Splunk IT Service Intelligence on Steroids

How to boost performance using the Splunk Metric Store

Oliver Hoppe

August 30, 2018 | Version 1.2

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

“Improvise, Adapt, Overcome.”

Clint Eastwood - Heartbreak Ridge (1986)

splunk> .conf18

Agenda

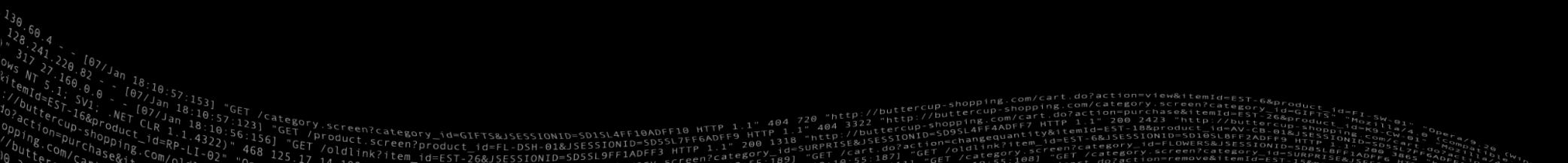
What we will talk about today

- Splunk IT Service Intelligence
  - Basic Concept of how to configure KPIs
  - Base Searches improve performance and reduce Cluster Load
- Splunk Metric Store
  - Using the Splunk Metric Store
- Splunk HTTP Event Collector
  - Using the Splunk HTTP Event Collector to ingest data
- Problem Statement
- Creating a Solution by bringing it all together
- Demo
- Conclusion
- Q&A

Splunk IT Service Intelligence (ITSI)

Introduction

- A monitoring and analytics solution that gives you visibility across IT and business services, and enables you to use AI to go from reactive to predictive IT.
- Proactively monitor your critical services and applications and use Deep Dives to effectively troubleshoot issues, identify the root cause of service degradations and enable alerting across multiple KPIs.
- Use the machine learning-powered service analyzer tree to leverage the right data and quickly determine the service or application origin of an incident and its root cause.
- Easily sift through vast amounts of events by filtering and sorting them based on priority. Additionally, trigger alerts, initiate remediation and automate incident workflows. Real-time anomaly detection also reduces alert fatigue.
- Customize visualizations of your IT services and key business metrics, and map KPIs to these visualizations to easily view the health and performance of what matters most.

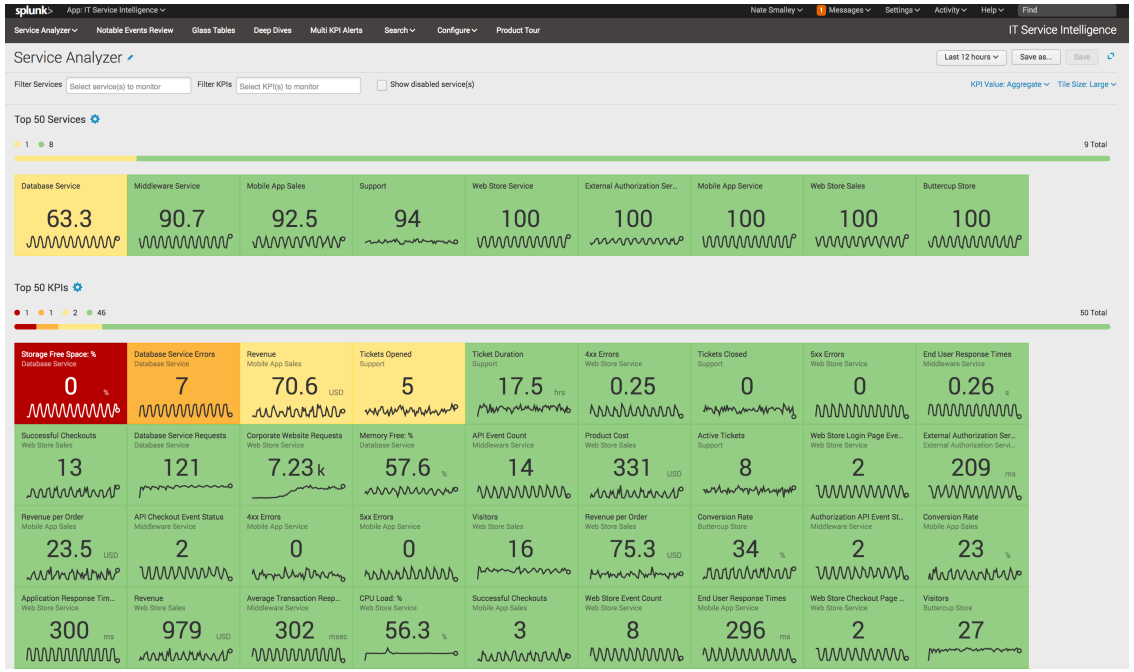


Splunk IT Service Intelligence (ITSI)

Basic Concept of how to configure KPIs

- KPIs are built by standard searches or data models
- Standard searches support any kind of data source and search query
- Each calculated KPI runs on a certain schedule and time span
- Adding more KPIs or building many KPIs with a frequent schedule and a long time span increases Cluster Utilization

Splunk IT Service Intelligence™

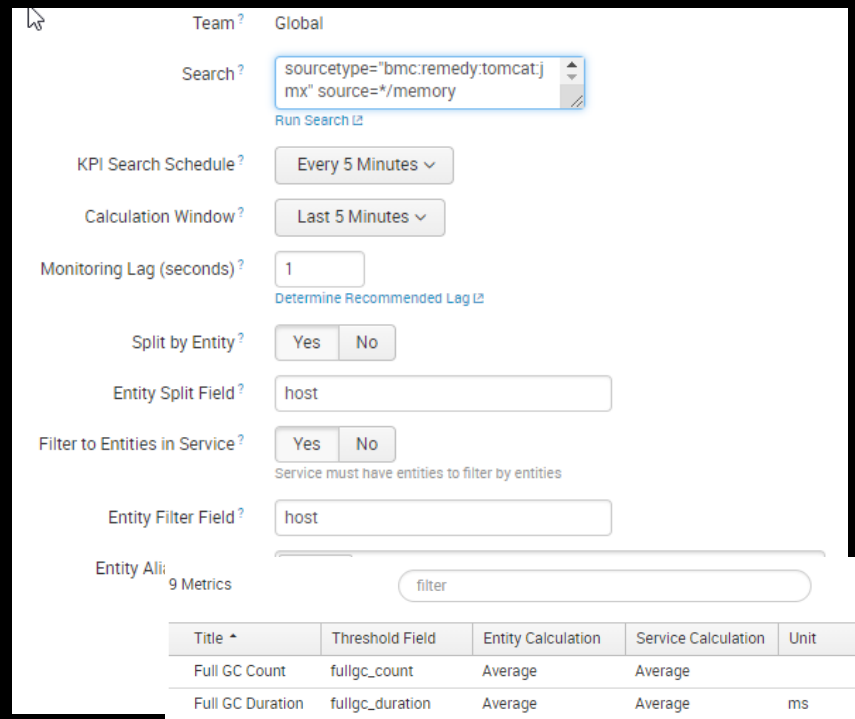


splunk> .conf18

Splunk IT Service Intelligence (ITSI)

Base searches improve performance and reduce Cluster Load

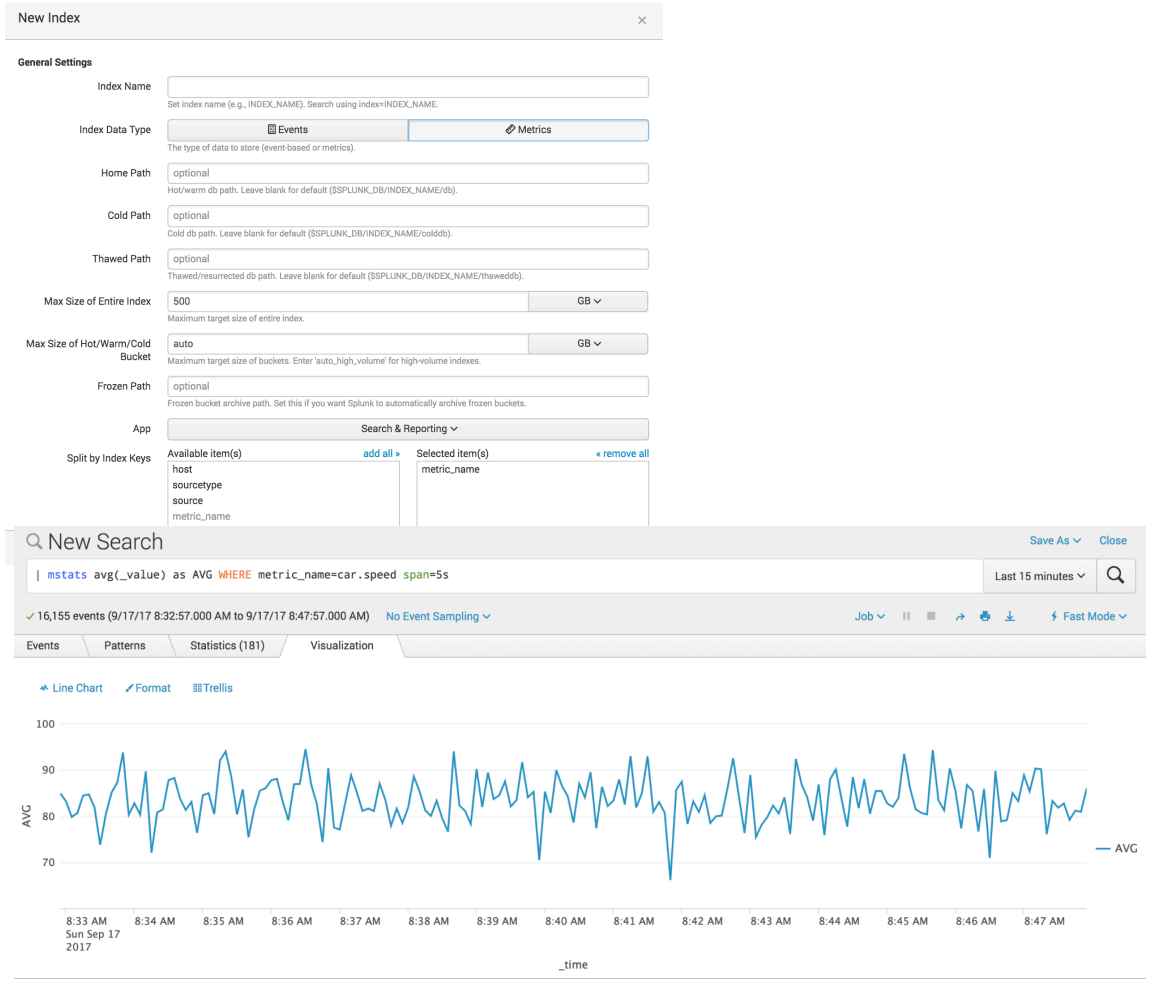
- Base Search reduce Cluster Load by using one common data input source to produce multiple KPIs during the same run
- Example: Building 4 different KPIs from an Apache HTTPD log for a highly utilized front end such as distinct users, average response time and count of 4XX and 5XX errors
  - Usually 4 searches
  - Using a Base Search brings a cluster load reduction of 75%
- ITSI prebuilt modules massively use Base Searches



splunk> .conf18

Splunk Metric Store

- Introduced as new Feature in Splunk 7.0
- Up to 200x faster than a traditional event index
- Follows the exact same definition and replication as an event index
- Consumes 150bytes of license per metric ingested
- Requires about 50% less disk storage space compared to storing the same payload in an events index



Splunk Metric Store

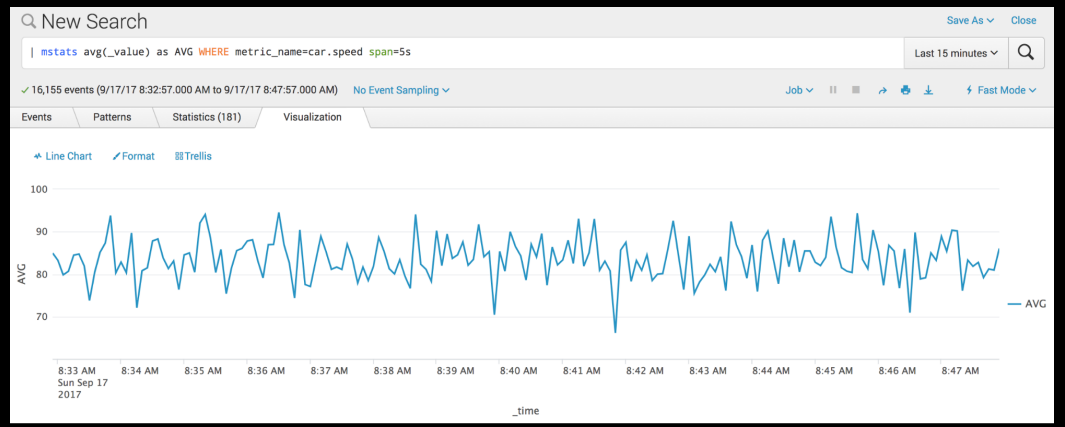
Using the Splunk Metric Store

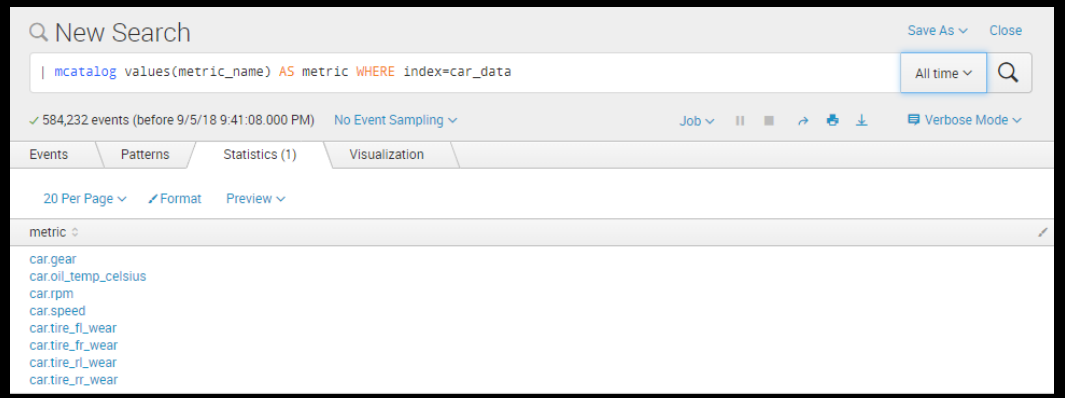
- mstats
  - Use the mstats command to analyze metrics
  - Example: | mstats avg(mycloud.compute.cpuIdle) WHERE index=my-metrics span=30s

    Returns the average value of the cpuIdle metric from my-metrics index and buckets them into 30s time spans
- mcatalog
  - Use the mcatalog command to search metrics data
  - Example: | mcatalog values(metric_name) WHERE index=my-metrics

    Returns all matric names of the metric index “my-metrics”
- mcollect
  - Use the mcollect command to convert events into metric data
  - Example: index=mycloud sourcetype=cpuStats
    | eval metric_name=mycloud.computecpuIdle
    | timechart span=60s min(cpuIdle) as minIdle by host
    | rename minIdle AS _value
    | mcollect index=my-metrics host

    Ingests the minimum idle cpu value for each host for each 60s bucket into the my-metrics index

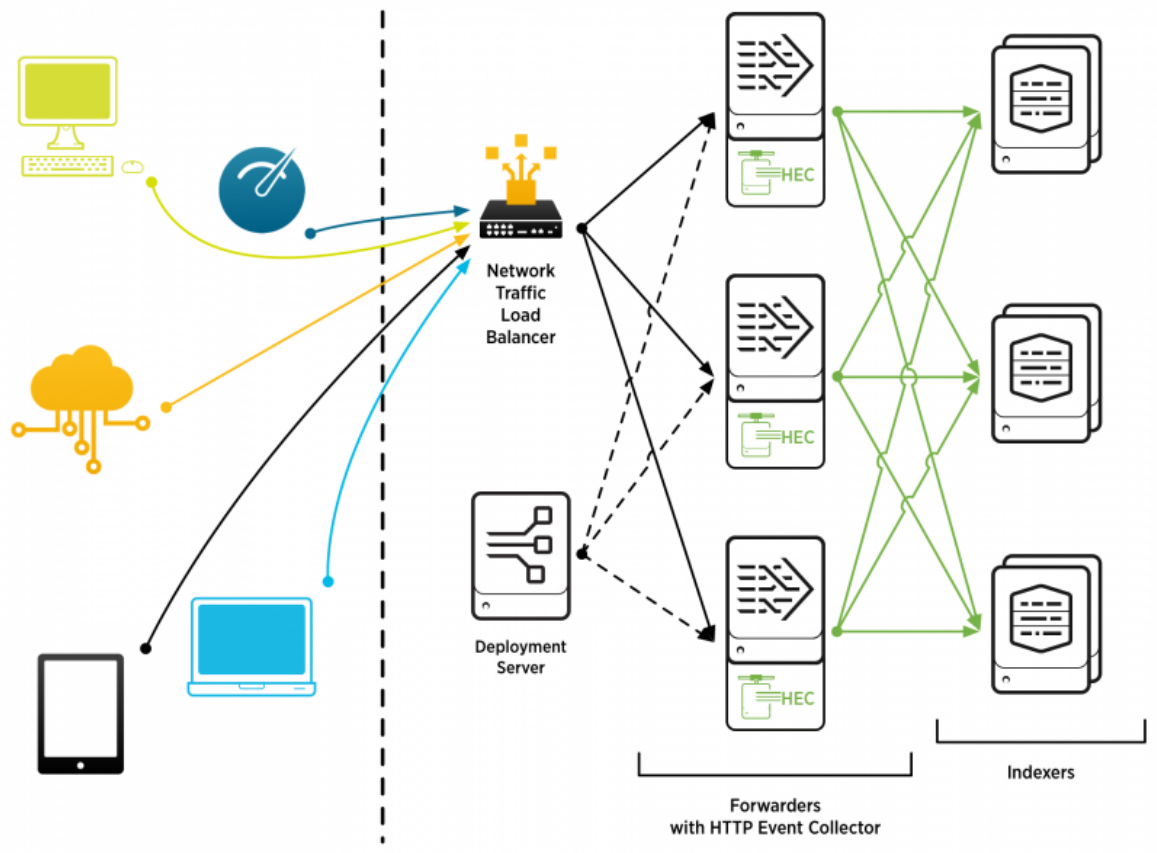




splunk> .conf18

Splunk HTTP Event Collector

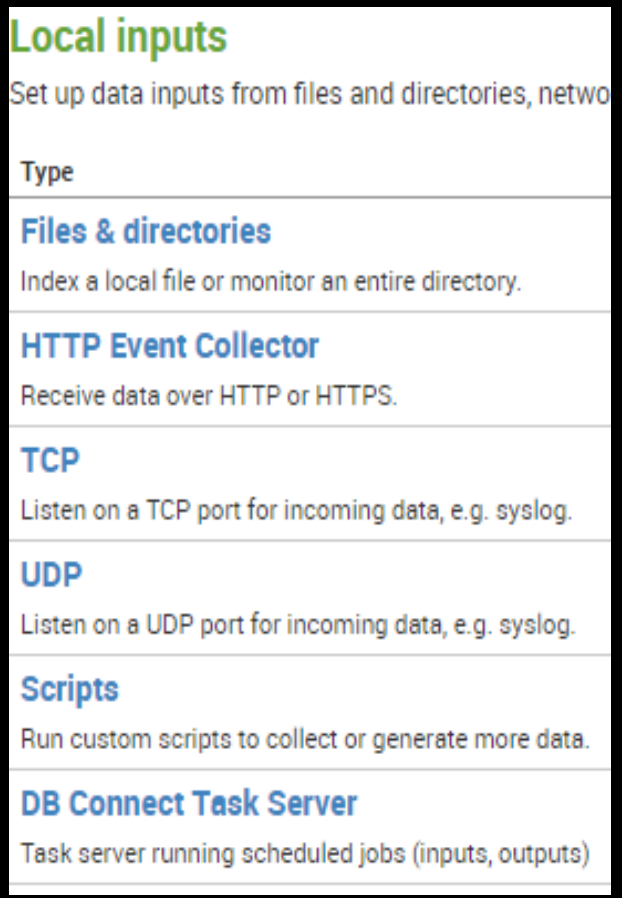
- Developer-friendly way to ingest data into Splunk
- Get data directly from the users browser-session or app usage
- Infrastructure metrics support for collectD and statsD
- Token based security for authentication
- Support of persistent queues to prevent data loss
- HTTPS for encryption is supported
- High-available and scalable due to its architecture
- The easiest way to ingest data from a streaming pipeline that should be transformed

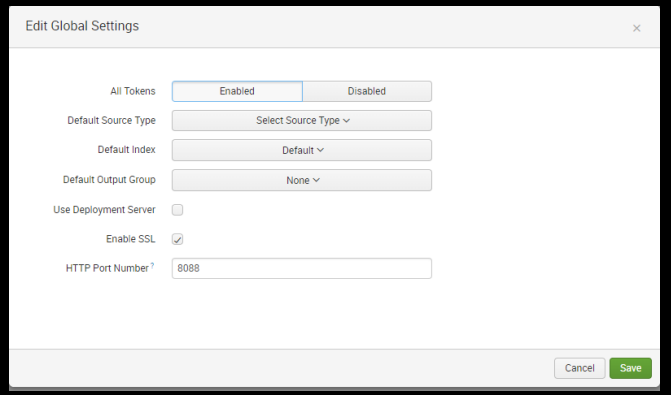


Splunk HTTP Event Collector

Using the Splunk HTTP Event Collector to ingest data

- Enabling the basic HTTP Event Collector via SplunkWeb takes just seconds
- Building a more complex HTTP Event Collector Cluster on intermediate Heavy Forwarders makes sense for bigger production deployments
- Sending an event to the HTTP Collector is easy
  - curl -k https://localhost:8088/services/collector -H "Authorization: Splunk b0221cd8-c4b4-465a-9a3c-273e3a75aa29“ -d '{"time": 1486683865.000, "event":"metric", "source":"disk", "host":"host_99", "fields":{"region":"us-west-1","datacenter":"us-west-1a","rack":"63","os":"Ubuntu16.10","arch":"x64","team":"LON","service":"6","service_version":"0","service_environment":"test","path":"/dev/sda1","fstype":"ext3","_value":1099511627776,"metric_name":"total"}}'



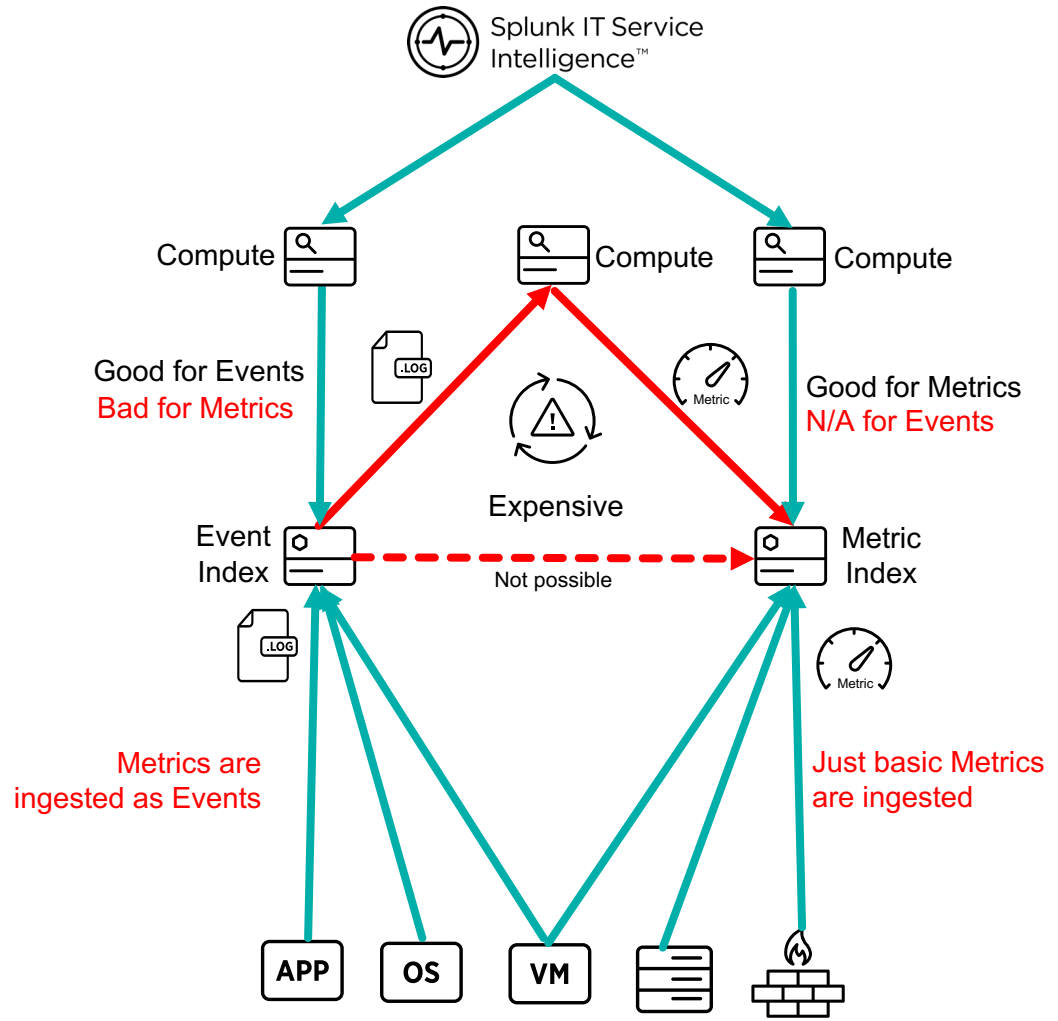


splunk> .conf18

Problem Statement

When your scheduler is about to die

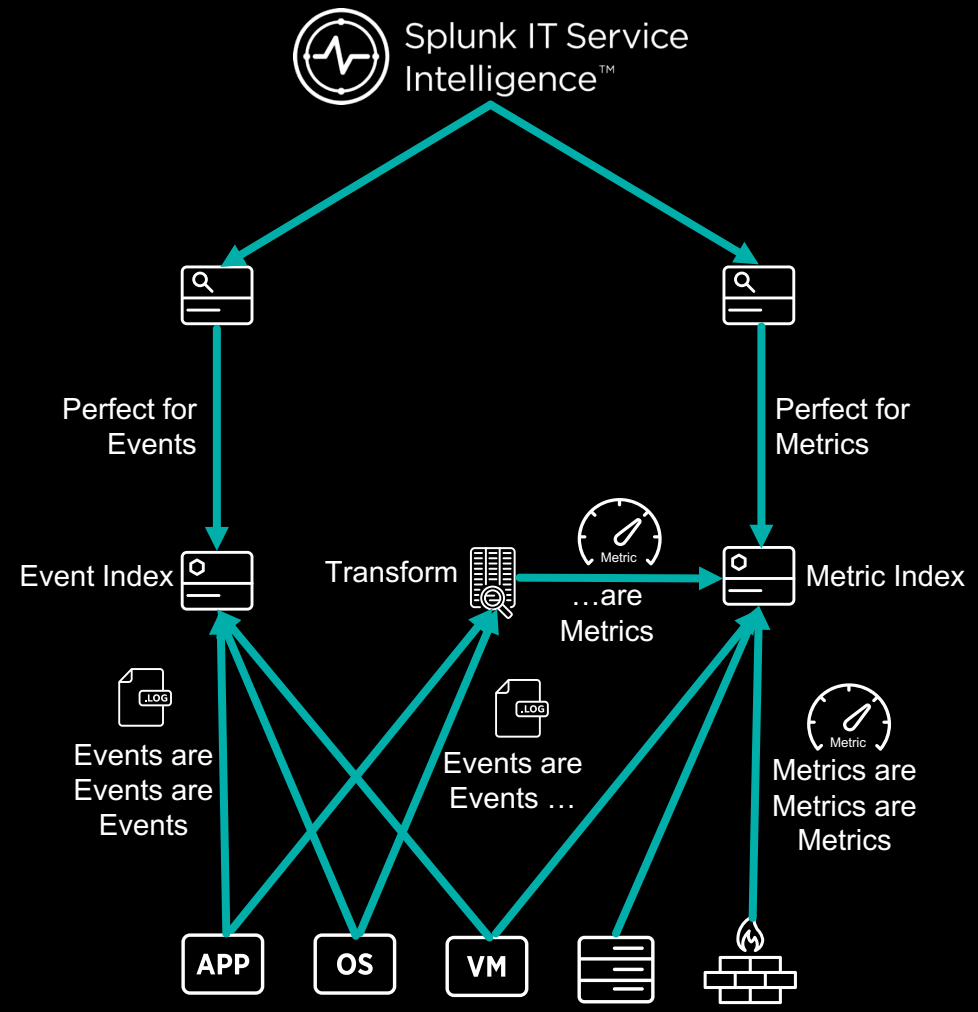
- Facts
  - Compute and IO are always limited in a centralized system
  - Persisting event data for the sake of transforming it into metric data using compute is fundamentally wrong
  - What looks like a performant centralized compute Cluster today might break tomorrow
  - There is no infinite scale
  - Not every metric is offered in a metric format
  - Transforming events into metrics at a central system does not scale long-term
- Assumptions (or hard to prove facts)
  - The amount of Data generated is exponentially growing
  - The need and acceptance of using data by the critical mass of every company is becoming reality
  - More data does not bring more value
    → more information brings more value
  - Decentralizing is the solution



Creating a Solution by Bringing it all Together

Plus adding a small piece

- Removing the compute to transform events into metrics from the central cluster massively improves performance and cluster utilization
- Transformation of events into metrics is ideally decentralized or at least decoupled from the central cluster
- The streaming processor of your choice will transform an event stream into a metric stream that can be ingested easily via the HTTP Event Collector



Demo

Conclusion

What we have seen and maybe learned today

- Running ITSI at scale requires
  - Need to ingest events for building KPIs that are naturally not representable as metrics
  - Need to ingest metrics for building KPIs that are already generated as metrics
  - Transform events into metrics where the Information is a metric but for various reasons is in an event format
- Combining different Splunk Products to achieve the goal is the right way
  - Splunk is a Data platform with several components; use and combine them
  - Complement the Product Suite where required with the technology that fits you best
- Improvise by adding more hardware for short term solutions
- Adapt by identifying the problem and defining a long term strategy
- Overcome by changing architecture

Q&A

Please ask here or approach me after the session

Thank You

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