**Xây dựng kênh dẫn dữ liệu để hỗ trợ phân tích dữ liệu luồng nhấp chuột với AWS**

**Tổng quan và mục tiêu**

Trong các khóa học AWS Academy khác nhau, bạn đã hoàn thành các phòng thí nghiệm thực hành. Bạn đã sử dụng các dịch vụ và tính năng AWS khác nhau để xây dựng nhiều giải pháp khác nhau.

Trong dự án này, bạn được thử thách sử dụng các dịch vụ AWS quen thuộc, cũng như các dịch vụ AWS có thể mới đối với bạn, để xây dựng giải pháp. Trong nhiều phần của dự án, hướng dẫn từng bước *không* được cung cấp. Đây là cố ý. Những phần cụ thể này của dự án nhằm thách thức bạn thực hành các kỹ năng mà bạn đã có được trong suốt kinh nghiệm học tập của mình trước dự án này. Trong một số trường hợp, bạn có thể được thử thách sử dụng các nguồn lực để độc lập học các kỹ năng mới.

Đến cuối dự án này, bạn sẽ có thể làm như sau:

* Triển khai kênh dẫn phân tích dữ liệu trên AWS hỗ trợ phân tích dữ liệu luồng nhấp chuột của trang web.
* Chuyển đổi dữ liệu luồng nhấp chuột trước khi nó đến lớp trực quan hóa.
* Sử dụng dịch vụ AWS để phân tích dữ liệu luồng nhấp chuột.
* Thiết kế cơ chế báo cáo bảng điều khiển để phân tích dữ liệu luồng nhấp chuột.
* Điều chỉnh quy trình phân tích dữ liệu.

**Môi trường phòng thí nghiệm và giám sát ngân sách của bạn**

**Môi trường này tồn tại lâu dài**. Khi bộ hẹn giờ phiên chạy đến 0:00, phiên sẽ kết thúc, nhưng mọi dữ liệu và tài nguyên bạn đã tạo trong tài khoản AWS sẽ được giữ lại. Tuy nhiên, lưu ý rằng địa chỉ IP công cộng của máy chủ web quán cà phê của bạn sẽ thay đổi. Nếu sau đó bạn khởi chạy một phiên mới (ví dụ: ngày hôm sau), bạn sẽ thấy rằng công việc của bạn vẫn còn trong môi trường phòng thí nghiệm, nhưng bạn sẽ cần truy cập EC2 và sử dụng địa chỉ IP công cộng mới để truy cập và duyệt trang web quán cà phê. Ngoài ra, tại bất kỳ thời điểm nào trước khi bộ hẹn giờ phiên đạt 0:00, bạn có thể chọn  **lại Start Lab** để kéo dài thời gian phiên phòng thí nghiệm.

**Quan trọng:** Theo dõi ngân sách phòng thí nghiệm của bạn trong giao diện phòng thí nghiệm. Khi bạn có một phiên phòng thí nghiệm đang hoạt động, thông tin ngân sách còn lại mới nhất đã biết sẽ hiển thị ở đầu màn hình này. Ngân sách còn lại mà bạn thấy có thể không phản ánh hoạt động tài khoản gần đây nhất của bạn. Nếu bạn vượt quá ngân sách phòng thí nghiệm, tài khoản phòng thí nghiệm của bạn sẽ bị vô hiệu hóa và tất cả tiến trình và tài nguyên sẽ bị mất. Nếu bạn vượt quá ngân sách, vui lòng liên hệ với nhà giáo dục của bạn để được hỗ trợ.

**Hạn chế dịch vụ AWS**

Trong môi trường phòng thực hành này, quyền truy cập vào các dịch vụ AWS và hành động dịch vụ bị hạn chế ở những dịch vụ cần thiết để hoàn thành hướng dẫn trong phòng thí nghiệm. Bạn có thể gặp lỗi nếu cố gắng truy cập các dịch vụ khác hoặc thực hiện các hành động ngoài các dịch vụ được mô tả trong phòng thực hành này.

**Kịch bản**

AnyCompany Café bán các mặt hàng tráng miệng và cà phê thông qua trang web của họ. Họ có quán cà phê ở nhiều thành phố trên khắp thế giới. Công ty muốn hiểu rõ hơn về doanh nghiệp bằng cách sử dụng dữ liệu về cách mọi người tương tác với trang web. Công ty có kế hoạch phân tích xu hướng dữ liệu luồng nhấp chuột để đưa ra quyết định thông minh hơn về nơi đầu tư. AnyCompany Café đã thuê công ty tư vấn của bạn để dẫn dắt nỗ lực này. Bạn là một chuyên gia tích hợp dữ liệu và làm việc với một nhóm kỹ sư dữ liệu, nhà phân tích dữ liệu và nhà phát triển web.

Dữ liệu nhật ký nhấp chuột từ trang web quán cà phê bao gồm một mục nhập cho mỗi nhấp chuột mà khách hàng tiềm năng thực hiện trong khi duyệt trang web. Nhiệm vụ của bạn là thiết kế và tạo một quy trình phân tích dữ liệu để thu thập dữ liệu luồng nhấp chuột. Bạn cũng sẽ tạo một bảng điều khiển phân tích để chủ quán cà phê có thể nhanh chóng quan sát hành vi của khách hàng. Nhóm muốn hiểu rõ hơn về hành vi của người dùng trên trang web và sau đó sử dụng những hiểu biết đó để tập trung nỗ lực quảng cáo. Công ty thậm chí có thể sử dụng dữ liệu để quyết định nơi mở thêm địa điểm.

**Yêu cầu giải pháp**

Giải pháp phải đáp ứng các yêu cầu sau:

* Thiết kế và tối ưu hóa chi phí giải pháp trước khi xây dựng.
* Đảm bảo rằng giải pháp là chức năng.
* Chuyển đổi dữ liệu.
* Nhập dữ liệu access\_log từ máy chủ web quán cà phê.
* Sử dụng dữ liệu nhật ký mô phỏng.
* Phân tích và trực quan hóa dữ liệu.
* Tạo thông tin chi tiết.

**Mẹo dự án phòng thí nghiệm**

Kiến thức về các lệnh Linux Bash sẽ hữu ích nhưng không bắt buộc. Các lệnh Bash được sử dụng trong phòng thí nghiệm này bao gồm cp, mv, cat, sudo, yum , wget và find. Bạn được khuyến khích làm quen với mục đích của mỗi lệnh nếu bạn chưa biết chúng.

**Mẹo:** Để biết thông tin về các lệnh Linux Bash, hãy xem [Trang Người đàn ông Linux trên linux.die.net](https://linux.die.net/man). Để tìm hiểu về một lệnh cụ thể, hãy nhập lệnh đó vào hộp tìm kiếm ở đầu trang.

Cuối cùng, bạn được khuyến khích tháo vát khi bạn hoàn thành dự án này. Ví dụ: tham khảo [Tài liệu AWS](https://docs.aws.amazon.com/) hoặc công cụ tìm kiếm nếu bạn cần câu trả lời cho câu hỏi kỹ thuật.

Khi bạn làm việc thông qua dự án, bạn sẽ tìm thấy các mẹo khác để giúp bạn hoàn thành dự án, cụ thể cho từng giai đoạn hoặc nhiệm vụ.

**Tiếp cận**

Bảng sau đây mô tả các giai đoạn của dự án:

| **Giai đoạn đánh số** | **Chi tiết** |
| --- | --- |
| 1 | Tạo sơ đồ kiến trúc và ước tính chi phí cho giải pháp. |
| 2 | Sử dụng AWS Cloud9 để truy cập máy chủ web. Quan sát nhật ký truy cập được tạo khi bạn duyệt trang web. |
| 3 | Cài đặt và cấu hình tác nhân Amazon CloudWatch và tệp cấu hình (httpd.conf) cho máy chủ web Apache để có thể thu thập và gửi nhật ký truy cập và lỗi đến CloudWatch. |
| 4 | Xác nhận rằng máy chủ web tạo nhật ký mà tác nhân CloudWatch thu thập và gửi đến CloudWatch. |
| 5 | Sử dụng nhật ký mô phỏng và đảm bảo rằng CloudWatch nhận được những nhật ký này. |
| 6 | Sử dụng CloudWatch Logs Insights để truy vấn nhóm nhật ký truy cập và tạo trực quan hóa. |
| 7 | Điều chỉnh quy trình phân tích để cung cấp thông tin chi tiết mới với dữ liệu vị trí địa lý của khách truy cập. Tạo bảng điều khiển mà chủ quán cà phê có thể sử dụng để đưa ra quyết định. |

**Giai đoạn 1: Lập kế hoạch thiết kế và dự toán chi phí**

Trong giai đoạn này, bạn sẽ lập kế hoạch thiết kế kiến trúc của mình. Đầu tiên, bạn sẽ tạo một sơ đồ kiến trúc.

Tiếp theo, bạn sẽ ước tính chi phí của giải pháp được đề xuất và trình bày ước tính cho nhà giáo dục của bạn. Bước đầu tiên quan trọng cho bất kỳ giải pháp nào là lập kế hoạch thiết kế và ước tính chi phí. Xem xét các thành phần khác nhau trong kiến trúc để điều chỉnh chi phí ước tính. Chi phí, cùng với các tính năng và hạn chế của các dịch vụ AWS cụ thể là những yếu tố quan trọng khi xây dựng giải pháp. Ví dụ, chúng có thể giúp xác định các thành phần giải pháp và một mẫu kiến trúc để sử dụng.

**Nhiệm vụ 1: Tạo sơ đồ kiến trúc**

1. Tạo một sơ đồ kiến trúc để minh họa những gì bạn dự định xây dựng. Xem xét cách bạn sẽ hoàn thành từng yêu cầu trong giải pháp. Đọc qua các giai đoạn trong tài liệu này để biết những dịch vụ và tính năng AWS nào mà công ty yêu cầu bạn sử dụng. Đảm bảo bao gồm các biểu tượng dịch vụ hoặc tài nguyên sau đây trong sơ đồ của bạn:
   * Nhóm nhật ký CloudWatch
   * CloudWatch Logs Insights
   * Bảng thông tin CloudWatch
   * Phiên bản môi trường AWS Cloud9 (Amazon EC2) cho máy chủ web
   * Vai trò AWS Identity and Access Management (IAM)
   * Amazon S3

**Tham khảo**

* + [Biểu tượng kiến trúc AWS](https://aws.amazon.com/architecture/icons/): Trang web này cung cấp các công cụ để vẽ sơ đồ kiến trúc AWS.
  + [Sơ đồ kiến trúc tham chiếu AWS](https://aws.amazon.com/architecture/reference-architecture-diagrams): Trang web này cung cấp danh sách các sơ đồ kiến trúc AWS cho các trường hợp sử dụng khác nhau. Bạn có thể muốn sử dụng các sơ đồ này làm tài liệu tham khảo.

1. Tùy chọn: Cân nhắc đưa các tài nguyên AWS bổ sung vào sơ đồ kiến trúc của bạn. Chẳng hạn:
   * Bộ chứa Amazon Simple Storage Service (Amazon S3) để lưu trữ bản sao nhật ký
   * Tính năng S3 Select để truy vấn nhật ký trong vùng lưu trữ, như một cơ chế thay thế để truy vấn dữ liệu trong tương lai

**Nhiệm vụ 2: Xây dựng dự toán chi phí**

Xây dựng ước tính chi phí cho thấy chi phí để chạy giải pháp ở Khu vực us-east-1 trong 12 tháng. Sử dụng [Bộ tính toán giá AWS](https://calculator.aws/) cho ước tính này.

**Tùy chọn: Thêm sơ đồ và ước tính chi phí vào bản trình bày**

* Thêm sơ đồ kiến trúc và ước tính chi phí của bạn vào các trang trình bày. Nhà giáo dục của bạn có thể muốn đánh giá thông tin này như là một phần của việc đánh giá công việc của bạn trong dự án này. Một mẫu bản trình bày được cung cấp.
* Ngoài ra, bạn nên chụp ảnh màn hình công việc của mình vào cuối mỗi nhiệm vụ hoặc giai đoạn để đưa vào bản trình bày hoặc tài liệu. Người hướng dẫn của bạn có thể sử dụng bài thuyết trình hoặc tài liệu để giúp đánh giá mức độ bạn hoàn thành các yêu cầu của dự án.

**Tham khảo**

* [Mẫu bản trình bày PowerPoint](https://aws-tc-largeobjects.s3.us-west-2.amazonaws.com/CUR-TF-200-ACCAP4-1-79925/capstone-4-clickstream/s3/Academy_Lab_Projects_Showcase_template.pptx)

Bạn đã tạo sơ đồ kiến trúc và ước tính chi phí cho giải pháp của mình.

**Giai đoạn 2: Phân tích website và xác nhận dữ liệu weblog**

Trang web quán cà phê đã tồn tại trong tài khoản AWS của bạn. Trong giai đoạn này, bạn sẽ phân tích các tài nguyên tài khoản hiện có. Sau đó, bạn sẽ xác minh rằng trang web quán cà phê có thể truy cập được. Cuối cùng, bạn sẽ quan sát cách dữ liệu nhấp chuột được tạo ra khi bạn truy cập trang web.

**Nhiệm vụ 1: Phân tích và hiểu môi trường phòng thí nghiệm**

Lần đầu tiên bạn khởi động môi trường phòng thực hành, tài nguyên sẽ được tạo cho bạn trong tài khoản AWS. Khi bạn đọc về từng tài nguyên, bạn được khuyến khích xác định vị trí và quan sát cài đặt của chúng bằng cách sử dụng Bảng điều khiển quản lý AWS.

Các tài nguyên bao gồm:

* Một đám mây riêng ảo ( VPC) có tên *Lab VPC* với một mạng con công cộng có tên *Public Subnet* trong đó. Mạng con có thể truy cập internet thông qua một cổng internet.
* Phiên bản AWS Cloud9 chạy dưới dạng phiên bản EC2 trong mạng con công cộng. AWS Cloud9 cung cấp môi trường phát triển tích hợp (IDE), thuận tiện để chỉnh sửa tệp và chạy các lệnh AWS Command Line Interface (AWS CLI).
* Nhóm bảo mật được liên kết với phiên bản EC2.
* Vai trò IAM có tên *CafeRole*, được gắn vào phiên bản EC2. Vai trò này cho phép truy cập từ phiên bản đến các dịch vụ AWS khác, chẳng hạn như AWS Systems Manager, Amazon S3 và CloudWatch.
* Phiên bản AWS Cloud9 cũng lưu trữ ứng dụng web cafe. Ứng dụng web chạy trên máy chủ web Apache httpd và PHP và nó sử dụng cơ sở dữ liệu MariaDB cũng đang chạy trên phiên bản.

**Quan trọng:**

* Khi phiên phòng thực hành của bạn kết thúc (vì bộ hẹn giờ chạy đến 0:00 hoặc bạn chọn **End Lab**), phiên bản AWS Cloud9 EC2 sẽ bị dừng. Tài nguyên mà bạn đã tạo trong tài khoản sẽ vẫn còn, nhưng bạn sẽ không thể truy cập ứng dụng web trong khi phiên bản bị dừng.
* Khi bạn khởi động lại phòng thực hành (ví dụ: ngày hôm sau), phiên bản EC2 AWS Cloud9 sẽ bị dừng và cần khởi động lại. Tham khảo [Dừng và bắt đầu phiên bản của bạn](https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/Stop_Start.html) để biết chi tiết.
* Bạn có thể đặt lại môi trường phòng thí nghiệm về cấu hình ban đầu bất kỳ lúc nào bằng cách chọn **Đặt lại** phía trên các hướng dẫn này.

**Cảnh báo:** Đặt lại môi trường sẽ *xóa mọi thứ bạn đã tạo trong tài khoản*.

**Nhiệm vụ 2: Sửa đổi nhóm bảo mật và xác minh rằng ứng dụng web tải**

Phiên bản AWS Cloud9 được tạo cho bạn chạy trên phiên bản EC2. Phiên bản đó cũng lưu trữ ứng dụng web café, chạy trên cổng TCP 80. Tuy nhiên, nhóm bảo mật được đính kèm với phiên bản chưa cho phép lưu lượng truy cập đến trên cổng đó. Trong tác vụ này, bạn sẽ giải quyết vấn đề đó.

1. Trong bảng điều khiển Amazon EC2, định vị phiên bản và điều chỉnh nhóm bảo mật sao cho nó cho phép lưu lượng TCP đến trên cổng 80 từ bất kỳ vị trí IPv4 nào đến trang web quán cà phê.
2. Sao chép địa chỉ IPv4 công khai của phiên bản EC2.

**Mẹo:** Nếu phiên bản không chạy, hãy khởi động ngay bây giờ để gán IP công cộng.

**Quan trọng:** Khi phiên phòng thực hành của bạn khởi động lại, địa chỉ IP công khai của phiên bản AWS Cloud9 sẽ thay đổi. Hãy chắc chắn rằng bạn đang sử dụng địa chỉ IPv4 công cộng mới nhất của phiên bản để truy cập trang web quán cà phê.

1. Dán địa chỉ IP vào tab trình duyệt mới và thêm /cafe vào cuối URL trước khi tải.

**Quan trọng:** Đảm bảo rằng bạn đang tải [*http://public-ip/cafe*](http://public-ip/cafe) chứ không  [*phải https://public-ip/cafe*](https://public-ip/cafe).

Trang web quán cà phê tải và hiển thị hình ảnh của các món tráng miệng mà quán cà phê cung cấp. Giữ trang này mở để sử dụng trong tác vụ sau.

**Nhiệm vụ 3: Quan sát và sao lưu access\_log httpd**

Bây giờ bạn biết rằng ứng dụng web có thể truy cập được, bạn sẽ theo dõi nhật ký ghi lại mỗi khi người dùng truy cập một trang trên trang web.

1. Trong bảng điều khiển AWS Cloud9, mở IDE cho môi trường AWS Cloud9 đã được tạo cho bạn.

**Lưu ý:** AWS Cloud9 IDE chạy trên cùng một phiên bản EC2 lưu trữ ứng dụng web cafe. Để biết thêm thông tin về IDE AWS Cloud9, hãy xem [Bắt đầu: Hướng dẫn cơ bản về AWS Cloud9](https://docs.aws.amazon.com/cloud9/latest/user-guide/tutorials-basic.html).

1. Trong thiết bị đầu cuối AWS Cloud9, sử dụng lệnh đuôi Linux để theo dõi các thay đổi trong thời gian thực đối với tệp access\_log của máy chủ web Apache .

**Mẹo:**

* + Giống như nhiều ứng dụng khác có thể được cài đặt trên Linux, máy chủ web Apache httpd lưu trữ nhật ký theo mặc định trong thư mục con của  */var/log*.
  + Với lệnh tail, bạn sẽ muốn sử dụng tham số -f. Để biết thêm thông tin về lệnh, hãy xem [trang hướng dẫn sử dụng Linux này để biết lệnh đuôi](https://man7.org/linux/man-pages/man1/tail.1.html).

1. Quay trở lại tab trình duyệt nơi ứng dụng web quán cà phê đang mở và làm mới trang.

**Mẹo:** Sắp xếp trang web quán cà phê và tab AWS Cloud9 IDE cạnh nhau để bạn có thể quan sát cách các mục nhập nhật ký được thực hiện khi duyệt trang web quán cà phê.

1. Xác nhận rằng mỗi lần bạn tải một trang trên trang web quán cà phê, một mục nhật ký mới được tạo trong access\_log. Hãy thử nhiều tác vụ khác nhau — làm mới trang chủ, truy cập trang **Menu** , gửi đơn đặt hàng hoặc xem lịch sử đơn hàng.
   * Khi bạn điều hướng trang web, hãy quan sát cách URL của trang thay đổi. Chẳng hạn:
     + Trang web chính ở [*http://public-ip/cafe*](http://public-ip/cafe).
     + Trang web menu ở [*http://public-ip/cafe/menu.php*](http://public-ip/cafe/menu.php).
     + Sau khi bạn đặt hàng, bạn sẽ được đưa đến một trang tại [*http://public-ip/cafe/processOrder.php*](http://public-ip/cafe/processOrder.php).
   * Đường dẫn của trang mà bạn tải được ghi lại trong mỗi mục nhật ký, cùng với các chi tiết khác. Bạn đang tạo và quan sát dữ liệu luồng nhấp chuột trong thời gian thực.
2. Để dừng lệnh đuôi, hãy nhấn Ctrl+C. Bạn được đưa trở lại dấu nhắc lệnh.
3. Để sao lưu tệp access\_log, hãy chạy lệnh sau:



sudo cp /var/log/httpd/access\_log /home/ec2-user/environment/initial\_access\_log

Bây giờ bạn đã hiểu rõ hơn về cách trang web quán cà phê hoạt động và cách tạo các mục nhật ký nhấp chuột.

**Giai đoạn 3: Cài đặt tác nhân CloudWatch và tạo tệp cấu hình**

Trong giai đoạn này, bạn sẽ cài đặt và cấu hình một tác nhân CloudWatch trên máy chủ web quán cà phê để bạn có thể sử dụng nhật ký CloudWatch để thu thập dữ liệu access\_log cũng như mọi dữ liệu error\_log.

**Tác vụ 1: Cài đặt tác nhân CloudWatch trên máy chủ web**

Để cài đặt tác nhân CloudWatch, hãy chạy lệnh sau trong thiết bị đầu cuối AWS Cloud9:



sudo yum cài đặt -y amazon-cloudwatch-agent

**Tác vụ 2: Tạo tệp cấu hình cho tác nhân CloudWatch**

1. Để tải xuống và cài đặt tệp cấu hình tác nhân CloudWatch, hãy chạy lệnh sau:



WGET https://aws-tc-largeobjects.s3.us-west-2.amazonaws.com/CUR-TF-200-ACCAP4-1-79925/capstone-4-clickstream/s3/config.json

sudo mv config.json /opt/aws/amazon-cloudwatch-agent/bin/

1. Để xem lại nội dung của tệp config.json, hãy chạy lệnh sau:



sudo cat /opt/aws/amazon-cloudwatch-agent/bin/config.json

Lưu ý rằng tệp này chứa thông tin cấu hình cho tác nhân, bao gồm:

* + Nhật ký nào cần thu thập từ máy chủ web quán cà phê
  + Nhóm nhật ký nào trong CloudWatch để đặt nhật ký
  + Thời gian lưu giữ nhật ký
  + Số liệu nào cần thu thập và cách tổng hợp chúng

**Lưu ý:** Tệp này được tạo bằng cách chạy sudo /opt/aws/amazon-cloudwatch-agent/bin/amazon-cloudwatch-agent-config-wizard và trả lời một loạt câu hỏi để cấu hình tác tử CloudWatch. Tuy nhiên, đối với dự án này, tệp cấu hình được cung cấp cho bạn.

**Nhiệm vụ 3: Định cấu hình định dạng nhật ký httpd.conf là JSON**

Bây giờ tác nhân CloudWatch đã được cài đặt và cấu hình, bạn sẽ sửa đổi định dạng mặc định của nhật ký truy cập và nhật ký lỗi mà máy chủ web tạo ra. Cụ thể, bạn sẽ cần thay đổi định dạng từ Định dạng nhật ký chung (CLF) mặc định sang định dạng JSON. Điều này sẽ cho phép tác nhân CloudWatch gửi tệp nhật ký đến các nhóm nhật ký CloudWatch.

1. Xác định vị trí tệp cấu hình của máy chủ web, được đặt tên là *httpd.conf* và sao lưu tệp.

**Mẹo:** Thư mục /etc và các thư mục con của nó là nơi phổ biến để các ứng dụng lưu trữ các tệp cấu hình. Nếu bạn không quen thuộc với cách làm việc với thiết bị đầu cuối Linux và sử dụng các lệnh, bạn có thể thấy hữu ích khi tra cứu chi tiết các lệnh cp, tìm và sudo trong [tài liệu tham khảo trang người đàn ông Linux](https://linux.die.net/man/).

1. Sửa đổi tệp *httpd.conf* để nhật ký lỗi và nhật ký truy cập sẽ được định dạng là JSON thay vì định dạng CLF mặc định:
   * Để tránh phải sử dụng trình soạn thảo văn bản đầu cuối (chẳng hạn như vi hoặc nano) để chỉnh sửa các tệp trong  *thư mục /etc/httpd/conf*, hãy chạy các lệnh sau:



ln -s /etc/httpd/conf /home/ec2-user/environment/httpdconf

**Lưu ý:** Lệnh trước tạo liên kết tượng trưng đến thư mục /*etc/httpd/conf* để bạn có thể xem các tệp trong  **cửa sổ** Môi trường AWS Cloud9.



sudo chown -R ec2-user /etc/httpd/conf

**Lưu ý:** Lệnh trước thay đổi quyền sở hữu các tệp trong thư mục /*etc/httpd/conf* để bạn có thể lưu các thay đổi khi sử dụng trình soạn thảo văn bản AWS Cloud9. Người dùng Linux ec2 không có quyền chỉnh sửa các tệp trong thư mục này.

* + Trong cửa sổ **Môi trường** của IDE, mở rộng thư mục **httpdconf** và mở tệp **httpd.conf** trong trình chỉnh sửa tệp.

1. Điều chỉnh cấu hình nhật ký lỗi:
   * Trong tệp *httpd.conf*, hãy chuyển đến dòng đọc ErrorLog "logs/error\_log"(khoảng dòng 182). Bình luận ra dòng.

**Mẹo:** Để nhận xét một dòng, hãy nhập ký tự # ở đầu dòng.

* + Sao chép và dán văn bản sau đây dưới dạng một dòng mới ngay sau dòng mà bạn vừa nhận xét:



ErrorLog "/var/log/www/error/error\_log"

* + Sao chép và dán văn bản sau đây dưới dạng dòng mới ngay sau dòng ErrorLog mà bạn vừa thêm:



ErrorLogFormat "{\"time\":\"%{%usec\_frac}t\", \"function\" : \"[%-m:%l]\", \"process\" : \"[pid%P]\" ,\"message\" : \"%M\"}"

1. Điều chỉnh cấu hình nhật ký truy cập:
   * Chuyển đến dòng có nội dung <IfModule log\_config\_module> (khoảng dòng 191). Chú thích dòng ***bên dưới nó***, có chứa LogFormat "%h %l %u %t \"%r\" %>s %b \"%{Referer}i\" \"%{User-Agent}i\"" kết hợp.
   * Sao chép và dán văn bản sau đây dưới dạng dòng mới ngay sau dòng chung LogFormat "%h %l %u %t \"%r\" %>s %b" vẫn còn:



LogFormat "{ \"time\":\"%{%Y-%m-%d}tT%{%T}t.%{msec\_frac}tZ\", \"process\":\"%D\", \"filename\":\"%f\", \"remoteIP\":\"%a\", \"host\":\"%V\", \"request\":\"%U\", \"query\":\"%q\",\"method\":\"%m\", \"status\":\"%>s\", \"userAgent\":\"%{User-agent}i\",\"referer\":\"%{Referer}i\"}" CloudWatch

* + Nhận xét toàn bộ phần logio\_module> <IfModule. Để làm điều này, hãy bình luận ba dòng chưa được nhận xét trước đó.
  + Chuyển đến dòng đọc CustomLog "nhật ký / access\_log" kết hợp (khoảng dòng 219).
  + Giữ dòng, nhưng sao chép và dán văn bản sau đây dưới dạng dòng mới ngay sau dòng CustomLog.



Đồng hồ đám mây CustomLog "/var/log/www/access/access\_log"

1. Lưu và đóng tệp.

**Tác vụ 4: Sử dụng tệp cấu hình cập nhật cho tác nhân CloudWatch**

1. Tạo thư mục nhật ký lỗi và truy cập mới để các vị trí thư mục mà bạn đã chỉ định trong tệp httpd.conf tồn tại trên máy chủ.

**Lưu ý:** Máy chủ web có thể tạo các tệp, nhưng trước tiên bạn phải xác định các thư mục.

1. Khởi động lại dịch vụ httpd (máy chủ web) để các thay đổi mà bạn đã thực hiện đối với cấu hình máy chủ web được tải.

**Mẹo:** Sử dụng lệnh systemctl kết hợp với lệnh sudo.

1. Khởi động tác tử CloudWatch để sử dụng tệp *config.json* mà bạn đã cài đặt.

**Mẹo:** Để biết thông tin về cách chạy lệnh amazon-cloudwatch-agent-ctl, hãy xem [Khởi động CloudWatch Agent bằng dòng lệnh](https://docs.aws.amazon.com/AmazonCloudWatch/latest/monitoring/install-CloudWatch-Agent-commandline-fleet.html#start-CloudWatch-Agent-EC2-commands-fleet).

1. Để xác nhận rằng tổng đài viên hiện đang chạy và hoạt động, hãy chạy lệnh sau:



Trạng thái dịch vụ Amazon-CloudWatch-Agent

Bây giờ bạn đã cài đặt và cấu hình tác nhân CloudWatch và tệp httpd.conf. Bạn cũng đã tạo thư mục nhật ký truy cập và lỗi.

**Giai đoạn 4: Kiểm thử tác nhân CloudWatch**

Trong giai đoạn này, bạn sẽ kiểm tra tác nhân CloudWatch để xác nhận rằng nó đang thu thập và gửi dữ liệu nhật ký truy cập máy chủ web và nhật ký lỗi tới CloudWatch.

**Nhiệm vụ 1: Quan sát tệp access\_log httdp**

1. Trong ứng dụng web quán cà phê, thực hiện nhiều hành động khác nhau, như bạn đã làm trước đây để tạo nhật ký.
2. Sử dụng kỹ thuật tương tự mà bạn đã sử dụng trong giai đoạn 2 nhiệm vụ 3 để quan sát các hành động khi chúng được thêm vào tệp access\_log. Hãy nhớ rằng bạn đã cấu hình máy chủ web để ghi vào một vị trí nhật ký khác và các mục nhật ký bây giờ sẽ được viết ở định dạng JSON thay vì Định dạng Nhật ký Chung (CLF) mặc định.

**Nhiệm vụ 2: Quan sát tập tin amazon-cloudwatch-agent.log**

1. Xác định vị trí của tệp *amazon-cloudwatch-agent.log* và sử dụng lệnh cat để xem nội dung của tệp.
2. Trong tệp, hãy tìm các dòng tương tự như sau:



[inputs.logfile] Đọc từ offset 10620 trong /var/log/www/error/error\_log

[inputs.logfile] Đọc từ offset 19850 trong /var/log/www/access/access\_log

Nhật ký đường ống [logagent] từ apache/error/i-0d378cc2f1bb61140(/var/log/www/error/error\_log) sang CloudwatchWatchLogs với lưu giữ 180

Nhật ký đường ống [logagent] từ apache/access/i-0d378cc2f1bb61140(/var/log/www/access/access\_log) sang CloudwatchWatchLogs với lưu giữ 180

Những dòng này cho biết rằng tác nhân CloudWatch đang đọc từ hai tệp nhật ký máy chủ web và đường ống (gửi) các nhật ký đó đến dịch vụ CloudWatch.

**Mẹo:** Bạn có thể yên tâm bỏ qua bất kỳ lỗi nào khi sử dụng đĩa ("/sys/kernel/debug/tracing"): các dòng bị từ chối quyền xuất hiện trong nhật ký.

**Nhiệm vụ 3: Đặt lệnh trong ứng dụng web và quan sát nhật ký CloudWatch**

1. Mở ứng dụng web quán cà phê tại http://<public-ip-address>/cafe

**Lưu ý:** Thay thế <địa chỉ IP công cộng> bằng địa chỉ IP công cộng của phiên bản máy chủ web.

**Mẹo:** Để làm cho các mục nhật ký hơi khác một chút, hãy sử dụng thiết bị di động để duyệt trang web.

1. Truy cập trang **Menu** và gửi đơn đặt hàng.
2. Xác minh rằng các hành động mới nhất của bạn đã được ghi lại trong CloudWatch:
   * Trong bảng điều khiển CloudWatch, tìm  **mục nhập** apache/access trong  **khu vực** Nhóm nhật ký.
   * Mở luồng nhật ký.
   * Mở rộng mục nhập đầu tiên để xem chi tiết nhật ký, ở định dạng JSON. Xác minh rằng các hành động mới nhất của bạn đã được ghi lại.

**Lưu ý:** Dữ liệu luồng nhấp chuột của bạn hiện đã có sẵn trong CloudWatch vì tác nhân CloudWatch trên máy chủ web ghi dữ liệu này vào nhật ký CloudWatch.

Bạn đã thử nghiệm thành công giải pháp ghi nhật ký luồng nhấp chuột mới với CloudWatch, sau khi xác nhận rằng nhật ký được nhận trong luồng nhật ký.

**Giai đoạn 5: Sử dụng nhật ký mô phỏng và đảm bảo rằng CloudWatch nhận được các mục nhập**

Trong giai đoạn này của dự án, bạn sẽ thay thế access\_log hiện có bằng một có dữ liệu mô phỏng. Nhật ký mô phỏng này có nhiều mục hơn bạn có thể tạo bằng cách tự duyệt trang web theo cách thủ công. Tập dữ liệu lớn hơn này sẽ giúp mô phỏng hoạt động nhật ký trong thế giới thực cho một trang web được truy cập thường xuyên.

Sau khi bạn thay thế tệp hiện có bằng dữ liệu mô phỏng này, bạn sẽ khởi động lại tác nhân CloudWatch để nó thu thập dữ liệu access\_log mới và gửi đến CloudWatch để phân tích thêm.

**Nhiệm vụ 1: Phân tích tệp nhật ký mô phỏng**

1. Xem lại tệp nhật ký mô phỏng:
   * Để xem một vài dòng đầu tiên của dữ liệu nhật ký mô phỏng, hãy chạy lệnh sau:



nhật ký mẫu mèo/access\_log.log | đầu

* + Để in đẹp dòng đầu tiên của tệp nhật ký để dễ đọc hơn, hãy chạy lệnh sau:



cat samplelogs/access\_log.log | head -1 | python -m json.tool

Lưu ý rằng định dạng của tệp khớp với định dạng của nhật ký mà máy chủ web của bạn đã tạo sau khi bạn thay đổi cài đặt LogFormat trong tệp httpd.conf. Việc kết hợp định dạng này là có chủ ý.

* + Để đếm số dòng trong tệp, hãy chạy lệnh sau:



nhật ký mẫu mèo/access\_log.log| WC -L

Như bạn có thể thấy, nhật ký này có rất nhiều dữ liệu trong đó.

**Nhiệm vụ 2: Sử dụng tệp nhật ký mới**

1. Để dừng dịch vụ tác nhân CloudWatch, hãy sử dụng lệnh systemctl Linux .

**Mẹo:** Để biết thêm thông tin, hãy xem [Dừng và khởi động lại CloudWatch Agent](https://docs.aws.amazon.com/AmazonCloudWatch/latest/monitoring/troubleshooting-CloudWatch-Agent.html#CloudWatch-Agent-troubleshooting-stopping-restarting).

1. Đặt nhật ký truy cập mô phỏng mới vào vị trí mà tác nhân CloudWatch dự kiến sẽ tìm thấy nhật ký đó.

**Quan trọng:** Nhật ký truy cập mô phỏng mới phải được đặt tên access\_log (*không* phải access\_log.log) để tác nhân sử dụng.

1. Khởi động lại dịch vụ tác nhân CloudWatch để dịch vụ này sẽ gửi các mục nhật ký mới đến CloudWatch.

**Tác vụ 3: Xác nhận rằng nhật ký mới xuất hiện trong nhóm nhật ký CloudWatch**

Bây giờ tệp nhật ký mới đã có sẵn và dịch vụ tác nhân CloudWatch đã được khởi động lại, bạn sẽ xác minh rằng nhật ký xuất hiện trong CloudWatch.

Trong bảng điều khiển CloudWatch, tìm nhóm nhật ký **apache/access**. Trong luồng nhật ký, hãy xác nhận rằng nhiều mục nhập xuất hiện trong luồng hơn bạn đã thấy trước đó.

**Mẹo:** Dấu hiệu cho thấy nhóm nhật ký hiện chứa dữ liệu nhật ký mới là nếu luồng có nhiều mục nhập có cùng dấu thời gian.

Bạn đã xác nhận rằng dữ liệu nhật ký mô phỏng hiện đang xuất hiện trong dòng nhật ký.

**Giai đoạn 6: Sử dụng CloudWatch Logs Insights để phân tích**

Bây giờ bạn đã xác nhận rằng tác nhân CloudWatch đang thu thập và gửi dữ liệu nhật ký mô phỏng đến CloudWatch, bạn sẽ sử dụng CloudWatch Logs Insights để chạy truy vấn trên dữ liệu access\_log.

Mục tiêu cụ thể của bạn trong giai đoạn này là sử dụng dữ liệu luồng nhấp chuột để xác định số lượng khách truy cập đã truy cập menu nhưng không mua bất cứ thứ gì.

**Nhiệm vụ 1: Xác định số lượng khách truy cập đã truy cập menu**

1. Trong ngăn điều hướng của bảng điều khiển CloudWatch, chọn **Logs Insights**.

**Mẹo:** Đối với tác vụ này và các tác vụ sau này, bạn có thể muốn tham khảo [Phân tích dữ liệu nhật ký bằng CloudWatch Logs Insights](https://docs.aws.amazon.com/AmazonCloudWatch/latest/logs/AnalyzingLogData.html).

1. Nhận số lượng người dùng đã truy cập menu:
   * Đặt khoảng thời gian của truy vấn để khớp với khoảng thời gian mà access\_log có dữ liệu.
   * Chạy truy vấn sau:



lĩnh vực @timestamp, remoteIP

| Yêu cầu bộ lọc = "/ cafe / menu.php"

| đếm (remoteIP) là Khách truy cập theo @timestamp

| Sắp xếp @timestamp ASC

Kết quả được tóm tắt trên tab **Nhật ký** với một thông báo theo định dạng sau (trong đó X và Y là các giá trị số):



Hiển thị 1000 bản ghi X phù hợp

Y ghi lại... Quét...

Dữ liệu chính cần lưu ý là số lượng bản ghi khớp (X) trong tổng số bản ghi đã được quét (Y).

* + Ghi lại thông tin này trong tệp văn bản mới trong IDE AWS Cloud9. Đặt tên cho tệp phase6-results.txt và lưu nó vào thư mục ~/environment. Ngoài ra, nếu người hướng dẫn của bạn đã yêu cầu bạn tạo một bài thuyết trình hoặc tài liệu để ghi lại công việc của bạn, hãy bao gồm thông tin ở đó.

1. Lưu truy vấn với các cài đặt sau:
   * **Tên truy vấn:** Nhập menu-khách truy cập
   * **Thư mục:** Chọn **Tạo mới** rồi nhập không phải kết quả địa lý

Lưu ý rằng truy vấn đã lưu xuất hiện trong bảng **Truy vấn**.

**Nhiệm vụ 2: Xác định số lượng khách truy cập đã mua hàng**

1. Xem số lượng người dùng đã mua hàng:
   * Đặt khoảng thời gian của truy vấn để khớp với khoảng thời gian mà access\_log có dữ liệu.
   * Chạy truy vấn từ tác vụ trước, nhưng thay đổi yêu cầu bộ lọc thành /cafe/processOrder.php
   * Ghi lại kết quả trong *kết quả giai đoạn6.txt tệp*, bản trình bày hoặc tài liệu của bạn.
2. Lưu truy vấn vào thư mục *không phải kết quả địa lý*. Sử dụng người mua cho tên truy vấn.

**Nhiệm vụ 3: Xác định số lượng khách truy cập đã truy cập menu nhưng không mua hàng**

1. Sử dụng kết quả của hai tác vụ trước để xác định có bao nhiêu người đã truy cập trang menu nhưng không mua hàng.

**Mẹo**: Công thức để có được số lượng này như sau:  *Số lượng khách truy cập đã truy cập menu nhưng không mua hàng = (Số lượng khách truy cập đã truy cập menu) - (Số lượng khách truy cập đã mua hàng)*

1. Tính tỷ lệ phần trăm những người đã truy cập trang menu và cũng đã mua hàng.
2. Ghi lại những kết quả này trong kết quả *giai đoạn6.txt tệp*, bản trình bày hoặc tài liệu của bạn.

Bạn đã sử dụng thành công các truy vấn trong CloudWatch Logs Insights để xác định có bao nhiêu khách truy cập trang web đã hoàn thành một số hành động nhất định.

**Giai đoạn 7: Điều chỉnh quy trình để cung cấp thông tin chuyên sâu mới**

Trong giai đoạn cuối cùng của dự án, bạn sẽ nâng cao dữ liệu trong quy trình của mình để cung cấp thông tin chi tiết mới cho chủ quán cà phê.

**Nhiệm vụ 1: Hiểu các yêu cầu**

Bạn gặp chủ quán cà phê và chia sẻ thông tin chi tiết mà bạn đã thu thập được từ dữ liệu luồng nhấp chuột. Chủ sở hữu rất ấn tượng với giải pháp mà bạn đã xây dựng cho đến nay, nhưng họ muốn biết thêm.

Chủ quán cà phê hỏi liệu bạn có thể sửa đổi giải pháp để cung cấp những điều sau đây không:

* Thông tin vị trí địa lý cho khách truy cập trang web.
* Bảng điều khiển trực quan để có được thông tin chi tiết. Bảng điều khiển phải bao gồm các thông tin sau:
  + Biểu *đồ hình tròn* hiển thị 10 thành phố có nhiều khách truy cập trang web nhất truy cập trang menu.
  + Bảng *nhật ký* hiển thị 10 thành phố có nhiều khách truy cập trang web nhất đã mua hàng.
  + Biểu *đồ hình tròn* hiển thị 10 khu vực có nhiều khách truy cập trang web nhất đã truy cập trang chính của trang web.
  + Biểu *đồ thanh* hiển thị 10 khu vực có nhiều khách truy cập trang web nhất đã mua hàng.
* Khả năng lưu trữ nhật ký truy cập trong Amazon S3 và khả năng truy vấn nhật ký bằng SQL

Thông tin vị trí địa lý và bảng điều khiển trực quan sẽ giúp chủ sở hữu phân tích xu hướng kinh doanh và xác định nơi mở địa điểm quán cà phê mới. Việc có thể lưu trữ dữ liệu trong Amazon S3 và sử dụng SQL để truy vấn dữ liệu sẽ mang lại sự linh hoạt khi sử dụng nhiều công cụ và dịch vụ khác nhau để phân tích dữ liệu trong tương lai.

Bạn có yêu cầu giải pháp mới. Các nhiệm vụ còn lại trong giai đoạn này cung cấp thêm chi tiết để giúp bạn đáp ứng các yêu cầu này.

**Nhiệm vụ 2: Sử dụng nhật ký mẫu bao gồm thông tin vị trí địa lý**

Trong tác vụ này, bạn sẽ thay thế access\_log hiện có bằng một tệp bao gồm thông tin vị trí địa lý.

1. Để phân tích access\_log mẫu bao gồm thông tin vị trí địa lý, hãy chạy các lệnh sau. Kết quả sẽ cung cấp thông tin về cấu trúc dữ liệu và số dòng trong tệp:



CD ~/môi trường

Đầu -1 Samplelogs/access\_log\_geo.log | Python -m json.tool

Nhật ký mẫu/access\_log\_geo.log mèo |WC -L

1. Thay thế access\_log hiện có trên máy chủ web bằng nhật ký mới có chứa thông tin vị trí địa lý. Đảm bảo rằng tác nhân CloudWatch sẽ thu thập dữ liệu và gửi đến CloudWatch.

Hãy nhớ lại rằng bạn đã thực hiện nhiệm vụ này trong giai đoạn 5, nhiệm vụ 2 và 3. Sự khác biệt lần này là bạn sẽ có thông tin vị trí địa lý trong access\_log mới và bạn cần xóa luồng nhật ký CloudWatch hiện có trước khi khởi động lại tác tử CloudWatch.

Ở cấp độ cao, các bước như sau:

* + Dừng tác nhân CloudWatch.
  + Sao chép nhật ký mới vào vị trí mà tác nhân CloudWatch dự kiến sẽ tìm thấy. Đảm bảo rằng nhật ký mới được đặt tên access\_log.
  + Để xác nhận rằng tệp access\_log đã được thay thế, hãy sử dụng lệnh sau:



sudo đầu -1 /var/log/www/access/access\_log

* + Xác minh rằng bạn thấy dữ liệu vị trí địa lý ("lat" và "lon") trong các mục nhật ký. Sau đây là một mục ví dụ:



{"time":"2023-02-22 03:38:18", "process":"5112", "filename":"/var/www/html/cafe", "remoteIP":"91.192.109.163", "host":"10.0.1.102", "request":"/cafe", "query":"" , "method":"GET", "status":"200", "useragent":"Mozilla/5.0 (Windows NT 10.0) AppleWebKit/531.2 (KHTML, như Gecko) Chrome/111.0.553.2.0.854.0", "referer":"-", "country":"ES", "region":"Andalusia" , "thành phố":"La Rinconada", "lat": "37.4867198", "lon" :"-5.9814868"}

1. Xóa nhóm nhật ký apache/access hiện có khỏi nhật ký CloudWatch, tạo lại nhóm nhật ký apache/access và khởi động lại dịch vụ tác nhân CloudWatch.

**Nhiệm vụ 3: Xây dựng bảng thông tin để quan sát dữ liệu vị trí địa lý trong CloudWatch Logs Insights**

**Tham khảo**

* [Trực quan hóa dữ liệu nhật ký trong biểu đồ](https://docs.aws.amazon.com/AmazonCloudWatch/latest/logs/CWL_Insights-Visualizing-Log-Data.html)

Trong tác vụ này, bạn sẽ xây dựng bảng thông tin CloudWatch để đáp ứng yêu cầu của chủ quán cà phê.

1. **Tạo bảng thông tin CloudWatch** có tên café-dashboard.

**Mẹo:** Trang [Sử dụng Bảng thông tin Amazon CloudWatch có](https://docs.aws.amazon.com/AmazonCloudWatch/latest/monitoring/CloudWatch_Dashboards.html) thể hữu ích.

1. **Thêm tiện ích biểu đồ hình tròn vào bảng điều khiển** để hiển thị thành phố nào có nhiều khách truy cập nhất vào trang menu của trang web quán cà phê. Sử dụng các cài đặt sau:
   * Loại widget: **Pie**
   * Nguồn dữ liệu: **Nhật ký**
   * Khoảng thời gian để truy vấn: **Tùy chỉnh** (Ngày bạn bắt đầu dự án đến hôm nay)
   * Nhóm nhật ký để truy vấn: **apache / access**
   * Truy vấn:



lĩnh vực từ xaIP, thành phố

| yêu cầu lọc = "/ cafe / menu.php"

| Số liệu thống kê count() như menuphổ biến theo thành phố

| Sắp xếp menuphổ biến desc

| Giới hạn 10

**Phân tích:** Truy vấn sẽ trả về các mục nhật ký với các trường remoteIP và city từ mỗi nhật ký, khi người dùng tải trang menu, được cung cấp thông qua bộ lọc của trường yêu cầu cho /cafe/menu.php. Truy vấn đếm số lượt truy cập vào trang menu cho mỗi thành phố và sau đó sắp xếp các thành phố đó theo thứ tự giảm dần (desc). Do đó, các thành phố có nhiều lượt truy cập nhất sẽ hiển thị ở đầu kết quả. Cuối cùng, truy vấn giới hạn kết quả trả về top 10.

* + Chạy truy vấn. Biểu đồ hình tròn xuất hiện.
  + Chọn **Tạo tiện ích**. Tiện ích được thêm vào bảng điều khiển.
  + Di chuột vào **nhóm Log: apache/access** title của widget và chọn biểu tượng chỉnh sửa.
  + Đổi tên tiện ích thành Thành phố truy cập menu nhiều nhất
  + Chọn **Áp dụng**.
  + Trên trang bảng thông tin, chọn **Lưu** ở góc trên cùng bên phải.

1. Sử dụng cách tiếp cận tương tự để **thêm tiện ích thứ hai**. Tiện ích sẽ hiển thị 10 thành phố có nhiều khách truy cập trang web nhất đã mua hàng:
   * Sử dụng bảng nhật ký.
   * Đặt tên cho tiện ích là Thành phố đặt nhiều đơn đặt hàng nhất.
   * Tiện ích phải bao gồm cùng khoảng thời gian với tiện ích đầu tiên.
   * Truy vấn sẽ lọc các yêu cầu đến trang /cafe/processOrder.php.
2. Sử dụng cách tiếp cận tương tự để **thêm tiện ích thứ ba**. Tiện ích phải hiển thị 10 khu vực có nhiều khách truy cập trang web nhất đã truy cập trang chính của trang web:
   * Sử dụng biểu đồ tròn.
   * Đặt tên tiện ích là Khu vực truy cập trang web nhiều nhất.
   * Tiện ích phải bao gồm cùng khoảng thời gian với tiện ích đầu tiên.
   * Truy vấn sẽ lọc các yêu cầu đến trang /cafe.
3. Sử dụng cách tiếp cận tương tự để **thêm tiện ích thứ tư**. Tiện ích sẽ hiển thị 10 khu vực có nhiều khách truy cập trang web nhất đã mua hàng:
   * Sử dụng biểu đồ thanh.
   * Đặt tên cho tiện ích là Khu vực đặt nhiều đơn đặt hàng nhất.
   * Tiện ích phải bao gồm cùng khoảng thời gian với tiện ích đầu tiên.
4. Sau khi bạn thêm tất cả bốn tiện ích vào bảng thông tin, hãy đảm bảo rằng bạn lưu bảng thông tin.

**Lưu ý:** Nếu người hướng dẫn yêu cầu bạn tạo bài thuyết trình hoặc tài liệu để ghi lại thành tích của bạn, hãy lưu truy vấn và ảnh chụp màn hình của bảng thông tin CloudWatch cuối cùng vào bài thuyết trình hoặc tài liệu của bạn.

**Tác vụ 4: Lưu tệp nhật ký vào vùng lưu trữ S3**

Trong nhiệm vụ này, bạn bắt đầu thực hiện yêu cầu cuối cùng, như được mô tả trong giai đoạn 7, nhiệm vụ 1.

Trong thiết bị đầu cuối AWS Cloud9, sử dụng AWS CLI để sao chép nhật ký truy cập vị trí địa lý từ thư mục /var/log/www/access hoặc thư mục ~/environment/samplelogs sang vùng lưu trữ S3 được cung cấp cho bạn trong môi trường phòng thực hành.

**Lưu ý:** Tên vùng lưu trữ chứa *logbucket*.

**Mẹo:** Tham [chiếu dòng lệnh AWS CLI dành cho Amazon S3](https://docs.aws.amazon.com/cli/latest/reference/s3/index.html#cli-aws-s3) có thể hữu ích để tham khảo.

**Nhiệm vụ 5: Sử dụng S3 Select để truy vấn nhật ký được lưu trữ trong Amazon S3**

Trong nhiệm vụ này, bạn hoàn thành yêu cầu cuối cùng của giải pháp.

Sử dụng S3 Select để truy vấn nhật ký hiện được lưu trữ trong Amazon S3. So sánh kết quả từ S3 Select với kết quả CloudWatch Insights.

**Mẹo:**  [Trang Lọc và truy xuất dữ liệu bằng Amazon S3 Select](https://docs.aws.amazon.com/AmazonS3/latest/userguide/selecting-content-from-objects.html) có thể hữu ích để tham khảo.

Bạn đã triển khai thành công dữ liệu vị trí địa lý trong giải pháp luồng nhấp chuột của mình. Bạn đã xây dựng một bảng điều khiển để nhận thông tin chi tiết từ dữ liệu mới và bạn đã điều chỉnh giải pháp để cung cấp sự linh hoạt để sử dụng các dịch vụ và tính năng khác để phân tích dữ liệu trong tương lai.

**Kết thúc phiên của bạn**

**Nhắc nhở:** Đây là một môi trường phòng thí nghiệm tồn tại lâu dài. Dữ liệu được giữ lại cho đến khi bạn sử dụng ngân sách được phân bổ hoặc đạt đến ngày kết thúc khóa học (tùy theo điều kiện nào xảy ra trước).

Để bảo toàn ngân sách của bạn khi bạn kết thúc trong ngày hoặc khi bạn kết thúc tích cực làm việc trên nhiệm vụ trong thời gian này, hãy làm như sau:

1. Ở đầu trang này, chọn **Kết thúc Phòng thí nghiệm**, sau đó chọn **Có** để xác nhận rằng bạn muốn kết thúc phòng thực hành.

Một bảng thông báo cho biết rằng phòng thí nghiệm đang chấm dứt.

**Lưu ý:** Chọn **End Lab** trong môi trường này sẽ *không* xóa tài nguyên bạn đã tạo. Họ sẽ vẫn ở đó vào lần tiếp theo bạn chọn Bắt đầu phòng thí nghiệm (ví dụ: vào một ngày khác).

1. Để đóng bảng điều khiển, hãy chọn **Đóng** ở góc trên bên phải.

© Năm 2023, Amazon Web Services, Inc. hoặc các chi nhánh của nó. Đã đăng ký Bản quyền.