

Aerospike

Trần Hữu Thúy

Ngày 7 tháng 3 năm 2020

Nội dung

1 Aerospike

2 Kiến trúc của Aerospike

3 Data Distribution

Tổng quan về Aerospike

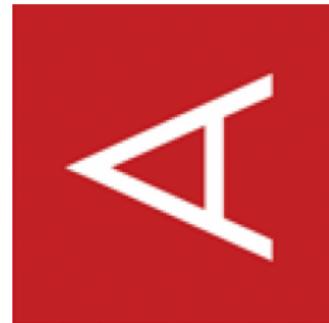
- Cơ sở dữ liệu mã nguồn mở NoSQL, được phát triển bởi công ty cùng tên.
- Cơ sở dữ liệu phân tán, có tính mở rộng.



Tổng quan về Aerospike

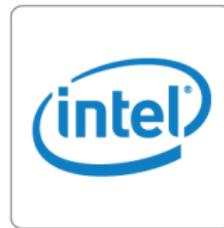
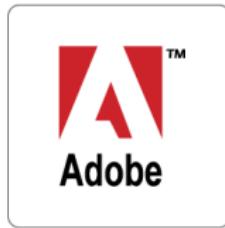
Mục tiêu

- Tạo ra các nền tảng ứng dụng web có tính linh hoạt và mở rộng.
- Cung cấp sức mạnh, ổn định như cơ sở dữ liệu truyền thống(cơ sở dữ liệu quan hệ).
- Cung cấp sự hiệu quả khi hoạt động.



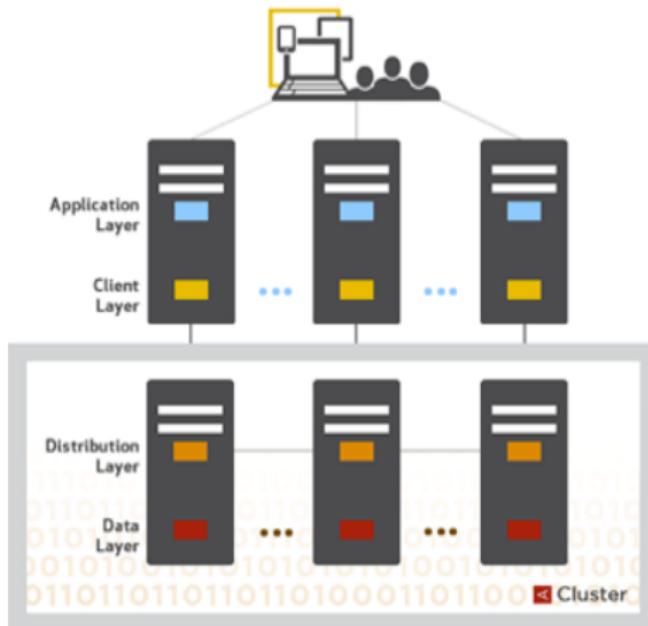
Tổng quan về Aerospike

Company



Kiến trúc

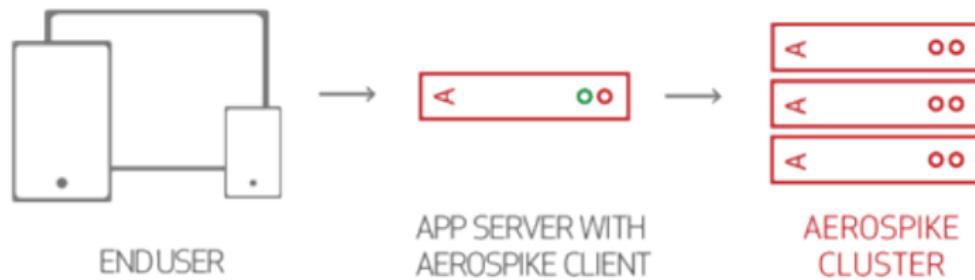
3 layer



Kiến trúc

Client Layer

- Bao gồm các thư viện mã nguồn mở, cài đặt Aerospike APIs, track nodes, và biết nơi dữ liệu được lưu trữ.



- Client giao tiếp với Cluster thông qua drivers trên C#, Java, Node.js,...
- Khi ứng dụng người dùng sử dụng Aerospike Client library như (Java client, ...). Nó bao gồm Aerospike Smart Client sử dụng giao thức để kết nối tới các server trên cluster.
- Khi code, không cần cài đặt, cấu hình quản lý conenction pooling kết nối trong cluster.

Kiến trúc

Distribution Layer

- Quản lý việc truyền thông trong cluter, nó tự động chuyển đổi, sao chép, đồng bộ hóa dữ liệu, tái cân bằng một cách thông minh và di chuyển dữ liệu.

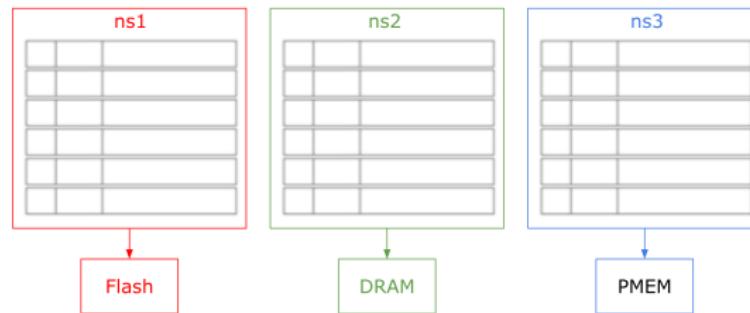


Hình: Sao chép dữ liệu factor=2

Kiến trúc

Data Storage Layer(tiếp)

- Kho lưu trữ dữ liệu theo kiểu key-value. Dữ liệu được chia thành các container, namespaces như đôi với cơ sở dữ liệu quan hệ.
- namespace tương tự như database của cơ sở dữ liệu quan hệ, và có thể lưu trữ trên Flash/SSD, DRAM, PMEM,..



Hình: Lưu trữ namespaces

Kiến trúc

Data Storage Layer(tiếp)

- Trong một namespace dữ liệu được chia thành các sets, mỗi set chứa nhiều record(bản ghi).



Hình: Tập các sets trong một namespace

Mỗi record bao gồm key, metadata, bins.

- **key** : định danh của duy nhất ở mỗi bản ghi, được băm để đánh địa chỉ của record.
- **metadata** : lưu thông tin bản ghi.
- **bins** : Các trường dữ liệu của bản ghi, các bin không cần phải có kiểu dữ liệu.

- Ta đã biết cơ bản về cách mỗi node lưu trữ dữ liệu.
- **Dữ liệu được phân chia cho mỗi node như nào ?**



Data Distribution

Aerospike sử dụng kiến trúc **Share Nothing** :

- Mỗi node trên cluster là giống nhau.
- Các node là ngang hàng.
- There is no single point of failure.

Thuật toán **Smart Partitions** phân tán dữ liệu trên tất cả các node.

END