splunk> .conf2017

Splunk & AWS

Gain real-time insights from your data at scale

Ray Zhu | Product Manager, AWS
Elias Haddad | Product Manager, Splunk

Forward-Looking Statements

During the course of this presentation, we may make forward-looking statements regarding future events or the expected performance of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results could differ materially. For important factors that may cause actual results to differ from those contained in our forward-looking statements, please review our filings with the SEC.

The forward-looking statements made in this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, this presentation may not contain current or accurate information. We do not assume any obligation to update any forward looking statements we may make. In addition, any information about our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only and shall not be incorporated into any contract or other commitment. Splunk undertakes no obligation either to develop the features or functionality described or to include any such feature or functionality in a future release.

Agenda

- Current Splunk ingestion landscape for AWS
- Current challenges
- New Solution
- Demo
- Q&A


Monitored by splunk>
Splunk Cloud Available Worldwide
amazon
web services™

Splunk Portfolio of AWS Solutions

End-to-End AWS Visibility

splunk> App for AWS

Available on Splunk Enterprise, Splunk Cloud and Splunk Light

AWS Integrations

AWS Lambda, IoT, Kinesis, EMR, EC2 Container Service

Self-deployed AMIs or SaaS on AWS Marketplace

splunk>enterprise

AMI on AWS Marketplace

splunk>cloud

SaaS Contract Billed through Marketplace

splunk> Insights for AWS Cloud Monitoring

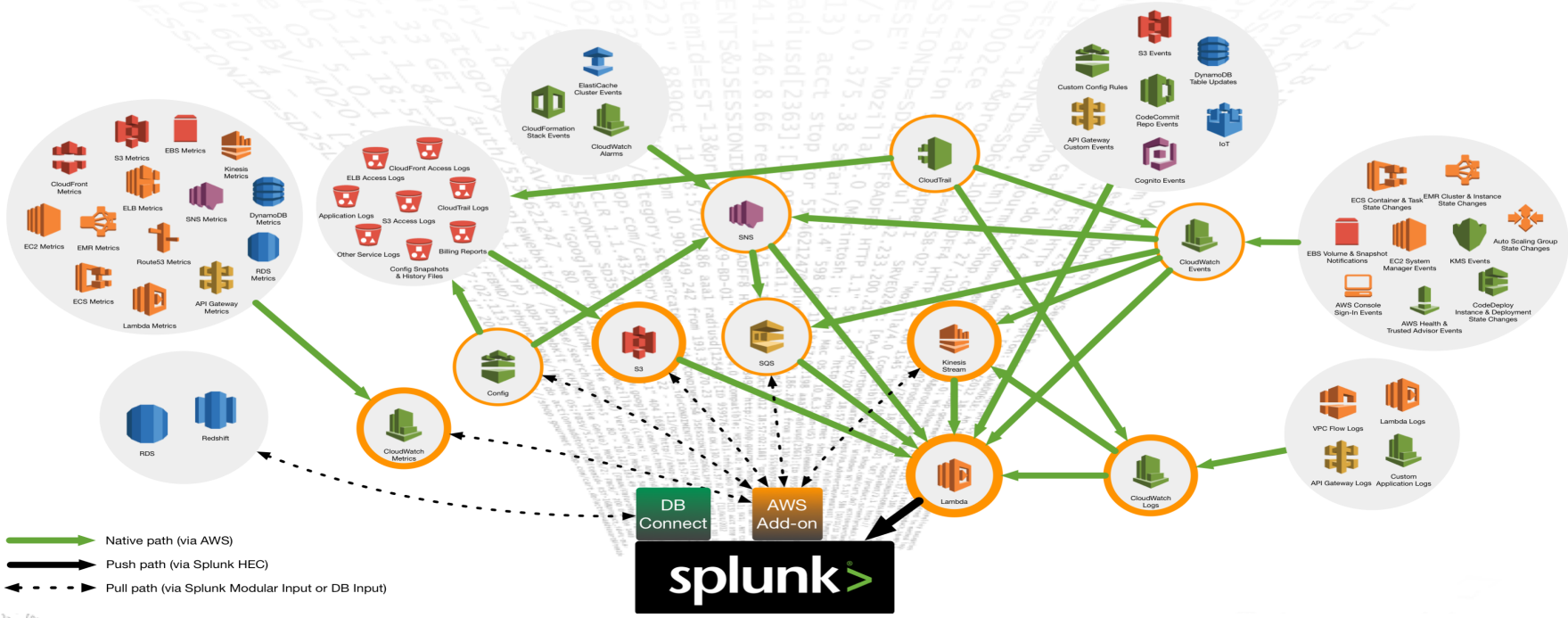
AMI on AWS Marketplace

AWS-based SaaS

splunk>cloud

Benefits of Splunk Enterprise as SaaS

Current Splunk GDI Landscape for AWS



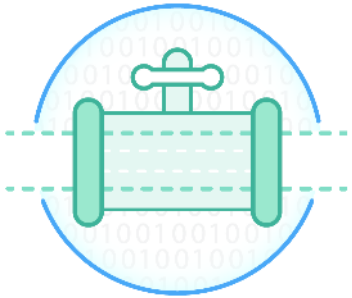
splunk> .conf2017

Challenges

- Reliability, scalability and fault tolerance
- Management overhead of data collection nodes
- Delayed event delivery due to poll based ingestion
- API throttling with poll based data ingestion

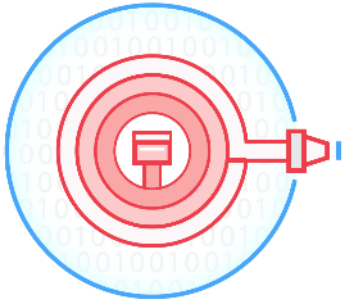
Need for New Solution

Amazon Kinesis



Kinesis Streams

Stores data as a continuous replayable stream for custom applications



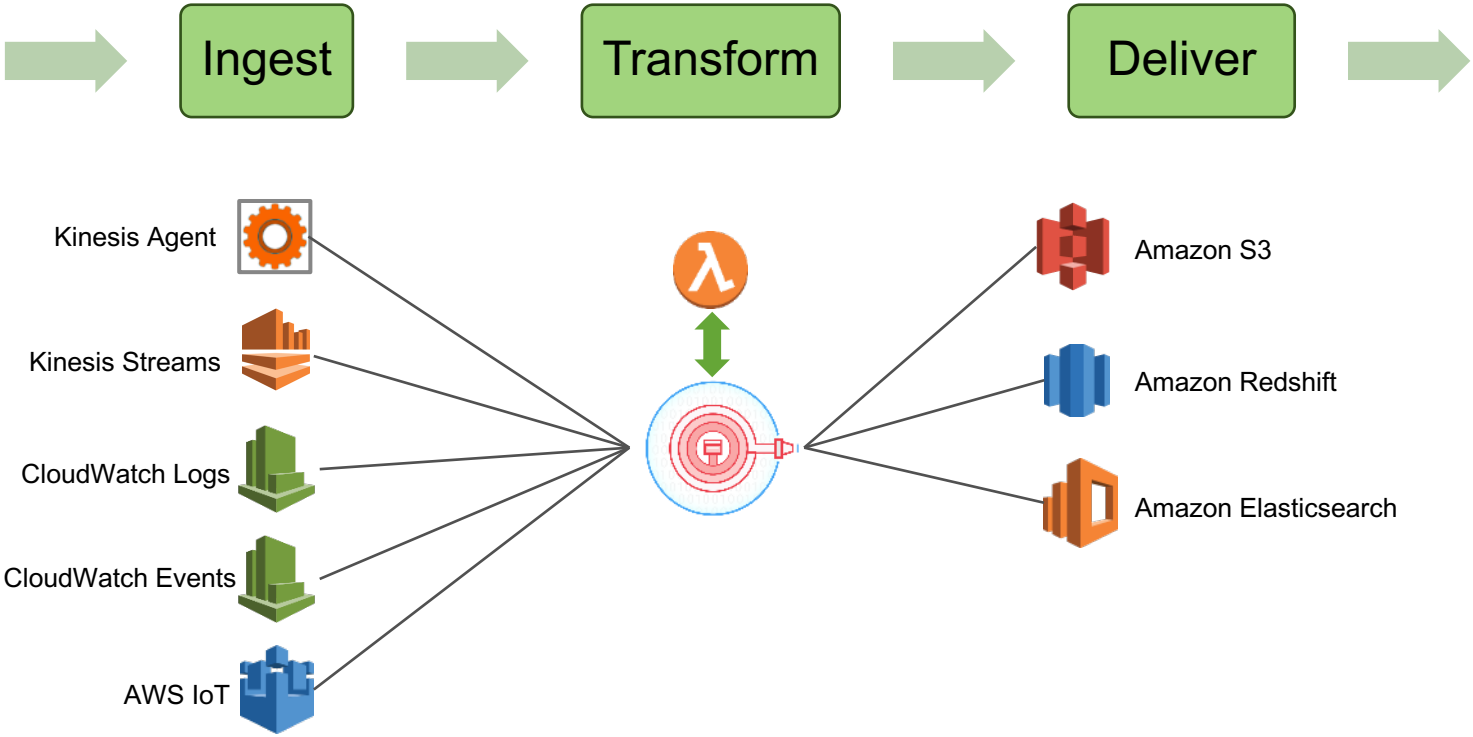
Kinesis Firehose

Load streaming data into Amazon S3, Amazon Redshift, and Amazon Elasticsearch Service

Kinesis Analytics

Analyze data streams using standard SQL queries

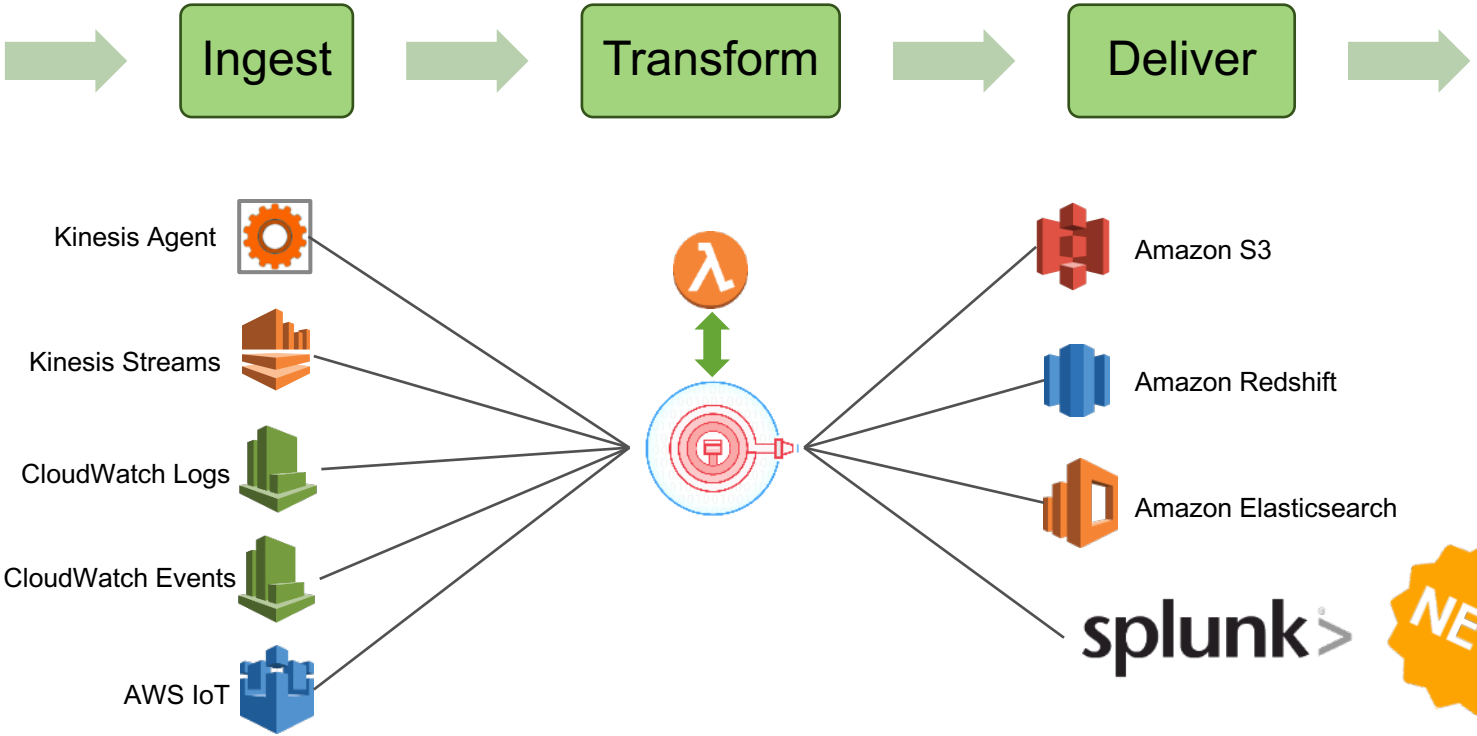
Current State of Kinesis Firehose



Our Answers to Challenges

- Reliability, scalability and fault tolerance challenges
  - Extremely reliable with underlying infrastructure operating in three different AZs
  - Extremely durable with three copies of same data in three different AZs
  - Temporarily holds and buffers data to absorb back pressure
  - Data backup to Amazon S3 upon failure
- Management overhead of data collection nodes in existing solution
  - Serverless with no resource provision or management overhead
- Delayed event delivery due to poll based ingestion
  - Push delivery with configurable buffer size and interval
- API throttling with poll based data ingestion
  - Horizontally scalable with no limit

Kinesis Firehose With Splunk Delivery



Kinesis Firehose Advantages

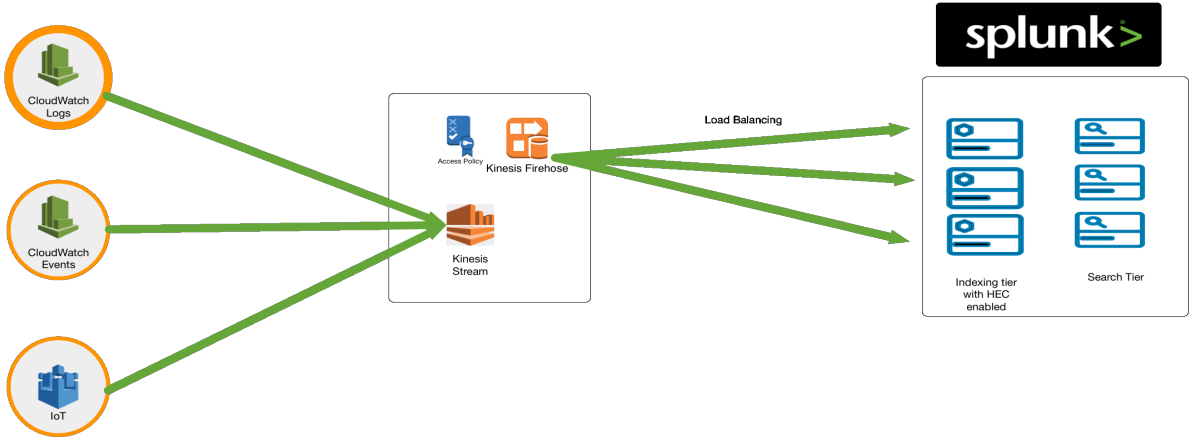
Why should I use Kinesis Firehose versus other ingestion mechanisms for Splunk?

splunk> .conf2017

Why Kinesis Firehose

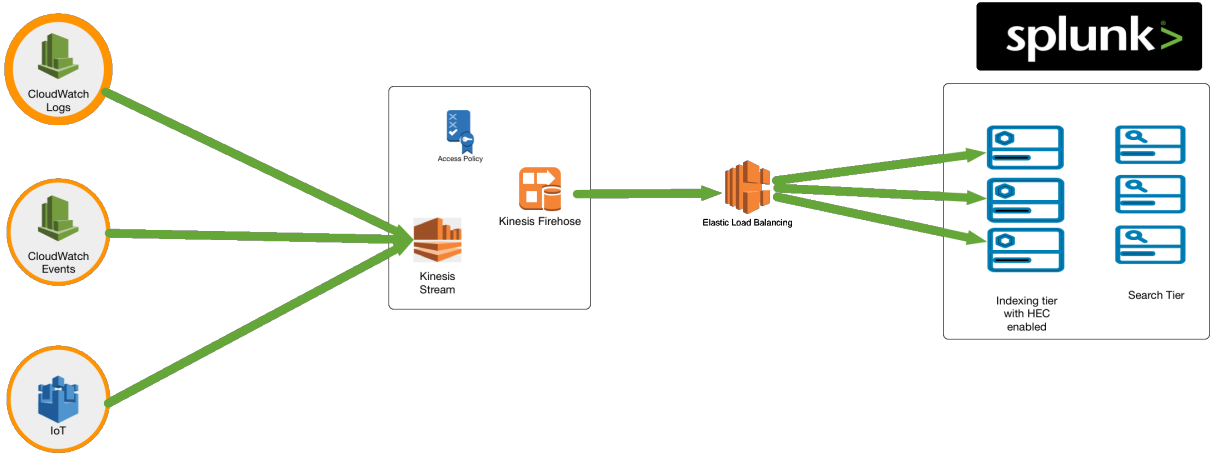
- Fully managed service with serverless architecture
- Bypass the need for setting up and managing heavy weight forwarder
- Extremely scalable and reliable
- Well integrated with various data sources
- Easy to use with no programming requirement
- Ability to transform raw data prior to sending it to Splunk
- Super low cost - $0.029 per GB of data ingested

Serverless and Scalable



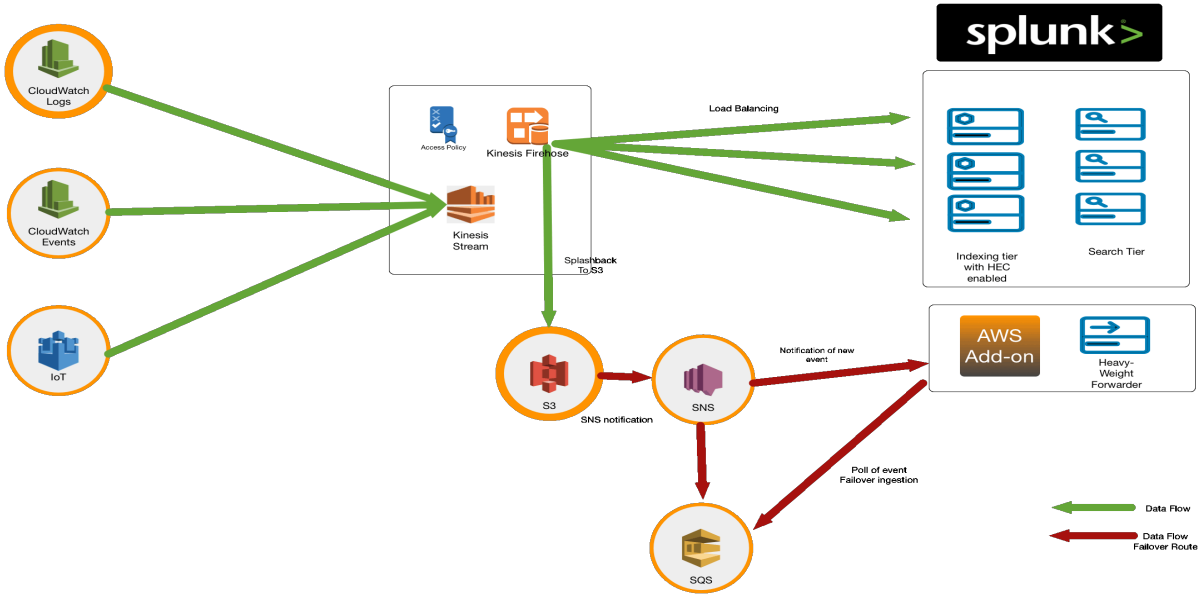
- Supports native balancing to indexing tier
- Supports Splunk Cloud and Splunk Enterprise

Serverless and Scalable



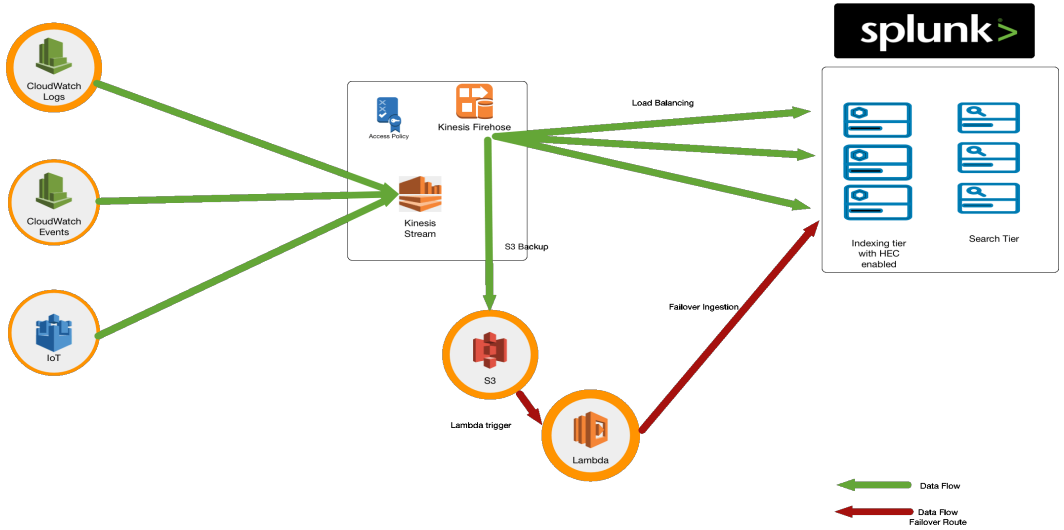
- Supports ELB and third party load balancers

Reliable AWS Add-on as Failover



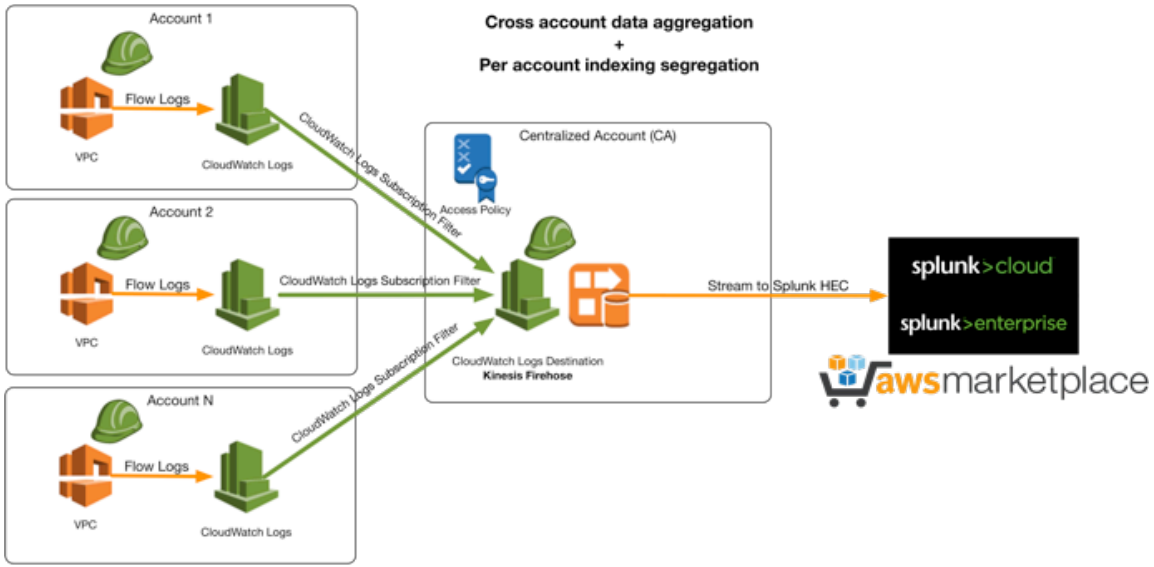
- Supports delivery acknowledgment. Un-acknowledged events can be persisted to S3 and re-ingested via alternative delivery mechanism.
- Un-delivered and un-acknowledged events can be ingested from S3 bucket using poll based mechanism (Splunk add-on for AWS)

Reliable Lambda to HEC as Failover



- Un-delivered and un-acknowledged events can be ingested from S3 using lambda for full push-based architecture.
- Lambda can be configured to push data to a failover HEC endpoint

Cross Account Delivery



- Consolidate VPC flow data from multiple account into one Firehose delivery stream
- Ability to route events to different indexes based on Lambda conditions

Kinesis Firehose Use Case

When should I use Kinesis Firehose versus other ingestion mechanisms for Splunk?

splunk> .conf2017

Supported Kinesis Firehose Data Sources

Here is a list of AWS Services supported by Kinesis Firehose

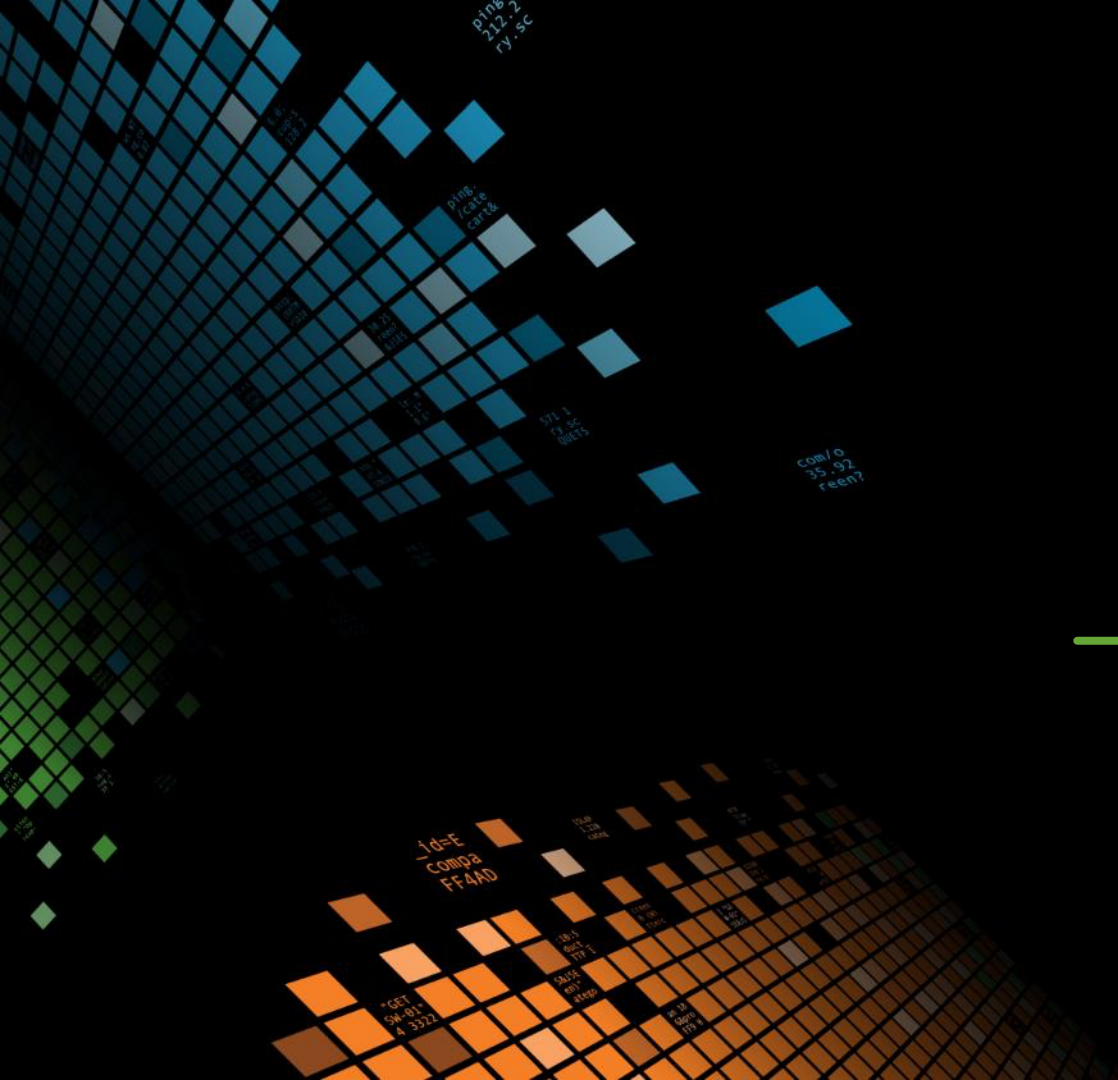
- AWS CloudWatch Logs
  - VPC Flow Logs
  - AWS Lambda Logs
- CloudWatch Events
  - AWS API Call Events (CloudTrail), Auto Scaling Events, AWS CodeBuild Events, AWS CodeCommit Events, AWS CodeDeploy Events, AWS CodePipeline Events, AWS Console Sign-in Events, Amazon EBS Events, Amazon EC2 Events, Amazon EC2 System Manager Events, Amazon EC2 System Manager Configuration Compliance Events, Amazon EC2 Maintenance Window Events, Amazon ECS EventsAmazon EMR Events, Amazon GameLift EventAWS Health Events, AWS KMS Events, Amazon Macie Events, Scheduled Events, Trusted Advisor Events
- AWS IoT
- Kinesis Streams

What Ingestion Mechanism Shall I Use?

Use Case	Kinesis Firehose	Splunk AWS Add-on
Supported Kinesis Firehose Data Sources	Preferred	-
Fault tolerance	Yes	Only SQS based S3 input
Guaranteed delivery and reliability	Yes	No
S3 Input	No	Yes
On-Prem Splunk with private IPs	No	Yes
Poll-based Data Collection (Firewall restrictions)	No	Yes

Kinesis Firehose Limits

- 20 Kinesis Firehose delivery streams per Region
- Default a maximum of 2,000 transactions/second, 5,000 records/second, and 5 MB/second
- Limits can be increased, but be careful not to increase past the incoming traffic amount. This can lead to small delivery batches to destinations, which is inefficient and can be costly.
- Please refer to the Kinesis Firehose documentation for instructions on how to increase limits: http://docs.aws.amazon.com/firehose/latest/dev/limits.html



Demo

In Summary

Splunk + AWS = Cloud Visibility

- Strong partnership with numerous product integrations

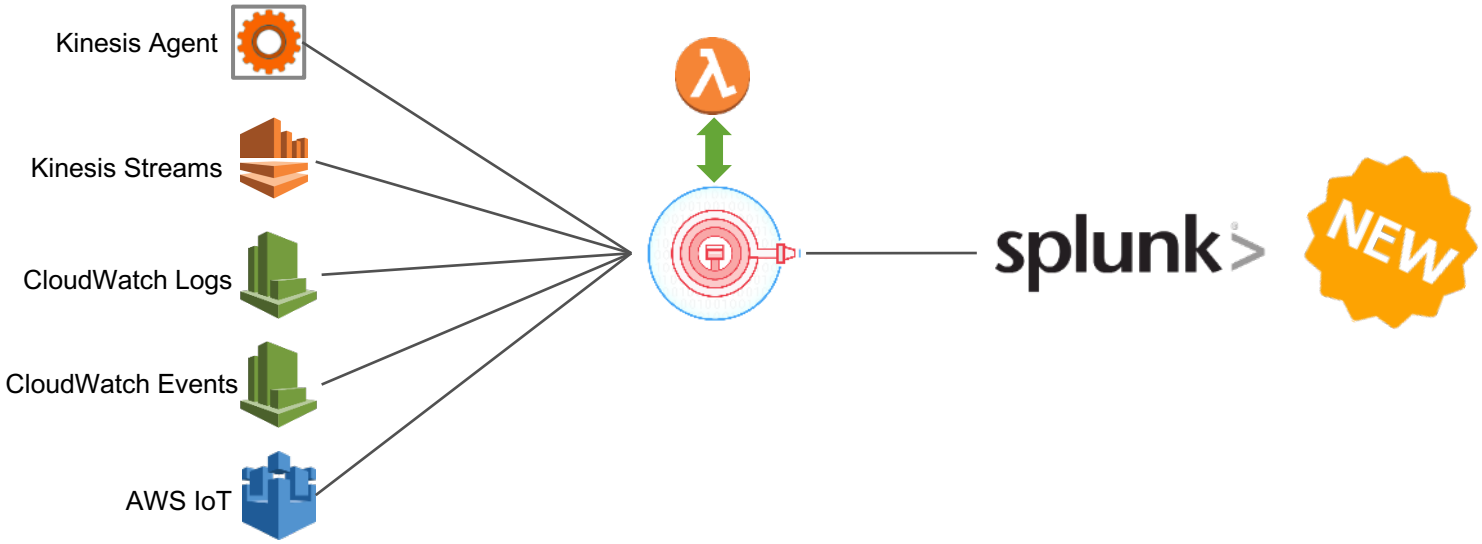
Current GDI for AWS data into Splunk

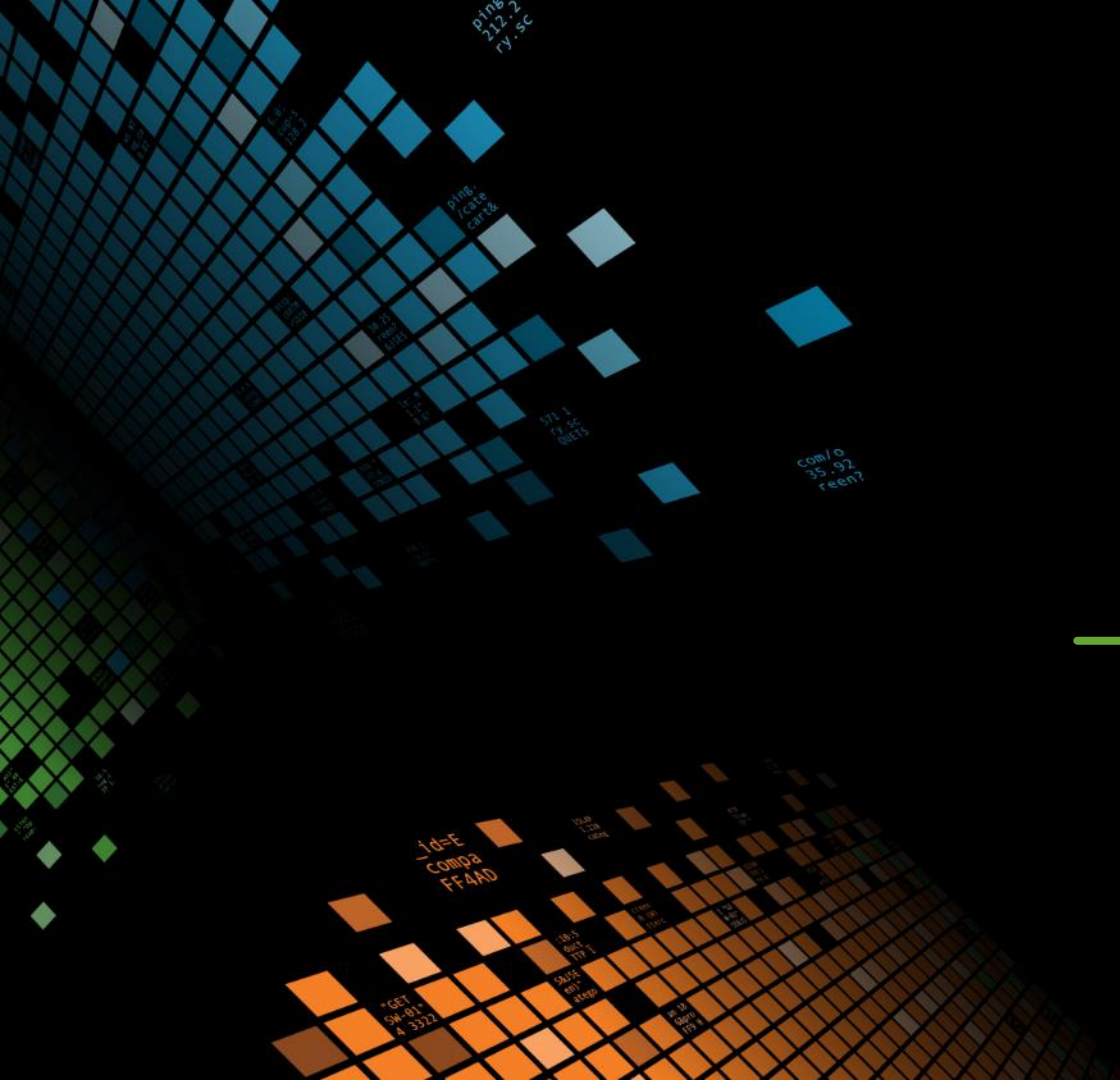
- HTTP Event Collector, AWS Add-on, DB Connect

Firehose Kinesis integration

- Addresses scalability and reliability concerns

Interested? Sign up for Beta





Q&A

splunk> .conf2017

Thank You

Don't forget to rate this session in the .conf2017 mobile app