



splunk® >

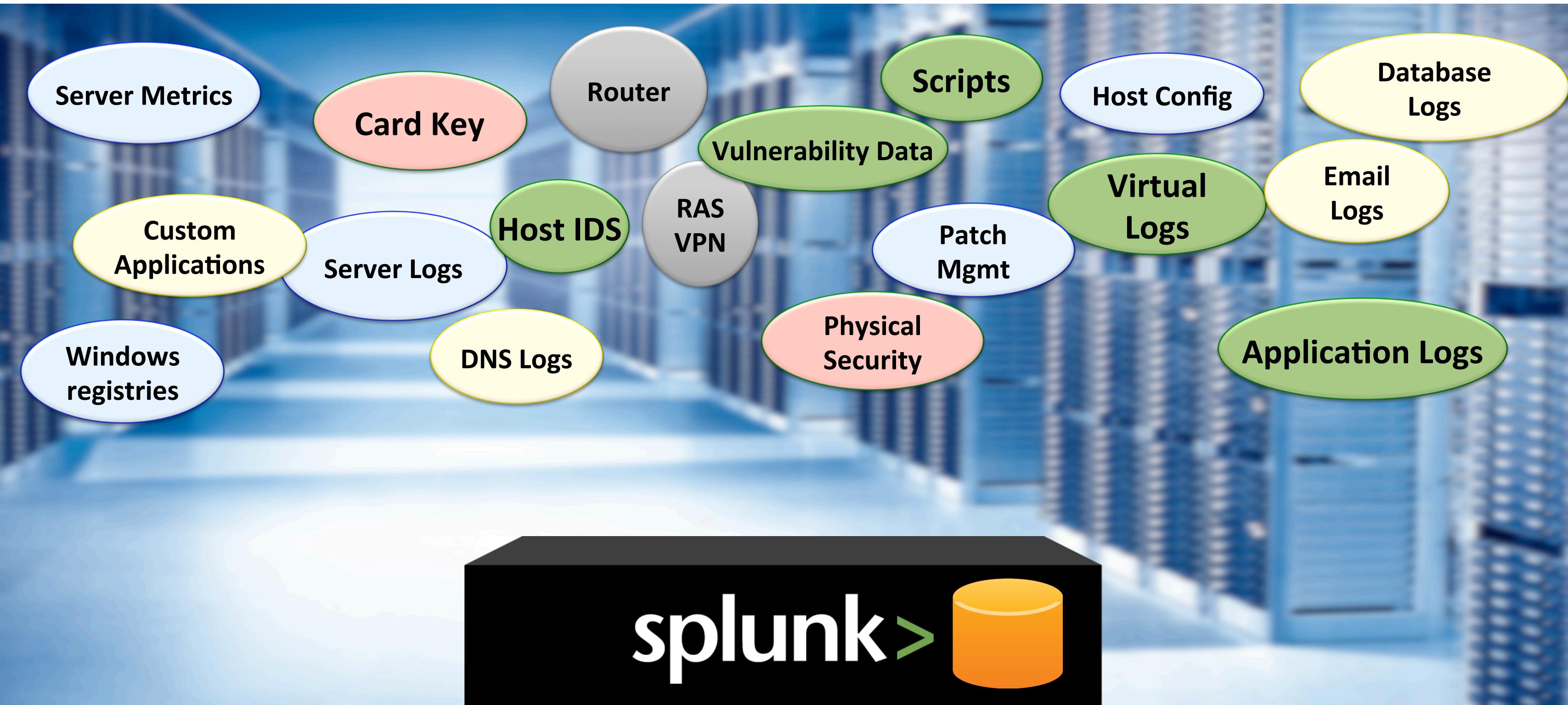
Overview

What does Splunk collect?

Machine Generated Data, not Human Generated Data

- Machine Data contains categorical record of all activity and behavior – customer behavior, user transactions, machine behavior

What does Splunk do?



What does Splunk Provide?

A common interface for all IT Data



search



alert



report



share

That provides Operational Intelligence



The Big Data Problem

What is Big Data?

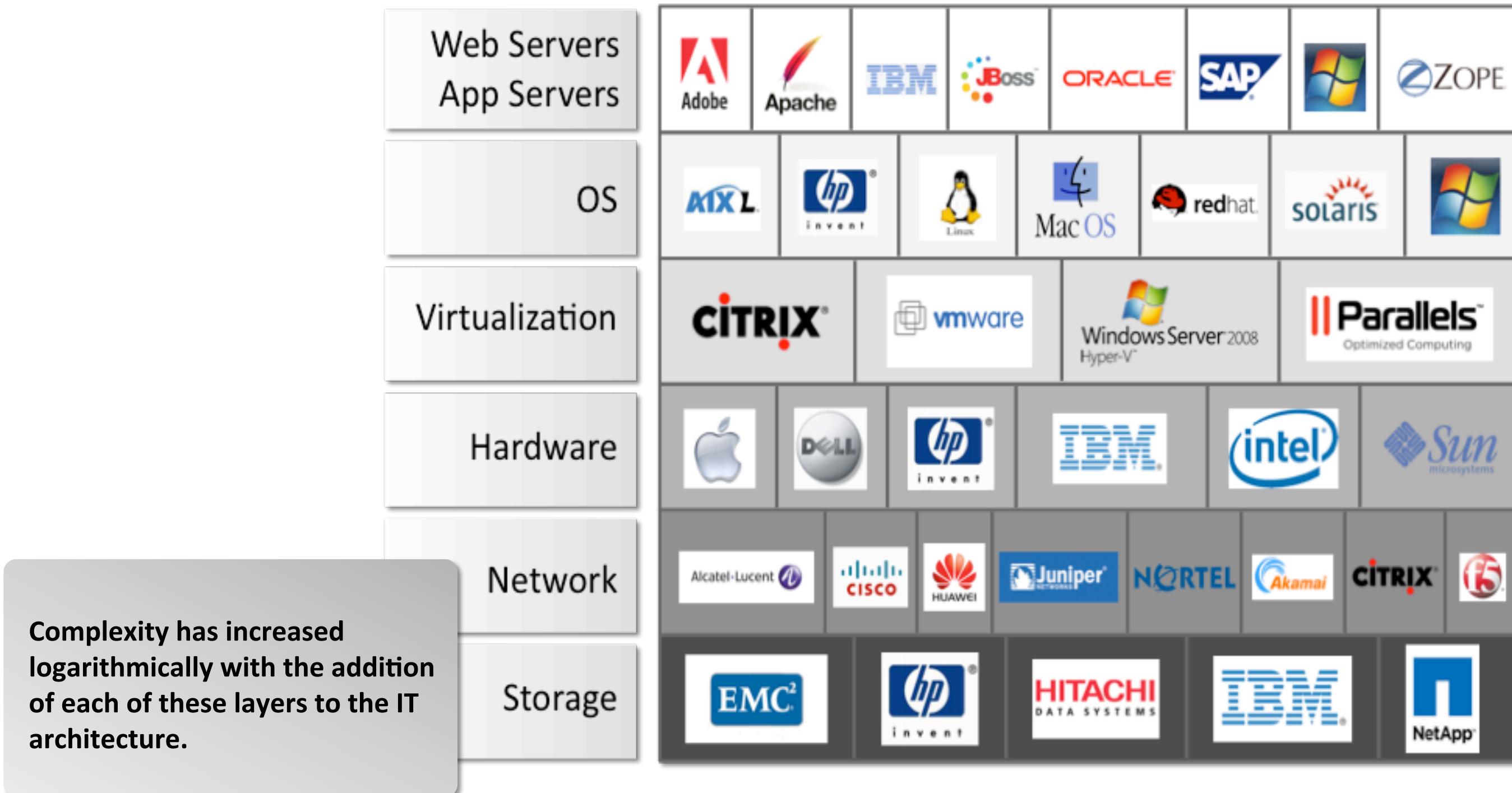
“...When the size of the data itself becomes part of the problem”

Mike Loukides

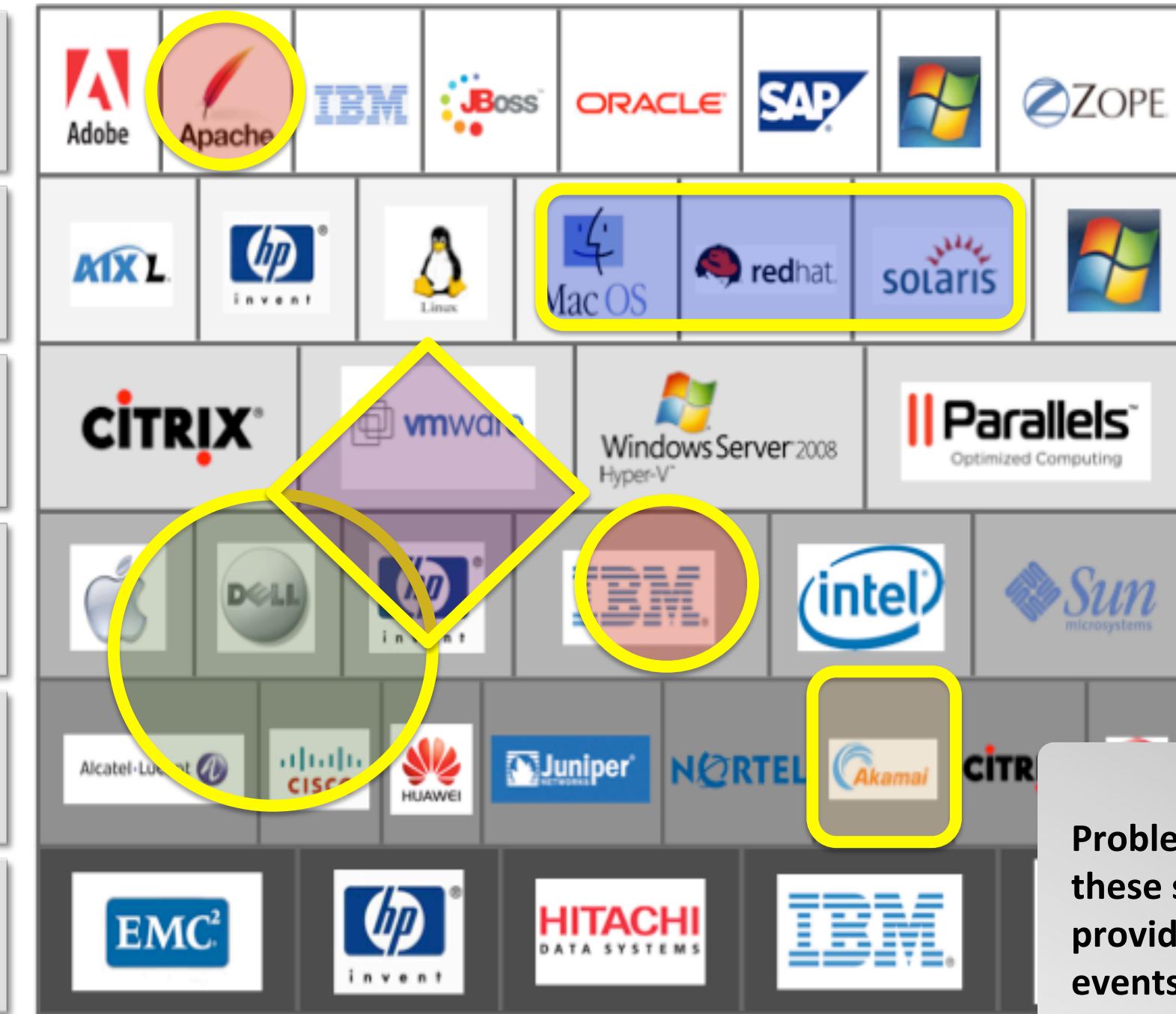
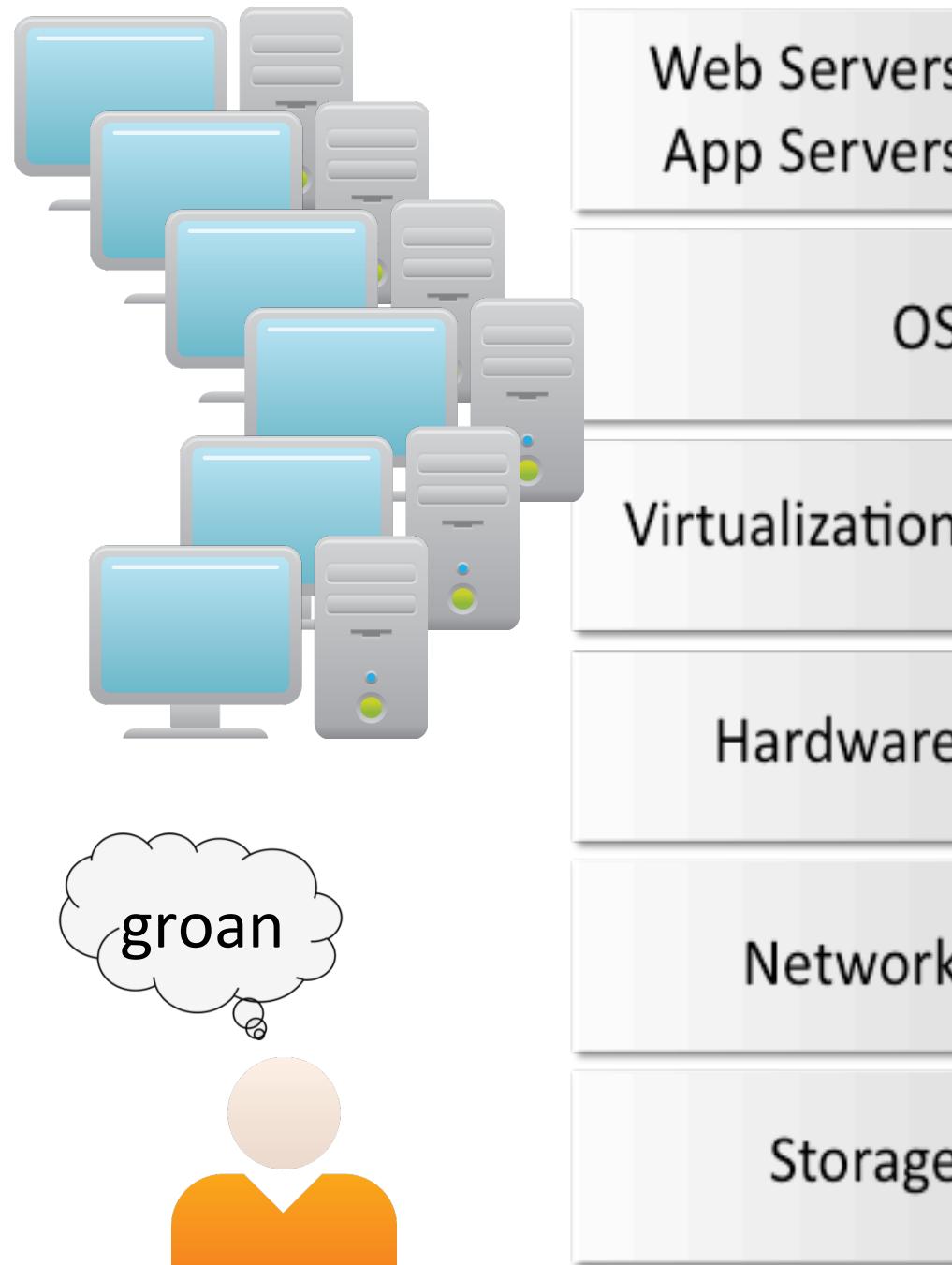
O'Reilly Radar

splunk®>

Machine data has its own “OSI model”

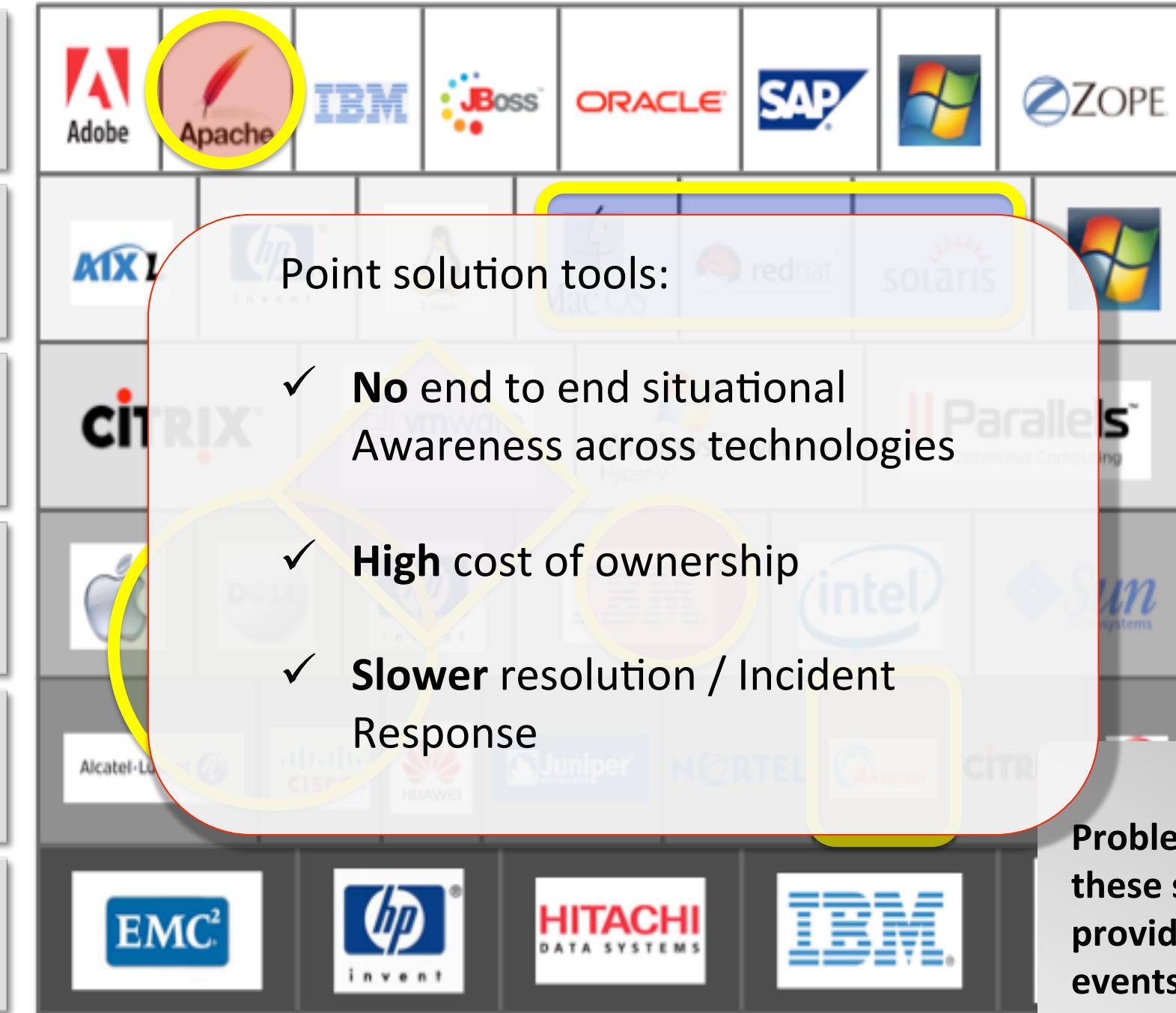
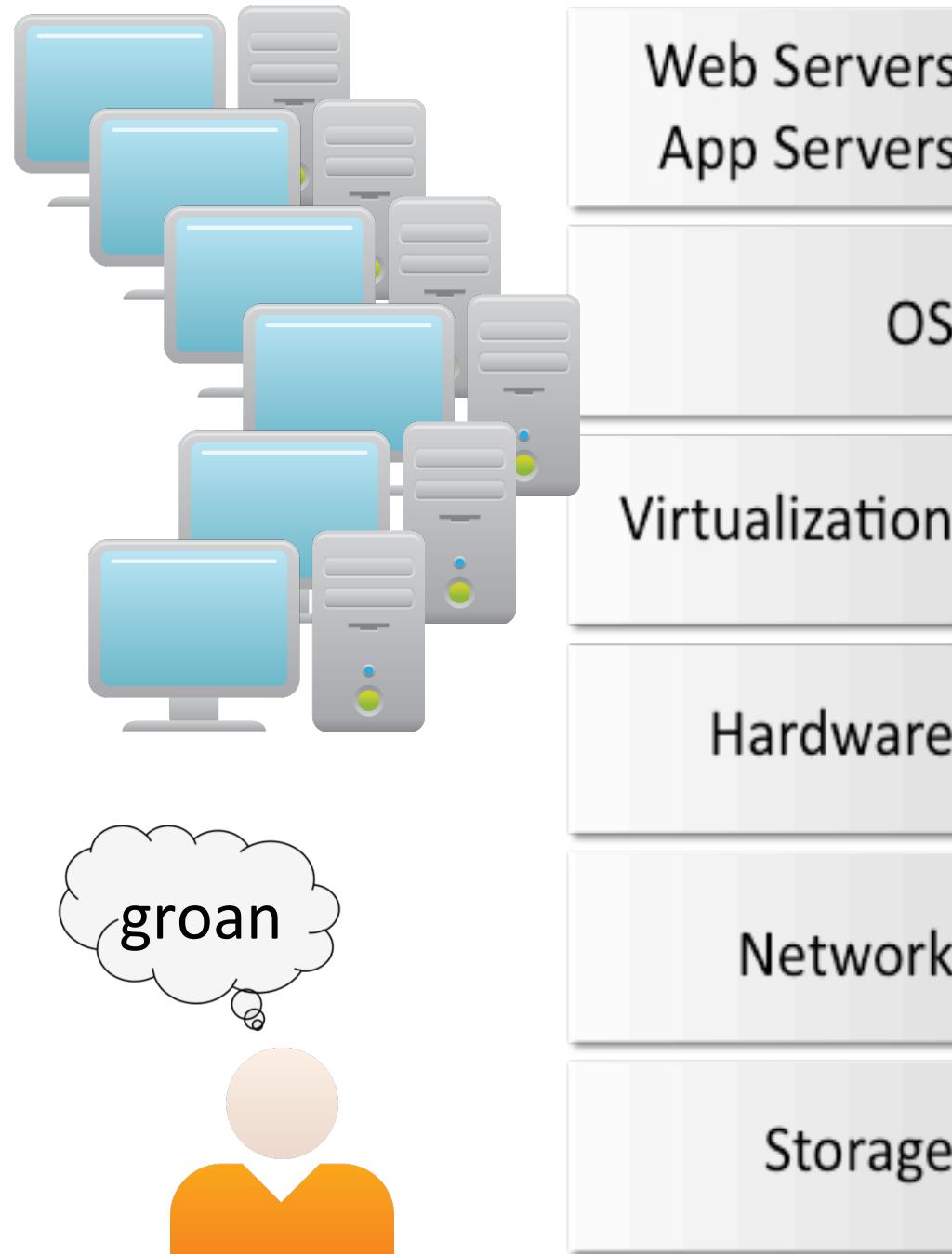


Point Solutions are Common – lots of consoles

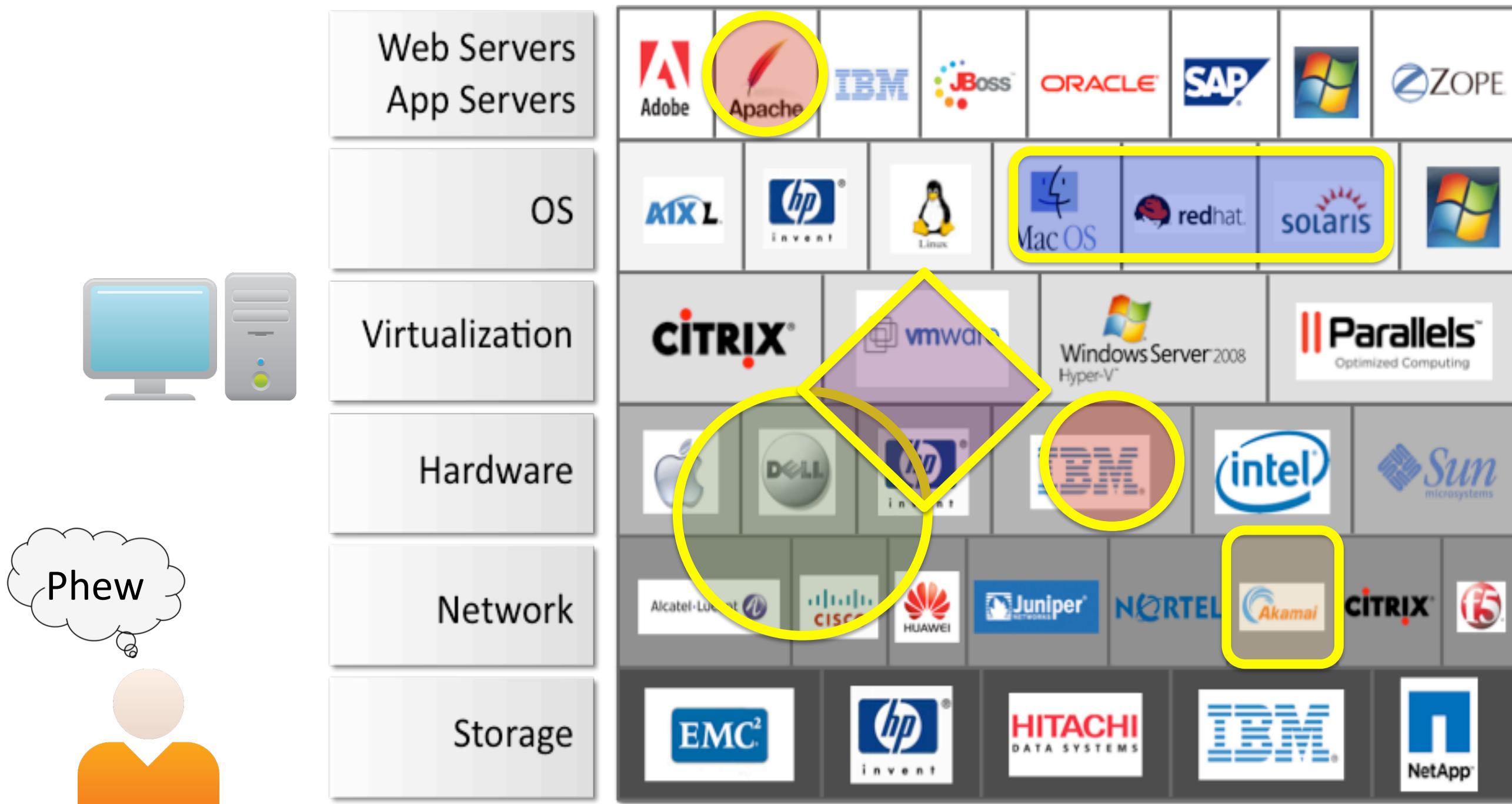


Problem is exasperated since these stovepipe tools do not provide the ability to correlate events between them.

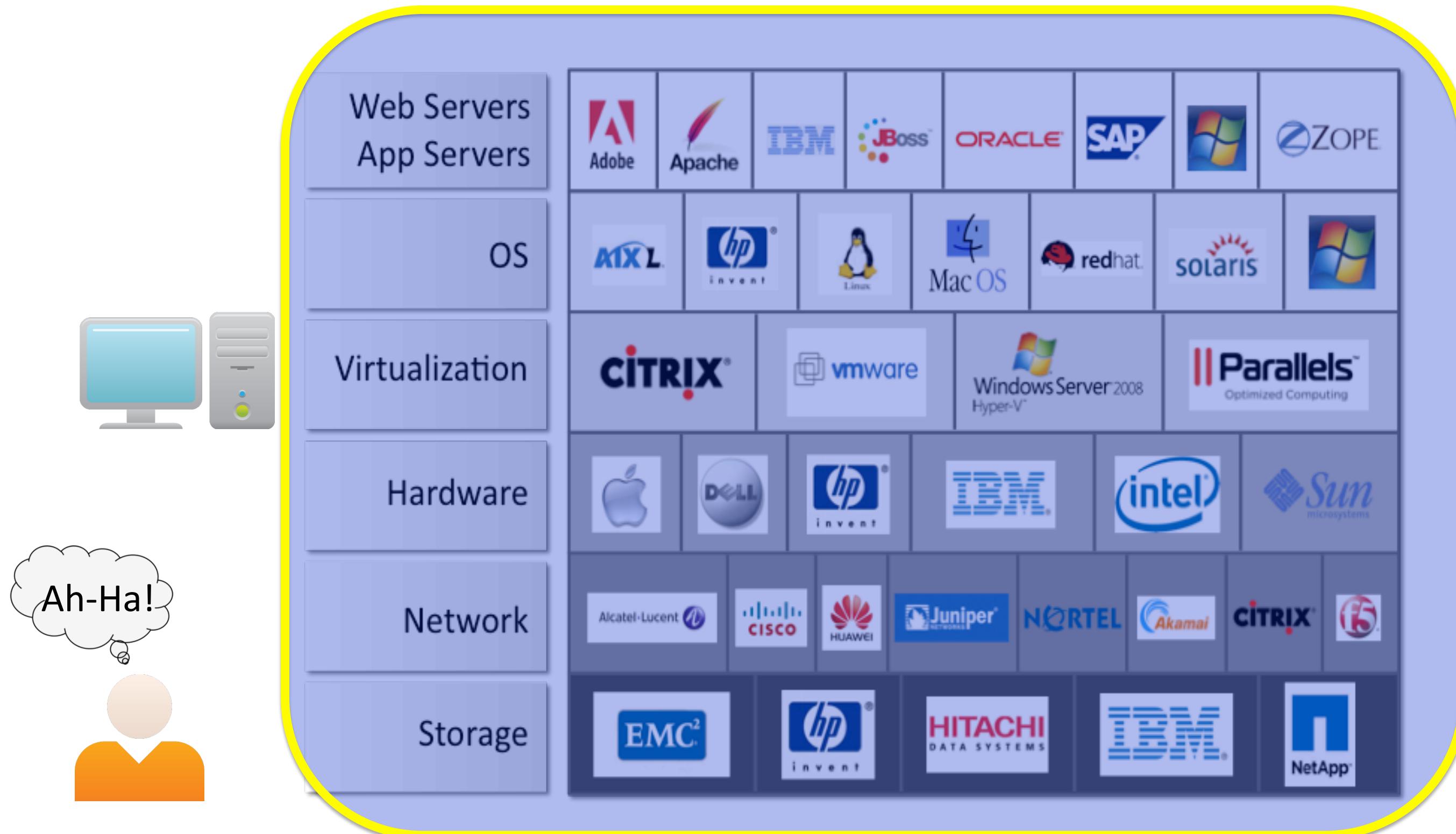
Point Solutions are Common – lots of consoles



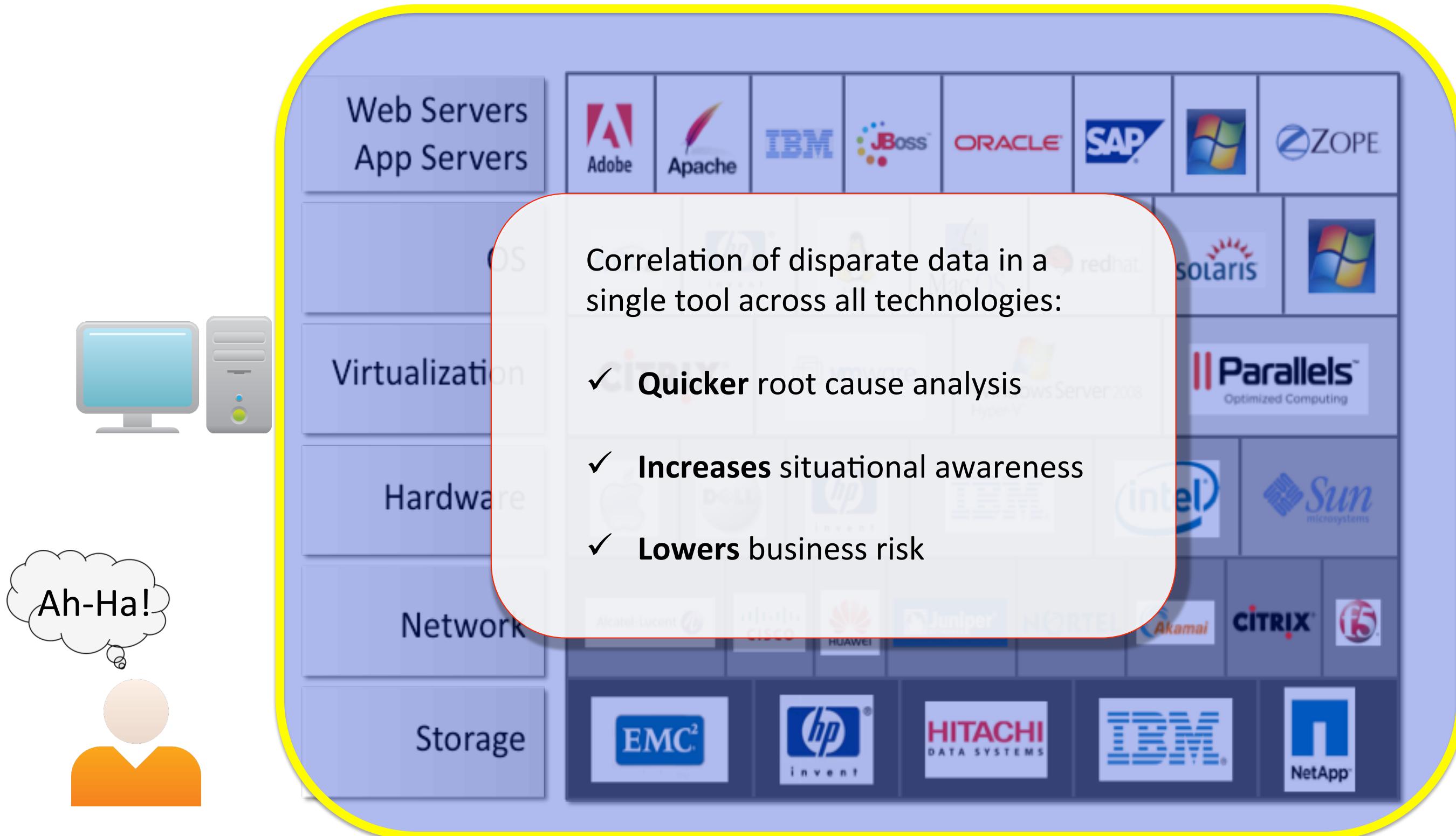
What is needed is a single method to access IT information. . .



Across the entire IT Architecture



Across the entire IT Architecture



See all IT and make IT useful

Finding your faults, just like Mom

Because ninjas are too busy

All batbelt no tights

Needle. Haystack. Found

It's like grep on steroids



Only cavemen use event viewer

Take the SH out of IT

Log management at the speed of thought

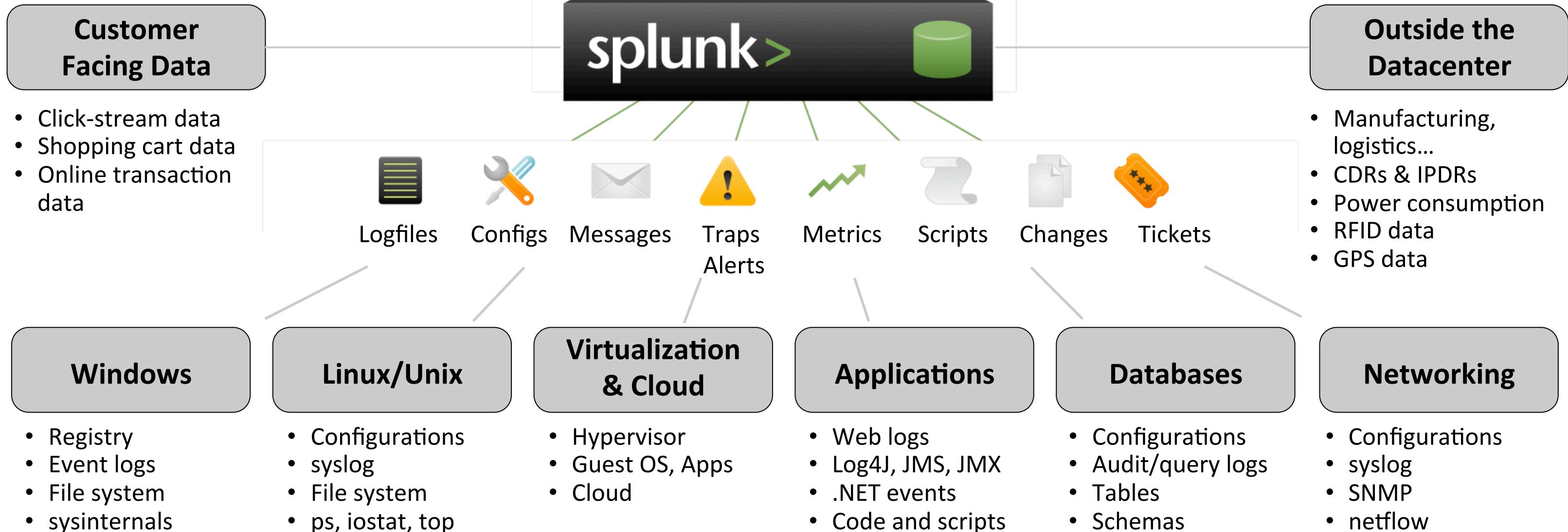
What can you do?



Over 2,900 enterprise customers use Splunk to gain better insight and visibility from their machine data. Why?

splunk®>

Index all data without parsers or connectors



Search using a powerful search language



splunk > Search

Summary Search Status Views Searches & Reports

Search | Actions

```
sourcetype="access_combined" | transaction JSESSIONID | where mvcount(clientip) > 1
```

Your search is paused.

Timeline: Scale: linear log

1 7:32:25 PM
Mon Jul 18
2011 7:32:30 PM 7:32:35 PM 7:32:40 P

50 fields | Pick fields

On Field discovery

Selected fields (3)
host (2)
source (1)
sourcetype (1)

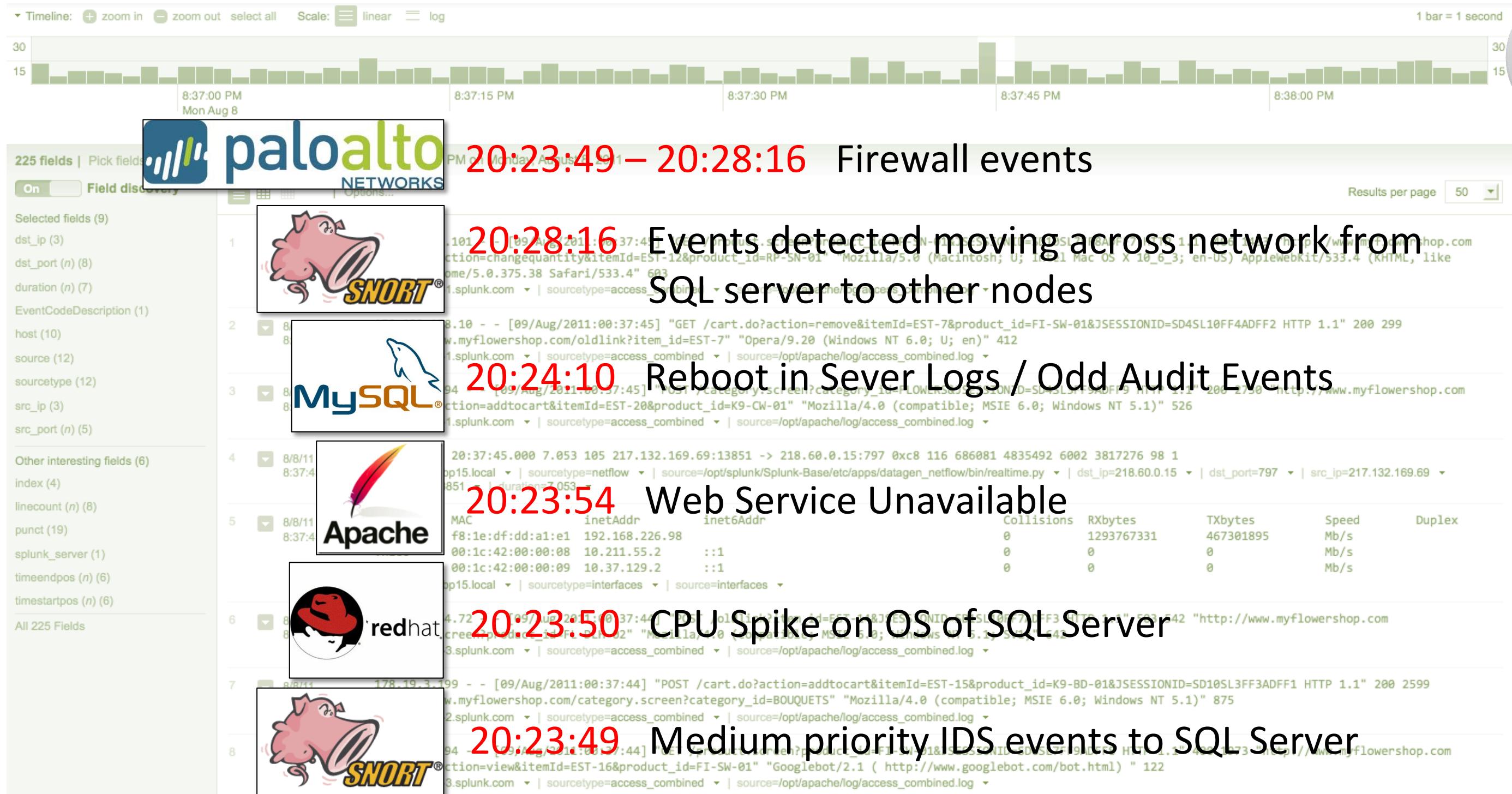
Other interesting fields (39)
action (1)
bytes (2)
call_flwrs_price (n) (1)
category_id (2)

365 results in the last 15 minutes (from 4:16:00 PM to 4:31:58 PM on Monday, August 8, 2011)

Overlay: None

	JSESSIONID	clientip
1	SD6SL3FF3ADFF5	10.103.4.4 233.77.49.94
2	SD2SL10FF6ADFF9	10.2.1.45 66.196.126.101
3	SD4SL1FF7ADFF6	187.231.45.62 192.168.11.2
4	SD2SL9FF7ADFF7	10.2.1.45 63.228.251.81
5	SD6SL1FF8ADFF4	10.2.1.44 10.103.40.24

Automatic Chronology



Alert in Real Time



splunk> Search

Summary Search Status Views Searches & Reports

Create Alert

— ① Save Search — ② Set Up Alert — ③ Define Actions —

Save the search before basing an alert on it.

Search name* SessionID Error

Search string* sourcetype="access_combined" | transaction JSESSIONID | where mvcount(clientip) > 1

Time range in the last 30 seconds (real-time)

Relative time syntax
rt-30s to rt

Time specifiers: y, mon, d, h, m, s [Learn more](#)

Share Keep search private
 Share as read-only to all users of current app
Additional permission settings available in Manager » Searches and Reports

Cancel **Next »**

Create Alert

transaction JSESSIONID | where mvcount(clientip) > 1

— ① Save Search — ② Set Up Alert — ③ Define Actions —

Condition If number of events
is greater than 0

Throttling After triggering the alert, don't trigger it again for 30 second(s)

Expiration After 24 hours

How long Alert manager keeps a record of each triggered alert.

Severity Medium

Cancel **« Back** **Next »**

Create Alert

— ① Save Search — ② Set Up Alert — ③ Define Actions —

Send email Enable

Splunk Alert: \$name\$
security_team@agency.gov

To send email you must set a valid MTA in [Email alert settings](#).

Include search results: as CSV

To send PDF's, [learn more](#) about PDF server.

Add to RSS Enable

RSS link displays after alert is created.

Run a script Enable

SSPLUNK_HOME/bin/scripts/ Script filename

Tracking Show triggered alerts in [Alert manager](#)

Cancel **« Back** **Finish »**

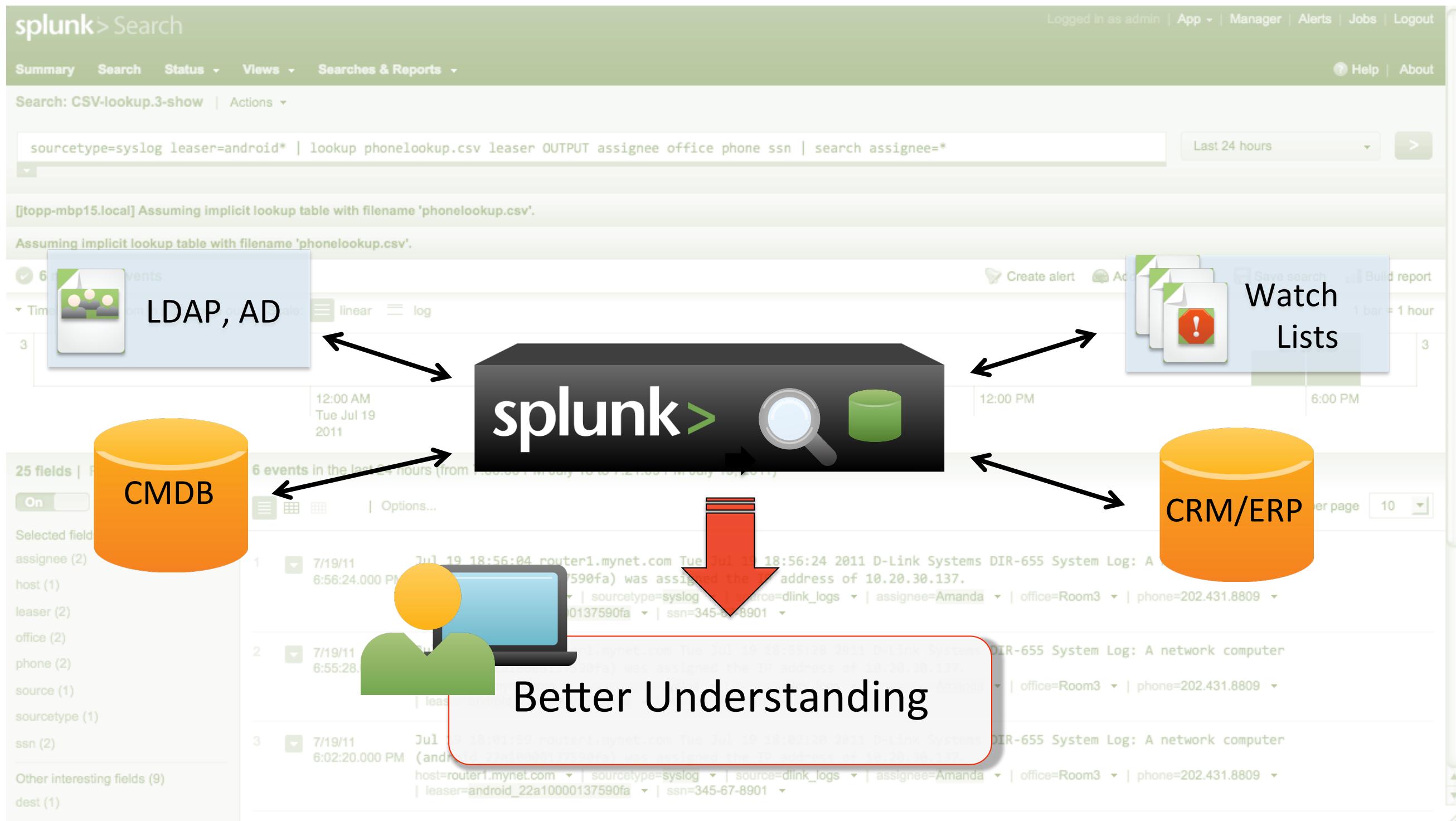
Define the Alert

action (1)
call_flwrs_price (n) (1)
category_id (2)

Alert Options

Who gets Altered

Enrich data from external sources



Report to any level

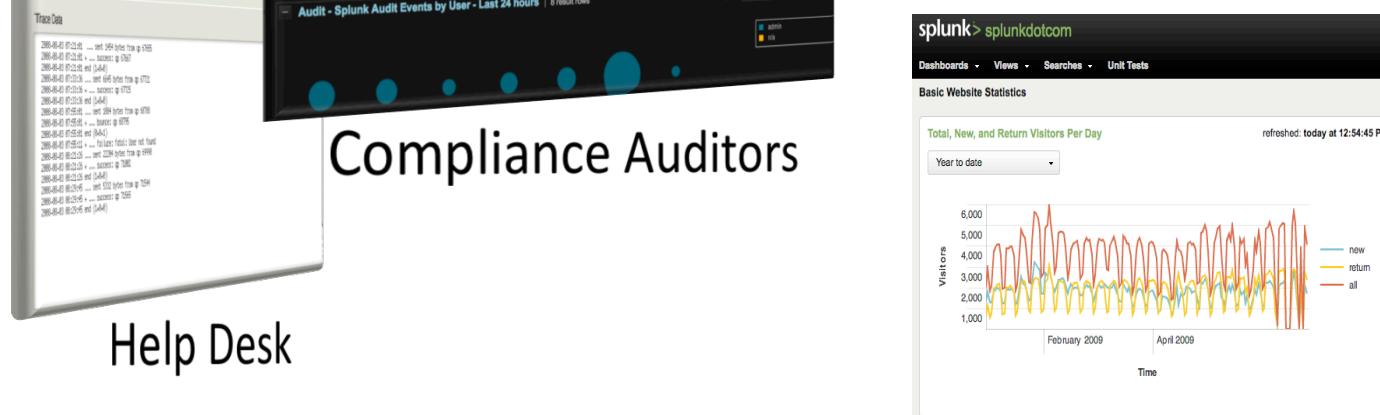


Support Multiple Use Cases

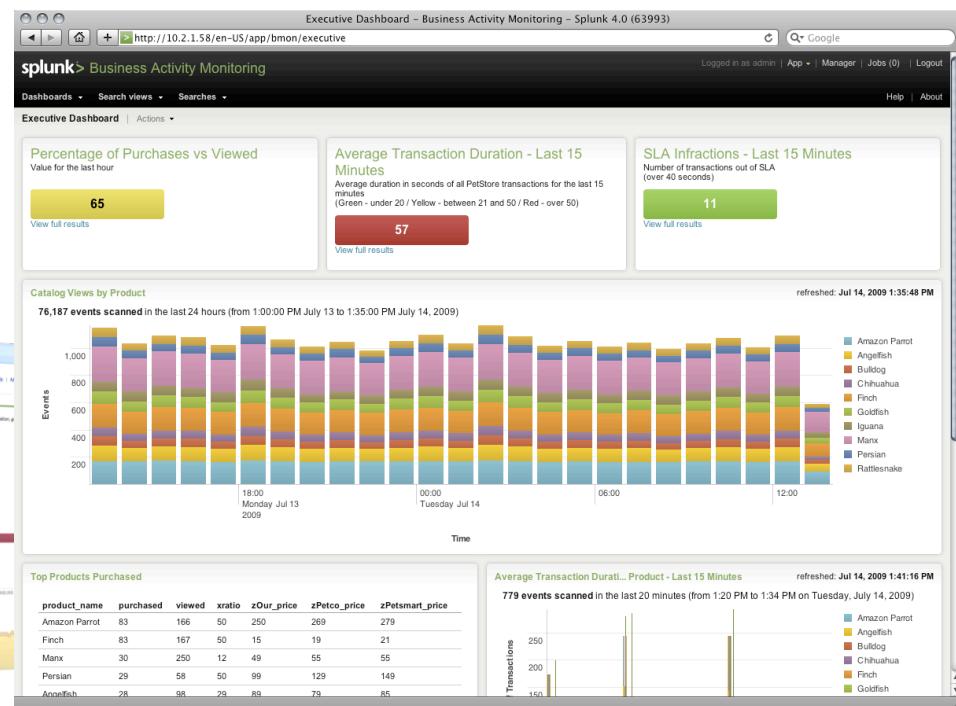


Compliance Auditors

Help Desk

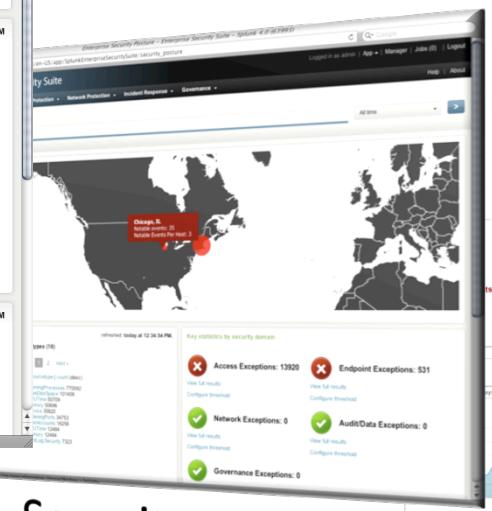


Website Managers



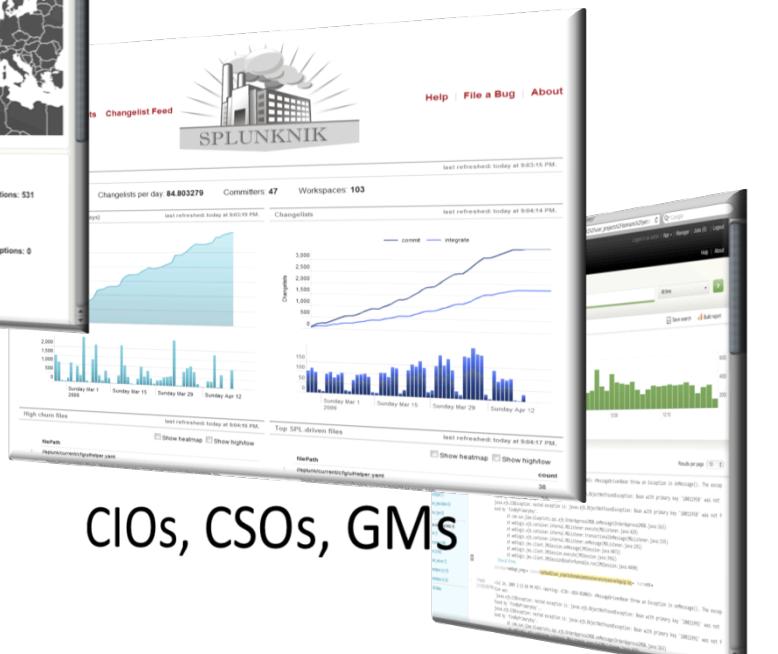
Server Teams

VPs of Infrastructure



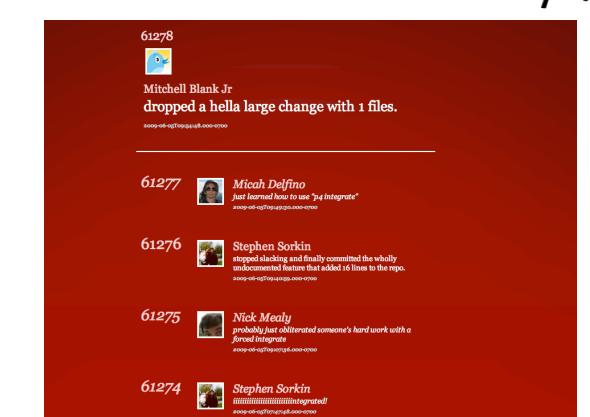
Security Teams

IT, Line of Business or Management



CIOs, CSOs, GMs

SysAdmins,
NW Admins, Developers



Mash up Web Apps

Delivering Operational Intelligence



Three Primary Capabilities

Search/Navigate

- Data drilldown
- “Needle in a haystack”
- Root cause analysis/troubleshooting
- Incident investigations

A screenshot of a Splunk search interface. The search bar at the top contains the text "splunk> error". Below the search bar is a table with columns for "Event Type", "Tags", "Source Types", "Hosts", and "Sources". The table shows 10 events, each with a timestamp, source, and log message. One event is highlighted in yellow. The log messages describe various system errors and events, such as "server1.sendmail[1449]: ... to-root, delay=00:00:00, xlabel=error, pri=35384, dn=5.6.8, stat-data format error" and "server1.sendmail[1449]: ... to-root, delay=00:00:00, xlabel=error, pri=34569, dn=5.6.8, stat-data format error".

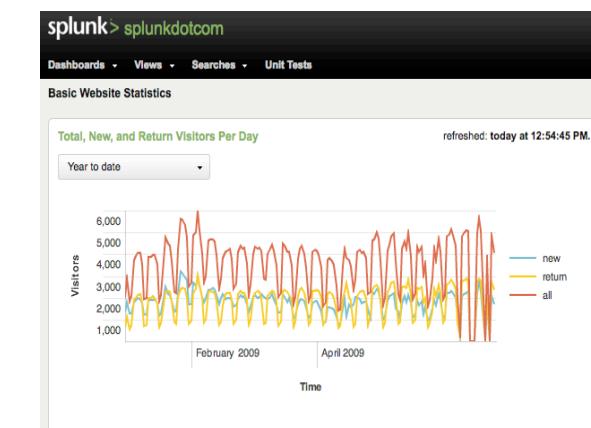
Real-time Visibility

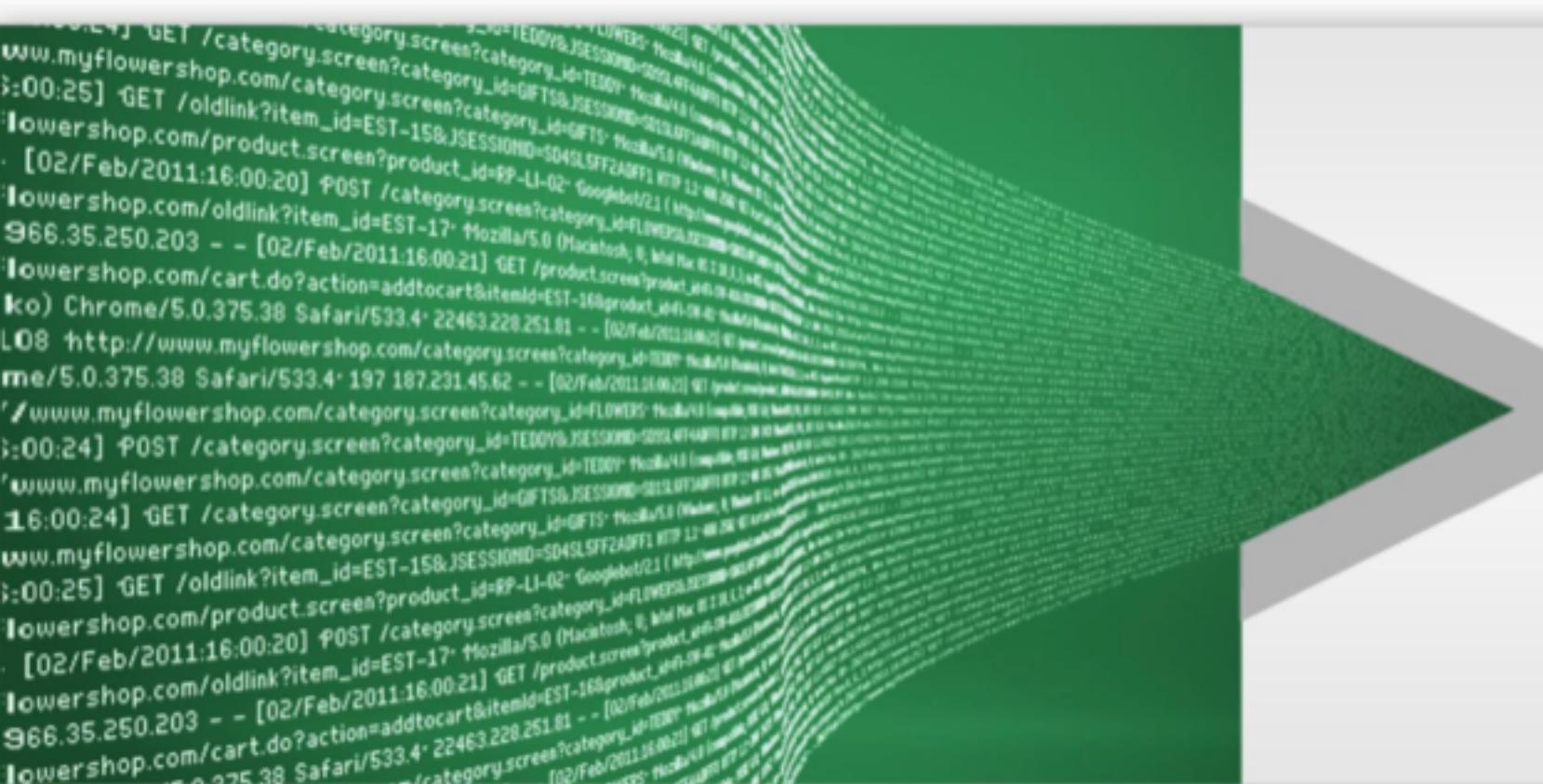
- Live dashboards
- Event correlation
- Monitoring and alerting
- Performance issues
- Transaction levels
- SLA tracking

A screenshot of a real-time visibility dashboard titled "Enterprise Security Posture - Enterprise Security Suite - Splunk 4.0 (63993)". The dashboard features a world map with a red dot over Chicago, IL, labeled "Total Events: 20". Below the map is a table of "Indexed security data" with columns for "Type" (40), "SourceTypes" (18), and "Key statistics by security domain". The table includes rows for "Access Exceptions: 13920", "Endpoint Exceptions: 531", "Network Exceptions: 0", and "Governance Exceptions: 0".

Historical Analytics

- Baseline and thresholds
- Trending
- Operational insights
- Historical patterns
- Compliance reports





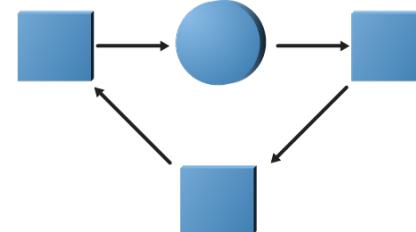
Why Splunk scales

“Splunk has been tackling [big data] with a unique solution that is generating a significant amount of commercial success”

David Menninger
VP & Research Director

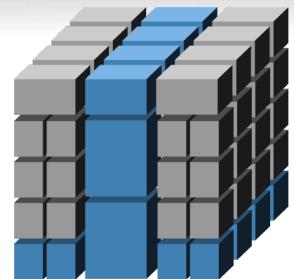


Databases are not suited for unstructured data



Relational Databases

- Financial records, manufacturing and logistical information, personnel data
- Data highly structured – database highly structured
- Inflexible schema, long deployment cycle



Multidimensional Databases

- Multidimensional data for business management and statistics
- Math computation strength – dense data
- Pivots data for flexible financial analysis
- Monthly reporting, not for real-time events



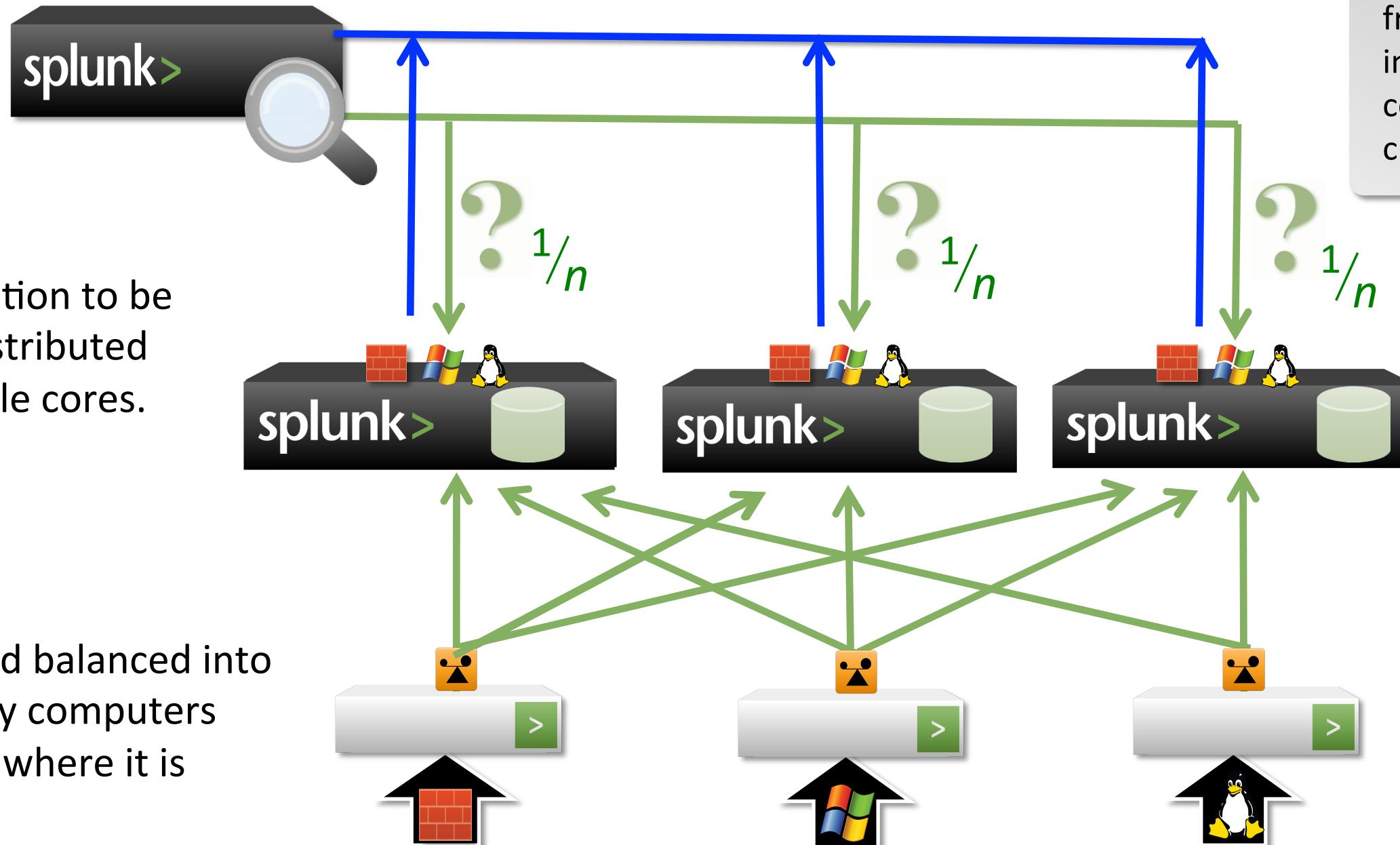
Machine Data Engine

- Time series unstructured data, with no predefined schema
- Generated by all IT systems, non-standard data, unpredictable formats
- Massive volume; fast navigation and correlation paramount

Distributed Search using Map Reduce



A ‘search’ (question to be answered) is distributed amongst multiple cores.

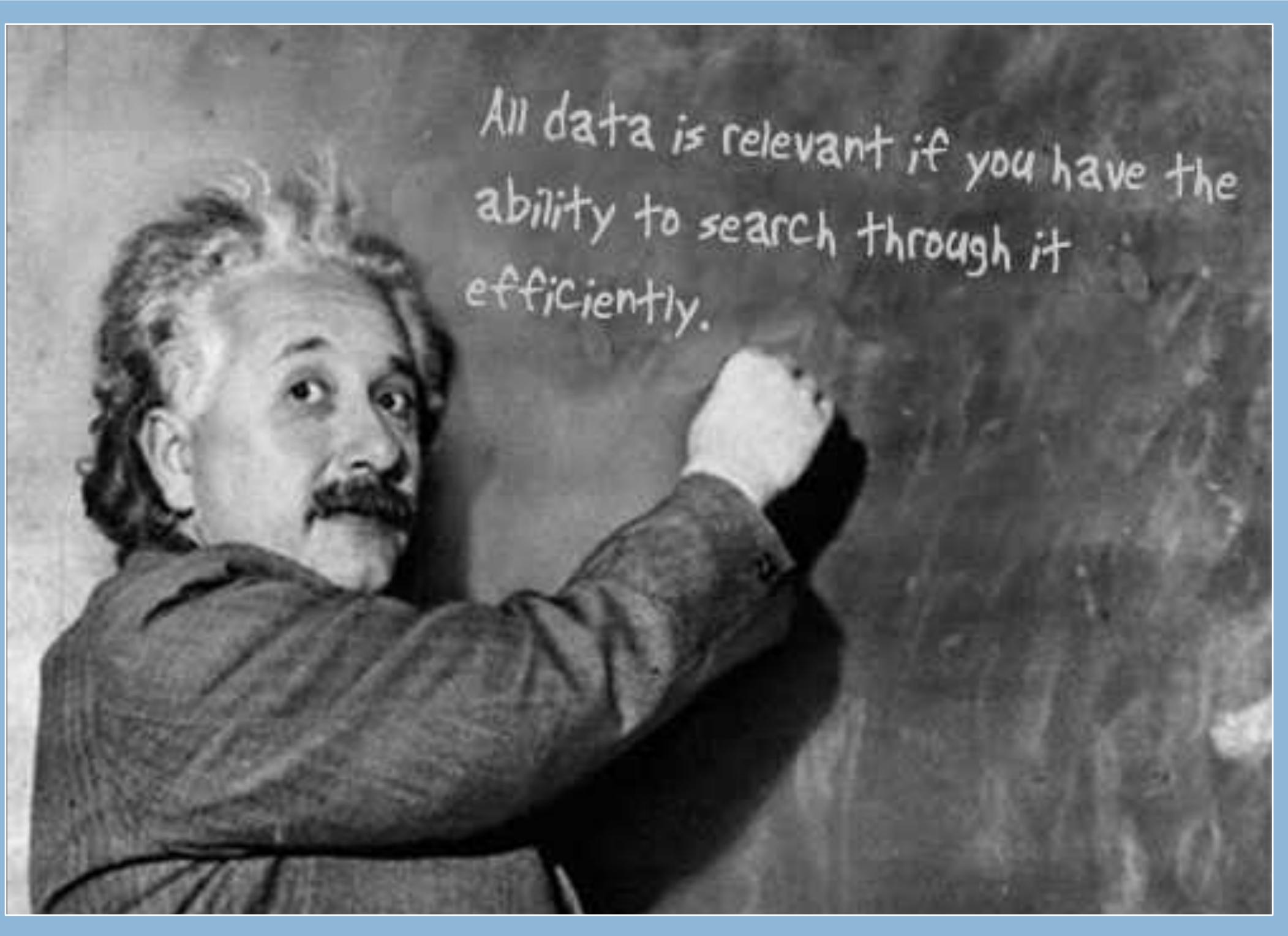


MapReduce is a software framework introduced by Google in 2004 to support distributed computing on large data sets on clusters of computers. - Wikipedia

Each Indexer processes a subset of the entire dataset and produces part of the overall answer back to the search head for “reduce”

Data is load balanced into commodity computers (indexers) where it is ‘mapped’.

Questions? Talk to a Splunk representative



Library of Congress

Anna Tant

Civilian Account Executive

Federal

atant@splunk.com

Free Download

Limited to 500mb/day

No alerting

www.splunk.com