Storage Validation – Are we ready for the coming Data storm?

Amritam Putatunda
Technical Marketing Engineer
Ixia Technologies Ltd
Evolution of Internet

- Static HTML
- SQL over HTTP
- Web2.0
- APPS
Network Infrastructure Validation: A model that is reaching maturity

- With complexity the network infrastructure validation and optimization became at first challenging and later complex.
- However with continuous investment the network validations has gained sufficient maturity.

- Functional validations
- Performance Validations

**Pre Deployment**

**Deployment**
- Integration Validations
- Optimization Validation

**Post Deployment**
- Change Managements
- Security Validation
The two sides of the story

The Network Infrastructure – The Blue eyed boy

Storage Infrastructure – The proverbial prodigal son
Identification of Blocking Point

- Strong validation frameworks have been formed for network infrastructures
- Substantial efforts are taken to identify blocking points in networks.

- Storage Infrastructures validations have till now been restricted to basic go-no-go testing
- Has been expected to work!!!
There is a need to test network infrastructure

- With growing complexities in applications and ever changing threat landscapes, network validations became a must.

- Application delivery Controller
- Server Load Balancers
- Deep Packet Inspection
- Firewalls/IPS/IDS
- Data leakage prevention Device
- Application aware routers/switches

Are you sure that only networking infrastructures were evolving?
Storage Validations – Is there a reason?

- SCSI Over FC
- (i)SCSI over FCoE
- iSCSI over TCP

- NFS1.0
- NFS2.0
- NFS3.0
- NFS4.0
- webNFS
- NFS4.1

- SMB1
- SMB2.0
- SMB2.1
- SMB3.0

- FC-4G
- 8G
- 16G
- 10G
- 100G
- 400G
- Ethernet 1G
Hardware Innovation's – Were we lacking here?

- Scale out and Scale up NAS technologies
- Reduction in cost of storage through continues innovations
- Breakthroughs in flash and solid state drive technologies.
- Innovations in SCSI and iSCSI array architectures
- Advancement in multi threading, multicore technologies
- Advancements in processing and offloading technologies.
Is your storage ready for the upcoming data storm?

- Is the filer best optimized for the type of applications it will be servicing?
- Can it simultaneously handle millions of parallel access?
- Are the SCSI arrays doing the best job of handling redirects, MPIO, failovers while servicing the hundreds of VM clients?
- Will the big data analytics work most efficiently with this particular DFS cluster?
- Is my cloud application optimized for my customers?
- Is this the best fit for my environment?

I do not know, but most probably I will throw hardware to the problem.
Deployments – It's all about money

- A lack or proper and strong validation framework impacts deployments.
- Worry free change managements
- Absence of relatable ROI frameworks

Vendor A

2012 2014 2016

$ $$$$ $$$$$

Vendor B

$ $$ $$

Vendor C

$$ $$ $$
A small shift – A Big Change
Application data pattern emulation

Eliminates additional variables from the environment

Helps calculate per application max performance of storage

Identify and optimize the millisecond delays in this part of the world

Helps isolating issues as the buck literally stops here

Vendor selection procedures – Finds what's best for a particular environment
### Steps in Application Emulation - Classification

<table>
<thead>
<tr>
<th>Category</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database</td>
<td>SQL, Oracle</td>
</tr>
<tr>
<td>Virtualization</td>
<td>VDI workloads, BYOD</td>
</tr>
<tr>
<td>Generic Filer Access</td>
<td>NAS shared filer accesses</td>
</tr>
<tr>
<td>Cloud Applications</td>
<td>Cloud Hosted Applications</td>
</tr>
<tr>
<td>Datacenter</td>
<td>Applications specific to data centers</td>
</tr>
<tr>
<td>Enterprise</td>
<td>Financial, Business critical apps</td>
</tr>
<tr>
<td>Others</td>
<td>Proprietary and other unique applications</td>
</tr>
</tbody>
</table>
Steps in Application Emulation – behaviour analysis

- Application classifications are generally the easier part. The complexities are introduced while:
  - Understanding application specific file/block size distributions
  - Analyzing sequential and randomness patterns of IO
  - Inspecting application specific data patterns – Realism.
Steps in application emulation – Protocol and user behavior analysis

- Access controls and access behaviors
- Caching abilities
- Control Plane
- Locks and leases
- Queuing, efficiencies
- Multiple Channels
- Concurrency
- Reliability and Resilience
- Load patterns
- Dynamism in User behaviors
Case Study: Video over SCSI

- A video write pattern on a SCSI disk.
  - Understanding alignments as a video write generally happens over a fixed block size irrespective of data content.
  - This creates holes throughout the disk
  - To understand max scalability of a video server
Case Study – Predictable analyzable data generation.

- Distributed file system has given us an efficient way to store, analyze and retrieve information from structured and unstructured data.
- However analysis efficiency measurements need analyzable and realistic data patterns generation.
Case Study: Cloud application validations

- Every other application is being hosted on cloud.
- Application behavior over cloud is mostly a guessing game.
- Applications failing on cloud is quite common.
- Even well-known applications did beta testing with real users.
Conclusions

Networking technology adaption in market is higher due to its smooth deployments and rapid adaptability.

With proper validation framework similar agility can be expected in storage.

Application relatable data points- Helps in understanding and speaking customers language.

Perfect harmony in an environment is possible by proper co-ordination and sync between all its elements.