**Cơ chế sao lưu và phục hồi dữ liệu**

Theo định nghĩa, Aerospike sao chép dữ liệu giữa các node của một cluster và giữa các trung tâm dữ liệu. Tuy nhiên, chính sách phổ biến trong các hoạt động của máy tính là bảo mật dữ liệu bằng cách sao lưu nó trong trường hợp xảy ra sự cố. Hầu hết các tổ chức đều có lịch trình và thủ tục để sao lưu. Các công cụ sao lưu và khôi phục Aerospike là asbackup và asrestore.

# **Sao lưu**

Tiện ích ‘asbackup’ được sử dụng để sao lưu namespace, set hoặc partition từ một cluster vào bộ nhớ cục bộ. Theo mặc định, các bản sao lưu không được song song hóa. Nếu một bản ghi được tạo hoặc cập nhật sau khi phân vùng của nó được sao lưu, thì bản sao lưu sẽ không phản ánh điều đó.

## **Cách sử dụng**

Hình thức cơ bản nhất của việc chạy sao lưu là chỉ cần chỉ định cluster để sao lưu ‘—host’, namespace để sao lưu ‘—namespace’, cũng như thư mục cục bộ cho các tệp sao lưu ‘—directory’. Giả sử rằng chúng ta có một cluster chứa một node có địa chỉ IP 1.2.3.4. Để sao lưu namespace ‘test’ trên cluster này vào thư mục ‘backup\_2015\_08\_24’, sử dụng lệnh sau:

$ asbackup --host 1.2.3.4 --namespace test --directory backup\_2015\_08\_24

## **Ước tính kích thước sao lưu**

Khi truyền tùy chọn dòng lệnh ‘—estimate’ sang asbackup (và bỏ qua --directory và --output-file), asbackup sẽ tạo một bản sao lưu thử nghiệm tạm thời gồm 10.000 bản ghi từ namespace. Sau đó, nó xuất ra, dựa trên các kích thước bản ghi được quan sát, ước tính kích thước trung bình của một bản ghi trong bản sao lưu. Để ước tính tổng kích thước của tệp sao lưu hoặc các tệp, hãy nhân kích thước này với số lượng bản ghi trong namespace và thêm 10% cho chỉ mục và chi phí.

Lưu ý: các bộ lọc cho mỗi bản ghi (filter-exp, modified-after, modified-before, no-ttl-only, after-digest, partition-list) và node-list không được tính đến trong ước tính và sử dụng các tùy chọn này sẽ không ảnh hưởng đến ước tính.

# **Khôi phục**

Tiện ích ‘asrestore’ giúp khôi phục các bản sao lưu được tạo bằng ‘asbackup’. Nếu namespace trên cluster đã chứa các bản ghi hiện có, chính sách ghi có thể định cấu hình sẽ xác định, bản ghi nào được ưu tiên hơn - các bản ghi trong namespace hoặc bản ghi từ bản sao lưu.

Nếu việc chèn bản ghi ban đầu không thành công, bộ lưu trữ sẽ thử chèn lại bản ghi 10 lần. Giữa các lần thử có 1 giây tạm dừng và lỗi được ghi vào nhật ký ở mức độ nghiêm trọng gỡ lỗi. Nếu bản ghi không được ghi sau 10 lần thử, quá trình khôi phục sẽ bị hủy bỏ với thông báo lỗi "Quá nhiều lỗi, đang bỏ cuộc". Các lỗi cụ thể khiến việc thử lại không được thực hiện là "tồn tại bản ghi" (khi tùy chọn --unique được sử dụng), "tạo không khớp" (trừ khi nếu tùy chọn - không tạo được sử dụng) và "tên người dùng hoặc mật khẩu không hợp lệ" .

Khi chạy với tùy chọn --directory, ‘asrestore’ mong đợi nhiều tệp sao lưu .asb ​​trong thư mục directory . Ngoài ra, --input-file giúp asrestore đọc bản sao lưu hoàn chỉnh từ một tệp duy nhất đã cho.

## **Cách sử dụng**

Hình thức cơ bản nhất của việc chạy khôi phục là chỉ cần chỉ định cụm để khôi phục ‘—host’ và thư mục cục bộ chứa các tệp sao lưu ‘—directory’. Giả sử rằng chúng ta có một cluster chứa một node có địa chỉ IP 1.2.3.4. Để khôi phục bản sao lưu từ thư mục backup\_2015\_08\_24, sử dụng lệnh sau:

$ asrestore --host 1.2.3.4 --directory backup\_2015\_08\_24

Theo mặc định, bản sao lưu được khôi phục về namespace mà nó đã được lấy từ đó. Tùy chọn --namespace có thể được sử dụng để khôi phục lại một namespace khác. Giả sử rằng bản sao lưu ở trên được lấy từ namespace ‘test’ và muốn khôi phục nó về namespace ‘prod’, dùng lệnh sau:

$ asrestore --host 1.2.3.4 --directory backup\_2015\_08\_24 --namespace test,prod

# **Sao lưu và phục hồi trên AWS (Amazon Web Services)**

Cách truyền thống để sao lưu và phục hồi là sử dụng Aerospike Tools:

* Để sao lưu:
  + Sao lưu bình thường bằng ‘asbackup’
  + Di chuyển tệp đến một vị trí an toàn.
* Để phục hồi: sử dụng ‘asrestore’ tải bản sao lưu lên hệ thống, khôi phục dữ liệu lại thành một cluster.

Tuy nhiên, cách này có một vài ưu/khuyết điểm cần lưu ý như:

|  |  |
| --- | --- |
| Ưu điểm | Khuyết điểm |
| Dễ sử dụng | Tốn thời gian |
| Ít chi phí | Load server bổ sung khi chạy |
| Chứa dữ liệu SMD | Lãng phí cluster |

Bên cạnh đó, Aerospike cũng có một cách khác cho việc sao lưu và phục hồi dữ liệu. Trong khi chạy, có thể sử dụng trực tiếp EBS hoặc thông qua các thiết bị ẩn, có thể tận dụng ảnh chụp nhanh của EBS. Điều này được thực hiện với các dữ liệu cấp khối, trong suốt đối với Aerospike.

* Để sao lưu: đơn giản chỉ cần chụp nhanh các khối dữ liệu trên EBS.
* Để phục hồi: Nếu các bản sao lưu được thực hiện thông qua ảnh chụp nhanh EBS, cần đưa các ảnh đó vào tập EBS. Khi cung cấp các ổ đĩa EBS thay thế, cần định cấu hình lại phiên bản thay thế mới chính xác như cách cấu hình phiên bản không thành công. Điều này có nghĩa là các ổ đĩa EBS được gắn ở cùng vị trí và cùng một số lượng ổ đĩa tạm thời được cung cấp.

Cách làm này cũng có các ưu/khuyết điểm như sau:

|  |  |
| --- | --- |
| Ưu điểm | Khuyết điểm |
| Nhanh chóng | Tốn kém hơn |
| Ít bị ảnh hưởng | Cần có các thành phần bổ sung khác |
| Dựa trên từng node riêng lẻ | Không tính đến SMD |

Giả sử phiên bản của bạn không thành công, nhưng bạn vẫn sử dụng EBS lâu dài. Trong trường hợp này, bạn có thể tách tập EBS khỏi bản sao bị lỗi và gắn lại vào bản sao thay thế.

1. Loại bỏ / Chấm dứt phiên bản không thành công, nhưng vẫn giữ volume EBS.
2. Triển khai Phiên bản mới, có cùng IP riêng với phiên bản cũ trong cùng một khu vực.
3. Đính kèm volume EBS từ Bước 1 vào phiên bản thay thế.
4. Khôi phục cấu hình cho node thay thế và khởi động lại node.

Khi node tham gia lại vào cluster, quá trình di chuyển sẽ xảy ra và dữ liệu sẽ được cân bằng lại.

<https://docs.aerospike.com/docs/tools/backup/index.html>

<https://docs.aerospike.com/docs/tools/backup/asbackup.html>

<https://docs.aerospike.com/docs/tools/backup/asrestore.html>

<https://docs.aerospike.com/docs/deploy_guides/aws/backup/index.html>