



開發人員指南

# Amazon 數據 Firehose



# Amazon 數據 Firehose: 開發人員指南

Copyright © 2024 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon 的商標和商業外觀不得用於任何非 Amazon 的產品或服務，也不能以任何可能造成客戶混淆、任何貶低或使 Amazon 名譽受損的方式使用 Amazon 的商標和商業外觀。所有其他非 Amazon 擁有的商標均為其各自擁有者的財產，這些擁有者可能附屬於 Amazon，或與 Amazon 有合作關係，亦或受到 Amazon 贊助。

# Table of Contents

.....	ix
什麼是 Amazon 數據 Firehose ? .....	1
重要概念 .....	1
資料流程 .....	2
設定 .....	4
註冊成為 AWS .....	4
選用：下載程式庫和工具 .....	4
建立 Firehose 串流 .....	6
來源、目的地和名稱 .....	6
記錄轉換和格式轉換 .....	8
目的地設定 .....	10
選擇 Amazon S3 作為目的地 .....	11
針對目的地，選擇 Amazon Redshift .....	13
選擇適合您目的地的 OpenSearch 服務 .....	18
為您的目的地選擇 OpenSearch 無伺服器 .....	20
為您的目的地選擇 HTTP 端點 .....	21
選擇 Datadog 作為目的地 .....	22
為您的目的地選擇 Honeycomb .....	24
選擇 Coralogix 作為您的目的地 .....	25
選擇 Dynatrace 作為目的地 .....	26
選擇您 LogicMonitor 的目的地 .....	28
選擇 Logz.io 作為目的地 .....	29
選擇 MongoDB Cloud 作為目的地 .....	30
選擇 New Relic 作為目的地 .....	31
選擇雪花作為您的目的地 .....	32
選擇 Splunk 做為目的地 .....	34
選擇 Splunk Observability Cloud 作為您的目的地 .....	36
選擇 Sumo Logic 作為目的地 .....	37
選擇 Elastic 作為目的地 .....	38
備份和進階設定 .....	39
備份設定 .....	39
進階設定 .....	40
緩衝提示 .....	42
測試您的 Firehose 串流 .....	45

必要條件 .....	45
使用 Amazon S3 作為目的地進行測試 .....	45
使用 Amazon Redshift 作為目的地進行測試 .....	46
使用 OpenSearch 服務作為目的地進行測試 .....	46
使用 Splunk 做為目的地進行測試 .....	47
將資料傳送至 Firehose 串流 .....	48
使用 Kinesis Data Streams 寫入 .....	48
使用 Amazon MSK 寫入 .....	50
使用 Amazon 數據 Firehose 代理程序編寫 .....	51
必要條件 .....	52
登入資料 .....	52
自訂登入資料提供者 .....	53
下載並安裝代理程式 .....	53
設定和啟動代理程式 .....	55
代理程式組態設定 .....	56
監控多個檔案目錄並寫入到多個串流 .....	60
使用代理程式預先處理資料 .....	60
代理程式 CLI 命令 .....	65
常見問答集 .....	65
使用 AWS SDK 寫入 .....	66
使用單一寫入作業 PutRecord .....	67
Batch 寫入作業使用 PutRecordBatch .....	67
使用 CloudWatch 記錄寫入 .....	68
記錄檔解壓縮 CloudWatch .....	68
記錄檔解壓縮後的訊息擷取 CloudWatch .....	68
啟用和停用解壓縮 .....	69
常見問答集 .....	65
使用 CloudWatch 事件撰寫 .....	72
使用 AWS IoT 編寫 .....	72
安全 .....	74
資料保護 .....	74
將 Kinesis Data Streams 視為資料來源的伺服器端加密 .....	75
使用 Direct PUT 或其他資料來源的伺服器端加密 .....	75
控制存取 .....	76
授予您的應用程式對 Amazon 資料 Firehose 資源的存取權 .....	77
授予 Amazon 數據 Firehose 訪問您的私有 Amazon MSK 集群 .....	78

允許 Amazon 數據 Firehose 承擔 IAM 角色 .....	78
授予 Amazon 數據 Firehose 訪問權限以 AWS Glue 進行數據格式轉換 .....	80
授予 Amazon 數據 Firehose 訪問 Amazon S3 目的地 .....	81
授予 Amazon 數據 Firehose 訪問亞 Amazon Redshift 目的地 .....	84
授予 Amazon 資料 Firehose 存取公共 OpenSearch 服務目的地 .....	88
授予 Amazon 資料 Firehose 存取 VPC 中的 OpenSearch 服務目的地 .....	91
授予 Amazon 資料 Firehose 存取公有 OpenSearch無伺服器目的地 .....	92
授予 Amazon 資料 Firehose 存取 VPC 中的 OpenSearch無伺服器目的地 .....	94
授予 Amazon 數據 Firehose 訪問 Splunk 目的地 .....	95
在 VPC 中存取 Splunk .....	98
訪問雪花或 HTTP 端點 .....	99
授予 Firehose 進入雪花目的地 .....	99
授予 Amazon 資料 Firehose 存取 HTTP 端點目的地 .....	101
從 Amazon MSK 跨帳戶交付 .....	104
跨帳戶交付至 Amazon S3 目的地 .....	107
跨帳戶交付至 OpenSearch服務目的地 .....	108
使用標籤控制存取 .....	109
監控 .....	112
合規驗證 .....	112
恢復能力 .....	112
災難復原 .....	113
基礎設施安全性 .....	113
VPC 端點 (PrivateLink) .....	113
安全最佳實務 .....	113
實作最低權限存取 .....	114
使用 IAM 角色 .....	114
在相依資源實作伺服器端加密 .....	114
用 CloudTrail 於監控 API 呼叫 .....	114
資料轉換 .....	115
資料轉換流程 .....	115
資料轉換和狀態模型 .....	115
Lambda 藍圖 .....	117
資料轉換失敗的處理方式 .....	118
Lambda 調用的持續時間 .....	119
來源記錄備份 .....	119
動態分割 .....	120

分割索引鍵 .....	120
使用內嵌剖析建立分割索引鍵 .....	121
使用 AWS Lambda 函數建立分割索引鍵 .....	122
用於動態分割的 Amazon S3 儲存貯體字首 .....	125
彙整資料的動態分割 .....	126
在將資料傳送至 S3 時新增新的行分隔符號 .....	127
如何啟用動態分割 .....	127
動態分割錯誤處理 .....	128
資料緩衝和動態分割 .....	128
記錄格式轉換 .....	129
記錄格式轉換要求 .....	129
選擇 JSON 還原序列化程式 .....	130
選擇序列化程式 .....	131
轉換輸入記錄格式 (主控台) .....	131
轉換輸入記錄格式 (API) .....	132
記錄格式轉換錯誤處理 .....	132
記錄格式轉換範例 .....	133
與 Managed Service in Apache Flink 整合 .....	134
資料交付 .....	135
資料交付格式 .....	135
資料交付頻率 .....	136
資料交付失敗的處理方式 .....	136
Amazon S3 物件名稱格式 .....	139
OpenSearch 服務目的地的索引輪替 .....	148
HTTP 端點目的地跨 AWS 帳戶和跨 AWS 區域交付 .....	149
重複的記錄 .....	149
如何暫停和繼續 Firehose 傳送串流 .....	149
了解 Firehose 如何處理派送失敗 .....	150
暫停「Firehose」傳送串流 .....	150
恢復 Firehose 交付流 .....	150
監控 .....	151
CloudWatch 警示的最佳實務 .....	151
以測 CloudWatch 量結果監視 .....	152
動態分割 CloudWatch 指標 .....	152
資料傳送 CloudWatch 指標 .....	153
資料導入指標 .....	162

API 層 CloudWatch 級測量結果 .....	168
資料轉換 CloudWatch 指標 .....	171
CloudWatch 日誌解壓縮測量結果 .....	171
格式轉換 CloudWatch 量度 .....	172
伺服器端加密 (SSE) CloudWatch 量度 .....	172
Amazon 數據 Firehose 的尺寸 .....	173
Amazon 數據 Firehose 使用指標 .....	173
訪問 CloudWatch Amazon 數據 Firehose 的指標 .....	174
使用 CloudWatch 記錄監控 .....	175
資料交付錯誤 .....	176
訪問 Amazon 數據 Firehose 的 CloudWatch 日誌 .....	209
監控代理程式運作狀態 .....	209
使用監控 CloudWatch .....	210
使用日誌記錄 Amazon 數據 Firehose API 調用 AWS CloudTrail .....	211
Amazon 數據 Firehose 信息 CloudTrail .....	211
範例：Amazon 資料 Firehose 記錄檔項目 .....	212
自訂 Amazon S3 字首 .....	217
timestamp 命名空間 .....	217
firehose 命名空間 .....	217
partitionKeyFromLambda 和 partitionKeyFromQuery 命名空間 .....	219
語義規則 .....	219
範例字首 .....	220
使用 Amazon 數據 Firehose AWS PrivateLink .....	222
適用於 Amazon 資料 Firehose 的 VPC 端點 (AWS PrivateLink) .....	222
使用接口 VPC 端點 ( AWS PrivateLink ) 用於 Amazon 數據 Firehose 件 .....	222
可用性 .....	225
標記您的 Firehose 串流 .....	227
標籤基本概念 .....	227
使用標記追蹤成本 .....	228
標籤限制 .....	228
使用 Amazon 數據 Firehose 火管 API 標記火管流 .....	229
教學課程：使用 Amazon 資料 Firehose 將 VPC 流程日誌導入 Splunk .....	230
故障診斷 .....	231
故障診斷 Amazon S3 .....	232
疑難排解 Amazon Redshift .....	232
Amazon OpenSearch 服務故障 .....	233

疑難排解 Splunk .....	234
疑難排解雪 .....	236
Firehose 交付串流建立失敗 .....	236
疑難排解 Firehose 端點可達性 .....	237
疑難排解 HTTP 端點 .....	237
CloudWatch 日誌 .....	238
疑難排解 MSK As A Source .....	240
Hose 建立失敗 .....	241
Hose 暫停 .....	241
Hose 預壓 .....	241
資料新鮮度錯誤 .....	242
MSK 叢集連線問題 .....	242
其他 .....	244
傳送串流無法做為 CloudWatch 記錄檔、 CloudWatch 事件或 AWS IoT 動作的目標使用 .....	245
資料新鮮度指標不斷增加或未發出 .....	245
Apache Parquet 的記錄格式轉換失敗 .....	246
雖然指標良好，但目的地卻沒有資料 .....	247
配額 .....	248
附錄 – HTTP 端點交付請求和回應規格 .....	251
要求格式 .....	251
回應格式 .....	254
範例 .....	257
文件歷史記錄 .....	258
AWS 詞彙表 .....	261



Amazon 數據 Firehose 以前被稱為 Amazon Kinesis Data Firehose

本文為英文版的機器翻譯版本，如內容有任何歧義或不一致之處，概以英文版為準。

# 什麼是 Amazon 數據 Firehose ？

Amazon 資料 Firehose 是一項全受管服務，可將即時[串流資料](#)交付到 Amazon 簡單儲存服務 (Amazon S3)、Amazon Redshift、Amazon 服務、亞馬遜 OpenSearch 無伺服器 OpenSearch 服務、Splunk 以及任何自訂 HTTP 端點或 HTTP 端點由支援的第三方服務提供者擁有的任何自訂 HTTP 端點或 HTTP 端點，包括資料多、重新遺物、Coralogix 和彈性。LogicMonitor 使用 Amazon 資料 Firehose，您不需要撰寫應用程式或管理資源。您可以將資料生產者設定為將資料傳送到 Amazon Data Firehose，它會自動將資料傳送到您指定的目的地。您也可以將 Amazon 資料 Firehose 設定為在交付資料之前轉換資料。

如需 AWS 大數據解決方案的詳細資訊，請參閱 [AWS 大數據](#)。如需 AWS 串流資料解決方案的詳細資訊，請參閱 [什麼是串流資料？](#)

## Note

請注意最新的 [Amazon MSK AWS 串流資料解決方案](#)，該解決方案提供 AWS CloudFormation 範本，讓資料流經生產者、串流儲存、消費者和目的地。

## 重要概念

當您開始使用 Amazon 資料 Firehose 時，您可以從瞭解下列概念中受益：

### Firehose 溪

Amazon 數據 Firehose 件的基礎實體。您可以通過創建一個 Firehose 流，然後將數據發送到該流來使用 Amazon 數據防火軟管。如需詳細資訊，請參閱 [建立 Firehose 串流](#) 及 [將資料傳送至 Firehose 串流](#)。

### record

您的資料生產者會傳送至 Firehose 串流的感興趣資料。記錄最大可達 1,000 KB。

### 資料產生來源

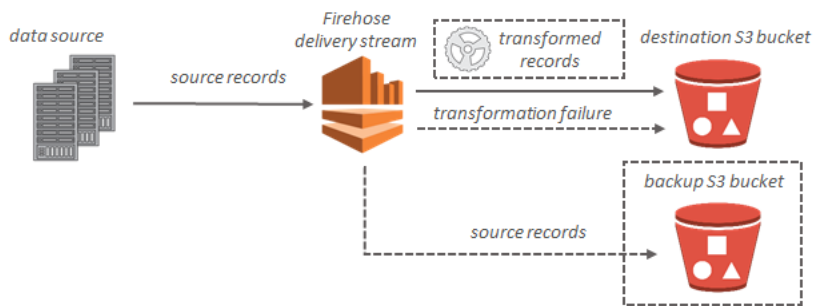
製作人將記錄發送到 Firehose 流。例如，將記錄資料傳送至 Firehose 串流的 Web 伺服器就是資料生產者。您也可以將 Firehose 串流設定為自動從現有 Kinesis 資料串流讀取資料，並將其載入目標。如需詳細資訊，請參閱 [將資料傳送至 Firehose 串流](#)。

## 緩衝區大小和緩衝間隔

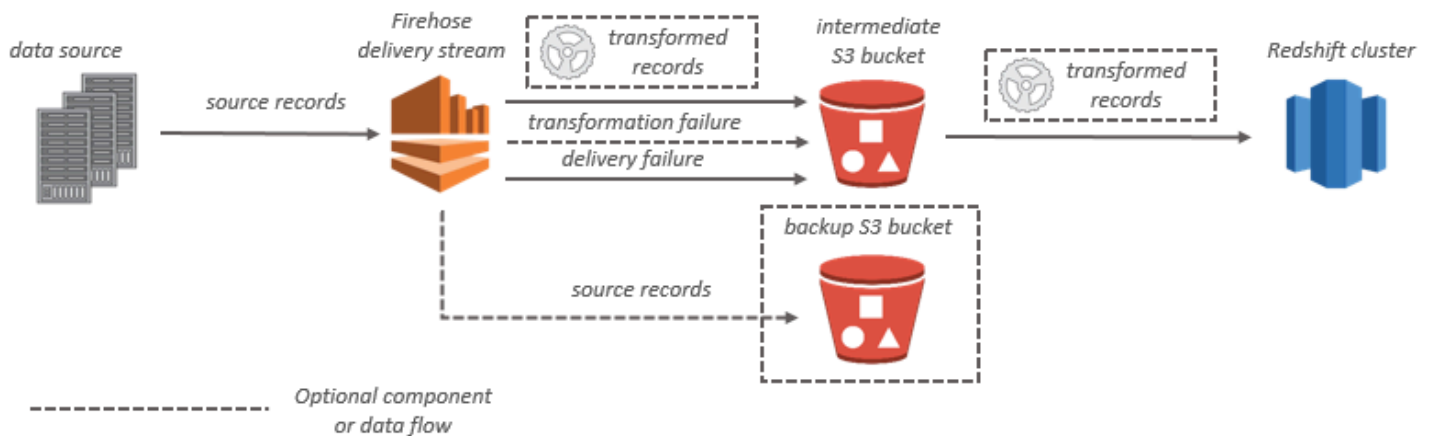
Amazon Data Firehose 會將傳入的串流資料緩衝為特定大小或一段時間，然後再將其傳送到目的地。Buffer Size 以 MB 為單位，並 Buffer Interval 且以秒為單位。

## 資料流程

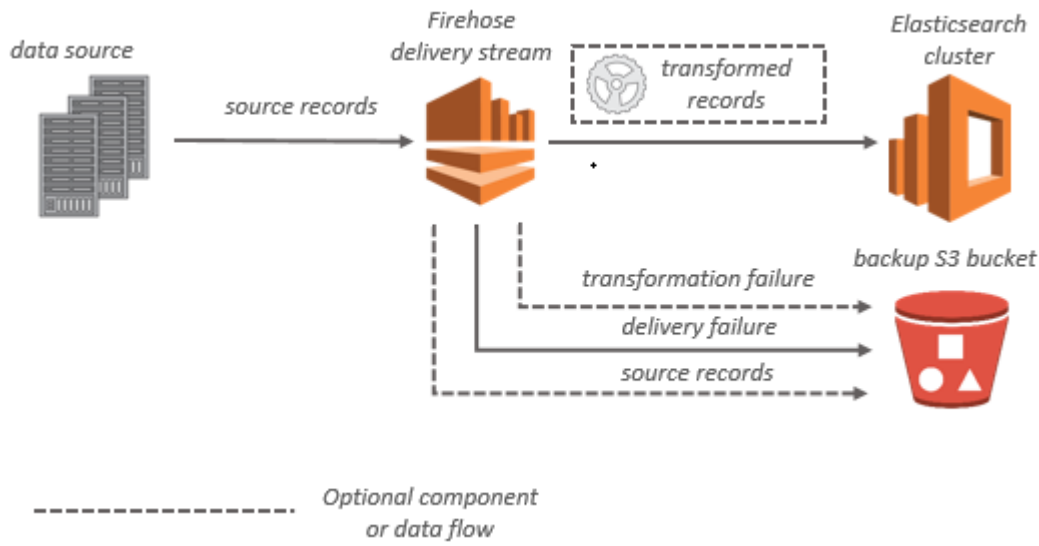
以 Amazon S3 目的地而言，串流資料都交付至您的 S3 儲存貯體。如果已啟用資料轉換，您可以選擇性將原始資料備份到另一個 Amazon S3 儲存貯體。



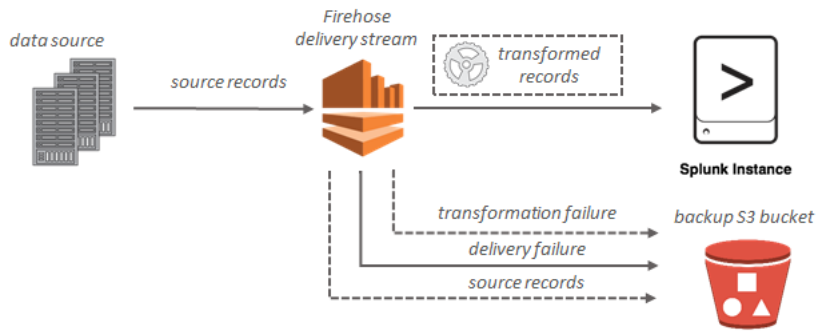
以 Amazon Redshift 目的地而言，串流資料會先交付至您的 S3 儲存貯體。然後，Amazon 數據 Firehose 發出 Amazon Redshift COPY 命令，將數據從 S3 存儲桶加載到 Amazon Redshift 集群。如果已啟用資料轉換，您可以選擇性將原始資料備份到另一個 Amazon S3 儲存貯體。



對於 OpenSearch 服務目的地，串流資料會傳遞到您的 OpenSearch 服務叢集，並且可以選擇性地同時備份到 S3 儲存貯體。



以 Splunk 目的地而言，串流資料會交付到 Splunk，然後可以選擇性同時備份到 S3 儲存貯體。



# 設置 Amazon 數據 Firehose

第一次使用 Amazon 資料 Firehose 之前，請先完成以下任務。

## 任務

- [註冊成為 AWS](#)
- [選用：下載程式庫和工具](#)

## 註冊成為 AWS

當您註冊 Amazon Web Services (AWS) 時，您的 AWS 帳戶將自動註冊為中的所有服務 AWS，包括 Amazon 數據 Firehose。您只需支付實際使用服務的費用。

如果您已經有 AWS 帳號，請跳至下一個工作。若您尚未擁有 AWS 帳戶，請使用下列程序建立帳戶。

### 註冊一個帳戶 AWS 戶

1. 開啟 <https://portal.aws.amazon.com/billing/signup>。
2. 請遵循線上指示進行。

部分註冊程序需接收來電，並在電話鍵盤輸入驗證碼。

當您註冊時 AWS 帳戶，會建立 AWS 帳戶根使用者一個。根使用者有權存取該帳戶中的所有 AWS 服務和資源。作為最佳安全實務，[將管理存取權指派給管理使用者](#)，並且僅使用根使用者來執行[需要根使用者存取權的任務](#)。

## 選用：下載程式庫和工具

下列程式庫和工具將協助您以程式設計方式和命令列使用 Amazon 資料 Firehose：

- [Firehose API 操作](#)是 Amazon 資料 Firehose 支援的基本操作集。
- [圍棋](#)，[Java](#)，[.NET](#)，[Node.js](#)，[Python](#) 和 [紅寶石](#)的 AWS SDK 包括 Amazon 數據 Firehose 支持和示例。

如果您的版本 AWS SDK for Java 不包含 Amazon 資料 Firehose 的範例，您也可以從[GitHub](#)下載最新的 AWS 開發套件。

- 該[AWS Command Line Interface](#)支持 Amazon 數據 Firehose。AWS CLI 可讓您從命令列控制多個 AWS 服務，並透過指令碼將其自動化。

# 建立 Firehose 串流

您可以使用AWS Management Console或 AWS SDK 將 Firehose 串流建立至您選擇的目的地。

您可以在建立 Firehose 串流之後隨時使用 Amazon 資料 Firehose 主控台或更新其設定。[UpdateDestination](#)當您的組態更新時，Firehose 串流會保持在ACTIVE狀態，而且您可以繼續傳送資料。更新的組態通常會在幾分鐘內生效。當您更新組態1之後，Firehose 串流的版本號碼會增加一定的值。並出現在交付的 Amazon S3 物件名稱中。如需詳細資訊，請參閱 [Amazon S3 物件名稱格式](#)。

下列主題說明如何建立 Firehose 串流：

## 主題

- [來源、目的地和名稱](#)
- [記錄轉換和格式轉換](#)
- [目的地設定](#)
- [備份和進階設定](#)
- [緩衝提示](#)

## 來源、目的地和名稱

1. 登入 AWS Management Console 並開啟 Amazon 資料 Firehose 主控台，網址為 <https://console.aws.amazon.com/firehose>
2. 選擇「建立 Firehose 串流」。
3. 為下列欄位輸入值：

### 來源

- **直接放置**：選擇此選項可建立製作者應用程式直接寫入的 Firehose 串流。目前，以下是 AWS 與 Amazon 資料 Firehose 中的 Direct PUT 整合的服務和代理程式和開放原始碼服務：
  - AWS SDK
  - AWS Lambda
  - AWS CloudWatch 日誌
  - AWS CloudWatch 活動
  - AWS 雲端指標串流

- AWS 物聯網
- AWS Eventbridge
- Amazon Simple Email Service
- Amazon SNS
- AWS WAF 網絡 ACL 記錄
- Amazon API Gateway - 存取日誌
- Amazon Pinpoint
- Amazon MSK 代理程式日誌
- Amazon Route 53 Resolver 查詢日誌
- AWS Network Firewall 警示記錄
- AWS Network Firewall 流程記錄
- Amazon Elasticache Redis SLOWLOG
- Kinesis 代理程式 (linux)
- Kinesis Tap (windows)
- Fluentbit
- Fluentd
- Apache Nifi
- Snowflake
- Kinesis 串流：選擇此選項可設定使用 Kinesis 資料串流作為資料來源的 Firehose 串流。然後，您可以使用 Amazon 資料 Firehose 輕鬆地從現有 Kinesis 資料串流讀取資料，並將其載入目的地。如需使用 Kinesis 資料串流做為資料來源的詳細資訊，請參閱[使用 Kinesis 資料串流寫入 Amazon 資料 Firehose](#)。
- Amazon MSK：選擇此選項可設定使用 Amazon MSK 做為資料來源的 Firehose 串流。然後，您可以使用 Firehose 輕鬆地從現有的 Amazon MSK 叢集讀取資料，並將其載入到指定的 S3 儲存貯體。如需有關使用 Amazon MSK 做為資料來源的詳細資訊，請參閱[使用 Amazon MSK 寫入 Amazon 資料 Firehose](#)。

## Firehose 目的地

Firehose 串流的目的地。Amazon 數據 Firehose 可以將數據記錄發送到各種目的地，包括 Amazon 簡單存儲服務 ( Amazon S3 )，亞馬遜 Redshift，亞馬遜 OpenSearch 服務以及您或任何第三方服務提供商擁有的任何 HTTP 端點。以下是支援的目的地：



- Amazon OpenSearch 無服務器
- Amazon Redshift
- Amazon S3
- Coralogix
- Datadog
- Dynatrace
- 彈性
- HTTP 端點
- Honeycomb
- Logic Monitor
- Logz.io
- MongoDB Cloud
- New Relic
- Splunk
- Splunk Observability Cloud
- Sumo Logic
- Snowflake

### Firehose 分流名稱

您的「Firehose」串流的名稱。

## 記錄轉換和格式轉換

設定 Amazon 資料 Firehose 以轉換和轉換您的記錄資料。

- 如果您選擇 Amazon MSK 做為交付串流的來源：
  1. 在透過 AWS Lambda 轉換來源記錄區段中，提供下列欄位的值：

### 資料轉換

若要建立不會轉換傳入資料的 Firehose 串流，請勿勾選「啟用資料轉換」核取方塊。

若要指定 Firehose 叫用的 Lambda 函數，並在傳送資料之前用來轉換內送資料，請勾選「啟用資料轉換」核取方塊。您可以使用其中一個 Lambda 藍圖，或選擇現有的 Lambda 函數，進而設定新的 Lambda 函數。您的 Lambda 函數必須包含 Firehose 所需的狀態模型。如需詳細資訊，請參閱 [Amazon 數據 Firehose 數據轉換](#)。

2. 在 Convert record format (轉換記錄格式) 區段中，提供以下欄位的值：

#### 記錄格式轉換

若要建立不轉換傳入資料記錄格式的 Firehose 串流，請選擇 [停用]。

若要轉換傳入記錄的格式，請選擇 Enabled (已啟用)，然後指定輸出格式。您需要指定一個 AWS Glue 表，其中包含您希望 Firehose 用於轉換記錄格式的結構描述。如需詳細資訊，請參閱 [記錄格式轉換](#)。

如需如何設定記錄格式轉換的範例 AWS CloudFormation，請參閱 [AWS:KinesisFirehose::DeliveryStream](#)。

- 如果您選擇 Apache Flink 的受管理服務或直接 PUT 作為傳送串流的來源，請在 [來源設定] 區段中：

1. 在 [轉換記錄] 底下，選擇下列其中一個選項：

- a. 如果您的目的地是 Amazon S3 或 Splunk，請在「解壓縮來源記錄 Amazon CloudWatch 日誌」區段中選擇「開啟解壓縮」。
- b. 在透過 AWS Lambda 轉換來源記錄區段中，提供下列欄位的值：

#### 資料轉換

若要建立不會轉換傳入資料的 Firehose 串流，請勿勾選「啟用資料轉換」核取方塊。

若要指定 Amazon 資料 Firehose 叫用的 Lambda 函數，並在交付傳入資料之前用於轉換傳入資料，請勾選啟用資料轉換核取方塊。您可以使用其中一個 Lambda 藍圖，或選擇現有的 Lambda 函數，進而設定新的 Lambda 函數。您的 Lambda 函數必須包含 Amazon 資料 Firehose 所需的狀態模型。如需詳細資訊，請參閱 [Amazon 數據 Firehose 數據轉換](#)。

2. 在 Convert record format (轉換記錄格式) 區段中，提供以下欄位的值：

#### 記錄格式轉換

若要建立不轉換傳入資料記錄格式的 Firehose 串流，請選擇 [停用]。

若要轉換傳入記錄的格式，請選擇 Enabled (已啟用)，然後指定輸出格式。您需要指定一個 AWS Glue 表，其中包含您希望 Amazon Data Firehose 用於轉換記錄格式的結構描述。如需詳細資訊，請參閱 [記錄格式轉換](#)。

如需如何設定記錄格式轉換的範例 AWS CloudFormation，請參閱 [AWS:KinesisFirehose::DeliveryStream](#)。

## 目的地設定

本主題會描述交付串流的目的地設定。如需緩衝提示的更多資訊，請參閱 [緩衝提示](#)。

### 主題

- [選擇 Amazon S3 作為目的地](#)
- [針對目的地，選擇 Amazon Redshift](#)
- [選擇適合您目的地的 OpenSearch 服務](#)
- [為您的目的地選擇 OpenSearch 無伺服器](#)
- [為您的目的地選擇 HTTP 端點](#)
- [選擇 Datadog 作為目的地](#)
- [為您的目的地選擇 Honeycomb](#)
- [選擇 Coralogix 作為您的目的地](#)
- [選擇 Dynatrace 作為目的地](#)
- [選擇您 LogicMonitor 的目的地](#)
- [選擇 Logz.io 作為目的地](#)
- [選擇 MongoDB Cloud 作為目的地](#)
- [選擇 New Relic 作為目的地](#)
- [選擇雪花作為您的目的地](#)
- [選擇 Splunk 做為目的地](#)
- [選擇 Splunk Observability Cloud 作為您的目的地](#)
- [選擇 Sumo Logic 作為目的地](#)
- [選擇 Elastic 作為目的地](#)

## 選擇 Amazon S3 作為目的地

您必須指定下列設定，才能使用 Amazon S3 做為 Firehose 串流的目的地。

- 請輸入下列欄位的值。

### S3 bucket (S3 儲存貯體)

選擇要交付串流資料您所擁有的 S3 儲存貯體。您可以建立新的 S3 儲存貯體或選擇現有的。

### 新行分隔符

您可以設定交付串流，以在交付到 Amazon S3 的物件中的記錄之間新增明細行分隔符號。若要這麼做，請選擇已啟用。若要在交付至 Amazon S3 的物件中的記錄之間新增行分隔符號，請選擇停用。如果您計劃使用 Athena 查詢具有彙總記錄的 S3 物件，請啟用此選項。

### 動態分割

選擇已啟用以啟用並設定動態分割。

### 多記錄解彙

此程序透過交付串流中的記錄進行剖析，並根據有效 JSON 或指定的新行分隔符號來分隔這些記錄。

如果您將多個事件、記錄或記錄彙總到單一 PutRecord 和 PutRecordBatch API 呼叫中，您仍然可以啟用和設定動態磁碟分割。透過彙總資料，當您啟用動態分割時，Amazon Data Firehose 會剖析記錄，並在每個 API 呼叫中尋找多個有效的 JSON 物件。當 Firehose 串流設定為 Kinesis 資料串流作為來源時，您也可以使用 Kinesis 製作者程式庫 (KPL) 中的內建彙總。資料分割功能在資料取消彙整後執行。因此，每個 API 呼叫中的每個記錄都可以交付到不同的 Amazon S3 字首。您也可以利用 Lambda 函數整合，在資料分割功能之前執行任何其他取消彙整或任何其他轉換。

#### Important

如果您的資料已彙整，則只有在執行資料取消彙整之後才能套用動態分割。因此，如果您對彙整資料啟用動態分割，則必須選擇已啟用以啟用多記錄取消彙整。

Firehose 串流會依照下列順序執行下列處理步驟：KPL (原始化) 解除彙總、JSON 或分隔符號解除彙總、Lambda 處理、資料分割、資料格式轉換以及 Amazon S3 交付。

## 多記錄解彙總類型

如果您啟用了多記錄解彙總，則必須指定 Firehose 解彙總資料的方法。使用下拉式功能表選擇 JSON 或分隔符號。

## 內嵌剖析

這是其中一種支援的機制，可動態分割繫結到 Amazon S3 的資料。若要使用內嵌剖析來動態分割資料，您必須指定要用作分割索引鍵的資料記錄參數，並為每個指定的分割索引鍵提供一個值。選擇已啟用以啟用並設定內嵌剖析。

### Important

如果您在上述步驟中指定 AWS Lambda 函數來轉換來源記錄，您可以使用此函數來動態分割繫結至 S3 的資料，而且您仍然可以使用內嵌剖析來建立分割索引鍵。透過動態分割，您可以使用內嵌剖析或 AWS Lambda 函數來建立分割金鑰。或者，您也可以同時使用內嵌剖析和 AWS Lambda 函數來建立分割金鑰。

## 動態分割索引鍵

您可以使用索引鍵和值欄位來指定要用作動態分割索引鍵的資料記錄參數，以及用來產生動態分割索引鍵值的 jq 查詢。Firehose 只支援 jq 1.6。您可以指定最多 50 個動態分割索引鍵。您必須為動態分區索引鍵值輸入有效的 jq 運算式，才能成功設定 Firehose 串流的動態磁碟分割。

## S3 儲存貯體的字首

啟用和設定動態分割時，您必須指定 Amazon Data Firehose 要交付分區資料的 S3 儲存貯體前置詞。

為了正確設定動態分割，S3 儲存貯體字首的數目必須與指定的分割索引鍵數目相同。

您可以使用內嵌剖析或指定的 AWS Lambda 函數來分割來源資料。如果您指定 AWS Lambda 函數為來源資料建立分割金鑰，則必須使用下列格式手動輸入 S3 儲存貯體前置詞值："L partitionKeyFrom lambda: keyID"。如果您使用內嵌剖析來指定來源資料的分割索引鍵，您可以使用下列格式手動輸入 S3 儲存貯體預覽值："partitionKeyFromquery: keyID"，或者選擇 [套用動態分割金鑰] 按鈕，使用動態分割金鑰/值組來自動產生 S3 儲存貯體前綴。使用內嵌剖析或 AWS Lambda 對資料進行分割時，您也可以使用 S3 儲存貯體前綴中使用下

列運算式表單：`!{命名空間：值}`，其中命名空間可以是查詢或 `partitionKeyFrom Lambda.partitionKeyFrom`

### S3 儲存貯體和 S3 錯誤輸出前置詞時區

在 [Amazon 簡單儲存服務物件的自訂首碼中](#)，選擇您要用於日期和時間的時區。依預設，Firehose 會以 UTC 格式新增時間前置字元。如果您想要使用不同的時區，可以變更 S3 前置字元中使用的時區。

#### 緩衝提示

Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。

### S3 壓縮

選擇 GZIP、Snappy、Zip 或 Hadoopp 相容的 Snappy 資料壓縮，或不壓縮資料。以 Amazon Redshift 作為目的地的交付串流無法使用 Snappy、Zip 和 Hadoopp 相容的 Snappy 壓縮功能。

### S3 檔案副檔名格式 (選用)

為交付到 Amazon S3 目的地儲存貯體的物件指定副檔名格式。如果啟用此功能，指定的副檔名將會覆寫資料格式轉換或 S3 壓縮功能 (例如 .拍板或 .gz) 所附加的預設副檔名。將此功能與資料格式轉換或 S3 壓縮搭配使用時，請確定您是否設定了正確的副檔名。副檔名必須以句點 (.) 開頭，且可包含允許的字元：`0-9a-z!-_.*"()`。副檔名不能超過 128 個字元。

### S3 加密

Firehose 支援使用 AWS Key Management Service (SSE-KMS) 的 Amazon S3 伺服器端加密，以便在 Amazon S3 中加密交付的資料。您可以選擇使用目的地 S3 儲存貯體中指定的預設加密類型，或使用您擁有的金鑰清單中的 AWS KMS 金鑰加密。如果您使用 AWS KMS 金鑰加密資料，則可以使用預設 AWS 受管金鑰 (`aws/s3`) 或客戶受管金鑰。如需詳細資訊，請參閱 [使用 K AWS MS 管理金鑰 \(SSE-KMS\) 使用伺服器端加密來保護資料](#)。

## 針對目的地，選擇 Amazon Redshift

本節說明使用 Amazon Redshift 作為您的 Firehose 串流目的地的設定。

根據您是擁有 Amazon Redshift 佈建的叢集，還是 Amazon Redshift Serverless 工作群組，選擇下列其中一個程序。

- [Amazon Redshift 佈建叢集](#)

- [Amazon Redshift Serverless 工作群組](#)

## Amazon Redshift 佈建叢集

本節說明使用 Amazon Redshift 佈建的叢集做為您的 Firehose 串流目的地的設定。

- 為下列欄位輸入值：

### 叢集

要放入複製的 S3 儲存貯體的 Amazon Redshift 叢集。將 Amazon Redshift 叢集設定為可公開存取，並解除封鎖 Amazon 資料 Firehose IP 地址。如需詳細資訊，請參閱 [授予 Amazon 數據 Firehose 訪問亞 Amazon Redshift 目的地](#)。

### 使用者名稱

有 Amazon Redshift 叢集存取許可的 Amazon Redshift 使用者。該使用者必須具備 Amazon Redshift INSERT 許可，才能夠將 S3 儲存貯體中的資料複製到 Amazon Redshift 叢集。

### 密碼

有權存取叢集的使用者的密碼。

### 資料庫

要放入複製資料的 Amazon Redshift 資料庫。

### 資料表

要放入複製資料的 Amazon Redshift 資料表。

### 資料欄

(選用) 要放入複製資料的表格特定資料欄。如果您的 Amazon S3 物件中定義的資料欄數量少於 Amazon Redshift 資料表中的欄位數量，請使用此選項。

### 中繼 S3 目的地

Firehose 會先將您的資料交付到 S3 儲存貯體，然後發出 Amazon Redshift COPY 命令，將資料載入您的 Amazon Redshift 叢集。指定要交付串流資料您所擁有的 S3 儲存貯體。建立新的 S3 儲存貯體，或是選擇您目前擁有的儲存貯體。



Firehose 將資料載入您的 Amazon Redshift 叢集後，不會從 S3 儲存貯體中刪除該資料。您可以使用生命週期配置管理 S3 儲存貯體中的資料。如需詳細資訊，請參閱《Amazon Simple Storage Service 使用者指南》中的[物件生命週期管理](#)。

### 中繼 S3 字首

(選用) 要使用 Amazon S3 物件的預設字首，請將此選項保留空白。對於交付的 Amazon S3 物件，Firehose 會自動使用 YYYY/MM/dd/HH "" UTC 時間格式的前綴。您可以新增到此字首前方。如需詳細資訊，請參閱 [Amazon S3 物件名稱格式](#)。

### COPY options (COPY 選項)

此處指的是可以在 Amazon Redshift COPY 命令中指定的參數。您可以根據自己的組態來指定這些參數。例如，如果啟用 Amazon S3 資料壓縮，則需要 GZIP ""。如果您的 S3 儲存貯體與 Amazon Redshift 叢集不在相同的 AWS 區域，則需要「REGION」。如需詳細資訊，請參閱《Amazon Redshift 資料庫開發人員指南》中的 [COPY](#)。

### COPY command (COPY 命令)

Amazon Redshift COPY 命令。如需詳細資訊，請參閱《Amazon Redshift 資料庫開發人員指南》中的 [COPY](#)。

### Retry duration (重試持續時間)

如果傳送至 Amazon Redshift 叢集的資料 COPY 失敗，Firehose 重試的持續時間 (0—7200 秒)。Firehose 每 5 分鐘重試一次，直到重試持續時間結束為止。如果您將重試持續時間設置為 0 (零) 秒，Firehose 在 COPY 命令失敗時不會重試。

### 緩衝提示

Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。

### S3 壓縮

選擇 GZIP、Snappy、Zip 或 Hadoopp 相容的 Snappy 資料壓縮，或不壓縮資料。以 Amazon Redshift 作為目的地的交付串流無法使用 Snappy、Zip 和 Hadoopp 相容的 Snappy 壓縮功能。

### S3 檔案副檔名格式 (選用)

S3 檔案副檔名格式 (選用) — 為交付到 Amazon S3 目的地儲存貯體的物件指定檔案副檔名格式。如果啟用此功能，指定的副檔名將會覆寫資料格式轉換或 S3 壓縮功能 (例如 .拍板或 .gz) 所附加的預設副檔名。將此功能與資料格式轉換或 S3 壓縮搭配使用時，請確定您是否設定了



正確的副檔名。副檔名必須以句點 (.) 開頭，且可包含允許的字元：0-9a-z! -\_.\*' ()。副檔名不能超過 128 個字元。

## S3 加密

Firehose 支援使用 AWS Key Management Service (SSE-KMS) 的 Amazon S3 伺服器端加密，以便在 Amazon S3 中加密交付的資料。您可以選擇使用目的地 S3 儲存貯體中指定的預設加密類型，或使用您擁有的金鑰清單中的 AWS KMS 金鑰加密。如果您使用 AWS KMS 金鑰加密資料，則可以使用預設 AWS 受管金鑰 (aws/s3) 或客戶受管金鑰。如需詳細資訊，請參閱[使用 K AWS MS 管理金鑰 \(SSE-KMS\) 使用伺服器端加密來保護資料](#)。

## Amazon Redshift Serverless 工作群組

本節說明使用 Amazon Redshift 無伺服器工作群組做為您的 Firehose 串流目的地的設定。

- 為下列欄位輸入值：

Workgroup name (工作群組名稱)

要將 S3 儲存貯體資料複製到的 Amazon Redshift Serverless 工作群組。將 Amazon Redshift 無伺服器工作群組設定為可公開存取，並將 Firehose IP 地址解除封鎖。如需詳細資訊，請參閱[連線至 Amazon Redshift Serverless](#) 中「連線至可公開存取的 Amazon Redshift Serverless 執行個體」一節以及[授予 Amazon 數據 Firehose 訪問亞 Amazon Redshift 目的地](#)。

使用者名稱

有 Amazon Redshift Serverless 工作群組存取許可的 Amazon Redshift 使用者。該使用者必須具備 Amazon Redshift INSERT 許可，才可以將 S3 儲存貯體中的資料複製至 Amazon Redshift Serverless 工作群組。

密碼

具有存取 Amazon Redshift Serverless 工作群組許可之使用者的密碼。

資料庫

要放入複製資料的 Amazon Redshift 資料庫。

資料表

要放入複製資料的 Amazon Redshift 資料表。

## 資料欄

(選用) 要放入複製資料的表格特定資料欄。如果您的 Amazon S3 物件中定義的資料欄數量少於 Amazon Redshift 資料表中的欄位數量，請使用此選項。

## 中繼 S3 目的地

亞馬遜資料 Firehose 會先將您的資料交付到 S3 儲存貯體，然後發出 Amazon Redshift COPY 命令，將資料載入您的 Amazon Redshift 無伺服器工作群組。指定要交付串流資料您所擁有的 S3 儲存貯體。建立新的 S3 儲存貯體，或是選擇您目前擁有的儲存貯體。

將資料載入 Amazon Redshift 無伺服器工作群組後，Firehose 不會刪除 S3 儲存貯體中的資料。您可以使用生命週期配置管理 S3 儲存貯體中的資料。如需詳細資訊，請參閱《Amazon Simple Storage Service 使用者指南》中的[物件生命週期管理](#)。

## 中繼 S3 字首

(選用) 要使用 Amazon S3 物件的預設字首，請將此選項保留空白。對於交付的 Amazon S3 物件，Firehose 會自動使用 YYYY/MM/dd/HH "" UTC 時間格式的前綴。您可以新增到此字首前方。如需詳細資訊，請參閱[Amazon S3 物件名稱格式](#)。

## COPY options (COPY 選項)

此處指的是可以在 Amazon Redshift COPY 命令中指定的參數。您可以根據自己的組態來指定這些參數。例如，如果啟用 Amazon S3 資料壓縮，則需要 GZIP ""。如果您的 S3 儲存貯體與 Amazon Redshift 無伺服器工作群組位於相同的 AWS 區域，則需要「REGION」。如需詳細資訊，請參閱《Amazon Redshift 資料庫開發人員指南》中的[COPY](#)。

## COPY command (COPY 命令)

Amazon Redshift COPY 命令。如需詳細資訊，請參閱《Amazon Redshift 資料庫開發人員指南》中的[COPY](#)。

## Retry duration (重試持續時間)

如果傳送給 Amazon Redshift 無伺服器工作群組的資料 COPY 失敗，Firehose 重試的持續時間 (0—7200 秒)。Firehose 每 5 分鐘重試一次，直到重試持續時間結束為止。如果您將重試持續時間設置為 0 (零) 秒，Firehose 在 COPY 命令失敗時不會重試。

## 緩衝提示

Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。

## S3 壓縮

選擇 GZIP、Snappy、Zip 或 Hadoopp 相容的 Snappy 資料壓縮，或不壓縮資料。以 Amazon Redshift 作為目的地的交付串流無法使用 Snappy、Zip 和 Hadoopp 相容的 Snappy 壓縮功能。

## S3 檔案副檔名格式 (選用)

S3 檔案副檔名格式 (選用) — 為交付到 Amazon S3 目的地儲存貯體的物件指定檔案副檔名格式。如果啟用此功能，指定的副檔名將會覆寫資料格式轉換或 S3 壓縮功能 (例如 .拍板或 .gz) 所附加的預設副檔名。將此功能與資料格式轉換或 S3 壓縮搭配使用時，請確定您是否設定了正確的副檔名。副檔名必須以句點 (.) 開頭，且可包含允許的字元：0-9a-z! -\_.\*' ()。副檔名不能超過 128 個字元。

## S3 加密

Firehose 支援使用 AWS Key Management Service (SSE-KMS) 的 Amazon S3 伺服器端加密，以便在 Amazon S3 中加密交付的資料。您可以選擇使用目的地 S3 儲存貯體中指定的預設加密類型，或使用您擁有的金鑰清單中的 AWS KMS 金鑰加密。如果您使用 AWS KMS 金鑰加密資料，則可以使用預設 AWS 受管金鑰 (aws/s3) 或客戶受管金鑰。如需詳細資訊，請參閱[使用 K AWS MS 管理金鑰 \(SSE-KMS\) 使用伺服器端加密來保護資料](#)。

## 選擇適合您目的地的 OpenSearch 服務

本節說明針對目的地使用 OpenSearch 服務的選項。

- 為下列欄位輸入值：

OpenSearch 服務網域

資料傳送至的 OpenSearch 服務網域。

索引

將資料索引編入 OpenSearch 服務叢集時要使用的 OpenSearch 服務索引名稱。

Index rotation (索引輪換)

選擇是否應輪換 OpenSearch 服務索引以及輪換頻率。如果啟用了索引輪換，Amazon Data Firehose 會將對應的時間戳記附加到指定的索引名稱並進行輪換。如需詳細資訊，請參閱[OpenSearch 服務目的地的索引輪替](#)。

## 類型

將資料索引到 OpenSearch 服務叢集時要使用的 OpenSearch 服務類型名稱。對於彈性搜索 7.x 和 OpenSearch 1.x，每個索引只能有一種類型。如果您嘗試為已有其他類型的現有索引指定新類型，Firehose 會在執行階段傳回錯誤。

針對 Elasticsearch 7.x，將此欄位保留空白。

### Retry duration (重試持續時間)

如果對 OpenSearch 服務叢集的索引要求失敗，Firehose 重試的持續時間 (0—7200 秒)。Firehose 每 5 分鐘重試一次，直到重試持續時間結束為止。如果您將重試持續時間設定為 0 (零) 秒，Firehose 不會在索引要求失敗時重試。

### DocumentID 類型

表示設定文件 ID 的方法。支援的方法是 Firefox 產生的文件 ID 和 OpenSearch 服務產生的文件 ID。未設定文件 ID 值時，FireHost 產生的文件 ID 是預設選項。OpenSearch 建議使用服務產生的文件 ID，因為它支援大量寫入作業 (包括記錄分析和可觀察性)，在 OpenSearch Service 網域消耗較少的 CPU 資源，進而改善效能。

### 目的地 VPC 連線能力

如果您的 OpenSearch 服務網域位於私人 VPC 中，請使用此區段來指定該 VPC。同時指定您希望 Amazon Data Firehose 在將資料傳送到您 OpenSearch 的服務網域時使用的子網路和子群組。您可以使用 Ser OpenSearch vice 網域所使用的相同安全性群組。如果您指定不同的安全性群組，請確定這些群組允許輸出 HTTPS 流量傳送至 OpenSearch 服務網域的安全性群組。此外，請確定 OpenSearch 服務網域的安全性群組允許您設定 Firehose 串流時指定的安全性群組的 HTTPS 流量。如果您對 Firehose 串流和 OpenSearch 服務網域使用相同的安全性群組，請確定安全性群組的輸入規則允許 HTTPS 流量。如需安全群組規則的詳細資訊，請參閱 Amazon VPC 文件中 OpenSearch 的[安全群組規則](#)。

#### Important

當您指定子網路以將資料傳送至私有 VPC 中的目的地時，請確定所選子網路中有足夠數量的可用 IP 位址。如果指定的子網路中沒有可用的可用 IP 位址，Firehose 無法在私有 VPC 中為資料傳遞建立或新增 ENI，而且傳遞將會降級或失敗。

## 緩衝提示

Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。

## 為您的目的地選擇 OpenSearch 無伺服器

本節說明在目的地使用 OpenSearch 無伺服器的選項。

- 為下列欄位輸入值：

### OpenSearch 無伺服器收集

資料傳送至的一組 OpenSearch 無伺服器索引的端點。

### 索引

將資料索引至 OpenSearch 無伺服器集合時使用的 OpenSearch 無伺服器索引名稱。

### 目的地 VPC 連線能力

如果您的 OpenSearch 無伺服器集合位於私人 VPC 中，請使用此區段來指定該 VPC。同時指定您希望 Amazon Data Firehose 在將資料傳送到無伺服器集合時使用的子網路和子群組。

### OpenSearch

#### Important

當您指定子網路以將資料傳送至私有 VPC 中的目的地時，請確定所選子網路中有足夠數量的可用 IP 位址。如果指定的子網路中沒有可用的可用 IP 位址，Firehose 無法在私有 VPC 中為資料傳遞建立或新增 ENI，而且傳遞將會降級或失敗。

### Retry duration (重試持續時間)

如果 OpenSearch 無伺服器收集的索引請求失敗，Amazon 資料 Firehose 重試的持續時間 (0—7200 秒)。Amazon 資料 Firehose 每 5 分鐘重試一次，直到重試持續時間結束為止。如果您將重試持續時間設定為 0 (零) 秒，則 Amazon Data Firehose 在索引請求失敗時不會重試。

## 緩衝提示

Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。

## 為您的目的地選擇 HTTP 端點

本節會描述使用 HTTP 端點作為目的地的選項。

### Important

如果您選擇 HTTP 端點作為目的地，則請檢閱並遵循 [附錄 – HTTP 端點交付請求和回應規格](#) 中的指示。

- 為下列欄位提供值：

#### HTTP 端點名稱 - 選用

指定 HTTP 端點的使用者易記名稱。例如 My HTTP Endpoint Destination。

#### HTTP 端點 URL

請使用下列格式指定 HTTP 端點的 URL：`https://xyz.httpendpoint.com`。該 URL 必須為 HTTPS URL。

#### 存取金鑰 - 選用

請聯絡端點擁有者以取得存取金鑰 (如果需要)，以便從 Firehose 將資料傳送至其端點。

#### 內容編碼

Amazon Data Firehose 會使用內容編碼來壓縮請求本文，然後再將請求傳送到目的地。選擇 GZIP 或已停用以啟用/停用請求的內容編碼。

#### Retry duration (重試持續時間)

指定 Amazon 資料 Firehose 重試將資料傳送到所選 HTTP 端點的時間長度。

傳送資料之後，Amazon 資料 Firehose 會先等待來自 HTTP 端點的確認。如果發生錯誤或確認未在確認逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持

續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon Data Firehose 將資料傳送到 HTTP 端點 (無論是初始嘗試還是重試)，都會重新啟動確認逾時計數器，並等待來自 HTTP 端點的確認。

即使重試持續時間過期，Amazon Data Firehose 仍會等待確認，直到收到確認或已到達確認逾時期間為止。如果確認逾時，Amazon 資料 Firehose 會判斷重試計數器是否還有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到確認或判斷已無重試時間為止。

如果您不希望 Amazon 資料 Firehose 重試傳送資料，請將此值設定為 0。

### 參數 - 選用

Amazon 數據 Firehose 在每個 HTTP 調用中包含這些鍵值對。這些參數可協助您識別和整理目的地。

### 緩衝提示

Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。

#### Important

對於 HTTP 端點目的地，如果您在 CloudWatch 記錄檔中看到來自目標端點的 413 個回應碼，請降低 Firehose 串流上的緩衝提示大小，然後再試一次。

## 選擇 Datadog 作為目的地

本節說明使用 Datadog 作為目的地的選項。[如需有關 Datadog 的詳細資訊，請參閱 https://docs.datadoghq.com/integrations/amazon\\_web\\_services/](https://docs.datadoghq.com/integrations/amazon_web_services/)。

- 為下列欄位提供值：

### HTTP 端點 URL

從下拉式功能表的下列選項中選擇 HTTP 端點 URL：

- Datadog 日誌 - US1
- Datadog 日誌 - US5



- Datadog 日誌 - EU
- Datadog 日誌 - GOV
- Datadog 指標 - US
- Datadog 指標 - EU

## API 金鑰

請聯絡資料多，以取得啟用資料從 Amazon 資料 Firehose 傳送至此端點所需的 API 金鑰。

## 內容編碼

Amazon Data Firehose 會使用內容編碼來壓縮請求本文，然後再將請求傳送到目的地。選擇 GZIP 或已停用以啟用/停用請求的內容編碼。

## Retry duration (重試持續時間)

指定 Amazon 資料 Firehose 重試將資料傳送到所選 HTTP 端點的時間長度。

傳送資料之後，Amazon 資料 Firehose 會先等待來自 HTTP 端點的確認。如果發生錯誤或確認未在確認逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon Data Firehose 將資料傳送到 HTTP 端點 (無論是初始嘗試還是重試)，都會重新啟動確認逾時計數器，並等待來自 HTTP 端點的確認。

即使重試持續時間過期，Amazon Data Firehose 仍會等待確認，直到收到確認或已到達確認逾時期間為止。如果確認逾時，Amazon 資料 Firehose 會判斷重試計數器是否還有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到確認或判斷已無重試時間為止。

如果您不希望 Amazon 資料 Firehose 重試傳送資料，請將此值設定為 0。

## 參數 - 選用

Amazon 數據 Firehose 在每個 HTTP 調用中包含這些鍵值對。這些參數可協助您識別和整理目的地。

## 緩衝提示

Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。



## 為您的目的地選擇 Honeycomb

本節說明使用 Honeycomb 作為目的地的選項。如需有關蜂巢的詳細資訊，請參閱 <https://docs.honeycomb.io/getting-data-in/度量標準/aws-cloudwatch-metrics>。

- 為下列欄位提供值：

### Honeycomb Kinesis 端點

請使用下列格式指定 HTTP 端點的 URL：`https://api.honeycomb.io/1/kinesis_events/{{dataset}}`

### API 金鑰

請聯絡蜂巢以取得啟用資料從 Amazon 資料 Firehose 傳送至此端點所需的 API 金鑰。

### 內容編碼

Amazon Data Firehose 會使用內容編碼來壓縮請求本文，然後再將請求傳送到目的地。選擇 GZIP 以啟用要求的內容編碼。這是 Honeycomb 目的地的建議選項。

### Retry duration (重試持續時間)

指定 Amazon 資料 Firehose 重試將資料傳送到所選 HTTP 端點的時間長度。

傳送資料之後，Amazon 資料 Firehose 會先等待來自 HTTP 端點的確認。如果發生錯誤或確認未在確認逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon Data Firehose 將資料傳送到 HTTP 端點 (無論是初始嘗試還是重試)，都會重新啟動確認逾時計數器，並等待來自 HTTP 端點的確認。

即使重試持續時間過期，Amazon Data Firehose 仍會等待確認，直到收到確認或已到達確認逾時期間為止。如果確認逾時，Amazon 資料 Firehose 會判斷重試計數器是否還有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到確認或判斷已無重試時間為止。

如果您不希望 Amazon 資料 Firehose 重試傳送資料，請將此值設定為 0。

### 參數 - 選用

Amazon 數據 Firehose 在每個 HTTP 調用中包含這些鍵值對。這些參數可協助您識別和整理目的地。

## 緩衝提示

Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。

## 選擇 Coralogix 作為您的目的地

本節會描述使用 Coralogix 作為目的地的選項。如需 Coralogix 的詳細資訊，請參閱 <https://coralogix.com/integrations/aws-firehose>。

- 為下列欄位提供值：

### HTTP 端點 URL

從下拉式功能表的下列選項中選擇 HTTP 端點 URL：

- Coralogix - US
- Coralogix - SINGAPORE
- Coralogix - IRELAND
- Coralogix - INDIA
- Coralogix - STOCKHOLM

### 私有金鑰

請聯絡 Coralogix 以取得啟用資料從 Amazon 資料 Firehose 傳送到此端點所需的私密金鑰。

### 內容編碼

Amazon Data Firehose 會使用內容編碼來壓縮請求本文，然後再將請求傳送到目的地。選擇 GZIP 以啟用要求的內容編碼。這是 Coralogix 目的地的建議選項。

### Retry duration (重試持續時間)

指定 Amazon 資料 Firehose 重試將資料傳送到所選 HTTP 端點的時間長度。

傳送資料之後，Amazon 資料 Firehose 會先等待來自 HTTP 端點的確認。如果發生錯誤或確認未在確認逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon Data Firehose 將資料傳送到 HTTP 端點 (無論是初始嘗試還是重試)，都會重新啟動確認逾時計數器，並等待來自 HTTP 端點的確認。

即使重試持續時間過期，Amazon Data Firehose 仍會等待確認，直到收到確認或已到達確認逾時期間為止。如果確認逾時，Amazon 資料 Firehose 會判斷重試計數器是否還有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到確認或判斷已無重試時間為止。

如果您不希望 Amazon 資料 Firehose 重試傳送資料，請將此值設定為 0。

### 參數 - 選用

Amazon 數據 Firehose 在每個 HTTP 調用中包含這些鍵值對。這些參數可協助您識別和整理目的地。

- `applicationName`：執行 Data Firehose 的環境
- `subsystemName`：資料 Firehose 整合的名稱
- `電腦名稱`：使用中的「Firehose」串流的名稱

### 緩衝提示

Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。

## 選擇 Dynatrace 作為目的地

本節會描述使用 Dynatrace 作為目的地的選項。如需詳細資訊，請參閱 <https://www.dynatrace.com/support/help/technology-support/cloud-platforms/amazon-web-services/整合/cloudwatch-metric-streams>。

- 為下列欄位提供值：

### HTTP 端點 URL

從下拉式功能表中選擇 HTTP 端點 URL (Dynatrace US、Dynatrace EU 或 Dynatrace Global)。

## API 記號

從 Amazon 數據 Firehose 生成數據交付所需的 Dynatrace API 令牌。如需詳細資訊，請參閱 <https://www.dynatrace.com/support/help/dynatrace-api/basics/dynatrace-api-authentication/>。

## API URL

提供您的 Dynatrace 環境的 API URL。

## 內容編碼

Amazon Data Firehose 會使用內容編碼來壓縮請求本文，然後再將請求傳送到目的地。選擇 GZIP 或已停用以啟用/停用請求的內容編碼。

## Retry duration (重試持續時間)

指定 Amazon 資料 Firehose 重試將資料傳送到所選 HTTP 端點的時間長度。

傳送資料之後，Amazon 資料 Firehose 會先等待來自 HTTP 端點的確認。如果發生錯誤或確認未在確認逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon Data Firehose 將資料傳送到 HTTP 端點 (無論是初始嘗試還是重試)，都會重新啟動確認逾時計數器，並等待來自 HTTP 端點的確認。

即使重試持續時間過期，Amazon Data Firehose 仍會等待確認，直到收到確認或已到達確認逾時期間為止。如果確認逾時，Amazon 資料 Firehose 會判斷重試計數器是否還有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到確認或判斷已無重試時間為止。

如果您不希望 Amazon 資料 Firehose 重試傳送資料，請將此值設定為 0。

## 參數 - 選用

Amazon 數據 Firehose 在每個 HTTP 調用中包含這些鍵值對。這些參數可協助您識別和整理目的地。

## 緩衝提示

Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。

## 選擇您 LogicMonitor 的目的地

本節說明用於目的地 LogicMonitor 的選項。如需詳細資訊，請參閱 <https://www.logicmonitor.com>。

- 為下列欄位提供值：

### HTTP 端點 URL

請使用下列格式指定 HTTP 端點的 URL：`https://ACCOUNT.logicmonitor.com`

### API 金鑰

請聯絡 LogicMonitor 以取得啟用從 Amazon 資料 Firehose 向此端點傳送資料所需的 API 金鑰。

### 內容編碼

Amazon Data Firehose 會使用內容編碼來壓縮請求本文，然後再將請求傳送到目的地。選擇 GZIP 或已停用以啟用/停用請求的內容編碼。

### Retry duration (重試持續時間)

指定 Amazon 資料 Firehose 重試將資料傳送到所選 HTTP 端點的時間長度。

傳送資料之後，Amazon 資料 Firehose 會先等待來自 HTTP 端點的確認。如果發生錯誤或確認未在確認逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon Data Firehose 將資料傳送到 HTTP 端點 (無論是初始嘗試還是重試)，都會重新啟動確認逾時計數器，並等待來自 HTTP 端點的確認。

即使重試持續時間過期，Amazon Data Firehose 仍會等待確認，直到收到確認或已到達確認逾時期間為止。如果確認逾時，Amazon 資料 Firehose 會判斷重試計數器是否還有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到確認或判斷已無重試時間為止。

如果您不希望 Amazon 資料 Firehose 重試傳送資料，請將此值設定為 0。

### 參數 - 選用

Amazon 數據 Firehose 在每個 HTTP 調用中包含這些鍵值對。這些參數可協助您識別和整理目的地。

## 緩衝提示

Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。

## 選擇 Logz.io 作為目的地

本節說明使用 Logz.io 作為目的地的選項。如需詳細資訊，請參閱 <https://logz.io/>。

### Note

在歐洲 (米蘭) 地區，Logz.io 不支援做為 Amazon 資料 Firehose 目的地。

- 為下列欄位提供值：

#### HTTP 端點 URL

請使用下列格式指定 HTTP 端點的 URL：`https://listener-aws-metrics-stream-<region>.logz.io/`。例如 `https://listener-aws-metrics-stream-us.logz.io/`。該 URL 必須為 HTTPS URL。

#### 內容編碼

Amazon Data Firehose 會使用內容編碼來壓縮請求本文，然後再將請求傳送到目的地。選擇 GZIP 或已停用以啟用/停用請求的內容編碼。

#### Retry duration (重試持續時間)

指定 Amazon 數據 Firehose 重試將數據發送到 Logz.io 的時間長度。

傳送資料之後，Amazon 資料 Firehose 會先等待來自 HTTP 端點的確認。如果發生錯誤或確認未在確認逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon Data Firehose 將資料傳送到 HTTP 端點 (無論是初始嘗試還是重試)，都會重新啟動確認逾時計數器，並等待來自 HTTP 端點的確認。

即使重試持續時間過期，Amazon Data Firehose 仍會等待確認，直到收到確認或已到達確認逾時期間為止。如果確認逾時，Amazon 資料 Firehose 會判斷重試計數器是否還有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到確認或判斷已無重試時間為止。

如果您不希望 Amazon 資料 Firehose 重試傳送資料，請將此值設定為 0。

#### 參數 - 選用

Amazon 數據 Firehose 在每個 HTTP 調用中包含這些鍵值對。這些參數可協助您識別和整理目的地。

#### 緩衝提示

Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。

## 選擇 MongoDB Cloud 作為目的地

本節會描述使用 MongoDB Cloud 作為目的地的選項。如需詳細資訊，請參閱 <https://www.mongodb.com>。

- 為下列欄位提供值：

#### MongoDB 領域 webhook URL

請使用下列格式指定 HTTP 端點的 URL：`https://webhooks.mongodb-realm.com`。該 URL 必須為 HTTPS URL。

#### API 金鑰

請聯絡 MongoDB 雲端，以取得啟用資料從 Amazon 資料快速傳遞到此端點所需的 API 金鑰。

#### 內容編碼

Amazon Data Firehose 會使用內容編碼來壓縮請求本文，然後再將請求傳送到目的地。選擇 GZIP 或已停用以啟用/停用請求的內容編碼。

#### Retry duration (重試持續時間)

指定 Amazon 資料 Firehose 重試將資料傳送至所選第三方供應商的時間長度。

傳送資料之後，Amazon 資料 Firehose 會先等待來自 HTTP 端點的確認。如果發生錯誤或確認未在確認逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon Data Firehose 將資料傳送到 HTTP 端點 (無論是初始嘗試還是重試)，都會重新啟動確認逾時計數器，並等待來自 HTTP 端點的確認。

即使重試持續時間過期，Amazon Data Firehose 仍會等待確認，直到收到確認或已到達確認逾時期間為止。如果確認逾時，Amazon 資料 Firehose 會判斷重試計數器是否還有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到確認或判斷已無重試時間為止。

如果您不希望 Amazon 資料 Firehose 重試傳送資料，請將此值設定為 0。

### 緩衝提示

Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。

### 參數 - 選用

Amazon 數據 Firehose 在每個 HTTP 調用中包含這些鍵值對。這些參數可協助您識別和整理目的地。

## 選擇 New Relic 作為目的地

本節會描述使用 New Relic 作為目的地的選項。如需詳細資訊，請參閱 <https://newrelic.com>。

- 為下列欄位提供值：

### HTTP 端點 URL

從下拉式功能表的下列選項中選擇 HTTP 端點 URL：

- New Relic 日誌 - US
- New Relic 指標 - US
- New Relic 指標 - EU



## API 金鑰

輸入 New Relic One 帳戶設定中的授權金鑰 (40 個字元的十六進位字串)。必須使用此 API 金鑰才能從 Firehose 將資料傳送至此端點。

## 內容編碼

Amazon Data Firehose 會使用內容編碼來壓縮請求本文，然後再將請求傳送到目的地。選擇 GZIP 或已停用以啟用/停用請求的內容編碼。

## Retry duration (重試持續時間)

指定 Amazon 資料 Firehose 重試將資料傳送到新遺物 HTTP 端點的時間長度。

傳送資料之後，Amazon 資料 Firehose 會先等待來自 HTTP 端點的確認。如果發生錯誤或確認未在確認逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon Data Firehose 將資料傳送到 HTTP 端點 (無論是初始嘗試還是重試)，都會重新啟動確認逾時計數器，並等待來自 HTTP 端點的確認。

即使重試持續時間過期，Amazon Data Firehose 仍會等待確認，直到收到確認或已到達確認逾時期間為止。如果確認逾時，Amazon 資料 Firehose 會判斷重試計數器是否還有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到確認或判斷已無重試時間為止。

如果您不希望 Amazon 資料 Firehose 重試傳送資料，請將此值設定為 0。

## 參數 - 選用

Amazon 數據 Firehose 在每個 HTTP 調用中包含這些鍵值對。這些參數可協助您識別和整理目的地。

## 緩衝提示

Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。

## 選擇雪花作為您的目的地

本節說明在目的地使用雪花的選項。

## 連線設定

- 為下列欄位提供值：

### 雪花帳戶網址

指定雪花提供的地區帳戶 URL。當啟用私人連結時，Firehose 不支援預覽中的無區域 URL。請參閱[雪花文件](#)，瞭解如何判斷您的帳戶 URL。例如：xy12345.us-east-1.aws.snowflakecomputing.com。請注意，不能指定端口號，而協議 ( https : // ) 是可選的。

### 使用者登入

指定要用於載入資料的雪花使用者。請確定使用者有權將資料插入 Snowflake 資料表。

### 私有金鑰

指定用於透過 Snowflake 驗證之 key pair 的使用者私密金鑰。請確定私密金鑰的格式為 PKCS8。請勿將 PEM 頁首和頁尾包含為私密金鑰的一部分。如果鍵被分成多行，請刪除換行符。

### Passphrase (密碼短語)

加密金鑰時，用來解密私密金鑰的密碼。如果私密金鑰未加密，請將此欄位保留空白。如需詳細資訊，請參閱[使用金鑰配對驗證與金鑰輪換](#)。

### 角色配置

使用預設的雪花角色 — 如果選取此選項，Firehose 將不會將任何角色傳遞給「雪花」。預設角色會假設載入資料。請確定預設角色具有將資料插入 Snowflake 資料表的權限。

使用自訂雪花角色 — 輸入 Firehose 將資料載入雪花表格時所要使用的非預設雪花角色。

### 雪花連接

選項為「私人」或「公用」。

### 私人 VPCE 識別碼 (選擇性)

Firehose 的 VPCE ID 可以與雪花私密連接。識別碼格式是相同的 `[##] .vp--[ID]`。如需詳細資訊，請參閱[雪AWS PrivateLink 花 \(S\)](#)。

**Note**

在公開預覽期間，請確定您的雪花網路允許存取 Firehose。若要取得存取權，請聯絡 AWS Support 將 Firehose VPC 端點新增至允許清單，或考慮停用 Snowflake 叢集上的網路原則。

## 資料庫組態

- 您必須指定下列設定，才能使用雪花作為 Firehose 傳送串流的目的地：
  - 雪花數據庫-雪花中的所有數據保持在數據庫中。
  - 雪花結構描述 — 每個資料庫都包含一或多個結構描述，這些結構描述是資料庫物件的邏輯群組，例如表格和視觀表
  - 雪花表-雪花中的所有數據都存儲在數據庫表中，邏輯結構為列和行的集合。

## 雪花表格的資料載入選項

- 使用 JSON 鍵作為列名
- 使用變體欄
  - 內容欄名稱 — 在表格中指定必須載入原始資料的欄名稱。
  - 中繼資料欄名稱 (選擇性) — 在表格中指定欄名稱，其中必須載入中繼資料資訊。

## 選擇 Splunk 做為目的地

本節說明使用 Splunk 做為目的地的選項。

**Note**

Firehose 會將資料傳送至使用 Classic Load Balancer 或 Application Load Balancer 載平衡器設定的 Splunk 叢集。

- 為下列欄位提供值：

## Splunk cluster endpoint (Splunk 叢集端點)

若要確定端點，請參閱 Splunk 文件中的[設定 Amazon 資料 Firehose 以將資料傳送到 Splunk 平台](#)。

## Splunk endpoint type (Splunk 端點類型)

在大多數情況下，請選擇 Raw endpoint。選擇您是 Event endpoint 否使用 AWS Lambda 以依事件類型將資料傳送至不同索引來預先處理資料。如需要使用哪個端點的相關資訊，請參閱 Splunk 說明文件中的[設定 Amazon 資料 Firehose 以將資料傳送到 Splunk 平台](#)。

## Authentication token (身分驗證字符)

若要設定可從 Amazon 資料 Firehose 接收資料的 Splunk 端點，請參閱 Splunk 說明文件中的 Amazon 資料 Firehose [Splunk 附加元件的安裝和組態概觀](#)。當您為此 Firehose 串流設定端點時，請儲存從 Splunk 取得的權杖，並將其新增至此處。

## HEC acknowledgement timeout (HEC 確認訊息逾時)

指定 Amazon 數據 Firehose 等待 Splunk 索引確認的時間長度。如果 Splunk 未在逾時到達前傳送通知，Amazon 資料 Firehose 會將其視為資料傳遞失敗。然後，Amazon 資料 Firehose 會根據您設定的重試持續時間值，將資料重試或備份到 Amazon S3 儲存貯體。

## Retry duration (重試持續時間)

指定 Amazon 資料 Firehose 重試將資料傳送至 Splunk 的時間長度。

發送數據後，Amazon 數據 Firehose 首先等待 Splunk 的確認。如果發生錯誤或確認未在確認逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon 資料 Firehose 將資料傳送至 Splunk (無論是初始嘗試還是重試)，都會重新啟動確認逾時計數器，並等待 Splunk 的確認。

即使重試持續時間過期，Amazon Data Firehose 仍會等待確認，直到收到確認或已到達確認逾時期間為止。如果確認逾時，Amazon 資料 Firehose 會判斷重試計數器是否還有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到確認或判斷已無重試時間為止。

如果您不希望 Amazon 資料 Firehose 重試傳送資料，請將此值設定為 0。

## 選擇 Splunk Observability Cloud 作為您的目的地

本節會描述使用 Splunk Observability Cloud 作為目的地的選項。如需詳細資訊，[請參閱 `get-data-in / connect/aws/aws-apiconfig.html # connect-to-aws-using the-splunk-observability-cloud-API`](https://docs.splunk.com/observability/en/gdi/connect/aws/aws-apiconfig.html#connect-to-aws-using-the-splunk-observability-cloud-API)。 <https://docs.splunk.com/observability/en/gdi/>

- 為下列欄位提供值：

### 雲端擷取端點 URL

您可以在 Splunk 可觀測性主控台的設定檔 > 組織 > 即時資料擷取端點中找到 Splunk 可觀測性雲端的即時資料擷取 URL。

### 存取記號

從 Splunk 可觀測性主控台內的「設定」>「存取記號」中複製具有 INGEST 授權範圍的 Splunk 可觀測性存取記號

### 內容編碼

Amazon Data Firehose 會使用內容編碼來壓縮請求本文，然後再將請求傳送到目的地。選擇 GZIP 或已停用以啟用/停用請求的內容編碼。

### Retry duration (重試持續時間)

指定 Amazon 資料 Firehose 重試將資料傳送到所選 HTTP 端點的時間長度。

傳送資料之後，Amazon 資料 Firehose 會先等待來自 HTTP 端點的確認。如果發生錯誤或確認未在確認逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon Data Firehose 將資料傳送到 HTTP 端點 (無論是初始嘗試還是重試)，都會重新啟動確認逾時計數器，並等待來自 HTTP 端點的確認。

即使重試持續時間過期，Amazon Data Firehose 仍會等待確認，直到收到確認或已到達確認逾時期間為止。如果確認逾時，Amazon 資料 Firehose 會判斷重試計數器是否還有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到確認或判斷已無重試時間為止。

如果您不希望 Amazon 資料 Firehose 重試傳送資料，請將此值設定為 0。

## 參數 - 選用

Amazon 數據 Firehose 在每個 HTTP 調用中包含這些鍵值對。這些參數可協助您識別和整理目的地。

### 緩衝提示

Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的目的地緩衝區大小因服務提供者而異。

## 選擇 Sumo Logic 作為目的地

本節會描述使用 Sumo Logic 作為目的地的選項。如需詳細資訊，請參閱 <https://www.sumologic.com>。

- 為下列欄位提供值：

### HTTP 端點 URL

請使用下列格式指定 HTTP 端點的 URL：`https://deployment name.sumologic.net/receiver/v1/kinesis/dataType/access token`。該 URL 必須為 HTTPS URL。

### 內容編碼

Amazon Data Firehose 會使用內容編碼來壓縮請求本文，然後再將請求傳送到目的地。選擇 GZIP 或已停用以啟用/停用請求的內容編碼。

### Retry duration (重試持續時間)

指定 Amazon 數據 Firehose 重試將數據發送到 Sumo 邏輯的時間。

傳送資料之後，Amazon 資料 Firehose 會先等待來自 HTTP 端點的確認。如果發生錯誤或確認未在確認逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon Data Firehose 將資料傳送到 HTTP 端點 (無論是初始嘗試還是重試)，都會重新啟動確認逾時計數器，並等待來自 HTTP 端點的確認。

即使重試持續時間過期，Amazon Data Firehose 仍會等待確認，直到收到確認或已到達確認逾時期間為止。如果確認逾時，Amazon 資料 Firehose 會判斷重試計數器是否還有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到確認或判斷已無重試時間為止。

如果您不希望 Amazon 資料 Firehose 重試傳送資料，請將此值設定為 0。

### 參數 - 選用

Amazon 數據 Firehose 在每個 HTTP 調用中包含這些鍵值對。這些參數可協助您識別和整理目的地。

### 緩衝提示

Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。建議的 Elastic 目的地緩衝區大小因服務提供者而異。

## 選擇 Elastic 作為目的地

本節會描述使用 Elastic 作為目的地的選項。

- 為下列欄位提供值：

### Elastic 端點 URL

請使用下列格式指定 HTTP 端點的 URL：`https://<cluster-id>.es.<region>.aws.elastic-cloud.com`。該 URL 必須為 HTTPS URL。

### API 金鑰

請聯絡彈性服務以取得啟用資料從 Amazon 資料快速傳遞至其服務所需的 API 金鑰。

### 內容編碼

Amazon Data Firehose 會使用內容編碼來壓縮請求本文，然後再將請求傳送到目的地。選擇 GZIP (預設選擇) 或已停用以啟用/停用請求的內容編碼。

### Retry duration (重試持續時間)

指定 Amazon 數據 Firehose 重試將數據發送到彈性的時間長度。

傳送資料之後，Amazon 資料 Firehose 會先等待來自 HTTP 端點的確認。如果發生錯誤或確認未在確認逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持



續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon Data Firehose 將資料傳送到 HTTP 端點 (無論是初始嘗試還是重試)，都會重新啟動確認逾時計數器，並等待來自 HTTP 端點的確認。

即使重試持續時間過期，Amazon Data Firehose 仍會等待確認，直到收到確認或已到達確認逾時期間為止。如果確認逾時，Amazon 資料 Firehose 會判斷重試計數器是否還有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到確認或判斷已無重試時間為止。

如果您不希望 Amazon 資料 Firehose 重試傳送資料，請將此值設定為 0。

### 參數 - 選用

Amazon 數據 Firehose 在每個 HTTP 調用中包含這些鍵值對。這些參數可協助您識別和整理目的地。

### 緩衝提示

Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其傳送到指定的目的地。Elastic 目的地的建議緩衝區大小為 1 MiB。

## 備份和進階設定

本主題說明如何設定 Firehose 串流的備份和進階設定。

### 備份設定

亞馬遜資料 Firehose 使用 Amazon S3 備份嘗試交付到您選擇的目的地的所有或失敗的資料。

#### Important

只有當 Firehose 串流的來源為直接 PUT 或 Kinesis Data Streams 時，才支援 Backup 設定。

如果您做了下列其中一個選擇，則可以為 Firehose 串流指定 S3 備份設定：

- 如果您將 Amazon S3 設定為 Amazon 資料 Firehose Firehose 串流的目的地，並選擇指定 AWS Lambda 函數來轉換資料記錄，或者選擇轉換交付串流的資料記錄格式。
- 如果您將 Amazon Redshift 設定為 Amazon 資料 Firehose 串流的目的地，並選擇指定 AWS Lambda 函數來轉換資料記錄。



- 如果您將以下任何服務設置為 Firehose 流的目的地：Amazon OpenSearch 服務，數據多，重啟，HTTP 端點，MongoDB 雲，新遺物 LogicMonitor，濺射或相撲邏輯。

以下是您的 Amazon 資料 Firehose 交付串流的備份設定：

- Amazon S3 中的來源記錄備份 - 如果 S3 或 Amazon Redshift 是您選取的目的地，此設定會指出您是要啟用來源資料備份，還是保持停用狀態。如果將任何其他支援的服務 (S3 或 Amazon Redshift 除外) 設為您選取的目的地，則此設定會指示您是要備份所有來源資料，還是僅備份失敗的資料。
- S3 備份儲存貯體-這是 Amazon 資料 Firehose 備份您資料的 S3 儲存貯體。
- S3 備份儲存貯體前綴-這是 Amazon 資料 Firehose 備份您資料的前置詞。
- S3 備份儲存貯體錯誤輸出字首 - 所有失敗的資料都會備份在此 S3 儲存貯體錯誤輸出字首中。
- 緩衝提示、備份的壓縮和加密-Amazon Data Firehose 使用 Amazon S3 備份嘗試傳送到您選擇目的地的所有或失敗的資料。Amazon 資料 Firehose 會先緩衝傳入的資料，然後再將其交付 (備份) 到 Amazon S3。您可以選擇 1-128 的緩衝區大小 MiBs 和 60—900 秒的緩衝區間隔。只要滿足其中一項條件，即會觸發資料交付至 Amazon S3 的動作。如果啟用資料轉換，緩衝區間隔會從 Amazon Data Firehose 接收轉換的資料到資料交付至 Amazon S3 的時間套用。如果資料傳遞到目的地落後於寫入 Firehose 串流的資料，Amazon 資料 Firehose 會動態提高緩衝區大小以 catch。此動作可確保所有資料皆能成功傳送至目的地。
- S3 壓縮-選擇 GZIP、快速、Zip 或 Hadoopp 相容的快速資料壓縮，或不壓縮資料。以 Amazon Redshift 作為目的地的交付串流無法使用 Snappy、Zip 和 Hadoopp 相容的 Snappy 壓縮功能。
- S3 檔案副檔名格式 (選用) — 為交付到 Amazon S3 目的地儲存貯體的物件指定檔案副檔名格式。如果啟用此功能，指定的副檔名將會覆寫資料格式轉換或 S3 壓縮功能 (例如 .拍板或 .gz) 所附加的預設副檔名。將此功能與資料格式轉換或 S3 壓縮搭配使用時，請確定您是否設定了正確的副檔名。副檔名必須以句點 (.) 開頭，且可包含允許的字元：0-9a-z! \_.\*' ()。副檔名不能超過 128 個字元。
- Firehose 支援使用 AWS Key Management Service (SSE-KMS) 的 Amazon S3 伺服器端加密，以便在 Amazon S3 中加密交付的資料。您可以選擇使用目的地 S3 儲存貯體中指定的預設加密類型，或使用您擁有的金鑰清單中的 AWS KMS 金鑰加密。如果您使用 AWS KMS 金鑰加密資料，則可以使用預設 AWS 受管金鑰 (aws/s3) 或客戶受管金鑰。如需詳細資訊，請參閱[使用 K AWS MS 管理金鑰 \(SSE-KMS\) 使用伺服器端加密來保護資料](#)。

## 進階設定

以下是您的 Amazon 資料 Firehose 交付串流的進階設定：

- 伺服器端加密-A Amazon 資料 Firehose 透過金 AWS 鑰管理服務 (AWS KMS) 支援 Amazon S3 伺服器端加密，以加密 Amazon S3 中的交付資料。如需詳細資訊，請參閱[使用 KMS AWS 受管金鑰 \(SSE-KMS\) 使用伺服器端加密來保護資料](#)。
- 錯誤記錄-A Amazon 資料 Firehose 會記錄與處理和交付相關的錯誤。此外，啟用資料轉換後，它可以記錄 Lambda 叫用，並將資料傳送錯誤傳送至 CloudWatch 記錄。如需詳細資訊，請參閱[使用 CloudWatch 記錄監控 Amazon 資料 Firehose](#)。

### Important

雖然可選，但強烈建議您在 Firehose 串流建立期間啟用 Amazon 資料火管錯誤記錄。此作法可確保在記錄處理或交付失敗時，您可以存取錯誤詳細資料。

- 許可-A Amazon 資料 Firehose 會針對 Firehose 串流所需的所有許可使用 IAM 角色。您可以選擇建立自動指派所需許可的新角色，或選擇為 Amazon Data Firehose 建立的現有角色。此角色可用來授與 Firehose 存取各種服務，包括 S3 儲存貯體、AWS KMS 金鑰 (如果啟用資料加密) 和 Lambda 函數 (如果啟用了資料轉換)。主控台可能建立一個含預留位置的角色。如需詳細資訊，請參閱[什麼是 IAM ?](#)。
- 標籤-您可以新增標籤來組織資 AWS 源、追蹤成本並控制存取權。

如果您在 `CreateDeliveryStream` 動作中指定標籤，Amazon Data Firehose 會對該 `firehose:TagDeliveryStream` 動作執行其他授權，以驗證使用者是否具有建立標籤的權限。如果您未提供此權限，則使用 IAM 資源標籤建立新 Firehose 交付串流的請求將會失敗，並顯示 `AccessDeniedException` 如下所示。

```
AccessDeniedException
```

```
User: arn:aws:sts::x:assumed-role/x/x is not authorized to perform:
  firehose:TagDeliveryStream on resource: arn:aws:firehose:us-east-1:x:deliverystream/
x with an explicit deny in an identity-based policy.
```

下列範例示範允許使用者建立交付串流和套用標籤的原則。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "firehose:CreateDeliveryStream",
      "Resource": "*",
    }
  ]
}
```

```
    },
    {
      "Effect": "Allow",
      "Action": "firehose:TagDeliveryStream",
      "Resource": "*",
    }
  ]
}
```

選擇備份和進階設定後，請檢閱您的選擇，然後選擇 [建立 Firehose 串流]。

新的 Firehose 串流在「建立」狀態下需要幾分鐘的時間，才能使用。Firehose 串流處於作用中狀態後，您就可以開始從您的製作者傳送資料給該串流。

## 緩衝提示

Amazon Data Firehose 會將記憶體中的傳入串流資料緩衝至特定大小 (緩衝大小) 和一段特定時間 (緩衝間隔)，然後再將其傳送到指定的目的地。當您想要將最佳大小的檔案交付到 Amazon S3，並從資料處理應用程式獲得更好的效能，或調整 Firehose 交付率以符合目的地速度時，可以使用緩衝提示。

您可以在建立新交付串流時設定緩衝大小和緩衝區間隔，或更新現有交付串流的緩衝大小和緩衝間隔。緩衝大小以 MB 為單位測量，緩衝間隔以秒為單位測量。不過，如果指定它們其中一個值，您也必須提供其他的值。滿足的第一個緩衝區條件會觸發 Firehose 傳送資料。如果您未設定緩衝值，則會使用預設值。

您可以透過 AWS Management Console、AWS Command Line Interface 或 AWS SDK 設定 Firehose 緩衝提示。對於現有串流，您可以使用主控台中的 [編輯] 選項或使用 API，以適合您使用案例的值重新設定緩衝提示。[UpdateDestination](#) 對於新串流，您可以使用主控台或使用 [CreateDeliveryStream](#) API 來設定緩衝提示，做為建立新串流的一部分。若要調整緩衝大小，請 `IntervalInSeconds` 在 [CreateDeliveryStream](#) 或 [UpdateDestination](#) API 的目標特定 `DestinationConfiguration` 參數中設定 `SizeInMBs` 和。

### Note

為了滿足即時使用案例的較低延遲，您可以使用零緩衝間隔提示。當您將緩衝間隔設定為零秒時，Firehose 不會緩衝資料，而且會在幾秒鐘內傳送資料。在您將緩衝提示變更為較低的值之前，請洽詢廠商以取得 Firehose 目的地的建議緩衝提示。

**Note**

當您設定少於 60 秒的緩衝時間間隔以提供較低的延遲時間時，Firehose 會針對 S3 目的地使用多部分上傳。由於 S3 目的地的多部分上傳，如果您選擇的緩衝時間間隔少於 60 秒，則 S3 PUT API 成本會有所增加。

如需目的地特定緩衝提示範圍和預設值，請參閱下表：

目的地	緩衝大小 (以 MB 為單位) (默認在括號中)	緩衝間隔 (以秒為單位) (預設值在括號中)
S3	一百二十八 (5)	0-900
Redshift	一百二十八 (5)	0-900
OpenSearch 無伺服器	一百	0-900
OpenSearch	一百	0-900
Splunk	一至五	0-60 (60)
Datadog	1-4	九百
Coralogix	1-64 (6)	九百
Dynatrace	一六四 (5)	九百
彈性	1	九百
Honeycomb	1-64 (15)	九百
端點	一六四 (5)	九百
LogicMonitor	一六四 (5)	九百
洛吉奧	一六四 (5)	九百

目的地	緩衝大小 (以 MB 為單位) (默認在括號中)	緩衝間隔 (以秒為單位) (預設值在括號中)
蒙古数据库	1-16 (5)	九百
新遺物	一六四 (5)	九百
Sumologic	—	九百
Splunk Observability Cloud	—	九百

# 使用範例資料測試您的 Firehose 串流

您可以使用擷取 AWS Management Console 取模擬股票股票代碼資料。主控台會在您的瀏覽器中執行指令碼，將範例記錄放入 Firehose 串流中。這可讓您測試 Firehose 串流的組態，而不必產生自己的測試資料。

以下為取自模擬資料的範例：

```
{"TICKER_SYMBOL": "QXZ", "SECTOR": "HEALTHCARE", "CHANGE": -0.05, "PRICE": 84.51}
```

請注意，當您的 Firehose 串流傳輸資料時，會收取標準 Amazon 資料 Firehose 費用，但產生資料時不會收取任何費用。若要停止產生這些費用，您可以隨時從主控台停止範例串流。

## 目錄

- [必要條件](#)
- [使用 Amazon S3 作為目的地進行測試](#)
- [使用 Amazon Redshift 作為目的地進行測試](#)
- [使用 OpenSearch 服務作為目的地進行測試](#)
- [使用 Splunk 做為目的地進行測試](#)

## 必要條件

在開始之前，請先建立「Firehose」串流。如需詳細資訊，請參閱 [建立 Firehose 串流](#)。

## 使用 Amazon S3 作為目的地進行測試

使用下列程序，使用亞馬遜簡單儲存服務 (Amazon S3) 做為目的地來測試您的 Firehose 串流。

若要使用 Amazon S3 測試 Firehose 串流

1. 打開 Firehose 主控台，網址為 <https://console.aws.amazon.com/firehose/>。
2. 選擇「Firehose」串流。
3. 在 Test with demo data (以示範資料測試) 底下，選擇 Start sending demo data (開始傳送示範資料) 以產生股票行情資料範例。
4. 按照螢幕上的指示，確認資料正交付到 S3 儲存貯體。請注意，根據您在儲存貯體的緩衝設定，新物件可能需要幾分鐘才會出現於您的儲存貯體中。

5. 測試完成後，選擇 Stop sending demo data (停止傳送示範資料) 以停止產生使用費。

## 使用 Amazon Redshift 作為目的地進行測試

請使用下列程序，使用 Amazon Redshift 作為目的地來測試您的 Firehose 串流。

使用 Amazon Redshift 測試 Firehose 流

1. 您的 Firehose 串流預期資料表會出現在您的 Amazon Redshift 叢集中。[透過 SQL 介面連線至 Amazon Redshift](#)，並執行下列陳述式，建立接收範例資料的資料表。

```
create table firehose_test_table
(
  TICKER_SYMBOL varchar(4),
  SECTOR varchar(16),
  CHANGE float,
  PRICE float
);
```

2. 打開 Firehose 主控台，[網址為 https://console.aws.amazon.com/firehose/](https://console.aws.amazon.com/firehose/)。
3. 選擇「Firehose」串流。
4. 編輯 Firehose 串流的目的地詳細資料，以指向新建立的資料表 firehose\_test\_table。
5. 在 Test with demo data (以示範資料測試) 底下，選擇 Start sending demo data (開始傳送示範資料) 以產生股票行情資料範例。
6. 按照螢幕上的指示，確認資料交付到表格中。請注意，根據緩衝設定，新的列可能需要幾分鐘才會出現於您的表格中。
7. 測試完成後，選擇 Stop sending demo data (停止傳送示範資料) 以停止產生使用費。
8. 編輯 Firehose 串流的目的地詳細資料，以指向另一個資料表。
9. (選用) 刪除 firehose\_test\_table 表格。

## 使用 OpenSearch 服務作為目的地進行測試

使用下列程序，以 Amazon OpenSearch 服務做為目的地測試您的 Firehose 串流。

使用「服務」測試 Firehose 串 OpenSearch 流

1. 打開 Firehose 主控台，[網址為 https://console.aws.amazon.com/firehose/](https://console.aws.amazon.com/firehose/)。

2. 選擇「Firehose」串流。
3. 在 Test with demo data (以示範資料測試) 底下，選擇 Start sending demo data (開始傳送示範資料) 以產生股票行情資料範例。
4. 依照螢幕上的指示，確認資料是否已傳送至您的 OpenSearch 服務網域。如需詳細資訊，請參閱 Amazon OpenSearch 服務開發人員指南中的搜尋 OpenSearch 服務網域中的文件。
5. 測試完成後，選擇 Stop sending demo data (停止傳送示範資料) 以停止產生使用費。

## 使用 Splunk 做為目的地進行測試

使用下列程序，使用 Splunk 做為目的地來測試您的 Firehose 串流。

使用 Splunk 測試 Firehose 串流

1. 打開 Firehose 主控台，網址為 <https://console.aws.amazon.com/firehose/>。
2. 選擇「Firehose」串流。
3. 在 Test with demo data (以示範資料測試) 底下，選擇 Start sending demo data (開始傳送示範資料) 以產生股票行情資料範例。
4. 確認資料交付到您的 Splunk 索引。Splunk 中的搜尋範例為 `sourcetype="aws:firehose:json"和 index="name-of-your-splunk-index"`。如需進一步了解如何在 Splunk 中搜尋事件，請參閱 Splunk 文件中的 [Search Manual](#)。

如果測試資料未出現於您的 Splunk 索引，則請檢查您的 Amazon S3 儲存貯體是否有失敗事件。另請參閱[資料未傳送至 Splunk](#)。

5. 測試結束後，選擇 Stop sending demo data (停止傳送示範資料) 以停止產生使用費。



## 將資料傳送至 Firehose 串流

您可以使用不同類型的來源將資料傳送到您的 Firehose 串流：您可以使用 Kinesis 資料串流、Kinesis 代理程式或 Amazon 資料 Firehose API 使用 SDK。AWS 您也可以使用 Amazon CloudWatch 日誌、CloudWatch 事件或 AWS IoT 做為資料來源。如果您是 Amazon 資料 Firehose 的新手，請花一些時間熟悉中 [什麼是 Amazon 數據 Firehose？](#) 提到的概念和術語。

### Note

某些 AWS 服務只能將訊息和事件傳送至位於相同地區的 Firehose 串流。如果您在設定 Amazon CloudWatch 日誌、CloudWatch 事件的目標時，Firehose 串流未顯示為選項，或者 AWS IoT，請確認您的 Firehose 串流與其他服務位於相同的區域。

### 主題

- [使用 Kinesis 資料串流寫入 Amazon 資料 Firehose](#)
- [使用 Amazon MSK 寫入 Amazon 數據 Firehose](#)
- [使用 Kinesis 代理程式寫入 Amazon 資料 Firehose](#)
- [使用 SDK 寫入 Amazon 數據 Firehose AWS](#)
- [使用日誌寫入 Amazon 數據 Firehose 件 CloudWatch](#)
- [使用事件寫入 Amazon 數據 Firehose 件 CloudWatch](#)
- [使用寫入 Amazon 數據 Firehose 件 AWS IoT](#)

## 使用 Kinesis 資料串流寫入 Amazon 資料 Firehose

您可以設定 Amazon Kinesis Data Streams，將資訊傳送至 Firehose 串流。

### Important

如果您使用 Kinesis Producer Library (KPL) 來寫入資料到 Kinesis 資料串流，則可使用彙整來合併您寫入至該 Kinesis 資料串流的記錄。如果您接著使用該資料串流作為 Firehose 串流的來源，Amazon Data Firehose 會先將記錄解彙總，然後再將記錄傳送到目的地。如果您將交付串流設定為轉換資料，Amazon Data Firehose 會先將記錄解除彙總，然後再將記錄傳送到這些記

錄。AWS Lambda如需詳細資訊，請參閱[使用 Kinesis Producer Library 開發 Amazon Kinesis Data Streams 生產者及彙整](#)。

1. 登入AWS Management Console並開啟 Amazon 資料 Firehose 主控台，網址為 <https://console.aws.amazon.com/firehose/>。
2. 選擇「建立 Firehose 串流」。在 Name and source (名稱與來源) 的頁面，提供下列欄位的值：

Firehose 分流名稱

您的「Firehose」串流的名稱。

來源

選擇 Kinesis 串流以設定使用 Kinesis 資料串流作為資料來源的 Firehose 串流。然後，您可以使用 Amazon Data Firehose 輕鬆讀取現有資料串流中的資料，並將其載入目的地。

若要使用 Kinesis 資料串流作為來源，在 Kinesis 串流清單中選擇現有串流，或選擇新建來建立新的 Kinesis 資料串流。建立新的串流之後，選擇重新整理來更新 Kinesis 串流清單。如果您有大量串流，使用 Filter by name (依名稱篩選) 來篩選清單。

#### Note

當您將 Kinesis 資料串流設定為 Firehose 串流的來源時，Amazon 資料 Firehose PutRecord 和PutRecordBatch操作都會停用。若要在此情況下將資料新增至 Firehose 串流，請使用 Kinesis Data Streams PutRecord 和PutRecords作業。

Amazon 資料 Firehose 開始從 Kinesis 串流的LATEST位置讀取資料。如需 Kinesis 資料串流位置的詳細資訊，請參閱[GetShardIterator](#)。Amazon 資料 Firehose 會針對每個碎片每秒呼叫 Kinesis Data Streams [GetRecords](#)作業一次。

可以從同一個 Kinesis 串流讀取多個 Firehose 串流。其他 Kinesis 應用程式 (取用者) 也可以讀取相同串流。來自任何 Firehose 串流或其他消費者應用程式的每次呼叫都會計入碎片的整體節流限制。為了避免受到調節限制，請小心規劃您的應用程式。如需有關 Kinesis Data Streams 限制的詳細資訊，請參閱 [Amazon Kinesis Data Streams 限制](#)。

3. 選擇 Next (下一步) 前往[記錄轉換和格式轉換](#)頁面。

# 使用 Amazon MSK 寫入 Amazon 數據 Firehose

您可以將 Amazon MSK 設定為將資訊傳送至 Firehose 串流。

1. 登入 AWS Management Console 並開啟 Amazon 資料 Firehose 主控台，網址為 <https://console.aws.amazon.com/firehose/>。
2. 選擇「建立 Firehose 串流」。

在頁面的選擇來源和目的地區段中，提供以下欄位的值：

## 來源

選擇 Amazon MSK 以設定使用 Amazon MSK 做為資料來源的 Firehose 串流。您可以在 MSK 佈建和 MSK 無伺服器叢集之間選擇。然後，您可以使用 Amazon 資料 Firehose 輕鬆地從特定 Amazon MSK 叢集和主題讀取資料，並將其載入指定的 S3 目的地。

## 目的地

選擇 Amazon S3 作為您的 Firehose 串流的目的地。

在頁面的來源設定區段中，提供以下欄位的值：

## Amazon MSK 叢集連線功能

根據您的叢集組態，選擇私有引導代理程式 (建議使用) 或公有引導代理程式選項。引導代理程式由 Apache Kafka 用戶端用作連線至叢集的起點。公共引導代理程序旨在從外部進行公共訪問 AWS，而私有引導代理程序旨在從內部訪問 AWS。如需有關 Amazon MSK 的詳細資訊，請參閱 [Amazon Managed Streaming for Apache Kafka](#)。

若要透過私有引導代理程式連線至佈建或無伺服器 Amazon MSK 叢集，叢集必須符合以下所有要求。

- 叢集必須為作用中。
- 叢集必須使用 IAM 作為其存取控制方法之一。
- IAM 存取控制方法必須啟用多 VPC 私有連線功能。
- 您必須向此叢集新增以資源為基礎的政策，該政策授予 Amazon 資料 Firehose 服務主體叫用 Amazon MS CreateVpcConnection K API 的權限。

若要透過公有引導代理程式連線至佈建的 Amazon MSK 叢集，叢集必須符合以下所有要求。

- 叢集必須為作用中。

- 叢集必須使用 IAM 作為其存取控制方法之一。
- 叢集必須為可公開存取。

### Amazon MSK 叢集

對於相同的帳戶案例，請指定您的 Firehose 串流將從中讀取資料的 Amazon MSK 叢集的 ARN。

如需跨帳戶案例，請參閱 [從 Amazon MSK 跨帳戶交付](#)。

### 主題

指定您希望交付串流從其中擷取資料的 Apache Kafka 主題。建立 Firehose 串流之後，您就無法更新此主題。

在頁面的 Firehose 串流名稱區段中，提供下列欄位的值：

#### Firehose 分流名稱

指定 Firehose 串流的名稱。

3. 接下來，您可以完成設定記錄轉換和記錄格式轉換的選用步驟。如需更多詳細資訊，請參閱 [記錄轉換和格式轉換](#)。

## 使用 Kinesis 代理程式寫入 Amazon 資料 Firehose

Amazon Kinesis 代理程式是獨立的 Java 軟體應用程式，可做為參考實作，顯示如何收集資料並將其傳送至 Firehose。代理程式會持續監控一組檔案，並將新資料傳送至您的 Firehose 交付串流。代理程式會顯示如何處理檔案輪替、檢查點和失敗時重試。它顯示了如何以可靠，及時和簡單的方式交付數據。它還顯示如何發出 CloudWatch 指標，以更好地監控和疑難排解串流程序。要了解更多信息，[aws/ amazon-kinesis-agent](#)。

根據預設，記錄會從各個檔案根據換行符號 ('\n') 字元進行剖析。不過，代理程式也可以設定為剖析多行記錄 (請參閱 [代理程式組態設定](#))。

您可以在以 Linux 為基礎的伺服器環境安裝代理程式，例如 Web 伺服器、日誌伺服器，及資料庫伺服器。安裝代理程式之後，請指定要監視的檔案和 Firehose 串流來設定代理程式。代理程式設定完成後，會持久地從檔案收集資料，並可靠地將其傳送至 Firehose 串流。

### 主題

- [必要條件](#)
- [登入資料](#)
- [自訂登入資料提供者](#)
- [下載並安裝代理程式](#)
- [設定和啟動代理程式](#)
- [代理程式組態設定](#)
- [監控多個檔案目錄並寫入到多個串流](#)
- [使用代理程式預先處理資料](#)
- [代理程式 CLI 命令](#)
- [常見問答集](#)

## 必要條件

- 您的作業系統必須是 Amazon Linux，或 Red Hat Enterprise Linux 版本 7 或更新版本。
- 2.0.0 版或更新版本代理程式執行時使用的是 JRE 1.8 版或更新版本。1.1.x 版代理程式執行時使用的是 JRE 1.7 或更新版本。
- 如果您使用 Amazon EC2 執行您的代理程式，則請啟動您的 EC2 執行個體。
- 您指定的 IAM 角色或 AWS 登入資料必須具有執行 Amazon 資料 Firehose [PutRecordBatch](#) 操作的權限，代理程式才能將資料傳送至您的 Firehose 串流。如果您啟用代理程式的 CloudWatch 監視，則還需要執行 CloudWatch [PutMetricData](#) 作業的權限。如需詳細資訊，請參閱 [Amazon 的使用 Amazon 資料 Firehose 控制存取 監控 Kinesis 代理程式運作狀態、以及身份驗證和存取控制 CloudWatch](#)。

## 登入資料

使用以下方法之一管理您的 AWS 憑證：

- 建立自訂登入資料提供者。如需詳細資訊，請參閱 [the section called “自訂登入資料提供者”](#)。
- 當您啟動 EC2 執行個體時，指定 IAM 角色。
- 設定代理程式時，請指定 AWS 憑證 (請參閱 [the section called “代理程式組態設定”](#) 底下組態資料表內的 `awsAccessKeyId` 和 `awsSecretAccessKey` 項目)。
- 編輯 `/etc/sysconfig/aws-kinesis-agent` 指定您的 AWS 區域和 AWS 存取金鑰。

- 如果您的 EC2 執行個體位於不同的 AWS 帳戶中，請建立 IAM 角色以提供 Amazon 資料 Firehose 服務的存取權。[在設定代理程式時指定該角色 \(請參 `assumeRoleARN` 和 `assumeRoleExternal I d`\)](#)。使用前述方法之一，指定其他帳戶中可擔任此角色的使用者的 AWS 憑證。

## 自訂登入資料提供者

您可以在下列組態設定中建立自訂憑證提供者，並將其類別名稱和 jar 路徑指定給 Kinesis 代理程式：`userDefinedCredentialsProvider.classname` 和 `userDefinedCredentialsProvider.location`。如需這兩個組態設定的說明，請參閱[the section called “代理程式組態設定”](#)。

若要建立自訂憑證提供者，請定義實作 `AWSCredentialsProvider` 介面的類別，如下列範例所示。

```
import com.amazonaws.auth.AWSCredentials;
import com.amazonaws.auth.AWSCredentialsProvider;
import com.amazonaws.auth.BasicAWSCredentials;

public class YourClassName implements AWSCredentialsProvider {
    public YourClassName() {
    }

    public AWSCredentials getCredentials() {
        return new BasicAWSCredentials("key1", "key2");
    }

    public void refresh() {
    }
}
```

您的類別必須有一個不帶引數的建構函數。

AWS 會定期調用重新整理方法以取得更新的憑證。如果您希望憑證提供者在其整個生命週期內提供不同的憑證，則請在此方法中包含重新整理憑證的程式碼。或者，如果您希望憑證提供者提供靜態 (不變更) 憑證，則可以將此方法保留為空。

## 下載並安裝代理程式

首先，連接至您的執行個體。如需詳細資訊，請參閱《Amazon EC2 Linux 執行個體使用者指南》中的[連線至您的執行個體](#)。如果您無法連線，請參閱《Amazon EC2 Linux 執行個體使用者指南》中的[對連線至您的執行個體進行疑難排解](#)。

然後，使用以下方法之一安裝代理程式。

- 從 Amazon Linux 儲存器安裝代理程式

此方法只適用於 Amazon Linux 執行個體。使用下列命令：

```
sudo yum install -y aws-kinesis-agent
```

2.0.0 版或更新版本代理程式安裝在具有 Amazon Linux 2 (AL2) 作業系統的電腦上。此代理程式版本需要 Java 1.8 或更新版本。如果所需的 Java 版本尚不存在，則代理程式安裝程序會安裝該版本。如需 Amazon Linux 2 的詳細資訊，請參閱 <https://aws.amazon.com/amazon-linux-2/>。

- 從 Amazon S3 儲存器安裝代理程式

此方法適用於 Red Hat Enterprise Linux 以及 Amazon Linux 2 執行個體，因為其會從公開的儲存器安裝代理程式。使用下列命令來下載並安裝最新版的代理程式版本 2.x.x：

```
sudo yum install -y https://s3.amazonaws.com/streaming-data-agent/aws-kinesis-agent-latest.amzn2.noarch.rpm
```

若要安裝特定版本的代理程式，請在命令中指定版本號碼。例如，下列命令會安裝代理程式 v 2.0.1。

```
sudo yum install -y https://streaming-data-agent.s3.amazonaws.com/aws-kinesis-agent-2.0.1-1.amzn1.noarch.rpm
```

如果您有 Java 1.7 且不想將其升級，則可下載與 Java 1.7 相容的 1.x.x 版代理程式。例如，若要下載 1.1.6 版代理程式，您可以使用以下命令：

```
sudo yum install -y https://s3.amazonaws.com/streaming-data-agent/aws-kinesis-agent-1.1.6-1.amzn1.noarch.rpm
```

使用以下命令可下載最新的 1.x.x 版代理程式：

```
sudo yum install -y https://s3.amazonaws.com/streaming-data-agent/aws-kinesis-agent-latest.amzn1.noarch.rpm
```

- 若要從 GitHub 存放庫設定代理程式

1. 首先，請確定您已安裝所需的 Java 版本，具體取決於代理程式版本。
2. 從 [aws-lab amazon-kinesis-agent](#) GitHub / 軟體庫下載代理程式。
3. 瀏覽到下載目錄並執行下列命令以安裝代理程式：

```
sudo ./setup --install
```

- 若要在 Docker 容器中設定代理程式

Kinesis 代理程式也可以透過 [amazonlinux](#) 容器基礎在容器中執行。使用以下 Dockerfile，然後執行 `docker build`。

```
FROM amazonlinux

RUN yum install -y aws-kinesis-agent which findutils
COPY agent.json /etc/aws-kinesis/agent.json

CMD ["start-aws-kinesis-agent"]
```

## 設定和啟動代理程式

### 設定和啟動代理程式

1. 開啟並編輯組態檔案 (如果使用預設檔案存取許可，即以超級使用者身分執行)：`/etc/aws-kinesis/agent.json`

在此組態檔案中，指定代理程式從中收集資料的檔案 (`filePattern`)，以及代理程式將資料傳送至的 Firehose 串流 (`deliveryStream`) 名稱。檔案名稱具備一種模式，可讓代理程式辨識檔案輪換。您可以輪換檔案或建立新的檔案，每秒不超過一次。代理程式會使用檔案建立時間戳記



來決定要追蹤哪些檔案，並將哪些檔案追蹤至 Firehose 串流。如果建立新檔案或輪換檔案的頻率超過每秒一次，將導致代理程式無法正確區分這些檔案。

```
{
  "flows": [
    {
      "filePattern": "/tmp/app.log*",
      "deliveryStream": "yourdeliverystream"
    }
  ]
}
```

預設 AWS 區域為 us-east-1。如果您使用不同的區域，請將 `firehose.endpoint` 設定新增至組態檔案，藉此指定您區域的端點。如需詳細資訊，請參閱 [代理程式組態設定](#)。

## 2. 手動啟動代理程式：

```
sudo service aws-kinesis-agent start
```

## 3. (選用) 設定代理程式在系統啟動時開始執行：

```
sudo chkconfig aws-kinesis-agent on
```

代理程式現在已做為系統服務在背景執行。它會持續監視指定的檔案，並將資料傳送至指定的 Firehose 串流。代理程式的活動記錄於 `/var/log/aws-kinesis-agent/aws-kinesis-agent.log`。

## 代理程式組態設定

代理程式支援兩種必要的組態設定 `filePattern` 和 `deliveryStream`，以及用於其他功能的選用組態設定。您可以在 `/etc/aws-kinesis/agent.json` 指定必要及選用的組態設定。

當您變更組態檔案時，必須使用下列命令停止及啟動代理程式：

```
sudo service aws-kinesis-agent stop
sudo service aws-kinesis-agent start
```

或者，您可以使用下列命令：

```
sudo service aws-kinesis-agent restart
```

以下是一般組態設定。

組態設定	描述
<code>assumeRoleARN</code>	使用者欲擔任角色的 Amazon Resource Name (ARN)。如需更多資訊，請參閱《IAM 使用者指南》中的 <a href="#">使用 IAM 角色跨 AWS 帳戶委派存取</a> 。
<code>assumeRoleExternalId</code>	選用的識別符決定誰可以擔任此角色。如需詳細資訊，請參閱《IAM 使用者指南》中的 <a href="#">如何使用外部 ID</a> 。
<code>awsAccessKeyId</code>	覆寫預設憑證的 AWS 存取金鑰 ID。此設定優先於所有其他登入資料供應商。
<code>awsSecretAccessKey</code>	覆寫預設憑證的 AWS 秘密金鑰。此設定優先於所有其他登入資料供應商。
<code>cloudwatch.emitMetrics</code>	啟用代理程式發出指標 (如 CloudWatch 果設定為 true)。 預設：true
<code>cloudwatch.endpoint</code>	的區域端點 CloudWatch。 預設：monitoring.us-east-1.amazonaws.com
<code>firehose.endpoint</code>	Amazon 數據 Firehose 的區域端點。 預設：firehose.us-east-1.amazonaws.com
<code>sts.endpoint</code>	適用於 AWS Security Token Service 的區域端點。 預設：https://sts.amazonaws.com
<code>userDefinedCredentialsProvider.classname</code>	如果您定義自訂登入資料提供者，請使用此設定提供其完整類別名稱。不要在類別名稱的末尾包含 <code>.class</code> 。
<code>userDefinedCredential</code>	如果您定義自訂登入資料提供者，請使用此設定來指定包含自訂登入資料提供者之 jar 的絕對路徑。代理程式也會在下列位置尋找 jar 檔案： <code>/usr/share/aws-kinesis-agent/lib/</code> 。

組態設定	描述
ialsProvider.location	

以下是流程組態設定。

組態設定	描述
aggregateRecordSizeBytes	<p>若要讓代理程式彙總記錄，然後在一次作業中將其放入 Firehose 串流，請指定此設定。將其設定為您希望在代理程式將彙總記錄放入 Firehose 串流之前具有的大小。</p> <p>預設值：0 (無彙總)</p>
dataProcessingOptions	<p>在傳送至 Firehose 串流之前，會套用至每個已剖析記錄的處理選項清單。此處理選項會在指定的資料夾執行。如需詳細資訊，請參閱 <a href="#">使用代理程式預先處理資料</a>。</p>
deliveryStream	[必填] 「Firehose」串流的名稱。
filePattern	<p>[必要] 需要代理程式監控的檔案部分。符合此模式的任何檔案將由代理程式自動挑選及監控。對於符合此模式的所有檔案，請將讀取許可授與 <code>aws-kinesis-agent-user</code> 。對於包含檔案的目錄，請將讀取和執行許可授與 <code>aws-kinesis-agent-user</code> 。</p> <div style="border: 1px solid #f08080; border-radius: 10px; padding: 10px; margin-top: 10px;"> <p><b>⚠ Important</b></p> <p>代理程式會挑選符合此模式的任何檔案。若要確保代理程式不會挑選意外的記錄，請小心選擇此模式。</p> </div>
initialPosition	<p>檔案開始進行剖析的初始位置。有效值為 <code>START_OF_FILE</code> 和 <code>END_OF_FILE</code> 。</p> <p>預設：<code>END_OF_FILE</code></p>
maxBufferAgeMillis	代理程式在將資料傳送至 Firehose 串流之前緩衝資料的時間上限 (以毫秒為單位)。

組態設定	描述
	<p>數值範圍：1,000–900,000 (1 秒到 15 分鐘)</p> <p>預設：60,000 (1 分鐘)</p>
maxBuffer SizeBytes	<p>代理程式會在將資料傳送至 Firehose 串流之前緩衝資料的大小上限 (以位元組為單位)。</p> <p>數值範圍：1–4,194,304 (4 MB)</p> <p>預設：4,194,304 (4 MB)</p>
maxBuffer SizeRecords	<p>代理程式會在將資料傳送至 Firehose 串流之前緩衝資料的記錄數目上限。</p> <p>數值範圍：1–500</p> <p>預設：500</p>
minTimeBe tweenFile PollsMillis	<p>代理程式輪詢和剖析檔案以找出新資料的時間間隔 (以毫秒為單位)。</p> <p>數值範圍：1 或以上</p> <p>預設：100</p>
multiLine StartPattern	<p>用於識別記錄開始處的模式。記錄是由符合模式的一列及不符合模式的任何幾列所組成。有效值為常規運算式。根據預設，每個新日誌檔中的新列會剖析為一筆記錄。</p>
skipHeaderLines	<p>代理程式剖析監控檔案開頭部分時略過的列數。</p> <p>數值範圍：0 或以上</p> <p>預設：0 (零)</p>
truncated RecordTer minator	<p>當記錄大小超過 Amazon Data Firehose 記錄大小限制時，代理程式用來截斷已剖析記錄的字串。(1,000 KB)</p> <p>預設：'\n' (換行符號)</p>

## 監控多個檔案目錄並寫入到多個串流

透過指定多個流程組態設定，您可以設定代理程式來監控多個檔案目錄，然後將資料傳送到多個串流。在下列組態範例中，代理程式會監控兩個檔案目錄，並將資料分別傳送至 Kinesis 資料串流和 Firehose 串流。您可以為 Kinesis 資料串流和 Amazon 資料 Firehose 指定不同的端點，這樣您的資料串流和 Firehose 串流就不需要位於同一個區域。

```
{
  "cloudwatch.emitMetrics": true,
  "kinesis.endpoint": "https://your/kinesis/endpoint",
  "firehose.endpoint": "https://your/firehose/endpoint",
  "flows": [
    {
      "filePattern": "/tmp/app1.log*",
      "kinesisStream": "yourkinesisstream"
    },
    {
      "filePattern": "/tmp/app2.log*",
      "deliveryStream": "yourfirehosedeliverystream"
    }
  ]
}
```

如需將代理程式與 Amazon Kinesis Data Streams 搭配使用的詳細資訊，請參閱[使用 Kinesis 代理程式寫入 Amazon Kinesis Data Streams](#)。

## 使用代理程式預先處理資料

代理程式可以預先處理從受監控檔案剖析的記錄，然後再將這些記錄傳送至 Firehose 串流。您可以將 `dataProcessingOptions` 組態設定新增到您的檔案流程以啟用此功能。可新增一個或多個處理選項，這些選項將依照指定的順序執行。

代理程式支援以下處理選項。由於代理程式是開放原始碼，因此您可進一步開發和擴展其處理選項。您可以從[Kinesis 代理程式](#)下載代理程式。

### 處理選項

#### SINGLELINE

移除換行字元、前方空格及結尾空格，藉此將多列記錄轉換為單列記錄。

```
{
  "optionName": "SINGLELINE"
}
```

## CSVTOJSON

將記錄從分隔符號區隔格式轉換為 JSON 格式。

```
{
  "optionName": "CSVTOJSON",
  "customFieldNames": [ "field1", "field2", ... ],
  "delimiter": "yourdelimiter"
}
```

### customFieldNames

[必要] 欄位名稱在每個 JSON 鍵值對中做為鍵。例如，如果您指定 ["f1", "f2"]，記錄「v1、v2」將轉換為 {"f1":"v1","f2":"v2"}。

### delimiter

在記錄做為分隔符號的字串。預設為逗號 (,)。

## LOGTOJSON

將記錄從日誌格式轉換為 JSON 格式。支援的日誌格式為 Apache Common Log、Apache Combined Log、Apache Error Log、以及 RFC3164 Syslog。

```
{
  "optionName": "LOGTOJSON",
  "logFormat": "logformat",
  "matchPattern": "yourregexpattern",
  "customFieldNames": [ "field1", "field2", ... ]
}
```

### logFormat

[必要] 日誌項目格式。以下是可能的值：

- COMMONAPACHELOG – Apache Common Log 格式。根據預設，每個日誌項目皆有以下模式：「%{host} %{ident} %{authuser} [%{datetime}] \"%{request}\" %{response} %{bytes}」。

- COMBINEDAPACHELOG – Apache Combined Log 格式。根據預設，每個日誌項目皆有以下模式：「`{host} {ident} {authuser} [{datetime}] \"%{request}\"`  
`{response} {bytes} {referrer} {agent}`」。
- APACHEERRORLOG – Apache Error Log 格式。根據預設，每個日誌項目皆有以下模式：  
「`{timestamp} [{module}:{severity}] [pid {processid}:tid`  
`{threadid}] [client: {client}] {message}`」。
- SYSLOG – RFC3164 Syslog 格式。根據預設，每個日誌項目皆有以下模式：  
「`{timestamp} {hostname} {program}[{processid}]: {message}`」。

### matchPattern

覆寫指定日誌格式的預設模式。如果日誌項目使用自訂格式，則使用此設定從日誌項目擷取值。若您指定 `matchPattern`，您也必須指定 `customFieldNames`。

### customFieldNames

自訂欄位名稱在每個 JSON 鍵值對中做為鍵。您可以使用此設定來定義從 `matchPattern` 擷取的值的欄位名稱，或覆寫預先定義的日誌格式的預設欄位名稱。

### Example : LOGTOJSON 組態

這裡提供一個 Apache Common Log 項目轉換為 JSON 格式的 LOGTOJSON 組態範例：

```
{
  "optionName": "LOGTOJSON",
  "logFormat": "COMMONAPACHELOG"
}
```

轉換前：

```
64.242.88.10 - - [07/Mar/2004:16:10:02 -0800] "GET /mailman/listinfo/hsdivision
HTTP/1.1" 200 6291
```

轉換後：

```
{"host":"64.242.88.10","ident":null,"authuser":null,"datetime":"07/
Mar/2004:16:10:02 -0800","request":"GET /mailman/listinfo/hsdivision
HTTP/1.1","response":"200","bytes":"6291"}
```

**Example**：使用自訂欄位的 LOGTOJSON 組態

以下是另一個 LOGTOJSON 組態範例：

```
{
  "optionName": "LOGTOJSON",
  "logFormat": "COMMONAPACHELOG",
  "customFieldNames": ["f1", "f2", "f3", "f4", "f5", "f6", "f7"]
}
```

使用此組態設定，前一個範例的相同 Apache Common Log 項目轉換為 JSON 格式如下：

```
{"f1":"64.242.88.10","f2":null,"f3":null,"f4":"07/Mar/2004:16:10:02 -0800","f5":"GET /
mailman/listinfo/hsdivision HTTP/1.1","f6":"200","f7":"6291"}
```

**Example**：轉換 Apache Common Log 項目

下列流程組態將 Apache Common Log 項目轉換為 JSON 格式的單列記錄：

```
{
  "flows": [
    {
      "filePattern": "/tmp/app.log*",
      "deliveryStream": "my-delivery-stream",
      "dataProcessingOptions": [
        {
          "optionName": "LOGTOJSON",
          "logFormat": "COMMONAPACHELOG"
        }
      ]
    }
  ]
}
```

**Example**：轉換多列記錄

以下流程組態剖析第一行從「[SEQUENCE=」開始的多列記錄。每筆記錄都會先轉換為單列記錄。然後，根據定位鍵分隔符號從記錄中擷取值。擷取的值會對應到指定的 customFieldNames 值以形成 JSON 格式的單列記錄。

```
{
  "flows": [
```



```

    {
      "filePattern": "/tmp/app.log*",
      "deliveryStream": "my-delivery-stream",
      "multilineStartPattern": "\\[SEQUENCE=",
      "dataProcessingOptions": [
        {
          "optionName": "SINGLELINE"
        },
        {
          "optionName": "CSVTOJSON",
          "customFieldNames": [ "field1", "field2", "field3" ],
          "delimiter": "\\t"
        }
      ]
    }
  ]
}

```

#### Example : 使用匹配模式的 LOGTOJSON 組態

以下是 Apache Common Log 項目轉換為 JSON 格式的 LOGTOJSON 組態範例，省略最後欄位 (位元組)：

```

{
  "optionName": "LOGTOJSON",
  "logFormat": "COMMONAPACHELOG",
  "matchPattern": "^(([\\d.]+) ([\\S+]) ([\\S+]) \\[[([\\w:/]+\\s[+\\-]\\d{4})\\]\\] \\\"(.+?)\\\" ([\\d]{3})\"",
  "customFieldNames": ["host", "ident", "authuser", "datetime", "request", "response"]
}

```

轉換前：

```

123.45.67.89 - - [27/Oct/2000:09:27:09 -0400] "GET /java/javaResources.html HTTP/1.0"
200

```

轉換後：

```

{"host":"123.45.67.89","ident":null,"authuser":null,"datetime":"27/Oct/2000:09:27:09-0400","request":"GET /java/javaResources.html HTTP/1.0","response":"200"}

```

## 代理程式 CLI 命令

在系統啟動時自動開始執行代理程式：

```
sudo chkconfig aws-kinesis-agent on
```

檢查代理程式的狀態：

```
sudo service aws-kinesis-agent status
```

停止代理程式：

```
sudo service aws-kinesis-agent stop
```

從這個位置讀取代理程式的日誌檔案：

```
/var/log/aws-kinesis-agent/aws-kinesis-agent.log
```

解除安裝代理程式：

```
sudo yum remove aws-kinesis-agent
```

## 常見問答集

是否有適用於 Windows 的 Kinesis 代理程式？

[適用於 Windows 的 Kinesis 代理程式](#) 是不同於適用於 Linux 平台的 Kinesis 代理程式的軟體。

為什麼 Kinesis 代理程式會減速和/或 **RecordSendErrors** 增加？

這通常是由於來自 Kinesis 的限流。檢查 Kinesis Data Streams

的 `WriteProvisionedThroughputExceeded` 量度或 Firehose 串流的 `ThrottledRecords` 量度。

這些指標中從 0 開始的任何增量，均表示需要提升串流限制。如需詳細資訊，請參閱 [Kinesis 資料串流限制](#) 和 [Firehose 串流](#)。

排除限流之後，請查看 Kinesis 代理程式是否設定為追蹤大量小型檔案。Kinesis 代理程式追蹤新檔案時會有延遲，因此 Kinesis 代理程式應追蹤少量較大的檔案。嘗試將日誌檔案合併至較大的檔案中。

## 為什麼我會遇到 `java.lang.OutOfMemoryError` 例外狀況？

Kinesis 代理程式沒有足夠的記憶體可以處理其目前的工作負載。嘗試增加 `/usr/bin/start-aws-kinesis-agent` 中的 `JAVA_START_HEAP` 和 `JAVA_MAX_HEAP` 並重新啟動代理程式。

## 為什麼我會遇到 `IllegalStateException : connection pool shut down` 例外狀況？

Kinesis 代理程式沒有足夠的連線可以處理其目前的工作負載。嘗試在位於 `/etc/aws-kinesis/agent.json` 的一般代理程式組態設定中增加 `maxConnections` 和 `maxSendingThreads`。這些欄位的預設值是可用執行期處理器的 12 倍。如需進階用戶端組態設定的詳細資訊，請參閱 [AgentConfiguration.java](#)。

## 如何使用 Kinesis 代理程式對另一個問題進行偵錯？

可以在 `/etc/aws-kinesis/log4j.xml` 中啟用 `DEBUG` 層級日誌。

## 我應該如何對 Kinesis Agent 進行設定？

`maxBufferSizeBytes` 越小，Kinesis 代理程式傳送資料的頻率就越高。這可能很好，因為這樣會減少記錄的交付時間，但也增加了 Kinesis 的每秒請求。

## 為什麼 Kinesis 代理程式傳送重複的日誌？

發生這種情況是由於檔案追蹤組態錯誤。請確保每個 `fileFlow's filePattern` 僅與一個檔案相符。如果正在 `copytruncate` 模式中使用 `logrotate` 模式下，也可能發生這種情況。嘗試將模式變更為預設模式，或建立模式以避免重複。如需有關處理重複記錄的詳細資訊，請參閱 [處理重複記錄](#)。

## 使用 SDK 寫入 Amazon 數據 Firehose AWS

[您可以使用 Amazon 數據 Firehose API，將數據發送到 Firehose 流使用](#)

[Java，.NET，Node.js，Python 或紅寶石的開發AWS套件](#)。如果您是 Amazon 資料 Firehose 的新手，請花一些時間熟悉 [什麼是 Amazon 數據 Firehose？](#) 提到的概念和術語。如需詳細資訊，請參閱 [Amazon Web Services 開發功能入門](#)。

這些範例不代表可立即生產的程式碼，無法檢查出所有可能的例外狀況，也不可視為任何潛在安全或效能疑慮的原因。

Amazon 資料防火管 API 提供兩種操作，用於將資料傳送到您的 Firehose 串流：[PutRecord](#)和 [PutRecordBatch](#) `PutRecord()` 在一次通話中發送一個數據記錄，並 `PutRecordBatch()` 可以在一個呼叫中發送多個數據記錄。

## 主題

- [使用單一寫入作業 PutRecord](#)
- [Batch 寫入作業使用 PutRecordBatch](#)

## 使用單一寫入作業 PutRecord

放置資料只需要 Firehose 串流名稱和位元組緩衝區 (<=1000 KB)。由於 Amazon 資料 Firehose 會先批次處理多筆記錄，再將檔案載入 Amazon S3，因此您可能需要新增記錄分隔符號。若要將資料一次放入 Firehose 串流中，請使用下列程式碼：

```
PutRecordRequest putRecordRequest = new PutRecordRequest();
putRecordRequest.setDeliveryStreamName(deliveryStreamName);

String data = line + "\n";

Record record = new Record().withData(ByteBuffer.wrap(data.getBytes()));
putRecordRequest.setRecord(record);

// Put record into the DeliveryStream
firehoseClient.putRecord(putRecordRequest);
```

如需更多程式碼內容，請參閱 AWS SDK 隨附的範本程式碼。如需要要求和回應語法的相關資訊，請參閱 [Firehose API 作業](#) 中的相關主題。

## Batch 寫入作業使用 PutRecordBatch

放置資料只需要 Firehose 串流名稱和記錄清單。由於 Amazon 資料 Firehose 會先批次處理多筆記錄，再將檔案載入 Amazon S3，因此您可能需要新增記錄分隔符號。若要將資料記錄批次放入 Firehose 串流中，請使用下列程式碼：

```
PutRecordBatchRequest putRecordBatchRequest = new PutRecordBatchRequest();
putRecordBatchRequest.setDeliveryStreamName(deliveryStreamName);
putRecordBatchRequest.setRecords(recordList);

// Put Record Batch records. Max No.Of Records we can put in a
// single put record batch request is 500
firehoseClient.putRecordBatch(putRecordBatchRequest);

recordList.clear();
```

如需更多程式碼內容，請參閱 AWS SDK 隨附的範本程式碼。如需要求和回應語法的相關資訊，請參閱 [Firehose API 作業](#) 中的相關主題。

## 使用日誌寫入 Amazon 數據 Firehose 件 CloudWatch

CloudWatch 日誌事件可以使用 CloudWatch 訂閱過濾器發送到 Firehose。如需詳細資訊，請參閱 [Amazon 資料 Firehose 的訂閱篩選器](#)。

CloudWatch 記錄事件會以壓縮的 gzip 格式傳送至 Firehose。如果您想要將解壓縮的記錄事件傳送至 Firehose 目的地，可以使用 Firehose 中的解壓縮功能來自動解壓縮記錄檔。CloudWatch

### Important

目前，Firehose 不支援將 CloudWatch 日誌傳遞到 Amazon OpenSearch 服務目的地，因為 Amazon 將多個日誌事件 CloudWatch 合併到一個 Firehose 記錄中，而且 Amazon OpenSearch 服務無法在一個記錄中接受多個日誌事件。作為替代方案，您可以考慮在 [CloudWatch 日誌中使用 Amazon OpenSearch 服務的訂閱過濾器](#)。

## 記錄檔解壓縮 CloudWatch

如果您使用 Firehose 傳送 CloudWatch 記錄檔，並且想要將解壓縮的資料傳送至您的傳送串流目的地，請使用 [Firehose 資料格式轉換 \(實木複合地板、ORC\) 或動態磁碟分割](#)。您必須為 Firehose 傳送串流啟用解壓縮功能。

您可以使用 AWS Management Console、AWS Command Line Interface 或 AWS SDK 啟用解壓縮。

### Note

如果您在串流上啟用解壓縮功能，請僅將該串流用於 CloudWatch 記錄訂閱篩選器，而不是針對付費記錄檔使用。如果您在用來擷取 CloudWatch 記錄和付費記錄的串流上啟用解壓縮功能，則 Firehose 的付費記錄擷取會失敗。此解壓縮功能僅適用於 CloudWatch 記錄檔。

## 記錄檔解壓縮後的訊息擷取 CloudWatch

當您啟用解壓縮時，您也可以選擇啟用訊息擷取。使用訊息擷取時，Firehose 會從解壓縮的記錄檔記錄中過濾掉所有中繼資料，例如擁有者、日誌群組、CloudWatch 日誌串流及其他中繼資料，並僅傳送

訊息欄位內的內容。如果您要將資料傳送至 Splunk 目的地，您必須開啟訊息擷取功能，Splunk 才能剖析資料。以下是有和沒有消息提取解壓縮後的示例輸出。

圖 1：解壓縮後沒有消息提取的樣本輸出：

```
{
  "owner": "111111111111",
  "logGroup": "CloudTrail/logs",
  "logStream": "111111111111_CloudTrail/logs_us-east-1",
  "subscriptionFilters": [
    "Destination"
  ],
  "messageType": "DATA_MESSAGE",
  "logEvents": [
    {
      "id": "31953106606966983378809025079804211143289615424298221568",
      "timestamp": 1432826855000,
      "message": "{\"eventVersion\":\"1.03\",\"userIdentity\":{\"type\":\"Root1\"}}"
    },
    {
      "id": "31953106606966983378809025079804211143289615424298221569",
      "timestamp": 1432826855000,
      "message": "{\"eventVersion\":\"1.03\",\"userIdentity\":{\"type\":\"Root2\"}}"
    },
    {
      "id": "31953106606966983378809025079804211143289615424298221570",
      "timestamp": 1432826855000,
      "message": "{\"eventVersion\":\"1.03\",\"userIdentity\":{\"type\":\"Root3\"}}"
    }
  ]
}
```

圖 2：使用消息提取解壓縮後的樣本輸出：

```
{"eventVersion":"1.03","userIdentity":{"type":"Root1"}}
{"eventVersion":"1.03","userIdentity":{"type":"Root2"}}
{"eventVersion":"1.03","userIdentity":{"type":"Root3"}}
```

## 啟用和停用解壓縮

您可以使用 AWS Management Console、AWS Command Line Interface 或 AWS SDK 來啟用和停用解壓縮。

## 使用啟用新資料串流的解壓縮功能 AWS Management Console

若要在新資料串流上啟用解壓縮，請使用 AWS Management Console

1. 登入 AWS Management Console 並開啟運動主控台，網址為 <https://console.aws.amazon.com/kinesis>。
2. 在導覽窗格中，選擇 Data Firehose。
3. 選擇 Create Delivery Stream (建立交付串流)。
4. 在「選擇源和目的地」下

### 交付串流來源

您的「Firehose」串流的來源。選擇下列其中一個來源：

- 直接放入 — 選擇此選項可建立製作者應用程式直接寫入的 Firehose 串流。目前，以下是 AWS 與 Firehose Direct PUT 整合的服務、代理商和開放原始碼服務：
- Kinesis 串流：選擇此選項可設定使用 Kinesis 資料串流作為資料來源的 Firehose 串流。然後，您可以使用 Firehose 輕鬆地從現有 Kinesis 資料串流讀取資料，並將其載入目標。如需詳細資訊，請參閱 [使用 Kinesis Data Streams 寫入至 Firehose](#)

### 目的地

Firehose 串流的目的地。選擇下列其中一項：

- Amazon S3
  - Splunk
5. 在 [傳送串流名稱] 底下，輸入串流的名稱。
  6. 在 [轉換記錄-選用] 下：
    - 在「解壓縮來自 Amazon CloudWatch 日誌的來源記錄」區段中，選擇「開啟解壓縮」。
    - 如果要在解壓縮後使用訊息擷取，請選擇 [開啟訊息擷取]。

## 使用啟用現有資料串流的解壓縮 AWS Management Console

如果您的 Firehose 串流具有 Lambda 函數來執行解壓縮，您可以使用 Firehose 解壓縮功能來取代它。在繼續之前，請檢閱 Lambda 函數程式碼，確認它只執行解壓縮或訊息擷取。Lambda 函數的輸出看起來應類似於上一節中的 [圖 1] 或 [圖 2] 所示的範例。如果輸出看起來相似，您可以使用下列步驟取代 Lambda 函數。

1. 以此[藍圖](#)取代您目前的 Lambda 函數。新的藍圖 Lambda 函數會自動偵測傳入資料是壓縮還是解壓縮。它僅在其輸入數據被壓縮時執行解壓縮。
2. 使用內建的 Firehose 選項來開啟解壓縮功能。
3. 啟用 Firehose 串流的 CloudWatch 指標 (如果尚未啟用)。監視測量結果 CloudWatchProcessorLambda\_IncomingCompressedData 並等待此測量結果變更為零。這會確認傳送至 Lambda 函數的所有輸入資料都已解壓縮，而且不再需要 Lambda 函數。
4. 移除 Lambda 資料轉換，因為您不再需要它來解壓縮串流。

## 使用停用解壓縮 AWS Management Console

若要停用資料串流上的解壓縮，請使用 AWS Management Console

1. 登入 AWS Management Console 並開啟運動主控台，網址為 <https://console.aws.amazon.com/kinesis>。
2. 在導覽窗格中，選擇 Data Firehose。
3. 選擇您要編輯的交付串流。
4. 選擇 Configuration (組態)。
5. 在「轉換記錄」區段中，選擇「編輯」。
6. 在「從 Amazon CloudWatch 日誌解壓縮來源記錄」下方，清除「開啟解壓縮」，然後選擇「儲存變更」。

## 常見問答集

如果在解壓縮過程中出現錯誤，源數據會發生什麼？

如果 Amazon Data Firehose 無法解壓縮記錄，則記錄會按原樣 (以壓縮格式) 傳送至您在交付串流建立期間指定的錯誤 S3 儲存貯體。除了記錄之外，交付的物件還包含錯誤代碼和錯誤訊息，這些物件將傳遞到名為的 S3 儲存貯體前綴decompression-failed。在記錄解壓失敗後，Firehose 會繼續處理其他紀錄。

如果在成功解壓縮後處理管道中出現錯誤，源數據會發生什麼？

如果 Amazon Data Firehose 在解壓縮後 (例如動態分割和資料格式轉換) 在處理步驟中發生錯誤，記錄會以壓縮格式傳送到您在交付串流建立時間期間指定的錯誤 S3 儲存貯體。除了記錄之外，傳送的物件也包含錯誤碼和錯誤訊息。



## 如果出現錯誤或異常，您將如何通知？

如果解壓縮期間發生錯誤或例外狀況，如果您設定 CloudWatch 記錄檔，Firehose 會將錯誤訊息記錄到 CloudWatch 記錄檔中。此外，Firehose 會將指標傳送至您可以監控的 CloudWatch 度量。您也可以根據 Firehose 發出的指標，選擇性地建立警示。

## 當put操作不來自 CloudWatch 日誌時會發生什麼情況？

如果客戶puts不來自 CloudWatch 記錄檔，則會傳回下列錯誤訊息：

```
Put to Firehose failed for AccountId: <accountID>, FirehoseName: <firehosename> because the request is not originating from allowed source types.
```

## Firehose 針對解壓縮功能發出哪些指標？

Firehose 會發出每個記錄的減壓指標。您應該選擇期間（1 分鐘），統計（總和），日期範圍以獲取DecompressedRecords失敗或成功或失DecompressedBytes敗或成功的數量。如需更多詳細資訊，請參閱 [CloudWatch 日誌解壓縮測量結果](#)。

## 使用事件寫入 Amazon 數據 Firehose 件 CloudWatch

您可以將目標新增 CloudWatch 至事件規則，將 Amazon 設定為將 CloudWatch 事件傳送至 Firehose 串流。

若要為將 CloudWatch 事件傳送至現有交付串流的事件規則建立目標

1. 請登入AWS Management Console並開啟 CloudWatch 主控台，[網址為 https://console.aws.amazon.com/cloudwatch/](https://console.aws.amazon.com/cloudwatch/)。
2. 選擇 Create rule (建立規則)。
3. 在「步驟 1：建立規則」頁面上，針對「目標」選擇「新增目標」，然後選擇「Firehose Firehose 串流」。
4. 對於 Firehose 串流，請選擇現有的 Amazon 資料 Firehose 交付串流。

如需有關建立 [CloudWatch 事件規則](#)的詳細資訊，請參閱 [Amazon CloudWatch 活動入門](#)。

## 使用寫入 Amazon 數據 Firehose 件 AWS IoT

您可以透過新增動作AWS IoT來設定將資訊傳送至 Firehose 串流。

## 若要建立將事件傳送至現有 Firehose 串流的動作

1. 在 AWS IoT 主控台中建立規則時，在建立規則頁面的設定一或多個動作下，選擇新增動作。
2. 選擇將訊息傳送至 Amazon Kinesis Firehose 串流。
3. 選擇 Configure action (設定動作)。
4. 在「串流名稱」中，選擇現有的「Firehose 串流」。
5. 在 Separator (分隔符號) 的部分，選擇欲插入記錄間的分隔符號字元。
6. 對於 IAM 角色名稱，選擇現有 IAM 角色或選擇建立新角色。
7. 選擇新增動作。

如需建立 AWS IoT 規則的詳細資訊，請參閱 [AWS IoT 規則教學課程](#)。

# Amazon 數據 Firehose 中的安全

雲安全 AWS 是最高的優先級。身為 AWS 客戶，您將受益於資料中心和網路架構，專為滿足最敏感安全性組織的需求而打造。

安全是 AWS 與您之間共同承擔的責任。[共同責任模型](#) 將此描述為雲端的安全和雲端內的安全：

- 雲端的安全性 — AWS 負責保護在 AWS 雲端中執行 AWS 服務的基礎架構。AWS 還為您提供可以安全使用的服務。第三方稽核人員定期檢測及驗證安全的效率也是我們 [AWS 合規計劃](#) 的一部分。要了解適用於 Data Firehose 的合規計劃，請參閱 [合規計劃的範圍AWS 服務](#)。
- 雲端中的安全性 — 您的責任取決於您使用的 AWS 服務。您也必須對資料敏感度、組織要求，以及適用法律和法規等其他因素負責。

本文件可協助您了解如何在使用 Data Firehose 時應用共同的責任模型。下列主題說明如何設定 Data Firehose 以符合您的安全性和合規性目標。您還將學習如何使用其他可以幫助您監控和保護 Data Firehose 資源的 AWS 服務。

## 主題

- [Amazon 數據 Firehose 中的數據保護](#)
- [使用 Amazon 資料 Firehose 控制存取](#)
- [監控 Amazon 數據 Firehose](#)
- [Amazon 資料 Firehose 的合規驗證](#)
- [Amazon 數據 Firehose 中的彈性](#)
- [Amazon 數據 Firehose 中的基礎設施](#)
- [Amazon 數據 Firehose 的安全最佳實踐](#)

## Amazon 數據 Firehose 中的數據保護

Amazon 數據 Firehose 使用 TLS 協議對所有傳輸中的數據進行加密。此外，對於在處理期間存放在臨時儲存中的資料，Amazon Data Firehose 會使用總和檢查碼驗證來加密資料，[AWS Key Management Service](#) 並驗證資料完整性。

如果您有敏感資料，則可以在使用 Amazon Data Firehose 時啟用伺服器端資料加密。您的做法取決於資料的來源。

**Note**

如果您在透過命令列介面或 API 存取時需要經 AWS 過 FIPS 140-2 驗證的加密模組，請使用 FIPS 端點。如需有關 FIPS 和 FIPS 端點的更多相關資訊，請參閱[聯邦資訊處理標準 \(FIPS\) 140-2 概觀](#)。

## 將 Kinesis Data Streams 視為資料來源的伺服器端加密

當您從資料生產者傳送資料到資料串流時，Kinesis Data Streams 會先使用 AWS Key Management Service (AWS KMS) 金鑰加密您的資料，然後再儲存靜態資料。當您的 Amazon 資料火管 Firehose 串流從您的資料串流讀取資料時，Kinesis Data Streams 會先解密資料，然後將資料傳送至 Amazon 資料 Firehose。Amazon 資料 Firehose 會根據您指定的緩衝提示來緩衝記憶體中的資料。再傳送到您的目的地，而不會將未加密的資料儲存為靜態資料。

如需如何為 Kinesis 資料串流啟用伺服器端加密的相關資訊，請參閱《Amazon Kinesis Data Streams 開發人員指南》中的[使用伺服器端加密](#)。

## 使用 Direct PUT 或其他資料來源的伺服器端加密

如果您使用或將資料傳送到 Firehose 串流 [PutRecordBatch](#)，[PutRecord](#) 或者如果您使用 AWS IoT Amazon CloudWatch 日誌或 CloudWatch 事件傳送資料，則可以使用 [StartDeliveryStreamEncryption](#) 操作開啟伺服器端加密。

若要停止 server-side-encryption，請使用該 [StopDeliveryStreamEncryption](#) 作業。

您也可以在建立 Firehose 串流時啟用 SSE。若要這麼做，請指定呼叫 [DeliveryStreamEncryptionConfigurationInput](#) 時間 [CreateDeliveryStream](#)。

當 CMK 屬於類型時 CUSTOMER\_MANAGED\_CMK，如果 Amazon Data Firehose 服務因為 a `KMSNotFoundException`、a、a 或 a `KMSAccessDeniedException` 而無法解密記錄 `KMSInvalidStateException`，則服務會等待最多 24 小時 (保留期)，以便您解決問題。`KMSDisabledException` 如果問題持續超過保留期間，服務會略過那些已經過保留期間且無法解密的記錄，然後捨棄資料。Amazon 資料 Firehose 提供下列四個 CloudWatch 指標，您可以用來追蹤四個 AWS KMS 例外狀況：

- `KMSKeyAccessDenied`
- `KMSKeyDisabled`
- `KMSKeyInvalidState`

- KMSKeyNotFound

如需這四個指標的詳細資訊，請參閱 [the section called “以測 CloudWatch 量結果監視”](#)。

#### Important

若要加密您的 Firehose 串流，請使用對稱 CMK。Amazon 數據 Firehose 不支持非對稱 CMK。如需有關對稱和非對稱 CMK 的詳細資訊，請參閱開發人員指南中的 [關於對稱和非對稱 CMK](#)。AWS Key Management Service

#### Note

當您使用 [客戶管理金鑰](#) (客戶 \_MANAGED\_CMK) 為您的 Firehose 交付串流啟用伺服器端加密 (SSE) 時，Firehose 服務會在使用您的金鑰時設定加密內容。由於此加密內容代表使用您 AWS 帳戶擁有的金鑰的情況，因此會將其記錄為您 AWS 帳戶的 AWS CloudTrail 事件記錄的一部分。此加密內容是 Firehose 服務所產生的系統。您的申請不應對 Firehose 服務設定的加密內容的格式或內容做任何假設。

## 使用 Amazon 資料 Firehose 控制存取

以下各節說明如何控制 Amazon 資料 Firehose 資源的存取和存取。它們涵蓋的資訊包括如何授予應用程式存取權限，以便將資料傳送至 Firehose 串流。他們還描述了如何授予 Amazon 數據 Firehose 訪問您的 Amazon 簡單存儲服務 (Amazon S3) 存儲桶，亞馬遜 Redshift 集群或亞馬遜 OpenSearch 服務集群的訪問權限，以及如果您使用數據多，Dynatrace，MongoDB，新遺物，Splunk 或相撲邏輯作為目的地所需的訪問許可。LogicMonitor 最後，您將在本主題指南中找到如何設定 Amazon Data Firehose，以便將資料傳送到屬於不同 AWS 帳戶的目的地。管理所有這些存取形式的技術是 AWS Identity and Access Management (IAM)。如需 IAM 的相關資訊，請參閱 [什麼是 IAM?](#)。

### 目錄

- [授予您的應用程式對 Amazon 資料 Firehose 資源的存取權](#)
- [授予 Amazon 數據 Firehose 訪問您的私有 Amazon MSK 集群](#)
- [允許 Amazon 數據 Firehose 承擔 IAM 角色](#)
- [授予 Amazon 數據 Firehose 訪問權限以 AWS Glue 進行數據格式轉換](#)
- [授予 Amazon 數據 Firehose 訪問 Amazon S3 目的地](#)

- [授予 Amazon 數據 Firehose 訪問亞 Amazon Redshift 目的地](#)
- [授予 Amazon 資料 Firehose 存取公共 OpenSearch 服務目的地](#)
- [授予 Amazon 資料 Firehose 存取 VPC 中的 OpenSearch 服務目的地](#)
- [授予 Amazon 資料 Firehose 存取公有 OpenSearch無伺服器目的地](#)
- [授予 Amazon 資料 Firehose 存取 VPC 中的 OpenSearch無伺服器目的地](#)
- [授予 Amazon 數據 Firehose 訪問 Splunk 目的地](#)
- [在 VPC 中存取 Splunk](#)
- [訪問雪花或 HTTP 端點](#)
- [授予 Firehose 進入雪花目的地](#)
- [授予 Amazon 資料 Firehose 存取 HTTP 端點目的地](#)
- [從 Amazon MSK 跨帳戶交付](#)
- [跨帳戶交付至 Amazon S3 目的地](#)
- [跨帳戶交付至 OpenSearch服務目的地](#)
- [使用標籤控制存取](#)

## 授予您的應用程式對 Amazon 資料 Firehose 資源的存取權

若要讓應用程式存取 Firehose 串流，請使用類似於此範例的政策。您可以修改 Action 部分，調整授予存取的個別 API 操作，或搭配 "firehose:\*"，將存取權授予給所有操作。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "firehose:DeleteDeliveryStream",
        "firehose:PutRecord",
        "firehose:PutRecordBatch",
        "firehose:UpdateDestination"
      ],
      "Resource": [
        "arn:aws:firehose:region:account-id:deliverystream/delivery-stream-name"
      ]
    }
  ]
}
```

```
]
}
```

## 授予 Amazon 數據 Firehose 訪問您的私有 Amazon MSK 集群

如果 Firehose 串流的來源是私有 Amazon MSK 叢集，請使用類似於此範例的政策。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Principal": {
        "Service": [
          "firehose.amazonaws.com"
        ]
      },
      "Effect": "Allow",
      "Action": [
        "kafka:CreateVpcConnection"
      ],
      "Resource": "cluster-arn"
    }
  ]
}
```

## 允許 Amazon 數據 Firehose 承擔 IAM 角色

本節描述授與 Amazon Data Firehose 存取權，以便從來源到目的地擷取、處理和交付資料的權限和政策。

### Note

如果您使用主控台建立 Firehose 串流，並選擇建立新角色的選項，請將必要的信任原則 AWS 附加至該角色。如果您希望 Amazon Data Firehose 使用現有的 IAM 角色，或者您自行建立角色，請將以下信任政策附加到該角色，以便 Amazon Data Firehose 可以承擔。編輯政策，以您的 **## ID** 取代 AWS 帳戶 ID。如需有關如何修改角色之信任關係的資訊，請參閱[修改角色](#)。

Amazon 資料 Firehose 針對交付串流處理和交付資料所需的所有許可使用 IAM 角色。請確定已將下列信任政策附加到該角色，以便 Amazon 資料 Firehose 可以承擔。

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Sid": "",
    "Effect": "Allow",
    "Principal": {
      "Service": "firehose.amazonaws.com"
    },
    "Action": "sts:AssumeRole",
    "Condition": {
      "StringEquals": {
        "sts:ExternalId": "account-id"
      }
    }
  }]
}
```

此政策使用 `sts:ExternalId` 條件內容金鑰來確保只有源自您 AWS 帳戶的 Amazon Data Firehose 活動才能擔任此 IAM 角色。如需防止未經授權使用 IAM 角色的詳細資訊，請參閱《IAM 使用者指南》中的[混淆代理問題](#)。

如果您選擇 Amazon MSK 做為 Firehose 串流的來源，則必須指定另一個 IAM 角色，以授予 Amazon 資料 Firehose 許可，以從指定的 Amazon MSK 叢集擷取來源資料。請確定已將下列信任政策附加到該角色，以便 Amazon 資料 Firehose 可以承擔。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Principal": {
        "Service": [
          "firehose.amazonaws.com"
        ]
      },
      "Effect": "Allow",
      "Action": "sts:AssumeRole"
    }
  ]
}
```



請確定這個角色授與 Amazon 資料 Firehose 許可以從指定的 Amazon MSK 叢集擷取來源資料，並授予以下許可：

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": [
      "kafka:GetBootstrapBrokers",
      "kafka:DescribeCluster",
      "kafka:DescribeClusterV2",
      "kafka-cluster:Connect"
    ],
    "Resource": "CLUSTER-ARN"
  },
  {
    "Effect": "Allow",
    "Action": [
      "kafka-cluster:DescribeTopic",
      "kafka-cluster:DescribeTopicDynamicConfiguration",
      "kafka-cluster:ReadData"
    ],
    "Resource": "TOPIC-ARN"
  }
]
```

## 授予 Amazon 數據 Firehose 訪問權限以 AWS Glue 進行數據格式轉換

如果您的 Firehose 串流執行資料格式轉換，Amazon 資料 Firehose 會參考儲存於中的資料表定義。AWS Glue 若要為 Amazon 資料 Firehose 提供必要的存取權限 AWS Glue，請在您的政策中新增以下聲明。如需如何尋找表格 ARN 的詳細資訊，請參閱[指定 AWS Glue 合資源 ARN](#)。

```
[{
  "Effect": "Allow",
  "Action": [
    "glue:GetTable",
    "glue:GetTableVersion",
    "glue:GetTableVersions"
  ],
  "Resource": "table-arn"
}]
```

```
}, {  
  "Sid": "GetSchemaVersion",  
  "Effect": "Allow",  
  "Action": [  
    "glue:GetSchemaVersion"  
  ],  
  "Resource": ["*"]  
}]
```

從結構描述登錄檔取得結構描述的建議原則沒有資源限制。如需詳細資訊，請參閱開發人員指 [AWS Glue 南中關於還原序列化程式的 IAM 範例](#)。

### Note

目前，AWS Glue 以色列 (特拉維夫)、亞太區域 (雅加達) 或中東 (阿拉伯聯合大公國) 地區不支援。如果您正在使用亞太區域 (雅加達) 區域或中東 (阿拉伯聯合大公國) 區域的 Amazon Data Firehose，請務必在目前支援的其 AWS Glue AWS Glue 中一個區域提供 Amazon Data Firehose 存取權。支援資料 Firehose 之間的跨區域互通性。AWS Glue 如需支援地 AWS Glue 區的詳細資訊，請參閱 <https://docs.aws.amazon.com/general/latest/gr/glue.html>

## 授予 Amazon 數據 Firehose 訪問 Amazon S3 目的地

當您使用 Amazon S3 目的地時，Amazon Data Firehose 會將資料交付到您的 S3 儲存貯體，並且可以選擇性地使用您擁有的 AWS KMS 金鑰進行資料加密。如果啟用了錯誤記錄，Amazon Data Firehose 也會將資料傳送錯誤傳送到您的 CloudWatch 日誌群組和串流。建立 Firehose 串流時，您必須具備 IAM 角色。Amazon Data Firehose 會擔任 IAM 角色，並取得指定儲存貯體、金鑰以及 CloudWatch 日誌群組和串流的存取權。

使用下列存取政策讓 Amazon 資料 Firehose 存取您的 S3 儲存貯體和 AWS KMS 金鑰。如果您未擁有 S3 儲存貯體，請新增 `s3:PutObjectAc1` 至 Amazon S3 動作清單。這可讓儲存貯體擁有者完全存取 Amazon 資料 Firehose 交付的物件。

```
{  
  "Version": "2012-10-17",  
  "Statement": [  
    {  
      "Effect": "Allow",  
      "Action": [  

```

```

        "s3:AbortMultipartUpload",
        "s3:GetBucketLocation",
        "s3:GetObject",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads",
        "s3:PutObject"
    ],
    "Resource": [
        "arn:aws:s3:::bucket-name",
        "arn:aws:s3:::bucket-name/*"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "kinesis:DescribeStream",
        "kinesis:GetShardIterator",
        "kinesis:GetRecords",
        "kinesis:ListShards"
    ],
    "Resource": "arn:aws:kinesis:region:account-id:stream/stream-name"
},
{
    "Effect": "Allow",
    "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey"
    ],
    "Resource": [
        "arn:aws:kms:region:account-id:key/key-id"
    ],
    "Condition": {
        "StringEquals": {
            "kms:ViaService": "s3.region.amazonaws.com"
        },
        "StringLike": {
            "kms:EncryptionContext:aws:s3:arn": "arn:aws:s3:::bucket-name/
prefix*"
        }
    }
},
{
    "Effect": "Allow",
    "Action": [

```

```

        "logs:PutLogEvents"
    ],
    "Resource": [
        "arn:aws:logs:region:account-id:log-group:log-group-name:log-stream:log-stream-name"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "lambda:InvokeFunction",
        "lambda:GetFunctionConfiguration"
    ],
    "Resource": [
        "arn:aws:lambda:region:account-id:function:function-name:function-version"
    ]
}
]
}

```

上述政策亦有允許存取 Amazon Kinesis Data Streams 的陳述式。如果您不使用 Amazon Kinesis Data Streams 作為資料來源，可移除該陳述式。如果您使用 Amazon MSK 做為來源，則可以使用下列方式取代該陳述式：

```

{
    "Sid": "",
    "Effect": "Allow",
    "Action": [
        "kafka:GetBootstrapBrokers",
        "kafka:DescribeCluster",
        "kafka:DescribeClusterV2",
        "kafka-cluster:Connect"
    ],
    "Resource": "arn:aws:kafka:{{mskClusterRegion}}:{{mskClusterAccount}}:cluster/{{mskClusterName}}/{{clusterUUID}}"
},
{
    "Sid": "",
    "Effect": "Allow",
    "Action": [
        "kafka-cluster:DescribeTopic",
        "kafka-cluster:DescribeTopicDynamicConfiguration",

```

```
    "kafka-cluster:ReadData"
  ],
  "Resource": "arn:aws:kafka:{{mskClusterRegion}}:{{mskClusterAccount}}:topic/
  {{mskClusterName}}/{{clusterUUID}}/{{mskTopicName}}"
},
{
  "Sid": "",
  "Effect": "Allow",
  "Action": [
    "kafka-cluster:DescribeGroup"
  ],
  "Resource": "arn:aws:kafka:{{mskClusterRegion}}:{{mskClusterAccount}}:group/
  {{mskClusterName}}/{{clusterUUID}}/*"
}
```

如需有關允許其他 AWS 服務存取您的 AWS 資源的詳細資訊，請參閱 IAM 使用者指南中的[建立角色以將許可委派給 AWS 服務](#)。

若要了解如何授予 Amazon 資料 Firehose 存取其他帳戶中 Amazon S3 目的地的權限，請參閱[the section called “跨帳戶交付至 Amazon S3 目的地”](#)。

## 授予 Amazon 數據 Firehose 訪問亞 Amazon Redshift 目的地

當您在使用 Amazon Amazon Redshift 目的地時授予 Amazon 資料 Firehose 的存取權時，請參閱以下內容。

### 主題

- [IAM 角色與存取政策](#)
- [VPC 存取 Amazon Redshift 佈建的叢集或 Amazon Redshift Serverless 工作群組](#)

## IAM 角色與存取政策

當您使用 Amazon Redshift 目的地時，Amazon 數據 Firehose 將數據作為中間位置交付到您的 S3 存儲桶。它可以選擇使用您擁有的密 AWS KMS 鑰進行數據加密。然後，亞馬遜數據 Firehose 將資料從 S3 儲存貯體載入到您的 Amazon Redshift 佈建的叢集或 Amazon Redshift 無伺服器工作群組。如果啟用了錯誤記錄，Amazon Data Firehose 也會將資料傳送錯誤傳送到您的 CloudWatch 日誌群組和串流。Amazon 資料 Firehose 會使用指定的 Amazon Redshift 使用者名稱和密碼來存取佈建的叢集或 Amazon Redshift 無伺服器工作群組，並使用 IAM 角色存取指定的儲存貯體、金鑰、CloudWatch 日誌群組和串流。建立 Firehose 串流時，您必須具備 IAM 角色。

使用下列存取政策讓 Amazon 資料 Firehose 存取您的 S3 儲存貯體和 AWS KMS 金鑰。如果您不擁有 S3 儲存貯體，請新增 `s3:PutObjectAcl` 至 Amazon S3 動作清單，以授予儲存貯體擁有者完全存取 Amazon Data Firehose 交付的物件。此政策亦有允許存取 Amazon Kinesis Data Streams 的陳述式。如果您不使用 Amazon Kinesis Data Streams 作為資料來源，可移除該陳述式。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:GetBucketLocation",
        "s3:GetObject",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads",
        "s3:PutObject"
      ],
      "Resource": [
        "arn:aws:s3:::bucket-name",
        "arn:aws:s3:::bucket-name/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey"
      ],
      "Resource": [
        "arn:aws:kms:region:account-id:key/key-id"
      ],
      "Condition": {
        "StringEquals": {
          "kms:ViaService": "s3.region.amazonaws.com"
        },
        "StringLike": {
          "kms:EncryptionContext:aws:s3:arn": "arn:aws:s3:::bucket-name/prefix*"
        }
      }
    }
  ],
}
```

```

    {
      "Effect": "Allow",
      "Action": [
        "kinesis:DescribeStream",
        "kinesis:GetShardIterator",
        "kinesis:GetRecords",
        "kinesis:ListShards"
      ],
      "Resource": "arn:aws:kinesis:region:account-id:stream/stream-name"
    },
    {
      "Effect": "Allow",
      "Action": [
        "logs:PutLogEvents"
      ],
      "Resource": [
        "arn:aws:logs:region:account-id:log-group:log-group-name:log-stream:log-stream-name"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "lambda:InvokeFunction",
        "lambda:GetFunctionConfiguration"
      ],
      "Resource": [
        "arn:aws:lambda:region:account-id:function:function-name:function-version"
      ]
    }
  ]
}

```

如需有關允許其他 AWS 服務存取您的 AWS 資源的詳細資訊，請參閱 IAM 使用者指南中的[建立角色以將許可委派給 AWS 服務](#)。

## VPC 存取 Amazon Redshift 佈建的叢集或 Amazon Redshift Serverless 工作群組

如果您的 Amazon Redshift 佈建的叢集或 Amazon Redshift Serverless 工作群組位於虛擬私有雲端 (VPC)，則其必須有公有 IP 地址，可供公開存取。此外，透過解除封鎖 Amazon 資料 Firehose IP 地址，授予亞馬遜資料防 Firehose 存取您的 Amazon Redshift 佈建叢集或亞馬遜 Redshift 無伺服器工作群組的存取權。Amazon 資料 Firehose 目前針對每個可用區域使用一個 CIDR 區塊：

- 13.58.135.96/27 適用於美國東部 (俄亥俄)
- 52.70.63.192/27 適用於美國東部 (維吉尼亞北部)
- 13.57.135.192/27 適用於美國西部 (加利佛尼亞北部)
- 52.89.255.224/27 適用於美國西部 (奧勒岡)
- 18.253.138.96/27 AWS GovCloud (美國東部)
- 52.61.204.160/27 為 AWS GovCloud (美國西部)
- 35.183.92.128/27 適用於加拿大 (中部)
- 40.176.98.192/27 適用於加拿大西部 (卡加利)
- 18.162.221.32/27 適用於亞太區域 (香港)
- 13.232.67.32/27 適用於亞太區域 (孟買)
- 18.60.192.128/27 適用於亞太區域 (海德拉巴)
- 13.209.1.64/27 適用於亞太區域 (首爾)
- 13.228.64.192/27 適用於亞太區域 (新加坡)
- 13.210.67.224/27 適用於亞太區域 (雪梨)
- 108.136.221.64/27 適用於亞太區域 (雅加達)
- 13.113.196.224/27 適用於亞太區域 (東京)
- 13.208.177.192/27 適用於亞太區域 (大阪)
- 52.81.151.32/27 適用於中國 (北京)
- 161.189.23.64/27 適用於中國 (寧夏)
- 16.62.183.32/27 適用於歐洲 (蘇黎世)
- 35.158.127.160/27 適用於歐洲 (法蘭克福)
- 52.19.239.192/27 適用於歐洲 (愛爾蘭)
- 18.130.1.96/27 適用於歐洲 (倫敦)
- 35.180.1.96/27 適用於歐洲 (巴黎)
- 13.53.63.224/27 適用於歐洲 (斯德哥爾摩)
- 15.185.91.0/27 適用於中東 (巴林)
- 18.228.1.128/27 適用於南美洲 (聖保羅)
- 15.161.135.128/27 適用於歐洲 (米蘭)
- 13.244.121.224/27 適用於非洲 (開普敦)
- 3.28.159.32/27 適用於中東 (阿拉伯聯合大公國)



- 51.16.102.0/27 適用於以色列 (特拉維夫)
- 16.50.161.128/27 適用於亞太區域 (墨爾本)

如需如何解鎖 IP 地址的詳細資訊，請參閱《Amazon Redshift 入門指南》中的[授予叢集存取步驟](#)。

## 授予 Amazon 資料 Firehose 存取公共 OpenSearch 服務目的地

當您使用 OpenSearch 服務目的地時，Amazon Data Firehose 會將資料交付到您的 OpenSearch 服務叢集，並同時將故障或所有文件備份到 S3 儲存貯體。如果啟用了錯誤記錄，Amazon Data Firehose 也會將資料傳送錯誤傳送到您的 CloudWatch 日誌群組和串流。Amazon 資料 Firehose 使用 IAM 角色來存取指定的 OpenSearch 服務網域、S3 儲存貯體、AWS KMS 金鑰以及 CloudWatch 日誌群組和串流。建立 Firehose 串流時，您必須具備 IAM 角色。

使用下列存取政策讓 Amazon 資料 Firehose 存取您的 S3 儲存貯體、OpenSearch 服務網域和 AWS KMS 金鑰。如果您不擁有 S3 儲存貯體，請新增 `s3:PutObjectAcl` 至 Amazon S3 動作清單，以授予儲存貯體擁有者完全存取 Amazon Data Firehose 交付的物件。此政策亦有允許存取 Amazon Kinesis Data Streams 的陳述式。如果您不使用 Amazon Kinesis Data Streams 作為資料來源，可移除該陳述式。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:GetBucketLocation",
        "s3:GetObject",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads",
        "s3:PutObject"
      ],
      "Resource": [
        "arn:aws:s3:::bucket-name",
        "arn:aws:s3:::bucket-name/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt",
```

```

        "kms:GenerateDataKey"
    ],
    "Resource": [
        "arn:aws:kms:region:account-id:key/key-id"
    ],
    "Condition": {
        "StringEquals": {
            "kms:ViaService": "s3.region.amazonaws.com"
        },
        "StringLike": {
            "kms:EncryptionContext:aws:s3:arn": "arn:aws:s3:::bucket-name/prefix*"
        }
    }
},
{
    "Effect": "Allow",
    "Action": [
        "es:DescribeDomain",
        "es:DescribeDomains",
        "es:DescribeDomainConfig",
        "es:ESHttpPost",
        "es:ESHttpPut"
    ],
    "Resource": [
        "arn:aws:es:region:account-id:domain/domain-name",
        "arn:aws:es:region:account-id:domain/domain-name/*"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "es:ESHttpGet"
    ],
    "Resource": [
        "arn:aws:es:region:account-id:domain/domain-name/_all/_settings",
        "arn:aws:es:region:account-id:domain/domain-name/_cluster/stats",
        "arn:aws:es:region:account-id:domain/domain-name/index-name*/_mapping/type-name",
        "arn:aws:es:region:account-id:domain/domain-name/_nodes",
        "arn:aws:es:region:account-id:domain/domain-name/_nodes/stats",
        "arn:aws:es:region:account-id:domain/domain-name/_nodes/*/stats",
        "arn:aws:es:region:account-id:domain/domain-name/_stats",
        "arn:aws:es:region:account-id:domain/domain-name/index-name*/_stats",

```

```

        "arn:aws:es:region:account-id:domain/domain-name/"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "kinesis:DescribeStream",
        "kinesis:GetShardIterator",
        "kinesis:GetRecords",
        "kinesis:ListShards"
    ],
    "Resource": "arn:aws:kinesis:region:account-id:stream/stream-name"
},
{
    "Effect": "Allow",
    "Action": [
        "logs:PutLogEvents"
    ],
    "Resource": [
        "arn:aws:logs:region:account-id:log-group:log-group-name:log-stream:log-stream-name"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "lambda:InvokeFunction",
        "lambda:GetFunctionConfiguration"
    ],
    "Resource": [
        "arn:aws:lambda:region:account-id:function:function-name:function-version"
    ]
}
]
}

```

如需有關允許其他 AWS 服務存取您的 AWS 資源的詳細資訊，請參閱 IAM 使用者指南中的[建立角色以將許可委派給 AWS 服務](#)。

要了解如何授予 Amazon 資料 Firehose 存取其他帳戶中 OpenSearch 服務叢集的權限，請參閱[the section called “跨帳戶交付至 OpenSearch 服務目的地”](#)。

## 授予 Amazon 資料 Firehose 存取 VPC 中的 OpenSearch 服務目的地

如果您的 OpenSearch 服務網域位於 VPC 中，請確定您將上一節所述的許可授予 Amazon 資料 Firehose。此外，您需要向 Amazon 數據 Firehose 授予以下許可，以使其能夠訪問 OpenSearch 服務域的 VPC。

- ec2:DescribeVpcs
- ec2:DescribeVpcAttribute
- ec2:DescribeSubnets
- ec2:DescribeSecurityGroups
- ec2:DescribeNetworkInterfaces
- ec2:CreateNetworkInterface
- ec2:CreateNetworkInterfacePermission
- ec2>DeleteNetworkInterface

### Important

建立交付串流後，請勿撤銷這些權限。如果您撤銷這些權限，則每當 OpenSearch 服務嘗試查詢或更新 ENI 時，Firehose 串流都會降級或停止向您的服務網域傳送資料。

### Important

當您指定子網路以將資料傳送至私有 VPC 中的目的地時，請確定所選子網路中有足夠數量的可用 IP 位址。如果指定的子網路中沒有可用的可用 IP 位址，Firehose 無法在私有 VPC 中為資料傳遞建立或新增 ENI，而且傳遞將會降級或失敗。

當您建立或更新交付串流時，您必須指定 Firehose 在將資料傳送至您的 OpenSearch 服務網域時使用的安全性群組。您可以使用 OpenSearch Service 網域所使用的相同安全性群組或不同的安全性群組。如果您指定不同的安全性群組，請確定它允許輸出 HTTPS 流量傳送至 OpenSearch 服務網域的安全性群組。同時確保 OpenSearch 服務網域的安全性群組允許您設定 Firehose 串流時指定的安全性群組的 HTTPS 流量。如果您對 Firehose 串流和 OpenSearch 服務網域使用相同的安全性群組，請確定安全群組輸入規則允許 HTTPS 流量。如需安全群組規則的詳細資訊，請參閱 Amazon VPC 文件中 OpenSearch 的[安全群組規則](#)。

## 授予 Amazon 資料 Firehose 存取公有 OpenSearch 無伺服器目的地

當您使用 OpenSearch 無伺服器目的地時，Amazon Data Firehose 會將資料交付到您的 OpenSearch 無伺服器集合，並同時將故障或所有文件備份到 S3 儲存貯體。如果啟用了錯誤記錄，Amazon Data Firehose 也會將資料傳送錯誤傳送到您的 CloudWatch 日誌群組和串流。Amazon Data Firehose 使用 IAM 角色來存取指定的 OpenSearch 無伺服器收集、S3 儲存貯體、AWS KMS 金鑰以及 CloudWatch 日誌群組和串流。建立 Firehose 串流時，您必須具備 IAM 角色。

使用下列存取政策讓 Amazon Data Firehose 存取您的 S3 儲存貯體、OpenSearch 無伺服器網域和 AWS KMS 金鑰。如果您不擁有 S3 儲存貯體，請新增 `s3:PutObjectAcl` 至 Amazon S3 動作清單，以授予儲存貯體擁有者完全存取 Amazon Data Firehose 交付的物件。此政策亦有允許存取 Amazon Kinesis Data Streams 的陳述式。如果您不使用 Amazon Kinesis Data Streams 作為資料來源，可移除該陳述式。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:GetBucketLocation",
        "s3:GetObject",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads",
        "s3:PutObject"
      ],
      "Resource": [
        "arn:aws:s3:::bucket-name",
        "arn:aws:s3:::bucket-name/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey"
      ],
      "Resource": [
        "arn:aws:kms:region:account-id:key/key-id"
      ],
      "Condition": {
```

```

        "StringEquals": {
            "kms:ViaService": "s3.region.amazonaws.com"
        },
        "StringLike": {
            "kms:EncryptionContext:aws:s3:arn": "arn:aws:s3:::bucket-name/
prefix*"
        }
    },
    {
        "Effect": "Allow",
        "Action": [
            "kinesis:DescribeStream",
            "kinesis:GetShardIterator",
            "kinesis:GetRecords",
            "kinesis:ListShards"
        ],
        "Resource": "arn:aws:kinesis:region:account-id:stream/stream-name"
    },
    {
        "Effect": "Allow",
        "Action": [
            "logs:PutLogEvents"
        ],
        "Resource": [
            "arn:aws:logs:region:account-id:log-group:log-group-name:log-stream:log-
stream-name"
        ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "lambda:InvokeFunction",
            "lambda:GetFunctionConfiguration"
        ],
        "Resource": [
            "arn:aws:lambda:region:account-id:function:function-name:function-
version"
        ]
    },
    {
        "Effect": "Allow",
        "Action": "aoss:APIAccessAll",
        "Resource": "arn:aws:aoss:region:account-id:collection/collection-id"
    }

```

```
    }
  ]
}
```

除了上述政策之外，您還必須將 Amazon Data Firehose 設定為在資料存取政策中指派下列最低許可：

```
[
  {
    "Rules": [
      {
        "ResourceType": "index",
        "Resource": [
          "index/target-collection/target-index"
        ],
        "Permission": [
          "aoss:WriteDocument",
          "aoss:UpdateIndex",
          "aoss>CreateIndex"
        ]
      }
    ],
    "Principal": [
      "arn:aws:sts::account-id:assumed-role/firehose-delivery-role-name/*"
    ]
  }
]
```

如需有關允許其他 AWS 服務存取您的 AWS 資源的詳細資訊，請參閱 IAM 使用者指南中的[建立角色以將許可委派給 AWS 服務](#)。

## 授予 Amazon 資料 Firehose 存取 VPC 中的 OpenSearch 無伺服器目的地

如果您的 OpenSearch 無伺服器集合位於 VPC 中，請確定您將上一節所述的許可授予 Amazon Data Firehose。此外，您需要授予 Amazon 資料 Firehose 下列權限，才能讓它存取您的 OpenSearch 無伺服器集合的 VPC。

- ec2:DescribeVpcs
- ec2:DescribeVpcAttribute
- ec2:DescribeSubnets
- ec2:DescribeSecurityGroups

- ec2:DescribeNetworkInterfaces
- ec2:CreateNetworkInterface
- ec2:CreateNetworkInterfacePermission
- ec2>DeleteNetworkInterface

### ⚠ Important

建立交付串流後，請勿撤銷這些權限。如果您撤銷這些權限，則每當 OpenSearch 服務嘗試查詢或更新 ENI 時，Firehose 串流都會降級或停止向您的服務網域傳送資料。

### ⚠ Important

當您指定子網路以將資料傳送至私有 VPC 中的目的地時，請確定所選子網路中有足夠數量的可用 IP 位址。如果指定的子網路中沒有可用的可用 IP 位址，Firehose 無法在私有 VPC 中為資料傳遞建立或新增 ENI，而且傳遞將會降級或失敗。

建立或更新交付串流時，您可以指定 Firehose 在將資料傳送至 OpenSearch 無伺服器集合時使用的安全性群組。您可以使用 OpenSearch 無伺服器集合所使用的相同安全性群組，或使用不同的安全性群組。如果您指定不同的安全性群組，請確定它允許輸出 HTTPS 流量傳送至 OpenSearch 無伺服器集合的安全性群組。此外，請確保 OpenSearch 無伺服器集合的安全性群組允許您設定 Firehose 串流時指定的安全性群組的 HTTPS 流量。如果您對 Firehose 串流和 OpenSearch 無伺服器集合使用相同的安全性群組，請確定安全群組輸入規則允許 HTTPS 流量。如需安全群組規則的詳細資訊，請參閱 Amazon VPC 文件中 OpenSearch 的[安全群組規則](#)。

## 授予 Amazon 數據 Firehose 訪問 Splunk 目的地

當您使用 Splunk 目的地時，Amazon 資料 Firehose 會將資料傳送到您的 Splunk HTTP 事件收集器 (HEC) 端點。它也會將該資料備份到您指定的 Amazon S3 儲存貯體，您可以選擇性地使用自己擁有的 AWS KMS 金鑰進行 Amazon S3 伺服器端加密。如果啟用了錯誤記錄，Firehose 會將資料傳送錯誤傳送到您的記 CloudWatch 錄串流。您也可以使用 AWS Lambda 資料轉換。

如果您使用 AWS 負載平衡器，請確定它是 Classic Load Balancer 或 Application Load Balancer。此外，啟用以持續時間為基礎的黏性工作階段，且針對 Classic Load Balancer 停用 Cookie 到期，且 Application Load Balancer 的到期時間上限 (7 天)。如需如何執行此作業的詳細資訊，請參閱 [Classic Load Balancer 或 Application Load Balancer 載平衡器的持續時間型工作階段黏著性](#)。



您在建立交付串流時必須擁有 IAM 角色。Firehose 會擔任 IAM 角色，並取得指定儲存貯體、金鑰、CloudWatch 記錄群組和串流的存取權。

使用下列存取政策啟用 Amazon 資料 Firehose 存取您的 S3 儲存貯體。如果您不擁有 S3 儲存貯體，請新增 `s3:PutObjectAcl` 至 Amazon S3 動作清單，以授予儲存貯體擁有者完全存取 Amazon Data Firehose 交付的物件。此政策也授予 Amazon 資料 Firehose 存取權以 CloudWatch 進行錯誤記錄和 AWS Lambda 資料轉換。此政策亦有允許存取 Amazon Kinesis Data Streams 的陳述式。如果您不使用 Amazon Kinesis Data Streams 作為資料來源，可移除該陳述式。Amazon 數據 Firehose 不使用 IAM 訪問 Splunk。存取 Splunk 時，它會使用您 HEC 符記。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:GetBucketLocation",
        "s3:GetObject",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads",
        "s3:PutObject"
      ],
      "Resource": [
        "arn:aws:s3:::bucket-name",
        "arn:aws:s3:::bucket-name/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey"
      ],
      "Resource": [
        "arn:aws:kms:region:account-id:key/key-id"
      ],
      "Condition": {
        "StringEquals": {
          "kms:ViaService": "s3.region.amazonaws.com"
        },
        "StringLike": {
```

```

        "kms:EncryptionContext:aws:s3:arn": "arn:aws:s3:::bucket-name/
prefix*"
    }
}
},
{
    "Effect": "Allow",
    "Action": [
        "kinesis:DescribeStream",
        "kinesis:GetShardIterator",
        "kinesis:GetRecords",
        "kinesis:ListShards"
    ],
    "Resource": "arn:aws:kinesis:region:account-id:stream/stream-name"
},
{
    "Effect": "Allow",
    "Action": [
        "logs:PutLogEvents"
    ],
    "Resource": [
        "arn:aws:logs:region:account-id:log-group:log-group-name:log-stream:*"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "lambda:InvokeFunction",
        "lambda:GetFunctionConfiguration"
    ],
    "Resource": [
        "arn:aws:lambda:region:account-id:function:function-name:function-
version"
    ]
}
]
}
}

```

如需有關允許其他 AWS 服務存取您的 AWS 資源的詳細資訊，請參閱 IAM 使用者指南中的[建立角色以將許可委派給 AWS 服務](#)。

## 在 VPC 中存取 Splunk

如果您的 Splunk 平台位在 VPC，則其必須有公有 IP 地址，可供公開存取。此外，透過解除封鎖 Amazon 資料 Firehose IP 地址，授予 Amazon 資料 Firehose 存取您的 Splunk 平台的權限。Amazon 資料 Firehose 目前使用下列 CIDR 區塊。

- 18.216.68.160/27, 18.216.170.64/27, 18.216.170.96/27 適用於美國東部 (俄亥俄)
- 34.238.188.128/26, 34.238.188.192/26, 34.238.195.0/26 適用於美國東部 (維吉尼亞北部)
- 13.57.180.0/26 適用於美國西部 (加利佛尼亞北部)
- 34.216.24.32/27, 34.216.24.192/27, 34.216.24.224/27 適用於美國西部 (奧勒岡)
- 18.253.138.192/26 AWS GovCloud (美國東部)
- 52.61.204.192/26 為 AWS GovCloud (美國西部)
- 18.162.221.64/26 適用於亞太區域 (香港)
- 13.232.67.64/26 適用於亞太區域 (孟買)
- 13.209.71.0/26 適用於亞太區域 (首爾)
- 13.229.187.128/26 適用於亞太區域 (新加坡)
- 13.211.12.0/26 適用於亞太區域 (雪梨)
- 13.230.21.0/27, 13.230.21.32/27 適用於亞太區域 (東京)
- 51.16.102.64/26 適用於以色列 (特拉維夫)
- 35.183.92.64/26 適用於加拿大 (中部)
- 40.176.98.128/26 適用於加拿大西部 (卡加利)
- 18.194.95.192/27, 18.194.95.224/27, 18.195.48.0/27 適用於歐洲 (法蘭克福)
- 34.241.197.32/27, 34.241.197.64/27, 34.241.197.96/27 適用於歐洲 (愛爾蘭)
- 18.130.91.0/26 適用於歐洲 (倫敦)
- 35.180.112.0/26 適用於歐洲 (巴黎)
- 13.53.191.0/26 適用於歐洲 (斯德哥爾摩)
- 15.185.91.64/26 適用於中東 (巴林)
- 18.228.1.192/26 適用於南美洲 (聖保羅)
- 15.161.135.192/26 適用於歐洲 (米蘭)

- 13.244.165.128/26 適用於非洲 (開普敦)
- 13.208.217.0/26 適用於亞太區域 (大阪)
- 52.81.151.64/26 適用於中國 (北京)
- 161.189.23.128/26 適用於中國 (寧夏)
- 108.136.221.128/26 適用於亞太區域 (雅加達)
- 3.28.159.64/26 適用於中東 (阿拉伯聯合大公國)
- 51.16.102.64/26 適用於以色列 (特拉維夫)
- 16.62.183.64/26 適用於歐洲 (蘇黎世)
- 18.60.192.192/26 適用於亞太區域 (海德拉巴)
- 16.50.161.192/26 適用於亞太區域 (墨爾本)

## 訪問雪花或 HTTP 端點

當目的地為 HTTP 端點或雪花狀態時，沒有特定於 Amazon 資料 Firehose 的 [AWS IP 位址範圍](#) 子集。

若要將 Amazon Data Firehose 新增至雪花網路政策中的允許清單或公開 HTTP 或 HTTPS 端點，請將所有目前的 [AWS IP 位址範圍](#) 新增至輸入規則。

### Note

通知不一定來自與其相關主題位於相同地 AWS 區的 IP 位址。您必須包含所有區域的 AWS IP 位址範圍。

## 授予 Firehose 進入雪花目的地

當您使用雪花作為目的地時，Firehose 會使用您的雪花帳戶 URL 將資料傳送至雪花帳戶。它也會將錯誤資料備份到您指定的 Amazon 簡單儲存體服務儲存貯體，您也可以選擇使用自己擁有的 AWS Key Management Service 金鑰進行 Amazon S3 伺服器端加密。如果啟用了錯誤記錄，Firehose 會將資料傳送錯誤傳送至您的記 CloudWatch 錄串流。

您在建立交付串流時必須擁有 IAM 角色。Firehose 會擔任 IAM 角色，並取得指定儲存貯體、金鑰和 CloudWatch 記錄群組和串流的存取權。使用下列存取政策來啟用 Firehose 存取您的 S3 儲存貯體。如

果您不擁有 S3 儲存貯體，請新增 `s3:PutObjectAcl` 至 Amazon 簡單儲存服務動作清單，以授予儲存貯體擁有者對 Firehose 交付的物件的完整存取權。此政策也會授予 Firehose 存取權以 CloudWatch 進行錯誤記錄。此政策亦有允許存取 Amazon Kinesis Data Streams 的陳述式。如果您不使用 Amazon Kinesis Data Streams 作為資料來源，可移除該陳述式。Firehose 不使用 IAM 來訪問雪花。對於訪問雪花，它使用您的雪花帳戶網址和 PrivateLink Vpce ID 在私人集群的情況下。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:GetBucketLocation",
        "s3:GetObject",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads",
        "s3:PutObject"
      ],
      "Resource": [
        "arn:aws:s3:::bucket-name",
        "arn:aws:s3:::bucket-name/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey"
      ],
      "Resource": [
        "arn:aws:kms:region:account-id:key/key-id"
      ],
      "Condition": {
        "StringEquals": {
          "kms:ViaService": "s3.region.amazonaws.com"
        },
        "StringLike": {
          "kms:EncryptionContext:aws:s3:arn": "arn:aws:s3:::bucket-name/prefix*"
        }
      }
    }
  ],
}
```

```
{
  "Effect": "Allow",
  "Action": [
    "kinesis:DescribeStream",
    "kinesis:GetShardIterator",
    "kinesis:GetRecords",
    "kinesis:ListShards"
  ],
  "Resource": "arn:aws:kinesis:region:account-id:stream/stream-name"
},
{
  "Effect": "Allow",
  "Action": [
    "logs:PutLogEvents"
  ],
  "Resource": [
    "arn:aws:logs:region:account-id:log-group:log-group-name:log-stream:*"
  ]
}
]
```

如需有關允許其他 AWS 服務存取您的 AWS 資源的詳細資訊，請參閱 IAM 使用者指南中的[建立角色以將許可委派給 AWS 服務](#)。

## 授予 Amazon 資料 Firehose 存取 HTTP 端點目的地

您可以使用 Amazon 數據 Firehose 將數據傳遞到任何 HTTP 端點目的地。Amazon Data Firehose 也會將該資料備份到您指定的 Amazon S3 儲存貯體，您可以選擇性地使用自己擁有的 AWS KMS 金鑰進行 Amazon S3 伺服器端加密。如果啟用了錯誤記錄，Amazon Data Firehose 會將資料傳送錯誤傳送到您的 CloudWatch 日誌串流。您也可以使用 AWS Lambda 資料轉換。

建立 Firehose 串流時，您必須具備 IAM 角色。Amazon Data Firehose 會擔任 IAM 角色，並取得指定儲存貯體、金鑰以及 CloudWatch 日誌群組和串流的存取權。

使用下列存取政策可讓 Amazon 資料 Firehose 存取您為資料備份指定的 S3 儲存貯體。如果您不擁有 S3 儲存貯體，請新增 `s3:PutObjectAcl` 至 Amazon S3 動作清單，以授予儲存貯體擁有者完全存取 Amazon Data Firehose 交付的物件。此政策也授予 Amazon 資料 Firehose 存取權以 CloudWatch 進行錯誤記錄和 AWS Lambda 資料轉換。此政策亦有允許存取 Amazon Kinesis Data Streams 的陳述式。如果您不使用 Amazon Kinesis Data Streams 作為資料來源，可移除該陳述式。

**⚠ Important**

Amazon 數據 Firehose 件不使用 IAM 訪問受支持的第三方服務提供商擁有的 HTTP 端點目的地，包括數據多，啟發，MongoDB，新遺物 LogicMonitor，濺射或相撲邏輯。若要存取受支援的第三方服務提供者所擁有的指定 HTTP 端點目的地，請聯絡該服務供應商以取得 API 金鑰或啟用從 Amazon Data Firehose 傳送至該服務的資料所需的存取金鑰。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:GetBucketLocation",
        "s3:GetObject",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads",
        "s3:PutObject"
      ],
      "Resource": [
        "arn:aws:s3:::bucket-name",
        "arn:aws:s3:::bucket-name/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey"
      ],
      "Resource": [
        "arn:aws:kms:region:account-id:key/key-id"
      ],
      "Condition": {
        "StringEquals": {
          "kms:ViaService": "s3.region.amazonaws.com"
        },
        "StringLike": {
```

```

        "kms:EncryptionContext:aws:s3:arn": "arn:aws:s3::bucket-name/
prefix*"
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kinesis:DescribeStream",
      "kinesis:GetShardIterator",
      "kinesis:GetRecords",
      "kinesis:ListShards"
    ],
    "Resource": "arn:aws:kinesis:region:account-id:stream/stream-name"
  },
  {
    "Effect": "Allow",
    "Action": [
      "logs:PutLogEvents"
    ],
    "Resource": [
      "arn:aws:logs:region:account-id:log-group:log-group-name:log-stream:*"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "lambda:InvokeFunction",
      "lambda:GetFunctionConfiguration"
    ],
    "Resource": [
      "arn:aws:lambda:region:account-id:function:function-name:function-
version"
    ]
  }
]
}

```

如需有關允許其他 AWS 服務存取您的 AWS 資源的詳細資訊，請參閱 IAM 使用者指南中的[建立角色以將許可委派給 AWS 服務](#)。



**⚠ Important**

目前 Amazon 資料 Firehose 不支援將資料傳送到 VPC 中的 HTTP 端點。

## 從 Amazon MSK 跨帳戶交付

如果您的案例是跨帳戶案例，而您要從 Firehose 帳戶建立交付串流 (例如，帳戶 B)，而您的來源是另一個 AWS 帳戶 (帳戶 A) 中的 MSK 叢集，則必須具備下列設定：

帳戶 A：

1. 在 Amazon MSK 主控台中，選擇佈建的叢集，然後選擇屬性。
2. 在網路設定下，選擇編輯並開啟多 VPC 連線。
3. 在安全性設定下，選擇編輯叢集。
  - a. 如果叢集尚未設定政策，請勾選包含 Firehose 服務主體和啟用 Firehose 跨帳戶 S3 交付。AWS Management Console 會自動產生具有適當權限的策略。
  - b. 如果叢集已設定政策，請將下列許可新增至現有政策：

```
{
  "Effect": "Allow",
  "Principal": {
    "AWS": "arn:aws:iam::arn:role/mskaasTestDeliveryRole"
  },
  "Action": [
    "kafka:GetBootstrapBrokers",
    "kafka:DescribeCluster",
    "kafka:DescribeClusterV2",
    "kafka-cluster:Connect"
  ],
  "Resource": "arn:aws:kafka:us-east-1:arn:cluster/D0-NOT-TOUCH-mskaas-
provisioned-privateLink/xxxxxxxx-2f3a-462a-ba09-xxxxxxxx-20" // ARN of the
cluster
},
{
  "Effect": "Allow",
  "Principal": {
    "AWS": "arn:aws:iam::arn:role/mskaasTestDeliveryRole"
  },
  "Action": [
```

```

    "kafka-cluster:DescribeTopic",
    "kafka-cluster:DescribeTopicDynamicConfiguration",
    "kafka-cluster:ReadData"
  ],
  "Resource": "arn:aws:kafka:us-east-1:arn:topic/D0-NOT-TOUCH-mskaas-
provisioned-privateLink/xxxxxxxx-2f3a-462a-ba09-xxxxxxxx-20/*"//topic of the
cluster
},
{
  "Effect": "Allow",
  "Principal": {
    "AWS": "arn:aws:iam::233450236687:role/mskaasTestDeliveryRole"
  },
  "Action": "kafka-cluster:DescribeGroup",
  "Resource": "arn:aws:kafka:us-east-1:arn:group/D0-NOT-TOUCH-mskaas-
provisioned-privateLink/xxxxxxxx-2f3a-462a-ba09-xxxxxxxx-20/*" //topic of
the cluster
},
}

```

4. 在 AWS 主體下，輸入帳戶 B 的主要 ID。
5. 在主題下，指定您希望交付串流從其中擷取資料的 Apache Kafka 主題。建立交付串流後，您便無法更新此主題。
6. 選擇 Save changes (儲存變更)

#### 帳戶 B：

1. 在 Firehose 主控台中，選擇「使用帳戶 B 建立交付串流」。
2. 在來源下，選擇 Amazon Managed Streaming for Apache Kafka。
3. 在來源設定下，針對 Amazon Managed Streaming for Apache Kafka 叢集，在帳戶 A 中輸入 Amazon MSK 叢集的 ARN。
4. 在主題下，指定您希望交付串流從其中擷取資料的 Apache Kafka 主題。建立交付串流後，您便無法更新此主題。
5. 交付串流名稱指定交付串流名稱。

在帳戶 B 中建立交付串流時，您必須具有 IAM 角色 (使用時預設建立 AWS Management Console)，該角色授予交付串流對已設定主題之跨帳戶 Amazon MSK 叢集的「讀取」存取權。

以下是由 AWS Management Console 設定的內容：

```
{
  "Sid": "",
  "Effect": "Allow",
  "Action": [
    "kafka:GetBootstrapBrokers",
    "kafka:DescribeCluster",
    "kafka:DescribeClusterV2",
    "kafka-cluster:Connect"
  ],
  "Resource": "arn:aws:kafka:us-east-1:arn:cluster/D0-NOT-TOUCH-mskaas-provisioned-privateLink/xxxxxxxx-2f3a-462a-ba09-xxxxxxxx-20/*" //topic of the cluster
},
{
  "Sid": "",
  "Effect": "Allow",
  "Action": [
    "kafka-cluster:DescribeTopic",
    "kafka-cluster:DescribeTopicDynamicConfiguration",
    "kafka-cluster:ReadData"
  ],
  "Resource": "arn:aws:kafka:us-east-1:arn:topic/D0-NOT-TOUCH-mskaas-provisioned-privateLink/xxxxxxxx-2f3a-462a-ba09-xxxxxxxx-20/mskaas_test_topic" //topic of the cluster
},
{
  "Sid": "",
  "Effect": "Allow",
  "Action": [
    "kafka-cluster:DescribeGroup"
  ],
  "Resource": "arn:aws:kafka:us-east-1:arn:group/D0-NOT-TOUCH-mskaas-provisioned-privateLink/xxxxxxxx-2f3a-462a-ba09-xxxxxxxx-20/*" //topic of the cluster
},
}
```

接下來，您可以完成設定記錄轉換和記錄格式轉換的選用步驟。如需詳細資訊，請參閱 [記錄轉換和格式轉換](#)。

## 跨帳戶交付至 Amazon S3 目的地

您可以使用 AWS CLI 或亞馬遜數據火管 API，在一個帳戶中使用不同 AWS 帳戶中的 Amazon S3 目的地建立 Firehose 串流。下列程序顯示設定帳戶 A 擁有的 Firehose 串流以將資料傳送至帳戶 B 所擁有的 Amazon S3 儲存貯體的範例。

1. 使用[授予 Firehose 存取 Amazon S3 目的地](#)中所述的步驟，在帳戶 A 下建立 IAM 角色。

### Note

存取政策中指定的 Amazon S3 儲存貯體在此範例中為 B 帳戶所有。請務必在存取政策中新增 `s3:PutObjectAcl` Amazon S3 動作清單，以便授予帳戶 B 對 Amazon Data Firehose 交付之物件的完整存取權。跨帳戶交付需要此許可。Amazon 數據 Firehose 將請求上的 `x-amz-acl` 標題設置為 `bucket-owner-full-control`。

2. 若要允許先前建立的 IAM 角色存取，請在 B 帳戶下建立 S3 儲存貯體政策。以下程式碼為儲存貯體政策的範例。如需詳細資訊，請參閱[使用儲存貯體政策和使用政策](#)。

```
{

  "Version": "2012-10-17",
  "Id": "PolicyID",
  "Statement": [
    {
      "Sid": "StmtID",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::accountA-id:role/iam-role-name"
      },
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:GetBucketLocation",
        "s3:GetObject",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads",
        "s3:PutObject",
        "s3:PutObjectAcl"
      ],
      "Resource": [
        "arn:aws:s3::bucket-name",
        "arn:aws:s3::bucket-name/*"
      ]
    }
  ]
}
```

```

    ]
  }
]
}

```

3. 使用您在步驟 1 中建立的 IAM 角色，在帳戶 A 下建立 Firehose 串流。

## 跨帳戶交付至 OpenSearch 服務目的地

您可以使用 AWS CLI 或 Amazon 資料 Firehose API，在具有不同帳戶的 OpenSearch 服務目的地的一個 AWS 帳戶中建立 Firehose 串流。下列程序顯示如何在帳戶 A 下建立 Firehose 串流，並將其設定為將資料傳送至帳戶 B 所擁有的 OpenSearch 服務目的地的範例。

1. 使用 [the section called “授予 Amazon 資料 Firehose 存取公共 OpenSearch 服務目的地”](#) 中所述的步驟，在 A 帳戶下建立 IAM 角色。
2. 若要允許從您在上一個步驟中建立的 IAM 角色存取，請在帳戶 B 下建立 OpenSearch 服務政策，以下 JSON 為範例。

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::Account-A-ID:role/firehose_delivery_role "
      },
      "Action": "es:ESHttpGet",
      "Resource": [
        "arn:aws:es:us-east-1:Account-B-ID:domain/cross-account-cluster/_all/_settings",
        "arn:aws:es:us-east-1:Account-B-ID:domain/cross-account-cluster/_cluster/stats",
        "arn:aws:es:us-east-1:Account-B-ID:domain/cross-account-cluster/roletest*/_mapping/roletest",
        "arn:aws:es:us-east-1:Account-B-ID:domain/cross-account-cluster/_nodes",
        "arn:aws:es:us-east-1:Account-B-ID:domain/cross-account-cluster/_nodes/stats",
        "arn:aws:es:us-east-1:Account-B-ID:domain/cross-account-cluster/_nodes/*/stats",
        "arn:aws:es:us-east-1:Account-B-ID:domain/cross-account-cluster/_stats",

```

```

        "arn:aws:es:us-east-1:Account-B-ID:domain/cross-account-cluster/roletest*/
_stats",
        "arn:aws:es:us-east-1:Account-B-ID:domain/cross-account-cluster/"
    ]
}
]
}

```

3. 使用您在步驟 1 中建立的 IAM 角色，在帳戶 A 下建立 Firehose 串流。當您建立 Firehose 串流時，請使用 AWS CLI 或 Amazon 資料 Firehose API，並指定 `ClusterEndpoint` 欄位而不是用 `DomainARN` 於 OpenSearch 服務。

### Note

若要在具有不同帳戶的 OpenSearch 服務目標的 AWS 帳戶中建立 Firehose 串流，您必須使用 AWS CLI 或 Amazon 資料 Firehose API。您無法使用建 AWS Management Console 立這種跨帳戶設定。

## 使用標籤控制存取

您可以在 IAM 政策中使用選用 `Condition` 元素 (或 `Condition` 區塊)，根據標籤金鑰和值微調對 Amazon Data Firehose 操作的存取。以下小節說明如何針對不同的 Amazon 資料 Firehose 作業執行此操作。如需使用 `Condition` 元素的詳細資訊，並了解該元素中可利用的運算子，請參閱 [IAM JSON 政策元素：Condition](#)。

### CreateDeliveryStream

若是 `CreateDeliveryStream` 操作，請使用 `aws:RequestTag` 條件金鑰。在下方範例中，`MyKey` 與 `MyValue` 分別表示標籤的金鑰和對應值。如需更多資訊，請參閱 [標籤基本概念](#)

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "firehose:CreateDeliveryStream",
      "Resource": "*",
      "Condition": {
        "StringEquals": {

```

```
        "aws:RequestTag/MyKey": "MyValue"
      }
    }
  ]
}
```

## TagDeliveryStream

若是 TagDeliveryStream 操作，請使用 `aws:TagKeys` 條件金鑰。在下方範例中，MyKey 為標籤金鑰範例。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "firehose:TagDeliveryStream",
      "Resource": "*",
      "Condition": {
        "ForAnyValue:StringEquals": {
          "aws:TagKeys": "MyKey"
        }
      }
    }
  ]
}
```

## UntagDeliveryStream

若是 UntagDeliveryStream 操作，請使用 `aws:TagKeys` 條件金鑰。在下方範例中，MyKey 為標籤金鑰範例。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "firehose:UntagDeliveryStream",
      "Resource": "*",
      "Condition": {
        "ForAnyValue:StringEquals": {
```

```

    "aws:TagKeys": "MyKey"
  }
}
]
}

```

## ListDeliveryStreams

ListDeliveryStreams 無法搭配使用以標籤為基礎的存取控制。

## 其他 Amazon 數據 Firehose 操作

對於、和以外的所有 Amazon 資料 Firehose 作業 CreateDeliveryStream、TagDeliveryStream、UntagDeliveryStream、ListDeliveryStreams，請使用 aws:RequestTag 條件金鑰。在下方範例中，MyKey 與 MyValue 分別表示標籤的金鑰和對應值。

ListDeliveryStreams 中，使用 firehose:ResourceTag 條件鍵根據該 Firehose 串流上的標籤來控制存取。

在下方範例中，MyKey 與 MyValue 分別表示標籤的金鑰和對應值。此原則僅適用於具有名稱為 MyKey 值為 MyValue 標籤的資料 Firehose 串流。如需有關根據資源標籤控制存取的詳細資訊，請參閱 IAM 使用者指南 [中的使用標籤控制 AWS 資源存取](#)。

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Deny",
      "Action": "firehose:DescribeDeliveryStream",
      "Resource": "*",
      "Condition": {
        "StringEquals": {
          "firehose:ResourceTag/MyKey": "MyValue"
        }
      }
    }
  ]
}

```



## 監控 Amazon 數據 Firehose

Amazon 資料 Firehose 為您的 Firehose 串流提供監控功能。如需詳細資訊，請參閱 [監控](#)。

## Amazon 資料 Firehose 的合規驗證

第三方稽核員會評估 Amazon Data Firehose 的安全性和合規性，做為多個 AWS 合規計劃的一部分。這些計劃包括 SOC、PCI、FedRAMP、HIPAA 等等。

如需特定合規方案範圍內的 AWS 服務清單，請參閱 [合規方案範圍內的 AWS 服務](#)。如需一般資訊，請參閱 [AWS 合規計劃](#)。

您可以使用下載第三方稽核報告 AWS Artifact。如需詳細資訊，請參閱 [在 AWS Artifact 中下載報表](#)。

使用 Data Firehose 時的合規責任取決於您資料的敏感性、公司的合規目標以及適用的法律和法規。如果您使用「資料 Firehose」須符合 HIPAA、PCI 或 FedRAMP 等標準，可提供相關資源以協助：  
AWS

- [安全性與合規性快速入門指南](#) — 這些部署指南討論架構考量，並提供在上部署以安全性和法規遵循為重點的基準環境的步驟。 AWS
- [建構 HIPAA 安全性與合規性白皮書 — 本白皮書](#) 說明公司如何使用建立符合 HIPAA 標準的應用 AWS 程式。
- [AWS 合規資源](#) — 此工作簿和指南集合可能適用於您的產業和所在地。
- [AWS Config](#) — 此 AWS 服務評估您的資源配置是否符合內部實踐，行業準則和法規。
- [AWS Security Hub](#) — 此 AWS 服務提供安全狀態的全面檢視，協助您檢查您 AWS 是否符合安全性產業標準和最佳做法。

## Amazon 數據 Firehose 中的彈性

AWS 全球基礎架構是圍繞區 AWS 域和可用區域建立的。AWS 區域提供多個實體分離和隔離的可用區域，這些區域透過低延遲、高輸送量和高度備援的網路連線。透過可用區域，您所設計與操作的應用程式和資料庫，就能夠在可用區域之間自動容錯移轉，而不會發生中斷。可用區域的可用性、容錯能力和擴充能力，均較單一或多個資料中心的傳統基礎設施還高。

如需區域和可用區域的相關 AWS 資訊，請參閱 [AWS 全域基礎結構](#)。

除了 AWS 全球基礎架構之外，Data Firehose 還提供多種功能，以協助支援您的資料恢復能力和備份需求。

## 災難復原

Amazon Data Firehose 以無伺服器模式執行，並透過執行自動遷移來處理主機降級、可用區域可用性和其他基礎設施相關問題。發生這種情況時，Amazon 資料防火軟管可確保 Firehose 串流能夠移轉而不會遺失任何資料。

## Amazon 數據 Firehose 中的基礎設施

作為一項受管服務，Amazon 資料 Firehose 受到 AWS 全球網路安全的保護。有關 AWS 安全服務以及如何 AWS 保護基礎結構的詳細資訊，請參閱[AWS 雲端安全](#)。若要使用基礎架構安全性的最佳做法來設計您的 AWS 環境，請參閱[安全性支柱架構良 AWS 好的架構中的基礎結構保護](#)。

您可以使用 AWS 已發佈的 API 呼叫透過網路存取 Firehose。使用者端必須支援下列專案：

- Transport Layer Security (TLS)。我們需要 TLS 1.2 並建議使用 TLS 1.3。
- 具備完美轉送私密(PFS)的密碼套件，例如 DHE (Ephemeral Diffie-Hellman)或 ECDHE (Elliptic Curve Ephemeral Diffie-Hellman)。現代系統(如 Java 7 和更新版本)大多會支援這些模式。

此外，請求必須使用存取金鑰 ID 和與 IAM 主體相關聯的私密存取金鑰來簽署。或者，您可以透過[AWS Security Token Service](#) (AWS STS) 來產生暫時安全憑證來簽署請求。

### Note

對於傳出的 HTTPS 請求，Amazon 資料 Firehose 會使用 HTTP 程式庫，該程式庫會自動選取目標端支援的最高 TLS 通訊協定版本。

## VPC 端點 (PrivateLink)

Amazon 資料 Firehose 為 VPC 端點提供支援 () PrivateLink。如需詳細資訊，請參閱[使用 Amazon 數據 Firehose AWS PrivateLink](#)。

## Amazon 數據 Firehose 的安全最佳實踐

Amazon Data Firehose 在您開發和實作自己的安全政策時，提供許多安全功能供您考慮。以下最佳實務為一般準則，並不代表完整的安全解決方案。這些最佳實務可能不適用或無法滿足您的環境需求，因此請將其視為實用建議就好，而不要當作是指示。

## 實作最低權限存取

授予許可時，您可以決定誰取得哪些 Amazon 資料 Firehose 資源的許可。您還需針對這些資源啟用允許執行的動作，因此，您只應授與執行任務所需的許可。對降低錯誤或惡意意圖所引起的安全風險和影響而言，實作最低權限存取是相當重要的一環。

## 使用 IAM 角色

生產者和用戶端應用程式必須具有有效的登入資料才能存取 Amazon Data Firehose 交付串流，而您的 Firehose 串流必須具有有效的登入資料才能存取目的地。您不應將 AWS 登入資料直接存放在用戶端應用程式或 Amazon S3 儲存貯體中。這些是不會自動輪換的長期憑證，如果遭到盜用，可能會對業務造成嚴重的影響。

相反地，您應該使用 IAM 角色來管理生產者和用戶端應用程式的臨時登入資料，以存取 Firehose 串流。使用角色時，您不必使用長期登入資料 (例如使用者名稱和密碼或存取金鑰) 來存取其他資源。

如需詳細資訊，請參閱《IAM 使用者指南》中的以下主題：

- [IAM 角色](#)
- [常見的角色方案：使用者、應用程式和服務](#)

## 在相依資源實作伺服器端加密

可以在 Amazon 資料 Firehose 體中加密靜態資料和傳輸中的資料。如需詳細資訊，請參閱 [Amazon 資料 Firehose 中的資料保護](#)。

## 用 CloudTrail 於監控 API 呼叫

Amazon 數據 Firehose 與該服務集成在一起 AWS CloudTrail，該服務可提供 Amazon 數據 Firehose 中的用戶，角色或 AWS 服務採取的操作記錄。

使用收集的資訊 CloudTrail，您可以判斷向 Amazon Data Firehose 提出的請求、提出請求的來源 IP 地址、提出請求的人員、提出請求的時間以及其他詳細資訊。

如需詳細資訊，請參閱 [the section called “使用日誌記錄 Amazon 數 Firehose API 調用 AWS CloudTrail”](#)。

# Amazon 數據 Firehose 數據轉換

Amazon 資料 Firehose 可以叫用您的 Lambda 函數來轉換傳入的來源資料，並將轉換後的資料交付到目的地。您可以在建立 Firehose 串流時啟用 Amazon 資料 Firehose 資料轉換。

## 資料轉換流程

當您啟用 Firehose 資料轉換時，Firehose 會緩衝傳入的資料。緩衝大小提示的範圍介於 0.2 MB 和 3MB 之間。除了 Splunk 和雪花之外，所有目的地的預設 Lambda 緩衝大小提示都是 1 MB。如果是「潑濺」和「雪花」，預設的緩衝提示是 256 KB。Lambda 緩衝間隔提示的範圍介於 0 到 900 秒之間。除了雪花以外的所有目的地，預設 Lambda 緩衝間隔提示為 60 秒。對於雪花，預設的緩衝提示間隔為 30 秒。若要調整緩衝大小，請使用呼 [ProcessorParameter](#) 叫的 `BufferSizeInMBs` 和 `IntervalInSeconds` 來設定 [CreateDeliveryStream](#) 或 [UpdateDestination](#) API 的 [ProcessingConfiguration](#) 參數。然後，Firehose 會使用同步叫用模式，以非同步方式呼叫指定的 Lambda 函數，以每個緩衝的批次方式來叫用 AWS Lambda。轉換後的資料會從 Lambda 傳送至 Firehose。然後，Firehose 會在達到指定的目的地緩衝大小或緩衝間隔時（以先發生者為準）將其發送到目的地。

### Important

在請求和回應方面，Lambda 同步調用模式的承載大小上限為 6 MB。因此，用來傳送請求至函式的緩衝區大小必須小於或等於 6 MB，且函式所傳回的回應也不得超過 6 MB。

## 資料轉換和狀態模型

來自 Lambda 的所有轉換記錄都必須包含下列參數，否則 Amazon 資料 Firehose 會拒絕這些參數，並將其視為資料轉換失敗。

對於 Kinesis Data Streams 和 Direct PUT：

`recordId`

記錄 ID 會在叫用期間從 Amazon 資料 Firehose 傳送至 Lambda。轉換記錄必須包含相同的記錄 ID。原始記錄的 ID 與轉換記錄的 ID 若有任何不符，就會視為資料轉換失敗。

## result

記錄的資料轉換狀態。可能的值包括：Ok (記錄轉換成功)、Dropped (處理邏輯刻意捨棄記錄)，以及 ProcessingFailed (記錄無法轉換)。如果記錄的狀態為Ok或Dropped，Amazon 資料 Firehose 會將其視為已成功處理。否則，Amazon 數據 Firehose 認為未成功處理。

## 資料

已轉換資料承載 (base64 編碼後)。

以下是一個示例 Lambda 結果輸出：

```
{
  "recordId": "<recordId from the Lambda input>",
  "result": "Ok",
  "data": "<Base64 encoded Transformed data>"
}
```

## 對於 Amazon MSK

### recordId

在呼叫期間，記錄識別碼會從 Firehose 傳送至 Lambda。轉換記錄必須包含相同的記錄 ID。原始記錄的 ID 與轉換記錄的 ID 若有任何不符，就會視為資料轉換失敗。

### result

記錄的資料轉換狀態。可能的值包括：Ok (記錄轉換成功)、Dropped (處理邏輯刻意捨棄記錄)，以及 ProcessingFailed (記錄無法轉換)。如果記錄的狀態為Ok或Dropped，Firehose 會認為該記錄已成功處理。否則，Firehose 會認為未成功處理。

### KafkaRecordValue

已轉換資料承載 (base64 編碼後)。

以下是一個示例 Lambda 結果輸出：

```
{
  "recordId": "<recordId from the Lambda input>",
  "result": "Ok",
  "kafkaRecordValue": "<Base64 encoded Transformed data>"
}
```

## Lambda 藍圖

這些藍圖示範如何建立和使用 AWS Lambda 函數來轉換 Amazon 資料 Firehose 資料串流中的資料。

若要查看主控台中可用的 AWS Lambda 藍圖

1. 請登入 AWS Management Console 並開啟 AWS Lambda 主控台，網址為 <https://console.aws.amazon.com/lambda/>。
2. 選擇 Create function (建立函式)，接著選擇 Use a blueprint (使用藍圖)。
3. 在「藍圖」欄位中，搜尋關鍵字 firehose 以尋找 Amazon 資料 Firehose Lambda 藍圖。

藍圖清單：

- 傳送至 Amazon 資料 Firehose 串流的處理記錄 (Node.js, Python)

此藍圖顯示了如何使用 AWS Lambda 處理 Firehose 資料串流中資料的基本範例。

最新發佈日期：2016 年 11 月。

發行說明：無。

- 傳送至 Firehose 的處理 CloudWatch 日誌

此藍圖已取代。如需處理傳送至 Firehose 之 CloudWatch 記錄的相關資訊，請參閱 [使用 CloudWatch 記錄檔寫入 Firehose](#)。

- 將系統日誌格式的 Amazon 數據 Firehose 流記錄轉換為 JSON ( Node.js )

此藍圖顯示如何將 RFC3164 Syslog 格式的輸入記錄轉換為 JSON。

最新發佈日期：2016 年 11 月。

發行說明：無。

若要查看中可用的藍圖 AWS Serverless Application Repository

1. 前往 [AWS Serverless Application Repository](#)。
2. 選擇瀏覽所有應用程式。
3. 在 Applications (應用程式) 欄位中，搜尋關鍵字 firehose。

您也可以在不使用藍圖的情況下建立 Lambda 函數。請參閱[開始使用 AWS Lambda](#)。

## 資料轉換失敗的處理方式

如果您的 Lambda 函數叫用因為網路逾時或您已達到 Lambda 叫用限制而失敗，Amazon 資料 Firehose 預設會重試呼叫三次。如果呼叫未成功，Amazon 資料 Firehose 就會略過該批次的記錄。略過的記錄會視為未處理成功。您可以使用或 [UpdateDestination](#) API 指定 [CreateDeliveryStream](#) 或覆寫重試選項。對於這種類型的故障，您可以將調用錯誤記錄到 Amazon CloudWatch 日誌。如需詳細資訊，請參閱 [使 CloudWatch 用日誌監控 Amazon 數據 Firehose](#)。

如果記錄的資料轉換狀態為 `ProcessingFailed`，Amazon 資料 Firehose 會將記錄視為未成功處理記錄。對於這種類型的故障，您可以從 Lambda 函數將錯誤 CloudWatch 日誌發送到 Amazon 日誌。如需詳細資訊，請參閱 [AWS Lambda 開發人員指南 AWS Lambda 中的存取 Amazon CloudWatch 日誌](#)。

如果資料轉換失敗，未成功處理的記錄會傳送到 `processing-failed` 資料夾中的 S3 儲存貯體。記錄的格式如下：

```
{
  "attemptsMade": "count",
  "arrivalTimestamp": "timestamp",
  "errorCode": "code",
  "errorMessage": "message",
  "attemptEndingTimestamp": "timestamp",
  "rawData": "data",
  "lambdaArn": "arn"
}
```

### `attemptsMade`

嘗試叫用請求的次數。

### `arrivalTimestamp`

Amazon 數據 Firehose 件收到該記錄的時間。

### `errorCode`

Lambda 傳回 HTTP 錯誤代碼。

### `errorMessage`

Lambda 傳回錯誤訊息。



## attemptEndingTimestamp

Amazon 數據 Firehose 件停止嘗試 Lambda 調用的時間。

## rawData

Base64 編碼記錄資料。

## lambdaArn

Lambda 函數的 Amazon Resource Name (ARN)。

## Lambda 調用的持續時間

Amazon 資料 Firehose 支援最多 5 分鐘的 Lambda 叫用時間。如果您的 Lambda 函數需要 5 分鐘以上才能完成，您會收到下列錯誤：Firehose 在呼叫 AWS Lambda 時遇到逾時錯誤。支援的函數逾時上限為 5 分鐘。

如需 Amazon 資料 Firehose 在發生此類錯誤時會執行哪些動作的相關資訊，請參閱[the section called “資料轉換失敗的處理方式”](#)。

## 來源記錄備份

Amazon Data Firehose 可以將所有未轉換的記錄同時備份到 S3 儲存貯體，同時將轉換後的記錄交付到目的地。您可以在建立或更新 Firehose 串流時啟用來源記錄備份。啟用來源記錄備份後，便不能停用。



# Amazon 數據 Firehose 中的動態分區

動態磁碟分割可讓您使用資料中的金鑰 (例如, `customer_id` 或 `transaction_id`) 持續分割 Firehose 中的串流資料, 然後將這些金鑰分組的資料交付到對應的 Amazon Simple Storage Service (Amazon S3) 前置詞。這可讓您使用各種服務 (例如 Amazon 雅典娜、Amazon EMR、亞馬遜 Amazon Redshift Spectrum 和亞馬遜) 對 Amazon S3 中的串流資料執行高效能且具成本效益的分析。QuickSight 此外, 在需要額外處理的使用案例中, AWS Glue 可以在動態分區的串流資料交付到 Amazon S3 之後, 執行更複雜的擷取、轉換和載入 (ETL) 任務。

分割資料可最大限度地減少掃描的資料量、最佳化性能, 並降低 Amazon S3 上的分析查詢成本。它還可以增加對資料的精細存取。傳統上使用 Firehose 串流來擷取資料並將其載入 Amazon S3。若要分割 Amazon S3 分析的串流資料集, 您需要在 Amazon S3 儲存貯體之間執行分割應用程式, 然後才能將資料提供給分析, 這可能會變得複雜或昂貴。

透過動態分割, Firehose 會使用動態或靜態定義的資料金鑰持續分組傳輸中的資料, 並依金鑰將資料交付至個別的 Amazon S3 前置詞。這減少了幾 time-to-insight 分鐘或幾小時。它還可以降低成本並簡化架構。

## 主題

- [分割索引鍵](#)
- [用於動態分割的 Amazon S3 儲存貯體字首](#)
- [彙整資料的動態分割](#)
- [在將資料傳送至 S3 時新增新的行分隔符號](#)
- [如何啟用動態分割](#)
- [動態分割錯誤處理](#)
- [資料緩衝和動態分割](#)

## 分割索引鍵

透過動態分割, 您可以根據分割索引鍵分割資料, 從串流 S3 資料建立目標資料集。分割索引鍵可讓您根據特定值篩選串流資料。例如, 如果您需要根據客戶 ID 和國家/地區篩選資料, 可以將 `customer_id` 的資料欄位指定為一個分割索引鍵, 而將 `country` 的資料欄位指定為另一個分割索引鍵。然後, 您可以指定運算式 (使用支援的格式) 來定義要傳送動態分割資料記錄的 S3 儲存貯體字首。

以下是建立分割索引鍵的支援方法:

- 內聯解析-此方法使用 Firehose 內置的支持機制，一個 [jq 解析器](#)，用於從 JSON 格式的數據記錄中提取用於分區的密鑰。目前，我們僅支援jq 1.6版本。
- AWS Lambda 函數-此方法使用指定的 AWS Lambda 函數來提取和返回分區所需的數據字段。

### ⚠ Important

啟用動態分割時，您必須至少設定下列一種方法，以分割您的資料。您可以設定其中一種方法來同時指定您的分割索引鍵，或同時指定兩個分割索引鍵。

## 使用內嵌剖析建立分割索引鍵

若要將內嵌剖析設定為串流資料的動態分割方法，您必須選擇要用作分割索引鍵的資料記錄參數，並為每個指定的分割索引鍵提供一個值。

讓我們來看看下面的示例數據記錄，看看如何使用內聯解析為它定義分區鍵：

```
{
  "type": {
    "device": "mobile",
    "event": "user_clicked_submit_button"
  },
  "customer_id": "1234567890",
  "event_timestamp": 1565382027,   #epoch timestamp
  "region": "sample_region"
}
```

例如，您可以選擇根據 `customer_id` 參數或 `event_timestamp` 參數對資料進行分割。這表示您希望在決定要交付記錄的 S3 字首時，使用每個記錄中的 `customer_id` 參數或 `event_timestamp` 參數值。您也可以選擇巢狀參數，就像使用運算式 `.type.device` 的 `device` 一樣。您的動態分割邏輯可以依賴於多個參數。

選取分割索引鍵的資料參數之後，您可以將每個參數映射至有效的 jq 運算式。下表顯示了參數到 jq 運算式的映射：

參數	jq 運算式
<code>customer_id</code>	<code>.customer_id</code>

參數	jq 運算式
device	.type.device
year	.event_timestamp  strftime("%Y")
month	.event_timestamp  strftime("%m")
day	.event_timestamp  strftime("%d")
hour	.event_timestamp  strftime("%H")

在執行階段，Firehose 會根據每筆記錄中的資料，使用上方的右欄來評估參數。

## 使用 AWS Lambda 函數建立分割索引鍵

對於壓縮或加密的資料記錄，或 JSON 以外任何檔案格式的資料，您可以使用整合的 AWS Lambda 函數搭配您自己的自訂程式碼來解壓縮、解密或轉換記錄，以擷取和傳回分割所需的資料欄位。這是現有轉換 Lambda 函數的擴充功能，現在可透過 Firehose 使用。您可以轉換、剖析和傳回資料欄位，然後使用相同的 Lambda 函數來進行動態分割。

以下是 Python 中的 Firehose 流處理 Lambda 函數的示例，該函數會重播從輸入到輸出的每個讀取記錄，並從記錄中提取分區鍵。

```
from __future__ import print_function
import base64
import json
import datetime

# Signature for all Lambda functions that user must implement
def lambda_handler(firehose_records_input, context):
    print("Received records for processing from DeliveryStream: " +
          firehose_records_input['deliveryStreamArn']
          + ", Region: " + firehose_records_input['region']
          + ", and InvocationId: " + firehose_records_input['invocationId'])

    # Create return value.
    firehose_records_output = {'records': []}

    # Create result object.
```

```
# Go through records and process them

for firehose_record_input in firehose_records_input['records']:
    # Get user payload
    payload = base64.b64decode(firehose_record_input['data'])
    json_value = json.loads(payload)

    print("Record that was received")
    print(json_value)
    print("\n")
    # Create output Firehose record and add modified payload and record ID to it.
    firehose_record_output = {}
    event_timestamp = datetime.datetime.fromtimestamp(json_value['eventTimestamp'])
    partition_keys = {"customerId": json_value['customerId'],
                      "year": event_timestamp.strftime('%Y'),
                      "month": event_timestamp.strftime('%m'),
                      "date": event_timestamp.strftime('%d'),
                      "hour": event_timestamp.strftime('%H'),
                      "minute": event_timestamp.strftime('%M')}

    # Create output Firehose record and add modified payload and record ID to it.
    firehose_record_output = {'recordId': firehose_record_input['recordId'],
                              'data': firehose_record_input['data'],
                              'result': 'Ok',
                              'metadata': { 'partitionKeys': partition_keys }}

    # Must set proper record ID
    # Add the record to the list of output records.

    firehose_records_output['records'].append(firehose_record_output)

# At the end return processed records
return firehose_records_output
```

以下是 Go 中的 Firehose 串流處理 Lambda 函式的範例，它會重播從輸入到輸出的每個讀取記錄，並從記錄擷取分割金鑰。

```
package main

import (
    "fmt"
```

```
"encoding/json"
"time"
"strconv"

"github.com/aws/aws-lambda-go/events"
"github.com/aws/aws-lambda-go/lambda"
)

type DataFirehoseEventData struct {
    CustomerId string `json:"customerId"`
}

func handleRequest(evnt events.DataFirehoseEvent) (events.DataFirehoseResponse, error)
{

    fmt.Printf("InvocationID: %s\n", evnt.InvocationID)
    fmt.Printf("DeliveryStreamArn: %s\n", evnt.DeliveryStreamArn)
    fmt.Printf("Region: %s\n", evnt.Region)

    var response events.DataFirehoseResponse

    for _, record := range evnt.Records {
        fmt.Printf("RecordID: %s\n", record.RecordID)
        fmt.Printf("ApproximateArrivalTimestamp: %s\n", record.ApproximateArrivalTimestamp)

        var transformedRecord events.DataFirehoseResponseRecord
        transformedRecord.RecordID = record.RecordID
        transformedRecord.Result = events.DataFirehoseTransformedStateOk
        transformedRecord.Data = record.Data

        var metaData events.DataFirehoseResponseRecordMetadata
        var recordData DataFirehoseEventData
        partitionKeys := make(map[string]string)

        currentTime := time.Now()
        json.Unmarshal(record.Data, &recordData)
        partitionKeys["customerId"] = recordData.CustomerId
        partitionKeys["year"] = strconv.Itoa(currentTime.Year())
        partitionKeys["month"] = strconv.Itoa(int(currentTime.Month()))
        partitionKeys["date"] = strconv.Itoa(currentTime.Day())
        partitionKeys["hour"] = strconv.Itoa(currentTime.Hour())
        partitionKeys["minute"] = strconv.Itoa(currentTime.Minute())
        metaData.PartitionKeys = partitionKeys
        transformedRecord.Metadata = metaData
    }
}
```

```
    response.Records = append(response.Records, transformedRecord)
}

return response, nil
}

func main() {
    lambda.Start(handleRequest)
}
```

## 用於動態分割的 Amazon S3 儲存貯體字首

當您建立使用 Amazon S3 做為目的地的 Firehose 串流時，您必須指定一個 Amazon S3 儲存貯體，其中 Firehose 要交付您的資料。Amazon S3 儲存貯體字首用於整理儲存在 Amazon S3 儲存貯體中的資料。Amazon S3 儲存貯體字首類似於目錄，可讓您在儲存貯體中對類似物件進行分組。

透過動態分割，您的分割資料會交付到指定的 Amazon S3 字首。如果您未啟用動態磁碟分割，則為 Firehose 串流指定 S3 儲存貯體前綴是可選的。但是，如果您選擇啟用動態分割，則必須指定 Firehose 向其提供分區資料的 S3 儲存貯體前綴。

在您啟用動態磁碟分割的每個 Firehose 串流中，S3 儲存貯體前綴值是由以該交付串流指定的分割金鑰為基礎的運算式所組成。再次使用上述資料記錄範例，您可以建立下列 S3 字首值，該值由以上定義的分割索引鍵為基礎的運算式組成：

```
"ExtendedS3DestinationConfiguration": {
  "BucketARN": "arn:aws:s3:::my-logs-prod",
  "Prefix": "customer_id={!{partitionKeyFromQuery:customer_id}}/
    device={!{partitionKeyFromQuery:device}}/
    year={!{partitionKeyFromQuery:year}}/
    month={!{partitionKeyFromQuery:month}}/
    day={!{partitionKeyFromQuery:day}}/
    hour={!{partitionKeyFromQuery:hour}}/"
}
```

Firehose 在運行時評估上述表達式。它會將符合相同評估後 S3 字首運算式的記錄分組到單一資料集中。Firehose 接著會將每個資料集傳送至評估後的 S3 前置詞。資料集傳遞至 S3 的頻率取決於 Firehose 串流緩衝區設定。因此，本範例中的記錄會交付至下列 S3 物件金鑰：

```
s3://my-logs-prod/customer_id=1234567890/device=mobile/year=2019/month=08/day=09/hour=20/my-delivery-stream-2019-08-09-23-55-09-a9fa96af-e4e4-409f-bac3-1f804714faaa
```

對於動態分割，您必須在 S3 儲存貯體字首中使用下列運算式格式：`!{namespace:value}`，其中命名空間可以是 `partitionKeyFromQuery` 和/或 `partitionKeyFromLambda`。如果您使用內嵌剖析來建立來源資料的分割索引鍵，則必須指定由以下格式指定之運算式所組成的 S3 儲存貯體字首值：`"partitionKeyFromQuery:keyID"`。如果您使用 AWS Lambda 函數為來源資料建立分割索引鍵，則必須指定由以下格式指定之運算式所組成的 S3 儲存貯體字首值：`"partitionKeyFromLambda:keyID"`。

### Note

您也可以使用配置單元樣式格式指定 S3 存儲桶前綴值，例如客戶 `_id=!{partitionKeyFrom查詢:客戶ID}`。

如需詳細資訊，請參閱[建立 Amazon Firehose 串流](#)和 Amazon S3 物件的[自訂首碼中的「為您的目的地選擇 Amazon S3」](#)。

## 彙整資料的動態分割

您可以將動態分割套用至彙整資料 (例如，將多個事件、日誌或彙整成單一 `PutRecord` 和 `PutRecordBatch` API 呼叫的記錄)，但必須先將此資料取消彙整。您可以通過啟用多記錄解彙總來分解數據-通過 Firehose 流中的記錄進行解析並將其分離的過程。多記錄取消彙整可以是 JSON 類型，這意味著記錄的分隔基於有效的 JSON 來執行。或者它可以是 Delimited 類型，這意味著根據指定的自定義分隔符執行記錄的分隔。此自訂分隔符號必須是 base-64 編碼字串。例如，如果您想要使用下列字串做為自訂分隔符號####，您必須以 base-64 編碼格式指定它，並將其轉換為 `IyMjIw==`

使用彙總資料時，當您啟用動態磁碟分割時，Firehose 會剖析記錄，並根據指定的多重記錄解彙總類型，在每個 API 呼叫中尋找有效的 JSON 物件或分隔記錄。

### Important

如果您的資料已彙整，則只有在您的資料第一次取消彙整時，才能套用動態分割。

**⚠ Important**

當您在 Firehose 中使用資料轉換功能時，會在資料轉換之前套用解彙總。進入 Firehose 的資料將依下列順序處理：解彙總 → 透過 Lambda 進行資料轉換 → 分割金鑰。

## 在將資料傳送至 S3 時新增新的行分隔符號

您可以啟用「新行分隔符號」，在交付到 Amazon S3 的物件中的記錄之間新增行分隔符號。這對於在 Amazon S3 中剖析物件很有幫助。當動態磁碟分割套用至彙總資料時，這也特別有用，因為多重記錄解彙總 (必須先套用至彙總資料，才能動態分割資料) 會在剖析程序中從記錄中移除新行。

## 如何啟用動態分割

您可以透過 Amazon 資料防火管管理主控台、CLI 或 API 為您的 Firehose 串流設定動態分割。

**⚠ Important**

只有在建立新 Firehose 串流時，才能啟用動態磁碟分割。您無法為尚未啟用動態磁碟分割的現有 Firehose 串流啟用動態磁碟分割。

如需如何在建立新 Firehose 串流時透過 Firehose 管理主控台啟用和設定動態磁碟分割的詳細步驟，請參閱[建立 Amazon Firehose 串流](#)。當您完成為 Firehose 串流指定目的地的任務時，請務必遵循「為您的目的地選擇 Amazon S3」一節中的步驟，因為目前只有使用 Amazon S3 做為目的地的 Firehose 串流才支援動態分割。

啟用作用中 Firehose 串流上的動態分割後，您可以透過新增或移除或更新現有的分割金鑰和 S3 前置碼運算式來更新組態。一旦更新，Firehose 就會開始使用新的金鑰和新的 S3 前置碼運算式。

**⚠ Important**

一旦您在 Firehose 串流上啟用動態磁碟分割，就無法在此 Firehose 串流中停用該磁碟分割。



## 動態分割錯誤處理

如果 Amazon Data Firehose 無法剖析 Firehose 串流中的資料記錄，或無法擷取指定的分區金鑰，或評估 S3 前置字元值中包含的運算式，則這些資料記錄會傳送至您在其中啟用動態分割的 Firehose 串流時必須指定的 S3 錯誤儲存貯體前綴。S3 錯誤儲存貯體前綴包含 Firehose 無法傳遞至指定 S3 目的地的所有記錄。這些記錄根據錯誤類型進行組織。除了記錄之外，傳送的物件還包括錯誤的相關資訊，以協助了解並解決錯誤。

如果您要為此 Firehose 串流啟用動態磁碟分割，則必須為 Firehose 串流指定 S3 錯誤儲存貯體前置詞。如果您不想為 Firehose 串流啟用動態分割，則可選擇指定 S3 錯誤儲存貯體前置詞。

## 資料緩衝和動態分割

Amazon Data Firehose 會將傳入的串流資料緩衝為特定大小並保持一段特定時間，然後再將其傳送到指定的目的地。您可以在建立新 Firehose 串流時設定緩衝區大小和緩衝區間隔，或更新現有 Firehose 串流的緩衝區大小和緩衝區間隔。緩衝區大小以 MB 為測量單位，緩衝區間隔以秒為測量單位。

啟用動態磁碟分割後，Firehose 會根據設定的緩衝提示 (大小和時間)，在內部緩衝屬於指定分割區的記錄，然後再將這些記錄傳送到 Amazon S3 儲存貯體。為了提供最大尺寸的對象，Firehose 在內部使用多級緩衝。因此，一批記錄的 end-to-end 延遲可能是配置緩衝提示時間的 1.5 倍。這會影響 Firehose 串流的資料新鮮度。

作用中分割區計數是交付緩衝區內的作用中分割區總數。例如，如果動態分割查詢每秒建構 3 個分割區，而且您的緩衝區提示組態會每 60 秒觸發交付，則平均而言，您就會有 180 個作用中分割區。如果 Firehose 無法將分割區中的資料傳送到目的地，則此分割區會在傳遞緩衝區中計為作用中，直到可以傳送為止。

根據記錄資料欄位和 S3 字首運算式將 S3 字首評估為新值時，會建立新的分割。會為每個作用中的分割建立新的緩衝區。具有相同評估 S3 字首的每個後續記錄都會交付到該緩衝區。一旦緩衝區達到緩衝區大小限制或緩衝時間間隔，Firehose 就會建立包含緩衝區資料的物件，並將其傳送到指定的 Amazon S3 前綴。一旦物件交付，該分割區和分割本身的緩衝區就會被刪除，並從作用中的分割計數中移除。一旦分別滿足每個分區的緩衝區大小或間隔，Firehose 將每個緩衝區數據作為單個對象提供。一旦作用中分割區數量達到每個交付串流 500 個上限，Firehose 串流中的其餘記錄就會傳送到指定的 S3 錯誤儲存貯體前綴。

# 在 Firehose 中轉換您的輸入記錄格式

Amazon 數據 Firehose 可以在將數據存儲在 Amazon S3 之前，將輸入數據的格式從 JSON 轉換為 [Apache 實木複合地板或阿帕奇 ORC](#)。Parquet 和 ORC 為單欄資料格式，相較於橫列導向格式 (如 JSON) 更可節省空間，並加快查詢速度。若您希望轉換的輸入格式並非 JSON，而是逗號分隔值 (CSV) 或結構化文字，則可先使用 AWS Lambda 將其轉換為 JSON。如需詳細資訊，請參閱 [資料轉換](#)。

## 主題

- [記錄格式轉換要求](#)
- [選擇 JSON 還原序列化程式](#)
- [選擇序列化程式](#)
- [轉換輸入記錄格式 \(主控台\)](#)
- [轉換輸入記錄格式 \(API\)](#)
- [記錄格式轉換錯誤處理](#)
- [記錄格式轉換範例](#)

## 記錄格式轉換要求

Amazon 資料 Firehose 需要以下三個元素來轉換記錄資料的格式：

- 用於讀取輸入資料 JSON 的反序列化程式 — 您可以選擇以下兩種反序列化程式之一：[阿帕奇蜂巢 JSON 或 OpenX JSON](#)。 [SerDe SerDe](#)

### Note

將多個 JSON 文件合併到同一筆記錄時，請確保您的輸入仍以支援的 JSON 格式顯示。JSON 文件的陣列不是有效的輸入。

例如，正確的輸入是：`{"a":1}{ "a":2}`

不正確的輸入是：`[{"a":1}, {"a":2}]`

- 決定如何解譯該資料的結構描述 – 使用 [AWS Glue](#) 在 AWS Glue Data Catalog 中建立結構描述。Amazon 資料 Firehose 接著會參考該結構描述，並使用該結構描述來解譯您的輸入資料。您可以使用相同的結構描述來設定 Amazon 資料 Firehose 和分析軟體。如需詳細資訊，請參閱《AWS Glue 開發人員指南》中的 [填入 AWS Glue Data Catalog](#)。

**Note**

在 AWS Glue Data Catalog 中建立的結構描述應該符合輸入資料結構。否則，已轉換資料將不會包含結構描述中未指定的屬性。如果您使用巢狀 JSON，則請在鏡像 JSON 資料結構的結構描述中使用 STRUCT 類型。有關如何使用 STRUCT 類型處理巢狀 JSON，請參閱[此範例](#)。

- [將數據轉換為目標單欄存儲格式 \( 實木複合地板或 ORC \) 的序列化程序-您可以選擇兩種類型的序列化器之一：ORC 或鑲木地板。 SerDe SerDe](#)

**Important**

如果啟用記錄格式轉換，則無法將 Amazon 數據 Firehose 目的地設置為 Amazon OpenSearch 服務，亞馬遜紅移或 Splunk。啟用格式轉換後，Amazon S3 是您唯一可用於 Firehose 串流的目的地。

即使您在將記錄傳送至 Amazon Data Firehose 之前彙總了記錄，也可以轉換資料的格式。

## 選擇 JSON 還原序列化程式

Serde 如果您的輸入 [JSON 包含下列格式的時間戳記](#)，請選擇 [OpenX JSON](#)：

- yyyy-MM-dd'T'HH:mm:ss[.S]'Z'，其中小數點最多可有 9 位數 – 例如：2017-02-07T15:13:01.39256Z。
- yyyy-[M]M-[d]d HH:mm:ss[.S]，其中小數點最多可有 9 位數 – 例如：2017-02-07 15:13:01.14。
- Epoch 秒 – 例如：1518033528。
- Epoch 毫秒 – 例如：1518033528123。
- 浮點 epoch 秒 – 例如：1518033528.123。

該 OpenX JSON SerDe 可以將句點 ( . ) 轉換為下劃線 ( \_ )。也可將 JSON 金鑰轉換為小寫，再加以還原序列。[如需有關此還原序列化器可透過 Amazon 資料 Firehose 使用的選項的詳細資訊](#)，請參閱 [OpenX。JsonSerDe](#)

如果您不確定要選擇哪個解串器，請使用 OpenX JSON SerDe，除非您有不支持的時間戳記。

如果您有先前列出的格式以外的時間戳記，請使用 [Apache Hive JSON SerDe](#)。選擇此還原序列化程式後，您可指定欲使用的時間戳記格式，方法是遵循 Joda-Time DateTimeFormat 格式字串的模式語法。如需詳細資訊，請參閱 [類別 DateTimeFormat](#)。

您亦可使用特殊值 `millis` 來剖析時間戳記 (epoch 毫秒)。如果您沒有指定格式，Amazon 資料 Firehose 預設會使 `java.sql.Timestamp::valueOf` 用。

配置單元 JSON SerDe 不允許以下內容：

- 欄位名稱內的句點 (.)。
- 類型為 `uniontype` 的欄位。
- 結構描述內有數值類型但屬於 JSON 中字串的欄位，例如，如果模式是 (一個 `int`)，而 JSON 是 `{"a":"123"}`，配置單元 SerDe 給出了一個錯誤。

配置單元 SerDe 不會將嵌套的 JSON 轉換為字符串。例如，若其中有 `{"a":{"inner":1}}`，則 `{"inner":1}` 不會視為字串。

## 選擇序列化程式

您選擇的序列化程式取決於您的商業需求。 [要了解有關兩個序列化程序選項的更多信息，請參閱 ORC SerDe 和 鑲木地板。 SerDe](#)

## 轉換輸入記錄格式 (主控台)

您可以在建立或更新 Firehose 串流時，在主控台上啟用資料格式轉換。啟用資料格式轉換後，Amazon S3 是唯一可為 Firehose 串流設定的目的地。此外，啟用格式轉換將停用 Amazon S3 壓縮。然而，轉換程序中會自動出現 Snappy 壓縮。Amazon 數據 Firehose 在這種情況下使用的 Snappy 的框架格式與 Hadoop 兼容。這表示您可使用 Snappy 壓縮的結果，並在 Athena 中查詢這些資料。 [有關 Hadoop 依賴的活潑框架格式，請參閱 `.java.BlockCompressorStream`](#)

啟用資料 Firehose 串流的資料格式轉換

1. 登入 AWS Management Console，然後開啟 Amazon 資料 Firehose 主控台，網址為 <https://console.aws.amazon.com/firehose/>。
2. 選擇要更新的 Firehose 串流，或依照中的步驟建立新的 Firehose 串流。 [建立 Firehose 串流](#)

3. 在 Convert record format (轉換記錄格式) 底下，將 Record format conversion (記錄格式轉換) 設定為 Enabled (已啟用)。
4. 選擇您想要的輸出格式。如需兩個選項的詳細資訊，請參閱 [Apache Parquet](#) 和 [Apache ORC](#)。
5. 選擇 AWS Glue 表格以指定來源記錄的結構描述。設定區域、資料庫、表格與表格版本。

## 轉換輸入記錄格式 (API)

如果您希望 Amazon 數據 Firehose 將輸入數據的格式從 JSON 轉換為鑲木地板或 ORC，請在擴展 DS3 或擴展 DS DestinationConfiguration 3 中指定可選 [DataFormatConversionConfiguration](#) 元素。DestinationUpdate 如果您指定 [DataFormatConversionConfiguration](#)，則會套用下列限制：

- 在中 [BufferingHints](#)，如果啟用記錄格式轉換，則無法設定 SizeInMBs 為小於 64 的值。格式轉換未啟用時，該值預設為 5，啟用後則變成 128。
- 您必須在 [延伸 DS3 CompressionFormat](#) 中 DestinationConfiguration 或在 [延伸 DS3](#) 中設定為 [DestinationUpdate](#) UNCOMPRESSEDCompressionFormat 的預設值為 UNCOMPRESSED。因此，您也可以將其保留在 [擴展 DestinationConfiguration](#) DS3 中未指定。序列化程序仍會壓縮這些資料，預設使用 Snappy 壓縮。Amazon 數據 Firehose 在這種情況下使用的 Snappy 的框架格式與 Hadoop 兼容。這表示您可使用 Snappy 壓縮的結果，並在 Athena 中查詢這些資料。[有關 Hadoop 依賴的活潑框架格式](#)，請參閱 [.java。BlockCompressorStream](#) 設定序列化程式時，您可選擇其他壓縮類型。

## 記錄格式轉換錯誤處理

當 Amazon 資料 Firehose 無法剖析或還原序列化記錄 (例如，當資料與結構描述不相符時)，它會使用錯誤前置詞將其寫入 Amazon S3。如果寫入失敗，Amazon 資料 Firehose 會永遠重試，阻止進一步的交付。對於每個失敗的記錄，Amazon 資料 Firehose 會使用下列結構描述撰寫 JSON 文件：

```
{
  "attemptsMade": long,
  "arrivalTimestamp": long,
  "lastErrorCode": string,
  "lastErrorMessage": string,
  "attemptEndingTimestamp": long,
  "rawData": string,
  "sequenceNumber": string,
  "subSequenceNumber": long,
  "dataCatalogTable": {
```

```
"catalogId": string,  
"databaseName": string,  
"tableName": string,  
"region": string,  
"versionId": string,  
"catalogArn": string  
}  
}
```

## 記錄格式轉換範例

如需如何設定記錄格式轉換的範例AWS CloudFormation，請參閱[AWS:DataFirehose::DeliveryStream](#)。

# 使用 Amazon Managed Service for Apache Flink

透過 Amazon Managed Service for Apache Flink，您可以使用 Java、Scala 或 SQL 來處理和分析串流資料。此服務可讓您針對串流來源撰寫和執程式碼，以執行時間序列分析、饋送即時儀表板，以及建立即時指標。

如需與適用於 Apache Flink 的 Amazon 受管服務整合的範例，請參閱[範例：寫入 Amazon 資料 Firehose](#)。

在本練習中，您會建立一個 Apache Flink 應用程式，該應用程式將 Kinesis 資料串流作為來源，並以 Firehose 串流做為接收器。使用接收器，您就可以驗證 Amazon S3 儲存貯體的應用程式輸出。

開始之前，請先設定必要的先決條件：

- [適用於 Apache Flink 應用程式的受管理服務元件](#)
- [完成練習的先決條件](#)



# Amazon 數據 Firehose 數據交付

資料傳送至 Firehose 串流後，資料會自動傳送至您選擇的目的地。

## Important

如果您使用 Kinesis Producer Library (KPL) 來寫入資料到 Kinesis 資料串流，則可使用彙整來合併您寫入至該 Kinesis 資料串流的記錄。如果您接著使用該資料串流作為 Firehose 串流的來源，Amazon Data Firehose 會先將記錄解彙總，然後再將記錄傳送到目的地。如果您將 Firehose 串流設定為轉換資料，Amazon 資料 Firehose 會先將記錄解除彙總，然後再傳送記錄。AWS Lambda 如需詳細資訊，請參閱《Amazon Kinesis Data Streams 開發人員指南》中的[使用 Kinesis Producer Library 開發 Amazon Kinesis Data Streams 生產者及彙整](#)。

## 主題

- [資料交付格式](#)
- [資料交付頻率](#)
- [資料交付失敗的處理方式](#)
- [Amazon S3 物件名稱格式](#)
- [OpenSearch 服務目的地的索引輪替](#)
- [HTTP 端點目的地跨 AWS 帳戶和跨 AWS 區域交付](#)
- [重複的記錄](#)
- [如何暫停和繼續 Firehose 傳送串流](#)

## 資料交付格式

對於將資料交付至 Amazon S3 Simple Storage Service (Amazon S3)，Firehose 會根據交付串流的緩衝組態連接多個傳入記錄。接著，該服務會將這些記錄作為 Amazon S3 物件，並交付至 Amazon S3。根據預設，Firehose 會連接不含任何分隔符號的資料。[如果您想要在記錄之間有新的行分隔符號，您可以啟用 Firehoseconsole 組態或 API 參數中的功能來新增行分隔符號。](#)

若要將資料傳送至 Amazon Redshift，Firehose 會先以先前所述的格式將傳入的資料交付到 S3 儲存貯體。然後，Firehose 會發出亞馬遜 Redshift COPY 命令，將資料從 S3 儲存貯體載入到 Amazon Redshift 佈建的叢集或 Amazon Redshift 無伺服器工作群組。確保 Amazon 資料 Firehose 將多個傳入記錄連接到 Amazon S3 物件後，可將 Amazon S3 物件複製到您的 Amazon Redshift 佈建叢集或



Amazon Redshift 無伺服器工作群組。如需詳細資訊，請參閱 [Amazon Redshift COPY 命令資料格式參數](#)。

對於將資料傳送至 OpenSearch 服務和 OpenSearch 無伺服器，Amazon Data Firehose 會根據 Firehose 串流的緩衝組態來緩衝傳入的記錄。然後，它會產生 OpenSearch 服務或 OpenSearch 無伺服器大量要求，將多筆記錄索引至您的 Ser OpenSearch vice 叢集或 OpenSearch 無伺服器集合。在將記錄傳送至 Amazon 資料 Firehose 之前，請確定您的記錄是 UTF-8 編碼並展平為單行 JSON 物件。此外，您的 OpenSearch Service 叢集的 `rest.action.multi.allow_explicit_index` 選項必須設定為 `true` (預設值)，才能使用每筆記錄設定的明確索引來批次處理要求。如需詳細資訊，請參閱 Amazon OpenSearch 服務開發人員指南中的 OpenSearch 服務 [設定進階選項](#)。

若要將資料傳送至 Splunk，Amazon 資料 Firehose 會串連您傳送的位元組。如果您的資料需要定界符號 (例如換行字元)，您必須自行插入。請確認將 Splunk 設定為剖析這類定界符號。

將資料傳送到受支援的第三方服務供應商所擁有的 HTTP 端點時，您可以使用整合式 Amazon Lambda 服務建立函數，將傳入記錄轉換為符合服務提供者的整合預期的格式。請聯絡您為目的地選擇 HTTP 端點的第三方服務供應商，以進一步了解其接受的記錄格式。

## 資料交付頻率

每個 Firehose 目的地都有自己的資料傳送頻率。如需詳細資訊，請參閱 [緩衝提示](#)。

## 資料交付失敗的處理方式

每個 Amazon 資料 Firehose 目的地都有自己的資料交付失敗處理。

### Amazon Simple Storage Service (Amazon S3)

資料交付至 S3 儲存貯體時，可能會因各種問題導致失敗。例如，儲存貯體可能不再存在，Amazon Data Firehose 假設的 IAM 角色可能無法存取儲存貯體、網路故障或類似事件。在這些情況下，Amazon 資料 Firehose 會持續重試長達 24 小時，直到交付成功為止。Amazon 數據 Firehose 件的最長數據存儲時間為 24 小時。如果超過 24 小時仍無法交付，資料將會遺失。

### Amazon Redshift

對於 Amazon Redshift 目的地，您可以在建立 Firehose 串流時指定重試持續時間 (0—7200 秒)。

將資料交付到您的 Amazon Redshift 佈建的叢集或 Amazon Redshift Serverless 工作群組可能會因多種原因而失敗。例如，您可能 Firehose 串流的叢集配置不正確、正在維護的叢集或工作群組，或是網路故障。在這些情況下，Amazon 資料 Firehose 會在指定的持續時間內重試，並略過該特定批次的 Amazon S3 物件。略過物件的資訊會傳送至 S3 儲存貯體，成為 `errors/` 資料夾中的資

訊清單檔案，以供手動回填使用。如需進一步了解如何使用資訊清單檔案來手動 COPY 資料，請參閱[使用資訊清單來指定資料檔案](#)。

## Amazon OpenSearch 服務和 OpenSearch 無伺服器

對於「OpenSearch 服務」和「OpenSearch 無伺服器」目的地，您可以在建立交付串流時指定重試持續時間 (0—7200 秒)。

資料傳遞至您的 Ser OpenSearch vice 叢集或 OpenSearch 無伺服器集合可能會因為多種原因而失敗。例如，您的 Firehose 串流可能有不正確的 OpenSearch 服務叢集或 OpenSearch 無伺服器集合組態、維護中的 OpenSearch 服務叢集或 OpenSearch 無伺服器集合、網路故障或類似事件。在這些情況下，Amazon Data Firehose 會在指定的持續時間內重試，然後略過該特定索引請求。略過的文件會傳輸至 S3 儲存貯體的 AmazonOpenSearchService\_failed/資料夾，以供手動回填使用。

對於 OpenSearch 服務，每個文件都具有以下 JSON 格式：

```
{
  "attemptsMade": "(number of index requests attempted)",
  "arrivalTimestamp": "(the time when the document was received by Firehose)",
  "errorCode": "(http error code returned by OpenSearch Service)",
  "errorMessage": "(error message returned by OpenSearch Service)",
  "attemptEndingTimestamp": "(the time when Firehose stopped attempting index request)",
  "esDocumentId": "(intended OpenSearch Service document ID)",
  "esIndexName": "(intended OpenSearch Service index name)",
  "esTypeName": "(intended OpenSearch Service type name)",
  "rawData": "(base64-encoded document data)"
}
```

對於 OpenSearch 無伺服器，每個文件都具有下列 JSON 格式：

```
{
  "attemptsMade": "(number of index requests attempted)",
  "arrivalTimestamp": "(the time when the document was received by Firehose)",
  "errorCode": "(http error code returned by OpenSearch Serverless)",
  "errorMessage": "(error message returned by OpenSearch Serverless)",
  "attemptEndingTimestamp": "(the time when Firehose stopped attempting index request)",
  "osDocumentId": "(intended OpenSearch Serverless document ID)",
  "osIndexName": "(intended OpenSearch Serverless index name)",
  "rawData": "(base64-encoded document data)"
}
```

```
}
```

## Splunk

當 Amazon 數據 Firehose 將數據發送到 Splunk 時，它會等待 Splunk 的確認。如果發生錯誤，或確認未在確認逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon 資料 Firehose 將資料傳送至 Splunk 時，無論是初始嘗試還是重試，都會重新啟動確認逾時計數器。並等待 Splunk 傳回確認訊息。即使重試持續時間過期，Amazon Data Firehose 仍會等待確認，直到收到確認或到達確認逾時為止。如果確認逾時，Amazon Data Firehose 會檢查以判斷重試計數器是否有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到確認或判斷已無重試時間為止。

未收到確認訊息，並非唯一的資料交付錯誤類型。如需其他資料交付錯誤類型的相關資訊，請參閱 [Splunk 資料交付錯誤](#)。如果您的重試期間大於 0，則任何資料交付錯誤都會觸發重試邏輯。

錯誤記錄例示如下。

```
{
  "attemptsMade": 0,
  "arrivalTimestamp": 1506035354675,
  "errorCode": "Splunk.AckTimeout",
  "errorMessage": "Did not receive an acknowledgement from HEC before the HEC acknowledgement timeout expired. Despite the acknowledgement timeout, it's possible the data was indexed successfully in Splunk. Amazon Data Firehose backs up in Amazon S3 data for which the acknowledgement timeout expired.",
  "attemptEndingTimestamp": 13626284715507,
  "rawData":
  "MiAyNTE2MjAyNzIyMDkgZW5pLTA1ZjMyMmQ1IDIxOC45Mi4xODguMjE0IDE3Mi4xNi4xLjE2NyAyNTIzMyAxNDMzID",
  "EventId": "49577193928114147339600778471082492393164139877200035842.0"
}
```

## HTTP 端點目的地

當 Amazon 資料 Firehose 將資料傳送到 HTTP 端點目的地時，它會等待來自此目的地的回應。如果發生錯誤，或回應未在回應逾時期間內送達，Amazon Data Firehose 會啟動重試持續時間計數器。直到重試持續時間過期之前，該服務都不會中斷重試作業。之後，亞馬遜數據 Firehose 將其視為數據交付失敗，並將數據備份到您的 Amazon S3 存儲桶。

每次 Amazon Data Firehose 將資料傳送到 HTTP 端點目的地時，無論是初始嘗試還是重試，都會重新啟動回應逾時計數器。然後，它會等待來自 HTTP 端點目的地的回應到達。即使重試持續時間過期，Amazon Data Firehose 仍會等待回應，直到收到回應或到達回應逾時為止。如果回應逾時，Amazon Data Firehose 會檢查重試計數器是否有剩餘時間。如果還有時間，將再次重試並重複執行邏輯，直到收到回應或判斷已無重試時間為止。

未收到回應，並非唯一的資料交付錯誤類型。如需其他資料交付錯誤類型的相關資訊，請參閱 [HTTP 端點資料交付錯誤](#)

錯誤記錄例示如下。

```
{
  "attemptsMade":5,
  "arrivalTimestamp":1594265943615,
  "errorCode":"HttpEndpoint.DestinationException",
  "errorMessage":"Received the following response from the endpoint destination.
  {\"requestId\": \"109777ac-8f9b-4082-8e8d-b4f12b5fc17b\", \"timestamp\": 1594266081268,
  \"errorMessage\": \"Unauthorized\"}",
  "attemptEndingTimestamp":1594266081318,
  "rawData\":\"c2FtcGx1IHJhdyBkYXRh\",
  "subsequenceNumber":0,
  "dataId\":\"49607357361271740811418664280693044274821622880012337186.0\"
}
```

## Amazon S3 物件名稱格式

當 Firehose 將資料傳送至 Amazon S3 時，S3 物件金鑰名稱會遵循以下格式 <evaluated prefix><suffix>，其中後置字元的格式<delivery stream name><delivery stream version><year><month><day><hour><minute><second><uuid><file extension><delivery stream version>為 ----- 從 1 開始，每次 Firehose 交付串流的組態變更都會增加 1。您可以變更交付串流的組態 (如 S3 儲存貯體名稱、緩衝提示、壓縮及解密操作)。您可以使用 Firehose 控制台或 [UpdateDestination](#) API 操作來完成此操作。

對於<evaluated prefix>，Firehose 會在格式中加入預設時間字 YYYY/MM/dd/HH 首。此前置詞會在值區中建立邏輯階層，其中每個正斜線 (/) 都會在階層中建立層級。您可以透過指定包含在執行階段評估之運算式的自訂前置詞來修改此結構。如需如何指定自訂前置詞的詳細資訊，請參閱 [Amazon 簡單儲存服務物件的自訂首碼](#)。

根據預設，用於時間前置字元的時區為 UTC，但您可以將其變更為您偏好的時區。例如，如果您想要使用日本標準時間而不是 UTC，則可以在 AWS Management Console 或 [API 參數設定 \(CustomTimeZone\)](#) 中將時區設定為亞洲/東京。下列清單包含 Firehose 中 S3 前置詞組態支援的時區：

## Firehose 中 S3 前置詞組態支援的時區

### Africa

```
Africa/Abidjan
Africa/Accra
Africa/Addis_Ababa
Africa/Algiers
Africa/Asmera
Africa/Bangui
Africa/Banjul
Africa/Bissau
Africa/Blantyre
Africa/Bujumbura
Africa/Cairo
Africa/Casablanca
Africa/Conakry
Africa/Dakar
Africa/Dar_es_Salaam
Africa/Djibouti
Africa/Douala
Africa/Freetown
Africa/Gaborone
Africa/Harare
Africa/Johannesburg
Africa/Kampala
Africa/Khartoum
Africa/Kigali
Africa/Kinshasa
Africa/Lagos
Africa/Libreville
Africa/Lome
Africa/Luanda
Africa/Lubumbashi
Africa/Lusaka
Africa/Malabo
Africa/Maputo
Africa/Maseru
```

Africa/Mbabane  
Africa/Mogadishu  
Africa/Monrovia  
Africa/Nairobi  
Africa/Ndjamena  
Africa/Niamey  
Africa/Nouakchott  
Africa/Ouagadougou  
Africa/Porto-Novo  
Africa/Sao\_Tome  
Africa/Timbuktu  
Africa/Tripoli  
Africa/Tunis  
Africa/Windhoek

## America

America/Adak  
America/Anchorage  
America/Anguilla  
America/Antigua  
America/Aruba  
America/Asuncion  
America/Barbados  
America/Belize  
America/Bogota  
America/Buenos\_Aires  
America/Caracas  
America/Cayenne  
America/Cayman  
America/Chicago  
America/Costa\_Rica  
America/Cuiaba  
America/Curacao  
America/Dawson\_Creek  
America/Denver  
America/Dominica  
America/Edmonton  
America/El\_Salvador  
America/Fortaleza  
America/Godthab  
America/Grand\_Turk  
America/Grenada

America/Guadeloupe  
America/Guatemala  
America/Guayaquil  
America/Guyana  
America/Halifax  
America/Havana  
America/Indianapolis  
America/Jamaica  
America/La\_Paz  
America/Lima  
America/Los\_Angeles  
America/Managua  
America/Manaus  
America/Martinique  
America/Mazatlan  
America/Mexico\_City  
America/Miquelon  
America/Montevideo  
America/Montreal  
America/Montserrat  
America/Nassau  
America/New\_York  
America/Noronha  
America/Panama  
America/Paramaribo  
America/Phoenix  
America/Port\_of\_Spain  
America/Port-au-Prince  
America/Porto\_Acre  
America/Puerto\_Rico  
America/Regina  
America/Rio\_Branco  
America/Santiago  
America/Santo\_Domingo  
America/Sao\_Paulo  
America/Scoresbysund  
America/St\_Johns  
America/St\_Kitts  
America/St\_Lucia  
America/St\_Thomas  
America/St\_Vincent  
America/Tegucigalpa  
America/Thule  
America/Tijuana

```
America/Tortola  
America/Vancouver  
America/Winnipeg
```

## Antarctica

```
Antarctica/Casey  
Antarctica/DumontDURville  
Antarctica/Mawson  
Antarctica/McMurdo  
Antarctica/Palmer
```

## Asia

```
Asia/Aden  
Asia/Almaty  
Asia/Amman  
Asia/Anadyr  
Asia/Aqtau  
Asia/Aqtobe  
Asia/Ashgabat  
Asia/Ashkhabad  
Asia/Baghdad  
Asia/Bahrain  
Asia/Baku  
Asia/Bangkok  
Asia/Beirut  
Asia/Bishkek  
Asia/Brunei  
Asia/Calcutta  
Asia/Colombo  
Asia/Dacca  
Asia/Damascus  
Asia/Dhaka  
Asia/Dubai  
Asia/Dushanbe  
Asia/Hong_Kong  
Asia/Irkutsk  
Asia/Jakarta  
Asia/Jayapura  
Asia/Jerusalem  
Asia/Kabul  
Asia/Kamchatka
```



Asia/Karachi  
Asia/Katmandu  
Asia/Krasnoyarsk  
Asia/Kuala\_Lumpur  
Asia/Kuwait  
Asia/Macao  
Asia/Magadan  
Asia/Manila  
Asia/Muscat  
Asia/Nicosia  
Asia/Novosibirsk  
Asia/Phnom\_Penh  
Asia/Pyongyang  
Asia/Qatar  
Asia/Rangoon  
Asia/Riyadh  
Asia/Saigon  
Asia/Seoul  
Asia/Shanghai  
Asia/Singapore  
Asia/Taipei  
Asia/Tashkent  
Asia/Tbilisi  
Asia/Tehran  
Asia/Thimbu  
Asia/Thimphu  
Asia/Tokyo  
Asia/Ujung\_Pandang  
Asia/Ulaanbaatar  
Asia/Ulan\_Bator  
Asia/Vientiane  
Asia/Vladivostok  
Asia/Yakutsk  
Asia/Yekaterinburg  
Asia/Yerevan

## Atlantic

Atlantic/Azores  
Atlantic/Bermuda  
Atlantic/Canary  
Atlantic/Cape\_Verde  
Atlantic/Faeroe

Atlantic/Jan\_Mayen  
Atlantic/Reykjavik  
Atlantic/South\_Georgia  
Atlantic/St\_Helena  
Atlantic/Stanley

## Australia

Australia/Adelaide  
Australia/Brisbane  
Australia/Broken\_Hill  
Australia/Darwin  
Australia/Hobart  
Australia/Lord\_Howe  
Australia/Perth  
Australia/Sydney

## Europe

Europe/Amsterdam  
Europe/Andorra  
Europe/Athens  
Europe/Belgrade  
Europe/Berlin  
Europe/Brussels  
Europe/Bucharest  
Europe/Budapest  
Europe/Chisinau  
Europe/Copenhagen  
Europe/Dublin  
Europe/Gibraltar  
Europe/Helsinki  
Europe/Istanbul  
Europe/Kaliningrad  
Europe/Kiev  
Europe/Lisbon  
Europe/London  
Europe/Luxembourg  
Europe/Madrid  
Europe/Malta  
Europe/Minsk  
Europe/Monaco  
Europe/Moscow

Europe/Oslo  
Europe/Paris  
Europe/Prague  
Europe/Riga  
Europe/Rome  
Europe/Samara  
Europe/Simferopol  
Europe/Sofia  
Europe/Stockholm  
Europe/Tallinn  
Europe/Tirane  
Europe/Vaduz  
Europe/Vienna  
Europe/Vilnius  
Europe/Warsaw  
Europe/Zurich

## Indian

Indian/Antananarivo  
Indian/Chagos  
Indian/Christmas  
Indian/Cocos  
Indian/Comoro  
Indian/Kerguelen  
Indian/Mahe  
Indian/Maldives  
Indian/Mauritius  
Indian/Mayotte  
Indian/Reunion

## Pacific

Pacific/Apia  
Pacific/Auckland  
Pacific/Chatham  
Pacific/Easter  
Pacific/Efate  
Pacific/Enderbury  
Pacific/Fakaofu  
Pacific/Fiji  
Pacific/Funafuti  
Pacific/Galapagos

```

Pacific/Gambier
Pacific/Guadalcanal
Pacific/Guam
Pacific/Honolulu
Pacific/Kiritimati
Pacific/Kosrae
Pacific/Majuro
Pacific/Marquesas
Pacific/Nauru
Pacific/Niue
Pacific/Norfolk
Pacific/Noumea
Pacific/Pago_Pago
Pacific/Palau
Pacific/Pitcairn
Pacific/Ponape
Pacific/Port_Moresby
Pacific/Rarotonga
Pacific/Saipan
Pacific/Tahiti
Pacific/Tarawa
Pacific/Tongatapu
Pacific/Truk
Pacific/Wake
Pacific/Wallis

```

您無法變更字尾欄位，除了 <file extension>。當您啟用資料格式轉換或壓縮時，Firehose 會根據組態附加檔案副檔名。下表說明 Firehose 附加的預設副檔名：

組態	副檔名
數據格式轉換：鑲木地板	. 鑲木地板
數據格式轉換：ORC	. 獸人
壓縮：	.gz
壓縮：拉鍊	.zip
壓縮：活潑	.snappy

組態	副檔名
壓縮：波多，活潑	. 赫納皮

您也可以在 Firehose 主控台或 API 中指定您偏好的副檔名。副檔名必須以句點 (.) 開頭，且可包含允許的字元：0-9a-z! -\_.\*' ()。副檔名不能超過 128 個字元。

#### Note

當您指定檔案副檔名時，它會覆寫 Firehose 在啟用[資料格式轉換](#)或壓縮時新增的預設檔案副檔名。

## OpenSearch 服務目的地的索引輪替

對於「OpenSearch 服務」目的地，您可以從下列五個選項之一指定以時間為基礎的索引循環選項：NoRotationOneHour、OneDay、OneWeek、或OneMonth。

根據您選擇的輪換選項，Amazon Data Firehose 會將 UTC 到達時間戳記的一部分附加到您指定的索引名稱。並依此輪換該時間戳記。下列範例會針對每個索引循環選項，顯示 OpenSearch Service 中產生的索引名稱，其中指定的索引名稱為，myindex而到達時間戳記為2016-02-25T13:00:00Z。

RotationPeriod	IndexName
NoRotation	myindex
OneHour	myindex-2016-02-25-13
OneDay	myindex-2016-02-25
OneWeek	myindex-2016-w08
OneMonth	myindex-2016-02

**Note**

透過 OneWeek 選項，Data Firehose 會使用 <YEAR>-w<WEEK NUMBER> 格式自動建立索引 (例如，2020-w33)，其中週數是使用 UTC 時間計算，並根據下列美國慣例來計算：

- 一週的開始日為週日
- 一年中的第一週是指當年包含週六的第一週

## HTTP 端點目的地跨 AWS 帳戶和跨 AWS 區域交付

Amazon 資料 Firehose 支援跨 AWS 帳戶將資料傳遞到 HTTP 端點目的地。Amazon 數據火管 Firehose 流和您選擇作為目的地的 HTTP 端點可以位於不同 AWS 的帳戶中。

Amazon 資料 Firehose 也支援跨 AWS 區域將資料傳遞到 HTTP 端點目的地。您可以將資料從某個區域的 Firehose 串流傳送到另一個 AWS 區 AWS 域的 HTTP 端點。您也可以將 HTTP 端點 URL 設定為您想要的目的地，將資料從 Firehose 串流傳遞至 AWS 區域以外的 HTTP 端點目的地，例如傳送至您自己的內部部署伺服器。在這些情況下，額外的資料傳輸費用會新增到您的交付成本中。如需詳細資訊，請參閱「隨需定價」頁面上的[資料傳輸](#)一節。

## 重複的記錄

Amazon 數據 Firehose 使用 at-least-once 語義進行數據交付。在某些情況下，例如資料交付逾時時，如果原始資料傳遞請求最終通過，Amazon Data Firehose 的交付重試可能會導致重複項目。這適用於 Amazon 資料 Firehose 支援的所有目標類型。

## 如何暫停和繼續 Firehose 傳送串流

在 Firehose 中設定交付串流後，串流來源中可用的資料會持續傳送至目的地。如果您遇到串流目的地暫時無法使用的情況 (例如，在規劃的維護操作期間)，您可能會想要暫停資料交付，並在目的地再次可用時繼續執行。下列各節說明如何完成此操作：

**Important**

當您使用下述方法暫停和繼續串流時，在繼續串流之後，您會看到很少有記錄傳送到 Amazon S3 中的錯誤儲存貯體，而其餘串流會繼續傳送到目的地。這是此方法的已知限制，因為在將多次重試追蹤為失敗之後，少數記錄先前無法傳遞至目的地。

## 了解 Firehose 如何處理派送失敗

在 Firehose 中設定交付串流時，對於許多目的地 (例如 OpenSearch Splunk 和 HTTP 端點)，您也會設定 S3 儲存貯體，在其中備份無法交付的資料。如需有關 Firehose 如何在傳送失敗時備份資料的詳細資訊，請參閱[資料傳遞失敗處理](#)。如需如何授與 S3 儲存貯體的存取權限 (其中可備份無法交付的資料) 的詳細資訊，請參閱[授與 Firehose 存取權給 Amazon S3 目的地](#)。當 Firehose (a) 無法將資料傳遞至串流目的地，且 (b) 無法將資料寫入備份 S3 儲存貯體以進行失敗交付時，它會有效地暫停串流交付，直到資料可以傳送到目的地或寫入備份 S3 位置為止。

## 暫停「Firehose」傳送串流

若要在 Firehose 中暫停串流傳遞，請先移除 Firehose 寫入 S3 備份位置以便失敗交付的許可。例如，如果您想要暫停具有 OpenSearch 目的地的傳送串流，可以透過更新權限來執行此操作。如需詳細資訊，請參閱[將 Firehose 存取權授予公共 OpenSearch 服務目的地](#)。

移除動作 `s3:PutObject` 的 "Effect": "Allow" 許可，並明確新增陳述式，該陳述式會針對用於備份失敗交付的 S3 儲存貯體的動作 `s3:PutObject` 套用 "Effect": "Deny" 許可。接下來，關閉串流目的地 (例如，關閉目標 OpenSearch 網域)，或移除 Firehose 寫入目的地的權限。若要更新其他目的地的許可，請查看[使用 Amazon Data Firehose 控制存取](#)中目的地的章節。完成這兩個動作之後，Firehose 將停止傳送串流，而且您可以使用 [Firehose 的 CloudWatch 指標](#) 來監控此項目。

### Important

當您在 Firehose 中暫停串流傳遞時，您需要確保串流的來源 (例如，在 Kinesis Data Streams 或適用於 Kafka 的受管理服務中) 設定為保留資料，直到串流傳遞恢復並將資料傳送至目的地為止。如果來源是「直接輸入」，則 Firehose 會保留資料 24 小時。如果您未繼續串流，並在資料保留期限到期之前交付資料，則可能會發生資料遺失。

## 恢復 Firehose 交付流

若要繼續傳遞，請先開啟目的地並確保 Firehose 具有將串流傳送至目的地的權限，將先前所做的變更還原至串流目的地。接下來，還原先套用於 S3 儲存貯體的許可變更，以備份失敗的交付。也就是說，套用動作 `s3:PutObject` 的 "Effect": "Allow" 許可，並對用於備份失敗交付的 S3 儲存貯體的動作 `s3:PutObject` 移除 "Effect": "Deny" 許可。最後，使用 [Firehose 的 CloudWatch 指標](#) 進行監視，以確認串流正在傳送至目的地。若要檢視和疑難排解錯誤，請使用適用於 [Firehose 的 Amazon CloudWatch 日誌監控功能](#)。

# 監控 Amazon 數據 Firehose

您可以使用以下功能監控 Amazon 數據 Firehose：

## 主題

- [CloudWatch 警示的最佳實務](#)
- [使 CloudWatch 用指標監控 Amazon 資料 Firehose](#)
- [訪問 CloudWatch Amazon 數據 Firehose 的指標](#)
- [使 CloudWatch 用日誌監控 Amazon 數據 Firehose](#)
- [訪問 Amazon 數據 Firehose 的 CloudWatch 日誌](#)
- [監控 Kinesis 代理程式運作狀態](#)
- [使用日誌記錄 Amazon 數據 Firehose API 調用 AWS CloudTrail](#)

## CloudWatch 警示的最佳實務

新增下列指標超過緩衝限制時的 CloudWatch 警示 (最多 15 分鐘)：

- `DeliveryToS3.DataFreshness`
- `DeliveryToSplunk.DataFreshness`
- `DeliveryToAmazonOpenSearchService.DataFreshness`
- `DeliveryToAmazonOpenSearchServerless.DataFreshness`
- `DeliveryToHttpEndpoint.DataFreshness`

此外，根據以下指標數學表達式建立警示。

- `IncomingBytes (Sum per 5 Minutes) / 300` 接近 `BytesPerSecondLimit` 的百分比。
- `IncomingRecords (Sum per 5 Minutes) / 300` 接近 `RecordsPerSecondLimit` 的百分比。
- `IncomingPutRequests (Sum per 5 Minutes) / 300` 接近 `PutRequestsPerSecondLimit` 的百分比。

我們建議為其建立警示的另一個指標是 `ThrottledRecords`。



如需關於警示進入 ALARM 狀態時的故障排除詳細資訊，請參閱[故障診斷](#)。

## 使 CloudWatch 用指標監控 Amazon 資料 Firehose

### Important

請務必在屬於您目的地的所有 CloudWatch 量度上啟用警示，以便及時識別錯誤。

Amazon 資料 Firehose 與 Amazon CloudWatch 指標整合，因此您可以收集、檢視和分析 Firehose 串流的 CloudWatch 指標。例如，您可以監控 IncomingBytes 和 IncomingRecords 標，以追蹤從資料生產者擷取到 Amazon Data Firehose 的資料。

Amazon 資料 Firehose 每分鐘收集和發佈 CloudWatch 指標。但是，如果傳入資料的突發僅發生幾秒鐘，則可能無法在一分鐘指標中完全擷取或顯示這些資料。這是因為 CloudWatch 指標會在一分鐘內從 Amazon 資料 Firehose 彙總一次。

為「Firehose」串流收集的指標是免費的。如需 Kinesis 代理程式指標的詳細資訊，請參閱 [監控 Kinesis 代理程式運作狀態](#)。

### 主題

- [動態分割 CloudWatch 指標](#)
- [資料傳送 CloudWatch 指標](#)
- [資料導入指標](#)
- [API 層 CloudWatch 級測量結果](#)
- [資料轉換 CloudWatch 指標](#)
- [CloudWatch 日誌解壓縮測量結果](#)
- [格式轉換 CloudWatch 量度](#)
- [伺服器端加密 \(SSE\) CloudWatch 量度](#)
- [Amazon 數據 Firehose 的尺寸](#)
- [Amazon 數據 Firehose 使用指標](#)

## 動態分割 CloudWatch 指標

如果啟用[動態磁碟分割](#)，AWS/Firehose 命名空間會包含下列量度。

指標	描述
ActivePartitionsLimit	Firehose 會在將資料傳送至錯誤值區之前，串流處理的作用中分割區數目上限。  單位：計數
PartitionCount	正在處理的分割區數目，換句話說，使用中的分割區計數。這個數字在 1 和分割區數量上限的 500 (預設值) 之間變化。  單位：計數
PartitionCountExceeded	此指標表示您是否超過分割數量上限。它根據是否違反限制發出 1 或 0。
JQProcessing.Duration	傳回在 JQ Lambda 函數中執行 JQ 運算式所花費的時間量。  單位：毫秒
PerPartitionThroughput	表示每個分割區正在處理的輸送量。此指標可讓您能監控每個分割區的輸送量。  單位: StandardUnit. BytesSecond
DeliveryToS3.ObjectCount	表示要交付到 S3 儲存貯體的物件數目。  單位：計數

## 資料傳送 CloudWatch 指標

AWS/Firehose 命名空間包含下列服務層級指標。如果您看到 BackupToS3.Success、DeliveryToS3.Success、DeliveryToSplunk.Success、DeliveryToAma 或 DeliveryToRedshift.Success 的平均值小幅下降，並不表示資料遺失。Amazon Data Firehose 會重試交付錯誤，並且在記錄成功交付到設定的目的地或備份 S3 儲存貯體之前不會向前移動。

### 主題

- [交付至 OpenSearch 服務](#)
- [交付至 OpenSearch無伺服器](#)
- [交付至 Amazon Redshift](#)
- [交付至 Amazon S3](#)
- [送到雪花](#)
- [交付至 Splunk](#)
- [交付至 HTTP 端點](#)

## 交付至 OpenSearch 服務

指標	描述
<code>DeliveryToAmazonOpenSearchService.Bytes</code>	指定期間內索引至 OpenSearch Service 的位元組數目。  單位：位元組
<code>DeliveryToAmazonOpenSearchService.DataFreshness</code>	Amazon 數據 Firehose 中最古老的記錄的年齡 ( 從進入 Amazon 數據 Firehose 到現在 )。任何超過此年齡的記錄都已交付給 OpenSearch 服務。  單位：秒
<code>DeliveryToAmazonOpenSearchService.Records</code>	指定期間內為 OpenSearch Service 編製索引的記錄數目。  單位：計數
<code>DeliveryToAmazonOpenSearchService.Success</code>	成功編製索引之記錄的總和與已嘗試之記錄的總和。
<code>DeliveryToS3.Bytes</code>	在指定期間內，交付至 Amazon S3 的位元組數目。只有在您為所有文件啟用備份時，Amazon 資料 Firehose 才會發出此指標。  單位：計數
<code>DeliveryToS3.DataFreshness</code>	Amazon 數據 Firehose 中最古老的記錄的年齡 ( 從進入 Amazon 數據 Firehose 到現在 )。任何比此存留期還要久

指標	描述
	<p>的記錄都已交付至 S3 儲存貯體。只有在您為所有文件啟用備份時，Amazon 資料 Firehose 才會發出此指標。</p> <p>單位：秒</p>
DeliveryToS3.Records	<p>在指定期間內，交付至 Amazon S3 的記錄數目。只有在您為所有文件啟用備份時，Amazon 資料 Firehose 才會發出此指標。</p> <p>單位：計數</p>
DeliveryToS3.Success	<p>所有 Amazon S3 Put 命令總和中的成功 Amazon S3 Put 命令總和。無論是僅針對失敗文件還是針對所有文件啟用備份，Amazon Data Firehose 一律會發出此指標。</p>
DeliveryToAmazonOpenSearchService.AuthFailure	<p>身分驗證和授權錯誤。驗證 OS/ES 叢集政策和角色許可。</p> <p>0 表示沒有問題。1 表示驗證失敗。</p>
DeliveryToAmazonOpenSearchService.DeliveryRejected	<p>交付拒絕錯誤。驗證 OS/ES 叢集政策和角色許可。</p> <p>0 表示沒有問題。1 表示交付失敗。</p>

## 交付至 OpenSearch 無伺服器

指標	描述
DeliveryToAmazonOpenSearchServerless.Bytes	<p>指定期間內索引至 OpenSearch 無伺服器的位元組數。</p> <p>單位：位元組</p>
DeliveryToAmazonOpenSearchServerless.DataFreshness	<p>Amazon 數據 Firehose 中最古老的記錄的年齡 ( 從進入 Amazon 數據 Firehose 到現在 )。任何超過此年齡的記錄都已傳遞至 OpenSearch 無伺服器。</p> <p>單位：秒</p>

指標	描述
DeliveryToAmazonOpenSearchServerless.Records	指定期間內，為 OpenSearch 無伺服器編製索引的記錄數目。  單位：計數
DeliveryToAmazonOpenSearchServerless.Success	成功編製索引之記錄的總和與已嘗試之記錄的總和。
DeliveryToS3.Bytes	在指定期間內，交付至 Amazon S3 的位元組數目。只有在您為所有文件啟用備份時，Amazon 資料 Firehose 才會發出此指標。  單位：計數
DeliveryToS3.DataFreshness	Amazon 數據 Firehose 中最古老的記錄的年齡（從進入 Amazon 數據 Firehose 到現在）。任何比此存留期還要久的記錄都已交付至 S3 儲存貯體。只有在您為所有文件啟用備份時，Amazon 資料 Firehose 才會發出此指標。  單位：秒
DeliveryToS3.Records	在指定期間內，交付至 Amazon S3 的記錄數目。只有在您為所有文件啟用備份時，Amazon 資料 Firehose 才會發出此指標。  單位：計數
DeliveryToS3.Success	所有 Amazon S3 Put 命令總和中的成功 Amazon S3 Put 命令總和。無論是僅針對失敗文件還是針對所有文件啟用備份，Amazon Data Firehose 一律會發出此指標。
DeliveryToAmazonOpenSearchServerless.AuthFailure	身分驗證和授權錯誤。驗證 OS/ES 叢集政策和角色許可。  0 表示沒有問題。1 表示發生驗證失敗。

指標	描述
DeliveryToAmazonOpenSearchServerless.DeliveryRejected	交付拒絕錯誤。驗證 OS/ES 叢集政策和角色許可。 0 表示沒有問題。1 表示交付失敗。

## 交付至 Amazon Redshift

指標	描述
DeliveryToRedshift.Bytes	在指定期間內，複製至 Amazon Redshift 的位元組數目。 單位：計數
DeliveryToRedshift.Records	在指定期間內，複製至 Amazon Redshift 的記錄數目。 單位：計數
DeliveryToRedshift.Success	所有 Amazon Redshift COPY 命令的總和中的成功 Amazon Redshift COPY 命令總和。
DeliveryToS3.Bytes	在指定期間內，交付至 Amazon S3 的位元組數目。 單位：位元組
DeliveryToS3.DataFreshness	Amazon 數據 Firehose 中最古老的記錄的年齡（從進入 Amazon 數據 Firehose 到現在）。任何比此存留期還要久的記錄都已交付至 S3 儲存貯體。 單位：秒
DeliveryToS3.Records	在指定期間內，交付至 Amazon S3 的記錄數目。 單位：計數
DeliveryToS3.Success	所有 Amazon S3 Put 命令總和中的成功 Amazon S3 Put 命令總和。

指標	描述
BackupToS3.Bytes	<p>在指定期間內，交付至 Amazon S3 供備份的位元組數目。Amazon 資料 Firehose 會在啟用備份至 Amazon S3 時發出此指標。</p> <p>單位：計數</p>
BackupToS3.DataFreshness	<p>Amazon 數據 Firehose 中最古老的記錄的年齡（從進入 Amazon 數據 Firehose 到現在）。任何比此存留期還要久的記錄都已交付至 Amazon S3 儲存貯體供備份。Amazon 資料 Firehose 會在啟用備份至 Amazon S3 時發出此指標。</p> <p>單位：秒</p>
BackupToS3.Records	<p>在指定期間內，交付至 Amazon S3 供備份的記錄數目。Amazon 資料 Firehose 會在啟用備份至 Amazon S3 時發出此指標。</p> <p>單位：計數</p>
BackupToS3.Success	<p>所有用於備份的 Amazon S3 Put 命令總和中的成功 Amazon S3 備份 Put 命令總和。Amazon 資料 Firehose 會在啟用備份至 Amazon S3 時發出此指標。</p>

## 交付至 Amazon S3

下表中的指標與 Amazon S3 是 Firehose 串流的主要目的地時交付有關。

指標	描述
DeliveryToS3.Bytes	<p>在指定期間內，交付至 Amazon S3 的位元組數目。</p> <p>單位：位元組</p>
DeliveryToS3.DataFreshness	<p>Amazon 數據 Firehose 中最古老的記錄的年齡（從進入 Amazon 數據 Firehose 到現在）。任何比此存留期還要久的記錄都已交付至 S3 儲存貯體。</p>

指標	描述
	單位：秒
DeliveryToS3.Records	在指定期間內，交付至 Amazon S3 的記錄數目。  單位：計數
DeliveryToS3.Success	所有 Amazon S3 Put 命令總和中的成功 Amazon S3 Put 命令總和。
BackupToS3.Bytes	在指定期間內，交付至 Amazon S3 供備份的位元組數目。Amazon Data Firehose 會在啟用備份時發出此指標 (只有在同時啟用資料轉換時才能執行)。  單位：計數
BackupToS3.DataFreshness	Amazon 數據 Firehose 中最古老的記錄的年齡 (從進入 Amazon 數據 Firehose 到現在)。任何比此存留期還要久的記錄都已交付至 Amazon S3 儲存貯體供備份。Amazon Data Firehose 會在啟用備份時發出此指標 (只有在同時啟用資料轉換時才能執行)。  單位：秒
BackupToS3.Records	在指定期間內，交付至 Amazon S3 供備份的記錄數目。Amazon Data Firehose 會在啟用備份時發出此指標 (只有在同時啟用資料轉換時才能執行)。  單位：計數
BackupToS3.Success	所有用於備份的 Amazon S3 Put 命令總和中的成功 Amazon S3 備份 Put 命令總和。Amazon Data Firehose 會在啟用備份時發出此指標 (只有在同時啟用資料轉換時才能執行)。



## 送到雪花

指標	描述
DeliveryToSnowflake.Bytes	<p>在指定時間段內傳遞給雪花的位元組數。</p> <p>單位：位元組</p>
DeliveryToSnowflake.DataFreshness	<p>Firehose 最古老紀錄的年齡（由進入到現在）。任何超過這個年齡的記錄都已經傳遞給雪花。</p> <p>單位：秒</p>
DeliveryToSnowflake.Records	<p>在指定期間內傳送至 Snowflake 的記錄數目。</p> <p>單位：計數</p>
DeliveryToSnowflake.Success	<p>已成功傳送的記錄總和，超過嘗試的記錄總和。</p>

## 交付至 Splunk

指標	描述
DeliveryToSplunk.Bytes	<p>在指定期間內，交付至 Splunk 的位元組數目。</p> <p>單位：位元組</p>
DeliveryToSplunk.DataAckLatency	<p>Amazon 資料 Firehose 傳送資料後，接收 Splunk 確認所需的大約持續時間。這項指標的上升趨勢或下降趨勢比絕對約略值更有用。上升趨勢可能表示 Splunk 索引器的索引和確認速度較慢。</p> <p>單位：秒</p>
DeliveryToSplunk.DataFreshness	<p>Amazon 數據 Firehose 中最古老的記錄的年齡（從進入 Amazon 數據 Firehose 到現在）。任何比此存留期還要久的記錄都已交付至 Splunk。</p>

指標	描述
	單位：秒
DeliveryToSplunk.Records	在指定期間內，交付至 Splunk 的記錄數目。 單位：計數
DeliveryToSplunk.Success	成功編製索引之記錄的總和與已嘗試之記錄的總和。
DeliveryToS3.Success	所有 Amazon S3 Put 命令總和中的成功 Amazon S3 Put 命令總和。啟用備份至 Amazon S3 時，會發出此指標。
BackupToS3.Bytes	在指定期間內，交付至 Amazon S3 供備份的位元組數目。Amazon 資料 Firehose 會在設定為備份所有文件時發出此指標。 單位：計數
BackupToS3.DataFreshness	Amazon 數據 Firehose 中最古老的記錄的年齡（從進入 Amazon 數據 Firehose 到現在）。任何比此存留期還要久的記錄都已交付至 Amazon S3 儲存貯體供備份。Amazon 資料 Firehose 會在設定為備份所有文件時發出此指標。 單位：秒
BackupToS3.Records	在指定期間內，交付至 Amazon S3 供備份的記錄數目。Amazon 資料 Firehose 會在設定為備份所有文件時發出此指標。 單位：計數
BackupToS3.Success	所有用於備份的 Amazon S3 Put 命令總和中的成功 Amazon S3 備份 Put 命令總和。Amazon 資料 Firehose 會在設定為備份所有文件時發出此指標。

## 交付至 HTTP 端點

指標	描述
<code>DeliveryToHttpEndpoint.Bytes</code>	成功交付至 HTTP 端點的位元組數目。 單位：位元組
<code>DeliveryToHttpEndpoint.Records</code>	成功交付至 HTTP 端點的記錄數目。 單位：計數
<code>DeliveryToHttpEndpoint.DataFreshness</code>	Amazon 數據 Firehose 件中最古老的記錄的年齡。 單位：秒
<code>DeliveryToHttpEndpoint.Success</code>	所有到 HTTP 端點之成功資料交付請求的總和 單位：計數
<code>DeliveryToHttpEndpoint.ProcessedBytes</code>	嘗試處理的位元組數目，包括重試。
<code>DeliveryToHttpEndpoint.ProcessedRecords</code>	包括重試在內的嘗試記錄數目。

## 資料導入指標

### 主題

- [透過 Kinesis Data Streams 進行資料擷取](#)
- [透過 Direct PUT 導入資料](#)
- [MSK 的資料擷取](#)

## 透過 Kinesis Data Streams 進行資料擷取

指標	描述
<code>DataReadFromKinesisStream.Bytes</code>	<p>當資料來源為 Kinesis 資料串流時，此指標指出從該資料串流讀取的位元組數目。此數目包含由於容錯移轉的重新讀取。</p> <p>單位：位元組</p>
<code>DataReadFromKinesisStream.Records</code>	<p>當資料來源為 Kinesis 資料串流時，此指標指出從該資料串流讀取的記錄數目。此數目包含由於容錯移轉的重新讀取。</p> <p>單位：計數</p>
<code>ThrottledDescribeStream</code>	<p>當資料來源為 Kinesis 資料串流時，調節 <code>DescribeStream</code> 操作的總次數。</p> <p>單位：計數</p>
<code>ThrottledGetRecords</code>	<p>當資料來源為 Kinesis 資料串流時，調節 <code>GetRecords</code> 操作的總次數。</p> <p>單位：計數</p>
<code>ThrottledGetShardIterator</code>	<p>當資料來源為 Kinesis 資料串流時，調節 <code>GetShardIterator</code> 操作的總次數。</p> <p>單位：計數</p>

## 透過 Direct PUT 導入資料

指標	描述
<code>BackupToS3.Bytes</code>	<p>在指定期間內，交付至 Amazon S3 供備份的位元組數目。當 Amazon S3 或亞馬遜紅移目的地啟用資料轉換時，Amazon 資料 Firehose 會發出此指標。</p>

指標	描述
	單位：位元組
BackupToS3.DataFreshness	Amazon 數據 Firehose 中最古老的記錄的年齡 ( 從進入 Amazon 數據 Firehose 到現在 )。任何比此存留期還要久的記錄都已交付至 Amazon S3 儲存貯體供備份。當 Amazon S3 或亞馬遜紅移目的地啟用資料轉換時，Amazon 資料 Firehose 會發出此指標。  單位：秒
BackupToS3.Records	在指定期間內，交付至 Amazon S3 供備份的記錄數目。當 Amazon S3 或亞馬遜紅移目的地啟用資料轉換時，Amazon 資料 Firehose 會發出此指標。  單位：計數
BackupToS3.Success	所有用於備份的 Amazon S3 Put 命令總和中的成功 Amazon S3 備份 Put 命令總和。當 Amazon S3 或亞馬遜紅移目的地啟用資料轉換時，Amazon 資料 Firehose 會發出此指標。
BytesPerSecondLimit	Firehose 串流在節流之前可擷取的目前每秒位元組數目上限。若要請求增加此限制，請前往 <a href="#">AWS 支援中心</a> 並選擇建立案例，然後選擇服務配額提升。
DataReadFromKinesisStream.Bytes	當資料來源為 Kinesis 資料串流時，此指標指出從該資料串流讀取的位元組數目。此數目包含由於容錯移轉的重新讀取。  單位：位元組
DataReadFromKinesisStream.Records	當資料來源為 Kinesis 資料串流時，此指標指出從該資料串流讀取的記錄數目。此數目包含由於容錯移轉的重新讀取。  單位：計數

指標	描述
DeliveryToAmazonOpenSearchService.Bytes	指定期間內索引至 OpenSearch Service 的位元組數目。 單位：位元組
DeliveryToAmazonOpenSearchService.DataFreshness	Amazon 數據 Firehose 中最古老的記錄的年齡 ( 從進入 Amazon 數據 Firehose 到現在 )。任何超過此年齡的記錄都已交付給 OpenSearch 服務。 單位：秒
DeliveryToAmazonOpenSearchService.Records	指定期間內為 OpenSearch Service 編製索引的記錄數目。 單位：計數
DeliveryToAmazonOpenSearchService.Success	成功編製索引之記錄的總和與已嘗試之記錄的總和。
DeliveryToRedshift.Bytes	在指定期間內，複製至 Amazon Redshift 的位元組數目。 單位：位元組
DeliveryToRedshift.Records	在指定期間內，複製至 Amazon Redshift 的記錄數目。 單位：計數
DeliveryToRedshift.Success	所有 Amazon Redshift COPY 命令的總和中的成功 Amazon Redshift COPY 命令總和。
DeliveryToS3.Bytes	在指定期間內，交付至 Amazon S3 的位元組數目。 單位：位元組
DeliveryToS3.DataFreshness	Amazon 數據 Firehose 中最古老的記錄的年齡 ( 從進入 Amazon 數據 Firehose 到現在 )。任何比此存留期還要久的記錄都已交付至 S3 儲存貯體。 單位：秒

指標	描述
DeliveryToS3.Records	在指定期間內，交付至 Amazon S3 的記錄數目。  單位：計數
DeliveryToS3.Success	所有 Amazon S3 Put 命令總和中的成功 Amazon S3 Put 命令總和。
DeliveryToSplunk.Bytes	在指定期間內，交付至 Splunk 的位元組數目。  單位：位元組
DeliveryToSplunk.DataAckLatency	Amazon 資料 Firehose 傳送資料後，接收 Splunk 確認所需的大約持續時間。這項指標的上升趨勢或下降趨勢比絕對約略值更有用。上升趨勢可能表示 Splunk 索引器的索引和確認速度較慢。  單位：秒
DeliveryToSplunk.DataFreshness	Amazon 數據 Firehose 中最古老的記錄的年齡（從進入 Amazon 數據 Firehose 到現在）。任何比此存留期還要久的記錄都已交付至 Splunk。  單位：秒
DeliveryToSplunk.Records	在指定期間內，交付至 Splunk 的記錄數目。  單位：計數
DeliveryToSplunk.Success	成功編製索引之記錄的總和與已嘗試之記錄的總和。
IncomingBytes	指定期間內成功嵌入傳遞串流的位元組數目。當資料擷取超過其中一個交付串流限制時，可能會進行限流。限流的資料將不會計算在 IncomingBytes 內。  單位：位元組

指標	描述
IncomingPutRequests	指定期間內成功 PutRecord 和 PutRecordBatch 要求的數目。  單位：計數
IncomingRecords	指定期間內成功擷取至傳遞串流的記錄數目。當資料擷取超過其中一個交付串流限制時，可能會進行限流。限流的資料將不會計算在 IncomingRecords 內。  單位：計數
KinesisMillisBehindLatest	當資料來源為 Kinesis 資料串流時，這項指標表示上次讀取記錄落後 Kinesis 資料串流中最新記錄的毫秒數。  單位：毫秒
RecordsPerSecondLimit	Firehose 串流在節流之前每秒可擷取的記錄數目上限。  單位：計數
ThrottledRecords	因為資料擷取超過其中一個 Firehose 串流限制而受到限制的記錄數目。  單位：計數

## MSK 的資料擷取

指標	描述
DataReadFromSource .Records	從來源 Kafka 主題讀取的記錄數目。  單位：計數
DataReadFromSource.Bytes	從來源 Kafka 主題讀取的位元組數目。  單位：位元組
SourceThrottled.Delay	來源 Kafka 叢集從來源 Kafka 主題傳回記錄的延遲時間。



指標	描述
	單位：毫秒
BytesPerSecondLimit	Firehose 將從來源 Kafka 主題的每個分割區讀取的目前輸送量限制。 單位：位元組/秒
KafkaOffsetLag	Firehose 從來源 Kafka 主題所讀取記錄的最大偏移量與來源 Kafka 主題可用記錄的最大偏移量之間的差異。 單位：計數
FailedValidation.Records	記錄驗證失敗的記錄數目。 單位：計數
FailedValidation.Bytes	記錄驗證失敗的位元組數目。 單位：位元組
DataReadFromSource .Backpressured	指出 Firehose 串流在從來源磁碟分割讀取記錄時延遲，這可能是因為 BytesPerSecondLimit 每個分割區已超過或正常傳送流程緩慢或已停止 單位：布林

## API 層 CloudWatch 級測量結果

AWS/Firehose 命名空間包含下列 API 層級指標。

指標	描述
DescribeDeliveryStream.Latency	每個 DescribeDeliveryStream 操作所需的時間，這是對指定的期間進行測量。 單位：毫秒

指標	描述
DescribeDeliveryStream.Requests	DescribeDeliveryStream 請求的總數。 單位：計數
ListDeliveryStreams.Latency	每個 ListDeliveryStreams 操作所需的時間，這是對指定的期間進行測量。 單位：毫秒
ListDeliveryStreams.Requests	ListFirehose 請求的總數。 單位：計數
PutRecord.Bytes	在指定期間內放入 Firehose 串流 PutRecord 的位元組數目。 單位：位元組
PutRecord.Latency	每個 PutRecord 操作所需的時間，這是對指定的期間進行測量。 單位：毫秒
PutRecord.Requests	PutRecord 請求的總數，它等於來自 PutRecord 操作的總數。 單位：計數
PutRecordBatch.Bytes	在指定期間內放入 Firehose 串流 PutRecordBatch 的位元組數目。 單位：位元組
PutRecordBatch.Latency	每個 PutRecordBatch 操作所需的時間，這是對指定的期間進行測量。 單位：毫秒

指標	描述
PutRecordBatch.Records	來自 PutRecordBatch 操作的記錄總數。 單位：計數
PutRecordBatch.Requests	PutRecordBatch 請求的總數。 單位：計數
PutRequestsPerSecondLimit	Firehose 串流在節流之前每秒可處理的最大 put 要求數。這個數字包括 PutRecord 和 PutRecordBatch 請求。 單位：計數
ThrottledDescribeStream	當資料來源為 Kinesis 資料串流時，調節 DescribeStream 操作的總次數。 單位：計數
ThrottledGetRecords	當資料來源為 Kinesis 資料串流時，調節 GetRecords 操作的總次數。 單位：計數
ThrottledGetShardIterator	當資料來源為 Kinesis 資料串流時，調節 GetShardIterator 操作的總次數。 單位：計數
UpdateDeliveryStream.Latency	每個 UpdateDeliveryStream 操作所需的時間，這是對指定的期間進行測量。 單位：毫秒
UpdateDeliveryStream.Requests	UpdateDeliveryStream 請求的總數。 單位：計數

## 資料轉換 CloudWatch 指標

如果啟用 Lambda 的資料轉換，AWS/Firehose 命名空間包含下列指標。

指標	描述
ExecuteProcessing.Duration	Firehose 執行的每個 Lambda 函數叫用所需的時間。 單位：毫秒
ExecuteProcessing.Success	成功 Lambda 函數調用的總和與 Lambda 函數調用總數之比。
SucceedProcessing.Records	在指定的期間內，成功處理的記錄數目。 單位：計數
SucceedProcessing.Bytes	在指定的期間內，成功處理的位元組記錄數目。 單位：位元組

## CloudWatch 日誌解壓縮測量結果

如果針對 CloudWatch 記錄傳送啟用解壓縮，AWS/Firehose 命名空間會包含下列度量。

指標	描述
OutputDecompressedBytes.Success	成功解壓縮資料 (位元組) 單位：位元組
OutputDecompressedBytes.Failed	失敗的解壓縮資料 (位元組) 單位：位元組
OutputDecompressedRecords.Success	解壓縮成功的記錄數目

指標	描述
	單位：計數
OutputDecompressedRecords.Failed	失敗的解壓縮記錄數目 單位：計數

## 格式轉換 CloudWatch 量度

如果啟用格式轉換，AWS/Firehose 命名空間包含下列指標。

指標	描述
SucceedConversion.Records	已成功轉換的記錄數目。 單位：計數
SucceedConversion.Bytes	已成功轉換的記錄大小。 單位：位元組
FailedConversion.Records	無法轉換的記錄數目。 單位：計數
FailedConversion.Bytes	無法轉換的記錄大小。 單位：位元組

## 伺服器端加密 (SSE) CloudWatch 量度

AWS/Firehose 命名空間包含下列與 SSE 相關的指標。

指標	描述
KMSKeyAccessDenied	服務遇到 KMSAccessDeniedException 交付串流的次數。

指標	描述
	單位：計數
KMSKeyDisabled	服務遇到 <code>KMSDisabledException</code> 交付串流的次數。 單位：計數
KMSKeyInvalidState	服務遇到 <code>KMSInvalidStateException</code> 交付串流的次數。 單位：計數
KMSKeyNotFound	服務遇到 <code>KMSNotFoundException</code> 交付串流的次數。 單位：計數

## Amazon 數據 Firehose 的尺寸

若要依 Firehose 串流篩選量度，請使用 `DeliveryStreamName` 維度。

## Amazon 數據 Firehose 使用指標

您可以使用 CloudWatch 使用量度來提供您帳戶資源使用情況的可見度。使用這些指標，在 CloudWatch 圖形和儀表板上視覺化您目前的服務使用情況。

服務配額使用量測量結果位於 `AWS/Usage` 命名空間中，每分鐘收集一次。

目前，此命名空間中唯一 CloudWatch 發佈的測量結果名稱是 `ResourceCount`。此指標會與維度 `Service`、`Class`、`Type` 和 `Resource` 一起發佈。

指標	描述
ResourceCount	您的帳戶中正在執行的特定資源數量。資源由與指標相關聯的維度定義。  此指標最有用的統計資料是 <code>MAXIMUM</code> ，它表示 1 分鐘期間內使用的最大資源數量。

下列維度可用來調整 Amazon 資料 Firehose 發佈的使用量指標。

維度	描述
Service	包含資源的 AWS 服務名稱。對於 Amazon 資料 Firehose 使用量指標，此維度的值為 Firehose。
Class	正在追蹤的資源類別。Amazon 資料 Firehose API 使用量指標使用此維度的 None 值為。
Type	正在追蹤的資源類型。目前，當服務維度為 Firehose 時，類型的唯一有效值為 Resource。
Resource	AWS 資源的名稱。目前，當服務維度為 Firehose 時，資源的唯一有效值為 DeliveryStreams。

## 訪問 CloudWatch Amazon 數據 Firehose 的指標

您可以使用 CloudWatch 主控台、命令列或 CloudWatch API 監控 Amazon 資料 Firehose 的指標。以下程序將說明如何使用這些不同的方法來存取指標。

使用 CloudWatch 主控台存取指標

1. [請在以下位置開啟 CloudWatch 主控台。](https://console.aws.amazon.com/cloudwatch/) <https://console.aws.amazon.com/cloudwatch/>
2. 在導覽列上，選擇一個區域。
3. 在導覽窗格中，選擇 指標。
4. 選擇 Firehose (Firehose) 命名空間。
5. 選擇「Firehose 串流量度」或「Firehose 指標」。
6. 選擇欲新增至圖表的指標。

若要使用存取量度 AWS CLI

使用 [列表公制](#)和 [get-metric-statistics](#) 命令。

```
aws cloudwatch list-metrics --namespace "AWS/Firehose"
```

```
aws cloudwatch get-metric-statistics --namespace "AWS/Firehose" \
```

```
--metric-name DescribeDeliveryStream.Latency --statistics Average --period 3600 \  
--start-time 2017-06-01T00:00:00Z --end-time 2017-06-30T00:00:00Z
```

## 使 CloudWatch 用日誌監控 Amazon 數據 Firehose

Amazon 資料 Firehose 與 Amazon CloudWatch 日誌整合，因此當 Lambda 呼叫資料轉換或資料交付失敗時，您可以檢視特定的錯誤日誌。您可以在建立 Firehose 串流時啟用 Amazon 資料 Firehose 錯誤記錄。

如果您在 Amazon 資料 Firehose 主控台中啟用 Amazon 資料 Firehose 錯誤記錄，則會代表您為 Firehose 串流建立日誌群組和對應的日誌串流。記錄群組名稱的格式為 `/aws/kinesisfirehose/delivery-stream-name`，其中 *delivery-stream-name* 是對應 Firehose 串流的名稱。DestinationDelivery 是建立並用來記錄與傳遞至主要目的地相關的任何錯誤的記錄資料流。只有在為目的地啟用 S3 備份時，才會建立另一個名為 BackupDelivery 的日誌串流。BackupDelivery 日誌串流可用來記錄與交付至 S3 備份相關的任何錯誤。

例如，如果您建立一個以 Amazon Redshift 做為目的地的 Firehose 串流 `MyStream ""`，並啟用 Amazon 資料 Firehose 錯誤記錄，則會代表您建立下列項目：一個名為的日誌群組 `aws/kinesisfirehose/MyStream` 和兩個名為的日誌串流。DestinationDelivery BackupDelivery 在此範例中，DestinationDelivery 將用於記錄與交付到 Amazon Redshift 目的地以及中繼 S3 目的地相關的任何錯誤。如果啟用 S3 備份，BackupDelivery 將用於記錄與交付到 S3 備份儲存貯體相關的任何錯誤。

您可以透過 API 或使用 `CloudWatchLoggingOptions` 組態啟 AWS CloudFormation 用 Amazon 資料 Firehose 錯誤記錄。AWS CLI 若要這樣做，請事先建立日誌群組和日誌資料流。我們建議您只為 Amazon 資料 Firehose 錯誤記錄保留該日誌群組和日誌串流。同時，請確認相關 IAM 政策具備 `"logs:putLogEvents"` 許可。如需詳細資訊，請參閱 [使用 Amazon 資料 Firehose 控制存取](#)。

請注意，Amazon 資料 Firehose 不保證所有交付錯誤日誌都會傳送至 CloudWatch 日誌。在交付失敗率很高的情況下，Amazon Data Firehose 會先取樣交付錯誤日誌，然後再將它們傳送到 CloudWatch 日誌。

傳送至 CloudWatch 記錄檔的錯誤記錄需支付額外費用。如需詳細資訊，請參閱 [Amazon CloudWatch 定價](#)。

### 目錄

- [資料交付錯誤](#)



## 資料交付錯誤

以下是每個 Amazon 資料 Firehose 目的地的資料傳送錯誤碼和訊息清單。每個錯誤訊息亦建議修復問題應採取的適當動作。

### 錯誤

- [Amazon S3 資料交付錯誤](#)
- [Amazon Redshift 資料交付錯誤](#)
- [雪花資料傳送錯誤](#)
- [Splunk 資料交付錯誤](#)
- [ElasticSearch 資料傳送錯誤](#)
- [HTTPS 端點資料交付錯誤](#)
- [Amazon OpenSearch 服務數據交付錯誤](#)
- [Lambda 調用錯誤](#)
- [Kinesis 調用錯誤](#)
- [Kinesis DirectPut 調用錯誤](#)
- [AWS Glue 調用錯誤](#)
- [DataFormatConversion 調用錯誤](#)

### Amazon S3 資料交付錯誤

亞馬遜數據 Firehose 件可以將以下與 Amazon S3 相關的錯誤發送到日誌。 CloudWatch

錯誤程式碼	錯誤訊息與資訊
S3.KMS.NotFoundException	「找不到提供的 AWS KMS 密鑰。如果您正在使用您認為具有正確角色的有效 AWS KMS 密鑰的內容，請檢查附加 AWS KMS 密鑰的帳戶是否存在問題。」
S3.KMS.RequestLimitExceeded	「在嘗試加密 S3 物件時，KMS 每秒請求超過限制。請增加每秒請求限制。」  如需詳細資訊，請參閱 AWS Key Management Service 開發人員指南中的 <a href="#">限制</a> 。

錯誤程式碼	錯誤訊息與資訊
S3.AccessDenied	「存取遭拒。確保所提供 IAM 角色的信任政策允許 Amazon Data Firehose 擔任該角色，而且存取政策允許存取 S3 儲存貯體。」
S3.AccountProblem	「您的 AWS 帳戶存在問題，導致操作無法成功完成。聯絡 AWS 支援。」
S3.AllAccessDisabled	「所提供的帳戶存取已停用。聯絡 S AWS support。」
S3.InvalidPayer	「所提供的帳戶存取已停用。聯絡 S AWS support。」
S3.NotSignedUp	「帳戶未註冊 Amazon S3。請註冊帳戶或使用不同的帳戶。」
S3.NoSuchBucket	「指定的儲存貯體不存在。建立儲存貯體或使用另一現有的儲存貯體。」
S3.MethodNotAllowed	「指定的方法不得使用此資源。請修改儲存貯體的政策，允許正確的 Amazon S3 操作許可。」
InternalError	「嘗試傳輸資料時發生內部錯誤。交付將被重試；如果錯誤仍然存在，則將報告 AWS 給解決。」
S3.KMS.KeyDisabled	「提供的 KMS 金鑰已停用。啟用金鑰或使用其他金鑰。」
S3.KMS.InvalidStateException	「提供的 KMS 金鑰處於無效狀態。請使用不同的金鑰。」
KMS.InvalidStateException	「提供的 KMS 金鑰處於無效狀態。請使用不同的金鑰。」
KMS.DisabledException	「提供的 KMS 金鑰已停用。請啟用金鑰或使用其他金鑰。」
S3.SlowDown	「指定儲存貯體的 Put 請求率太高。增加 Firehose 串流緩衝區大小或減少來自其他應用程式的 put 要求。」

錯誤程式碼	錯誤訊息與資訊
S3.SubscriptionRequired	「呼叫 S3 時存取被拒絕。確保傳入的 IAM 角色和 KMS 金鑰 (如有提供) 具有 Amazon S3 訂閱。」
S3.InvalidToken	「提供的記號格式錯誤或以其他方式無效。請檢查提供的憑證。」
S3.KMS.KeyNotConfigured	「KMS 金鑰未設定。設定您的 KMS MasterKey ID，或停用 S3 儲存貯體的加密。」
S3.KMS.AsymmetricCMKNotSupported	「Amazon S3 支援對稱 CMK。您無法使用非對稱 CMK 來加密 Amazon S3 中的資料。要獲取 CMK 的類型，請使用 KMS DescribeKey 操作。」
S3.IllegalLocationConstraintException	「Firehose 目前使用 s3 全球端點將資料交付到設定的 s3 儲存貯體。已設定 s3 儲存貯體的區域不支援 s3 全球端點。請在與 s3 儲存貯體相同的區域中建立 Firehose 串流，或在支援全球端點的區域中使用 s3 儲存貯體。」
S3.InvalidPrefixConfigurationException	「用於時間戳記評估的自定義 s3 字首無效。檢查您的 s3 字首包含一年中當前日期和時間的有效運算式。」
DataFormatConversion.MalformedData	「記號之間發現非法字符。」

## Amazon Redshift 資料交付錯誤

Amazon 數據 Firehose 件可以將以下與亞馬遜 RedShift 相關的錯誤發送到日誌。CloudWatch

錯誤程式碼	錯誤訊息與資訊
Redshift.TableNotFound	「找不到欲載入資料的表格。請確認該表格已存在。」

錯誤程式碼	錯誤訊息與資訊
	找不到應從 S3 複製至 Amazon Redshift 內的資料表目的地。請注意，Amazon 數據 Firehose 不會創建 Amazon Redshift 表，如果它不存在。
Redshift. SyntaxError	「COPY 命令包含語法錯誤。請重試命令。」
Redshift. AuthenticationFailed	「所提供的使用者名稱和密碼驗證失敗。請提供有效的使用者名稱和密碼。」
Redshift. AccessDenied	「存取遭拒。確保所提供 IAM 角色的信任政策允許 Amazon 資料 Firehose 擔任該角色。」
Redshift. S3BucketAccessDenied	「COPY 命令無法存取 S3 儲存貯體。請確認所提供 IAM 角色的存取政策允許存取 S3 儲存貯體。」
Redshift. DataLoadFailed	「載入資料至表格失敗。請檢查 STL_LOAD_ERRORS 系統表格以了解詳細資訊。」
Redshift. ColumnNotFound	「COPY 命令的欄位不存在於表格中。請指定有效的欄位名稱。」
Redshift. DatabaseNotFound	「找不到 Amazon Redshift 目的地設定值中指定的資料庫或 JDBC URL。請指定有效的資料庫名稱。」
Redshift. IncorrectCopyOptions	<p>「所提供的 COPY 選項互相衝突或過多。部分選項不相容於特定組合。請查看 COPY 命令參考資料以取得詳細資訊。」</p> <p>如需詳細資訊，請參閱《Amazon Redshift 資料庫開發人員指南》中的 <a href="#">Amazon Redshift COPY 命令</a>。</p>
Redshift. MissingColumn	「表格結構描述將欄位定義為無預設值的非空白值，且未包含在欄位清單中。請排除此欄位、確認載入的資料一律提供此欄位的值，或新增此資料表的預設值至 Amazon Redshift 結構描述。」

錯誤程式碼	錯誤訊息與資訊
Redshift. ConnectionFailed	「連線至指定的 Amazon Redshift 叢集失敗。請確定安全設定允許 Amazon 資料 Firehose 連線、Amazon Redshift 目的地組態或 JDBC 網址中指定的叢集或資料庫正確無誤，且叢集可用。」
Redshift. ColumnMismatch	「COPY 命令的 jsonpaths 數量和目的地表格的欄位數量應相符。請重試命令。」
Redshift. IncorrectOrMissingRegion	「Amazon Redshift 嘗試使用錯誤的區域的端點來存取 S3 儲存貯體。請於 COPY 命令選項中指定正確的區域值，或確認 S3 儲存貯體與 Amazon Redshift 資料庫位於相同區域。」
Redshift. IncorrectJsonPathsFile	「所提供的 jsonpaths 檔案未支援 JSON 格式。請重試命令。」
Redshift. MissingS3File	「Amazon Redshift 所需的一個或多個 S3 檔案已自 S3 儲存貯體移除。請檢查 S3 儲存貯體的政策來移除 S3 檔案自動刪除。」
Redshift. InsufficientPrivilege	「使用者不具備將資料載入表格的許可。請檢查 Amazon Redshift INSERT 許可的使用者許可。」
Redshift. ReadOnlyCluster	「無法執行該查詢，因系統正在調整規模。請稍後再嘗試查詢。」
Redshift. DiskFull	「資料無法載入，因為磁碟已滿。請增加 Amazon Redshift 叢集的容量，或刪除未使用的資料以釋放磁碟空間。」
InternalError	「嘗試傳輸資料時發生內部錯誤。交付將被重試; 如果錯誤仍然存在，則將報告 AWS 給解決。」
Redshift. ArgumentNotSupported	「COPY 命令包含不支持的選項。」

錯誤程式碼	錯誤訊息與資訊
Redshift. AnalyzeTableAccess Denied	「存取遭拒。從 S3 複製到 Redshift 失敗，因為分析資料表只能由資料表或資料庫擁有者完成。」
Redshift. SchemaNotFound	「找不到 Amazon Redshift 目標配置中指定 DataTableName 的模式。請指定有效的結構描述名稱。」
Redshift. ColumnSpecifiedMoreThanOnce	「在列列表中指定了一個以上的列。確保刪除重複的列。」
Redshift. ColumnNotNullWithoutDefault	「有一個無預設值的非空白欄位，且未包含在欄位清單中。請確認此類欄位包含在欄位清單中。」
Redshift. IncorrectBucketRegion	「Redshift 嘗試使用與叢集不同區域的儲存貯體。請指定與叢集同一區域的儲存貯體。」
Redshift. S3SlowDown	「S3 的高請求率。降低速率以避免限流。」
Redshift. InvalidCopyOptionForJson	「請將 auto 或有效的 S3 路徑用於 json 複製選項。」
Redshift. InvalidCopyOptionJSONPathFormat	「COPY 失敗，出現錯誤 \"無效的 JSONPath 格式。陣列索引超出範圍\"。請糾正 JSONPath 運算式。」

錯誤程式碼	錯誤訊息與資訊
Redshift. InvalidCopyOptionRBACaclNotAllowed	「COPY 失敗，出現錯誤 \「在未啟用許可傳播時，無法使用 RBAC acl 架構。」
Redshift. DiskSpaceQuotaExceeded	「由於磁盤空間配額超出，交易中止。釋放磁碟空間或請求提升結構描述的配額。」
Redshift. ConnectionsLimitExceeded	「超出使用者的連線限制。」
Redshift. SslNotSupported	「由於伺服器不支援 SSL，因此與指定的 Amazon Redshift 叢集的連線失敗。請檢查您的叢集設定。」
Redshift. HoseNotFound	「hose 已被刪除。請檢查您的 hose 狀態。」
Redshift. Delimiter	「copyCommand 中的 copyOptions 分隔符號無效。確保它屬於單一字元。」
Redshift. QueryCancelled	「使用者已經取消了 COPY 操作。」
Redshift. CompressionMismatch	「hose 設定為 UNCOMPRESSED，但 copyOption 包括壓縮格式。」
Redshift. EncryptionCredentials	「ENCRYPTED 選項需要以下格式的憑證：'aws_iam_role=...;master_symmetric_key=...' or 'aws_access_key_id=...;aws_secret_access_key=...[;token=...];master_symmetric_key=...'」

錯誤程式碼	錯誤訊息與資訊
Redshift.InvalidCopyOptions	「無效的 COPY 組態選項。」
Redshift.InvalidMessageFormat	「COPY 命令包含無效字符。」
Redshift.TransactionIdLimitReached	「已達交易 ID 限制。」
Redshift.DestinationRemoved	「請確認是否存在 redshift 目的地，並且在 Firehose 組態中已正確設定。」
Redshift.OutOfMemory	「Redshift 叢集的記憶體不足。請確保叢集有足夠的容量。」
Redshift.Cannot Fork Process	「Redshift 叢集的記憶體不足。請確保叢集有足夠的容量。」
Redshift.SslFailure	「握手期間 SSL 連線已關閉。」
Redshift.Resize	「Redshift 叢集正在調整大小。當叢集調整大小時，Firehose 將無法傳送資料。」
Redshift.ImproperQualifiedName	「限定名稱不正確 (帶點的名稱太多)。」
Redshift.InvalidJsonPathFormat	「無效的 JSONPath 格式。」



錯誤程式碼	錯誤訊息與資訊
Redshift. TooManyConnections Exception	「與 Redshift 的連線太多。」
Redshift. PSQLErrorException	「PS 從 Redshift QLErrorException 觀察。」
Redshift. DuplicateSecondsSpecification	「以日期/時間格式中的秒數規格重複。」
Redshift. RelationCouldNotBeOpened	「遇到 Redshift 錯誤，無法打開關係。檢查指定資料庫的 Redshift 日誌。」
Redshift. TooManyClients	「從 Redshift 遇到太多客戶端例外狀況。如果有多個生產者同時寫入資料庫，請重新檢視資料庫的連線上限。」

## 雪花資料傳送錯誤

Firehose 可以將下列與雪花相關的錯誤傳送至記錄檔。 CloudWatch

錯誤程式碼	錯誤訊息與資訊
Snowflake. .InvalidUrl	「Firehose 無法連接到雪花。請確保在雪花目標配置中正確指定了帳戶 URL。」
Snowflake. .InvalidUser	「Firehose 無法連接到雪花。請確保在雪花目標配置中正確指定了用戶。」
Snowflake. .InvalidRole	「指定的雪花角色不存在或未被授權。請確定角色已授與指定的使用者」

錯誤程式碼	錯誤訊息與資訊
Snowflake .InvalidTable	「提供的表不存在或未授權」
Snowflake .InvalidSchema	「提供的模式不存在或未授權」
Snowflake .InvalidDatabase	「提供的數據庫不存在或未被授權」
Snowflake .InvalidPrivateKeyOrPassphrase	「指定的私鑰或密碼無效。請注意，提供的私鑰應該是有效的 PEM RSA 私鑰」
Snowflake .MissingColumns	「由於輸入有效負載中缺少列，插入請求被拒絕。確保為所有不可為空的列指定了值」
Snowflake .ExtraColumns	「插入請求被拒絕，由於額外的列。不應指定表中不存在的列」
Snowflake .InvalidInput	「由於輸入格式無效，交付失敗。確保提供的輸入有效負載採用 JSON 格式可接受」
Snowflake .IncorrectValue	「由於輸入有效負載中的數據類型不正確，傳遞失敗。確保輸入有效負載中指定的 JSON 值符合雪花表定義中聲明的數據類型」

## Splunk 資料交付錯誤

Amazon 數據 Firehose 可以將以下與 Splunk 相關的錯誤發送到日誌。CloudWatch

錯誤程式碼	錯誤訊息與資訊
<code>Splunk.ProxyWithoutStickySessions</code>	「如果您在 Amazon 資料 Firehose 和 HEC 節點之間有代理伺服器 (ELB 或其他)，則必須啟用黏性工作階段才能支援 HEC ACK。」
<code>Splunk.DisabledToken</code>	「HEC 符記未啟用。請啟用符記以將資料交付至 Splunk。」
<code>Splunk.InvalidToken</code>	「無效的 HEC 符記。使用有效的 HEC 令牌更新 Amazon 數據 Firehose 件。」
<code>Splunk.InvalidDataFormat</code>	「資料格式化不正確。欲理解如何正確格式化原生或事件 HEC 端點的資料，請參閱 <a href="#">Splunk Event Data</a> 。」
<code>Splunk.InvalidIndex</code>	「HEC 符記或輸入的索引無效。請檢查您的索引設定並再試一次。」
<code>Splunk.ServerError</code>	「將資料傳送至 Splunk 失敗，因為 HEC 節點的伺服器出現錯誤。如果 Amazon 數據 Firehose 中的重試持續時間大於 0，Amazon 數據 Firehose 將重試發送數據。如果所有重試都失敗，Amazon 數據 Firehose 將數據備份到 Amazon S3。」
<code>Splunk.DisabledAck</code>	「HEC 符記的 indexer 確認未啟用。請啟用 indexer 確認並再試一次。如需詳細資訊，請參閱 <a href="#">Enable indexer acknowledgement</a> 。」
<code>Splunk.AckTimeout</code>	「HEC 確認逾時過期之前，未收到 HEC 的確認。儘管確認逾時，資料仍可能已成功索引至 Splunk。亞馬遜數據 Firehose 件備份到確認超時過期的 Amazon S3 數據中。」
<code>Splunk.MaxRetriesFailed</code>	「無法傳送資料至 Splunk，或無法接收確認。請檢查您的 HEC 運作狀態並再試一次。」
<code>Splunk.ConnectionTimeout</code>	「連線至 Splunk 逾時。這可能是暫時性錯誤，請求將會重試。如果所有重試都失敗，Amazon 數據 Firehose 將數據備份到 Amazon S3。」

錯誤程式碼	錯誤訊息與資訊
Splunk.InvalidEndpoint	「無法連線至 HEC 終端節點。確保 HEC 端點 URL 有效且可從 Amazon 數據 Firehose 件訪問。」
Splunk.ConnectionClosed	「無法將資料傳送至 Splunk，因為連線失敗。這可能是暫時性錯誤。增加 Amazon 資料 Firehose 組態中的重試持續時間可能會防止此類暫時性故障。」
Splunk.SSLUnverified	「無法連線至 HEC 終端節點。主機與同儕提供的憑證不符。請確認該憑證和主機有效。」
Splunk.SSLHandshake	「無法連線至 HEC 終端節點。請確認該憑證和主機有效。」
Splunk.URLNotFound	「在 Splunk 伺服器上找不到請求的 URL。請檢查 Splunk 叢集並確認其設定正確。」
Splunk.ServerError.ContentTooLarge	「傳送至 Splunk 的資料失敗，因為伺服器錯誤顯示 statusCode : 413，訊息：您的使用者端傳送的請求太大。請參閱 SPLUNK 文件以設定 max_content_length。」
Splunk.IndexerBusy	「將資料傳送至 Splunk 失敗，因為 HEC 節點的伺服器出現錯誤。確保 HEC 端點或 Elastic Load Balancer 可連接並且狀態良好。」
Splunk.ConnectionRecycled	「從 Firehose 與 Splunk 的連線已回收。將重試交付。」
Splunk.AcknowledgmentsDisabled	「無法獲得 POST 的確認。確保已在 HEC 端點上啟用了確認。」
Splunk.InvalidHecResponseCharacter	「在 HEC 響應中發現無效字符，請務必檢查服務和 HEC 組態。」

## ElasticSearch 資料傳送錯誤

Amazon 數據 Firehose 件可以將以下 ElasticSearch 錯誤發送到 CloudWatch 日誌。

錯誤程式碼	錯誤訊息與資訊
ES.AccessDenied	「存取遭拒。確保提供與 Firehose 相關聯的 IAM 角色不會被刪除。」
ES.ResourceNotFound	「指定的 AWS 彈性搜索域不存在。」

## HTTPS 端點資料交付錯誤

Amazon 數據 Firehose 件可以將以下 HTTP 端點相關錯誤發送到日誌。CloudWatch 如果這些錯誤都不符合您遇到的問題，則預設錯誤如下：「嘗試傳送資料時發生內部錯誤。交付將被重試；如果錯誤仍然存在，則將報告 AWS 給解決。」

錯誤程式碼	錯誤訊息與資訊
HttpEndpoint.RequestTimeout	在收到回應之前交付逾時，將會重試。如果此錯誤仍然存在，請聯絡 AWS Firehose 服務團隊。
HttpEndpoint.ResponseTooLarge	「從端點收到的回應尺寸過大。請聯絡端點的擁有者以解決此問題。」
HttpEndpoint.InvalidResponseFromDestination	「從指定端點收到的回應無效。請聯絡端點的擁有者以解決此問題。」
HttpEndpoint.DestinationException	「從端點目的地接收到以下回應。」

錯誤程式碼	錯誤訊息與資訊
<code>HttpEndpoint.ConnectionFailed</code>	「無法連線至目的地端點。請聯絡端點的擁有者以解決此問題。」
<code>HttpEndpoint.ConnectionReset</code>	「無法維持與端點的連線。請聯絡端點的擁有者以解決此問題。」
<code>HttpEndpoint.ConnectionReset</code>	「難以維持與端點的連線。請聯絡端點的所有者。」
<code>HttpEndpoint.ResponseReasonPhraseExceededLimit</code>	「從端點收到的回應原因片語超過設定的 64 個字元限制。」
<code>HttpEndpoint.InvalidResponseFromDestination</code>	「從端點收到的回應無效。如需詳細資訊，請參閱 Firehose 文件中的 HTTP 端點疑難排解。原因：「
<code>HttpEndpoint.DestinationException</code>	「交付至端點失敗。如需詳細資訊，請參閱 Firehose 文件中的 HTTP 端點疑難排解。收到狀態碼的回應」
<code>HttpEndpoint.InvalidStatusCode</code>	「收到無效的回應狀態碼。」
<code>HttpEndpoint.SSLHandshakeFailure</code>	「無法與端點完成 SSL 握手。請聯絡端點的擁有者以解決此問題。」

錯誤程式碼	錯誤訊息與資訊
<code>HttpEndpoint.SSLHandshakeFailure</code>	「無法與端點完成 SSL 握手。請聯絡端點的擁有者以解決此問題。」
<code>HttpEndpoint.SSLFailure</code>	「無法與端點完成 TLS 握手。請聯絡端點的擁有者以解決此問題。」
<code>HttpEndpoint.SSLHandshakeCertificatePathFailure</code>	「由於憑證路徑無效，無法與端點完成 SSL 交握。請聯絡端點的擁有者以解決此問題。」
<code>HttpEndpoint.SSLHandshakeCertificatePathValidationFailure</code>	「由於認證路徑驗證失敗，無法與端點完成 SSL 握手。請聯絡端點的擁有者以解決此問題。」
<code>HttpEndpoint.MakeRequestFailure.IllegalUriException</code>	「由於 URI 中輸入無效，HttpEndpoint 請求失敗。請確保輸入 URI 中的所有字符都有效。」
<code>HttpEndpoint.MakeRequestFailure.IllegalCharacterInHeaderValue</code>	「由於非法響應錯誤，HttpEndpoint 請求失敗。標頭值中的非法字元 '\n'。」

錯誤程式碼	錯誤訊息與資訊
HttpEndpoint.IllegalResponseFailure	「由於非法響應錯誤，HttpEndpoint 請求失敗。HTTP 消息不得包含多個內容類型標頭。」
HttpEndpoint.IllegalMessageStart	「由於非法響應錯誤，HttpEndpoint 請求失敗。非法 HTTP 消息啟動。如需詳細資訊，請參閱 Firehose 文件中的 HTTP 端點疑難排解。」

## Amazon OpenSearch 服務數據交付錯誤

對於 OpenSearch 服務目的地，Amazon 資料 Firehose 會在服務傳回 CloudWatch 日誌時將錯誤傳送至日誌。OpenSearch

除了可能從 OpenSearch 叢集傳回的錯誤之外，您可能會遇到下列兩個錯誤：

- 嘗試將資料傳送至目 OpenSearch 的地服務叢集期間發生驗證/授權錯誤。這可能是由於任何權限問題和/或間歇性地在修改 Amazon Data Firehose 目標 OpenSearch 服務網域組態時發生。請檢查叢集原則和角色許可。
- 由於身份驗證/授權失敗，數據無法傳遞到目的地 OpenSearch 服務集群。這可能是由於任何權限問題和/或間歇性地在修改 Amazon Data Firehose 目標 OpenSearch 服務網域組態時發生。請檢查叢集原則和角色許可。

錯誤程式碼	錯誤訊息與資訊
OS.AccessDenied	「存取遭拒。確保所提供 IAM 角色的信任政策允許 Firehose 擔任該角色，並且存取政策允許存取 Amazon OpenSearch 服務 API。」
OS.AccessDenied	「存取遭拒。確保所提供 IAM 角色的信任政策允許 Firehose 擔任該角色，並且存取政策允許存取 Amazon OpenSearch 服務 API。」
OS.AccessDenied	「存取遭拒。確保提供與 Firehose 相關聯的 IAM 角色不會被刪除。」
OS.AccessDenied	「存取遭拒。確保提供與 Firehose 相關聯的 IAM 角色不會被刪除。」



錯誤程式碼	錯誤訊息與資訊
OS.ResourceNotFound	「指定的 Amazon OpenSearch 服務域不存在。」
OS.ResourceNotFound	「指定的 Amazon OpenSearch 服務域不存在。」
OS.AccessDenied	「存取遭拒。確保所提供 IAM 角色的信任政策允許 Firehose 擔任該角色，並且存取政策允許存取 Amazon OpenSearch 服務 API。」
OS.RequestTimeout	「向 Amazon OpenSearch 服務叢集或 OpenSearch 無伺服器收集的請求逾時。確保叢集或集合具有足夠的容量來容納目前的工作負載。」
OS.ClusterError	「Amazon OpenSearch 服務集群返回了一個未指定的錯誤。」
OS.RequestTimeout	「對 Amazon OpenSearch 服務集群的請求超時。確保叢集具有足夠的容量來容納目前的工作負載。」
OS.ConnectionFailed	「連接到 Amazon OpenSearch 服務集群或 OpenSearch 無伺服器集合時出現問題。確保叢集或集合狀態良好且可存取。」
OS.ConnectionReset	「無法保持與 Amazon OpenSearch 服務集群或 OpenSearch 無伺服器集合的連接。請聯絡叢集或集合的擁有者以解決此問題。」
OS.ConnectionReset	「無法維護與 Amazon OpenSearch 服務集群或 OpenSearch 無伺服器集合的連接。確保叢集或集合狀態良好，並且具有足夠的容量來容納目前的工作負載。」
OS.ConnectionReset	「無法維護與 Amazon OpenSearch 服務集群或 OpenSearch 無伺服器集合的連接。確保叢集或集合狀態良好，並且具有足夠的容量來容納目前的工作負載。」
OS.AccessDenied	「存取遭拒。確保 Amazon OpenSearch 服務叢集上的存取政策授予對已設定 IAM 角色的存取權。」

錯誤程式碼	錯誤訊息與資訊
OS.ValidationException	「OpenSearch 叢集傳回 ES ServiceException。其中一個原因是群集已升級到 OS 2.x 或更高版本，但軟管仍然配置了 TypeName 參數。透過將設定為空字串 TypeName 來更新軟管規劃，或將端點變更為支援 Type 參數的叢集。」
OS.ValidationException	「成員必須滿足正則運算式模式： <code>[a-z][a-z0-9\\-]+</code> 」
OS.JsonParseException	「Amazon 服 OpenSearch 務集群返回了一個 JsonParseException. 確保放入的資料是有效的。」
OS.AmazonOpenSearchServiceParseException	「Amazon 服 OpenSearch 務集群返回了一個 AmazonOpenSearchServiceParseException. 確保放入的資料是有效的。」
OS.ExplicitIndexInBulkNotAllowed	「確保在 Amazon 服務集群上休息動作多。OpenSearch
OS.ClusterError	「Amazon OpenSearch 服務集群或 OpenSearch 無服務器集合返回一個未指定的錯誤。」
OS.ClusterBlockException	「叢集傳回了 ClusterBlockException. 它可能會超載。」
OS.InvalidARN	「提供的 Amazon OpenSearch 服務 ARN 無效. 請檢查您的 DeliveryStream 配置。」
OS.MalformedData	「一個或多個記錄格式錯誤。請確保每條記錄都是單個有效的 JSON 對象，並且不包含換行符。」
OS.InternalError	「嘗試交付資料時發生內部錯誤。交付將被重試; 如果錯誤仍然存在，將報告 AWS 給解決。」

錯誤程式碼	錯誤訊息與資訊
OS.AliasWithMultipleIndicesNotAllowed	「別名有一個以上的索引與它相關聯。確保別名只有一個與其關聯的索引。」
OS.UnsupportedVersion	「Amazon OpenSearch 服務 6.0 目前不支持 Amazon 數據 Firehose 件。請聯繫 <a href="#">AWS support</a> 以獲取更多信息。」
OS.CharacterConversionException	「一個或多個記錄包含無效字符。」
OS.InvalidDomainNameLength	「域名稱長度不在有效的作業系統限制範圍內。」
OS.VPCDomainNotSupported	「目前不支援 VPC 中的 Amazon OpenSearch 服務網域。」
OS.ConnectionError	「http 服務器意外關閉了連接，請驗證 Amazon OpenSearch 服務集群或 OpenSearch 無服務器集合的運行狀態。」
OS.LargeFieldData	「Amazon OpenSearch 服務叢集中止了請求，因為它包含大於允許的欄位資料。」
OS.BadGateway	「Amazon OpenSearch 服務集群或 OpenSearch 無服務器集合中止了響應的請求：502 錯誤的網關。」
OS.ServiceException	「從 Amazon OpenSearch 服務集群或 OpenSearch 無服務器集合收到錯誤。如果叢集或集合位於 VPC 後面，請確保網路組態允許連線。」
OS.GatewayTimeout	「Firehose 在連接到 Amazon OpenSearch 服務叢集或 OpenSearch 無伺服器集合時遇到逾時錯誤。」
OS.MalformedData	「Amazon 資料防火管不支援 Firehose 記錄內的 Amazon OpenSearch 服務批量 API 命令。」

錯誤程式碼	錯誤訊息與資訊
OS.ResponseEntryCountMismatch	「來自批量 API 的響應包含的條目多於傳送的記錄數目。確保每條記錄只包含一個 JSON 對象，並且沒有換行符。」

## Lambda 調用錯誤

Amazon 數據 Firehose 件可以將以下 Lambda 調用錯誤發送到 CloudWatch 日誌。

錯誤程式碼	錯誤訊息與資訊
Lambda.AssumeRoleAccessDenied	「存取遭拒。確保所提供 IAM 角色的信任政策允許 Amazon 資料 Firehose 擔任該角色。」
Lambda.InvokeAccessDenied	「存取遭拒。請確認存取政策允許存取 Lambda 函數。」
Lambda.JsonProcessingException	「Lambda 函數回傳記錄時發生錯誤剖析。確保返回的記錄遵循 Amazon 數據 Firehose 所要求的狀態模型。」 如需詳細資訊，請參閱 <a href="#">資料轉換和狀態模型</a> 。
Lambda.InvokeLimitExceeded	「Lambda 並行執行超過限制。請增加並行執行的限制。」 如需詳細資訊，請參閱《AWS Lambda 開發人員指南》中的 <a href="#">AWS Lambda 限制</a> 。
Lambda.DuplicatedRecordId	「多個回傳記錄具有相同的記錄 ID。請確認 Lambda 函數回傳的每一記錄均有獨特記錄 ID。」 如需詳細資訊，請參閱 <a href="#">資料轉換和狀態模型</a> 。
Lambda.MissingRecordId	「一個或多個記錄 ID 未回傳。請確認 Lambda 函數會回傳所有接收到的記錄 ID。」 如需詳細資訊，請參閱 <a href="#">資料轉換和狀態模型</a> 。

錯誤程式碼	錯誤訊息與資訊
Lambda.ResourceNotFound	「指定的 Lambda 函數不存在。請使用另一現有的函數。」
Lambda.InvalidSubnetIDException	「Lambda 函數 VPC 設定指定的子網路 ID 無效。請確認子網路 ID 有效。」
Lambda.InvalidSecurityGroupIDException	「Lambda 函數 VPC 設定指定的安全群組 ID 無效。請確認安全群組 ID 有效。」
Lambda.SubnetIPAddressLimitReachedException	<p>「無法AWS Lambda 為 Lambda 函數設定 VPC 存取權，因為一或多個已設定的子網路沒有可用的 IP 位址。請提高 IP 地址的限制。」</p> <p>如需有關 Amazon VPC 配額的詳細資訊，請參閱《Amazon VPC 使用者指南》中的 <a href="#">Amazon VPC 限制 - VPC 和子網路</a>。</p>
Lambda.ENILimitReachedException	<p>「AWS Lambda 無法在 VPC 中建立彈性網路介面 (ENI)，指定為 Lambda 函數組態的一部分，因為已達到網路介面的限制。請提高網路介面的限制。」</p> <p>如需詳細資訊，請參閱《Amazon VPC 使用者指南》<a href="https://docs.aws.amazon.com/vpc/latest/userguide/VPC_Appendix_Limits.html#vpc-limits-enis">https://docs.aws.amazon.com/vpc/latest/userguide/VPC_Appendix_Limits.html#vpc-limits-enis</a>的 Amazon VPC 限制 - 網路介面。</p>
Lambda.FunctionTimedOut	Lambda 函數調用已逾時。增加 Lambda 函數中的「逾時」設定。如需詳細資訊，請參閱 <a href="#">設定函數逾時</a> 。

錯誤程式碼	錯誤訊息與資訊
Lambda FunctionError	<p>這可能是由於以下任何錯誤造成的：</p> <ul style="list-style-type: none"> <li>無效的輸出結構。檢查您的功能，並確保輸出是所需的格式。此外，請確認已處理的記錄包含 Dropped、Ok 或 ProcessingFailed 的有效結果狀態。</li> <li>已成功調用 Lambda 函數，但傳回錯誤結果。</li> <li>Lambda 無法解密環境變數，因為 KMS 存取被拒絕。請檢查函數的 KMS 金鑰設定以及金鑰政策。如需詳細資訊，請參閱<a href="#">金鑰存取疑難排解</a>。</li> </ul>
Lambda FunctionRequestTimedOut	<p>在叫用 Lambda 時，Amazon 資料 Firehose 在請求逾時組態錯誤之前遇到請求未完成。重新檢視 Lambda 程式碼，以檢查 Lambda 程式碼是否已準備好在超出設定的逾時後執行。如果已就緒，請考慮調整 Lambda 組態設定，包括記憶體、逾時。如需詳細資訊，請參閱<a href="#">設定 Lambda 函數選項</a>。</p>
Lambda TargetServerFailedToRespond	<p>Amazon 數據 Firehose 件遇到錯誤。呼叫 AWS Lambda 服務時，目標伺服器無法回應錯誤。</p>
Lambda InvalidZipFileException	<p>Amazon 數據 Firehose 調用 Lambda 函數 InvalidZipFileException 時遇到。檢查您的 Lambda 函數組態設定和 Lambda 程式碼壓縮檔案。</p>
Lambda InternalServerError	<p>Amazon 數據 Firehose 在調用 AWS Lambda 服務 InternalServerError 時遇到。Amazon 數據 Firehose 件將重試發送數據固定次數。您可以使用 CreateDeliveryStream 或 UpdateDestination API 來指定或覆寫重試選項。如果錯誤仍然存在，請聯絡 AWS Lambda 支援團隊。</p>
Lambda ServiceUnavailable	<p>Amazon 數據 Firehose 調用 AWS Lambda 服務 ServiceUnavailable Exception 時遇到。Amazon 數據 Firehose 件將重試發送數據固定次數。您可以使用 CreateDeliveryStream 或 UpdateDestination API 來指定或覆寫重試選項。如果錯誤仍然存在，請聯絡 AWS Lambda 支援。</p>
Lambda InvalidSecurityToken	<p>由於無效的安全記號，無法調用 Lambda 函數。跨分割 Lambda 調用不支援跨分割 Lambda 調用。</p>

錯誤程式碼	錯誤訊息與資訊
Lambda InvocationFailure	這可能是由於以下任何錯誤造成的： <ul style="list-style-type: none"> <li>• Amazon 數據 Firehose 件在調用 AWS Lambda 時遇到錯誤。操作將重試；若錯誤仍持續出現，將會回報至 AWS 以尋求解決方案。」</li> <li>• Amazon 數據 Firehose 件遇到了InvalidStateException 來自 Lambda 的 KMS。Lambda 無法解密環境變數，因為使用的 KMS 金鑰對於解密而言處於無效狀態。請檢查 Lambda 函數的 KMS 金鑰。</li> <li>• Amazon 數據 Firehose 件遇到了 AWS LambdaException 來自 Lambda 的。Lambda 無法初始化提供的容器映像。驗證映像。</li> <li>• Amazon 數據 Firehose 件在調用 AWS Lambda 時遇到超時錯誤。支援的函數逾時上限為 5 分鐘。如需詳細資訊，請參閱<a href="#">資料轉換執行持續時間</a>。</li> </ul>
Lambda JsonMappingException	Lambda 函數回傳記錄時發生錯誤剖析。確認資料欄位採用 base-64 編碼。

## Kinesis 調用錯誤

Amazon 資料 Firehose 可以將下列 Kinesis 叫用錯誤傳送到日誌。 CloudWatch

錯誤程式碼	錯誤訊息與資訊
Kinesis.AccessDenied	「呼叫 Kinesis 時存取被拒絕。確保所使用 IAM 角色上的存取政策允許存取適當的 Kinesis API。」
Kinesis.ResourceNotFound	「Firehose 未能從串流中讀取。如果 Firehose 與 Kinesis Stream 相連，則該串流可能不存在，或碎片可能已合併或分割。如果 Firehose 是屬於 DirectPut 類型的話，就可能已經不存在了。」
Kinesis.SubscriptionRequired	「呼叫 Kinesis 時存取被拒絕。確保為 Kinesis 串流存取傳遞的 IAM 角色具有 AWS Kinesis 訂閱。」

錯誤程式碼	錯誤訊息與資訊
Kinesis.T hrottling	「呼叫 Kinesis 時遇到限流錯誤。這可能是因為其他應用程式呼叫與 Firehose Firehose 串流相同的 API，或者您已經建立了太多與來源相同的 Kinesis 串流的 Firehose 串流。」
Kinesis.T hrottling	「呼叫 Kinesis 時遇到限流錯誤。這可能是因為其他應用程式呼叫與 Firehose Firehose 串流相同的 API，或者您已經建立了太多與來源相同的 Kinesis 串流的 Firehose 串流。」
Kinesis.A ccessDenied	「呼叫 Kinesis 時存取被拒絕。確保所使用 IAM 角色上的存取政策允許存取適當的 Kinesis API。」
Kinesis.A ccessDenied	「嘗試在基礎 Kinesis 串流上呼叫 api 時，存取遭到拒絕。確保 IAM 角色已傳播且有效。」
Kinesis.K MS.Access DeniedExc eption	「Firehose 無法存取用於加密/解密 Kinesis 串流的 KMS 金鑰。請授予 Firehose 交付角色對金鑰的存取權。」
Kinesis.K MS.KeyDisabled	「Firehose 無法從來源 Kinesis 串流讀取，因為用於加密/解密的 KMS 金鑰已停用。啟用金鑰，以便讀取可以繼續。」
Kinesis.K MS.Invalid StateExc eption	Firehose 無法從來源 Kinesis Streams 讀取，因為用來加密 Kinesis Streams 的 KMS 金鑰處於無效狀態。」
Kinesis.K MS.NotFound Exception	「Firehose 無法從來源 Kinesis 串流讀取，因為找不到用來加密的 KMS 金鑰。」

## Kinesis DirectPut 調用錯誤

Amazon 資料 Firehose 可以將下列 Kinesis DirectPut 叫用錯誤傳送到日誌。 CloudWatch



錯誤程式碼	錯誤訊息與資訊
Firehose.KMS.AccessDeniedException	「Firehose 無法存取 KMS 金鑰。請檢查金鑰政策。」
Firehose.KMS.InvalidStateException	「Firehose 無法解密資料，因為用於加密資料的 KMS 金鑰處於無效狀態。」
Firehose.KMS.NotFoundException	「Firehose 無法解密資料，因為找不到用於加密資料的 KMS 金鑰。」
Firehose.KMS.KeyDisabled	「Firehose 無法解密資料，因為用於加密資料的 KMS 金鑰已停用。啟用金鑰，以便可以繼續進行交付。」

## AWS Glue 調用錯誤

Amazon 數據 Firehose 可以將以下 AWS Glue 調用錯誤發送到 CloudWatch 日誌。

錯誤程式碼	錯誤訊息與資訊
DataFormatConversion.InvalidSchema	「結構描述無效。」
DataFormatConversion.EntityNotFound	「找不到指定的資料表/資料庫。請確保資料表/資料庫存在，並且結構描述組態中提供的值是正確的，尤其是在大小寫方面。」

錯誤程式碼	錯誤訊息與資訊
DataFormatConversion.InvalidInput	「無法從 glue 中找到匹配的模式。請確保具有提供的目錄 ID 的指定資料庫存在。」
DataFormatConversion.InvalidInput	「無法從 glue 中找到匹配的模式。請確保傳遞的 ARN 格式正確。」
DataFormatConversion.InvalidInput	「無法從 glue 中找到匹配的模式。請確保提供的 catalogId 有效。」
DataFormatConversion.InvalidVersionId	「無法從 glue 中找到匹配的模式。請確保資料表的指定版本存在。」
DataFormatConversion.NonExistentColumns	「無法從 glue 中找到匹配的模式。請確保資料表設定了包含目標列的非空儲存描述符。」
DataFormatConversion.AccessDenied	「假設角色時拒絕存取。請確保在資料格式轉換組態中指定的角色已授予 Firehose 服務承擔該角色的許可。」
DataFormatConversion.ThrottledByGlue	「呼叫 Glue 時遇到限流錯誤。可以提高請求速率限制，或降低透過其他應用程式呼叫 Glue 的當前速率。」

錯誤程式碼	錯誤訊息與資訊
DataFormatConversion.AccessDenied	「呼叫 Glue 時存取被拒絕。請確保在資料格式轉換設定中指定的角色具有必要的許可。」
DataFormatConversion.InvalidGlueRole	「無效的角色。請確保在資料格式轉換組態中指定的角色存在。」
DataFormatConversion.InvalidGlueRole	「包含在請求中的安全性記號無效。確保提供與 Firehose 相關聯的 IAM 角色不會被刪除。」
DataFormatConversion.GlueNotAvailableInRegion	「AWS Glue 在您指定的地區尚未提供，請指定其他地區。」
DataFormatConversion.GlueEncryptionException	「擷取主金鑰時發生錯誤。確保金鑰存在並具有正確的存取許可。」
DataFormatConversion.SchemaValidationTimeout	「從 Glue 檢索資料表時超時。如果您有大量的 Glue 資料表版本，請新增「glue:GetTableVersion」權限 (建議) 或刪除未使用的資料表版本。如果您在 Glue 中沒有大量的表格，請聯絡 Sup AWS port 人員。」
DataFirehose.InternalError	「從 Glue 檢索資料表時超時。如果您有大量的 Glue 資料表版本，請新增「glue:GetTableVersion」權限 (建議) 或刪除未使用的資料表版本。如果您在 Glue 中沒有大量的表格，請聯絡 Sup AWS port 人員。」

錯誤程式碼	錯誤訊息與資訊
DataFormatConversion.GlueEncryptionException	「擷取主金鑰時發生錯誤。確保金鑰存在且狀態正確。」

## DataFormatConversion 調用錯誤

Amazon 數據 Firehose 可以將以下 DataFormatConversion 調用錯誤發送到 CloudWatch 日誌。

錯誤程式碼	錯誤訊息與資訊
DataFormatConversion.InvalidSchema	「結構描述無效。」
DataFormatConversion.ValidationException	「欄位名稱和類型必須是非空字符串。」
DataFormatConversion.ParseError	「遇到格式錯誤的 JSON。」
DataFormatConversion.MalformedData	「資料與結構描述不相符。」
DataFormatConversion.MalformedData	「JSON 金鑰的長度不得大於 262144」

錯誤程式碼	錯誤訊息與資訊
DataFormatConversion.MalformedData	「資料無法解碼為 UTF-8。」
DataFormatConversion.MalformedData	「記號之間發現非法字符。」
DataFormatConversion.InvalidTypeFormat	「類型格式無效。檢查類型語法。」
DataFormatConversion.InvalidSchema	「無效的結構描述。請確保欄位名稱中沒有特殊字符或空格。」
DataFormatConversion.InvalidRecord	「記錄不符合結構描述。一個或多個映射鍵對於映射無效<string, string>。」
DataFormatConversion.MalformedData	「輸入 JSON 在頂層包含一個基本類型。頂層必須是物件或陣列。」
DataFormatConversion.MalformedData	「輸入 JSON 在頂層包含一個基本類型。頂層必須是物件或陣列。」

錯誤程式碼	錯誤訊息與資訊
DataFormatConversion.MalformedData	「記錄為空或僅包含空格。」
DataFormatConversion.MalformedData	「遇到無效字元。」
DataFormatConversion.MalformedData	「遇到無效或不支持的時間戳格式。如需支援的時間戳記格式，請參閱《Firehose 開發人員指南》。」
DataFormatConversion.MalformedData	「在資料中找到了一個標量類型，但在結構描述上指定了一個複雜的類型。」
DataFormatConversion.MalformedData	「資料與結構描述不相符。」
DataFormatConversion.MalformedData	「在資料中找到了一個標量類型，但在結構描述上指定了一個複雜的類型。」
DataFormatConversion.ConversionFailureException	"ConversionFailureException"

錯誤程式碼	錯誤訊息與資訊
DataFormatConversion.CustomerErrorException	"DataFormatConversionCustomerErrorException"
DataFormatConversion.CustomerErrorException	"DataFormatConversionCustomerErrorException"
DataFormatConversion.MalformedData	「資料與結構描述不相符。」
DataFormatConversion.InvalidSchema	「結構描述無效。」
DataFormatConversion.MalformedData	「資料與結構描述不相符。一個或多個日期的格式無效。」
DataFormatConversion.MalformedData	「資料包含不支援的高度巢狀 JSON 結構。」

錯誤程式碼	錯誤訊息與資訊
DataFormatConversion.EntityNotFound	「找不到指定的資料表/資料庫。請確保資料表/資料庫存在，並且結構描述組態中提供的值是正確的，尤其是在大小寫方面。」
DataFormatConversion.InvalidInput	「無法從 glue 中找到匹配的模式。請確保具有提供的目錄 ID 的指定資料庫存在。」
DataFormatConversion.InvalidInput	「無法從 glue 中找到匹配的模式。請確保傳遞的 ARN 格式正確。」
DataFormatConversion.InvalidInput	「無法從 glue 中找到匹配的模式。請確保提供的 catalogId 有效。」
DataFormatConversion.InvalidVersionId	「無法從 glue 中找到匹配的模式。請確保資料表的指定版本存在。」
DataFormatConversion.NonExistentColumns	「無法從 glue 中找到匹配的模式。請確保資料表設定了包含目標列的非空儲存描述符。」
DataFormatConversion.AccessDenied	「假設角色時拒絕存取。請確保在資料格式轉換組態中指定的角色已授予 Firehose 服務承擔該角色的許可。」



錯誤程式碼	錯誤訊息與資訊
DataFormatConversion.ThrottledByGlue	「呼叫 Glue 時遇到限流錯誤。可以提高請求速率限制，或降低透過其他應用程式呼叫 Glue 的當前速率。」
DataFormatConversion.AccessDenied	「呼叫 Glue 時存取被拒絕。請確保在資料格式轉換設定中指定的角色具有必要的許可。」
DataFormatConversion.InvalidGlueRole	「無效的角色。請確保在資料格式轉換組態中指定的角色存在。」
DataFormatConversion.GlueNotAvailableInRegion	「AWS Glue 在您指定的地區尚未提供，請指定其他地區。」
DataFormatConversion.GlueEncryptionException	「擷取主金鑰時發生錯誤。確保金鑰存在並具有正確的存取許可。」
DataFormatConversion.SchemaValidationTimeout	「從 Glue 檢索資料表時超時。如果您有大量的 Glue 資料表版本，請新增「glue:GetTableVersion」權限 (建議) 或刪除未使用的資料表版本。如果您在 Glue 中沒有大量的表格，請聯絡 Sup AWS port 人員。」
DataFirehose.InternalError	「從 Glue 檢索資料表時超時。如果您有大量的 Glue 資料表版本，請新增「glue:GetTableVersion」權限 (建議) 或刪除未使用的資料表版本。如果您在 Glue 中沒有大量的表格，請聯絡 Sup AWS port 人員。」

錯誤程式碼	錯誤訊息與資訊
DataForma tConversion.Malfor medData	「一個或多個欄位的格式不正確。」

## 訪問 Amazon 數據 Firehose 的 CloudWatch 日誌

您可以使用 Amazon 資料 Firehose 主控台或主控台，檢視與 Amazon 資料 Firehose 資料交付失敗相關的錯誤日誌。CloudWatch 以下程序將說明如何使用下列兩種方法來存取錯誤日誌。

若要使用 Amazon 資料 Firehose 主控台存取錯誤日誌

1. 登入 AWS Management Console 並開啟 Firehose 主控台，網址為 <https://console.aws.amazon.com/firehose>
2. 在導覽列上，選擇「AWS 區域」。
3. 選擇「Firehose」串流名稱，即可前往「Firehose 串流詳細資料」頁面。
4. 選擇 Error Log (錯誤日誌)，檢視資料交付失敗相關的錯誤日誌清單。

使用 CloudWatch 主控台存取錯誤記錄

1. [請在以下位置開啟 CloudWatch 主控台。](https://console.aws.amazon.com/cloudwatch/) <https://console.aws.amazon.com/cloudwatch/>
2. 在導覽列上，選擇一個區域。
3. 在導覽窗格中，選擇日誌。
4. 選擇日誌群組和日誌串流，以檢視資料交付失敗相關的錯誤日誌清單。

## 監控 Kinesis 代理程式運作狀態

Kinesis 代理程式發行命名空間為的自訂 CloudWatch 指標。AWS KinesisAgent 它有助於評估代理程式是否運作良好、依指定將資料提交至 Amazon Data Firehose，以及在資料生產者上消耗適當數量的 CPU 和記憶體資源。

傳送的記錄數和位元組數等度量對於瞭解代理程式將資料提交至 Firehose 串流的速率非常有用。當這些指標滑落至預期閾值以下特定百分比或滑落至零，可能表示設定有問題、網路出現錯誤或代理程式運

作狀態不佳。諸如主機 CPU 和記憶體消耗量與代理程式錯誤計數器等指標，均顯示資料產生來源的資源使用情況，並提供潛在的設定或主機錯誤等洞見。最後，代理程式亦會記錄服務例外狀況，以協助調查代理程式的問題。

代理程式指標，會在代理程式的 `cloudwatch.endpoint` 的阻態所指定的區域中回報。如需詳細資訊，請參閱 [代理程式組態設定](#)。

對從多個 Kinesis 代理程式發佈的 Cloudwatch 指標進行彙整或合併。

Kinesis 代理程式發出的指標將收取一筆名目費用，此功能預設為啟用。如需詳細資訊，請參閱 [Amazon CloudWatch 定價](#)。

## 使用監控 CloudWatch

Kinesis 代理程式會將下列指標傳送至 CloudWatch。

指標	描述
BytesSent	指定期間內傳送至 Firehose 串流的位元組數目。  單位：位元組
RecordSendAttempts	呼叫於指定期間內PutRecordBatch 嘗試的記錄 (不論第一次嘗試或是重試) 數量。  單位：計數
RecordSendErrors	呼叫於指定期間內PutRecordBatch 回傳失敗狀態的記錄 (包括重試) 數量。  單位：計數
ServiceErrors	呼叫於指定期間內PutRecordBatch 導致服務錯誤 (調節錯誤除外) 數量。  單位：計數

## 使用日誌記錄 Amazon 數 Firehose API 調用 AWS CloudTrail

Amazon 數據 Firehose 與該服務集成在一起 AWS CloudTrail，該服務可提供 Amazon 數據 Firehose 中的用戶，角色或 AWS 服務採取的操作記錄。CloudTrail 將 Amazon 資料 Firehose 的所有 API 呼叫擷取為事件。擷取的呼叫包括來自 Amazon 資料 Firehose 主控台的呼叫，以及對 Amazon 資料 Firehose API 作業的程式碼呼叫。如果您建立追蹤，您可以啟用持續交付 CloudTrail 事件到 Amazon S3 儲存貯體，包括 Amazon 資料 Firehose 的事件。如果您未設定追蹤，您仍然可以在 [事件歷程記錄] 中檢視 CloudTrail 主控台中最近的事件。使用收集的資訊 CloudTrail，您可以判斷向 Amazon Data Firehose 提出的請求、提出請求的來源 IP 地址、提出請求的人員、提出請求的時間以及其他詳細資訊。

若要進一步了解 CloudTrail，包括如何設定和啟用它，請參閱[AWS CloudTrail 使用者指南](#)。

### Amazon 數據 Firehose 信息 CloudTrail

CloudTrail 在您創建 AWS 帳戶時，您的帳戶已啟用。當 Amazon Data Firehose 中發生受支援的事件活動時，該活動會與事件歷史記錄中的其他 AWS 服務 CloudTrail 事件一起記錄在事件中。您可以檢視、搜尋和下載 AWS 帳戶的最新事件。如需詳細資訊，請參閱[檢視具有事 CloudTrail 事件記錄的事件](#)。

如需 AWS 帳戶中持續的事件記錄 (包括 Amazon 資料 Firehose 的事件)，請建立追蹤。追蹤可 CloudTrail 將日誌檔交付到 Amazon S3 儲存貯體。根據預設，當您在主控台中建立追蹤時，追蹤會套用至所有 AWS 區域。追蹤記錄來自 AWS 分區中所有區域的事件，並將日誌檔傳送到您指定的 Amazon S3 儲存貯體。此外，您還可以設定其他 AWS 服務，以進一步分析 CloudTrail 記錄中收集的事件資料並採取行動。如需詳細資訊，請參閱下列內容：

- [建立追蹤的概觀](#)
- [CloudTrail 支援的服務與整合](#)
- [設定的 Amazon SNS 通知 CloudTrail](#)
- [從多個區域接收 CloudTrail 記錄檔並從多個帳戶接收 CloudTrail 記錄檔](#)

Amazon 資料 Firehose 支援將下列動作記錄為記 CloudTrail 錄檔中的事件：

- [CreateDeliveryStream](#)
- [DeleteDeliveryStream](#)
- [DescribeDeliveryStream](#)
- [ListDeliveryStreams](#)
- [ListTagsForDeliveryStream](#)

- [TagDeliveryStream](#)
- [StartDeliveryStreamEncryption](#)
- [StopDeliveryStreamEncryption](#)
- [UntagDeliveryStream](#)
- [UpdateDestination](#)

每一筆事件或日誌專案都會包含產生請求者的資訊。身分資訊可協助您判斷下列事項：

- 要求是使用根使用者登入資料還是 AWS Identity and Access Management (IAM) 使用者登入資料提出。
- 提出該請求時，是否使用了特定角色或聯合身分使用者的暫時安全憑證。
- 請求是否由其他 AWS 服務提出。

如需詳細資訊，請參閱[CloudTrail 使用 userIdentity 元素](#)。

## 範例：Amazon 資料 Firehose 記錄檔項目

追蹤是一種組態，可讓事件以日誌檔的形式傳遞到您指定的 Amazon S3 儲存貯體。CloudTrail 記錄檔包含一或多個記錄項目。事件代表來自任何來源的單一請求，包括有關請求的操作，動作的日期和時間，請求參數等信息。CloudTrail 日誌文件不是公共 API 調用的有序堆棧跟踪，因此它們不會以任何特定順序顯示。

下列範例顯示示

範CreateDeliveryStream、DescribeDeliveryStreamListDeliveryStreams、UpdateDestination 的 CloudTrail 記錄項目。

```
{
  "Records": [
    {
      "eventVersion": "1.02",
      "userIdentity": {
        "type": "IAMUser",
        "principalId": "AKIAIOSFODNN7EXAMPLE",
        "arn": "arn:aws:iam::111122223333:user/CloudTrail_Test_User",
        "accountId": "111122223333",
        "accessKeyId": "AKIAI44QH8DHBEXAMPLE",
        "userName": "CloudTrail_Test_User"
      }
    },
  ],
}
```

```
"eventTime":"2016-02-24T18:08:22Z",
"eventSource":"firehose.amazonaws.com",
"eventName":"CreateDeliveryStream",
"awsRegion":"us-east-1",
"sourceIPAddress":"127.0.0.1",
"userAgent":"aws-internal/3",
"requestParameters":{"
  "deliveryStreamName":"TestRedshiftStream",
  "redshiftDestinationConfiguration":{"
    "s3Configuration":{"
      "compressionFormat":"GZIP",
      "prefix":"prefix",
      "bucketARN":"arn:aws:s3:::firehose-cloudtrail-test-bucket",
      "roleARN":"arn:aws:iam::111122223333:role/Firehose",
      "bufferingHints":{"
        "sizeInMBs":3,
        "intervalInSeconds":900
      },
      "encryptionConfiguration":{"
        "kMSEncryptionConfig":{"
          "aWSKMSKeyARN":"arn:aws:kms:us-east-1:key"
        }
      }
    },
    "clusterJDBCURL":"jdbc:redshift://example.abc123.us-
west-2.redshift.amazonaws.com:5439/dev",
    "copyCommand":{"
      "copyOptions":"copyOptions",
      "dataTableName":"dataTable"
    },
    "password":"","
    "username":"","
    "roleARN":"arn:aws:iam::111122223333:role/Firehose"
  }
},
"responseElements":{"
  "deliveryStreamARN":"arn:aws:firehose:us-
east-1:111122223333:deliverystream/TestRedshiftStream"
},
"requestID":"958abf6a-db21-11e5-bb88-91ae9617edf5",
"eventID":"875d2d68-476c-4ad5-bbc6-d02872cfc884",
"eventType":"AwsApiCall",
"recipientAccountId":"111122223333"
},
```

```
{
  "eventVersion": "1.02",
  "userIdentity": {
    "type": "IAMUser",
    "principalId": "AKIAIOSFODNN7EXAMPLE",
    "arn": "arn:aws:iam::111122223333:user/CloudTrail_Test_User",
    "accountId": "111122223333",
    "accessKeyId": "AKIAI44QH8DHBEXAMPLE",
    "userName": "CloudTrail_Test_User"
  },
  "eventTime": "2016-02-24T18:08:54Z",
  "eventSource": "firehose.amazonaws.com",
  "eventName": "DescribeDeliveryStream",
  "awsRegion": "us-east-1",
  "sourceIPAddress": "127.0.0.1",
  "userAgent": "aws-internal/3",
  "requestParameters": {
    "deliveryStreamName": "TestRedshiftStream"
  },
  "responseElements": null,
  "requestID": "aa6ea5ed-db21-11e5-bb88-91ae9617edf5",
  "eventID": "d9b285d8-d690-4d5c-b9fe-d1ad5ab03f14",
  "eventType": "AwsApiCall",
  "recipientAccountId": "111122223333"
},
{
  "eventVersion": "1.02",
  "userIdentity": {
    "type": "IAMUser",
    "principalId": "AKIAIOSFODNN7EXAMPLE",
    "arn": "arn:aws:iam::111122223333:user/CloudTrail_Test_User",
    "accountId": "111122223333",
    "accessKeyId": "AKIAI44QH8DHBEXAMPLE",
    "userName": "CloudTrail_Test_User"
  },
  "eventTime": "2016-02-24T18:10:00Z",
  "eventSource": "firehose.amazonaws.com",
  "eventName": "ListDeliveryStreams",
  "awsRegion": "us-east-1",
  "sourceIPAddress": "127.0.0.1",
  "userAgent": "aws-internal/3",
  "requestParameters": {
    "limit": 10
  },
}
```

```
    "responseElements": null,
    "requestID": "d1bf7f86-db21-11e5-bb88-91ae9617edf5",
    "eventID": "67f63c74-4335-48c0-9004-4ba35ce00128",
    "eventType": "AwsApiCall",
    "recipientAccountId": "111122223333"
  },
  {
    "eventVersion": "1.02",
    "userIdentity": {
      "type": "IAMUser",
      "principalId": "AKIAIOSFODNN7EXAMPLE",
      "arn": "arn:aws:iam::111122223333:user/CloudTrail_Test_User",
      "accountId": "111122223333",
      "accessKeyId": "AKIAI44QH8DHBEXAMPLE",
      "userName": "CloudTrail_Test_User"
    },
    "eventTime": "2016-02-24T18:10:09Z",
    "eventSource": "firehose.amazonaws.com",
    "eventName": "UpdateDestination",
    "awsRegion": "us-east-1",
    "sourceIPAddress": "127.0.0.1",
    "userAgent": "aws-internal/3",
    "requestParameters": {
      "destinationId": "destinationId-000000000001",
      "deliveryStreamName": "TestRedshiftStream",
      "currentDeliveryStreamVersionId": "1",
      "redshiftDestinationUpdate": {
        "roleARN": "arn:aws:iam::111122223333:role/Firehose",
        "clusterJDBCURL": "jdbc:redshift://example.abc123.us-
west-2.redshift.amazonaws.com:5439/dev",
        "password": "",
        "username": "",
        "copyCommand": {
          "copyOptions": "copyOptions",
          "dataTableName": "dataTable"
        }
      },
      "s3Update": {
        "bucketARN": "arn:aws:s3:::firehose-cloudtrail-test-bucket-update",
        "roleARN": "arn:aws:iam::111122223333:role/Firehose",
        "compressionFormat": "GZIP",
        "bufferingHints": {
          "sizeInMBs": 3,
          "intervalInSeconds": 900
        }
      }
    }
  },
}
```



```

        "encryptionConfiguration":{
            "kMSEncryptionConfig":{
                "aWSKMSKeyARN":"arn:aws:kms:us-east-1:key"
            }
        },
        "prefix":"arn:aws:s3:::firehose-cloudtrail-test-bucket"
    }
},
"responseElements":null,
"requestID":"d549428d-db21-11e5-bb88-91ae9617edf5",
"eventID":"1cb21e0b-416a-415d-bbf9-769b152a6585",
"eventType":"AwsApiCall",
"recipientAccountId":"111122223333"
},
{
    "eventVersion":"1.02",
    "userIdentity":{
        "type":"IAMUser",
        "principalId":"AKIAIOSFODNN7EXAMPLE",
        "arn":"arn:aws:iam::111122223333:user/CloudTrail_Test_User",
        "accountId":"111122223333",
        "accessKeyId":"AKIAI44QH8DHBEXAMPLE",
        "userName":"CloudTrail_Test_User"
    },
    "eventTime":"2016-02-24T18:10:12Z",
    "eventSource":"firehose.amazonaws.com",
    "eventName":"DeleteDeliveryStream",
    "awsRegion":"us-east-1",
    "sourceIPAddress":"127.0.0.1",
    "userAgent":"aws-internal/3",
    "requestParameters":{
        "deliveryStreamName":"TestRedshiftStream"
    },
    "responseElements":null,
    "requestID":"d85968c1-db21-11e5-bb88-91ae9617edf5",
    "eventID":"dd46bb98-b4e9-42ff-a6af-32d57e636ad1",
    "eventType":"AwsApiCall",
    "recipientAccountId":"111122223333"
}
]
}

```

## 適用於 Amazon S3 物件的自訂字首

交付到 Amazon S3 的物件會遵循的 [名稱格式](#) <evaluated prefix><suffix>。您可以指定包含在執行階段評估的運算式的自訂前置詞。您指定的自訂前置詞將會覆寫的預設首碼 YYYY/MM/dd/HH。

自訂字首可使用下列形式的運算式：`!{namespace: value}`，其中 namespace 可為以下項目之一，如下列部分所述。

- firehose
- timestamp
- partitionKeyFromQuery
- partitionKeyFromLambda

如果字首結尾為斜線，看起來會像是 Amazon S3 儲存貯體內的資料夾。如需詳細資訊，請參閱 [Amazon 資料 FirehoseDeveloper 指南中的 Amazon S3 物件名稱格式](#)。

### timestamp 命名空間

此命名空間的有效值為有效 [Java DateTimeFormatter](#) 字串的字串。例如，在 2018 年，表達式！`{timestamp:yyyy}` 的評估結果為 2018。

評估時間戳記時，Firehose 會使用正在寫入之 Amazon S3 物件中所包含之最舊記錄的近似到達時間戳記。

依預設，時間戳記為 UTC。但是，您可以指定您喜歡的時區。例如，如果您想要使用日本標準時間而不是 UTC，則可以在 AWS Management Console 或 API 參數設定 ([CustomTimeZone](#)) 中將時區設定為亞洲/東京。若要查看支援的時區清單，請參閱 [Amazon S3 物件名稱格式](#)。

若您在同一字首表達式內多次使用 timestamp 命名空間，每個執行個體的評估結果都是相同的時間。

### firehose 命名空間

此命名空間可使用兩個值：`error-output-type` 和 `random-string`。下表說明如何使用這兩個值。

**firehose** 命名空間值

Conversion (轉換)	描述	範例輸入	範例輸出	備註
error-output-type	<p>根據交付串流的組態和失敗原因，評估為下列其中一個字串：{處理失敗、失敗、溢出 AmazonOpenSearchService 失敗、 、}。format-conversion-failed http-endpoint-failed</p> <p>若您在同一表達式內多次使用此值，每個執行個體的評估結果都是相同的錯誤字串。</p>	myPrefix/ result={! firehose: error-output-type} /!{timestamp:yyyy/ MM/dd}	myPrefix/ result=processing-failed/2018/08/03	此 error-output-type 值只能在 ErrorOutputPrefix 欄位中使用。
random-string	<p>評估結果為 11 個字元的隨機字串。若您在同一表達式內多次使用此值，每個執行個體的評估結果為新的隨機字串。</p>	myPrefix/ !{firehose:random-string}/	myPrefix/ 046b6c7f-0b/	<p>可用於這兩種字首類型。</p> <p>您可將其放在格式字串的開頭，以取得 Amazon S3 達到極高輸送量有時需要的隨機字首。</p>

# partitionKeyFromLambda 和 partitionKeyFromQuery 命名空間

對於[動態分割](#)，您必須在 S3 儲存貯體字首中使用下列運算式格式：`!{namespace:value}`，其中命名空間可以是 `partitionKeyFromQuery` 或 `partitionKeyFromLambda`，也可以是兩者。如果您使用內嵌剖析來建立來源資料的分割索引鍵，則必須指定由以下格式指定之運算式所組成的 S3 儲存貯體字首值：`"partitionKeyFromQuery:keyID"`。如果您使用 AWS Lambda 函數為來源資料建立分割索引鍵，則必須指定由以下格式指定之運算式所組成的 S3 儲存貯體字首值：`"partitionKeyFromLambda:keyID"`。如需詳細資訊，請參閱[建立 Amazon 資料 Firehose Delivery 串流中的「為您的目的地選擇 Amazon S3」](#)。

## 語義規則

下列規則適用於 Prefix 及 ErrorOutputPrefix 表達式。

- 以 timestamp 命名空間而言，不在單引號內的任何字元都將納入評估。換言之，值欄位中任何以單引號逸出的字串都將依照字面意思處理。
- 如果您指定不包含時間戳記命名空間運算式的前置詞，Firehose 會將運算式附加！`{timestamp:yyyy/MM/dd/HH/}` 至欄位中的 Prefix 值。
- 序列 `!{ 僅出現於 !{namespace:value} 表達式。`
- 僅在 Prefix 不具表達式時，ErrorOutputPrefix 才可為零。在此情況下，Prefix 評估為 `<specified-prefix>yyyy/MM/DDD/HH/`，而 ErrorOutputPrefix 評估為 `<specified-prefix><error-output-type>YYYY/MM/DDD/HH/`。DDD 代表某年某日。
- 若您指定 ErrorOutputPrefix 的表達式，務必納入至少一個 `!{firehose:error-output-type}` 執行個體。
- Prefix 無法納入 `!{firehose:error-output-type}`。
- Prefix 或 ErrorOutputPrefix 評估後都不能超過 512 個字元。
- 若目的地為 Amazon Redshift，Prefix 必定不能具備運算式，ErrorOutputPrefix 必須為零。
- 當目的地是 Amazon OpenSearch 服務或 Splunk，並且沒有 ErrorOutputPrefix 指定時，Firehose 會使用該 Prefix 欄位來輸入失敗的記錄。
- 當目的地為 Amazon S3 時，Amazon S3 目的地組態中的 Prefix 和 ErrorOutputPrefix 會分別用於成功的記錄和失敗的記錄。如果您使用 AWS CLI 或 API，則可以透過其自己的 Prefix 和 ErrorOutputPrefix 使用 `ExtendedS3DestinationConfiguration` 來指定 Amazon S3 備份組態。

- 當您使用 AWS Management Console 並將目標設定為 Amazon S3 時，Firehose 會分別使用目標組態 `ErrorOutputPrefix` 中的 `Prefix` 和來處理成功記錄和失敗的記錄。如果您指定字首但沒有錯誤字首，Firehose 會自動將錯誤前置字元設定為 `!{firehose:error-output-type}/`。
- 當您與 API `ExtendedS3DestinationConfiguration` 搭配使用時，或者 AWS CloudFormation，如果您指定 `S3BackupConfiguration`，Firehose 不會提供預設值 `ErrorOutputPrefix`。AWS CLI
- 建立 `ErrorOutputPrefix` 運算式時 `partitionKeyFromLambda`，您無法使用和 `partitionKeyFromQuery` 命名空間。

## 範例字首

### Prefix 及 ErrorOutputPrefix 範例

輸入	評估的字首 (於 2018 年 8 月 27 日上午 10:30 UTC)
Prefix : 未指定  ErrorOutputPrefix : myFirehoseFailures/!{firehose:error-output-type}/	Prefix: 2018/08/27/10  ErrorOutputPrefix : myFirehoseFailures/processing-failed/
Prefix: !{timestamp:yyyy/MM/dd}  ErrorOutputPrefix : 未指定	無效輸入 : Prefix 具有表達式時，ErrorOutputPrefix 不可為零
Prefix: myFirehose/DeliveredYear=!{timestamp:yyyy}/anyMonth/rand=!{firehose:random-string}  ErrorOutputPrefix : myFirehoseFailures/!{firehose:error-output-type}/!{timestamp:yyyy}/anyMonth/!{timestamp:dd}	Prefix: myFirehose/DeliveredYear=2018/anyMonth/rand=5abf82daaa5  ErrorOutputPrefix : myFirehoseFailures/processing-failed/2018/anyMonth/10
Prefix: myPrefix/year=!{timestamp:yyyy}/month=!{times	Prefix: myPrefix/year=2018/month=07/day=06/hour=23/

輸入	評估的字首 (於 2018 年 8 月 27 日上午 10:30 UTC)
<pre>tamp:MM}/day={!{timestamp:dd}/ hour={!{timestamp:HH}/  ErrorOutputPrefix : myErrorPrefix/ year={!{timestamp:yyyy}/month=! {timestamp:MM}/day={!{timesta mp:dd}/hour={!{timestamp:HH}/! {firehose:error-output-type}</pre>	<pre>ErrorOutputPrefix : myErrorPrefix/ year=2018/month=07/day=06/hour= 23/processing-failed</pre>
<pre>Prefix: myFirehosePrefix/  ErrorOutputPrefix : 未指定</pre>	<pre>Prefix: myFirehosePrefix/2 018/08/27/  ErrorOutputPrefix : myFirehos ePrefix/processing-failed/2 018/08/27/</pre>

# 使用 Amazon 數據 Firehose AWS PrivateLink

## 適用於 Amazon 資料 Firehose 的 VPC 端點 (AWS PrivateLink)

您可以使用界面 VPC 端點來防止 Amazon VPC 和 Amazon 數據 Firehose 之間的流量離開 Amazon 網絡。介面 VPC 端點不需要網際網路閘道、NAT 裝置、VPN 連接或是 AWS Direct Connect 連線。介面虛擬私人雲端端點採用這項 AWS 技術 AWS PrivateLink，可透過 elastic network interface 與 Amazon VPC 中的私有 IP，在 AWS 服務之間進行私有通訊。如需詳細資訊，請參閱 [Amazon Virtual Private Cloud](#)。

## 使用接口 VPC 端點 ( AWS PrivateLink ) 用於 Amazon 數據 Firehose 件

若要開始使用，請建立一個介面 VPC 端點，以便讓 Amazon VPC 資源中的 Amazon 資料 Firehose 流量開始流經介面 VPC 端點。建立端點時，您可以將端點政策附加到該端點上，以控制對 Amazon Data Firehose 的存取。有關使用政策控制從 VPC 端點到 Amazon Data Firehose 的存取的詳細資訊，請參閱使用 [VPC 端點控制對服務的存取](#)。

以下範例顯示如何在 VPC 中設定 AWS Lambda 函數，並建立 VPC 端點，以允許該功能與 Amazon Data Firehose 服務安全地通訊。在此範例中，您使用政策允許 Lambda 函數列出目前區域中的 Firehose 串流，但不能描述任何 Firehose 串流。

### 建立 VPC 端點

1. 登入 AWS Management Console，並在 <https://console.aws.amazon.com/vpc/> 開啟 Amazon VPC 主控台。
2. 在 VPC 儀表板中選擇 Endpoints (端點)。
3. 選擇 Create Endpoint (建立端點)。
4. 在服務名稱清單中，選擇 `com.amazonaws.your_region.kinesis-firehose`。
5. 選擇要在其中建立端點的 VPC 和一或多個子網路。
6. 選擇要與端點關聯的一或多個安全群組。
7. 對於 Policy (政策)，選擇 Custom (自訂) 並貼上以下政策：

```
{
```

```
"Statement": [
  {
    "Sid": "Allow-only-specific-PrivateAPIs",
    "Principal": "*",
    "Action": [
      "firehose:ListDeliveryStreams"
    ],
    "Effect": "Allow",
    "Resource": [
      "*"
    ]
  },
  {
    "Sid": "Allow-only-specific-PrivateAPIs",
    "Principal": "*",
    "Action": [
      "firehose:DescribeDeliveryStream"
    ],
    "Effect": "Deny",
    "Resource": [
      "*"
    ]
  }
]
```

## 8. 選擇 建立端點。

建立要與 Lambda 函數搭配使用的 IAM 角色

1. 開啟位於 <https://console.aws.amazon.com/iam/> 的 IAM 主控台。
2. 在左側導覽窗格中，選擇角色，然後選擇建立角色。
3. 在選取可信實體的類型下，保留預設選項 AWS 服務。
4. 在 Choose the service that will use this role (選擇將使用此角色的服務) 下，選擇 Lambda (Lambda)。
5. 選擇 Next: Permissions (下一步：許可)。
6. 在政策清單中，搜尋和新增名為 AWSLambdaVPCLambdaAccessExecutionRole 和 AmazonDataFirehoseReadOnlyAccess 的兩個政策。



**⚠ Important**

以下是範例。您可能需要對生產環境採用更嚴格的政策。

7. 選擇下一步：標籤。在本練習中，您不需要新增標籤。選擇 下一步：檢閱。
8. 輸入角色名稱，然後選擇建立角色。

### 在 VPC 中建立 Lambda 函數

1. 前往 <https://console.aws.amazon.com/lambda/> 開啟 AWS Lambda 主控台。
2. 選擇 建立函數。
3. 選擇 Author from scratch (從頭開始撰寫)。
4. 輸入函數名稱，然後將 Runtime (執行時間) 設定為 Python 3.6。
5. 在 Permissions (許可) 下，展開 Choose or create an execution role (選擇或建立執行角色)。
6. 在 Execution role (執行角色) 清單中，選擇 Use an existing role (使用現有角色)。
7. 在 Existing role (現有角色) 清單中，選擇您以上建立的角色。
8. 選擇 建立函式。
9. 在 Function code (函數程式碼) 中貼上以下程式碼。

```
import json
import boto3
import os
from botocore.exceptions import ClientError

def lambda_handler(event, context):
    REGION = os.environ['AWS_REGION']
    client = boto3.client(
        'firehose',
        REGION
    )
    print("Calling list_delivery_streams with ListDeliveryStreams allowed
policy.")
    delivery_stream_request = client.list_delivery_streams()
    print("Successfully returned list_delivery_streams request %s." % (
        delivery_stream_request
    ))
```

```
describe_access_denied = False
try:
    print("Calling describe_delivery_stream with DescribeDeliveryStream
denied policy.")
    delivery_stream_info =
client.describe_delivery_stream(DeliveryStreamName='test-describe-denied')
except ClientError as e:
    error_code = e.response['Error']['Code']
    print ("Caught %s." % (error_code))
    if error_code == 'AccessDeniedException':
        describe_access_denied = True

if not describe_access_denied:
    raise
else:
    print("Access denied test succeeded.")
```

10. 在 Basic settings (基本設定) 下，將逾時設定為 1 分鐘。
11. 在 Network (網路) 下，選擇以上端點建立所在的 VPC，然後選擇子網路與您在建立安全群組時與該端點建立關聯的安全群組。
12. 請在頁面頂端附近選擇儲存。
13. 選擇 測試。
14. 輸入事件名稱，然後選擇建立。
15. 再次選擇 Test (測試)。這可讓函數得以執行。在執行結果出現後，展開 Details (詳細資訊) 並將日誌輸出與函數程式碼做比較。成功的結果會顯示區域內的 Firehose 串流清單，以及下列輸出結果：

```
Calling describe_delivery_stream.
```

```
AccessDeniedException
```

```
Access denied test succeeded.
```

## 可用性

目前支援介面 VPC 端點的區域如下：

- 美國東部 (俄亥俄)
- 美國東部 (維吉尼亞北部)

- 美國西部 (加利佛尼亞北部)
- 美國西部 (奧勒岡)
- 亞太區域 (孟買)
- 亞太區域 (首爾)
- 亞太區域 (新加坡)
- 亞太區域 (雪梨)
- 亞太區域 (東京)
- 亞太區域 (香港)
- 加拿大 (中部)
- 加拿大西部 (卡加利)
- 中國 (北京)
- 中國 (寧夏)
- 歐洲 (法蘭克福)
- 歐洲 (愛爾蘭)
- 歐洲 (倫敦)
- Europe (Paris)
- 南美洲 (聖保羅)
- AWS GovCloud (美國東部)
- AWS GovCloud (美國西部)
- 歐洲 (西班牙)
- 中東 (阿拉伯聯合大公國)
- 亞太區域 (雅加達)
- 亞太區域 (大阪)
- 以色列 (特拉維夫)

# 在 Amazon 數據 Firehose 中標記您的 Firehose 流

您可以將自己的中繼資料以標籤的形式指派給在 Amazon 資料 Firehose 中建立的 Firehose 串流。標籤是您為串流所定義的索引鍵值組。使用標籤是管理 AWS 資源和組織資料 (包括帳單資料) 的簡單而強大的方式。

## 主題

- [標籤基本概念](#)
- [使用標記追蹤成本](#)
- [標籤限制](#)
- [使用 Amazon 數據 Firehose 火管 API 標記火管流](#)

## 標籤基本概念

您可以使用 Amazon 資料 Firehose API 來完成下列任務：

- 將標籤新增至「Firehose」串流。
- 列出「Firehose」串流的標籤。
- 從「Firehose」串流移除標籤。

您可以使用標籤來分類您的 Firehose 串流。例如，您可以依目的、擁有者或環境來分類 Firehose 串流。由於您定義了每個標籤的金鑰和值，您可以建立一組自訂的類別，以符合您的特定需求。例如，您可以定義一組標籤，以協助您依擁有者和相關應用程式追蹤 Firehose 串流。

下列為數個標籤的範例：

- Project: *Project name*
- Owner: *Name*
- Purpose: Load testing
- Application: *Application name*
- Environment: Production

如果您在 `CreateDeliveryStream` 動作中指定標籤，Amazon Data Firehose 會對該 `firehose:TagDeliveryStream` 動作執行其他授權，以驗證使用者是否具有建立標籤的

權限。如果您未提供此權限，則使用 IAM 資源標籤建立新 Firehose 交付串流的請求將會失敗，AccessDeniedException 如下所示。

#### AccessDeniedException

```
User: arn:aws:sts::x:assumed-role/x/x is not authorized to perform:
  firehose:TagDeliveryStream on resource: arn:aws:firehose:us-east-1:x:deliverystream/x
  with an explicit deny in an identity-based policy.
```

下列範例示範允許使用者建立交付串流和套用標籤的原則。

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "firehose:CreateDeliveryStream",
      "Resource": "*",
    },
    {
      "Effect": "Allow",
      "Action": "firehose:TagDeliveryStream",
      "Resource": "*",
    }
  ]
}
```

## 使用標記追蹤成本

您可以使用標籤來分類和追蹤 AWS 成本。當您將標籤套用至 AWS 資源 (包括 Firehose 串流) 時，您的 AWS 成本分配報告會包含依標籤彙總的使用量和成本。套用代表商業類別的標籤 (例如成本中心、應用程式名稱或擁有者)，即可整理多個服務的成本。如需詳細資訊，請參閱《AWS Billing 使用者指南》中的[將成本分配標籤用於自訂帳單報告](#)。

## 標籤限制

下列限制適用於 Amazon 資料 Firehose 中的標籤。

## 基本限制

- 每項資源 (串流) 的標籤數上限為 50。
- 標籤金鑰與值皆區分大小寫。
- 您無法變更或編輯已刪除串流的標籤。

## 標籤鍵限制

- 每個標籤鍵都必須是唯一的。如果您新增具有已使用金鑰的標籤，則新的標籤會覆寫現有金鑰值對。
- 標籤金鑰開頭不能為 `aws:`，因為此字首保留供 AWS 使用。AWS 會代表您建立開頭為此字首的標籤，但您無法加以編輯或刪除。
- 標籤鍵的長度必須介於 1 到 128 個 Unicode 字元之間。
- 標籤鍵必須包含下列字元：Unicode 字母、數字、空格以及下列特殊字元：`_ . / = + - @`。

## 標籤值限制

- 標籤值的長度必須介於 0 到 255 個 Unicode 字元之間。
- 標籤值可以空白。否則，它們必須包含下列字元：Unicode 字母、數字、空格以及下列任何特殊字元：`_ . / = + - @`。

## 使用 Amazon 數據 Firehose 火管 API 標記火管流

您可以在呼叫 [CreateDeliveryStream](#) 以建立新 Firehose 串流時指定標籤。對於現有的交付串流時，您可以使用以下三種操作來新增、列出和移除標籤：

- [TagDeliveryStream](#)
- [ListTagsForDeliveryStream](#)
- [UntagDeliveryStream](#)

# 教學課程：使用 Amazon 資料 Firehose 將 VPC 流程日誌導入 Splunk

如需教學課程，請參閱[使用 Amazon 資料 Firehose 將 VPC 流程日誌導入 Splunk](#)。

# Amazon 數據 Firehose 故障

如果 Firehose 在傳送或處理資料時發生錯誤，則會重試，直到設定的重試持續時間到期為止。如果重試持續時間在成功交付資料之前結束，Firehose 會將資料備份到設定的 S3 備份儲存貯體。如果目的地是 Amazon S3，交付失敗或交付到備份 S3 儲存貯體失敗，Firehose 會持續重試，直到保留期結束為止。至於 DirectPut 交付串流，Firehose 會保留 24 小時的紀錄。對於資料來源為 Kinesis 資料串流的交付串流，您可以依照 [變更資料保留期間](#) 中所述變更保留期間。

如果資料來源是 Kinesis 資料串流，Firehose 會無限期地重試下列作業：DescribeStream、和。GetRecords GetShardIterator

如果交付串流使用 DirectPut，請檢查 IncomingBytes 和 IncomingRecords 指標，以查看是否有傳入流量。如果您使用的是 PutRecord 或 PutRecordBatch，請確定您擷取例外狀況並重試。我們建議搭配指數退避的重試原則，以進行抖動和多次重試。此外，如果您使用 PutRecordBatch API，請確定您的程式碼會檢查回應 [FailedPutCount](#) 中的值，即使 API 呼叫成功也是如此。

如果交付串流使用 Kinesis 資料串流作為其來源，請檢查來源資料串流的 IncomingBytes 和 IncomingRecords 指標。此外，請確定將針對交付串流發出 DataReadFromKinesisStream.Bytes 和 DataReadFromKinesisStream.Records 指標。

如需使用追蹤傳送錯誤的相關資訊 CloudWatch，請參閱 [the section called “使用 CloudWatch 記錄監控”](#)。

## 問題

- [故障診斷 Amazon S3](#)
- [疑難排解 Amazon Redshift](#)
- [Amazon OpenSearch 服務故障](#)
- [疑難排解 Splunk](#)
- [疑難排解雪](#)
- [疑難排解 Firehose 端點可達性](#)
- [疑難排解 HTTP 端點](#)
- [疑難排解 MSK As A Source](#)
- [其他](#)



## 故障診斷 Amazon S3

若資料未交付至您的 Amazon Simple Storage Service (Amazon S3) 儲存貯體，請確認下列事項。

- 檢查 Firehose IncomingBytes 和 IncomingRecords 指標，確認資料已成功傳送至您的 Firehose 串流。如需詳細資訊，請參閱 [使 CloudWatch 用指標監控 Amazon 資料 Firehose](#)。
- 如果啟用使用 Lambda 進行資料轉換，請檢查火管 ExecuteProcessingSuccess 量度，以確定 Firehose 已嘗試叫用您的 Lambda 函數。如需詳細資訊，請參閱 [使 CloudWatch 用指標監控 Amazon 資料 Firehose](#)。
- 檢查火管 DeliveryToS3.Success 指標，確定 Firehose 已嘗試將資料放入 Amazon S3 儲存貯體。如需詳細資訊，請參閱 [使 CloudWatch 用指標監控 Amazon 資料 Firehose](#)。
- 若尚未啟用錯誤記錄，請啟用這項功能並檢查傳送失敗的錯誤記錄。如需詳細資訊，請參閱 [使 CloudWatch 用日誌監控 Amazon 數據 Firehose](#)。
- 如果您在日誌中看到錯誤消息，指出「調用 Amazon S3 服務 InternalServerError 時遇到了 Firehose。操作將重試；如果錯誤仍然存在，請聯繫 S3 以解決問題。」，這可能是因為 S3 中單一分割區的請求率大幅提高。您可以優化 S3 前綴設計模式以緩解此問題。如需詳細資訊，請參閱 [最佳實務設計模式：最佳化 Amazon S3 效能](#)。如果這樣無法解決問題，請聯絡 Sup AWS port 以取得進一步協助。
- 請確定 Firehose 串流中指定的 Amazon S3 儲存貯體仍然存在。
- 若已啟用 Lambda 資料轉換功能，請確認您在交付串流中指定的 Lambda 函數仍然存在。
- 請確定 Firehose 串流中指定的 IAM 角色可以存取 S3 儲存貯體和 Lambda 函數 (如果已啟用資料轉換)。此外，請確定 IAM 角色可存取記錄群組和記錄串流，以檢查錯誤記錄。CloudWatch 如需詳細資訊，請參閱 [授予 Amazon 數據 Firehose 訪問 Amazon S3 目的地](#)。
- 若您正使用資料轉換功能，請確認 Lambda 函數的回應從未回傳大小超過 6 MB 的乘載。如需詳細資訊，請參閱 [Amazon 資料 FirehoseData 轉換](#)。

## 疑難排解 Amazon Redshift

如果資料未交付到您的 Amazon Redshift 佈建叢集或 Amazon Redshift Serverless 工作群組，請檢查下列各項。

資料會先交付至您的 S3 儲存貯體，再載入 Amazon Redshift。如果資料未傳送至您的 S3 儲存貯體，請參閱 [故障診斷 Amazon S3](#)。

- 檢查 Firehose `DeliveryToRedshift.Success` 指標，確定 Firehose 已嘗試將資料從 S3 儲存貯體複製到 Amazon Redshift 佈建的叢集或 Amazon Redshift 無伺服器工作群組。如需詳細資訊，請參閱 [使 CloudWatch 用指標監控 Amazon 資料 Firehose](#)。
- 若尚未啟用錯誤記錄，請啟用這項功能並檢查傳送失敗的錯誤記錄。如需詳細資訊，請參閱 [使 CloudWatch 用日誌監控 Amazon 數據 Firehose](#)。
- 請查看 Amazon Redshift `STL_CONNECTION_LOG` 表格，看看 Firehose 是否可以成功建立連線。本表格可供您依使用者名稱檢視連線和狀態。如需詳細資訊，請參閱《Amazon Redshift 資料庫開發人員指南》中的 [STL\\_CONNECTION\\_LOG](#)。
- 如果上述檢查顯示連線已建立，則請查看 Amazon Redshift `STL_LOAD_ERRORS` 資料表以驗證 COPY 失敗原因。如需詳細資訊，請參閱《Amazon Redshift 資料庫開發人員指南》中的 [STL\\_LOAD\\_ERRORS](#)。
- 請確定 Firehose 串流中的 Amazon Redshift 組態是否正確且有效。
- 請確定 Firehose 串流中指定的 IAM 角色可以存取 Amazon Redshift 複製資料的 S3 儲存貯體，以及用於資料轉換的 Lambda 函數 (如果啟用了資料轉換)。此外，請確定 IAM 角色可存取記錄群組和記錄串流，以檢查錯誤記錄。CloudWatch 如需詳細資訊，請參閱 [授予 Amazon 數據 Firehose 訪問亞 Amazon Redshift 目的地](#)。
- 如果您的 Amazon Redshift 佈建的叢集或 Amazon Redshift 無伺服器工作群組位於虛擬私有雲端 (VPC) 中，請確定叢集允許從 Firehose IP 地址進行存取。如需詳細資訊，請參閱 [授予 Amazon 數據 Firehose 訪問亞 Amazon Redshift 目的地](#)。
- 請確認 Amazon Redshift 佈建的叢集或 Amazon Redshift Serverless 工作群組可公開使用。
- 若您正使用資料轉換功能，請確認 Lambda 函數的回應從未回傳大小超過 6 MB 的乘載。如需詳細資訊，請參閱 [Amazon 資料 FirehoseData 轉換](#)。

## Amazon OpenSearch 服務故障

如果資料未傳送至您的 OpenSearch 服務網域，請檢查下列項目。

您可以將資料同時備份到 Amazon S3 儲存貯體。如果資料未傳送至您的 S3 儲存貯體，請參閱 [故障診斷 Amazon S3](#)。

- 檢查 Firehose `IncomingBytes` 和 `IncomingRecords` 指標，確認資料已成功傳送至您的 Firehose 串流。如需詳細資訊，請參閱 [使 CloudWatch 用指標監控 Amazon 資料 Firehose](#)。
- 如果啟用使用 Lambda 進行資料轉換，請檢查 `ExecuteProcessingSuccess` 量度，以確定 Firehose 已嘗試叫用您的 Lambda 函數。如需詳細資訊，請參閱 [使 CloudWatch 用指標監控 Amazon 資料 Firehose](#)。

- 檢查 Firehose `DeliveryToAmazonOpenSearchService.Success` 量度，確定 Firehose 已嘗試將資料索引至 OpenSearch 服務叢集。如需詳細資訊，請參閱 [使 CloudWatch 用指標監控 Amazon 資料 Firehose](#)。
- 若尚未啟用錯誤記錄，請啟用這項功能並檢查傳送失敗的錯誤記錄。如需詳細資訊，請參閱 [使 CloudWatch 用日誌監控 Amazon 數據 Firehose](#)。
- 請確定交付串流中的 OpenSearch Service 設定正確且有效。
- 若已啟用 Lambda 資料轉換功能，請確認您在交付串流中指定的 Lambda 函數仍然存在。此外，請確定 IAM 角色可存取記錄群組和記錄串流，以檢查錯誤記錄。CloudWatch 如需詳細資訊，請參閱 [FirehoseAccess 閱授與公用 OpenSearch 服務目的地](#)。
- 確保在交付串流中指定的 IAM 角色可以存取您的 OpenSearch 服務叢集、S3 備份儲存貯體和 Lambda 函數 (如果已啟用資料轉換)。此外，請確定 IAM 角色可存取記錄群組和記錄串流，以檢查錯誤記錄。CloudWatch 如需詳細資訊，請參閱 [FirehoseAccess 閱授與公用 OpenSearch 服務目的地](#)。
- 若您正使用資料轉換功能，請確認 Lambda 函數的回應從未回傳大小超過 6 MB 的乘載。如需詳細資訊，請參閱 [Amazon 資料 FirehoseData 轉換](#)。
- Amazon 數據防火安全性地不支持將日誌交付到 Amazon OpenSearch 服務目的地，因為 Amazon 將多個 CloudWatch 日誌事件 CloudWatch 合併到一個 Firehose 記錄中，並且 Amazon OpenSearch 服務無法在一個記錄中接受多個日誌事件。作為替代方案，您可以考慮在 [CloudWatch 日誌中使用 Amazon OpenSearch 服務的訂閱過濾器](#)。

## 疑難排解 Splunk

如果資料未傳送至您的 Splunk 終端節點，請確認以下步驟。

- 如果您的 Splunk 平台位於 VPC 中，請確定 Firehose 可以存取它。如需詳細資訊，請參閱 [存取 VPC 內的 Splunk](#)。
- 如果您使用 AWS 負載平衡器，請確定它是 Classic Load Balancer 或 Application Load Balancer。此外，啟用以持續時間為基礎的黏性工作階段，且針對 Classic Load Balancer 停用 Cookie 到期，且 Application Load Balancer 的到期時間上限 (7 天)。如需如何執行此作業的詳細資訊，請參閱 [Classic Load Balancer 或 Application Load Balancer 載平衡器的持續時間型工作階段黏著性](#)。
- 檢閱 Splunk 平台的要求。Firehose 的 Splunk 附加元件需要 Splunk 平台 6.X 版或更新版本。如需詳細資訊，請參閱 [Amazon Kinesis Firehose 的 Splunk 附加元件](#)。
- 如果 Firehose 和 HTTP 事件收集器 (HEC) 節點之間有代理伺服器 (Elastic Load Balancing 或其他)，請啟用黏性工作階段以支援 HEC 確認 (ACK)。

- 確認您使用的是有效的 HEC 符記。
- 確認 HEC 符記已啟用。請參閱[啟用和停用事件收集器符記](#)。
- 檢查傳送到 Splunk 的資料是否正確格式化。如需詳細資訊，請參閱[HTTP 事件收集器格式化事件](#)。
- 確認 HEC 符記和輸入事件已設定有效索引。
- 如果上傳至 Splunk 失敗原因在於 HEC 節點伺服器錯誤，請求會自動重試。如果重試失敗，資料會備份到 Amazon S3。檢查您的資料是否出現在 Amazon S3，若有則表示重試失敗。
- 確認您已在 HEC 符記上啟用 indexer 確認功能。如需詳細資訊，請參閱[啟用 indexer 確認功能](#)。
- 增加 Firehose 交付串流HECAcknowledgmentTimeoutInSeconds的 Splunk 目的地組態中的值。
- 在 Firehose 交付串流DurationInSecondsRetryOptions的 Splunk 目的地組態中，增加下方的值。
- 檢查您的 HEC 運作狀態。
- 若您正使用資料轉換功能，請確認 Lambda 函數的回應從未回傳大小超過 6 MB 的乘載。如需詳細資訊，請參閱[Amazon 資料 FirehoseData 轉換](#)。
- 請確認名為 ackIdleCleanup 的 Splunk 參數已設定為 true。預設為 false。欲將此參數設定為 true，請執行下列動作：
  - 以[受管 Splunk 雲端部署](#)而言，請提交使用 Splunk 支援入口網站的案例。此時，請尋求 Splunk 協助啟用 HTTP 事件收集器、將 inputs.conf 內的 ackIdleCleanup 設定為 true，並建立或修改負載平衡器以使用此附加元件。
  - 以[分散式 Splunk Enterprise 部署](#)而言，請在 inputs.conf 檔案中將 ackIdleCleanup 參數設定為 true。若是 \*nix 使用者，這個檔案位於 \$SPLUNK\_HOME/etc/apps/splunk\_httpinput/local/ 底下。若是 Windows 使用者，此檔案位於 %SPLUNK\_HOME%\etc\apps\splunk\_httpinput\local\ 底下。
  - 以[單一執行個體 Splunk Enterprise 部署](#)而言，請在 inputs.conf 檔案中將 ackIdleCleanup 參數設定為 true。若是 \*nix 使用者，這個檔案位於 \$SPLUNK\_HOME/etc/apps/splunk\_httpinput/local/ 底下。若是 Windows 使用者，此檔案位於 %SPLUNK\_HOME%\etc\apps\splunk\_httpinput\local\ 底下。
- 請確定 FirehoseDelivery 串流中指定的 IAM 角色可以存取 S3 備份儲存貯體和 Lambda 函數進行資料轉換 (如果已啟用資料轉換)。此外，請確定 IAM 角色可存取 CloudWatch 記錄群組和記錄串流，以檢查錯誤記錄。如需詳細資訊，請參閱[授 FirehoseAccess 予 Splunk 目的地](#)。
- 請參閱[Amazon Kinesis Firehose 的 Splunk 附加元件故障診斷](#)。

## 疑難排解雪

本節說明使用 Snowflake 作為目的地時的常見疑難排解步驟

### Firehose 交付串流建立失敗

如果將資料傳遞至已 PrivateLink 啟用雪花叢集的串流的傳遞串流建立失敗，表示 Firehose 無法存取 VPCE-ID。這可能是由於以下原因之一：

- 不正確的虛擬伺服器識別碼。確認沒有印刷錯誤。
- Firehose 在預覽中不支援無區域雪花網址。使用雪花帳戶定位器提供 URL。如需詳細資訊，請參閱 [雪花文件](#)。
- 確認 Firehose 傳送串流是在與「雪花區域」相同的 AWS 區域中建立。
- 如果問題仍然存在，請聯繫 AWS 支持。

### 傳送失敗

如果資料未傳送到 Snowflake 資料表，請檢查下列項目。雪花傳遞失敗的資料將連同錯誤碼和承載對應的錯誤訊息一起傳送至 S3 錯誤儲存貯體。以下是一些常見的錯誤情況。如需錯誤代碼的完整清單，請參閱 [雪花資料傳送錯誤](#)。

- 錯誤代碼：雪花。DefaultRoleMissing：表示建立傳遞串流時未設定雪花角色。如果未配置「雪花」角色，那么請確保您將默認角色設置為「雪花」用戶指定的角色。
- 錯誤代碼：雪花。ExtraColumns：表示由於輸入有效負載中的額外欄，因此拒絕插入至 Snowflake。不應指定表格中不存在的資料行。請注意，雪花欄名稱是區分大小寫的。如果儘管資料行存在於資料表中，但傳遞失敗並出現此錯誤，請確定輸入裝載中的資料行名稱大小寫與資料表定義中宣告的資料行名稱相符。
- 錯誤代碼：雪花。MissingColumns：表示由於輸入有效負載中缺少資料行，因此拒絕插入至 Snowflake。確保為所有不可為空的列指定了值。
- 錯誤代碼：雪花。InvalidInput：當 Firehose 無法將提供的輸入裝載解析為有效的 JSON 格式時，可能會發生這種情況。確保 json 有效負載格式正確，沒有額外的雙引號，引號，轉義字符等。目前 Firehose 僅支持單個 JSON 項目作為記錄有效負載，不支持 JSON 數組。
- 錯誤代碼：雪花。InvalidValue：指出由於輸入有效負載中的資料類型不正確而傳送失敗。請確定輸入裝載中指定的 JSON 值符合 Snowflake 資料表定義中宣告的資料類型。
- 錯誤代碼：雪花。InvalidTableType：表示不支援在傳遞串流中設定的資料表類型。如需支援的資料表、資料行和資料類型，請參閱 Snowpipe 串流的 [限制](#) 中的限制。



**Note**

基於任何原因，如果在建立傳遞串流之後，Snowflake 目的地上的資料表定義或角色權限有所變更，Firehose 可能需要幾分鐘的時間才能偵測到這些變更。如果您因此看到傳送錯誤，請嘗試刪除並重新建立傳送串流。

## 疑難排解 Firehose 端點可達性

如果 Firehose API 遇到逾時，請執行下列步驟來測試端點可達性：

- 檢查 API 請求是否從 VPC 中的主機發出。來自 VPC 的所有流量都需要設定 Firehose VPC 端點。如需詳細資訊，請參閱[搭 AWS PrivateLink 配使用 Firehose](#)。
- 如果流量來自公用網路或 VPC，且 Firehose VPC 端點已在特定子網路中設定，請從主機執行下列命令以檢查網路連線。您可以在 Firehose 端點[和配額上找到 Firehose 端點](#)。
- 使用 traceroute 或之類的工具 tcping 來檢查網路設置是否正確。如果失敗，請檢查您的網路設定：

例如：

```
tracert firehose.us-east-2.amazonaws.com
```

或

```
tcping firehose.us-east-2.amazonaws.com 443
```

- 如果網路設定正確且下列命令失敗，請檢查 [Amazon CA \(認證授權單位\)](#) 是否位於信任鏈中。

例如：

```
curl firehose.us-east-2.amazonaws.com
```

如果上述命令成功，請再次嘗試 API 以查看是否存在從 API 返回的響應。

## 疑難排解 HTTP 端點

本節說明使用 Amazon 資料 Firehose 將資料傳送到一般 HTTP 端點目的地和合作夥伴目的地 (包括資料多、重建、MongoDB、新遺物、Splunk 或 Sumo 邏輯) 時的常見疑難排解步驟。LogicMonitor 就本

節而言，所有適用的目的地均稱為 HTTP 端點。請確定 Firehose 交付串流中指定的 IAM 角色可以存取 S3 備份儲存貯體和 Lambda 函數進行資料轉換 (如果已啟用資料轉換)。此外，請確定 IAM 角色可存取記錄群組和記錄串流，以檢查錯誤記錄。CloudWatch 如需詳細資訊，請參閱[將 Firehose 存取權授與 HTTP 端點目的地](#)。

### Note

本節中的資訊不適用於下列目的地：Splunk、OpenSearch 服務、S3 和 Redshift。

## CloudWatch 日誌

強烈建議您啟用 [Firehose 的 CloudWatch 記錄](#) 功能。只有在交付到目的地時發生錯誤時，才會發佈日誌。

### 目的地例外狀況

ErrorCode: HttpEndpoint.DestinationException

```
{
  "deliveryStreamARN": "arn:aws:firehose:us-east-1:123456789012:deliverystream/ronald-test",
  "destination": "custom.firehose.endpoint.com...",
  "deliveryStreamVersionId": 1,
  "message": "The following response was received from the endpoint destination.
413: {\"requestId\": \"43b8e724-dbac-4510-adb7-ef211c6044b9\", \"timestamp\": 1598556019164, \"errorMessage\": \"Payload too large\"}",
  "errorCode": "HttpEndpoint.DestinationException",
  "processor": "arn:aws:lambda:us-east-1:379522611494:function:httpLambdaProcessing"
}
```

目標例外狀況表示 Firehose 能夠與您的端點建立連線並發出 HTTP 請求，但未收到 200 個回應碼。並非 200 的 2xx 個回應也會導致目的地例外狀況。Amazon Data Firehose 會將 CloudWatch 從設定的端點收到的回應代碼和截斷的回應承載記錄記錄到日誌。由於 Amazon Data Firehose 會記錄回應代碼和有效負載而無需修改或解釋，因此端點必須提供拒絕 Amazon Data Firehose HTTP 交付請求的確切原因。以下是針對這些例外狀況的最常見疑難排解建議：

- 400：表示您因 Amazon 資料 Firehose 設定錯誤而傳送錯誤的請求。請確認您有正確的 [url](#)、[一般屬性](#)、[內容編碼](#)、[存取金鑰](#)，以及目的地 [緩衝提示](#)。請參閱有關所需組態的目的地特定文件。

- 401：表示您為 Firehose 串流設定的存取金鑰不正確或遺失。
- 403：表示您為 Firehose 串流設定的存取金鑰沒有將資料傳遞至已設定端點的權限。
- 413：表示 Amazon Data Firehose 傳送至端點的請求承載資料過大，無法處理。嘗試將[緩衝提示降低](#)到目的地的建議大小。
- 429：表示 Amazon 資料 Firehose 傳送請求的速率高於目的地所能處理的速率。透過增加緩衝時間和/或增加緩衝大小 (但仍在目的地限制內) 來微調緩衝提示。
- 5xx：表示目的地發生問題。Amazon 數據 Firehose 服務仍然正常工作。

### ⚠ Important

**重要事項：**雖然這些是常見的疑難排解建議，但特定端點可能有不同的原因提供回應代碼，並且應先遵循端點特定建議。

## 無效的回應

ErrorCode: HttpEndpoint.InvalidResponseFromDestination

```
{
  "deliveryStreamARN": "arn:aws:firehose:us-east-1:123456789012:deliverystream/ronald-test",
  "destination": "custom.firehose.endpoint.com...",
  "deliveryStreamVersionId": 1,
  "message": "The response received from the specified endpoint is invalid. Contact the owner of the endpoint to resolve the issue. Response for request 2de9e8e9-7296-47b0-bea6-9f17b133d847 is not recognized as valid JSON or has unexpected fields. Raw response received: 200 {\"requestId\": null}\",
  "errorCode": "HttpEndpoint.InvalidResponseFromDestination",
  "processor": "arn:aws:lambda:us-east-1:379522611494:function:httpLambdaProcessing"
}
```

無效的回應例外狀況表示 Amazon 資料 Firehose 收到來自端點目的地的無效回應。回應必須符合[回應規格](#)，否則 Amazon Data Firehose 會將交付嘗試視為失敗，並會重新傳送相同的資料，直到超過設定的重試持續時間為止。即使回應狀態為 200，Amazon 資料 Firehose 也會將不遵循回應規格的回應視為失敗。如果您要開發與 Amazon Data Firehose 相容的端點，請遵循回應規格以確保資料成功交付。

以下是一些常見的無效回覆類型，以及修正這些問題的方法：



- 無效的 JSON 或非預期欄位：表示回應無法正確還原序列化為 JSON，或具有非預期的欄位。請確認回應未進行內容編碼。
- 遺漏 RequestId：表示回應不包含「requestId 項目」。
- RequestId 不相符：表示回應中的 requestId 項目不符合傳出的 requestId ID。
- 缺失時間戳記：表示回應不包含時間戳記欄位。時間戳記欄位必須是數字，而不是字串。
- 缺失內容類型標題：表示回應不包含「內容類型：應用程式/json」標頭。沒有接受的其他內容類型。

### Important

**重要事項：**Amazon 資料 Firehose 只能將資料傳送到符合 Firehose 請求和**回應**規格的端點。如果您要將目的地設定為第三方服務，請確保您使用的是正確的 Amazon Data Firehose 相容端點，該端點可能與公用擷取端點不同。[例如，數據多的 Amazon 數據 Firehose 端點是 https://aws-kinesis-http-intake](https://aws-kinesis-http-intake)。 <https://api.datadoghq.com/>

## 其他常見錯誤

下面列有其他錯誤代碼和定義。

- 錯誤代碼: HttpEndpoint. RequestTimeout-指示端點所花費的時間超過 3 分鐘來回應。如果您是目的地的擁有者，則請縮短目的地端點的回應時間。如果您不是目的地的擁有者，請聯絡擁有者並詢問是否可以採取任何措施來縮短回應時間 (減少緩衝提示，以減少每個請求處理的資料)。
- 錯誤代碼: HttpEndpoint. ResponseTooLarge-表示回應太大。回應必須小於 1 MiB (包括標頭)。
- 錯誤代碼: HttpEndpoint. ConnectionFailed-表示無法與設定的端點建立連線。這可能是因為設定的網址中有錯字、Amazon Data Firehose 無法存取端點，或者端點花費太長時間來回應連線要求所致。
- 錯誤代碼: HttpEndpoint. ConnectionReset-指示已建立連線，但端點已重設或過早關閉。
- 錯誤代碼: HttpEndpoint.SSL HandshakeFailure-表示 SSL 交涉無法成功與設定的端點完成。

## 疑難排解 MSK As A Source

本節會描述使用 MSK As Source 時的常見疑難排解步驟

**Note**

如需疑難排解處理、轉換或 S3 交付問題，請參閱前面的章節

## Hose 建立失敗

如果您使用 MSK As Source 的 hose 建立失敗，請檢查以下內容

- 檢查來源 MSK 叢集是否處於「作用中」狀態。
- 如果您正在使用私有連線，則請確認[叢集上的私有連結已開啟](#)

如果您使用的是公有連線，則請確認[叢集上的公有存取已開啟](#)

- 如果您正在使用私有連線，則請確認您新增[可讓 Firehose 建立私有連結的資源型政策](#)。另請參閱：[MSK 跨帳戶許可](#)
- 確認來源組態中的角色具有[從叢集主題擷取資料的許可](#)
- 確保您的 VPC 安全群組允許[叢集啟動程序伺服器所使用的連接埠](#)上的傳入流量

## Hose 暫停

如果您的 hose 處於懸掛狀態，請檢查以下內容

- 檢查來源 MSK 叢集是否處於「作用中」狀態。
- 檢查來源主題是否存在。如果主題被刪除並重新創建，則還必須刪除並重新創建 Firehose Firehose 流。

## Hose 預壓

當超過 BytesPerSecondLimit 每個分割區或正常傳送流程緩慢或停止時，DataReadFromSource.Backpressured 的值將會是 1。

- 如果您正在擊中，BytesPerSecondLimit 請檢查 DataReadFromSource.Bytes 度量並請求提高限制。
- 檢查 CloudWatch 日誌、目的地指標、資料轉換指標和格式轉換指標，以識別瓶頸。

## 資料新鮮度錯誤

資料新鮮度似乎有誤

- Firehose 會根據使用記錄的時間戳記來計算資料新鮮度。為了確保在生產者記錄保留在 Kafka 的代理日誌中時正確記錄此時間戳記，請將 Kafka 主題時間戳記類型組態設定為 `message.timestamp.type=LogAppendTime`。

## MSK 叢集連線問題

下列程序說明如何驗證 MSK 叢集的連線能力。如需設定 Amazon MSK 用戶端的詳細資訊，請參閱 Amazon 管理串流開發人員指南中的 Amazon MSK 開發人員指南中的[開始使用亞馬遜 MSK](#)。

驗證與 MSK 叢集的連線

1. 創建一個基於 UNIX 的 ( 最好是 AL2 ) Amazon EC2 實例。如果您的叢集上僅啟用 VPC 連線，請確保您的 EC2 執行個體在相同的 VPC 中執行。一旦執行個體可用，即可安全殼層連線 如需詳細資訊，請參閱 Amazon EC2 Linux 執行個體使用者指南中的[本教學課程](#)。
2. 執行下列命令，使用 Yum 套件管理員來安裝 Java。如需詳細資訊，請參閱 Amazon Corretto 8 使用者[指南中的安裝說明](#)。

```
sudo yum install java-1.8.0
```

3. 執行下列命令來安裝[AWS 用戶端](#)。

```
curl "https://awscli.amazonaws.com/awscli-exe-linux-x86_64.zip" -o "awscliv2.zip"  
unzip awscliv2.zip  
sudo ./aws/install
```

4. 通過運行以下命令下載阿帕奇卡夫卡客戶端 2.6\* 版本。

```
wget https://archive.apache.org/dist/kafka/2.6.2/kafka_2.12-2.6.2.tgz  
tar -xzf kafka_2.12-2.6.2.tgz
```

5. 前往 `kafka_2.12-2.6.2/libs` 目錄，然後執行下列命令以下載 Amazon MSK IAM JAR 檔案。

```
wget https://github.com/aws/aws-msk-iam-auth/releases/download/v1.1.3/aws-msk-iam-auth-1.1.3-all.jar
```

6. 在卡夫卡 `bin/client.properties` 文件夾中創建文件。
7. 以您在 Firehose 中使用的角色 ARN 取 `awsRoleArn` 代，`SourceConfiguration` 並驗證憑證位置。允許您的用 AWS 戶端使用者擔任角色 `awsRoleArn`。AWS 用戶端使用者將嘗試擔任您在此處指定的角色。

```
[ec2-user@ip-xx-xx-xx-xx bin]$ cat client.properties
security.protocol=SASL_SSL
sasl.mechanism=AWS_MSK_IAM
sasl.jaas.config=software.amazon.msk.auth.iam.IAMLoginModule required
  awsRoleArn="<role arn>" awsStsRegion="<region name>";
sasl.client.callback.handler.class=software.amazon.msk.auth.iam.IAMClientCallbackHandler
awsDebugCreds=true
ssl.truststore.location=/usr/lib/jvm/java-1.8.0-
openjdk-1.8.0.342.b07-1.amzn2.0.1.x86_64/jre/lib/security/cacerts
ssl.truststore.password=changeit
```

8. 運行以下 Kafka 命令列出主題。如果您的連線是公用的，請使用公用端點啟動程序伺服器。如果您的連線是私人的，請使用私有端點啟動程序伺服器。

```
bin/kafka-topics.sh --list --bootstrap-server <bootstrap servers> --command-config
bin/client.properties
```

如果要求成功，您應該會看到類似下列範例的輸出。

```
[ec2-user@ip-xx-xx-xx-xx kafka_2.12-2.6.2]$ bin/kafka-topics.sh --list --bootstrap-
server <bootstrap servers> --command-config bin/client.properties

[xxxx-xx-xx 05:49:50,877] WARN The configuration 'awsDebugCreds' was supplied but
isn't a known config. (org.apache.kafka.clients.admin.AdminClientConfig)
[xxxx-xx-xx 05:49:50,878] WARN The configuration 'ssl.truststore.location' was
supplied but isn't a known config.
(org.apache.kafka.clients.admin.AdminClientConfig)
[xxxx-xx-xx 05:49:50,878] WARN The configuration 'sasl.jaas.config' was supplied
but isn't a known config. (org.apache.kafka.clients.admin.AdminClientConfig)
[xxxx-xx-xx 05:49:50,878] WARN The configuration
'sasl.client.callback.handler.class' was supplied but isn't a known config.
(org.apache.kafka.clients.admin.AdminClientConfig)
[xxxx-xx-xx 05:49:50,878] WARN The configuration 'ssl.truststore.password' was
supplied but isn't a known config.
(org.apache.kafka.clients.admin.AdminClientConfig)
```

```
[xxxx-xx-xx 05:50:21,629] WARN [AdminClient clientId=adminclient-1] Connection to
node...
__amazon_msk_canary
__consumer_offsets
```

9. 如果您在執行上一個指令碼時遇到任何問題，請確認您提供的啟動程序伺服器可透過指定的連接埠存取。要做到這一點，你可以下載並使用 telnet 或類似的實用程序，如下面的命令。

```
sudo yum install telnet
telnet <bootstrap servers><port>
```

如果請求成功，你會得到下面的輸出。這表示您可以在本機 VPC 中連線至 MSK 叢集，而啟動程序伺服器在指定的連接埠上運作良好。

```
Connected to ..
```

10. 如果要求不成功，請檢查 VPC [安全性群組](#) 上的輸入規則。例如，您可以在輸入規則上使用下列屬性。

```
Type: All traffic
Port: Port used by the bootstrap server (e.g. 14001)
Source: 0.0.0.0/0
```

重試上一個步驟所示的 telnet 連線。如果您仍然無法連線，或 Firehose 連線仍然失敗，請聯絡 [AWS 支援人員](#)。

## 其他

### 主題

- [傳送串流無法做為 CloudWatch 記錄檔、CloudWatch 事件或 AWS IoT 動作的目標使用](#)
- [資料新鮮度指標不斷增加或未發出](#)
- [Apache Parquet 的記錄格式轉換失敗](#)
- [雖然指標良好，但目的地卻沒有資料](#)

## 傳送串流無法做為 CloudWatch 記錄檔、 CloudWatch 事件或 AWS IoT 動作的目標使用

某些 AWS 服務只能將訊息和事件傳送至位於相同 AWS 地區的 Firehose 交付串流。確認您的 Firehose 交付串流與其他服務位於相同的區域。

### 資料新鮮度指標不斷增加或未發出

資料新鮮度是衡量資料目前在交付串流中的程度。這是交付串流中最舊資料記錄的時間，從 Firehose 擷取資料到目前的時間為止。Firehose 提供可用來監控資料新鮮度的指標。若要識別指定目的地的資料新鮮度指標，請參閱 [the section called “以測 CloudWatch 量結果監視”](#)。

如果您啟用所有事件或所有文件的備份，請監控兩個不同的資料新鮮度指標：一個針對主要目的地，另一個則針對備份。

如果未發出資料新鮮度指標，這表示交付串流沒有作用中的交付。資料交付完全封鎖時或沒有傳入資料時，就會發生這種情況。

如果資料新鮮度指標不斷增加，這表示資料交付落後。這種情況可能是由於下列其中一個原因而發生的。

- 目的地無法處理交付率。如果 Firehose 由於高流量而遇到暫時性錯誤，則交付可能會落後。這可能發生在 Amazon S3 以外的目的地（這可能發生在 OpenSearch 服務，Amazon Redshift 或 Splunk）。確定您的目的地有足夠的容量來處理傳入流量。
- 目的地很慢。如果 Firehose 遇到高延遲時間，資料傳遞可能會落後。監控目的地的延遲指標。
- Lambda 函數緩慢。這可能會導致資料交付率低於交付串流的資料擷取率。如果可能的話，請提高 Lambda 函數的效率。例如，如果此函數執行網路 IO，請使用多個執行緒或非同步 IO 來增加並行性。此外，請考慮增加 Lambda 函數的記憶體大小，以便 CPU 配置可以相應地增加。這可能會導致更快的 Lambda 調用。如需設定 Lambda 函數的相關資訊，請參閱 [設定 AWS Lambda 函數](#)
- 資料交付期間發生失敗。如需如何使用 Amazon CloudWatch 日誌監控錯誤的相關資訊，請參閱 [the section called “使用 CloudWatch 記錄監控”](#)。
- 如果交付串流的資料來源是 Kinesis 資料串流，則可能會發生調節。檢查 `ThrottledGetRecords`、`ThrottledGetShardIterator` 和 `ThrottledDescribeStream` 指標。如果有多個取用者附加到 Kinesis 資料串流，請考慮下列情況：
  - 如果 `ThrottledGetRecords` 和 `ThrottledGetShardIterator` 指標很高，建議您增加針對資料串流佈建的碎片數目。

- 如果高 `ThrottledDescribeStream`，建議您將 `kinesis:listshards` 權限新增至中設定的角色 [KinesisStreamSourceConfiguration](#)。
- 目的地的緩衝提示很低。這可能會增加 Firehose 需要到達目的地的往返次數，這可能會導致遞送落後。請考慮增加緩衝提示的值。如需詳細資訊，請參閱 [BufferingHints](#)。
- 當錯誤頻繁發生時，高重試持續時間可能會導致交付落後。請考慮減少重試持續時間。此外，監控錯誤並嘗試減少錯誤。如需如何使用 Amazon CloudWatch 日誌監控錯誤的相關資訊，請參閱 [the section called “使用 CloudWatch 記錄監控”](#)。
- 如果目的地是 Splunk，而且 `DeliveryToSplunk.DataFreshness` 很高，但 `DeliveryToSplunk.Success` 看起來不錯，Splunk 叢集可能忙碌中。如果可能的話，釋出 Splunk 叢集。或者，請聯絡 Sup AWS port 部門，並要求增加 Firehose 用來與 Splunk 叢集通訊的頻道數量。

## Apache Parquet 的記錄格式轉換失敗

如果您採用包含 `Set` 類型的 DynamoDB 資料，透過 Lambda 將其串流至交付串流，並使用將記錄格式轉換為 Apache 實木複合地板，就會發生這種情況。AWS Glue Data Catalog

當 AWS Glue 爬行者程式為 DynamoDB 設定的資料類型 (`StringSet`、和 `BinarySet`) 編製索引時 `NumberSet`，會將它們分別儲存在資料目錄中的形式為 `SET<STRING>`、`SET<BIGINT>`、和 `SET<BINARY>`。但是，為了使 Firehose 將數據記錄轉換為 Apache 實木複合地板格式，它需要 Apache 蜂巢數據類型。因為集合類型不是有效的 Apache Hive 資料類型，轉換會失敗。若要讓轉換正常運作，請以 Apache Hive 資料類型更新資料目錄。您可以在資料目錄中將 `set` 變更為 `array` 來這麼做。

**array** 在資料目錄中將一個或多個資 AWS Glue 料類型從變更 **set** 為

1. 請登入 AWS Management Console 並開啟 AWS Glue 主控台，網址為 <https://console.aws.amazon.com/glue/>。
2. 在左側窗格的 Data catalog (資料目錄) 標題下，選擇 Tables (資料表)。
3. 在資料表清單中，選擇您需要修改一或多個資料類型的資料表名稱。您將會移至該資料表的詳細資訊頁面。
4. 選擇詳細資訊頁面右上角的編輯結構描述按鈕。
5. 在 Data type (資料類型) 欄中，選擇第一個 `set` 資料類型。
6. 在 Column type (欄類型) 下拉式清單中，將類型從 `set` 變更為 `array`。

7. 在ArraySchema欄位中array<binary>，根據您案例的適當資料類型輸入array<string>、或。array<int>
8. 選擇更新。
9. 重複上述步驟，將其他 set 類型轉換為 array 類型。
10. 選擇儲存。

## 雖然指標良好，但目的地卻沒有資料

如果沒有資料擷取問題，且針對交付串流發出的指標看起來不錯，但是您在目的地看不到資料，請檢查讀取器邏輯。確認您的讀者正確解析出所有資料。



# Amazon 數據 Firehose 配額

Amazon 數據 Firehose 件具有以下配額。

- 使用 Amazon MSK 做為 Firehose 串流的來源，每個 Firehose 串流的預設配額為每個磁碟分割 10 MB /秒讀取輸送量和10MB 的最大記錄大小。您可以使用[增加服務配額](#)來要求增加每個分割區每秒 10 MB/秒的讀取輸送量的預設配額。
- 使用 Amazon MSK 做為 Firehose 串流的來源時，如果啟用 Lambda，則記錄大小上限為 6Mb；如果停用 AWS Lambda，則最大記錄大小為 10Mb。AWS Lambda 將其傳入記錄上限為 6 MB，而 Amazon 資料 Firehose 體會將 6Mb 以上的記錄轉送至錯誤 S3 儲存貯體。如果 Lambda 已停用，Firehose 會將其傳入的記錄限制為 10 MB。如果 Amazon 資料火管從 Amazon MSK 接收到大於 10MB 的記錄大小，則 Amazon 資料 Firehose 會將此記錄傳送到 S3 錯誤儲存貯體，並將 Cloudwatch 指標發送到您的帳戶。如需 AWS Lambda 限制的詳細資訊，請參閱：<https://docs.aws.amazon.com/lambda/latest/dg/gettingstarted-limits.html>。
- 啟用交付串流上的[動態分割區](#)時，可為該交付串流建立 500 個作用中分割區的預設配額。作用中分割區計數是交付緩衝區內的作用中分割區總數。例如，如果動態分割查詢每秒建構 3 個分割區，而且您的緩衝區提示組態會每 60 秒觸發交付，則平均而言，您就會有 180 個作用中分割區。在分割區中交付資料後，此分割區就不再處於作用中狀態。您可以使用[Amazon 資料 Firehose 限制表單](#)，請求將此配額增加至每個指定交付串流最多 5000 個作用中分區。如果您需要更多分割區，則可以建立更多交付串流，並將作用中分割區分散到這些交付串流之中。
- 啟用交付串流上的[動態分割](#)時，每個作用中分割區支援每秒 1 GB 的輸送量上限。
- 每個帳號在每個地區的 Firehose 傳送串流數量有以下配額：
  - 美國東部 (維吉尼亞北部)、美國東部 (俄亥俄)、美國西部 (奧勒岡)、歐洲 (愛爾蘭)、亞太區域 (東京)：5,000 個交付串流
  - 歐洲 (法蘭克福)、歐洲 (倫敦)、亞太區域 (新加坡)、亞太區域 (雪梨)、亞太區域 (首爾)、亞太區域 (孟買)、AWS GovCloud (美國西部)、加拿大 (西部)、加拿大 (中部)：2,000 個交付串流
  - 歐洲 (巴黎)、歐洲 (米蘭)、歐洲 (斯德哥爾摩)、亞太區域 (香港)、亞太區域 (大阪)、南美洲 (聖保羅)、中國 (寧夏)、中國 (北京)、中東 (巴林)、AWS GovCloud (美國東部)、非洲 (開普敦)：500 個送貨流
  - 歐洲 (蘇黎世)、歐洲 (西班牙)、亞太區域 (海德拉巴)、亞太區域 (雅加達)、亞太區域 (墨爾本)、中東 (阿拉伯聯合大公國)、以色列 (特拉維夫)、加拿大西部 (卡加利)、加拿大 (中部)：100 個送貨串流
- 如果您超過此限制，則呼叫 [CreateDeliveryStream](#) 會導致 `LimitExceededException` 例外狀況。若要增加此配額，您可以使用 [Service Quotas](#) (如果您的區域提供的話)。如需有關使用

Service Quotas 的詳細資訊，請參閱[請求配額提升](#)。如果您的區域無法使用 Service Quotas，您可以使用 [Amazon 資料 Firehose 限制表單](#) 申請增加。

- 當 Direct PUT 設定為資料來源時，每個 Firehose 串流會針對[PutRecord](#)和[PutRecordBatch](#)要求提供下列組合配額：
  - 對於美國東部 (維吉尼亞北部)、美國西部 (奧勒岡) 和歐洲 (愛爾蘭)：500,000 筆記錄/秒、2,000 個請求/秒和 5 MiB/秒。
  - 美國東部 (俄亥俄)、美國西部 (加利佛尼亞北部)、AWS GovCloud (美國東部)、AWS GovCloud (美國西部)、亞太區域 (香港)、亞太區域 (孟買)、亞太區域 (首爾)、亞太區域 (新加坡)、中國 (北京)、中國 (寧夏)、亞太區域 (雪梨)、亞太區域 (東京)、加拿大 (中部)、加拿大西部 (卡加利)、歐洲 (歐洲)、法蘭克福巴黎、歐洲 (斯德哥爾摩)、中東 (巴林)、南美洲 (聖保羅)、非洲 (開普敦) 和歐洲 (米蘭)：每秒 100,000 筆記錄、每秒 1,000 次請求，以及每秒 1 MB

若要要求提高配額，請使用 [Amazon 資料 Firehose 限制表單](#)。這三個配額會按比例提高。例如，如果您將美國東部 (維吉尼亞北部)、美國西部 (奧勒岡) 或歐洲 (愛爾蘭) 中的輸送量配額提升至 10 MiB/秒，則另外兩個配額會提升至 4,000 個請求/秒，以及 1,000,000 筆記錄/秒。

#### Important

如果提高的配額遠高於執行中的流量，則會導致傳送至目的地的批次變得極小，既缺乏效率且可能使目的地服務產生更多成本。提高配額時，請務必符合目前執行中的流量，如果流量增加，應進一步提高配額。

#### Important

請注意，較小的資料記錄可能會帶來更高的成本。[Firehose 擷取定價](#)是根據您傳送至服務的資料記錄數量，乘以每筆記錄的大小四捨五入至最接近的 5KB (5120 位元組)。因此，對於相同的傳入資料 (位元組) 量，如果有更多的傳入記錄，則產生的成本會更高。例如，如果傳入資料量總計為 5 MiB，則與使用 1,000 筆記錄傳送相同數量的資料相比，傳送超過 5,000 筆記錄的 5 MiB 資料會花費更高。如需詳細資訊，請參閱[AWS 計算機](#)中的 Amazon 資料 Firehose。

**Note**

將 Kinesis Data Streams 設定為資料來源時，此配額不適用，而且 Amazon 資料 Firehose 無限制地上下擴展。

- 每個 Firehose 串流會儲存最多 24 小時的資料記錄，以防傳送目的地無法使用且來源為 DirectPut 何。如果來源是 Kinesis Data Streams (KDS) 且目的地無法使用，則會根據您的 KDS 組態保留資料。
- 在 base64 編碼之前，傳送至 Amazon 資料 Firehose 的記錄大小上限為 1,000 KiB。
- [PutRecordBatch](#) 操作每次呼叫可處理的上限為 500 個記錄或 4 MiB，以較小者為準。此配額無法變更。
- 下列操作每秒最多可提供五次調用 (這是硬性限制)：[CreateDeliveryStream](#)、[DeleteDeliveryStream](#)、[DescribeDeliveryStream](#)、[ListDeliver](#)
- 緩衝提示間隔介於 60 秒到 900 秒。
- 對於從 Amazon 數據 Firehose 件交付到 Amazon Redshift，只支持可公開訪問的 Amazon Redshift 集群。
- Amazon Redshift 和 OpenSearch 服務交付的重試持續時間範圍從 0 秒到 7,200 秒。
- Firehose 支持彈性搜索版本 1.5，2.3，5.1，5.3，5.5，5.6，以及所有 6.\* 和 7.\* 版本和 Amazon OpenSearch 服務 2.x 最高 2.11。
- 當目的地是 Amazon S3、Amazon Redshift 或 OpenSearch 服務時，Amazon 數據 Firehose 允許每個碎片最多 5 次未完成的 Lambda 調用。對於 Splunk，配額為每個碎片 10 個未完成的 Lambda 調用。
- 您可以使 CUSTOMER\_MANAGED\_CMK 類型的 CMK 來加密最多 500 個交付串流。

## 附錄 – HTTP 端點交付請求和回應規格

為了讓 Amazon 資料 Firehose 能夠成功地將資料交付到自訂 HTTP 端點，這些端點必須接受請求並使用特定 Amazon Data Firehose 請求和回應格式傳送回應。本節說明 Amazon 資料 Firehose 服務傳送至自訂 HTTP 端點之 HTTP 請求的格式規格，以及 Amazon 資料 Firehose 服務預期的 HTTP 回應格式規格。在 Amazon 資料 Firehose 逾時該請求之前，HTTP 端點有 3 分鐘的時間來回應請求。Amazon 資料 Firehose 會將不遵守正確格式的回應視為交付失敗。

### 主題

- [要求格式](#)
- [回應格式](#)
- [範例](#)

## 要求格式

### 路徑和 URL 參數

這些是由您作為單一 URL 欄位的一部分直接進行設定。Amazon 數據 Firehose 將它們按配置發送而無需修改。僅 https 目的地受支援。系統會在交付串流組態期間套用 URL 限制。

#### Note

目前，HTTP 端點資料交付僅支援連接埠 443。

### HTTP 標頭 – X-Amz-Firehose-Protocol-Version

此標頭用於指示請求/回應格式的版本。目前僅支援 1.0 版本。

### HTTP 標頭 – X-Amz-Firehose-Request-Id

此標頭的值是不透明的 GUID，可用於偵錯和重複資料刪除目的。如果可能，端點實作應該記錄此標頭的值，以滿足成功和不成功的請求。在相同請求的多次嘗試之間，請求 ID 保持不變。

### HTTP 標頭 – Content-Type

Content-Type 標頭的值永遠是 application/json。

## HTTP 標頭 – Content-Encoding

Firehose 串流可以設定為在傳送要求時使用 GZIP 壓縮主體。啟用此壓縮時，根據標準實務，Content-Encoding 標頭的值會設定為 gzip。如果未啟用壓縮，則 Content-Encoding 標頭完全不存在。

## HTTP 標頭 – Content-Length

以標準方式對其進行使用。

## HTTP 標頭 – X-Amz-Firehose-Source-Arn :

Firehose 串流的 ARN 以 ASCII 字串格式表示。ARN 對區域編碼，AWS 帳戶 ID 和串流名稱。例如 `arn:aws:firehose:us-east-1:123456789:deliverystream/testStream`。

## HTTP 標頭 – X-Amz-Firehose-Access-Key

此標頭帶有 API 金鑰或其他憑證。您可以在建立或更新交付串流時建立或更新 API 金鑰 (也稱為授權記號)。Amazon 數據 Firehose 將訪問密鑰的大小限制為 4096 字節。Amazon 數據 Firehose 不會嘗試以任何方式解釋此密鑰。設定的金鑰會逐字複製到此標頭的值中。

內容可以是任意的，並且可能代表 JWT 記號或 ACCESS\_KEY。如果端點需要多欄位憑證 (例如，使用者名稱和密碼)，則所有欄位的值都應以端點理解的格式 (JSON 或 CSV) 儲存在單一存取金鑰中。如果原始內容是二進位，則此欄位可以是 base-64 編碼。Amazon 資料 Firehose 不會修改和/或編碼已設定的值，並按原樣使用內容。

## HTTP 標頭 – X-Amz-Firehose-Common-Attributes

該標頭包含與整個請求和/或請求中的所有記錄有關的共同屬性 (中繼資料)。這些是您在建立 Firehose 串流時直接設定的。系統會將此屬性的值編碼為具有下列結構描述的 JSON 物件：

```
"$schema": http://json-schema.org/draft-07/schema#
properties:
  commonAttributes:
    type: object
    minProperties: 0
    maxProperties: 50
    patternProperties:
      "^.{1,256}$":
        type: string
        minLength: 0
```

```
maxLength: 1024
```

範例如下：

```
"commonAttributes": {  
  "deployment -context": "pre-prod-gamma",  
  "device-types": ""  
}
```

## 主體 – 大小上限

壓縮之前，主體大小上限由您設定，最大可達 64 MiB。

## 主體 – 結構描述

主體包含具有下列 JSON 結構描述 (以 YAML 寫入) 的單一 JSON 文件：

```
"$schema": http://json-schema.org/draft-07/schema#  
  
title: FirehoseCustomHttpsEndpointRequest  
description: >  
  The request body that the Firehose service sends to  
  custom HTTPS endpoints.  
type: object  
properties:  
  requestId:  
    description: >  
      Same as the value in the X-Amz-Firehose-Request-Id header,  
      duplicated here for convenience.  
    type: string  
  timestamp:  
    description: >  
      The timestamp (milliseconds since epoch) at which the Firehose  
      server generated this request.  
    type: integer  
  records:  
    description: >  
      The actual records of the Firehose stream, carrying  
      the customer data.  
    type: array  
    minItems: 1
```

```
maxItems: 10000
items:
  type: object
  properties:
    data:
      description: >
        The data of this record, in Base64. Note that empty
        records are permitted in Firehose. The maximum allowed
        size of the data, before Base64 encoding, is 1024000
        bytes; the maximum length of this field is therefore
        1365336 chars.
      type: string
      minLength: 0
      maxLength: 1365336

required:
  - requestId
  - records
```

範例如下：

```
{
  "requestId": "ed4acda5-034f-9f42-bba1-f29aea6d7d8f",
  "timestamp": 1578090901599
  "records": [
    {
      "data": "aGVsbG8="
    },
    {
      "data": "aGVsbG8gd29ybGQ="
    }
  ]
}
```

## 回應格式

出現錯誤時的預設行為

如果回應無法符合下列要求，Firehose 伺服器會將其視為具有 500 個狀態碼且沒有主體的狀態碼。

## 狀態碼

HTTP 狀態碼必須在 2XX、4XX 或 5XX 的範圍內。

Amazon 數據 Firehose 服務器不遵循重定向 ( 3XX 狀態碼 )。只將回應碼 200 視為成功將記錄交付至 HTTP/EP。會將回應碼 413 (大小已超出) 視為永久失效，如果已設定，則不會將記錄批次傳送至錯誤儲存貯體。會將所有其他回應碼視為可重試的錯誤，並受限於稍後解釋的輪詢重試演算法。

## 標頭 – 內容類型

唯一可接受的內容類型是應用程式/json。

## HTTP 標頭 – Content-Encoding

不能使用 Content-Encoding。必須對主體進行解壓縮。

## HTTP 標頭 – Content-Length

如果回應有主體，則 Content-Length 標頭必須存在。

## 主體 – 大小上限

回應主體的大小必須為小於或等於 1 MiB。

```
"$schema": http://json-schema.org/draft-07/schema#

title: FirehoseCustomHttpsEndpointResponse

description: >
  The response body that the Firehose service sends to
  custom HTTPS endpoints.
type: object
properties:
  requestId:
    description: >
      Must match the requestId in the request.
    type: string

  timestamp:
    description: >
      The timestamp (milliseconds since epoch) at which the
      server processed this request.
    type: integer
```



```

errorMessage:
  description: >
    For failed requests, a message explaining the failure.
    If a request fails after exhausting all retries, the last
    Instance of the error message is copied to error output
    S3 bucket if configured.
  type: string
  minLength: 0
  maxLength: 8192
required:
  - requestId
  - timestamp

```

範例如下：

```

Failure Case (HTTP Response Code 4xx or 5xx)
{
  "requestId": "ed4acda5-034f-9f42-bba1-f29aea6d7d8f",
  "timestamp": "1578090903599",
  "errorMessage": "Unable to deliver records due to unknown error."
}
Success case (HTTP Response Code 200)
{
  "requestId": "ed4acda5-034f-9f42-bba1-f29aea6d7d8f",
  "timestamp": 1578090903599
}

```

## 錯誤回應處理

在所有錯誤情況下，Amazon Data Firehose 伺服器都會使用指數退回演算法重新嘗試相同批次記錄的交付。重試會使用初始退刀時間 (1 秒) 且抖動係數為 (15%) 來取消重試，並使用新增抖動的公式 ( $\text{initial-backoff-time} * (\text{乘數} (2) ^ \text{retry\_count})$ ) 來取消每次後續重試。退避時間受限於 2 分鐘的間隔上限。例如，在第 N 次重試時，退避時間是  $= \text{MAX}(120\text{sec}, (1 * (2^n)) * \text{random}(0.85, 1, 15))$ 。

在上一個方程式中所指定的參數可能會發生變更。請參閱 AWS Firehose 文件，了解指數退避演算法中使用的確切初始退避時間、最大退避時間、乘數和抖動百分比。

在每次後續重試嘗試中，傳送記錄的存取金鑰和/或目的地可能會根據 Firehose 串流的更新組態而變更。Amazon 資料 Firehose 服務會以最佳方式在重試期間使用相同的請求識別碼。可將這最後一

個功能用於 HTTP 端點伺服器的重複資料刪除目的。如果在允許的最長時間（根據 Firehose 串流配置）之後仍未傳遞請求，則可以根據串流配置選擇性地將批次記錄傳送到錯誤存儲桶。

## 範例

CWLog 來源請求的範例：

```
{
  "requestId": "ed4acda5-034f-9f42-bba1-f29aea6d7d8f",
  "timestamp": 1578090901599,
  "records": [
    {
      "data": {
        "messageType": "DATA_MESSAGE",
        "owner": "123456789012",
        "logGroup": "log_group_name",
        "logStream": "log_stream_name",
        "subscriptionFilters": [
          "subscription_filter_name"
        ],
        "logEvents": [
          {
            "id": "01234567890123456789012345678901234567890123456789012345",
            "timestamp": 1510109208016,
            "message": "log message 1"
          },
          {
            "id": "01234567890123456789012345678901234567890123456789012345",
            "timestamp": 1510109208017,
            "message": "log message 2"
          }
        ]
      }
    }
  ]
}
```

## 文件歷史記錄

下表說明 Amazon 資料 Firehose 文件的重要變更。

變更	描述	變更日期
Amazon Kinesis Data Firehose 現在被稱為 Amazon 數據 Firehose	Amazon Kinesis Data Firehose 已重新命名為 Amazon 數據 Firehose 件。請參閱 <a href="#">什麼是 Amazon 數據 Firehose ?</a>	2024年2月9 日
添加雪花作為目的地	您可以使用 Snowflake 作為目的地建立交付串流。請參閱 <a href="#">the section called “選擇雪花作為您的目的地”</a> 。	2024 年 1 月 19 日
增加了日誌的自動解壓縮 CloudWatch	您可以在新串流或現有串流上啟用解壓縮功能，將解壓縮的 CloudWatch 記錄檔資料傳送至 Firehose 目的地。請參閱 <a href="#">the section called “使用 CloudWatch 記錄寫入”</a> 。	2023 年 12 月 15 日
新增 Splunk Observability Cloud 作為目的地	您可以使用 Splunk 可觀測性雲端作為目的地來建立 Firehose 串流。請參閱 <a href="#">the section called “選擇 Splunk Observability Cloud 作為您的目的地”</a> 。	2023 年 10 月 3 日
已新增 Amazon Managed Streaming for Apache Kafka 作為資料來源	您現在可以將 Amazon MSK 設定為將資訊傳送至 Firehose 串流。請參閱 <a href="#">the section called “使用 Amazon MSK 寫入”</a> 。	2023 年 9 月 26 日
增加了對文檔 ID 類型的 OpenSearch 服務目的地支持	如果 OpenSearch 服務是您的 Firehose 串流的目的地，DocumentId 類型會指出設定文件 ID 的方法。支援的方法是 Firehose 產生的文件 ID 和 OpenSearch 服務產生的文件 ID。請參閱 <a href="#">the section called “目的地設定”</a> 。	2023 年 5 月 10 日
已新增支援動態分割	增加了對 Amazon 數據 Firehose 中對流數據進行持續動態分區的支持。請參閱 <a href="#">動態分割</a> 。	2021 年 8 月 31 日

變更	描述	變更日期
新增自訂字首的相關主題。	新增相關主題，其中指出為交付至 Amazon S3 的資料建置自訂字首時，能夠使用的運算式。請參閱 <a href="#">自訂 Amazon S3 字首</a> 。	2018 年 12 月 20 日
添加了新的 Amazon 數據 Firehose 教程	已新增教學課程，示範如何透過 Amazon 資料 Firehose 將 Amazon VPC 流程日誌傳送至 Splunk。請參閱 <a href="#">教學課程：使用 Amazon 資料 Firehose 將 VPC 流程日誌導入 Splunk</a> 。	2018 年 10 月 30 日
增加了四個新的 Amazon 數據 Firehose 區域	新增巴黎、孟買、聖保羅和倫敦區域。如需詳細資訊，請參閱 <a href="#">Amazon 數據 Firehose 配額</a> 。	2018 年 6 月 27 日
增加了兩個新的 Amazon 數據 Firehose 區域	新增首爾和蒙特婁區域。如需詳細資訊，請參閱 <a href="#">Amazon 數據 Firehose 配額</a> 。	2018 年 6 月 13 日
新的 Kinesis Streams 做為來源功能	新增 Kinesis 串流作為「資料 Firehose」串流記錄的可能來源。如需詳細資訊，請參閱 <a href="#">來源、目的地和名稱</a> 。	2017 年 8 月 18 日
更新主控台文件	已更新 Firehose 串流建立精靈。如需詳細資訊，請參閱 <a href="#">建立 Firehose 串流</a> 。	2017 年 7 月 19 日
新資料轉換	您可以將 Amazon 資料 Firehose 設定為在資料交付之前轉換您的資料。如需詳細資訊，請參閱 <a href="#">Amazon 數據 Firehose 數據轉換</a> 。	2016 年 12 月 19 日
新 Amazon Redshift COPY 複製重試	您可以設定亞馬遜資料 Firehose，以便在失敗時重試複製命令到您的 Amazon Redshift 叢集。如需詳細資訊，請參閱 <a href="#">建立 Firehose 串流</a> 、 <a href="#">Amazon 數據 Firehose 數據交付</a> 及 <a href="#">Amazon 數據 Firehose 配額</a> 。	2016 年 5 月 18 日

變更	描述	變更日期
新的 Amazon 數據 Firehose 的地，Amazon OpenSearch 服務	您可以使用 Amazon OpenSearch 服務作為目的地創建 Firehose 流。如需詳細資訊，請參閱 <a href="#">建立 Firehose 串流</a> 、 <a href="#">Amazon 數據 Firehose 數據交付</a> 及 <a href="#">授予 Amazon 資料 Firehose 存取公共 OpenSearch 服務目的地</a> 。	2016 年 4 月 19 日
新增增強型 CloudWatch 指標和疑難排解功能	更新 <a href="#">監控 Amazon 數據 Firehose</a> 和 <a href="#">Amazon 數據 Firehose 故障</a> 。	2016 年 4 月 19 日
全新增強型 Kinesis 代理程式	已更新 <a href="#">使用 Kinesis 代理程式寫入 Amazon 資料 Firehose</a> 。	2016 年 4 月 11 日
全新 Kinesis 代理程式	新增了 <a href="#">使用 Kinesis 代理程式寫入 Amazon 資料 Firehose</a> 。	2015 年 10 月 2 日
初始版本	Amazon 資料 Firehose 開發人員指南的初始版本。	2015 年 10 月 4 日

# AWS 詞彙表

如需最新的 AWS 術語，請參閱《AWS 詞彙表 參考》中的 [AWS 詞彙表](#)。