
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

DAT219

What's new with Amazon DocumentDB

Vin Yu

Senior Product Manager
AWS

Rashim Gupta

Senior Manager, Product Management
AWS

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Amazon DocumentDB overview

What’s new

Amazon DocumentDB Elastic Clusters

Demo

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon DocumentDB overview

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

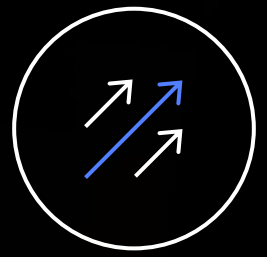
What is Amazon DocumentDB?

Amazon DocumentDB is a **scalable**, **highly durable**, and **fully managed** database service for operating mission-critical MongoDB API-based workloads

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon DocumentDB

SCALABLE, HIGHLY DURABLE, AND FULLY MANAGED DATABASE SERVICE FOR OPERATING MONGODB API-BASED WORKLOADS



Fast and scalable

- Auto scaling storage and I/O
- Storage scales to 64 TiB
- Scale compute in minutes
- Scale out to 15 replicas for millions of reads
- Globally distributed



Fully managed

- Built-in high availability
- Backups enabled by default
- Durable by default
- Security best practices by default
- Automatic patching
- Monitoring and alerting



MongoDB API compatible

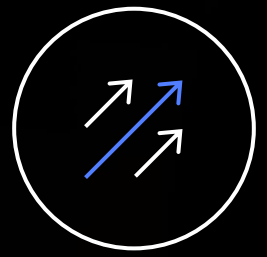
- Applications, drivers, and tools can be used with little or no change
- Supports hundreds of APIs, operators, and stages
- Continually working backward from customers to deliver the capabilities they need

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon DocumentDB

SCALABLE, HIGHLY DURABLE, AND FULLY MANAGED DATABASE SERVICE FOR OPERATING MONGODB API-BASED WORKLOADS



Fast and scalable

- Auto scaling storage and I/O
- Storage scales to 64 TiB
- Scale compute in minutes
- Scale out to 15 replicas for millions of reads
- Globally distributed



Fully managed

- Built-in high availability
- Backups enabled by default
- Durable by default
- Security best practices by default
- Automatic patching
- Monitoring and alerting



MongoDB API compatible

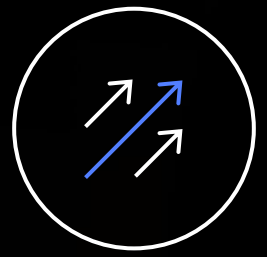
- Applications, drivers, and tools can be used with little or no change
- Supports hundreds of APIs, operators, and stages
- Continually working backward from customers to deliver the capabilities they need

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon DocumentDB

SCALABLE, HIGHLY DURABLE, AND FULLY MANAGED DATABASE SERVICE FOR OPERATING MONGODB API-BASED WORKLOADS



Fast and scalable

- Auto scaling storage and I/O
- Storage scales to 64 TiB
- Scale compute in minutes
- Scale out to 15 replicas for millions of reads
- Globally distributed



Fully managed

- Built-in high availability
- Backups enabled by default
- Durable by default
- Security best practices by default
- Automatic patching
- Monitoring and alerting



MongoDB API compatible

- Applications, drivers, and tools can be used with little or no change
- Supports hundreds of APIs, operators, and stages
- Continually working backward from customers to deliver the capabilities they need

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

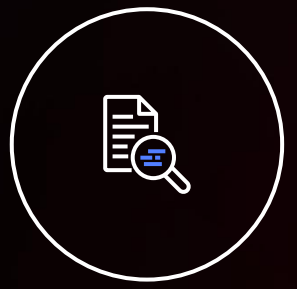
When should you use a document database?



JSON data



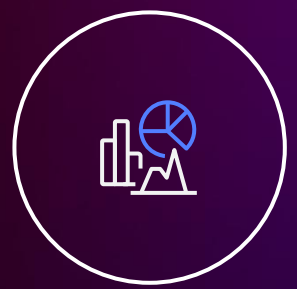
Flexible schema for fast iteration



Ad hoc query capabilities



Flexible indexing



Operational and analytics workloads

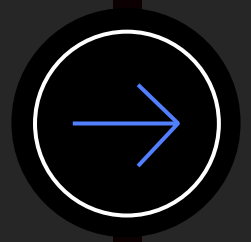
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

When should you use a document database?



Document use case: gaming user profiles

```
{
  userid: 181276,
  username: "Anna1942",
  name: {first: "Anna",
         last: "Kent"}
}
```



```
{
  userid: 181276,
  username: "Anna1942",
  name: {first: "Anna",
         last: "Kent"},
  ExplodingSnails: {
    hi_score: 3185400,
    global_rank: 5139,
    bonus_levels: true
  },
  promotions: ["new user","5%","snail lover"]
}
```

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Use cases



Web publishing



Product catalog



Content management



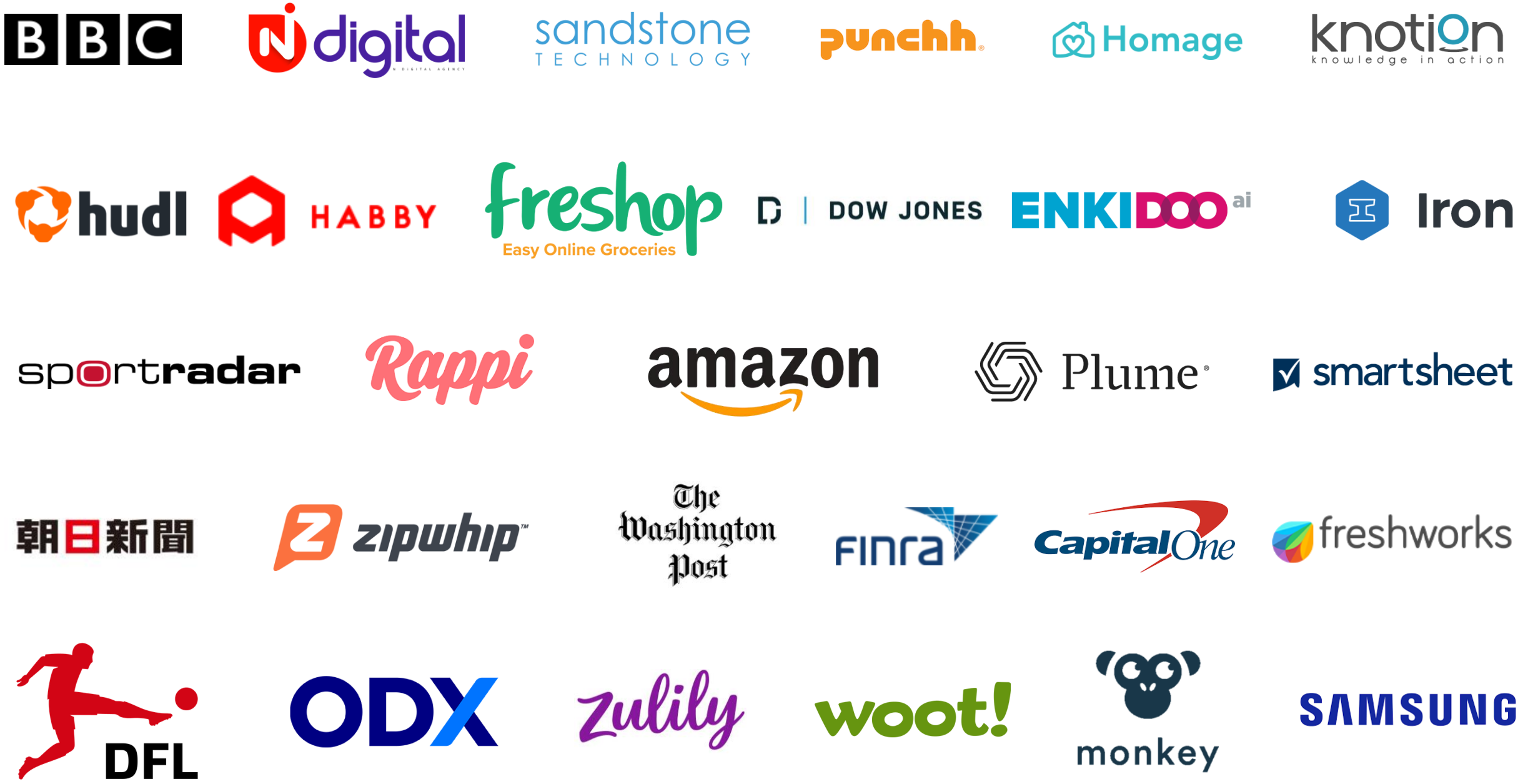
Regulatory documents



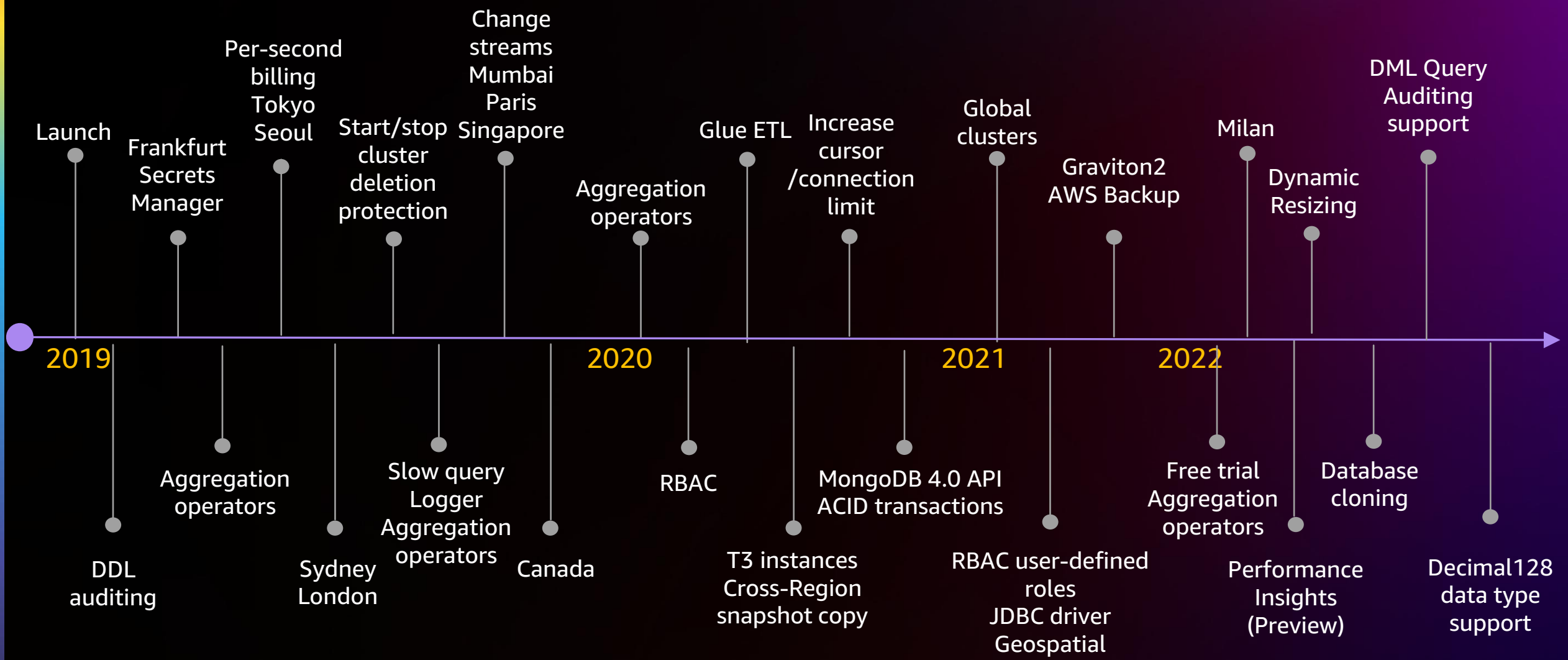
Gaming

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.


BBC
digital
sandstone TECHNOLOGY
punchh
Homage
knotion knowledge in action
hudl
HABBY
freshop Easy Online Groceries
DOW JONES
ENKIDOO ai
Iron
sportradar
Rappi
amazon
Plume
smartsheet
朝日新聞
zipwhip
The Washington Post
FINRA
CapitalOne
freshworks
DFL
ODX
Zulily
woot!
monkey
SAMSUNG

Working backwards from customers



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

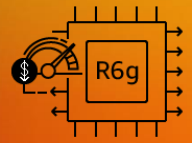
What’s new in 2022

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

New Regions and instance availability



Graviton2 instances (t4g, r6g) availability



Asia Pacific (Seoul)

South America (São Paulo)

Europe (Paris)

Region launches

Milan, Italy

Beijing, China

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Query capabilities

Geospatial indexes

$nearSphere, $geoNear, $minDistance, $maxDistance, $geoWithin, $geoIntersects

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Query capabilities

Geospatial indexes

$nearSphere, $geoNear, $minDistance, $maxDistance, $geoWithin, $geoIntersects

Aggregation operators

$mergeObjects and $reduce

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Query capabilities

Geospatial indexes

$nearSphere, $geoNear, $minDistance, $maxDistance, $geoWithin, $geoIntersects

Aggregation operators

$mergeObjects and $reduce

Decimal128

Provides 128 bits of decimal representation supporting 34 decimal digits of precision

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Enterprise management features

Dynamic volume resizing

Storage space automatically increases up to a maximum size of 64 tebibytes (TiB) when you need more storage and decreases automatically when you delete data.

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Enterprise management features

Dynamic volume resizing

Storage space automatically increases up to a maximum size of 64 tebibytes (TiB) when you need more storage and decreases automatically when you delete data.

Fast database cloning

Create clones to enable fast creation of a new cluster that has the same data as the original. Database cloning is faster than restoring a snapshot and requires no additional space at the time of creation.

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Enterprise management features

Dynamic volume resizing

Storage space automatically increases up to a maximum size of 64 tebibytes (TiB) when you need more storage and decreases automatically when you delete data.

Fast database cloning

Create clones to enable fast creation of a new cluster that has the same data as the original. Database cloning is faster than restoring a snapshot and requires no additional space at the time of creation.

Query auditing (now with DML)

DocumentDB will record Data Manipulation Language (DML) events to Amazon CloudWatch Logs for monitoring and analysis. This includes events corresponding to insert(), update(), delete(), find(), aggregates, and more.

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Enterprise management features

Dynamic volume resizing

Storage space automatically increases up to a maximum size of 64 tebibytes (TiB) when you need more storage and decreases automatically when you delete data.

Fast database cloning

Create clones to enable fast creation of a new cluster that has the same data as the original. Database cloning is faster than restoring a snapshot and requires no additional space at the time of creation.

Query auditing (now with DML)

DocumentDB will record Data Manipulation Language (DML) events to Amazon CloudWatch Logs for monitoring and analysis. This includes events corresponding to insert(), update(), delete(), find(), aggregates, and more.

Performance Insights (preview)

Performance tuning and monitoring feature that helps you quickly assess the load on your database and determine when and where to take action.

With one click, you can add a fully managed performance monitoring solution to your Amazon DocumentDB instances.

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Enterprise management features

Performance Insights (preview)

Performance tuning and monitoring feature that helps you quickly assess the load on your database and determine when and where to take action.

With one click, you can add a fully managed performance monitoring solution to your Amazon DocumentDB instances

- When was there unexpected load on my database?
- Which queries had the most load impact on my databases?
- Which applications or clients are causing load on my instance?
- Which queries are consuming the most CPU on my instance?
- What is consuming high I/O on my instance?

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

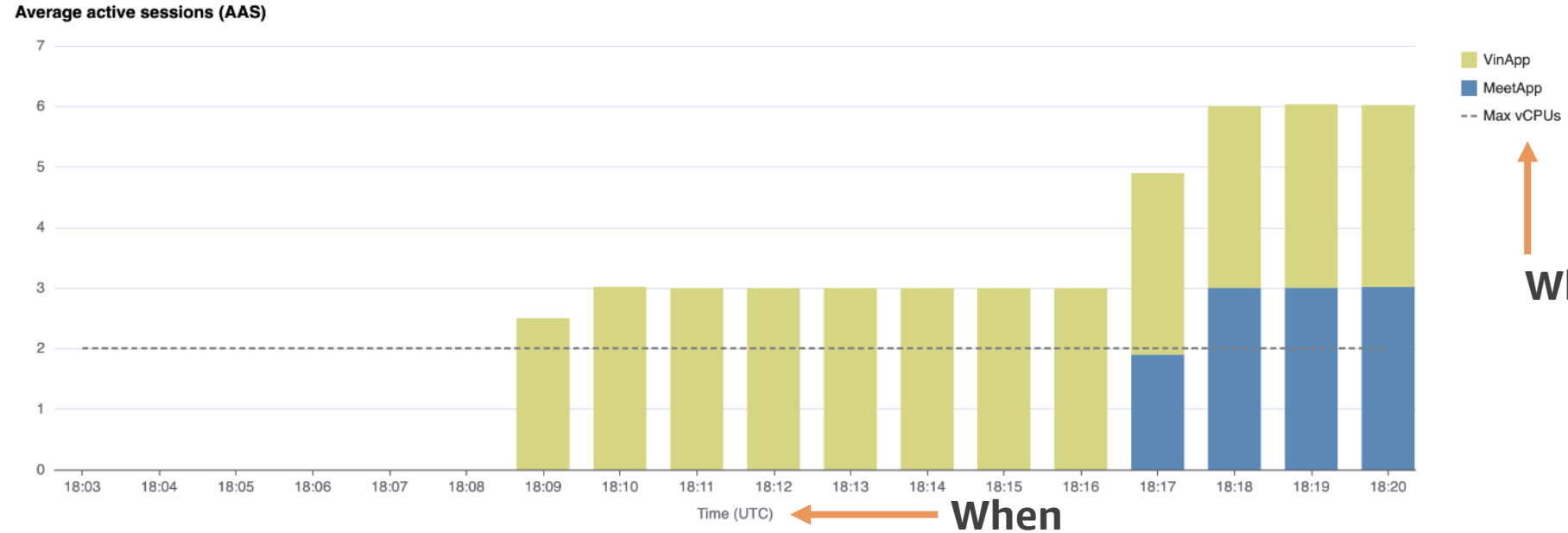
Dashboard
Clusters
Performance Insights
Snapshots
Subnet groups
Parameter groups
Event Subscriptions
Events
What's New 38
Tutorials

Database load

Current activity measured in average active sessions (AAS)

Bar | Slice by application

Show max vCPU



Top waits | Top queries | Top hosts | Top databases | Top applications

Top queries (2)

Top Queries by Application

Find query statements

Load by application (AAS)	Query statements	Support ID
2.58	{"find":"customers","filter":{"FirstName":"?"},"sort":{"key":{"$numberInt":"?"}}...	8E777F3719C7B1BE802117F56C6FECF5DA89C56C
0.66	{"find":"customers","filter":{"FirstName":"Tim"},"sort":{"key":{"$numberInt":"1"...	F9BE014B441F2D8944DA692E01DD8EB68834F3E4
0.66	{"find":"customers","filter":{"FirstName":"Tim"},"sort":{"key":{"$numberInt":"1"...	E82CFA68F5F76C3DF5201D7CA70988FE2E75A11B
0.66	{"find":"customers","filter":{"FirstName":"Tim"},"sort":{"key":{"$numberInt":"1"...	3E775FAAA5FB66AA4471FAC6BC6EFA3BA5CB0A2B

Introducing…
Amazon DocumentDB Elastic Clusters!

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Elastic Cluster

Why Elastic Clusters?

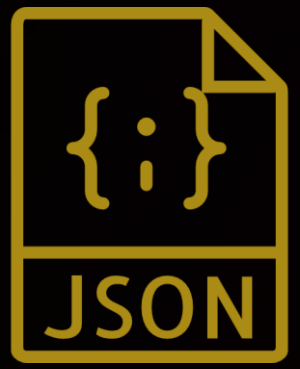
Elastic Cluster capabilities

How it works

Best practices

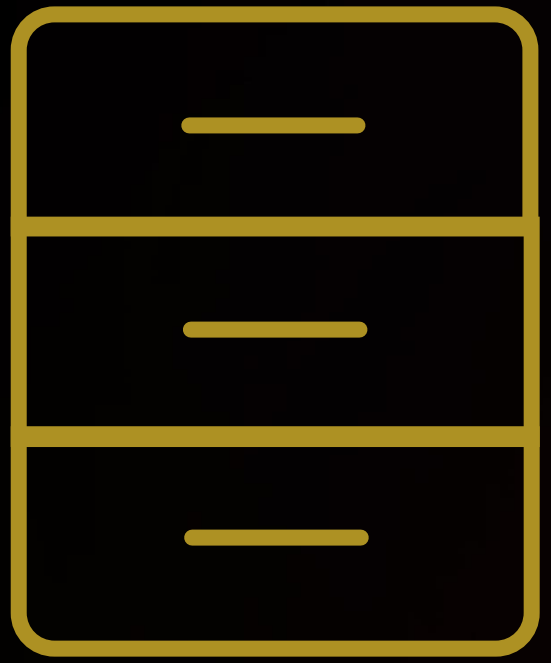
Demo

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



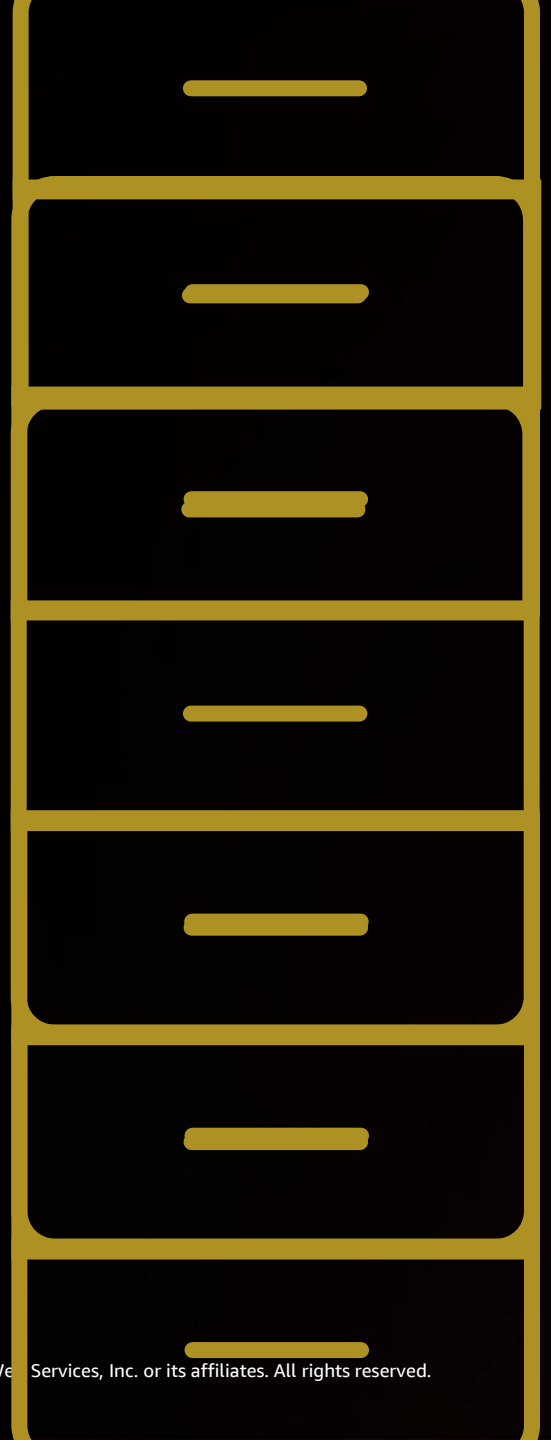


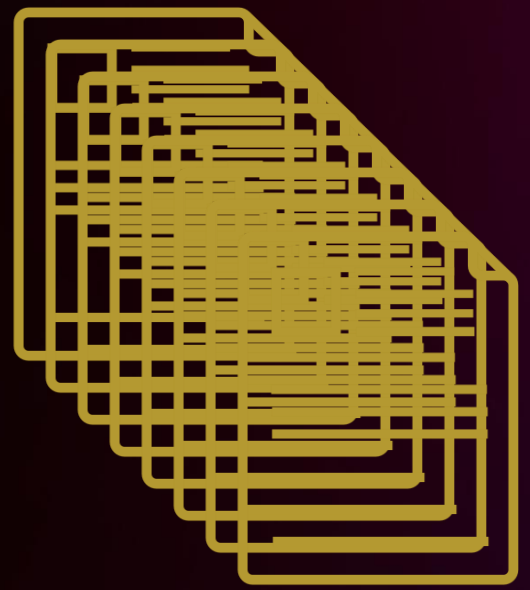
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



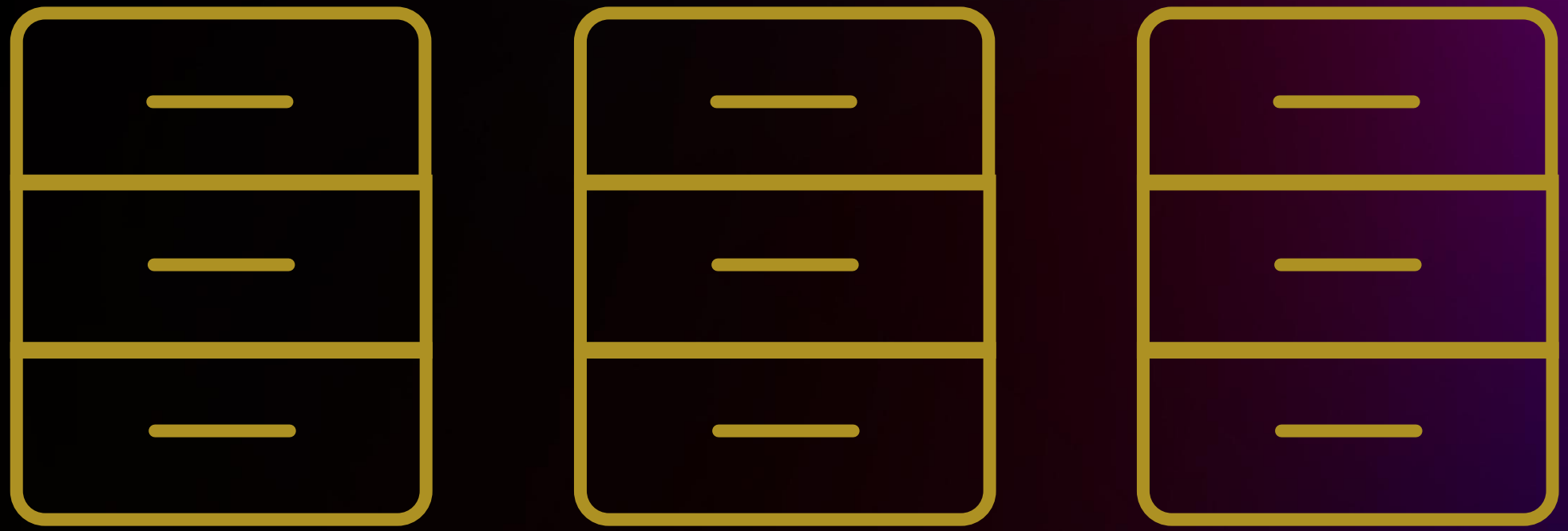


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.





© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon DocumentDB Elastic Clusters

NEW!



New type of cluster (Elastic Cluster) for Amazon DocumentDB that supports workloads with millions of reads/writes per second and petabytes of storage capacity

Elastic Clusters also simplify how developers interact with Amazon DocumentDB; no need to choose, manage, or upgrade instances!



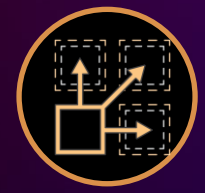
Compatible with MongoDB APIs for sharding



Millions of reads and writes



Up to 300,000 connections



Easy and fast scaling

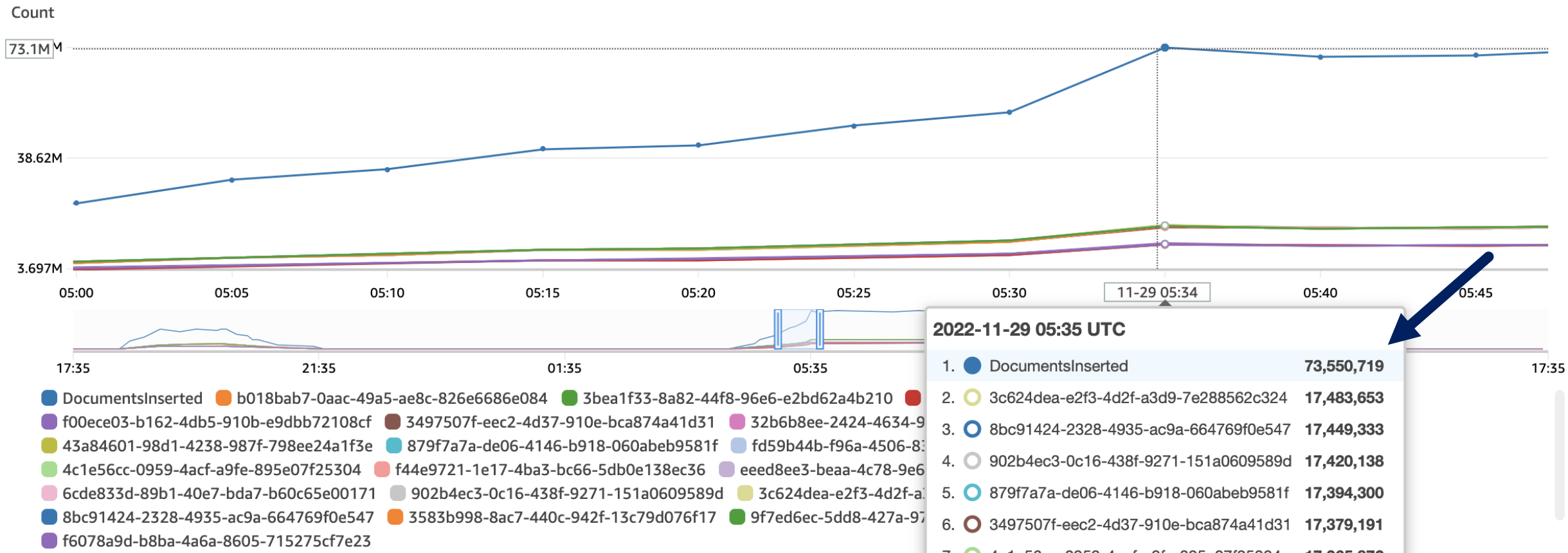
aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Elastic-Cluster-Stretch-5

1h 3h 12h 1d 3d 1w Custom 1m Actions

DocumentsInserted



2022-11-29 05:35 UTC

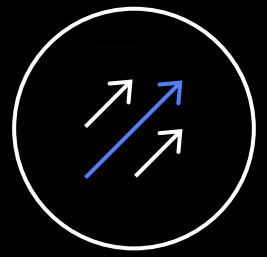
1.	DocumentsInserted	73,550,719
2.	3c624dea-e2f3-4d2f-a3d9-7e288562c324	17,483,653
3.	8bc91424-2328-4935-ac9a-664769f0e547	17,449,333
4.	902b4ec3-0c16-438f-9271-151a0609589d	17,420,138
5.	879f7a7a-de06-4146-b918-060abeb9581f	17,394,300
6.	3497507f-eec2-4d37-910e-bca874a41d31	17,379,191
7.	4c1e56cc-0959-4acf-a9fe-895e07f25304	17,365,872
8.	9f7ed6ec-5dd8-427a-9780-d161723b6b8d	17,348,492
9.	38fdca80-fe87-456d-82bd-78851fceafbd	17,323,626
10.	3ca88e3b-16f1-44f0-b719-5337bbbf322c	17,275,018

ResourceCount

No unit

Amazon DocumentDB Elastic Clusters

FAST, SCALABLE, AND FULLY MANAGED MONGODB API-COMPATIBLE DATABASE SERVICE THAT PROVIDES VIRTUALLY LIMITLESS SCALE



Fast and scalable

- Scales easily to millions of reads/writes per second and PBs of storage capacity
- Scaling operations complete in as little as minutes
- I/O auto scaling
- Most queries complete in milliseconds



Fully managed

- Built-in high availability
- Snapshot backups
- Durable by default
- Security best practices by default
- Automatic patching
- Monitoring and alerting



MongoDB API compatible

- Applications, drivers, and tools can be used with little or no change
- Supports hundreds of APIs, operators, and stages
- Continually working backward from customers to deliver the capabilities they need

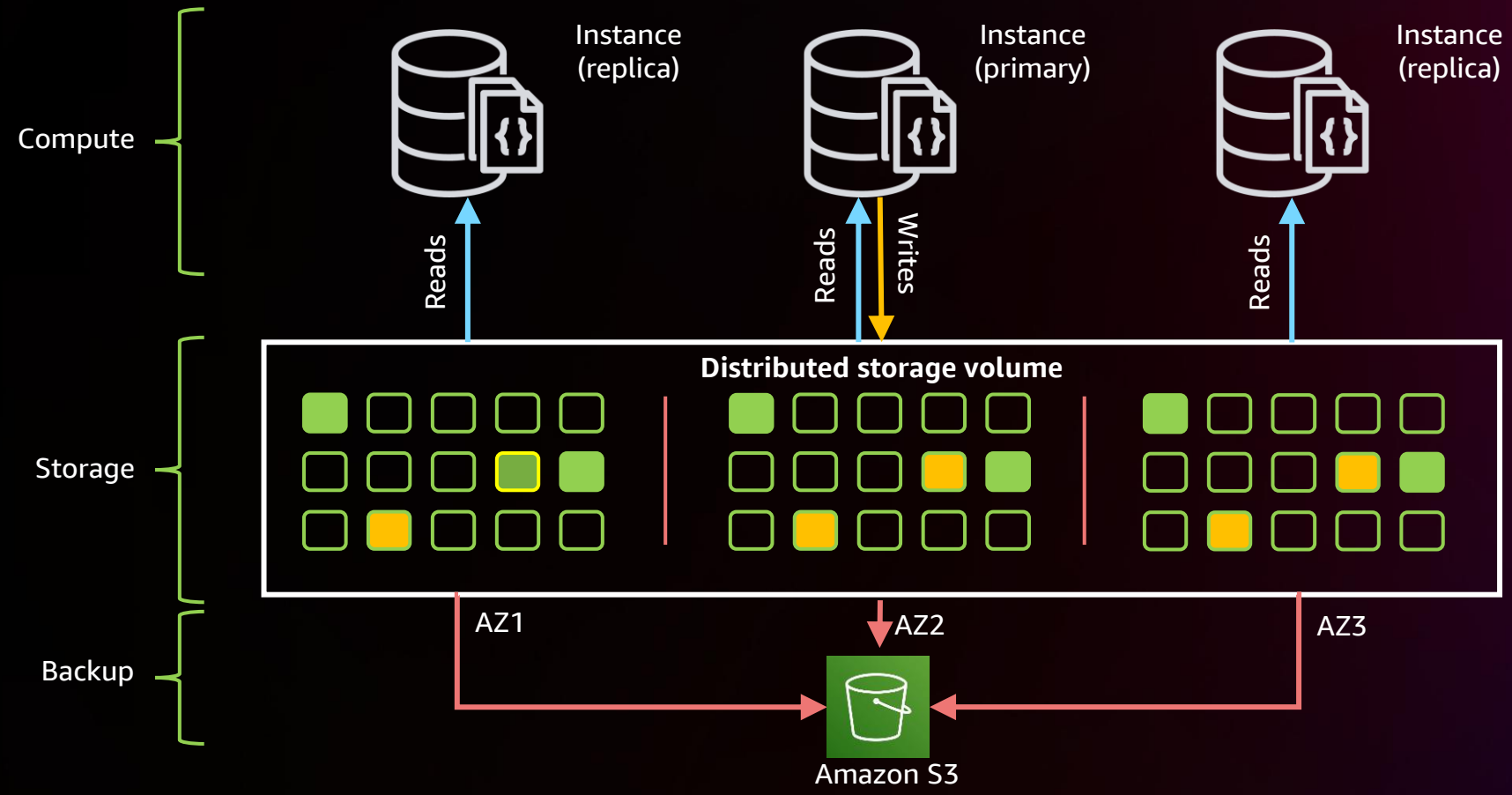
aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How it works

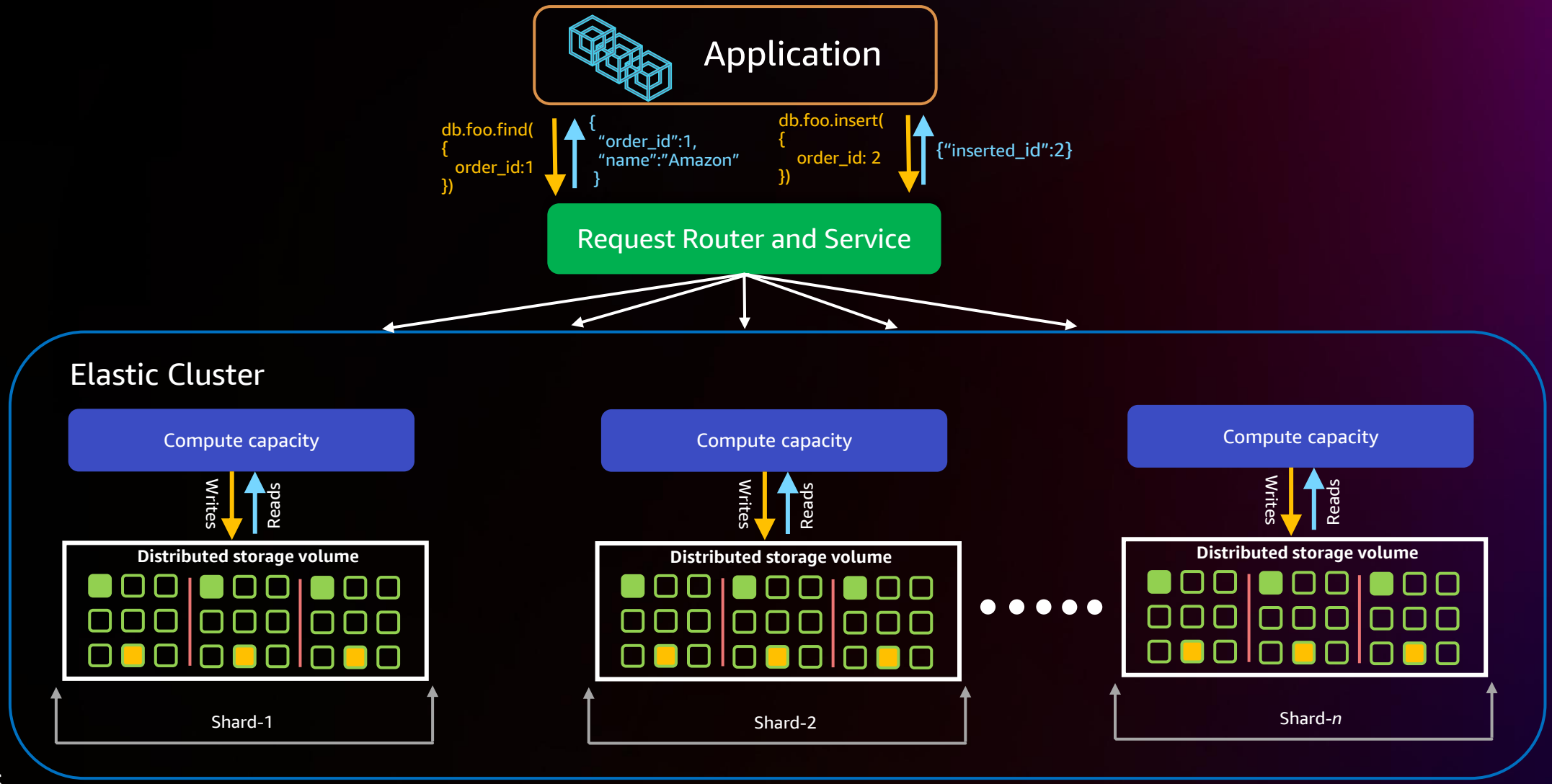
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon DocumentDB cloud-native architecture



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Elastic Cluster architecture



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Example – Employee dataset

{"name": "Vin", "lastname": "Yu", "role": "AWS Product Manager", "user_id": 1, "phone": "2223333"}

{"name": "Meet", "lastname": "Bhagdev", "role": "AWS Product Manager", "user_id": 2, "phone": "3334444"}

{"name": "Rashim", "lastname": "Gupta", "role": "AWS Product Manager","user_id": 3, "phone": "4445555"}

{"name": "Karthik", "lastname": "Vijayraghavan", "role": "AWS Solutions Architect","user_id": 4, "phone": "6667777"}

{"name": "Srini", "lastname": "Margasahayam", "role": "AWS Solutions Architect", "user_id": 5, "phone": "7778888"}

{"name": "Andrew", "lastname": "Whitaker", "role": "AWS Software Engineer", "user_id": 6, "phone": "8889999"}

{"name": "Sean", "lastname": "Oczkowski", "role": "AWS Software Engineer", "user_id": 7, "phone": "9991111"}

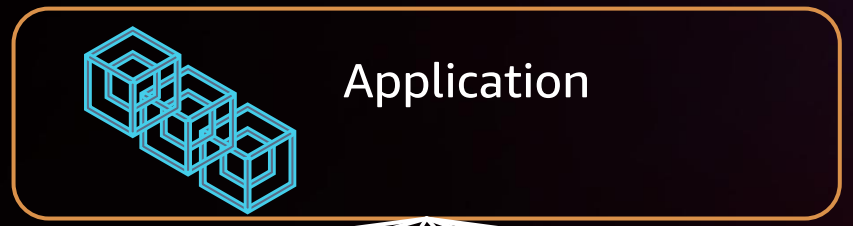
{"name": "Kari", "lastname": "Lai", "role": "AWS Marketing Manager", "user_id": 8, "phone": "1112222"}

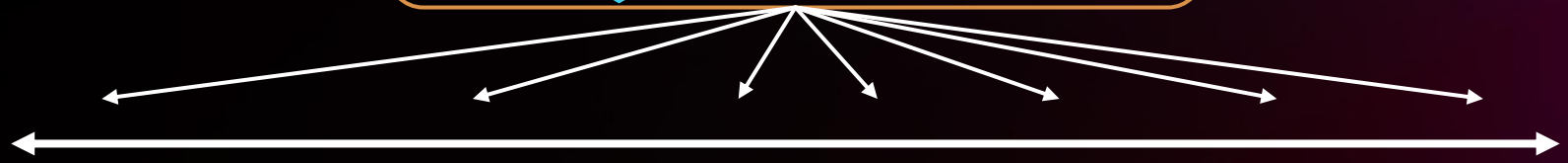
{"name": "Joy", "lastname": "Wang", "role": "AWS GTM Specialist", "user_id": 9, "phone": "1112222"}

Collection – Employees
Shard key – user_id

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Shard_key = user_id





Shard0	Shard1	Shard2	Shard*n*
Vin	Rashim	Srini	Andrew
Meet	Karthik	Kari	Sean

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Insert additional data

{"name": ”Abhilash", "lastname": ”Koppula”, "role": ”AWS Software Engineer", ”user_id": 10, "phone": "2223333"}

{"name": ”Sam", "lastname": ”Dwarkanath”, "role": ”General Manager", ”user_id": 11, "phone": "3334444"}

{"name": ”Cody”, "lastname": ”Allen”, "role": "AWS Solutions Architect", "user_id": 12, "phone": "4445555"}

{"name": ”Andy", "lastname": ”Jassy”, "role": ”CEO", "user_id": 13, "phone": "6667777"}

{"name": ”Adam", "lastname": "Selipsky”, "role": "AWS CEO", "user_id": 14, "phone": "7778888"}

{"name": ”Brandon", "lastname": ”Goldstein”, "role": "AWS Engineer", "user_id": 15, "phone": "8889999"}

{"name": ”Gitesh", "lastname": ”Tyagi”, "role": "AWS Engineer", "user_id": 16, "phone": "9991111"}

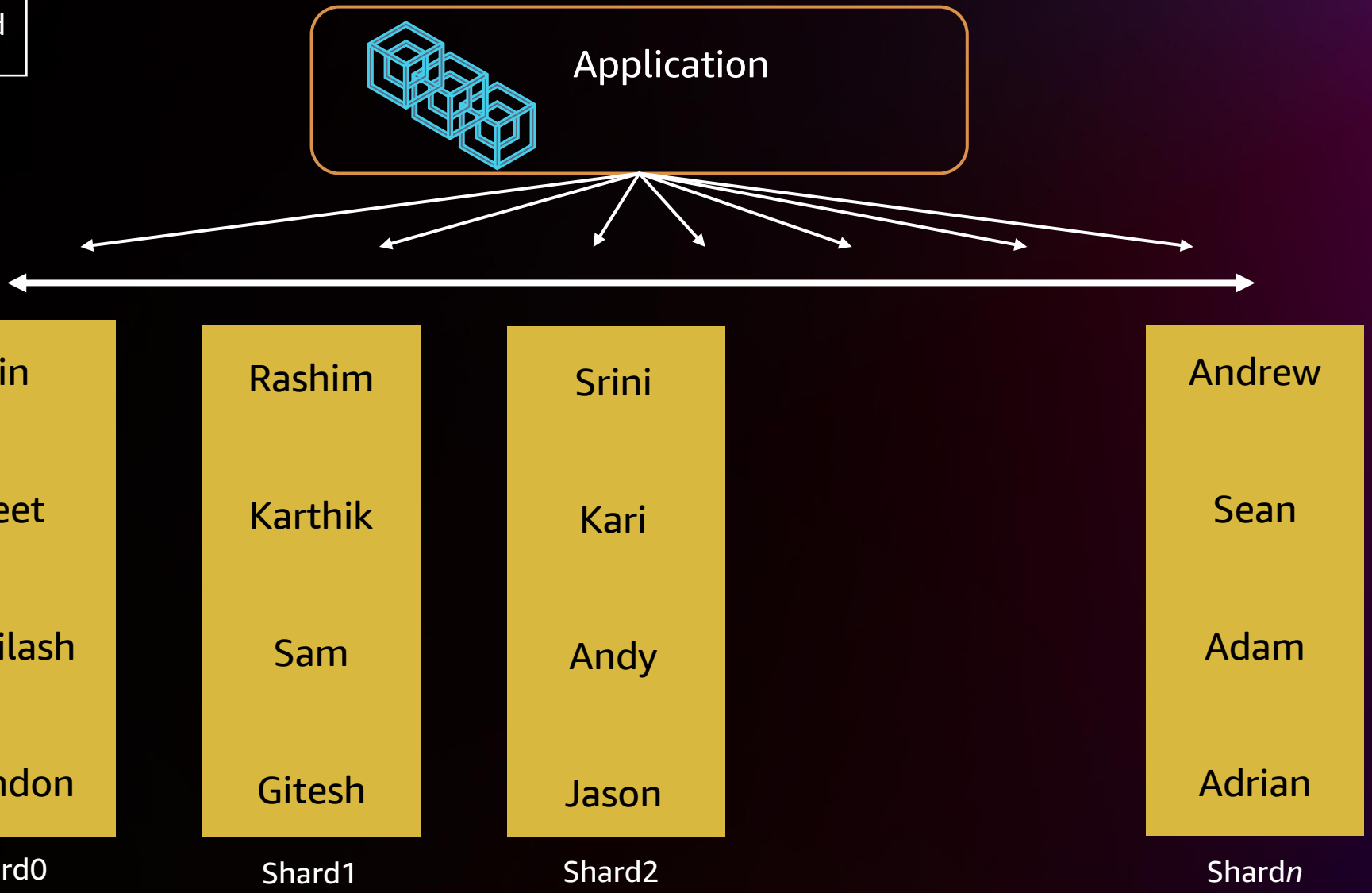
{"name": ”Jason", "lastname": ”Plank”, "role": "AWS GTM Specialist", "user_id": 17, "phone": "1112222"}

{"name": ”Adrian", "lastname": ”Chan”, "role": "AWS Software Engineer", "user_id": 18, "phone": "1112222"}

Collection – Employees
Shard key – user_id

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Shard_key = user_id

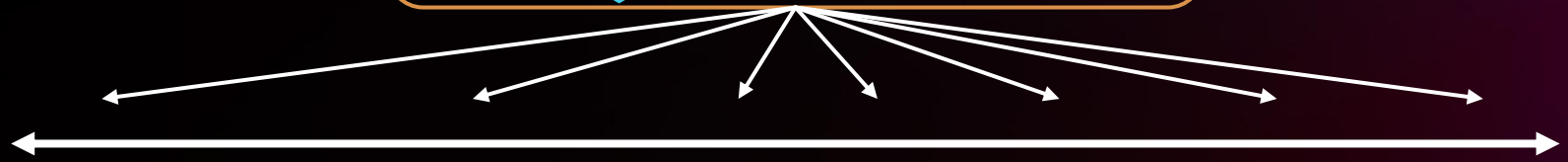


Amazon DocumentDB uses Hash Sharding to partition your data across underlying shards

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Shard_key = user_id





Shard0	Shard1	Shard2	Shard3	Shard*n*
Vin	Rashim	Srini	Brandon	Andrew
Meet	Karthik	Kari	Gitesh	Sean
Abhilash	Sam	Andy	Jason	Adam
				Adrian

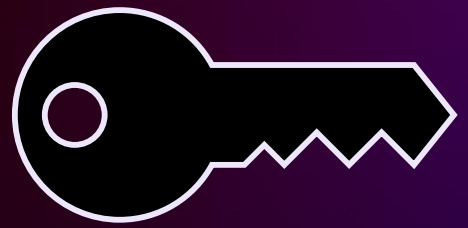
Amazon DocumentDB automatically repartitions data when you scale out by adding additional capacity

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What makes a good shard key?

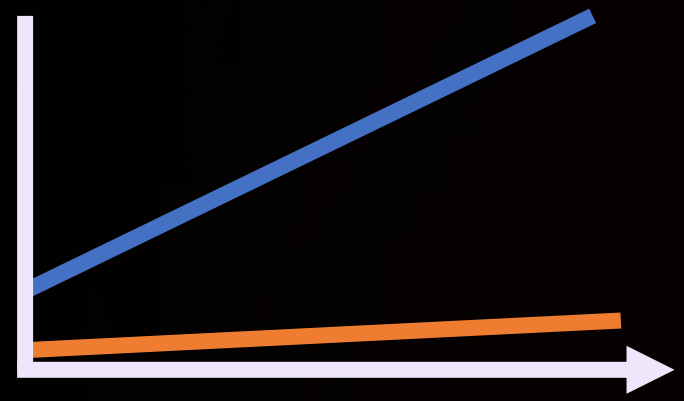
1. Evenly distributes storage



2. Evenly distributes load

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Challenge 1: Uneven data distribution

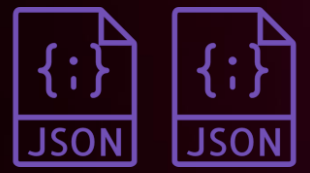


Shard 1

Shard 2

Shard key = State





"CA"

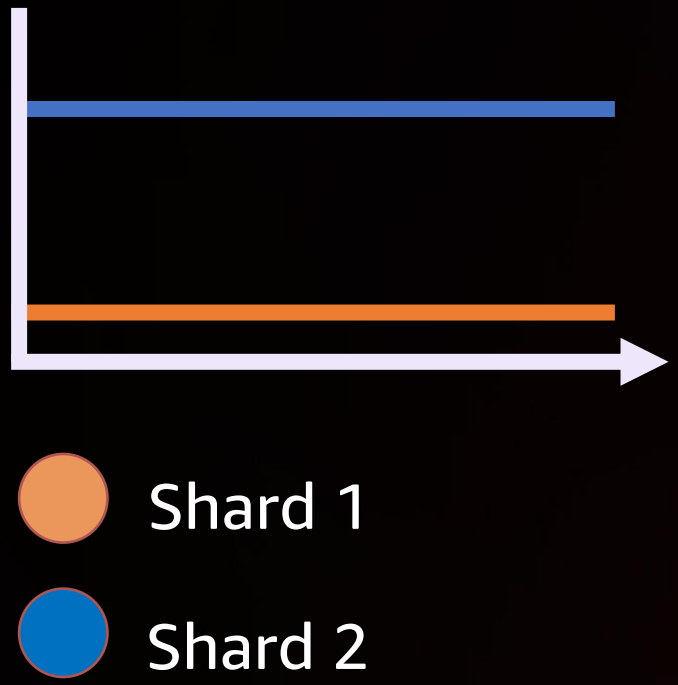


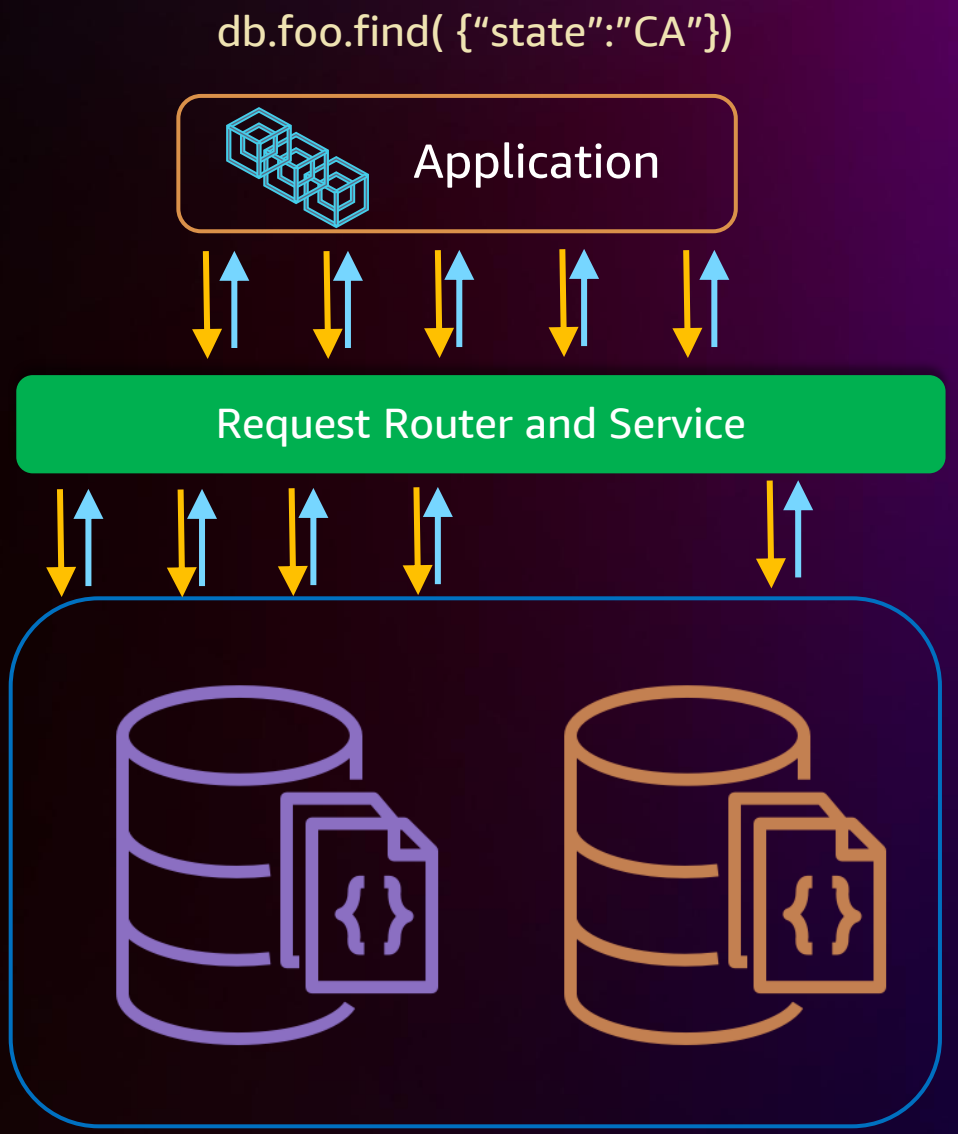
"WA"



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Challenge 2: Uneven load distribution





aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Choosing a shard key

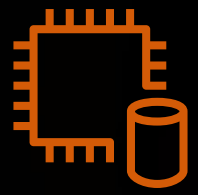
1. **High cardinality** (large number of unique values) –
   Good examples are employeeId, customerId, sessionId, orderId

2. **Shard key frequency and usage patterns –** Choose a shard key that aligns with the types of queries

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How is it priced?

No upfront costs and only pay for what you use with easy-to-understand predictable pricing



Compute – vCPUs per hour



Storage – GB per month



Backups – GB per month

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon DocumentDB free trial



What's included?

750 hours – t3.medium instance (2 vCPUs, 4GB RAM)

5 GB storage

30 million I/Os

5 GB backup

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Vin Yu

vin-yu

vyyu@amazon.com

Rashim Gupta

rashimgupta

rashimg@amazon.com



Please complete the session survey in the **mobile app**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.