

# **AWS Architecture Training S3**

Amazon Simple Storage Service Amazon Amazon Module 6 Amazon S3 MapReduce

> ©2012 Amazon Web Services May not be reused or redistributed without permission

#### **Module Overview**

- S3 Features
- S3 Concepts
- S3 Namespace
- S3 Pricing
- S3 Use Cases



#### **Amazon Simple Storage Service**

- Amazon S3: Simple Storage Service
- Highly scalable data storage in-the-cloud
- Programmatic access via web services API
- Is a Web Store, not a file system
  - Optimized for WORM
  - Eventually consistent
- Fast, highly available and durable
- Economical



#### EBS and S3 -What's the Difference?

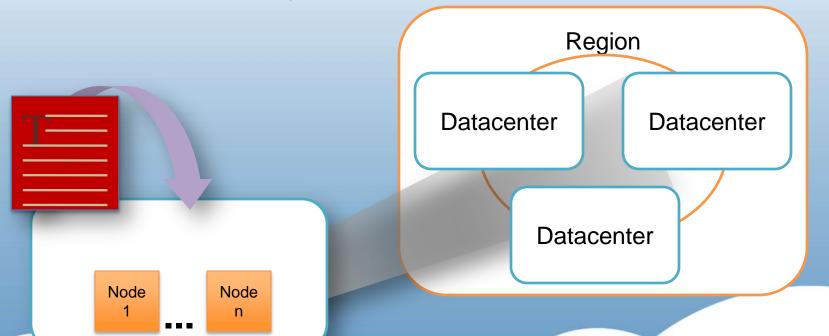
	EBS Standard	Amazon S3
Paradigm	File system	Object store
Performance	Very, very fast	Very fast
Redundancy	Within data center	Across data centers
Security	Visible only to your EC2	Public Key / Private Key
Pricing	\$0.10/GB/month allocated	\$0.125/GB/month stored
Access from the Net?	No(2)	Yes(1)
Typical use case	It's a disk drive	Write once, read many

- (1) Only with proper credentials, unless ACLs are world readable
- (2) Accessible from Net if mounted to server and set up as FTP, etc.



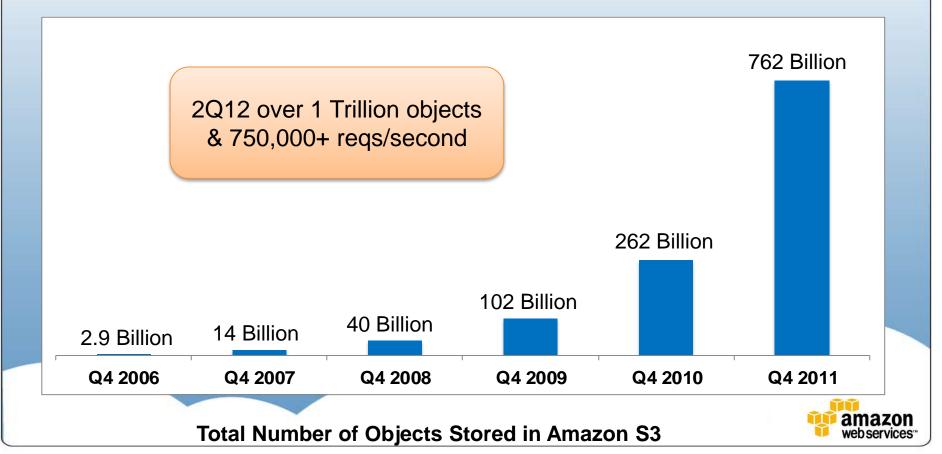
## Designed for 99.99999999% Durability

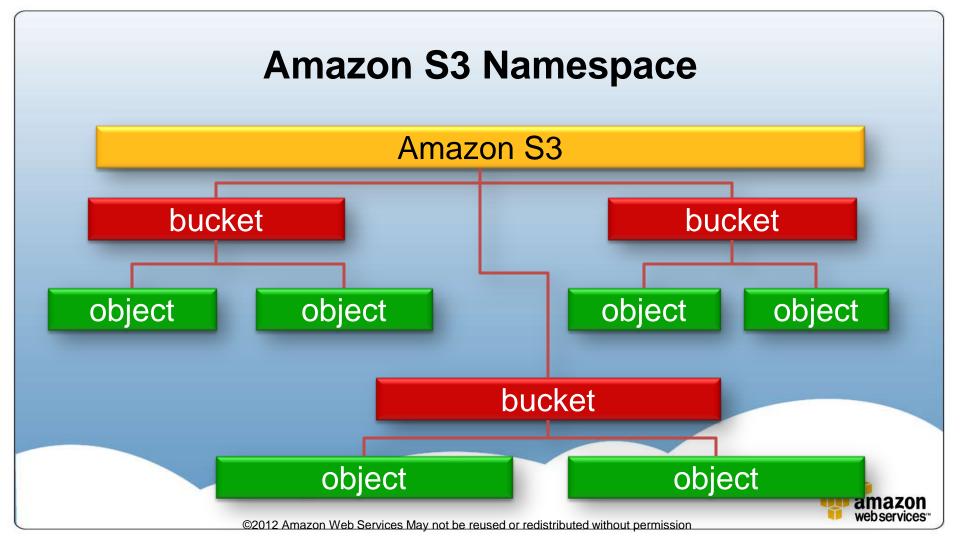
Note: Conceptual drawing only. Actual number of nodes & datacenters may vary

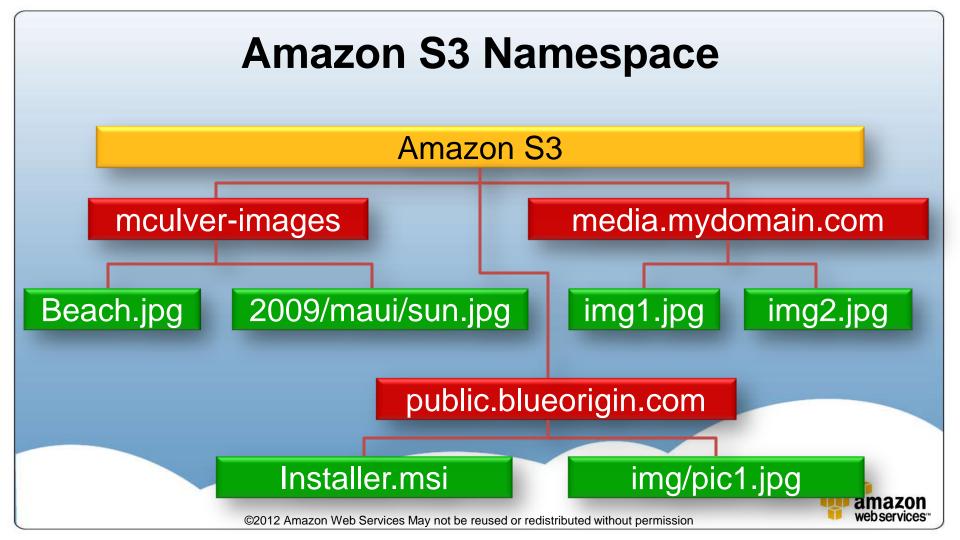












## Reduced Redundancy Storage Saves \$\$

- Designed to provide 99.99% durability and 99.99% availability of objects over a given year
- This durability level corresponds to an average annual expected loss of 0.01% of objects
- Designed to sustain the loss of data in a single facility



## S3 Pricing has Three Dimensions

Storage		Data Transfer		Requests		
Tier	Normal	RRS	Tier	Price	Type	Price
First 1 TB/mo	\$0.125/GB	\$0.093/GB	All transfer in	Free	PUT, COPY, POST, or LIST	\$0.01 per 1,000
Next 49 TB	\$0.111/GB	\$0.083/GB	First 1 GB out	\$0.00/GB	GET and all other requests	\$0.01 per 10,000
Next 450 TB	\$0.095/GB	\$0.073/GB	<= 10 TB out	\$0.12/GB		
Next 500 TB	\$0.090/GB	\$0.063/GB	Next 40 TB out	\$0.09/GB		
Next 4000 TB	\$0.080/GB	\$0.053/GB	Next 100 TB out	\$0.07/GB		
> 5000 TB	\$0.055/GB	\$0.037/GB	> 350 TB out	\$0.05/GB		



### **Access Control Lists**

- Control access to bucket or to object
- No inheritance from bucket to object
- Set at create, write, or any time
- List of up to 100 grants by:
  - Owner
  - Email
  - Any AWS account
  - Anyone
- Permissions:
  - READ Object or bucket
  - WRITE Create, overwrite, delete objects in bucket
  - READ\_ACP Read ACL for bucket or object
  - WRITE\_ACP Overwrite ACL for bucket or object
  - FULL CONTROL All of the above





### **Bucket Policy**

# Allow Jason or Mark to do anything in my S3 bucket when coming from the Office internal network

```
{"Statement":[{
          "Effect":"Allow",
          "Principal":{"AWS":["4649-6425", "5243-0045"]},
          "Action":"*",
          "Resource":"/mybucket/*",
          "Condition":{
                "IpAddress":{"AWS:SourceIp":"176.13.0.0/12"}
}}]}
```



### **Server Side Encryption**

SSE automatically encrypts data at rest within Amazon S3. Customer benefits include being:

- Simple Requires only an additional header in a PUT request
- Durable Highly Durable Keys = Highly durable data
- Strong Data encrypted using AES-256
- Self Managed No burden of managing a key store
- Secure Decryption requires simultaneous access to 3 components



# Simple: New PUT header

**PUT** 

x-amz-server-side-encryption

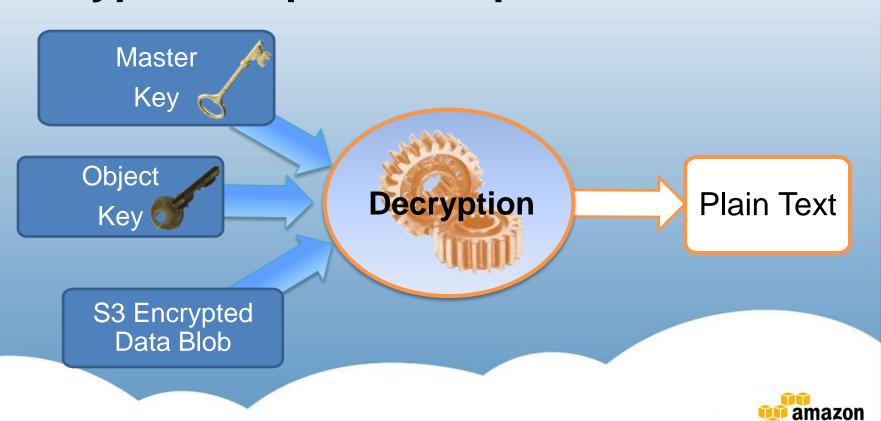
Amazon S3

**GET** 



## **SSE High Level Design Customer Data Encrypt** (PUT) **Encrypted Data** Random Key Generator **Encrypt** Encrypted (per-object) per-object Key Key Mgmt Service **Rotated Monthly**

# Decryption requires 3 separate factors



### **S3 Multipart Uploads**

- improved throughput upload parts in parallel
- **Quick recovery from any network issues -** Smaller part size minimizes the impact of restarting a failed upload due to a network error.
- Pause and resume object uploads
  - upload object parts over time
  - Once you initiate a multipart upload there is no expiry
  - must explicitly complete or abort the multipart upload
- Begin an upload before you know the final object size can upload as you are create.



#### **Even More Amazon S3 (not so simple, is it?)**

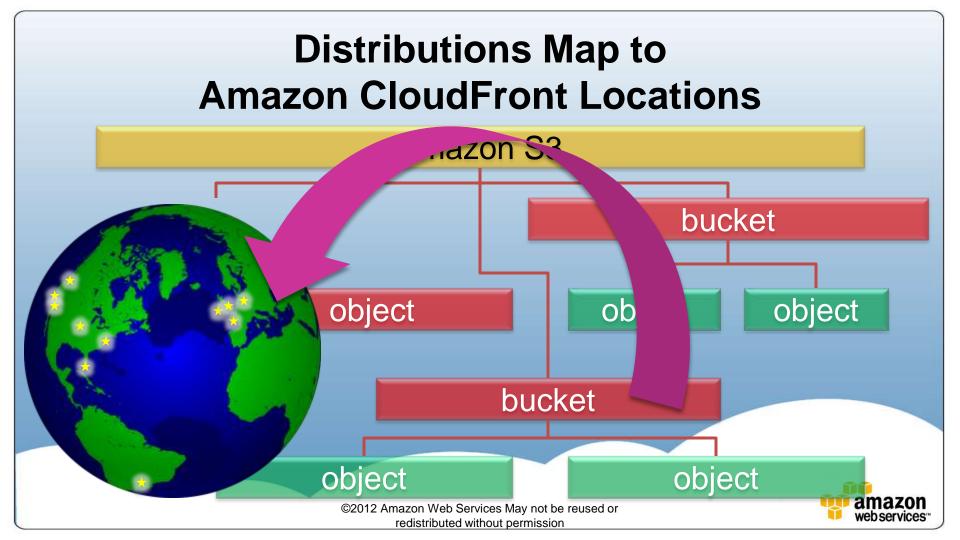
- Website support
- Versioning
- Browser Upload to S3
- Logging
- Requestor Pays
- Signed URLs
- Bittorrent support
- i IAM
- Meta-data
- Multi-object delete



# **AWS Import/Export**

- Get your data into AWS faster load it onto a portable storage device and ship it to an Amazon data center
- Faster than Internet transfer and more cost effective than upgrading your connectivity
- Use cases: data migration, offsite backup, direct data interchange, disaster recovery





#### **Amazon CloudFront**

Web service for content delivery

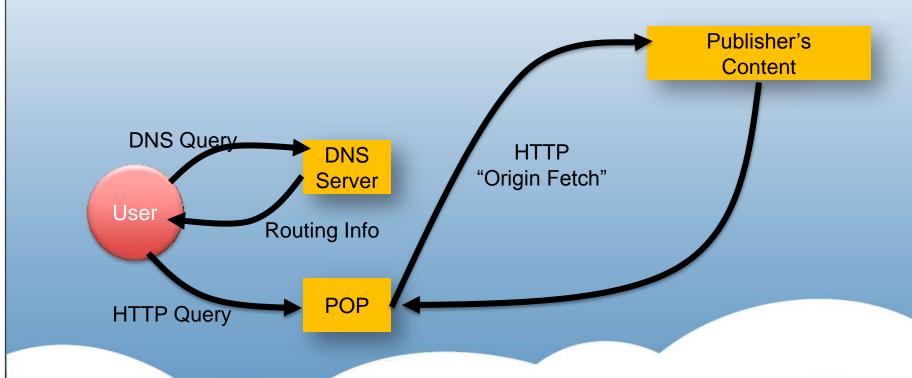
Edge caching of any object (S3 or server

- Pay-as-you-go pricing
  - No up-front contracts
  - No long-term commitments
  - Volume discounts
- 33 Edge locations:
  - US (16), Europe (10), Asia (6), South America (1)
- Easy to use:
  - Create distribution
  - Issue links to objects





#### **How Amazon CloudFront Works**



©2012 Amazon Web Services May not be reused or redistributed without permission



#### **Common S3 Use Cases**

- Media Sharing
- Media/Software Distribution
- Backup (Server & PC)
- Online Storage
- Application Storage





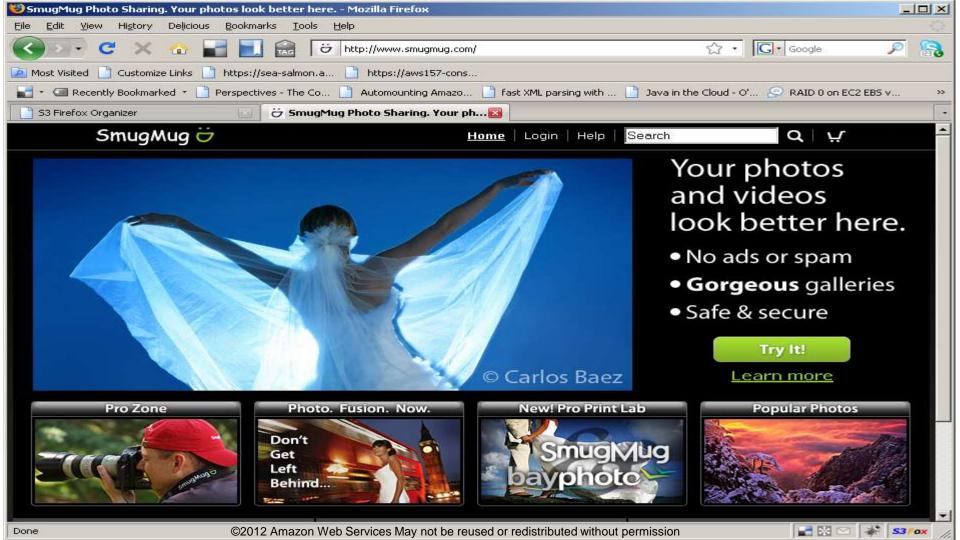
# **European Space Agency**

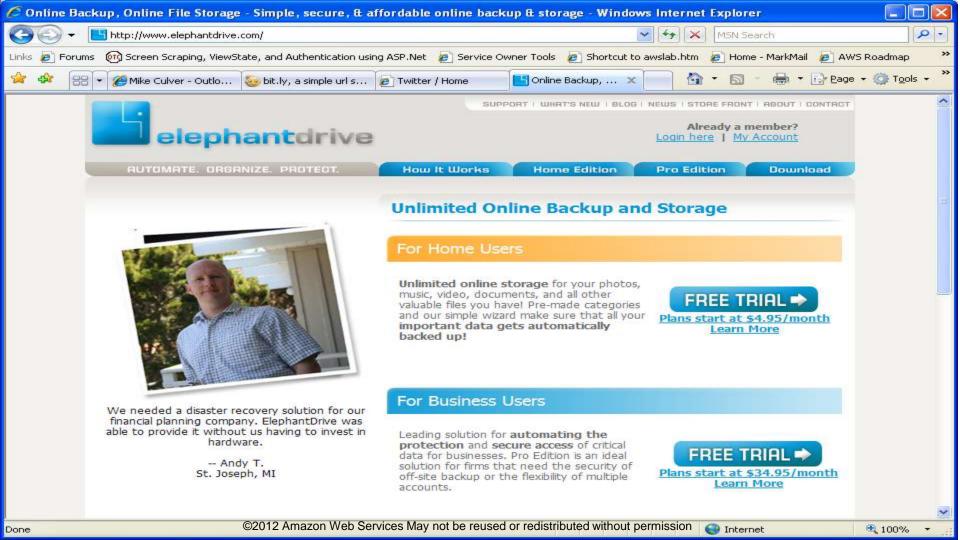
(\*) Source: http://www.esa.int/esaMI/ESRIN SITE/index.html

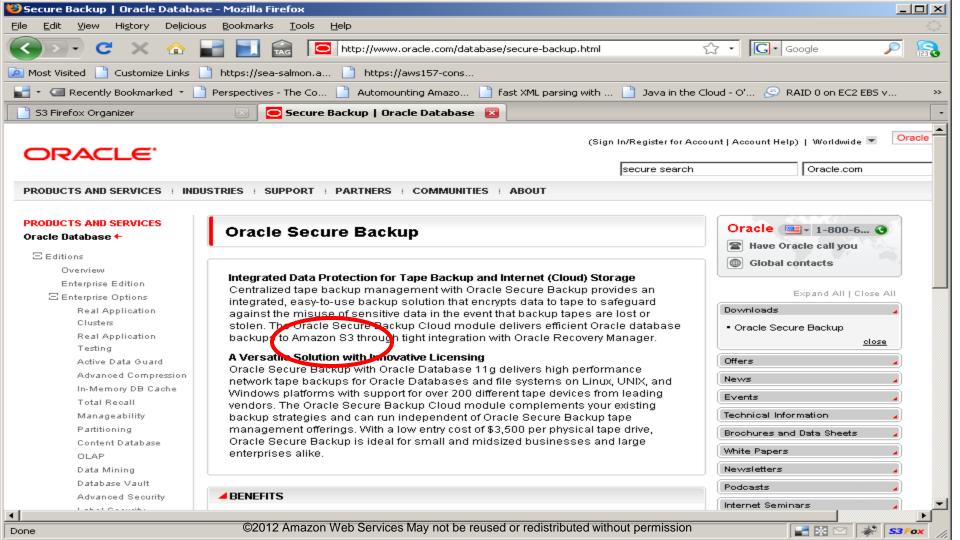
- Amazon S3 stores and serves satellite images and data to the community
- ESRIN: center of the European Space Agency for the Earth Observation Programs (\*)
  - Generate satellite images of Earth
  - Consumers include UN agencies, European Commission, NGOs, and Governments











#### **Amazon S3 Backup Solutions**

Commercial Packages

DataCastle <u>www.datacastlecorp.com</u>

Moonwalk

Zmanda <u>www.zmanda.com</u>

TwinStrata <u>www.twinstrata.com</u>

Cirtas <u>www.cirtas.com</u>Tarmin www.tarmin.com

Cloudberry
 <a href="http://cloudberrylab.com/?page=cloud-berry-backup">http://cloudberrylab.com/?page=cloud-berry-backup</a>

Cloudbuddy <u>www.mycloudbuddy.com</u>

Open Source Sync project

S3 Sync <u>www.s3sync.net/wiki</u>

Quillen <u>www.sourceforge.net/projects/quillen</u>

S3 Tools <u>www.s3tools.org/s3cmd</u>

S3Dav www.sourceforge.net/projects/s3dav



