



SC11 TORQUE Update

Ken Nielson and David Beer

GET TORQUED
TORQUE 4.0

Why TORQUE 4.0

TORQUE is a descendant of PBS (Portable Batch System) developed 1993-1994

- What was happening with computers in 1993 and 1994
 - 1993 - Commercial providers allowed to sell Internet connections to individuals
 - 1993 - Intel releases P5-based Pentium processors with 60 MHz and 66 MHz versions
 - 1993 - Novell purchased Digital Research, DR-DOS
 - 1993 - Windows NT 3.1 released which supported 32-bit programs
 - 1993 – MS-DOS 6.0 released
 - 1994 – Intel releases 90 MHz and 100 MHz Pentium Processors
 - 1994 – Motorola releases the 68060 processor
 - 1994 - Linus Torvalds releases version 1.0 of Linux Kernel.
 - 1994 – IBM releases PC-DOS 6.3

Why TORQUE 4.0

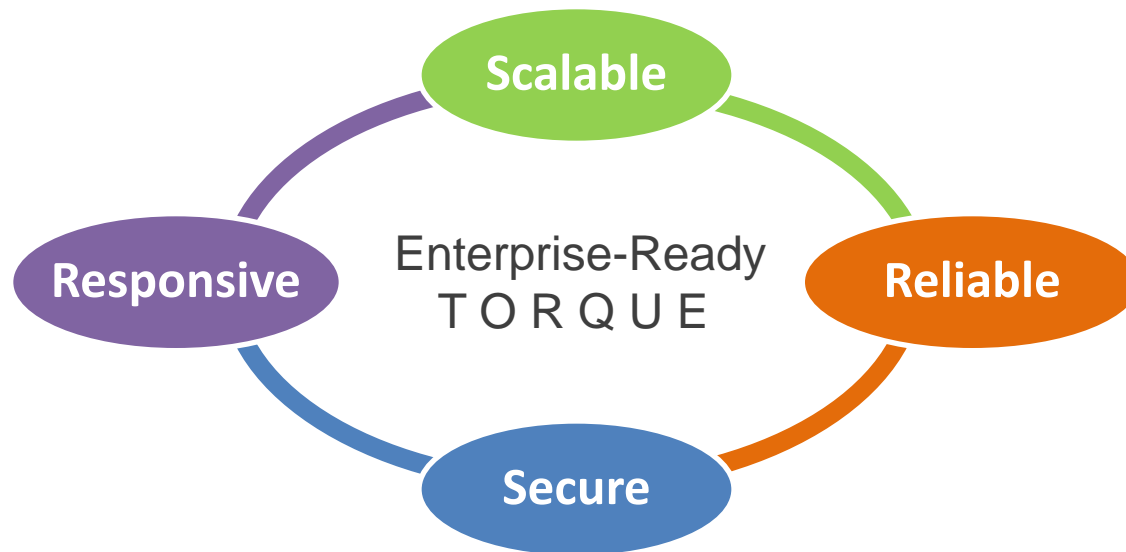
Program practices of 1993-1994

- Memory limited
 - Keep code size small
- Slow and unreliable network connections
- No threads for Unix or Linux
- Disk storage slow



TORQUE 4.0 Goals

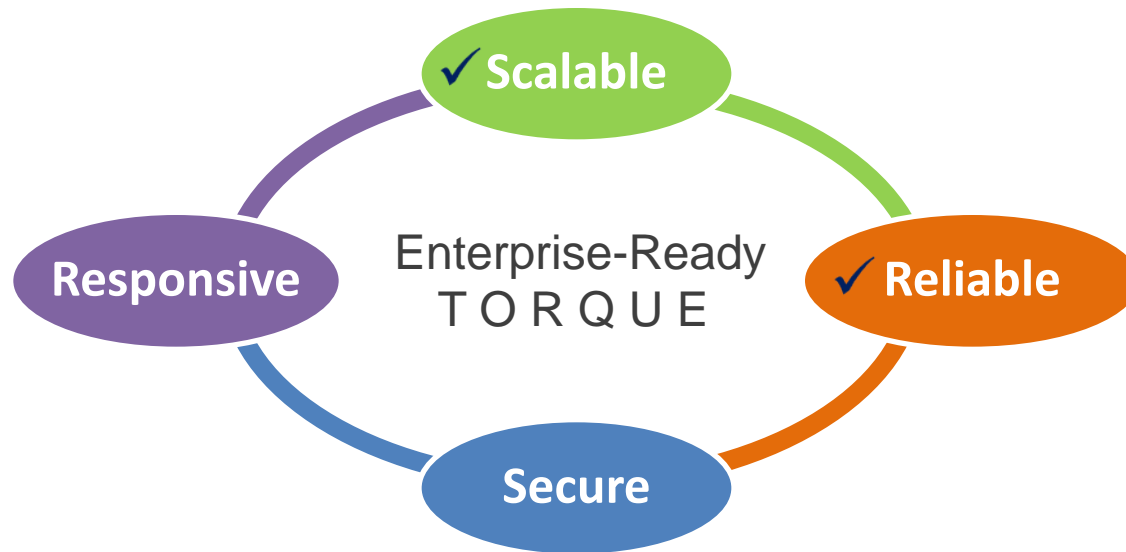
- Scalability
 - Support 10s of thousands of hosts in a cluster
 - Support jobs using 10s of thousands of hosts
- Improved Communications
- Higher Throughput and Response
- Improved Reliability
- Enhanced Security



TORQUE 4.0 Advancements

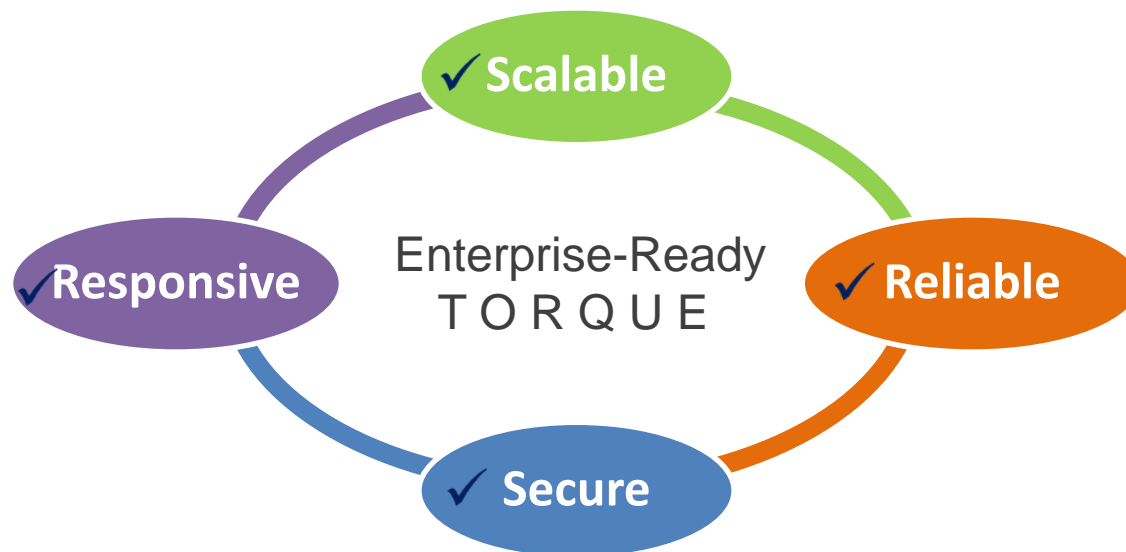
- Extended scalability for petaflop and beyond
 - Up on 7510 hosts
 - Reduce network overhead and bottlenecks
- All TCP-based communication improves reliability

```
PBS_Server;Svr;PBS_Server;LOG_ERROR::stream_eof,  
connection to kmn is bad, remote service may be  
down, message may be corrupt, or connection may  
have been dropped remotely (Premature end of  
message). setting node state to down
```



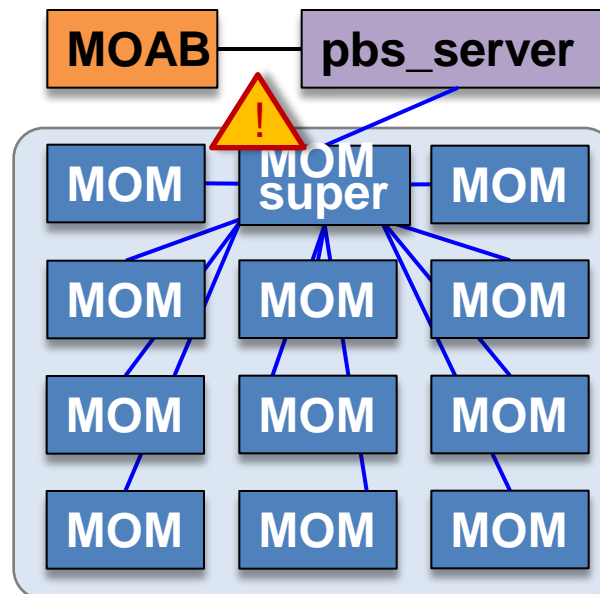
TORQUE 4.0 Advancements

- Multi-threading improves response and reliability
 - Improved responsiveness, no waiting for qstat
 - increases job through-put
 - Slow transactions do not stop system
 - Optimized internal algorithms
- Enhanced control and security over users



Job Radix

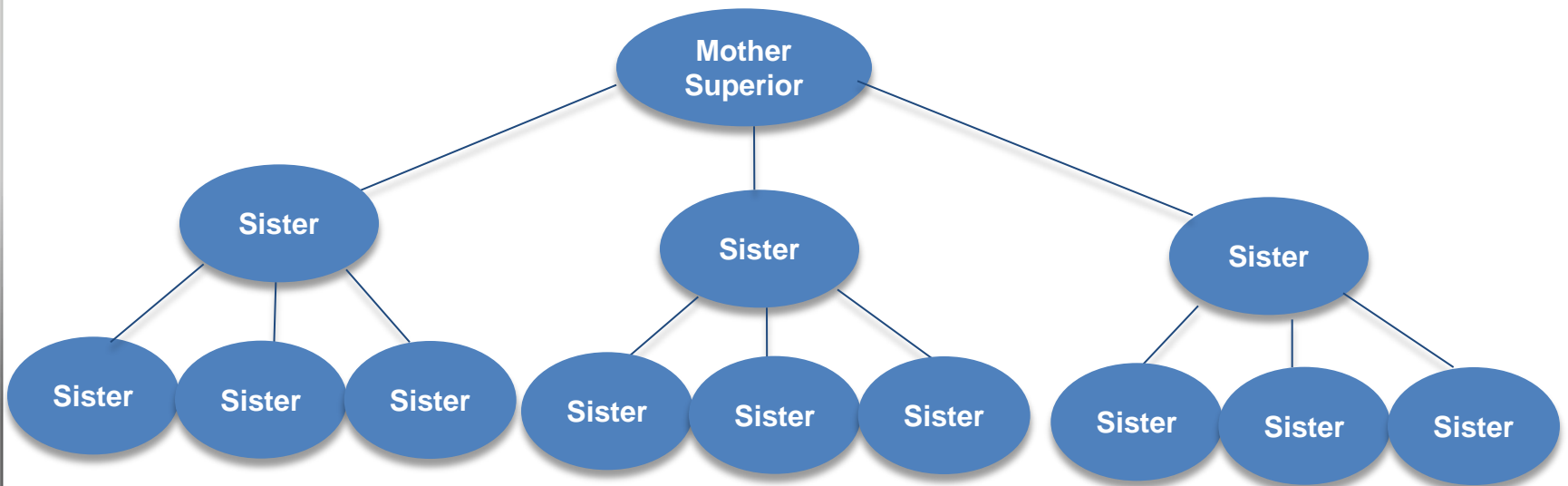
- Problem: all sister nodes communicate directly with Mother Superior
 - Causes communication failures when a node becomes saturated
 - Jobs are needlessly lost
 - Communication bottlenecks at single point



Scalability – Job Radix

- Job Radix Solution: **\$ qsub script.sh -W job_radix=3**
 - States that each MOM should communicate directly with 3 other MOMs
 - No longer a single stress point, each node communicates with a maximum of 3 other MOMs

Scalability – Job Radix=3



Enhanced Client and Security Management

- New authorization daemon (trqauthd) gives administrators better control over users
 - Eliminated security loophole
 - Users cannot run jobs as another user
 - trqauthd runs as “root” (replaced pbs_iff with “sticky” bit)



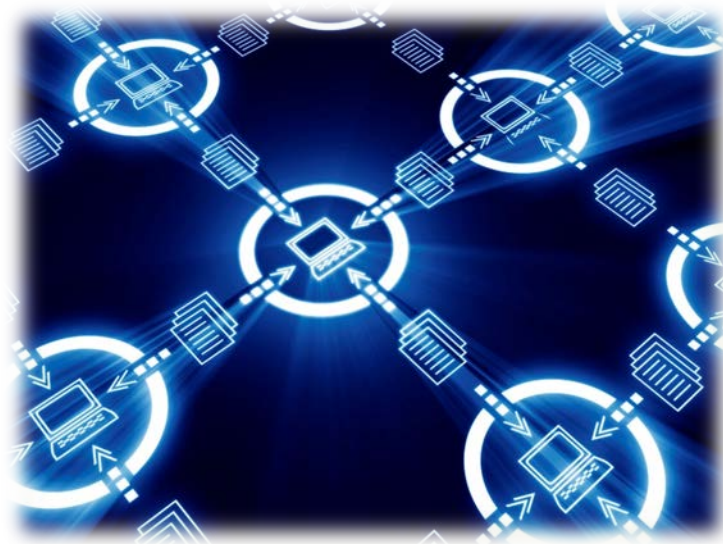
MOM Hierarchy

- Goals
 - Improve efficiency of MOM to server update communications
 - Reduce network congestion
 - Improve cluster status reliability
 - No more MOMS going down when nothing is wrong



MOM Hierarchy

- <TORQUE_HOME>server_priv/mom_hierarchy
 - Specifies where each MOM should send status updates. The updates get propagated from there to pbs_server.
 - Self-healing – Describes how retries should happen when the preferred connection can't be obtained.
 - Requires the server to process larger amounts of data, but with a lower number of connections, resulting in better performance.
 - Administrator can set up hierarchy to mirror actual network topology, resulting in optimum performance.



MOM Hierarchy

Sample file:

```
<path>
```

```
  <level>node1,node2,node3</level>
```

```
  <level>node4,node5,node6</level>
```

```
  ...
```

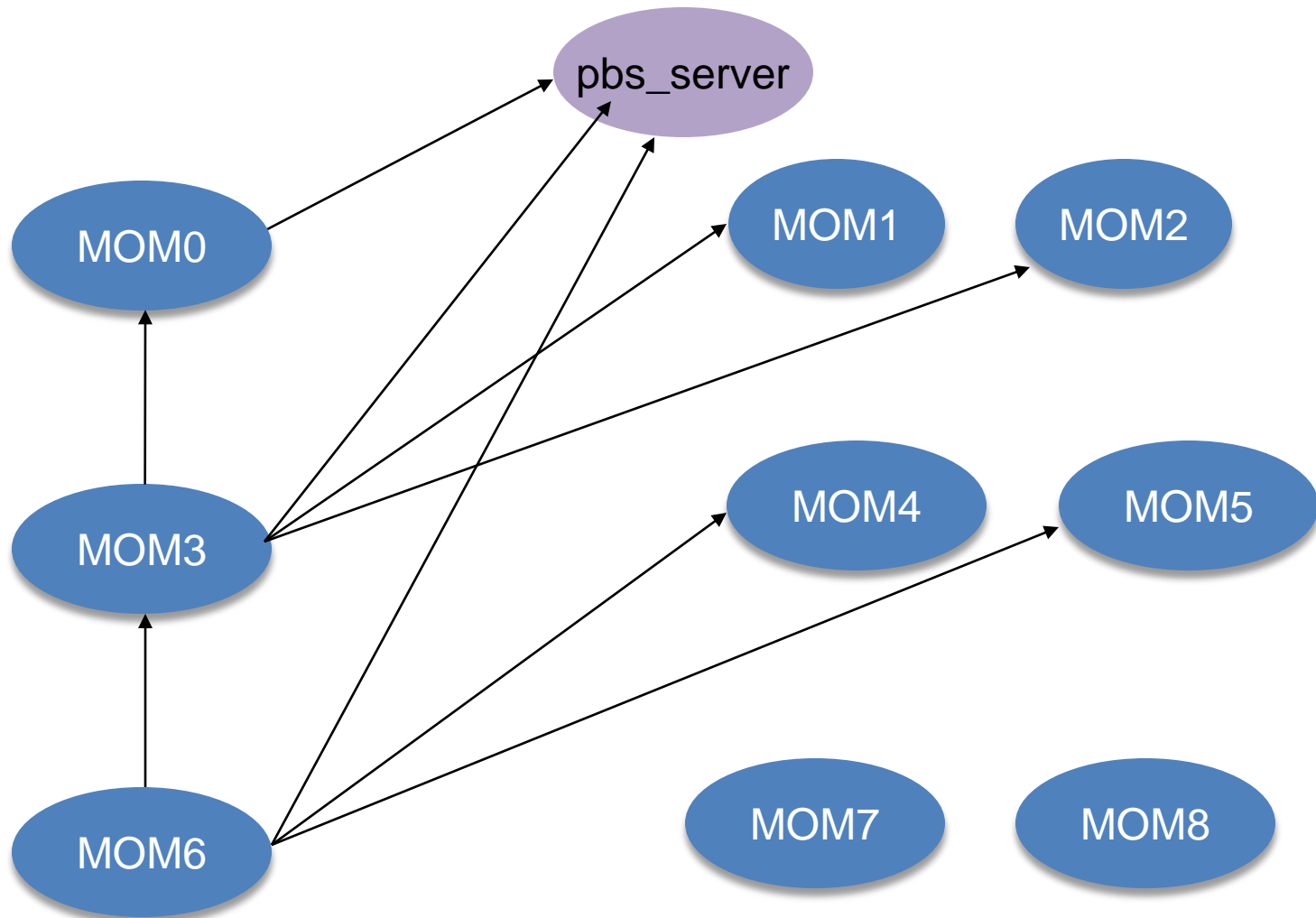
```
</path>
```

```
<path>
```

```
...
```

Note: You can specify multiple paths. Not all nodes need to be in all paths but nodes can be in multiple paths if desired.

MOM Hierarchy



Future Feature Wish List

- SMP-SMT and Job placement

TORQUE 4.0

- Scheduled for Release in Q1 2012
- We would love you to test it and give us feedback
- Download using subversion
 - `svn co svn://clusterresources.com/torque/trunk`

TORQUE 4.0

- Adaptive Computing representatives are floating around to answer any questions you may have
- Please ask...



FREE T-Shirt

Don't forget to scan your badge for your

FREE GET TORQUED T-Shirt



Available Only TODAY!!!



Adaptive

COMPUTING