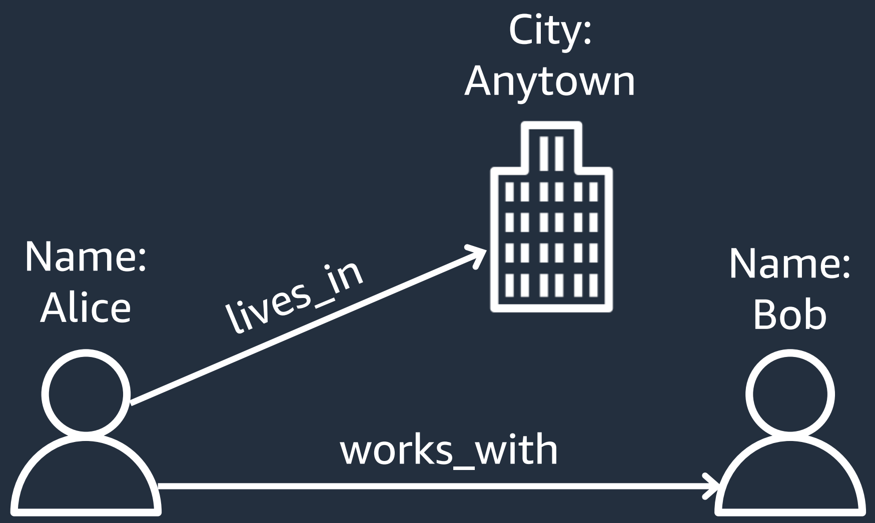
Graph Database introduction, deep-dive and demo with Amazon Neptune

Dave Bechberger (he/him)

Sr. Graph Architect
Amazon Neptune

© 2022, Amazon Web Services, Inc. or its affiliates.

Modern graphs

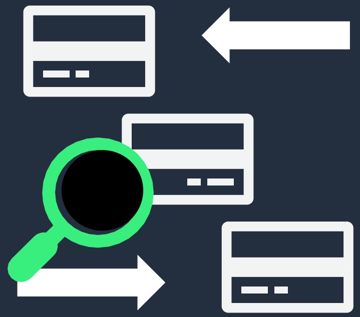


Graphs model data based on **relationships**

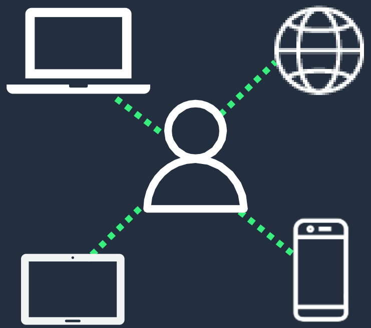
Graphs explore relationships and patterns in **connected data**

© 2022, Amazon Web Services, Inc. or its affiliates.

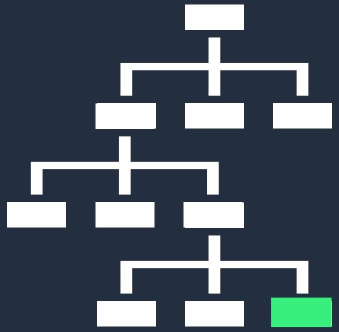
Customers are excited to use graphs



Fraud detection



Identity resolution



Knowledge organization (knowledge graph)

Security graphs

© 2022, Amazon Web Services, Inc. or its affiliates.

Common graph business problems

- We need to get better at detecting fraud
- My customers want better or more personalized recommendations
- We need to connect our siloed data sources
- We have multiple websites/applications and we need to link customer identities in these systems
- Our machine learning algorithms need improvement

© 2022, Amazon Web Services, Inc. or its affiliates.

Not so easily recognized graph problems

- Where are the risks in my IT Infrastructure/supply chain?
- Where did this data it come from?
- Why don’t my search results relate to my question?
- How does person X have access to information Y?
- How is this IAM role being used in my cloud infrastructure?

© 2022, Amazon Web Services, Inc. or its affiliates.

Graphs solve the Where, Why, and How

- **Where** are the risks in my IT Infrastructure/supply chain?
- **Where** did this data it come from?
- **Why** don’t my search results relate to my question?
- **How** does person X have access to information Y?
- **How** is this IAM role being used in my cloud infrastructure?

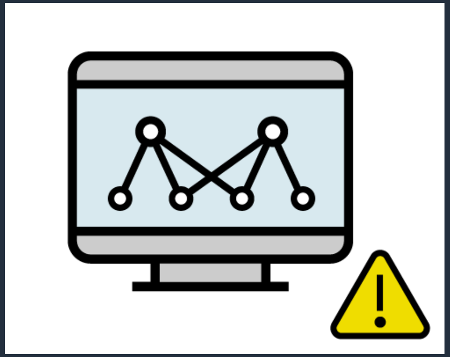
These questions:

- Navigate (variably) connected structure
- Filter or compute a result on the basis of the *strength*, *weight,* or *quality* of relationships
- Require traversing an unknown number of connections

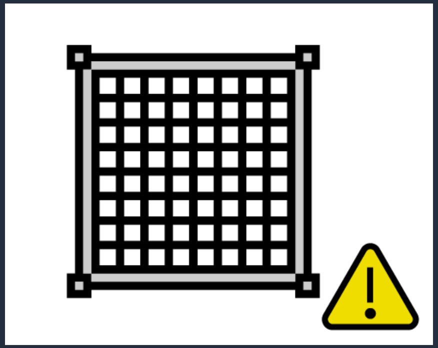
© 2022, Amazon Web Services, Inc. or its affiliates.

Challenges with highly connected data

Unnatural for querying



Inefficient processing



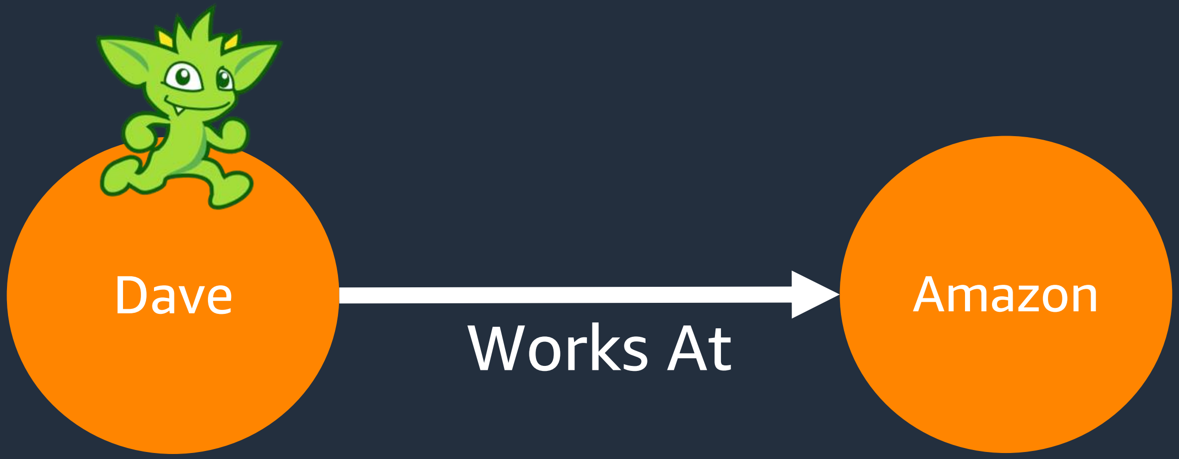
Rigid schema inflexible for changing data

© 2022, Amazon Web Services, Inc. or its affiliates.

Query languages

DESIGNED TO MOVE THROUGH DATA

Graph query languages are optimized to use connections to move through a network



Relational queries work by combining sets of data.

Person
Dave

$\bowtie$

Company
Amazon

=

Person	Company
Dave	Amazon

© 2022, Amazon Web Services, Inc. or its affiliates.

Efficient Processing

GRAPH DATABASES ARE OPTIMIZED FOR PROCESSING CONNECTED DATA

Let’s look at the example:

Dave – Works At -> Amazon

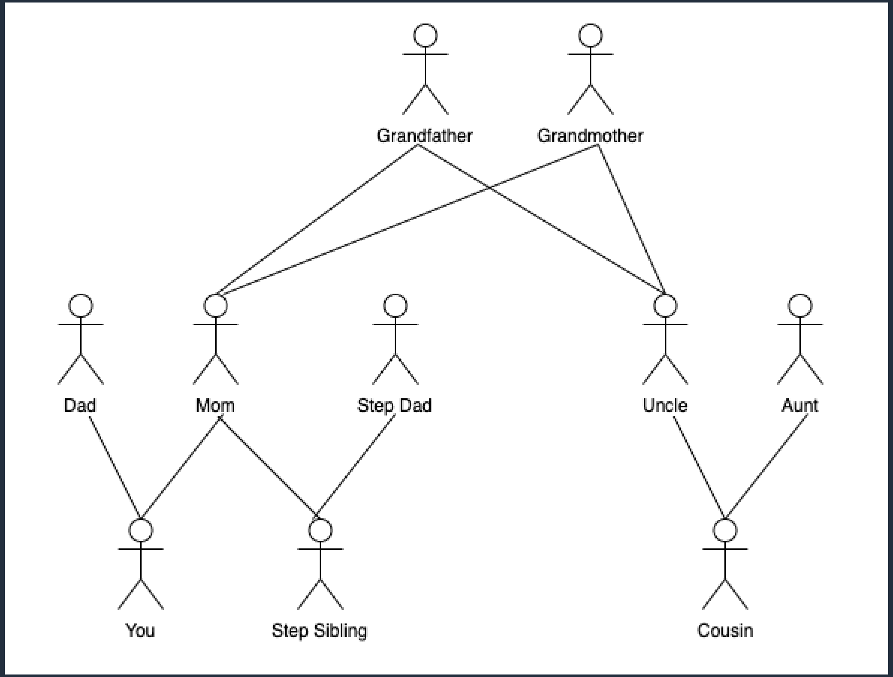
In a Graph database, the ***Works At*** connection is data, when needed the connection is retrieved

In a Relational databases, the ***Works At*** connection is metadata, when needed the connection must be calculated.

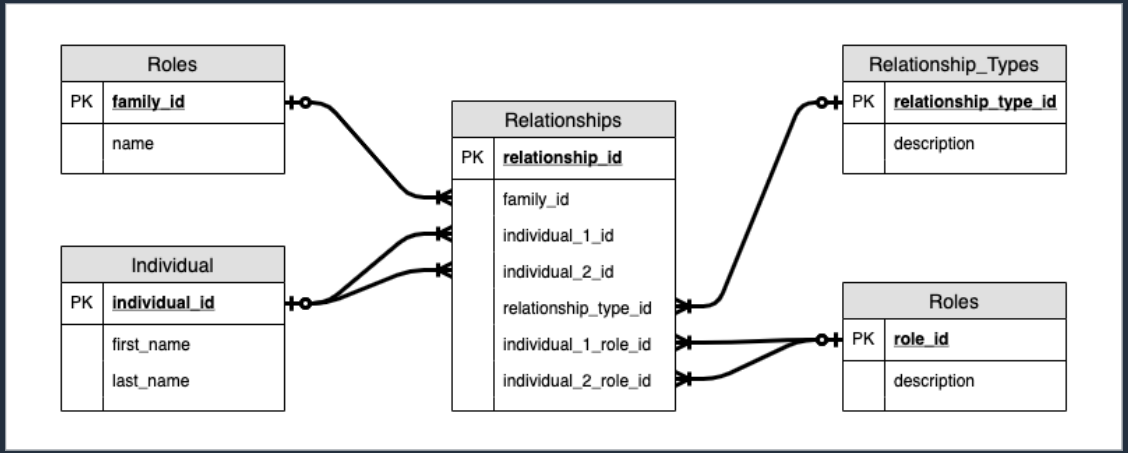
© 2022, Amazon Web Services, Inc. or its affiliates.

Schema flexibility

MAKING ADDING NEW DATA SOURCES EASY



vs



Bonus - Graphs are easier to understand by new and/or non-technical people

© 2022, Amazon Web Services, Inc. or its affiliates.

Why use a graph database service?

Traditional Databases are not optimized for connections

Self-managed solutions become complex

Evolution at Scale

© 2022, Amazon Web Services, Inc. or its affiliates.

Amazon Neptune

© 2022, Amazon Web Services, Inc. or its affiliates.

Amazon Neptune

Fully managed, purpose-built graph database in the cloud

Cost-effective

No hardware management

Instant provisioning

Scaling

Security & compliance

- Optimized to store and map billions of relationships
- Enables real-time navigation of connections with millisecond query response time
- Supports open standard query languages openCypher, Gremlin, and SPARQL

© 2022, Amazon Web Services, Inc. or its affiliates.

Leading graph models and languages

Property Graph

Resource Description Framework (RDF)

Some of our customers

Amazon Neptune

Customers across different verticals and use cases use Amazon Neptune in production today

LexisNexis FINRA SIEMENS AstraZeneca LIFEOMIC freshworks

amazon alexa THOMSON REUTERS FACTSET SAMSUNG Blackfynn Rappi

Uber ATG HUUUGE NBCUniversal NETFLIX asurion COX AUTOMOTIVE

PaySense intuit Pearson MARINUS ANALYTICS noonum audible FOR BUSINESS

© 2022, Amazon Web Services, Inc. or its affiliates.

Under the hood of Neptune

© 2022, Amazon Web Services, Inc. or its affiliates.

Architecture

Application Layer

Compute Layer

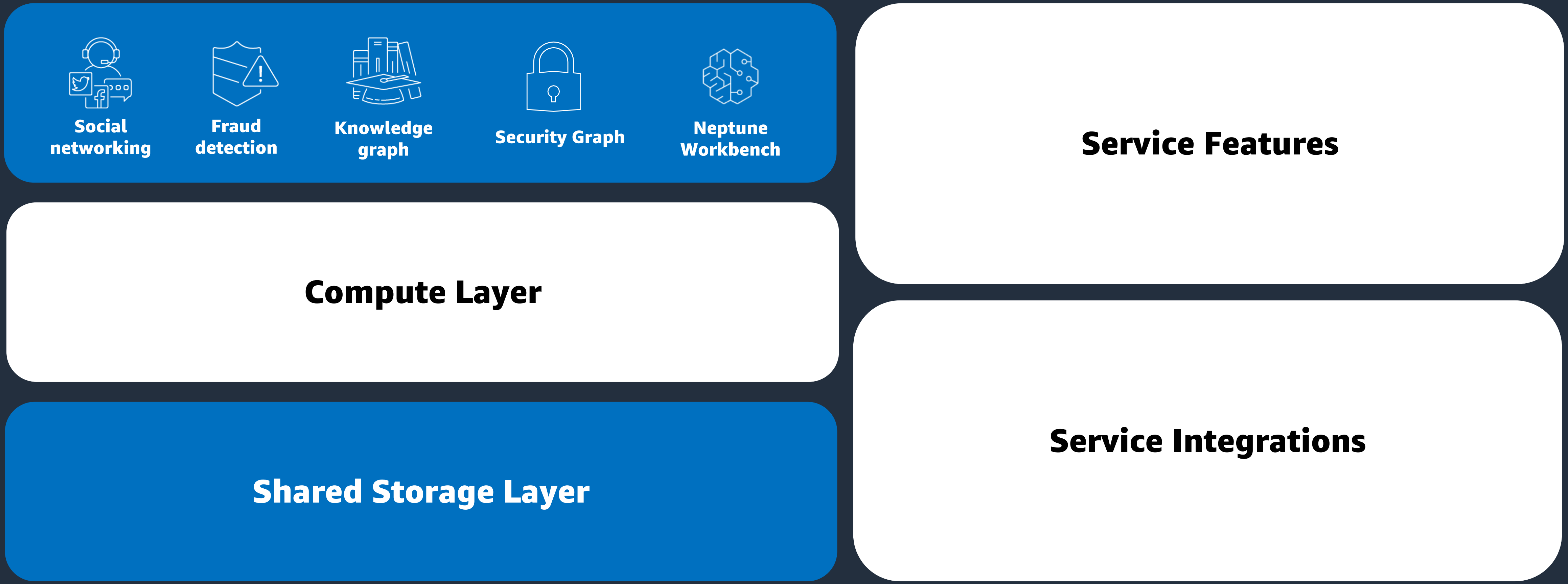
Shared Storage Layer

Service Features

Service Integrations

© 2022, Amazon Web Services, Inc. or its affiliates.

Architecture



© 2022, Amazon Web Services, Inc. or its affiliates.

Architecture

Social networking

Fraud detection

Knowledge graph

Security Graph

Neptune Workbench

R6g R5d R5 R4

T4g T3

1 writer and up to 15 read replicas

Shared Storage Layer

Service Features

Service Integrations

© 2022, Amazon Web Services, Inc. or its affiliates.

Architecture

Social networking

Fraud detection

Knowledge graph

Security Graph

Neptune Workbench

R6g R5d R5 R4 T4g T3

1 writer and up to 15 read replicas

6 copies of data across 3 AZs

Up to 128 TiB

Automated backup and restore

Database fast clone

Service Features

Service Integrations

© 2022, Amazon Web Services, Inc. or its affiliates.

Architecture

Social networking

Fraud detection

Knowledge graph

Security Graph

Neptune Workbench

R6g R5d R5 R4

T4g T3

1 writer and up to 15 read replicas

6 copies of data across 3 AZs

Up to 128 TiB

Automated backup and restore

Database fast clone

Bulk load from S3

Neptune Streams

Status Endpoint

Query Profile and Explain

Read replica auto scaling

Service Integrations

© 2022, Amazon Web Services, Inc. or its affiliates.

Architecture

Social networking

Fraud detection

Knowledge graph

Security Graph

Neptune Workbench

R6g R5d R5 R4

T4g T3

1 writer and up to 15 read replicas

6 copies of data across 3 AZs

Up to 128 TiB

Automated backup and restore

Database fast clone

Bulk load from S3

Neptune Streams

Status Endpoint

Query Profile and Explain

Read replica auto scaling

AWS Backup

Neptune ML

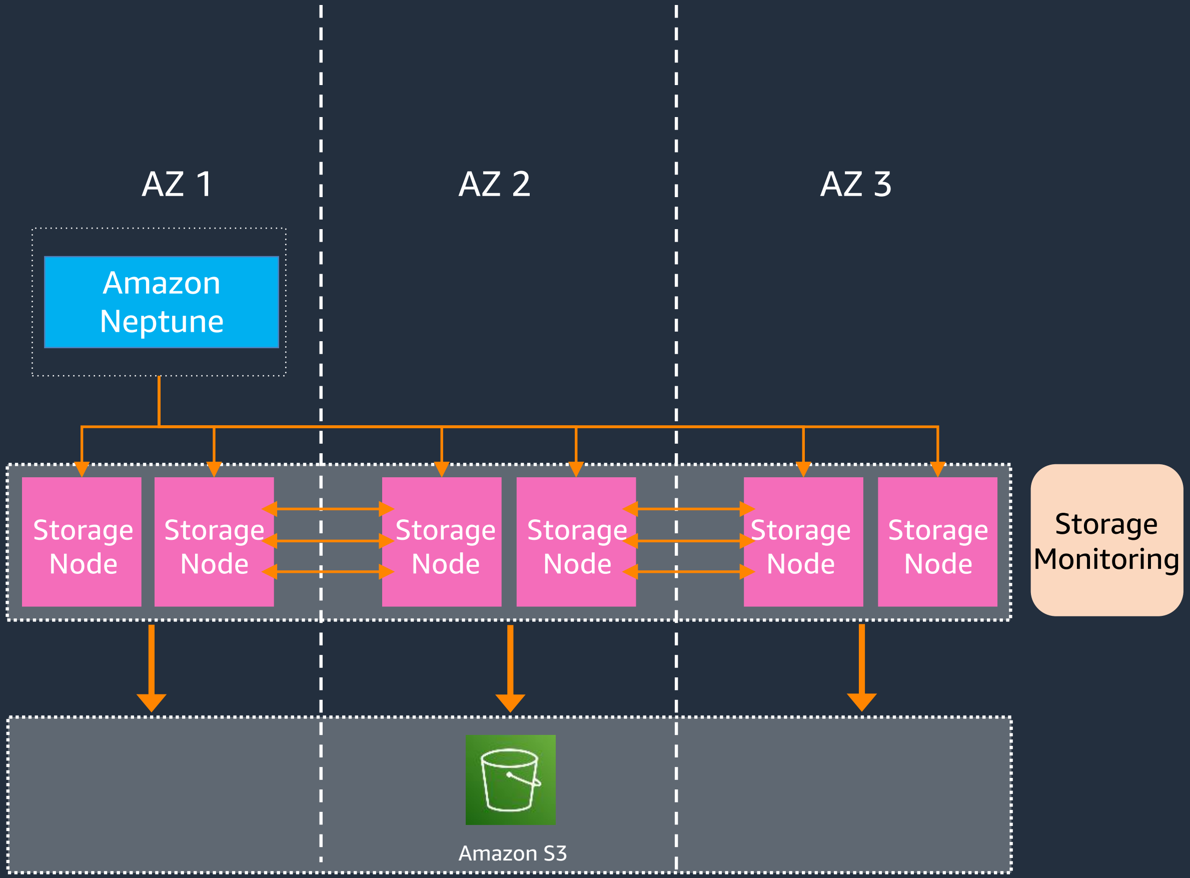
Amazon OpenSearch

© 2022, Amazon Web Services, Inc. or its affiliates.

Cloud-native storage

- Data is replicated 6 times across 3 AZs
- Continuous backup to Amazon S3

  Built for 11 9s durability
- 10 GB segments as unit of repair or hotspot rebalance
- Quorum system for read/write; latency tolerant
- Storage volume automatically grows up to 128 TiB



© 2022, Amazon Web Services, Inc. or its affiliates.

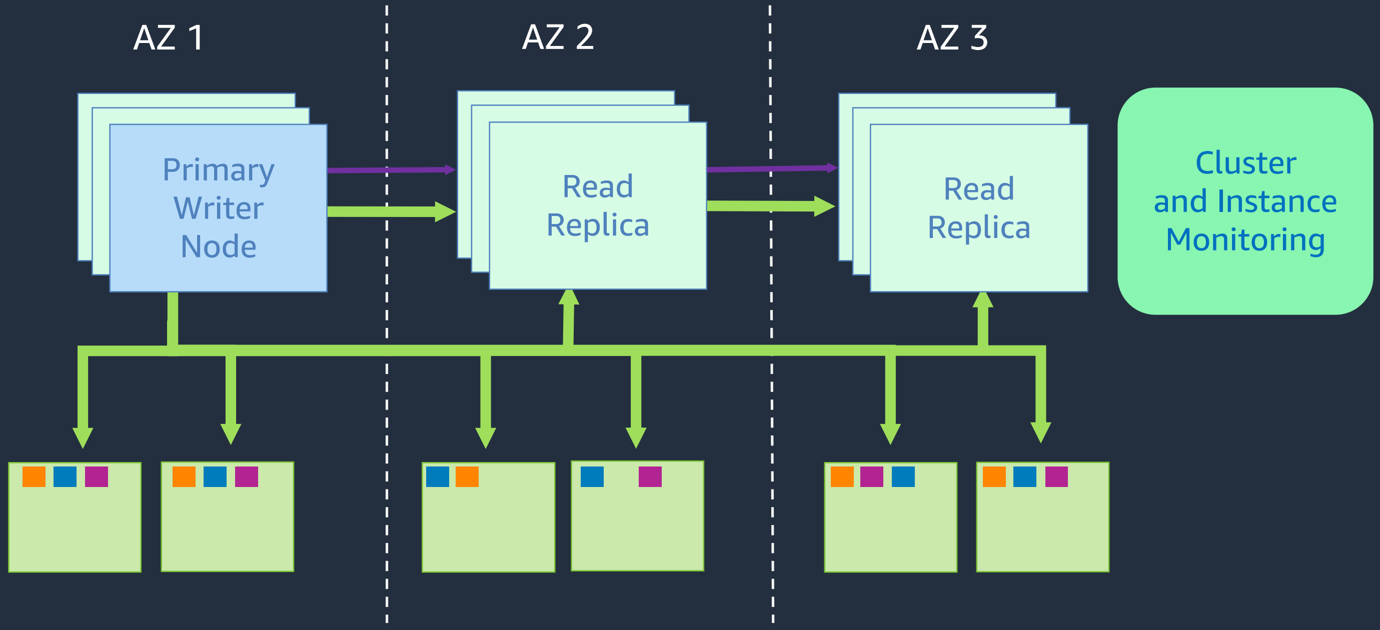
Read replicas

Availability

- Failing database nodes are automatically detected and replaced
- Failing database processes are automatically detected and recycled
- Replicas are automatically promoted to primary if needed (failover)
  - Customer-specifiable failover order

Performance

- Customer applications can scale out read traffic across read replicas
- Read balancing across read replicas
  - Use reader endpoint



© 2022, Amazon Web Services, Inc. or its affiliates.

Traffic distribution

Writer endpoint (Cluster endpoint)

- Always points to the current primary

Reader endpoint

- Distributes requests to available read replicas
- No fairness/round robin guarantee (often sufficient though)
- Gremlin WebSocket and OC Bolt connections are sticky

Custom request distribution strategy might be beneficial

- Fairness
- Optimizing cache locality, different timeouts
- Be aware of failovers!

© 2022, Amazon Web Services, Inc. or its affiliates.

Caching

Buffer Cache

- Uses the instance's memory
- Stores graph components
- Always on and enabled
- Monitor usage with CloudWatch metric BufferCacheHitRatio

Lookup Cache

- Uses R5d's NVMe-based SSD
- Stores property values (strings) and RDF literals
- Available on R5d instances only
- Use cases: frequent return/usage of large number of property values and RDF literals

Query Results Cache

- Uses the instance's memory
- Stores query results
- Can be enabled per-instance, except for t instance types
- Use cases: pagination, repeatable queries

© 2022, Amazon Web Services, Inc. or its affiliates.

Security

Network

- Network isolation – VPC only

Encryption

- Encryption at rest with AWS KMS
- Encryption in transit at all times (SSL only)

IAM

- IAM policies
- IAM Database Authentication
- Action based access control

© 2022, Amazon Web Services, Inc. or its affiliates.

Backup and restore

Automated Backups

- Daily automated backups during backup window
- Full storage volume snapshot
- Taken from read replica if possible
- Retention period: 1-35 days

Manual Snapshots

- Back up entire database instance
- Can be shared with other AWS accounts
- Can copy between regions

Restoring from a database snapshot (automated or manual)

- Creates new database instance
- Apply custom parameter groups and security groups after restore

Point-in-Time restore

- Restore from database instance (not snapshot)
- Creates new database instance
- Choose “Latest restorable time” or specify custom data and time
- 1 second granularity

© 2022, Amazon Web Services, Inc. or its affiliates.

Monitoring

AWS CloudTrail

- Log all Neptune API calls

Event notifications

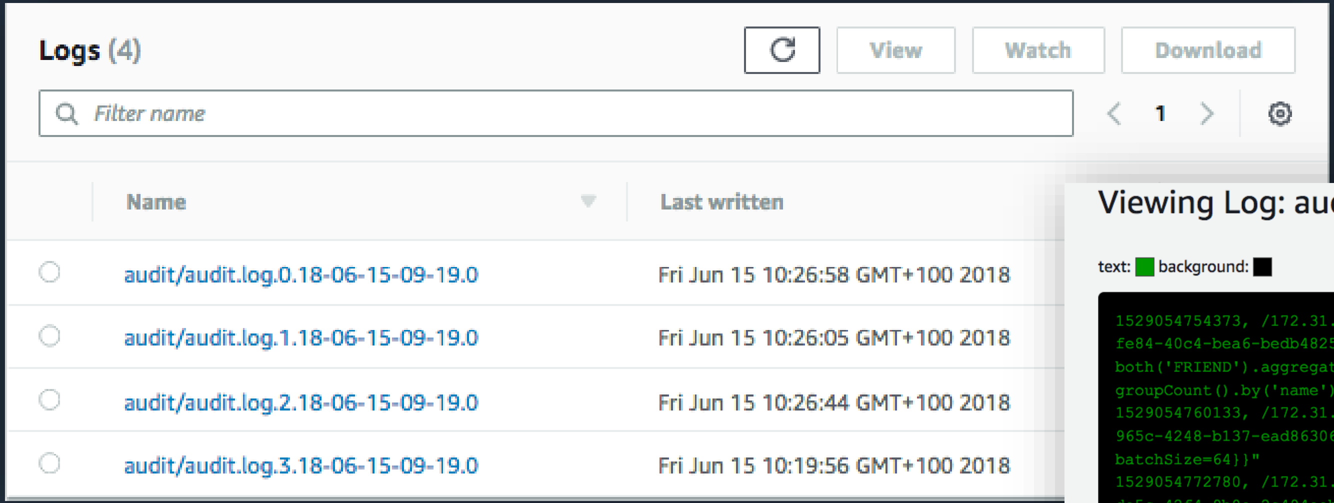
- Create SNS subscription via CLI or SDK
- Sources: `db-instance | db-cluster | db-parameter-group | db-security-group | db-snapshot | db-cluster-snapshot`

Amazon CloudWatch

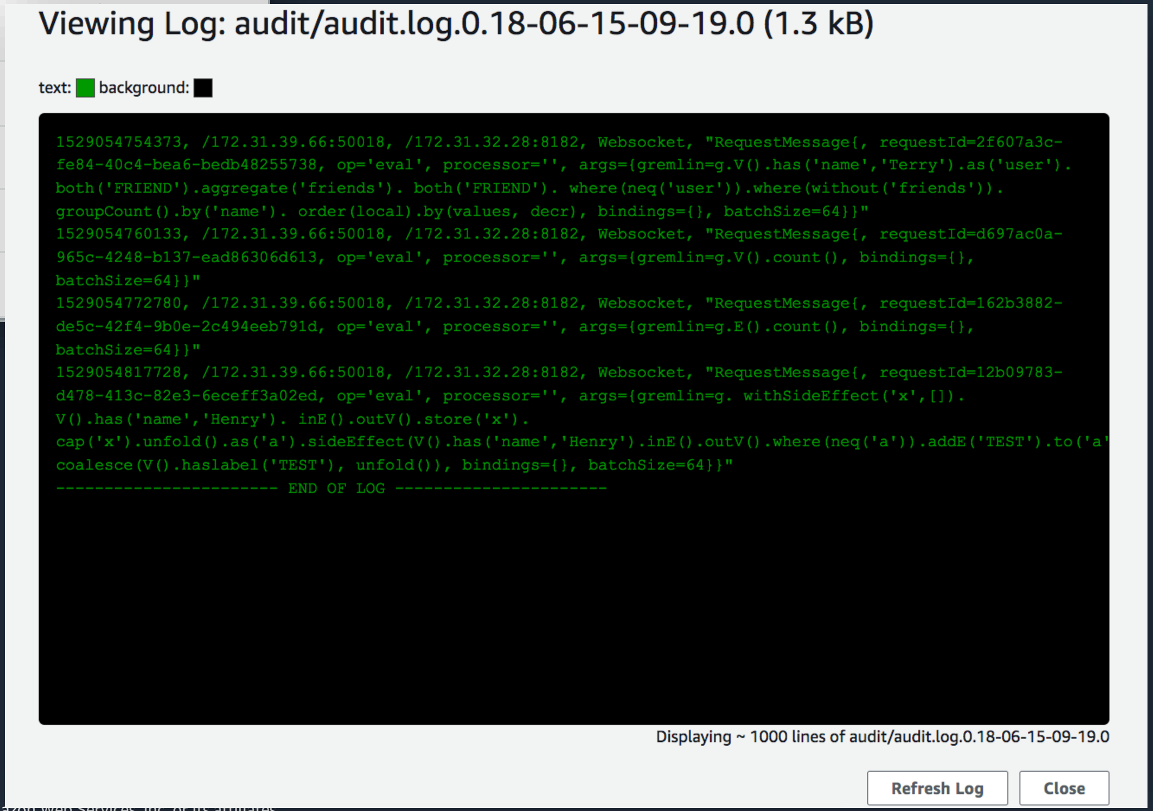
- CPU and memory utilization
- Query throughput
- Query success/error rates
- Read vs. write throughput
- Storage size

© 2022, Amazon Web Services, Inc. or its affiliates.

Audit Logs



Name	Last written
audit/audit.log.0.18-06-15-09-19.0	Fri Jun 15 10:26:58 GMT+100 2018
audit/audit.log.1.18-06-15-09-19.0	Fri Jun 15 10:26:05 GMT+100 2018
audit/audit.log.2.18-06-15-09-19.0	Fri Jun 15 10:26:44 GMT+100 2018
audit/audit.log.3.18-06-15-09-19.0	Fri Jun 15 10:19:56 GMT+100 2018



Viewing Log: audit/audit.log.0.18-06-15-09-19.0 (1.3 kB)

text: 🟩 background: ⬛

```
1529054754373, /172.31.39.66:50018, /172.31.32.28:8182, Websocket, "RequestMessage{, requestId=2f607a3c-
fe84-40c4-bea6-bedb48255738, op='eval', processor='', args={gremlin=g.V().has('name','Terry').as('user').
both('FRIEND').aggregate('friends'). both('FRIEND'). where(neq('user')).where(without('friends')).
groupCount().by('name'). order(local).by(values, decr), bindings={}, batchSize=64}}"
1529054760133, /172.31.39.66:50018, /172.31.32.28:8182, Websocket, "RequestMessage{, requestId=d697ac0a-
965c-4248-b137-ead86306d613, op='eval', processor='', args={gremlin=g.V().count(), bindings={},
batchSize=64}}"
1529054772780, /172.31.39.66:50018, /172.31.32.28:8182, Websocket, "RequestMessage{, requestId=162b3882-
de5c-42f4-9b0e-2c494eeb791d, op='eval', processor='', args={gremlin=g.E().count(), bindings={},
batchSize=64}}"
1529054817728, /172.31.39.66:50018, /172.31.32.28:8182, Websocket, "RequestMessage{, requestId=12b09783-
d478-413c-82e3-6eceff3a02ed, op='eval', processor='', args={gremlin=g. withSideEffect('x',[]).
V().has('name','Henry'). inE().outV().store('x').
cap('x').unfold().as('a').sideEffect(V().has('name','Henry').inE().outV().where(neq('a')).addE('TEST').to('a'
coalesce(V().haslabel('TEST'), unfold()), bindings={}, batchSize=64}}"
----------------------- END OF LOG ----------------------
```

Displaying ~ 1000 lines of audit/audit.log.0.18-06-15-09-19.0

- Timestamp
- ServerHost
- ClientHost
- ConnectionType
- RequestMessage

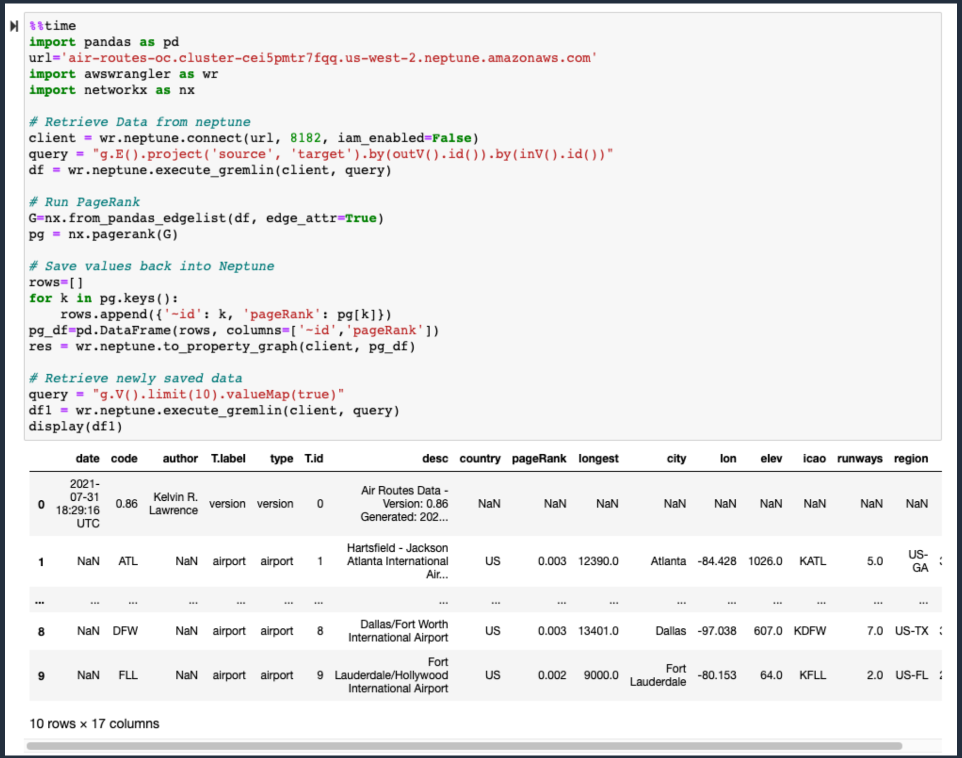
© 2022, Amazon Web Services, Inc. or its affiliates.

What’s new

© 2022, Amazon Web Services, Inc. or its affiliates.

Python Integration for graph analytics

OPEN SOURCE PYTHON INTEGRATION TO EASILY READ AND WRITE DATA STORED IN NEPTUNE



```
%%time
import pandas as pd
url='air-routes-oc.cluster-cei5pmtr7fqq.us-west-2.neptune.amazonaws.com'
import awswrangler as wr
import networkx as nx

# Retrieve Data from neptune
client = wr.neptune.connect(url, 8182, iam_enabled=False)
query = "g.E().project('source', 'target').by(outV().id()).by(inV().id())"
df = wr.neptune.execute_gremlin(client, query)

# Run PageRank
G=nx.from_pandas_edgelist(df, edge_attr=True)
pg = nx.pagerank(G)

# Save values back into Neptune
rows=[]
for k in pg.keys():
    rows.append({'~id': k, 'pageRank': pg[k]})
pg_df=pd.DataFrame(rows, columns=['~id','pageRank'])
res = wr.neptune.to_property_graph(client, pg_df)

# Retrieve newly saved data
query = "g.V().limit(10).valueMap(true)"
df1 = wr.neptune.execute_gremlin(client, query)
display(df1)
```

	date	code	author	T.label	type	T.id	desc	country	pageRank	longest	city	lon	elev	icao	runways	region
0	2021-07-31 18:29:16 UTC	0.86	Kelvin R. Lawrence	version	version	0	Air Routes Data - Version: 0.86 Generated: 202...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	NaN	ATL	NaN	airport	airport	1	Hartsfield - Jackson Atlanta International Air...	US	0.003	12390.0	Atlanta	-84.428	1026.0	KATL	5.0	US-GA
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
8	NaN	DFW	NaN	airport	airport	8	Dallas/Fort Worth International Airport	US	0.003	13401.0	Dallas	-97.038	607.0	KDFW	7.0	US-TX
9	NaN	FLL	NaN	airport	airport	9	Fort Lauderdale/Hollywood International Airport	US	0.002	9000.0	Fort Lauderdale	-80.153	64.0	KFLL	2.0	US-FL

10 rows × 17 columns

Manipulate data using Pandas DataFrames

Read/write Neptune data using openCypher, Gremlin, or SPARQL in any Python environment

Use with popular open-source Python tools

Run analytics and algorithm workloads using iGraph, NetworkX and any other Python libraries of your choice

Get started with sample application notebooks

Two analytics tutorials for fraud detection and customer 360 graph applications

© 2022, Amazon Web Services, Inc. or its affiliates.

openCypher for Amazon Neptune

DEVELOPERS CAN NOW USE OPENCYPHER, A POPULAR GRAPH QUERY LANGUAGE, WITH AMAZON NEPTUNE, PROVIDING THEM THE MOST CHOICE TO BUILD OR MIGRATE GRAPH APPLICATIONS

openCypher

Declarative

A popular declarative query language. Provides users a familiar structure to compose queries for graph applications.

Accelerated learning curve

It's SQL inspired syntax allows customers to draw on their SQL knowledge to help power their businesses with graph applications

No Extra Costs

There are no additional charges for using openCypher with Neptune, eliminating expensive enterprise licensing costs.

Works on existing property graph

Neptune's openCypher support is compatible with our customers' existing property graphs, customers do not need to create new graph databases.

© 2022, Amazon Web Services, Inc. or its affiliates.

Neptune ML

Make predictions on graph data without ML expertise
Automatically choose and train the best ML model for your workload, enabling you to make ML-based predictions on graph data in hours instead of weeks

Use state-of-the-art graph machine learning
Up to 50% more accurate, GNNs are state-of-the-art ML purpose-built to use the relationships in graphs based on research from Stanford University

Scale to large datasets
For graph applications with billions of relationships in knowledge graphs, fraud detection, or product recommendations

Support for custom models
Define your own custom model implementations with Python

New!

Support for SPARQL
Make predictions from both Gremlin and SPARQL queries

New!

© 2022, Amazon Web Services, Inc. or its affiliates.

Fine grained access control for data plane actions

SUPPORT FOR FINE GRAINED ACCESS CONTROL FOR DATA PLANE ACTIONS WHEN USING IAM AUTHENTICATION

Access Control

Provide fine grained access to users accessing Neptune data plane APIs for graph-data actions such as reading, writing, and deleting data from the graph

Default with IAM Auth

Simply enable IAM Authentication on Neptune engine release version 1.2.0.0 and start using FGAC

Manage all data plane APIs

Separate policies to provide access to any data plane API.

© 2022, Amazon Web Services, Inc. or its affiliates.

Auto Scale Read Replicas

ADD AND REMOVE READ REPLICAS TO YOUR NEPTUNE CLUSTER TO MEET APPLICATION DEMANDS

Configure Minimum and Maximum Capacity

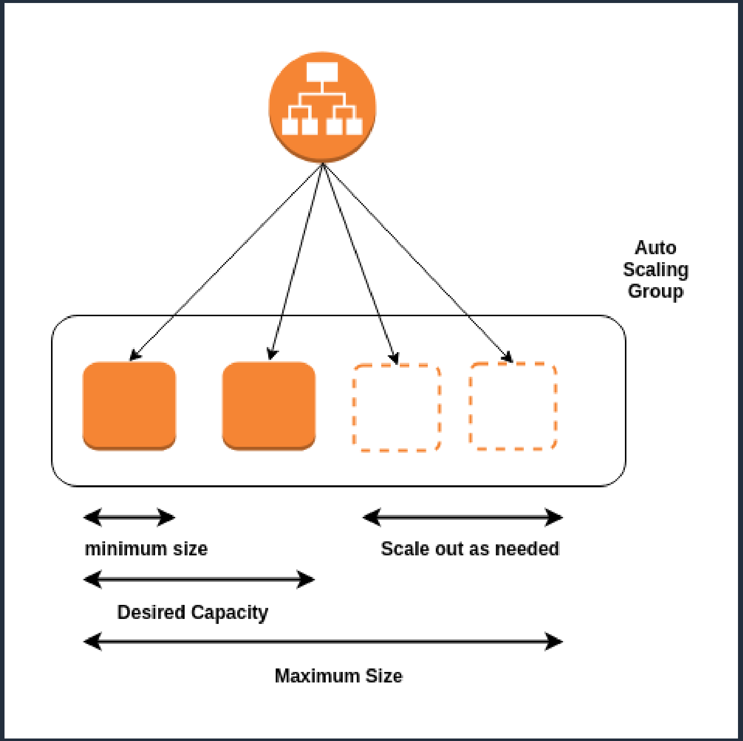
Scale out to as many as 15 read replicas

Define scaling threshold

Based on CloudWatch Metrics

Automate scaling activities

Eliminates operational burden of manual scaling



© 2022, Amazon Web Services, Inc. or its affiliates.

Amazon Neptune Global Database

DEPLOY NEPTUNE CLUSTERS ACROSS MULTIPLE AWS REGIONS FOR FAST CROSS-REGION DISASTER RECOVERY AND LOW-LATENCY GLOBAL READS

Disaster Recovery

Maintain business continuity in the event of regional outages with fast global failover to secondary AWS Regions

Low Latency Reads

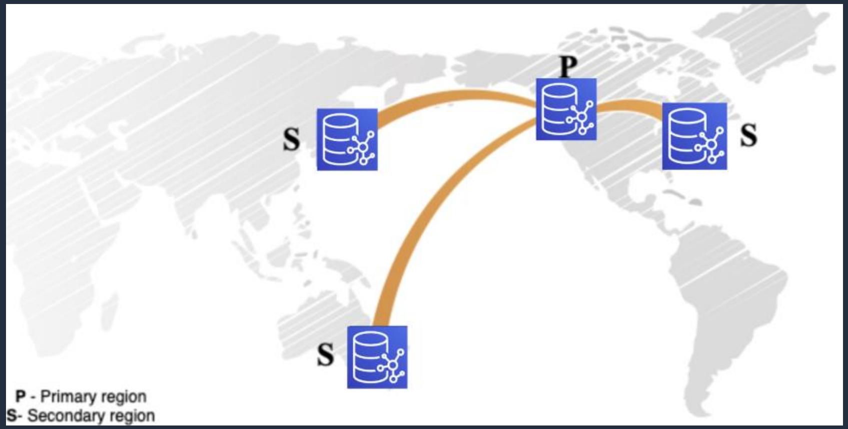
Connect to the Neptune cluster closest to your applications.

Fast cross region migrations

Migrate primary clusters to new AWS Region

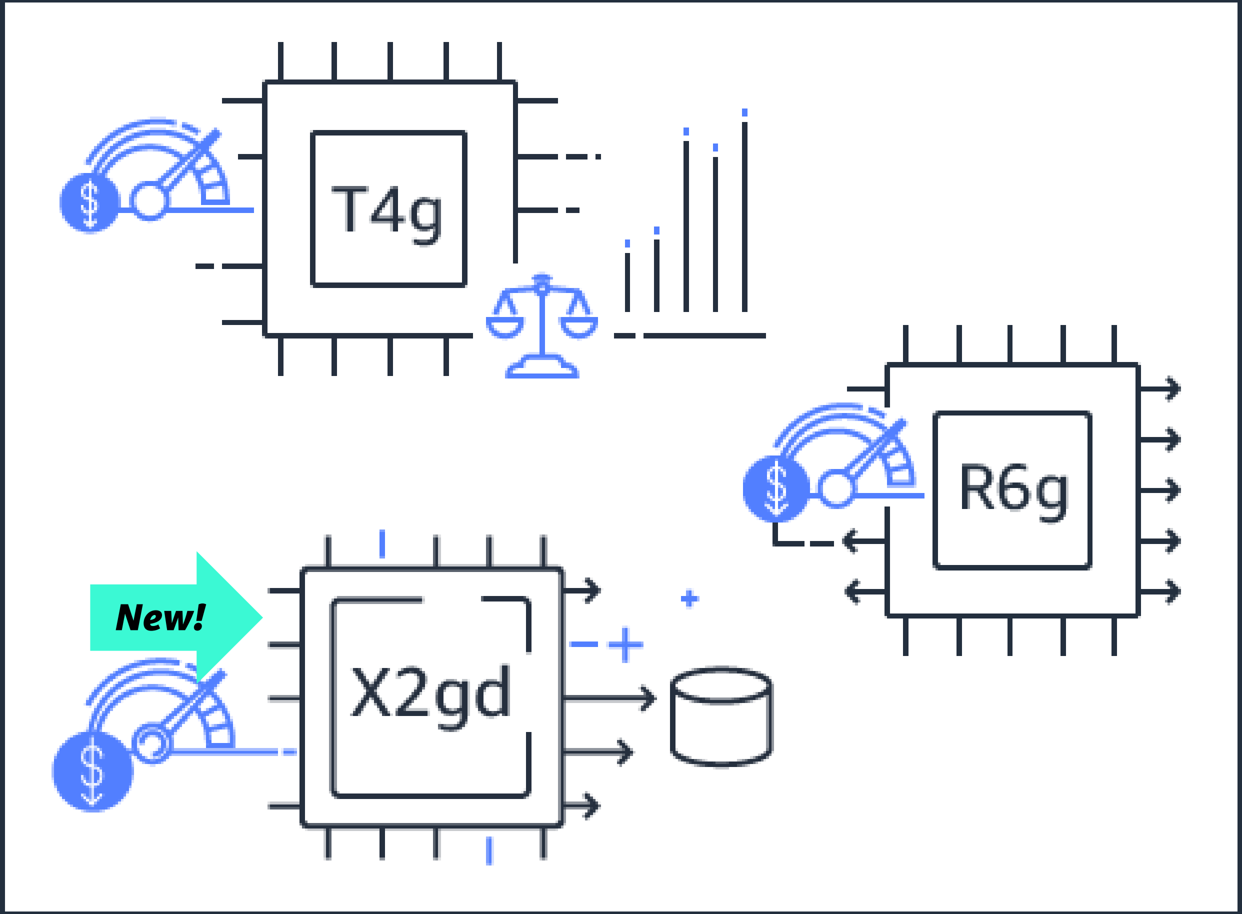
Low replication lag

Fast replication between AWS Regions



© 2022, Amazon Web Services, Inc. or its affiliates.

Graviton2 for Neptune



Better price performance

Improve query latency at a lower cost in comparison to x86-based instances of equivalent instance size

Inherit benefits of the AWS Nitro System

The AWS Nitro System offers private networking and fast local storage

Supports T4g, R6g, and X2g instance types

Provision low-cost burstable performance workloads using T4g instances, for development and testing use cases. Use the memory-optimized R6g instance for production workloads

© 2022, Amazon Web Services, Inc. or its affiliates.

Free Trial - Build graph applications for free!

IF YOUR ORGANIZATION HAS NEVER CREATED AN AMAZON NEPTUNE CLUSTER, NOW YOU CAN GET STARTED FOR FREE

Freedom of Choice

Freedom to get started

Freedom from expensive licenses

© 2022, Amazon Web Services, Inc. or its affiliates.

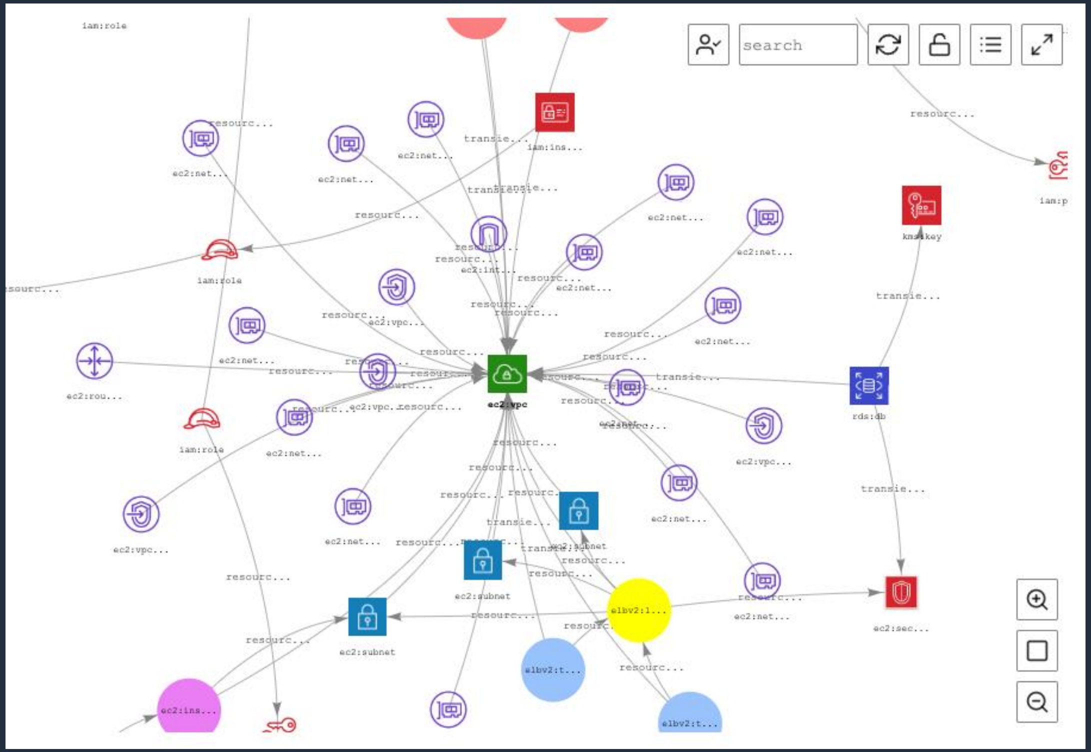
Demo

© 2022, Amazon Web Services, Inc. or its affiliates.

Demo - Altimeter

Altimeter – open source project from Tableau that graphs AWS resources

Blog Post - Graph your AWS resources with Amazon Neptune



© 2022, Amazon Web Services, Inc. or its affiliates.

Resources

© 2022, Amazon Web Services, Inc. or its affiliates.

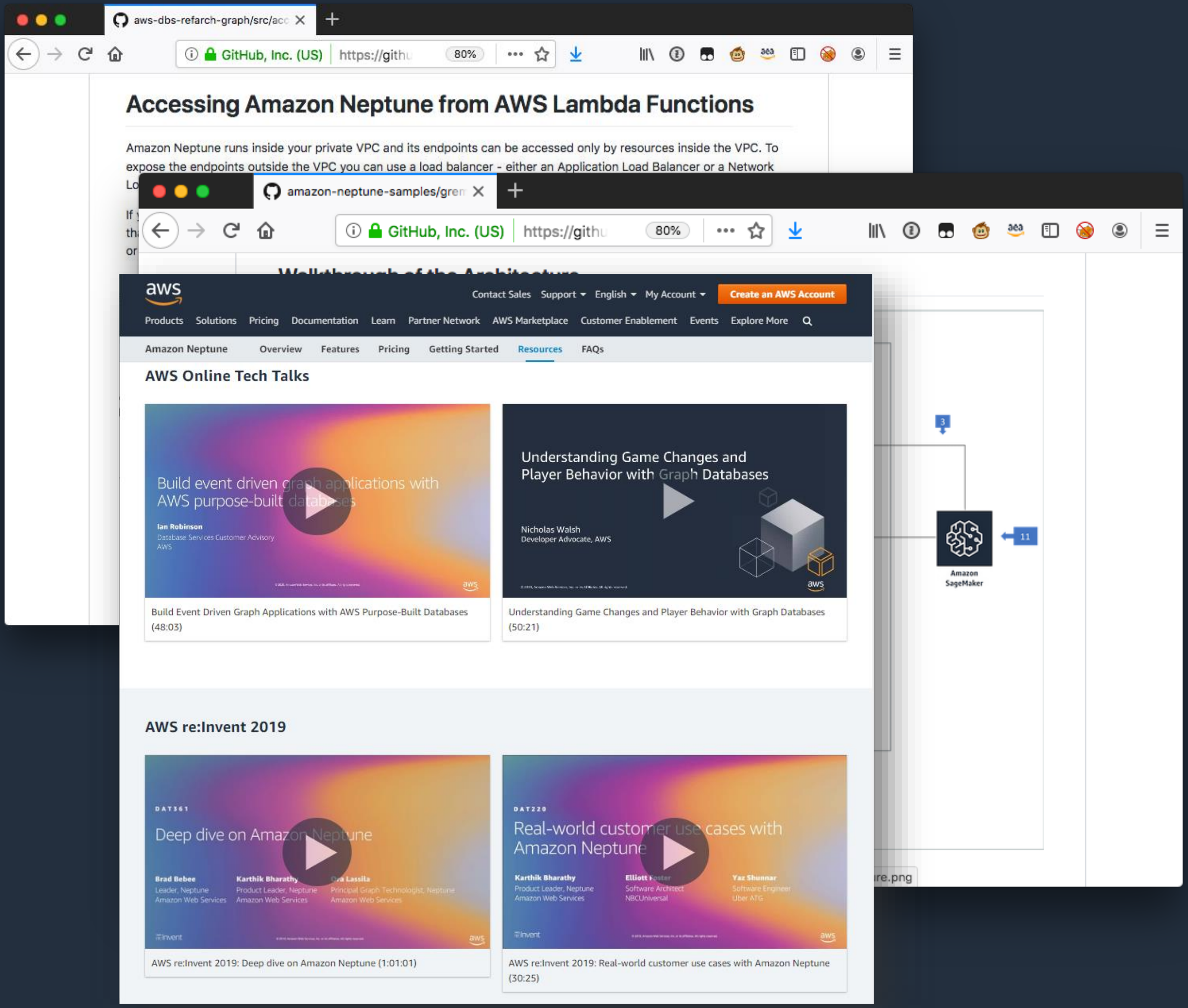
Additional resources

[Neptune Notebooks/Graph Notebook]

[Neptune Reference Architectures]

[Neptune Sample Applications]

[Use Cases, Videos, Blogs, Code ….]



Accessing Amazon Neptune from AWS Lambda Functions

Amazon Neptune runs inside your private VPC and its endpoints can be accessed only by resources inside the VPC. To expose the endpoints outside the VPC you can use a load balancer - either an Application Load Balancer or a Network

AWS Online Tech Talks

Build Event Driven Graph Applications with AWS Purpose-Built Databases (48:03)

Understanding Game Changes and Player Behavior with Graph Databases (50:21)

AWS re:Invent 2019

AWS re:Invent 2019: Deep dive on Amazon Neptune (1:01:01)

AWS re:Invent 2019: Real-world customer use cases with Amazon Neptune (30:25)

© 2022, Amazon Web Services, Inc. or its affiliates.

Thank you!

Dave Bechberger

Twitter: @bechbd

© 2022, Amazon Web Services, Inc. or its affiliates.