



Cloud Cost Management at Scale

Summer ESIP 2024

Ben Williams
EED-3 Solution Engineer
benjamin.j.williams@nasa.gov



Cloud Cost Management at Scale

- Motivation
- Our Approach
 - Estimation
 - Monitoring
 - Optimization
 - Protection

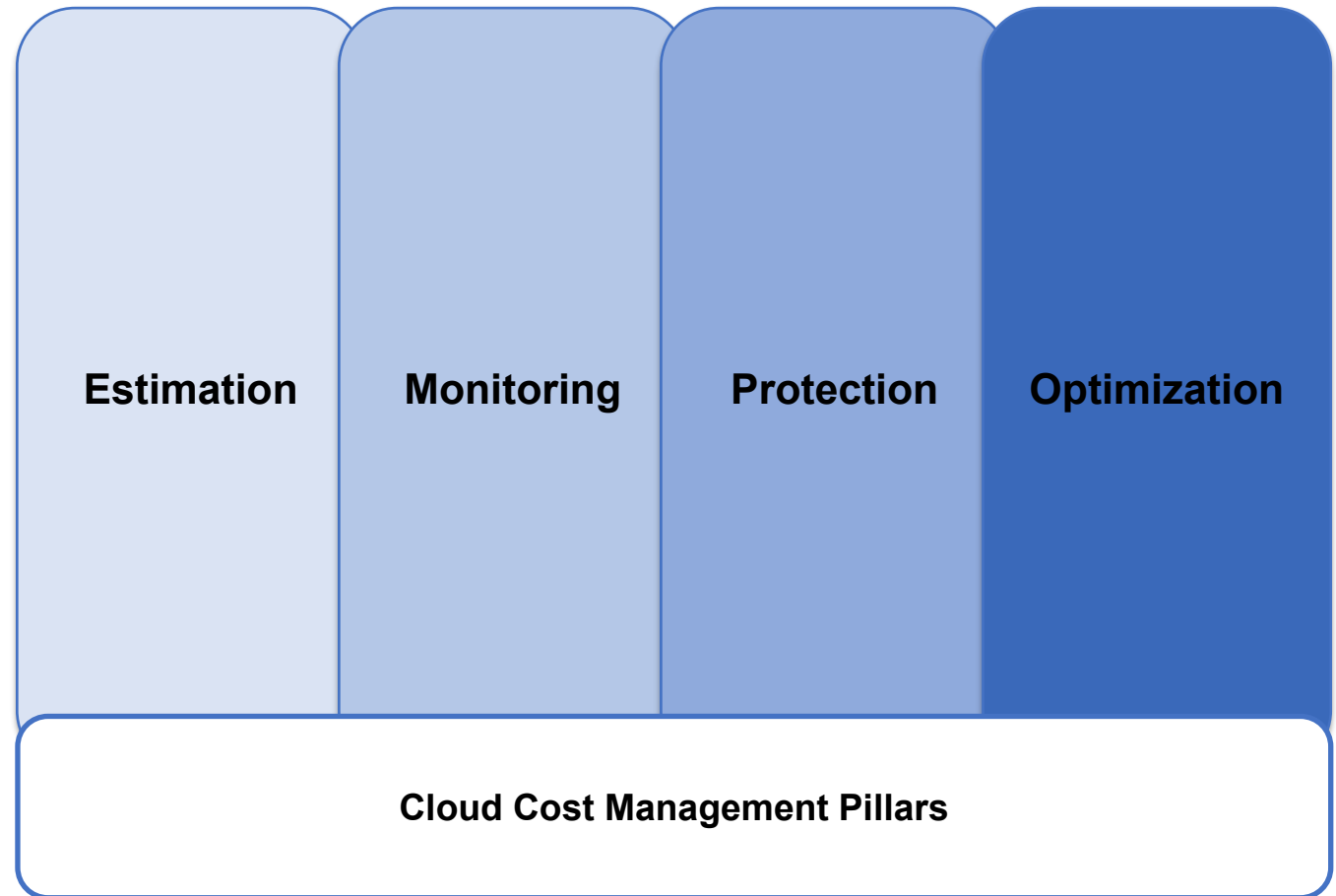
Motivation

- The NASA Earthdata Cloud (EDC) is a large and complex cloud computing ecosystem
- With our scale and growth, we had to level-up our approach to cloud cost management...
- Cost Threats:
 - Unintentional - Ex: infinite loops in dev/test
 - Unexpected - Ex: demand surge during natural disaster
 - Malicious - Ex: DDoS

EDC Attribute	Scale Metric
Developer Accounts	over 350
AWS Accounts	over 250
Published Endpoints (https://...)	over 400
S3 Object Storage (Petabytes)	over 100
EC2 Instances	over 1,000
AWS Services in use	over 65

Cloud Cost Management Approach

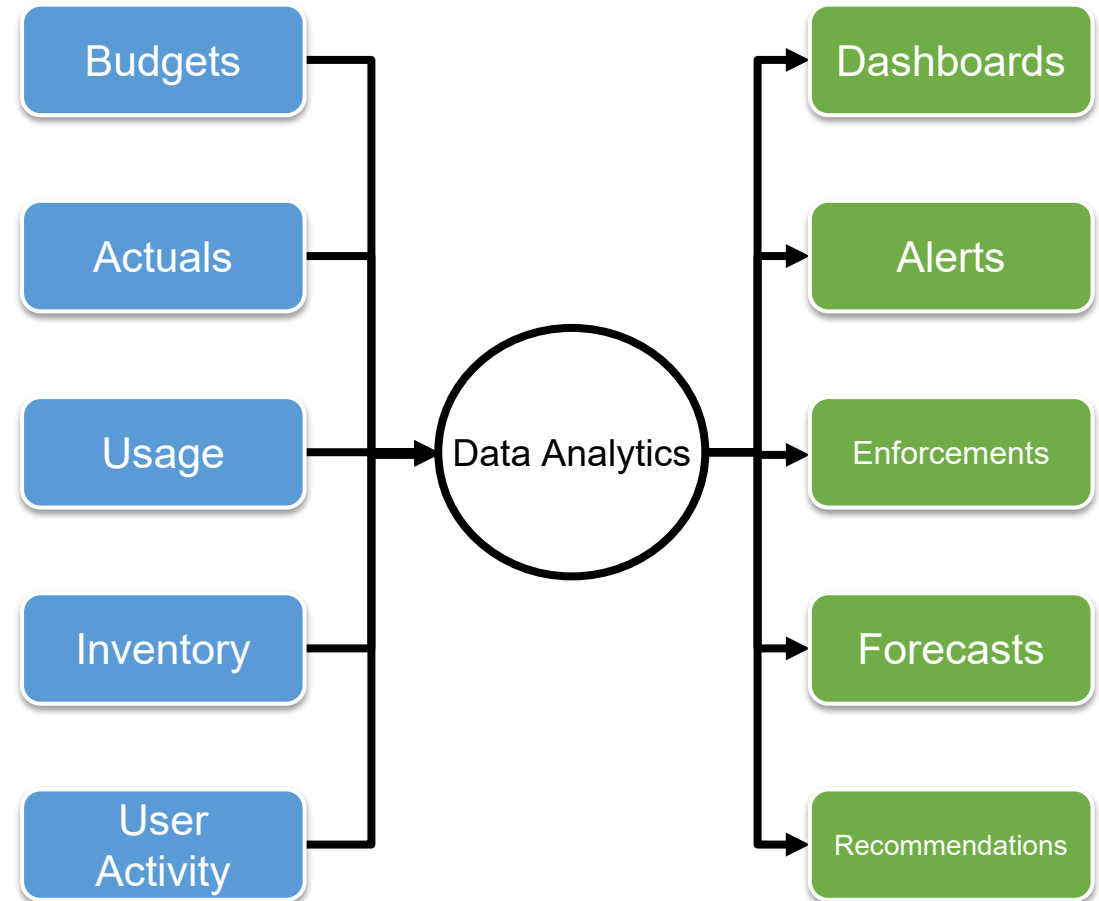
- **Estimation**
 - Processes and tooling to generate increasingly accurate estimates of future cloud costs
- **Monitoring**
 - Observability and analytics over historical and near-real-time cloud cost actuals
- **Protection**
 - Analytics and automation to avoid cloud cost overruns
- **Optimization**
 - Methods for identifying cost reduction opportunities and mitigate cloud cost waste



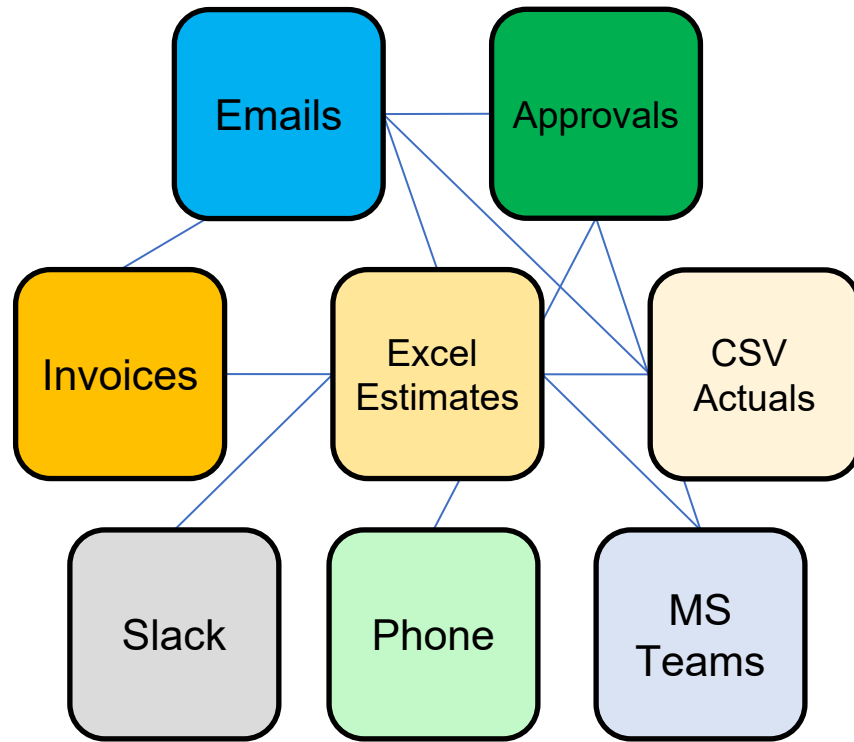
Cost Intelligence and Event Management (CIEM)

Treating cost data just like any other data source

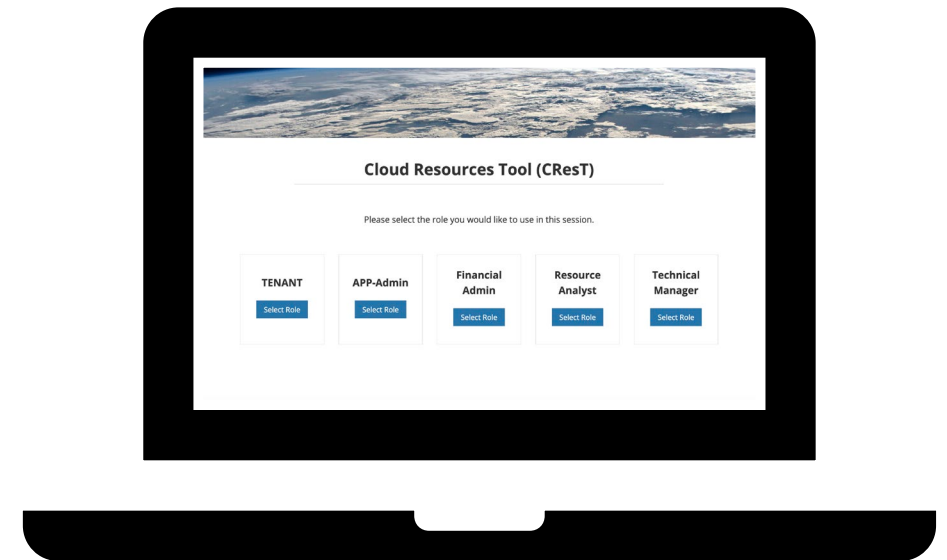
- elevate community practice of cloud cost optimization to the same level of tooling and scrutiny as cyber security
ex: Security Information and Event Management (SIEM)
- Cost analytics, recommendations, and alerts
 - **Comparative analytics:**
 - cost actuals vs. forecasts
 - trends and outliers
 - cost vs. 'sibling' app cost
 - **Anomaly Detection:**
 - cost / usage spikes
 - capacity overruns
 - **Cost Efficient Configuration Compliance:**
 - flag / alert for S3 vs. S3-IA usage
 - over-provisioned instances / volumes



Estimation: Cloud Resources Tool (CResT)

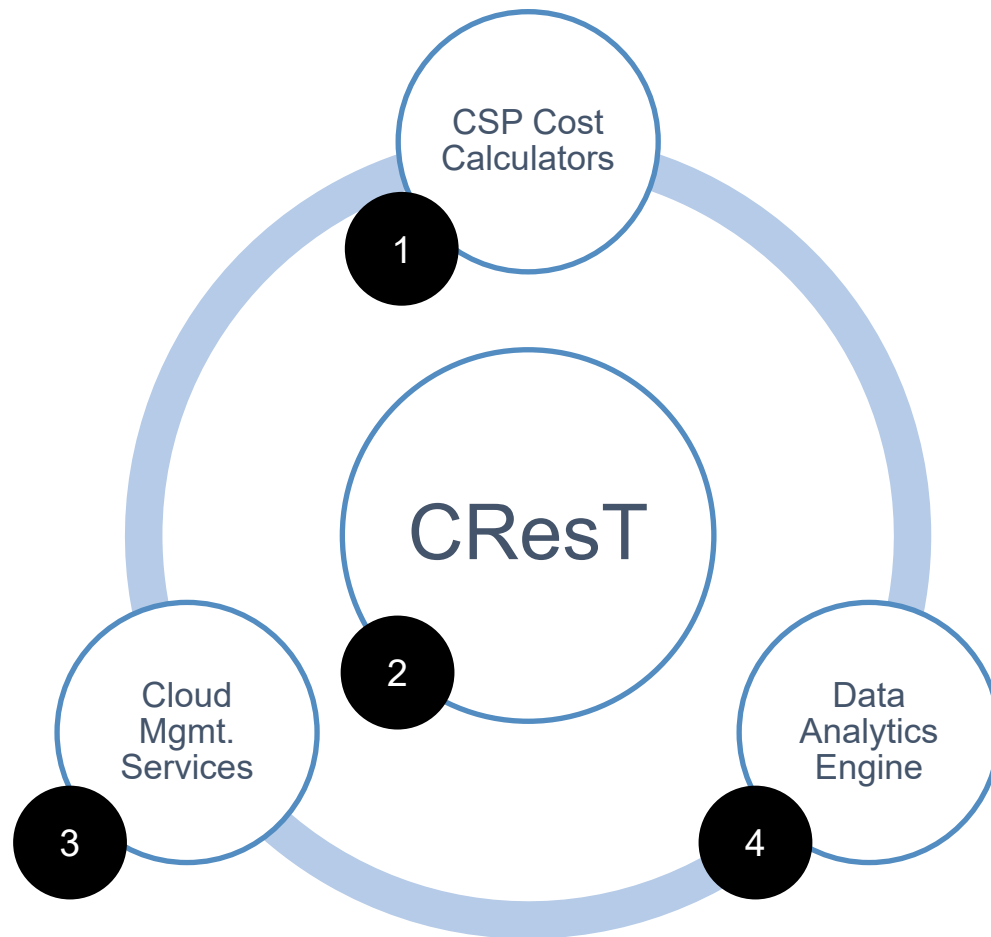


Where we started



Where we are today

Estimation: Cloud Resources Tool (CResT)



1. Basis of estimate

AWS Pricing Calculator

(<https://calculator.aws>)

2. Budget estimate and approval workflows

CResT

3. Automated cost overrun protection & budget tracking

Kion

(<https://kion.io>)

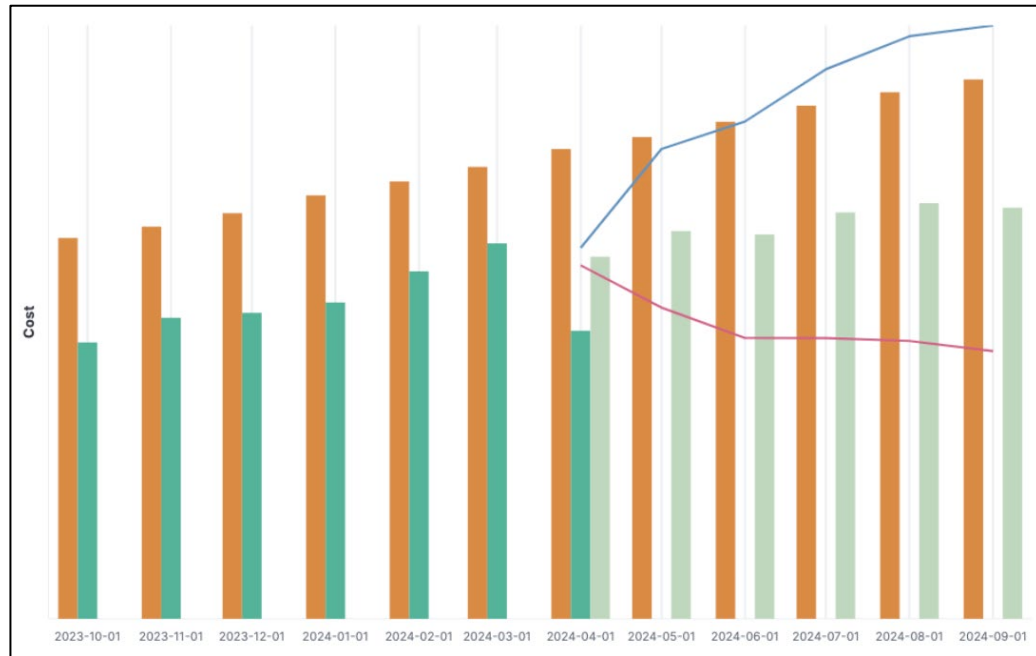
4. Cost analytics & optimization

ElasticSearch

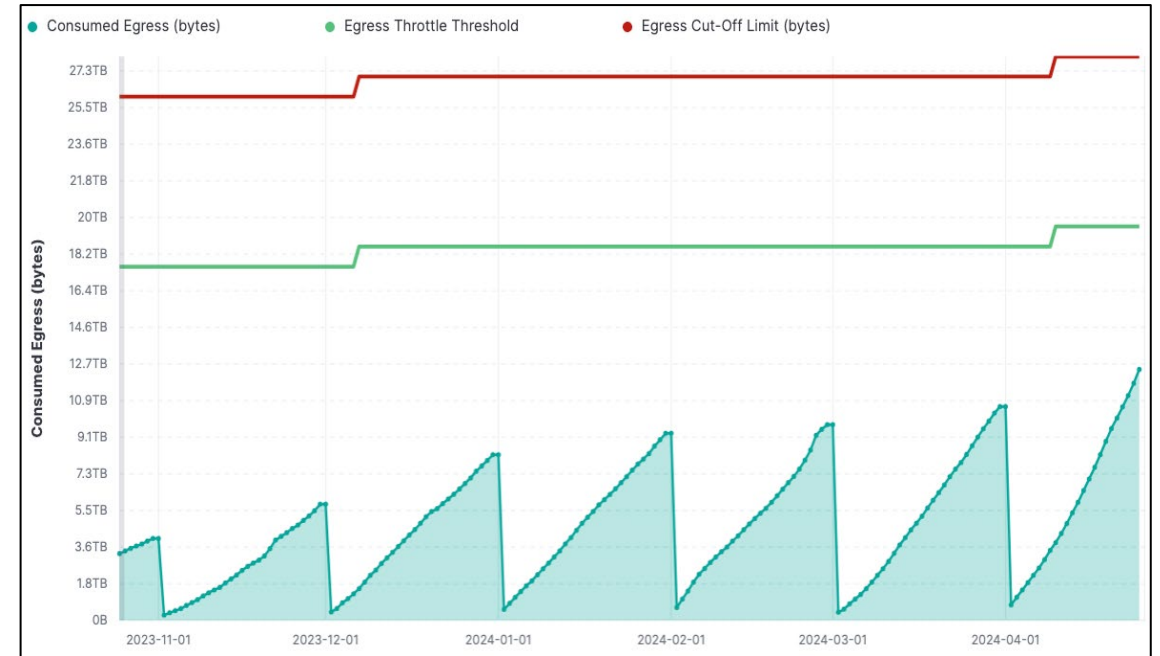
(<https://www.elastic.co>)

Monitoring: Budget vs. Actuals

Integrated views enabled by CResT



ML-derived cost forecasts vs. approved budgets



Egress monitoring and rate-limiting protections

Protection: Automated Cost Overrun Preventions

Back-stop safety net for rapidly incurred costs

- **Overall Account Budget Limits:**
 - **Freeze Spend:** disable spawning of new resources
 - **Circuit Breaker:** suspend compute / network resources
 - while ensuring no permanent loss of science data
- **Egress Distribution Limits:**
 - **Throttle:** rate limit distribution prior to reaching approved egress caps
 - **Cut-Off:** suspend egress distribution networking at approved egress caps

Both with push notifications in advance of automated actions

These automations have helped reduce the impact of major unexpected cost surges

...but have also triggered for less urgent cost events

Active Investigation Areas:

- Adaptive / anomaly-based triggers
- Programmatic inclusion of management reserves
- Finer-grained and targeted rate-limiting
- Build-time cost warnings / alerts

Optimization: Prioritize efforts by cost drivers

S3 Storage:

- Storage Class Tiering

EC2 Instances:

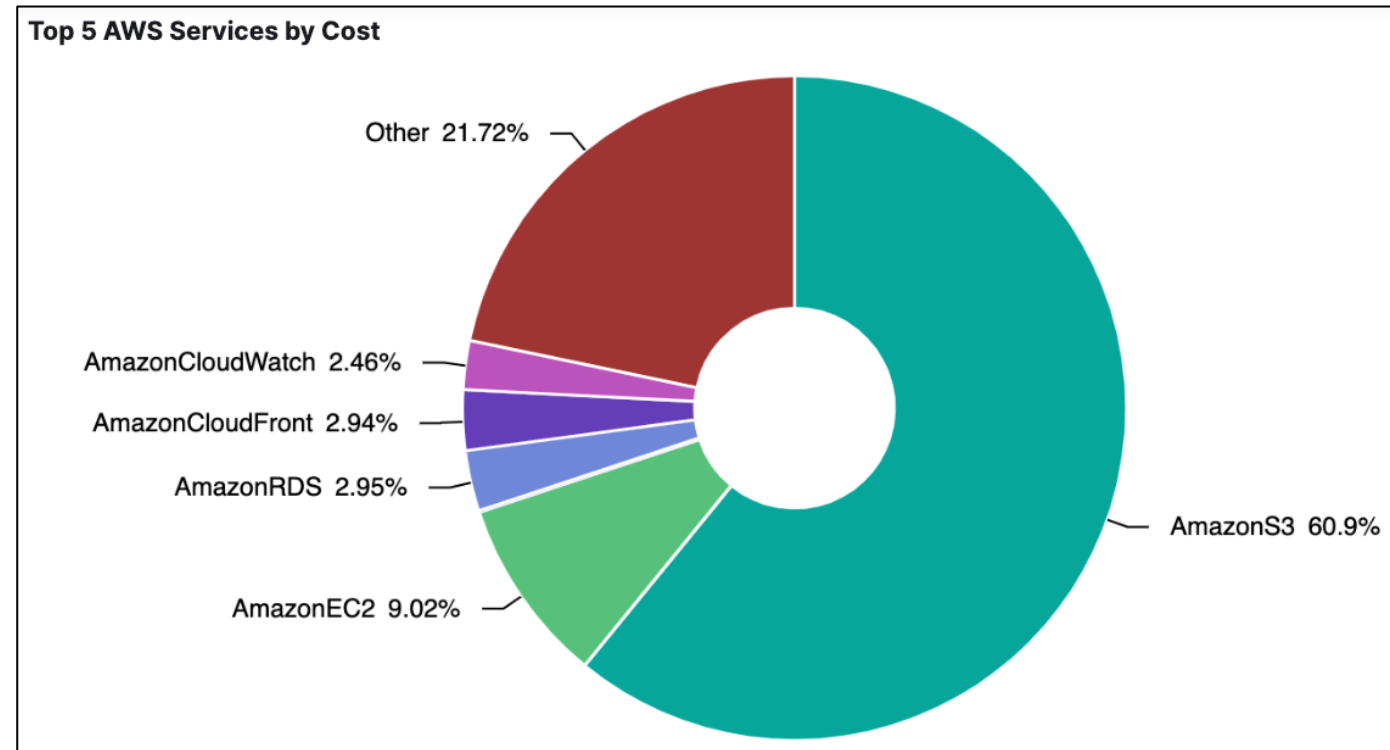
- Reserved Instances | Savings Plans
- Right-Sizing
- Off-Hours Suspension (non-prod)

RDS Instances:

- Reserved Instances
- Right-Sizing

CloudFront:

- Discount programs research institutions



Cost Distribution by AWS Service – past 12 months

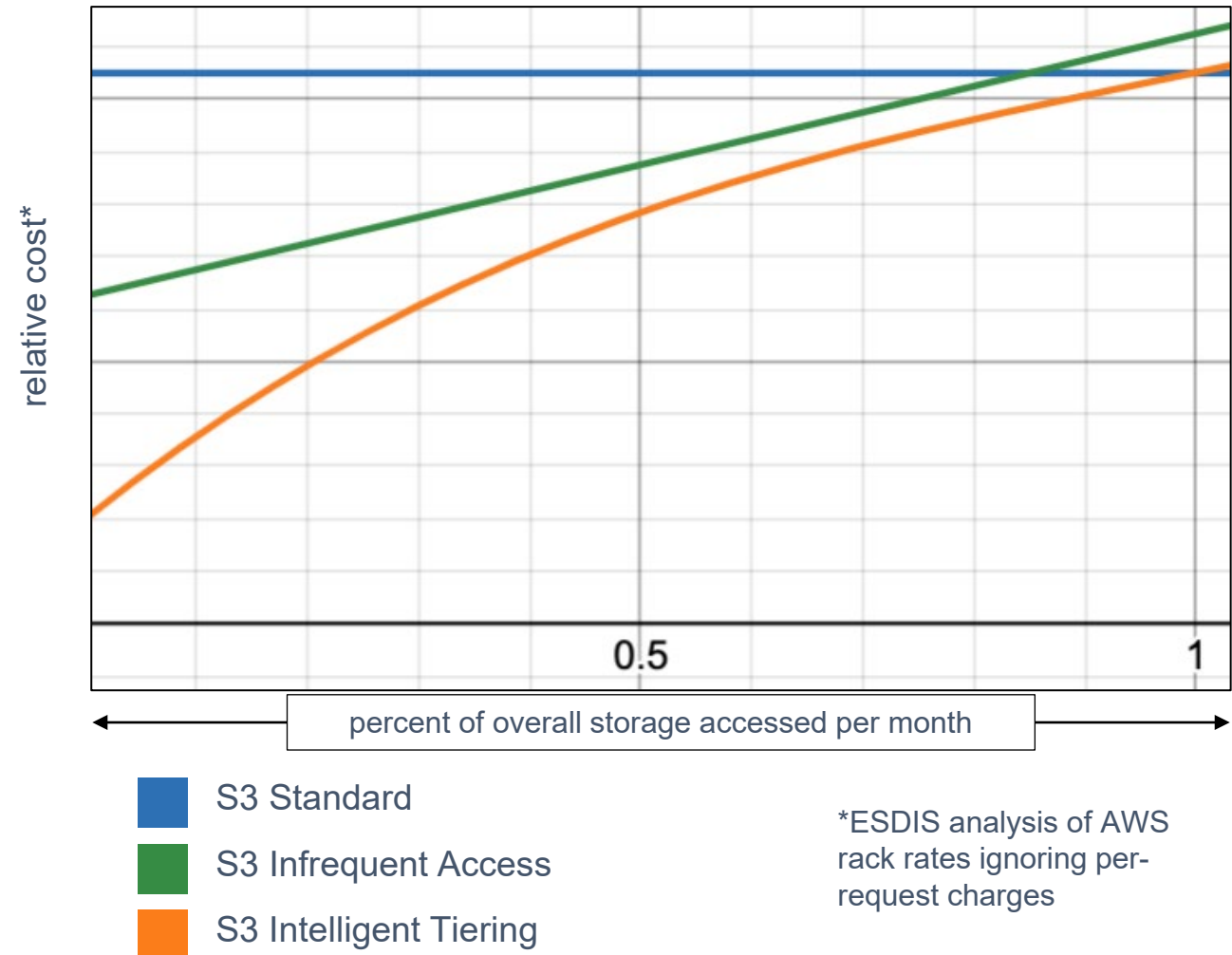
Optimization: Focus – S3 Storage Classes

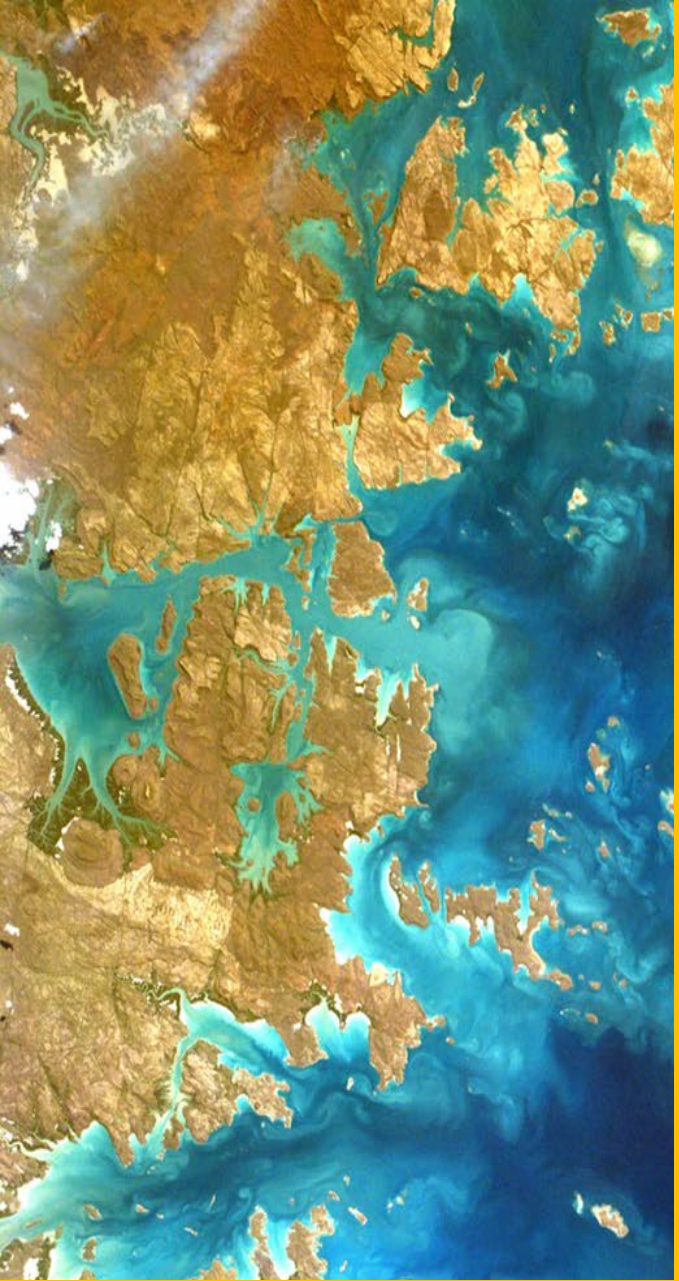
AWS provides multiple object storage tiers with varying cost contributors

- S3 Standard
- S3 Infrequent Access
- S3 Intelligent Tiering
- ...and many more

Along with tools to help profile and analyze storage tiers for best cost efficiency

- Storage Lens
- Storage Class Analyzer





This work was supported by NASA/GSFC under
Raytheon Company contract number
80GSFC21CA001