Panasas NetValidator

Curtis Anderson – Software Architect

4/18/2018

Storage Performance Depends Upon Network Configuration

To feed the beast, both storage and networking need to be optimally configured

Where “optimal” for one must take into account “optimal” for the other as well

Both storage and networking, when viewed in isolation, look fine

But, there’s an “impedance mismatch” limiting performance, or a latent problem waiting to take production down when an innocuous failure happens

Failures are easy to find, limitations or bottlenecks or potentials are not

Yet everything needs to be running optimally to get the most out of the HPC cluster storage

Subtle Network Problems are Hard to Find

The MTU on an inter-switch-link is not quite big enough

It all works, but B/W and CPU load are worse because of packet fragmentation caused by link

Both of the cables in a LAG go to the same line card in a switch

It all works now, but if that line card fails, storage traffic comes to a halt

A routing problem has your HPC storage traffic going thru a bottleneck

No flags are up, but a new link got plugged in and effective throughput suddenly tanked

Networks constantly change: “as designed” was good, “as exists” maybe not

A reconfiguration affects more than it should have, when/who/how will detect that it did?

NetValidator – Validating the Network Against the Storage

Check the live network configuration against the needs of PanFS storage

Uses FreeBSD tools and SNMP from the PanFS realm outward toward the client systems

Develops a topology of just the switches and cables the PanFS traffic flows through

Applies AI rules to check for misconfigurations and mismatches between needs and actuals

Able to raise a PanFS GUI alert immediately after any cable PanFS traffic flows thru is affected

Flags recommendations and potential problems for human review

Just a diagnostic, all operations are read-only and repair is left to human judgement

Part of our focus on stability and low TCO while delivering performance

In “research mode” now, moving toward release at some point in the future