# splunk>®

Matthew David Levy
Arib Alimuddin Patel
Shesh Nath Mishra
Ashwin Sethu Baskaran

# Content

- Introduction
- Terminologies
- Features of System
- Data Model
- Operations Supported
- Languages Supported
- Implementation Concepts
- Applications

# We stare at data all day.

# WTH is Big Data?!

larger **than small data?**

smaller than giant data?

# some cool sauce for DBAs?

Aaaahhh, no.

a simple way to describe a  massive problem

*or opportunity depending on your p.o.v.

# Big data comes out of machines

Machine-generated data is one of the fastest growing, most complex and most valuable segments of big data

GPS,
RFID,
Hypervisor,
Web Servers,
Email, Messaging
Clickstreams, Mobile,
Telephony, IVR, Databases,
Sensors, Telematics, Storage,
Servers, Security Devices, Desktops

Volume | Velocity | Variety | Variability

building a service?

you are a producer and consumer of data

using an app?

# Data! Good!

DATE/TIME

DEVICE

LAT/LONG

BATTERY STRENGTH

```
2011-11-06    11:57:31,65,00027d27-ae02-627d-a79a-fa0004d3a347,40.75496,-
              73.963853,60
2011-11-06    12:17:32,65,00027d27-ae02-627d-a79a-fa0004d3a347,40.755001,-
              73.963886,70
2011-11-06    12:37:38,65,00027d27-ae02-627d-a79a-fa0004d3a347,40.754982,-
              73.963849,75
2011-11-06    12:57:34,65,00027d27-ae02-627d-a79a-fa0004d3a347
              73.963883,85
2011-11-06    13:17:35,65,00027d27-ae02-627d-a79a-fa0004d3a347,40.754941,-
              73.9639,90
2011-11-06    13:37:36,65,00027d27-ae02-627d-a79a-fa0004d3a347,40.754948,-
              73.963874,90
2011-11-06    13:57:37,65,00027d27-ae02-627d-a79a-fa0004d3a347,40.754931,-
              73.963892,95
2011-11-06    14:17:38,50,00027d27-ae02-627d-a79a-
              fa0004d3a347,40.755232,-73.963522,100
2011-11-06    14:37:33,65,00027d27-ae02-627d-a79a-fa0004d3a347,40.754979,-
              73.9639,100
```

All this data can be pretty cool and empowering except one little

# PROBLEM

# A lot of it looks like this

```
0,1
13/Apr/2011 08:52:53,Info,Teardown,ASA-session-6-302014,TCP,
192.168.2.16,192.168.1.6,(empty),(empty),1100,43025,43025_tc
p,
(empty),0,1
13/Apr/2011 08:52:55,Info,Teardown,ASA-session-6-302014,TCP,
192.168.2.75,192.168.1.6,(empty),(empty),1048,135,epmap,(empty
),  0,1
13/Apr/2011 08:52:55,Info,Teardown,ASA-session-6-302014,TCP,
192.168.2.75,192.168.1.6,(empty),(empty),1049,43025,43025_tc
p,
(empty),0,1
13/Apr/2011 08:52:55,Info,Teardown,ASA-session-6-302014,TCP,
192.168.2.75,192.168.1.6,(empty),(empty),1051,135,epmap,(empty
),  0,1
13/Apr/2011 08:52:55,Info,Teardown,ASA-session-6-302014,TCP,
192.168.2.75,192.168.1.6,(empty),(empty),1052,43025,43025_tc
p
```

and we’re expected to talk to  it like this

```
selec  (selec  max(answer.answer  fro  answe  wher  answer.member_id in
t      member_i) fro  team_member mwherr projeect_i  (in (        project_i
from selec  d     wherm Bussiness_stream='Upstream and stage='Appraise'
project_id in  e    '  project_i  fro  projecteand wher lect subteam&lt;&gt; )
(selectpage_id=page.page_id as  m    a(selec  max(avgscore)      task_projec
)where task_project.project_i thhin t    project_i  fro  projectextr )
where subteam=1 )and task_pro(select project_i din      mproject_i  fro
project  where  stage='Appraise' Business_stre(select  d        m    )
and                         as bmax, (selec  max(answ'Upstre  answe
task_project.page_id=page.page_i  as     t  (selectfrom and     r    and
task_project where project_id=1  where datamax, avg(avgscore)  page_id) from
answer.page_id=page.page_id)   fro  task_project where      not
avg(avgscore)  (select projectextra where subteam=1)         in
task_project.page_id=page.page_i  as           (selec  avg(avgscore  fro
d)  task_project where      not in companyavg,project_i  fro)  projectextrm
where subteam=1) and      (select     project_id from  a    wher
Business_stream ='Upstream')  (select task_project.page_id=page.page_id)
and businessavg, page.*    page,riverorde  wher  page.category_name='Busines
from  Boundaries' and stage_name='Appraise'   e    s
riverorder.category_name=page.category_na  orde  by
riverorder.riverorder,page.order_  selec  (selec  max(answer.answer  from
answer where answer.member_id in ( t    member_id) from         wher
project_id in (  select project_id   project_team_members         e
```

It could be better.

yes? Better is good!

# Splunk >

- Splunk brings color and life to your data!
- Powerful platform for analyzing machine data.
- World of technology & World of business.
- Power and Versatility

# Splunk >

- Search
- Monitor
- Analyze
- Report

# Use Cases >

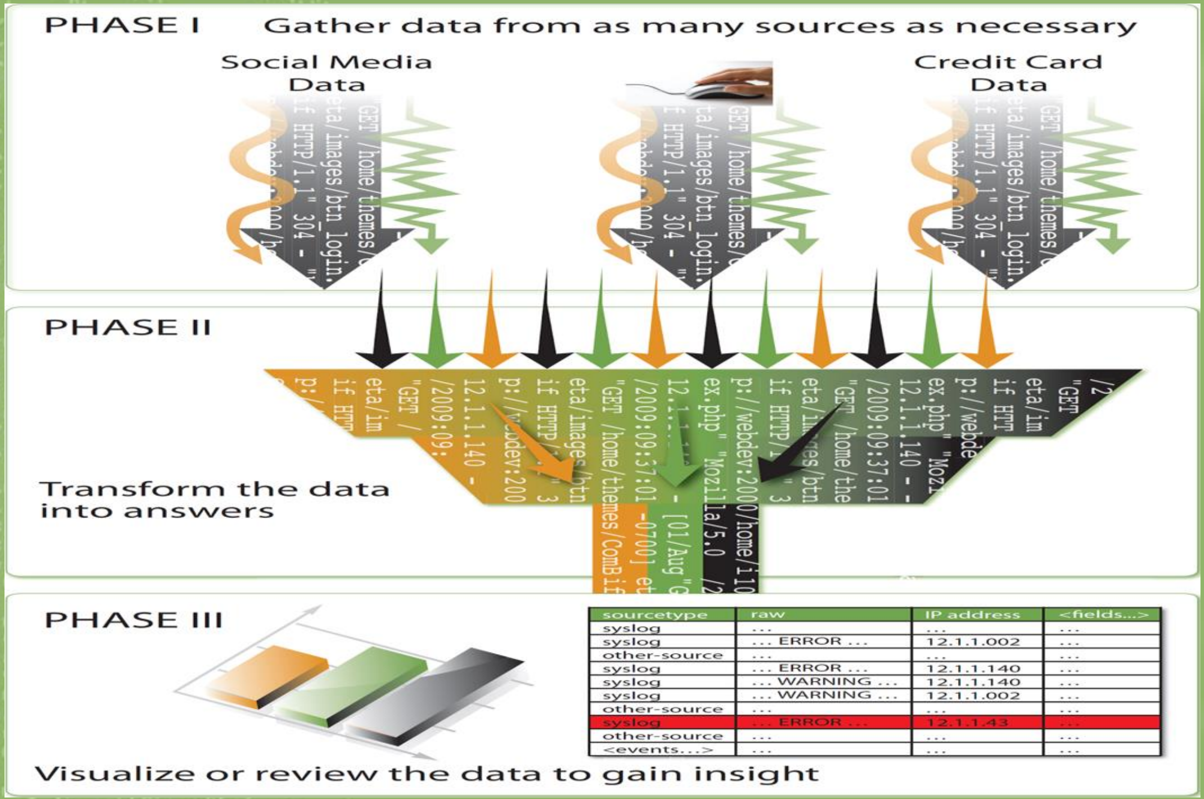
Google for log files

Splunk to the rescue in the data center

# Use Cases >

Splunk to the rescue in the Marketing department

# Phases >



PHASE I Gather data from as many sources as necessary

PHASE II

Transform the data into answers

PHASE III

| sourcetype | raw | IP address | <fields...> |
|---|---|---|---|
| syslog | ... | ... | ... |
| syslog | ... ERROR ... | 12.1.1.002 | ... |
| other-source | ... | ... | ... |
| syslog | ... ERROR ... | 12.1.1.140 | ... |
| syslog | ... WARNING ... | 12.1.1.140 | ... |
| syslog | ... WARNING ... | 12.1.1.002 | ... |
| other-source | ... | ... | ... |
| syslog | ... ERROR ... | 12.1.1.43 | ... |
| other-source | ... | ... | ... |
| <events...> | ... | ... | ... |

Visualize or review the data to gain insight

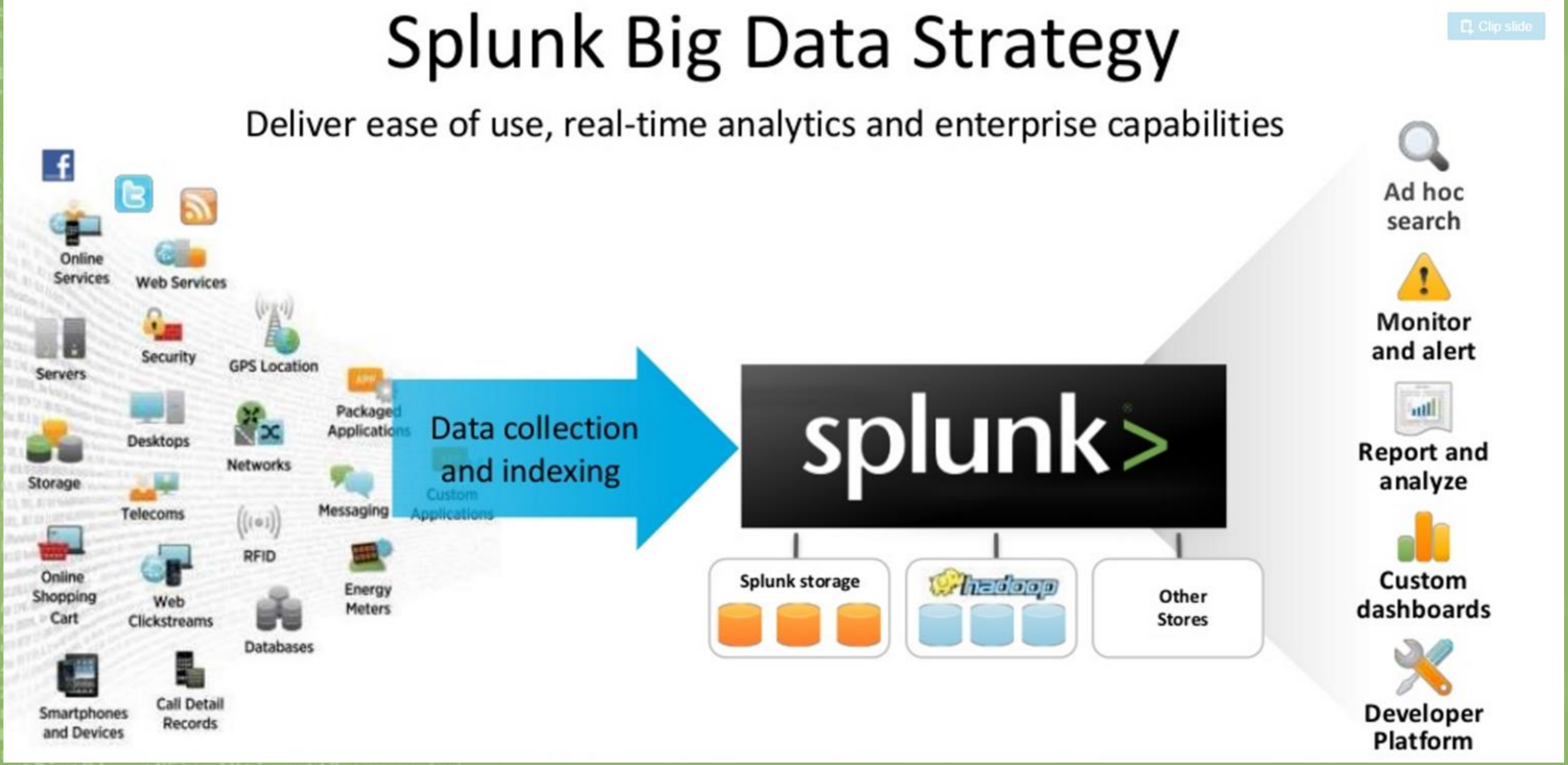
# The (Brief) Story of Splunk >

- Erik Swan and Rob Das in 2002
- "How do you solve problems in your infrastructure?"
- Troubleshoot IT problems and retrieve data by traditional means.
- Spelunking -> Splunk

# Products >

- Splunk Enterprise
- Splunk Storm
- Hunk
- Splunk Light
- Google with Splunk
- Splunkbase

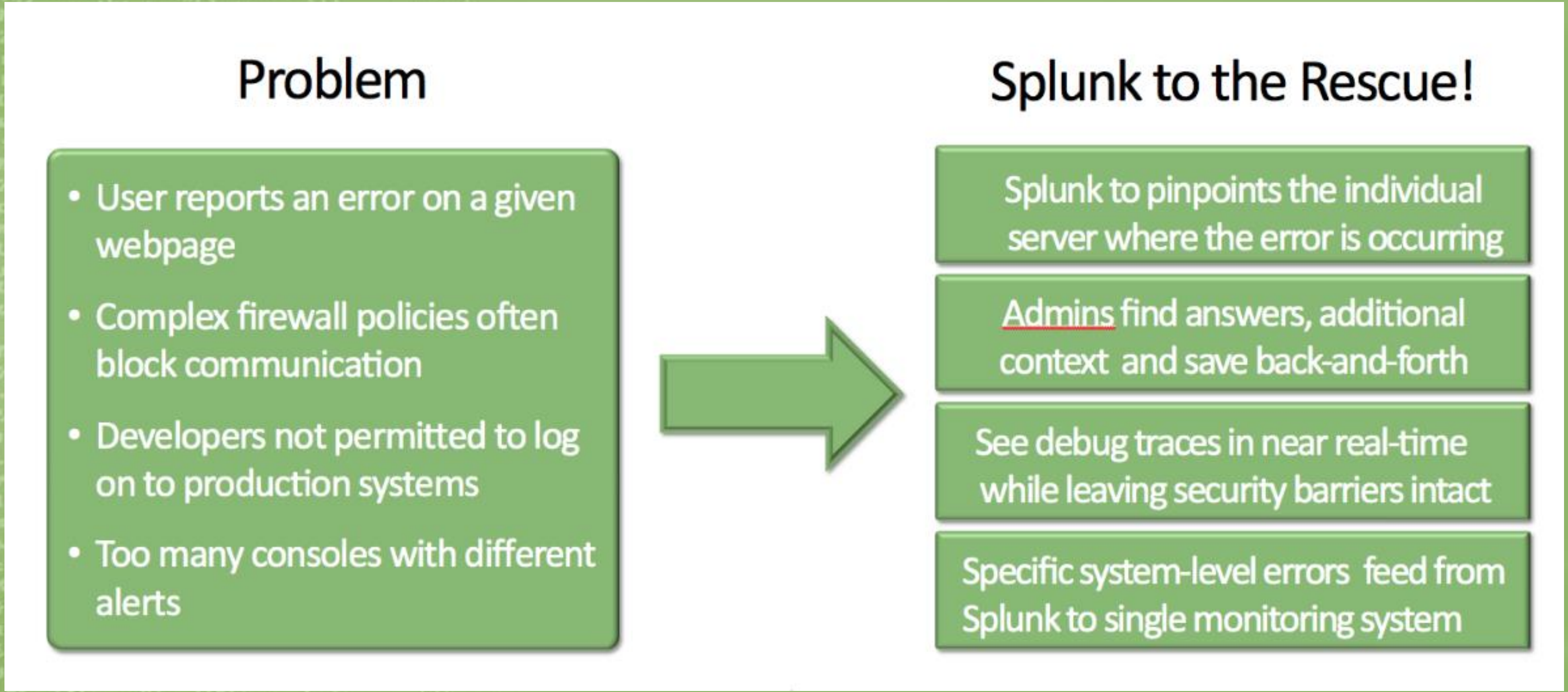
# Applications >



## Splunk Big Data Strategy

Deliver ease of use, real-time analytics and enterprise capabilities

# Solving Problems with Splunk



## Problem

- User reports an error on a given webpage
- Complex firewall policies often block communication
- Developers not permitted to log on to production systems
- Too many consoles with different alerts

## Splunk to the Rescue!

Splunk to pinpoints the individual server where the error is occurring

Admins find answers, additional context  and save back-and-forth

See debug traces in near real-time while leaving security barriers intact

Specific system-level errors  feed from Splunk to single monitoring system

# Solutions with Splunk >

- Converts logs to visual graphs and reports
- Identify and resolve issues faster.
- No separate database requirements.
- Supports any format and any amount of data.
- Simple to implement and scale
- Continually index all of your IT data in real time.
- Automatically discover useful information embedded in your data.
- Set up alerts.
- Proactively review your IT systems.

# Innovation with Splunk >

- Splunk has a mission of making machine data accessible across an organization by identifying data patterns, providing metrics, diagnosing problems and providing intelligence for business operations.

- Splunk is a horizontal technology used for application management, security and compliance, as well as business and web analytics.

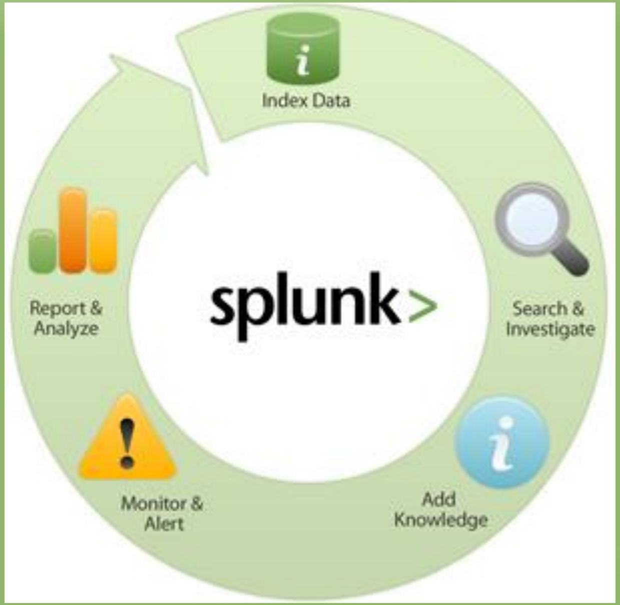
- As of early 2016, Splunk has over 10,000 customers worldwide.

# Operational Intelligence >

- Gain deeper understanding of customers
- Reveal important patterns and analytics
- Event & Detection
- Leverage live feeds & historical data
- Deploy solution quickly and provide flexibility

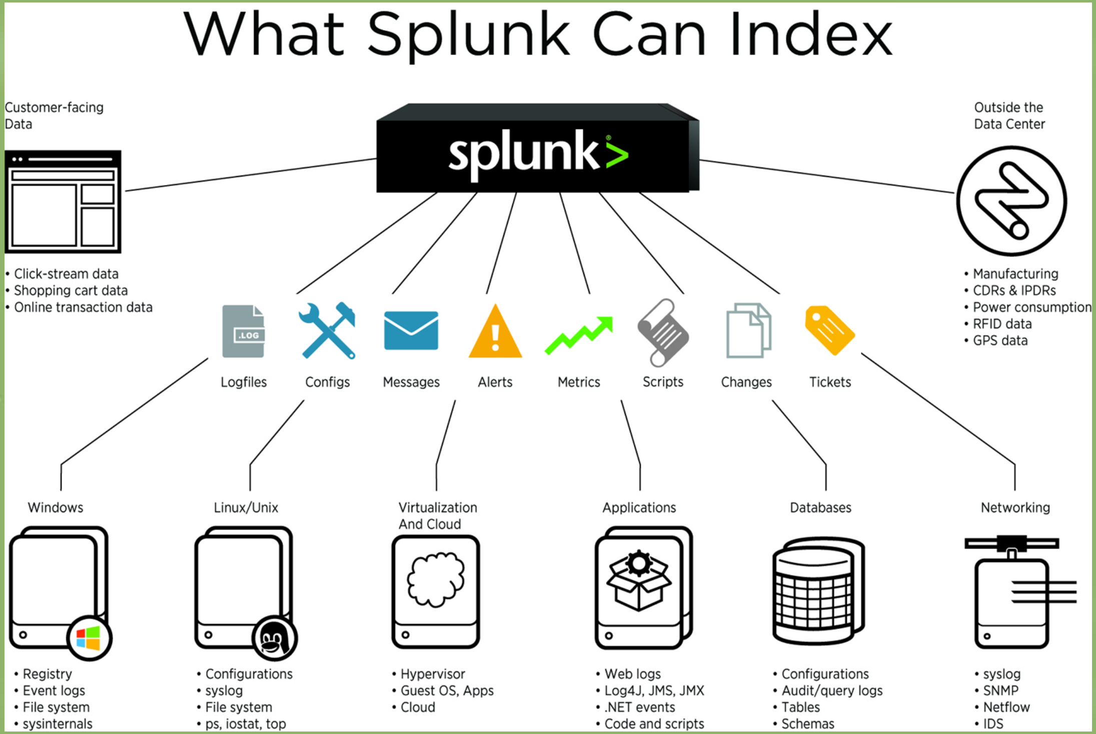
# Features >

- Collect and Index Data
- Search and Investigate
- Correlate and Analyze
- Visualize and Report
- Monitor and Alert



# Collect and Index Data >

- Index Anything, In Real Time
- Getting Data In
- Schema-on-the-Fly
- Time-Based Event Chronology



## What Splunk Can Index

Customer-facing Data

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

Outside the Data Center

- Manufacturing
- CDRs & IPDRs
- Power consumption
- RFID data
- GPS data

Windows

- Registry
- Event logs
- File system
- sysinternals

Linux/Unix

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

Virtualization And Cloud

- Hypervisor
- Guest OS, Apps
- Cloud

Applications

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

Databases

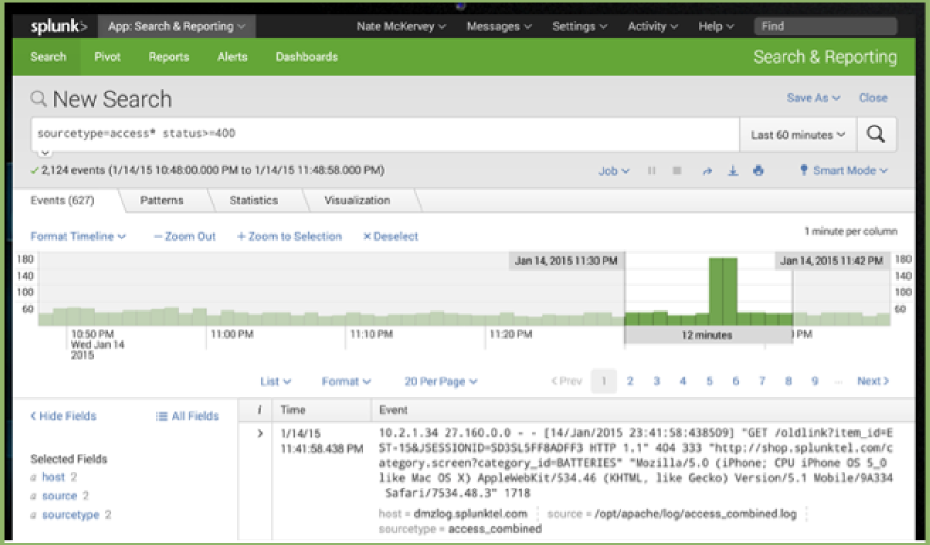
- Configurations
- Audit/query logs
- Tables
- Schemas

Networking

- syslog
- SNMP
- Netflow
- IDS

# Search and Investigate >

- Powerful search, analysis and visualization.
- Splunk Search Processing Language (SPL™)
- Transaction Search
- Interactive Results
- Data Sampling



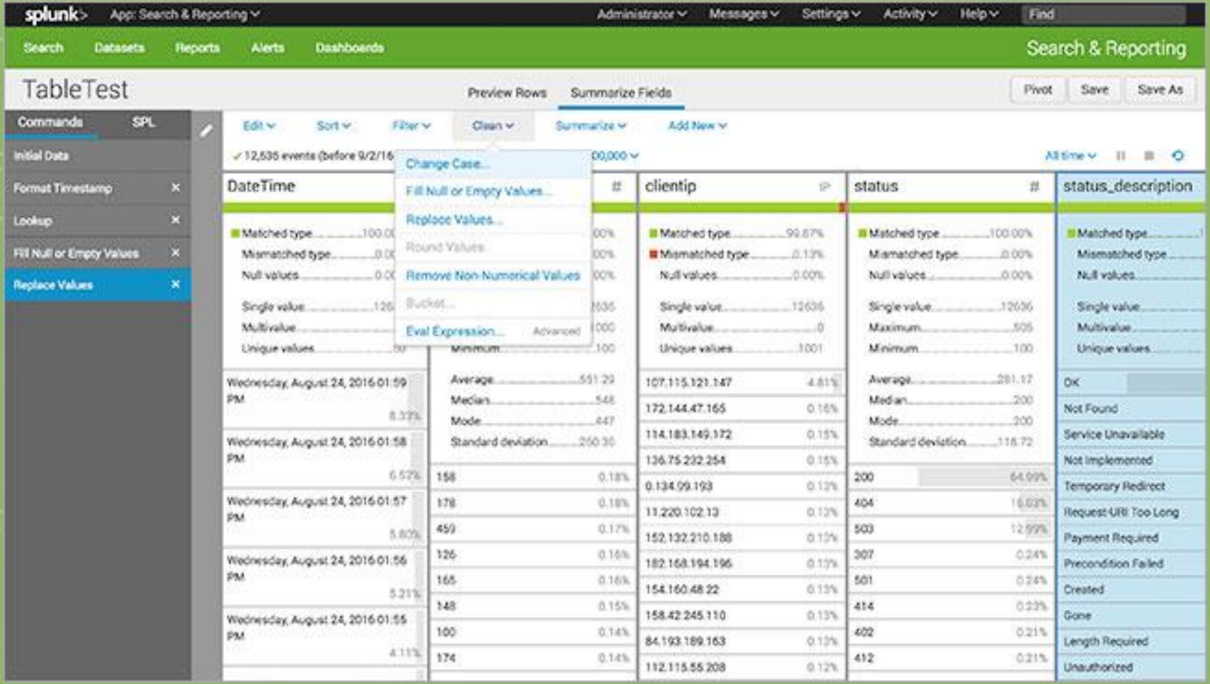
```
sourcetype=access* status>=400
```

✓ 2,124 events (1/14/15 10:48:00.000 PM to 1/14/15 11:48:58.000 PM)

| i | Time | Event |
|---|---|---|
| > | 1/14/15<br>11:41:58.438 PM | 10.2.1.34 27.160.0.0 - - [14/Jan/2015 23:41:58:438509] "GET /oldlink?item_id=EST-15&JSESSIONID=SD3SL5FF8ADFF3 HTTP 1.1" 404 333 "http://shop.splunktel.com/category.screen?category_id=BATTERIES" "Mozilla/5.0 (iPhone; CPU iPhone OS 5_0 like Mac OS X) AppleWebKit/534.46 (KHTML, like Gecko) Version/5.1 Mobile/9A334 Safari/7534.48.3" 1718<br>host = dmzlog.splunktel.com source = /opt/apache/log/access_combined.log sourcetype = access_combined |

# Correlate and Analyze >

- Machine Learning
- Correlate Complex Events
- Event Pattern Detection
- Datasets



TableTest

✓ 12,535 events (before 9/2/16 … 00,000

DateTime

Matched type ........ 100.0(
Mismatched type ........ 0.0(
Null values ........ 0.0(

Single value ........ 126
Multivalue ........
Unique values ........ 50

| | |
|---|---|
| Wednesday, August 24, 2016 01:59 PM | 8.33% |
| Wednesday, August 24, 2016 01:58 PM | 6.52% |
| Wednesday, August 24, 2016 01:57 PM | 5.80% |
| Wednesday, August 24, 2016 01:56 PM | 5.21% |
| Wednesday, August 24, 2016 01:55 PM | 4.11% |

#

00%
00%
00%

2636
000
Minimum ........ 100

Average ........ 551.29
Median ........ 548
Mode ........ 447
Standard deviation ........ 250.35

| | |
|---|---|
| 158 | 0.18% |
| 178 | 0.18% |
| 459 | 0.17% |
| 126 | 0.16% |
| 165 | 0.16% |
| 148 | 0.15% |
| 100 | 0.14% |
| 174 | 0.14% |

clientip

🟩 Matched type ........ 99.87%
🟥 Mismatched type ........ 0.13%
Null values ........ 0.00%

Single value ........ 12636
Multivalue ........ 0
Unique values ........ 1001

| | |
|---|---|
| 107.115.121.147 | 4.81% |
| 172.144.47.165 | 0.16% |
| 114.183.149.172 | 0.15% |
| 136.75.232.254 | 0.15% |
| 0.134.99.193 | 0.13% |
| 11.220.102.13 | 0.13% |
| 152.132.210.188 | 0.13% |
| 182.168.194.196 | 0.13% |
| 154.160.48.22 | 0.13% |
| 158.42.245.110 | 0.13% |
| 84.193.189.163 | 0.13% |
| 112.115.55.208 | 0.12% |

status

🟩 Matched type ........ 100.00%
Mismatched type ........ 0.00%
Null values ........ 0.00%

Single value ........ 12636
Maximum ........ 505
Minimum ........ 100

Average ........ 281.17
Median ........ 200
Mode ........ 200
Standard deviation ........ 118.72

| | |
|---|---|
| 200 | 64.99% |
| 404 | 16.60% |
| 503 | 12.99% |
| 307 | 0.24% |
| 501 | 0.24% |
| 414 | 0.23% |
| 402 | 0.21% |
| 412 | 0.21% |

status_description

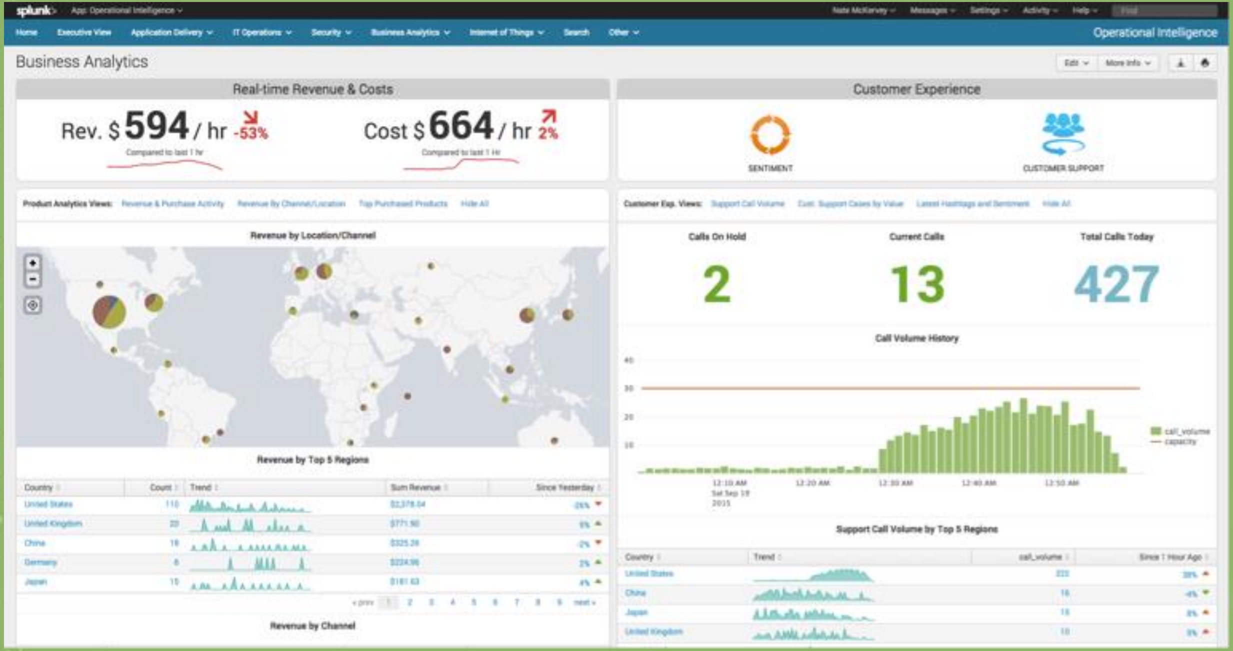
🟩 Matched type ........
Mismatched type ........
Null values ........

Single value ........
Multivalue ........
Unique values ........

OK
Not Found
Service Unavailable
Not Implemented
Temporary Redirect
Request URI Too Long
Payment Required
Precondition Failed
Created
Gone
Length Required
Unauthorized

# Visualize and Report >

- Visualizations
- Dashboards
- Automate and Share Reports



Business Analytics

Real-time Revenue & Costs

Rev. $594 / hr -53%
Compared to last 1 hr

Cost $664 / hr 2%
Compared to last 1 Hr

| Calls On Hold | Current Calls | Total Calls Today |
|---|---|---|
| 2 | 13 | 427 |

Revenue by Top 5 Regions

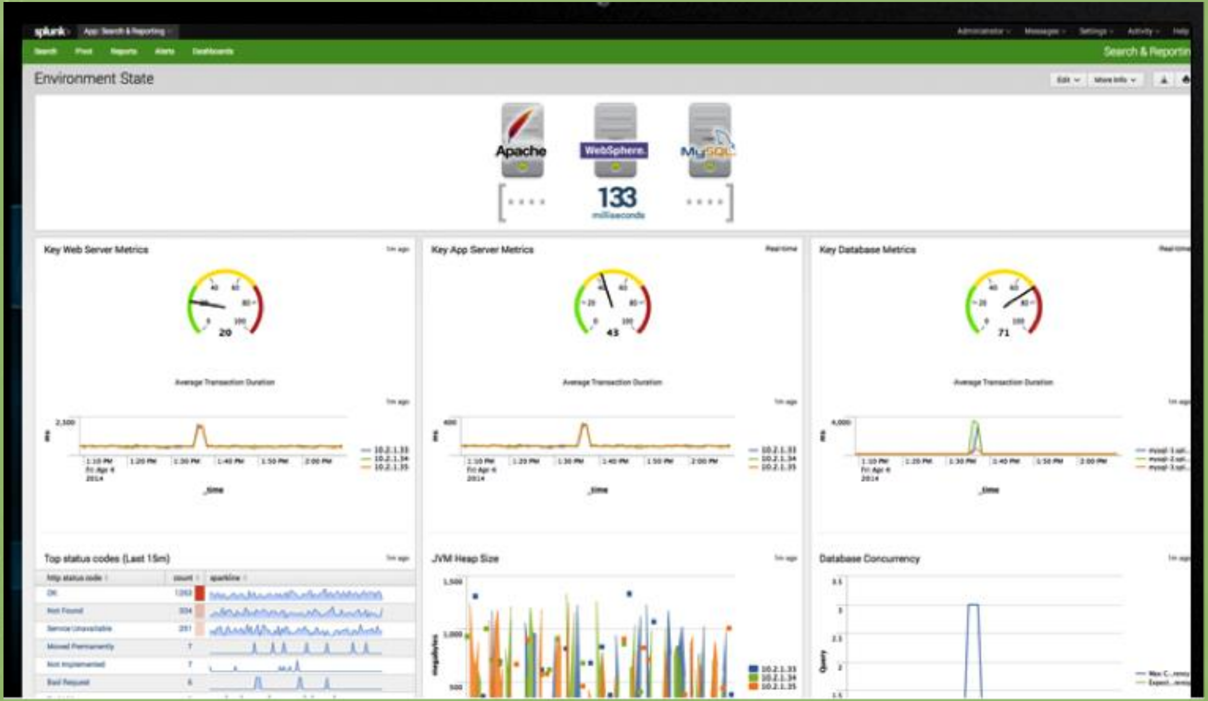
| Country | Count | Trend | Sum Revenue | Since Yesterday |
|---|---|---|---|---|
| United States | 110 | | $2,378.04 | -26% ▼ |
| United Kingdom | 22 | | $771.90 | 6% ▲ |
| China | 18 | | $325.28 | -2% ▼ |
| Germany | 8 | | $224.96 | 2% ▲ |
| Japan | 15 | | $181.63 | 4% ▲ |

Support Call Volume by Top 5 Regions

| Country | Trend | call_volume | Since 1 Hour Ago |
|---|---|---|---|
| United States | | 222 | 39% ▲ |
| China | | 18 | -4% ▼ |
| Japan | | 18 | 8% ▲ |
| United Kingdom | | 10 | 9% ▲ |

# Monitor and Alert >

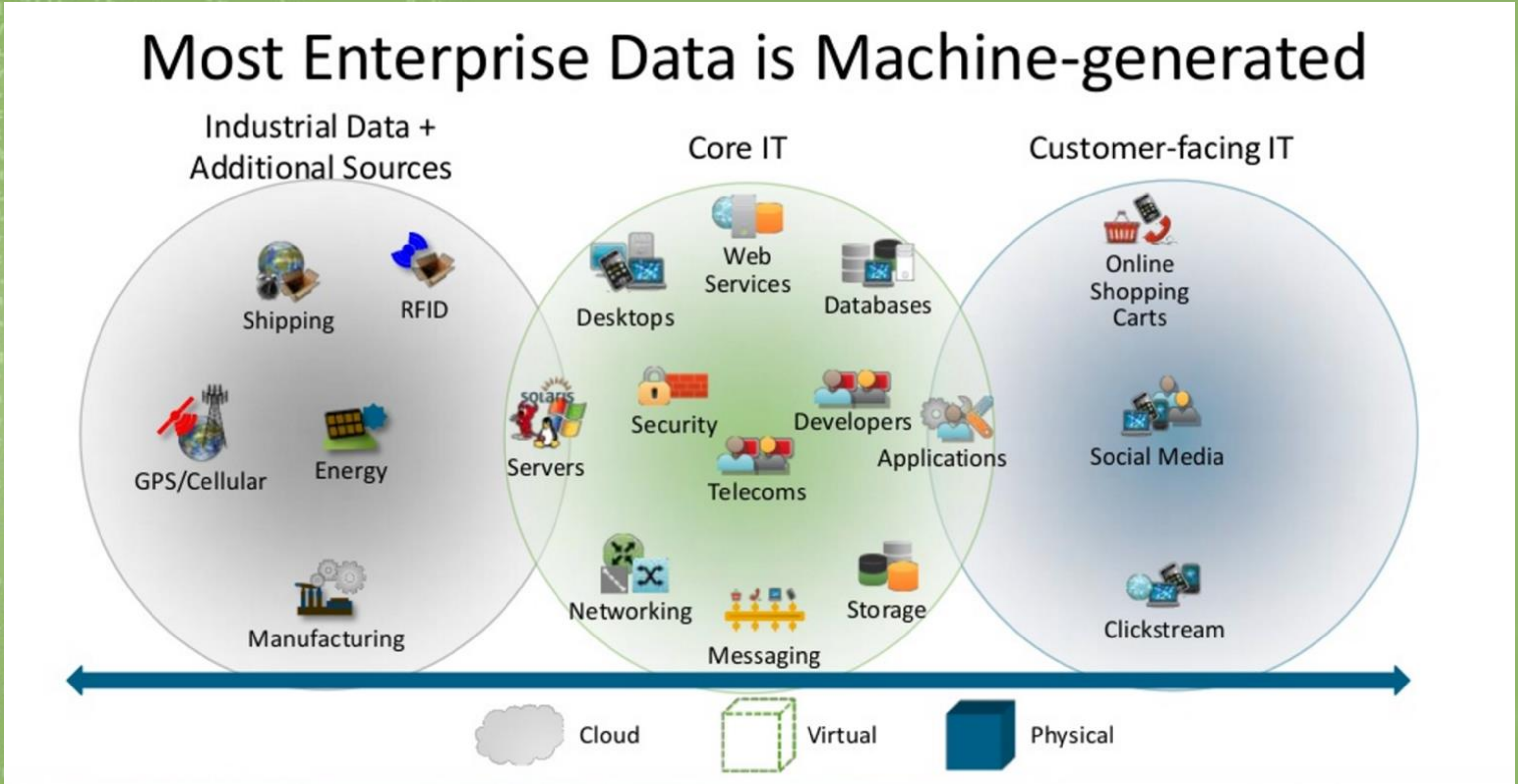
- Monitor Events and KPIs
- Proactive Alerting
- Access from Anywhere



Top status codes (Last 15m)

| http status code | count | sparkline |
|---|---|---|
| OK | 1,263 | |
| Not Found | 304 | |
| Service Unavailable | 251 | |
| Moved Permanently | 7 | |
| Not Implemented | 7 | |
| Bad Request | 6 | |

# Data Model >



## Most Enterprise Data is Machine-generated

Industrial Data + Additional Sources

Shipping

RFID

GPS/Cellular

Energy

Manufacturing

Core IT

Web Services

Desktops

Databases

Servers

Security

Developers

Applications

Telecoms

Networking

Storage

Messaging

Customer-facing IT

Online Shopping Carts

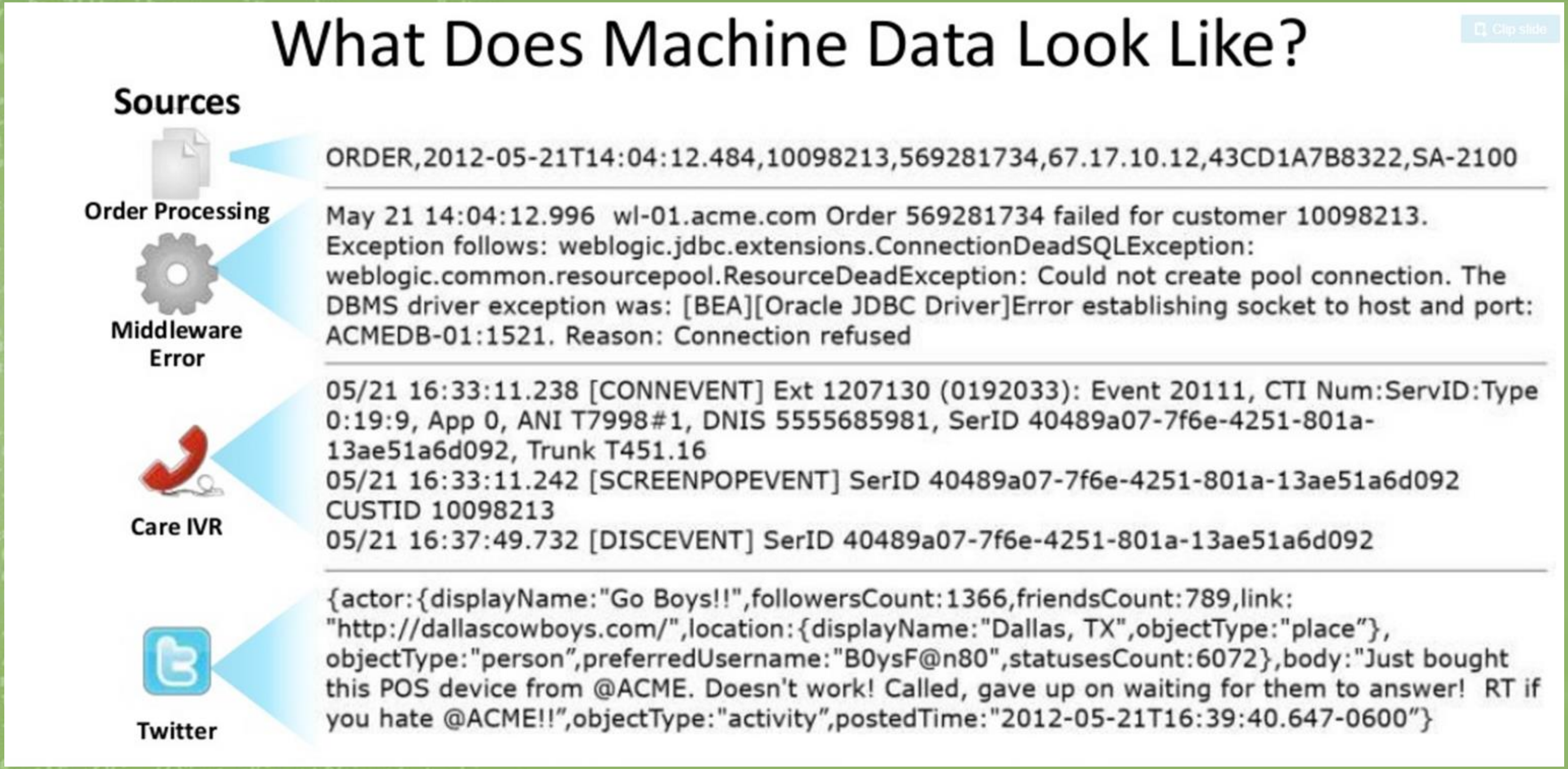
Social Media

Clickstream

Cloud

Virtual

Physical



# What Does Machine Data Look Like?

**Sources**

**Order Processing**

```
ORDER,2012-05-21T14:04:12.484,10098213,569281734,67.17.10.12,43CD1A7B8322,SA-2100
```

**Middleware Error**

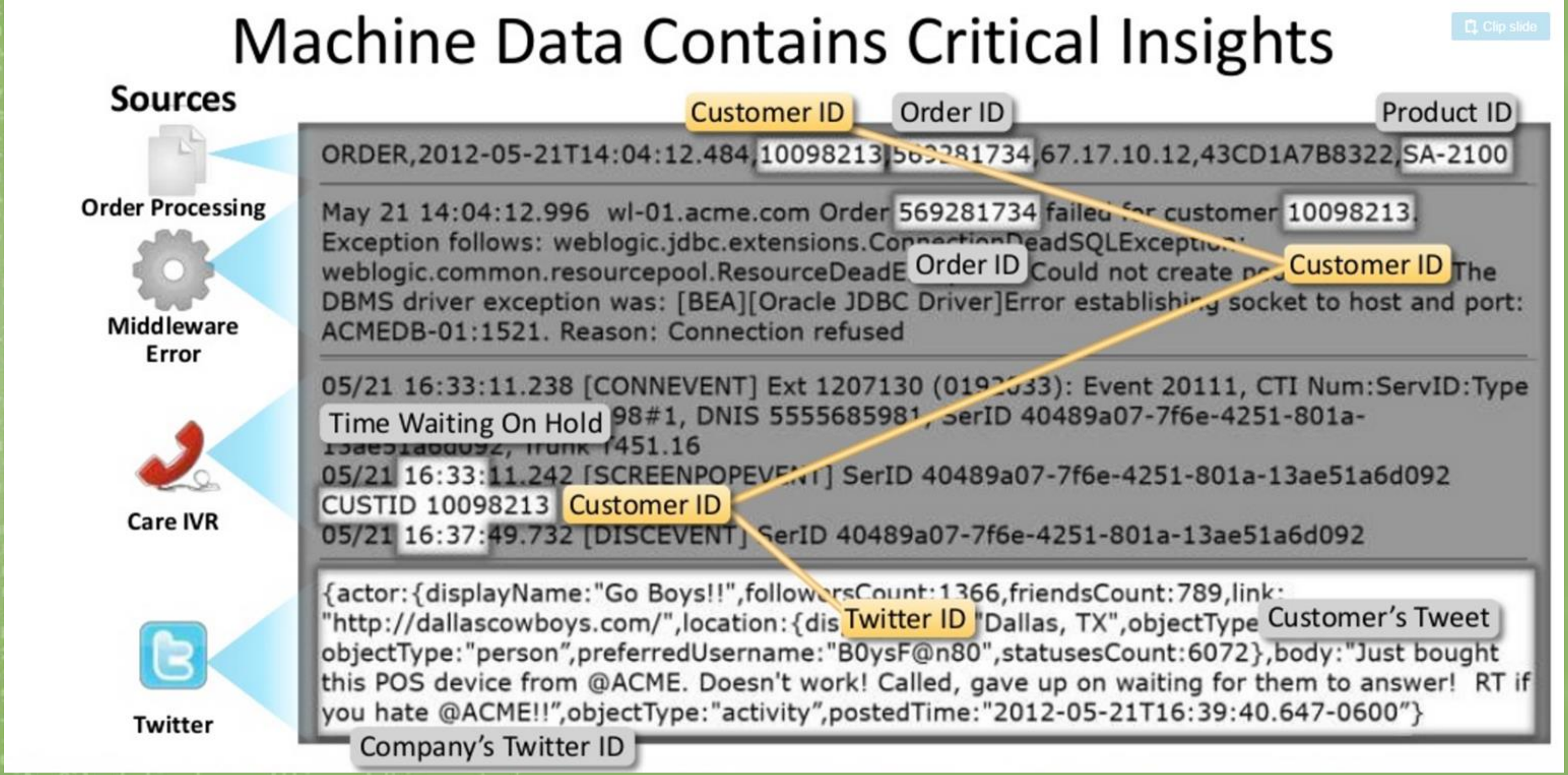
```
May 21 14:04:12.996  wl-01.acme.com Order 569281734 failed for customer 10098213.
Exception follows: weblogic.jdbc.extensions.ConnectionDeadSQLException:
weblogic.common.resourcepool.ResourceDeadException: Could not create pool connection. The
DBMS driver exception was: [BEA][Oracle JDBC Driver]Error establishing socket to host and port:
ACMEDB-01:1521. Reason: Connection refused
```

**Care IVR**

```
05/21 16:33:11.238 [CONNEVENT] Ext 1207130 (0192033): Event 20111, CTI Num:ServID:Type
0:19:9, App 0, ANI T7998#1, DNIS 5555685981, SerID 40489a07-7f6e-4251-801a-
13ae51a6d092, Trunk T451.16
05/21 16:33:11.242 [SCREENPOPEVENT] SerID 40489a07-7f6e-4251-801a-13ae51a6d092
CUSTID 10098213
05/21 16:37:49.732 [DISCEVENT] SerID 40489a07-7f6e-4251-801a-13ae51a6d092
```

**Twitter**

```
{actor:{displayName:"Go Boys!!",followersCount:1366,friendsCount:789,link:
"http://dallascowboys.com/",location:{displayName:"Dallas, TX",objectType:"place"},
objectType:"person",preferredUsername:"B0ysF@n80",statusesCount:6072},body:"Just bought
this POS device from @ACME. Doesn't work! Called, gave up on waiting for them to answer!  RT if
you hate @ACME!!",objectType:"activity",postedTime:"2012-05-21T16:39:40.647-0600"}
```



# Machine Data Contains Critical Insights

```
ORDER,2012-05-21T14:04:12.484,10098213,569281734,67.17.10.12,43CD1A7B8322,SA-2100
```

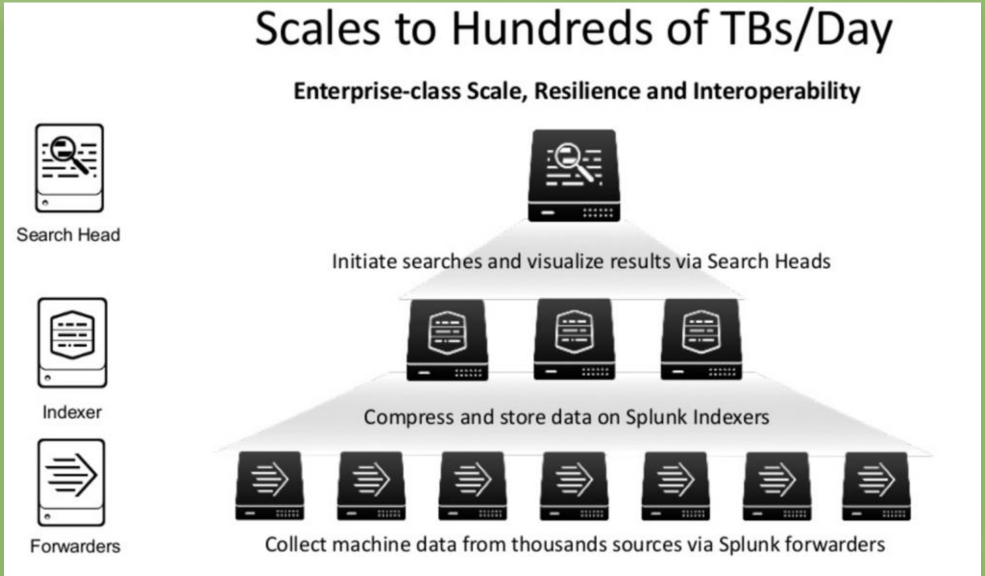
```
May 21 14:04:12.996  wl-01.acme.com Order 569281734 failed for customer 10098213.
Exception follows: weblogic.jdbc.extensions.ConnectionDeadSQLException:
weblogic.common.resourcepool.ResourceDeadE          Could not create            The
DBMS driver exception was: [BEA][Oracle JDBC Driver]Error establishing socket to host and port:
ACMEDB-01:1521. Reason: Connection refused
```

```
05/21 16:33:11.238 [CONNEVENT] Ext 1207130 (0192033): Event 20111, CTI Num:ServID:Type
                     98#1, DNIS 5555685981, SerID 40489a07-7f6e-4251-801a-
13ae51a6d092, Trunk T451.16
05/21 16:33:11.242 [SCREENPOPEVENT] SerID 40489a07-7f6e-4251-801a-13ae51a6d092
CUSTID 10098213
05/21 16:37:49.732 [DISCEVENT] SerID 40489a07-7f6e-4251-801a-13ae51a6d092
```

```
{actor:{displayName:"Go Boys!!",followersCount:1366,friendsCount:789,link:
"http://dallascowboys.com/",location:{dis            "Dallas, TX",objectType
objectType:"person",preferredUsername:"B0ysF@n80",statusesCount:6072},body:"Just bought
this POS device from @ACME. Doesn't work! Called, gave up on waiting for them to answer!  RT if
you hate @ACME!!",objectType:"activity",postedTime:"2012-05-21T16:39:40.647-0600"}
```

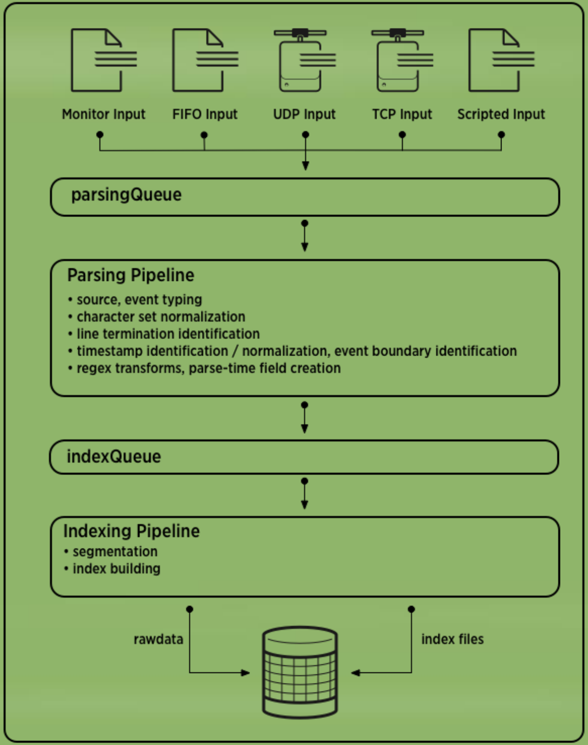
# Data Model >

- What is Indexing
- Indexes Supported
- Indexing Data


Scales to Hundreds of TBs/Day
Enterprise-class Scale, Resilience and Interoperability
Search Head
Initiate searches and visualize results via Search Heads
Indexer
Compress and store data on Splunk Indexers
Forwarders
Collect machine data from thousands sources via Splunk forwarders

# Event processing and the data pipeline

- Configures character set encoding.
- Configures line breaking for multi-line events.
- Identifies event timestamps.
- Extracts a set of useful standard fields.
- Segments events.
- Dynamically assigns metadata to events, if specified.


Monitor Input
FIFO Input
UDP Input
TCP Input
Scripted Input
parsingQueue
Parsing Pipeline
• source, event typing
• character set normalization
• line termination identification
• timestamp identification / normalization, event boundary identification
• regex transforms, parse-time field creation
indexQueue
Indexing Pipeline
• segmentation
• index building
rawdata
index files

# The Search & Reporting application

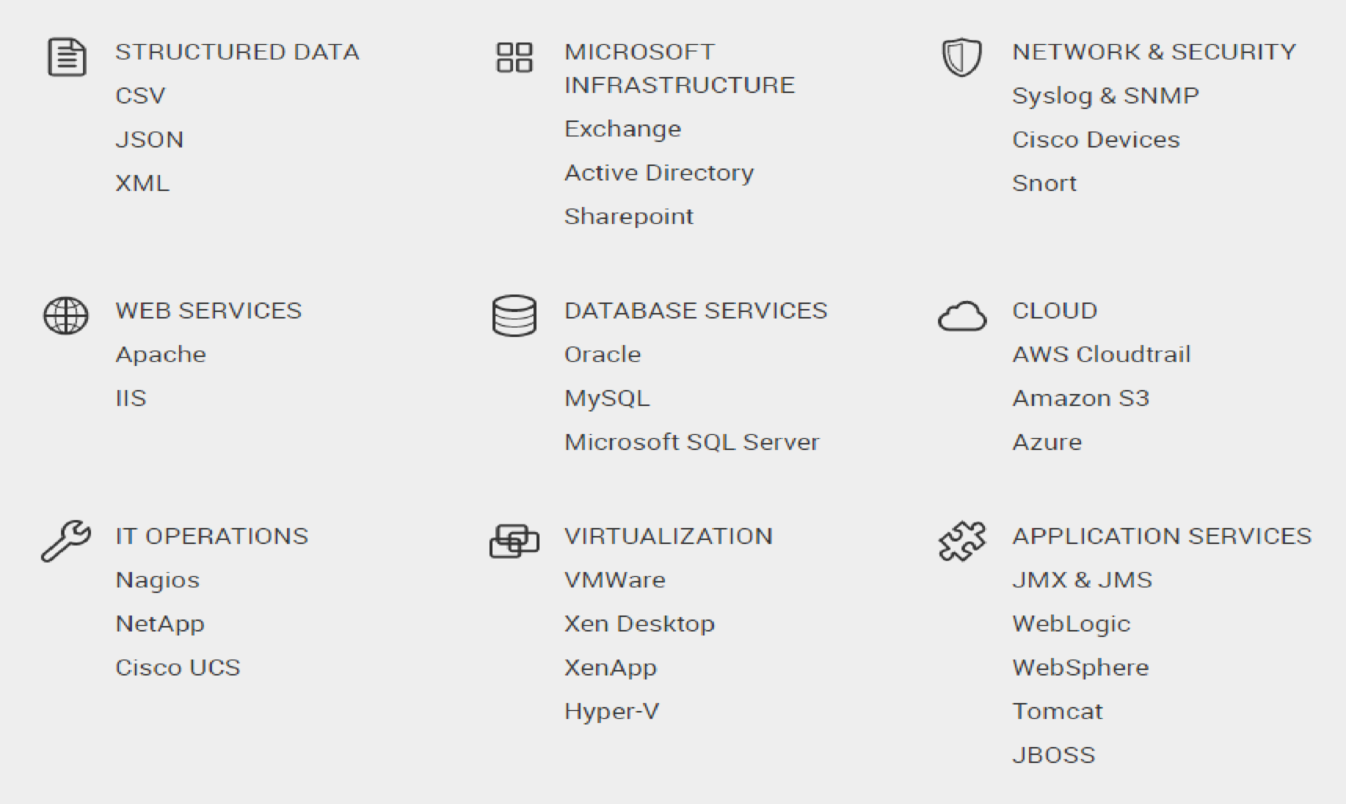
It is the primary interface for using the Splunk software

It can be used to

- Run searches
- Save reports
- Create dashboards.

# Uploading Data

- Adding the Data
  - The data is processed and transformed into a series of individual events that you can view, search, and analyze.
- Types of data
  - The Splunk platform accepts any type of data.
  - event logs
  - web logs
  - live application logs
  - network feeds



**STRUCTURED DATA**

- CSV
- JSON
- XML

**MICROSOFT INFRASTRUCTURE**

- Exchange
- Active Directory
- Sharepoint

**NETWORK & SECURITY**

- Syslog & SNMP
- Cisco Devices
- Snort

**WEB SERVICES**

- Apache
- IIS

**DATABASE SERVICES**

- Oracle
- MySQL
- Microsoft SQL Server

**CLOUD**

- AWS Cloudtrail
- Amazon S3
- Azure

**IT OPERATIONS**

- Nagios
- NetApp
- Cisco UCS

**VIRTUALIZATION**

- VMWare
- Xen Desktop
- XenApp
- Hyper-V

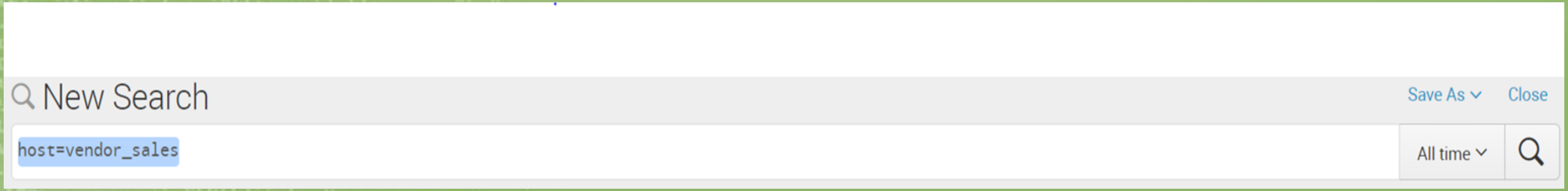
**APPLICATION SERVICES**

- JMX & JMS
- WebLogic
- WebSphere
- Tomcat
- JBOSS

# Where is the data stored?

- Indexing
- Events
- Events are stored in the index as a group of files that fall into two categories:
  - Raw data, which is the data that you add to the Splunk deployment. The raw data is stored in a compressed format.
  - Index files, which include some metadata files that point to the raw data.
- These files reside in sets of directories, called buckets, that are organized by age.

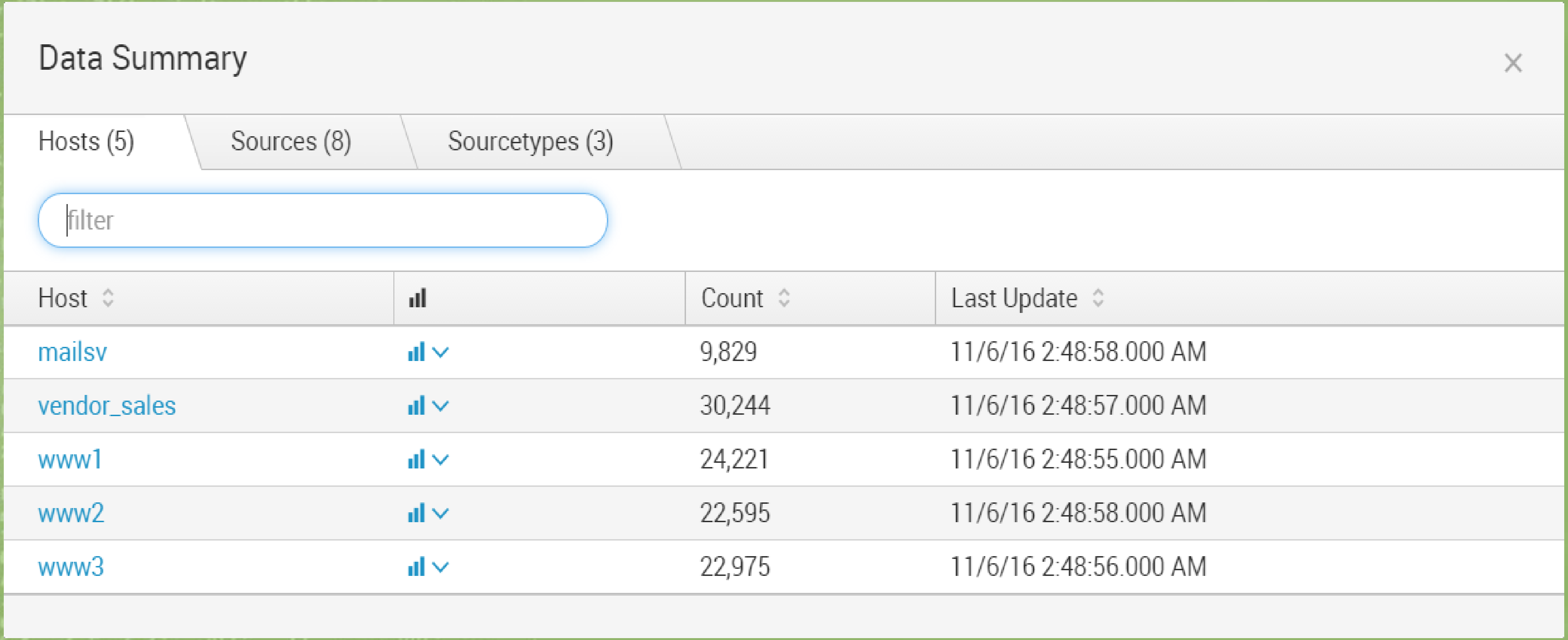
# Searching the data



```
host=vendor_sales
```

- host=vendor_sales
- source="tutorialdata.zip:.\\www1/access.log“
- source="tutorialdata.zip:.\\vendor_sales/vendor_sales.log“
- sourcetype="www1/secure"

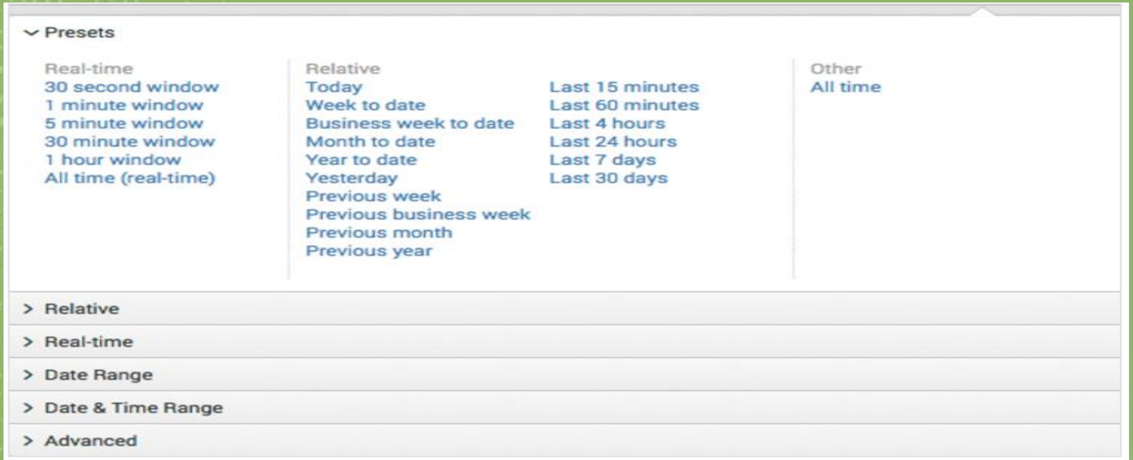
# Searching the Data-Data Summary Dialog Box



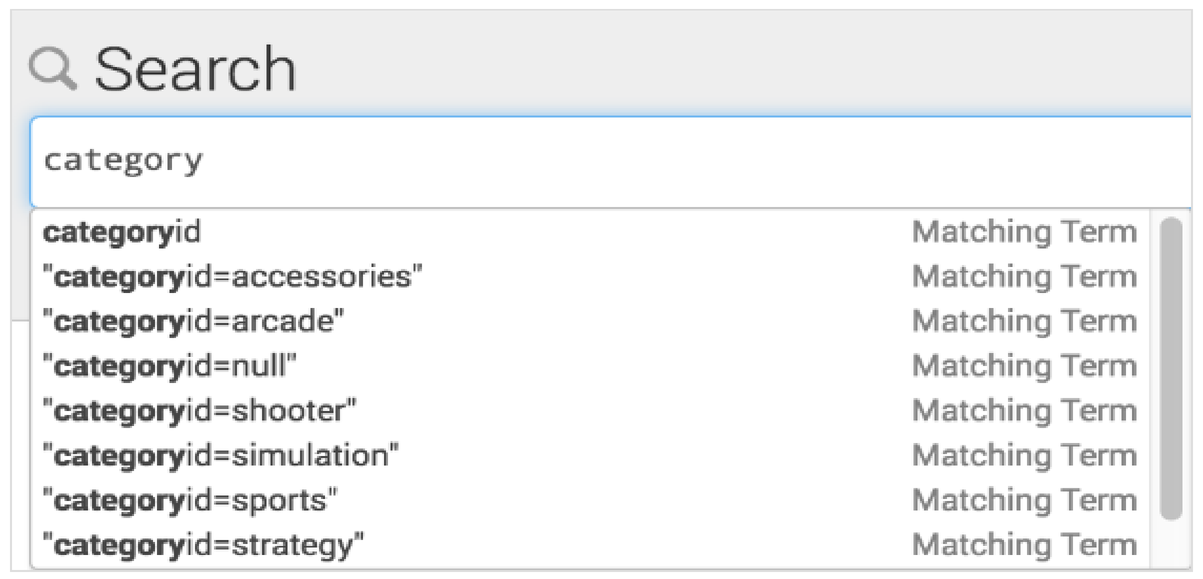
| Host | | Count | Last Update |
|---|---|---|---|
| mailsv | | 9,829 | 11/6/16 2:48:58.000 AM |
| vendor_sales | | 30,244 | 11/6/16 2:48:57.000 AM |
| www1 | | 24,221 | 11/6/16 2:48:55.000 AM |
| www2 | | 22,595 | 11/6/16 2:48:58.000 AM |
| www3 | | 22,975 | 11/6/16 2:48:56.000 AM |

# Specifying time ranges

- Optimize Searches
- Troubleshoot an issue



# Search Assistant



Search

category

**category**id Matching Term

"**category**id=accessories" Matching Term

"**category**id=arcade" Matching Term

"**category**id=null" Matching Term

"**category**id=shooter" Matching Term

"**category**id=simulation" Matching Term

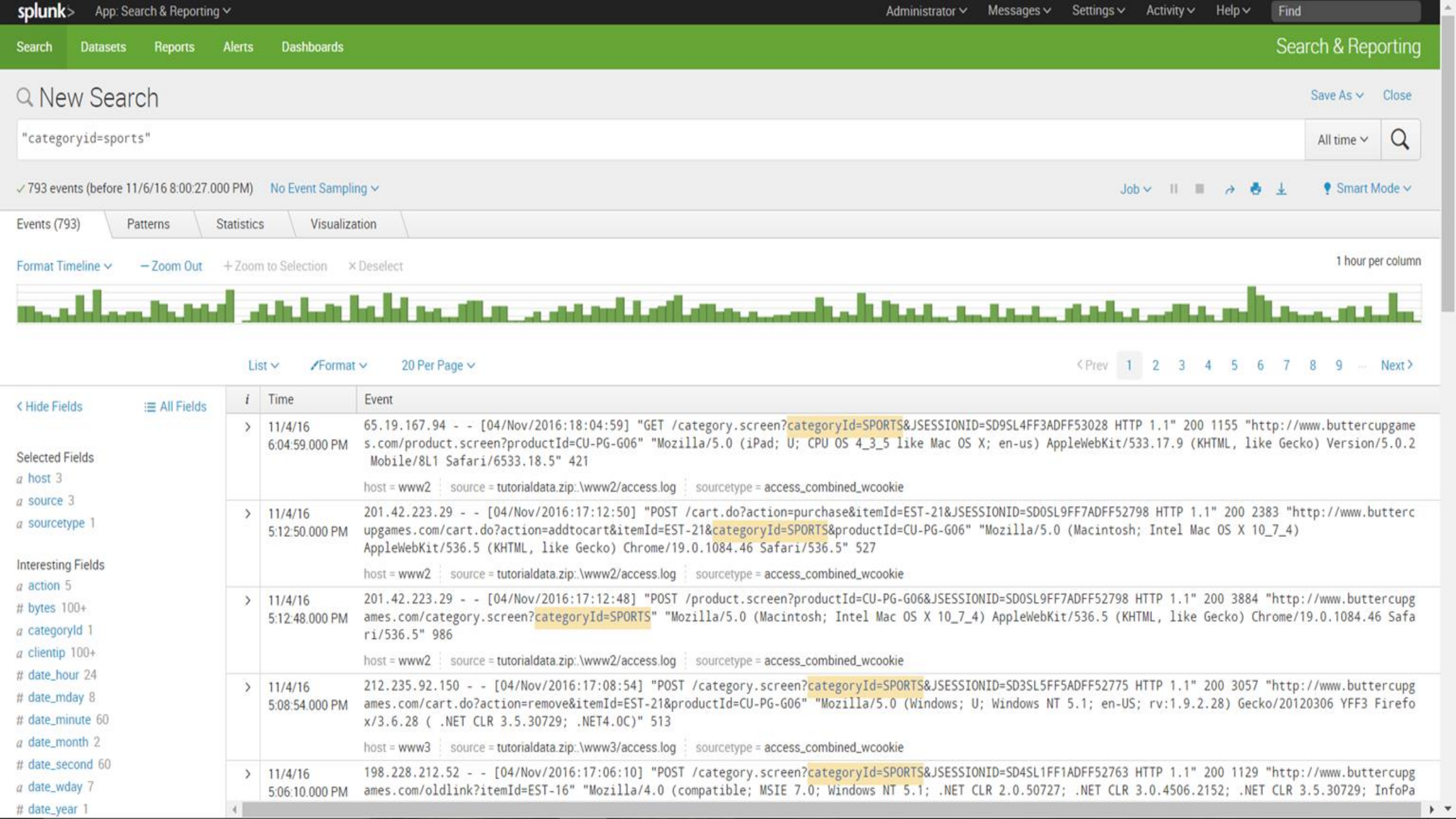
"**category**id=sports" Matching Term

"**category**id=strategy" Matching Term

# Understanding Searches

Below the Search bar are four tabs:

- Events
- Patterns
- Statistics
- Visualizations.



"categoryid=sports"

✓ 793 events (before 11/6/16 8:00:27.000 PM) No Event Sampling

Selected Fields

- a host 3
- a source 3
- a sourcetype 1

Interesting Fields

- a action 5
- # bytes 100+
- a categoryId 1
- a clientip 100+
- # date_hour 24
- # date_mday 8
- # date_minute 60
- a date_month 2
- # date_second 60
- a date_wday 7
- # date_year 1

| Time | Event |
|---|---|
| 11/4/16 6:04:59.000 PM | 65.19.167.94 - - [04/Nov/2016:18:04:59] "GET /category.screen?categoryId=SPORTS&JSESSIONID=SD9SL4FF3ADFF53028 HTTP 1.1" 200 1155 "http://www.buttercupgames.com/product.screen?productId=CU-PG-G06" "Mozilla/5.0 (iPad; U; CPU OS 4_3_5 like Mac OS X; en-us) AppleWebKit/533.17.9 (KHTML, like Gecko) Version/5.0.2 Mobile/8L1 Safari/6533.18.5" 421<br>host = www2 source = tutorialdata.zip:.\www2/access.log sourcetype = access_combined_wcookie |
| 11/4/16 5:12:50.000 PM | 201.42.223.29 - - [04/Nov/2016:17:12:50] "POST /cart.do?action=purchase&itemId=EST-21&JSESSIONID=SD0SL9FF7ADFF52798 HTTP 1.1" 200 2383 "http://www.buttercupgames.com/cart.do?action=addtocart&itemId=EST-21&categoryId=SPORTS&productId=CU-PG-G06" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_7_4) AppleWebKit/536.5 (KHTML, like Gecko) Chrome/19.0.1084.46 Safari/536.5" 527<br>host = www2 source = tutorialdata.zip:.\www2/access.log sourcetype = access_combined_wcookie |
| 11/4/16 5:12:48.000 PM | 201.42.223.29 - - [04/Nov/2016:17:12:48] "POST /product.screen?productId=CU-PG-G06&JSESSIONID=SD0SL9FF7ADFF52798 HTTP 1.1" 200 3884 "http://www.buttercupgames.com/category.screen?categoryId=SPORTS" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_7_4) AppleWebKit/536.5 (KHTML, like Gecko) Chrome/19.0.1084.46 Safari/536.5" 986<br>host = www2 source = tutorialdata.zip:.\www2/access.log sourcetype = access_combined_wcookie |
| 11/4/16 5:08:54.000 PM | 212.235.92.150 - - [04/Nov/2016:17:08:54] "POST /category.screen?categoryId=SPORTS&JSESSIONID=SD3SL5FF5ADFF52775 HTTP 1.1" 200 3057 "http://www.buttercupgames.com/cart.do?action=remove&itemId=EST-21&productId=CU-PG-G06" "Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.9.2.28) Gecko/20120306 YFF3 Firefox/3.6.28 ( .NET CLR 3.5.30729; .NET4.0C)" 513<br>host = www3 source = tutorialdata.zip:.\www3/access.log sourcetype = access_combined_wcookie |
| 11/4/16 5:06:10.000 PM | 198.228.212.52 - - [04/Nov/2016:17:06:10] "POST /category.screen?categoryId=SPORTS&JSESSIONID=SD4SL1FF1ADFF52763 HTTP 1.1" 200 1129 "http://www.buttercupgames.com/oldlink?itemId=EST-16" "Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 5.1; .NET CLR 2.0.50727; .NET CLR 3.0.4506.2152; .NET CLR 3.5.30729; InfoPa |

# Use fields to search

When searching for fields, we use the syntax **fieldname=fieldvalue.**

**Search for successful purchases**

- sourcetype=access_* status=200 action=purchase

**Search for unsuccessful purchases**

- sourcetype=access_* status!=200 action=purchase

**Search for errors**

- (error OR fail* OR severe) OR (status=404 OR status=500 OR status=503)

**Search for sales of a specific product**

- sourcetype=access_* status=200 action=purchase categoryId=simulation

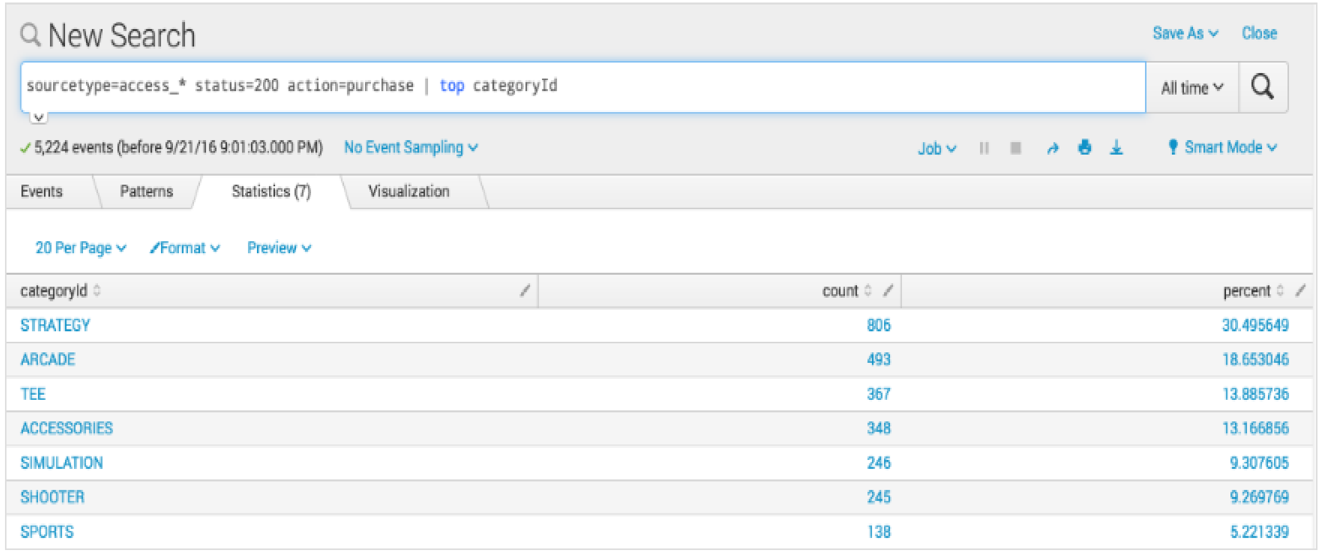
# Pipe and Commands

- The pipe character (|) indicates that you are about to use a command.
- The results of the search to the left of the pipe are used as the input to the command to the right of the pipe.

```
sourcetype=access_* status=200 action=purchase
sourcetype=access_* status=200 action=purchase | top
sourcetype=access_* status=200 action=purchase | top categoryId
```

# Statistics Tab

- The **top** command is a transforming command.



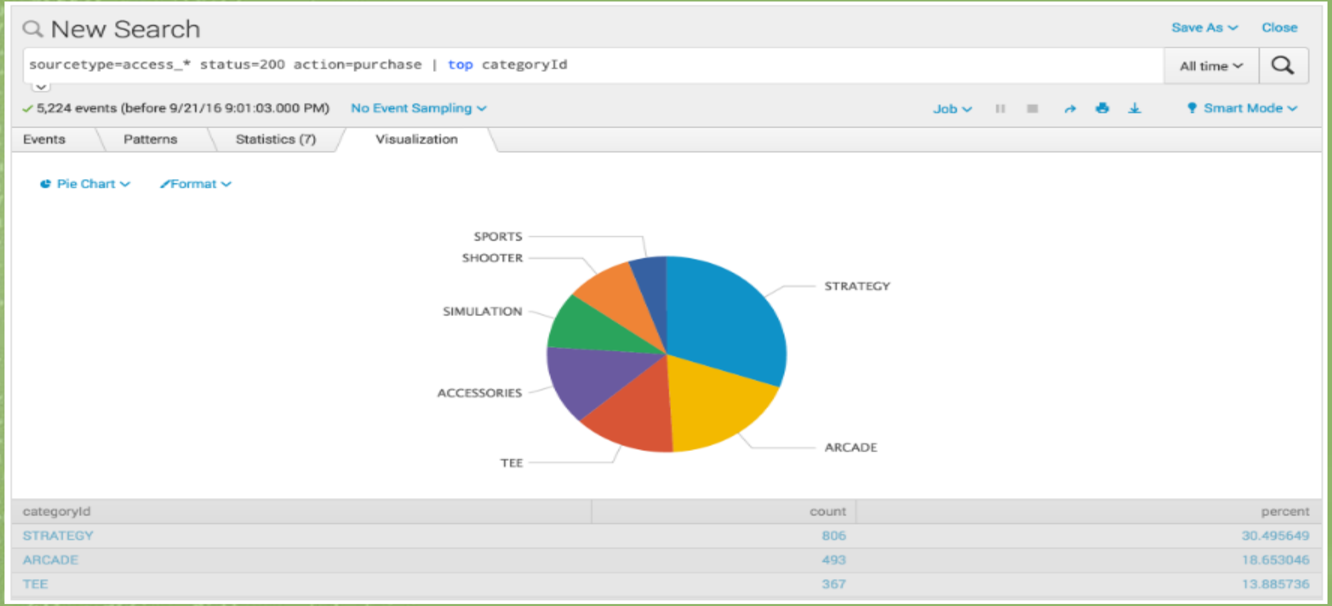
```
sourcetype=access_* status=200 action=purchase | top categoryId
```

✓ 5,224 events (before 9/21/16 9:01:03.000 PM)

| categoryId | count | percent |
|---|---|---|
| STRATEGY | 806 | 30.495649 |
| ARCADE | 493 | 18.653046 |
| TEE | 367 | 13.885736 |
| ACCESSORIES | 348 | 13.166856 |
| SIMULATION | 246 | 9.307605 |
| SHOOTER | 245 | 9.269769 |
| SPORTS | 138 | 5.221339 |

# Visualization

- Gives a graphical representation to the data.



```
sourcetype=access_* status=200 action=purchase | top categoryId
```

✓ 5,224 events (before 9/21/16 9:01:03.000 PM)

| categoryId | count | percent |
|---|---|---|
| STRATEGY | 806 | 30.495649 |
| ARCADE | 493 | 18.653046 |
| TEE | 367 | 13.885736 |

# Reports and Dashboards

- Reports are created whenever we save a search. We provide time ranges
- Dashboards are views that are made up of panels.
- The panels can contain modules such as search boxes, fields, charts, tables, and lists..

# SplunkBase

- Extend the Power of Splunk with Apps and Add-ons
- Splunkbase has 1000+ apps and add-ons from Splunk and it’s partners and it’s community.
- An app or add-on for almost any data source and user need.
- Apps or add-on belonging to below categories:

  DevOps (Example: Splunk App for Jenkins)

  IT Operations (Example: Alert Manager)

  Security, Fraud & Compliance (Example :Splunk Add-on for Oracle Database)

  Business Analytics (Example: Splunk Datasets Add-on)

  IoT & Industrial Data (Example: Machine Learning Toolkit)

  Utilities (Example: Splunk Add-on for Microsoft Windows)

# The Splunk REST API

- Exposes an API method for every feature in the product
- Whatever you can do in the UI – you can do through the API
- Index, Search, Visualize, Manage
- API is RESTful
- Endpoints are served by splunkd
- Requests are GET, POST, and DELETE HTTP methods
- Responses are Atom XML & JSON
- Versioning as of Splunk 5.0
- Search results can be output in CSV/JSON/XML

# SDKs Overview

- Stay true to the semantics of the particular language
- Provide implementation that feels natural to the developer
  - E.g. Project, build, IDE (where applicable) support
  - Cover REST API endpoints based on use cases of language
- Namespaces
  - *owner*: splunk username (defaults to current user)
  - *app*: app context (defaults to default app)
  - *sharing*: user | app | global | system

# Splunk has SDKs available for

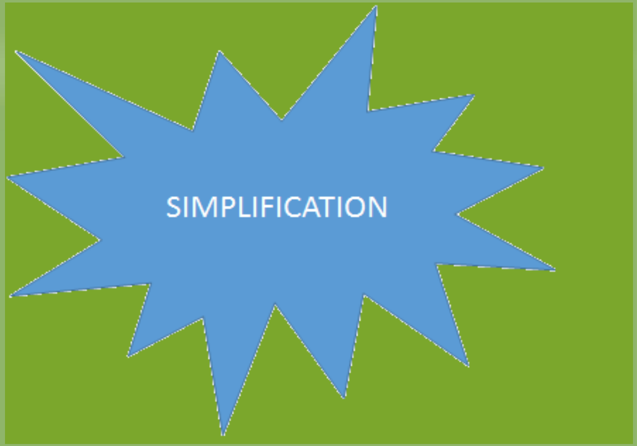
- Python
- Java
- Javascript
- PHP
- Ruby
- C#

# What can we do using the SDK

- Integrate with third party tools
- Log directly to Splunk
- Integrate search results into your application
- Extract data for archiving
- Build a UI on the web stack of  your own choice.

# What the Splunk SDKs do for you

- Handling HTTP access
- Authentication
- Managing namespaces
- Simplifying access to REST endpoint
- Building the correct URL for an endpoint
- Displaying simplified output for searches
- Over 160 endpoints that provide access to almost every feature of Splunk



SIMPLIFICATION

# How to Use SDK

- Connecting to Splunk using Java SDK and printing list of Users

```
// Create a map of arguments and add login parameters
ServiceArgs loginArgs = new ServiceArgs();
loginArgs.setUsername("admin");
loginArgs.setPassword("changeme");
loginArgs.setHost("localhost");
loginArgs.setPort(8089);
// Create a Service instance and log in with the argument map
Service service = Service.connect(loginArgs);
for (User user : service.getUsers().values())
        System.out.println(user.getName());
```

# Success Stories

## Challenges

•Difficulties monitoring impact of its Workforce Identity Access Management deployment on the business

•Problems prioritizing issues due to high volume of Remedy tickets caused by the new system

•Restricted ability to effectively map key performance indicators to critical service areas

•Lack of proactive service management

# Data Sources

- Application and DB logs
- Infrastructure metrics
- Network metrics
- Remedy
- Enabler services

# Business Impacts

- Glass table visualizations enable rapid and proactive issue resolution
- Custom KPIs empower teams across the business
- Proactive addressing of issues
- Improved visibility of open tickets, active status of tickets and number of impacted users

# Challenges

SurveyMonkey

- Needed a flexible way to drill down into site data
- Associate web activity with business results
- Reduce or eliminate multiple site analysis tools
- Better manage and integrate new acquisitions and products

# Data Sources

- Apache, clickstream logs
- Server, desktop, database and application activity logs
- Java applications and application servers
- .Net applications and servers
- System metrics

# Business Impacts

- Easier integration of data flows from acquired companies
- Streamlined foreign site expansion thanks to improved localized content and SEO optimization
- Increased ease and effectiveness of A/B site testing
- Reduced licensing costs by 45 percent
- Optimized site performance and resource allocation due to real-time error reporting and exception monitoring
- Improved user experience

# Challenges

- Previous business analytics solution was inflexible and unable to generate real-time insights
- Cumbersome manual analysis of data slowed down marketing efforts
- Lack of operational visibility
- Maintaining competitive advantage over local markets

# Data Sources

10 types of self-developed point-of-sale data:

- Product pricing
- Product categorization
- Product inventory
- Statistics about best sellers
- Seasonal trends

Promotional campaign data

CRM data

Sales tax data

Store financials

Employee work schedules

# Business Impact

- Real-time insights into business processes for better informed decisions
- Data analysis cycle reduced from days to minutes, leading to significant cost and time savings
- Lead time for promotional campaigns reduced by 80 percent
- Continued high level of customer service and optimized customer experience
- Operational resources freed up for greater overall productivity and efficiency

# Challenges

- Inability to get real-time data analysis
- Needed scalable solution for new mobile platform
- Required insights into customer behavior for strategic marketing planning

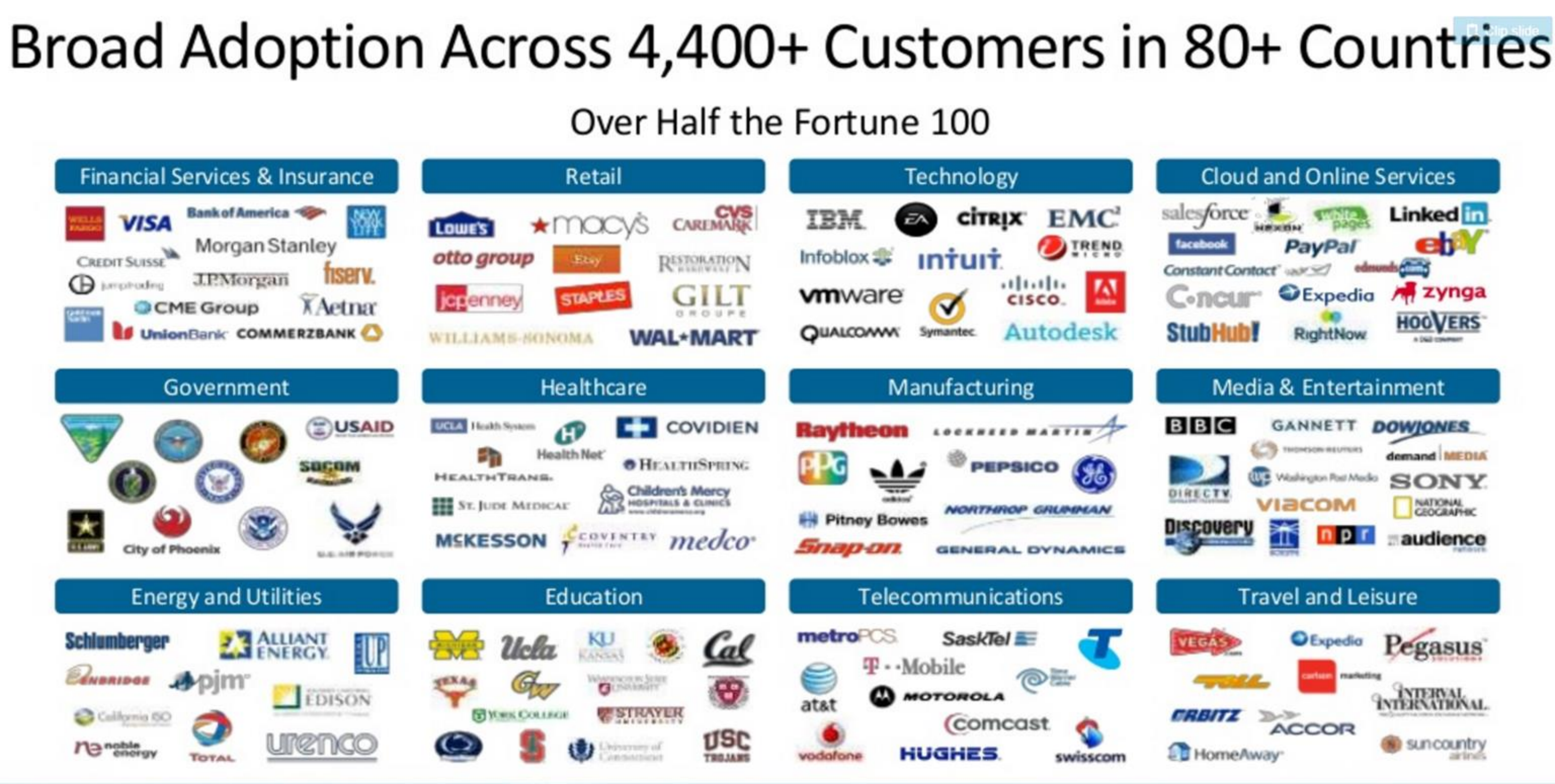
# Data Sources

- Online shopping/e-commerce web logs and web application server logs
- Shopping TV CTR log
- Mobile service web application logs
- Mobile device local application logs
- Internal lookup databases (products, customers)

# Business Impacts

- Improved operational efficiencies
- Integrated results from both web and mobile data sources
- ROI – cost savings of 50 percent over prior solution
- Time savings of 24 hours over previous weblogger data analysis solution
- Maximized marketing efforts from real-time insights into customer behavior
- Faster incident response times
- DevOps collaboration

# Popularity >



## Broad Adoption Across 4,400+ Customers in 80+ Countries

Over Half the Fortune 100

Financial Services & Insurance

Retail

Technology

Cloud and Online Services

Government

Healthcare

Manufacturing

Media & Entertainment

Energy and Utilities

Education

Telecommunications

Travel and Leisure

# THANK YOU!!!