Scalable, efficient and environmentally sustainable Long Term Digital Preservation of scientific datasets in the ARCHIVER project

Matthew Addis
Arkivum
3 May 2023
PV2023

arkivum
Bringing archived data to life

Agenda

- ARCHIVER
- Scalable LTDP in the Cloud
- Environmental Sustainability

ARCHIVER

ARCHIVER: The Project

***Focus**: Archiving and Data Preservation Services using cloud services available via the European Open Science Cloud (EOSC)*

***Procurement R&D budget**: 3.4M euro; Total Budget: 4.8M*

***Starting Date**: 1st of January 2019*

***Duration**: 42 Months*

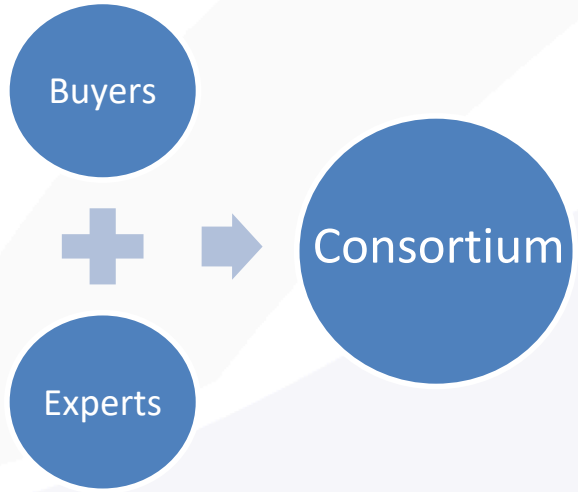
***Coordinator**: CERN (Lead Procurer)*

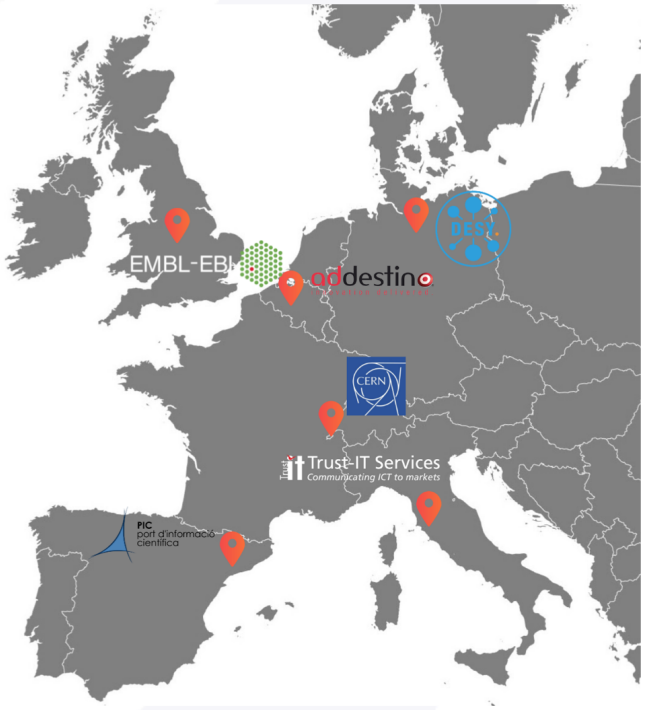
European Commission

EUROPEAN OPEN SCIENCE CLOUD

Buyers Group (BG) - Public organisations committing funds to contribute to a joint-R&D-procurement, research data use cases and R&D testing effort

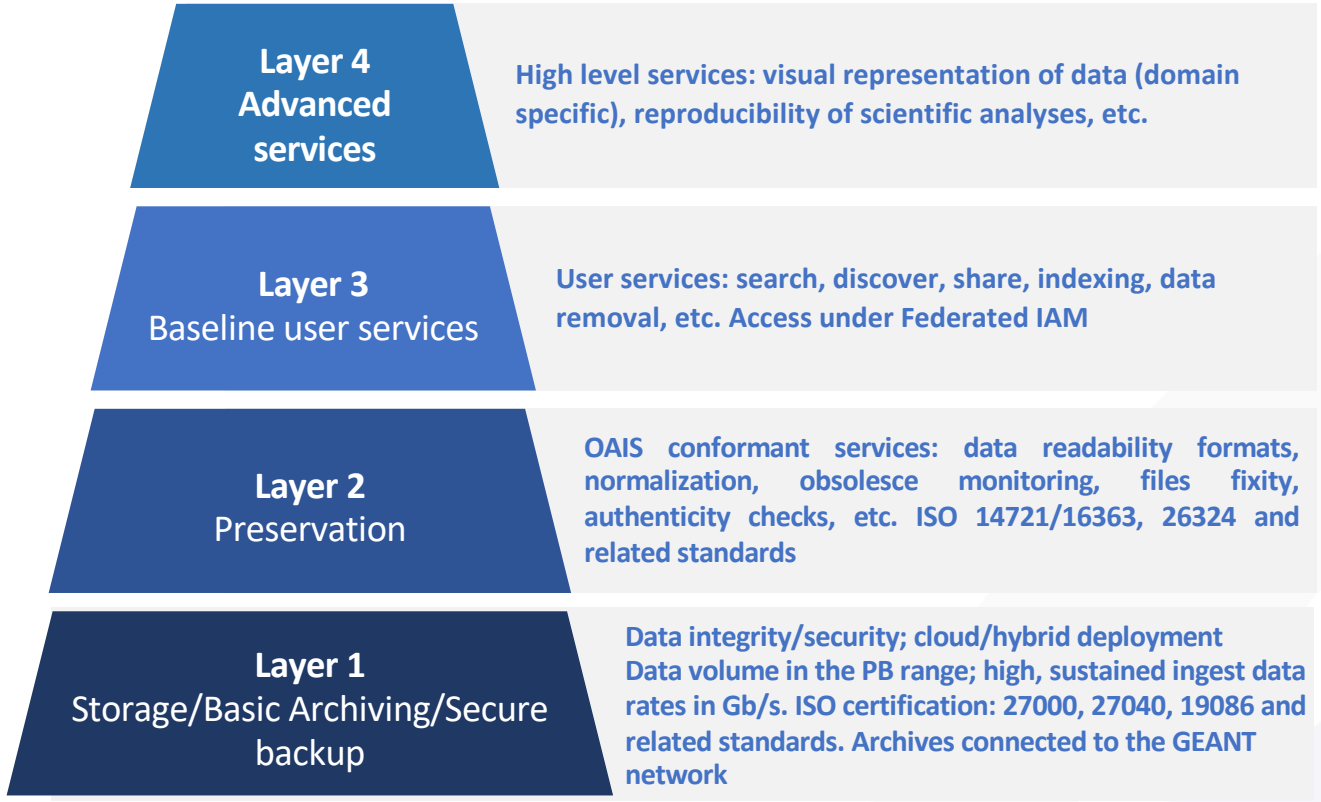
Experts - Partner organisations bringing expertise in requirement assessment and promotion activities

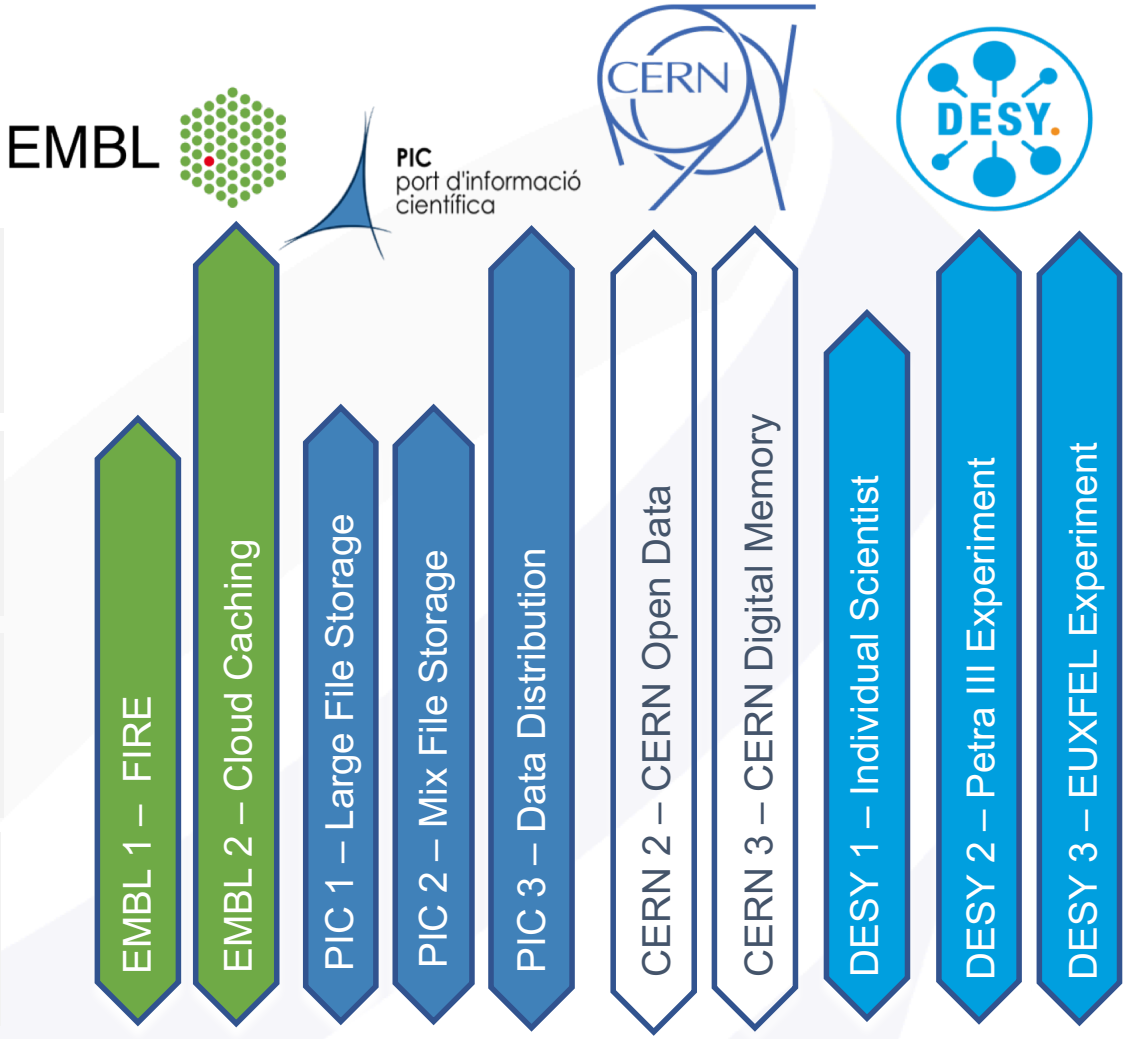




ARCHIVER: Buyers Group

ARCHIVER: Requirements and Use Cases





Scientific use cases deployments documented at: https://www.archiver-project.eu/deployment-scenarios

ARCHIVER "current state of the art" report in the context of the EOSC: https://doi.org/10.5281/zenodo.3618215

End Result

- Pilots of LTDP in the cloud at scale
- Ingest and Preservation rates of 100TB+ per day
- Scientific Data, Documents, Images, Video

Winner of the DPC 2022 award:

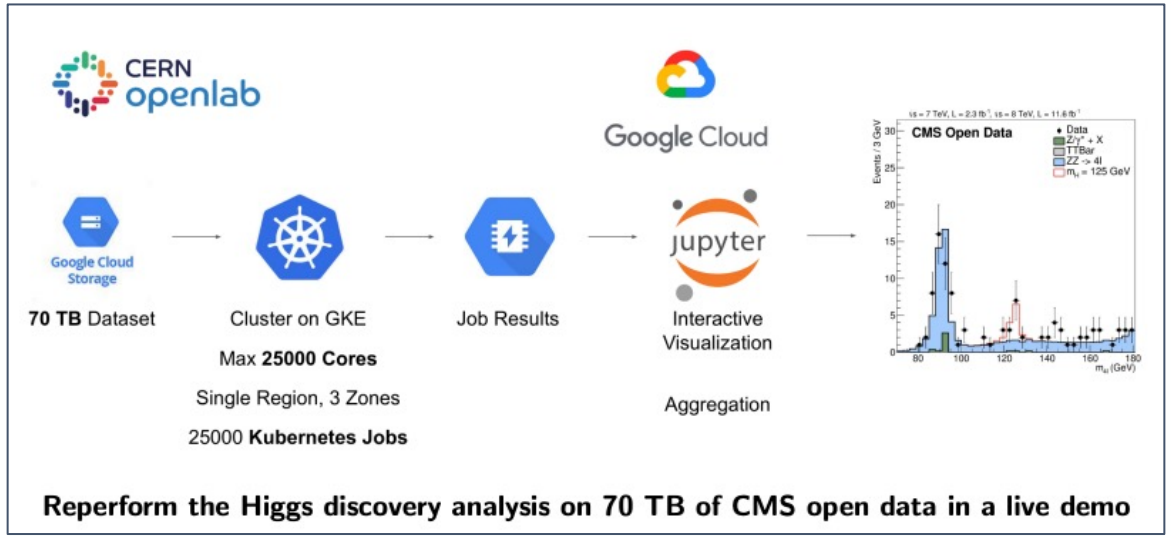
The International Council on Archives Award for Collaboration and Cooperation

https://zenodo.org/record/7691976

Arkivum Approach

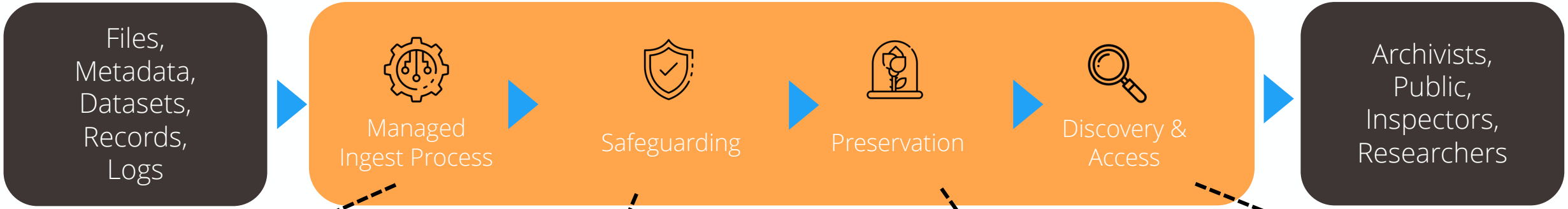
Partnership with Google

- GCP provides a highly scalable infrastructure (compute, storage, networking)
- Connected to GÉANT and NRENs across Europe
- GCP is widely used as a cloud platform for scientific data processing
- Low carbon footprint and good environmental credentials
- Discounts for education/research
- Arkivum is a Google Reseller
- Combined solution in ARCHIVER



https://indico.cern.ch/event/1008656/contributions/4232849/attachments/2192041/3705047/physics_analysis_google_meeting-16.02.2021.pdf

Arkivum SaaS Solution



Managed Ingest Process

- Multiple upload mechanisms
- Data encryption in transit
- Fixity checks for data integrity
- Data validation
- Metadata extraction
- Chain of custody
- Application integrations

Safeguarding

- Data stored in multiple locations
- Tiered storage
- Data immutability
- Automated data checks
- Retention Management
- Encryption at rest
- Data escrow with a third-party

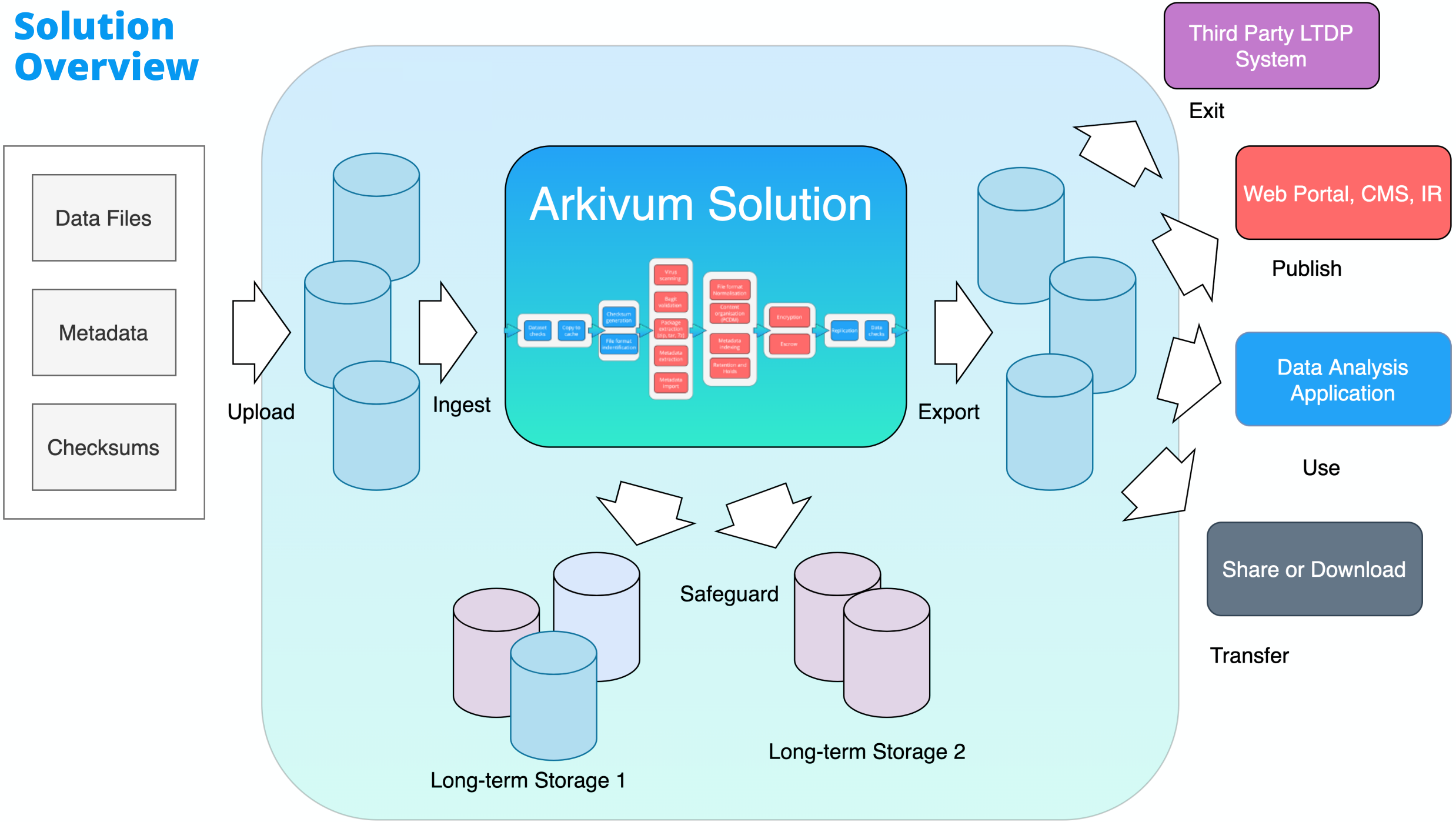
Preservation

- Normalisation to create preservation and access copies
- Archival Information Packages (AIP) in line with OAIS
- Manage preservation policies through Format Policy Registry
- Data renormalization on demand

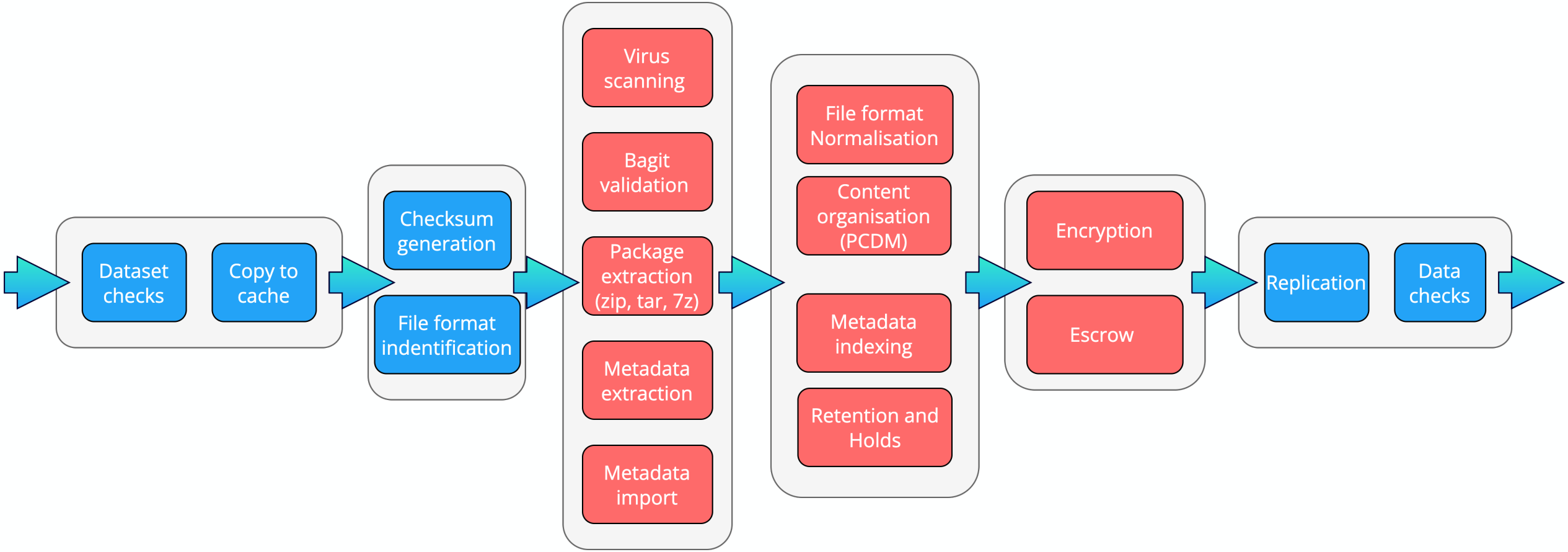
Discovery & Access

- User access provided via online dashboard
- Access control managed through integration with customer IT
- Advanced content search, leveraging metadata
- Browse, explore and navigate
- Download and export
- Application Integrations

Solution Overview



Content Processing Workflows



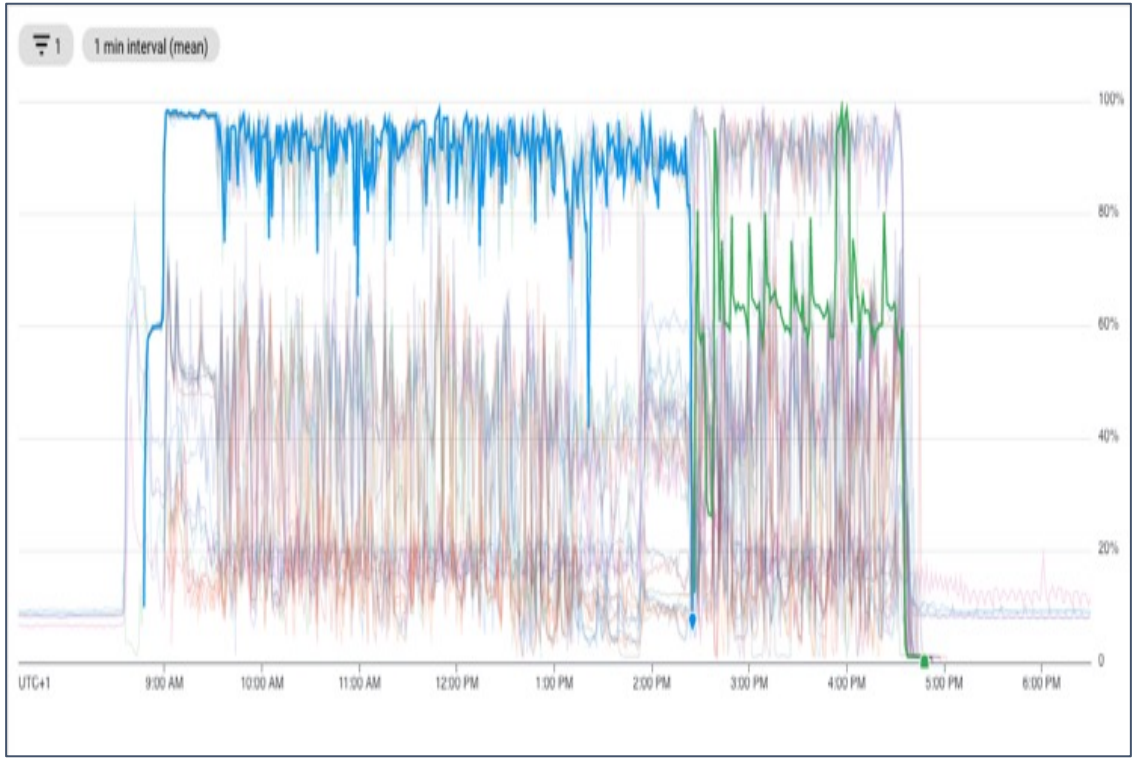
Serverless Computing: Scalability, Performance, Efficiency

- Scalable and cost-effective archiving workflows and processing
  - Kubernetes and autoscaling (keda)
  - Scale to zero as well as autoscaling for peak loads (pods and nodes)
  - Pre-emptible nodes to reduce costs (up to 70% lower)
  - Terraform and Ansible for provisioning
  - Rancher, Prometheus, Grafana, Kibana for monitoring and analytics

- Microservices
  - Checksum, virus scan, file format identification, caching, replication, unpack …
  - Stateless and able to run in parallel (can process 100TB+ per day)
  - Fault tolerant and can be restarted at any time
  - All jobs recorded and tracked
  - Communication via async messaging using Kafka
  - State held in a clustered Mongo database

Spot Instances

- Up to 90% cheaper than normal VMs
- Used on demand, e.g. in a Kubernetes cluster
- Can be reclaimed at any time!
  - 30 seconds notice

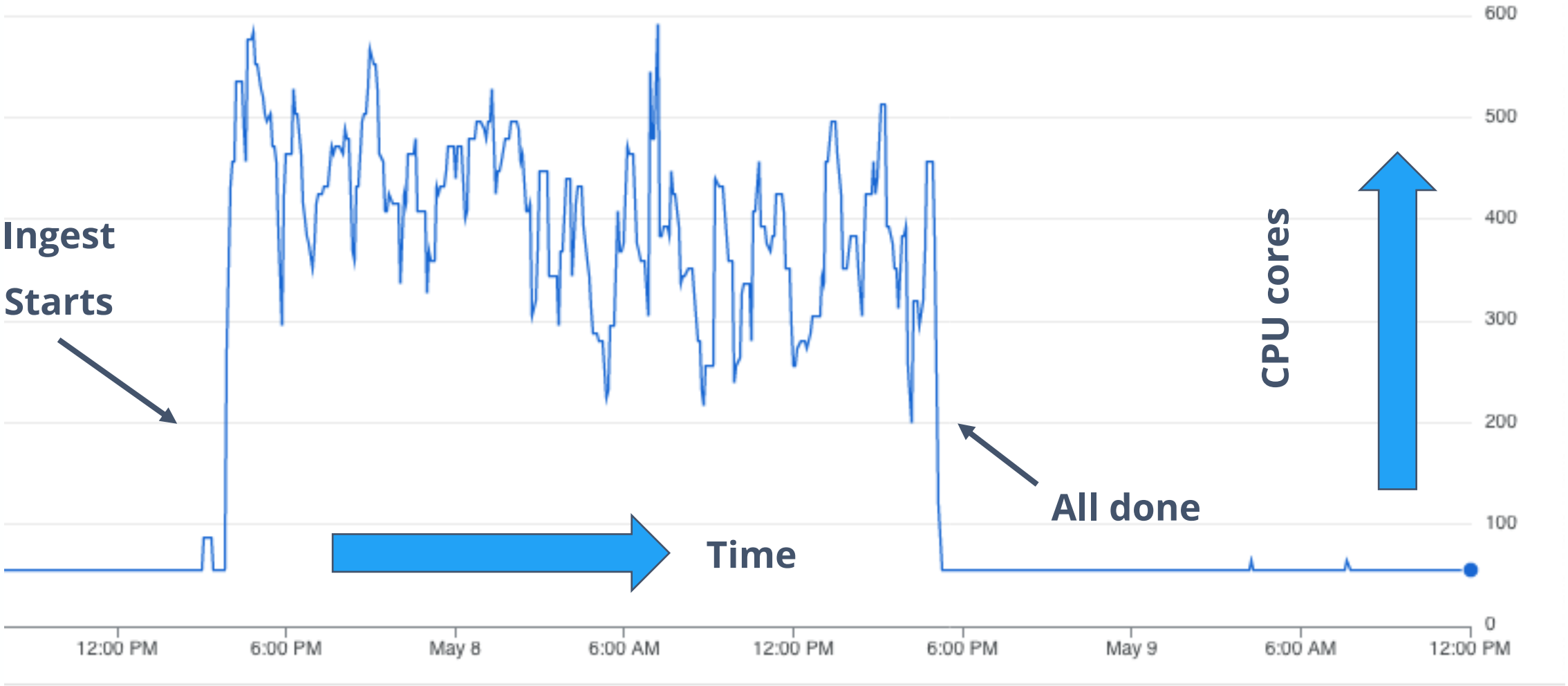


- Use resources that would otherwise be idle
- Increases utilization for the Cloud provider
- Reduce costs for LTDP

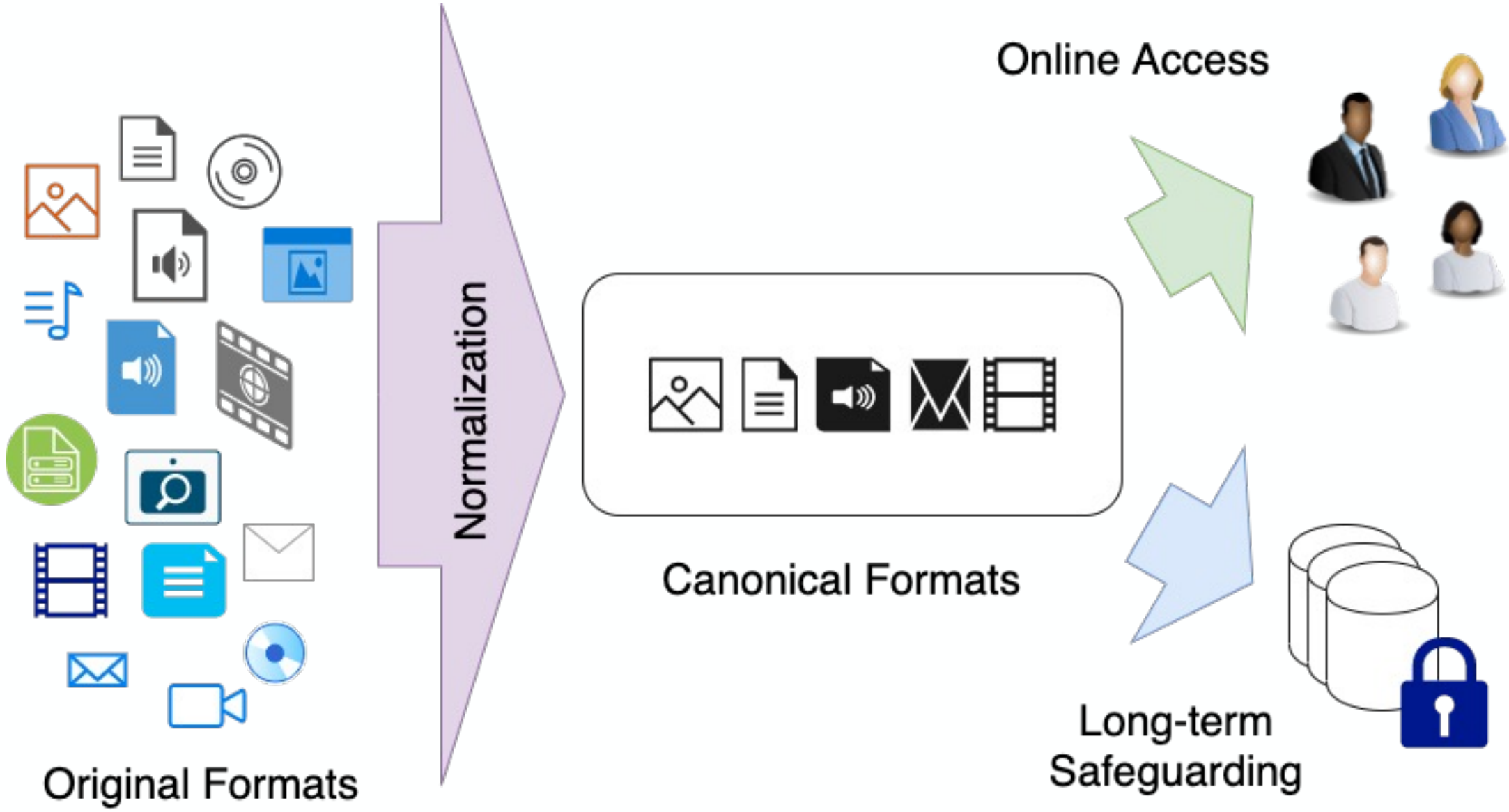
- But not easy: long running processes, timeouts, interdependent microservices ...

Results

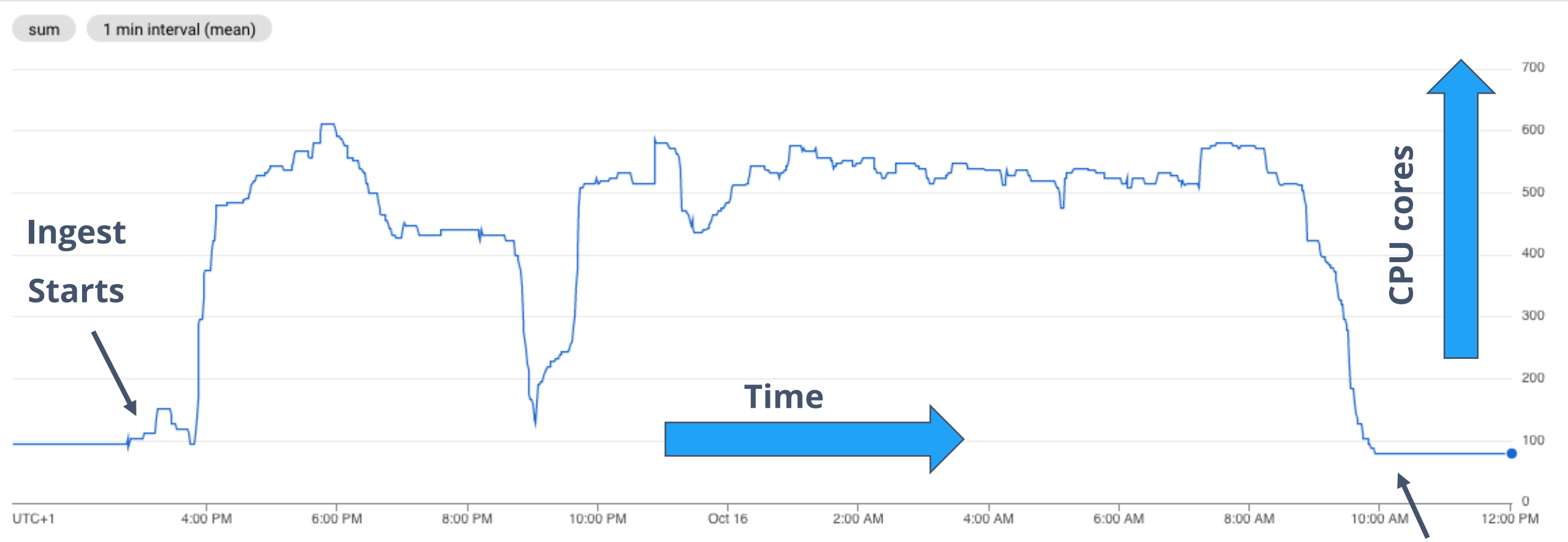
Example: Ingest and Archiving of Astronomy Datasets



440,000 image files (total 25TB) ingested as 2780 big data bags in approx 24hrs

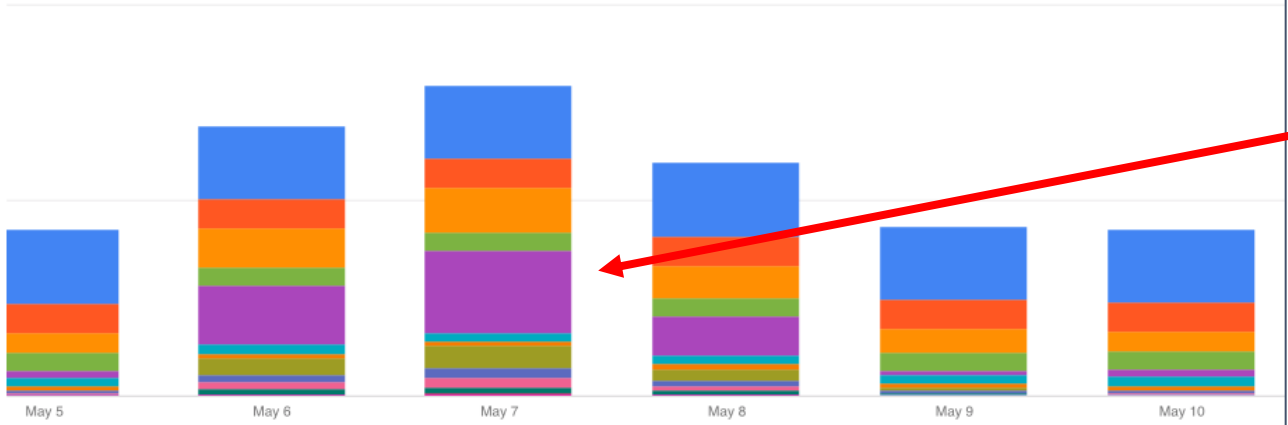

Online Access
Normalization
Canonical Formats
Original Formats
Long-term
Safeguarding

Example: Preserving Documents



- 349k GOV DOCS files in batches of 1000 files, including Archivematica to generate AIPs and DIPs
- 142k normalised files (Office file formats -> PDF/A)

GCP: Resource Consumption and Costs



SKU	Service	SKU ID	Usage	Cost	Discounts
Custom Instance Core running in Frankfurt	Compute Engine	47BE-44D0-C86F	6,720 hour	£195.89	–
Custom Instance Ram running in Frankfurt	Compute Engine	6971-C36E-2B52	20,160 gibibyte hour	£78.75	–
Spot Preemptible N2D AMD Instance Core running in Frankfurt	Compute Engine	E58E-BBDF-1B8A	11,978.32 hour	£77.70	–
SSD backed PD Capacity in Frankfurt	Compute Engine	378C-9DAC-0A5A	520.26 gibibyte month	£76.03	–
N1 Predefined Instance Core running in Frankfurt	Compute Engine	A9C0-BADB-1C34	1,680 hour	£49.02	–
N1 Predefined Instance Ram running in Frankfurt	Compute Engine	5C8C-3C2D-B331	6,300 gibibyte hour	£24.63	–
Spot Preemptible N2D AMD Custom Instance Core running in Frankfurt	Compute Engine	72BE-1F1F-40E8	3,107.99 hour	£20.16	–
Storage PD Capacity in Frankfurt	Compute Engine	7A7B-EA46-2897	356.3 gibibyte month	£12.25	–
Zonal Kubernetes Clusters	Kubernetes Engine	6B92-A835-08AB	168 hour	£12.03	–

Extra GCP resources consumed during an ingest and preservation job

Type of resource consumed (spot instances)

ID	Scenario	Characteristics	Original Test Case	Extra processing steps	Cost/TB ($)	CPU core-hours per TB
IN1	DESY bulk ingest	Large ingests (10,000+ files) of simulated scientific data. Average file size 500MB.	50TB. 100,000 files. Largely homogenous ingest of HDF5 data files.	None	4.93	42
IN2	EBI bulk ingest	Genomics data with a wide range of file sizes (few kB up to 100s GB). Ingest in batches with a batch size ranging from 500GB to 34TB.	146 ingests. Most ingests were 500GB in size consisting of approx. 100 files each. File sizes range from few kB to over 100GB. A few of the ingests were large, e.g. one of the ingests was 34TB, with some files over 600GB in size.	None	1.57	43
IN3	PIC bulk ingest (Star)	Large number of relatively small files. Ingest in small batches of approx. 100 files.	10,000 ingests. 7TB. 2M files. Majority of files were images in the 5-10MB range.	None	10.1	NA
IN4	PIC bulk ingest (Calibrated)	Large number of medium sized files. Ingest in small batches of approx. 100 files.	As above for PIC Star dataset, but file sizes are larger and total data volume is 70TB.	None	3.27	412
IN7	Video files (Archivematica)	Ingest of single video files from CERN digital memory datasets.	247 files. 6.5TB in total.	Files processed using Archivematica (video normalised for access). Metadata extraction from each video and indexing of this metadata using Elastic Search.	46.2	788
IN8	Mixed small files. British Library digbooks and GOV DOCS	British Library jpeg image files. GOV DOCs wide range of document formats (Word, PDF, PPT, HTML etc.).	British Library: 216,000 files. 88GB. GOV DOCS: 347,000 files. 291GB.	Metadata extraction from each file and indexing of this metadata using Elastic Search.	46.8	4,685
IN9	IN8 plus (1) AIP and DIP generation using Archivematica and (2) AIP safeguarding.	IN8 plus normalised files produced by Archivematica.	British Library: 216,000 files. 88GB. GOV DOCS: 347,000 files. 291GB. Normalised GOV DOCS files: 142,000 files. 91GB	Files processed using Archivematica (document file types normalised for preservation and for access). Metadata extraction from each file and indexing of this metadata using Elastic Search.	675	51,204

GCP: Reporting of Gross Carbon Emissions

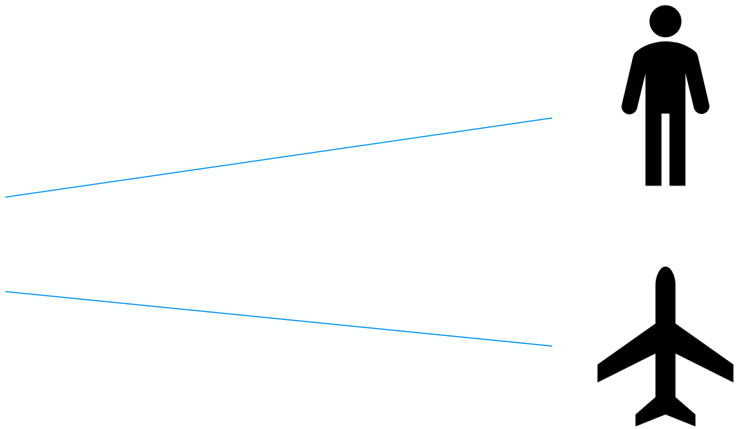
Carbon Footprint Methodology

- Get resource consumption and carbon emissions from Google reports
  - CPU resource consumption over 5 months (core-hours)
  - Storage consumption over 5 months (GB-months)
  - Gross emissions over 5 months per resource type (kgCO2 eq)
- Calculate metrics
  - kgCO2 eq per core-hour for compute
  - kgCO2 eq per TB-year for storage
- Measure resource consumption for specific preservation workflows (storage, compute)
  - Large files, small files, inside bagit bags, big ingests, lots of small ingests
  - File format identification, checksum generation, metadata extraction, replication etc.
  - Additional processing using Archivematica on-demand
- Calculate carbon emissions
  - kgCO2 eq per TB of data ingested for different scenarios

Gross Carbon Emissions From GCP Energy Consumption

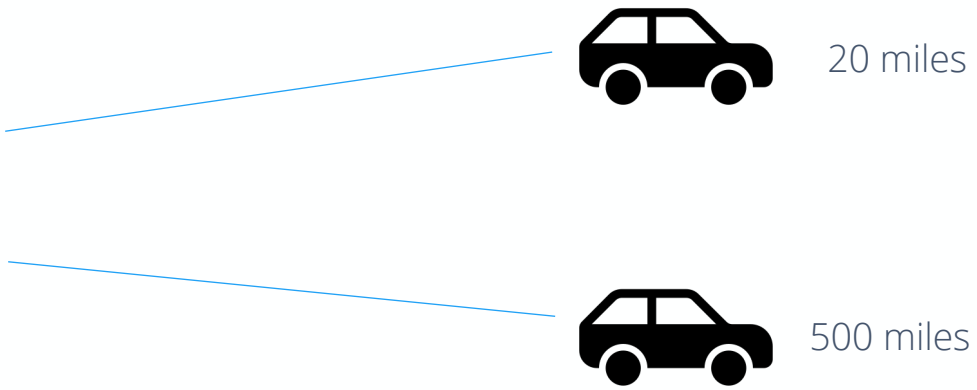
Large Astronomy Research Datasets

	GCP Frankfurt
1 PB data stored for 1 year	7800 kgCO2 eq
1 PB ingest of large image files	1600 kgCO2 eq



Large collections of office files

	GCP Frankfurt
1M office files stored for 1 year	5.5 kgCO2 eq
Ingest of 1M office files.	140 kgCO2 eq

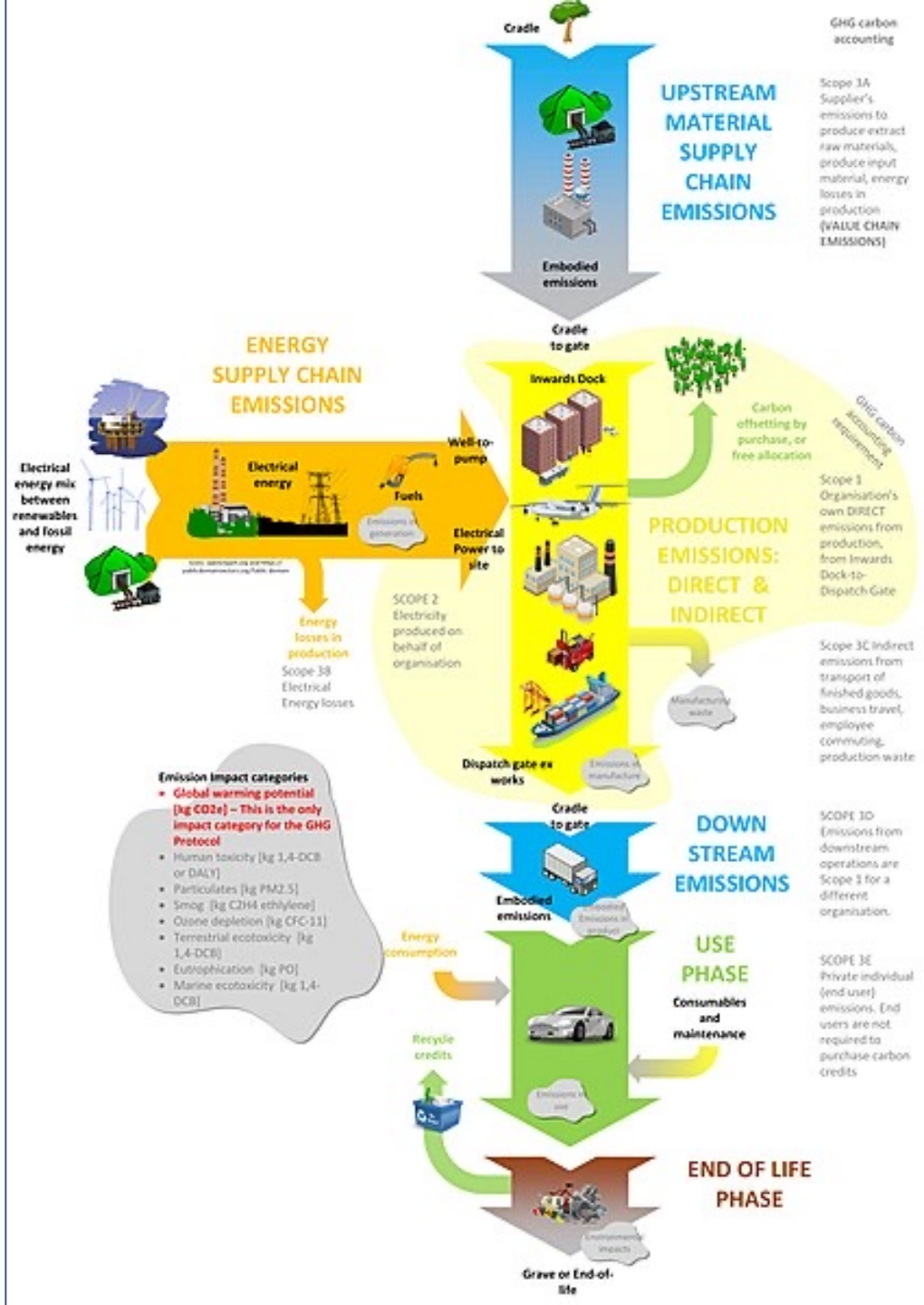


The net carbon emissions were zero!

Embodied Footprint

- Life Cycle Assessment (LCA)
- ISO 14040
- Cradle to Grave
- Raw materials, manufacturing, transportation, use, disposal, recycling

- The use stage (carbon footprint from energy used in the cloud) is a small part of the ICT lifecycle



https://en.wikipedia.org/wiki/Life-cycle_assessment#/media/File:Life_cycle_analysis_and_GHG_carbon_accounting.jpg

Embodied Footprint of Servers and Storage

- Storage
  - SEF (Storage Embodied Factor): kgCO2eq per GB
  - HDD lifetime 4-6 years
  - Hard Drive ~ 5 kgCO2eq per TB per year

- Servers
  - Cloud server lifetime 4-6 years
  - Cloud server utilization 50 – 65%
  - 1 core-hour ~ 0.5 gCO2eq

https://arxiv.org/pdf/2207.10793.pdf

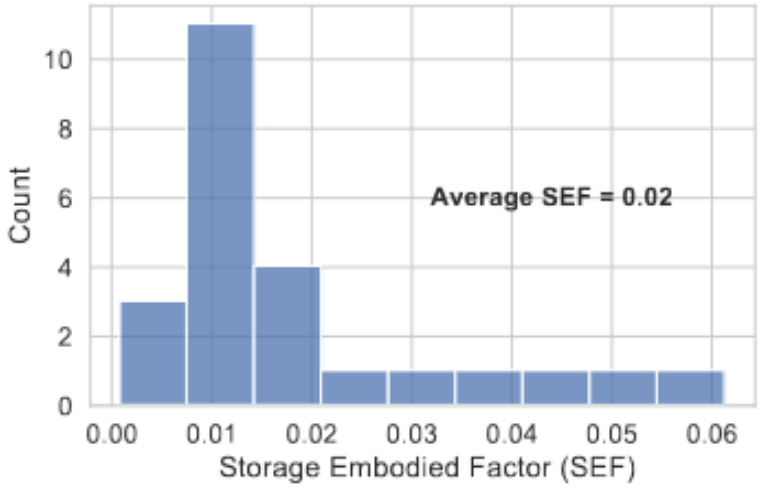


Figure 6: Distribution of estimated storage embodied factor for 24 Hard Drive Disks (HDDs).

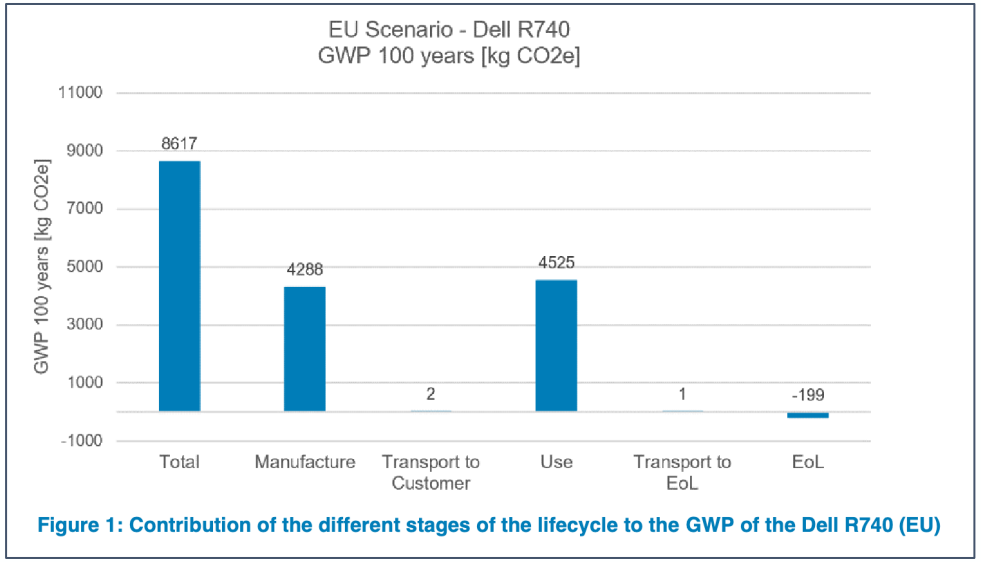


Figure 1: Contribution of the different stages of the lifecycle to the GWP of the Dell R740 (EU)

https://corporate.delltechnologies.com/content/dam/digitalassets/active/en/unauth/data-sheets/products/servers/lca_poweredge_r740.pdf

Overall Carbon Emissions

Large Astronomy Research Datasets

	Gross Emissions GCP Frankfurt (measured)	Embodied Footprint (estimated)
1 PB data stored for 1 year	7800 kgCO2 eq	4000 kgCO2 eq
1 PB ingest of large image files	1600 kgCO2 eq	200 kgCO2 eq

Large collections of office files

	Gross Emissions GCP Frankfurt (measured)	Embodied Footprint (estimated)
1M office files stored for 1 year	5.5 kgCO2 eq	4 kgCO2 eq
Ingest of 1M office files.	140 kgCO2 eq	25 kgCO2 eq

The net carbon emissions from energy use were zero!

Summary

Summary

- LTDP in the cloud using serverless computing
- 100TB per day throughput, PB scale storage
- K8s with scale-zero and spot instances to minimize costs
- Benchmarked resource usage and costs for different LTDP scenarios
- Quantified carbon footprint using LCA analysis and GCP emissions reporting

Ingest	**$2.2 per TB**
	0.3 kgCO_2eq
Long-term storage	**$30 per TB-year**
	5 kgCO_2eq per TB-year

MAGIC Telescope Calibrated tarbags estimated costs for GCP Finland using cold storage

More Info

- ARCHIVER Whitepaper:
  - https://zenodo.org/record/7691976
- Webinar:
  - https://arkivum.com/webinar-recording-archiver-sustainable-preservation-of-petabyte-scale-scientific-data/
- Demo and Tutorial
  - https://www.archiver-project.eu/sites/default/files/ARKIVUMWebinar-Presentation20220623.pdf
  - https://www.youtube.com/watch?v=312I6ljVvz8
- Arkivum solution on EOSC
  - https://marketplace.eosc-portal.eu/services/arkivum-digital-archiving-and-preservation-solution

arkivum
Bringing archived data to life

QUESTIONS?

Matthew Addis

matthew.addis@arkivum.com

https://orcid.org/0000-0002-3837-2526