Storage Scalability in Hybrid Cloud and Multicloud Environments

Live Webcast
June 11, 2020
10:00 am PT
Today’s Presenters

Pekon Gupta  
Solutions Architect  
SMART Modular

Rob Cone  
Senior Principal Engineer  
Intel

Piyush Chaudhary  
Senior Technical Staff Member  
IBM
SNIA Legal Notice

- The material contained in this presentation is copyrighted by the SNIA unless otherwise noted.
- Member companies and individual members may use this material in presentations and literature under the following conditions:
  - Any slide or slides used must be reproduced in their entirety without modification
  - The SNIA must be acknowledged as the source of any material used in the body of any document containing material from these presentations.
- This presentation is a project of the SNIA.
- Neither the author nor the presenter is an attorney and nothing in this presentation is intended to be, or should be construed as legal advice or an opinion of counsel. If you need legal advice or a legal opinion please contact your attorney.
- The information presented herein represents the author's personal opinion and current understanding of the relevant issues involved. The author, the presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.

NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.
SNIA-At-A-Glance

- 185 industry leading organizations
- 2,000 active contributing members
- 50,000 IT end users & storage pros worldwide
What We Do

Educate vendors and users on cloud storage, data services and orchestration

Support & promote business models and architectures: OpenStack, Software Defined Storage, Kubernetes, Object Storage

Understand Hyperscaler requirements. Incorporate them into standards and programs

Collaborate with other industry associations
Agenda

- Hybrid Cloud Multicloud Environments
  - Different ways to leverage the cloud
  - Data placement across multiple clouds
  - Examples of Microservices in the cloud

- Hybrid Cloud Multicloud Storage Requirement
  - Multi-cloud storage scaling
  - Security and data protection
  - Managing complexity
  - Data virtualization
Digitally Transform - Today

60%

of global GDP will be digitized by 2022, with growth in every industry driven by digitally enhanced offerings¹

Rapid, disruptive change is the new normal. Laying a modern foundation is critical and storage is a major part of this transition

Different Ways to Leverage Cloud

1. Enterprise Today
2. Lift and Shift
3. Integration of Microservices

- On Demand Cloud Storage / Backup
- Migrate Enterprise Applications to Cloud
- Deeply Integrate Cloud Architecture and Framework
Multiple Clouds Drive Business Success

Cloud architecture is the most efficient and agile computing solution, regardless of where it is deployed.
Applications are Being Re-Architected

The cost of data movement, latency and functionality determines where data is placed.
The ability to use multiple CSP can help further optimize enterprise solutions, but at a cost of complexity.
Netflix Transition from Optical Disk to Largest Hybrid Cloud

Netflix Apps

Content Cache

Content Management and Netflix Studios

Netflix Use of Hybrid SW Evolution

Multiple Years to rewrite their SW and fully migrate to cloud

Built on microservices to allow them independent development

https://netflixtechblog.com/tagged/cloud-architecture
Netflix Hybrid Built on Microservices

Purpose built API’s allow Microservices to change without effecting each other.

Each microservices access and uses data in a variety of mechanism
• S3, Hadoop, Rocksdb, etc
Generic Example - Microservices Building Blocks

East-to-West Processing
- Machine learning
- Inference
- Databases
- Big Data
- Caching
- Routing
- Metadata

North-to-South Processing

Microservices
- Advertisements
- Containers
- Customer data
- APIs
- Update Inventory
- Connect Customer
- Allocate
- Run Analytics
- Local Pricing
- Error mgnt
- Administration

Kubernetes

docker

Results
- Interactive Analytics
- Alerts
- Enterprise Applications
- Security
- Better SLA's
- Databases
- Smart Ads
- Secure Data
- File object, block

Databases
Big Data
Data Lake
Performant Storage

Generic Example - Multicloud Architecture with Microservices

Modularity allows for better placement and efficiency of microservices to provide better SLA’s.
Multi-cloud storage scaling

- Cloud deployments provide infrastructure agility
- Public clouds can provide on-demand nearly unlimited scaling
- API driven deployments improve speed and developer productivity
- As storage capacity scales so does the need for performance
- Reduces time-to-solution and allows for rapid innovation
- Customers already need Exabyte+ capacity and 1+ million volumes
- Applications need hundreds of GB/s and millions of IOPs of filesystem performance
- Allows rapid deployment, scaling and de-commissioning of storage
Enterprise Applications Need an Ecosystem

- Native encryption & secure erase
- Backup & archive
- Authentication, authorization & audit
- Regulatory compliance
- Data privacy
- Agile deployment & elasticity
- Movement of data
- Data ingest & egress
- Compute & data workflows
- Data privacy
Storage Requirements in a Hybrid Multicloud

- **Security**
  - Authentication
  - Authorization
  - Audit

- **Data Protection**
  - Data-at-rest
  - Data-in-motion
  - Secure erase

There is a need to extend trust beyond the confines of a single application or data center. Guaranteeing security and data protection in a hybrid multicloud needs strong open standards and collaboration.
Storage Requirements in a Hybrid Multicloud

- **Data Privacy**
  - European General Data Protection Regulation (GDPR)
  - California Consumer Privacy Act (CCPA)

- **Backup & Archive**
  - Replication is **not** backup
  - Needs to address regulatory requirements
  - Immutability and WORM

Data privacy cannot be an afterthought anymore. It's at the forefront of customers’ requirements and CSPs and storage solutions need to address this urgently.
Data Workflow Automation for Hybrid Multiclouds

- Deploy an enterprise data plane across the hybrid multicloud infrastructure
  - Ensure requisite data is available when needed with appropriate security
  - Apply enterprise policies across all storage infrastructure

- Need to automate the movement of data between on-premises and clouds
  - Must account for data gravity
  - Proper safeguards needed for security and egress costs

- Needs to be integrated into compute workflows
  - Simulation, HPC, technical computing
  - Analytics, artificial intelligence
How to Manage All this Complexity?

- **Service catalogs**
  - Curated services and APIs approved for use within an enterprise
  - Ensures security and compliance requirements are met while improving agility
  - Hides complexity and delegates authority to business units

- **CSI & storage operators**
  - Enterprizes are adopting container platforms aggressively
  - Container Storage Interface (CSI) provides a unified storage interface for various Container Orchestrator Systems (COs)
  - Storage operators can automate the deployment and interaction with the CSI drivers
  - Automate the provisioning, backup, cloning, snapshots, grow and shrink volumes, etc.
Data Virtualization

Data virtualization is an approach to data management that allows an application to retrieve and manipulate data without requiring technical details about the data, such as how it is formatted at source, or where it is physically located, and can provide a single customer view (or single view of any other entity) of the overall data. -- Wikipedia (link)

- Data gravity dictates that data isn’t moved if possible
- Composable applications are given access to data as needed
- Workload is pushed to where data is stored
- Minimizes data copies and the resulting entropy
- Preserves a single “source-of-truth” for the enterprise
Summary

- Cloud infrastructure provides the agility, flexibility and scaling needed to meet the explosive storage growth requirements.
- Data security, data privacy and data protection concerns are inhibiting the adoption of hybrid multicloud deployments by enterprise applications.
- Extension of the enterprise on-premises data plane to the multicloud should eliminate barriers for broader adoption of hybrid multicloud infrastructure by complex enterprise workflows.
- Service catalogs and data virtualization can be key tools to deal with the complexity of hybrid multicloud infrastructure.
- Investing in microservices can help optimize workloads, gain better efficiency in a multicloud environment.
SNIA Resources

- Webcasts:
  - Kubernetes in the Cloud (Part 1)
  - Kubernetes in the Cloud (Part 2)
  - Kubernetes in the Cloud (Part 3) Stateful Kubernetes

- Upcoming Webcast: June 23, 2020: A Multi-tenant Multi-cluster Kubernetes Datapocalypse is Coming
After This Webcast

- Please rate this webcast and provide us with feedback
- This webcast and a copy of the slides will be available at the SNIA Educational Library [https://www.snia.org/educational-library](https://www.snia.org/educational-library)
- A Q&A from this webcast will be posted to the SNIA Cloud blog: [www.sniacloud.com/](http://www.sniacloud.com/)
- Follow us on Twitter @SNIACloud
Thank you!

Q & A