AWS INNOVATE

AWS Storage & Data Migration

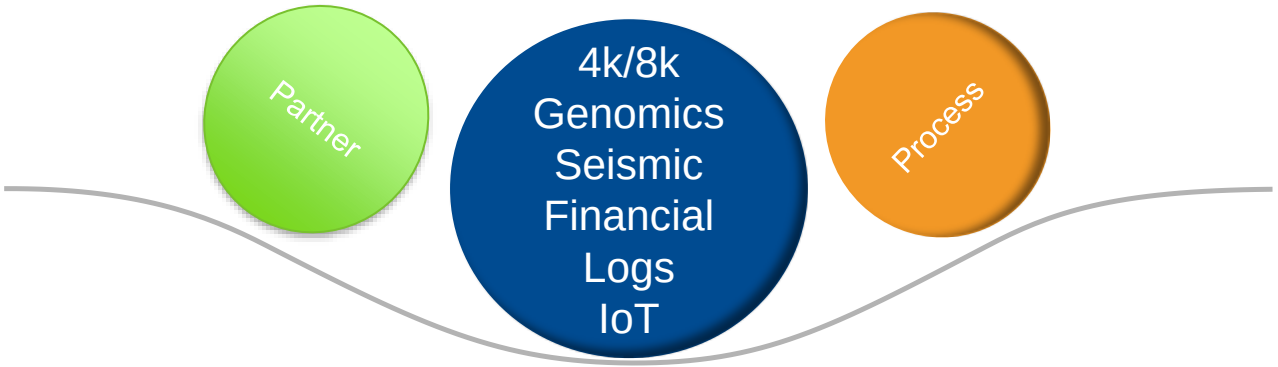
Matt Nowina – AWS Solutions Architect

May 10, 2017

Data has gravity

AWS INNOVATE

…easier to move processing to the data



amazon web services | intel


ARE YOU SURE THIS IS HOW WE GET DATA INTO THE CLOUD?
IT

The AWS Storage Portfolio

AWS INNOVATE

Amazon EBS
(persistent)

Amazon EC2
Instance Store
(ephemeral)

Block

Amazon EFS

File

Amazon S3

Amazon Glacier

Object

Cloud Data Migration

Snow* data
transport
family

Storage
Gateway

Direct
Connect

3rd Party
Connectors

Transfer
Acceleration

Kinesis Firehose

amazon web services | intel

The AWS Storage Portfolio

AWS INNOVATE

Amazon EBS
(persistent)

Amazon EC2
Instance Store
(ephemeral)

Block

Amazon EFS

File

Amazon S3

Amazon Glacier

Object

Cloud Data Migration

Snow* data
transport
family

Storage
Gateway

Direct
Connect

3rd Party
Connectors

Transfer
Acceleration

Kinesis Firehose

amazon web services | intel

AWS Data Transfer Services

AWS INNOVATE

Moving large batches offline	AWS Snowball & Snowmobile
Edge computing	AWS Snowball Edge
Augmenting on-prem with cloud	Amazon EFS & AWS Storage Gateway
Using a dedicated network	AWS Direct Connect
Integrating existing software	3rd party connectors
Moving over long distances	S3 Transfer Acceleration
Streaming data	Amazon Kinesis

amazon web services | intel

AWS INNOVATE

The Snow Family:

large batches or “edge” scenarios

amazon web services | intel

AWS Snow Family

AWS INNOVATE

Snowball

Petabyte-scale
data migration

Snowball Edge

Compute & Storage for
Hybrid/Edge workloads

Snowmobile

Exabyte-scale data
migration

amazon web services | intel

AWS Snowball

AWS INNOVATE

Petabyte-scale data transport

E-ink shipping label

- Rugged 8.5G impact case
- Rain and dust resistant
- Data encryption end-to-end
- 80 TB capacity/10G network

amazon web services | intel

How Snowball moves data

AWS INNOVATE

Create a job

Connect the Snowball

Copy data to the Snowball

Your data moved to Amazon S3

In transit to you

Delivered to you

Delivered to AWS

At AWS

Job created

Job completed

amazon web services | intel

AWS Snowball Edge

AWS INNOVATE

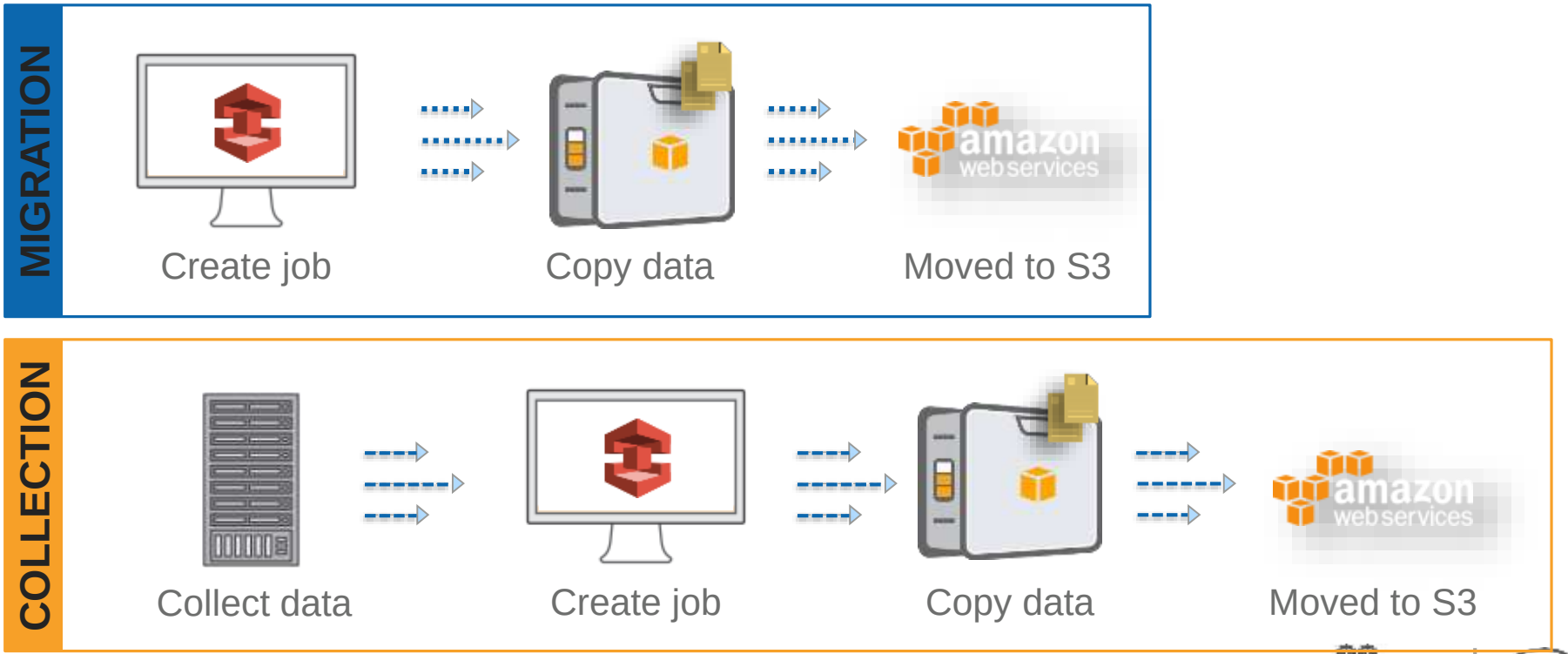
Petabyte-scale hybrid device with onboard compute and storage

- 100 TB local storage
- Local compute equivalent to an Amazon EC2 m4.4xlarge instance
- 10GBase-T, 10/25Gb SFP28, and 40Gb QSFP+ copper, and optical networking
- Ruggedized and rack-mountable

amazon web services | intel

Snowball Edge: Hybrid capabilities

AWS INNOVATE

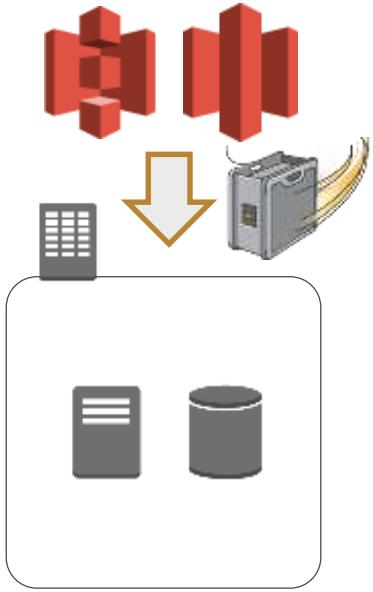


amazon web services | intel

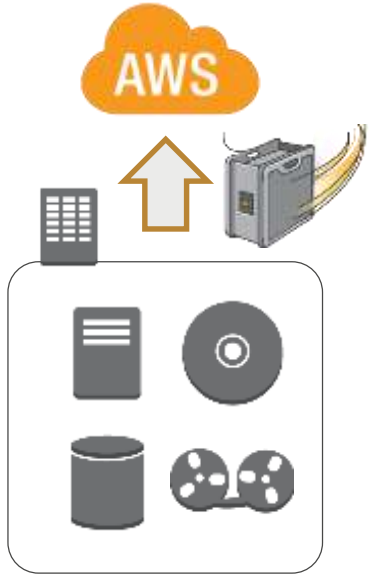
When to use AWS Snowball

AWS INNOVATE

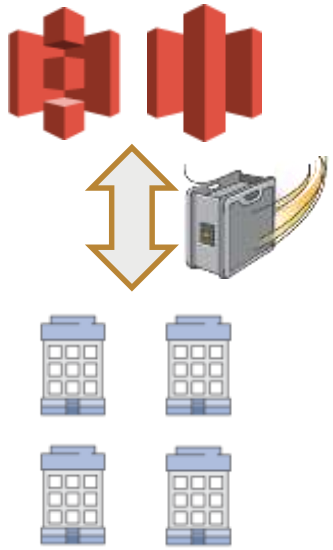
Cloud Migration



Disaster Recovery



Datacenter Decommission



Content Distribution

amazon web services | intel

Scripps Networks Interactive

AWS INNOVATE

Active archive transport and archival for digital content provider

scripps networks interactive

Problem Statement:

- Need storage platform to manage active archive content
- Existing content repository too large to migrate via available network-based ingest methods

Use of AWS:

- S3 and Snowball for massively scalable ingest
- S3 for storage, Glacier for content archive
- Snowball to securely transport existing media content from on-premises storage and tape vault

Business Benefits:

- Petabyte-scale data transport without increased network costs
- Massive scalability and elasticity
- Lower TCO for active archive storage

amazon web services | intel

AWS INNOVATE

Storage Gateway:

augmenting existing on-prem storage with cloud

amazon web services | intel

Storage Gateway hybrid storage solutions

AWS INNOVATE

Use standard storage protocols to access AWS storage services

Amazon CloudWatch

AWS Identity and Access Management (IAM)

AWS CloudTrail

AWS Key Management Service (KMS)

amazon web services | intel

Storage Gateway hybrid storage solutions

AWS INNOVATE

Use standard storage protocols to access AWS storage services

Amazon CloudWatch

AWS Identity and Access Management (IAM)

AWS CloudTrail

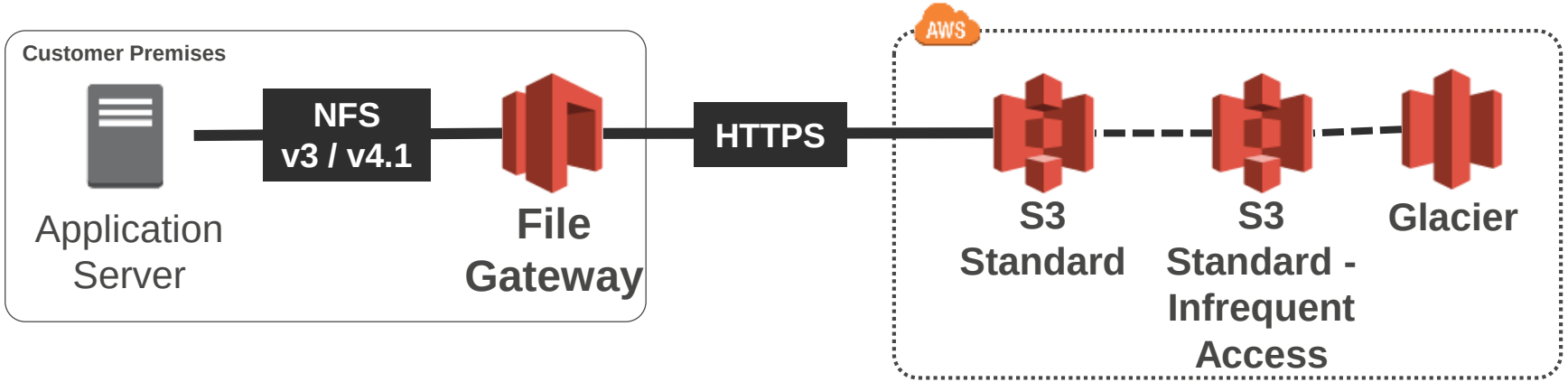
AWS Key Management Service (KMS)

amazon web services | intel

File gateway

AWS INNOVATE

On-premises file storage maintained as objects in Amazon S3



Data stored and retrieved from your S3 buckets

One-to-one mapping from files-to-objects

File metadata stored in object metadata

Bucket access managed by IAM role you own and manage

Use S3 Lifecycle Policies, versioning, or CRR to manage data

amazon web services | intel

Enabling cloud workloads

AWS INNOVATE

Move data to AWS storage for Big Data, cloud bursting, or migration

moderna
messenger therapeutics

"Storage Gateway has the promise to transform the way we move data into the cloud. The NFS interface lets us **easily integrate data files from analytical instruments**, and the transparent S3 storage lets us easily connect our cloud-based applications and leverage the **powerful storage capabilities of S3**.

With Storage Gateway, we can now **unleash the full power of AWS on our instrument data**."

amazon web services | intel

Enabling cloud workloads

AWS INNOVATE

Move data to AWS storage for Big Data, cloud bursting, or migration

moderna
messenger therapeutics

"Storage Gateway has the promise to transform the way we move data into the cloud. The NFS interface lets us **easily integrate data files from analytical instruments**, and the transparent S3 storage lets us easily connect our cloud-based applications and leverage the **powerful storage capabilities of S3**.

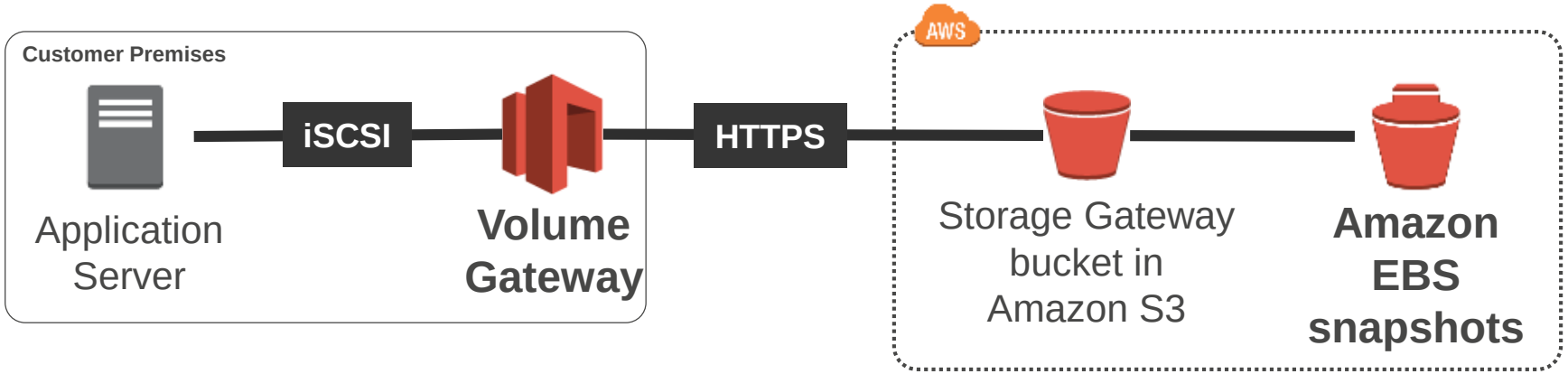
With Storage Gateway, we can now **unleash the full power of AWS on our instrument data**."

amazon web services | intel

Volume gateway

AWS INNOVATE

On-premises volume storage backed by Amazon S3 with EBS snapshots



Block storage in S3 accessed via the volume gateway

Compression of data in-transit and at-rest

Backup on-premises volumes to EBS snapshots

Create on-premises volumes from EBS snapshots

Up to 1PB of total volume storage per gateway

amazon web services | intel

Tiering storage: S3, Glacier, and EBS

AWS INNOVATE

Easily add AWS storage to your on-premises environment

JustGiving™

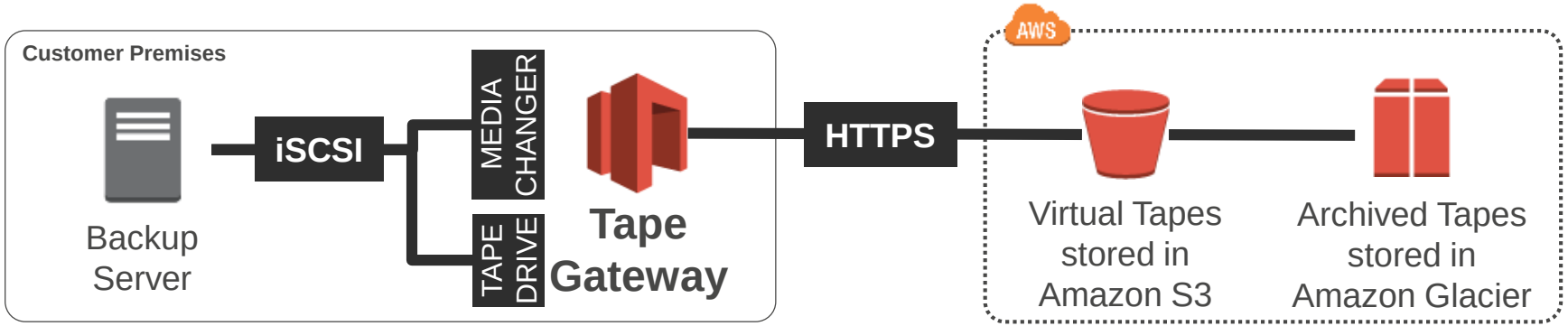
"Storage Gateway is at the **core of our disaster recovery and business continuity (BCM) processes**, handling our Co-Lo'd OLTP and OLAP off site data backups, as well as our in-office BCM. It **works transparently**, in a lights out way, archiving off to a separate AWS account with a simple grandfather-father-son snapshot plan in place".

amazon web services | intel

Tape gateway

AWS INNOVATE

Virtual tape storage in Amazon S3 and Glacier with VTL management



Virtual tape storage in S3 and Glacier accessed via tape gateway

Data compressed in-transit and at-rest

Up to 1 PB total tape storage per gateway, unlimited archive capacity

Supports leading backup applications:

VERITAS DELL EMC veeam Hewlett Packard Enterprise Microsoft System Center Data Protection Manager 2012 arcserve amazon web services intel

Backup, archive, disaster recovery

Cost effective storage in AWS with local or cloud restore

AWS INNOVATE

Southern OREGON UNIVERSITY

“**Tapes are a headache**, prone with hardware failures, offsite storage costs, and constant maintenance needs. Storage Gateway provided **the most cost-effective and simple alternative**. We even got disaster recovery by using a **bi-coastal data center**”.

amazon web services | intel

AWS INNOVATE

Amazon EFS with DirectConnect (DX): *Working over a dedicated network*

amazon web services | intel

Amazon Elastic File System (EFS)

AWS INNOVATE

Provides simple, scalable, highly available & durable file storage in the cloud

Petabyte scale file system distributed across an unconstrained number of storage servers in multiple Availability Zones (AZs)

Elastic capacity, automatically growing & shrinking as you add & remove files

amazon web services | intel

Amazon Elastic File System (EFS)

AWS INNOVATE

Standard file system interface & semantics (NFSv4.0 & 4.1)

Shared storage

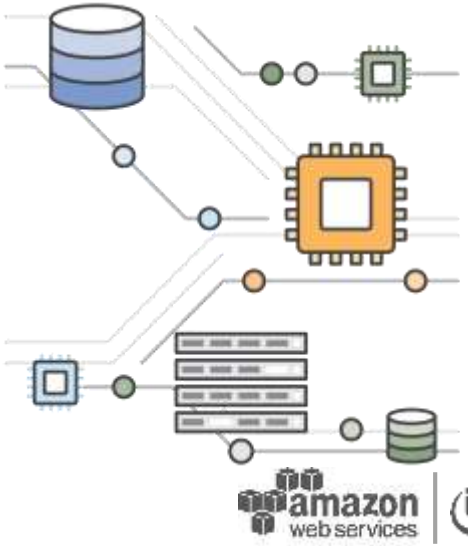
Highly available & highly durable

Consistent low latency

Strong read-after-write consistency

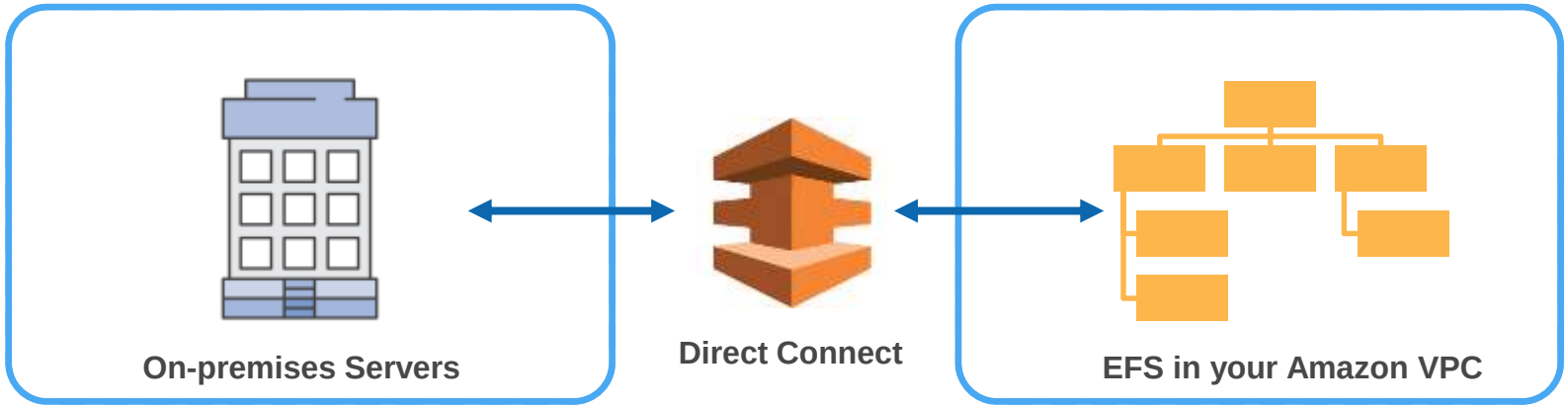
Elastic capacity

Fully managed



Access EFS file system via AWS DX

AWS INNOVATE



amazon web services | intel

Three scenarios for working with file data across on-premises environments and EFS

AWS INNOVATE

Migration

- Move entire data set permanently to EFS
- Access data from EC2 instances

Bursting

- Move data set temporarily to EFS
- Access data from EC2 instances and process
- Move data back on premises once processed

Backup / DR

- Maintain copy of entire data set on EFS
- Restore data to on premises storage or (for DR) access the data from failed-over applications running on EC2 instances

Tiering

amazon web services | intel

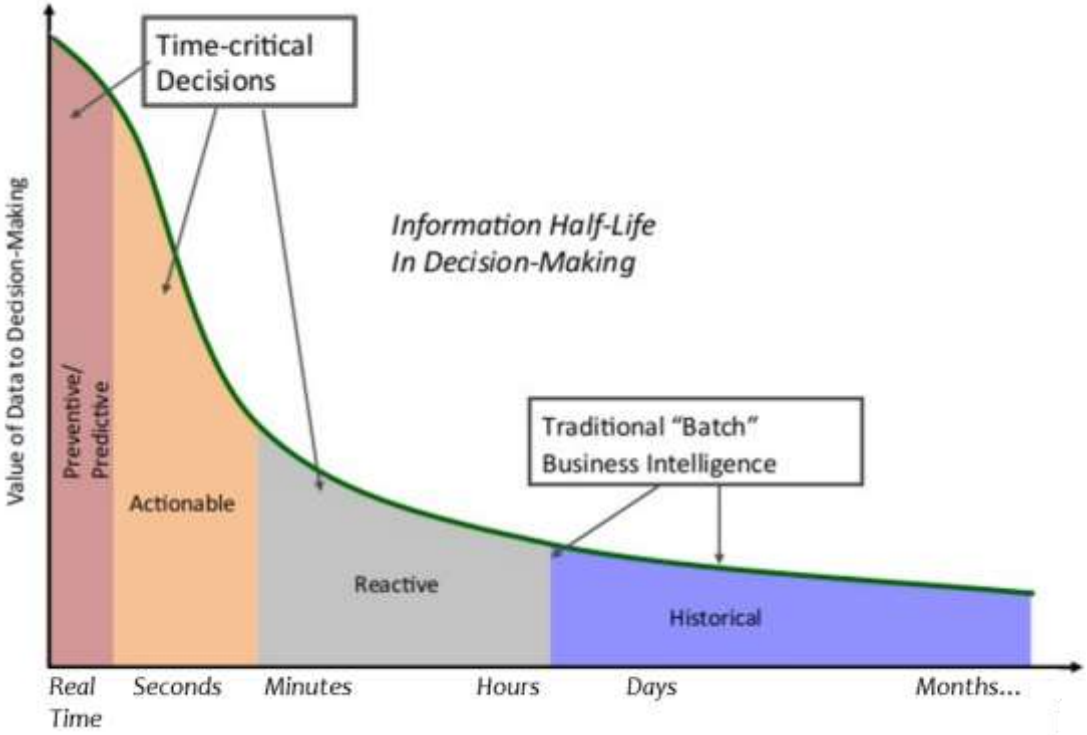
AWS INNOVATE

Amazon Kinesis:
streaming data

amazon web services | intel

Time Value of ~~Money~~ Data

AWS INNOVATE



amazon web services | intel

Streaming Data Scenarios

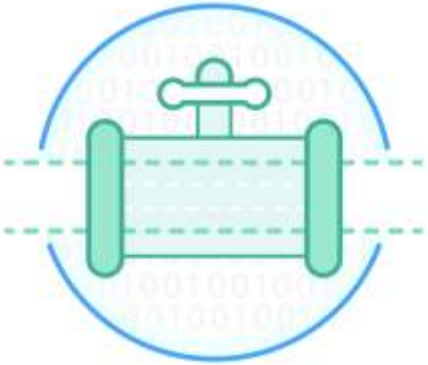
AWS INNOVATE

Scenarios	Accelerated Ingest-Transform-Load	Continual Metrics Generation	Responsive Data Analysis
Data Types	IT logs, applications logs, social media / clickstreams, sensor or device data, market data		
Ad/Marketing Tech	Publisher, bidder data aggregation	Advertising metrics like coverage, yield, conversion	Analytics on user engagement with ads, optimized bid / buy engines
IoT	Sensor, device telemetry data ingestion	IT operational metrics dashboards	Sensor operational intelligence, alerts, and notifications
Gaming	Online customer engagement data aggregation	Consumer engagement metrics for level success, transition rates, CTR	Clickstream analytics, leaderboard generation, player-skill match engines
Consumer Engagement	Online customer engagement data aggregation	Consumer engagement metrics like page views, CTR	Clickstream analytics, recommendation engines

amazon web services | intel

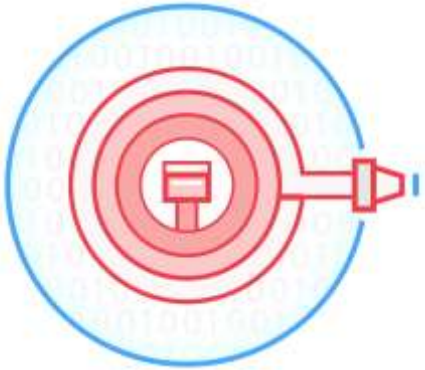
Amazon Kinesis

AWS INNOVATE



Kinesis Streams

Stores data as a continuous replayable stream for custom applications



Kinesis Firehose

Load streaming data into Amazon S3, Amazon Redshift, and Amazon Elasticsearch Service

Kinesis Analytics

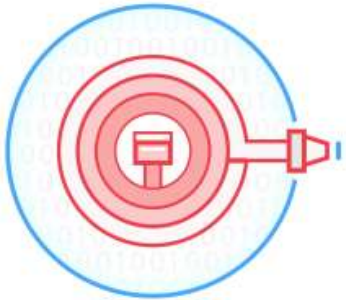
Analyze data streams using standard SQL queries

amazon web services | intel

Match the Services and Scenarios

AWS INNOVATE

1. **Accelerated Ingest-Transform-Load**

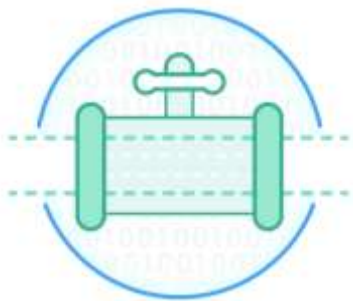


Kinesis Firehose

2. **Continual Metrics Generation**

Kinesis Analytics

3. **Responsive Data Analysis**



Kinesis Streams

amazon web services | intel

Kinesis Streams

AWS INNOVATE

Send clickstream data to Kinesis Streams

Kinesis Streams stores and exposes clickstream data for processing

Custom application built on Kinesis Client Library or Spark Streaming makes real-time content recommendations

Readers see personalized content suggestions

mlbam

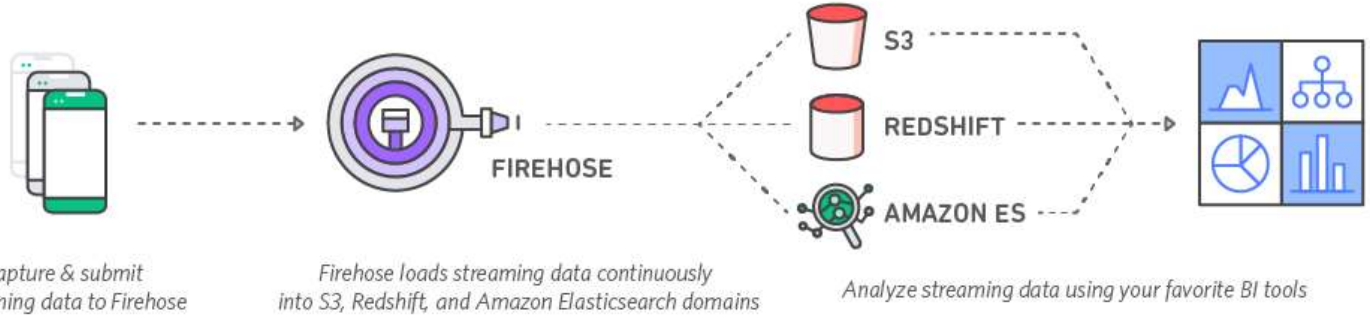
スシロー SUSHIRO

DataXu

amazon web services | intel

Kinesis Firehose

AWS INNOVATE



hello

realtor.com®

LiveRamp. an acxiom company

amazon web services | intel

AWS INNOVATE

S3 Transfer Acceleration:
moving large files over long distances

amazon web services | intel

Amazon S3 transfer acceleration

AWS INNOVATE

Change your endpoint, not your code

Leverages 59 global edge locations

Optimized protocols

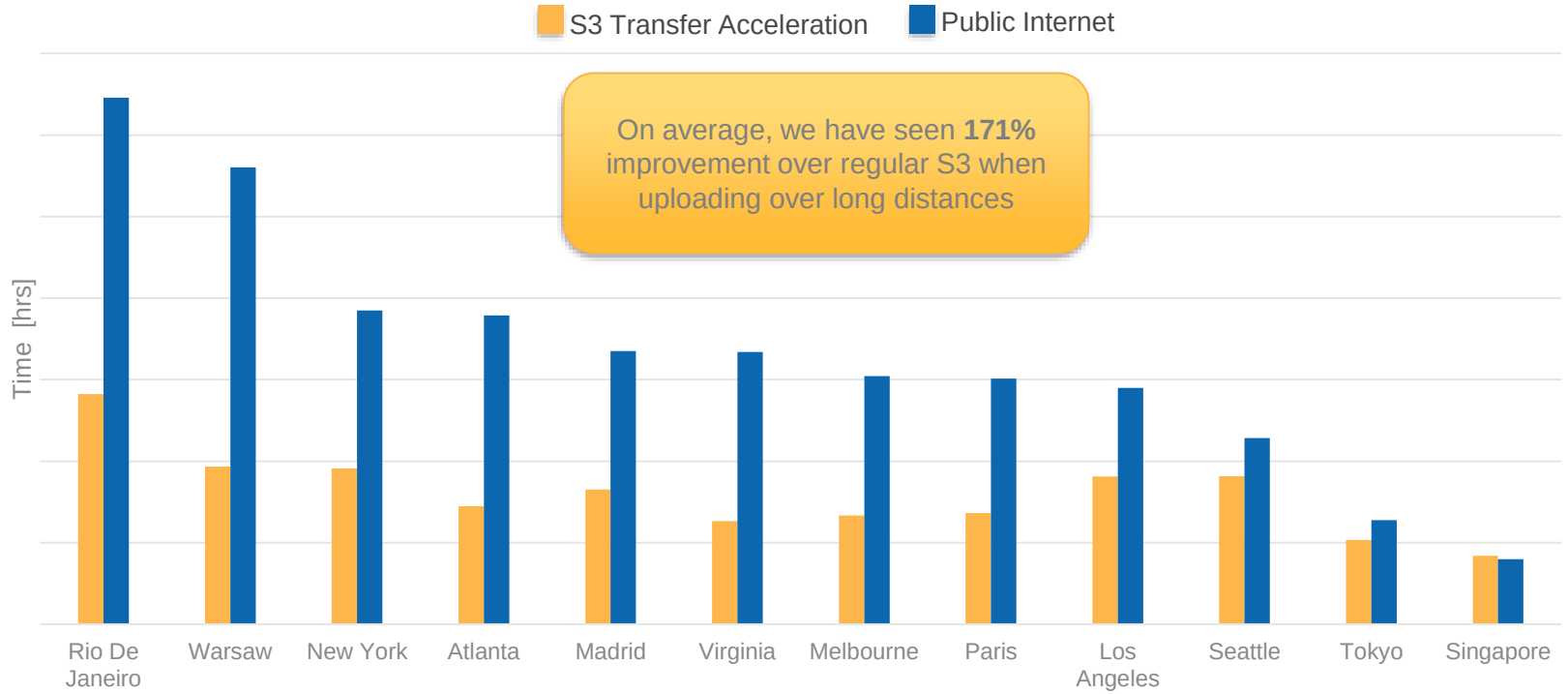
No firewall exceptions

No client software required

amazon web services | intel

How fast is S3 Transfer Acceleration?

AWS INNOVATE



500 GB upload from these edge locations to a bucket in Singapore

amazon web services | intel

Use Case: Media Upload

AWS INNOVATE

- Growing 30% per year
- Typical Friday during football season – over 35 hours of video every minute is uploaded

Data in: User > S3-TA > S3 > Transcode > Redshift
Data out: S3 > CloudFront > User

S3-TA = >20% increase in upload and encoding speeds

hudl

> *S3 transfer acceleration reduces the average time it takes for us to ingest videos from our global user base by almost half. This gives our customers the ability to edit and share videos sooner where speed is a critical factor.* - Brian Kaiser, CTO

amazon web services | intel

Summary: Hybrid Cloud Storage

AWS INNOVATE

	Offline	Online
Migration	Snowball, Snowball Edge, Snowmobile	Storage Gateway, EFS over DX
Bursting		Storage Gateway, S3 Transfer Acceleration
Tiering	Snowball Edge	Storage Gateway
Backup	Snowball	Storage Gateway, 3rd Parties, EFS over DX

amazon web services | intel

AWS INNOVATE

Partner Tiering & Migration Solutions:
working with what's already there

amazon web services | intel

Storage Partner Solutions

AWS INNOVATE

Technology Solutions vetted by the AWS Storage Competency Program

Primary Storage

Solutions that leverage file, block, object, and streamed data formats as an extension to on-premises storage

- AVERE
- ctera
- NetApp
- panzura
- SoftNAS
- ZADARA STORAGE

Backup and Recovery

Solutions that leverage Amazon S3 for durable data backup

- COMMVAULT
- Cloudberry
- ctera
- druva
- DELL EMC
- IBM
- NetApp
- VERITAS

Archive

Solutions that leverage Amazon Glacier for durable and cost-effective long-term data backup

- COMMVAULT
- NetApp
- sonian

BCDR

Solutions that utilize AWS to enable recovery strategies focused on RTO and RPO requirements

- CloudEndure
- CLOUDVELOX
- druva
- NetApp
- Zerto

aws.amazon.com/backup-recovery/partner-solutions/ Note: Represents a sample of storage partners

Note: Dell-EMC, IBM and Veritas have solutions and are working towards competency requirements

amazon web services | intel

Thank you!

Matt Nowina
mnowina@amazon.com

amazon web services | intel