redis

Deborah A., Reema D., Stella L., Yingying P.

Overview

- Introduction
  - Key Concepts
  - Architecture Overview
  - Features & Methods
- Setup Options
  - Licensing
  - Pricing
  - Cloud-Based Options
  - Market Metrics
- Comparison to Other Databases
  - Advantages and Disadvantages
  - Use Cases
- Current User
- Demo

Source:
https://techcrunch.com/2018/05/14/you-can-now-try-smart-compose-in-the-new-gmail/

Question

Is it possible for a database to meet all Cap Theorem requirements: Consistency, Availability, and Partition Tolerance?

Bottom Line Upfront

Whenever data has to be written and retrieved quickly, Redis offers a good alternative to traditional relational databases.

Introduction

redis

- Remote Dictionary Server
- Open source NoSQL database management system written in C
- No Declarative Query language.Provides proprietary API for building and issuing queries.
- Supports several modules including graph modules
- It is fast due to its in-memory data persistence
- Schemaless structure
- Scales vertically and horizontally

History

2009

- Developed by Salvatore Sanfilippo
- Built it to improve the scalability by developing a real time web log analyzer

May 2013

- Sponsored by Pivotal Software

June 2015

- Sponsored by Redis Lab

October 2018

- Redis 5.0 was released introducing RediStream

Key Concepts

Claims to be ACID compliant, although it is mostly BASE complaint:

- **Atomicity**: provides transaction related commands including WATCH which ensures operations on the database are indivisible and irreducible
- **Consistency**: Only permitted writes are allowed to be performed through the validation provided by redis
- **Isolation:** Being single-threaded each command or transaction using WATCH,EXEC is thereby isolated
- **Durability**: Can be configured to respond to a client write to confirm that a write operation has been written to disk

Architecture

- Built in clustering capabilities
    - Portions of the database are shared among multiple servers
    - A server is called a node
        - Node can be a primary (master) or secondary (slave)
        - Each node is responsible for only its own set of data
    - Cluster manager is a independent layer that monitors the overall health of the cluster and the network
    - Redis runs on Two TCP Ports

- Supports both scaling vertically and scaling horizontally
- Allows database to be stored not only in RAM but also on Flash memory such as a Solid-State Drive (SSD)

Features

- Eventually consistent
- Ensures durability by using append-only file for every write operation and taking snapshots of the database
- Each node of the cluster uses a proxy to provide multi-threaded (asynchronous) communication between client and node
- Supports multi-availability zone deployments
- Does not support complex aggregates, joins or mapReduce functionalities
- Does not have in-built support for SQL queries

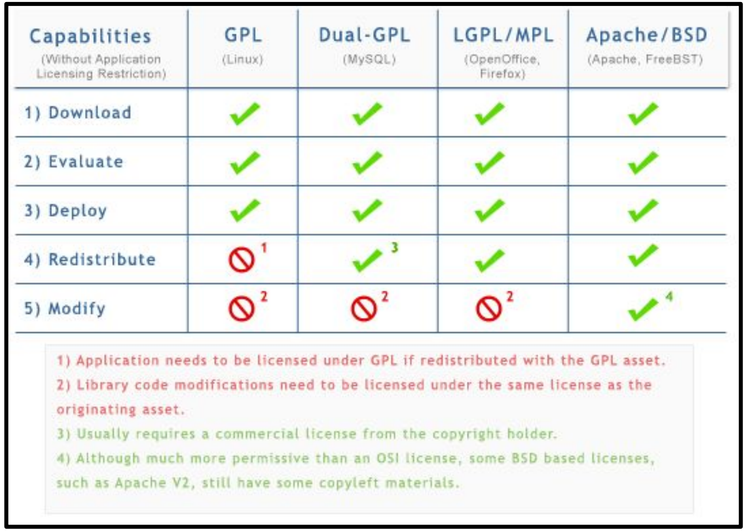
Specific Characteristics

- **Intelligent caching**: used to cache data between an application and a backend store or another relational database
- **Multi-modal nature** : data can be represented for multiple use cases simultaneously
- **Publisher/Subscriber mechanism** : fast moving data generated by thousands and millions of sensors are ingested, analyzed and processed.
- **RediSearch module** : provides features such as scoring,filtering and automatic suggestions based on search(indexes)
- **Native json handling**: JSON-formatted data can be used directly without needing to be transformed
- **Support for graph modules**

Licensing and Pricing

Open-source software released under a BSD 3-clause license

BSD licenses are a family of permissive free software licenses, imposing minimal restrictions on the use and distribution of covered software.



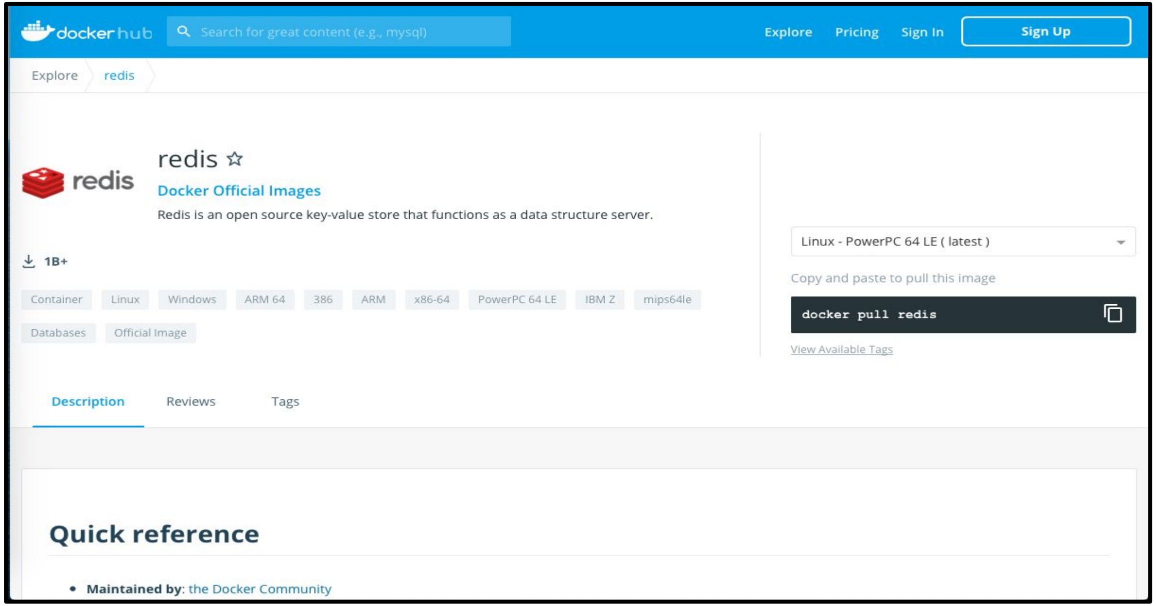
Capabilities (Without Application Licensing Restriction)	GPL (Linux)	Dual-GPL (MySQL)	LGPL/MPL (OpenOffice, Firefox)	Apache/BSD (Apache, FreeBST)
1) Download	✓	✓	✓	✓
2) Evaluate	✓	✓	✓	✓
3) Deploy	✓	✓	✓	✓
4) Redistribute	⃠[1]	✓[3]	✓	✓
5) Modify	⃠[2]	⃠[2]	⃠[2]	✓[4]

1) Application needs to be licensed under GPL if redistributed with the GPL asset.
2) Library code modifications need to be licensed under the same license as the originating asset.
3) Usually requires a commercial license from the copyright holder.
4) Although much more permissive than an OSI license, some BSD based licenses, such as Apache V2, still have some copyleft materials.

Source: http://www.bitsandbuzz.com/article/which-open-source-license/

Redis: Docker Containers

- Redis is free to download, install and use on local machines. Developers usually implement Redis through Docker containers.



redis ☆

Docker Official Images

Redis is an open source key-value store that functions as a data structure server.

```
docker pull redis
```

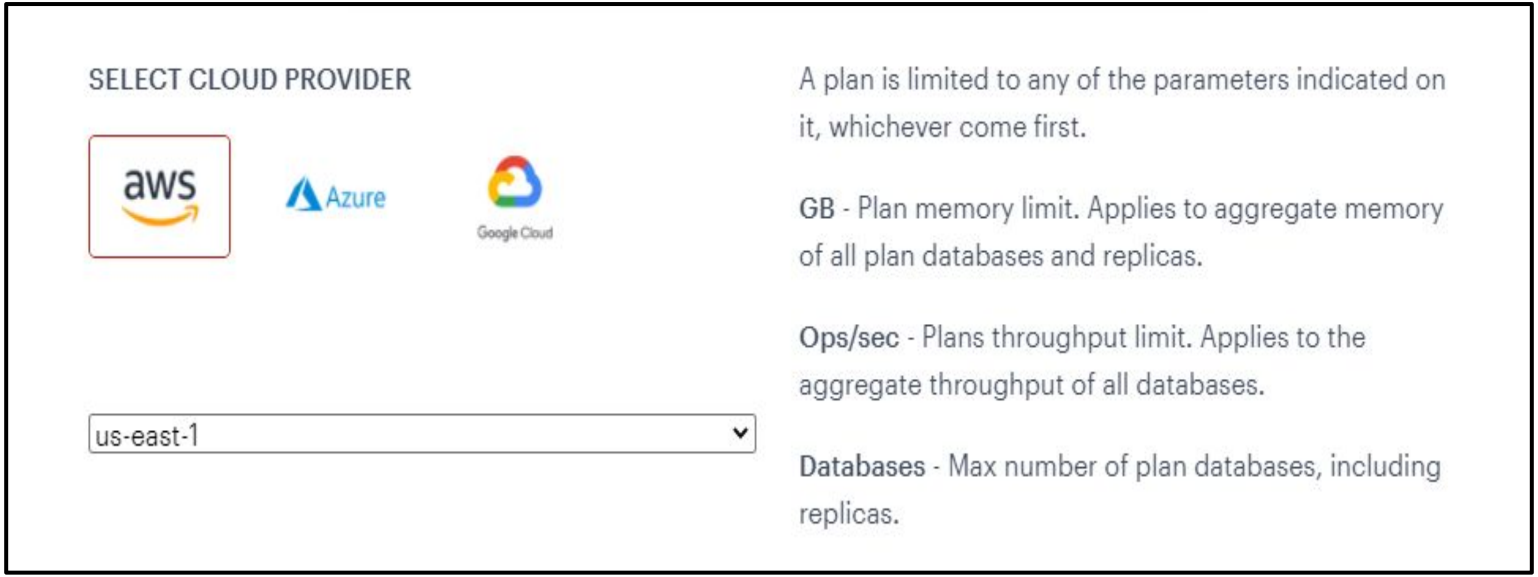
Quick reference

- **Maintained by**: the Docker Community

Source: https://hub.docker.com/_/redis

Cloud-based

- Redis can be automatically sized, scaled, upgraded, and operated on Microsoft Azure, Google Cloud, or Amazon Web Services



SELECT CLOUD PROVIDER

us-east-1

A plan is limited to any of the parameters indicated on it, whichever come first.

GB - Plan memory limit. Applies to aggregate memory of all plan databases and replicas.

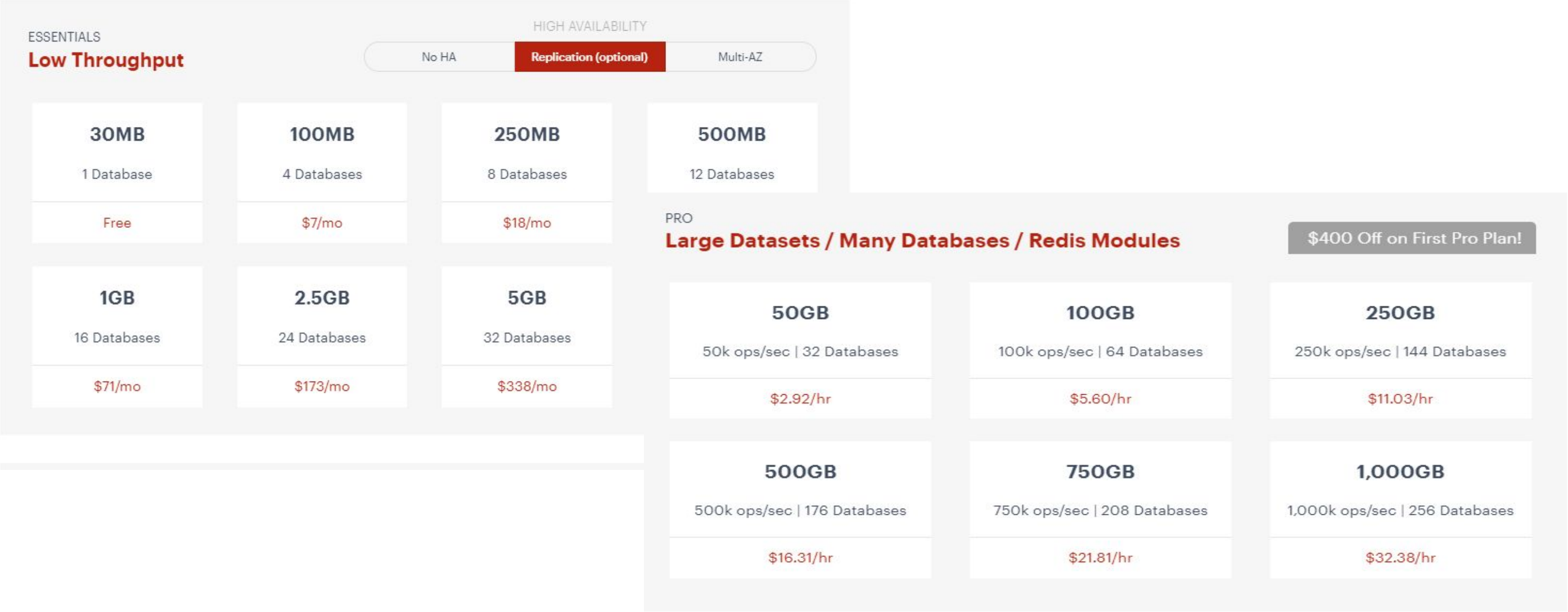
Ops/sec - Plans throughput limit. Applies to the aggregate throughput of all databases.

Databases - Max number of plan databases, including replicas.

Source: https://redislabs.com/redis-enterprise-software/overview/

Cloud Pricing

- Basic or Professional: Pay as you use



ESSENTIALS

Low Throughput

HIGH AVAILABILITY

No HA [✓ Replication (optional)] Multi-AZ

Size	Databases	Price
30MB	1 Database	Free
100MB	4 Databases	$7/mo
250MB	8 Databases	$18/mo
500MB	12 Databases	
1GB	16 Databases	$71/mo
2.5GB	24 Databases	$173/mo
5GB	32 Databases	$338/mo

PRO

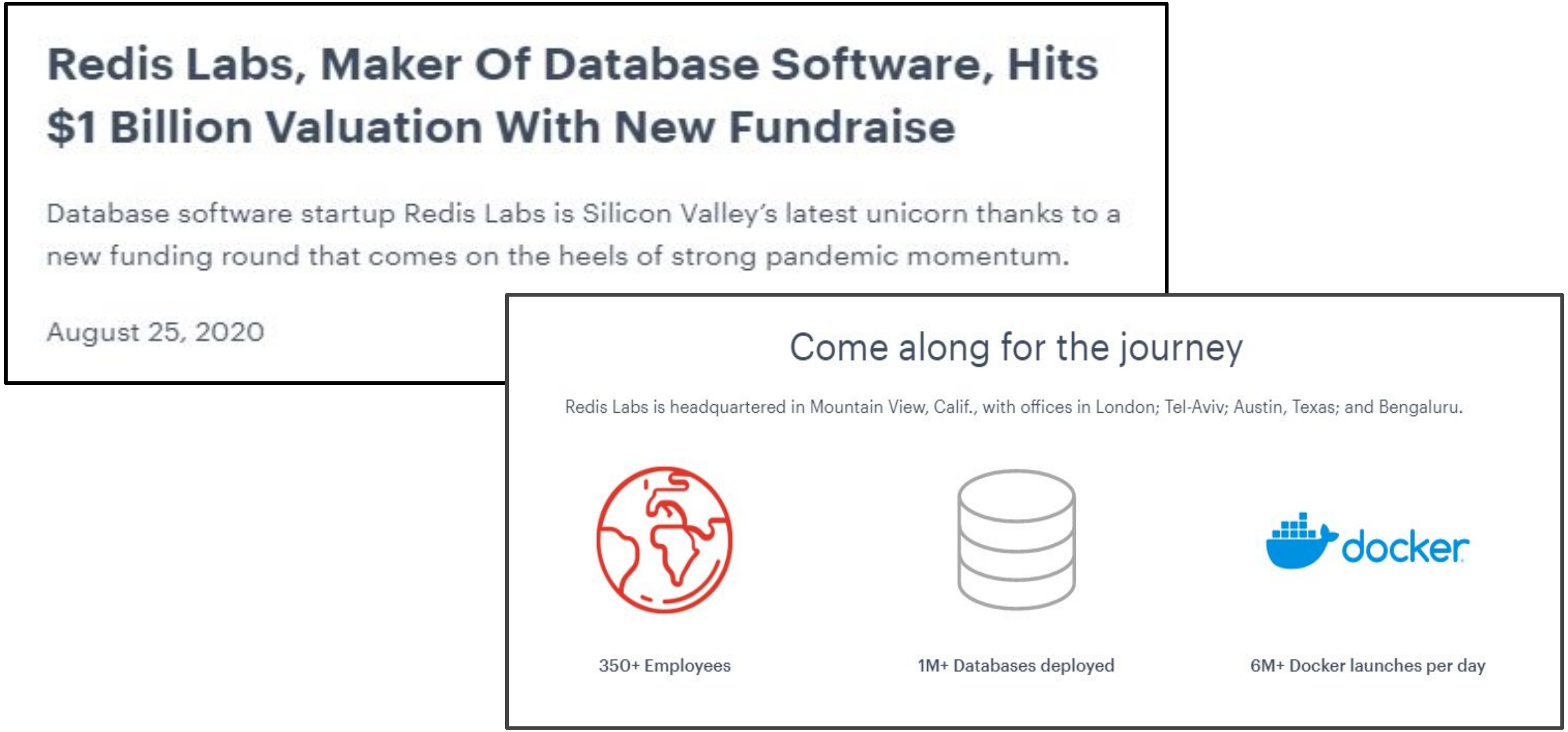
Large Datasets / Many Databases / Redis Modules

$400 Off on First Pro Plan!

Size	Throughput / Databases	Price
50GB	50k ops/sec | 32 Databases	$2.92/hr
100GB	100k ops/sec | 64 Databases	$5.60/hr
250GB	250k ops/sec | 144 Databases	$11.03/hr
500GB	500k ops/sec | 176 Databases	$16.31/hr
750GB	750k ops/sec | 208 Databases	$21.81/hr
1,000GB	1,000k ops/sec | 256 Databases	$32.38/hr

Source: https://redislabs.com/redis-enterprise-software/overview/

Market Metrics



Redis Labs, Maker Of Database Software, Hits $1 Billion Valuation With New Fundraise

Database software startup Redis Labs is Silicon Valley's latest unicorn thanks to a new funding round that comes on the heels of strong pandemic momentum.

August 25, 2020

Come along for the journey

Redis Labs is headquartered in Mountain View, Calif., with offices in London; Tel-Aviv; Austin, Texas; and Bengaluru.

350+ Employees

1M+ Databases deployed

6M+ Docker launches per day

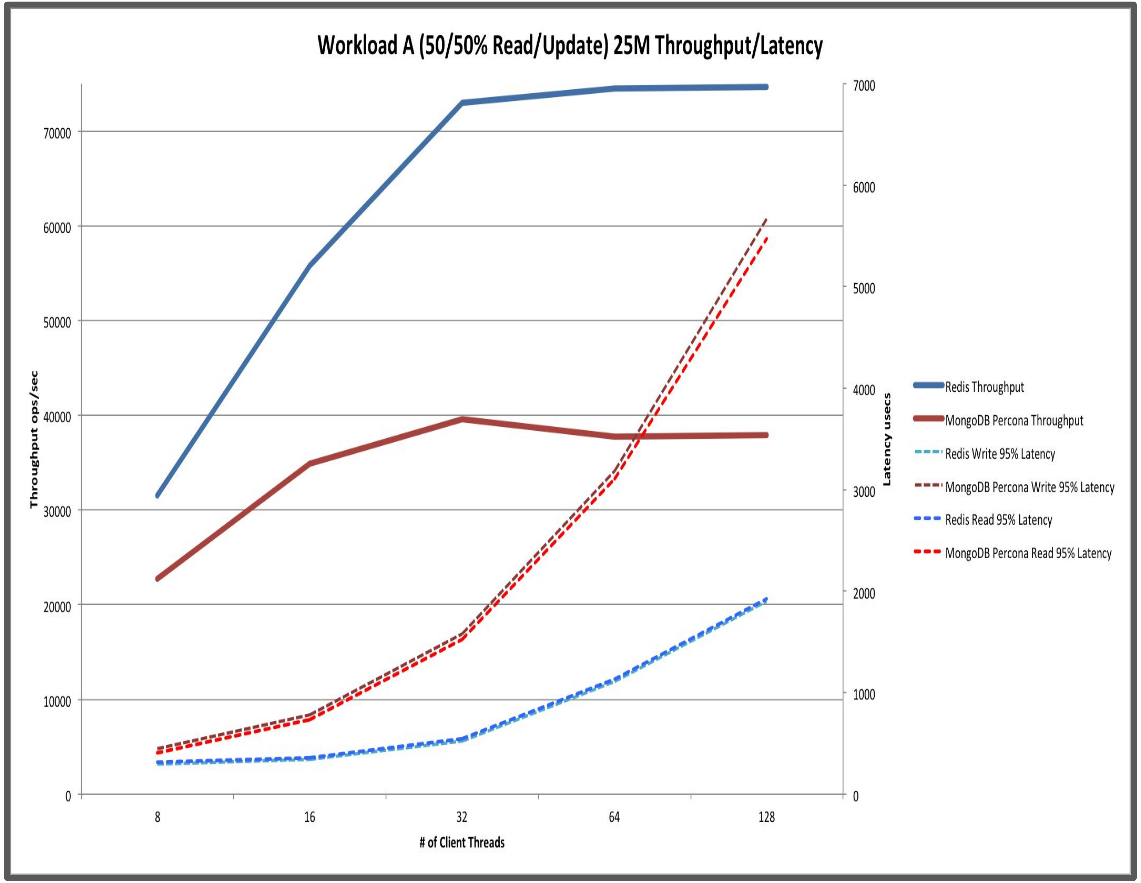
Source: https://redislabs.com/redis-enterprise-software/overview/

Comparison

How does Redis compare to other databases?

Capabilities	Redis	MongoDB	Neo4J	MYSQL
Operational Database				✓
Graph Database	✓		✓	
Document Database	✓	✓		
MapReduce	No	yes	No	No
Shema-full, schema-less. Schema-mix	schema-less	schema-less	schema-free and schema-optional	Shema-full
SQL	No	Read-only SQL queries via the MongoDB Connector for BI	No	Yes
APIs	proprietary protocol	proprietary protocol using JSON	Bolt protocol Cypher query language Java API	ADO.NET JDBC ODBC Proprietary native API
Foreign keys	No	No	No	Yes

Features	Redis	MongoDB	Neo4J	MYSQL
User and Role & Record Security	Simple password-based access control	Access rights for users and roles	Users, roles and permissions. Pluggable authentication with supported standards (LDAP, Active Directory, Kerberos)	Users with fine-grained authorization concept
Consistency concepts	Eventual Consistency Strong eventual consistency with CRDTs	Eventual Consistency Immediate Consistency	Causal and Eventual Consistency configurable in Causal Cluster setup Immediate Consistency in stand-alone mode	Immediate Consistency
ACID Transaction	Optimistic locking, atomic execution of commands blocks and scripts	Multi-document ACID Transactions with snapshot isolation	ACID	ACID
Durability	yes	yes	yes	yes
Sharding	✓	✓	None	horizontal partitioning,（sharding with MySQL Cluster or MySQL Fabric）



- For MongoDB, CPU was saturated by 32 threads onwards. Greater than 300% usage with single-digit idle percentages.
- For Redis, CPU utilization never crossed 95%.
- For Redis, by 64 threads and above, runs failed often with read-timeout exceptions.

Advantages

Redis is always a good choice if you need simple data to be available quickly. Its standard data backups features allowed an increased **read speed**: data can be retrieved from different instances. **Scaling** also works well when using Redis, both horizontally and vertically. If the Redis servers memory is insufficient, the software has built-in virtual memory management. This stores data on the hard drive. Redis itself is written in C, but there are clients for almost every programming language.

Advantages of Redis as a Cache

- A well-known caching solution that excels at it.
- Redis isn't a plain cache solution – it has an advanced data structures that provide many powerful ways to save and query data that can't be achieved with a vanilla key-value cache.
- Redis is fairly simple to set up, use and learn.
- Redis provides persistence that you can opt to set up, so cache warming in the event of a crash is hassle free.

Disadvantages

If you are working with **complex data** which requires equally extensive query options, Redis is not the right choice. Generally, data in a key-value store can only be accessed via its keys. In addition, Redis (like other in-memory databases) requires an **extremely large amount of memory** which can be expensive. You should also expect to provide more memory than the data actually uses.

- It doesn’t have inbuilt encryption on the wire.
- No role-based account control (RBAC).
- There isn’t a seamless, mature clustering solution.
- Can be a pain to deploy in large-scale cloud deployments.

Use Cases vs Not Use Cases

Redis will be a better fit for:

- Non-structured data
- Simple and quick lookups.
- Caching layer
- Improves performance for scalability issues

Redis will not be a better fit for:

- Data will not be frequently queried
- Write operation is not much
- The amount of data is small

Current Users

Where is redis currently being used?

Key Industries/Customers

craigslist
stackoverflow
新浪微博
weibo.com
digg
GitHub
flickr

Case Study: HealthStream

- HealthStream is dedicated to providing development training to healthcare professionals, such as:
    - training and learning management
    - talent management
    - performance assessment
    - credentialing
    - simulation-based training programs

Case Study 1: HealthStream Benefits/Challenges

- **Challenge:** Needs an easy-to-deploy, blazing-fast datastore that could enhance user responsiveness with low operational overhead and high availability

- **Solution**: HealthStream selected Redis Enterprise because it's demonstrated superb performance, high availability, and operational simplicity, in powering a rollout of new courses.

Case Study 2: Coffee Meets Bagel

- Coffee Meets Bagel differentiates itself from other dating apps in that by limiting the number of highly curated matches to its users every day at noon.

- **Challenge:** The platform has attracted 3 million users around the world since its launch in 2012
    - Company is hoping to expand its international presence and processing upwards of 1TB of data an hour

Case Study 2: Coffee Meets Bagel Benefits

- Automated scaling, failover, cluster creation, persistence, and high availability helps Coffee Meets Bagel focus on application innovation

- 24/7 support, automated operations, significant cost savings on cloud spend, and cloud-agnostic consistency

- auto-scaling capabilities have eliminated days of manual effort for scaling during high-traffic times

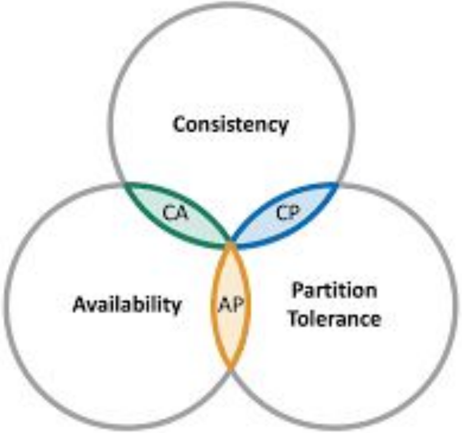
- Non-traditional database structure

CoffeeMeetsBagel

Conclusion

redis is:

- A multi-modal database system, with key-value at its core
- Performs really well for read operations, scales fast and initial cost of setting up is low
- Has no single point of failure as the database is distributed across the cluster



References

"The Best Redis Experience." *Redis Labs*, Redis Labs, 19 Aug. 2020, redislabs.com/.

"Redis for Dummies". John Wiley & Sons Limited Edition 2019

"Wikipedia". https://en.wikipedia.org/wiki/Redis

https://www.ionos.com/digitalguide/hosting/technical-matters/in-memory-databases/

https://www.ionos.com/digitalguide/online-marketing/web-analytics/dwh-what-is-a-data-warehouse/#c91531

Demo