



Big Data for Government Symposium

<http://www.ttcus.com>



@TECHTrain

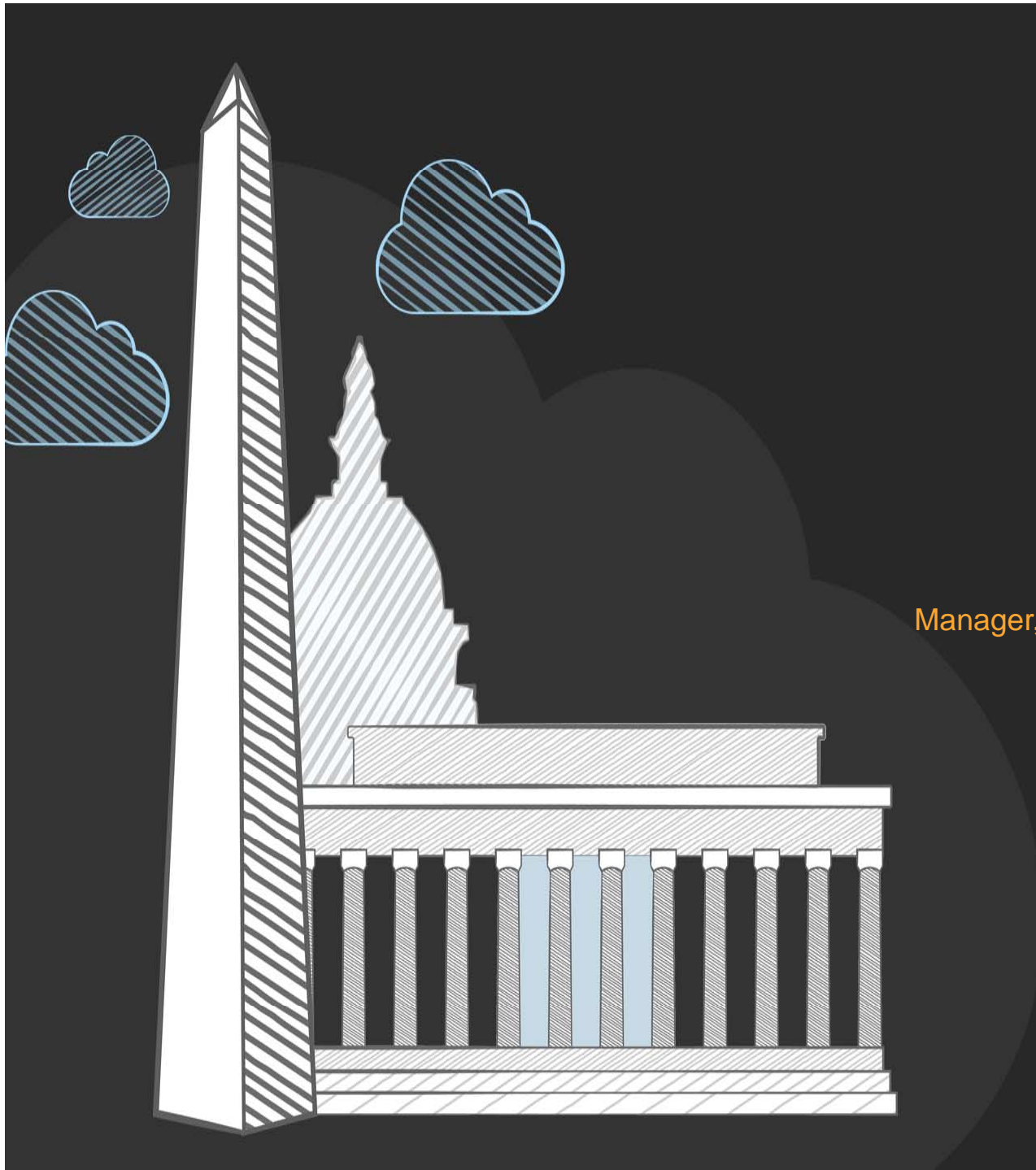


TTC™

Technology Training Corporation

Linkedin/Groups:
Technology Training
Corporation





Big Data on AWS

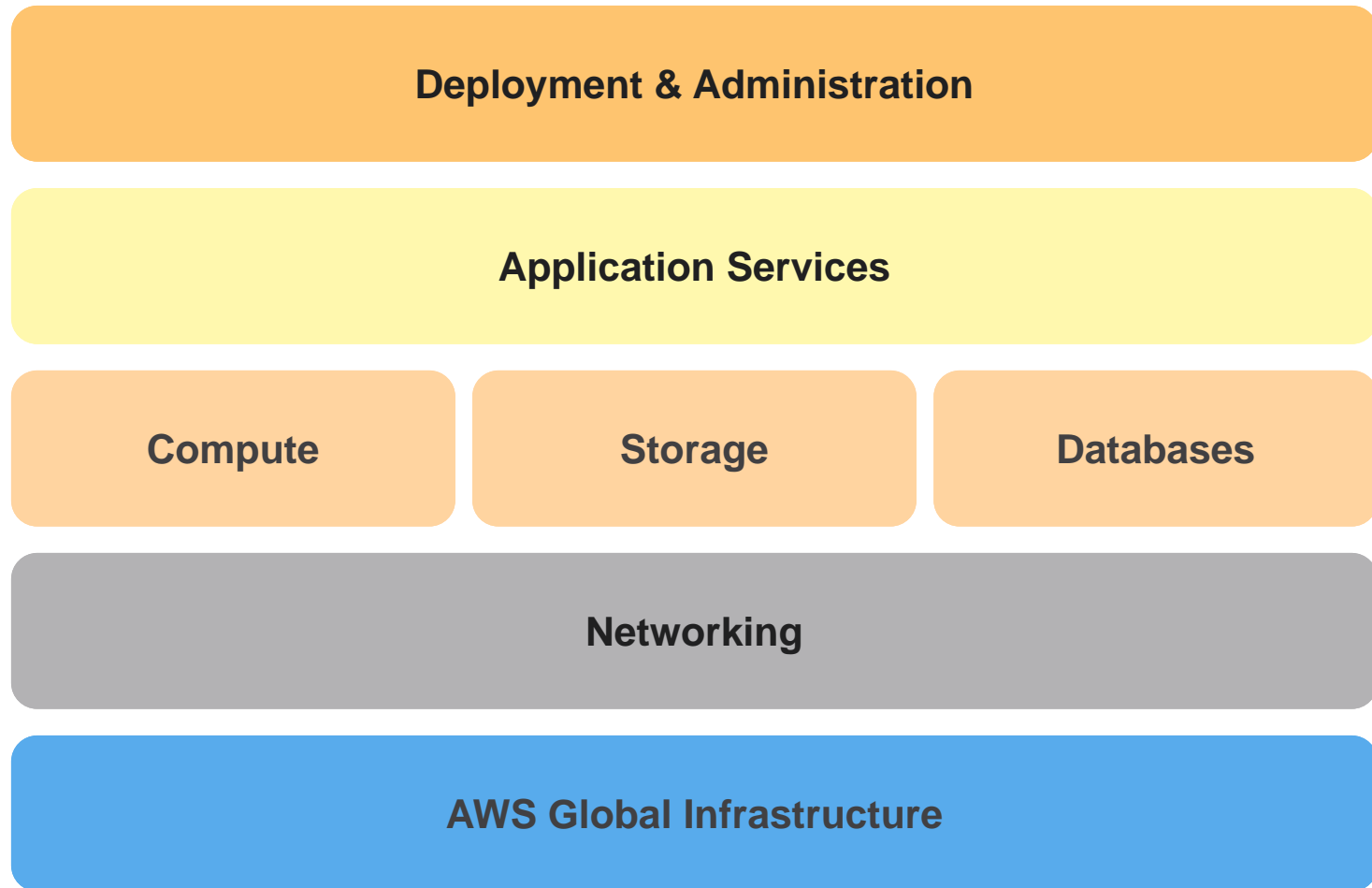
Tim Bixler

Manager, U.S. Federal Solutions Architecture

Principal Solutions Architect



AWS Platform



AWS Worldwide Public Sector



AWS Global Infrastructure

10 Regions

including GovCloud US

25 Availability Zones

51 Edge Locations

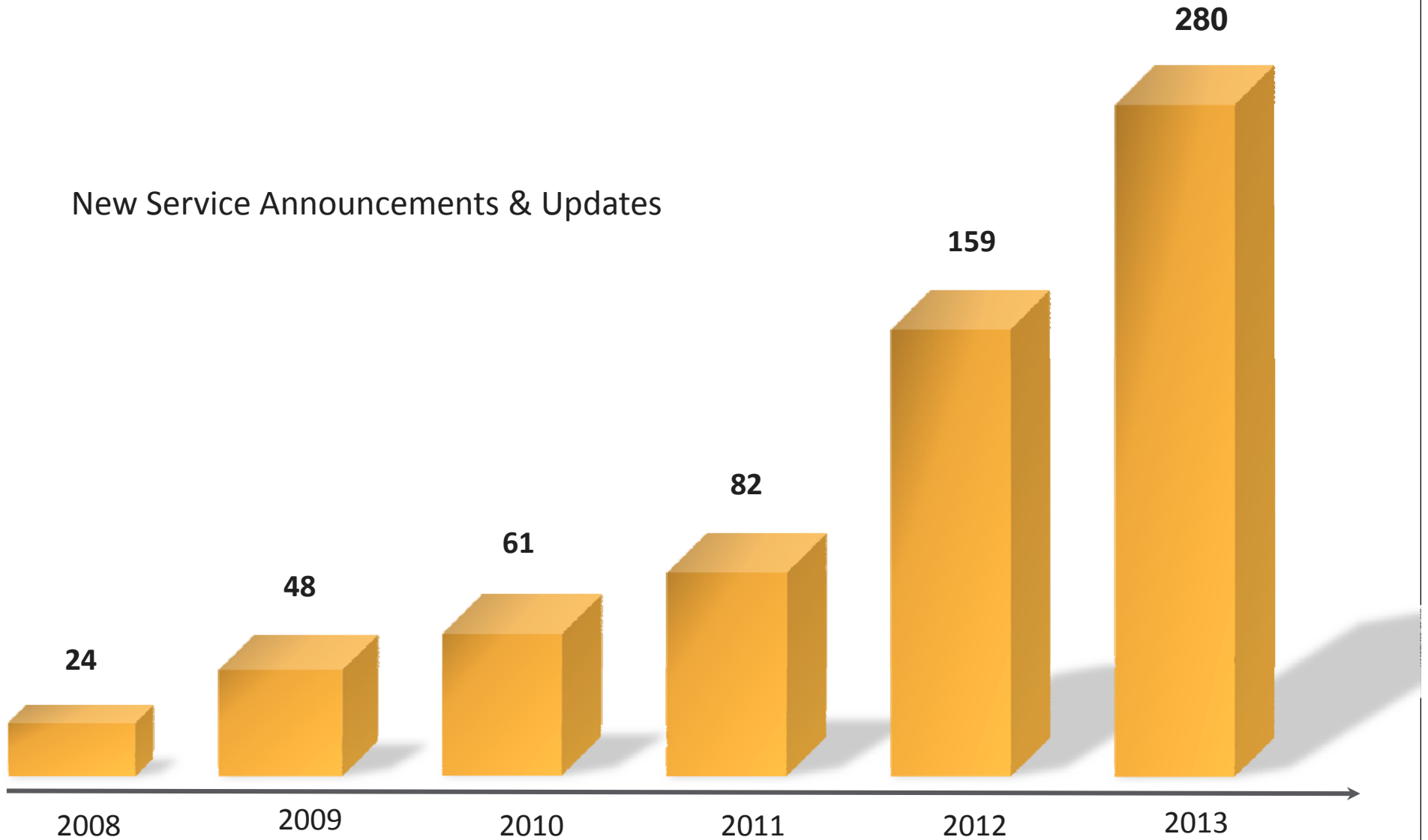


Each day AWS adds the equivalent server capacity to power
Amazon when it was a global, **\$7B** enterprise

(circa 2004)

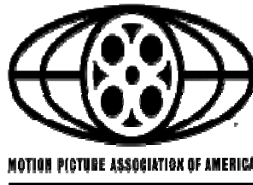
AWS Pace of Innovation

New Service Announcements & Updates



Built for Enterprise Security Requirements

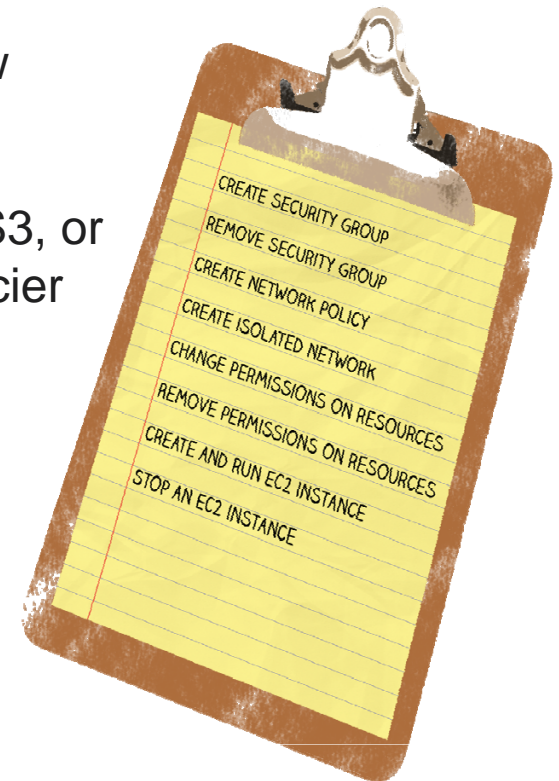
Government and Industry Certifications and Accreditations



AWS CloudTrail - AWS API call logging
for governance & compliance

Log and review
user activity

Store data in S3, or
archive to Glacier

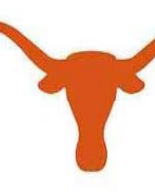


“AWS is the overwhelming market share leader, with **more than five times the compute capacity** in use than the aggregate total of the other 14 providers.”



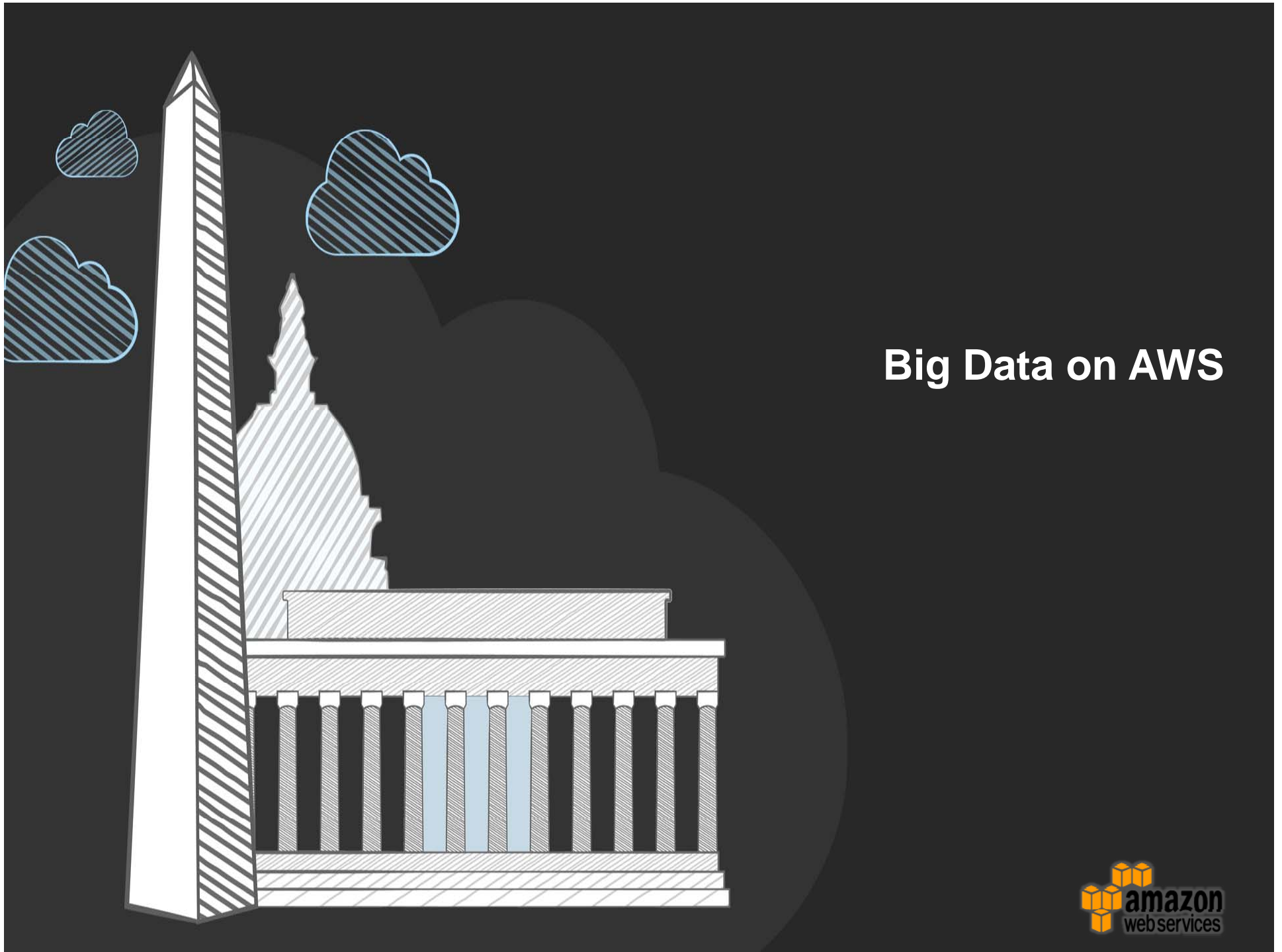
Gartner “Magic Quadrant for Cloud Infrastructure as a Service,” Lydia Leong, Douglas Toombs, Bob Gill, Gregor Petri, Tiny Haynes, May 28, 2014. This Magic Quadrant graphic was published by Gartner, Inc. as part of a larger research note and should be evaluated in the context of the entire report. The Gartner report is available at <http://aws.amazon.com/resources/analyst-reports/>. Gartner does not endorse any vendor, product or service depicted in its research publications, and does not advise technology users to select only those vendors with the highest ratings. Gartner research publications consist of the opinions of Gartner’s research organization and should not be construed as statements of fact. Gartner disclaims all warranties, expressed or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.

Used by Government Agencies & Educational Institutions Worldwide



AWS Partners Focused on Public Sector





Big Data on AWS

Generation



Collection & storage



Analytics & computation



Collaboration & sharing



AWS Worldwide Public Sector



Generation



Collection & storage



Analytics & computation



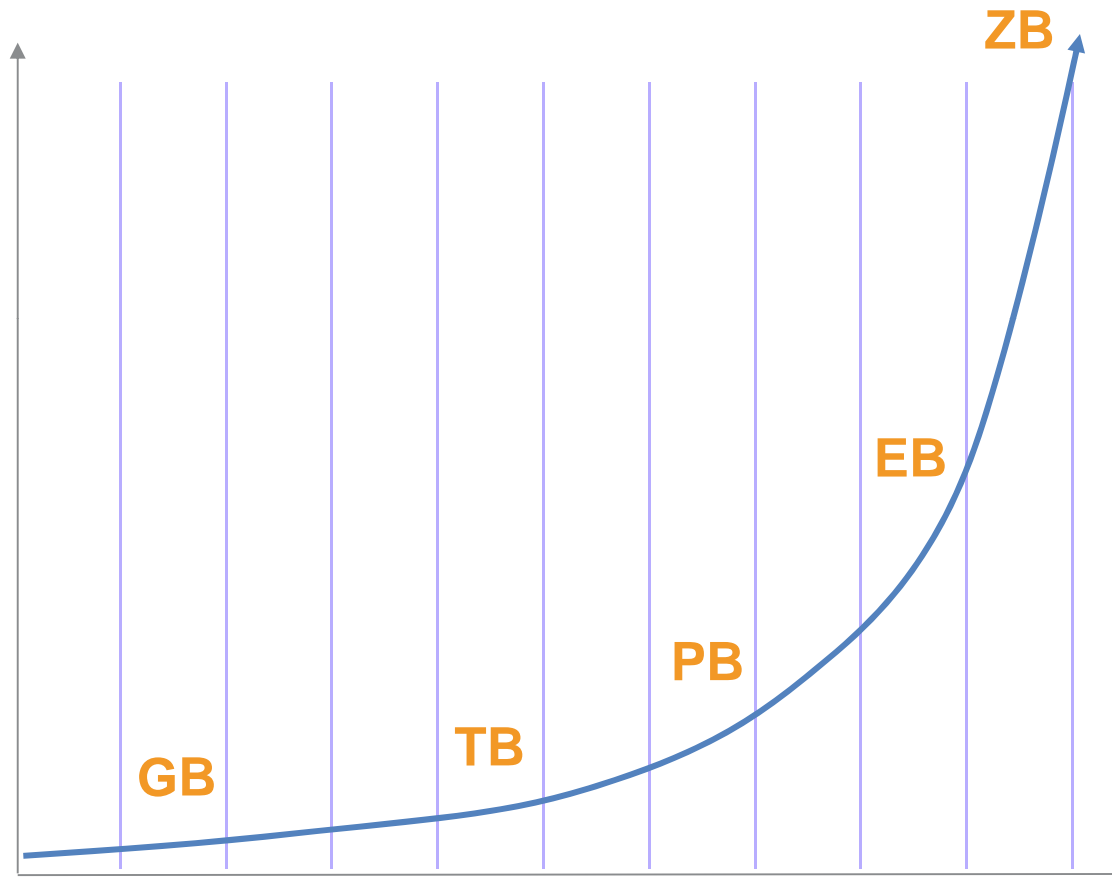
Collaboration & sharing



AWS Worldwide Public Sector



Big Data: Unconstrained data growth



95% of the 1.2 zettabytes of data in the digital universe is unstructured

70% of of this is user-generated content

Unstructured data growth explosive, with estimates of compound annual growth (CAGR) at 62% from 2008 – 2012.

Source: IDC



AWS Worldwide Public Sector



Lower cost,
higher throughput



Generation



Collection & storage



Analytics & computation



Collaboration & sharing



AWS Worldwide Public Sector



Lower cost,
higher throughput



Generation



Collection & storage

Highly
constrained



Analytics & computation

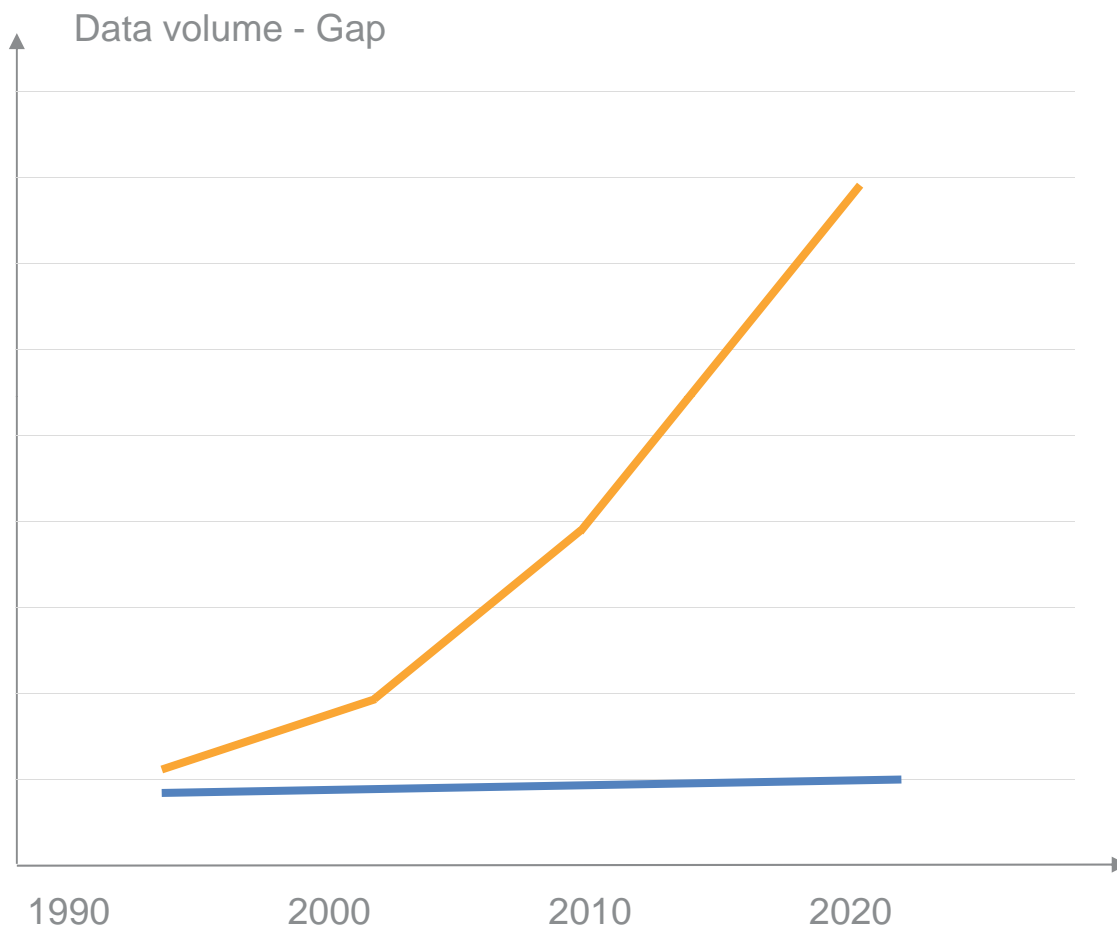


Collaboration & sharing



AWS Worldwide Public Sector





● Generated data

● Available for analysis



Source: User Survey Analysis: Key Trends Shaping the Future of Data Center Infrastructure Through 2011
2012 Worldwide Business Analytics Software 2012–2016 Forecast and 2011 Vendor Shares

AWS Worldwide Public Sector



Elastic and highly scalable

+

No upfront capital expense

+

Only pay for what you use

+

Available on-demand

=

Remove
constraints



AWS Worldwide Public Sector



Accelerated

Generation



Collection & storage



Analytics & computation



Collaboration & sharing



AWS Worldwide Public Sector



Big Data

Technologies and techniques for working productively with data, at any scale.



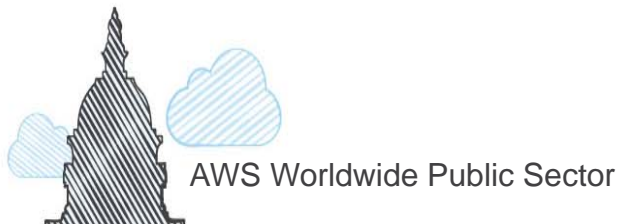
Big data and AWS Cloud computing

Big data

Potentially massive datasets

Cloud computing

Massive, virtually unlimited capacity



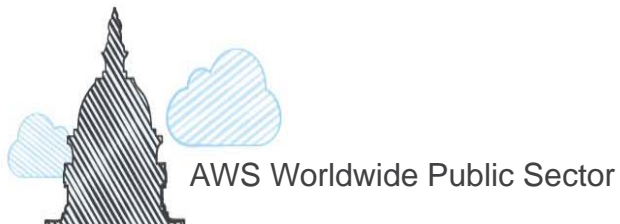
Big data and AWS Cloud computing

Big data

Iterative, experimental style of data manipulation and analysis

Cloud computing

Iterative, experimental style of infrastructure deployment/usage



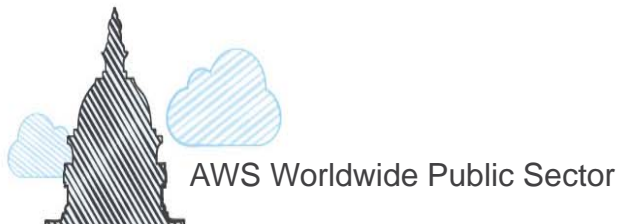
Big data and AWS Cloud computing

Big data

Frequently not steady-state workload; peaks and valleys

Cloud computing

At its most efficient with highly variable workloads



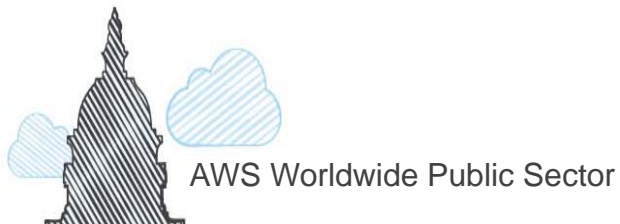
Big data and AWS Cloud computing

Big data

Absolute performance not as critical as “time to results”;
shared resources are a bottleneck

Cloud computing

Parallel compute projects allow
each workgroup to have more
autonomy, get faster results



Lower costs

Ease of use



Lower costs

Only pay for what you
use

No capital investment

Pay as you go



Programmable

Integrate with existing
tools

Zero admin

Easy to configure

Ease of use



Use the right tools



Amazon
S3



Amazon
Kinesis



Amazon
DynamoDB



Amazon
Redshift



Amazon
Elastic
MapReduce



AWS Worldwide Public Sector





Amazon
S3

- High scalable object store
- 99.999999999% durability
- AES-256 bit encryption
- Data lifecycle management





Amazon
Kinesis

- Real-time processing
- High throughput
- Elastic
- Integrates with EMR, S3, Redshift, DynamoDB





Amazon
DynamoDB

- NoSQL database
- Seamless scalability
- Low admin
- Single digit millisecond latency





Amazon
Redshift

- Relational data warehouse
- Massively parallel
- Petabyte scale
- Fully managed
- Low cost (\$1K/TB/Year with 3 year Reservation)





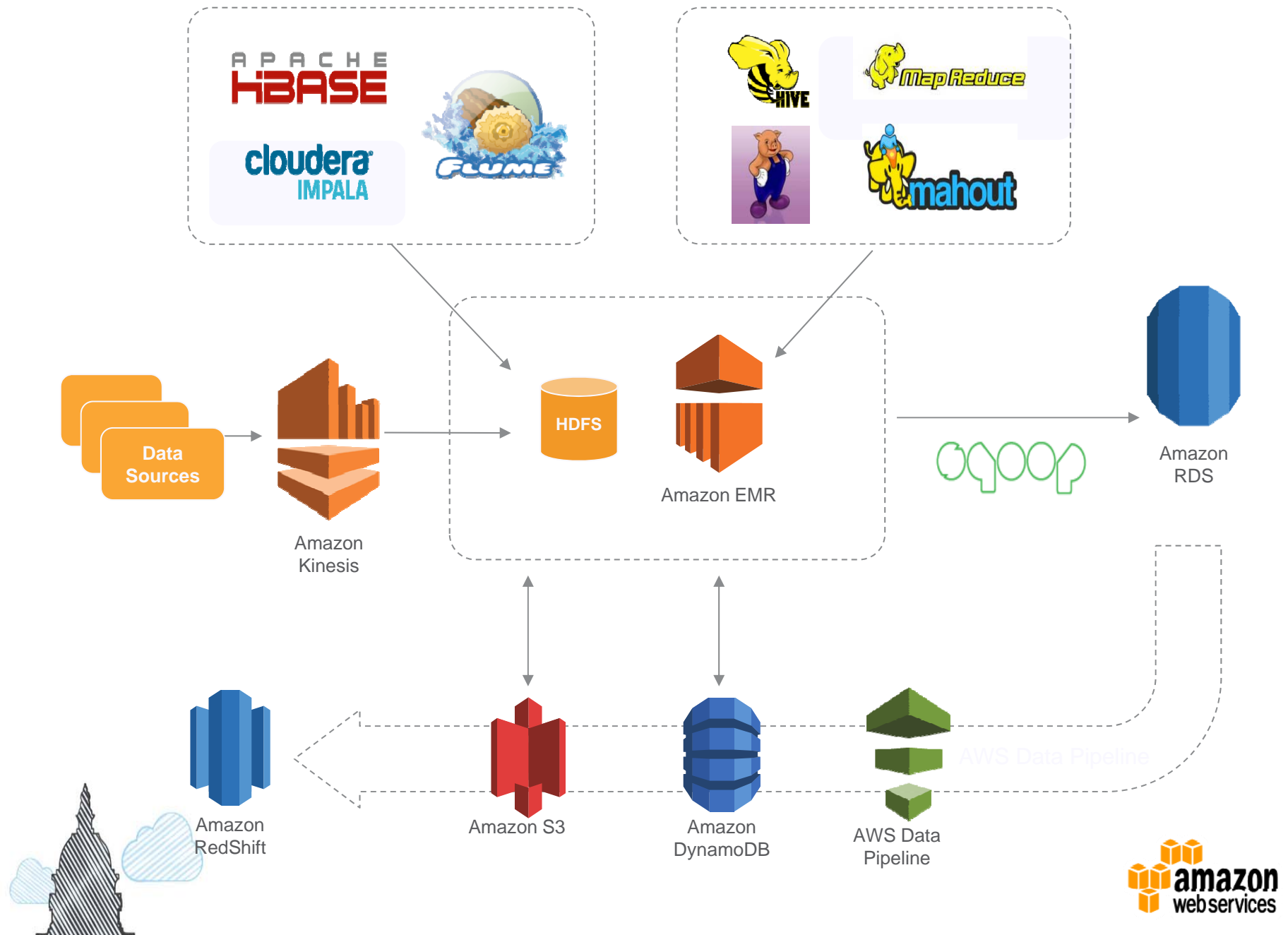
Amazon
Elastic
MapReduce

- Managed Hadoop clusters
- MapReduce, Hive, Pig, Impala, HBase, Spark, Accumulo, etc.
- Integrates with S3, DynamoDB, Redshift, Data Pipeline, Kinesis



Data management

Hadoop Ecosystem analytical tools



Generation



Collection & storage



Analytics & computation



Collaboration & sharing



AWS Worldwide Public Sector



Generation



Collection & storage



Analytics & computation



Collaboration & sharing



AWS Worldwide Public Sector



AWS
Direct Connect



AWS
Storage Gateway



AWS
Import/ Export



Amazon
Kinesis



S3



Amazon EMR



Amazon
Glacier



Amazon
Redshift



Amazon
DynamoDB



Amazon
RDS



Generation



Collection & storage



Analytics & computation



Collaboration & sharing



AWS Worldwide Public Sector



Generation



Collection & storage



Analytics & computation



Collaboration & sharing



AWS Worldwide Public Sector



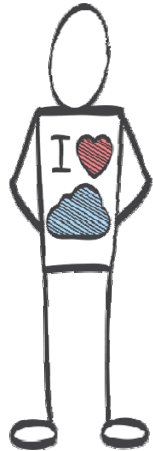
**The right tools.
At the right scale.
At the right time.**



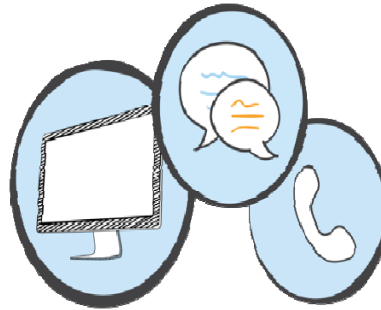
AWS is here to help



Solution
Architects



Professional
Services



Premium
Support



AWS Partner
Network (APN)

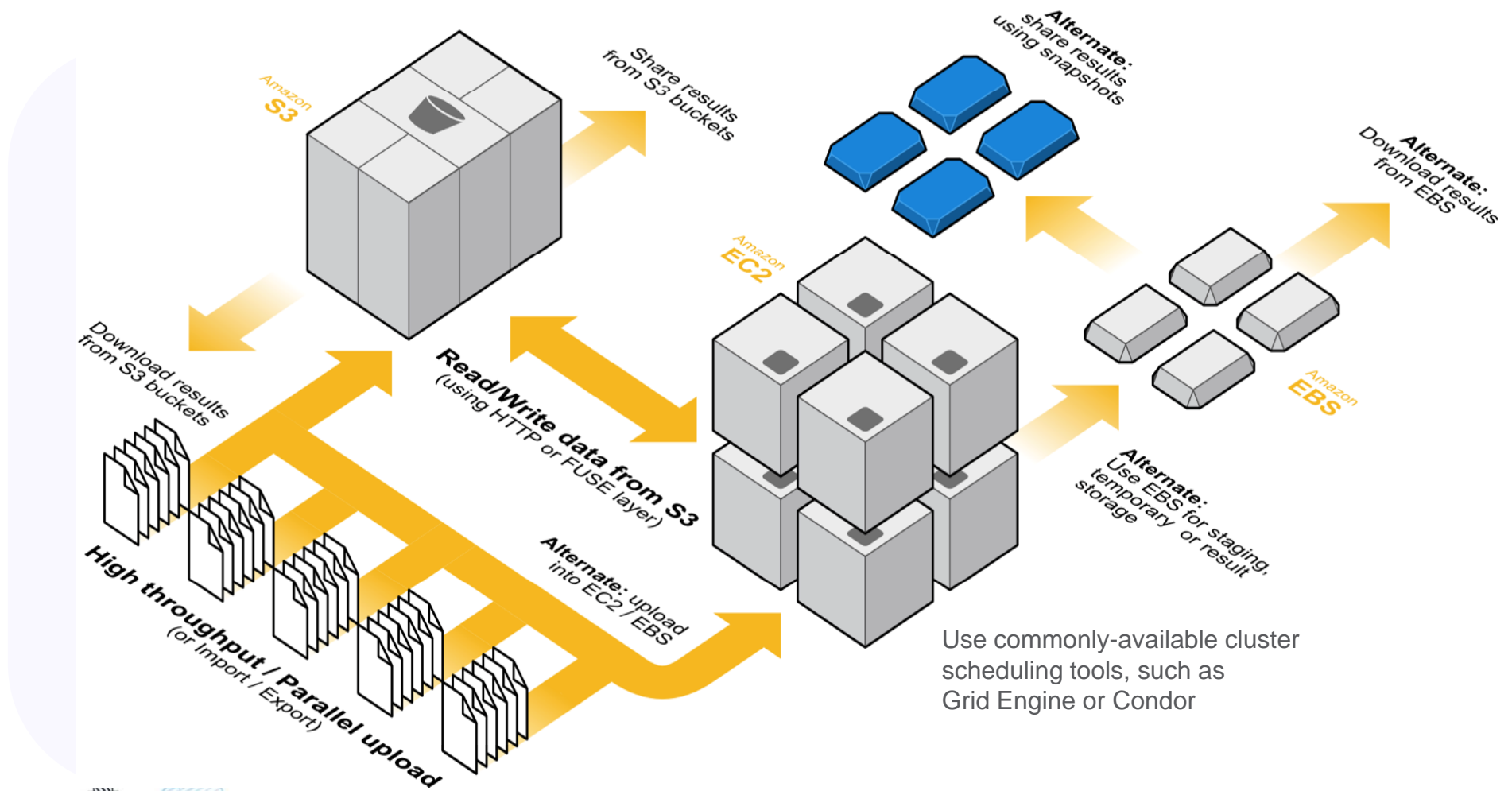


AWS Worldwide Public Sector



AWS Architecture Center

Processing large amounts of parallel data using a scalable cluster



<https://aws.amazon.com/architecture/>



Big Data & HPC. Powered by the AWS Cloud.



"It used to take us 12 months to deliver a software product. Since moving to the AWS Cloud, we're now delivering in 4 months, from concept to production"

Mark Field, CTO and Vice President, Life Technologies.



[Read the Story »](#)



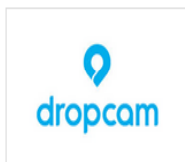
Life Technologies

[Read the Story »](#)



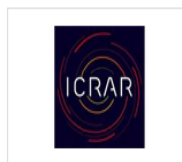
LinkedIn

[Read the Story »](#)



Dropcam

[Learn More »](#)



ICRAR

[Read the Story »](#)



Centers for Disease Control and Prevention (CDC)



Channel4

[Watch the Video »](#)



Yelp

[Read the Story »](#)



Nokia

[Watch the Video »](#)



AWS Worldwide Public Sector

Big Data Case Studies

Learn from other AWS customers

aws.amazon.com/solutions/case-studies/big-data



AWS Marketplace

AWS Online Software Store

aws.amazon.com/marketplace

Shop the big data category



AWS Worldwide Public Sector





<http://aws.amazon.com/marketplace>



AWS Worldwide Public Sector

AWS Public Data Sets

Free access to big data sets

aws.amazon.com/publicdatasets



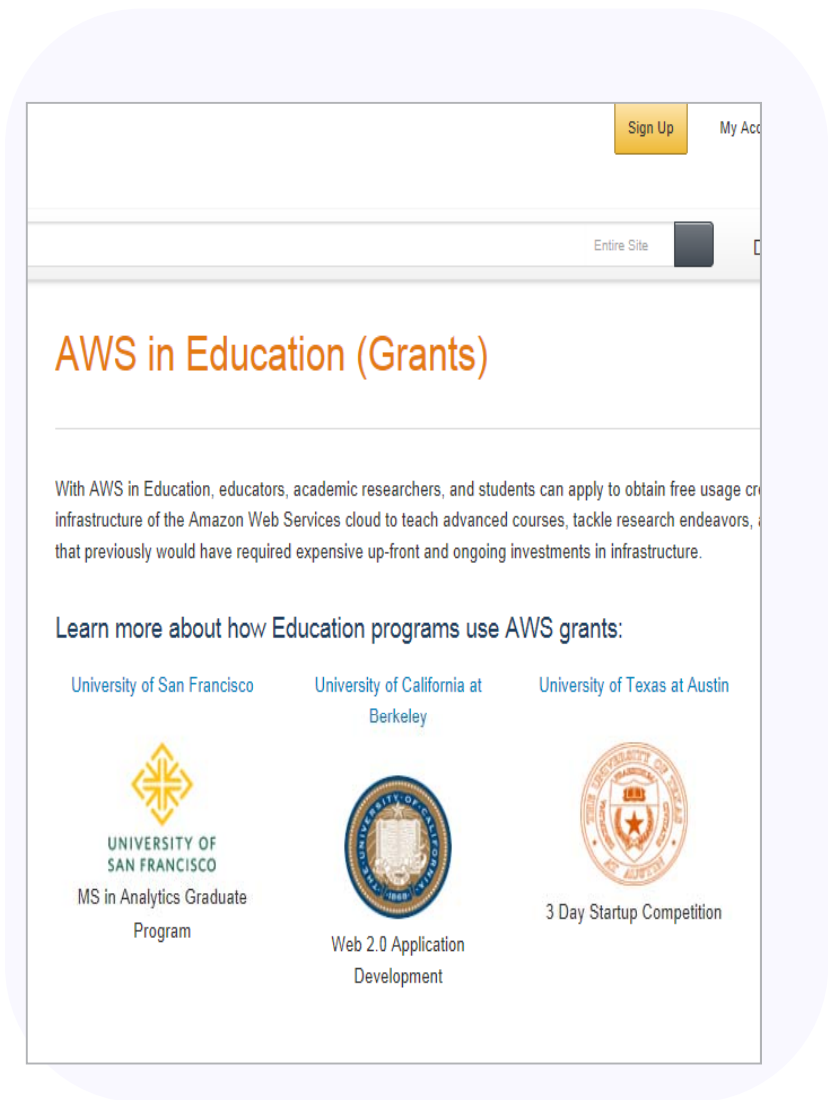
AWS Grants Program

AWS in Education

aws.amazon.com/grants



AWS Worldwide Public Sector





AWS Worldwide Public Sector

AWS Big Data Test Drives

APN Partner-provided labs

aws.amazon.com/testdrive/bigdata



AWS Training & Events

Webinars, Bootcamps,
and Self-Paced Labs

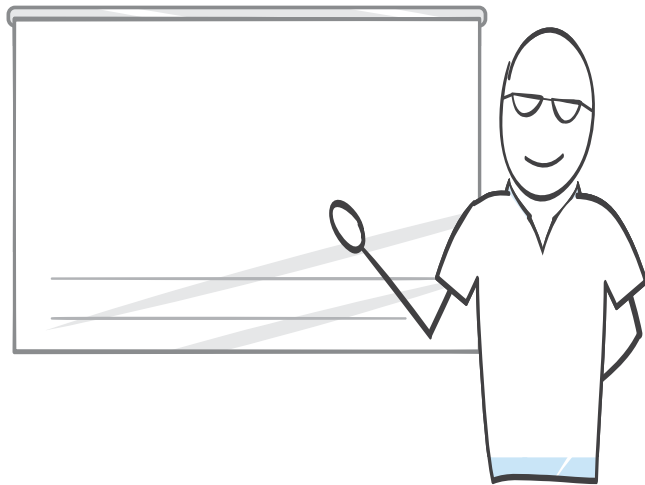
aws.amazon.com/events

<https://aws.amazon.com/training>



AWS Worldwide Public Sector





Big Data to AWS

Brand **new** course on Big Data

aws.amazon.com/training/course-descriptions/bigdata



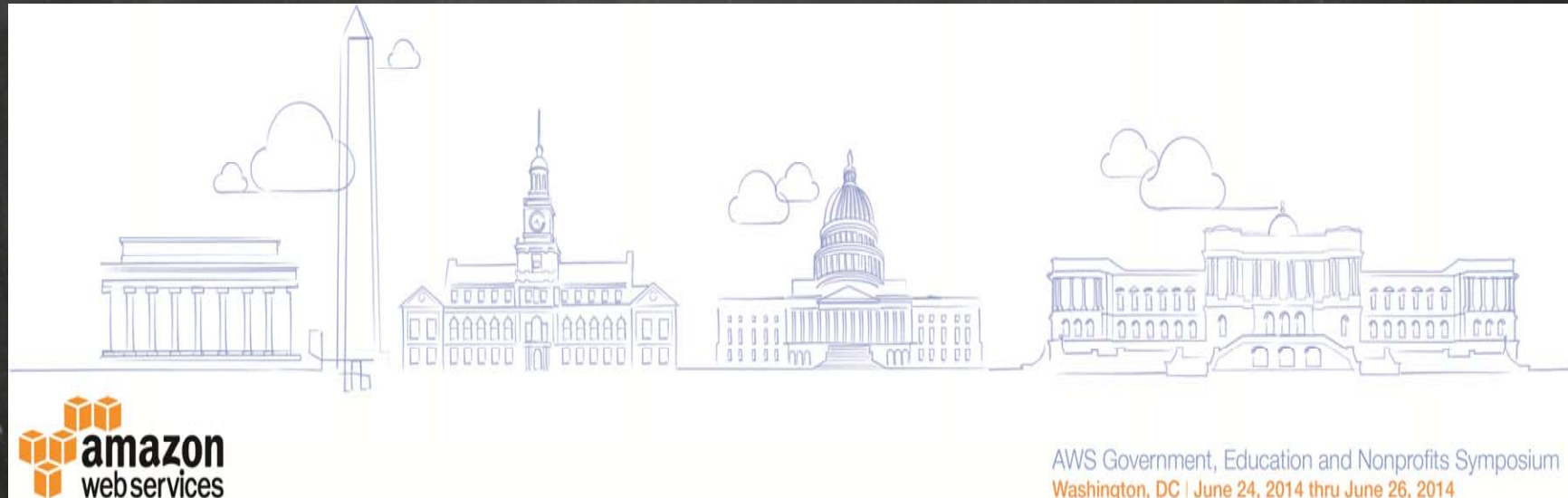
AWS Worldwide Public Sector



SAVE THE DATE:

AWS Government, Education, and Nonprofit Symposium

Washington, DC | June 24 - 26, 2014



<http://aws.amazon.com/events/wwps/symposia/washington/>





Thank you!

Tim Bixler

Manager, U.S. Federal Solutions Architecture

Principal Solutions Architect

tbixler@amazon.com