splunk> .conf2017

Automation of Event Correlation and Clustering With Built-In Machine Learning Algorithms in Splunk IT Service Intelligence (ITSI)

Ross Lazerowitz | Product Manager, ITSI

Vineetha Bettaiah | Data Scientist, ITSI

09/08/2017 | Washington, DC

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> .conf2017

You Need an Approach That…

Provides easy and seamless access to all data of any type and volume

1 Delivers **service context** to prioritize investigation

2 Understands **time-based behavior** based on historical patterns

3 **Helps you find what's broken quickly** with human-scale actionable alerts

The Three Pillars of Monitoring Data

ITSI needs to be able to handle all of this in order to be “The Backbone of IT Monitoring”

Logs

Metrics

Events

The Three Pillars of Monitoring Data

ITSI needs to be able to handle all of this in order to be “The Backbone of IT Monitoring”

Logs

Metrics

Events

What the Heck is an (IT) Event Anyway?

- For the purposes of this talk when we say “Event” we are referring to Events in the IT sense not the Splunk sense.
- Self descriptive message that tells a user that something happened.
- Usually contain some sort of title, severity, and description.
- Used to determine in the moment health.
- Often very noisey.
- Think alarm data coming out of tools like Nagios, Solarwinds, APM, Netcool, etc.

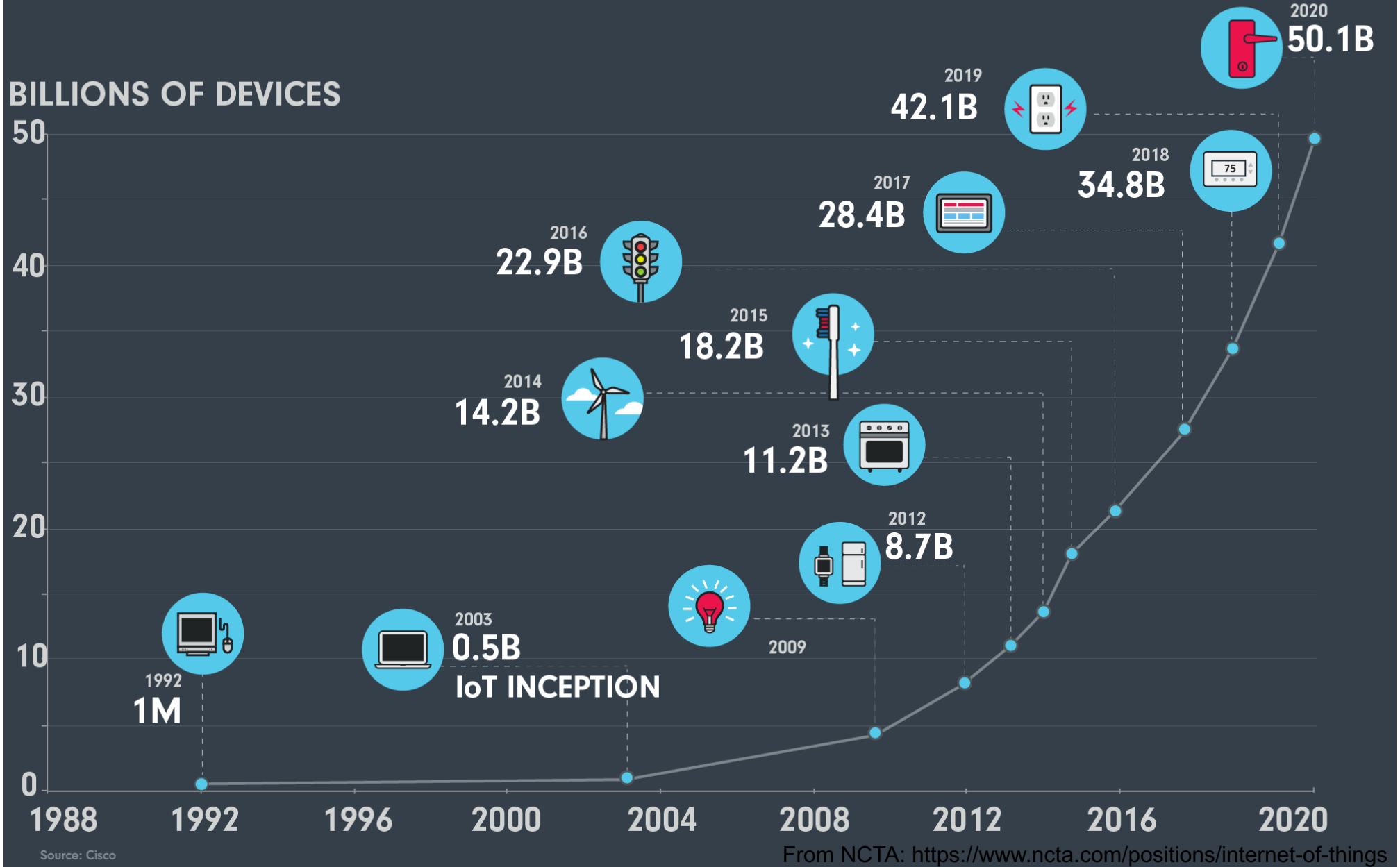
Example Event

Nagios Health Check

1502642822 src_host="splunk_sh-01" omd_site ="SJC" perfdata="SERVICEPERFDATA" name="check_dhcp" severity="OK" attempt="1" statetype="HARD" executiontime="0.000" latency="0.000" reason="OK: Received 1 DHCPOFFER(s), max lease time = 600 sec." result="OK"

GROWTH IN THE INTERNET OF THINGS

THE NUMBER OF CONNECTED DEVICES WILL EXCEED **50 BILLION** BY 2020



From NCTA: https://www.ncta.com/positions/internet-of-things

.conf2017

The Road to ITSI Event Analytics

ITSI 2.1

ITSI supports Notable Events.

.conf 2016

ITSI releases the Policy Engine. Users can curate policies that reduce the noise in events and take automated action.

.conf 2017

ITSI releases Smart Mode. ITSI can now use machine learning to reduce noise in events.

splunk> .conf2017

Splunk ITSI for Event Analytics

Simplify Your Operations With Artificial Intelligence and Service Context

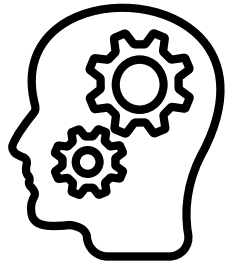
Service Context

Find and fix the most important issues

Contextualize and prioritize

Reduce time-to-resolution on business-critical services

Artificial Intelligence



Transform IT operations with machine learning

Separate valuable signal in noise

Enable IT with intelligence for **data-driven decisions**

Scalable Platform

Get a full view of your IT environment

Respond collaboratively and simplify operations

Share customized insights across the enterprise to enable **business-centric IT**

splunk> .conf2017

Under the hood

Vineetha Bettaiah

splunk> .conf2017

Objective

To identify all ideal groups of similar events

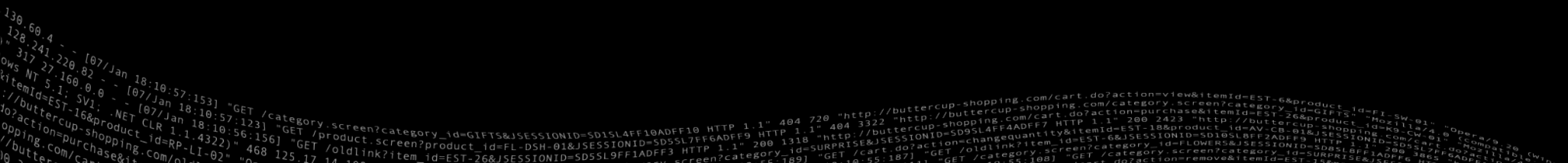
- Input – Set of events
- Features – Each event has a number of attribute fields (topological, numerical, text, etc)
- Output – List of ideal groups with characteristic metadata of each group

Group 1 Metadata

name = "check_dhcp
severity = “OK”
reason = “OK: Received 1 DHCPOFFER(s)

Group 2 Metadata

src_host="splunk_sh-01
statetype="HARD"

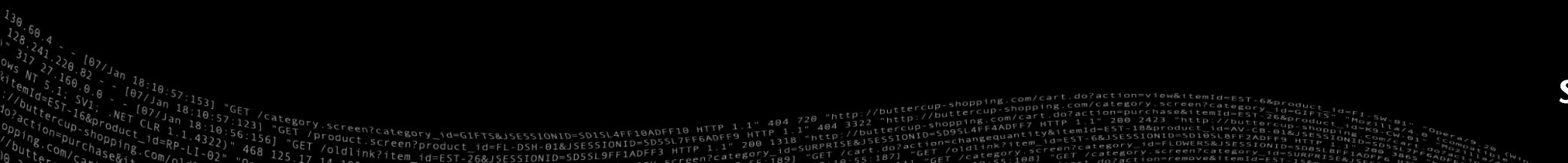


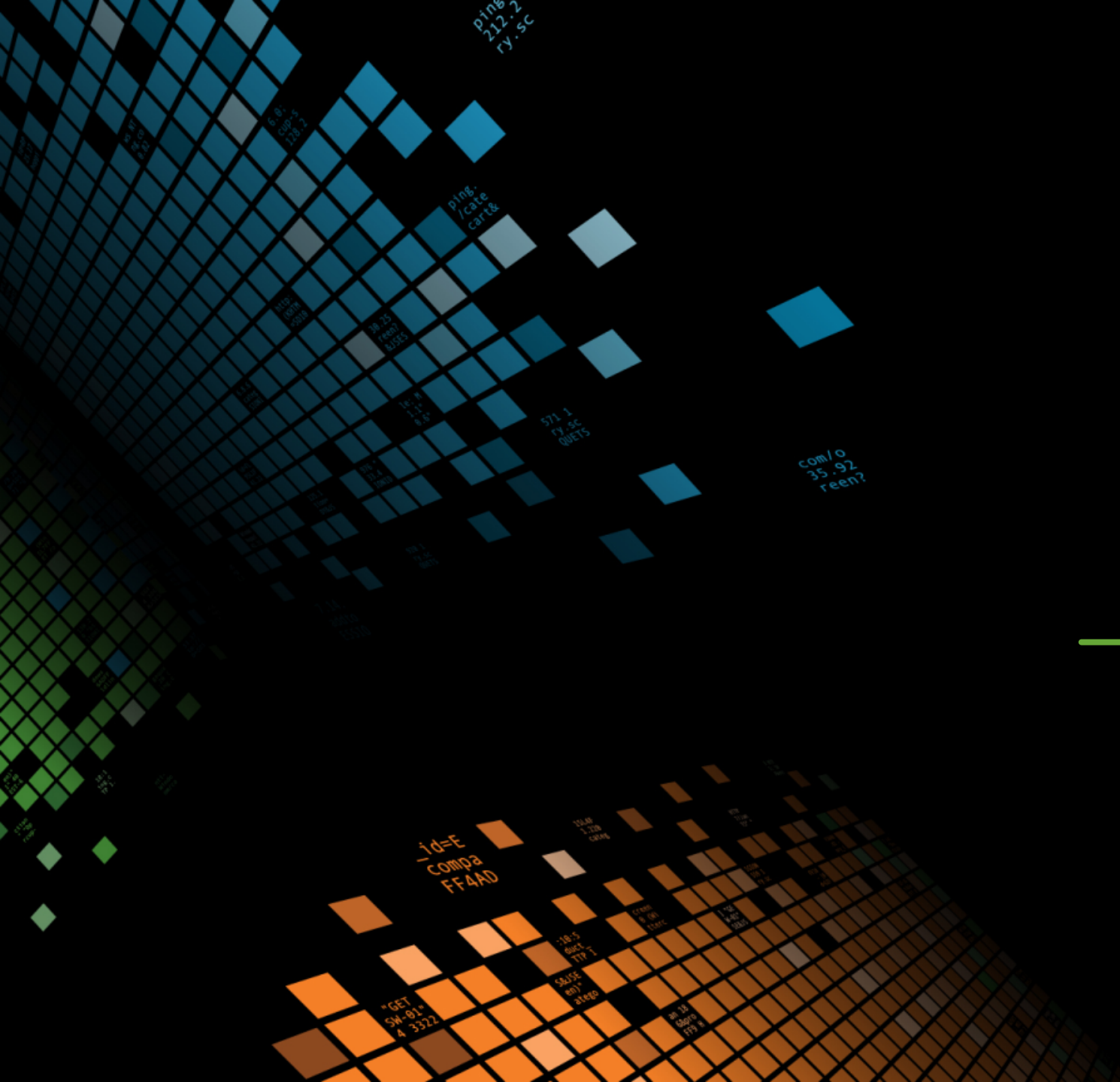
splunk> .conf2017

This is not easy

Big O is not your friend

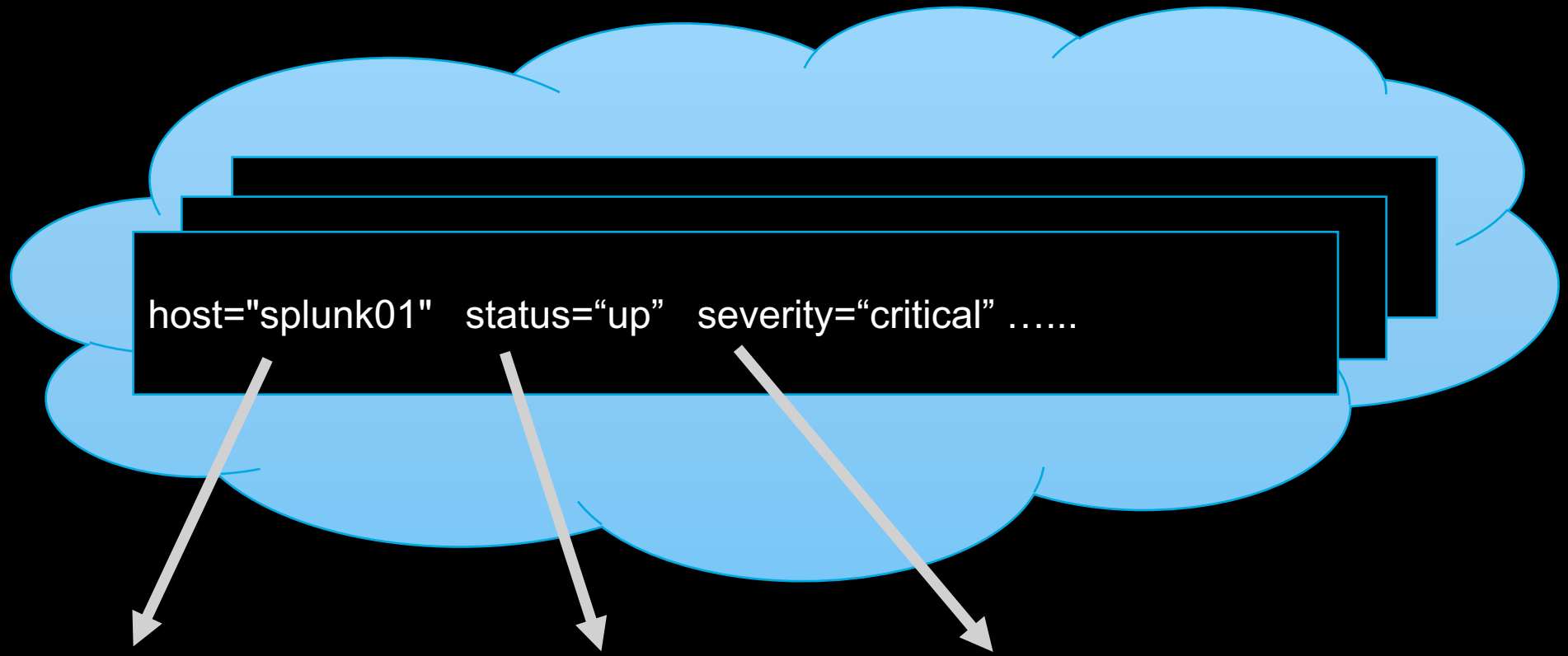
- The computational complexity is exponential
  - Number of event feature fields = 78
  - Number of possible feature values = 4
  - Number of possible groupings = 2^{80}
- An event can belong to many groups
- Brute force is not a practical approach for real time data





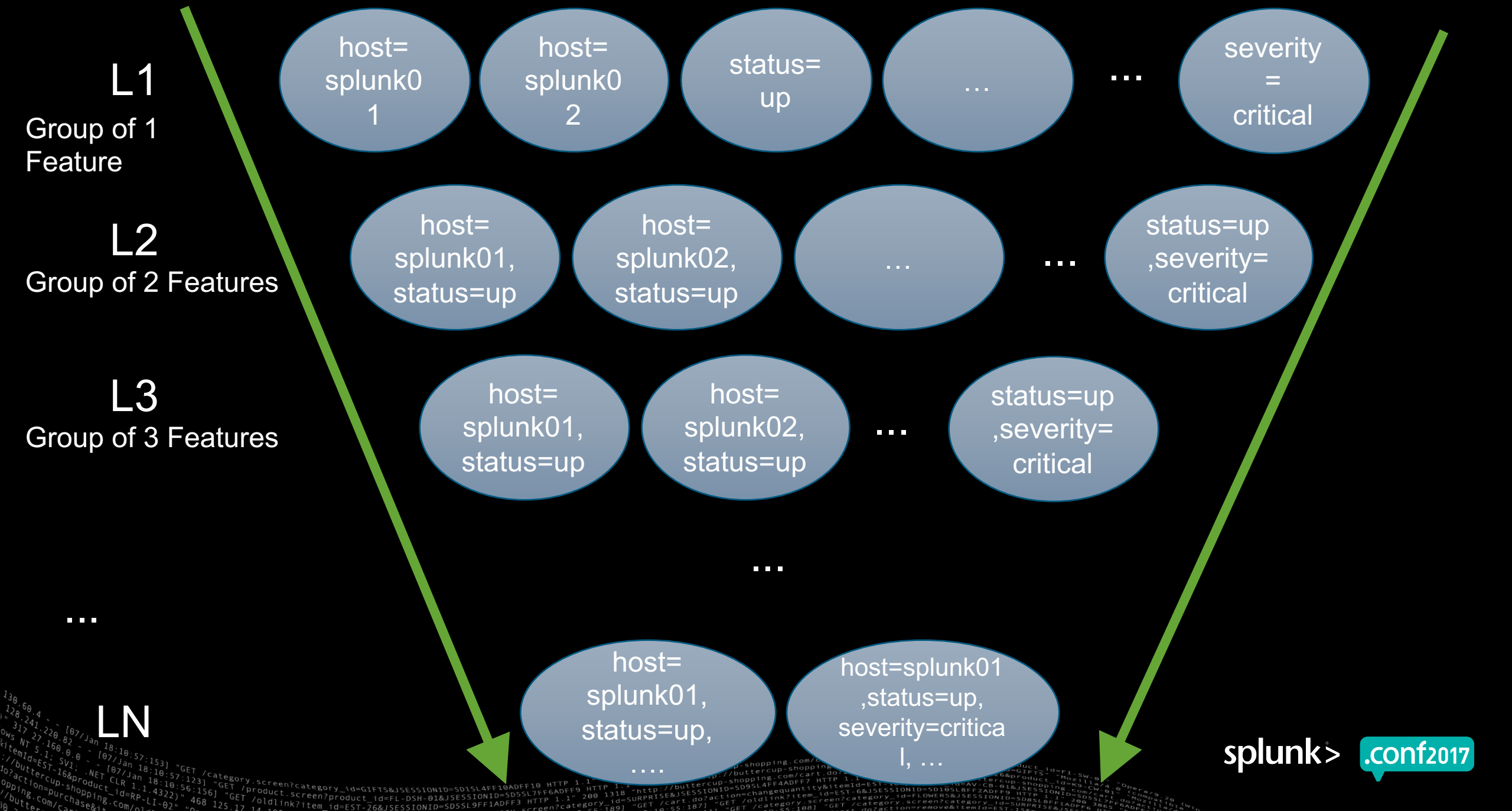
Approach and Algorithm

Step 1- Feature Extraction



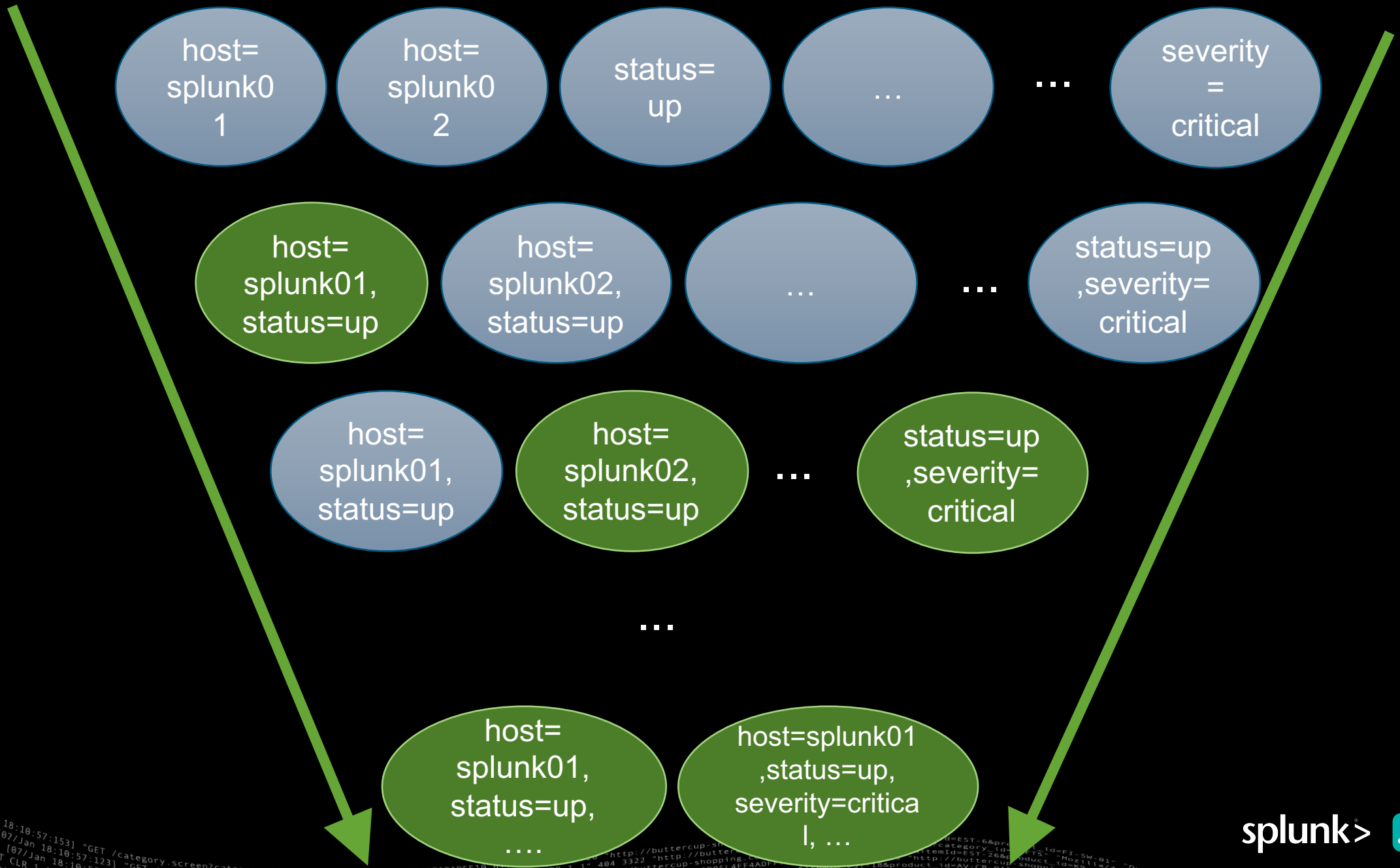
Step 2 – Reverse Pyramid Clustering

splunk> .conf2017

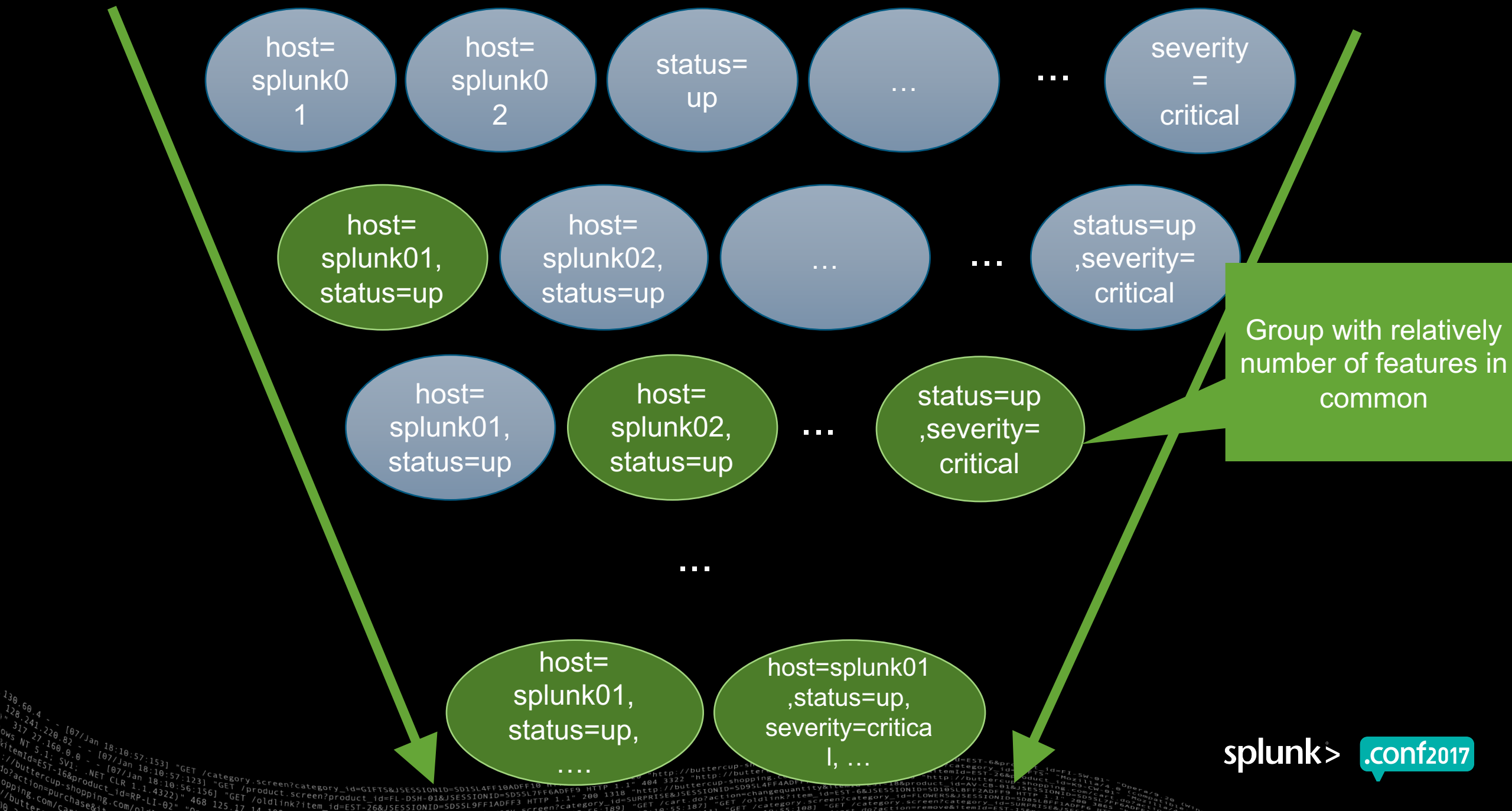


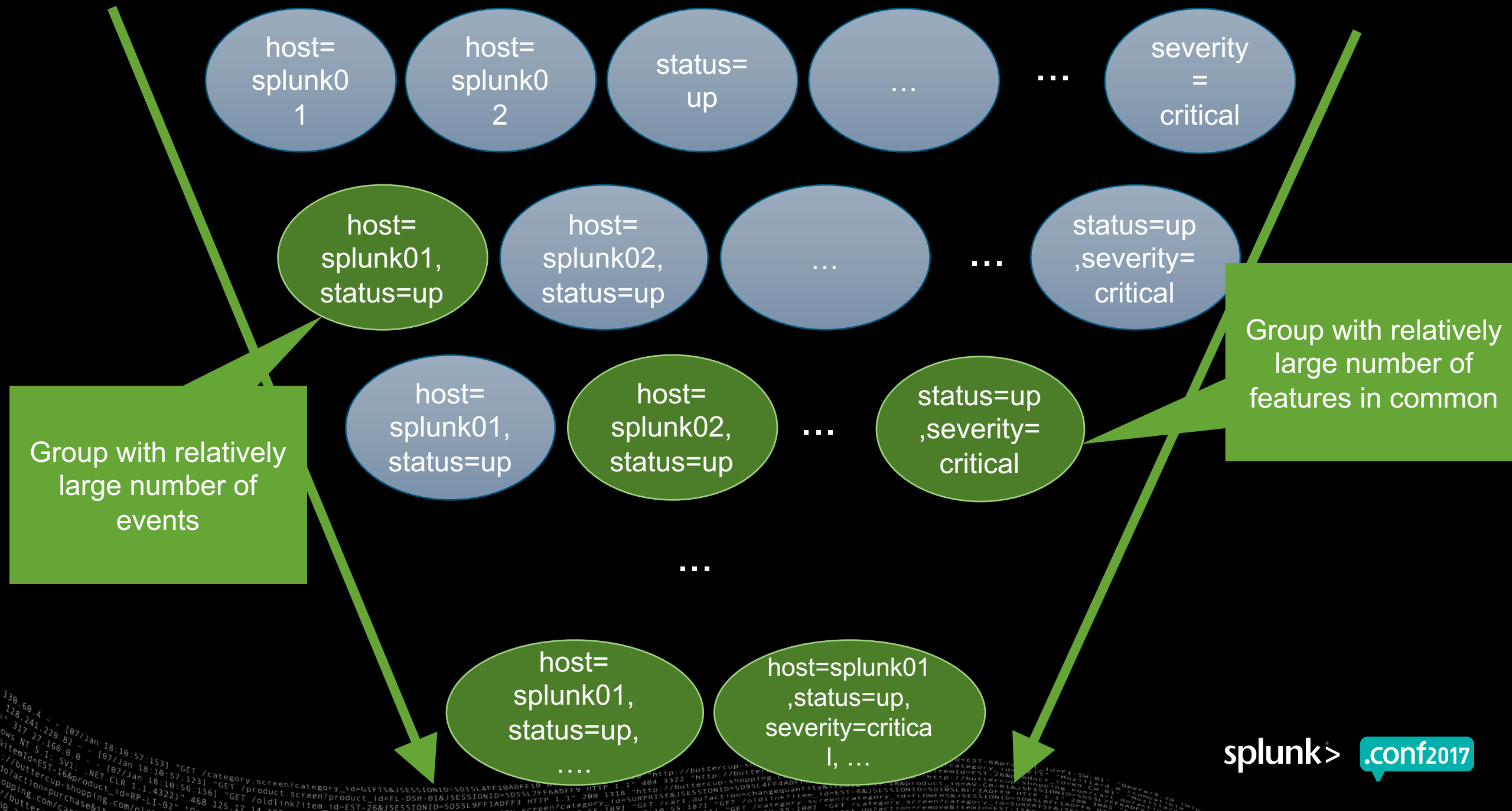
splunk> .conf2017

Step 3 – Backtracking



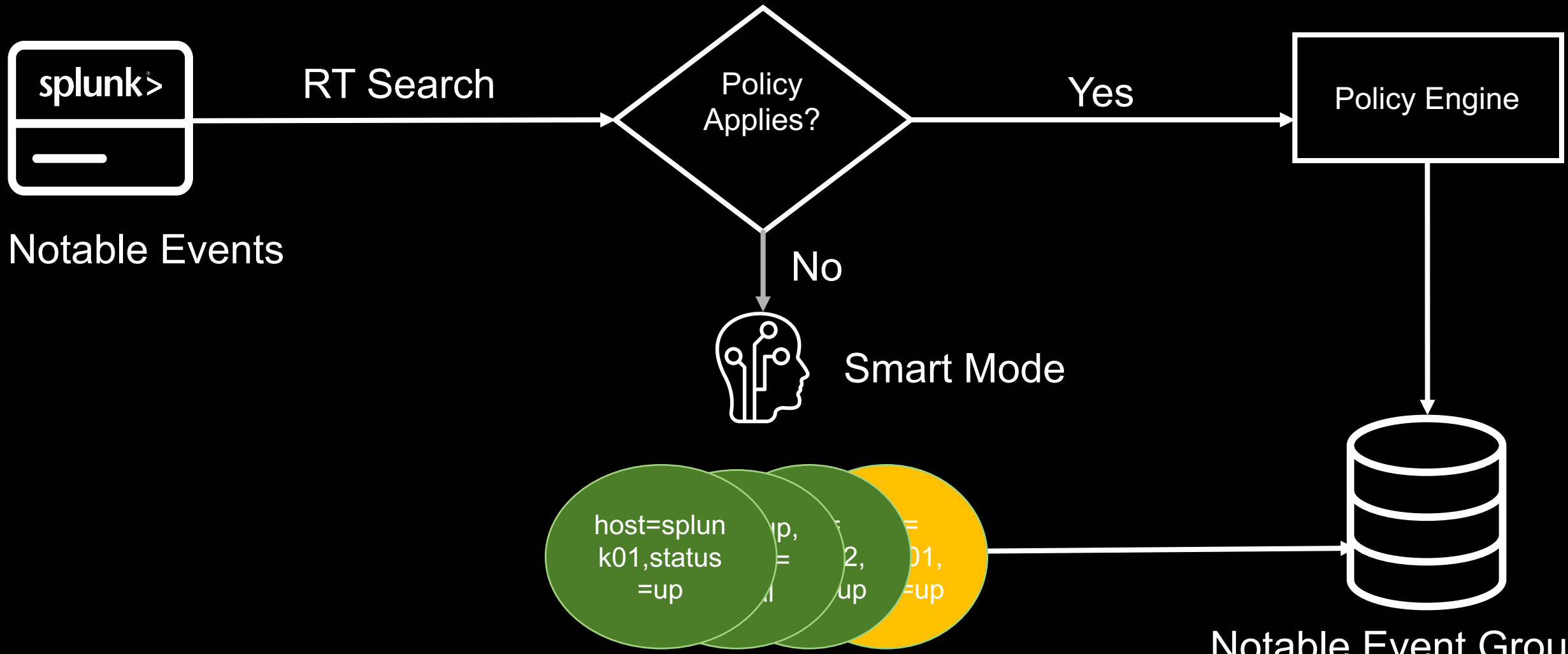
splunk> .conf2017

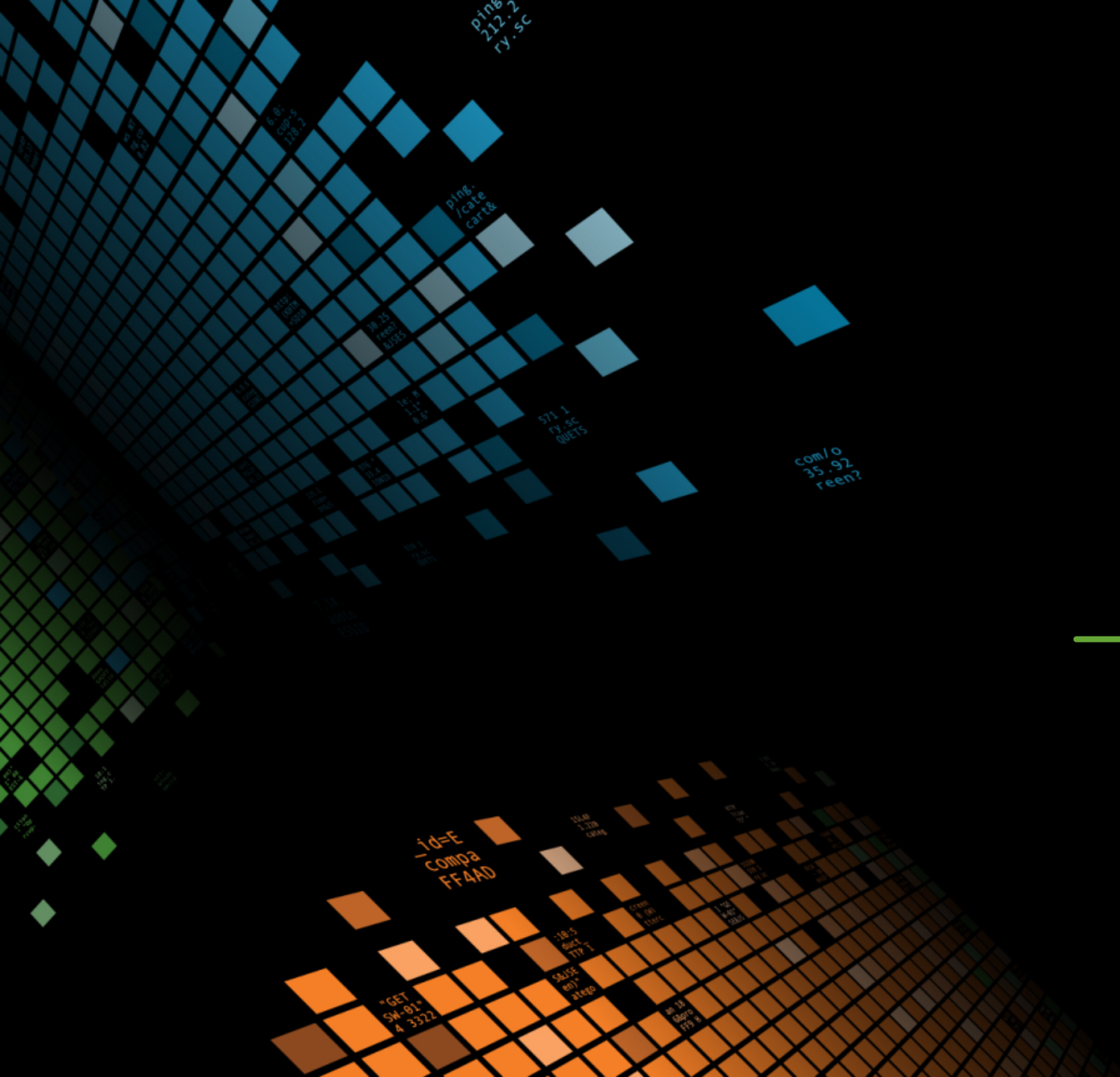




splunk> .conf2017

Let's run in real time!





Splunk Demo

Presented by Buttercup Splunker

splunk> .conf2017

The principles behind a successful ML feature

This is where the subtitle goes

Explainability

The feature should be able to clearly explain to the analyst any inferences made.

Configurability

The feature should be easy to configure and tune by users of different technical abilities.

Extensibility

The feature should be extensible so that it can be quickly iterated upon.

splunk> .conf2017

The ITSI way

This is where the subtitle goes

Explainability

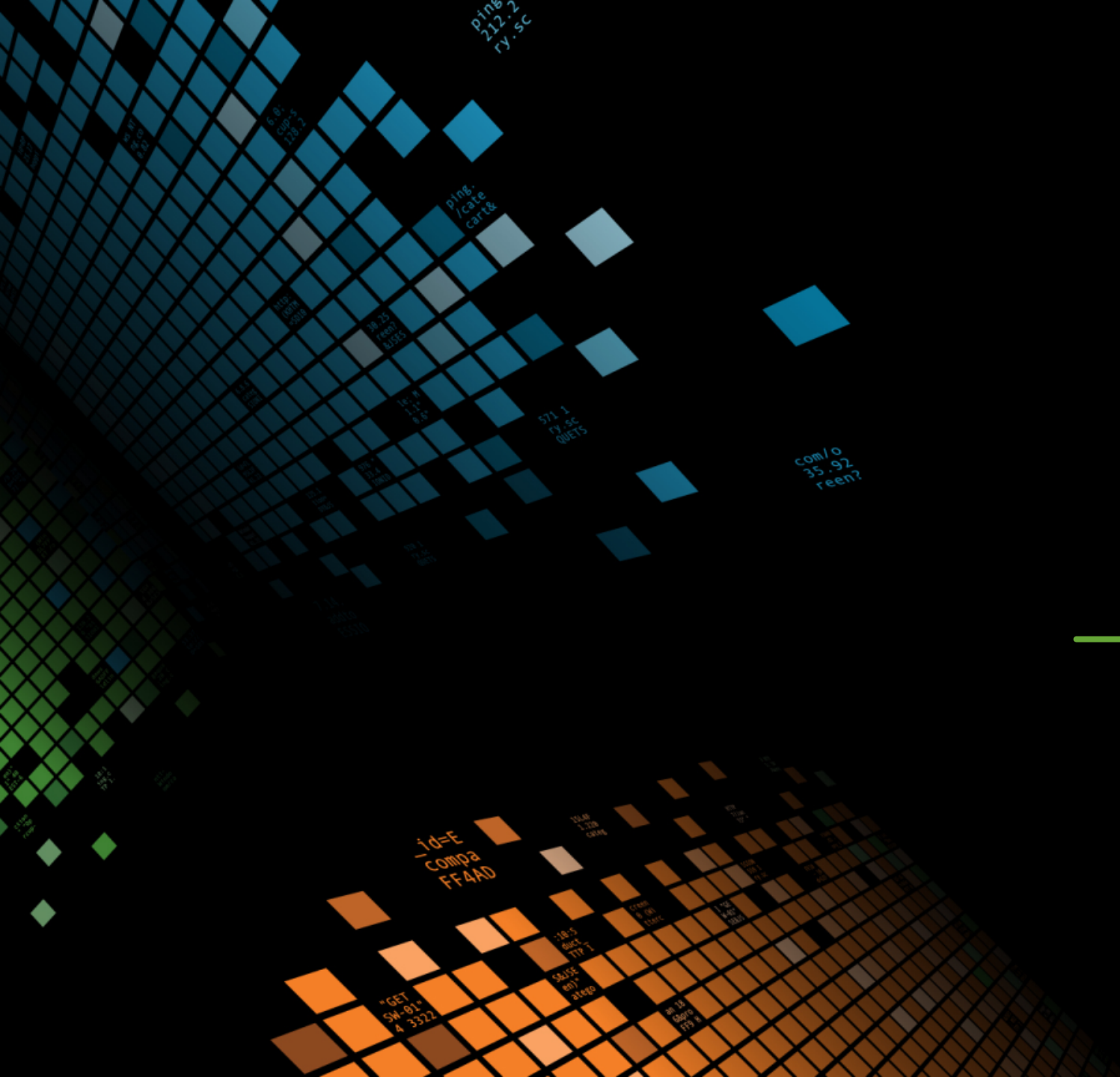
Every group that is discovered includes a human readable explanation of how it came to be.

Configurability

ITSI provides a UI that lets non-Data Scientists tune the algorithm's parameters.

Extensibility

Smart Mode is built on top of a Machine Learning Engine, on which additional "Factors" will be added over time.



Q&A

Participant name | Role

Participant name | Role

splunk> .conf2017