.65@o2ib 0s(-317s) opc 1102
    1 12345-192.168.1.8@o2ib 1s(-316s) opc 101
    1 12345-192.168.1.8@o2ib 0s(-317s) opc 400
```
MDS Trace

• Another option is to view trace information
• Tool in the process of being open-sourced for release

```bash
#lctl set_param debug="+rpctrace";
#lctl dk > /dev/null;
#sleep 60;
#lctl dk > /tmp/rpctrace;
#lctl set_param debug="-rpctrace";
# grep "procesed" /tmp/rpctrace
```
Job Stats

[root@mds1 ~]# for i in open mknod close; do
    echo "Operation: ${i}";
    lctl get_param *.*.job_stats | tail -n +3 | grep -e "job_id:" -e "${i}:" | \paste - - | column -t | sort -rnk 7 | head -n 5; done

Operation: open
- job_id: Cluster1-805873
  open: { samples: 223420362, unit: reqs }
- job_id: Cluster1-905871
  open: { samples: 218590178, unit: reqs }
- job_id: Login1
  open: { samples: 106275731, unit: reqs }
- job_id: Cluster2-13052
  open: { samples: 35191949, unit: reqs }
- job_id: Cluster3-298778
  open: { samples: 2652679, unit: reqs }

Operation: mknod
- job_id: Cluster1-805873
  mknod: { samples: 11830192, unit: reqs }
- job_id: Cluster1-905871
  mknod: { samples: 6254144, unit: reqs }
- job_id: Login1
  mknod: { samples: 289857, unit: reqs }
- job_id: Cluster2-13052
  mknod: { samples: 224907, unit: reqs }
- job_id: Cluster3-298778
  mknod: { samples: 63118, unit: reqs }

Operation: close
- job_id: Cluster1-805873
  close: { samples: 212332472, unit: reqs }
- job_id: Cluster1-905871
  close: { samples: 211580169, unit: reqs }
- job_id: Login1
  close: { samples: 106275721, unit: reqs }
- job_id: Cluster2-13052
  close: { samples: 34925470, unit: reqs }
- job_id: Cluster3-298778
  close: { samples: 2633952, unit: reqs }
<table>
<thead>
<tr>
<th>Children</th>
<th>Self</th>
<th>Shared Object</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.48%</td>
<td>0.01%</td>
<td>[kernel]</td>
<td>[k] system_call_fastpath</td>
</tr>
<tr>
<td>6.34%</td>
<td>6.28%</td>
<td>libc-2.12.so</td>
<td>[.] _int_malloc</td>
</tr>
<tr>
<td>6.27%</td>
<td>0.09%</td>
<td>[kernel]</td>
<td>[k] ll_file_io_generic</td>
</tr>
<tr>
<td>5.95%</td>
<td>0.01%</td>
<td>[kernel]</td>
<td>[k] cl_io_loop</td>
</tr>
<tr>
<td>4.28%</td>
<td>0.01%</td>
<td>libc-2.12.so</td>
<td>[.] __read_nocancel</td>
</tr>
<tr>
<td>4.26%</td>
<td>0.01%</td>
<td>[kernel]</td>
<td>[k] sys_read</td>
</tr>
<tr>
<td>4.23%</td>
<td>0.02%</td>
<td>[kernel]</td>
<td>[k] vfs_read</td>
</tr>
<tr>
<td>4.17%</td>
<td>0.01%</td>
<td>[kernel]</td>
<td>[k] ll_file_read</td>
</tr>
<tr>
<td>3.86%</td>
<td>0.01%</td>
<td>[kernel]</td>
<td>[k] ll_file_aio_read</td>
</tr>
</tbody>
</table>
Underlying Connectivity

[root@mds1 ~]# ibqueryerrors -s PortXmitWait

Errors for 0x78a0eb5e92b5 "Switch GUID=0x00078a0ebe5e92b5"
  GUID 0x78a0eb5e92b5 port ALL: [SymbolErrorCounter == 65535] [PortRcvErrors == 20] [LocalLinkIntegrityErrors == 2]
  GUID 0x78a0eb5e92b5 port 18: [PortRcvErrors == 20] [LocalLinkIntegrityErrors == 2]
  GUID 0x78a0eb5e92b5 port 22: [SymbolErrorCounter == 65535]
  GUID 0x78a0eb5e92b5 port 24: [SymbolErrorCounter == 65535]
  GUID 0x78a0eb5e92b5 port 25: [SymbolErrorCounter == 65535]
  GUID 0x78a0eb5e92b5 port 27: [SymbolErrorCounter == 65535]

[root@sns-mds1 ~]# ip -s link show eth0
2: eth0: <BROADCAST,MULTICAST,SLAVE,UP,LOWER_UP> mtu 9000 qdisc mq master bond0 state UP qlen 1000
  link/ether 00:00:00:00:00:00 brd ff:ff:ff:ff:ff:ff
  RX: bytes packets errors dropped overrun mcast
    1003357586616 1726131960 0 0 0 0
  TX: bytes packets errors dropped carrier colls
    130365111489692 1823776347 0 0 0 0
Proactive Options

- perf
- strace
- Process IO information: /proc/${PID}/io
- Darshan
Conclusion

• Which client is causing the issue
• What is the client doing
• Fix the immediate issue
• Work to keep it from happening again
Resources

- http://www.brendangregg.com/perf.html
- https://www.olcf.ornl.gov/kb_articles/darshan-basics/
- https://linux.die.net/man/8/ip
- https://linux.die.net/man/8/iblinkinfo
This work was supported by the United States Department of Defense (DoD) and used resources of the Computational Research and Development Programs at Oak Ridge National Laboratory.