AWS re:Invent

DAT323-R1

What's new with Amazon ElastiCache

Ruchita Arora
Product Manager
Amazon Web Services

Itay Maoz
Engineering Manager
Amazon Web Services

Nir Mashkowski
Product Manager
Amazon Web Services

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Real-time applications

Amazon ElastiCache

What’s new in 2019



Real-time applications need low latency and high throughput



Mobile apps

Web applications

E-commerce

Gaming or location apps

Users:	1 million+
Data volume:	TBs – PB
Locality:	Global
Performance:	Microseconds
Request rate:	Millions
Access:	Mobile, IoT, and devices
Scale:	Up/down, in/out
Economics:	Pay as you go
Developer access:	Open APIs

All Things Distributed

Werner Vogels' weblog on building scalable and robust distributed systems.

A one size fits all database doesn't fit anyone

By Werner Vogels on 21 June 2018 10:00 AM | Permalink | Comments (10)

A common question that I get is why do we offer so many database products? The answer for me is simple: Developers want their applications to be well architected and scale effectively. To do this, they need to be able to use multiple databases and data models within the same application.

Seldom can one database fit the needs of multiple distinct use cases. The days of the one-size-fits-all monolithic database are behind us, and developers are now building highly distributed applications using a multitude of purpose-built databases. Developers are doing what they do best: breaking complex applications into smaller pieces and then picking the best tool to solve each problem. The best tool for a job usually differs by use case.

For decades because the only database choice was a relational database, no matter the shape or function of the data in the application, the data was modeled as relational. Instead of the use case driving the requirements for the database, it was the other way around. The database was driving the data model for the application use case. Is a relational database purpose-built for a normalized schema and to enforce referential integrity in the database? Absolutely, but the key point here is that not all application data models or use cases match the relational model.



Contact Info

Werner Vogels
CTO - Amazon.com

werner@allthingsdistributed.com

Data categories and common use cases



Relational

Referential integrity, ACID transactions, and schema-on-write

CRM, Finance



Key-value

Low-latency, key lookups with high throughput and fast ingestion of data

Real-time bidding, shopping cart, social, product catalog, and customer preferences



Document

Indexing and storing documents with support for query on any attribute

Content management, personalization, and mobile



In-memory

Microseconds latency, key-based queries, and data structures

Leaderboards, real-time analytics, and caching



Graph

Creating and navigating relationships between data easily and quickly

Fraud detection, social networking, and recommendation engine

Real-time applications need low latency and high concurrency



Caching



Real-time analytics



Gaming leaderboards



Geospatial



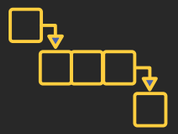
Media streaming



Session store



Chat apps



Message queues



Machine learning

Memcached



ElastiCache for
Memcached

Fast

Open source

No persistence

Very established

Easy to scale

Slab allocator

Supports strings, objects

Multi-threaded

In-memory
key-value data store

Redis



Fast
<1 ms latency for most commands

Open source

Simple

Highly available
Replication

Atomic operations
Supports transactions

In-memory key-value store

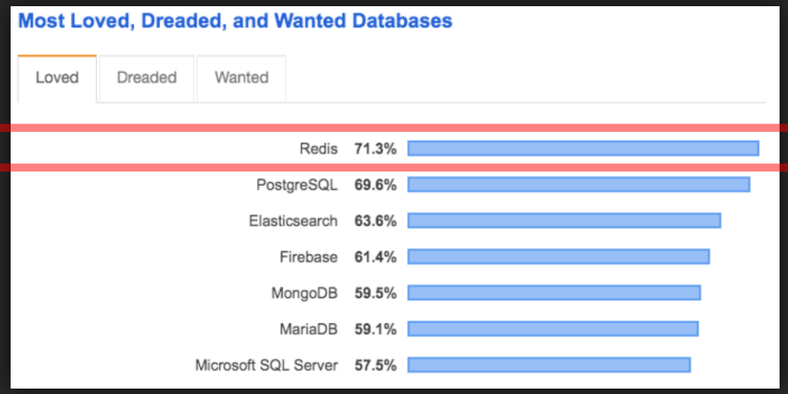
Powerful
~200 commands, Lua scripting, geospatial, and pub/sub

Various data structures
Strings, lists, hashes, sets, sorted sets, streams, bitmaps, and HyperLogLogs

Backup/restore
Enables snapshotting

Redis—#1 key-value store, most-loved database





* StackOverflow 2019 developer survey

Rank Nov 2019	Rank Oct 2019	Rank Nov 2018	DBMS	Database Model	Score Nov 2019	Score Oct 2019	Score Nov 2018
1.	1.	1.	Redis	Key-value, Multi-model	145.24	+2.32	+1.06
2.	2.	2.	Amazon DynamoDB	Multi-model	61.37	+1.19	+7.56
3.	3.	↑ 4.	Microsoft Azure Cosmos DB	Multi-model	31.98	+0.65	+9.94
4.	4.	↓ 3.	Memcached	Key-value	25.12	-0.78	-4.63
5.	5.	5.	Hazelcast	Key-value, Multi-model	7.82	+0.11	-1.45
6.			etcd	Key-value	7.15		
7.	7.	↑ 9.	Aerospike	Key-value	6.19	+0.42	+0.42
8.	↓ 6.	↓ 6.	Ehcache	Key-value	6.15	+0.12	-0.56
9.	↓ 8.	↓ 7.	Riak KV	Key-value	5.67	+0.11	-0.92
10.	↓ 9.	↓ 8.	OrientDB	Multi-model	5.38	+0.25	-0.41

* https://db-engines.com rankings for key-value store

Amazon ElastiCache

Amazon ElastiCache offers fully managed Redis and Memcached. Seamlessly deploy, run, and scale popular open source compatible in-memory data stores. Build data-intensive apps or improve the performance of your existing apps by retrieving data from high throughput and low latency in-memory data stores. Amazon ElastiCache is a popular choice for gaming, ad-tech, financial services, healthcare, and IoT apps.





ElastiCache for Memcached



ElastiCache for Redis

Introducing Amazon ElastiCache for Redis

Fully managed, Redis compatible, in-memory data store in the cloud



Extreme performance

In-memory data store and cache for sub-millisecond response times



Fully managed

AWS manages all hardware and software setup, configuration, monitoring



Easily scalable

Read scaling with replicas

Write and memory scaling with sharding

Online scaling



Reliable

Multi-AZ

Deep monitoring

Automatic failover



Secure and compliant

Amazon VPC

HIPAA, PCI, FedRAMP

Encryption at-rest and in-transit

Authentication



Redis compatible

Redis 5.0.5 Support

Redis clients compatible

Amazon ElastiCache for Redis customers

Using for **caching layer** for real-time analytics

A/B testing on Expedia's website with **~200M messages daily**

Using for **caching layer** for real-time data computation for **4 million daily bookings**

Grab estimates savings of 30%–40% on manpower with AWS



User session data.

Upgraded to optimized R5 instances to get 2/3rds CPU drop and handle traffic surge without scaling

Session state for their applications

Help developers build container-based stateless applications on AWS

Amazon ElastiCache for Redis: What's New - 2019

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

Amazon ElastiCache for Redis: What's new - 2019

Performance



Enhanced I/O

T3 Instance Support

Security



Auth Token Rotation

Customer Managed Master Keys

Rename Command

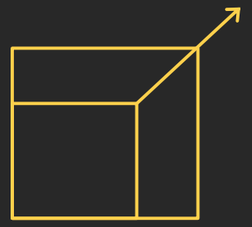
Fully Managed



Online migration from self-hosted Redis on Amazon Elastic Compute Cloud (Amazon EC2)

Self-service Maintenance updates

Scalability



Scale up / down

Reader endpoints

OSS compatibility



Redis 5.0.5

Memcached 1.5.16

Online migration from self-hosted Redis on Amazon EC2



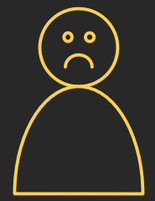
Self-managing Redis is challenging





Difficult to manage

Manage server provisioning, software patching, setup, configuration, and backups



Hard to make highly available

Need to implement fast error detection and remediation



Difficult to scale

Online scaling can be error prone and replication performance needs to be monitored



Expensive

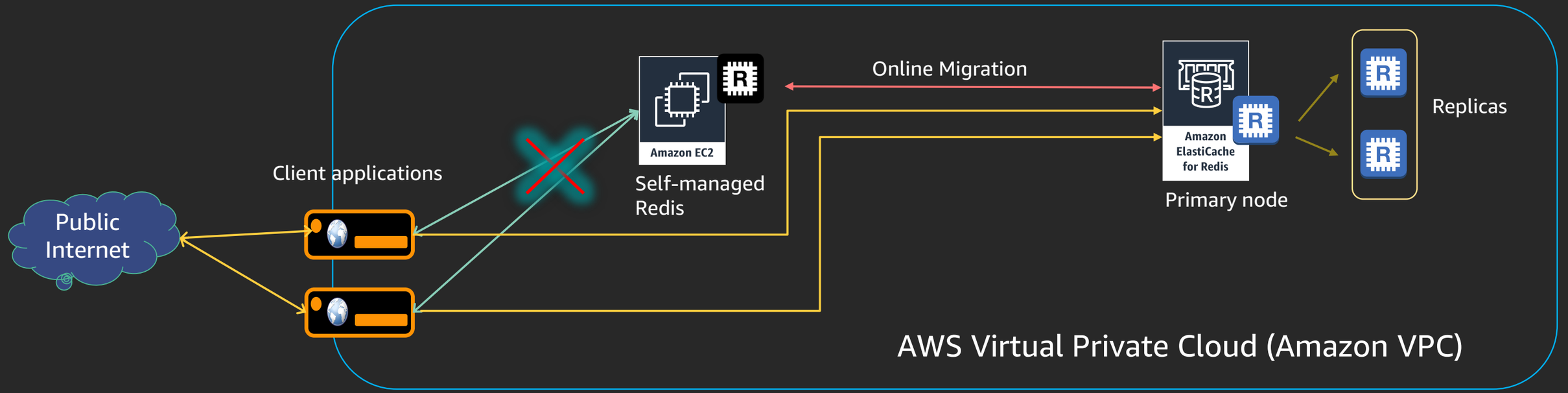
Invest in people, processes, hardware, and software

Online Migration to Amazon ElastiCache

Amazon ElastiCache data migration feature supports migrating your data from self-hosted Redis on Amazon EC2 to fully managed Amazon ElastiCache cluster-mode disabled configuration

Amazon ElastiCache replicates your Redis data in real time.

As soon as the data migration completes, update your client applications to use Amazon ElastiCache.



Demo: Amazon EC2 migration tool and performance

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

Performance



Our performance journey



Our performance journey started in Q4 2018, when we launched M5 and R5 instances, optimized for Amazon ElastiCache

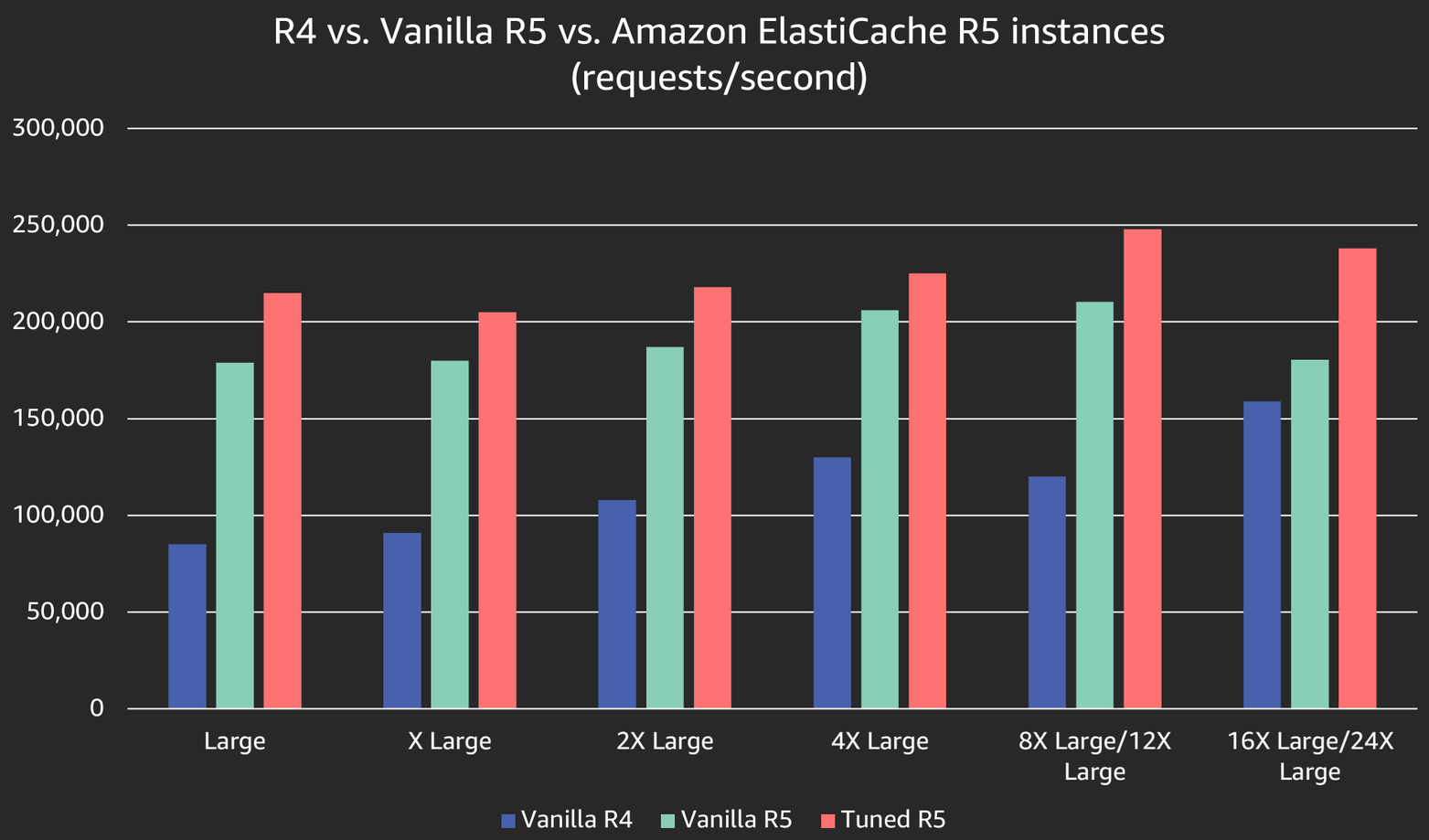


Amazon ElastiCache for Redis

We are constantly looking for ways to give our customers the best performance and the best price/performance

This includes evaluating different hardware options and optimizing them

M5 and R5 optimized instances



Scale up to **170 TB** of **in-memory capacity**

AWS **Nitro** System, dedicated **hardware** and lightweight hypervisor, delivers performance indistinguishable from **bare metal**

Custom Intel Xeon **Scalable** processors up to 3.1 GHz and AVX-512

Our performance journey - cont.



We didn’t stop there and continued to work on performance improvements



Amazon ElastiCache for Redis

Redis has single thread architecture

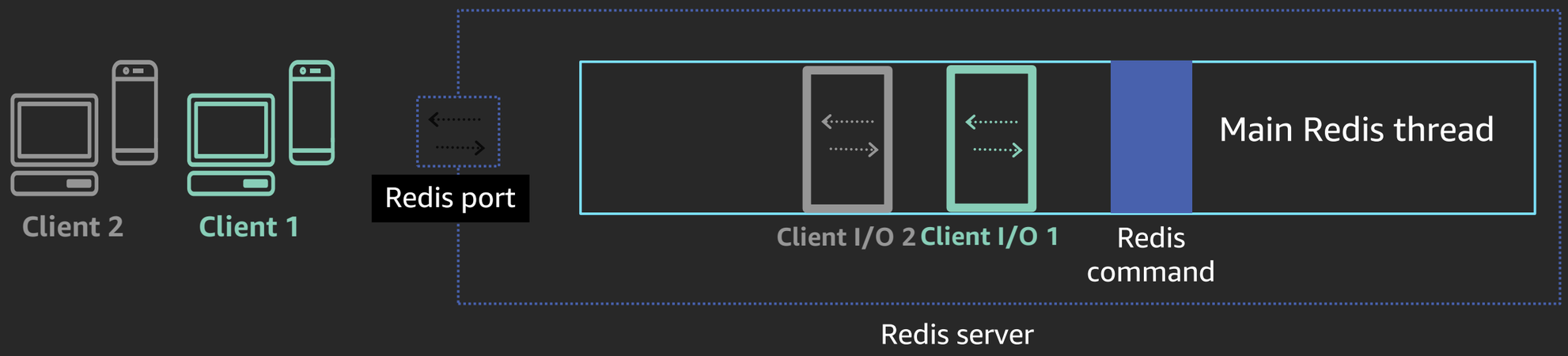


- Simplicity
- No race conditions, no synchronizations
- Easy to understand and extend
- Easy to support rich functionality
- Share-nothing architecture – scale by sharding
- Improves cache coherency

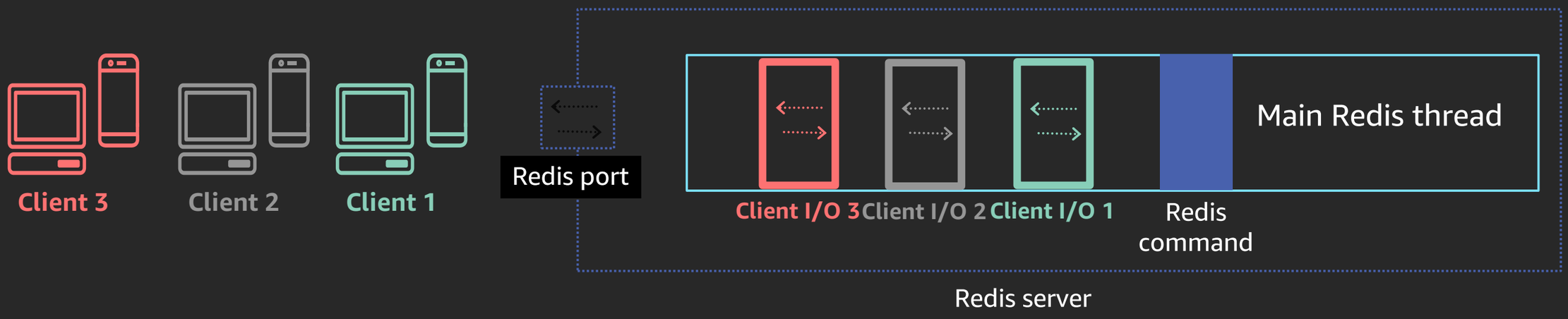
Redis connection processing: Sequential



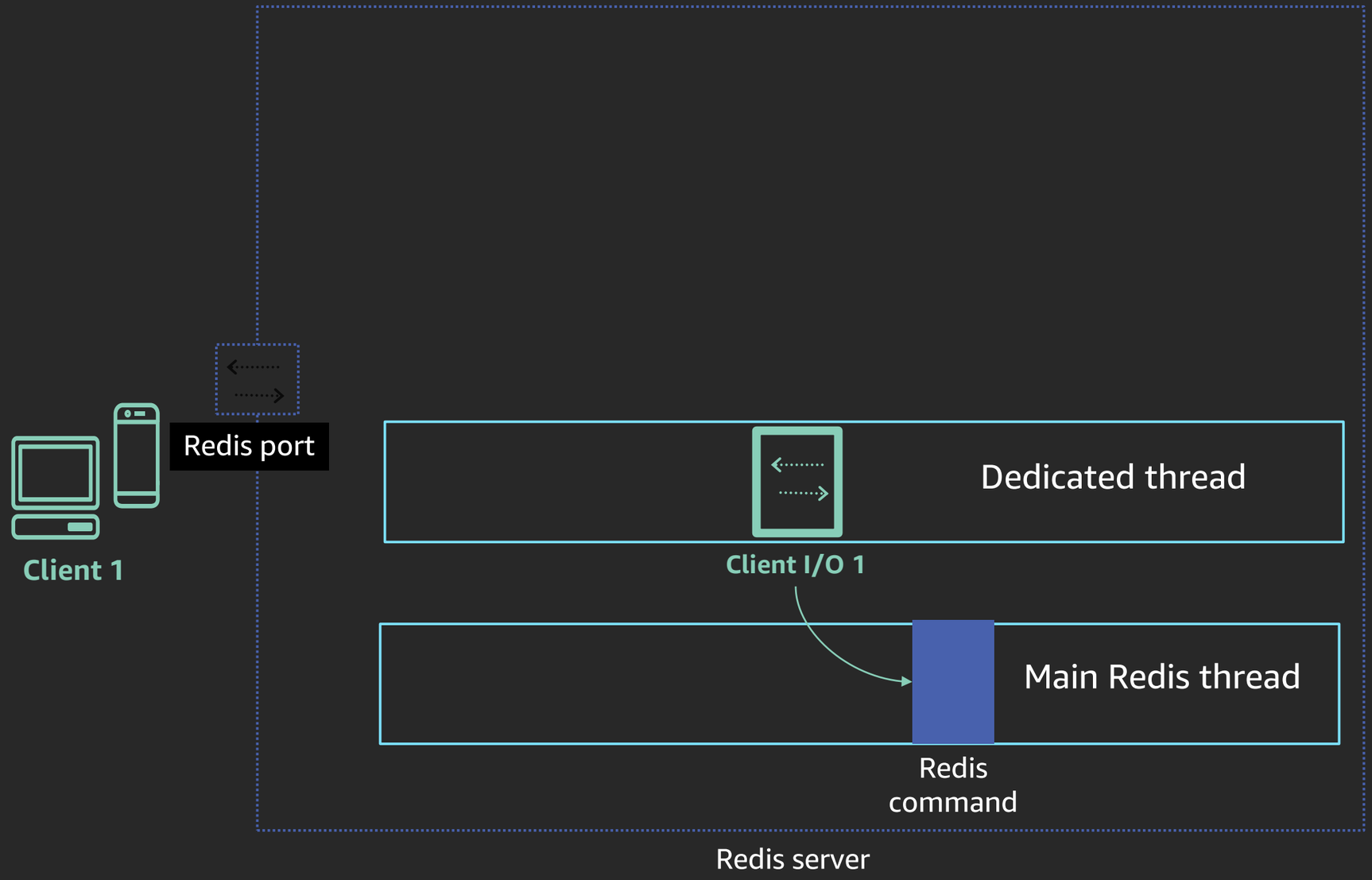
Redis connection processing: Sequential



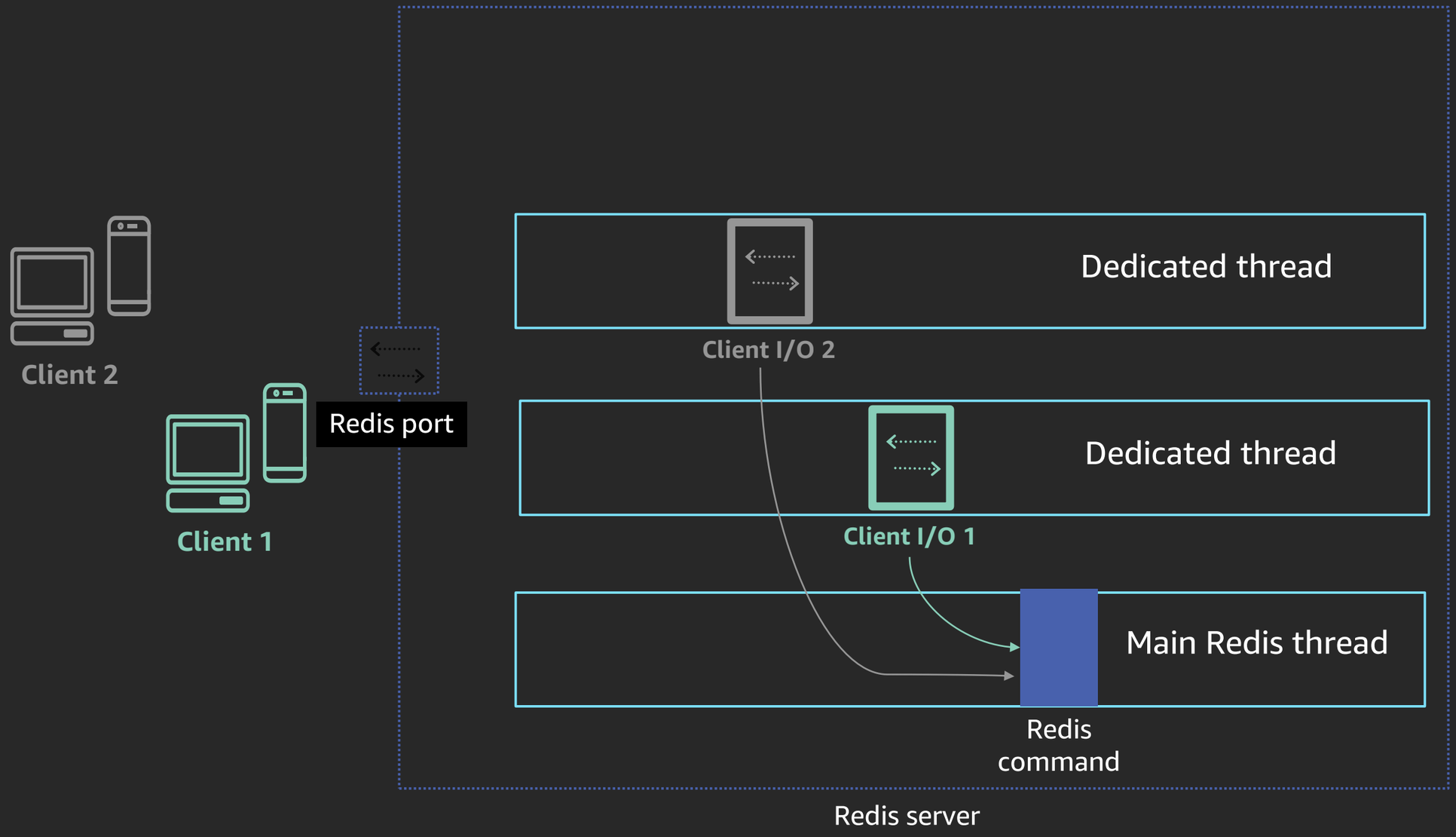
Redis connection processing: Sequential



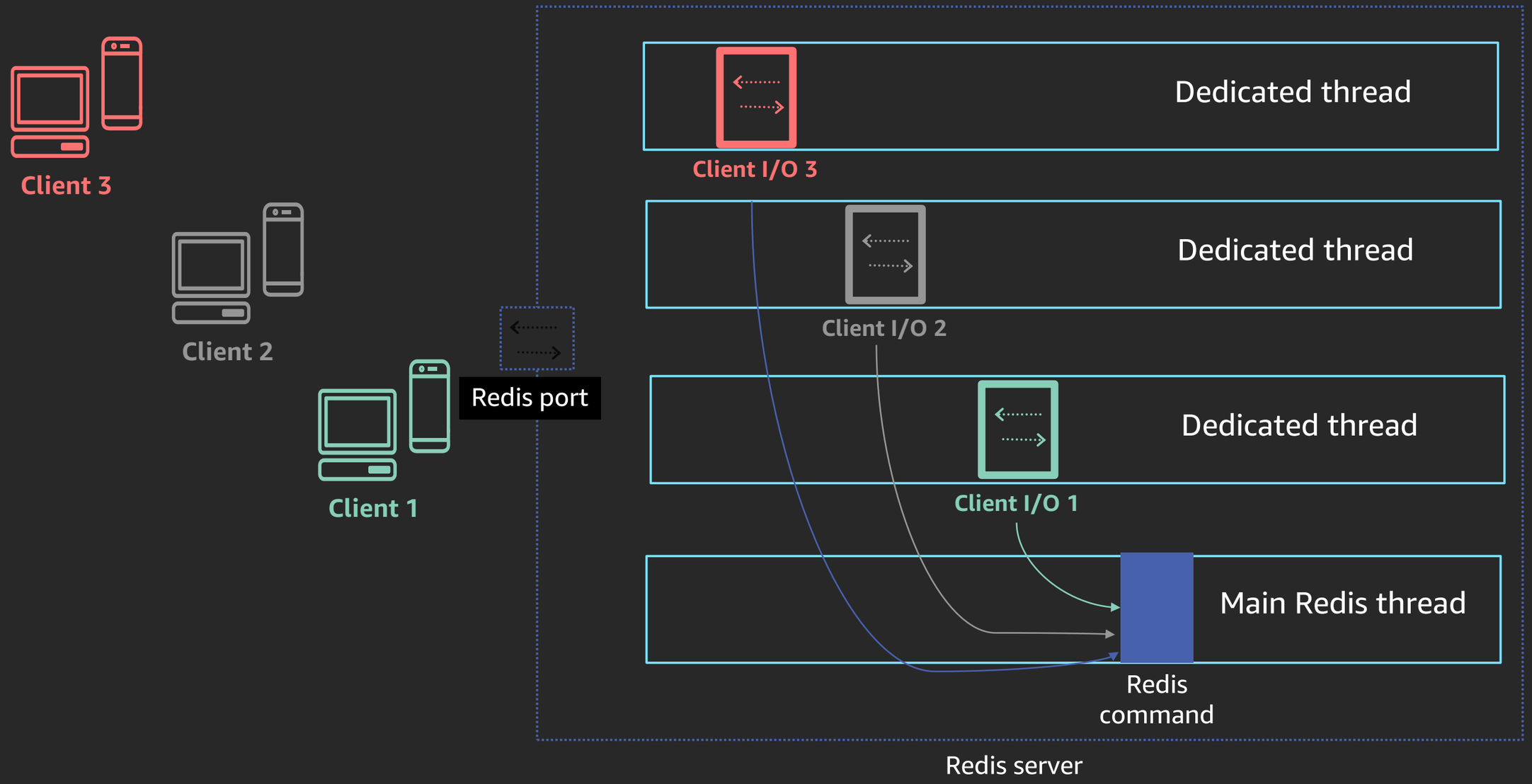
Redis connection processing: Parallel



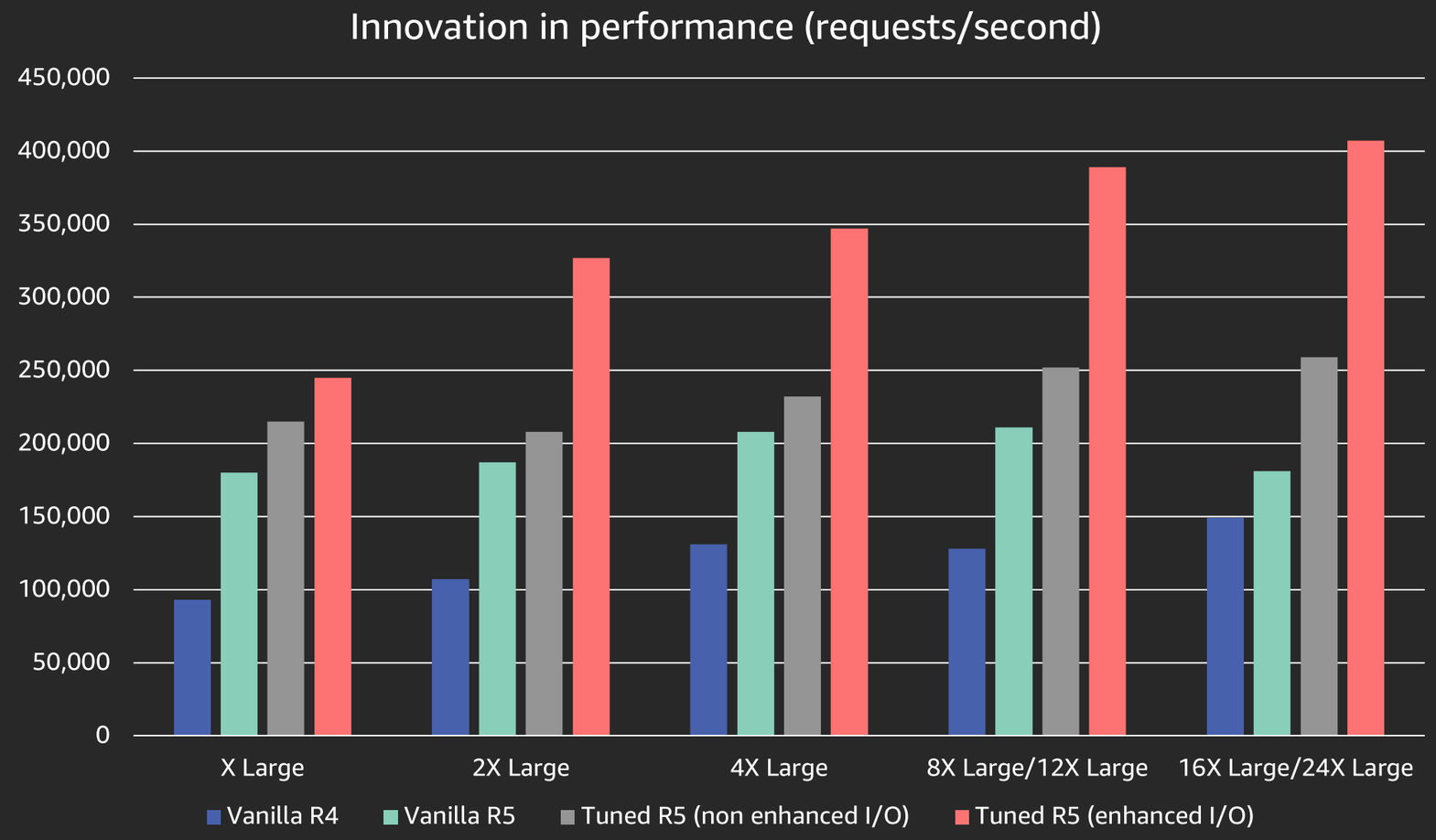
Redis connection processing: Parallel



Redis connection processing: Parallel



Optimized M5 and R5 instances + Enhanced I/O



Scale up to **170 TB** of **in-memory capacity**

AWS **Nitro** System, dedicated **hardware** and lightweight hypervisor, delivers performance indistinguishable from **bare metal**

Custom Intel Xeon **Scalable** processors up to 3.1 GHz and AVX-512

Dynamic network processing to **Enhance I/O**

T3 support



Ideal for **entry-level**, small, and medium Amazon ElastiCache workloads that may also experience temporary spikes in use

T3s have **ability to handle burst CPU usage** and can sustain **traffic spikes of 3x requests per second** vs. baseline traffic for short periods of time

Up to **36% better baseline performance and price/performance** over T2 nodes for Redis and Memcached

Online scale-up from T2 to T3

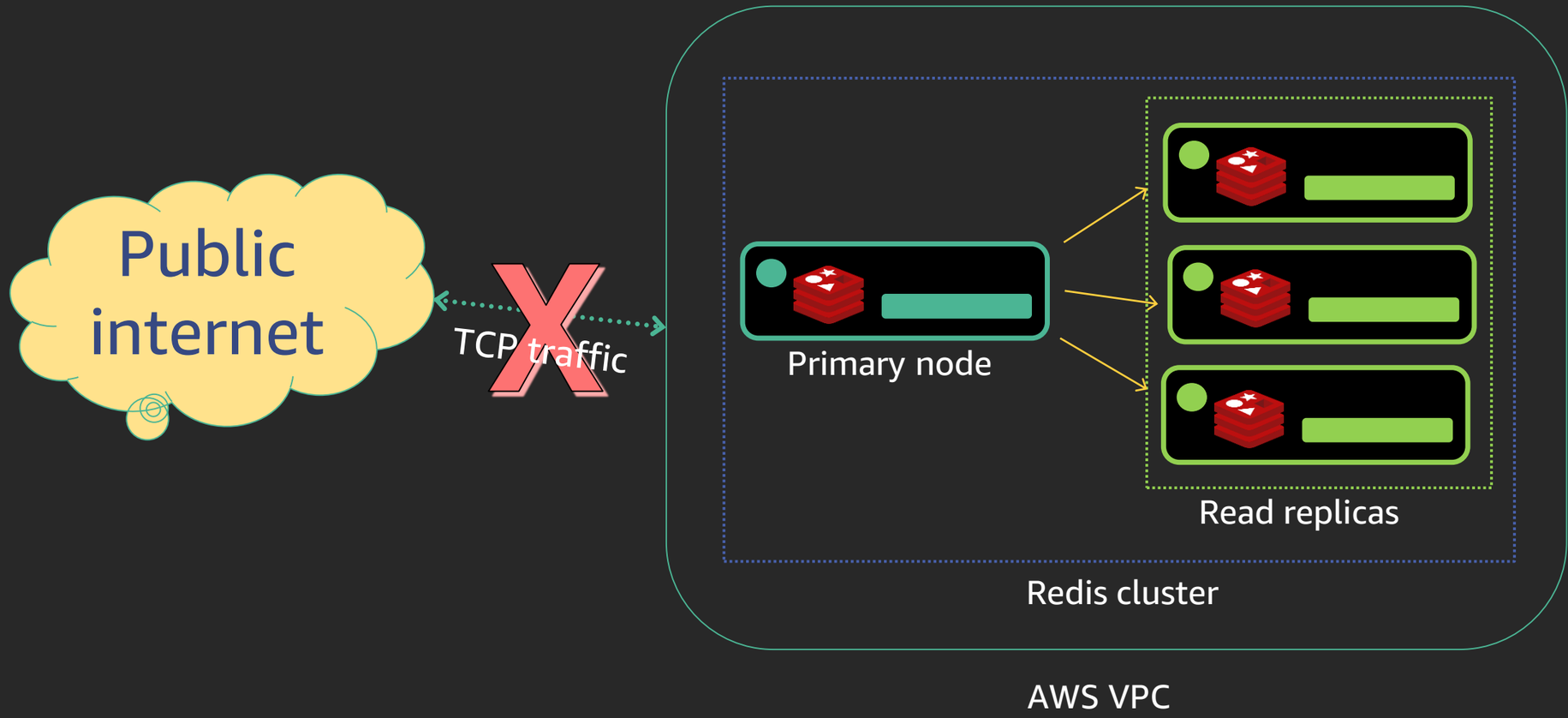
Free tier eligible; same pricing as T2

Security



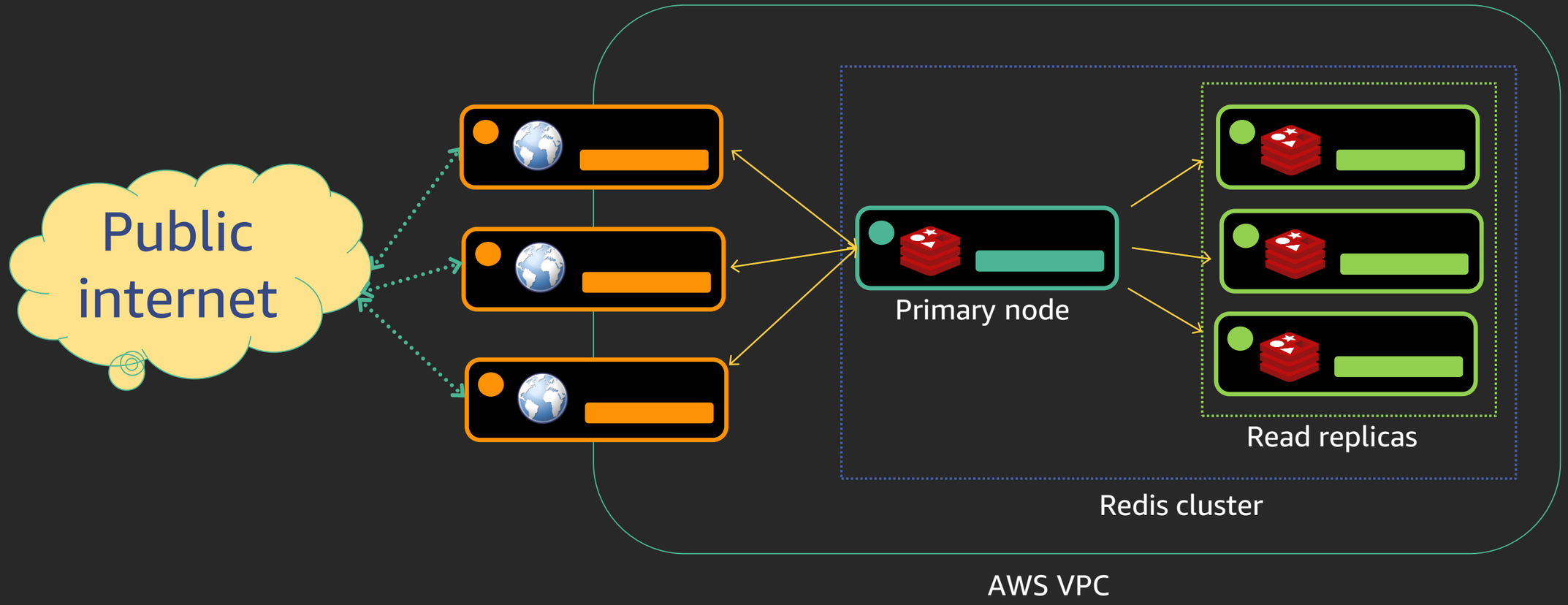
Secure by design

Amazon ElastiCache cluster default configuration



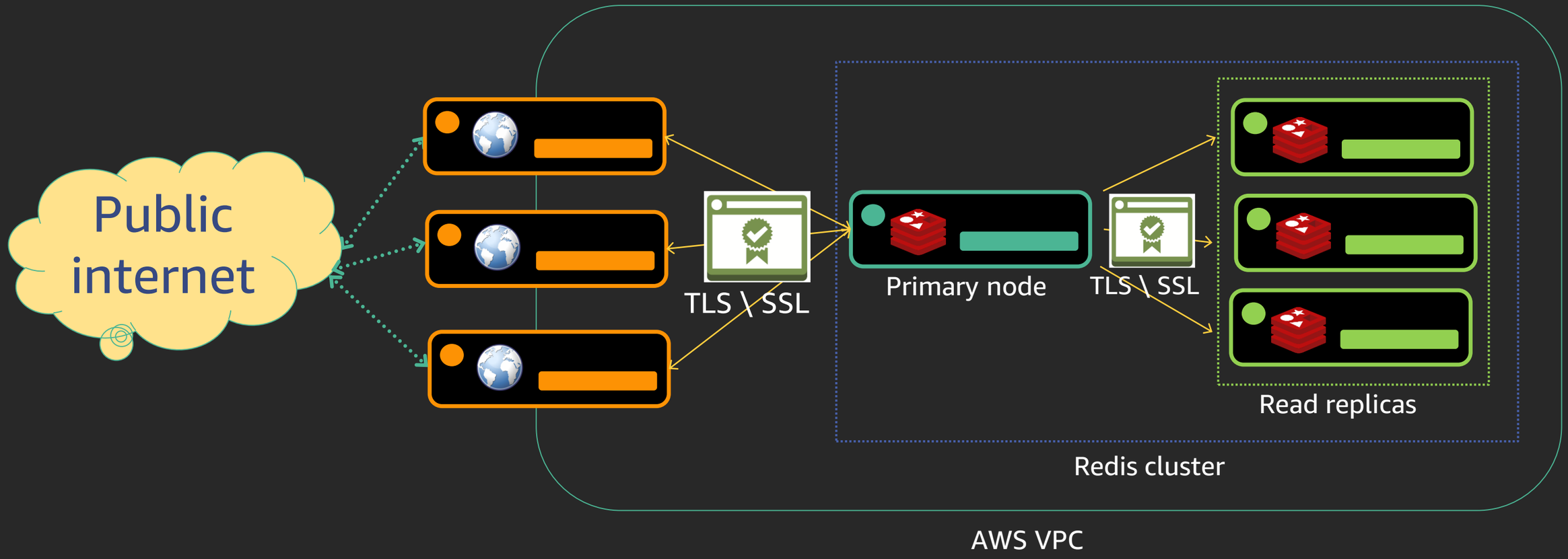
Access patterns

Allow access to the cluster through web frontends



Protect your cluster and data with TLS

Encrypt traffic in and out of your cluster



Security update: What's new

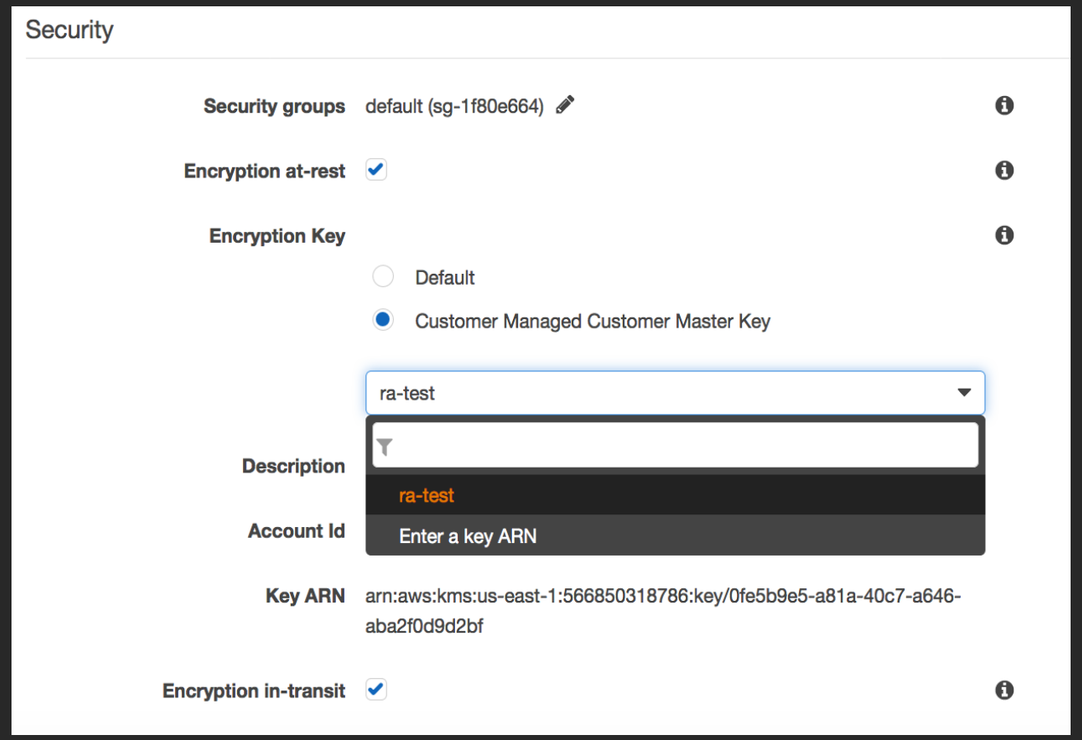
Customer managed Customer Master Keys (CMK) for encryption

Our customers had asked us for more **control** and **flexibility** to meet their security requirements

Customer managed CMKs are CMKs in **your AWS account** that **you create, own, and manage**

Encryption at rest using customer managed CMKs in **AWS Key Management Service (KMS)**

Amazon ElastiCache for Redis encrypts **all data on disk,** including service backups stored in Amazon Simple Storage Service (Amazon S3) with your encryption key



Security update: What's new

Modifying Redis Authentication Tokens



Authentication key rotation

Before this feature, authentication token was set once only during creation and can't be changed

Improve security by allowing **rotation of tokens**

Two-step process to **modify authentication tokens without interrupting clients**

Add brand-new tokens to existing clusters that were previously set up without authentication tokens

Supported on encryption in-transit enabled clusters

Redis 5.0.5 onwards

Security update: What's new

Rename Commands



Rename commands

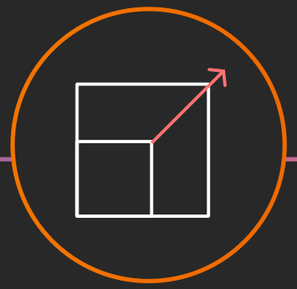
Prevent accidental misuse of potentially intrusive or expensive commands

Fully managed, online, no manual intervention, no downtime

Name changes **persist through reboots and replacements**

Redis 5.0.3 onwards

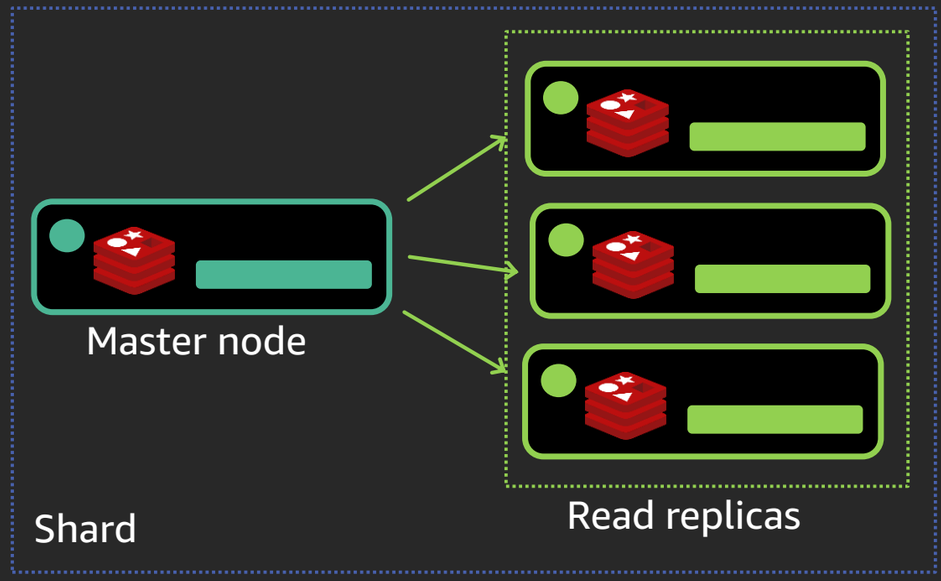
Scalability



Amazon ElastiCache scaling

Non-cluster mode enabled Redis clusters

One master node with up to 5 read replicas



Redis cluster: cluster mode disabled

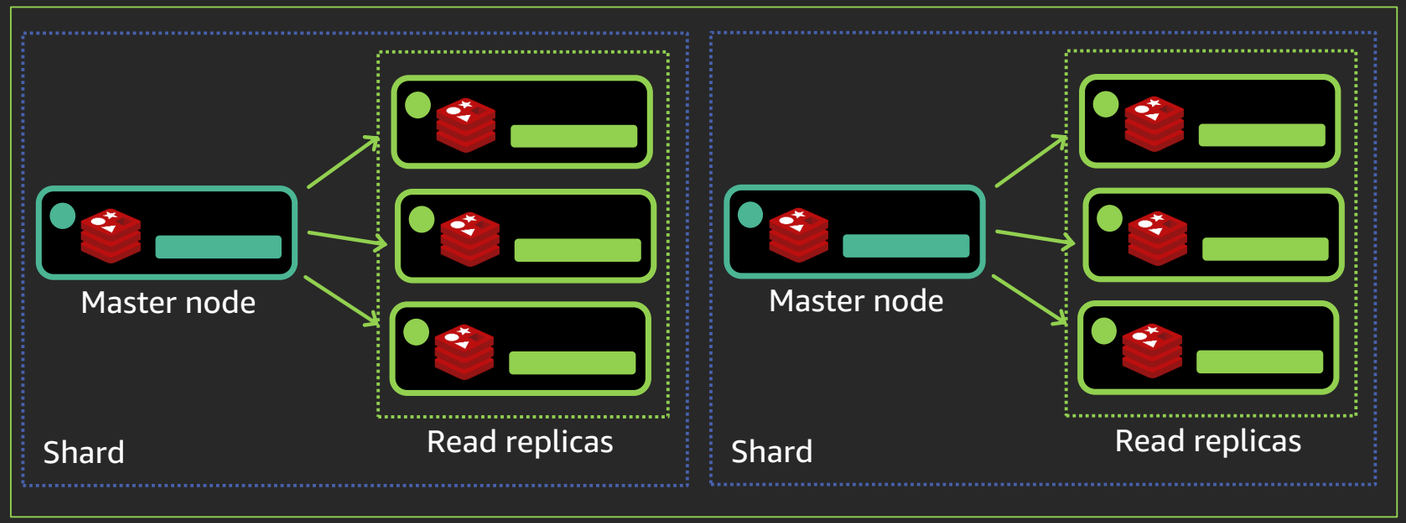
Cluster mode enabled Redis clusters

Several master nodes containing shards of the overall data in the cluster

Each **shard** can have up to 5 read replicas

A cluster can contain up to 250 nodes regardless of node type

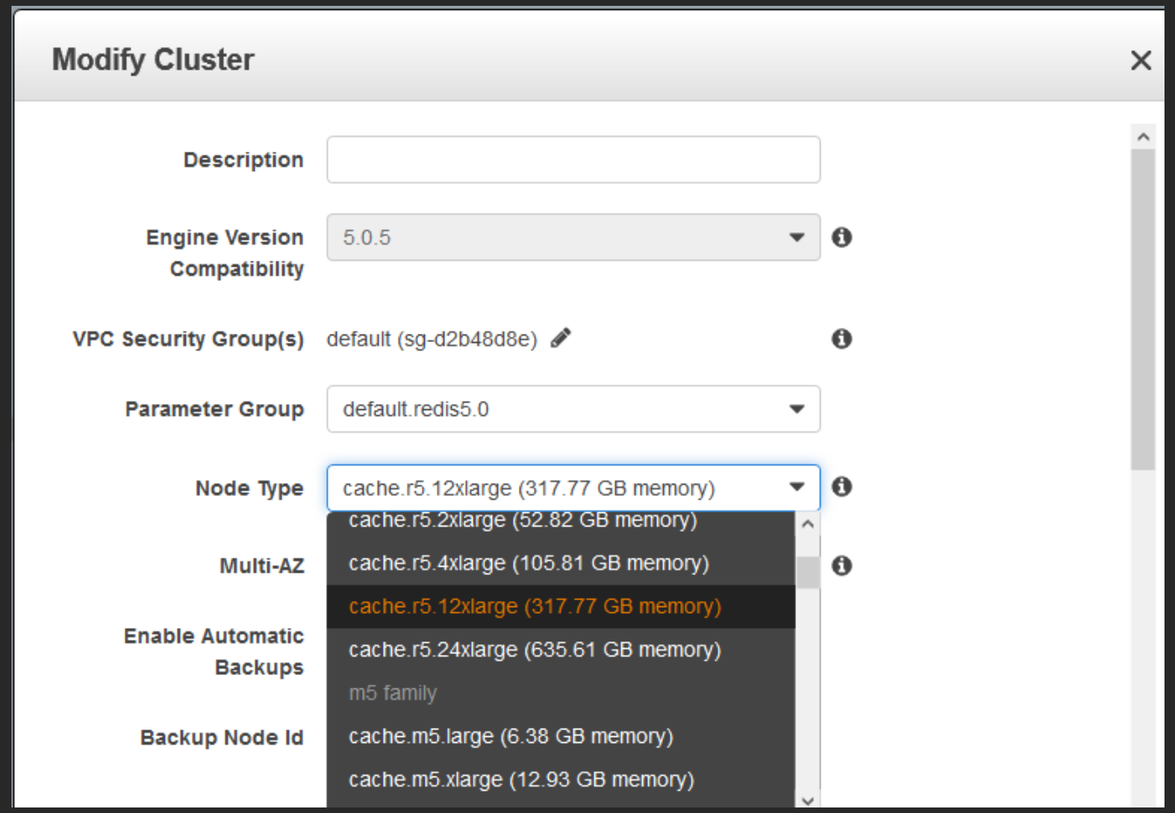
Supports up to 170 TB with online scale in/out



Redis cluster: cluster mode enabled

Scalability update: What's new

Online scale up and down

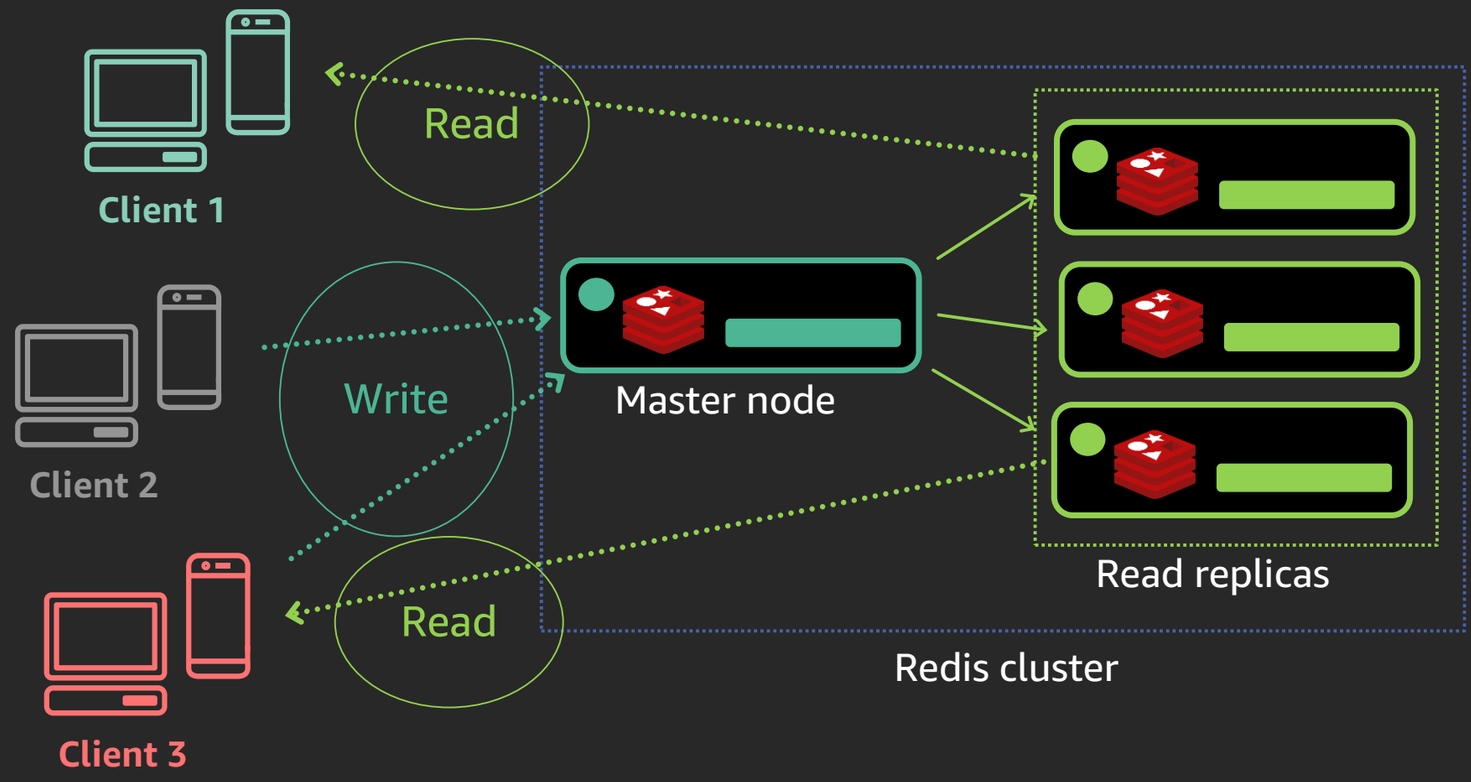


- Scale **up** to meet your peak traffic
- Scale **down** to right size and save cost
- Fully **online**, cluster continues to serve **reads and writes**
- Amazon ElastiCache for Redis 5.0.5
- Cluster Mode enabled, Cluster Mode disabled (auto failover enabled)
- All shards scale up/down

Scalability update: What's new

Reader endpoint for Cluster Mode disabled

Scaling reads to reduce the load on the master node



Scalability update: What's new

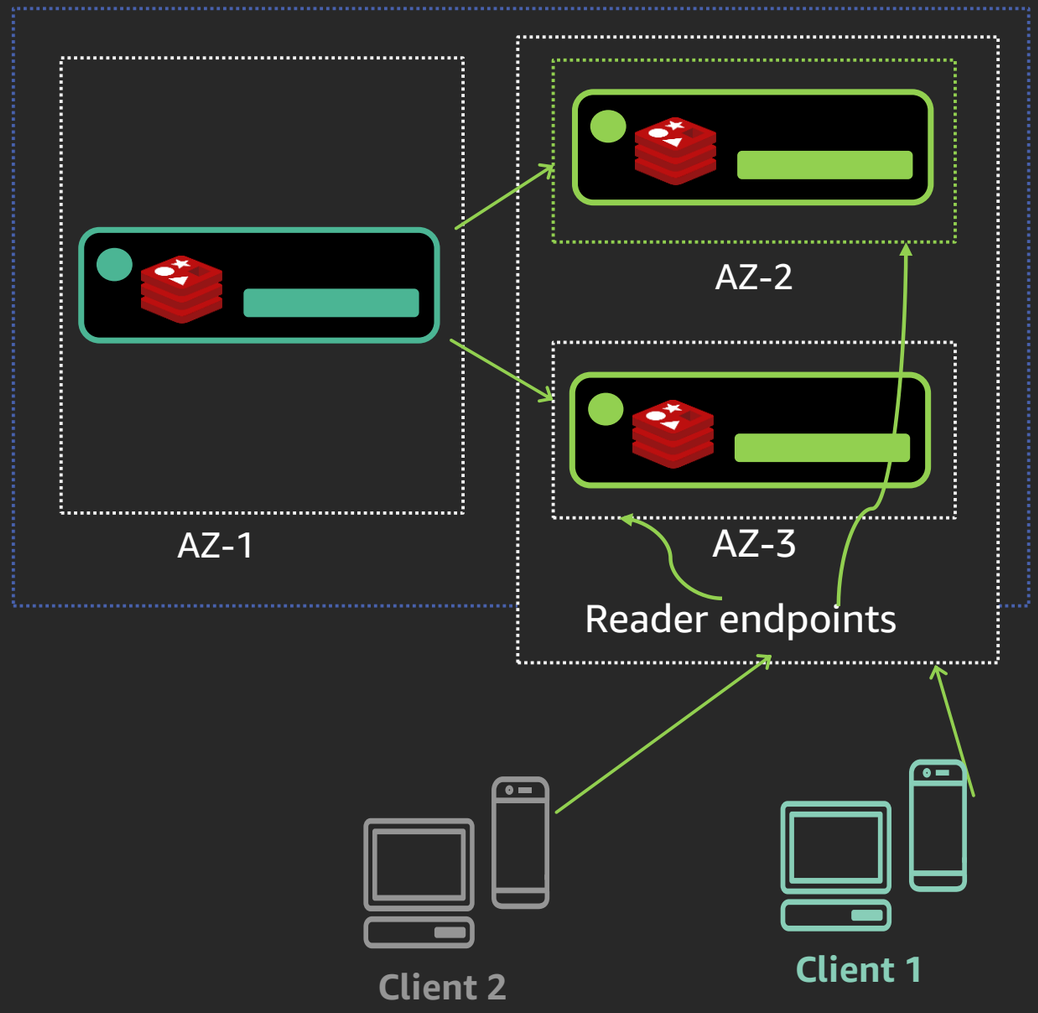
Reader endpoint

Direct all your read traffic via single reader endpoint

Divide incoming connections to the endpoint between all read replicas

Add or remove replicas **without interrupting client or making application changes**

Cluster Mode disabled



Fully managed



Self-serve update: What's new

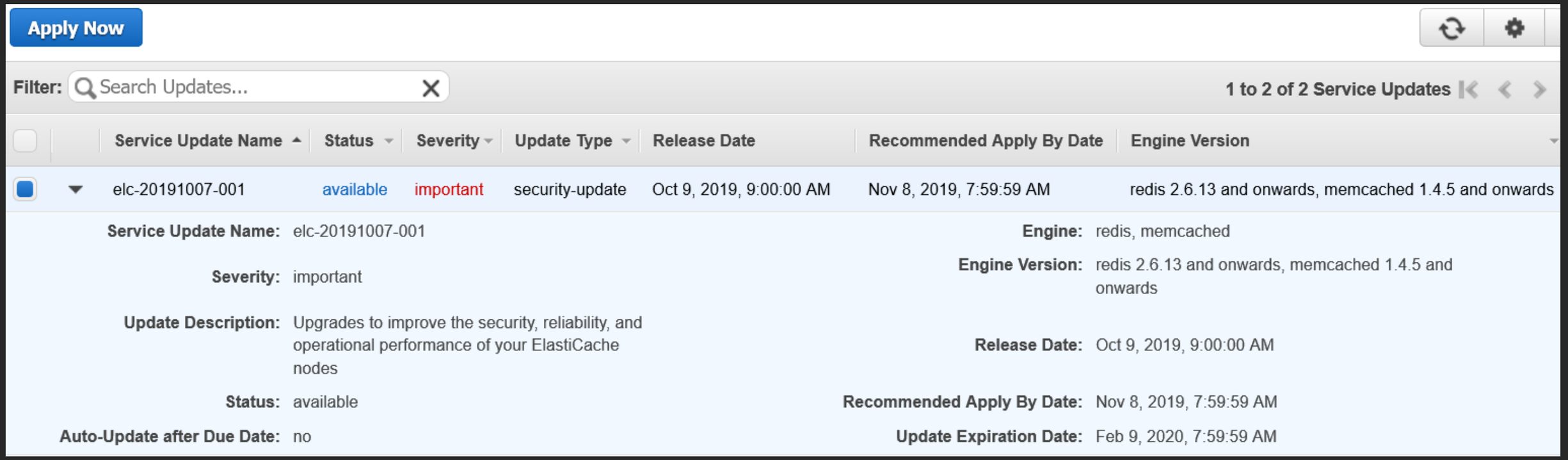
Self-service security updates

Get **notified** when service updates are available via Amazon Simple Notification Service (Amazon SNS), AWS Personal Health Dashboard, Amazon CloudWatch events, or email.

You **control** when. Apply service updates at the time of your choosing.

Track the progress in real-time.

Online updates for auto failover enabled clusters



Amazon ElastiCache for Redis: What's new - 2019

Performance



Enhanced I/O

T3 Instance Support

Security



Auth Token Rotation

Customer Managed Master Keys

Rename Command

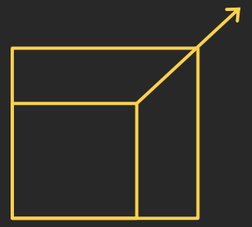
Fully Managed



Online migration from self-hosted Redis on Amazon Elastic Compute Cloud (Amazon EC2)

Self-service Maintenance updates

Scalability



Scale up / down

Reader endpoints

OSS compatibility

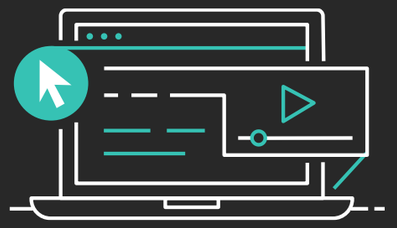


Redis 5.0.5

Memcached 1.5.16

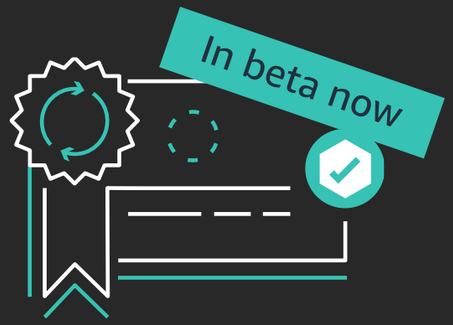
Learn databases with AWS Training and Certification

Resources created by the experts at AWS to help you build and validate database skills



25+ free digital training courses cover topics and services related to databases, including:

- Amazon Aurora
- Amazon Neptune
- Amazon DocumentDB
- Amazon DynamoDB
- Amazon ElastiCache
- Amazon Redshift
- Amazon RDS



Validate expertise with the new **AWS Certified Database - Specialty** beta exam

Visit aws.training

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws training and certification

Thank you!

Ruchita Arora: ruchitaa@amazon.com

Itay Maoz: itay@amazon.com

Nir Mashkowski: nmmashko@amazon.com

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws



Please complete the session survey in the mobile app.

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.