splunk>

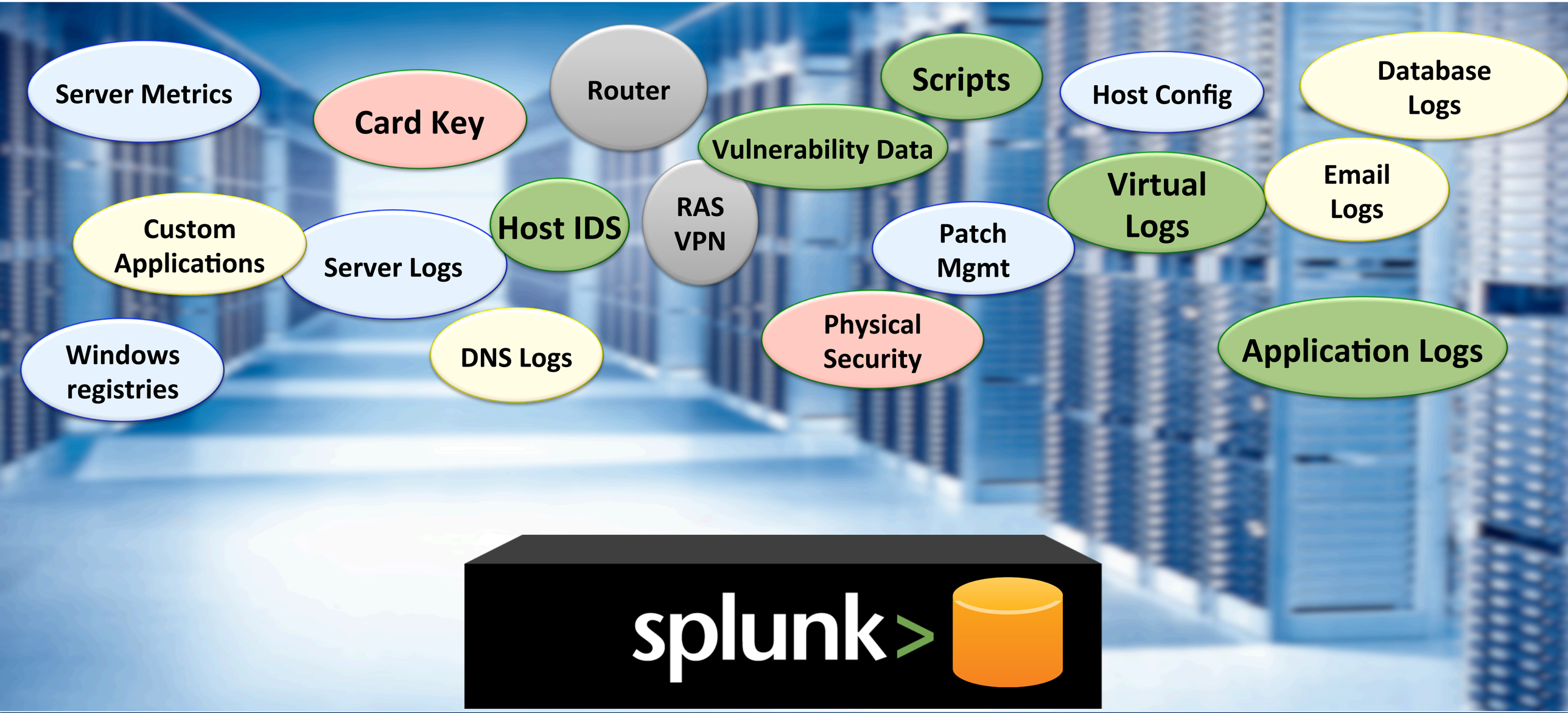
Overview

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?

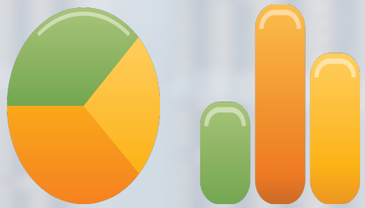


What does Splunk Provide?

A common interface for all IT Data

search

alert



report

share

That provides Operational Intelligence

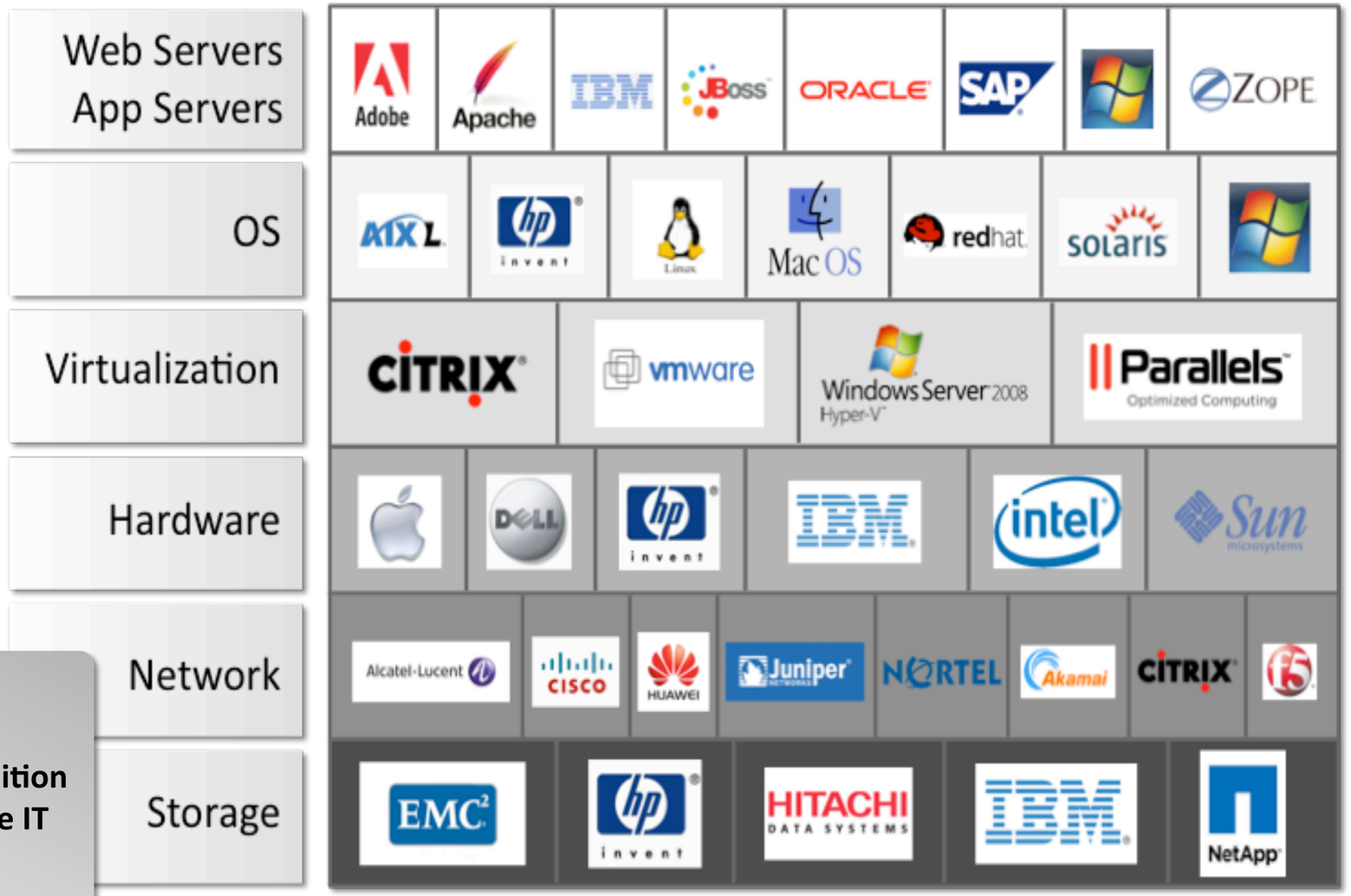
What is Big Data?

“…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”



Complexity has increased logarithmically with the addition of each of these layers to the IT architecture.

Point Solutions are Common – lots of consoles

Web Servers
App Servers

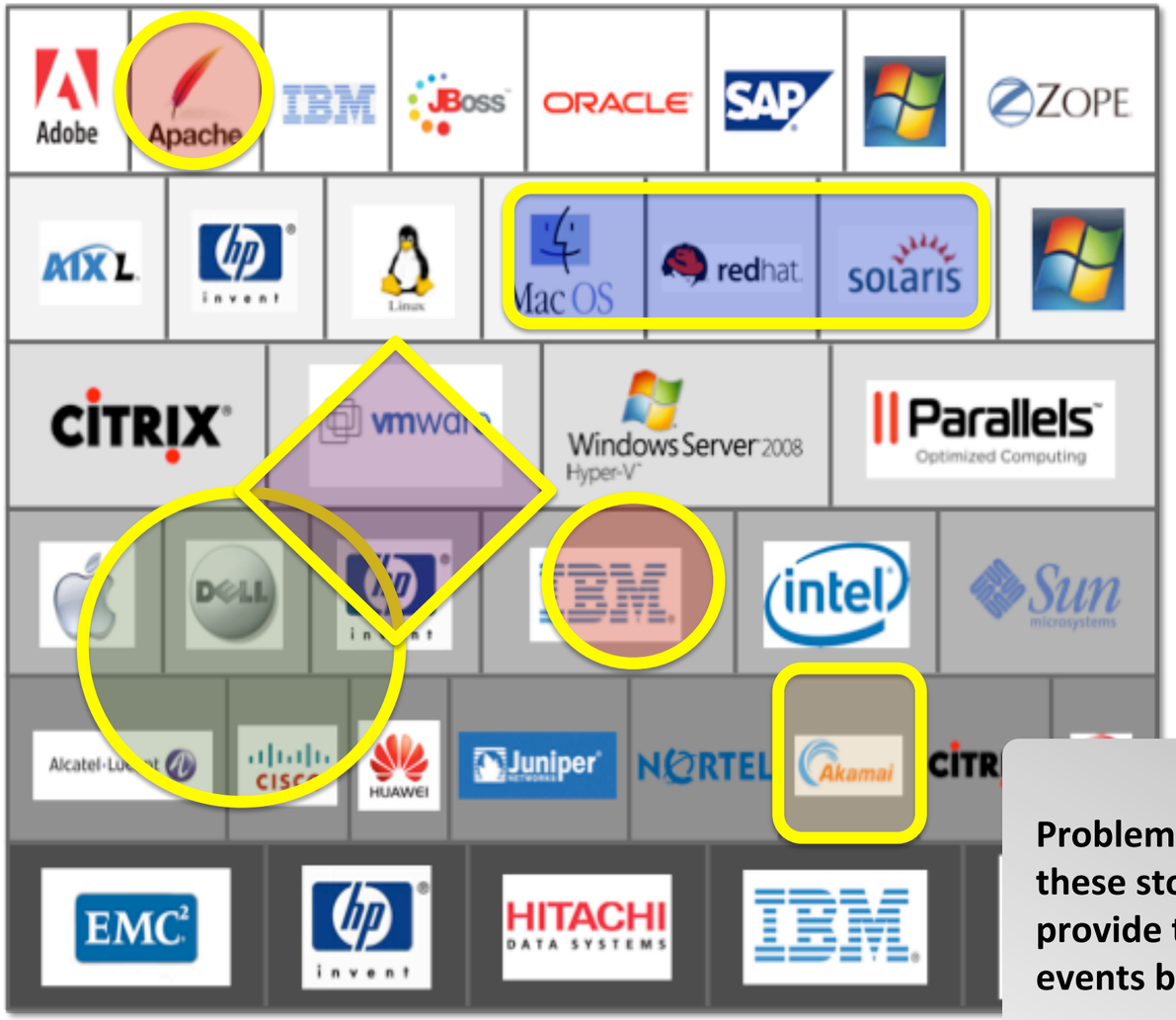
OS

Virtualization

Hardware

Network

Storage



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

Point Solutions are Common – lots of consoles

Web Servers
App Servers

OS

Virtualization

Hardware

Network

Storage

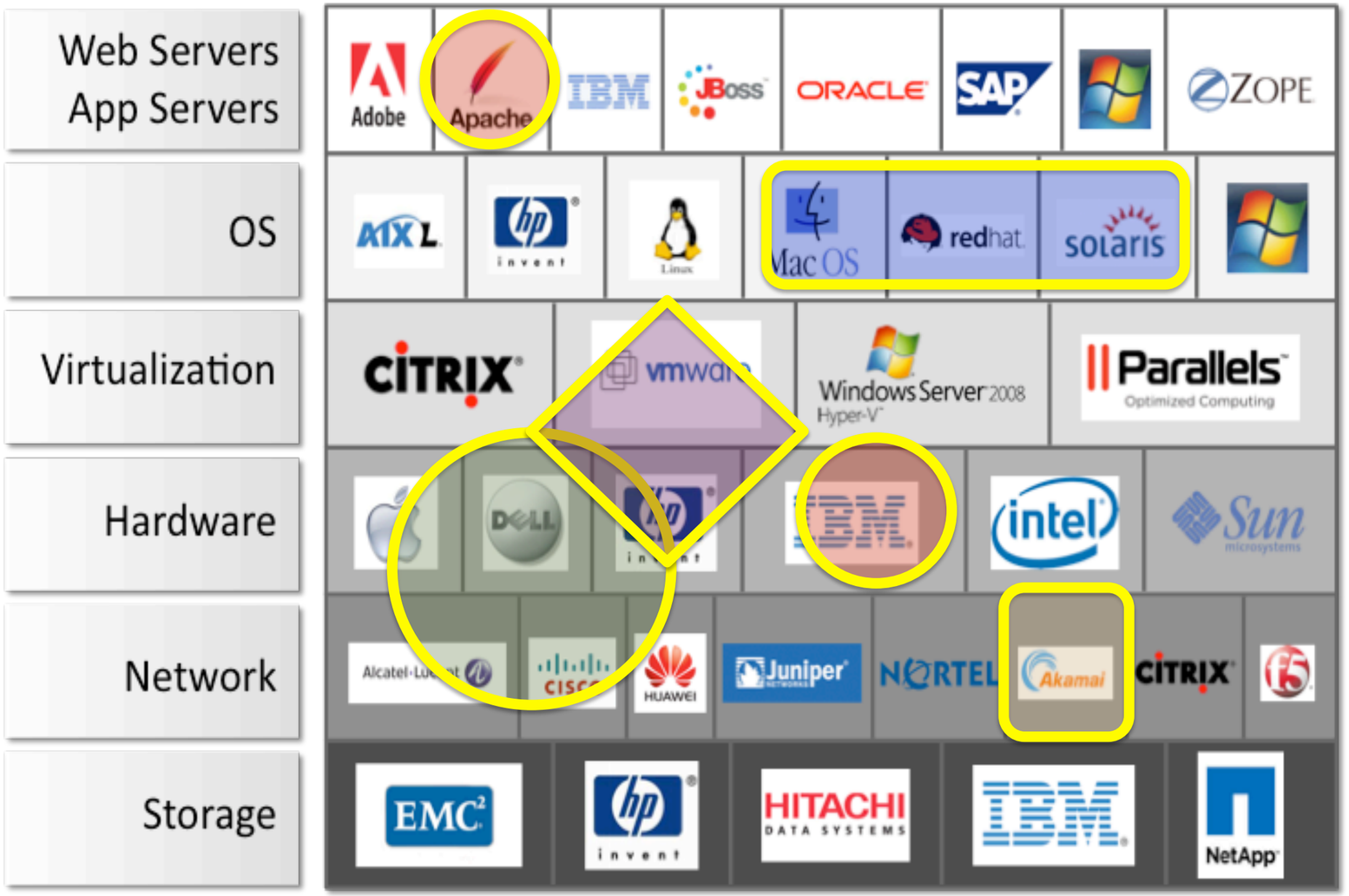
Point solution tools:

- ✓ **No** end to end situational Awareness across technologies
- ✓ **High** cost of ownership
- ✓ **Slower** resolution / Incident Response

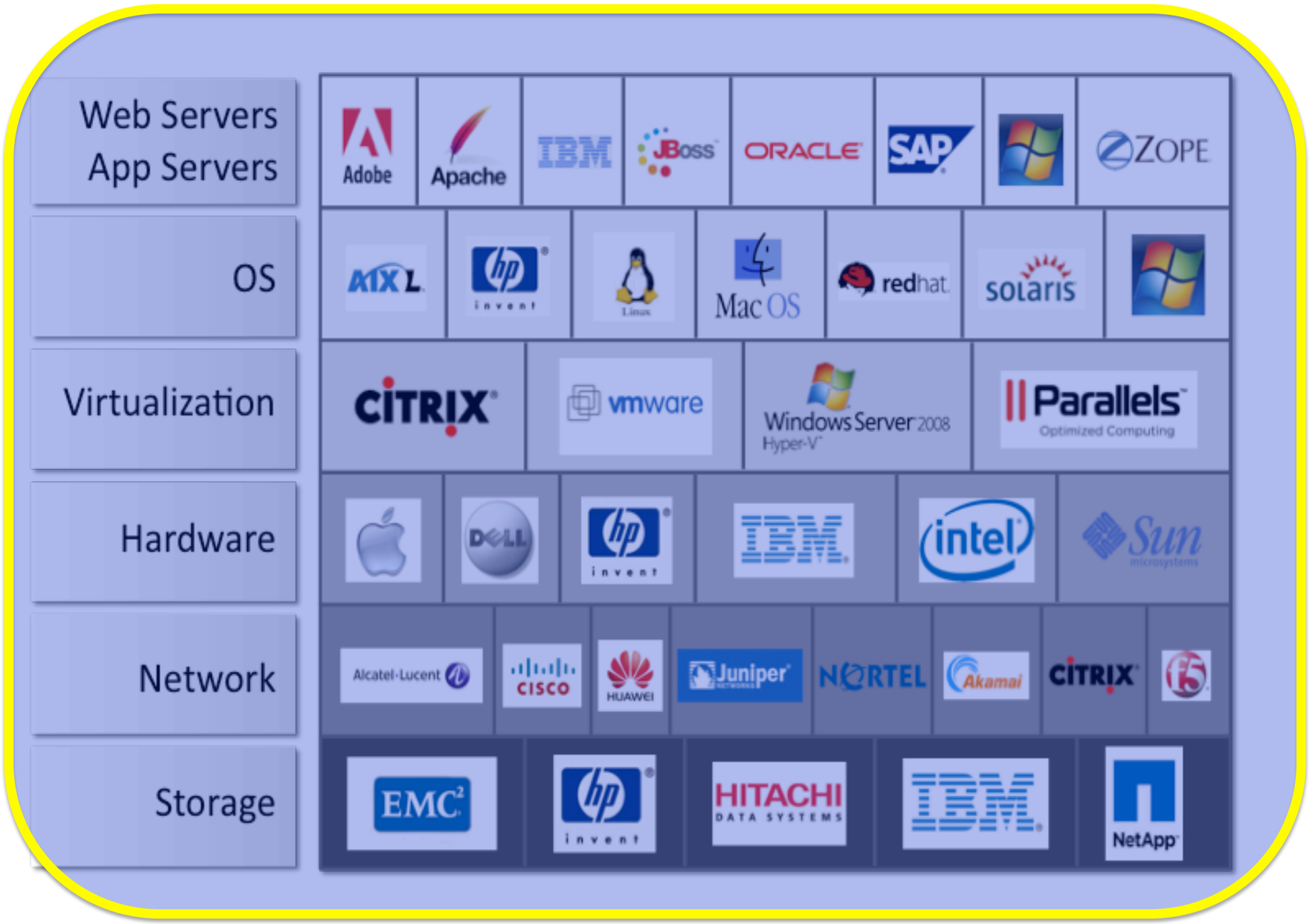
Problem is exasperated since these stovepipe tools do not provide the ability to correlate events between them.

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

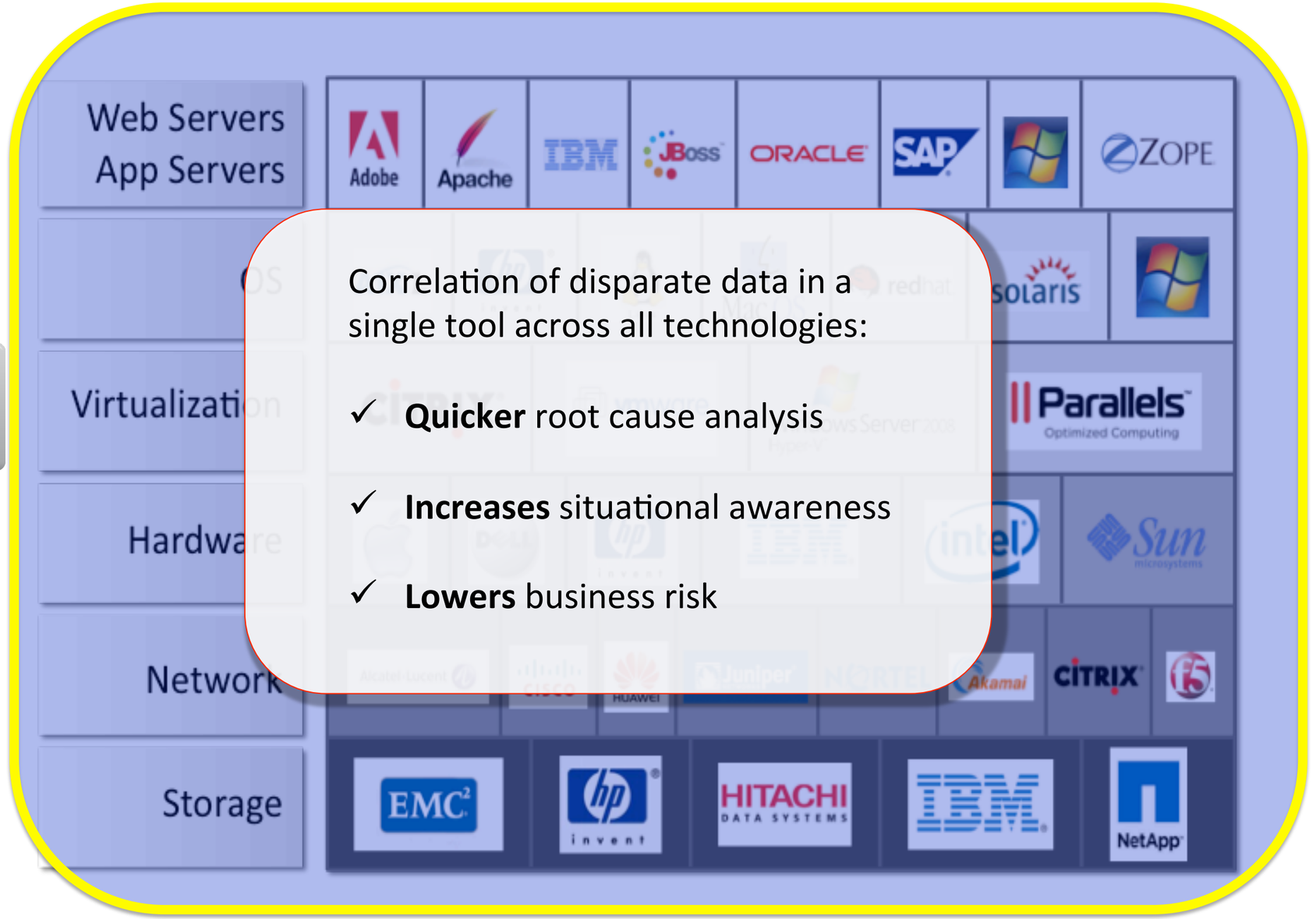
Phew



Across the entire IT Architecture



Across the entire IT Architecture



Correlation of disparate data in a single tool across all technologies:

- ✓ **Quicker** root cause analysis
- ✓ **Increases** situational awareness
- ✓ **Lowers** business risk

See all IT and make IT useful

Finding your faults, just like Mom

Because ninjas are too busy

All batbelt no tights

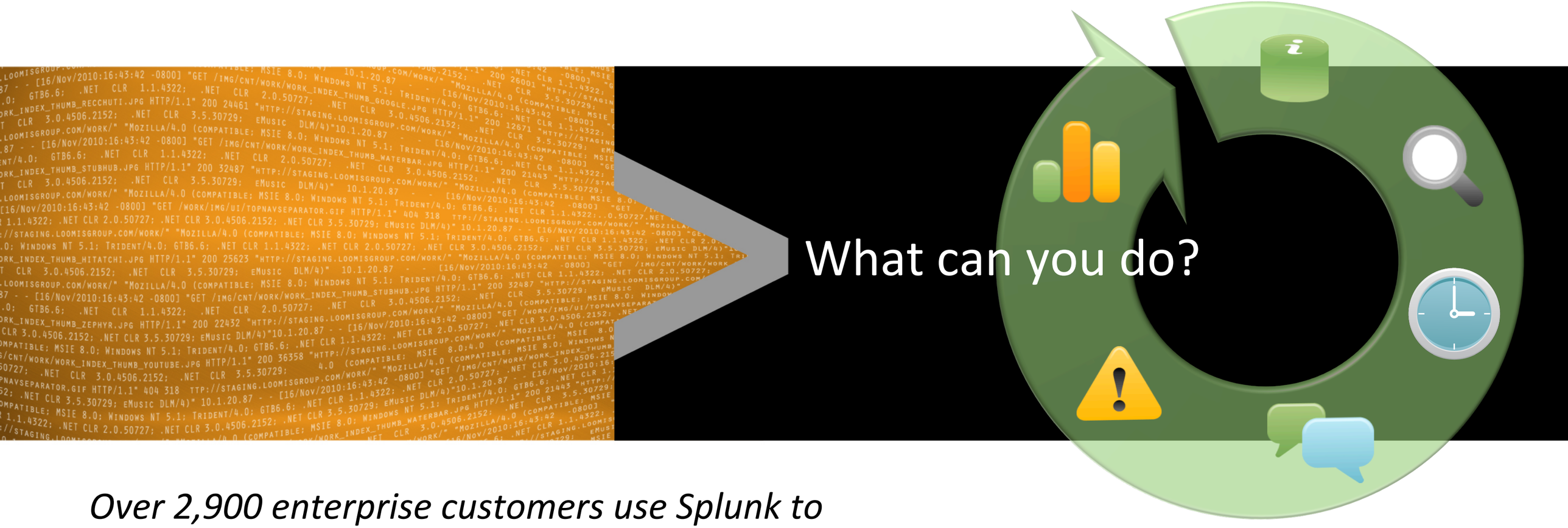
Needle. Haystack. Found

It's like grep on steroids

Only cavemen use event viewer

Take the SH out of IT

Log management at the speed of thought



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

Customer Facing Data

- Click-stream data
- Shopping cart data
- Online transaction data

splunk>

Outside the Datacenter

- Manufacturing, logistics…
- CDRs & IPDRs
- Power consumption
- RFID data
- GPS data

Logfiles | Configs | Messages | Traps Alerts | Metrics | Scripts | Changes | Tickets

Windows

- Registry
- Event logs
- File system
- sysinternals

Linux/Unix

- Configurations
- syslog
- File system
- ps, iostat, top

Virtualization & Cloud

- Hypervisor
- Guest OS, Apps
- Cloud

Applications

- Web logs
- Log4J, JMS, JMX
- .NET events
- Code and scripts

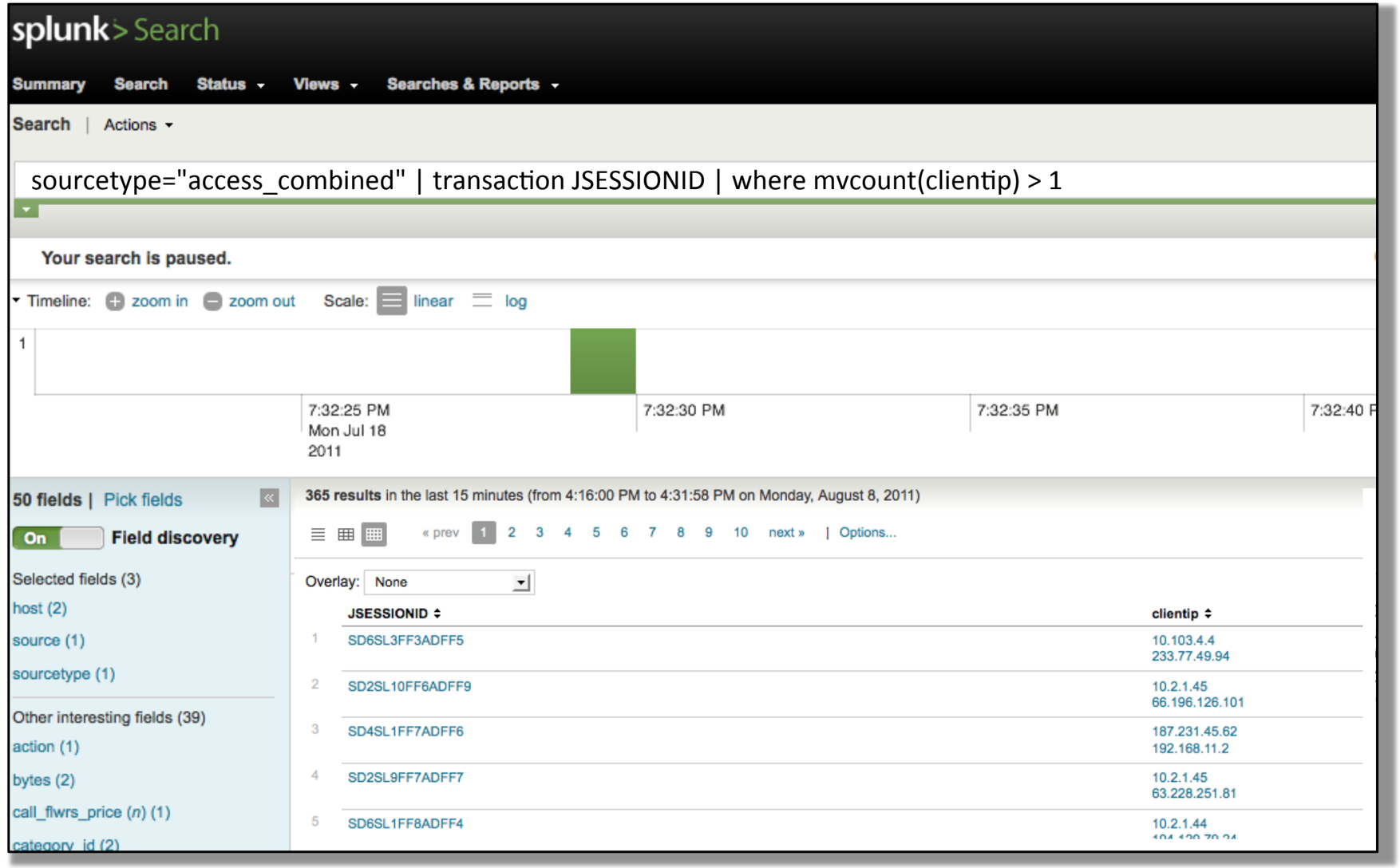
Databases

- Configurations
- Audit/query logs
- Tables
- Schemas

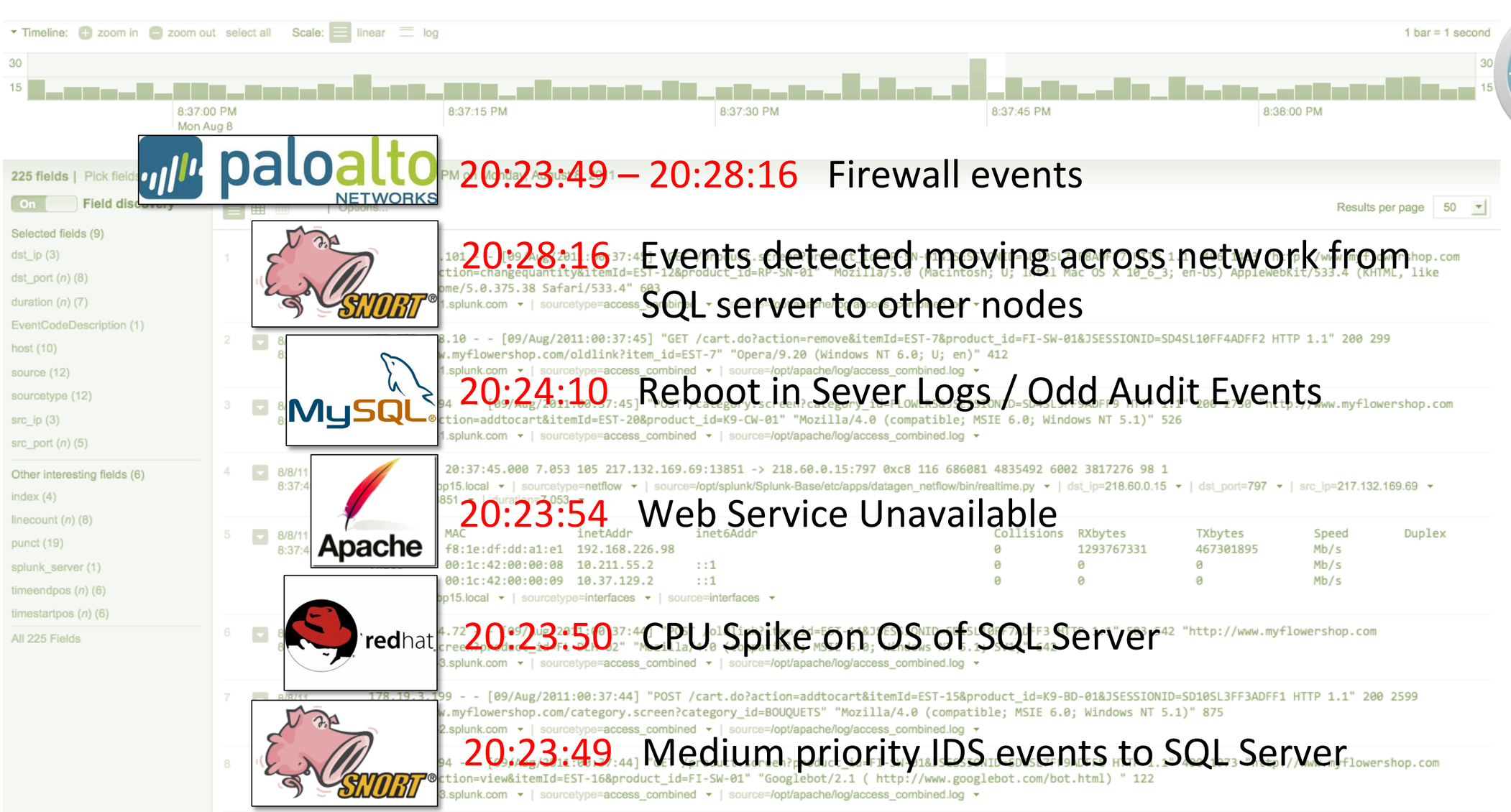
Networking

- Configurations
- syslog
- SNMP
- netflow

Search using a powerful search language



Automatic Chronology



paloalto NETWORKS

20:23:49 – 20:28:16 Firewall events

SNORT®

20:28:16 Events detected moving across network from SQL server to other nodes

MySQL®

20:24:10 Reboot in Sever Logs / Odd Audit Events

Apache

20:23:54 Web Service Unavailable

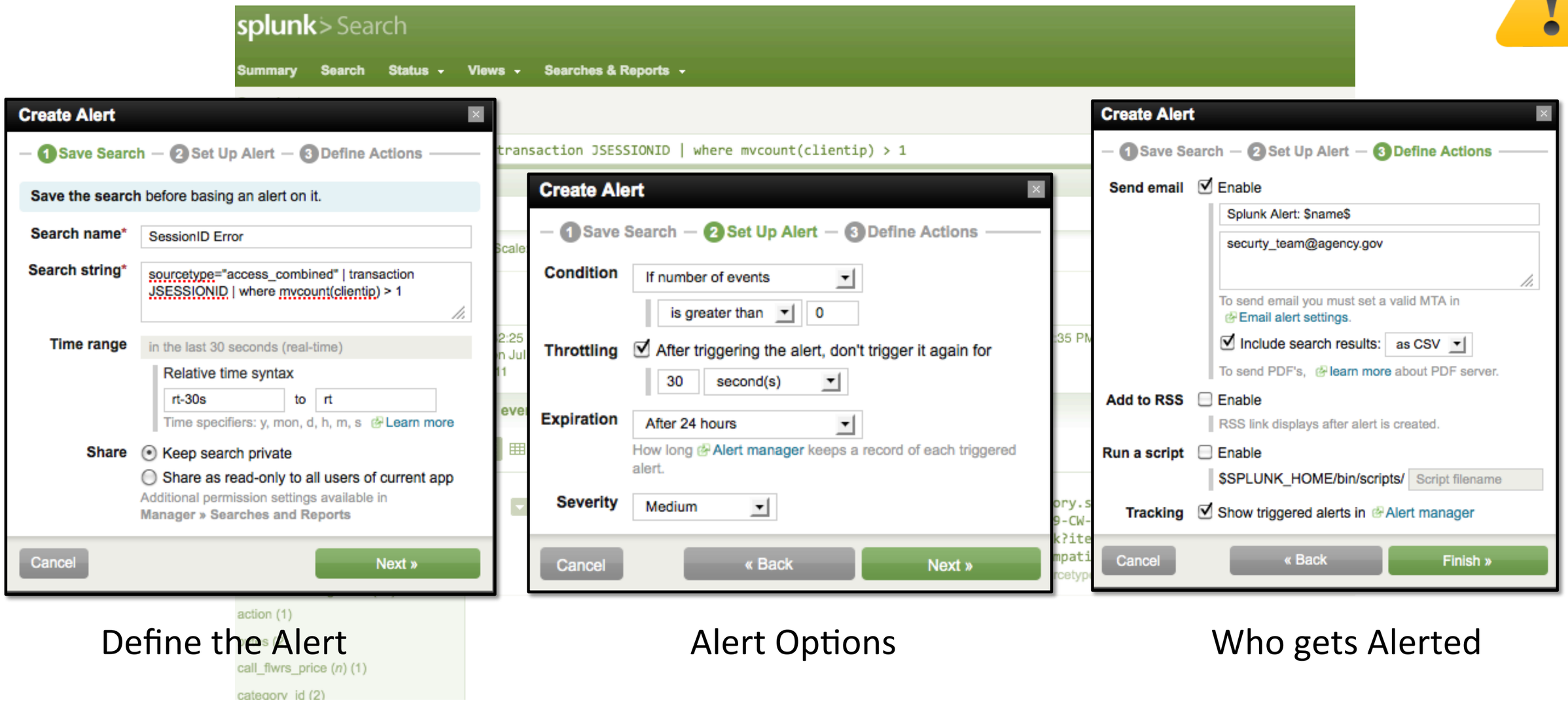
redhat

20:23:50 CPU Spike on OS of SQL Server

SNORT®

20:23:49 Medium priority IDS events to SQL Server

Alert in Real Time

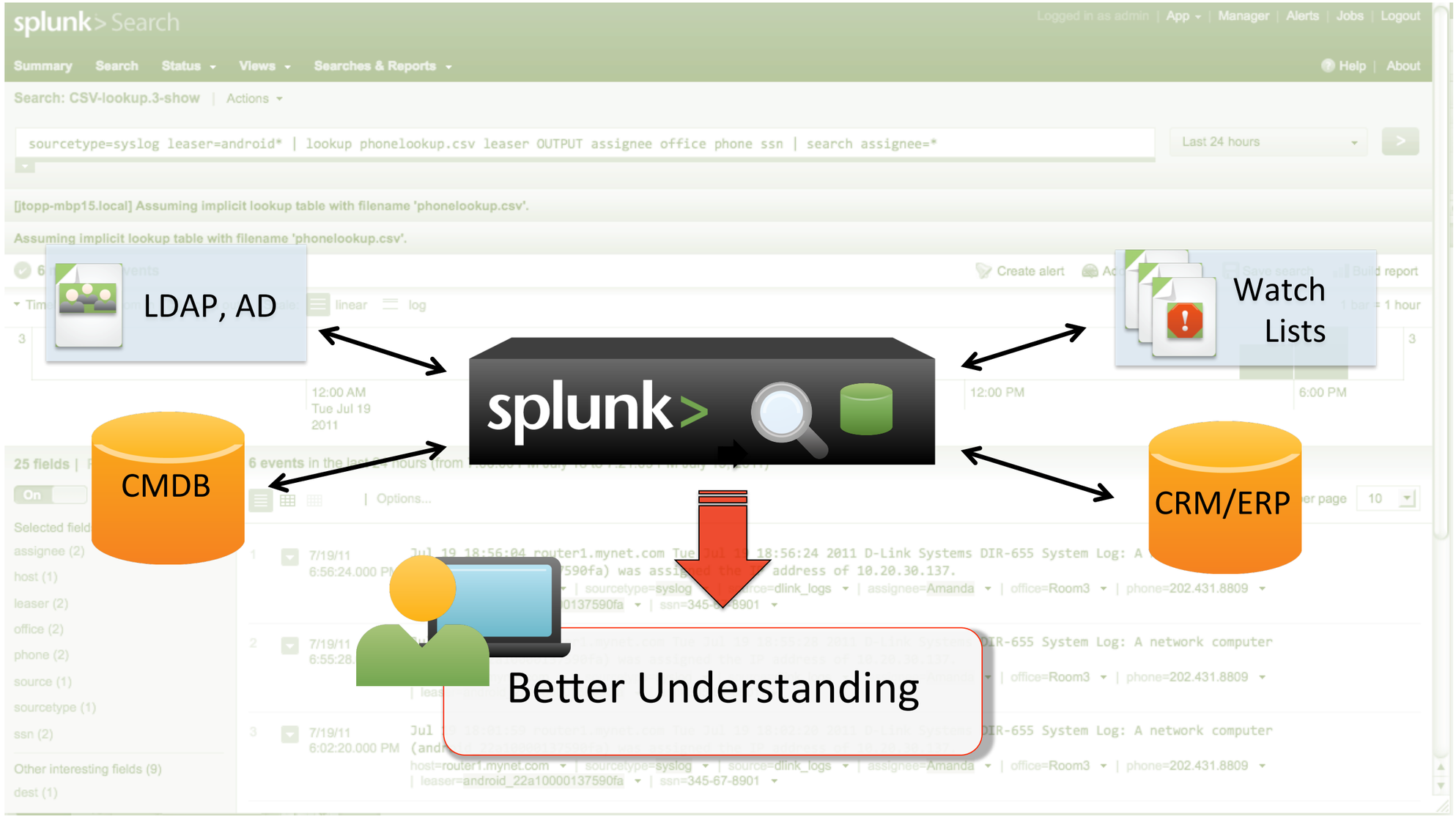


Define the Alert

Alert Options

Who gets Alerted

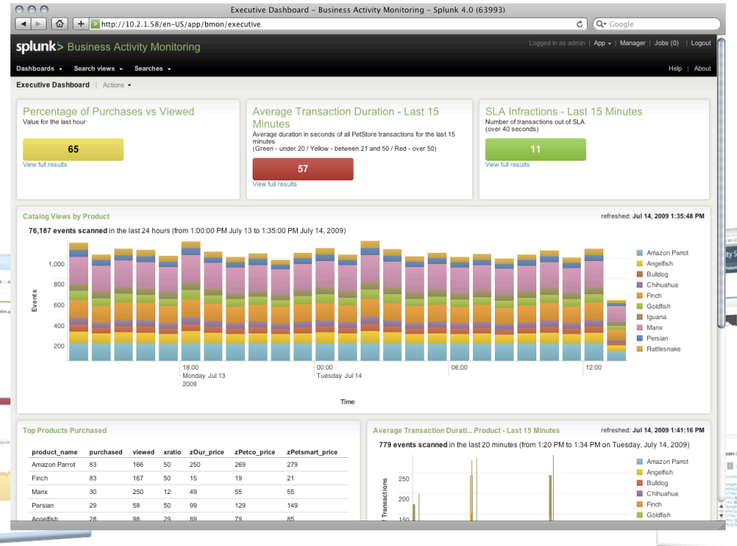
Enrich data from external sources



Report to any level

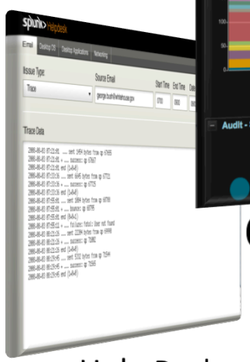
Support Multiple Use Cases

IT, Line of Business or Management



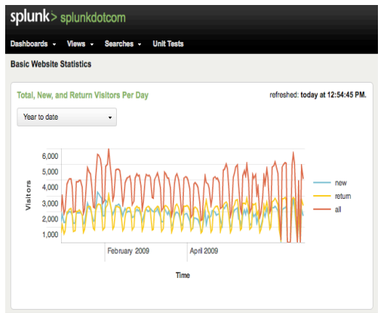
VPs of Infrastructure

Server Teams



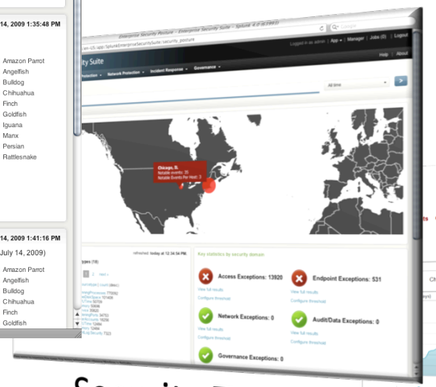
Help Desk

Compliance Auditors



Website Managers

Mash up Web Apps



Security Teams

CIOs, CSOs, GMs

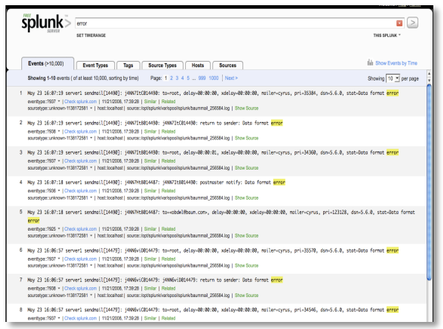
SysAdmins,
NW Admins, Developers

Delivering Operational Intelligence

Three Primary Capabilities

Search/Navigate

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

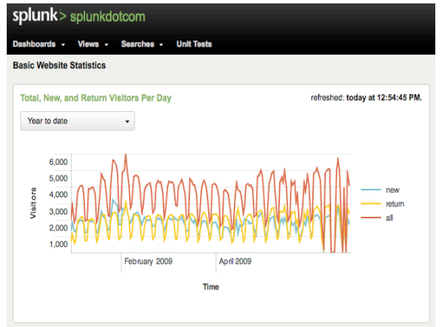


Real-time Visibility

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

Historical Analytics

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



Why Splunk scales

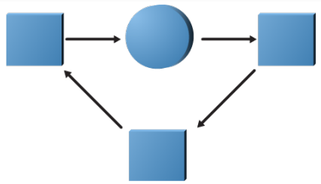
"Splunk has been tackling [big data] with a unique solution that is generating a significant amount of commercial success"

David Menninger
VP & Research Director

VENTANA RESEARCH

splunk>

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

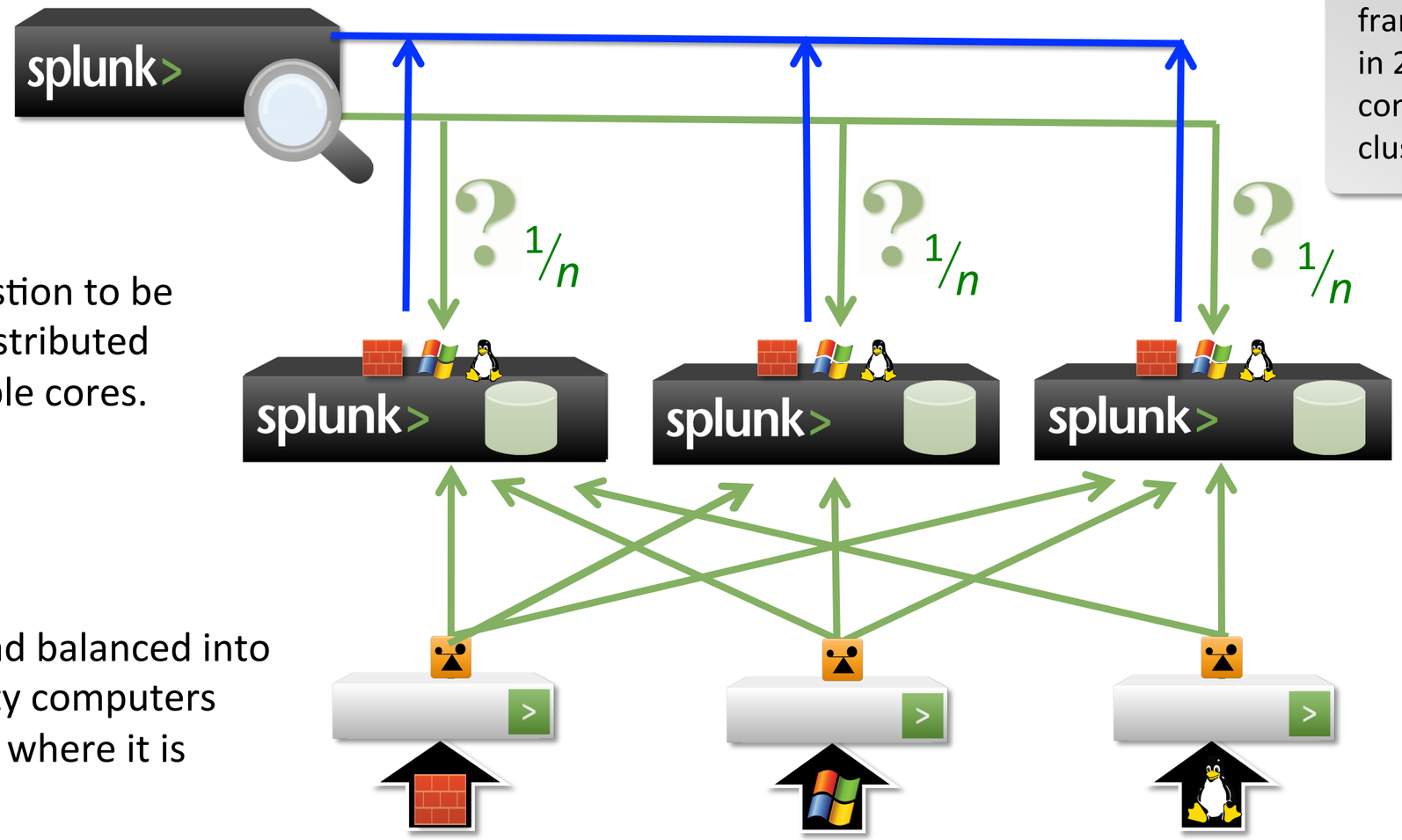
Machine Data Engine

- 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

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

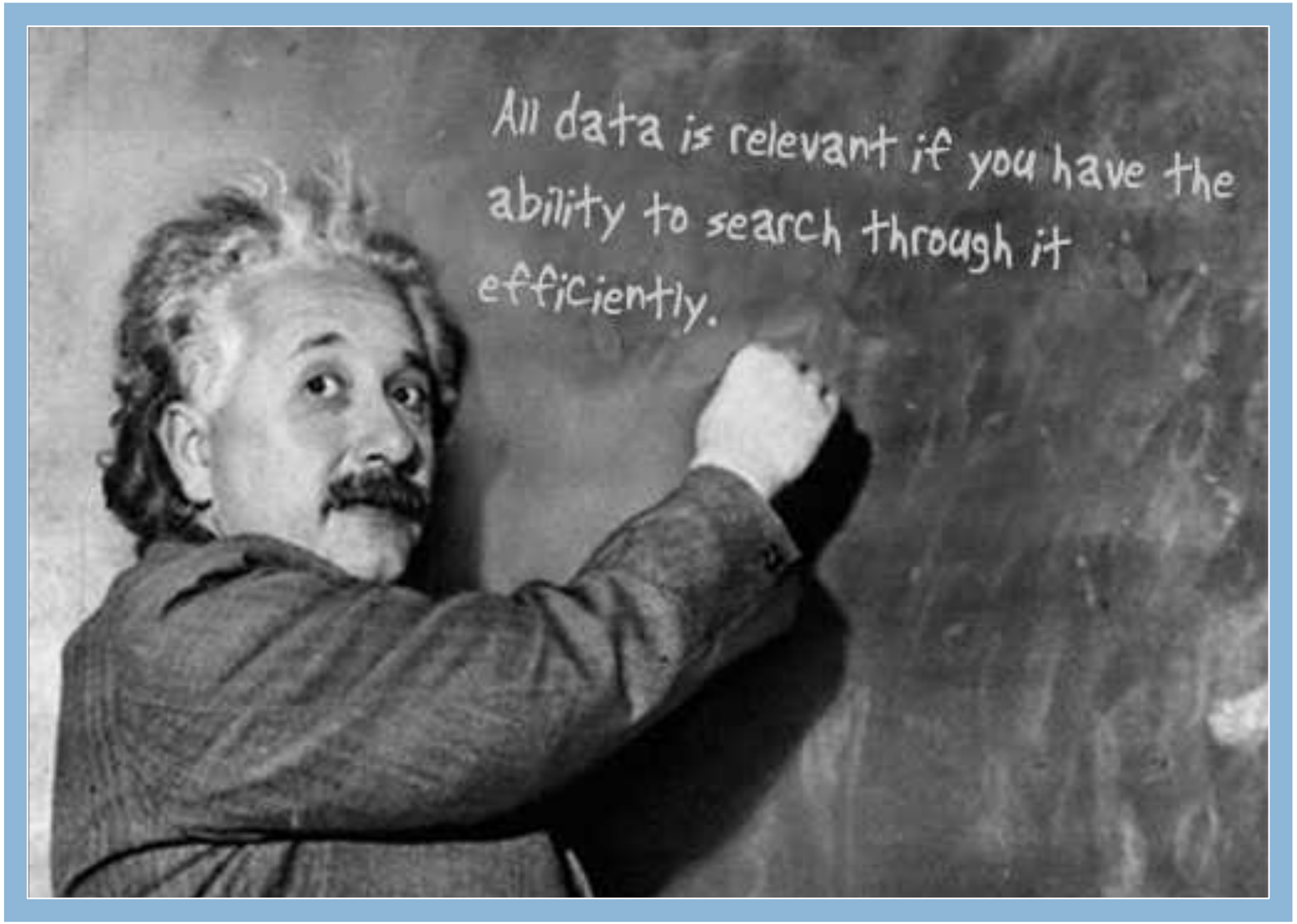
A ‘search’ (question to be answered) is distributed amongst multiple cores.



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