
AWS
re:Invent

STG207

Better, faster, and lower-cost storage: Optimizing Amazon S3

Matt Sidley

Senior Product Manager
Amazon S3

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Geospatial or lunar imagery

Compliance records

Analytics

Internet of Things (IoT) sensor data

Medical imagery and records

Media master files

Data lakes

Customer call-center records

Digital record preservation

Mobile sync and storage

Home-recording video

Origin storage for CDN

Seismic and reservoir simulation data

Pharmaceutical study data

Durable backups

DNA sequences

Surveillance video/closed-circuit television

Media assets

ML training data

Financial transaction records

Website hosting

Meteorological and environmental research

User-generated content

Autonomous vehicle data

Mapping data

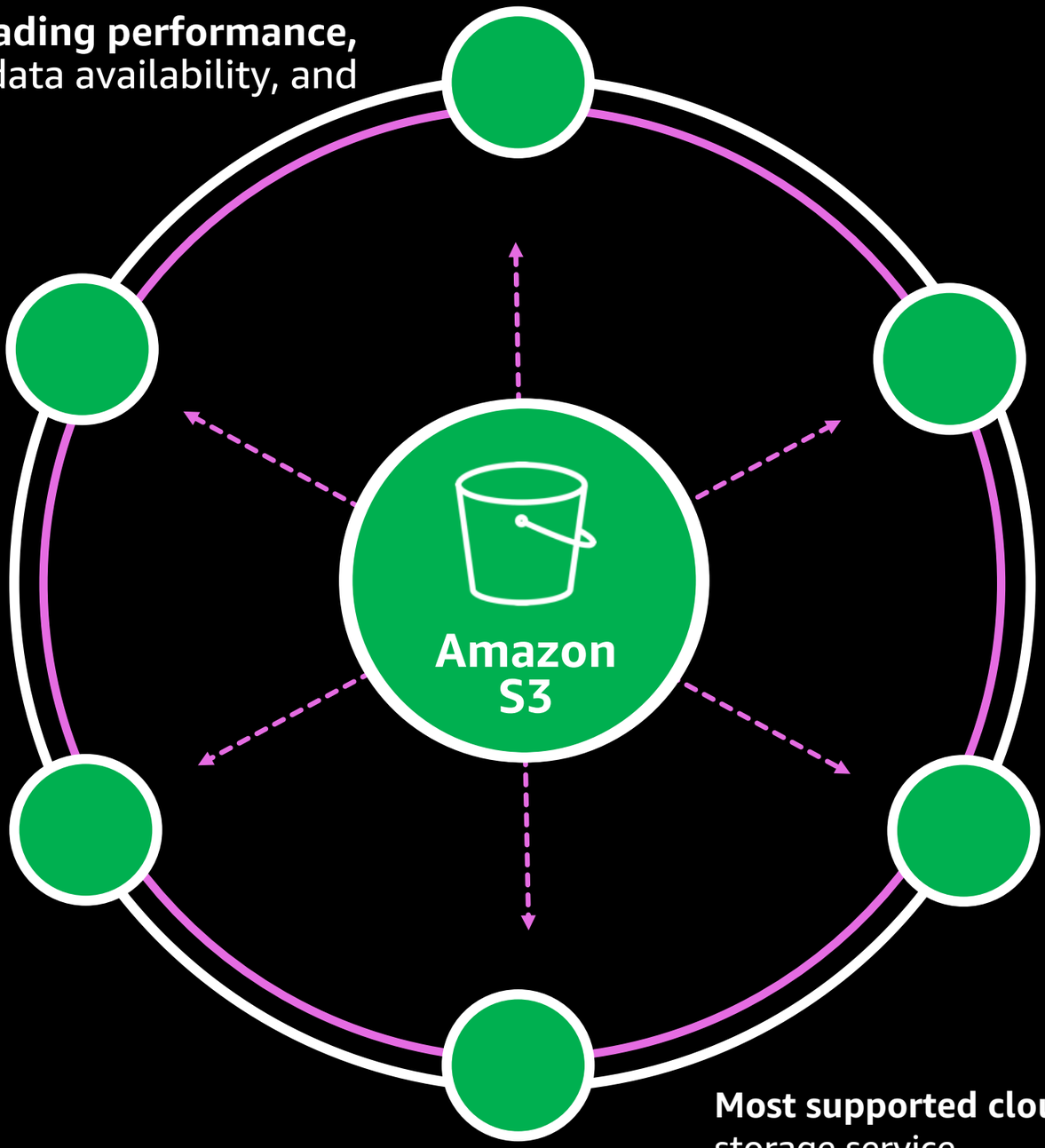
Log files

Oil and gas topography

Industry-leading performance, scalability, data availability, and durability

Wide range of cost-effective storage classes

Unmatched security, compliance, and audit capabilities

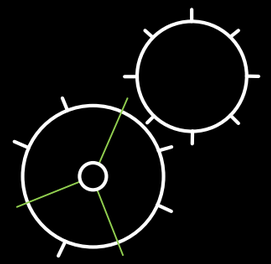


Easily manage data and access controls

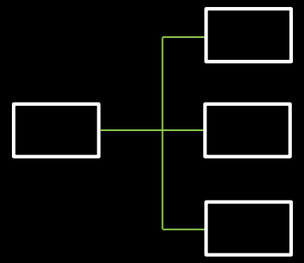
Query-in-place services for analytics

Most supported cloud storage service

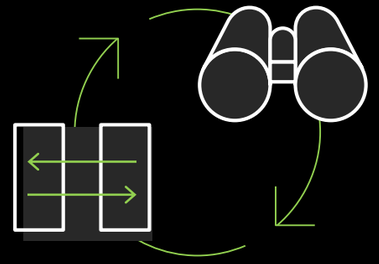
Pillars of optimizing storage performance and costs using Amazon S3



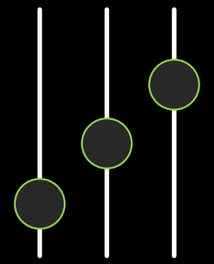
Application requirements



Data organization

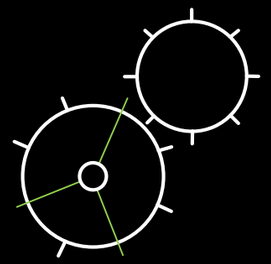


Understand, analyze, and optimize

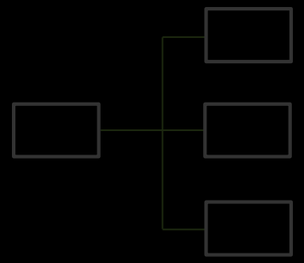


Continuous right sizing

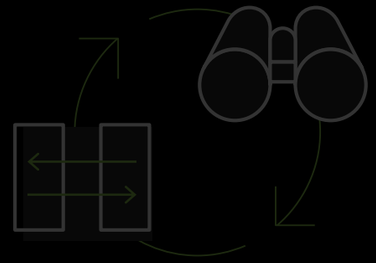
Pillars of optimizing storage performance and costs using Amazon S3



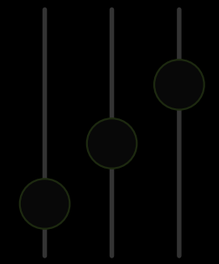
Application requirements



Data organization

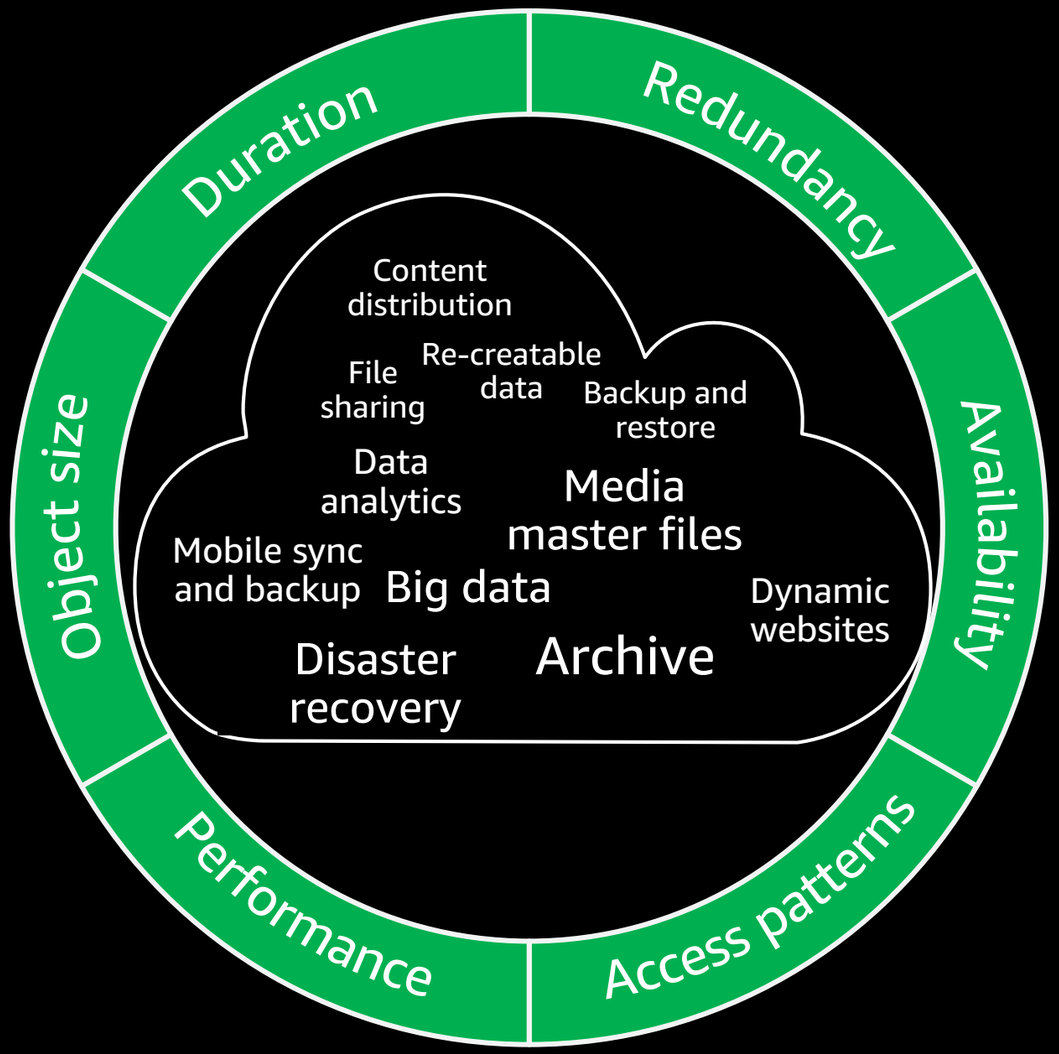


Understand, analyze, and optimize

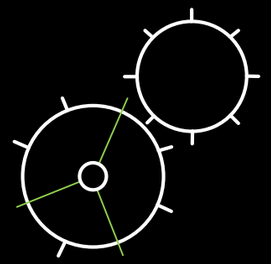


Continuous right sizing

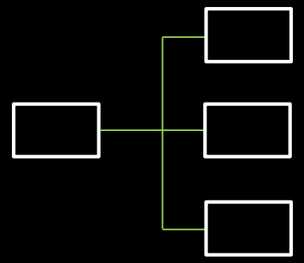
Define application requirements



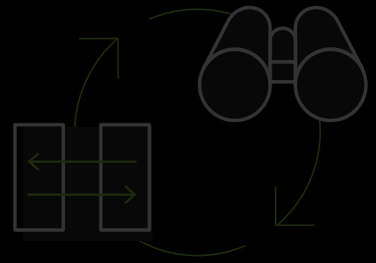
Pillars of optimizing storage performance and costs using Amazon S3



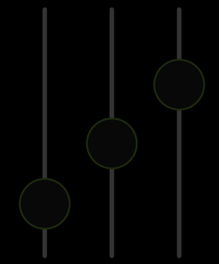
Application requirements



Data organization



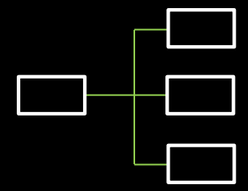
Understand, analyze, and optimize



Continuous right sizing



Organize data by buckets, prefixes, and object tags

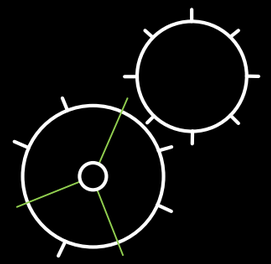




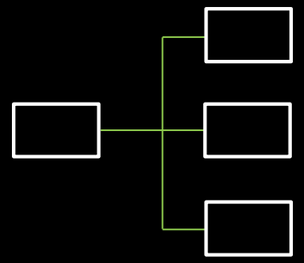
Account

Buckets	Prefixes	Object Tags
 Business Unit 1	Unit1/internal/usage/ Unit1/internal2/billing/usage Unit1/internal3/free/usage	Department = Finance Team = Business Reporting Classification = Confidential
 Business Unit 2	Unit2/model1/inference Unit2/model2/inference Unit2/predictions/prod	Department = Marketing Team = Data Science Classification = Confidential

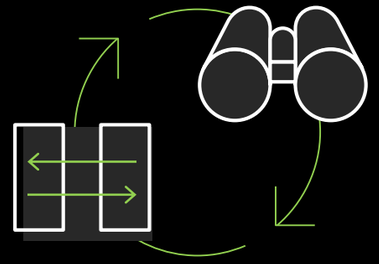
Pillars of optimizing storage performance and costs using Amazon S3



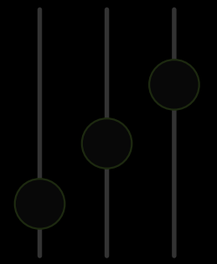
Application requirements



Data organization



Understand, analyze, and optimize

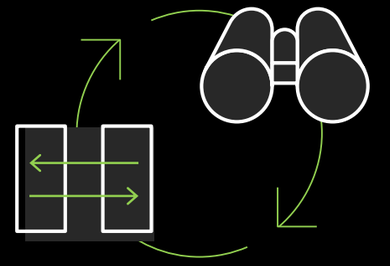


Continuous right sizing





Understand, analyze, and optimize



Amazon S3 Storage Lens



Amazon S3 Storage Class Analysis



Amazon S3 Inventory



Introducing Amazon S3 Storage Lens



Metrics – 17 usage metrics, 14 activity metrics (31 total)

Aggregation levels – organization, account, region, storage class, bucket, and prefix

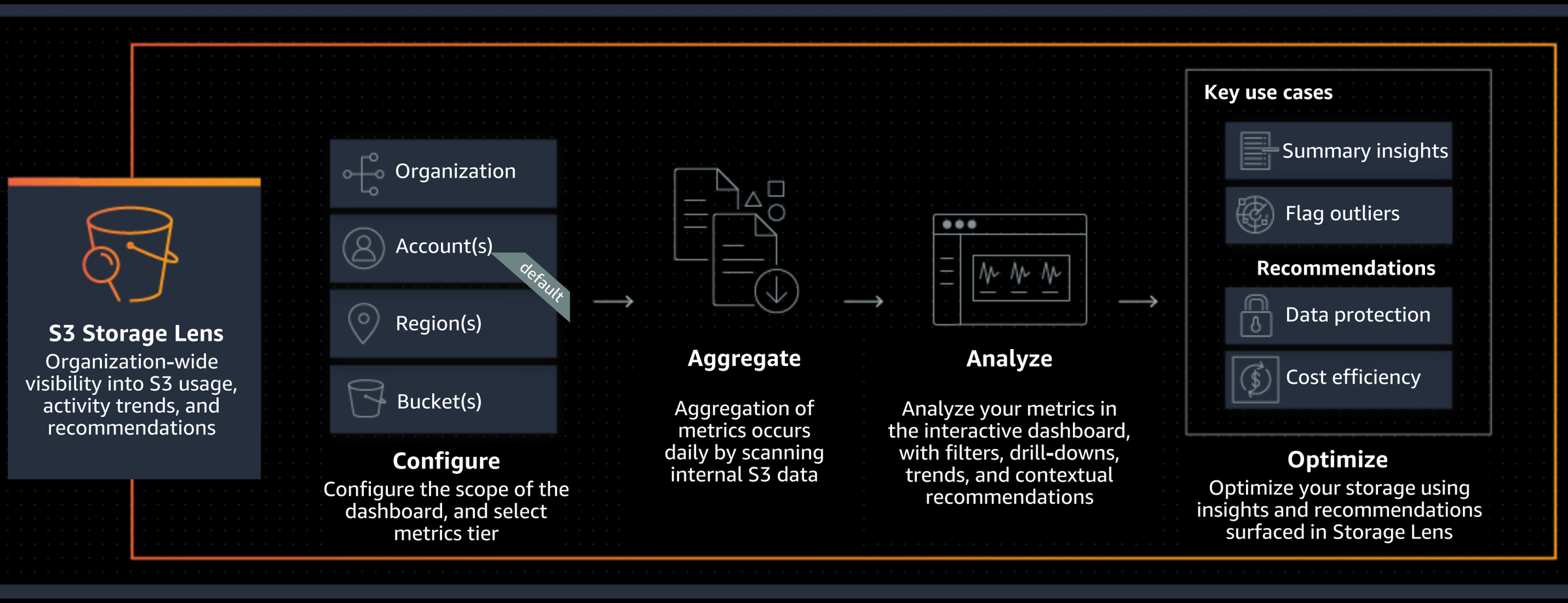
Refresh – daily updates (through previous day)

Data retention – up to 15 months

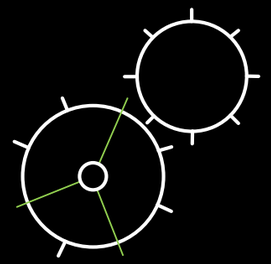
Visualization – interactive dashboard in the Amazon S3 console (export to Amazon S3 bucket also available)

Recommendations – metrics-driven contextual recommendations to highlight best practices

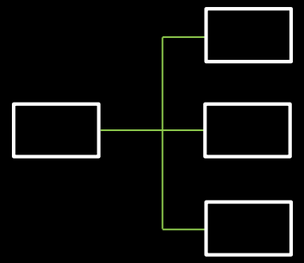
Amazon S3 Storage Lens: How it works



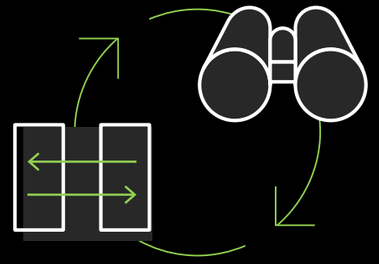
Pillars of optimizing storage performance and costs using Amazon S3



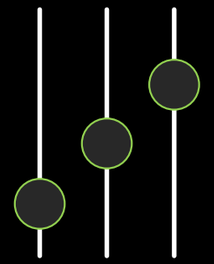
Application requirements



Data organization



Understand, analyze, and optimize



Continuous right sizing

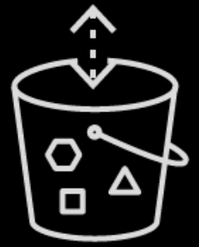
Your choice of Amazon S3 storage classes



S3 Standard



S3 Intelligent-Tiering



S3 Standard-IA



S3 Glacier



S3 Glacier Deep Archive



S3 One Zone-IA



S3 Outposts

AWS Region ≥ 3 AZ

- Active, frequently accessed data
- Milliseconds access

- Data with changing access patterns
- Opt in for Automatic Archiving
- Milliseconds access
- No retrieval fees
- Object monitoring fee
- Minimum storage duration

- Infrequently accessed data
- Milliseconds access
- Retrieval fee per GB
- Minimum storage duration
- Minimum object size

- Archive data
- In minutes and hours
- Retrieval fee per GB
- Minimum storage duration
- Minimum object size

- Long term Archive data
- Select hours
- Retrieval fee per GB
- Minimum storage duration
- Minimum object size

AWS AZ

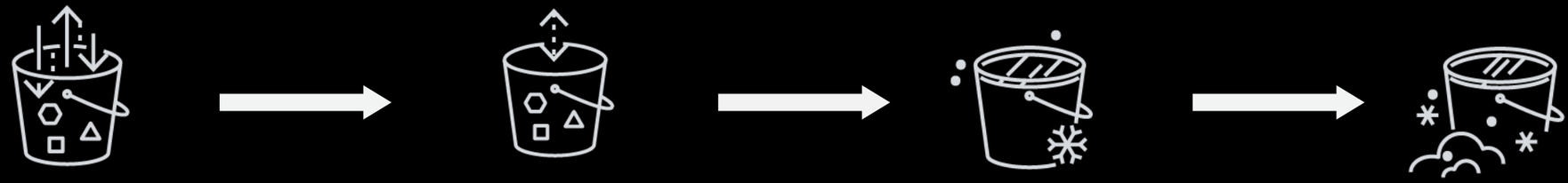
- Re-creatable, less accessed data
- Milliseconds access
- Retrieval fee per GB
- Minimum storage duration
- Minimum object size

AWS Outposts

- On-premises data
- Milliseconds access
- Encrypted with SSE-S3

Continuous right sizing to optimize your storage costs

Storage Classes and Lifecycle Management for predictable access patterns



S3 Standard

S3 Standard-IA

S3 Glacier

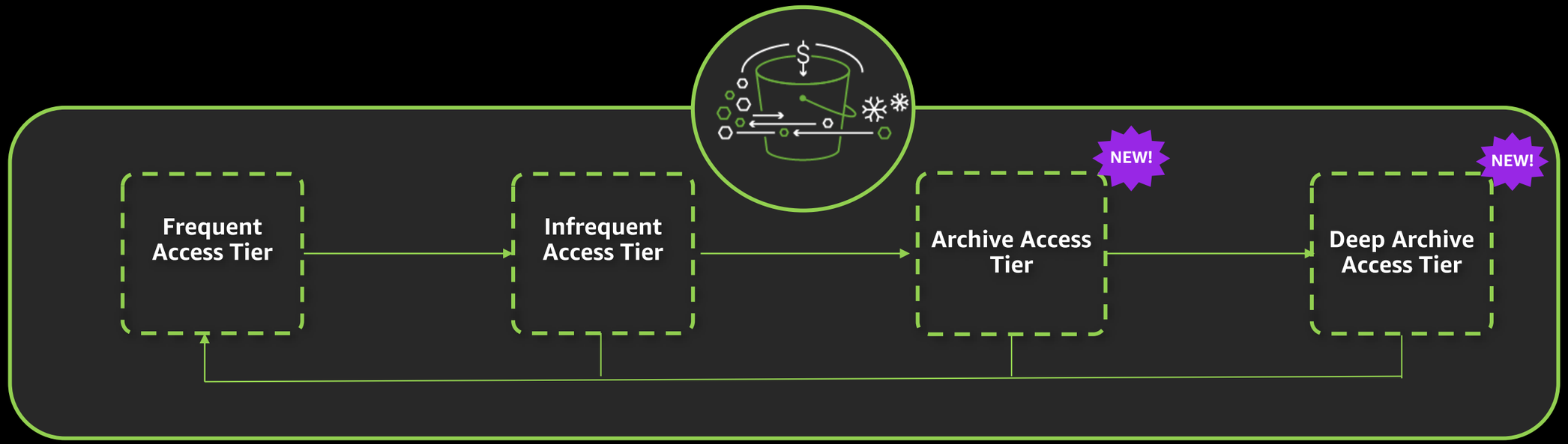
S3 Glacier
Deep Archive

Frequent

Access frequency

Archive

S3 Intelligent-Tiering for changing or unknown access patterns



Set S3 Lifecycle policy to tier and expire storage



Use S3 Storage Class Analysis to identify storage age groups that are less frequently accessed

S3 Lifecycle policy to tier to lower-cost storage classes and expire objects; Fine-tune analysis by bucket, prefix, or object tag

For example, transition objects older than 180 days to S3 Glacier, and objects older than 365 days to S3 Glacier Deep Archive

Use Storage Classes and Lifecycle Management for predictable access patterns

Amazon S3 Intelligent-Tiering Storage Class



Only cloud storage that delivers **automated cost savings**

Monitors and optimizes costs at a granular object level

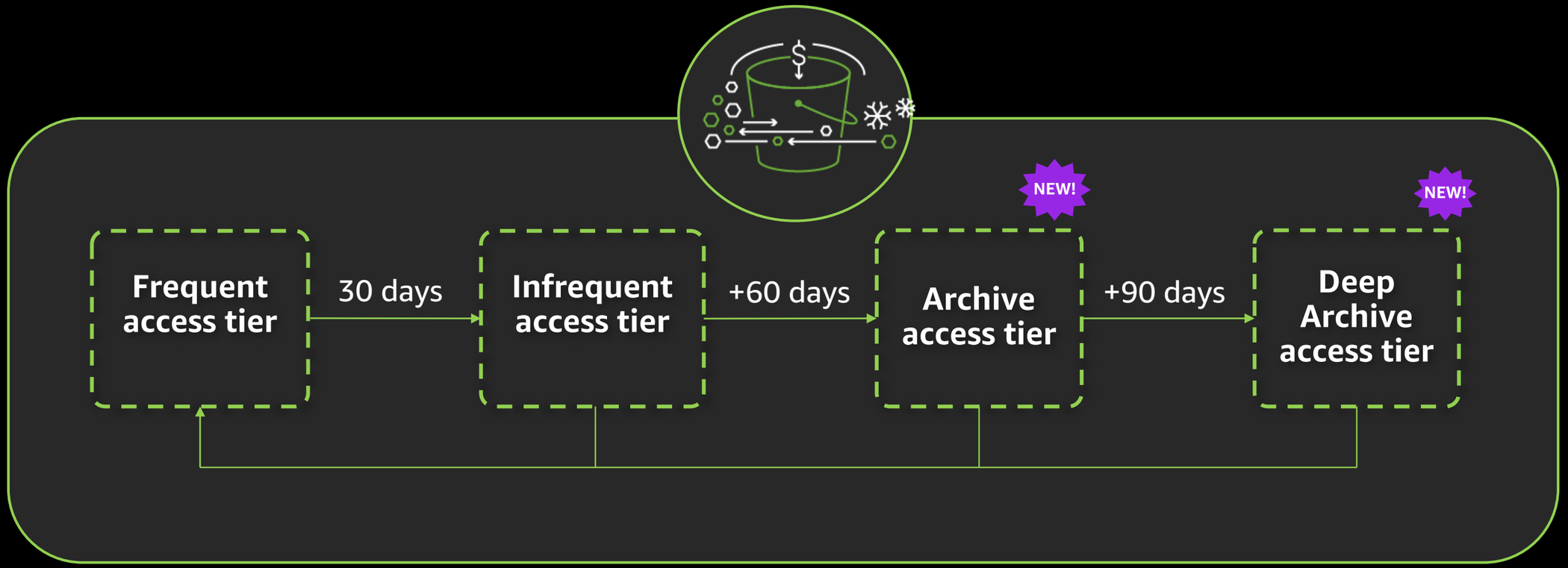
Moves objects between **four access tiers** for a small monthly monitoring and automation fee

Two low latency access tiers for frequent and infrequent access and two **new optional archive access tiers** designed for access in minutes and hours

No operational overhead, **no** lifecycle fees, and no retrieval fees

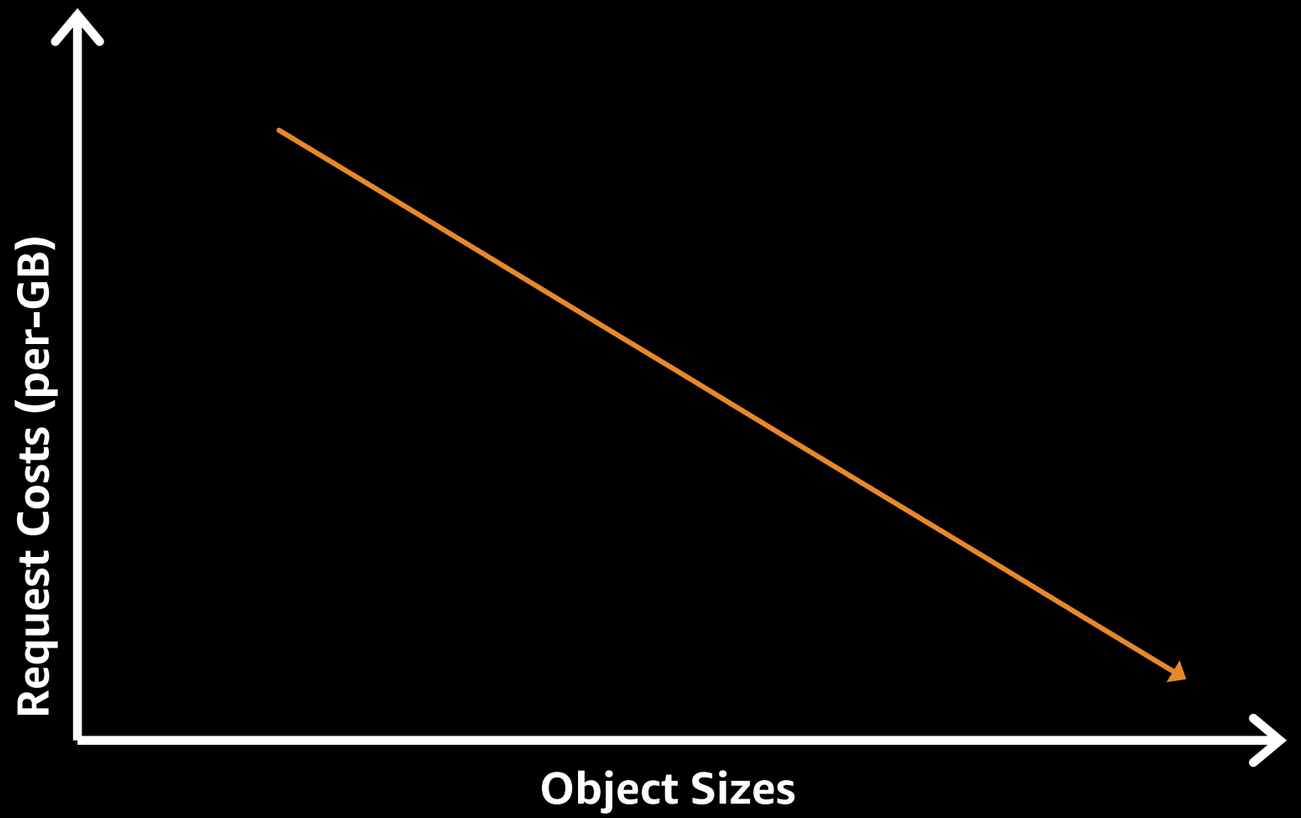
Designed for 99.9% availability and 99.999999999% (11 9's) of durability

S3 Intelligent-Tiering in action



Use S3 Intelligent-Tiering for changing or unknown access patterns

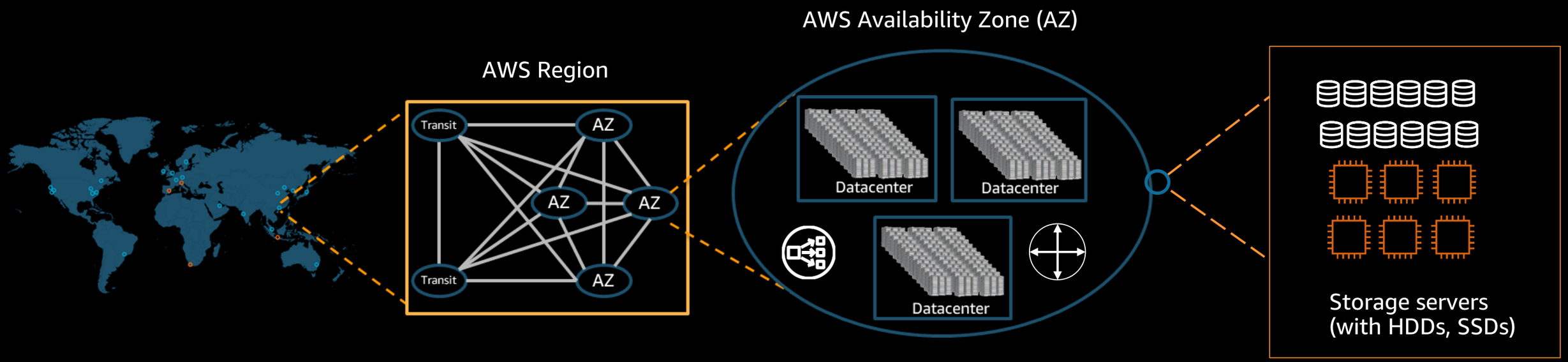
Relationship between object size and total request costs



- Batch data on ingest
- Use tools like Amazon EMR to batch existing data into larger objects

Batch small objects into larger objects to reduce number of requests

Amazon S3 internals: 10,000 ft view

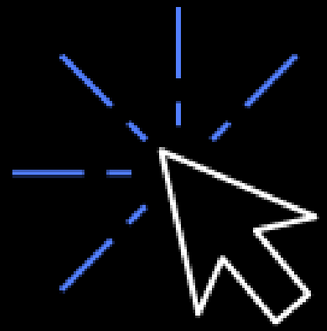


Availability zones, LBs, racks, switches, storage servers, disks, …

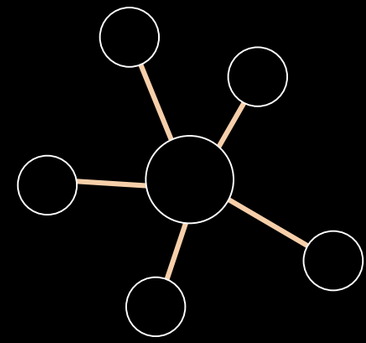
Thinking about Amazon S3



Trillions of objects and millions of requests per second

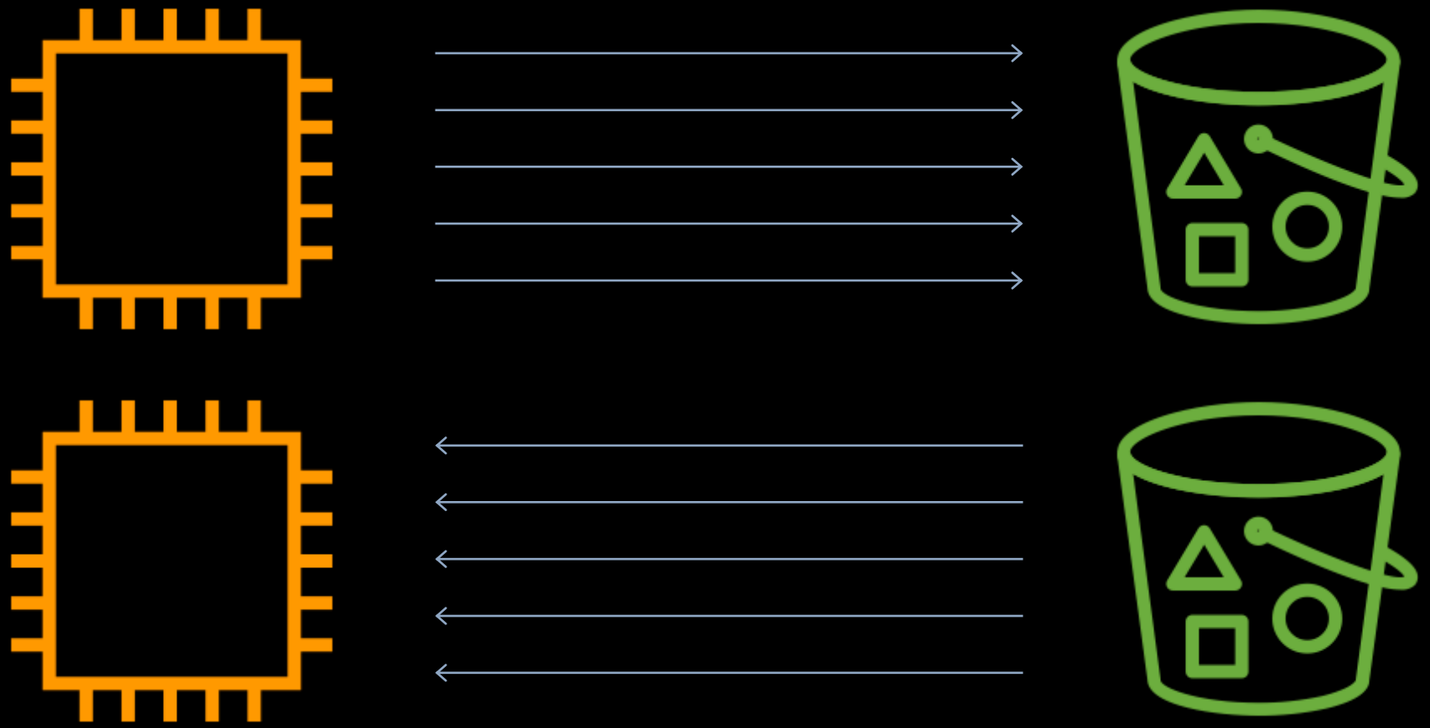


Simple GET/PUT interface to your objects

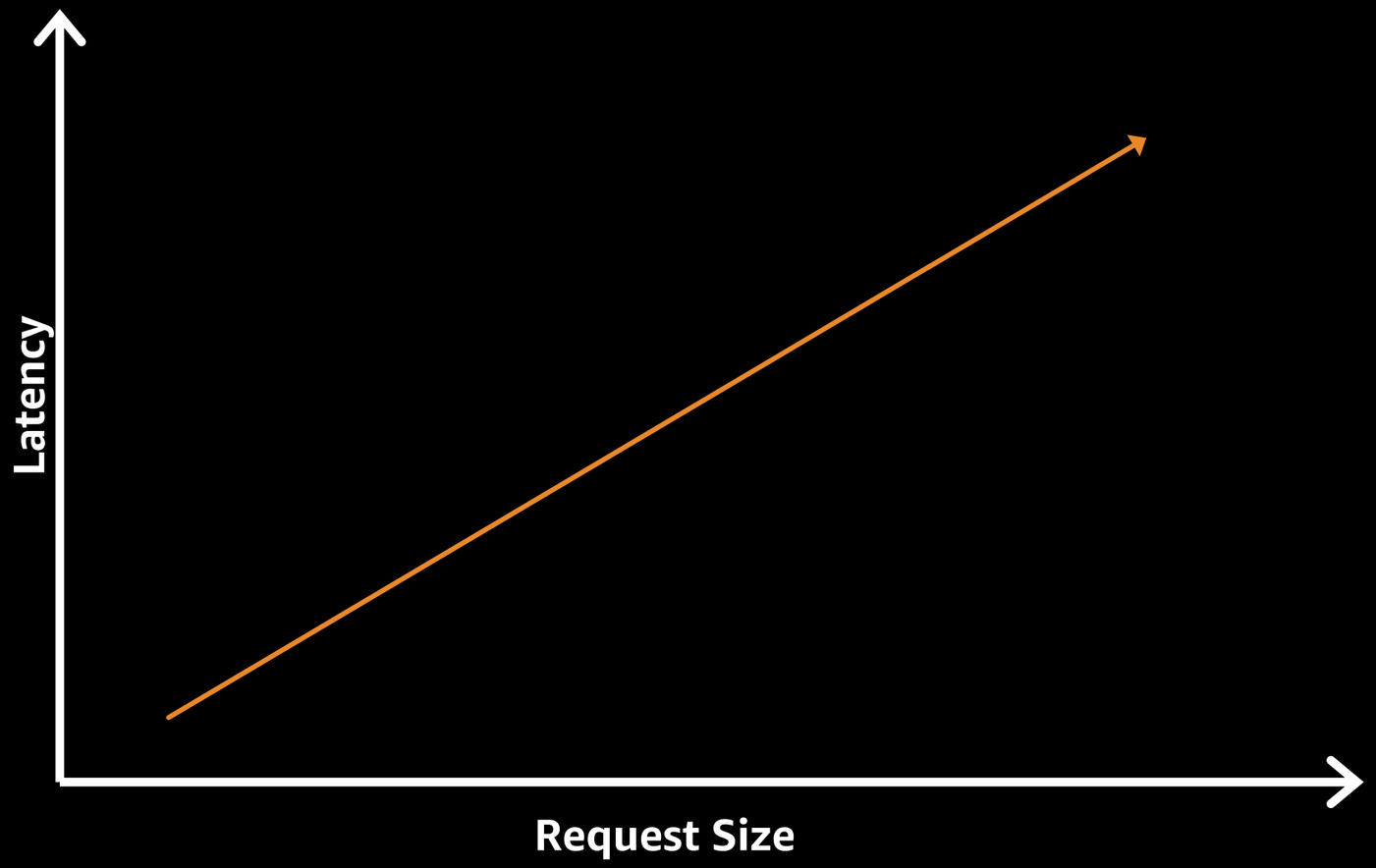


Distributed service with multiple endpoints

Horizontal scaling is key to maximizing performance



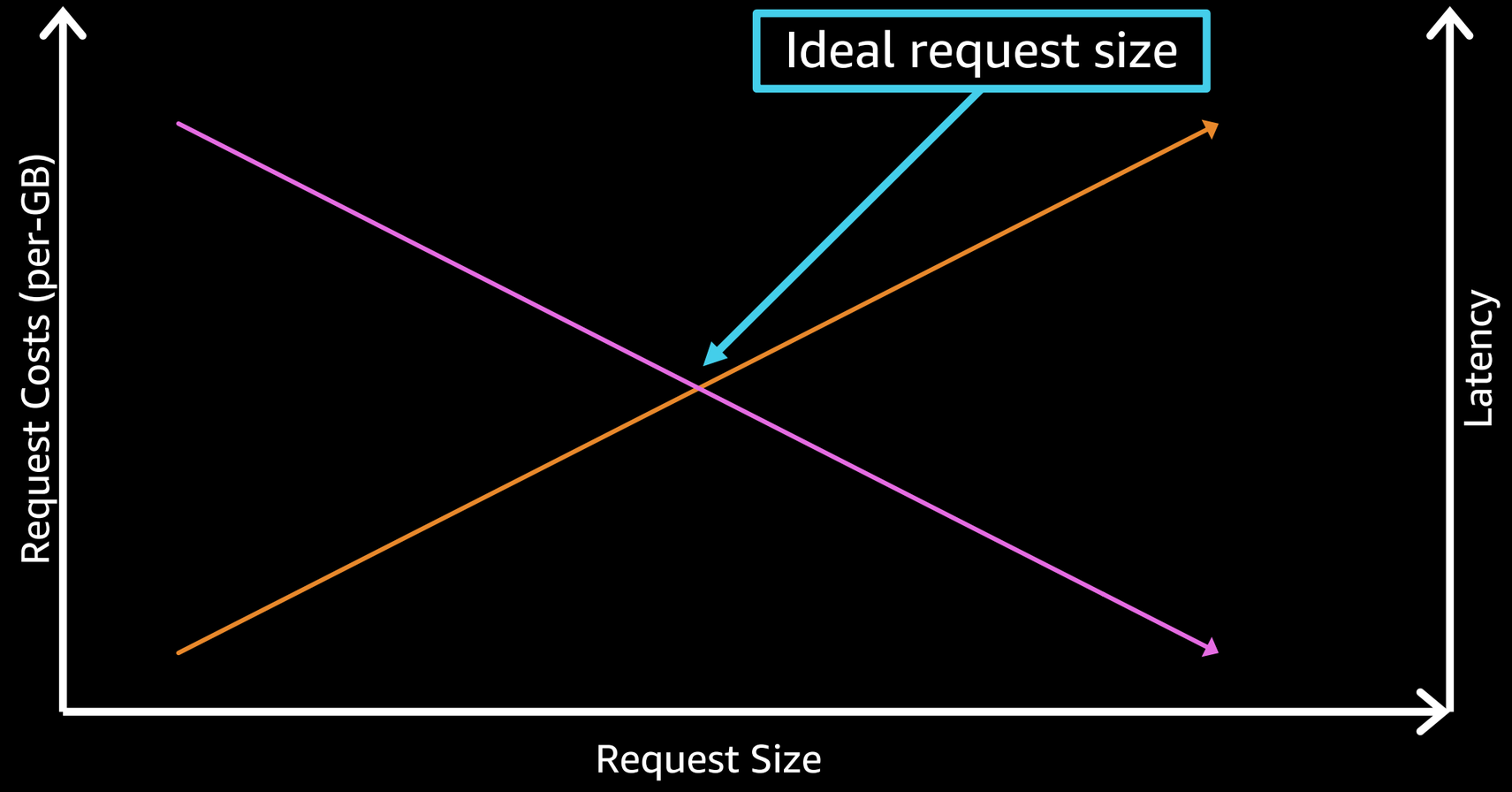
Relationship between latency and request size



- Use multi-part upload to parallelize large objects when uploading
- Use byte-range fetches to parallelize large objects when retrieving

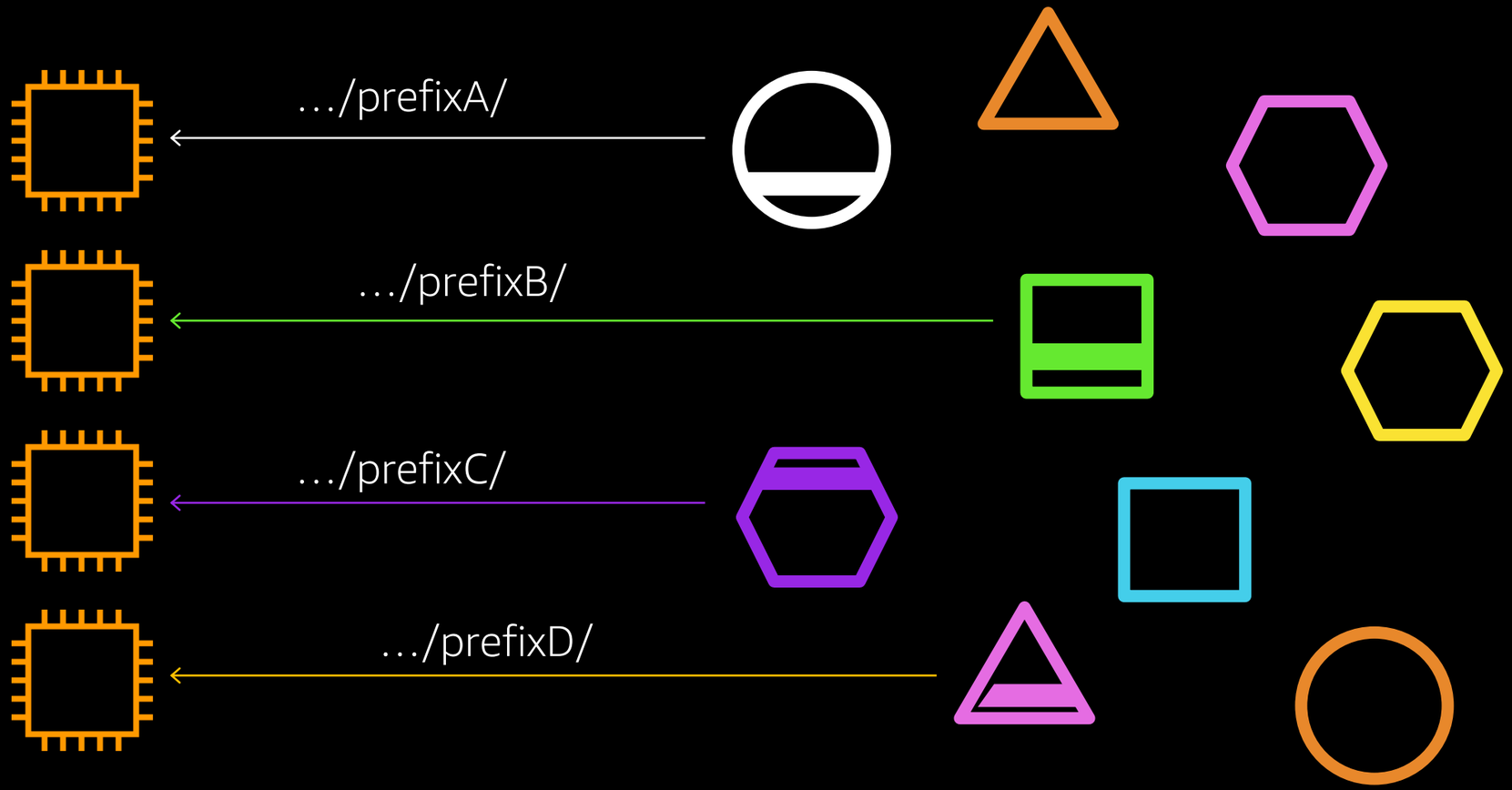
Parallelize large object uploads and downloads to reduce latency and improve throughput

Optimal request sizes for performance and cost



Find the optimal object and request size for your application

Spread load over the key namespace



Case study: Optimizing performance and cost for a social media data lake



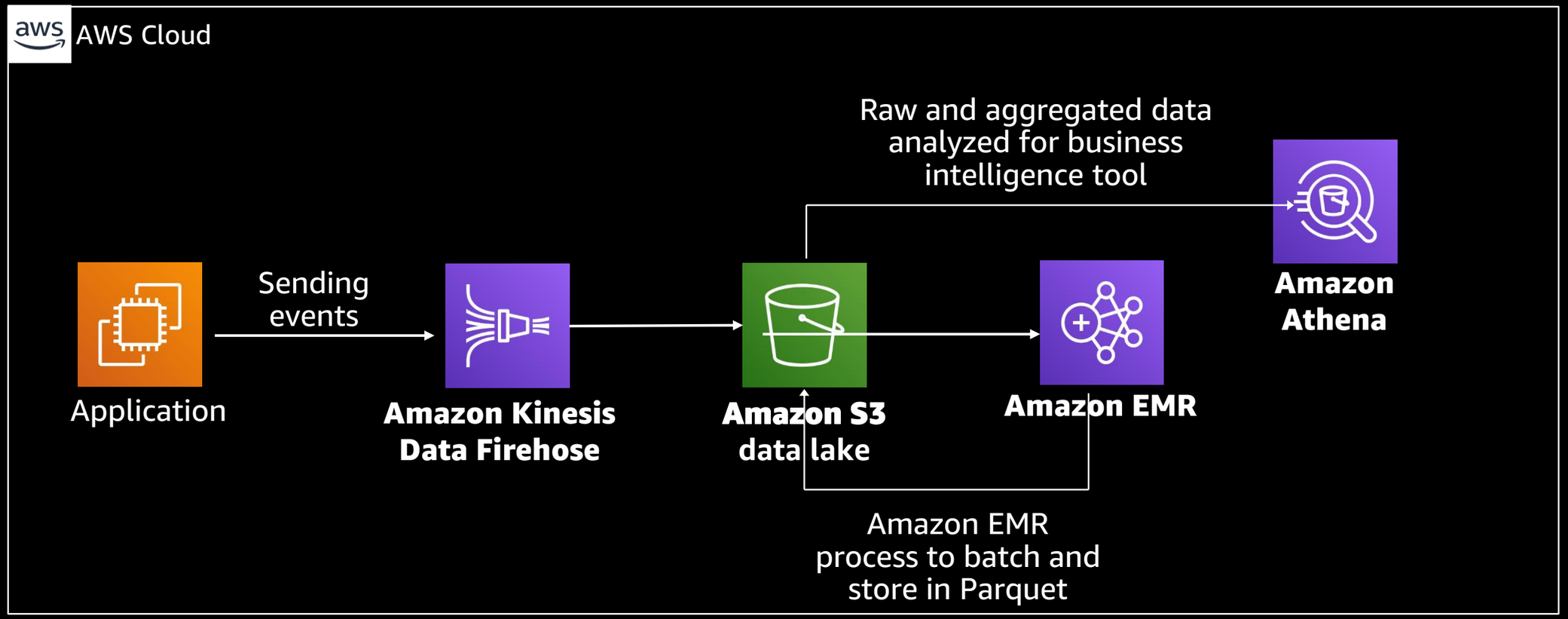
AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

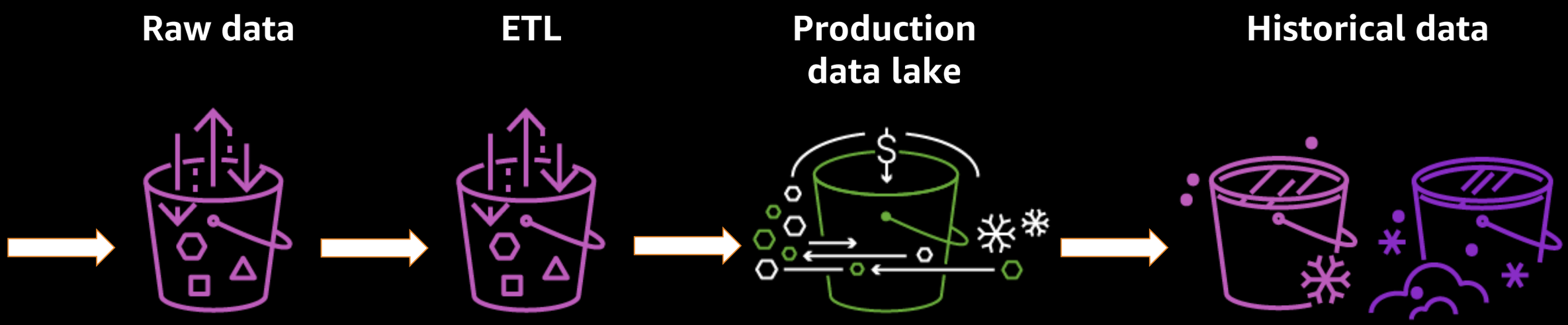
Case study: social media platform

Workload: Social media platform processing over a billion events daily

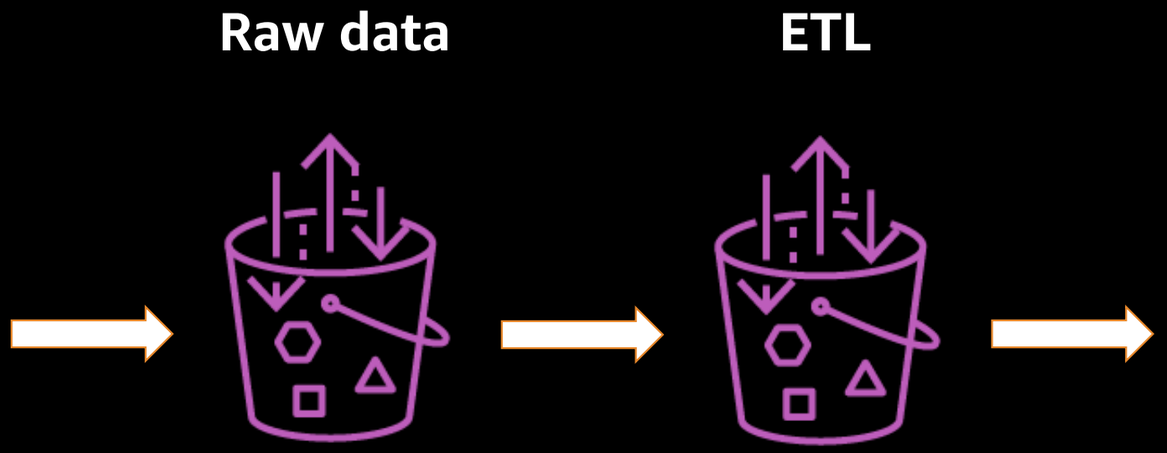
- Peaking at hundreds of thousands of events per second
- Storing many petabytes of data



Case study: the data pipeline



Case study: raw data and ETL



Recommendations

- Batch data on ingest
- Spread load across your key space
- Use S3 Standard or S3 Intelligent-Tiering for the lowest cost

Case study: production data lake

Production data lake



Recommendations

- Using multi-part upload and byte-range fetches
- Use S3 Intelligent-Tiering for automatic cost optimization

Case study: historical data

Historical data



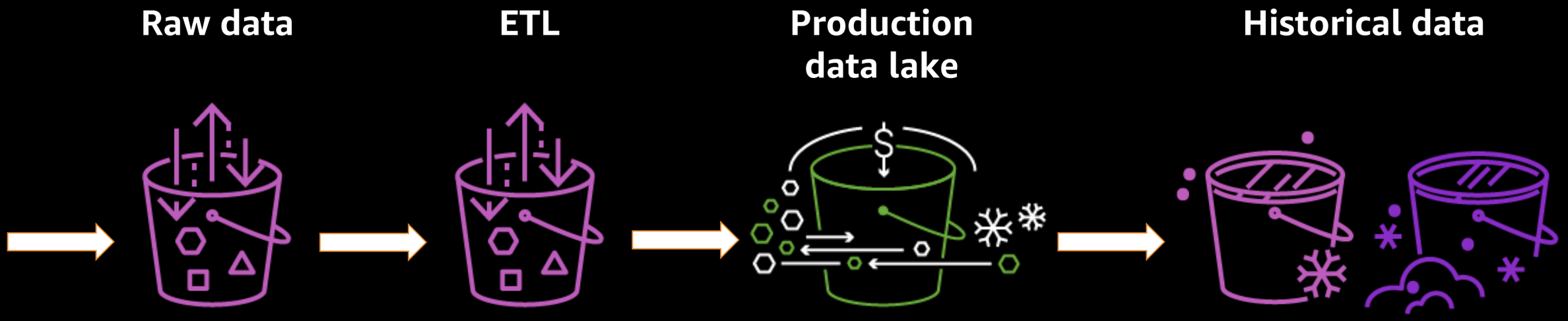
Recommendations

- Use the S3 Glacier storage classes or S3 Intelligent-Tiering Archive access tiers
- Choice depends on your restore performance requirements

Multiple restore speeds

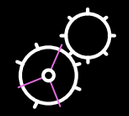
	S3 Glacier / INT Archive Tier	S3 Glacier Deep Archive / INT Deep Archive Tier
Expedited	Typically 1-5 minutes	Not available
Standard	Typically 3-5 hours	Typically within 12 hours
Bulk	Typically 5-12 hours	Typically within 48 hours

Case study: optimize each stage of the data lake

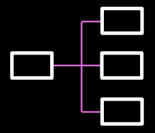


Optimize storage cost and performance for all stages of a data lake workflow

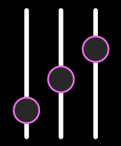
Putting it all together



Understand your application requirements



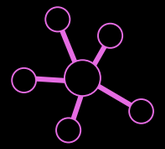
Organize your data using buckets, prefixes, and S3 Object Tags



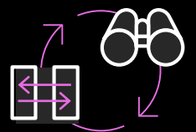
Monitor your data by using tools like S3 Storage Lens



Choose the right storage class for the lowest price



Scale horizontally to maximize performance



Optimize cost and performance for all stages of your workload

Thank you!

Matt Sidley

Senior Product Manager
Amazon S3

Twitter: @msidley

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Please complete the session survey

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.