



Splunking Big Data

Vadim Roussin

Regional Sales Director, Emerging Markets

March 22, 2012

AGENDA

- WHO WE ARE
- SPLUNK IN RUSSIA
- WHY DOES ANYONE (EVERYONE!) NEED A SPLUNK?
- WHAT ARE WE? OUTSIDE THE BOX (GARTNER BOXES)
- BIG DATA WITH SPLUNK
- INDUSTRY REGOGNITION
- CASE STUDIES
- CONCLUSION

WHO WE ARE

- Yesterday: garage startup founded in 2004
- Today:
 - Over 400 employees
 - Over \$100 million in revenue
 - 3500 customers in 84 countries (112% growth YOY)
- Tomorrow:
 - Public company
 - Next Google

2011 HIGHLIGHTS

- Company

- Achieved 175% year-over-year (Q2) revenue growth in EMEA
- Surpassed 430 employees, with eight offices worldwide

- Developers

- 52 new Splunkbase Apps – 166 total Apps now available for download
- New partner Apps for : Bluecoat, F5, Palo Alto, Citrix
- New partner-created Apps: Centrify Active Directory Integration for Splunk and Solera DeepSee App for Splunk

SPLUNK IN RUSSIA

- REPRESENTATIVE CUSTOMERS
 - YOTA
 - LUKOIL
 - ROSNEFT
- PROJECTS: financial services, manufacturing, telecom
- PARTNERS: three in Russia and Belarus and growing

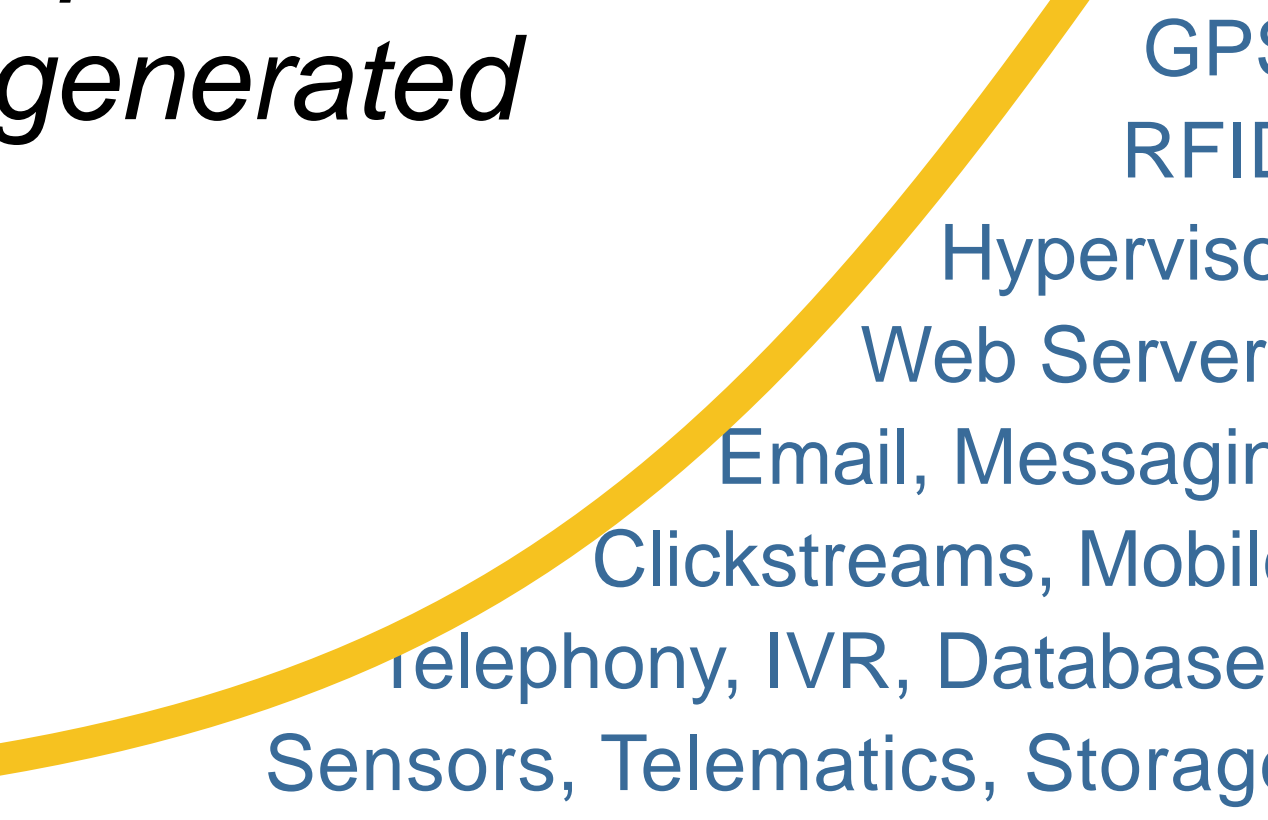
What is Big Data?

...When the size of the data itself becomes part of the problem...*

* Mike Loukides – O'Reilly Radar

World's Digital Data Growing Exponentially

Most enterprise data is machine-generated



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

“Between 2009 and 2020 digital data will grow 44x at a CAGR of 45%”
IDC 2010 Storage Market View

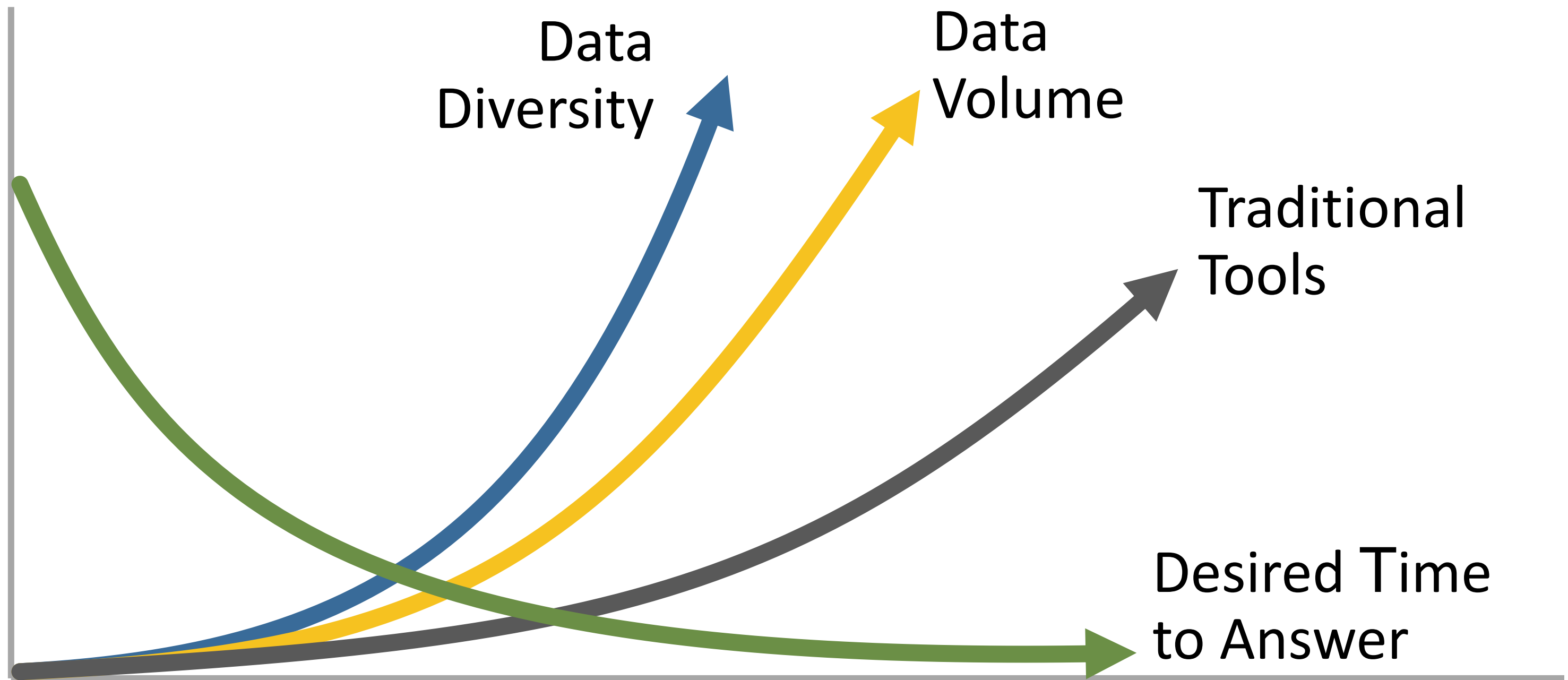
“Unstructured data accounts for more than 90% of the digital universe”
IDC 2011 Digital Universe Study: Extracting Value from Chaos

Most Enterprise Data is Machine-generated

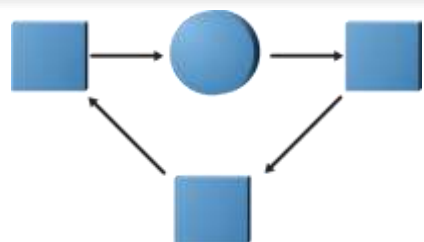
- Machine data is fastest growing, most complex, most valuable segment of Big Data
- Contains categorical record of all activity and behavior – customer behavior, user transactions, machine behavior
- **Value from data largely untapped – extremely difficult to process and analyze by traditional methods or in a timely manner**

Servers, Security devices, Desktops

It's Not Just About the Bigness

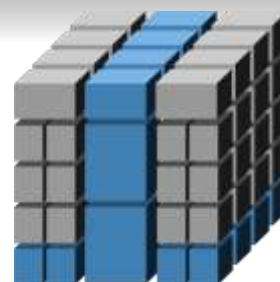


A Comparison of Technologies



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

What About Open Source Big Data Tools

Frameworks of API's and loosely coupled projects present significant issues

Logistics

- Requires senior engineering resources
- Time
- Commitment

Technology

- No easy way to get data in
- Requires some structure
- No easy way to get data out/visualize
- No enterprise integration w/ SSO, LDAP, etc.

People Issues

- Shortage of Talent
- Data Scientists
- Experts on non-traditional technologies



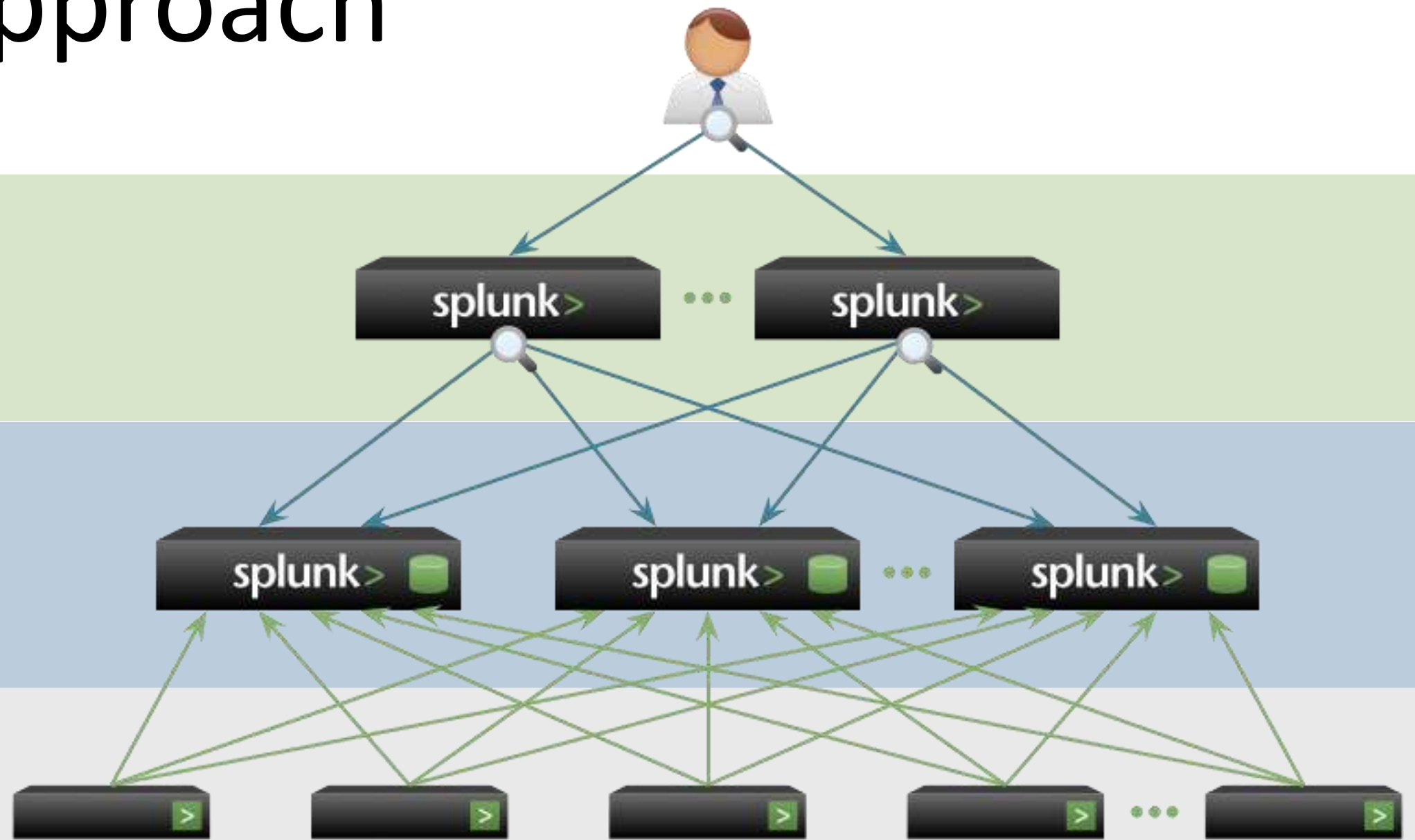
Make machine data accessible, usable
and valuable to everyone.

The Splunk Approach

Flexible Reporting
and Analysis

Real-time
Indexing

Universal
Collection



Rapid time-to-deploy: hours or days

splunk>

Splunk Collects and Indexes Any Machine Data

Customer Facing Data

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



Outside the Datacenter

- Manufacturing, logistics...
- 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 & 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

Splunk Collects and Indexes Any Machine Data

Customer Facing Data

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

Windows

- Registry
- Event logs
- File system
- sysinternals

Linux/Unix

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

Virtualization

- Hypervisor
- Cloud

Applications

- Web logs
- Log4J, JMS, JMX
- Code and scripts

Databases

- Configurations
- Audit/query logs
- Tables
- Schemas

Outside the Datacenter

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

Networking

- Configurations
- syslog
- SNMP
- netflow



Any amount, any location, any source.

Logfiles Configs Messages Traps Metrics Scripts Changes Tickets



No upfront schema



No custom connectors

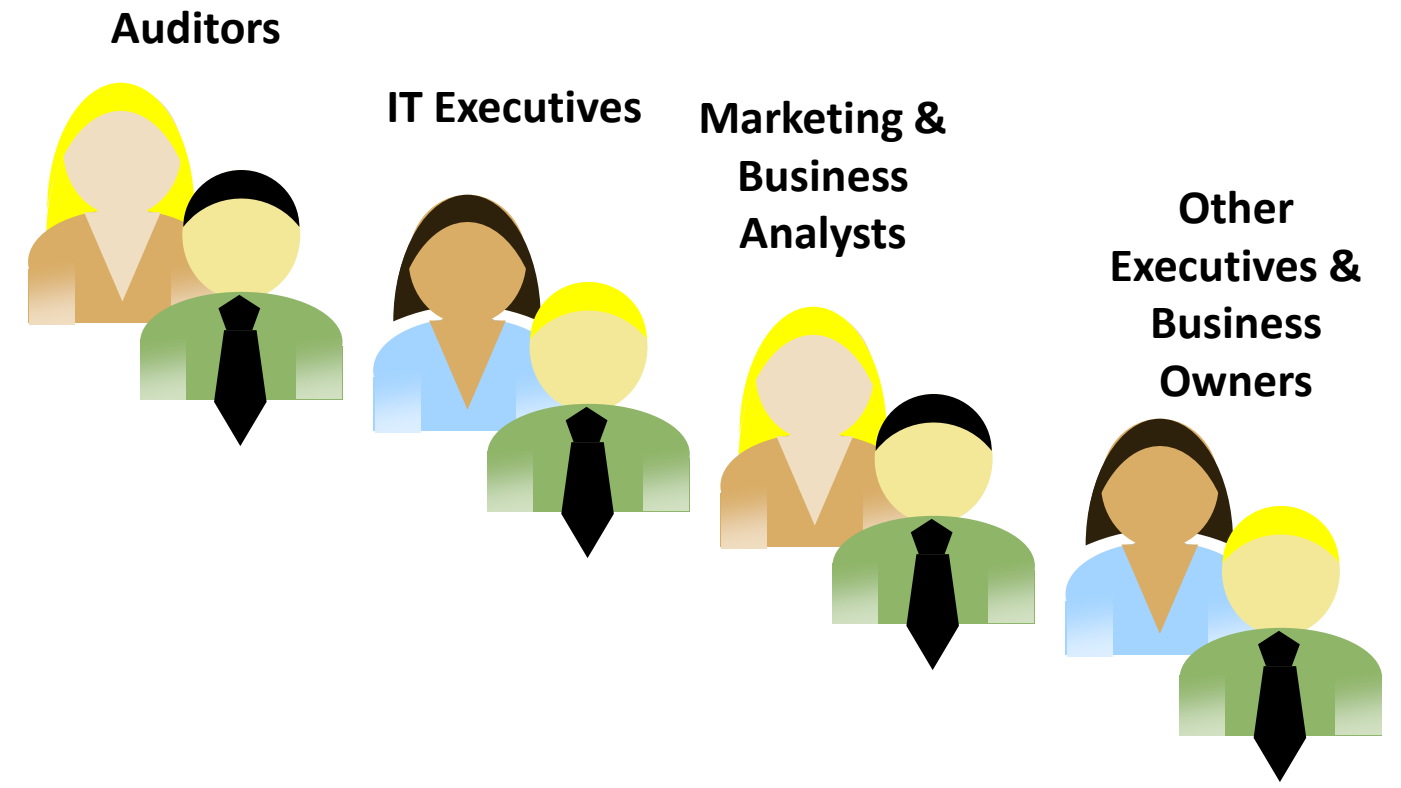
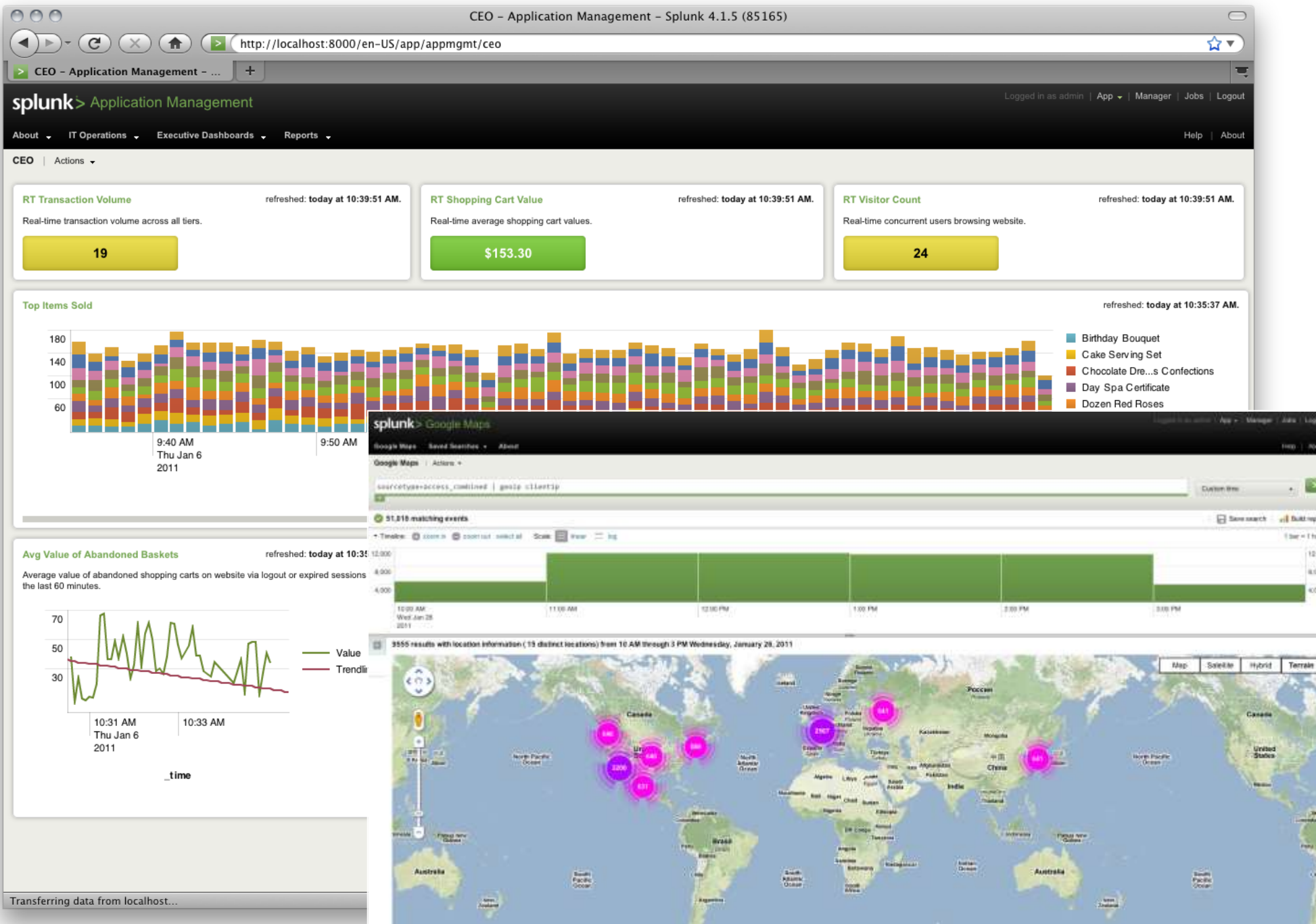


No RDBMS



No need to filter/forward

Create and Share Dashboards in Minutes

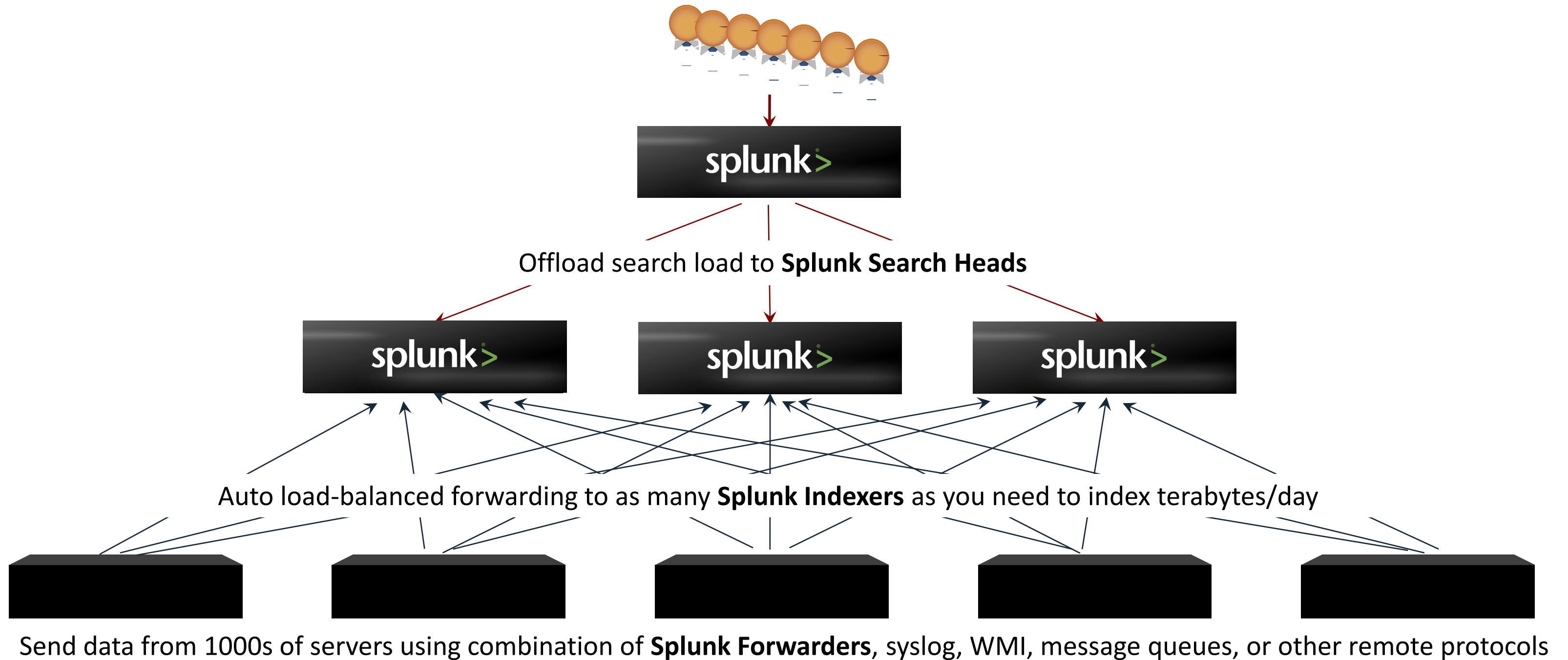


Deliver new levels of visibility and insight for IT and the business from operational data

End-to-end Integrated Solution



Massive Linear Scalability to Tens of TBs/Day

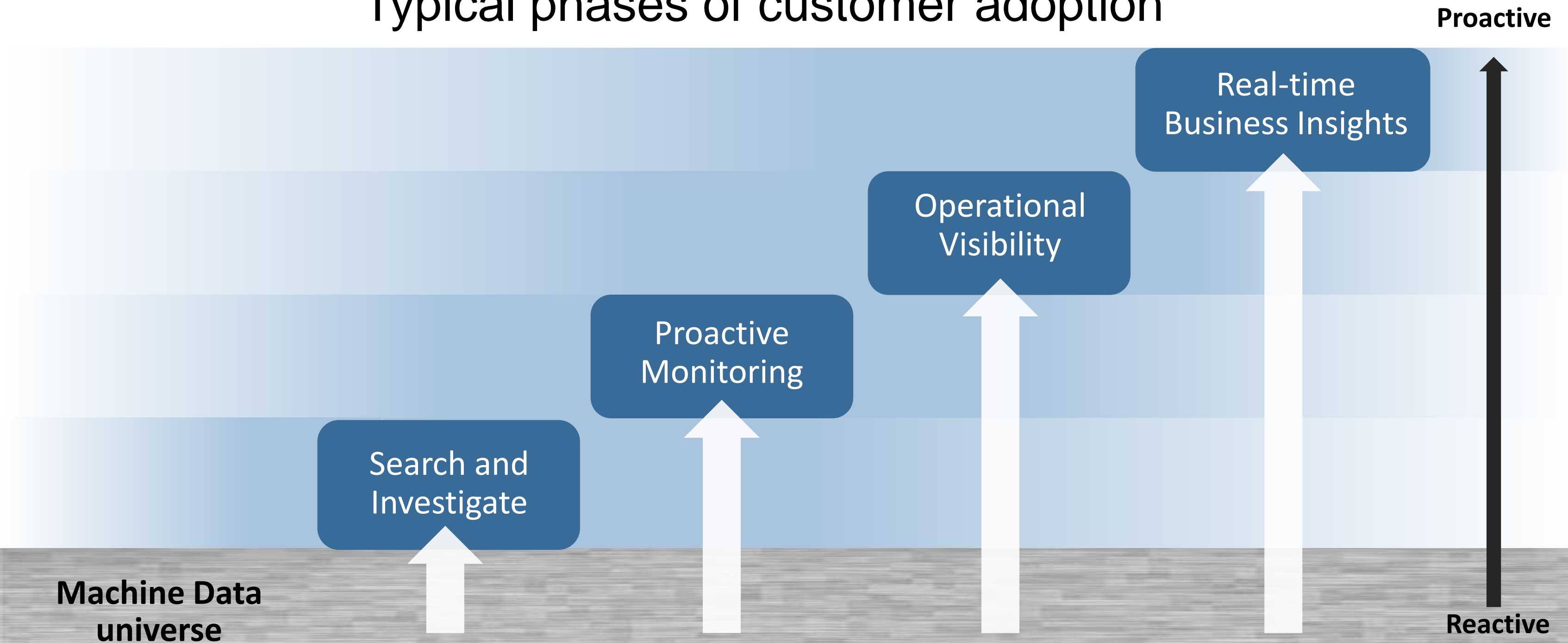


Splunk is Used Across IT and the Business








Providing Operational Intelligence from Machine Data

Typical phases of customer adoption



Splunking Big Data

Customer	Data Volume (per day)
Leading Social Gaming Company	12 TB
	6 TB
	4 TB
	1.2 TB
	900 GB
	800 GB

What Makes Splunk Unique?

- Integrated, end-to-end solution
- Rapid time-to-value
- Real-time and historical analysis in one system
- Secure, role-based access controls
- Scales efficiently to any data volume using commodity hardware
- Proven with over 2,900 enterprise customers



Over 3500 enterprise customers use Splunk to gain better insight and visibility from their machine data. Why?



Operational Intelligence Across the Business

“ We have taken application performance troubleshooting for 97,000 customers to the next level. ”

“ The fact that we had a data treasure chest was not obvious till Splunk came in to the picture. ”



- Provided higher service levels
- Providing usage metrics and analytics for their business
- Now offering new services: reporting on customer email campaigns



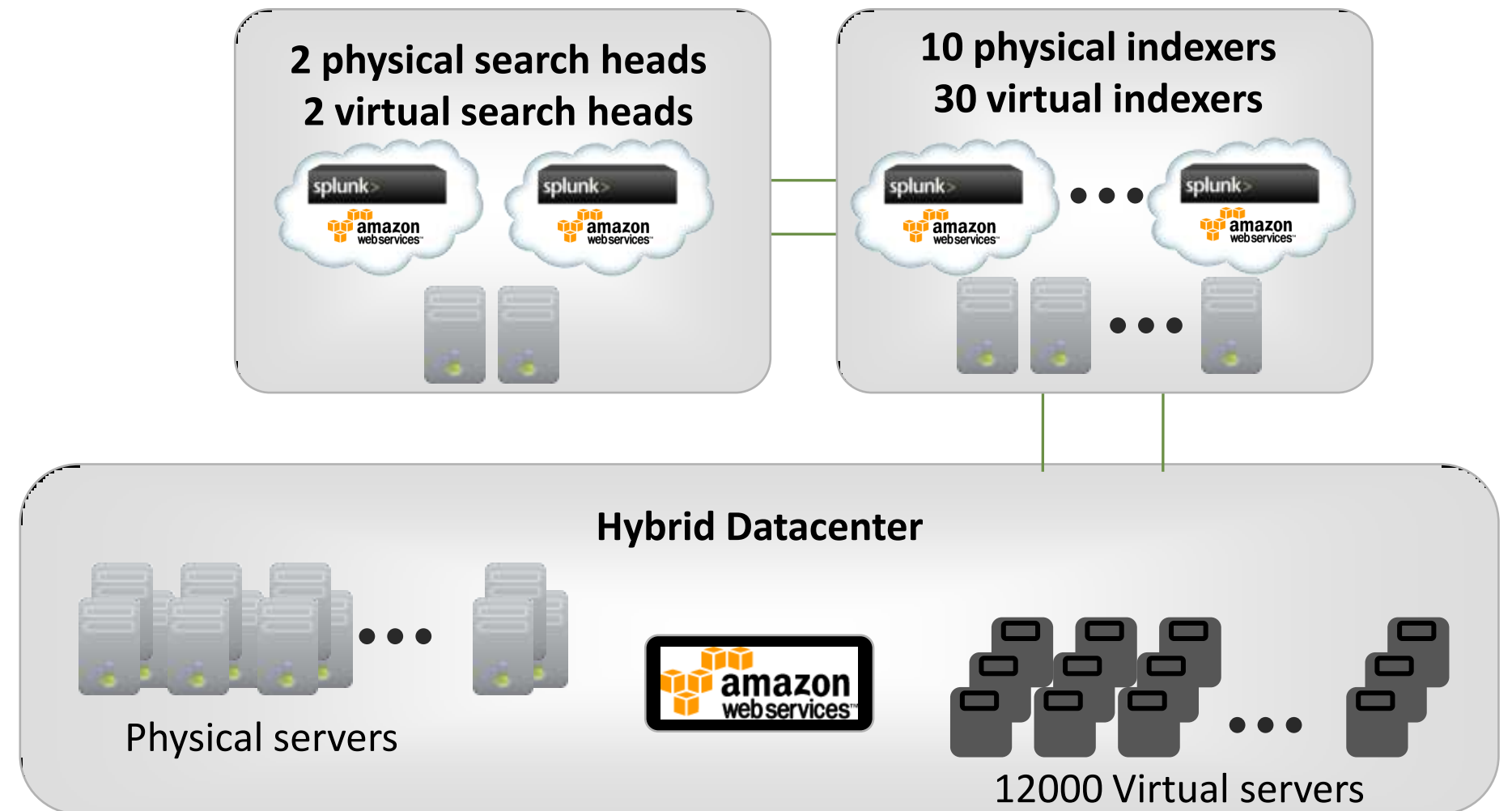
Narayan Bharadwaj
Director, Product Management

Social Gaming Usage and Performance

“Splunk indexes up to 12TB in a single day and helps us manage our enormous scaled out physical and cloud infrastructure.”

Architect

Social Gaming Company

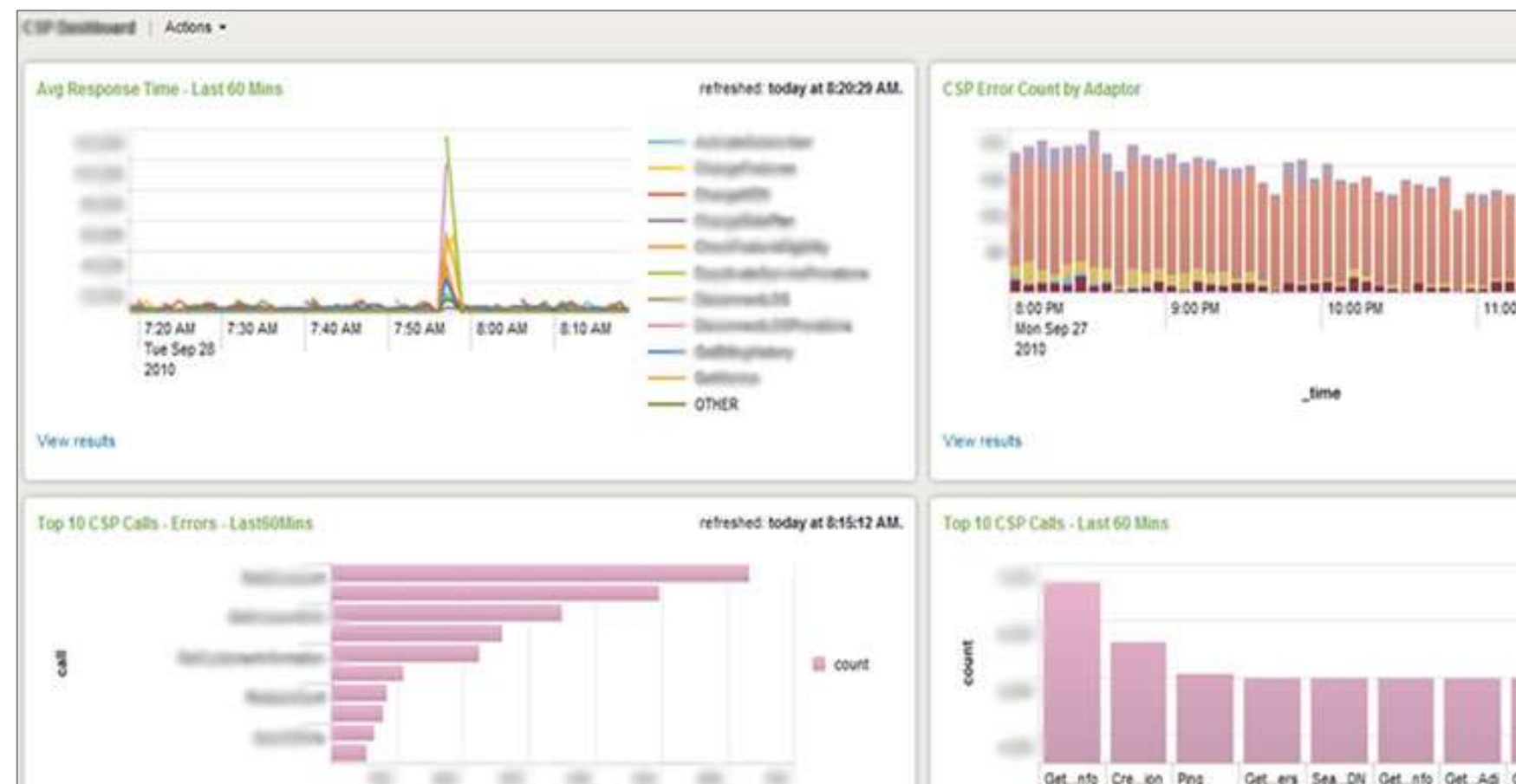


- Splunk used for managing many thousands on Amazon instances as well as physical infrastructure
- Used for troubleshooting, security, compliance infrastructure management and analytics

Business Analytics with Real-time Dashboards

“I built a business analytics dashboard for my manager in 5 minutes and he was sold.”

“Splunk lets us build dashboards to compare and correlate whatever we want—nothing else lets us do that.”



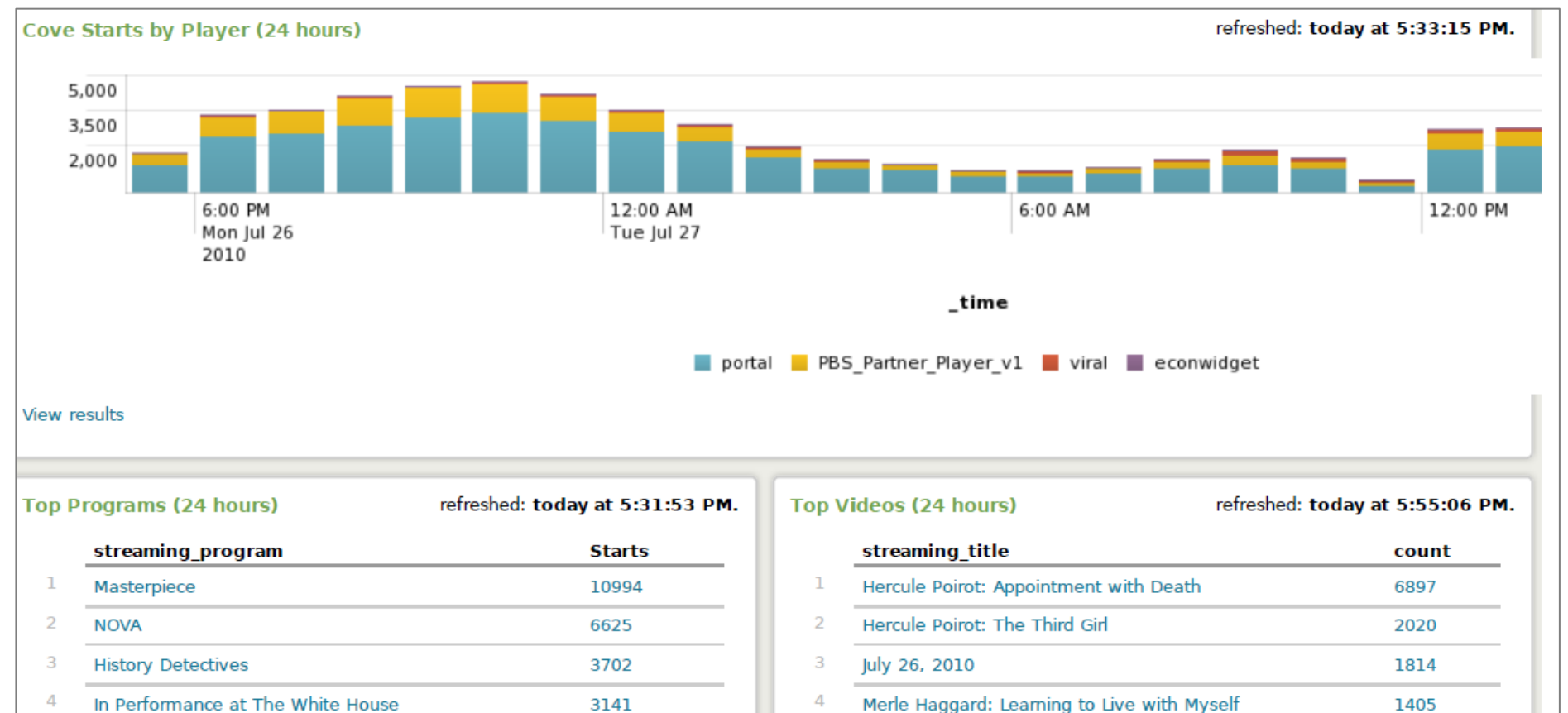
- Correlated F5, firewalls and malware for complete security posture
- Informed capacity planning
- Delivered executive dashboards look at activations by minute, by channel, by market



Roberto Quezada
IT Operations Analyst

New Web Intelligence Using Splunk

“After 6 months of effort with other products, only Splunk gave us the business reports about our web-based digital assets that we needed.”



Sondra Russell
Online Metrics Analyst

- Programming popularity
- Royalties
- Abandonment rates
- Views by player
- Errors

Telecoms Call Revenue and Cost Analysis

Business Systems



Tariffs Database

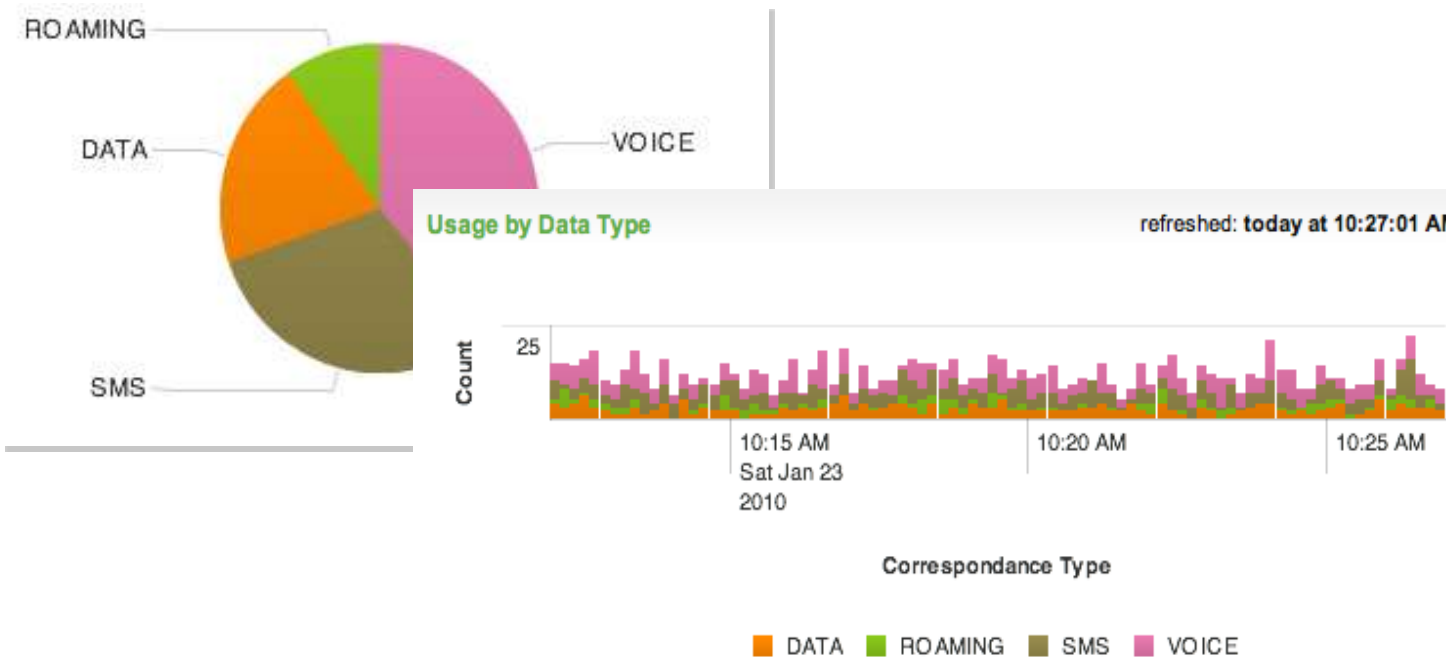
Operational Systems



Call Detail Records



Real-time visibility of revenue and service metrics



Lowest-cost providers per call

Call ID	Date/Time	ISDN	Switch	Carrier	Location	Price	Total Cost
1	1/23/10 2010/01/23 10:32:46	000630	558777090829965	2070090608	66924843904	0.2250	
2	1/23/10 2010/01/23 10:32:46	002746	019324412996455	2349997681	30027096900	0.1870	

Recent Press and Analyst Quotes

“The result of Splunk’s efforts is that even web developers can use its products to extract meaningful business insights from machine-generated data”

“Splunk is ahead of the game when it comes to democratizing big data”



Derrick Harris

Jun 8 2011

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



David Menninger

VP & Research Director



Thank You!
Questions?