Công ty An ninh mạng Viettel

Báo cáo về Redis

Sinh viên: Nguyễn Đan Trường

*MỤC LỤC*

[Chương 1. Tổng quan về Redis 3](#_Toc160539747)

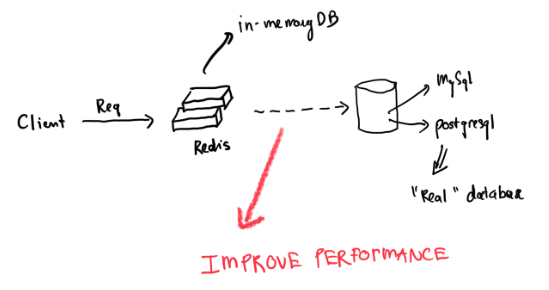
# Tổng quan về Redis

Redis là một in-memory database, lưu trữ dữ liệu dưới dạng key - value.

Redis thường được "đặt phía trước" main database (MySQL, PostgreSQL) để cải thiện hiệu năng hệ thống.

Dữ liệu thường được lưu trong redis là:

* Dữ liệu ít thay đổi nhưng được truy xuất thường xuyên
* Dữ liệu ít quan trọng và vẫn trong quá trình phát triển.



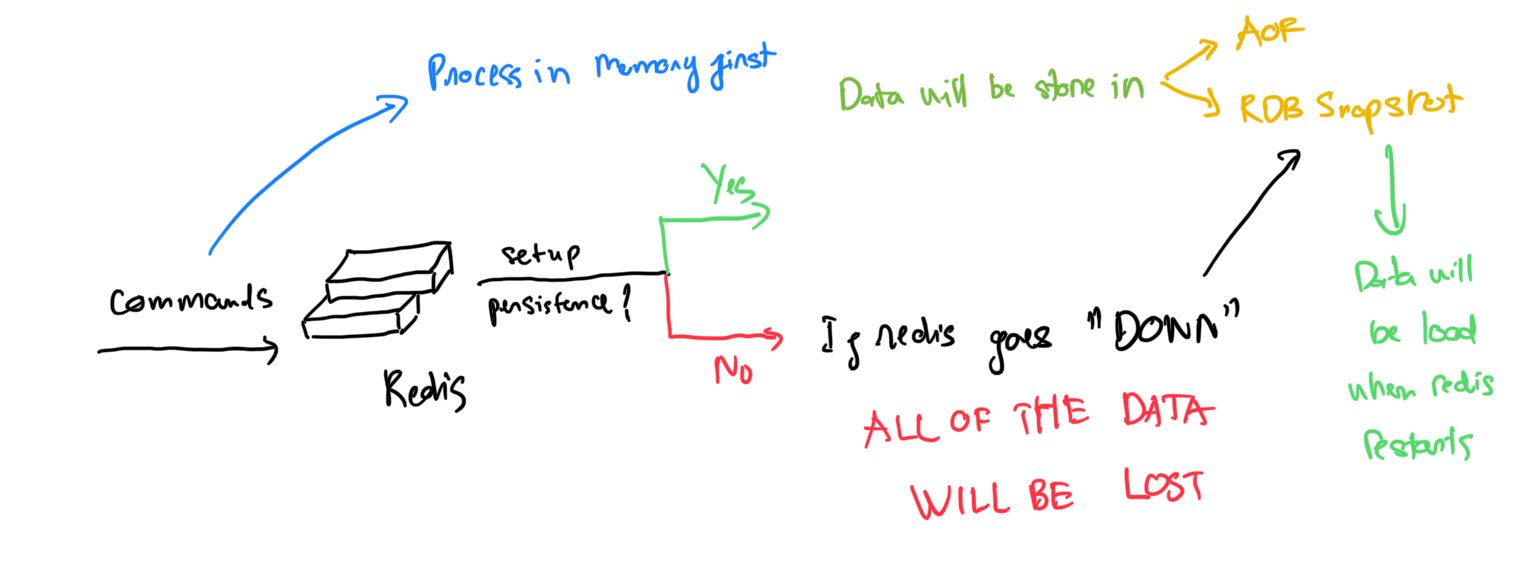
## Redis Architectures

**Single Redis Instance**

Đơn thuần là chỉ sử dụng duy nhất 1 Redis instance hoặc cluster. Cách làm này rất dễ triển khai xong nhược điểm dễ thấy nhất đó là nếu instance duy nhất này bị "sập" thì client sẽ không thể truy cập vào cache dẫn tới hiệu năng của hệ thống bị ảnh hưởng.

Các lệnh gửi đến redis sẽ được xử lí ở trong memory. Nếu setup persistence cho redis thì sẽ có các tiến trình con chạy trong những khoảng thời gian cố định để lưu dữ liệu của redis vào RDB snapshot hoặc AOF (append-only files)

Nếu không lưu dữ liệu vào snapshot hoặc AOF thì khi redis "sập" toàn bộ dữ liệu sẽ mất sạch, ngược lại nếu lưu thì khi redis restarts thì dữ liệu sẽ được phục hồi trở lại.



**RDB**

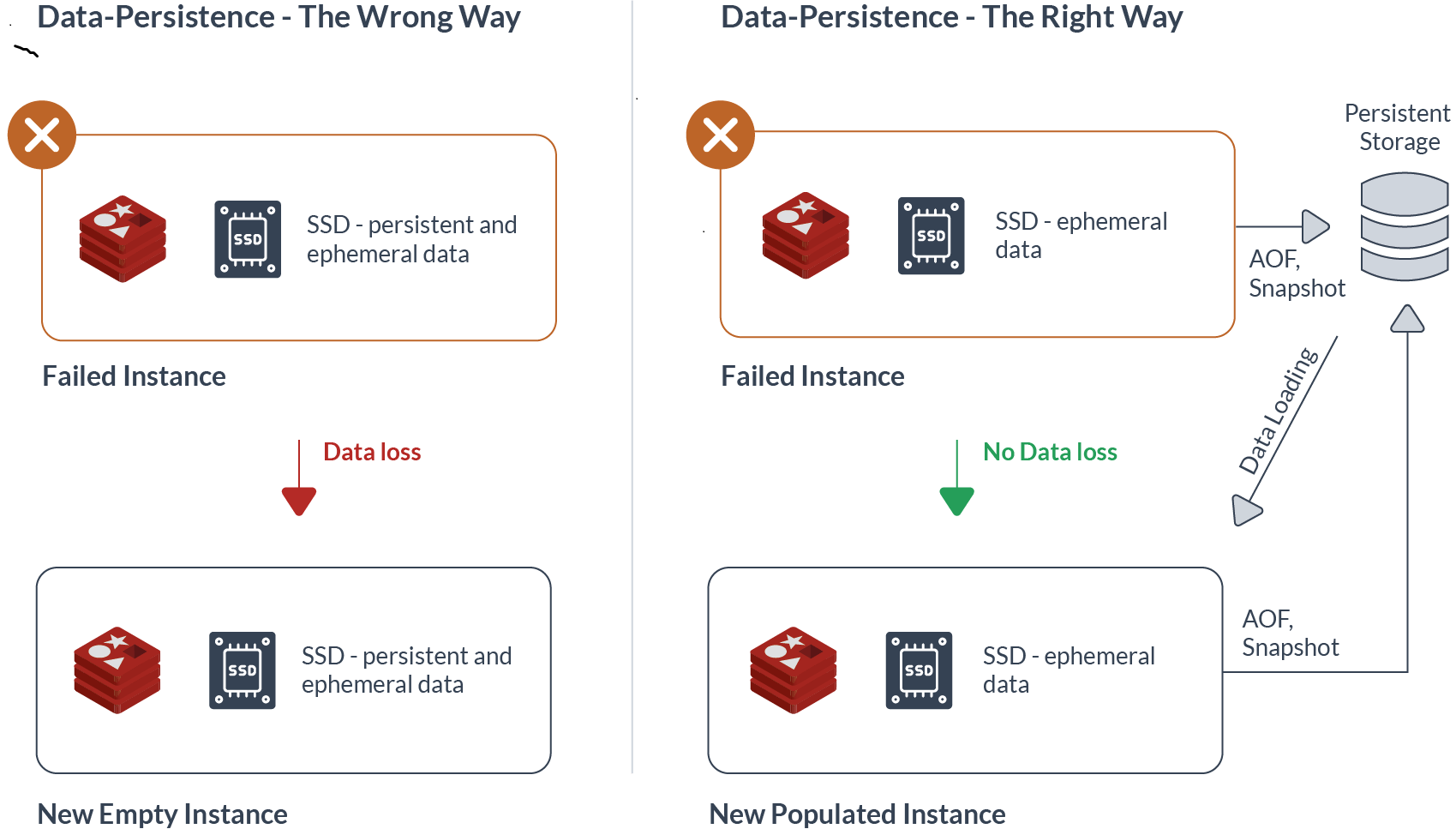
RDB thực hiện tạo và sao lưu snapshot của DB vào ổ cứng sau mỗi khoảng thời gian nhất định.

Ưu điểm:

* RDB cho phép người dùng lưu các version khác nhau của DB, rất thuận tiện khi có sự cố xảy ra.
* Bằng việc lưu trữ data vào 1 file cố định, người dùng có thể dễ dàng chuyển data đến các data centers khác nhau, hoặc chuyển đến lưu trữ trên Amazon S3.
* RDB giúp tối ưu hóa hiệu năng của Redis. Tiến trình Redis chính sẽ chỉ làm các công việc trên RAM, bao gồm các thao tác cơ bản được yêu cầu từ phía client như thêm/đọc/xóa, trong khi đó 1 tiến trình con sẽ đảm nhiệm các thao tác disk I/O. Cách tổ chức này giúp tối đa hiệu năng của Redis. Khi restart server, dùng RDB làm việc với lượng data lớn sẽ có tốc độ cao hơn là dùng AOF.

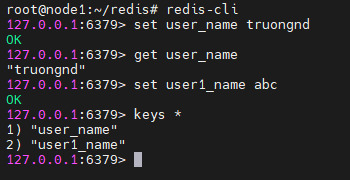
Nhược điểm:

* RDB không phải là lựa chọn tốt nếu bạn muốn giảm thiểu tối đa nguy cơ mất mát dữ liệu. Thông thường người dùng sẽ set up để tạo RDB snapshot 5 phút 1 lần (hoặc nhiều hơn). Do vậy, trong trường hợp có sự cố, Redis không thể hoạt động, dữ liệu trong những phút cuối sẽ bị mất.
* RDB cần dùng fork() để tạo tiến trình con phục vụ cho thao tác disk I/O. Trong trường hợp dữ liệu quá lớn, quá trình fork() có thể tốn thời gian và server sẽ không thể đáp ứng được request từ client trong vài milisecond hoặc thậm chí là 1 second tùy thuộc vào lượng data và hiệu năng CPU.



Như minh họa ở trên, trong trường hợp không có bản sao của tập dữ liệu còn lại trong RAM, Redis Enterprise sẽ tìm bản sao cuối cùng của tập dữ liệu trong các thiết bị được kết nối mạng được kết nối với nút không thành công và sử dụng nó để điền phân đoạn Redis trên phiên bản đám mây mới.

Sử dụng công cụ redis-cli và tạo key:

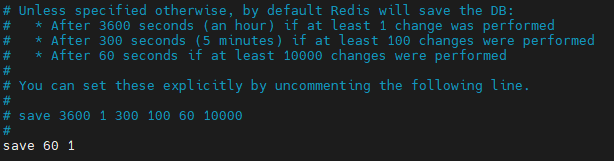


Đường dẫn thư mục làm việc, lưu trữ file rdb:

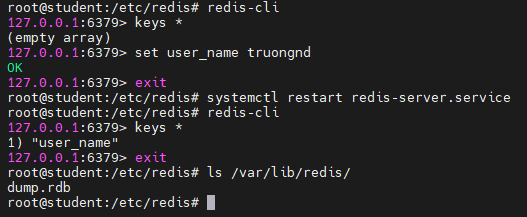




Theo mặc định, Redis snapshot dữ liệu và lưu nó vào file dump.rdp. Bạn có thể cấu hình Redis cứ sau N giây nếu có ít nhất M thay đổi trong tập dữ liệu thì sẽ snapshot hoặc có thể sử dụng lệnh SAVE hoặc BGSAVE



Ở đây Redis sẽ auto dump dữ liệu mỗi 60s nếu có ít nhất 1 key thay đổi.



**AOF (Append Only File)**

AOF lưu lại tất cả các thao tác write mà server nhận được, các thao tác này sẽ được chạy lại khi restart server hoặc tái thiết lập dataset ban đầu.

Ưu điểm

* Sử dụng AOF sẽ giúp đảm bảo dataset được bền vững hơn so với dùng RDB. Người dùng có thể config để Redis ghi log theo từng câu query hoặc mỗi giây 1 lần.
* Redis ghi log AOF theo kiểu thêm vào cuối file sẵn có, do đó tiến trình seek trên file có sẵn là không cần thiết. Ngoài ra, kể cả khi chỉ 1 nửa câu lệnh được ghi trong file log (có thể do ổ đĩa bị full), Redis vẫn có cơ chế quản lý và sửa chữa lối đó (redis-check-aof).
* Redis cung cấp tiến trình chạy nền, cho phép ghi lại file AOF khi dung lượng file quá lớn. Trong khi server vẫn thực hiện thao tác trên file cũ, 1 file hoàn toàn mới được tạo ra với số lượng tối thiểu operation phục vụ cho việc tạo dataset hiện tại. Và 1 khi file mới được ghi xong, Redis sẽ chuyển sang thực hiện thao tác ghi log trên file mới.

Nhược điểm

* File AOF thường lớn hơn file RDB với cùng 1 dataset.
* AOF có thể chậm hơn RDB tùy theo cách thức thiết lập khoảng thời gian cho việc sao lưu vào ổ cứng. Tuy nhiên, nếu thiết lập log 1 giây 1 lần có thể đạt hiệu năng tương đương với RDB.
* Developer của Redis đã từng gặp phải bug với AOF (mặc dù là rất hiếm), đó là lỗi AOF không thể tái tạo lại chính xác dataset khi restart Redis. Lỗi này chưa gặp phải khi làm việc với RDB bao giờ.

Mỗi phương thức đều có ưu/nhược điểm riêng, và có lẽ cần nhiều thời gian làm việc với Redis cũng như tùy theo ứng dụng mà đưa ra lựa chọn thích hợp. Nhiều người chọn AOF bới nó đảm bảo toàn vẹn dữ liệu rất tốt, nhưng Redis developer lại khuyến cáo nên dùng cả RDB, bởi nó rất thuận tiện cho việc backup database, tăng tốc độ cho quá trình restart cũng như tránh được bug của AOF.

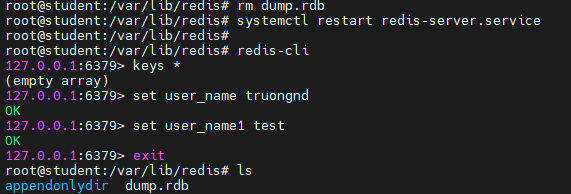
Cũng cần lưu ý thêm rằng, Redis cho phép không sử dụng tính năng lưu trữ thông tin trong ổ cứng (không RDB, không AOF), đồng thời cũng cho phép dùng cả 2 tính năng này trên cùng 1 instance. Tuy nhiên khi restart server, Redis sẽ dùng AOF cho việc tái tạo dataset ban đầu, bới AOF sẽ đảm bảo không bị mất mát dữ liệu tốt hơn là RDB.

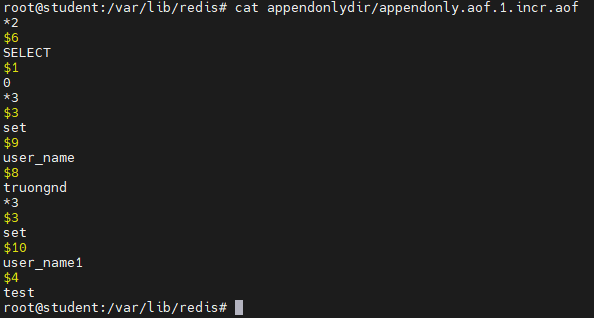
Để cấu hình AOF bạn cần thêm dòng sau vào file config:





Thực hiện restart lại redis-server và sử dụng lệnh SET để tạo key, ta thấy được redis tạo ra 1 folder là appendonlydir:





Với cấu hình này, mỗi khi redis nhận được lệnh thay đổi tập dữ liệu, nó sẽ thêm nó vào AOF. Khi bạn khởi động lại Redis, nó sẽ phát lại AOF để xây dựng lại database.

Kể từ Redis 7.0.0, Redis sử dụng cơ chế AOF nhiều phần. Nghĩa là, tệp AOF đơn gốc được chia thành tệp cơ sở (nhiều nhất là một) và tệp gia tăng (có thể có nhiều hơn một). Tệp cơ sở biểu thị ảnh chụp nhanh ban đầu (định dạng RDB hoặc AOF) của dữ liệu hiện tại khi AOF được viết lại. Các tệp gia tăng chứa các thay đổi gia tăng kể từ khi tệp AOF cơ sở cuối cùng được tạo. Tất cả các tệp này được đặt trong một thư mục riêng và được theo dõi bởi tệp kê khai.

**Cài đặt Redis:**

|  |
| --- |
| curl -fsSL https://packages.redis.io/gpg | sudo gpg --dearmor -o /usr/share/keyrings/redis-archive-keyring.gpg  echo "deb [signed-by=/usr/share/keyrings/redis-archive-keyring.gpg] https://packages.redis.io/deb $(lsb\_release -cs) main" | sudo tee /etc/apt/sources.list.d/redis.list  sudo apt-get update  sudo apt-get install redis |