Gain Control of Your Data Flow Using Stream Processing

Collect, process and deliver data with a single solution

Thor Taylor | Director, Product Management – Streaming

Joey Echeverria | Senior Principal Software Engineer - Streaming

October 2018

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.

splunk> .conf18

Gain Control of Your Data Flow Using Stream Processing

THOR TAYLOR

Director, Product Management - Streaming

JOEY ECHEVERRIA

Senior Principal Software Engineer - Streaming

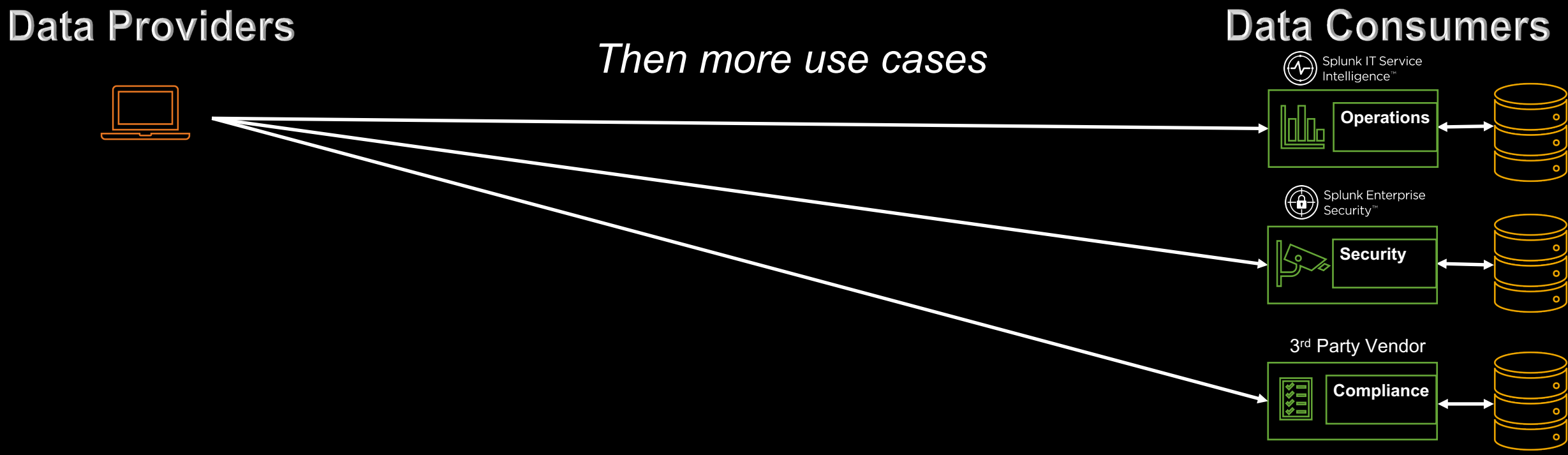
splunk> .conf18

Product Discussion

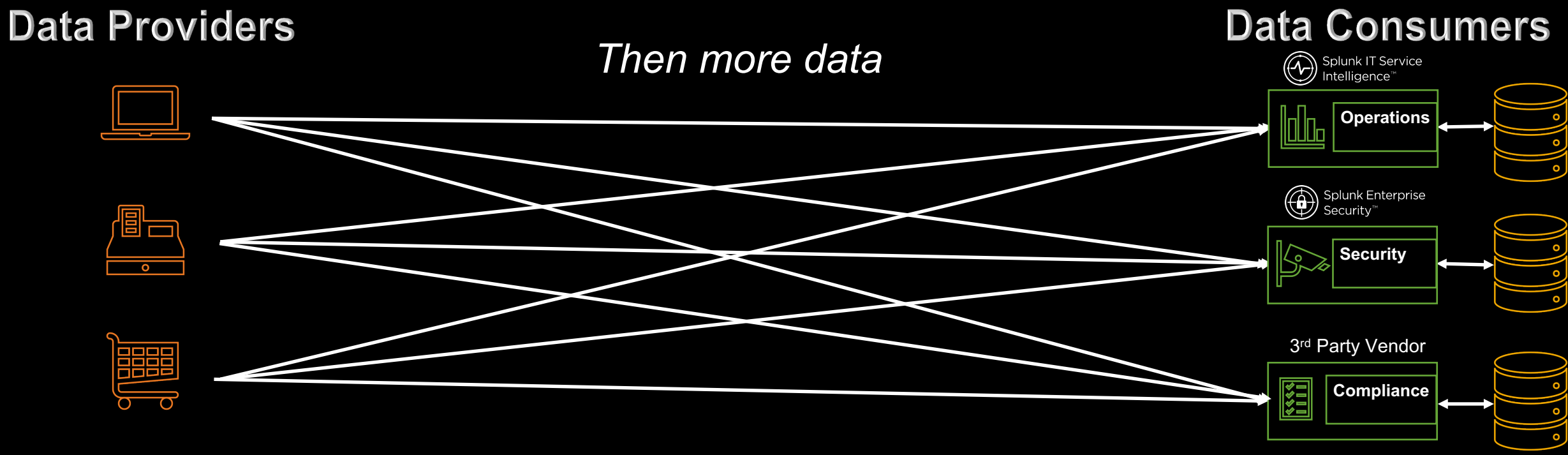
Overview of market discussion and product direction

Learning from our Past

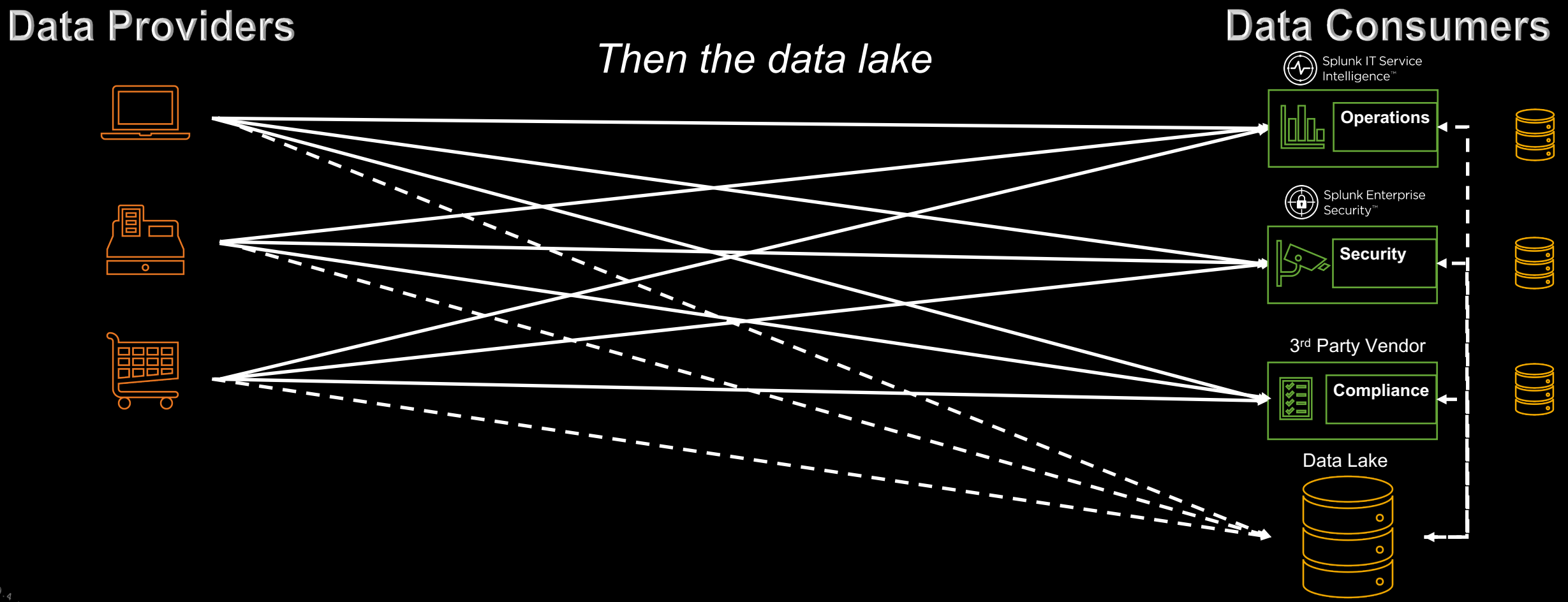
Learning from our Past



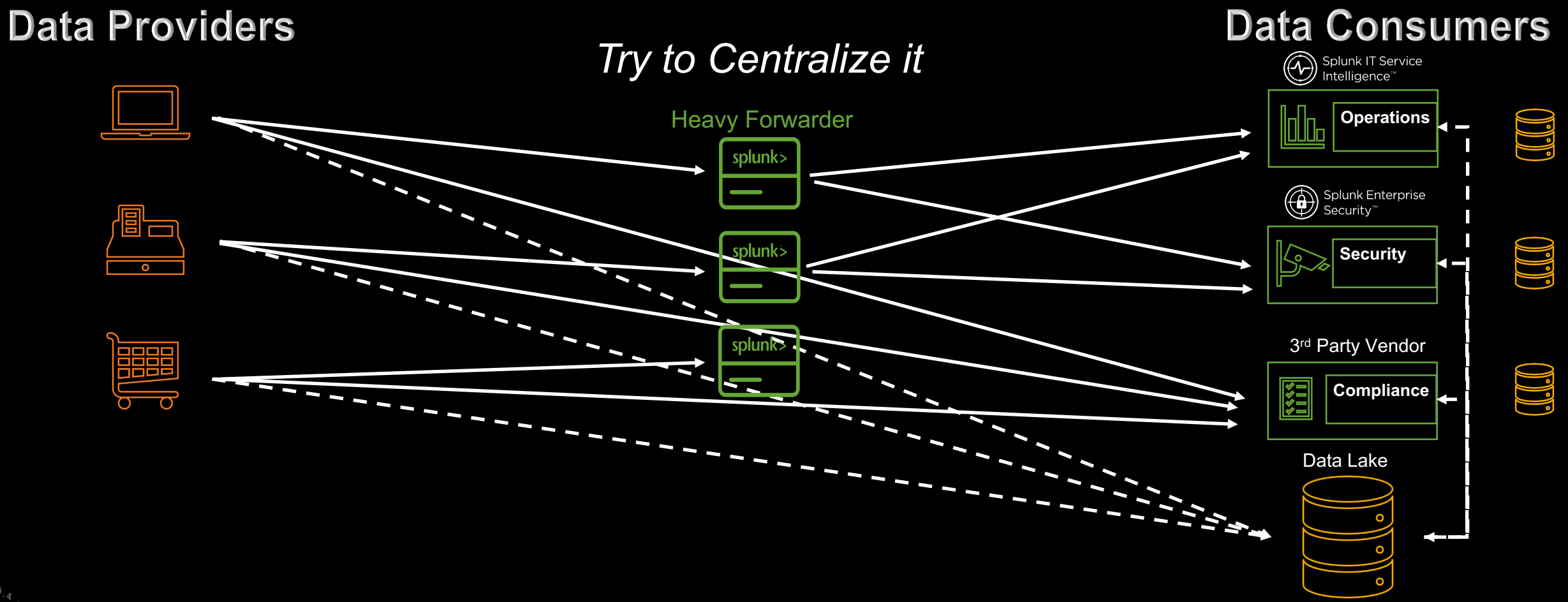
Learning from our Past



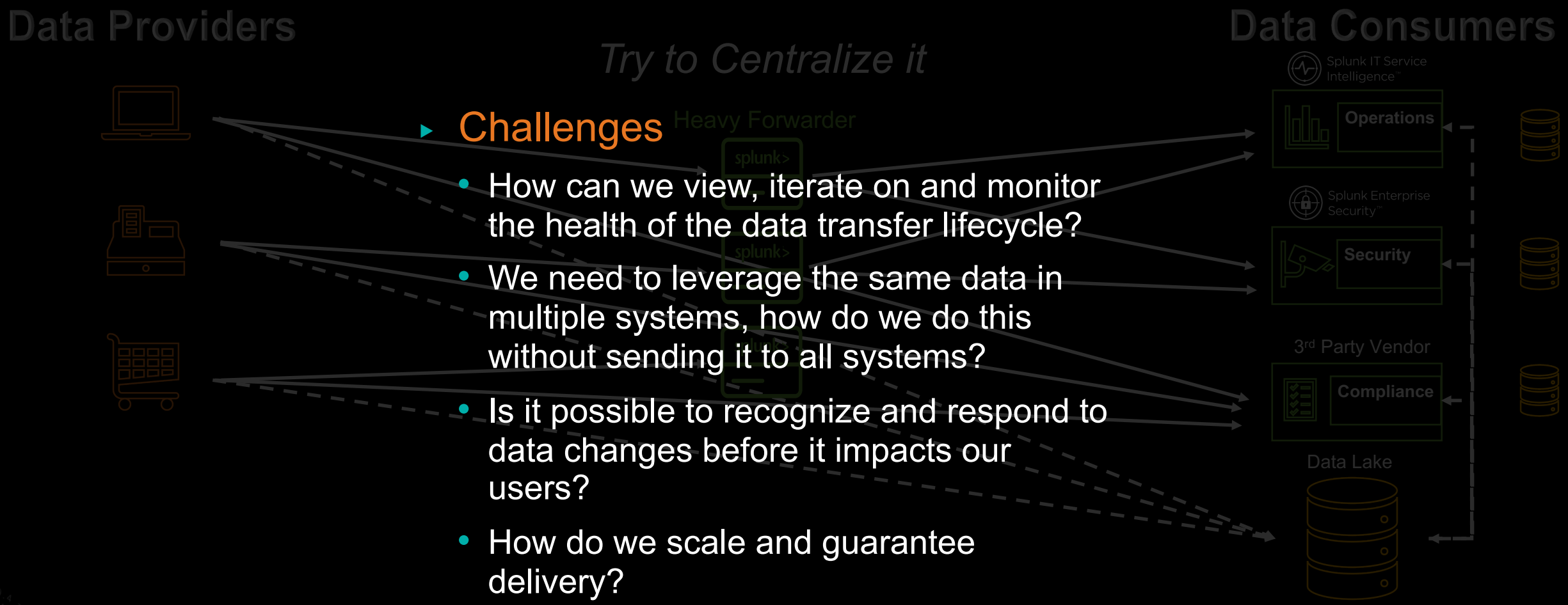
Learning from our Past



Learning from our Past



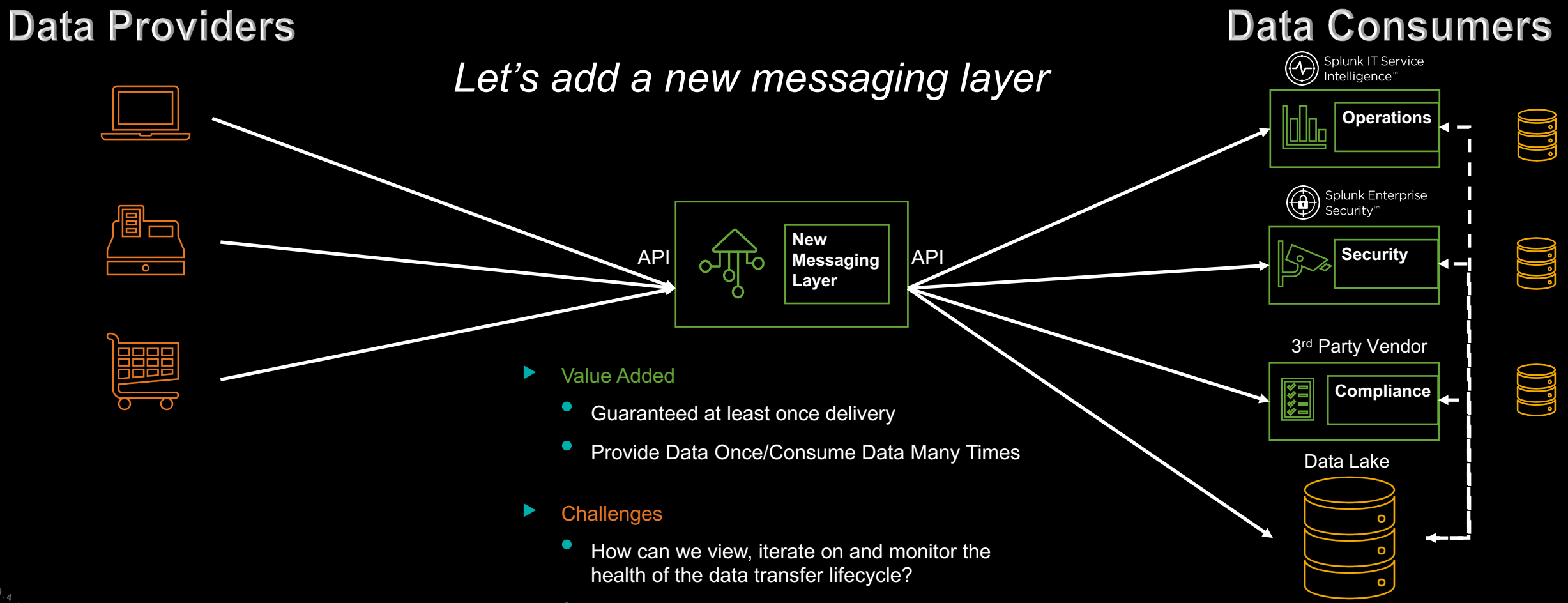
Learning from our Past



- Challenges
  - How can we view, iterate on and monitor the health of the data transfer lifecycle?
  - We need to leverage the same data in multiple systems, how do we do this without sending it to all systems?
  - Is it possible to recognize and respond to data changes before it impacts our users?
  - How do we scale and guarantee delivery?

Solving the Problem

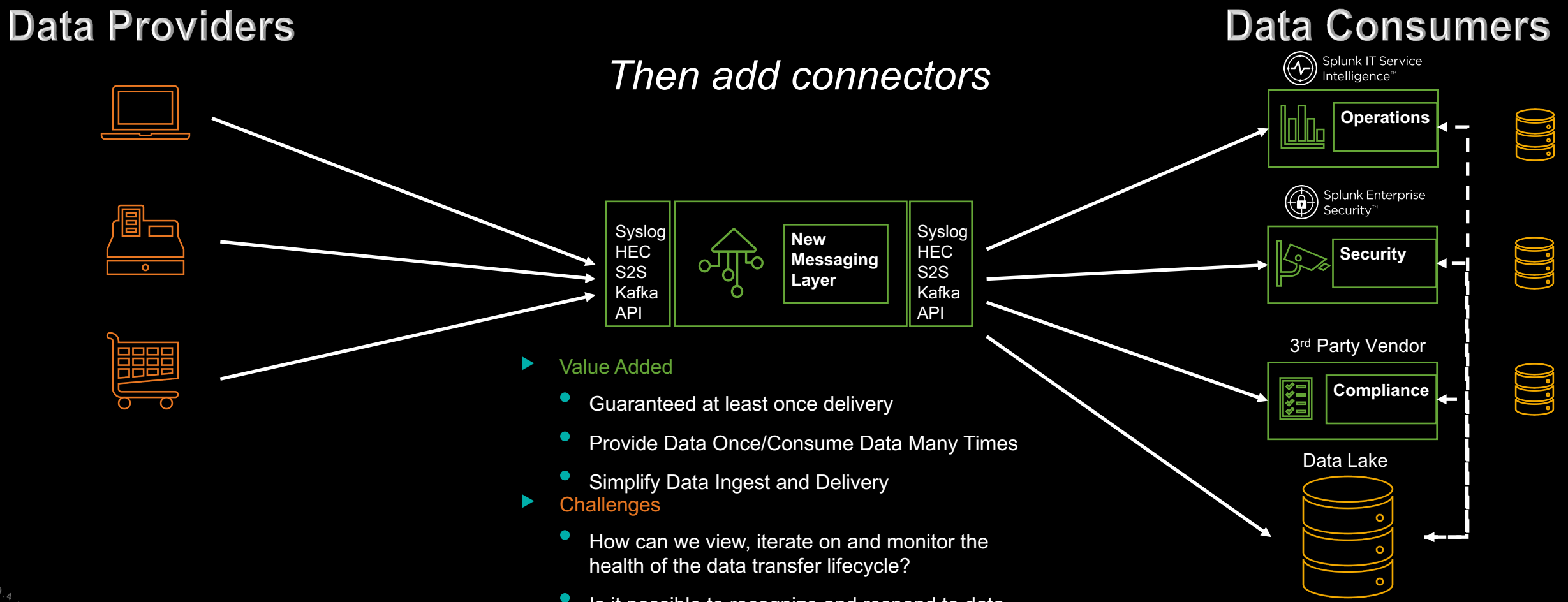
Send Data to Many Systems and Reduce Data Loss/Latency



- Value Added
  - Guaranteed at least once delivery
  - Provide Data Once/Consume Data Many Times
- Challenges
  - How can we view, iterate on and monitor the health of the data transfer lifecycle?
  - Is it possible to recognize and respond to data changes before it impacts our users?

Solving the Problem

Simplify getting data into both Splunk and non-Splunk systems

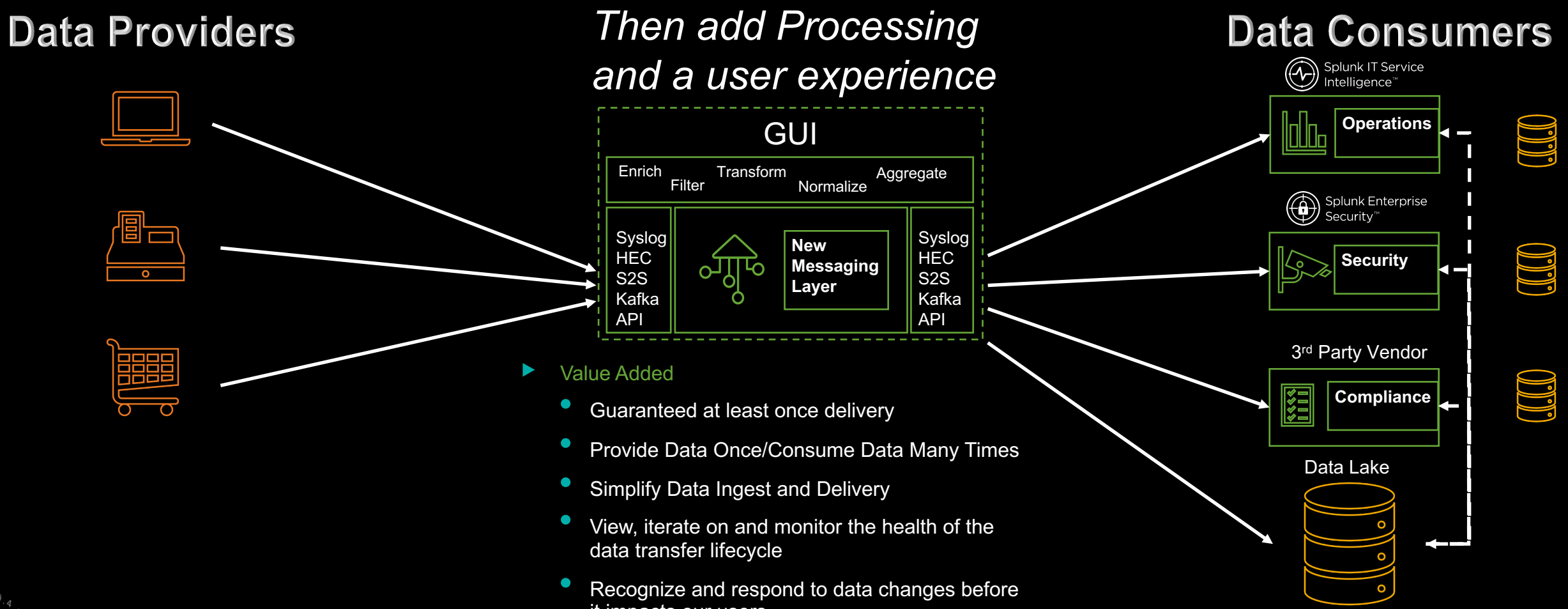


- Value Added
  - Guaranteed at least once delivery
  - Provide Data Once/Consume Data Many Times
  - Simplify Data Ingest and Delivery
- Challenges
  - How can we view, iterate on and monitor the health of the data transfer lifecycle?
  - Is it possible to recognize and respond to data changes before it impacts our users?

splunk> .conf18

Solving the Problem

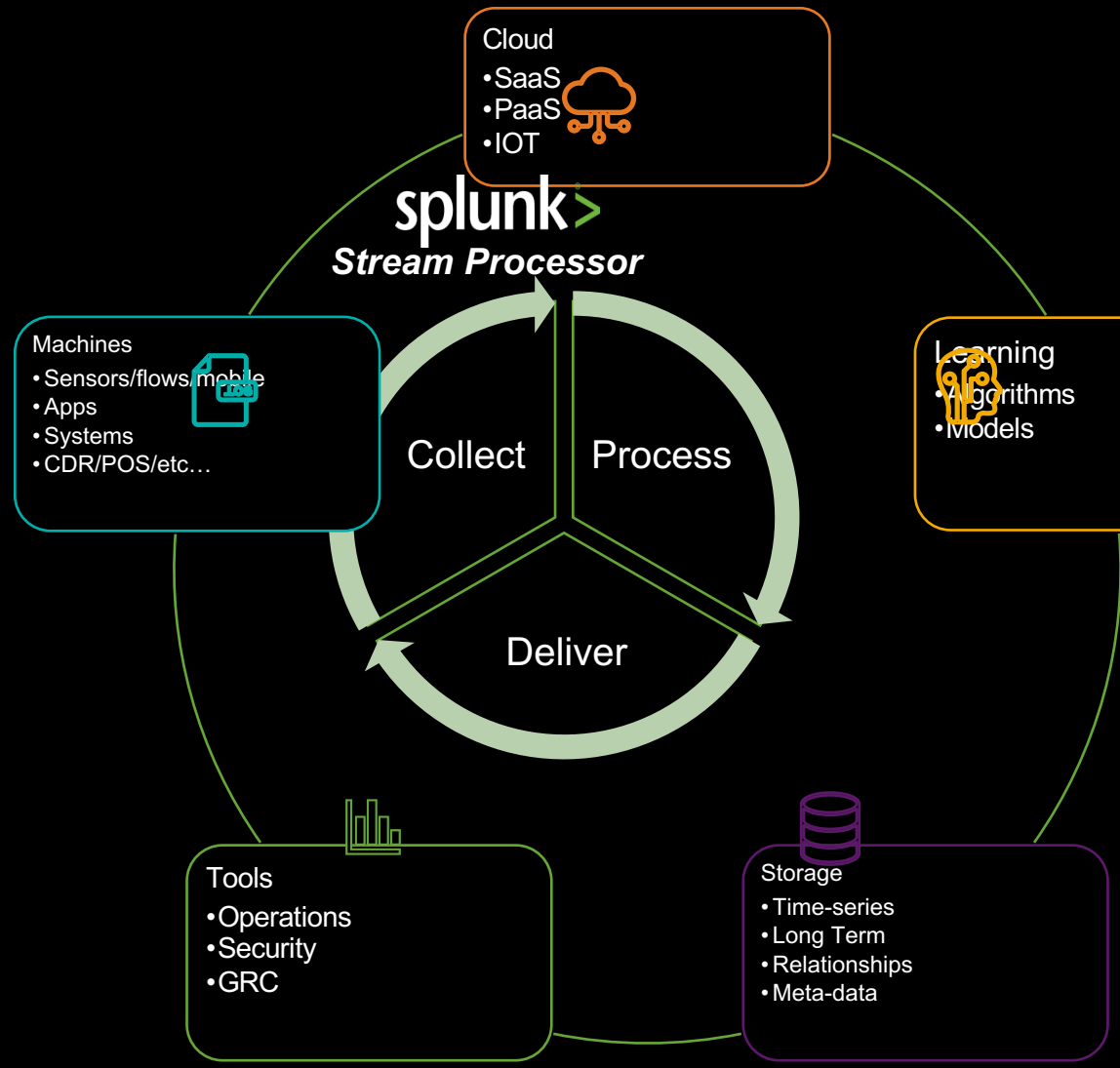
Ability to iterate on data transfer lifecycle



- Value Added
  - Guaranteed at least once delivery
  - Provide Data Once/Consume Data Many Times
  - Simplify Data Ingest and Delivery
  - View, iterate on and monitor the health of the data transfer lifecycle
  - Recognize and respond to data changes before it impacts our users

Splunk Data Stream Processor

Shifting the way we think about data ingest



Guaranteed at least once delivery

- Ability to view, iterate on and monitor the health of the data transfer lifecycle
- Simplify getting data into both Splunk and non-Splunk systems
- Improve the ability to recognize and respond to data changes

Architecture Discussion

Overview of the platform and API's of the product

Stateful Stream Processing

An Overview

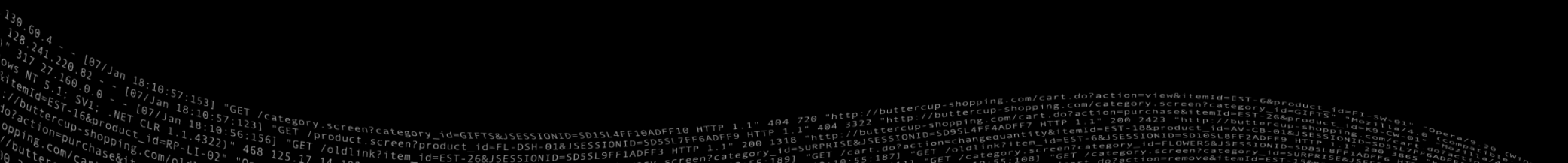
Data is processed *in stream* (e.g. during ingest)

Data delivery guarantees

- [x] At-least once
- [ ] At-most once
- [ ] Exactly once

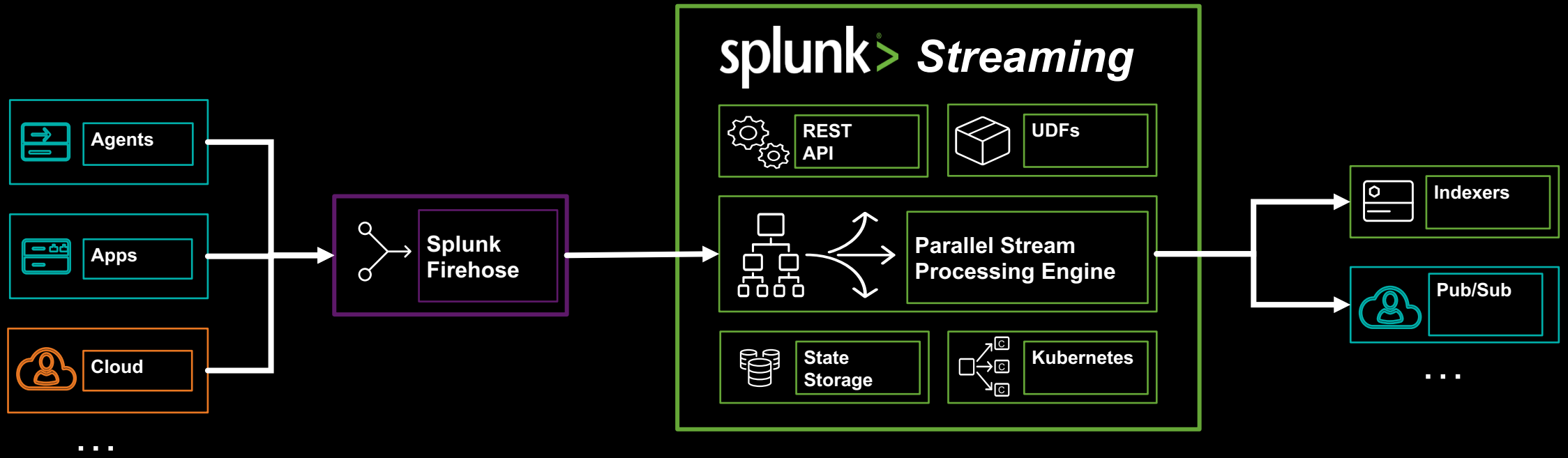
Stateful operators

- [x] Data sources
- [x] Data sinks
- [ ] Aggregation



splunk> .conf18

Architecture



User Defined Functions SDK

Java API

Fully Integrated

All built-in functions built using the same SDK

Function Types:

- Streaming operators
- Scalar functions
- Sources
- Sinks
- Aggregation functions

Demo

Four demos designed to showcase our capabilities for beta

Splunk Data Stream Processor

Beta 2018 Targets

Simplify Data Acquisition

Connect to anything from pub/sub to push/pull data providers

Maintain Control over the Data Flow

Route data where you want and only deliver data you want

Monitor and Respond to Data Drift

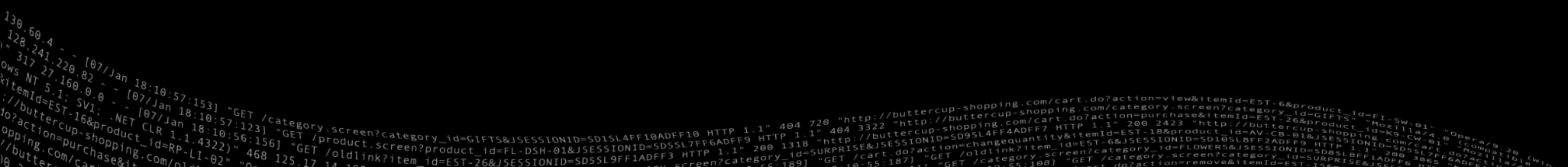
See what's going on before data gets to your Splunk Index

Meeting Demands through Extensibility

Customize capabilities and meet use cases using our API's

Data Ingestion Sessions

Wednesday	2:15 – 3:00	FN1913 - Old Meets New: Syslog and Splunk Connect for Kafka	Scott Haskell, Principal SE Architect, Splunk Mark Bonsack Staff Sales Engineer, Splunk
	2:15 – 3:00	FN1313 – Taming GDI: The Wild World of “Getting Data Into” Splunk	Peter Chen, Principal Software Engineer, Splunk Blaine Wastell, Product Management Director, Splunk
	3:15 – 4:00	FN1185 - Unleashing Data Ingestion From Apache Kafka	Scott Haskell, Principal SE Architect, Splunk Donald Tregonning, Senior Software Engineer, Splunk Ran Xie, Software Engineer, Splunk
	3:15 – 4:00	IT1647 - A Container Adventure: Scaling and Monitoring Kubernetes Logging Infrastructure	Matthew Modestino, ITOA Practitioner, Splunk David Baldwin, Principal Product Manager, Splunk Gimi Liang, Senior Software Engineer, Splunk
	4:30 – 5:15	FN1211 - Don't Miss the Bus — Splunking Kafka at Scale	Scott Haskell, Principal SE Architect, Splunk Ken Chen, Principal Software Engineer, Splunk Donald Tregonning, Senior Software Engineer, Splunk
	4:30 – 5:15	FN1919 - Gain Control of Your Data Flow Using Stream Processing	Thor Taylor, Director of Product Management, Splunk Joey Echeverria, Senior Principal Software Engineer, Splunk
Thursday	12:15 – 1:00	FN1729 - Splunk DB Connect Deep Dive: Beyond the Basics	Denis Vergnes, Principal Software Engineer, Splunk Tyler Muth, Analytics Architect, Splunk



splunk> .conf18

Thank You

Don't forget to rate this session in the .conf18 mobile app