

# Deep Dive on Amazon Aurora

Jeremy Bendat Mike Gallagher



# Agenda



- Why AWS
- Why Corplnfo
- The Roadmap to Aurora
- Aurora Case Study



### Intro to AWS





WWW.CORPINFO.COM ©2016CORPINFO

# What sets AWS apart?





**Experience** 

Building and managing cloud since 2006



**Service Breadth & Depth** 

40+ services to support any cloud workload



**Pace of Innovation** 

History of rapid, customer-driven releases



**Global Footprint** 

12 regions, 30 availability zones, 53 edge locations



**Pricing Philosophy** 

51 proactive price reductions to date



**Ecosystem** 

8,000+ SIs and ISVs; 2,000+ Marketplace products

# **Experience with Operational Reliability**



We are driven to remove any causes of failure.

Our goal is to make our operational performance indistinguishable from perfect.

- ✓ We have spent over a decade building the world's most reliable, secure, scalable, and cost-effective infrastructure.
- ✓ Service SLAs between 99.9% and 100% availability. Amazon S3 maintains a durability of 99.99999999%.
- ✓ Availability Zones exist on isolated fault lines, flood plains, and electrical grids to substantially reduce the chance of simultaneous failure.
- ✓ The AWS Service Health Dashboard provides 24/7 visibility in the real-time operational status of all services around the globe.

# **Experience with Cloud Security**



AWS provides the same, familiar approaches to security that companies have been using for decades with increased visibility, control, and auditability.

#### Visibility

View your entire infrastructure with a click

Deep insight with AWS CloudTrail

#### Control

You have sole authority on where data is stored

Shared responsibility model

#### **Auditability**

3<sup>rd</sup> party validation

SOC 1 / SOC 2 / SOC 3 SSAE 16 / ISAE 3402 PCI DSS Level 1 DIACAP & FISMA ISO 27001 / 9001 / 13485 ISO/TS 16949 FedRAMP (SM) FISMA
HIPAA
ITAR
MPAA
CSA
FIPS 140-2

Based on our experience, I believe that we can be even more secure in the AWS cloud than in our own data centers.

# Service Breadth & Depth



Enterprise Applications



Virtual Desktops



Collaboration and Sharing

#### **Analytics**



Hadoop



Real-time Streaming Data



Data Warehouse



**Data Pipelines** 

#### **App Services**



Queuing & Notifications



Workflow



App streaming



Transcoding



Email



Search

#### **Deployment & Management**



One-click web app deployment



Dev/ops resource management



Resource Templates

#### **Mobile Services**



Identity



Sync



Mobile Analytics



**Push Notifications** 

Administration & Security

**Platform Services** 



Identity Management



Access Control



Usage Auditing



Key Storage



Monitoring and

Core Services



Compute (VMs, Auto-scaling and Load Balancing)



Storage (Object, Block and Archival)



CDN



Databases (Relational, NoSQL, Caching)



Networking (VPC, DX, DNS)

Infrastructure



Regions



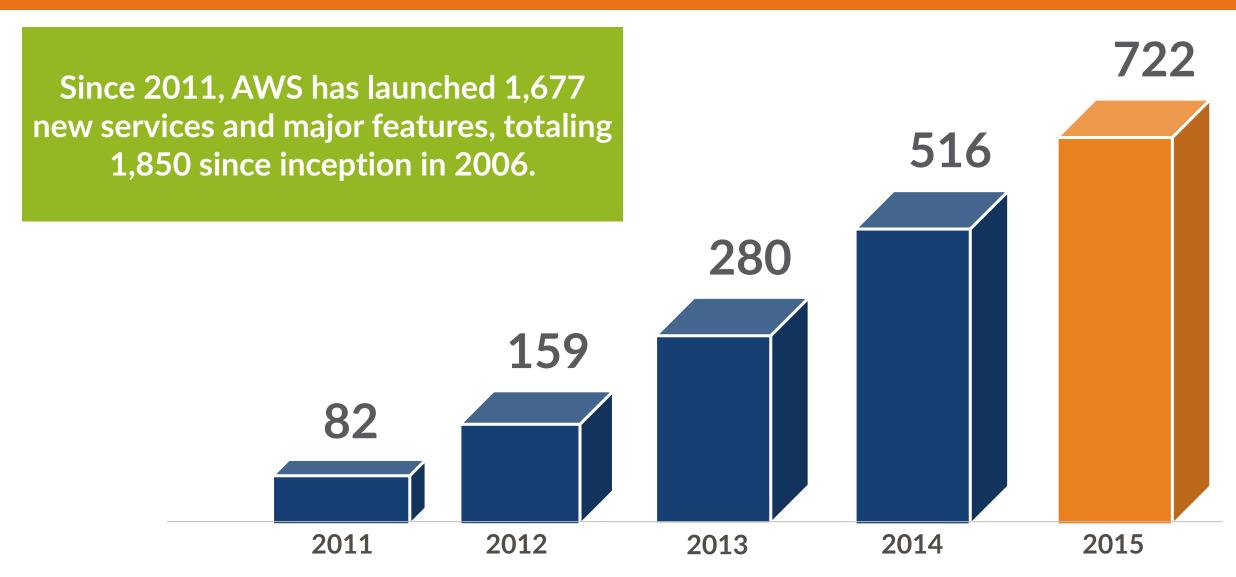
**Availability Zones** 



Points of Presence

# Rapid Pace of Innovation





#### **AWS Price Reductions**



AWS has announced price reductions

51X times since our inception in 2006. Recent price drops included...

34%

Amazon ElastiCache reduces prices for cache nodes by an average of 34%

March 26, 2014

51%

Amazon S3 reduces prices for Standard and Reduced Redundancy Storage, by an average of 51%

March 26, 2014

20%

Amazon Route 53 lowers prices for both standard queries and latency-based routing queries by 20%

July 31, 2014

\* As of January 2016

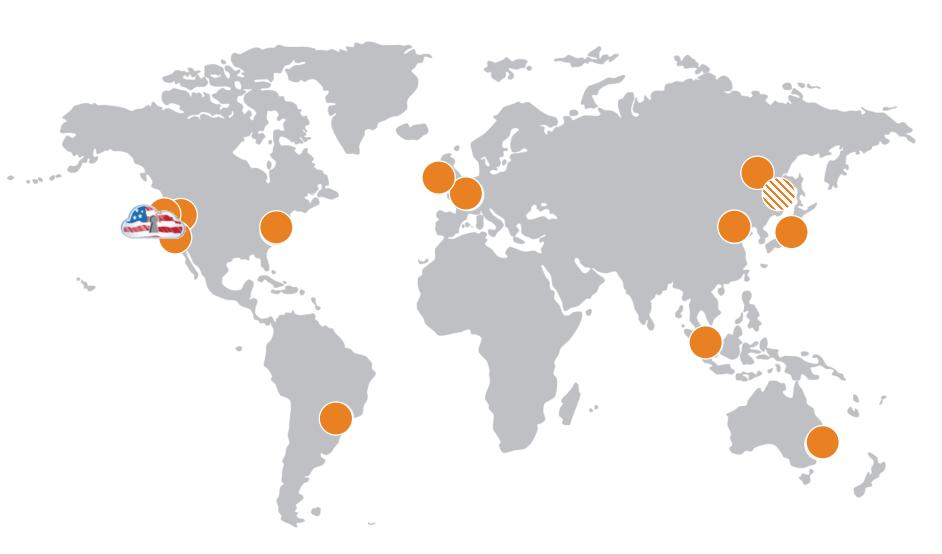
### **AWS Global Infrastructure**



12 Regions

30 Availability Zones

53
Edge
Locations



# **Pricing Philosophy**



#### High volume / low margin businesses are in our core DNA

Our economies of scale provide us with lower costs

Pricing model choice to support variable and stable workloads

Save more money as you grow bigger

51 price reductions since 2006

On-demand
Reserved
Spot

Tiered pricing
Volume discounts
Custom pricing

### **AWS** is Leader & Visionary



# Gartner Magic Quadrant for Cloud Infrastructure as a Service, Worldwide

#### Source: Gartner (May 2015)

Gartner "Magic Quadrant for Cloud Infrastructure as a Service, Worldwide," Lydia Leong, Douglas Toombs, Bob Gill, May 18, 2015. 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:

. 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 or other designation. 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.

Amazon Web Services Microsoft VMware IBM (SoftLayer)

Figure 1. Magic Quadrant for Cloud Infrastructure as a Service, Worldwide

COMPLETENESS OF VISION

Source: Gartner (May 2015)

©2016CORPINFO

As of May 2015



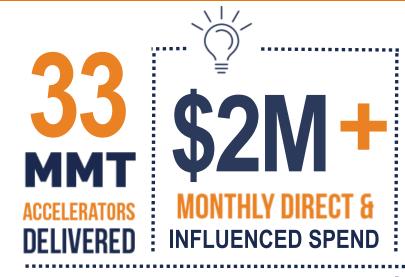
#### Partner Network

PREMIER CONSULTING PARTNER



#### **30+ Years of Customer Success**





REBORN IN THE CLOUD THROUGH PARTNERSHIP WITH AWS





AWS AUDITED MANAGED SERVICES

CORPINFO IS A GREAT EXAMPLE OF HOW AN ESTABLISHED TECHNOLOGY PROVIDER IS REIN-VENTING THEIR SERVICES BUSINESS WITH AWS

- TERRY WISE, AWS VP WW ALLIANCES

FOUNDED IN
1983

8 STILL UNDER FOUNDING
OWNERSHIP



100+ TEAM LOCATED ACROSS THE US WITH SOCAL AND US WEST FOCUS

# **AWS Premier Consulting Partner**

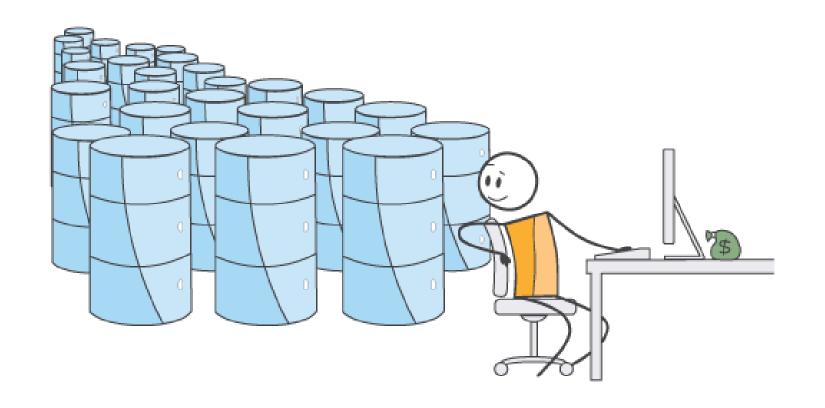




### PREMIER CONSULTING PARTNER

Corplnfo is one of the top APN Consulting partners globally, with extensive experience deploying customer solutions on AWS and a strong bench of trained and certified technical consultants.





# **Amazon** Aurora

WWW.CORPINFO.COM ©2016CORPINFO



# CorpInfo will review your entire AWS environment:

- Current Region
- Networking (VPC) & VPC Design
- VPN Gateways or Direct Connects
- Authentication & Authorization (IAM)
- Existing AWS infrastructure pain points
- Business application(s) connecting to the database





- Database monitoring performance metrics
- Internal support tickets
- Databases on the MySQL server
- Read replicas
- Master/slave replication relationships
- Issues converting MyISAM to Innodb
- Functionality for the MySQL database engine to confirm Aurora compatibility





Migration Roadmap

Target Platform

**My SQL Pain Points** 

Scale

**Stability** 

**Performance** 

Security

Management

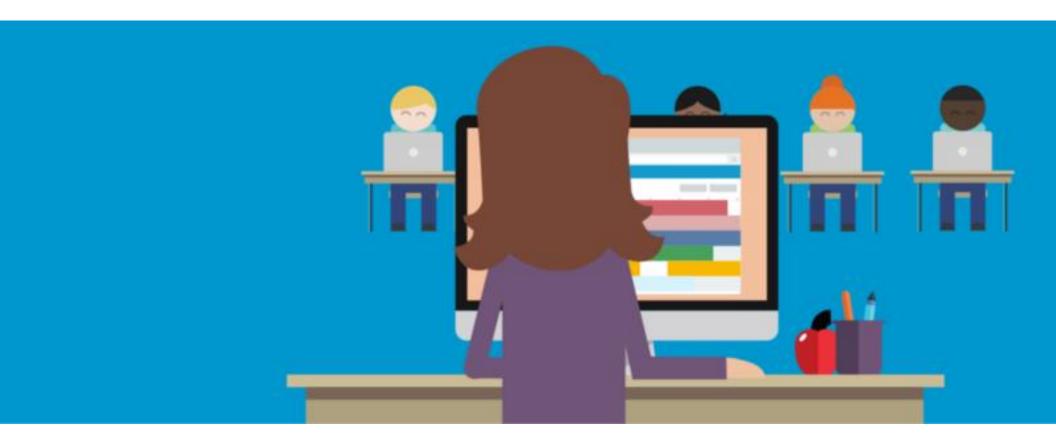
High Availability



# GoGuardian Case Study





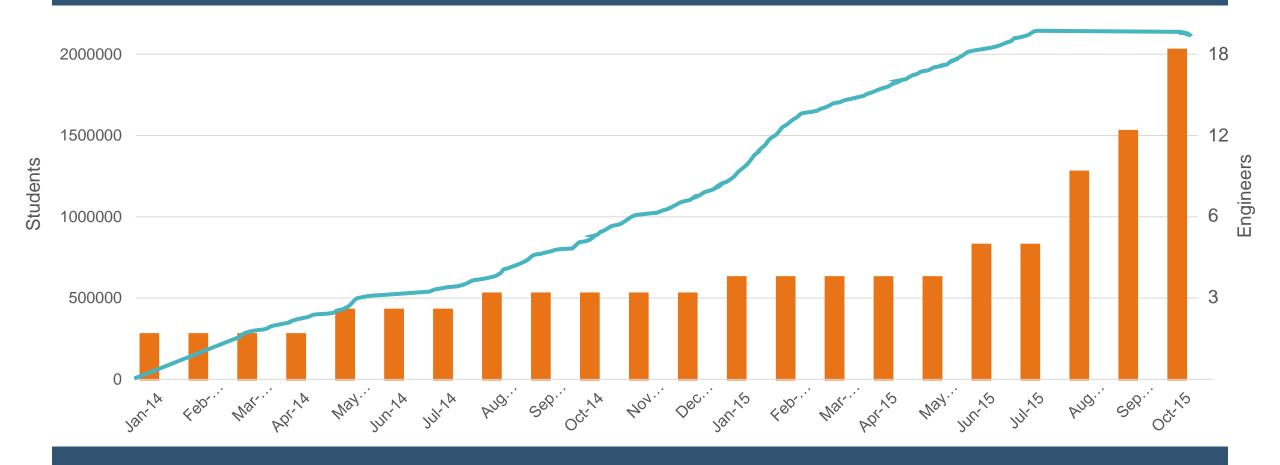


A smarter way to manage classroom technology. GoGuardian provides Chromebook management solutions that keep students safe online and make teaching easier

WWW.CORPINFO.COM ©2016CORPINFO





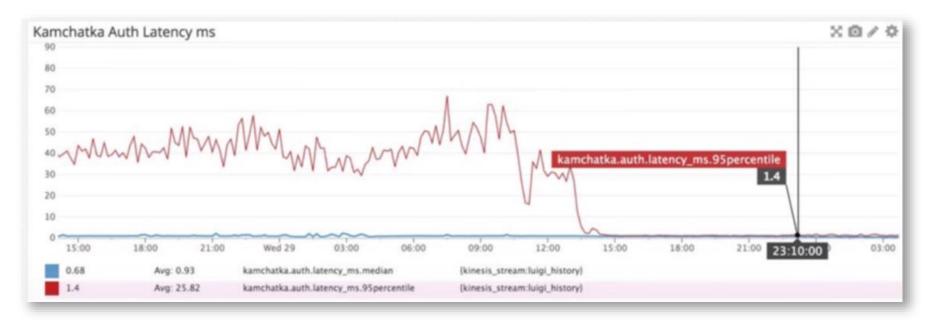


Crazy product growth, not so crazy team growth



#### The Aurora Performance Boost is Real:

- Kinesis combined with Aurora provides incredible throughput
- Easily add more sinks, pipe to more datastores & stream analytics
- Add more sources, e.g. demo data







scale without worrying about expanding server capacity

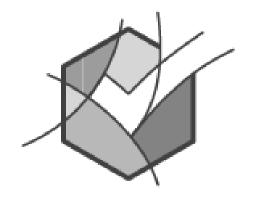
66

CorpInfo & AWS gives us the toolset for collecting, manipulating, and analyzing data to make it useful for our company and customers.

#### Additional AWS Resources











#### **Security**

aws.amazon.com / security

#### Compliance

aws.amazon.com / compliance

#### **White Papers**

aws.amazon.com/ whitepapers

#### **Architecture**

aws.amazon.com / architecture

# CORPINFO

Founded in 1983, Corplnfo is a leading technology firm providing Cloud Consulting Services, Infrastructure Solutions, and Managed Services. We use our experience to ensure that clients have the best technical solutions to solve their business challenges and deliver value for their organization. With a team of AWS certified solutions architects we support customers on the journey to the cloud and in unlocking the benefits AWS has to offer. We optimize the value of IT investments by thinking creatively to solve specific challenges while laying the groundwork for future growth and flexibility.

------

LOS ANGELES DALLAS IRVINE HOUSTON PHOENIX

WWW.CORPINFO.COM ©2016CORPINFO



#### **Amazon Aurora**

Puneet Agarwal, Solutions Architect Amazon Web Services



#### Relational databases were not designed for the cloud

SQL

Transactions

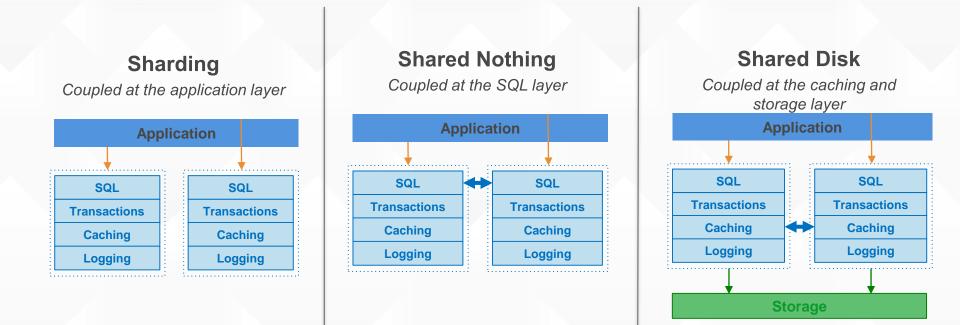
Caching

Logging

Multiple layers of functionality all in a monolithic stack



#### Not much has changed in last 20 years



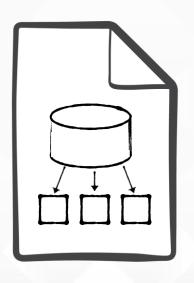
Even when you scale it out, you're still replicating the same stack

#### This is a problem.

For cost. For flexibility. And for availability.



#### Reimagining the relational database



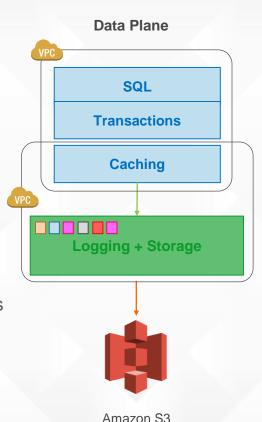
#### What if you were inventing the database today?

- You wouldn't design it the way we did in 1970.
- You'd build something
  - ✓ that can scale out ....
  - ✓ that is self-healing ....
  - ✓ that leverages existing AWS services ...



#### A service-oriented architecture applied to the database

- Moved the logging and storage layer into a multi-tenant, scale-out databaseoptimized storage service
- Integrated with other AWS services like
   Amazon EC2, Amazon VPC, Amazon
   DynamoDB, Amazon SWF, and Amazon
   Route 53 for control plane operations
- Integrated with Amazon S3 for continuous backup with 99.99999999 durability



#### **Control Plane**



Amazon DynamoDB



Amazon SWF



Amazon Route 53



#### Meet Amazon Aurora .....

#### Databases reimagined for the cloud



- Speed and availability of high-end commercial databases
- ☑ Simplicity and cost-effectiveness of open source databases
- ☑ Drop-in **compatibility** with MySQL
- ☑ Simple pay as you go pricing

Delivered as a managed service



#### **Customers are using Amazon Aurora**



#### Common customer use cases



**Fastest growing service in AWS history** 

Web and mobile

Content management

E-commerce, retail

Internet of Things

Search, advertising

BI and analytics

Games, media



#### **Expedia: On-line travel marketplace**



World's leading online travel company, with a portfolio that includes 150+ travel sites in 70 countries.

- Real-time business intelligence and analytics on a growing corpus of on-line travel market place data.
- Current SQL server based architecture is too expensive. Performance degrades as data volume grows.
- Cassandra with Solr index requires large memory footprint and hundreds of nodes, adding cost.

- Aurora meets scale and performance requirements with much lower cost.
- 25,000 inserts/sec with peak up to 70,000. 30ms average response time for write and 17ms for read, with 1 month of data.



## **PG&E:** Large public utility



One of the largest combination natural gas and electric utilities in the United States with approximately 16 million customers in 70,000-square-mile service area in northern and central California.

- Servicing high traffic surge during power events had always been a problem.
- Availability is critical when databases are down, it adversely affects service to gas and electrical customers.

- Being able to create multiple database replicas with millisecond latency allows them handle large surges in traffic and still give customers timely, upto-date information during a power event.
- Amazon Aurora, with 6-way replication, self healing storage and automatic instance repair, provides the availability and reliability needed for mission critical applications.



## ISCS: Insurance claims processing



Provides policy management, claim, billing solutions for casualty and property and insurance organizations

- Have been using Oracle and SQL server for operational and warehouse data
- Cost and maintenance of traditional commercial database has become the biggest expenditure and maintenance headache.

- The cost of a "more capable" deployment on Aurora has proven to be about 70% less than ISCS's SQL Server deployments.
- Eliminated backup window with Aurora's continuous backup; exploiting linear scaling with number of connections; continuous upload to Redshift using Aurora read replicas.



#### Alfresco: Enterprise content management



Leading the convergence of Enterprise
Content Management and Business
Process Management. More than 1,800
organizations in 195 countries rely on
Alfresco, including leaders in financial
services, healthcare, and the public sector.

- Scaling Alfresco document repositories to billions of documents
- Support user applications that require subsecond response times

- Scaled to 1 billion documents with a throughput of 3 million per hour, which is 10 times faster than their current environment.
- Moving from large data centers to cost-effective management with AWS and Aurora.



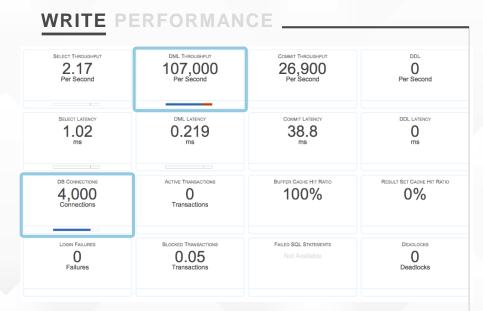
#### **Amazon Aurora is fast**

"When we ran Alfresco's workload on Aurora, we were blown away to find that Aurora was 10X faster than our MySQL environment" said John Newton, Founder and CTO of Alfresco. "Speed matters in our business and Aurora has been faster, cheaper, and considerably easier to use than MySQL"



#### SQL benchmark results

- MySQL Sysbench
- R3.8XL with 32 cores and 244 GB RAM



· 4 client machines with 1,000 threads each



Single client with 1,000 threads



#### Writes scale with table count

Tables	Amazon Aurora	MySQL I2.8XL Iocal SSD	MySQL I2.8XL RAM disk	RDS MySQL 30K IOPS (single AZ)
10	60,000	18,000	22,000	25,000
100	66,000	19,000	24,000	23,000
1,000	64,000	7,000	18,000	8,000
10,000	54,000	4,000	8,000	5,000

- Write-only workload
- 1,000 connections
- Query cache (default on for Amazon Aurora, off for MySQL)





#### Writes scale with connection count

Connections	Amazon Aurora	RDS MySQL 30K IOPS (single AZ)
50	40,000	10,000
500	71,000	21,000
5,000	110,000	13,000

**UP TO** 

8x

**FASTER** 

- OLTP Workload
- Variable connection count
- 250 tables
- Query cache (default on for Amazon Aurora, off for MySQL)



#### How Do we achieve these results?

#### DO LESS WORK \_\_\_\_\_

BE MORE EFFICIENT

Do fewer IOs

Minimize network packets

Cache prior results

Offload the database engine

Process asynchronously

Reduce latency path

Use lock-free data structures

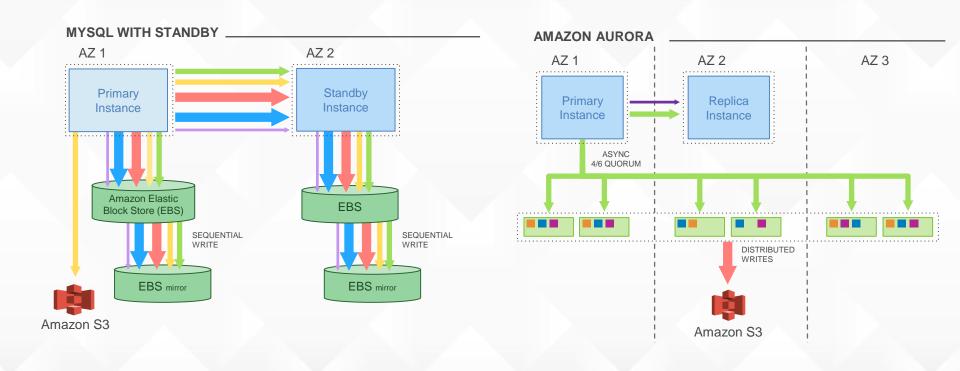
Batch operations together



## 10 traffic patterns: MySQL vs. Aurora

Binlog

Log records



Data



## 10 volume: MySQL vs. Aurora

#### 100GB Database / 1M Sysbench transactions

Workload	MySQL w/ 30K PIOS	Aurora	
Read Only	24,814	0	0.00%
Write Only	7,387,798	158,323	2.21%
OLTP	7,722,684	201,292	2.61%
R/W: 50/50	23,753,366	364,032	1.55%





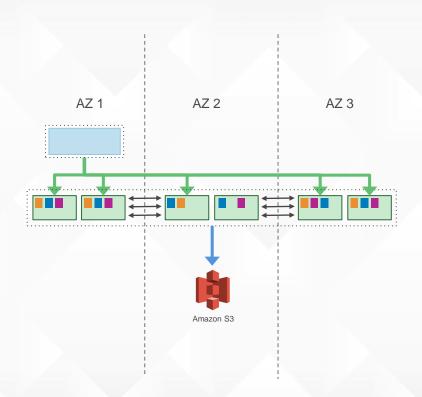
## Amazon Aurora is highly available

"Using Amazon Aurora, we can run many replicas with millisecond latency. This means during a power event we can handle large surges in traffic and still give our customers timely, up-to-date information. In addition, spreading these replicas across multiple AWS Availability Zones with automatic failover gives us confidence that our databases will be there when we need them." - Edward Wong, Solutions Architect at PG&E



## Highly available storage

- Six copies of data; quorum system for read/write; latency tolerant
- Background scrubbing; CRC on the wire & on disk
- Peer to peer gossip replication for catchup and recovery
- Continuous back to S3 as a quorum set member
- 10GB segments as unit of repair or hotspot rebalance

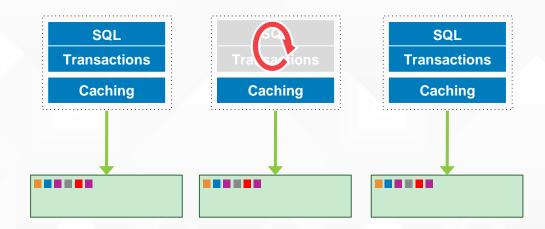




#### Survivable caches

- We moved the cache out of the database process
- Cache remains warm in the event of a database restart
- Lets you resume fully loaded operations much faster
- Instant crash recovery + survivable cache = quick and easy recovery from DB failures

Caching process is outside the DB process and remains warm across a database restart





#### **Instant crash recovery**

#### Traditional databases \_\_\_\_\_

- Have to replay logs since the last checkpoint
- Single-threaded in MySQL; requires a large number of disk accesses

Crash at T<sub>0</sub> requires a re-application of the SQL in the redo log since last checkpoint

**Checkpointed Data** 

Redo Log

#### **Amazon Aurora**

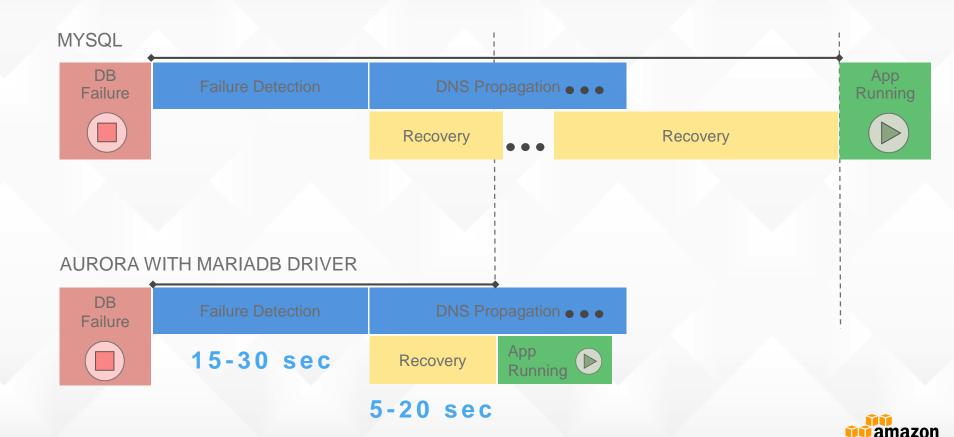
- Underlying storage replays redo records on demand as part of a disk read
- Parallel, distributed, asynchronous

Crash at T<sub>0</sub> will result in redo logs being applied to each segment on demand, in parallel, asynchronously





## Faster, more predictable failover



# Simulate failures using SQL

• To cause the failure of a component at the database node:

ALTER SYSTEM CRASH [{INSTANCE | DISPATCHER | NODE}]

To simulate the failure of disks:

ALTER SYSTEM SIMULATE percent\_failure DISK failure\_type IN

[DISK index | NODE index] FOR INTERVAL interval

To simulate the failure of networking:

ALTER SYSTEM SIMULATE percent\_failure NETWORK failure\_type [TO {ALL | read\_replica | availability\_zone}] FOR INTERVAL interval



### **Amazon Aurora is easy to use**

"Amazon Aurora's new user-friendly monitoring interface made it easy to diagnose and address issues. Its performance, reliability and monitoring really shows Amazon Aurora is an enterprise-grade AWS database." – Mohamad Reza, Information Systems Officer at United Nations



## Simplify storage management



Up to 64TB of storage – auto-incremented in 10GB units

- Continuous, incremental backups to Amazon S3
- Instantly create user snapshots—no performance impact
- Automatic storage scaling up to 64 TB—no performance impact
- Automatic restriping, mirror repair, hot spot management, encryption



## Simplify data security

- ☑ Encryption to secure data at rest
  - AES-256; hardware accelerated
  - All blocks on disk and in Amazon S3 are encrypted
  - Key management via AWS KMS
- SSL to secure data in transit
- ☑ Network isolation via Amazon VPC by default
- ☑ No direct access to nodes

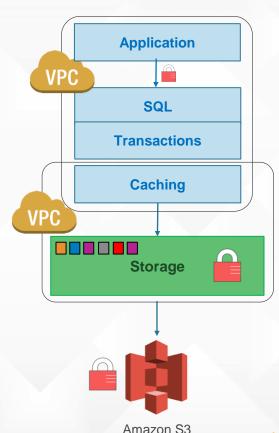






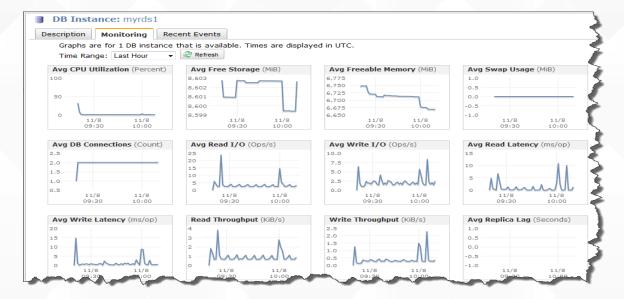








## Simplify monitoring with AWS console



#### **CloudWatch RDS Metrics**

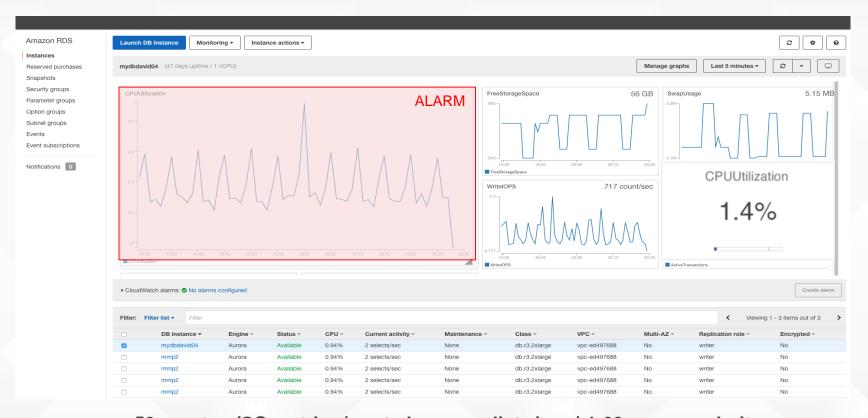
- CPU utilization
- Storage
- Memory
- Swap usage
- DB connections
- I/O (read and write)
- Latency (read and write)
- Throughput (read and write)
- Replica lag
- Many more

#### **CloudWatch Alarms**

Similar to on-premises custom monitoring tools



### **Advanced monitoring**



50+ system/OS metrics | sorted process list view | 1-60 sec granularity alarms on specific metrics | egress to CloudWatch Logs | integration with 3<sup>rd</sup>-party took webservices

# Delivered as a managed service



# If you host your databases on-premises

App optimization

Scaling

High availability

Database backups

DB s/w patches

DB s/w installs

OS patches

OS installation

Server maintenance

Rack and stack



# If you host your databases on-premises

App optimization

Scaling

High availability

Database backups

DB s/w patches

DB s/w installs

OS patches

OS installation

Server maintenance

Rack and stack



# If you host your databases in Amazon EC2

App optimization Scaling High availability Database backups DB s/w patches DB s/w installs OS patches

OS installation

Server maintenance

Rack and stack





## If you host your databases in Amazon EC2

App optimization Scaling High availability Database backups DB s/w patches DB s/w installs OS patches

OS installation

Server maintenance

Rack and stack





# If you choose a managed DB service

App optimization

Scaling

High availability

Database backups

DB s/w patches

DB s/w installs

OS patches

**OS** installation

Server maintenance

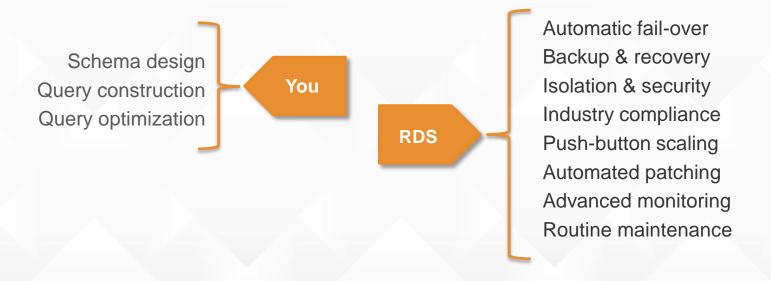
Rack and stack





#### Simplify database management

Amazon RDS takes care of your time-consuming database management tasks, freeing you to focus on your applications and business





# Amazon Aurora saves you money



#### Enterprise grade, open source pricing

	vCPU	Mem	Hourly Price
db.r3.large	2	15.25	\$0.29
db.r3.xlarge	4	30.5	\$0.58
db.r3.2xlarge	8	61	\$1.16
db.r3.4xlarge	16	122	\$2.32
db.r3.8xlarge	32	244	\$4.64

- Storage consumed, up to 64 TB, is \$0.10/GB-month
- IOs consumed are billed at \$0.20 per million I/O
- · Prices are for Virginia

#### Simple pricing

- No licenses
- No lock-in
- Pay only for what you use

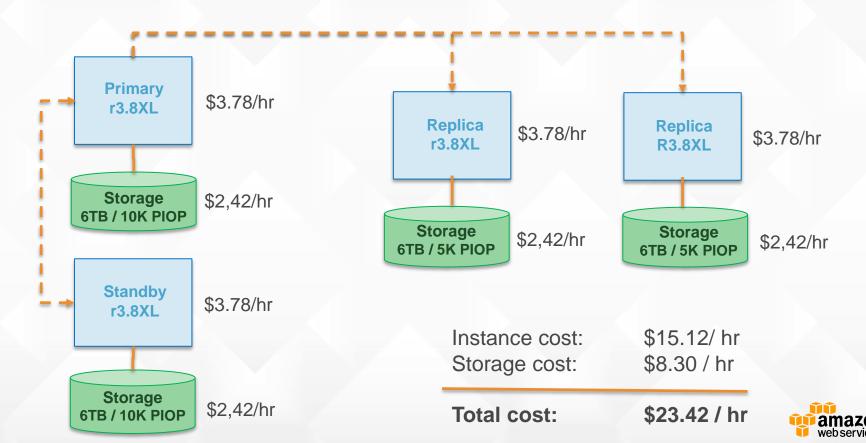
#### **Discounts**

- 44% with a 1-year RI
- 63% with a 3-year RI



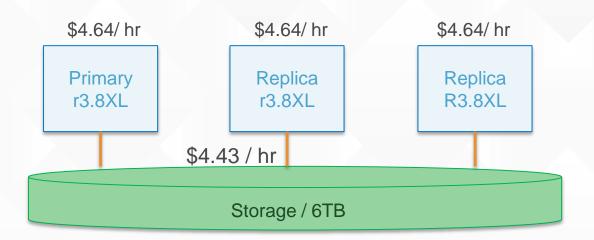
# Cost of ownership: Aurora vs. MySQL

MySQL configuration hourly cost



# Cost of ownership: Aurora vs. MySQL Aurora configuration hourly cost

- No idle standby instance
- Single shared storage volume
- No POIPs pay for use IO
- Reduction in overall IOP



Instance cost: \$13.92/ hr Storage cost: \$4.43 / hr

Total cost: \$18.35 / hr

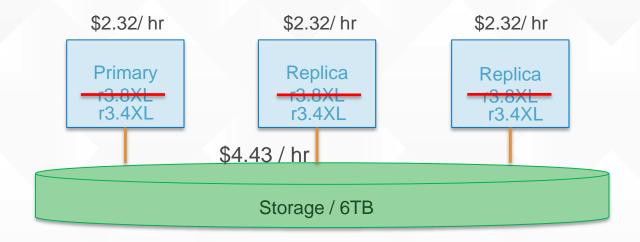
21.6% Savings



<sup>\*</sup>At a macro level Aurora saves over 50% in storage cost compared to RDS MySQL.

# Cost of ownership: Aurora vs. MySQL Further opportunity for saving

- Use smaller instance size
- Pay-as-you-go storage



Instance cost: Storage cost:

\$6.96/ hr \$4.43 / hr

Total cost:

\$11.39 / hr

51.3% Savings



- 1. Average IOPs is 50% of Max IOPs
- 2. 50% savings from shipping logs vs. full pages



# Migration to Aurora is easy





AWS
Database Migration
Service















Start your first migration in 10 minutes or less

Keep your apps running during the migration

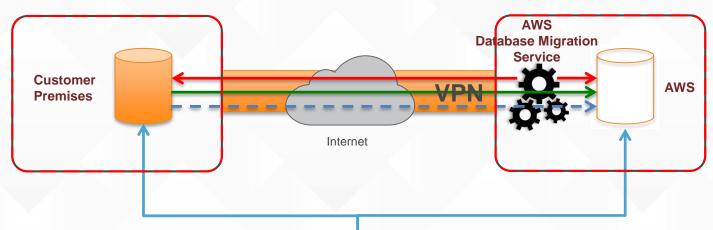
Replicate within, to or from Amazon EC2 or RDS

Move data to the same or different database

engine



# Keep your apps running during the migration



- Start a replication instance
- Connect to source and target databases
- Select tables, schemas, or databases

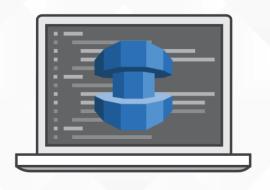


**Application Users** 

Let AWS Database Migration Service create tables, load data, and keep them in sync

Switch applications over to the target at your convenience





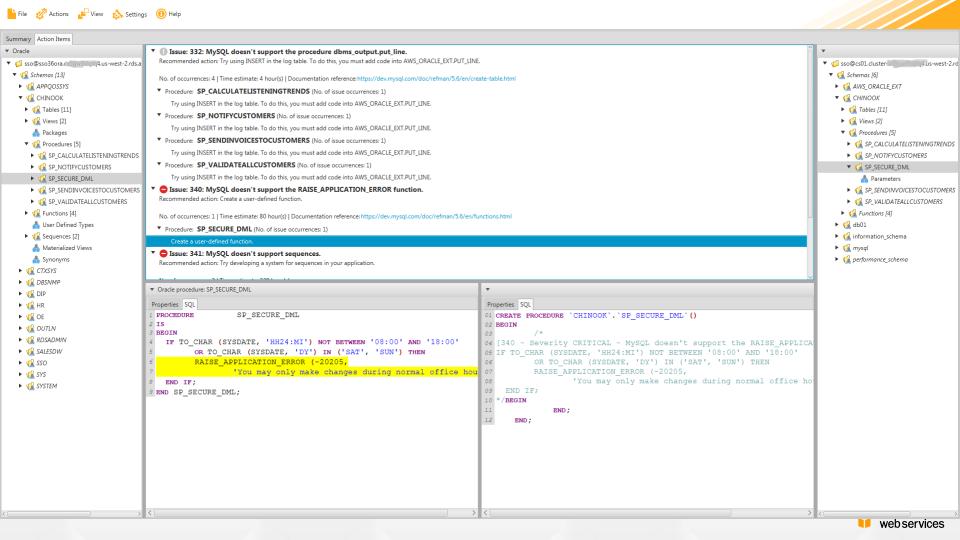
AWS Schema Conversion Tool Migrate off Oracle and SQL Server

Move your tables, views, stored procedures and DML to MySQL, MariaDB, and Amazon Aurora

Know exactly where manual edits are needed

Download at <u>aws.amazon.com/dms</u>





#### Well established eco-system



"We ran our compatibility test suites against Amazon Aurora and everything just worked." - Dan Jewett, Vice President of Product Management at Tableau















Q&A

