

Multi-Cloud data management in real-world usage

Paul Speciale

Chief Product Officer, Scality

Scality Multi-Cloud Data Management Solution

- Scality introduced Zenko as an Apache 2.0 open-source project in 2017 to help solve emerging problems in multi-cloud data management - on www.zenko.io and on <https://github.com/scality/Zenko>
- Now ~4 years of practical multi-cloud deployment experience with customers



The screenshot shows the Zenko website homepage. At the top is a navigation bar with the Zenko logo (a cat face) and the word "ZENKO". To the right of the logo are links: "What is Zenko?", "Developers", "Docs", "Blog", "Community", and "Try Zenko". The main content area features a large illustration with a cat face in the center, surrounded by various cloud provider logos (AWS, Azure, Google Cloud, etc.) and icons representing data management and APIs. Below the illustration, the text "ZENKO.IO" and "www.zenko.io" are displayed. A "Watch on YouTube" button is also visible. At the bottom of the main content area, the text "CLOUD FREEDOM FOR DEVELOPERS" is written in large, bold, white letters, followed by "THE OPEN SOURCE WAY TO CONTROL YOUR DATA ACROSS MULTIPLE CLOUDS" in smaller white letters.

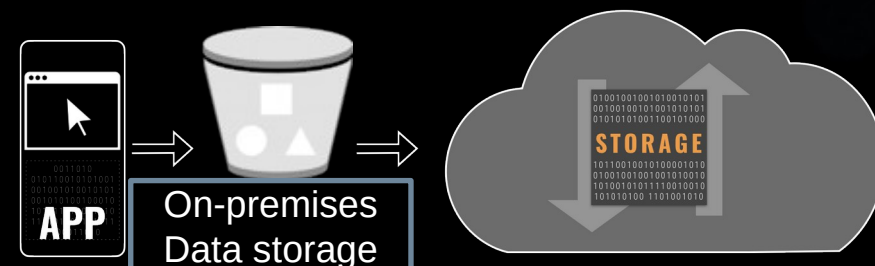
Cloud data management in practical real-world use

- The 3 key use-cases
 - Cloud archiving
 - Cloud data bursting
 - Cloud Disaster Recovery (D/R)
- Commercial usage
 - DataOps / Analytics cloud service
 - Data storage solutions

Customer focus: 3 key solutions for cloud data

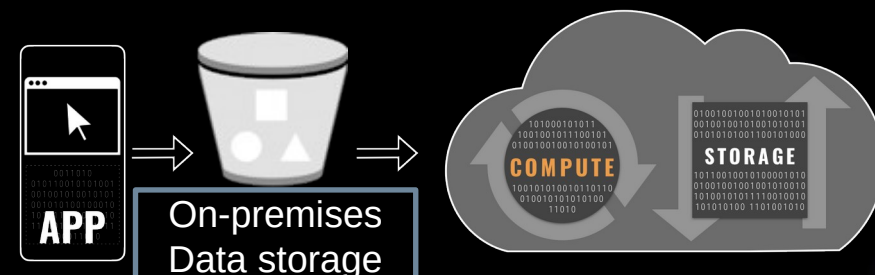
On-premises to cloud archive

- Offload cold/dormant data to cloud long-term archives (Azure Cold Storage, AWS Glacier or other)
- Solution: Lifecycle tiering policies - move data on data age/date rules



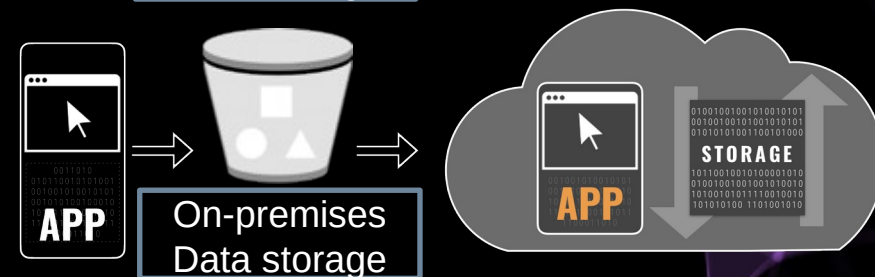
On-premises to cloud for compute bursting

- Move data to cloud compute services: AI/ML, analytics, transcoding, CDN, Video Indexing & others
- Solution: replication or tiering policies - based on tags/filters



On-premises to cloud for Disaster Recovery (D/R)

- Maintain a copy data in cloud for disaster recovery & business continuance in the event of site failure or outage
- Solution: continuous replication to cloud(s) to two copies with low RPO



CLOUD D/R IN LIFE SCIENCES (BIOPHARMA RESEARCH)

Major US Pharmaceuticals Company

BIG DATA FOR BIOPHARMA GENOMICS DATA

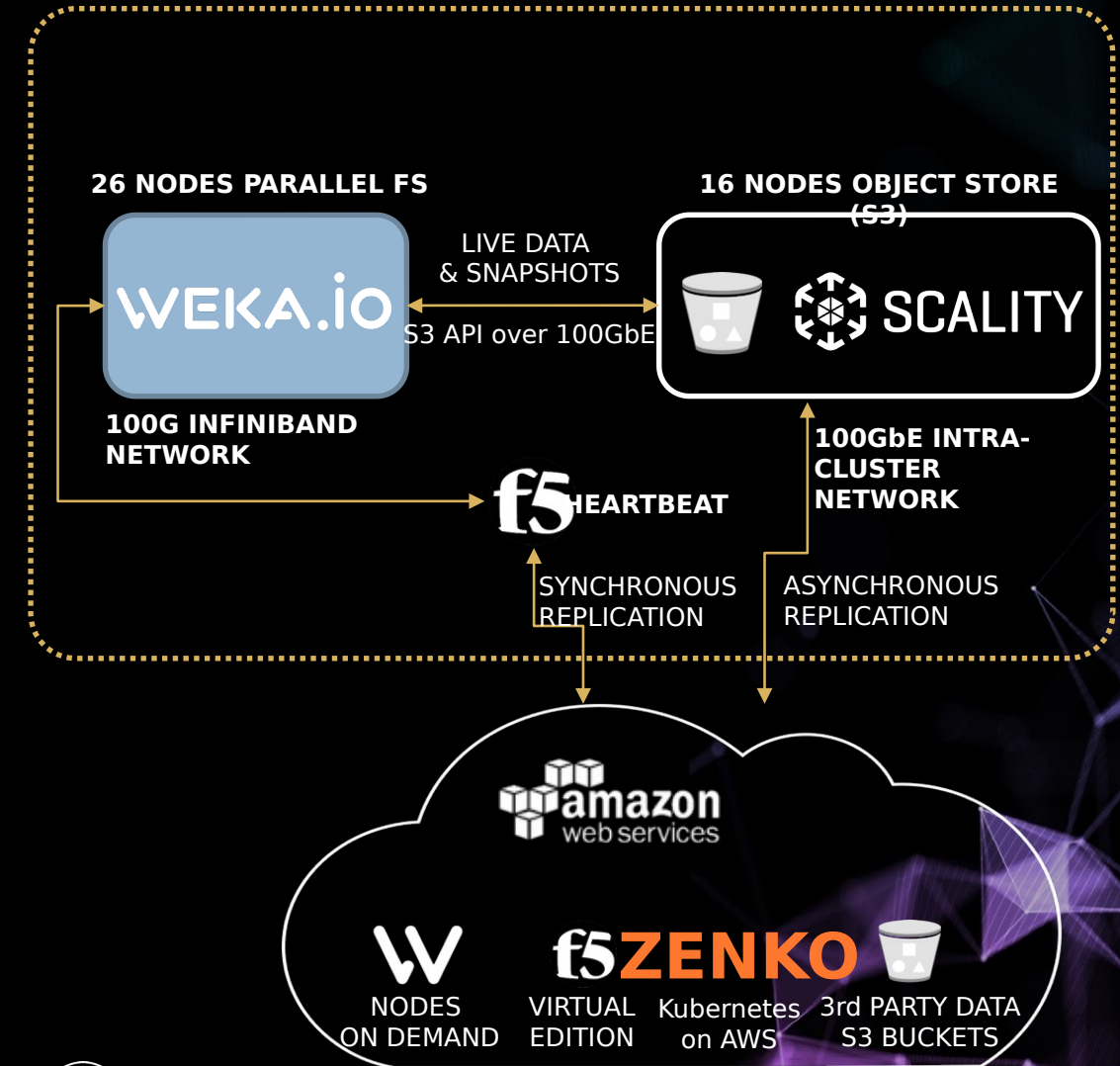
- high-performance flash file system solution Weka.io - physical on-premises and virtual in AWS EC2
- 10PB object store (Scality) for on-premises persistent storage

BUSINESS CONTINUANCE- ZENKO FOR CLOUD D/R

- Replicates data from RING to AWS S3 via “out-of-band” updates
- Data failover to cloud after primary site outage & manage failback once primary site is restored

CUSTOMER VALUE FOR CLOUD D/R

- Time to market vs 2nd physical DC
- Reduced Capex vs. second full physical deployment



The challenge: Multi-cloud data bursting in broadcast

A PREMIUM 24/7 MEDIA NETWORK

Bloomberg Media is the world's largest business news organization, reaching a premium audience of 91 million. The company delivers business and financial news across every platform and every time zone.

THE TECHNICAL CHALLENGE

To build a global system for file, object and cloud storage that would serve all of Bloomberg's storage and playback needs worldwide was a challenge:

- Cloud egress is costly and must be managed
- Storing content on multiple cloud providers has many benefits if managed properly.
- Public cloud storage doesn't guarantee 100% uptime.
- Global content lifecycle policies enable flexible, standardized and scalable data management.

Bloomberg

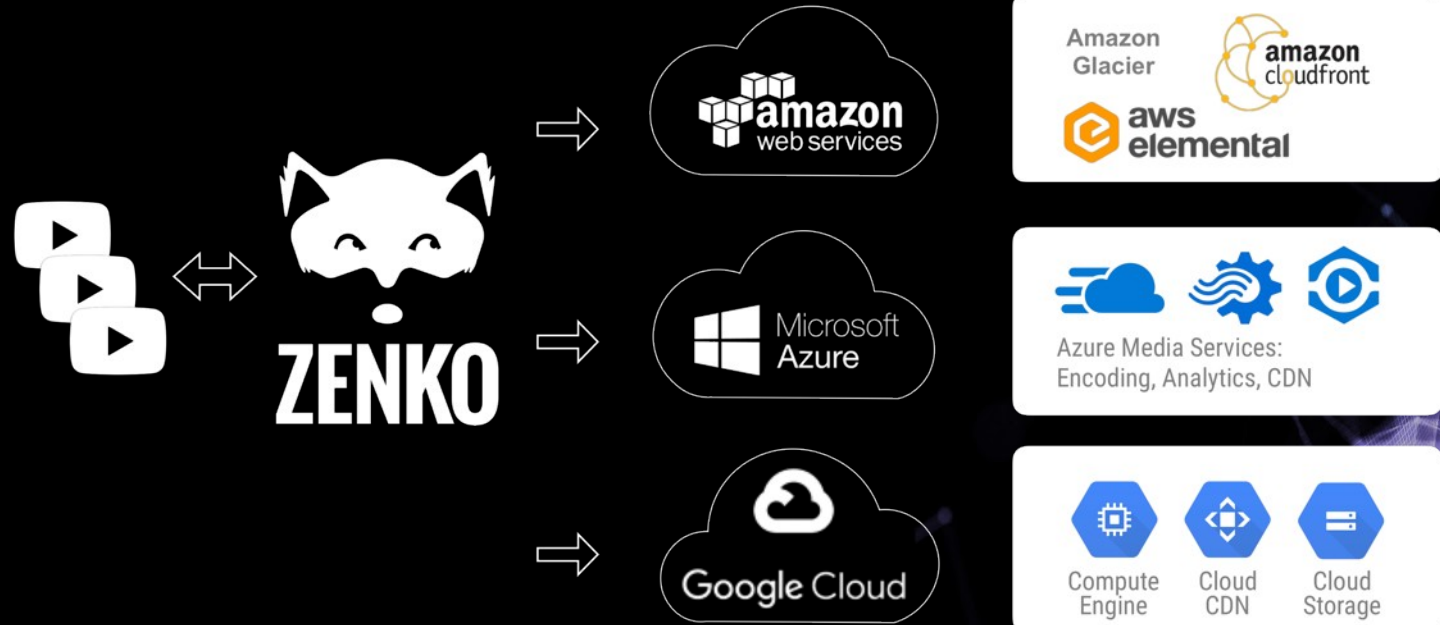
The multi-cloud CDN solution

OBJECT STORAGE FOR MEDIA ASSETS

- Data centers in London & NYC

ZENKO: MEDIA WORKFLOWS TO MULTIPLE PUBLIC CLOUDS

- Replicates media files from on-premises to public clouds for transcoding & CDN services
- Simultaneous transfers “in” to 15+ regions in AWS, Azure & Google
- Avoids egress charges to move data between clouds



“We have made very good progress in standing up not only a product, but really a whole infrastructure solution that can add the scale we need for all the changes coming in the future.”

– Bloomberg Cloud CTO

Commercial use of Zenko based multi-cloud solutions

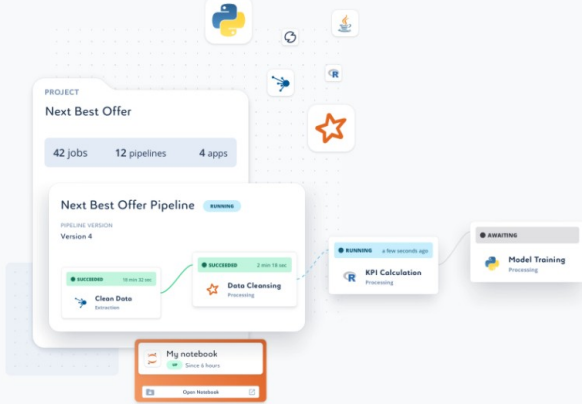
- Saagie Data Orchestrator
 - Cloud-based DataOps & Analytics Platform
 - ETL & visualize data workflows across clouds
- Fujifilm Object Archive
 - S3 object to tape solution
 - Extended Zenko with AWS Glacier API
- Object Storage
 - Scality's own ARTESCA lightweight, cloud-native object storage solution

Saagie helps Data Engineers achieve great things with company data

The Saagie DataOps Platform brings together the most popular technologies so you can deliver and run data projects quickly, easily and reliably.

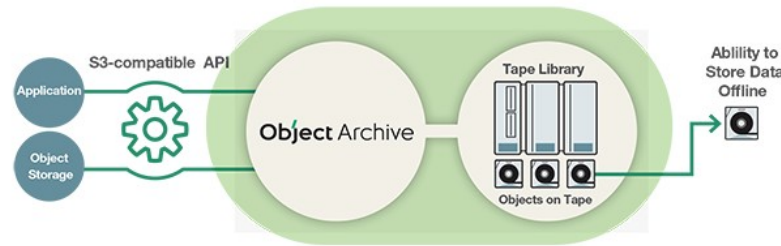
Extract, prepare, process and visualize data while getting it to the teams that need it in a fraction of the usual time. And once it's live, it won't let you down.

[Saagie for Data Engineers →](#)



S3-Compatible Read/Write to Tape!

Few tape storage devices in existence can work seamlessly with object storage and the cloud. Fujifilm, the world's largest manufacturer of data tape, has a new tape format for object storage called **OTFormat**. **OTFormat** allows for objects and metadata to be efficiently written and read to and from tape in native form.



Summary

- Customers are solving real-world problems using multi-cloud data management
- 3 core use-cases have been the focus for the last 2-3 years
- New requirements are emerging for **advanced/user-defined data management policies**
 - Please join our 10-minute lightning talk session this afternoon for an introduction
 - **User-Defined Cloud Workflow Policies: today 4:05 pm - 4:15 pm PDT.**



Thank You

Paul Speciale

Chief Product Officer, Scality
paul.speciale@scality.com



#sodacon202

1