HITACHI Inspire the Next

# Object Storage: An Intelligent Content Hub

#### PosAm TechDays February 2021

Andrej Gursky Solutions Consultant, Hitachi Vantara

© Hitachi Mantara LLC 2020. All Rights Reserved

#### Data is the New Oil

- Applications have evolved and require performant, scalable and secure storage to support them
- Frozen data is of no value to you

You need a modern solution to identify, utilize and monetize your data







#### Large cloud workloads look promising



- Customer organizations are bringing together data from across the businesses to:
- Make it available for their analytics pipeline
- Support better and make timely business decisions
- Develop and deliver new, revenue-generating initiatives faster

HITACHI

**Inspire the Next** 

#### But data growth remains the biggest challenge





access to data, they shouldn't be able to see

HITACHI

**Inspire the Next** 

# Ripple effects of data growth on large cloud workloads



Designed for high-performance, real-time analytics at a large scale

Becomes challenging and expensive to manage with added compute and storage capacity

Impacts on-time delivery of new services to the business

Additional resources become costly and wasteful



# Conventional approaches of handling data growth





#### Accessing just 10% of a 3PB public cloud data lake every month -> Access costs of around \$1M dollars per year



You require an approach that allows you to optimize and scale cloud workloads, while also simplifying management and ensuring that data is always available to the analytics pipeline.

HITACH Inspire the Next

#### Better way: Seamless integration of object storage HITACHI into cloud workloads

Seamlessly integrate smart object storage capabilities into the existing infrastructure, thus making object storage an integral part of the data lake fabric itself



Unified, efficient management of data across the entire data lake

Support for on-premises, cloud, and multicloud storage strategies

Specific capabilities to optimize data lakes built in cloud-native environments, such as Hadoop and Splunk



# Hitachi Content Platform (HCP) Object Storage from Hitachi

## What is Object Storage?

- Aggregate, manage, protect, and use content
- Just like we move photos from devices to a PC...
  - Simplify management and use
  - Software organizes, protects, shares photos
- IT uses object stores for similar reasons
  - Consolidate and organize
  - Protect and preserve
  - Manage and search







## **Comparison to Block Storage**

- Bucket of bits
- No meaning attached
  - One-bit string no different from the next
  - No data intelligence in storage
- Operations are against gross collections of bits, without knowledge of what the bits represent to the organisation.



### **Comparison to Block Storage**

- Objects take form
  - Not just a string of bits
- From storage subsystem to software system – called an object store
- Object storage itself has knowledge about data
- Now able to perform complex functions against the data



HITACH

**Inspire the Next** 

#### What is an Object?







#### System Metadata

- Filename: NZ219983.JPG
- Created: January 4, 2012 ٠
- Last modified: March 23, 2012

#### **Custom Metadata**

- Subject: Castle Point Lighthouse
- Place Taken: New Zealand
- Category: Travel
- Allow sharing: Yes

#### What is Metadata?

- Describing the object •
  - Help you to find the right one
  - Tells you what it is
  - The specifications
  - Used where and when
  - Access permissions
- Any and all objects •
  - Different attributes per object
  - Adding attributes later







#### Think of Metadata Like the Label on a Can

- End-user self service
- Automation
- Regulatory compliance
- Marketing
- Disposition
- Procedures
- Branding



# All this metadata about a can of soup!

Imagine what can be said about a picture, movie, contract, or report...

HITACH

**Inspire the Next** 

#### **The Power of Metadata**

Over time, metadata can be more valuable than the data itself

- What if data could describe itself?
- Think of how that could:
  - Simplify management
  - Facilitate protection
  - Serve as automation
  - Reduce costs
  - Mitigate risk
- That's the power of an object

- I was created by John Doe on January 1, 2011.
- Only the finance department can look at me.
- I have not been accessed for 6 months.
- Save your money, I don't need to be replicated.
- I contain information about a merger, so you should retain me for 10 years in case you are investigated.
- Delete me on January 1, 2021 to free up some storage space and reduce legal exposure.





Hi there, I am a

### **Object Storage**



#### By design: Self managing, self monitoring, self healing

- Intelligence
- Scale out
- Ultra-low touch
- Long-term online data storage
- Optimized for unstructured data
- Global access-ready



## **HCP Scalability**

- 80 nodes
- 160 racks
- >1EB per site
- >6EB (Geo-EC)
- S11: max 3.2PB per node
- S31: max 15.1PB per node
- >100B objects per system





#### **HCP Deployment Options**



HCP VM (vmware & KVM)



HCP Appliance



HCP for Cloud Scale



#### **Unrivaled Security and Protection**





Data Protection and Compliance

Access, Security, Efficiency

© Hitachi Vantara LLC 2020. All Rights Reserved.



#### **Unrivaled Data Resiliency**

#### **HCP is Backup Free**





- 1 Million times more available than AWS S3
- Exceptional data protection

## **Multitenancy: One Platform for All Data**



- Tenant segregates
  management
- Namespace segregates space
- Up to 1000 tenants and 10000 namespaces

Example of Namespace adress https://t1n1.t1.hcp.posam.sk https://t1n2.t1.hcp.posam.sk https://t2n3.t2.hcp.posam.sk

#### Example of object representation

https://t5n1.t5.hcp.posam.sk/rest/images/dc-bratislava.jpg https://t5n1.t5.hcp.posam.sk/hs3/images/dc-bratislava.jpg https://api.t5n1.t5.hcp.posam.sk/swift/v1/t5/t5n1/images/dc-bratislava.jpg



## Lower Costs And Reduce Total Cost Of Ownership Inspire the Next





# Hitachi Content Platform Use Cases

© Hitachi Vantara LLC 2020. All Rights Reserved.

### **HCP: The Digital Swiss Army Knife**





#### Primary Use Cases for HCP

#### HITACHI **Inspire the Next**





# WORKLOAD OPTIMIZATION USE CASE: Hadoop

© Hitachi Vantara LLC 2020. All Rights Reserved.

#### **Big Data Becomes A Big Mess**





Do you want to be responsible for not delivering the information that your business needs?

#### What is the Problem with Hadoop?



#### What is the Problem with Hadoop?





- Hadoop is economical and efficient when it contains active data, and compute is properly utilized.
- As data ages, access frequency decreases, resulting in "cold" data
- Cold data results in idle compute
- Nodes are added to increase storage capacity for new data (+ CPU, RAM, power, cooling, rack space, etc.)
- This process continues leading to a wasteful and expensive imbalance of compute and storage.
- Users of Hadoop estimate that 60-80% of their HDFS data is cold.



"Decoupling compute and storage is proving to be useful in big data deployments. It provides increased resource utilization, increased flexibility, and lower costs."

(IDC, "Five Benefits of Decoupling Compute and Storage for Big Data Deployments")

#### What About Public Cloud?

HITACHI Inspire the Next



Remote Storage Target

- · Public cloud gets expensive as you continue to store and access the data over time
- Whose issue is it when you lose data?
- How do you handle an application that becomes unavailable?
- How do you ensure compliance and handle regulatory audits?
- How will it perform and cost at scale?

32

## **Offload Methods are not Seamless**





- Offloading data with S3A removes files from HDFS
- The application layer must first copy data from HDFS to S3A
- Applications must then delete copied files from HDFS to free storage resources

## **Offload Methods are not Seamless**





- Application changes are required to access offloaded data.
- Data is no longer accessible via HDFS
- Data is read from S3A for which independent access controls must be implemented

#### Lumada Data Optimizer for Hadoop





- Seamless management and access to tiered data via HDFS with no disruptions
- Data management and security remain with HDFS, so no changes required to security practices
- No need to change data paths or configurations to access tiered data
- Highly-available fault tolerant design protects against volume failures

## Purpose-built for Hadoop Distributed File System

#### **Reduce Hadoop Total Cost Of Ownership**





#### What's your saving potential?



WINTER CORPORATION

HITACHI

**Inspire the Next** 



RICHARD WINTER FEBRUARY 2020

*For Hitachi Vantara Account Executives* DO NOT DISTRIBUTE EXTERNALLY

> WinterCorp www.wintercorp.com 42 DERBY LANE, TYNGSBORO, MA 672-695-1800

EXPERTS IN ANALYTIC DATA MANAGEMENT AT SCALE



# Why HCP?

© Hitachi Vantara LLC 2020. All Rights Reserved.

#### **Dedicated Solutions Validation Team**



#### Numerous Solutions with ISV Partners

- Hitachi product expertise
- 80 ISV partners, and growing
- 200 certified applications, and growing
- All major industries
- Proven joint partner solutions
- <u>Comprehensive compatibility guide</u>



#### **Real World Success**



© Hitachi Vantara LLC 2020. All Rights Reserved.

HITACHI

41

# Thank You

HITACHI Inspire the Next

## HITACHI Inspire the Next