Intro to Splunk for DBAs

Holly Willey

Sr. Sales Engineer, Splunk

splunk>

Disclaimer

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 the 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.

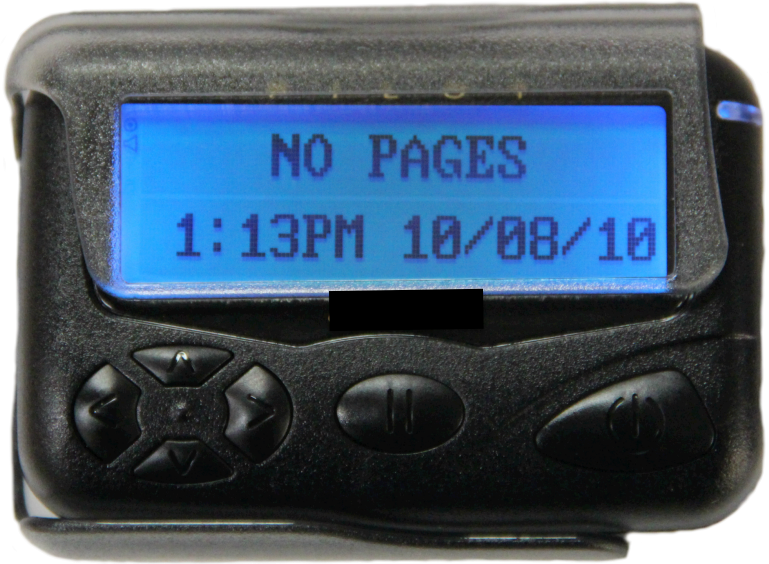
In addition, any information about our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only and shall not, be incorporated into any contract or other commitment. Splunk undertakes no obligation either to develop the features or functionality described or to include any such feature or functionality in a future release.

Understanding the Mindset

DBAs are responsible for the most important data repositories – mission critical systems

RDBMS Failure without backup or standby – worst case? Business IP lost = **business failure**

What Matters



Agenda

- Indexing & Searching
- Architecture
- Demo

Indexing & Searching

splunk>

Rise of Polyglot Persistence

E-commerce Platform

Session Data

Completed Orders

Product Catalog

Customer Social Graph

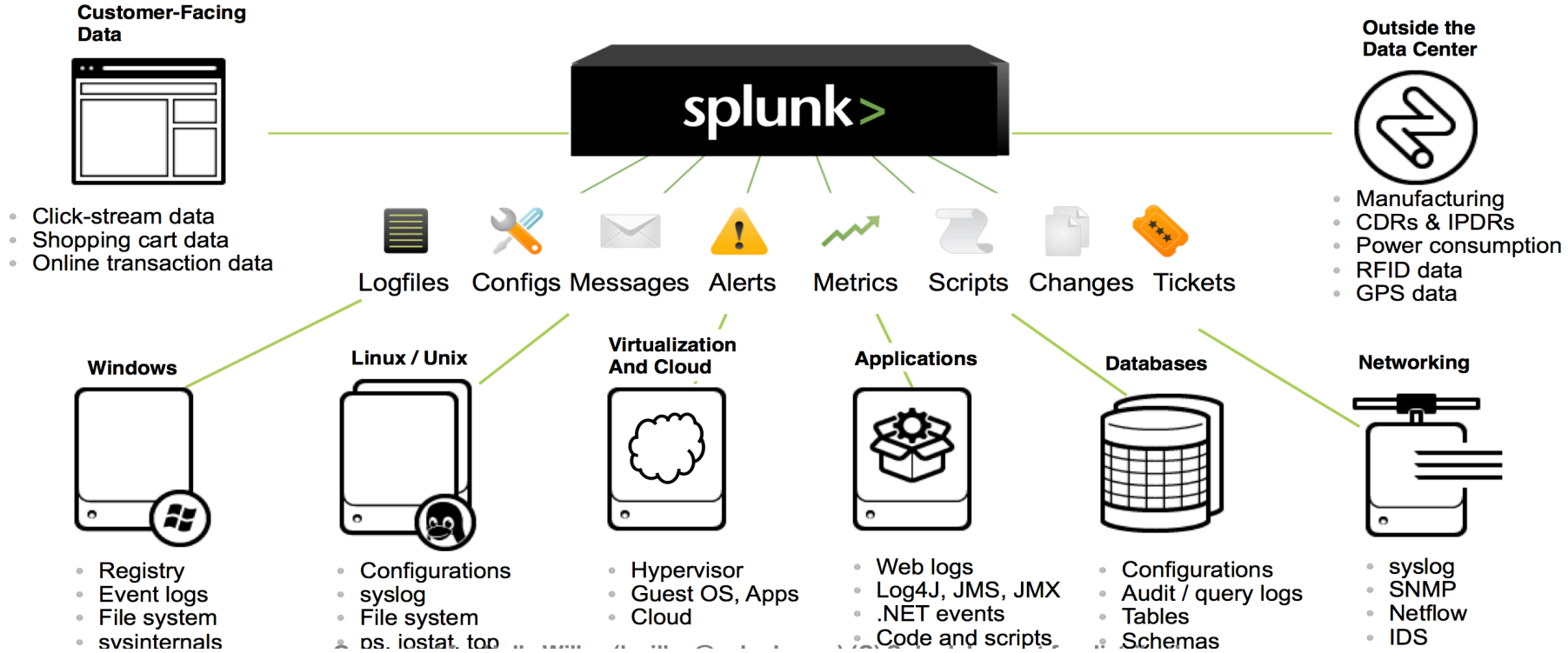
Key-Value store

Document store

RDBMS

Graph store

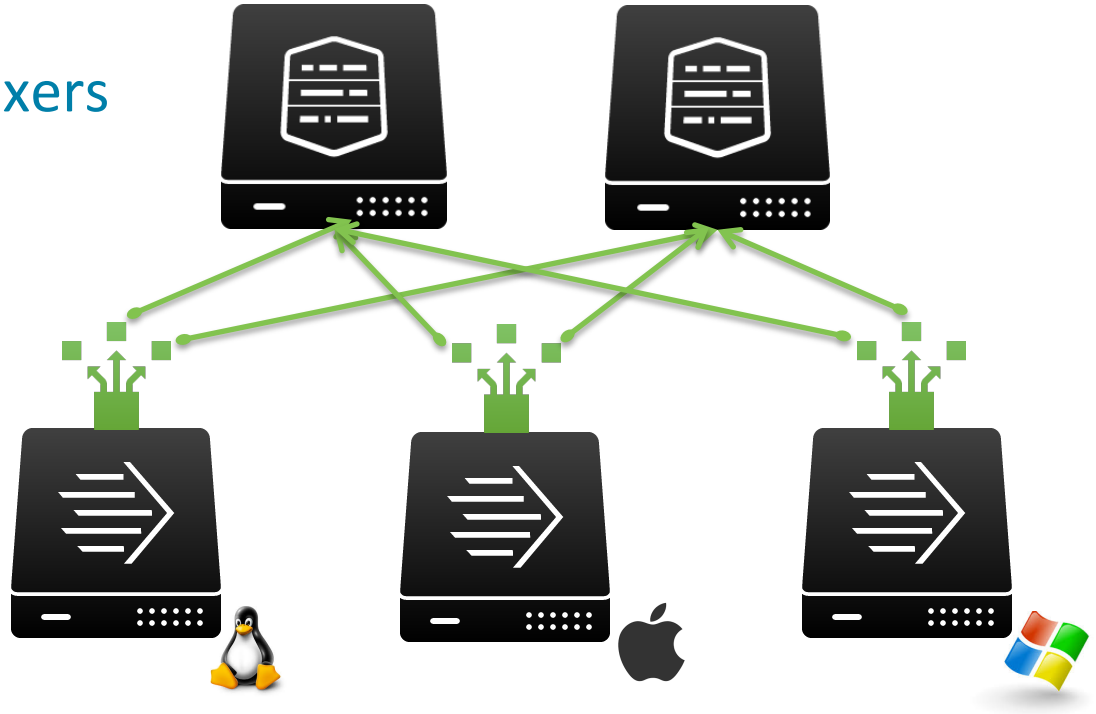
Collect & Index Machine Data



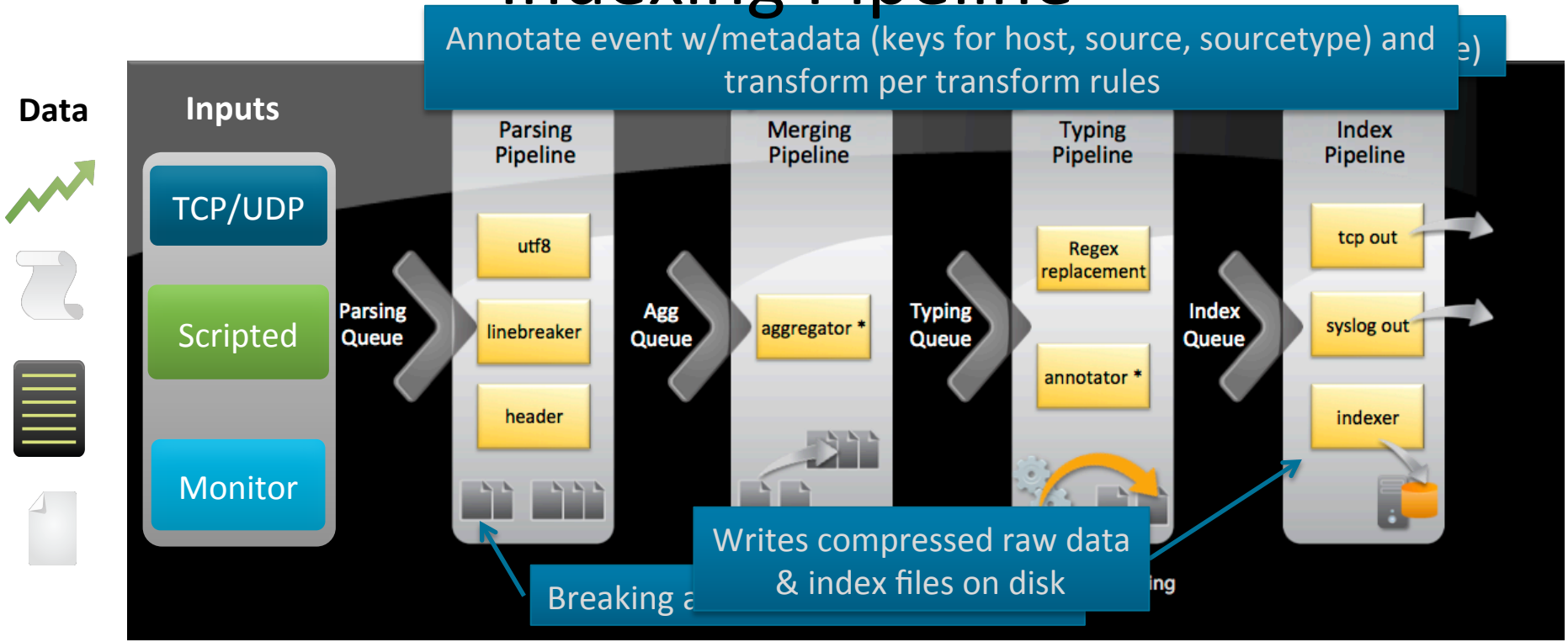
Forwarders & Indexers

Indexers

Forwarders
with load balancing



Indexing Pipeline



Search Processing Language (SPL)

SQL	SPL
`SELECT *` `FROM mytable`	`source=mytable`

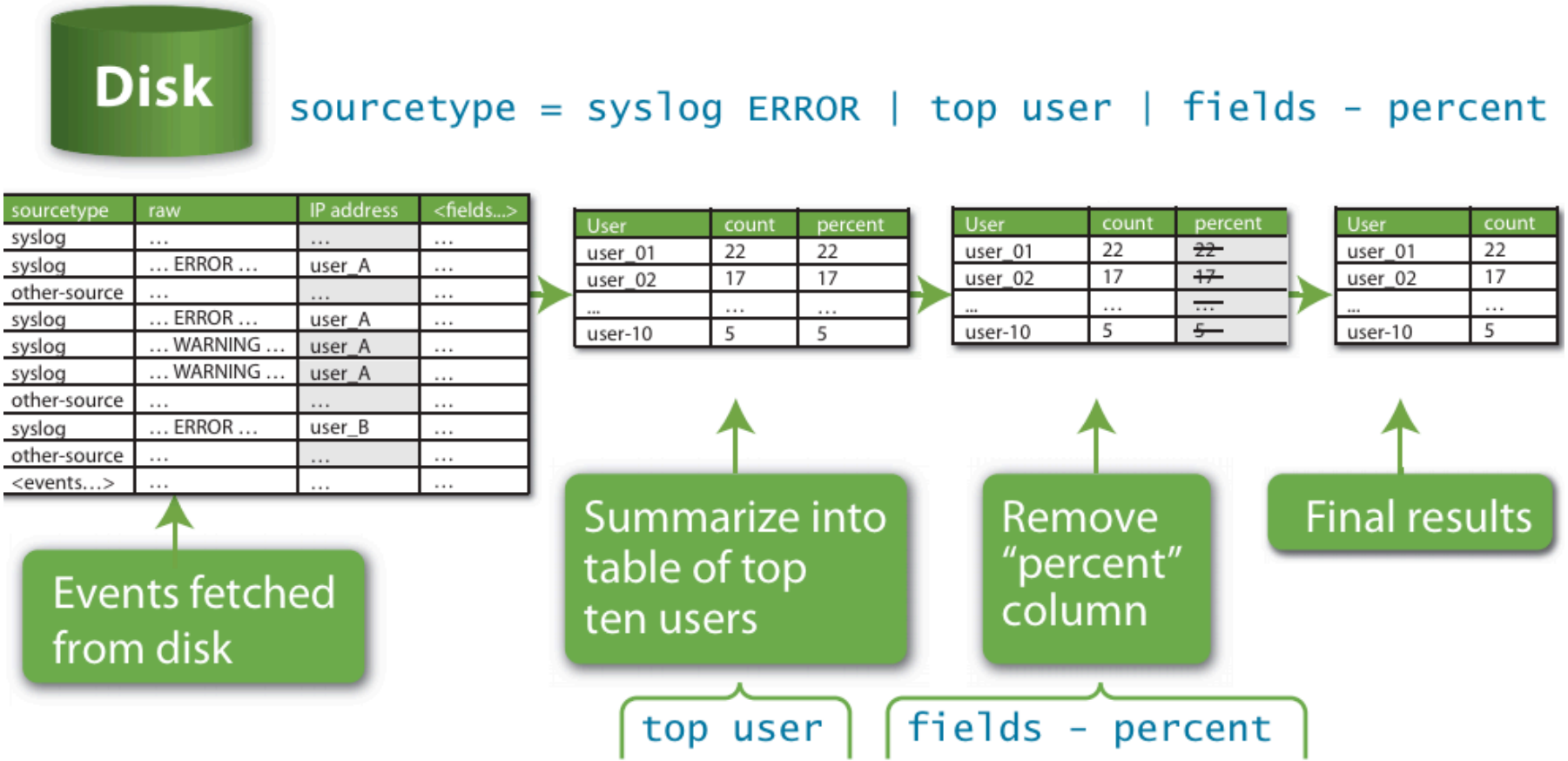
`SELECT *`
`FROM mytable`
`WHERE mycolumn=5`
	`source=mytable mycolumn=5`

`SELECT mycolumn1, mycolumn2` `FROM mytable`	`source=mytable` `| FIELDS mycolumn1, mycolumn2`

`SELECT *`
`FROM mytable`
`WHERE (mycolumn1="true" OR`
`mycolumn2="red") AND`
`mycolumn3="blue"`
	`source=mytable`
	`AND (mycolumn1="true" OR`
	`mycolumn2="red")`
	`AND mycolumn3="blue"`

Popular SPL Commands

Command	Description
dedup	Removes subsequent results matching a specified criteria
head/tail	Returns the first/last number *n* of specified results
top/rare	Displays the most/least common values of a field
timechart	Create a time series chart and corresponding table of statistics
transaction	Groups search results into transactions


Disk
sourcetype = syslog ERROR | top user | fields - percent
sourcetype raw IP address <fields...>
syslog … …
syslog … ERROR … user_A …
other-source … … …
syslog … ERROR … user_A …
syslog … WARNING … user_A …
syslog … WARNING … user_A …
other-source … … …
syslog … ERROR … user_B …
other-source … … …
<events…> … … …
User count percent
user_01 22 22
user_02 17 17
… … …
user-10 5 5
User count percent
user_01 22 22
user_02 17 17
… … …
user-10 5 5
User count
user_01 22
user_02 17
… …
user-10 5
Events fetched from disk
Summarize into table of top ten users
Remove "percent" column
Final results
top user
fields - percent

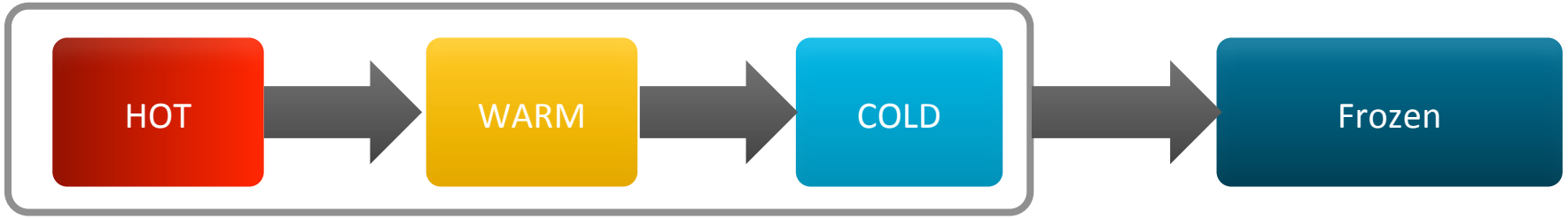
Terminology

RDBMS	Splunk
Query	Search
Table/View	Search Results
Index	Index
Row	Result/Event
Column	Field
Database/Schema	Index/App

Architecture

splunk>

Rotation & Retention



- Hot – Newest buckets of data that are still open for write
- Warm – Recent data but closed for writing (read only)
- Cold – Oldest data, commonly on cheaper, slower storage
- Frozen – No longer searchable, deleted or commonly archived data

HOT: Now to -3h	index	raw events
	index	raw events
	index	raw events
	index	raw events
	index	raw events
	index	raw events
59-Warm: -3 to -5h	index	raw events
…-Warm: -5 to -21h	index	raw events
	index	raw events
	index	raw events
49-Warm: -21 to -24h	index	raw events
48-Cold: -2d to -9d	index	raw events
…-Cold: -9d to -90d	index	raw events

Load Balancer

Cluster Members
+ 1 Captain

Search Head Cluster

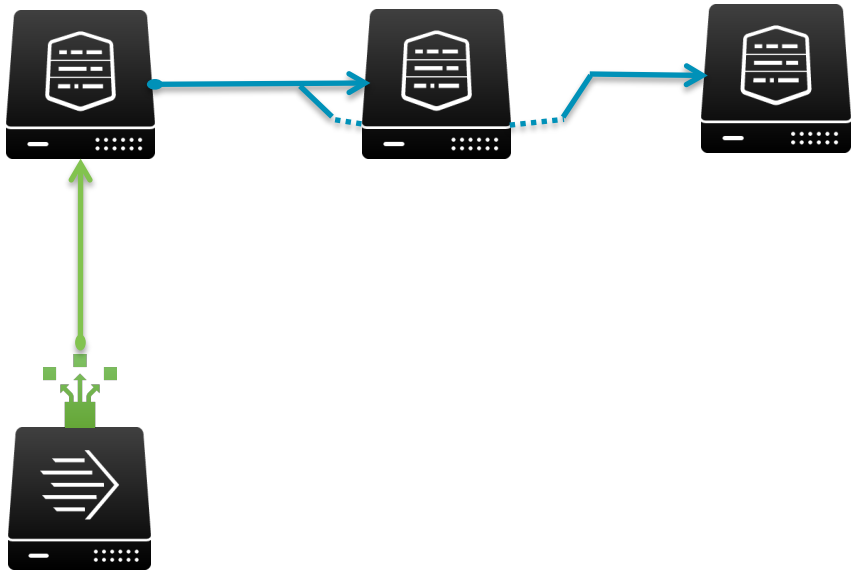
Peer Nodes (Search Peers)
+ 1 Master Node

Indexer Cluster

Forwarders
with
Load Balancing

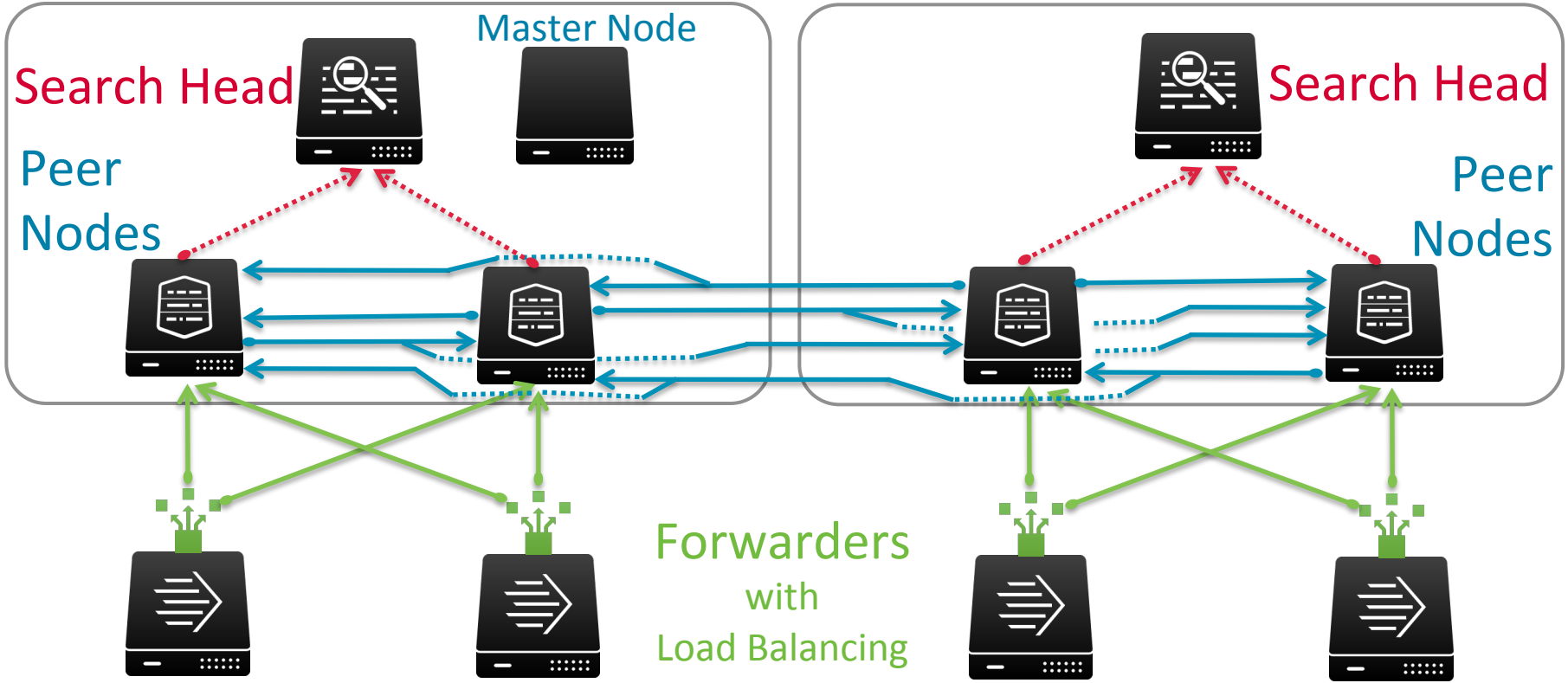
Replication Factor = 3

Index Cluster
Peer Nodes



Forwarder

Multisite Clusters



Compliance Requirements

Time series nature of Splunk indexing is uniquely suited to offloading logs

Ideal for centralized, consolidated retention & analysis of:

- Standard audit records
- Fine grained auditing trails
- Listener logs
- Alert logs

Demo

splunk>

Apps

- Splunk Add-on for Oracle Database
  https://splunkbase.splunk.com/app/1910/
- DB Connect
  https://splunkbase.splunk.com/app/2686/
- Oracle WebLogic App for Splunk
  https://splunkbase.splunk.com/app/1340/

Resources

- Real-Time Oracle 11g Log File Analysis
  https://pmdba.files.wordpress.com/2013/12/real-time-oracle-11g-log-file-analysis2.pdf
- Search Reference – Splunk for SQL Users
  http://docs.splunk.com/Documentation/Splunk/latest/SearchReference/SQLtoSplunk
- Exploring Splunk book in iOS / Kindle / PDF versions
  http://www.splunk.com/goto/book
- Quick reference guide
  http://www.splunk.com/web_assets/pdfs/secure/Splunk_Quick_Reference_Guide.pdf
- Splunk answers
  http://answers.splunk.com
- Splunk blogs
  http://blogs.splunk.com
- Splunk education
  http://www.splunk.com/view/education/SP-CAAAAH9
- Free eTraining
  https://inter.viewcentral.com/reg/splunk/elearning

THANK YOU

splunk>