

AWS re:Invent

NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

STG215

Preserving and maximizing the value of digital media assets using Amazon S3

Joe Fitzgerald

General Manager,
Amazon S3 Glacier
AWS



Patrick Scott

Director, Product Management,
Asset Management
Warner Bros. Discovery



Jay Brown

Director, Tech. Software
& Systems Architecture
Warner Bros. Discovery



Scott Gutterman

Senior Vice President Digital
Operations & Broadcast Technology
PGA TOUR



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Managing
exponential
data growth

Preserving
your most
priceless assets

Maximizing
the value of
archival data to
unlock business
value

Exploring customer stories from changemakers who are driving innovation: **Warner Bros., CNN, and PGA Tour**

1

Managing **exponential data growth**

MUSIC



LOGS



EVENTS



THIRD PARTY



101 ZB
of data created
and replicated in
the world in 2022



IOT



TV/
MOVIES



APPS

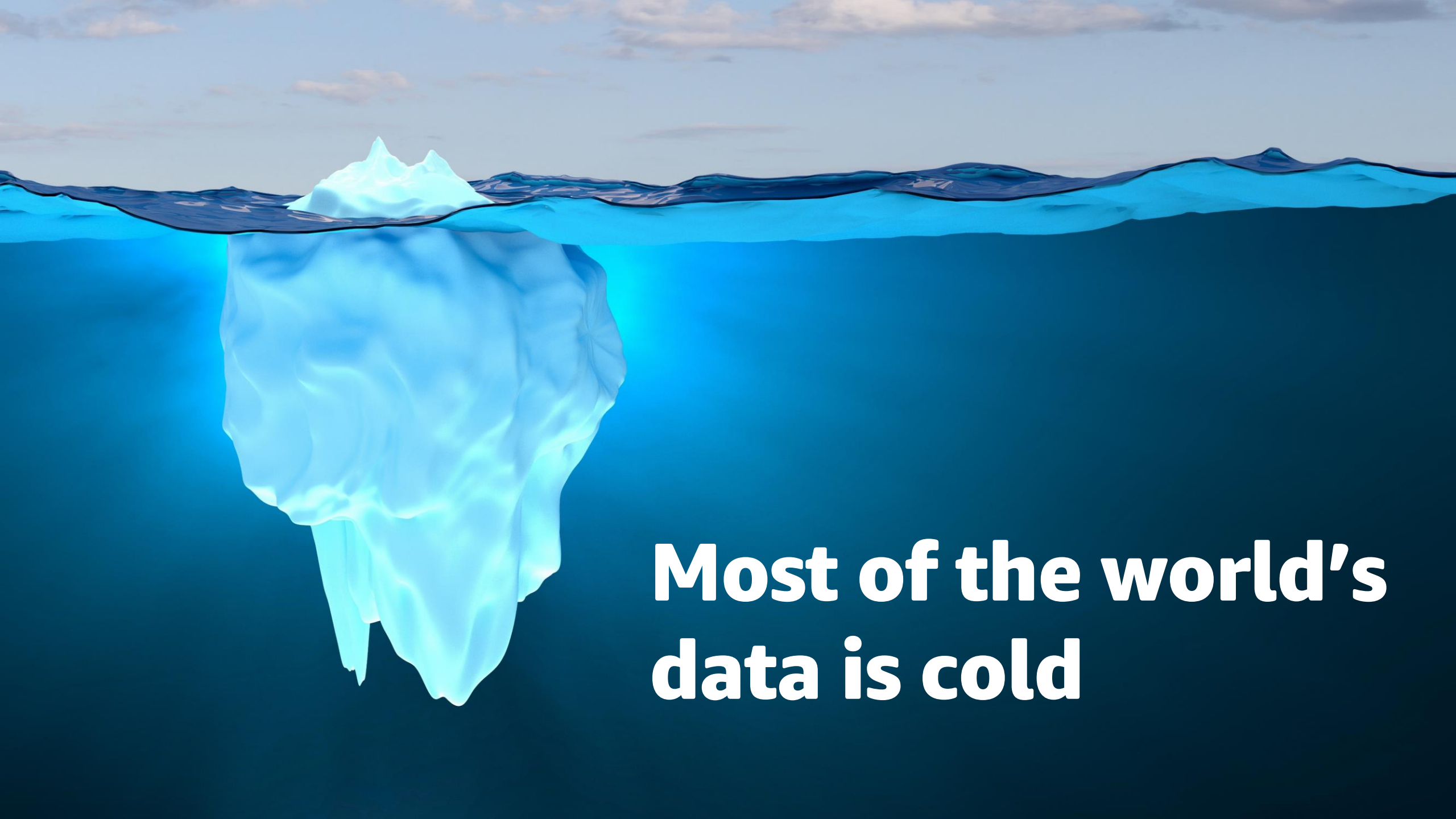


SENSORS



MOBILE





**Most of the world's
data is cold**

Amazon S3 archival storage classes



S3 Glacier Instant Retrieval

Fastest access to
archive storage



S3 Glacier Flexible Retrieval

Flexible retrieval options



S3 Glacier Deep Archive

Lowest storage cost
in the cloud



\$750 million in savings

S3 Intelligent-Tiering

Automatic archiving based on
the last access of your objects

2

Preserving your **most priceless assets**

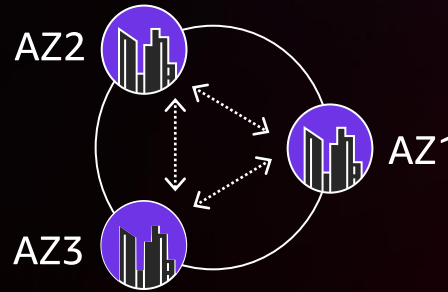


Amazon S3 – A safe place to store your data



30 Regions

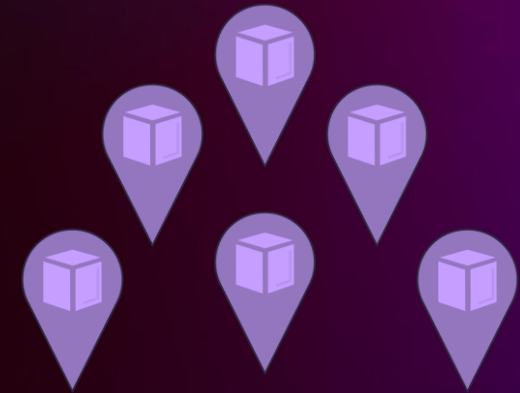
Geographic locations with multiple AZs, separated physically by miles and isolated



96 Availability Zones (AZs)

Objects stored across multiple devices spanning a minimum of 3 AZs

Multi-AZ storage classes function normally if an AZ is lost



Highly durable

Amazon S3 is Designed for 11 9s of durability to sustain concurrent device failures and data in the event of an entire AZ loss

AWS is architected to be one of the most flexible and secure cloud computing environments available today

S3 performs
**over 4 billion checksum
computations**
per second

3

Maximizing the value of archival data



1 PB

restored every day from
S3 Glacier

Available Now

Amazon S3 Glacier up to 10x restore throughput increase

UP TO 90% FASTER RETRIEVALS FROM S3 GLACIER FLEXIBLE RETRIEVAL, S3 GLACIER DEEP ARCHIVE



Available at no additional cost

Supports restore requests at a rate of up to 1,000 transactions per second, per account in an AWS Region

Ideal for restoring backups, responding to audit requests, retraining machine learning models, and performing analytics on historical data

Significantly reduces the restore completion time for datasets composed of small objects



Unlocking the potential of your data



Open up new opportunities



Create agility



Invent new customer experiences

Recap

Managing
exponential
data growth

Preserving
your most
priceless assets

Maximizing
the value of
archival data to
unlock business
value

Amazon S3 customers



Preserving “the Stuff that Dreams are Made of”

Patrick Scott (he/him)

Director, Product Management, Asset Management
Warner Bros. Discovery





WARNER BROS. DISCOVERY



OWN

 **Discovery**



MAGNOLIA
NETWORK



HBO

HBOmax

TruTV
CHANNEL



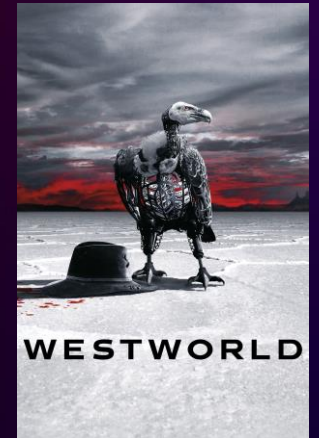
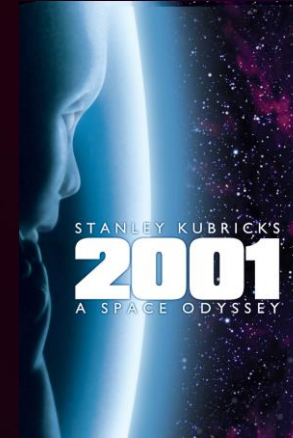
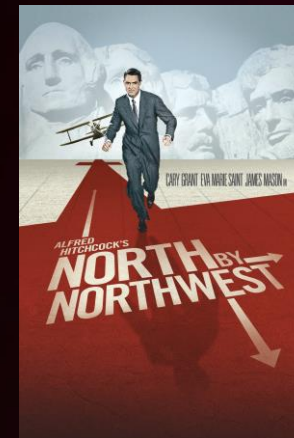
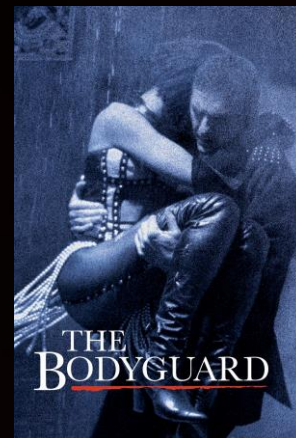
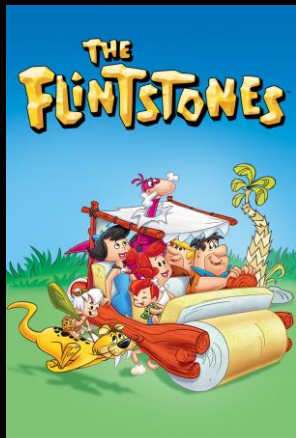
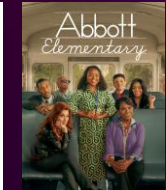
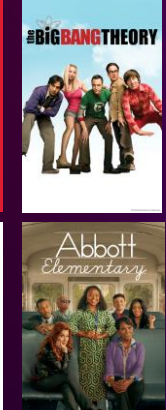
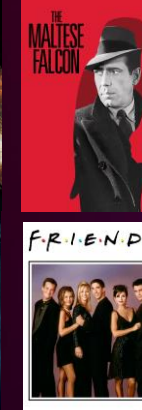
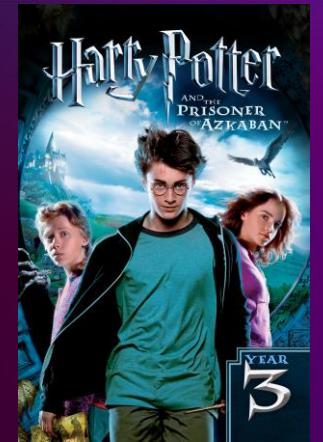
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



100 years of history



The Warner Brothers -- Harry, Jack, Sam, and Albert



- Millions of physical masters
 - 4 million reels of film prints and camera negatives
 - Video and LTO tapes, audio masters, stills, etc
- Thousands of hours of born-digital content being created daily
- Large bodies of content that fall outside of standard specs and media types





The Challenge

Preserve the content of the studio to ensure continuity of the business' assets

- Retain the best sources available today
- Pair innovative technology with expertly archived content to allow the catalog to meet the needs of future formats and business cases



The Assignment

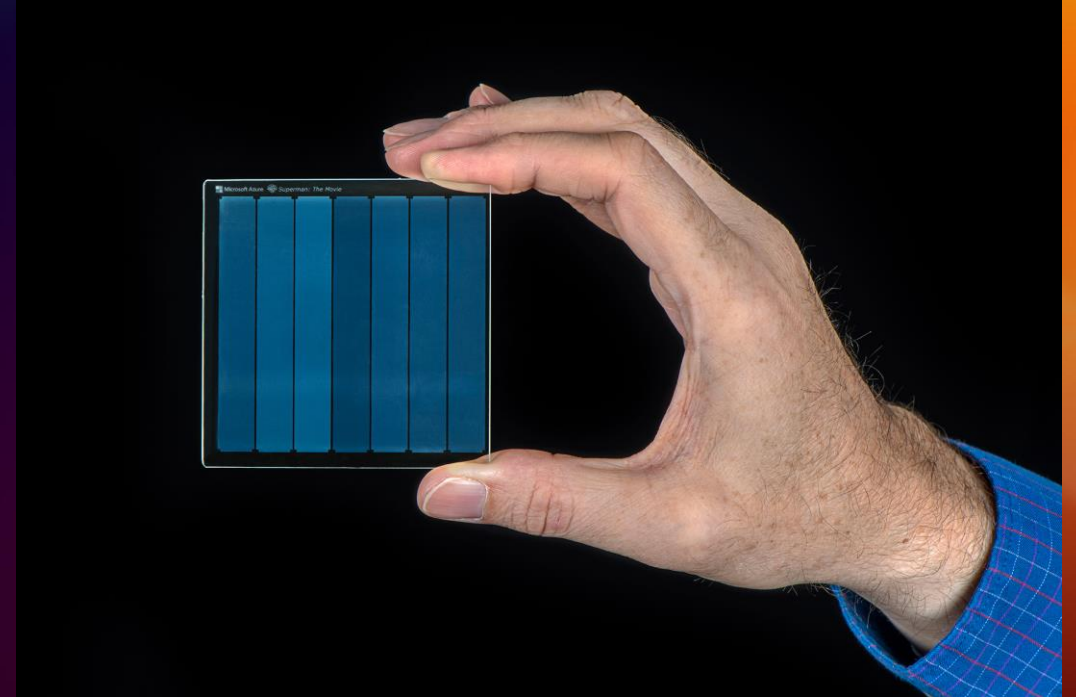
Ensure a lifespan for all assets for the life of the studio + 300 years

- Digitize all physical masters (e.g., content originally created on tape or film)
- Migrate all digitized physical and born-digital content into a MAM system for review, indexing, and discovery

Digitization



Innovation



Journey to the cloud

2014

Decision to move to digital preservation; first at-risk assets digitized and stored with on-premises storage

2014–2020

Continuous growth of storage with on-premises partners

2019

New MAM provider engaged, opening up path to cloud

2020

With assistance from AWS Partner, TapeArk, 25 PB of content lifted from on premises to AWS

2020–2022

COVID... Digital Archive team able to work remotely with AWS functionality, allowing them to move forward with content migration to the cloud and deployment of new MAM

Present

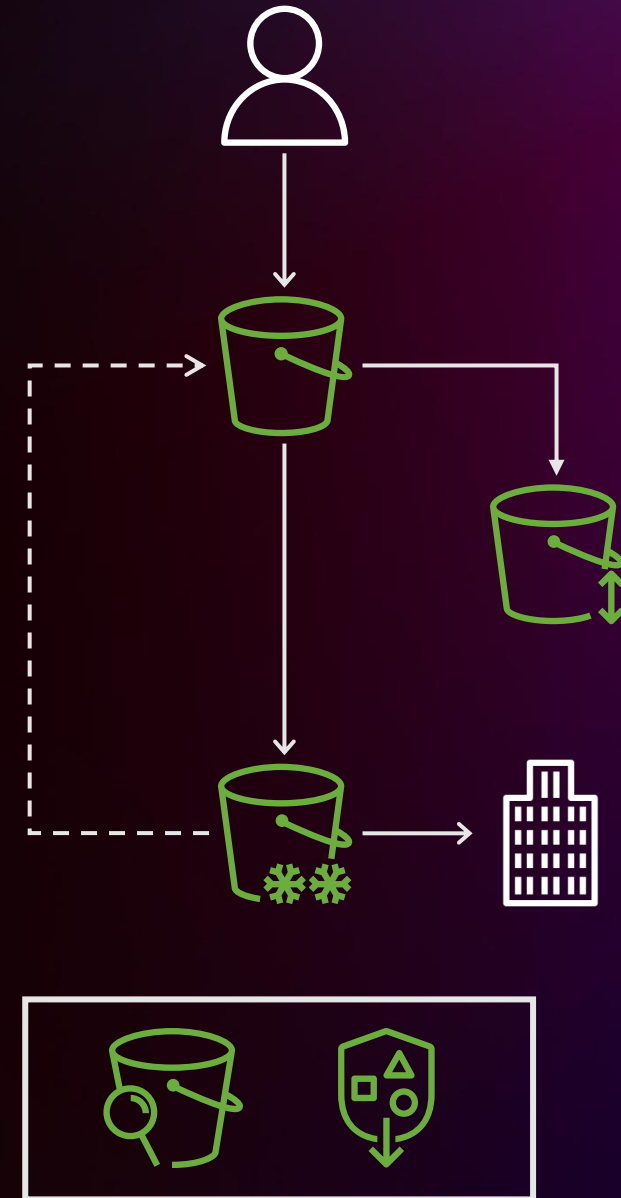
Growing collaboration with AWS, exploring further opportunities

Vault architecture



Vault ingest workflow

- Media package uploaded via API
- Media package initially stored in **S3 Standard**
- Package validated using hash comparisons and XML schema validation tools
- Proxy transcode created and stored in **S3 Standard-Infrequent Access**
- Audio-visual QC performed on proxy
- Metadata finalized and package marked ready for ingestion in Vault
- Package secured to **S3 Glacier Deep Archive**
- Package replicated to secondary site (DR)
- Original package deleted from **S3 Standard**



What's working...

- Clarity on costs
- Room to grow – no more “we’re out of room” challenges
- Simplified and expanded ingest
- Fixity innovation
- Improved integrations

What's next...

- Additional libraries
- New, larger, more complex formats and specs
- Better reporting and metrics
- Applied machine learning

Thank you!

Patrick Scott

Director, Product
Management, Asset Mgmt
Warner Bros. Discovery



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

CNN Content Preservation Archive

Jay Brown (he/him)

Director, Tech. Software & Systems Architecture,
Warner Bros. Discovery

Harry Schaefer (he/him)

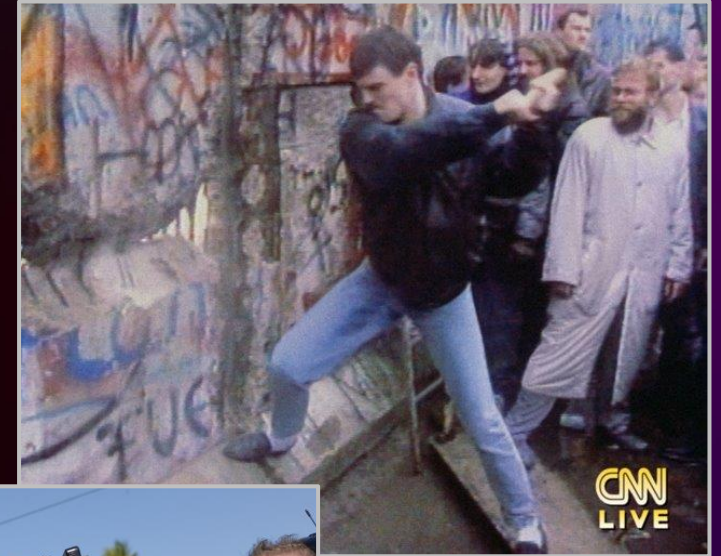
Principal Architect,
Warner Bros. Discovery



CNN video library & archive by the numbers

Archive size

- Approaching 30 PB
- Over 3 million digital video files containing over 1 million hours of video
- SD and HD broadcast "house" formats
- Raw feeds and edited assets
- Average file size of 6–8 GB
- Daily growth of 5 TB on average
- Annual growth of 1.5–2 PB



CNN video library & archive by the numbers

Archive usage

- Supports daily news production and long-form production, as well as content licensing
- Assets "born digital" since 2005, older assets continue to be digitized
- Daily restores of 1–2 TB, hundreds of files
- 25% of restores are for content less than 90 days old
- 25% of restores are for content greater than 5 years old
- Restore SLA measured in minutes



Where we were



- Centralized, on-premises storage in single location
- Offsite backup media (offline)
- Serving 4 major production centers, dozens of bureaus worldwide
- Thousands of users
- No real DR plan

Where we thought we were going

**Cloud archive
as backup/BCP
solution**

**Gradual move
to the cloud**

(plus buy more
on-premises storage)

**Cloud as primary
storage in the future**

The Atlanta Journal-Constitution

[Local News](#)[Georgia Politics](#)[Events](#)[Opinion](#)[Sports](#)[EPaper](#)[Newsletters](#)

Thank you for subscribing.

[Choose from a variety of newsletters.](#)

AT&T sells, will lease CNN Center



PIVOT!



Where we needed to go

Cloud archive as
primary storage
AND backup/BCP
solution

Immediate move to
the cloud, before
move out, and
before current
storage fills up

Cloud as primary
storage **NOW**,
supporting
production
workloads

Challenges



PIVOT FROM
BCP/BACKUP TO
PRIMARY ARCHIVE



MIGRATE OLD,
WRITE NEW
SIMULTANEOUSLY



GROWING
5 TB / DAY, PLUS
26+ PB TO
MIGRATE BY 2024



ENCRYPTION AND
SECURITY



NO NEGATIVE
IMPACT TO
PRODUCTION!!!

Decision: Amazon S3 & tiering



MULTI-REGION
STRATEGY

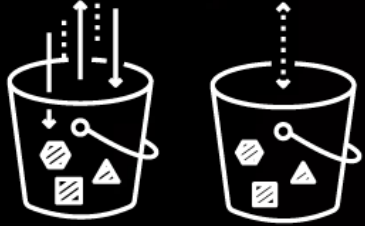


PRIMARY AND
SECONDARY COPIES



CROSS-REGION
REPLICATION

Tiering strategy – Goldilocks



S3 & S3 Standard-IA

Outside of our budget



S3 Glacier Deep Archive

Doesn't meet our SLA for primary copy
(but good for our backup)



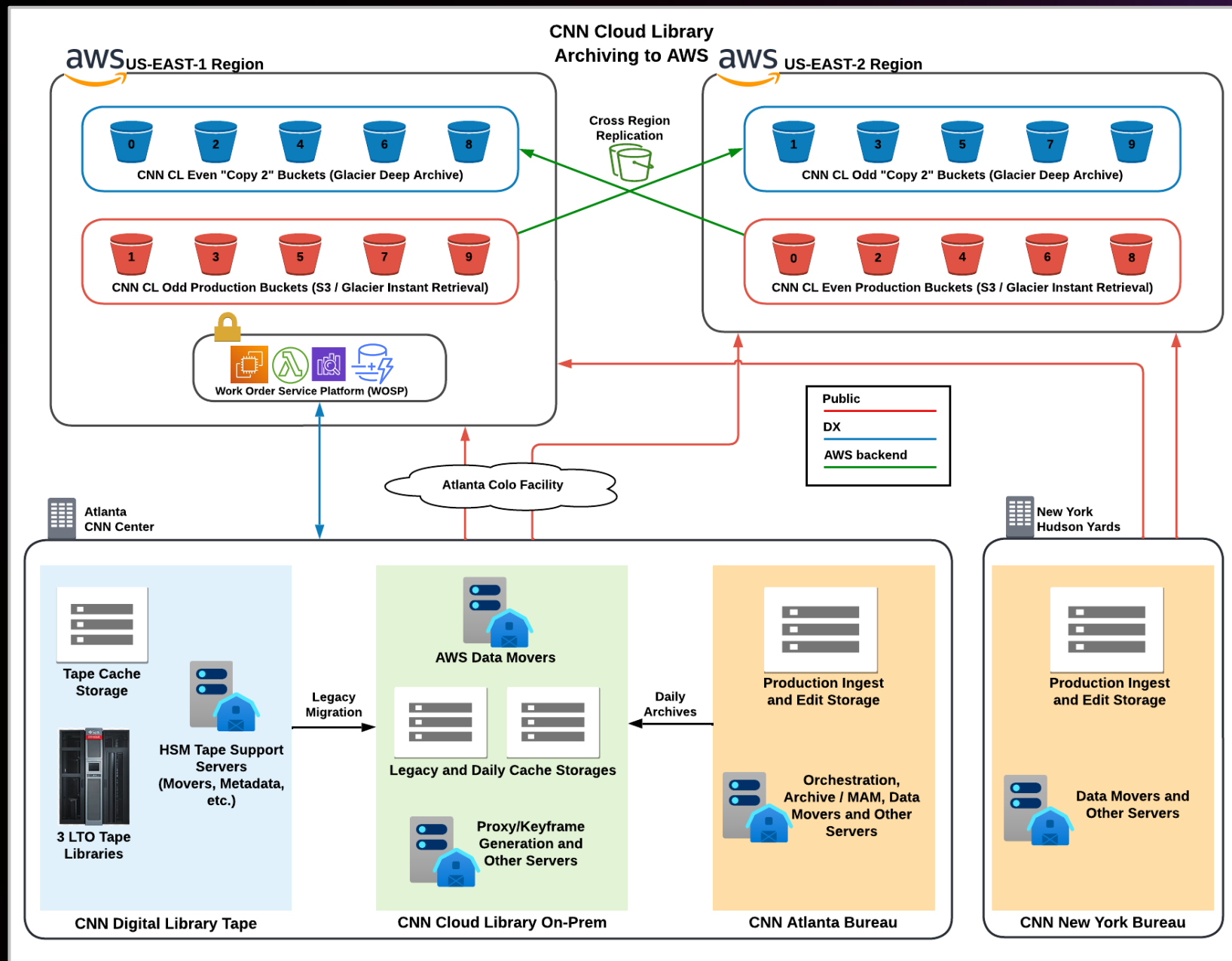
S3 Glacier Flexible Retrieval with expedited retrieval

Just about right

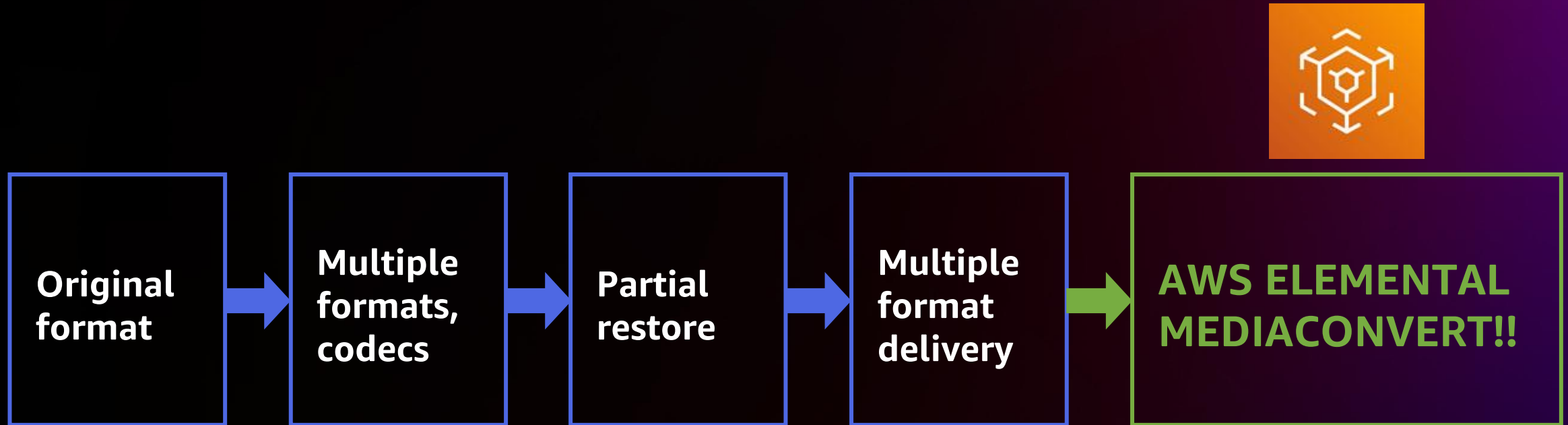


S3 GLACIER INSTANT RETRIEVAL!!

- Extra time required for data retrieval
- Similar cost profile to S3 Glacier Flexible Retrieval for our workload
- Simple, instant access with S3 Glacier Instant Retrieval



Additional Benefit: Transcode



Where we are today

**Stopped writing
content to our
on-premises library**

**80%+
migration
complete**

**Decommissioning
planning underway**

(not media migration or
expansion planning)

**Supporting
complete
production
workload and SLA
from S3/S3 Glacier**

Thank you!

Jay Brown (he/him)

Director, Tech. Software
& Systems Architecture,
Warner Bros. Discovery

Harry Schaefer (he/him)

Principal Architect,
Warner Bros. Discovery



PGA TOUR and Media Asset Management in the Cloud

Scott Gutterman (he/him)

Senior Vice President Digital Operations and
Broadcast Technology,
PGA TOUR



PGA TOUR and MAM in the cloud

2 main goals:

- 1 Preserve the history of the PGA TOUR
- 2 Enable quick and easy access to PGA TOUR video assets to a large and growing group of constituents



PGA TOUR and MAM in the cloud

IT WAS A 10-YEAR JOURNEY FROM OUR INITIAL SYSTEM TO MAM IN THE CLOUD



- 2010** Launch of the first PGA TOUR media asset management system on premises
- 2011** Ingested just under 5,000 hours of new content
- 2013** Provided access to CBS and NBC
- 2015** Launched MAM over the Internet (access from anywhere)
- 2016** Full Shotlink API integration to enhance metadata
- 2019** Integrated with IPB Curator and Adobe Premier
- 2021** Ingested over 15,000 hours of new content
- 2022** Fully deployed to AWS



Why AWS?

Only AWS met our needs:

- **Scalability**
Some weeks we have 6 live streams to ingest other weeks we have 20
- **AWS Elemental MediaConnect**
Could deliver on big weeks and even bigger events like Every Shot Live
- **Expertise**
Best in class professional services and thousands of developers on AWS
- **Single facility vulnerability**
We need to retire our on-premises data center and MAM for business continuity and security



Official Cloud Provider
of the PGA TOUR

PGA TOUR and MAM in the cloud

By the numbers:

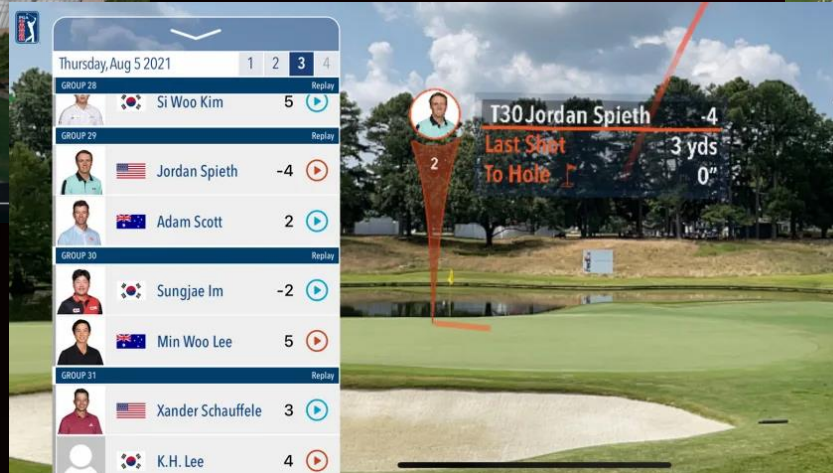
- 177k assets in the archive
- Over 200k hours of content
- 8.2M log entries (metadata tags)
- 13 petabytes of content stored in AWS



PGA TOUR and MAM in the cloud

Where to next?

- Creating a PGA TOUR Media Lake for all content types
- AI and ML applications – creating content from the assets based on context
- Generating more opportunities – faster for fans, players, and rights holders



Thank you!

Scott Gutterman

Senior Vice President Digital
Operations & Broadcast Technology,
PGA TOUR



Official Cloud Provider
of the PGA TOUR



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you



Customers in every industry archive data in AWS



Media &
Entertainment



Gaming



Healthcare &
Life Sciences



Financial
Services



Power &
Utilities



Energy



Manufacturing



Retail



Telecom



Automotive



Education



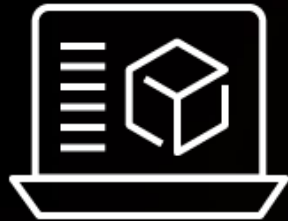
Government

Key takeaways

- 1 Archiving data in AWS unlocks cost savings opportunities
- 2 AWS protects your data and makes you more agile
- 3 Potential for data to reinvent your business is limitless

Continue your AWS Storage learning

**Build a
learning plan**



Set your AWS Storage
Learning Plans via
AWS Skill Builder

**Increase your
knowledge**



Use our **Ramp-Up Guides**
to build your storage
knowledge

**Earn AWS
Storage badges**



Demonstrate your
knowledge by achieving
digital badges

aws.training/storage



Thank you!



Please complete the session survey in the **mobile app**

