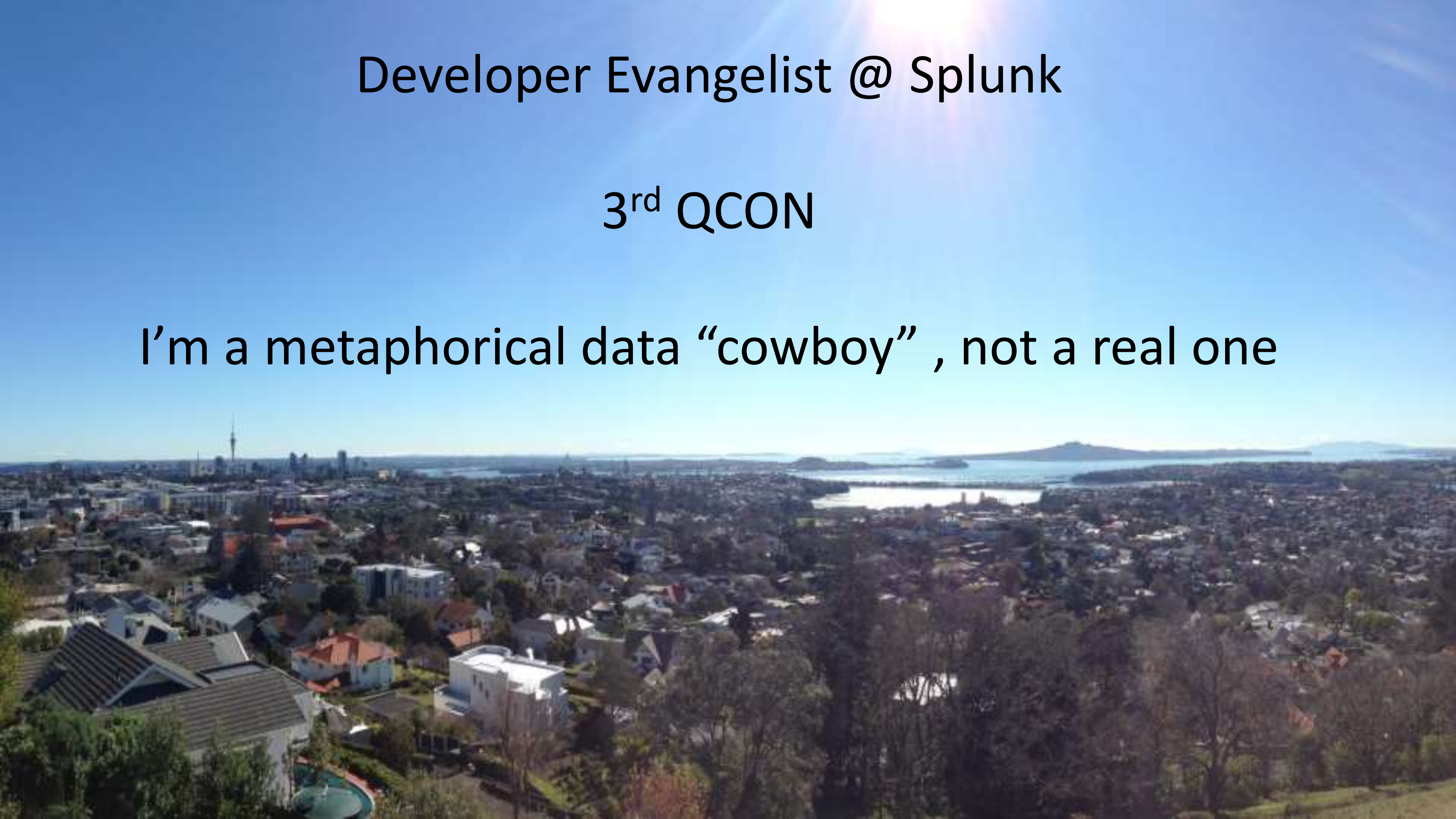
Wrangling Data at the IOT Rodeo

Damien Dallimore

ddallimore@splunk.com

@damiendallimore



Developer Evangelist @ Splunk

3rd QCON

I’m a metaphorical data “cowboy” , not a real one

splunk>
Welcome to the UK.
unk>

The cowboy metaphor

Data wrangling / lassoing (capturing)

Data needs harnessing (bring under control for analysis)

Data might need a little grooming (clean, filter)

Data might need branding (categorizing / labeling / enrichment)

Data corralling (correlation)

Data stabling (securing)

Data needs to go to the rodeo (a platform)

Make data useful = **be a data cowboy**

MACHINE DATA IS EVERYWHERE

BIG DATA COMES FROM MACHINES

Volume | Velocity | Variety | Variability

GPS,
RFID,
Hypervisor,
Web Servers,
Email, Messaging
Clickstreams, Mobile,
Telephony, IVR, Databases,
Sensors, Telematics, Storage,
Servers, Security Devices, Desktops

The IOT Revolution (or rather Evolution)

Internet of Documents

Internet of Commerce

Internet of People

Internet of APIs

Internet of Mobile

Internet of Things

CISCO

"Cisco estimates that 50 billion devices and objects will be connected to the Internet by 2020. Yet today, more than 99 percent of things in the physical world remain unconnected."

Google

"Google made another long-term big bet with the $3.2 billion buyout of Nest last January 13. This is a calculated move for Google to get into the Internet of Things revolution."

Gartner

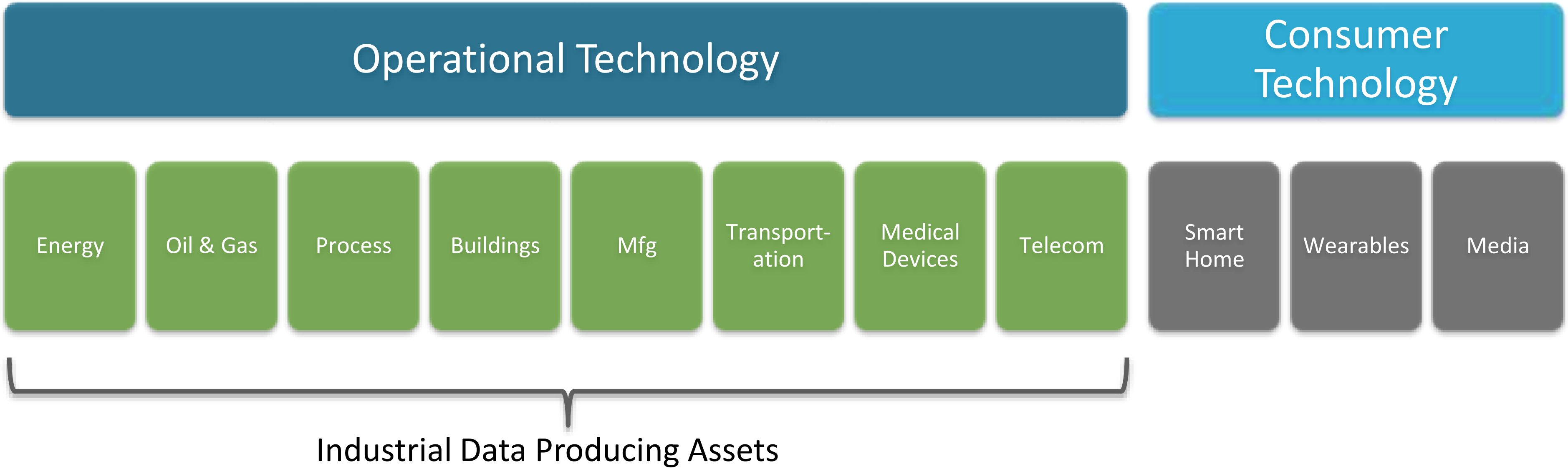
"By 2020 IoT product and service suppliers will generate incremental revenue exceeding $300 billion, mostly in services."

3SCALE

"In 2014 we expect these trends to continue with the number of public and private APIs climbing to between 100,000 and 200,000."

What is this IOT data, is it these things ?

The landscape is much, much vaster



Operational Technology

Consumer Technology

Energy

Oil & Gas

Process

Buildings

Mfg

Transportation

Medical Devices

Telecom

Smart Home

Wearables

Media

Industrial Data Producing Assets

Succeeding with IOT data

IOT data is already being generated
And we are already capturing this data

The key challenge will be in turning this into something genuinely useful. This is the opportunity.

Enable the developers & data domain experts
Give them the platforms and tools to be productive
This leads to ECOSYSTEM

How can Splunk help ?

Splunk can help you become an IOT data cowboy

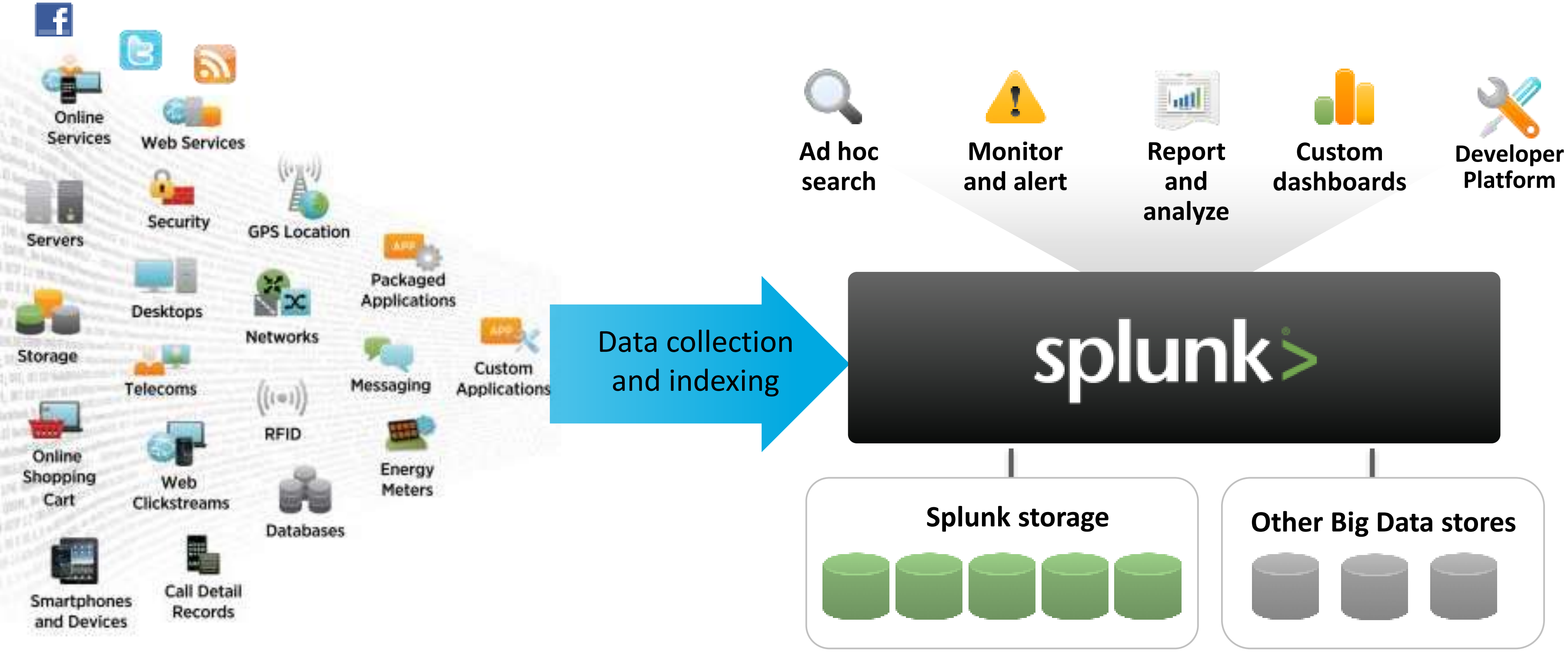
Wrangle – Collect the data

Harness – Search over the data / Correlate

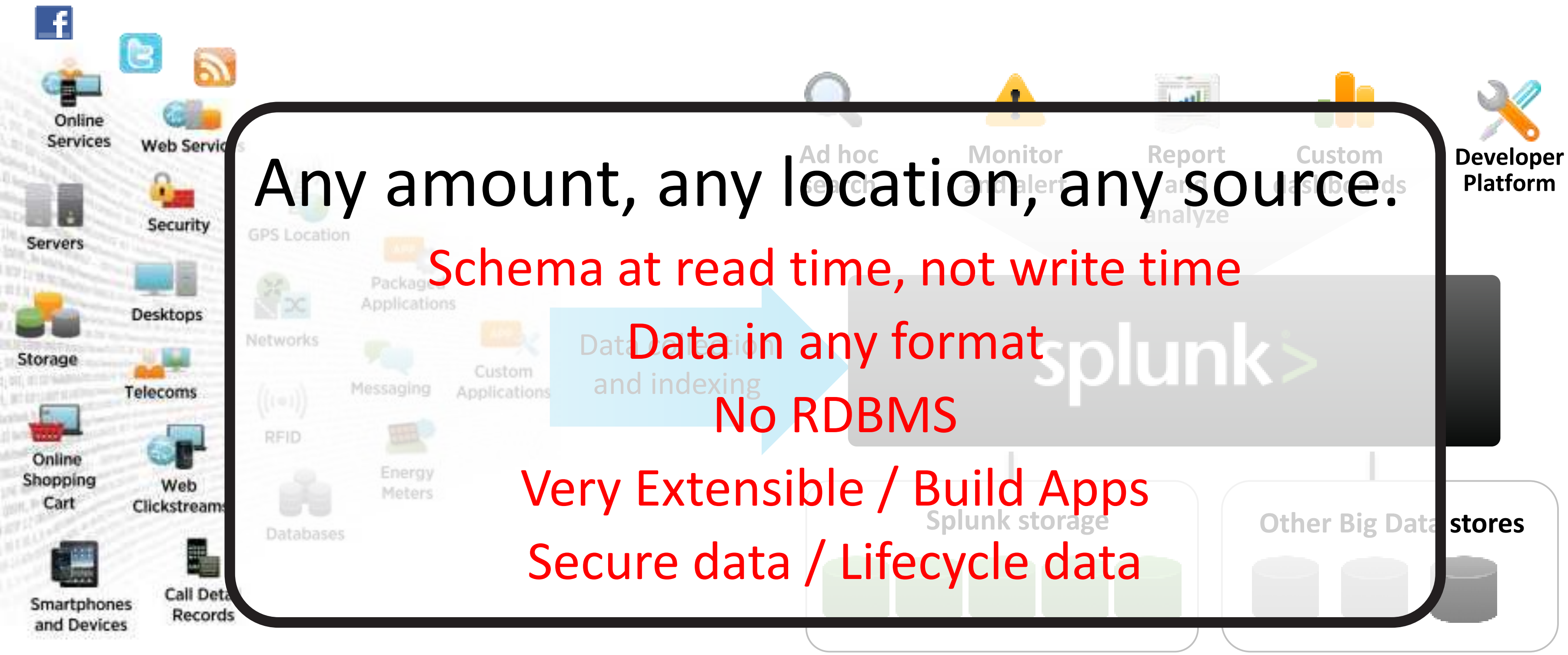
Show at the Rodeo – Visualize the data/Alerting

Provide a platform for Developers to build IOT Apps

Platform for machine data



Platform for machine data



Any amount, any location, any source.

Schema at read time, not write time

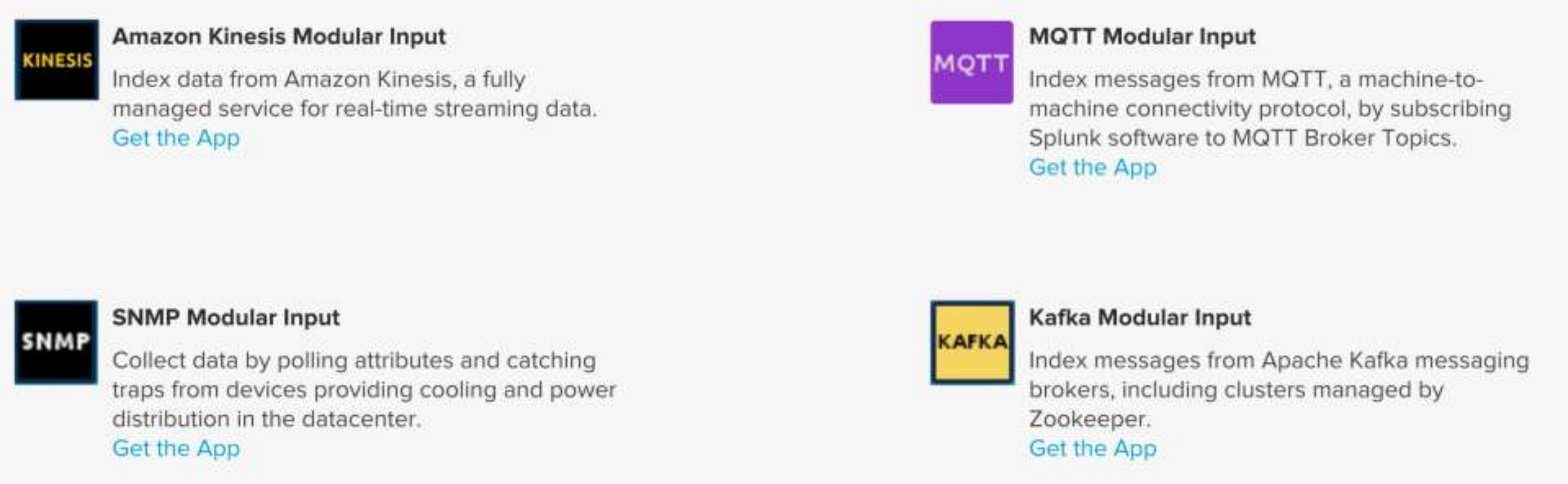
Data in any format

No RDBMS

Very Extensible / Build Apps

Secure data / Lifecycle data

Wrangling



KINESIS

Amazon Kinesis Modular Input

Index data from Amazon Kinesis, a fully managed service for real-time streaming data.

Get the App

MQTT

MQTT Modular Input

Index messages from MQTT, a machine-to-machine connectivity protocol, by subscribing Splunk software to MQTT Broker Topics.

Get the App

SNMP

SNMP Modular Input

Collect data by polling attributes and catching traps from devices providing cooling and power distribution in the datacenter.

Get the App

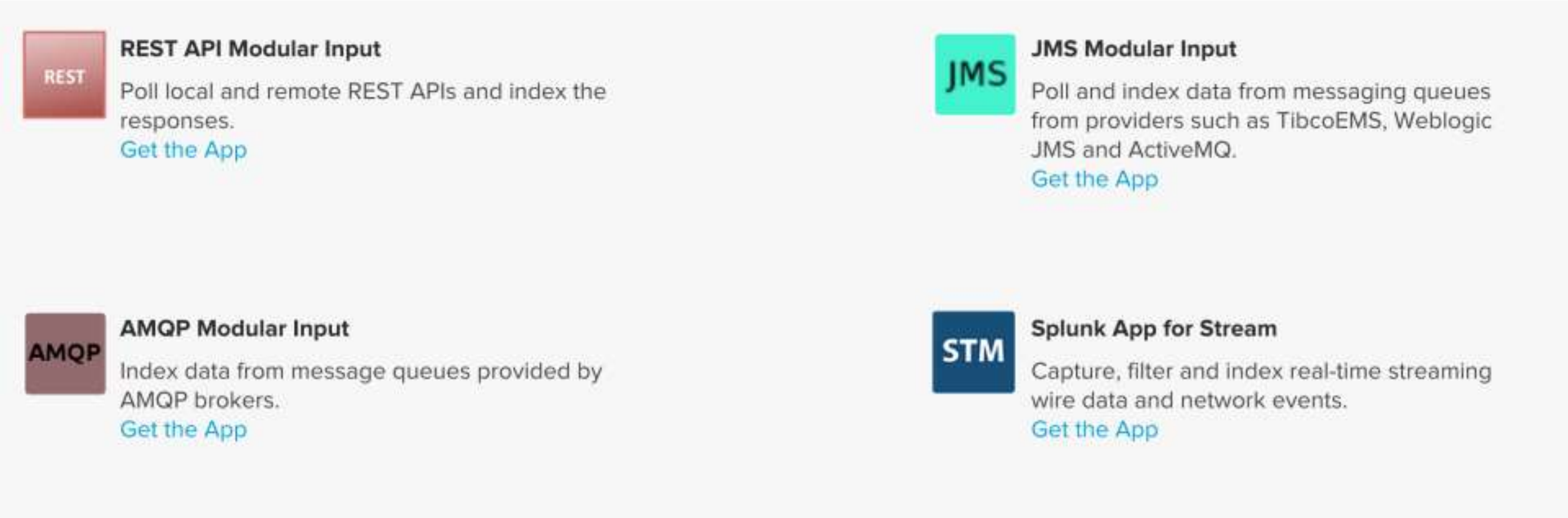
KAFKA

Kafka Modular Input

Index messages from Apache Kafka messaging brokers, including clusters managed by Zookeeper.

Get the App

Wrangling



REST API Modular Input

Poll local and remote REST APIs and index the responses.

Get the App

JMS Modular Input

Poll and index data from messaging queues from providers such as TibcoEMS, Weblogic JMS and ActiveMQ.

Get the App

AMQP Modular Input

Index data from message queues provided by AMQP brokers.

Get the App

Splunk App for Stream

Capture, filter and index real-time streaming wire data and network events.

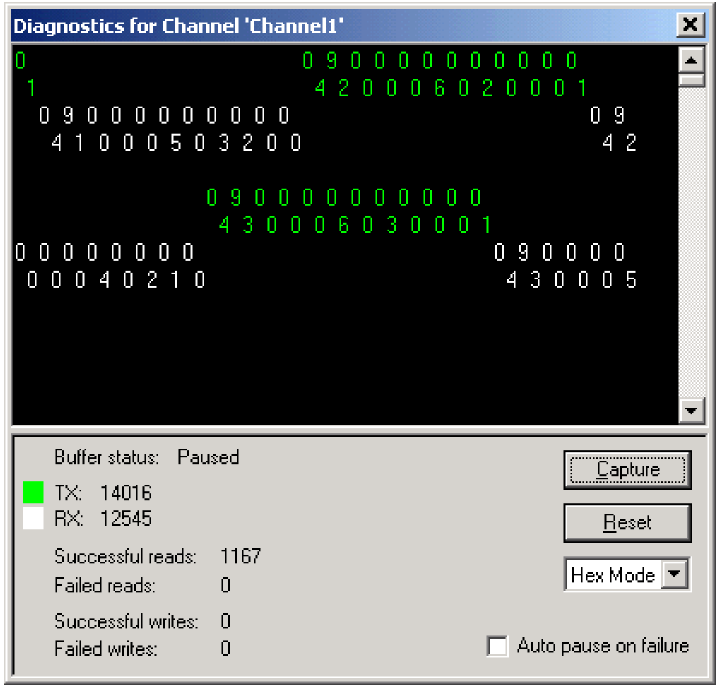
Get the App

Wrangling

Amazon Kinesis	Splunk Stream
MQTT	Kafka
JMS	AMQP
REST APIs	SNMP

COAP anybody ? Any other sources ?

Kepware Industrial Data Forwarder for Splunk



Diagnostics for Channel 'Channel1'

```
0                        0 9 0 0 0 0 0 0 0 0 0 0 0
 1                        4 2 0 0 0 6 0 2 0 0 0 1
  0 9 0 0 0 0 0 0 0 0 0 0                         0 9
   4 1 0 0 0 5 0 3 2 0 0                           4 2

                0 9 0 0 0 0 0 0 0 0 0 0 0
                 4 3 0 0 0 6 0 3 0 0 0 1
0 0 0 0 0 0 0 0                         0 9 0 0 0 0
 0 0 0 4 0 2 1 0                         4 3 0 0 0 5
```

Buffer status: Paused

🟩 TX: 14016

⬜ RX: 12545

Successful reads: 1167

Failed reads: 0

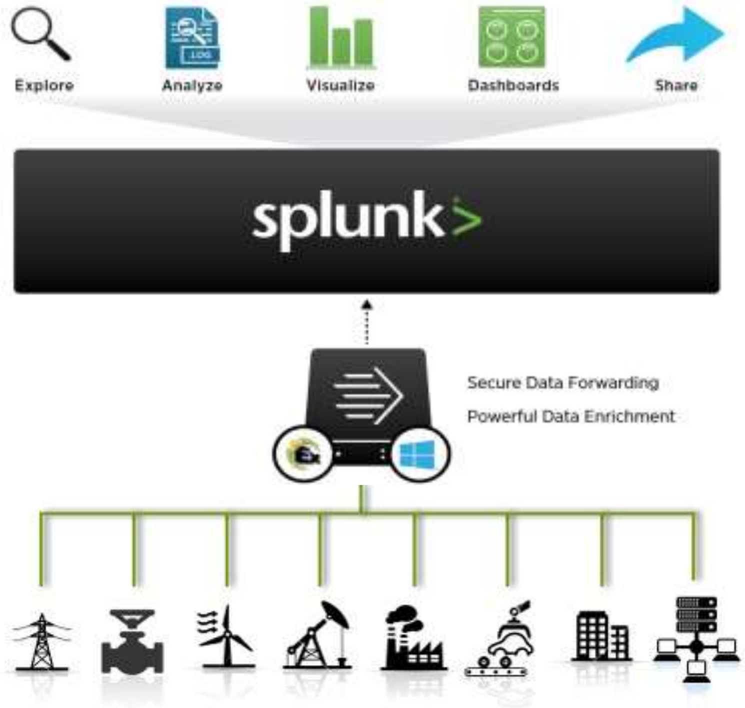
Successful writes: 0

Failed writes: 0

Hex Mode

☐ Auto pause on failure

Proprietary and legacy data translation



Real-time streaming data collection from 150+ industrial protocols

Examples of Kepware Supported Commercial and Proprietary Protocols

- ABB
- Allen-Bradley
- Analog Devices
- Aromat
- AutomationDirect
- Beckhoff
- Bristol
- Contrex
- Cutler-Hammer
- Fisher
- Fuji
- GE
- Honeywell
- Mettler-Toledo
- Mitsubishi
- Omron
- Opto 22
- Philips
- SattBus
- Scanivalve
- Siemens
- Simatic
- Sixnet
- SquareD
- Telemecanique
- Thermo Westronics
- Toshiba
- Toyopuc
- Triconex
- Wago
- WeatherBug
- Weatherford
- Yokogawa

Examples of Open Protocols

- BACnet IP
- Enron Modbus
- Modbus ASCII Serial
- Modbus Plus
- Modbus RTU Serial
- Modbus TCP/IP
- ODBC
- OPC DA
- OPC UA
- OPC XML-DA

Harnessing

```
05/27/2014T10:24:17GMT applicationId="safetyObs" eventType="safety" assetID="CV1002384-1045"
employeeId="114635" jobSite="PLEC-2014-GC" observationId="184568-451124-256" observation="Control Valve handle
extracted to manual position. No lockout/tagout or other tag visible. Process is running." observationCriticality="5"
imageId="PLEC-2014-GC-184568-451124-256" imageUri="https://mybucket.s3.amazonaws.com/PLEC-2014-GC-184568-
451124-256.png"
```

```
1543541, workorder, bsic, 78544, pipefitting, CV1002384, "install manual bleed bypass", 04/13/2014, 05/21/2014, 25663,
complete
```

```
05/22/2014 03:17:31 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/22/2014 04:21:45 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/22/2014 06:35:39 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/22/2014 07:40:29 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
```

Some data from a technician

Safety Observation Application

```
05/27/2014T10:24:17GMT applicationId="safetyObs" eventType="safety" assetID="CV1002384-1045"
employeeId="114635" jobSite="PLEC-2014-GC" observationId="184568-451124-256" observation="Control Valve handle
extracted to manual position. No lockout/tagout or other tag visible. Process is running." observationCriticality="5"
imageId="PLEC-2014-GC-184568-451124-256" imageUri="https://mybucket.s3.amazonaws.com/PLEC-2014-GC-184568-
451124-256.png"
```

```
1543541, workorder, bsic, 78544, pipefitting, CV1002384, "install manual bleed bypass", 04/13/2014, 05/21/2014, 25663,
complete
```

```
05/22/2014 03:17:31 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/22/2014 04:21:45 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/22/2014 06:35:39 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/22/2014 07:40:29 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
```

Some data from a work order

```
05/27/2014T10:24:17GMT applicationId="safetyObs" eventType="safety" assetID="CV1002384-1045"
employeeId="114635" jobSite="PLEC-2014-GC" observationId="184568-451124-256" observation="Control Valve handle
extracted to manual position. No lockout/tagout or other tag visible. Process is running." observationCriticality="5"
imageId="PLEC-2014-GC-184568-451                                        onaws.com/PLEC-2014-GC-184568-
451124-256.png"
```

CMMS (Work Order) Application

```
1543541, workorder, bsic, 78544, pipefitting, CV1002384, "install manual bleed bypass", 04/13/2014, 05/21/2014, 25663,
complete
```

```
05/22/2014 03:17:31 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/22/2014 04:21:45 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/22/2014 06:35:39 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/22/2014 07:40:29 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
```

Some data from a “thing”

```
05/27/2014T10:24:17GMT applicationId="safetyObs" eventType="safety" assetID="CV1002384-1045" employeeId="114635" jobSite="PLEC-2014-GC" observationId="184568-451124-256" observation="Control Valve handle extracted to manual position. No lockout/tagout or other tag visible. Process is running." observationCriticality="5" imageId="PLEC-2014-GC-184568-451124-256" imageUri="https://mybucket.s3.amazonaws.com/PLEC-2014-GC-184568-451124-256.png"
```

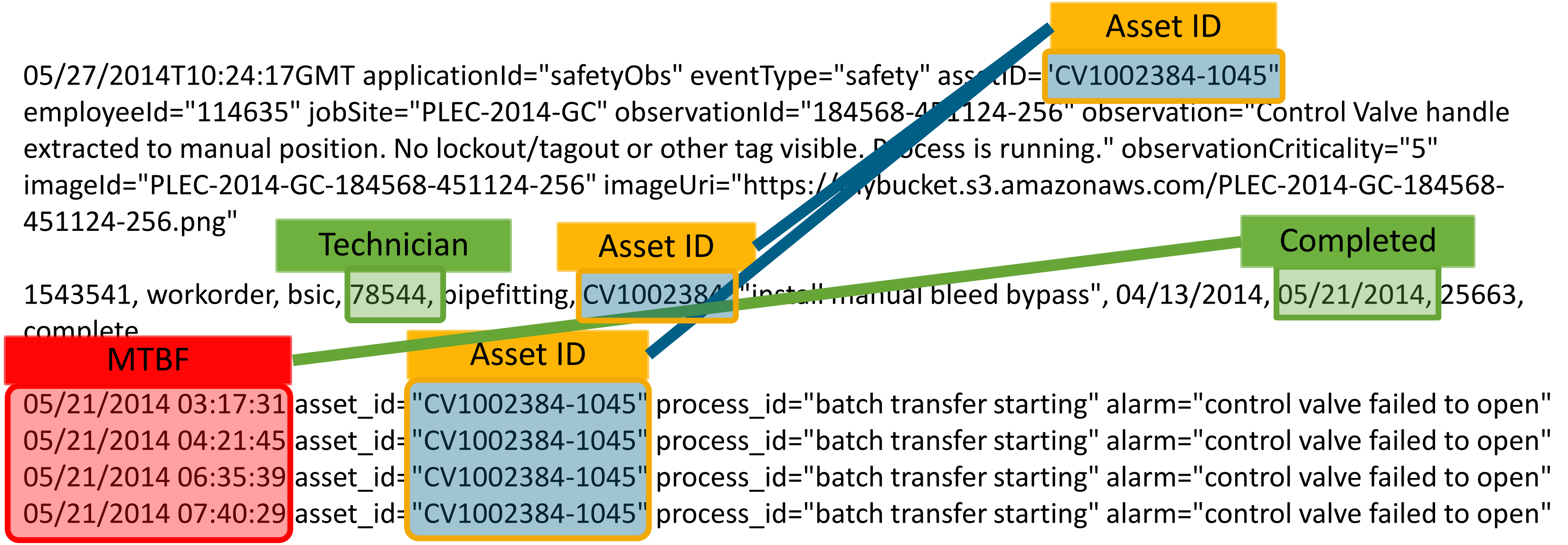
```
1543541, workorder, bsic, 78544, pipe
pass", 04/13/2014, 05/21/2014, 25663, complete
```

SCADA Event and Alarm Logs

```
05/22/2014 03:17:31 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/22/2014 04:21:45 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/22/2014 06:35:39 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/22/2014 07:40:29 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
```

splunk> listen to your data

Correlate the data– Make New Discoveries

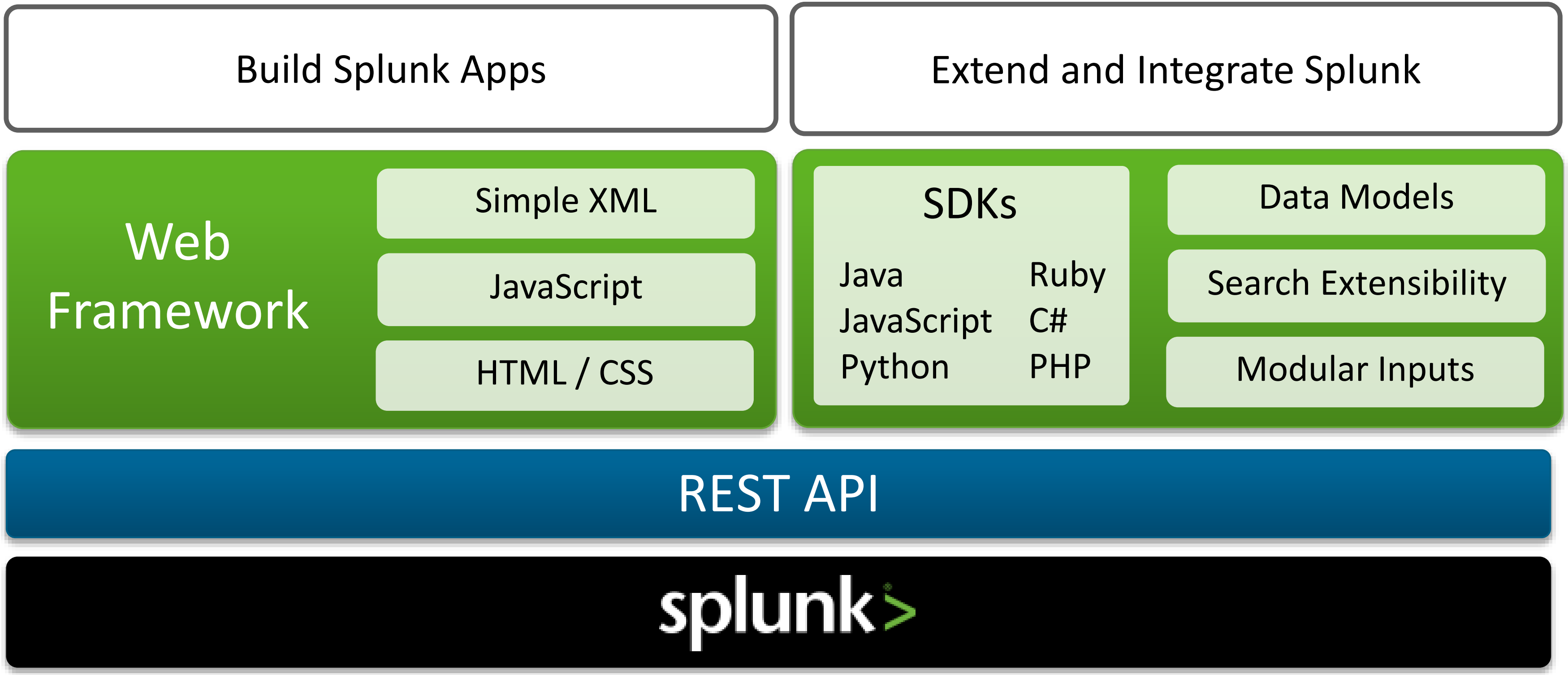


```
05/27/2014T10:24:17GMT applicationId="safetyObs" eventType="safety" assetID='CV1002384-1045"
employeeId="114635" jobSite="PLEC-2014-GC" observationId="184568-451124-256" observation="Control Valve handle
extracted to manual position. No lockout/tagout or other tag visible. Process is running." observationCriticality="5"
imageId="PLEC-2014-GC-184568-451124-256" imageUri="https://mybucket.s3.amazonaws.com/PLEC-2014-GC-184568-
451124-256.png"

1543541, workorder, bsic, 78544, pipefitting, CV1002384, "install manual bleed bypass", 04/13/2014, 05/21/2014, 25663,
complete

05/21/2014 03:17:31 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/21/2014 04:21:45 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/21/2014 06:35:39 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
05/21/2014 07:40:29 asset_id="CV1002384-1045" process_id="batch transfer starting" alarm="control valve failed to open"
```

Extensive Platform (Rodeo) for Developers



Build Splunk Apps

Extend and Integrate Splunk

Web Framework

- Simple XML
- JavaScript
- HTML / CSS

SDKs

Java	Ruby
JavaScript	C#
Python	PHP

- Data Models
- Search Extensibility
- Modular Inputs

REST API

splunk>

Splunk IOT Demos

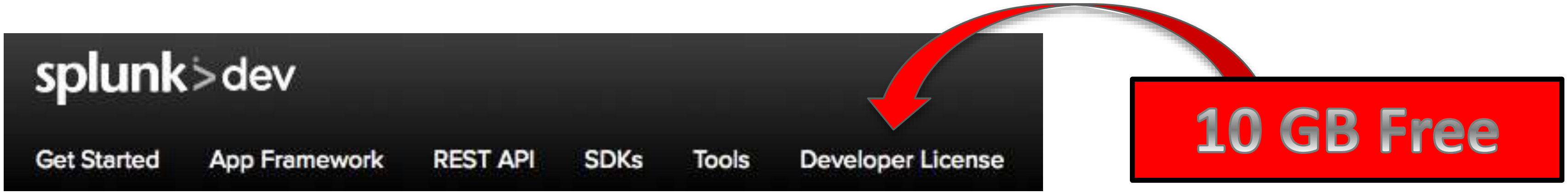
Splunk options

Splunk> Enterprise : Free to download and use. Index 500 MB/day.

Splunk> Cloud : Premium, cloud hosted. Full Enterprise stack.100% uptime.

Splunk> Sandbox : Spin up a cloud instance in minutes. Load in data.

Hunk> : Splunk for data in Hadoop HDFS, MongoDB , other datastores (Neo4J)



10 GB Free

SPLUNK.COM/IOT

APPS.SPLUNK.COM

DEV.SPLUNK.COM

BE AN IOT DATA COWBOY