Amazon Web Services
Cloud Computing in Action

Jeff Barr – jbarr@amazon.com

Who am I?

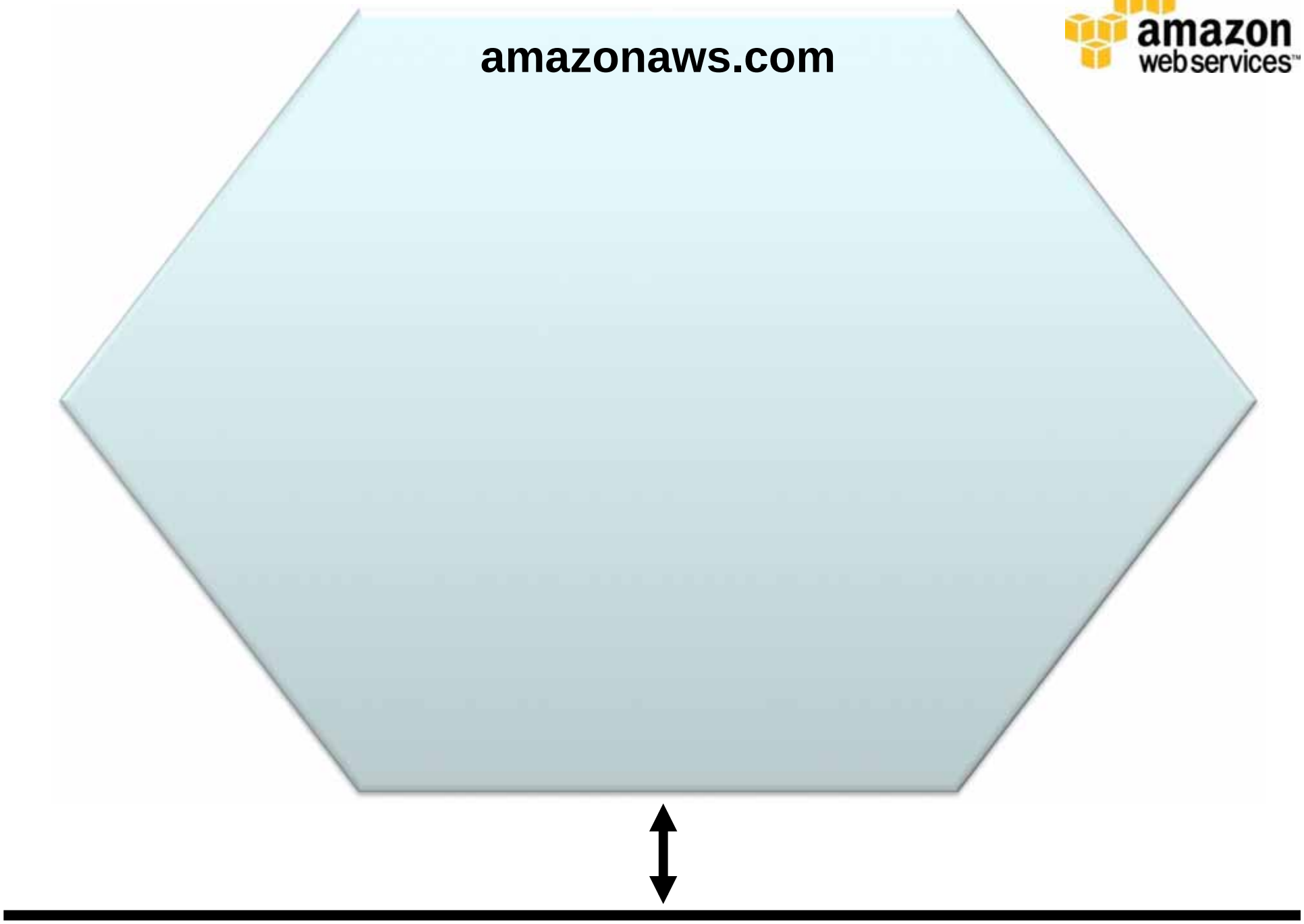
- Software development background
- Programmable applications and sites
- Microsoft Visual Basic and .Net Teams
- Startup / venture consultant
- 6 Years with Amazon:
  - Developer
  - Evangelist
  - Blogger

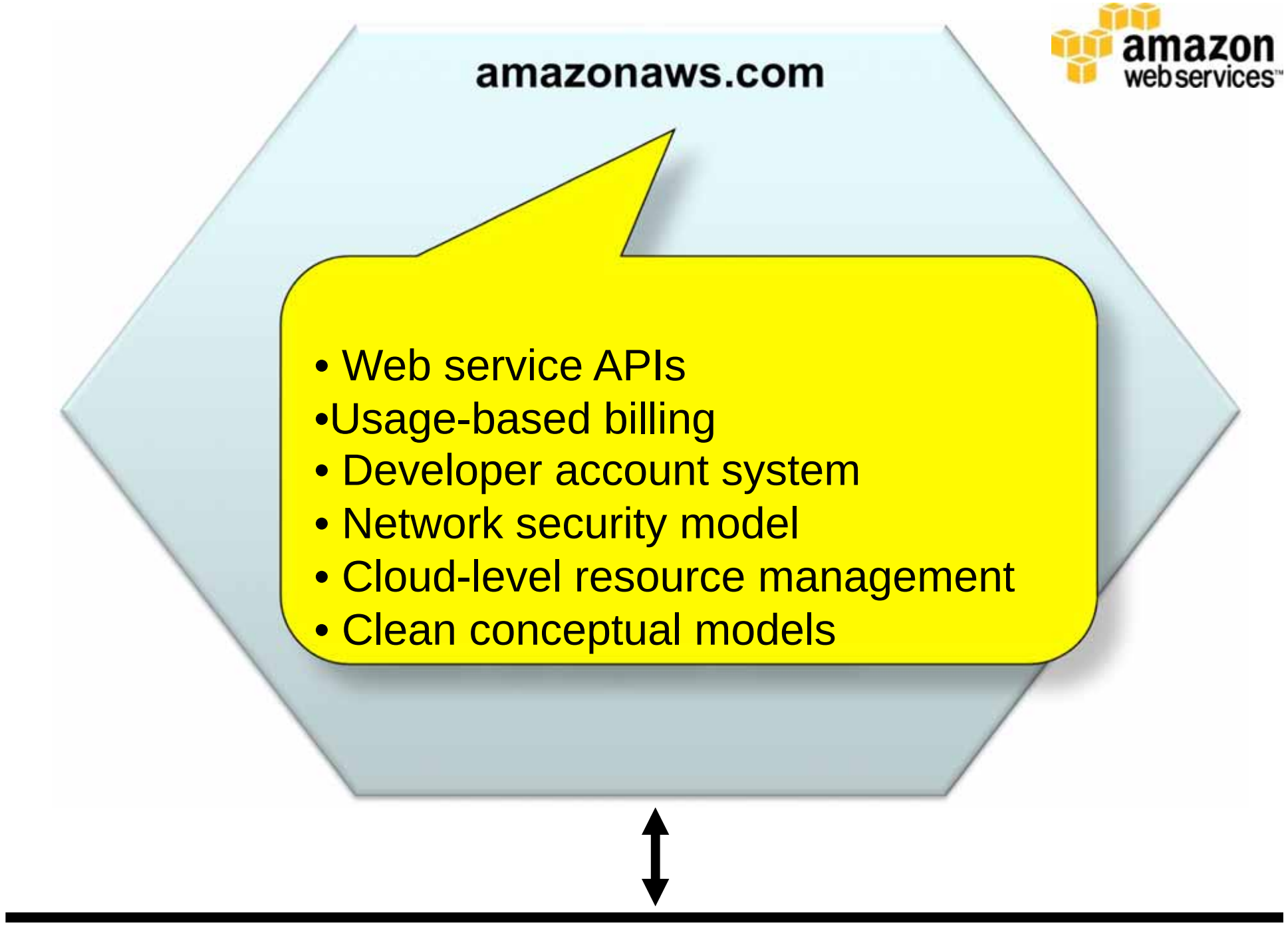
amazon web services™

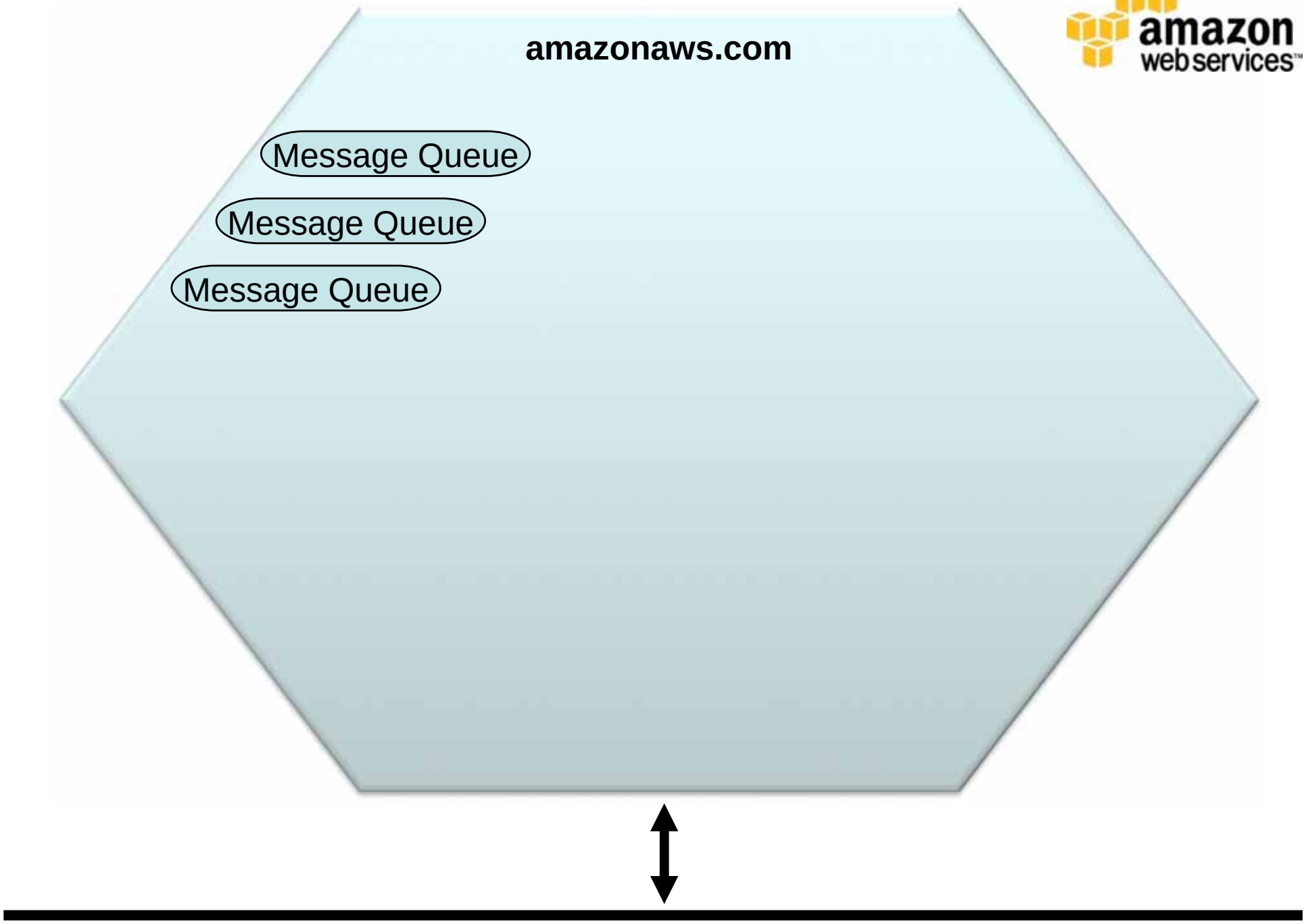
Our Offering

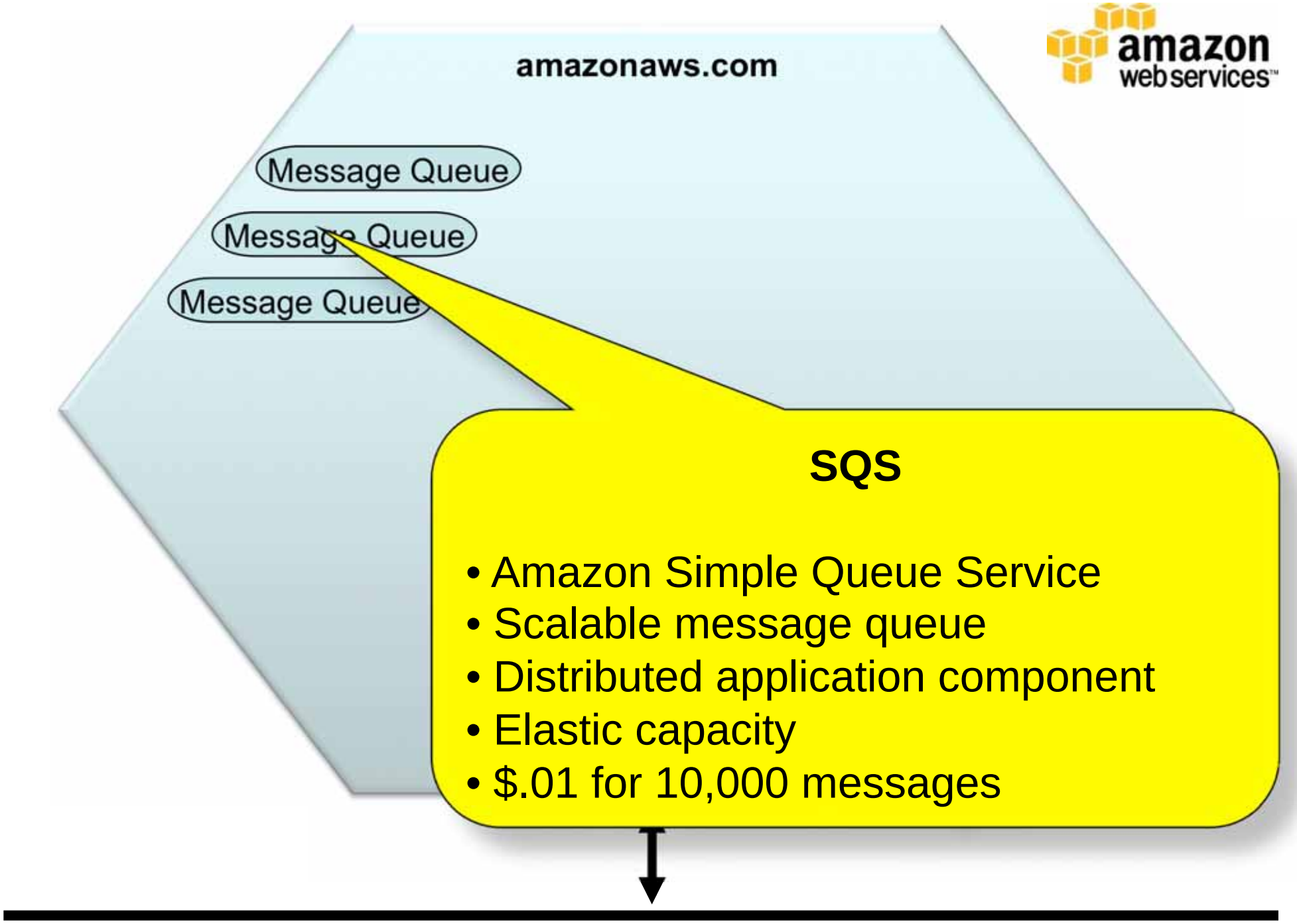
- AWS – Amazon Web Services
- Third major line of business
- Access to proven Amazon infrastructure
- Highly scalable & reliable
- Redundant & geographically dispersed
- Economical
- Metered, pay-as-you-go access

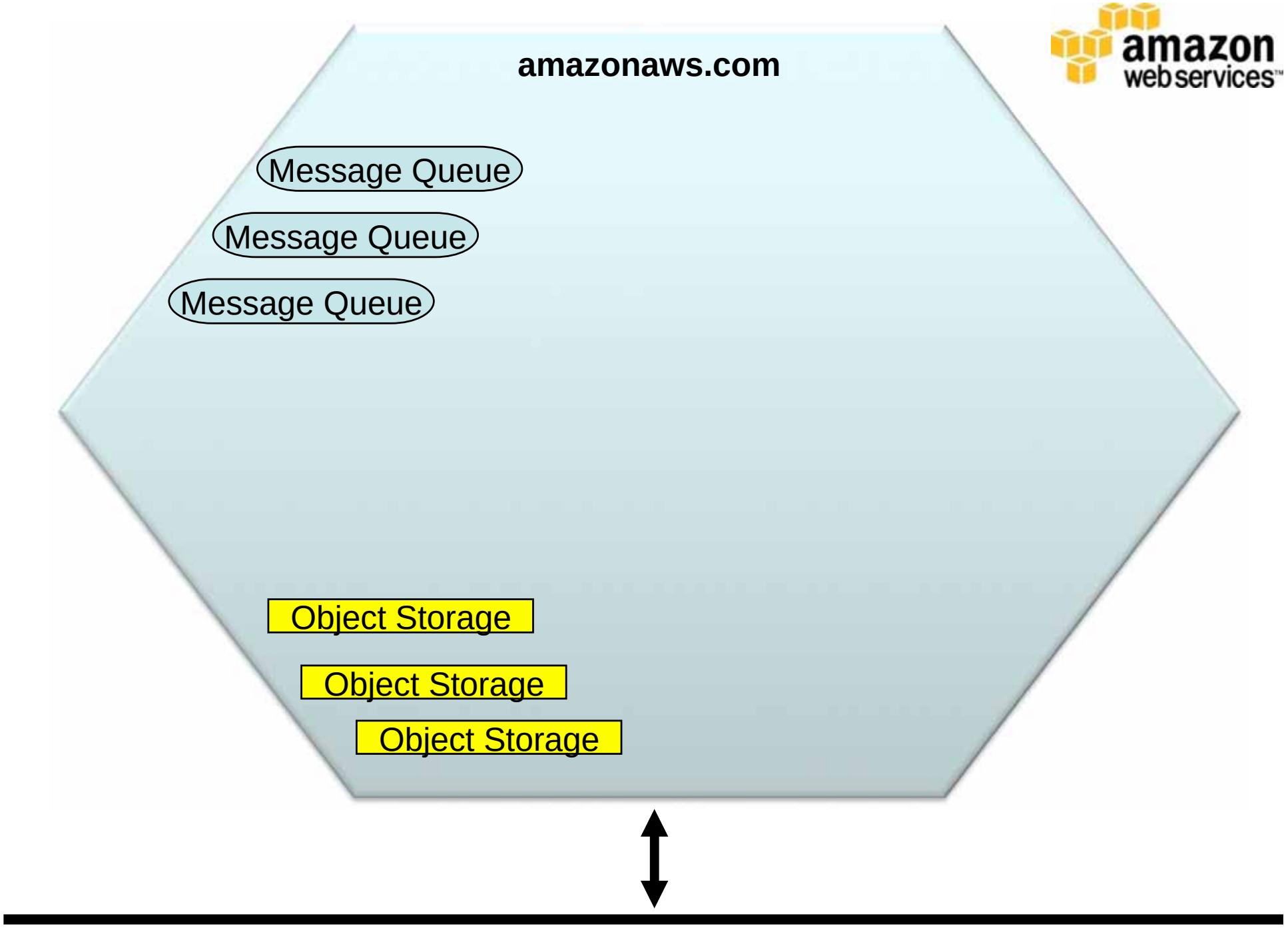
Cloud Computing Components

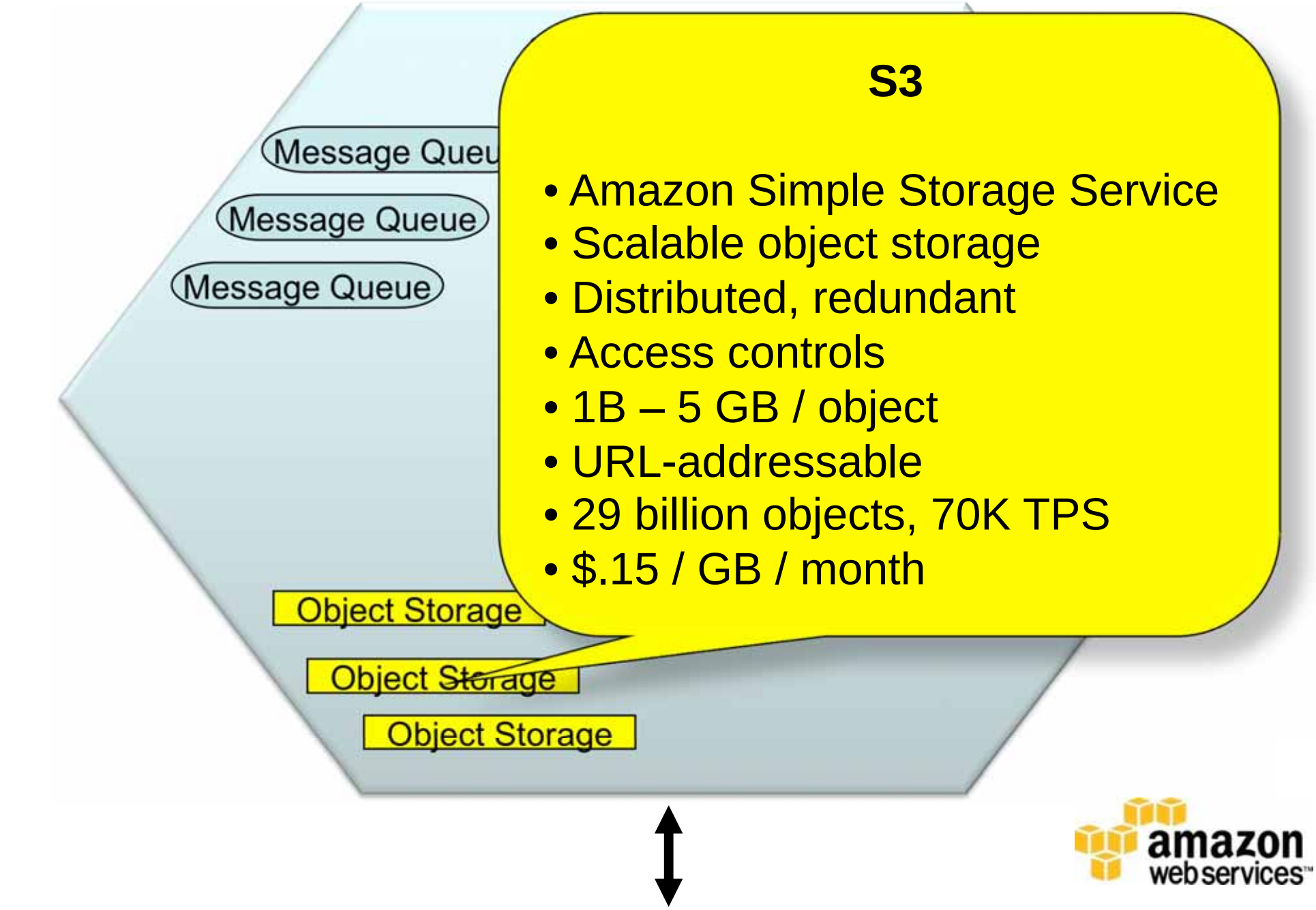

amazonaws.com
amazon
web services


amazonaws.com
amazon web services
• Web service APIs
•Usage-based billing
• Developer account system
• Network security model
• Cloud-level resource management
• Clean conceptual models

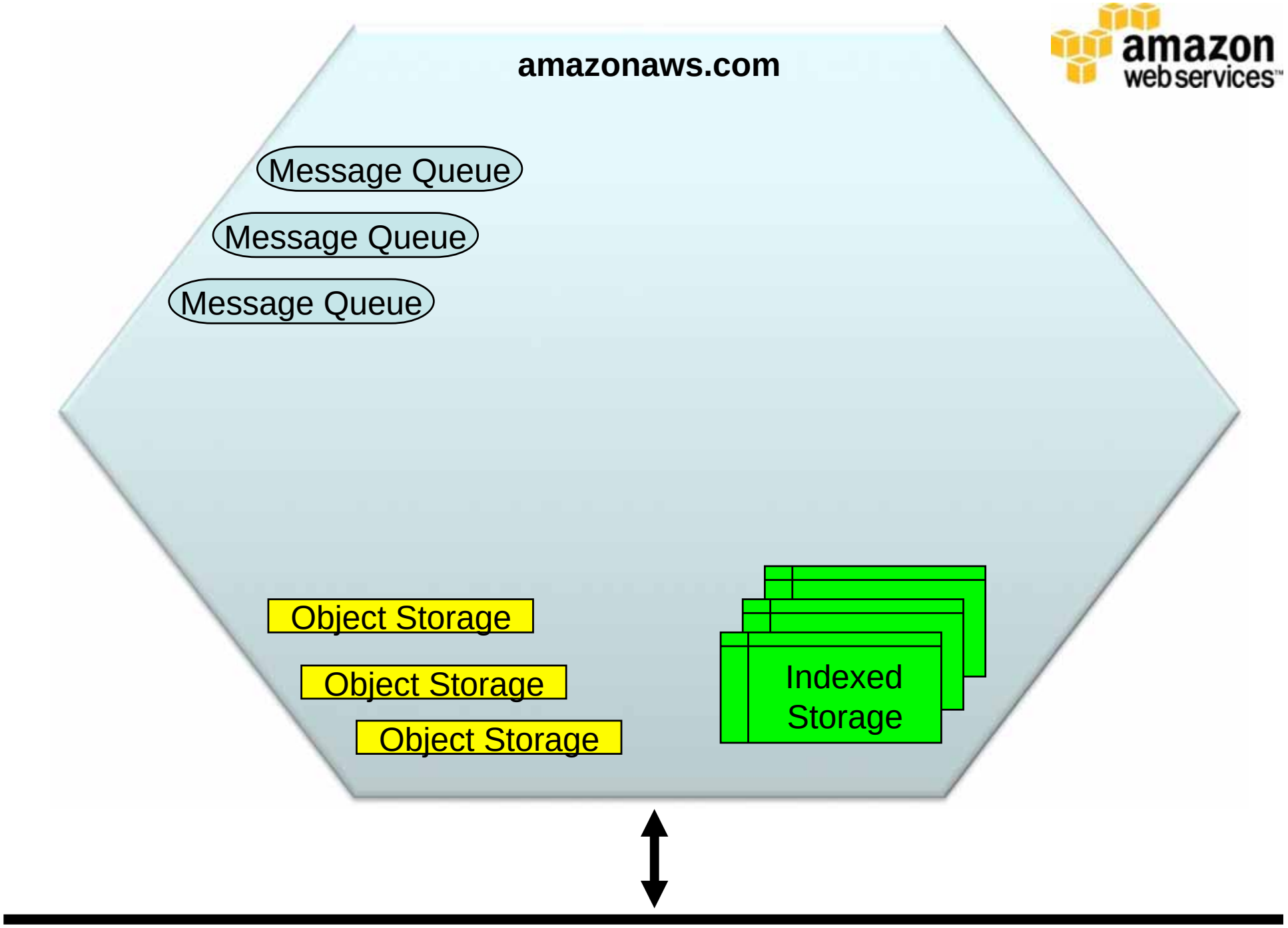

amazonaws.com
Message Queue
Message Queue
Message Queue

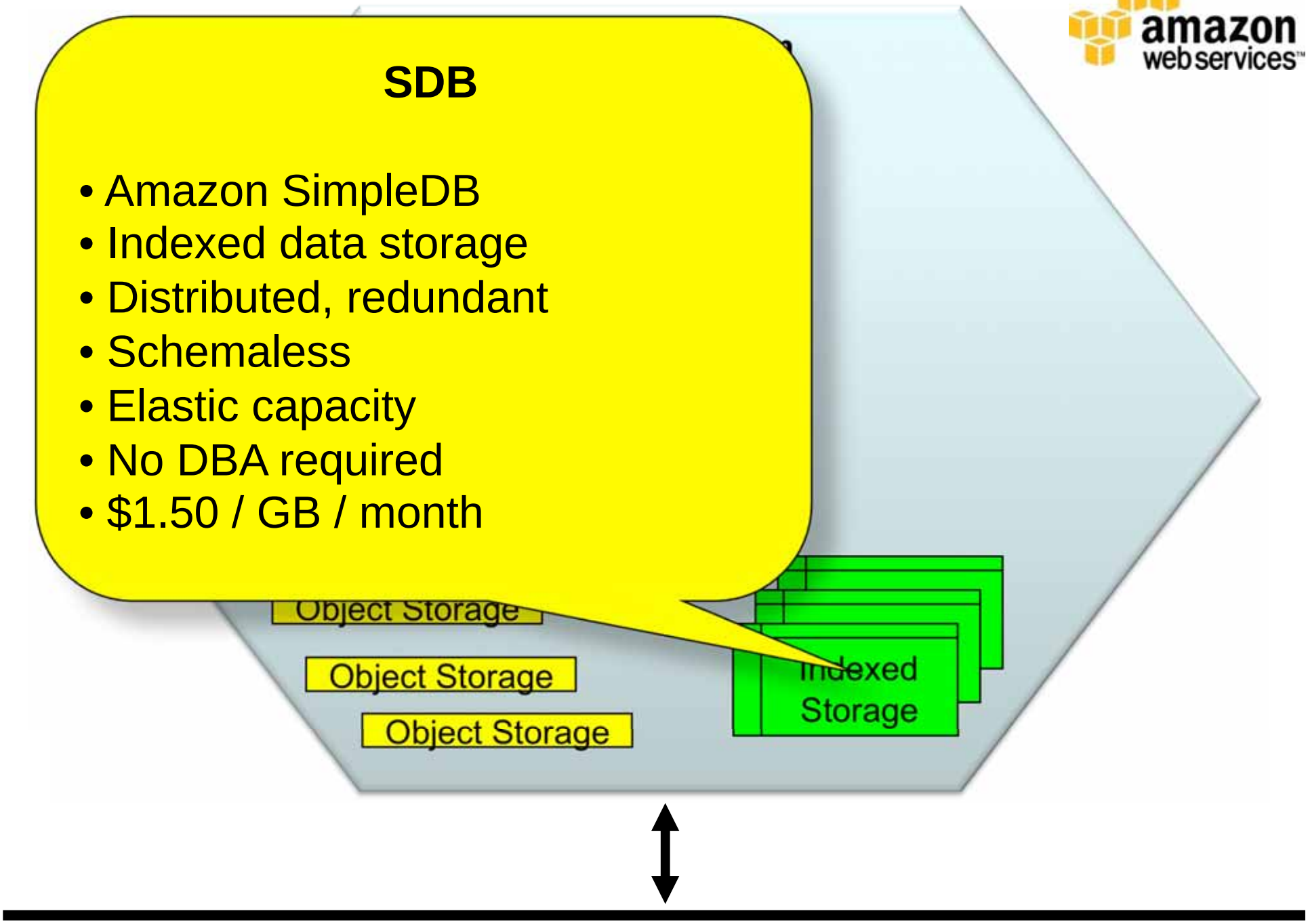

amazonaws.com
Message Queue
Message Queue
Message Queue
SQS
• Amazon Simple Queue Service
• Scalable message queue
• Distributed application component
• Elastic capacity
• $.01 for 10,000 messages

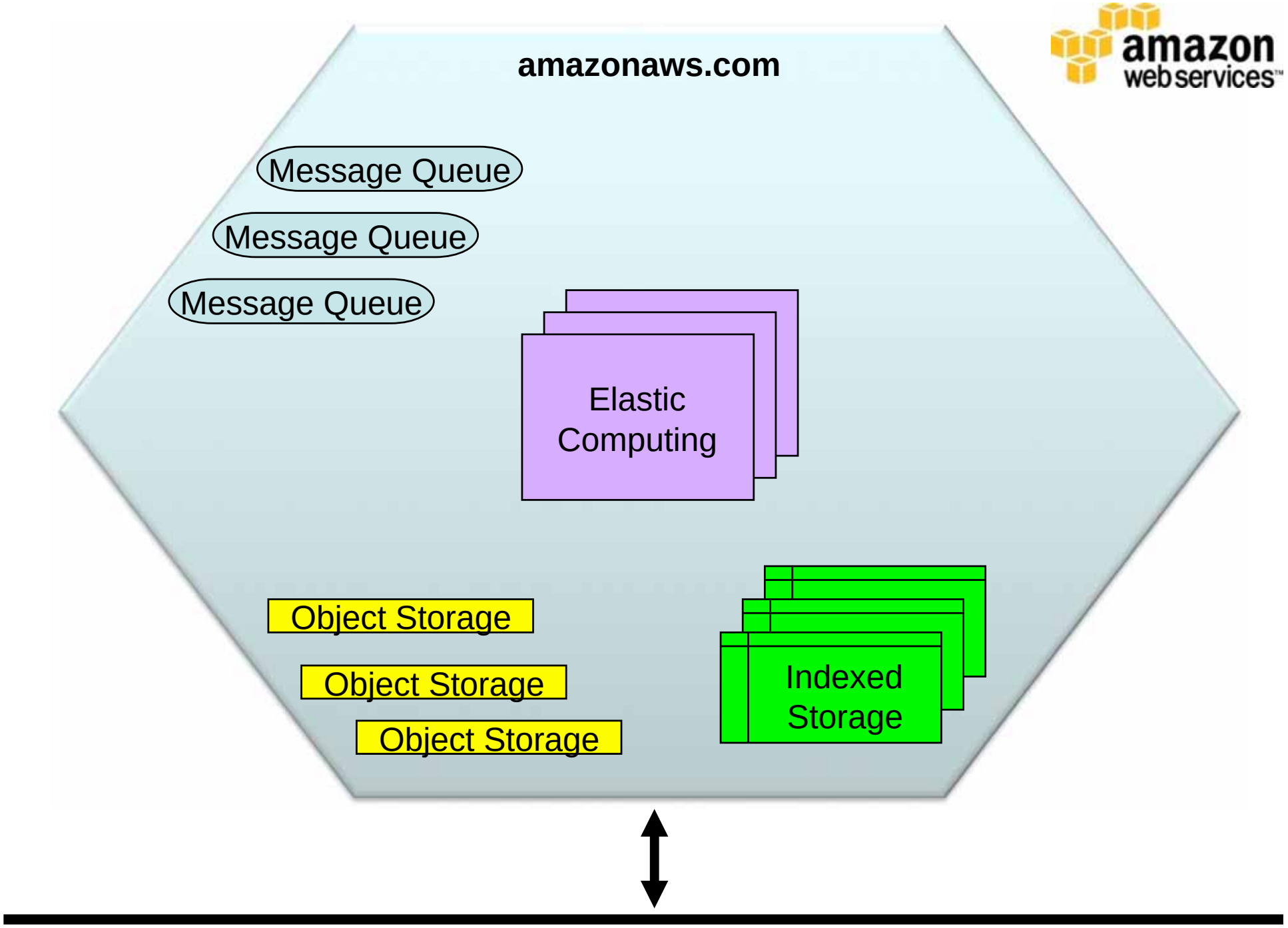

amazonaws.com
Message Queue
Message Queue
Message Queue
Object Storage
Object Storage
Object Storage

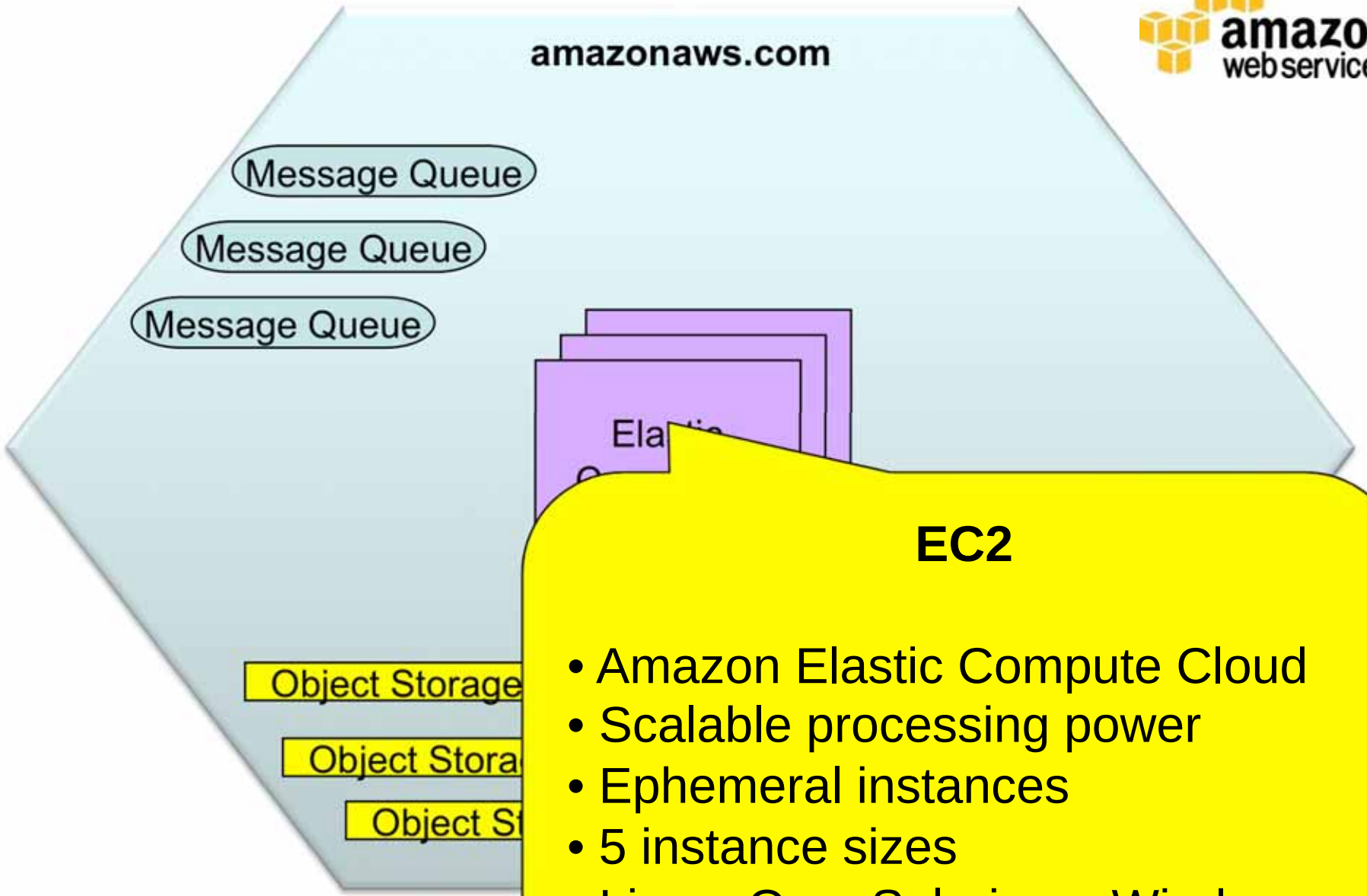


amazon web services™


amazonaws.com
Message Queue
Message Queue
Message Queue
Object Storage
Object Storage
Object Storage
Indexed Storage


SDB
• Amazon SimpleDB
• Indexed data storage
• Distributed, redundant
• Schemaless
• Elastic capacity
• No DBA required
• $1.50 / GB / month
Object Storage
Object Storage
Object Storage
Indexed
Storage


amazonaws.com
amazon web services
Message Queue
Message Queue
Message Queue
Elastic Computing
Object Storage
Object Storage
Object Storage
Indexed Storage



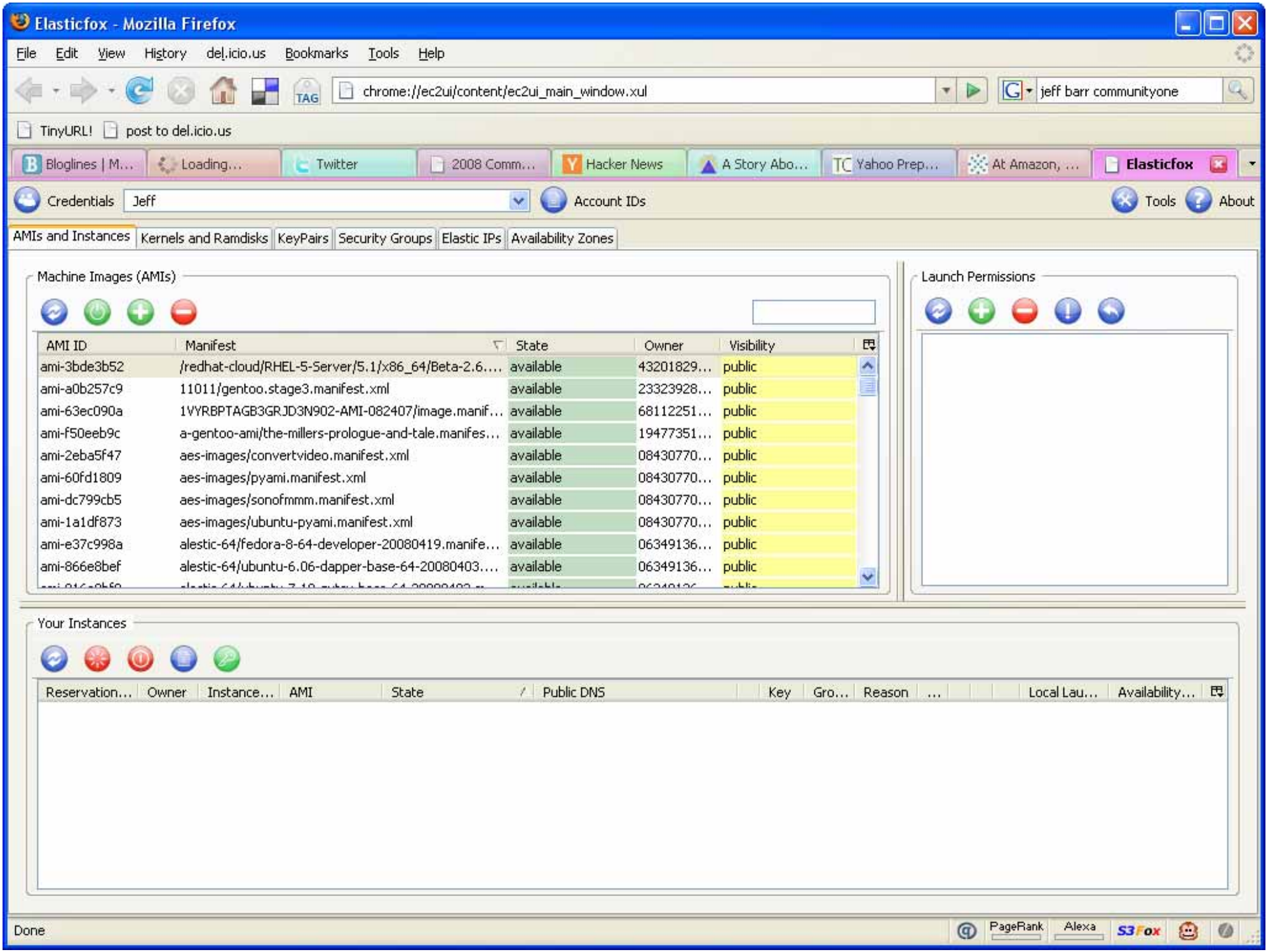
amazon web services

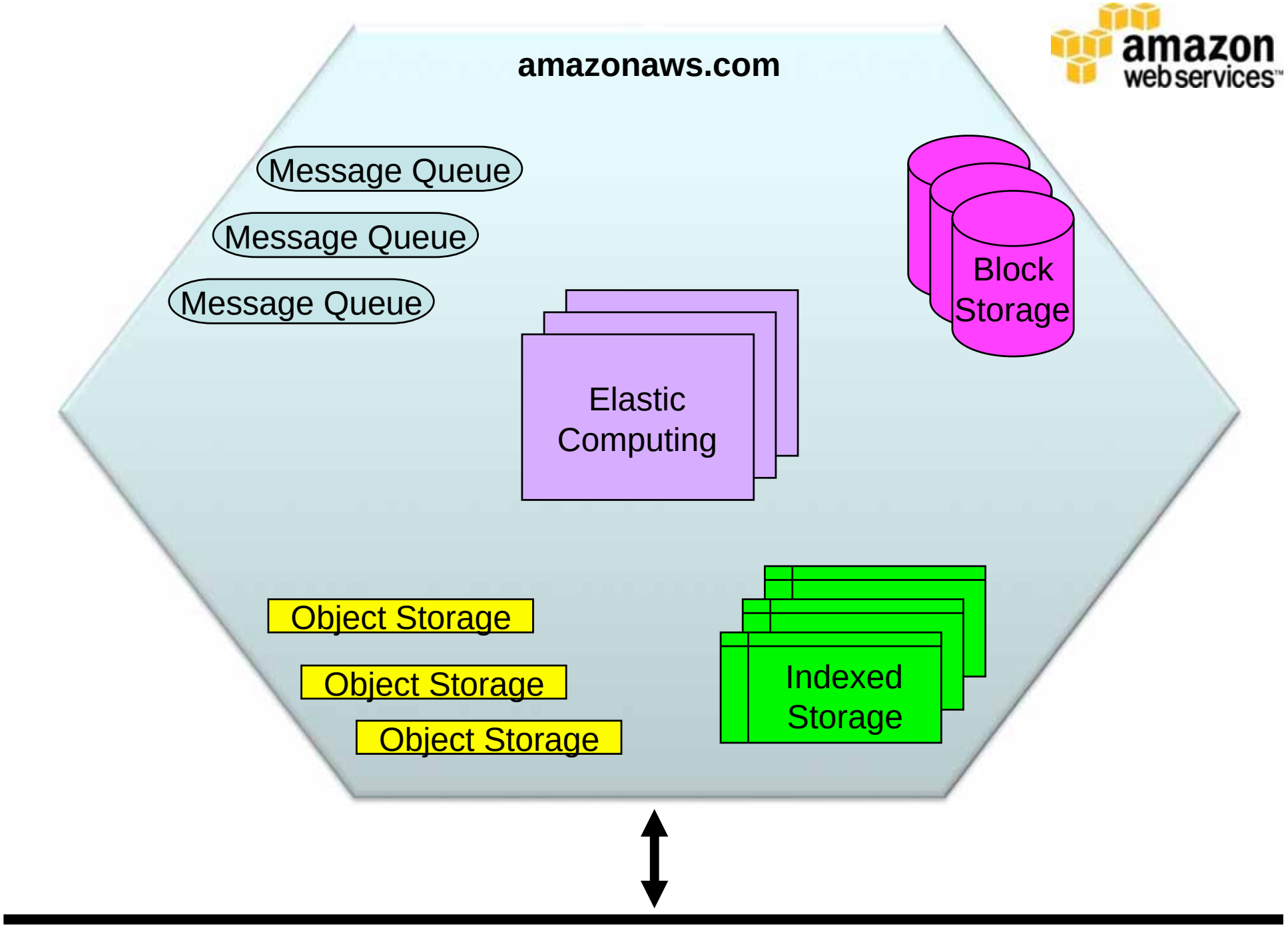
EC2

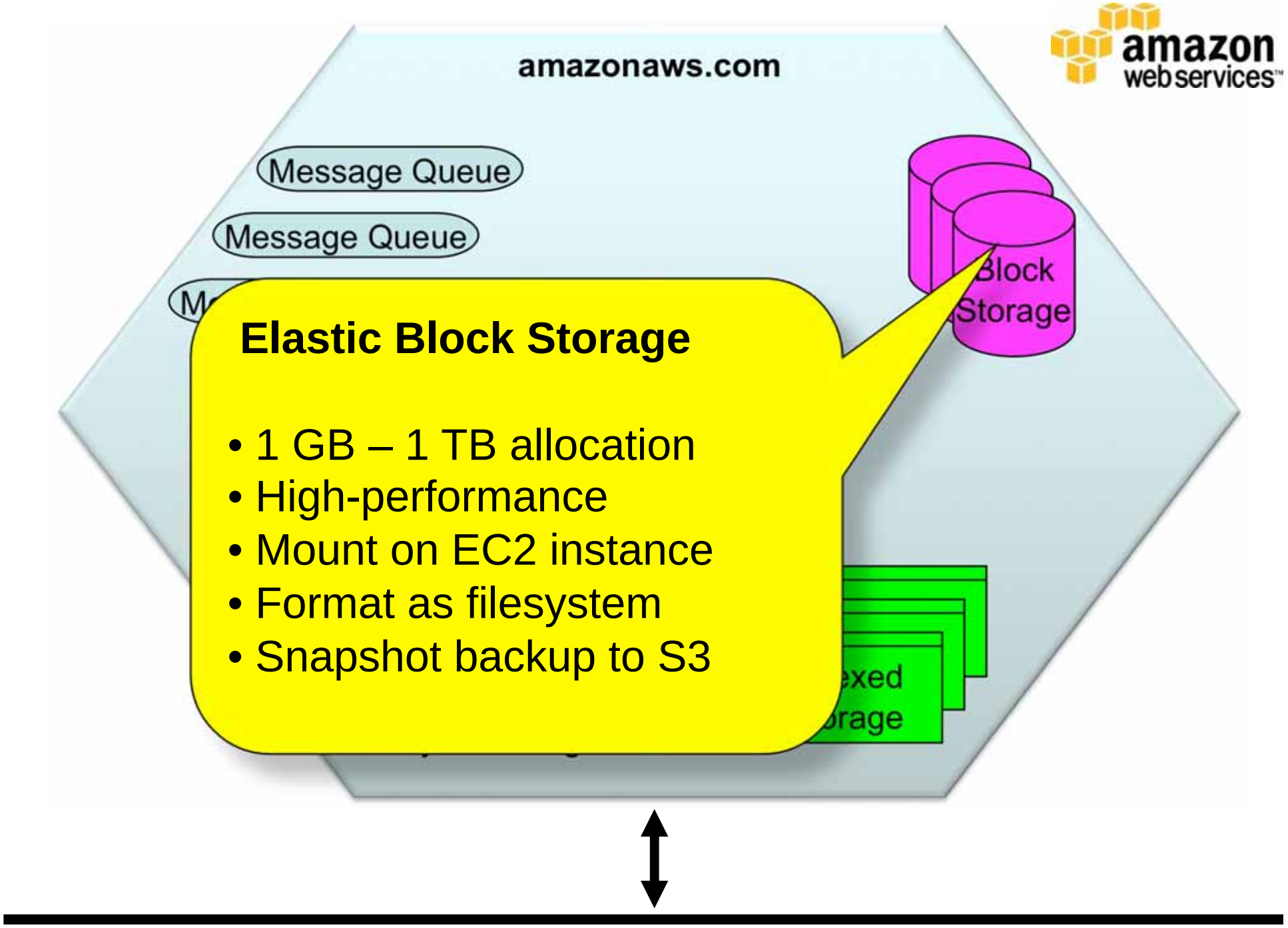
- Amazon Elastic Compute Cloud
- Scalable processing power
- Ephemeral instances
- 5 instance sizes
- Linux, OpenSolaris, or Windows
- Root-level access
- $0.10 - $0.80 / instance / hour

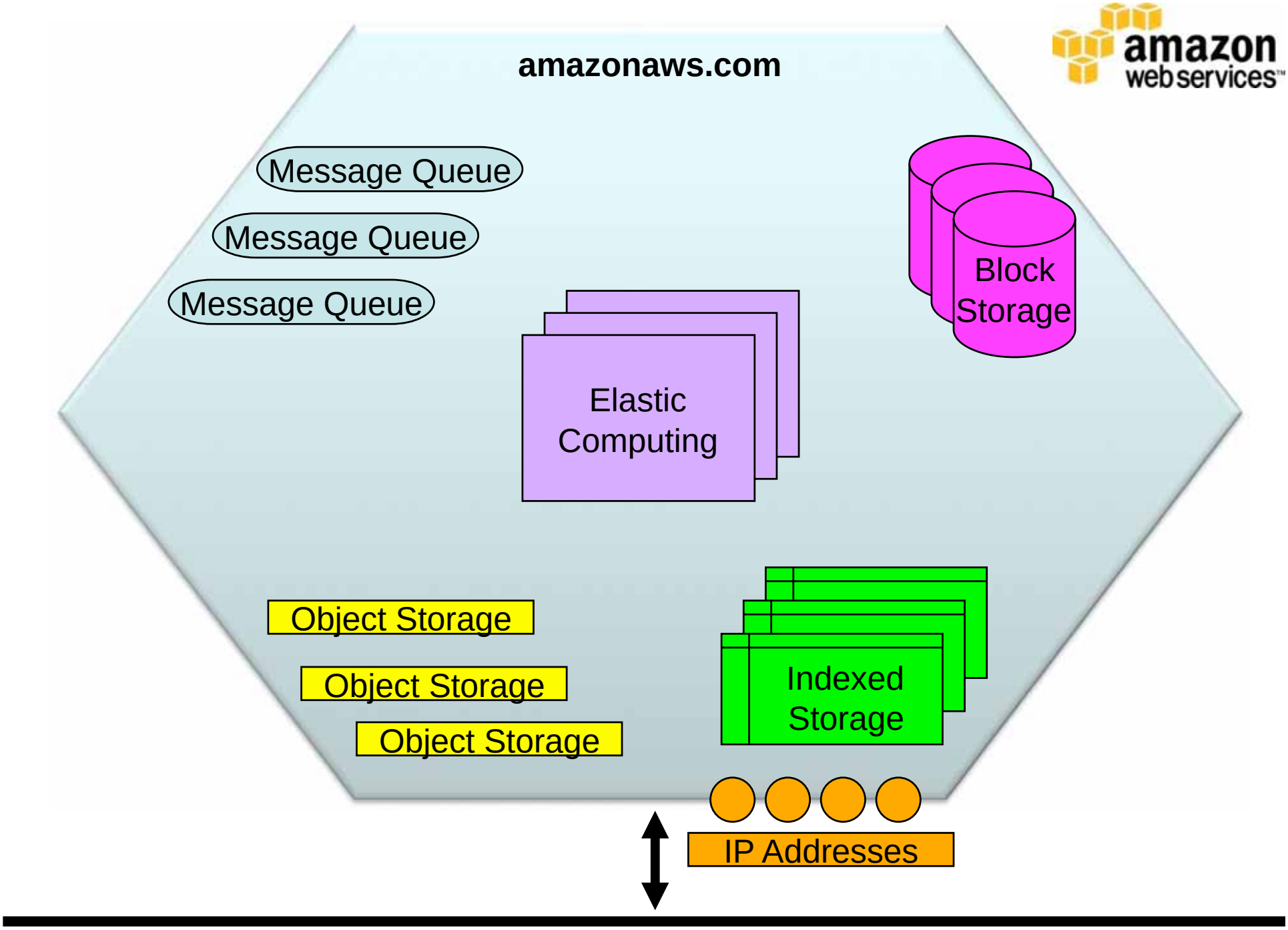
EC2 Instance Types

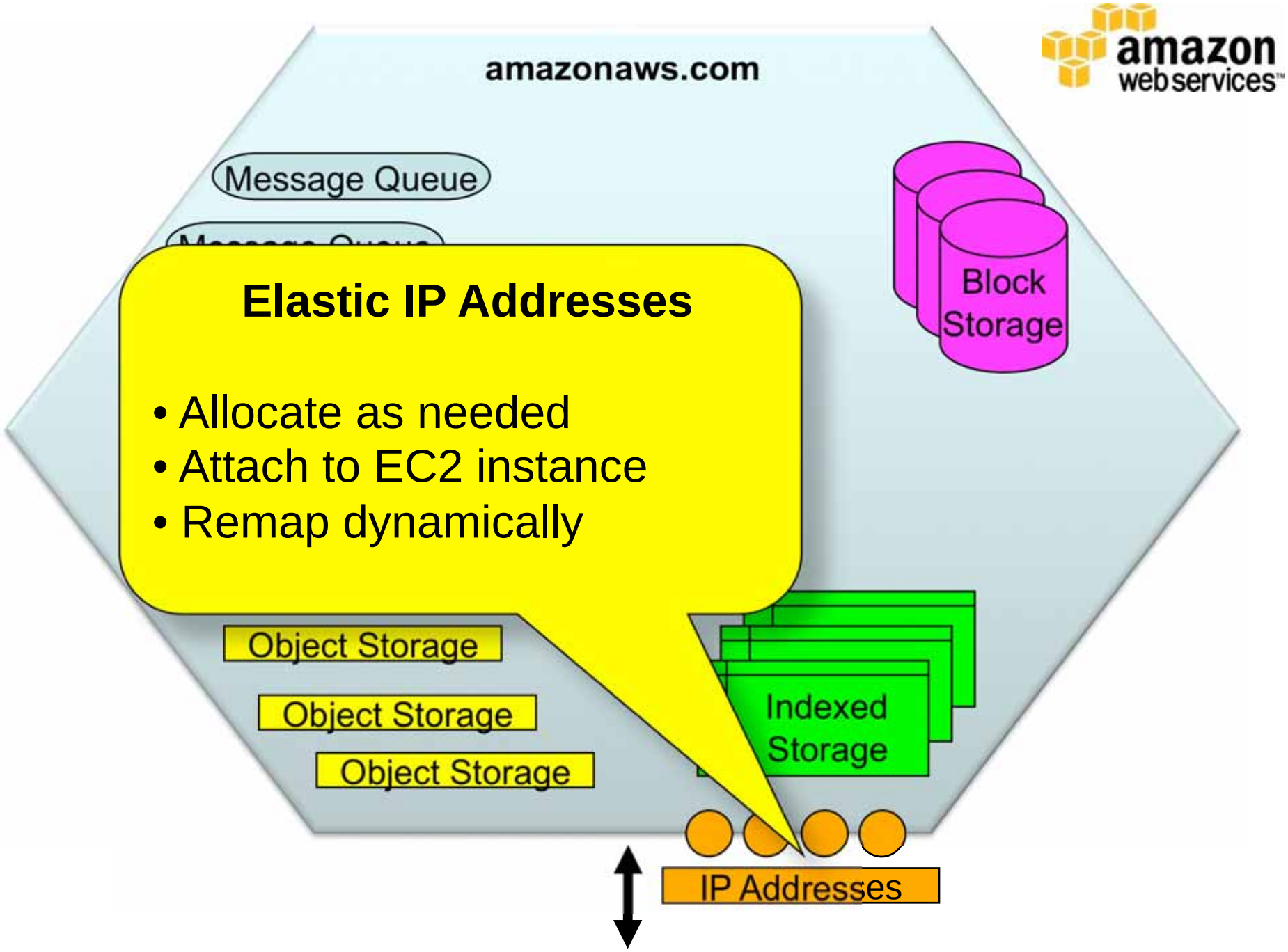
Small	Large	Extra Large	High CPU Medium	High CPU Large
1.7 GB RAM,	7.5 GB RAM	15 GB RAM	1.7 GB RAM,	7 GB RAM
1 EC2 Compute Unit (1 @ 1)	4 EC2 Compute Units (2 @ 2)	8 EC2 Compute Units (4 @ 2)	5 EC2 Compute Units (2 @ 2.5)	20 EC2 Compute Units (8 @ 2.5)
160 GB disk	850 GB disk	1690 GB disk	350 GB disk	1690 GB disk
32-bit	64-bit	64-bit	32-bit	64-bit
$0.10/Hr	$0.40/Hr	$0.80/Hr	$0.20/Hr	$0.80/Hr

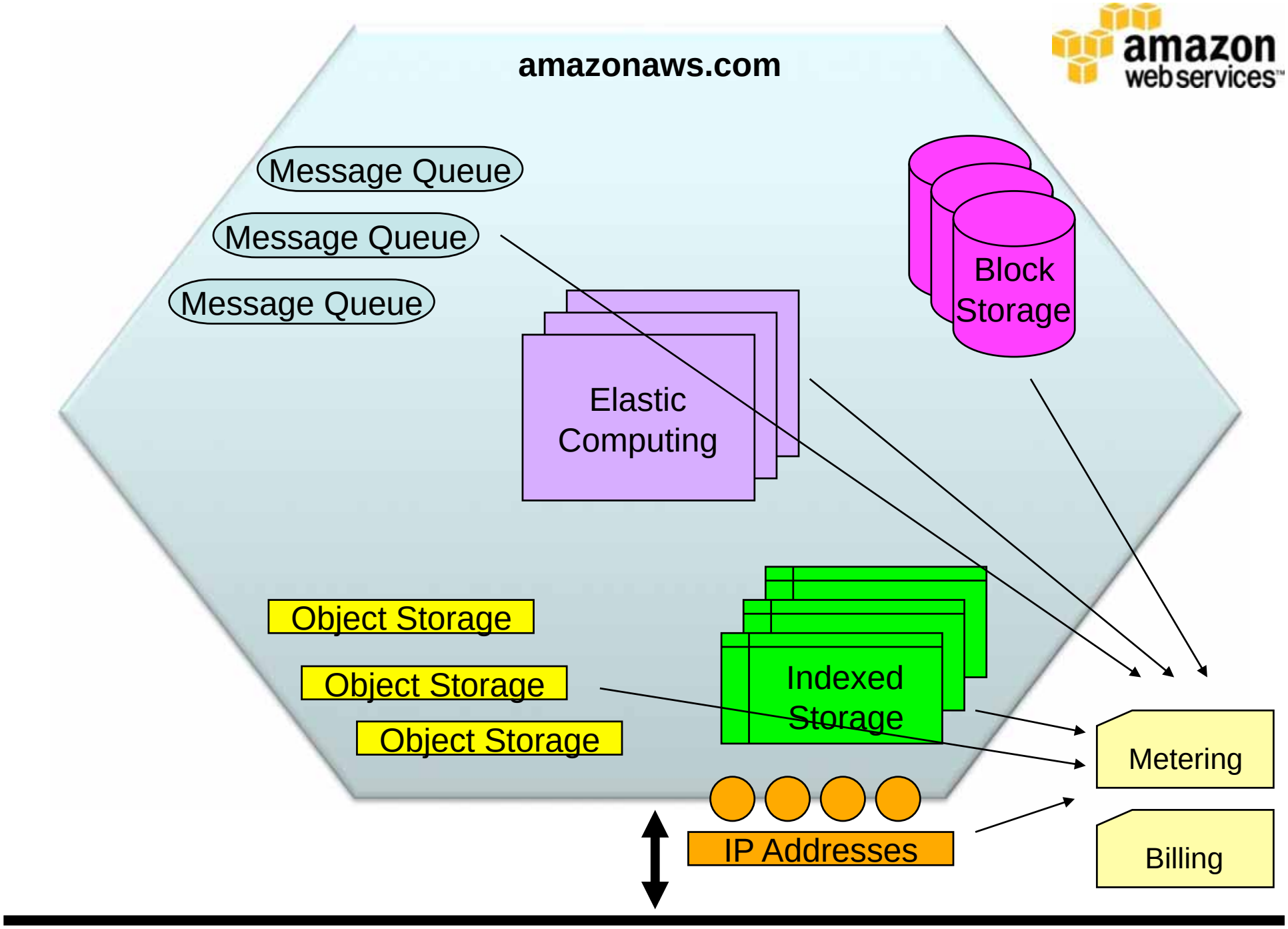

Elasticfox - Mozilla Firefox
File Edit View History del.icio.us Bookmarks Tools Help
chrome://ec2ui/content/ec2ui_main_window.xul
jeff barr communityone
TinyURL! post to del.icio.us
Bloglines | M... Loading... Twitter 2008 Comm... Hacker News A Story Abo... Yahoo Prep... At Amazon, ... Elasticfox
Credentials Jeff Account IDs Tools About
AMIs and Instances Kernels and Ramdisks KeyPairs Security Groups Elastic IPs Availability Zones
Machine Images (AMIs)
AMI ID Manifest State Owner Visibility
ami-3bde3b52 /redhat-cloud/RHEL-5-Server/5.1/x86_64/Beta-2.6.... available 43201829... public
ami-a0b257c9 11011/gentoo.stage3.manifest.xml available 23323928... public
ami-63ec090a 1VYRBPTAGB3GRJD3N902-AMI-082407/image.manif... available 68112251... public
ami-f50eeb9c a-gentoo-ami/the-millers-prologue-and-tale.manifes... available 19477351... public
ami-2eba5f47 aes-images/convertvideo.manifest.xml available 08430770... public
ami-60fd1809 aes-images/pyami.manifest.xml available 08430770... public
ami-dc799cb5 aes-images/sonofmmm.manifest.xml available 08430770... public
ami-1a1df873 aes-images/ubuntu-pyami.manifest.xml available 08430770... public
ami-e37c998a alestic-64/fedora-8-64-developer-20080419.manife... available 06349136... public
ami-866e8bef alestic-64/ubuntu-6.06-dapper-base-64-20080403.... available 06349136... public
Launch Permissions
Your Instances
Reservation... Owner Instance... AMI State Public DNS Key Gro... Reason ... Local Lau... Availability...
Done
PageRank Alexa S3Fox

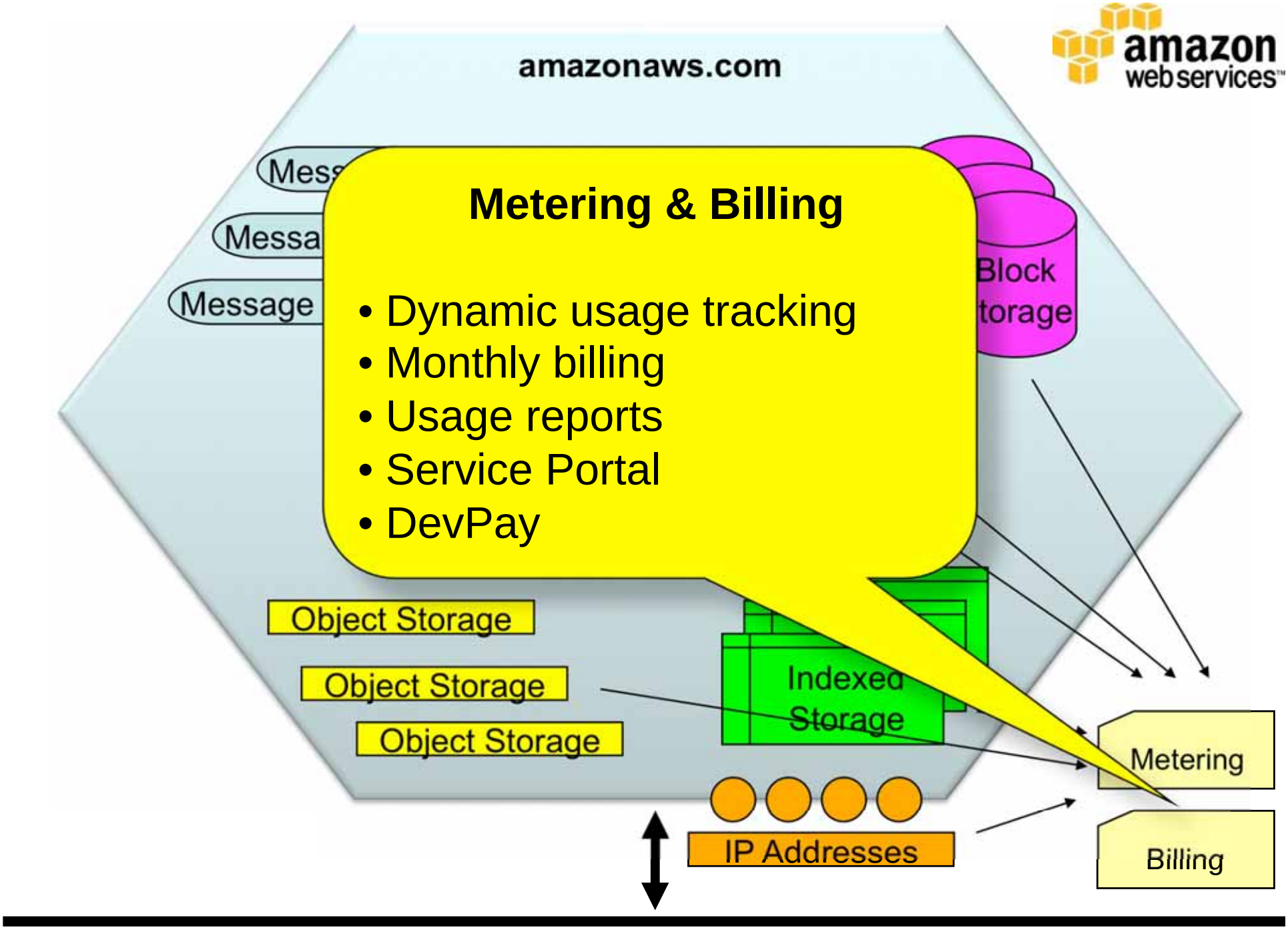

amazonaws.com
amazon web services
Message Queue
Message Queue
Message Queue
Block Storage
Elastic Computing
Object Storage
Object Storage
Object Storage
Indexed Storage


amazonaws.com
Message Queue
Message Queue
Block
Storage
Elastic Block Storage
• 1 GB – 1 TB allocation
• High-performance
• Mount on EC2 instance
• Format as filesystem
• Snapshot backup to S3

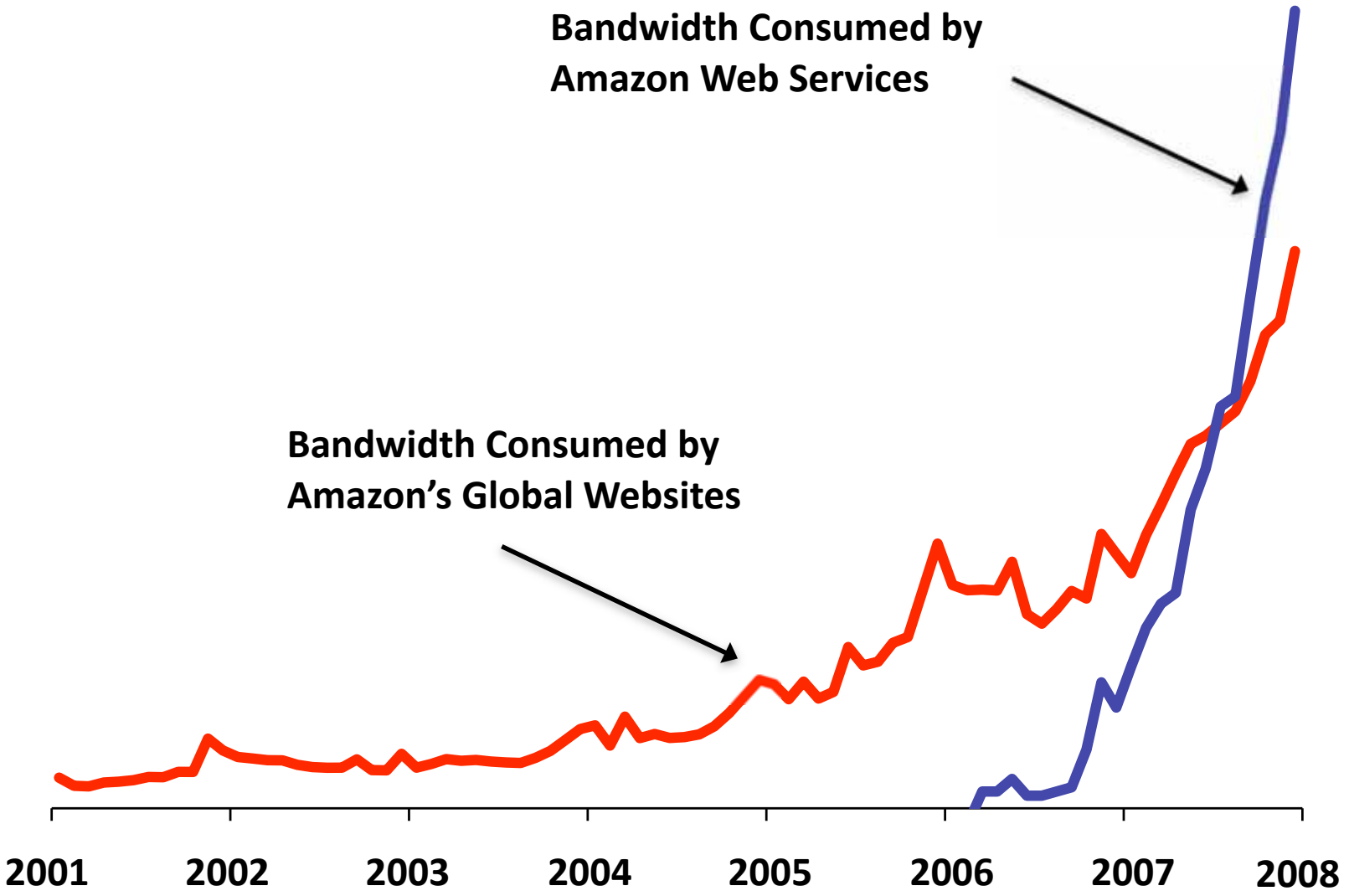

amazonaws.com
amazon web services
Message Queue
Message Queue
Message Queue
Block Storage
Elastic Computing
Object Storage
Object Storage
Object Storage
Indexed Storage
IP Addresses


amazonaws.com
Message Queue
Block Storage
Elastic IP Addresses
• Allocate as needed
• Attach to EC2 instance
• Remap dynamically
Object Storage
Object Storage
Object Storage
Indexed Storage
IP Addresses

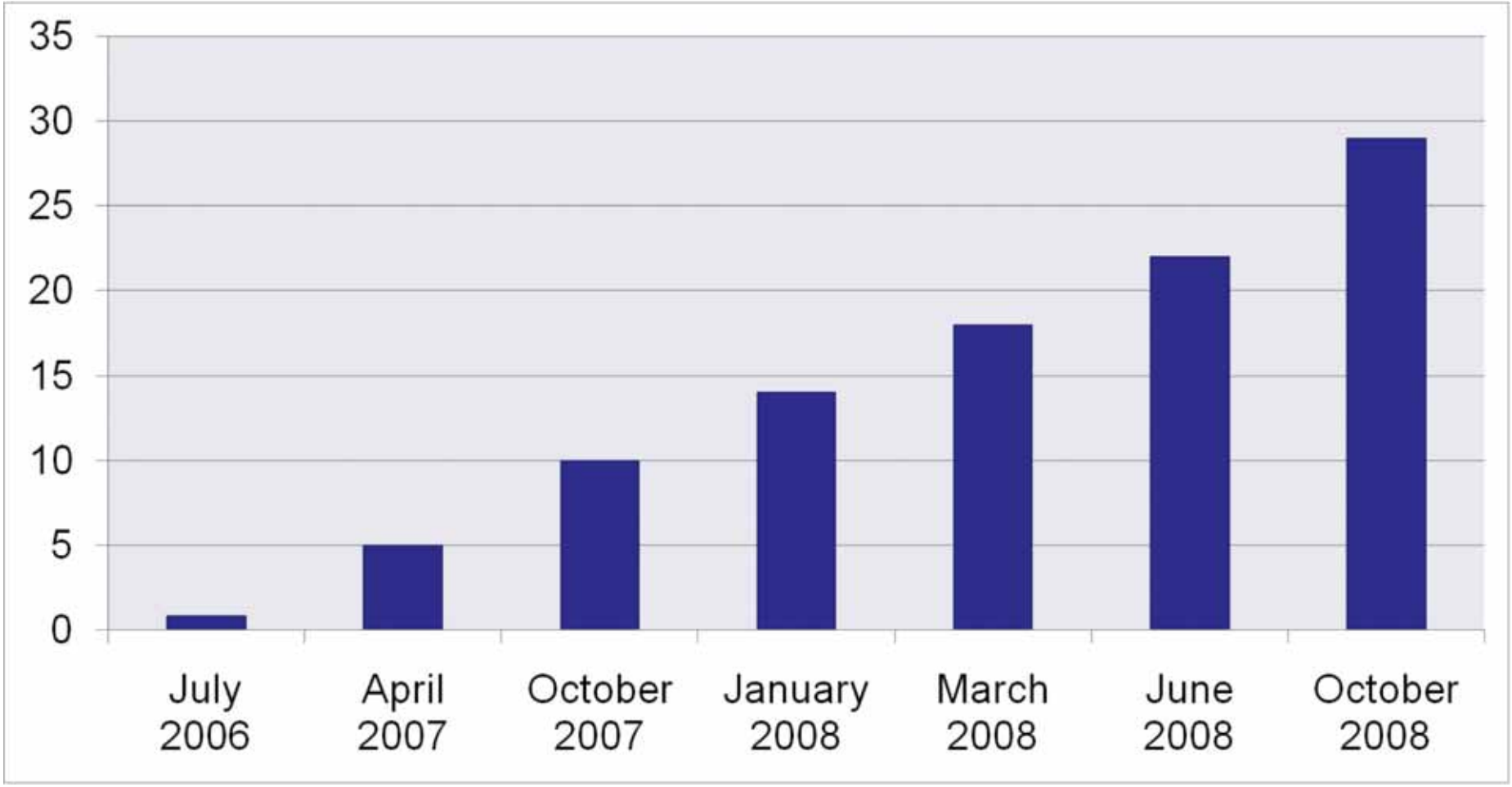

amazonaws.com
amazon web services
Message Queue
Message Queue
Message Queue
Block Storage
Elastic Computing
Object Storage
Object Storage
Object Storage
Indexed Storage
Metering
IP Addresses
Billing


amazonaws.com
amazon web services
Mess
Messa
Message
Block storage
Metering & Billing
• Dynamic usage tracking
• Monthly billing
• Usage reports
• Service Portal
• DevPay
Object Storage
Object Storage
Object Storage
Indexed Storage
Metering
IP Addresses
Billing

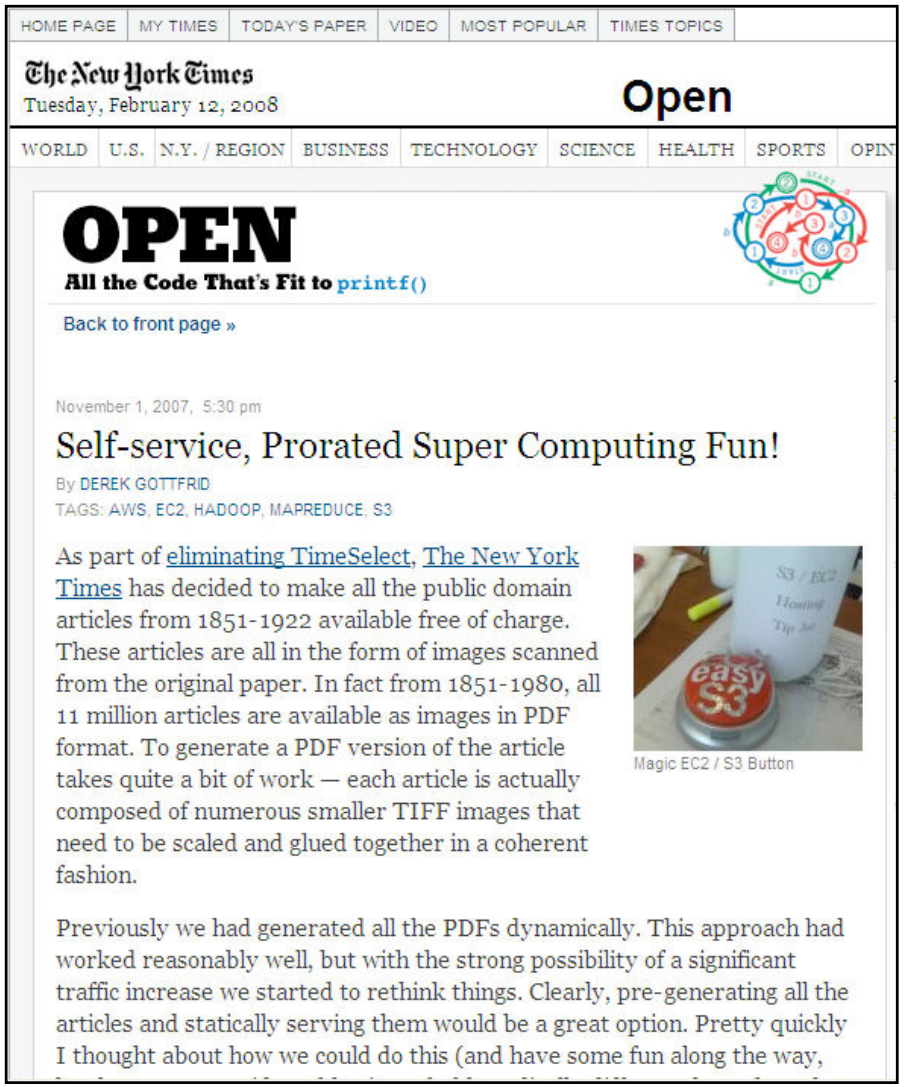
AWS
in Action


Bandwidth Consumed by
Amazon Web Services
Bandwidth Consumed by
Amazon's Global Websites
2001
2002
2003
2004
2005
2006
2007
2008

Billions of S3 Objects



“Prorated Super Computing Fun”
The New York Times



HOME PAGE | MY TIMES | TODAY'S PAPER | VIDEO | MOST POPULAR | TIMES TOPICS

The New York Times
Tuesday, February 12, 2008

Open

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPIN

OPEN

All the Code That's Fit to printf()

Back to front page »

November 1, 2007, 5:30 pm

Self-service, Prorated Super Computing Fun!

By DEREK GOTTFRID
TAGS: AWS, EC2, HADOOP, MAPREDUCE, S3

As part of eliminating TimeSelect, The New York Times has decided to make all the public domain articles from 1851-1922 available free of charge. These articles are all in the form of images scanned from the original paper. In fact from 1851-1980, all 11 million articles are available as images in PDF format. To generate a PDF version of the article takes quite a bit of work — each article is actually composed of numerous smaller TIFF images that need to be scaled and glued together in a coherent fashion.

Magic EC2 / S3 Button

Previously we had generated all the PDFs dynamically. This approach had worked reasonably well, but with the strong possibility of a significant traffic increase we started to rethink things. Clearly, pre-generating all the articles and statically serving them would be a great option. Pretty quickly I thought about how we could do this (and have some fun along the way,

- 1851-1922 Articles
- TIFF -> PDF
- Input: 11 Million Articles (4TB of data)
- What did they do?
  - 100 EC2 Instances for 24 hours
  - All data on S3
  - Output: 1.5 TB of Data
  - Hadoop, iText, JetS3t

amazon web services™

Washington Post

- Hillary Clinton’s public record
- How to process 11,000 pages of data
- Developer suggests EC2
- 200 instance EC2 job running 12 hours later
- 11 hours to OCR and index

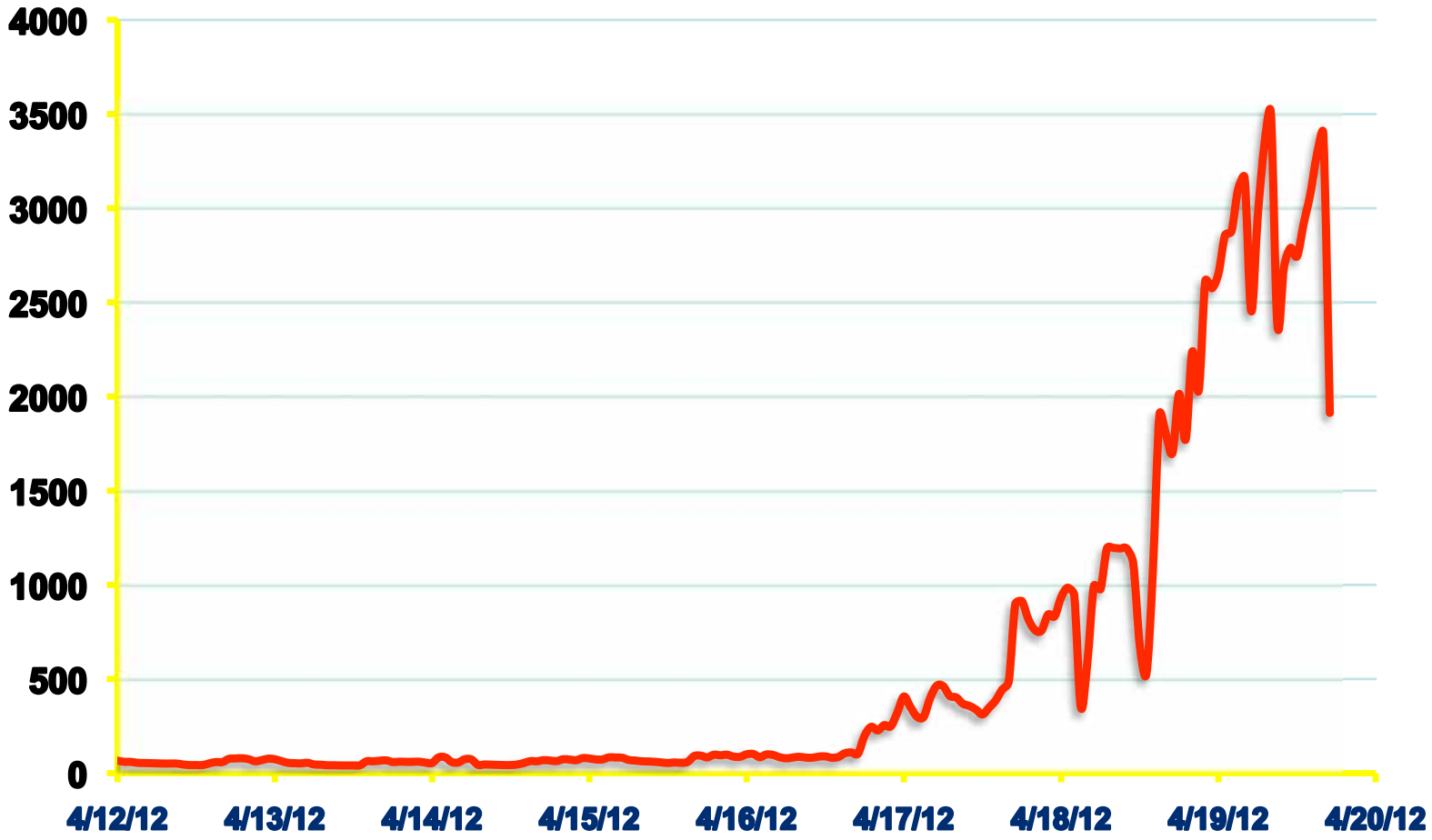
“This has been a big internal success: the politics dept hadn't thought that a 17k pdf could be made searchable, let alone in 12 hours. So it's definitely changing the understanding of what's possible and encouraged them to think big…”

Animoto

- Dynamic Music Video Generator
- Upload images
- Upload music
- Generate music video
- Facebook Application

50 Servers to 3500 Servers in 3 Days

animoto beta



AWS Security White Paper

- Certifications
  - SOX
  - SAS 70 Type II
  - HIPAA
- Physical security
- System security
- Backups

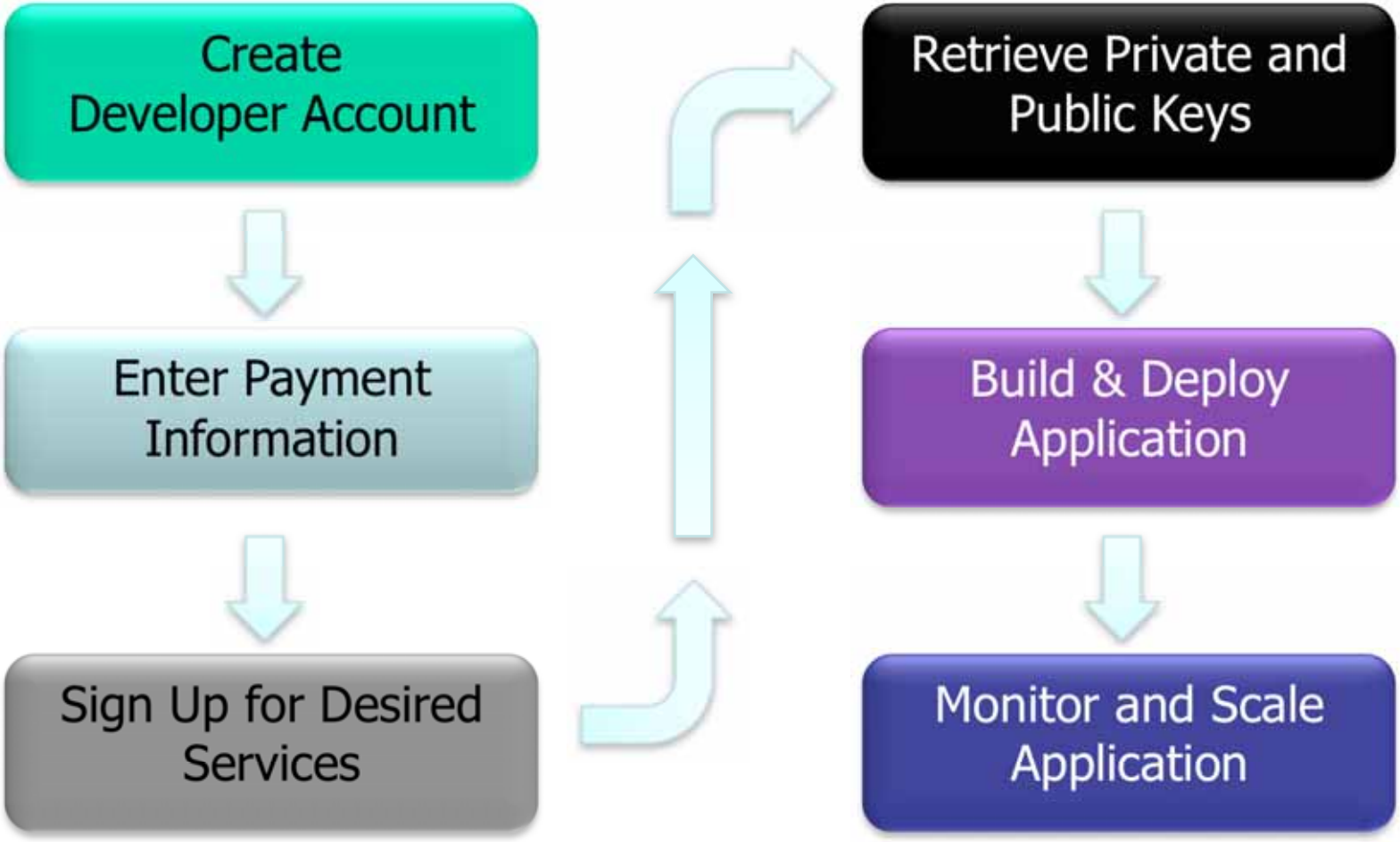
amazon web services™

Amazon Web Services:
Overview of Security Processes

September 2008
(Please consult http://aws.amazon.com/ for the latest version of this paper)

Amazon Web Services (AWS) delivers a highly scalable cloud computing platform with high availability and dependability, and the flexibility to enable customers to build a wide range of applications. The issues of end-to-end security and end-to-end privacy within the cloud computing world are more sophisticated than within a single data center not facing the Internet. Ensuring the confidentiality, integrity, and availability of customer's systems and data is of the utmost importance to AWS, as is maintaining trust and confidence. This document is intended to answer customer questions such as "How does AWS help me ensure my data is secure?" Specifically, AWS physical and operational security processes are described for network and infrastructure under AWS' management, as well as service-specific security implementations.

Getting Started With AWS



Important AWS Sites

- AWS Home Page
  aws.amazon.com
- AWS Blog:
  aws.typepad.com
- AWS Service Health Dashboard:
  status.aws.amazon.com
- AWS Forums:
  aws.amazon.com/forums
- AWS Zone:
  www.awszone.com

Q&A

Thank You!

Jeff Barr – jbarr@amazon.com