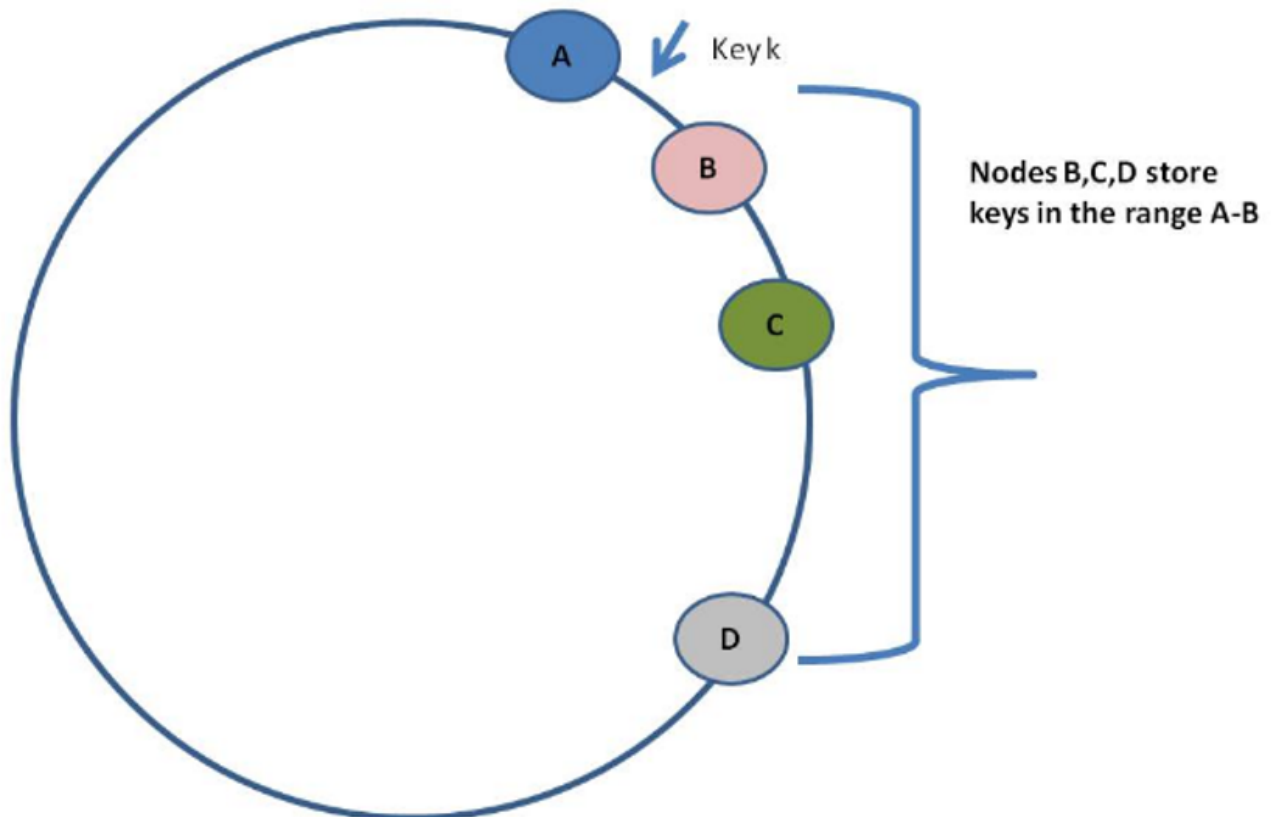


Replication dữ liệu trong Cassandra

Để đảm bảo tính sẵn sàng và liên tục trong Cassandra, mỗi đối tượng dữ liệu có thể được nhân bản và lưu giữ trên nhiều máy chủ. Nếu một trong các máy chủ lưu một phiên bản dữ liệu bị lỗi hoặc là phiên bản cũ, không phải là phiên bản được cập nhật dữ liệu mới nhất, Cassandra có cơ chế đồng bộ để luôn đảm bảo các thao tác đọc sẽ luôn trả về dữ liệu mới nhất. Đồng thời với việc này Cassandra tiến hành thao tác sửa lỗi đọc (read repair) là tiến trình ngầm để cập nhật trạng thái mới nhất cho tất cả các máy chủ lưu trữ nhân bản của dữ liệu.

Cassandra tổ chức các node máy chủ thành cụm theo định dạng vòng tròn và dữ liệu được phân tán theo vòng tròn này theo bảng hàm băm nhất quán (Distributed consistent hashing). Nếu mỗi dữ liệu của Cassandra được sao lưu trên N node, khi một khóa k được quyết định sẽ lưu vào một node nào đó, node đó sẽ được coi là node điều phối. Node điều phối có nhiệm vụ phân phối bản ghi đấy cho N-1 node còn lại theo nguyên tắc: từ node điều phối, đi theo chiều kim đồng hồ, dữ liệu sẽ được ghi lên 2 node tiếp theo được gặp.



Hình trên mô tả khi khóa k được xác định là sẽ ghi vào node B, node B sẽ đóng vai trò điều phối, luân chuyển khóa đấy cho 2 node tiếp theo là node C và node D. Như vậy, node D sẽ lưu trữ các khóa nằm trong vùng (A; D]. Danh sách các khóa trong vùng này được gọi là danh sách liên kết của node D. Việc đưa các giá trị của khóa k sang các node khác áp dụng cho tất cả các tác vụ ghi, cập nhật hay xóa.

Vì việc quyết định số lượng node được luân chuyển ngay lập tức mỗi khi có tác vụ ghi diễn ra ảnh hưởng trực tiếp đến mức độ nhất quán của hệ thống. Trong cấu hình của Cassandra Apache ta có một chỉ số "`replication_factor`" và "`w`". Chỉ số "`replication_factor`" sẽ được cài đặt ngay khi khởi tạo một `key_space`, đó là số lượng node trong vòng sẽ được dùng để sao lưu dữ liệu. Chỉ số "`w`" khi cấu hình Cassandra là số lượng node trả về kết quả khi thực hiện tác vụ ghi bắt buộc để tác vụ đấy được coi là thành công. Xét trên hình 8, khi ta đặt `replication_factor` = 3 và `w` = 2, khi khóa k được ghi vào thì cần phải có ít nhất 2 node trong 3 node B, C, D phản hồi lại ghi thành công thì tác vụ đấy mới được coi là thành công. Việc cài đặt chỉ số "`w`" cho ta thấy mức độ chi phí ta có thể bỏ ra để đảm bảo tính nhất quán của dữ liệu ngay lập tức.

Việc nhân bản dữ liệu cũng ảnh hưởng đến mức độ nhất quán của hệ thống. Mức độ nhất quán xét trên cả 2 phương diện đó là đọc và ghi dữ liệu. Để duy trì mức độ nhất quán của dữ liệu, Cassandra cung cấp cho người dùng nhiều mức độ nhất quán của các tác vụ đọc và ghi. Từ mức độ cao nhất đến thấp nhất, ta có thể điều chỉnh mức nhất quán dựa vào hai tham số cấu hình là "`w`" và "`r`" cùng với chỉ số "`replication_factor`". Trong đó, "`w`" là số node trả về khi ghi thành công, "`r`" là số node trả về khi đọc thành công. Nếu như tính nhất quán là sự ưu tiên, ta có thể đặt "`w`" và "`r`" sao cho đảm bảo

$$w + r > replication_factor$$

Và nên đảm bảo "`w`" hoặc "`r`" luôn nhỏ hơn `replication_factor` để cho được độ trễ tốt hơn. Giả sử như `replication_factor` = 3, vậy 2 giá trị của "`w`" và "`r`" tốt nhất sẽ là 2. Nghĩa là mỗi khi đọc và ghi dữ liệu, cần ít nhất 2 node trả về giá trị thì tác vụ đó coi là thành công. Và khi tác vụ đọc hoặc ghi thực hiện, sẽ luôn đảm bảo sẽ được thực hiện trên dữ liệu mới nhất mà tác vụ trước đó đã thực hiện.

Revision #1

Created 26 June 2021 03:53:30 by Laptrinh.vn

Updated 26 June 2021 03:55:35 by Laptrinh.vn