Intelligent Architecture for the Data-Driven Business
QoS: provisioning performance in multi-tenant cloud storage

Panelists:
Dr. Alea Fairchild, Director, The Constantia Institute
Jay Prassl, VP Marketing, SolidFire
Matthew Wallace, Director Product Development, ViaWest
Agenda

• Panel discussion:
  • Alea Fairchild: Performance issues in Cloud Storage QoS
  • Jay Prassl: Technology trends for QoS
  • Jason Carolan: Usage of QoS by CSPs

• Questions for the panel
Performance Issues in Storage QoS

• Short-term goals: Performance via integrating flash into its storage architectures

• Near term issues: How to make the performance that each application receives more predictable, taking into consideration management of the system's I/O queues and the relative allocation of flash and spinning disks.

• Next steps: Designate and assign a minimum IOPS rate and priority to each workload for performance.
QoS: Why do I need granular IOPS support?

- **Cost** – Compare SSD to standard disk in terms of performance, and then add management costs.
- **Differentiation** – Related to cost; service providers that can offer a more granular service stand more chance of getting business as the cloud computing landscape becomes more competitive.
- **Service-based** – Technology refreshes can mean retesting against new hardware; specifying against a service means hardware can be replaced/refreshed as long as it continues to deliver the same service level.
QoS and in-house SSD

• What is the need for QoS for all hybrid storage?
  – Unfortunately, even all-SSD systems are susceptible to the noisy neighbor problem, just at a different pressure level.
• As disks, and SSDs, keep growing, the inefficiencies of dedicating resources to workloads have become untenable.
• QoS controls at the storage systems should ideally implemented at the VM level
  – This will allow us to cram more workloads onto fewer resources while still delivering an appropriate level of performance to each.
QoS for CSP

Command and control
- Storage QoS will likely hold some appeal to cloud providers because the providers have even less control over what the virtual machines they host do than the managers of corporate data centers.

Business differentiation
- QoS mechanisms should also include not only performance floors but also ceilings, so cloud providers can offer different levels of performance at different prices...and prove it.
Business innovations using QoS for performance enhancement

• Disaster Recovery/ Workload orchestration
• Service definitions for backup, archive, and other storage classes
• Cross cloud mobility
Why don’t you run enterprise apps in the cloud?

• Current cloud architecture delivers inconsistent and variable performance (Effect of noisy neighbor)
• Inability to efficiently scale performance (storage sprawl)
• Unable to throttle performance independent of capacity
• Low levels of transparency (no visibility into systems)
• Perception of unreliability
Advancing the way the world uses the cloud.

**Applications**
- Test / Development
- Backup / Archive

**Performance Sensitive Apps**
- Oracle / SAP
- Hadoop / NoSQL

Application Value / Margin

- Low
- Med
- High

IOPS

Performance Sensitivity

Applications

- $$$$ (High)
- $$ (Med)
- $ (Low)
Cloud 2.0 – Where we are headed

- Evolution beyond Test/Dev drove 2nd phase of virtualization market growth
- Cloud market is at the inflection point, ready for production applications
IaaS – The Move towards Enterprise Apps

- Customers: SMBs, Developers
- Workloads: Non-Production, Test Dev, Public Cloud

Source: 451 Research

Focus:
Greater Traction with Mgt. Products & Vertical Offerings

Customers:
Enterprise Adoption & Usage

Workloads:
Production & Mission Critical Focus, Hybrid/Private Clouds

2011: $681
2012:
2013:
2014:
2015:
2016: $6,847

CAGR 49%

Source: 451 Research
Purpose Built for the Cloud

Quality of Service (QoS) is an Architecture

- All-SSD Platform
  - Deliver consistent latency for every IO
- True Scale-out Architecture
  - Linear performance gains as system scales
- RAID-less Data Protection
  - Predictable performance in any failure condition
- Balanced Load Distribution
  - Eliminate hot spots that create unpredictable IO latency
- Fine-grain Quality of Service Control
  - Guarantee volume performance
- Performance Virtualization
  - Deliver performance resources independent of capacity and on demand
The services stack perspective

- We see a foundation of extremely reliable Colocation that builds towards more complicated, high-touch services.
- Customers like to move up the stack.
- The more application and data management you do, the more likely you need to worry about storage performance.
24 Data Centers
450k Square Feet
55 MW UPS Power
15 Redundant NOCs

26 Unique Carriers

Robust Cloud & Managed Services
What Cloud Customers are Solving For

- Low Entry Costs / Shared Risk
  - Consumption-based utilization
  - Lower upfront costs
- Metered Costs provide Transparency
  - CFOs love this
- Ease of Control / Time to Market
  - Quick to enable Cloud services
  - No more phone calls, waiting
- Network Delivery
  - Generally Internet-based
  - Enables multiple devices, access technologies, anytime, anywhere
- Flexible and Elastic
  - Grow as I need, massive compute / storage for variable terms
  - Cost effective Disaster Recovery
- Leverage Expertise
  - How many security experts, database experts, etc does an Org have?
  - Hardened operational services – providers have excellent infrastructure / facilities!
  - Managed services moving to the cloud
Evolving into the Cloud with Enterprise Apps

Managed Servers & Colocation

Private Cloud

Community Cloud

Unified Fabric Backplane
Enterprise Apps in the cloud – What we see

If clouds can’t predict or smooth out I/O spikes the noisy neighbor crushes performance
Enterprise Apps in the cloud – What we see

Predictable I/O with consistent performance
No more noisy neighbor crushing database performance
Tackling Noisy Neighbors – Rate Limits

- Network not necessarily a great fit for storage
- Imperfect mapping – block size, protocols, compression
- Limiting access to an imprecise pool – how many IOPs does an array produce?

Image Credit: http://commons.wikimedia.org/wiki/File:Kitchen_Funnel.jpg
Tackling Noisy Neighbors – Priority

- You can sort access into lanes
- You can still have a traffic jam in any lane
- If the HOV lane gets overfull, it can just cut off people in other lanes
- “Trickle-down noisy neighbors”
Tackling Noisy Neighbors – Go Dedicated

• Performance you need
• All to yourself
• but...
• CapEx
• Agility – not cloudy
• Commitment
• Not necessarily predictable scaling

Image Credit: http://www.flickr.com/photos/30624565@N08/5000744753/
Tackling Noisy Neighbors – Tiering

• Moving workloads between faster and slower storage
• Can’t handle short-duration hotspots
• Noisy neighbor intra-tier
• Noisy neighbors displace other tenants to lower tiers

Image Credit: http://www.flickr.com/photos/vshioshvili/229207037/
So...
What do you need for reliable QoS?
Predictable disk performance

**Traditional Disk**
- Cache-hit vs Cache-miss
- Get stuck waiting on a spinning disk
- Variance

**SSD**
- Very predictable reads
- Write-anywhere with no effect on performance

Image Credit: http://www.flickr.com/photos/rjgmc28/6128959419/
Going Raid-less

• Bigger RAID sets adding data risk to traditional raid (bit rot without read-after write)

• Mitigating effect of drive failures

• Rebuilds – to quote one vendor
  – “With RAID 5/RAID50, read performance decreases by 30% and random write performance decreases by 50% when a set is degraded.”

• Leading to new data protection schemes
  – Mozy
  – Ring-based object store a la Swift
  – Solidfire Double-Helix
Level Scalability

- Don’t want to overprovision at the start
- Need to be able to scale up
- Can get caught up with many arrays or step functions

Image Credit: http://www.flickr.com/photos/40819389@N04/3757324473/
SLAs are hard

- SLA difficulty is an exponential function
- Multi-tenant SLAs are harder since tenants impact each other
- Therefore offering a very strong SLA for storage requires a huge amount of consistency, availability, and performance assurance
- Sample 2,000 IOPs Requirement – Pricing Exercise

- Traditional shared SAN - ~$4000/mo
- Dedicated Array - ~$4500-5000/mo
- SSD Array w/ QoS - $343/mo

Image Credit: http://www.flickr.com/photos/68751915@N05/6355336039/
Dr. Alea Fairchild
+32 495 228 150
Twitter: @Afairch
Skype: alea.fairchild
www.constantiainstitute.org

Jay Prassl
720-432-7141
Twitter: @jprassl
www.solidfire.com

Matt Wallace
720-891-1029
Twitter: @mattwallace
www.viawest.com
Intelligent Architecture for the Data-Driven Business