splunk> listen to your data™

SOLUTIONS GUIDE

The Splunk Guide to Operational Intelligence

Making Machine-generated Data Accessible, Usable and Valuable to Everyone

Most Challenging, Fastest Growing Segment of Big Data

"Big Data" is a class of information processing problem that, due to the volume, velocity, variety and complexity of the data, requires different approaches to support analytics to derive cost-effective, timely, business-relevant insight. (Gartner June 2012)

"Big data" is a term applied to data sets whose size is beyond the ability of commonly used software tools to capture, manage, and process the data within a tolerable elapsed time. Big data sizes are a constantly moving target, as of 2012 ranging from a few dozen terabytes to many petabytes of data in a single data set.

Examples include web logs, RFID, sensor networks, social networks, social data, Internet text and documents, Internet search indexing, call detail records, astronomy, atmospheric science, genomics, biogeochemical, biological, and other complex and often interdisciplinary scientific research, military surveillance, medical records, photography archives, video archives, and large-scale e-commerce.

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

The Machine Data Opportunity

All your IT applications, systems and infrastructure generate data every millisecond of every day. This **machine data** contains a definitive record of user transactions, customer behavior, machine behavior, security threats, fraudulent activity and more. It's also dynamic, unstructured and non-standard and makes up the majority of the data in your organization.

Machine data is an incredibly valuable resource, but organizations rarely get the value they need from it. Existing data analysis, management and monitoring solutions are simply not engineered for this type of data.

Take information management. Data warehouses and relational database management systems are based on rigid schemas and designed for structured, consistent data. They provide historical analysis but not real-time visibility. Enterprise Search is designed for human-generated data, such as documents and Web pages. This data is very different than machine data. Machine data has an order of magnitude greater volume and diversity than traditional, structured data.

IT management tools and security information and event management on the other hand are siloed and designed for one level of the organization. They provide a narrow view of the underlying data and are hard-wired for specific data types and sources. Or they monitor across systems, with serious gaps in the data they collect. They also don't provide historical context.

The fact is finding a better way to sift, distill and understand the vast amounts of machine data can transform how IT organizations manage, secure and audit IT. It can also provide valuable insights for the business on trends and behaviors of their customers and services.

While most companies don't realize it, machine data is the fastest growing, most complex and yet most valuable segment of big data. All websites, communications, networking and complex IT infrastructures generate massive streams of data every second of every day, in an array of unpredictable formats that are difficult to process and analyze by traditional methods or in a timely manner.

What is Machine Data?

- Machine-generated data is one of the fastest growing, most complex and most valuable segments of big data.
- It's all of the data generated by the technology infrastructure and systems that power the enterprise. This includes live data from packaged and custom applications, app servers, web servers, databases, networks, virtual machines, telecoms equipment, OS's and more.
- Every activity leaves a trace in machine data. The data contains a definitive record of your customer behavior, use transactions, application behavior, service levels and so on.

The Machine Data Challenge

Machine data holds critical operational insights into user behavior, security risks, capacity consumption, service levels, fraudulent activity, customer experience and much more.

Making use of this data however, presents real challenges:

- Machine data is generated by a multitude of disparate sources; correlating meaningful events across these is complex
- The data is unstructured and difficult to fit into a pre-defined schema
- Machine data is high-volume and time-series based requiring new approaches for management and analysis
- The most valuable insights from this data are often needed in real time

Existing business intelligence and data warehouse solutions are simply not engineered for this type of high-volume, dynamic and unstructured data. Emerging open source technologies can provide part of the answer, but typically require extensive and time-consuming integration with other open source projects.

Today's agile enterprises can't wait. Key stakeholders across the organization need to keep pace and adapt to rapidly changing business environments. They need a technology that supports real-time data discovery, ad hoc reports and rapid analysis. A solution that can give them answers as fast as they think of questions.

Operational Intelligence in Action: Using Machine Data to Instrument the Enterprise

What is Operational Intelligence?

Operational intelligence refers to a category of methods and technology for gaining visibility into your business and discovering insights for IT as well as for your entire enterprise.

Operational intelligence is not an outgrowth of business intelligence (BI) but a new approach based on sources of information not typically in the purview of BI solutions. Behind every IT infrastructure, behind the systems that run your business, are massively growing streams of machine-generated data. Leading organizations realize that this data can be incredibly valuable for better running not only IT but also other parts of the business. Operational intelligence is designed to specifically address this opportunity.

Why does it Matter?

Operational intelligence enables organizations to:

- Gain a deeper understanding using all relevant information, especially from machine data
- Reveal important patterns and analytics derived from correlating events from many sources
- Reduce the time between an important event and its detection
- Leverage live feeds and historical data to make sense of what is happening now, to find trends and anomalies, and to make more effective decisions based on that information
- Quickly deploy a solution and deliver the flexibility needed by organizations today and in the future—the ability to provide ad hoc reports, answer questions, and add new data sources

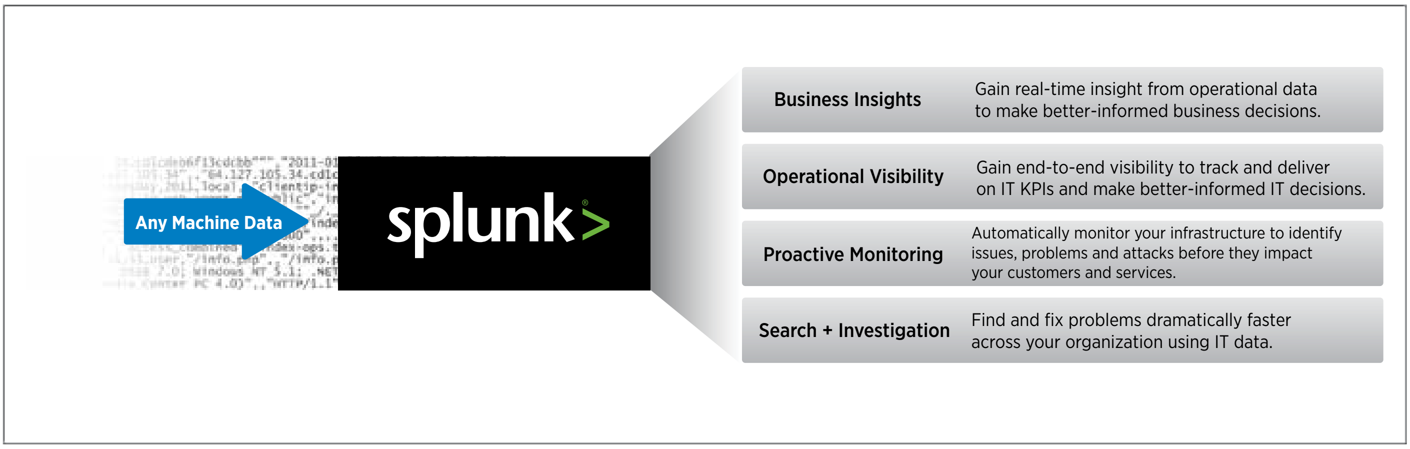
Splunking Machine Data

Splunk is the engine for machine data. It was developed to solve the whole machine data challenge and collects, indexes and harnesses your unstructured, time-series machine data. Splunk can read data from just about any source imaginable, such as network traffic, Web servers, custom applications, application servers, hypervisors, GPS systems, stock market feeds, social media and preexisting structured databases. It delivers a real-time understanding of what's happening and deep analysis of what's happened across your IT systems and infrastructure. It turns your machine data into the insights you need to make informed decisions.

What makes Splunk different?

Everything in one solution. Splunk is an integrated, end-to-end solution. It collects and organizes your machine data in one place. Once in Splunk you can search, browse, navigate, correlate, analyze and visualize your data in real-time. The goal is to make it easy - simply point Splunk at your data and start using it immediately.

Ease of deployment, ease of use. One person can download and implement Splunk in hours, rather than having a team of people take months or even years to deploy a solution. You can connect to your data in a few clicks and create powerful dashboards with a few more.


Any Machine Data
splunk>
Business Insights
Gain real-time insight from operational data to make better-informed business decisions.
Operational Visibility
Gain end-to-end visibility to track and deliver on IT KPIs and make better-informed IT decisions.
Proactive Monitoring
Automatically monitor your infrastructure to identify issues, problems and attacks before they impact your customers and services.
Search + Investigation
Find and fix problems dramatically faster across your organization using IT data.

Real-time and historical analysis. Search and analyze live streaming and terabytes of historically indexed data from one place. Splunk automatically monitors your data for trends and specific patterns of activity or behavior. Then notifies the people that need to know immediately.

Sophisticated search and visualization. Powerful search, drilldown and reporting capabilities meet the needs of novice users and expert analysts alike. Easy-to-create dashboards put critical insights from your machine data into the hands of the people who need it.

Proven scale on commodity hardware. Download and run Splunk on a single server in under 5 minutes. The same software can be scaled out across the largest global infrastructures, indexing tens of terabytes of data a day.

Role-based security and data signing. Underlying everything Splunk does is a robust security model, providing secure data handling, role-based access controls, auditability and assurance of data integrity.

Developing on Splunk

Splunk's developer platform provides open APIs that enable developers to plug into Splunk's map/reduce data processing pipeline, storage technology and management facilities. More information: http://dev.splunk.com.

What About Emerging Big Data Technologies?

Emerging big data technologies underscore the growing awareness and interest in solving the Big Data challenge. Splunk has been solving "big data" problems since our inception in 2004, so we understand the value that enterprises and organizations can gain from harnessing the tremendous value contained in their data.

Address storage only. Technologies such as Hadoop, Cassandra and HBase, address the storage problem only. They require integration to other "projects" to provide a complete solution, typically lots of engineering/consulting resources and subsequently long, multi-month/year implementation cycles.

They are batch-based. Hadoop-based approaches are typically batch. Splunk combines the ability to analyze patterns across real-time and historical data. We look at machine data as big data with a purpose, so have built an end-to-end solution accordingly.

Does Splunk integrate to Hadoop? We announced an integration to Hadoop and recently completed our first Beta phase. We are working on the next Beta and you can register for more information on our website here: www.splunk.com/bigdata.

Why are you integrating with Hadoop? Splunk Enterprise is widely deployed and proven. The new offering will bring the real-time enterprise capabilities and ease-of-use of Splunk Enterprise to the Hadoop ecosystem. In turn, Splunk users will be able to leverage Hadoop for archiving and specialized batch analytical processing.

Powerful benefits derived from Splunk's approach include:

- Dramatically better service levels — MTTR, reduced downtime
- Dramatically better productivity — accomplish in minutes what used to take hours or days
- Better manage agile infrastructures you're deploying — VMs, private clouds, public cloud
- Better manage risk — detect and investigate security incidents; demonstrate compliance
- Delivering new intelligence for IT and the business.

Customer Success with Splunk

With thousands of licensed customers, many Splunking terabytes of data per day, our users are the best example of massive machine data in action.

Expedia

Expedia, the world's largest online travel company, initially used Splunk to avoid website outages, saving them millions of dollars in lost revenue. They quickly expanded their use of Splunk and within 10 months were monitoring 98% of their infrastructure. Today, over 2,700 users at Expedia use Splunk to gain real-time insights of not only their IT infrastructure, but also online bookings, performance of air-travel coupons and optimizing SEM.

"Splunk provides real-time visibility and insights across a wide range of critical areas from server and application health and performance monitoring to bookings trends, coupon use and deal analysis. It's where we go first to perform rapid real-time analysis on tens of terabytes of unstructured, time-sensitive machine data."

Eddie Satterly, Sr. Director Infrastructure Architecture & Engineering, Global Infrastructure Systems

Salesforce.com

Salesforce.com, the industry-leading enterprise cloud computing company, uses Splunk to mine the large quantities of data generated from across its entire technology stack. Salesforce.com has over 500 users of Splunk dashboards from IT users monitoring customer experience to product managers performing analytics on new services like 'Chatter.'

"The fact that we had a data treasure chest was not obvious until Splunk came in to the picture. With Splunk, we have taken application troubleshooting for 97,000 customers to the next level. Splunk has augmented our ability to make data-driven decisions."

Narayan Bharadwaj, Director Product Management, Salesforce.com

NPR

NPR, the award winning, multimedia news organization reaching 26.8 million listeners per week, uses Splunk to gain better visibility and insight of their digital asset infrastructure.

NPR initially used Splunk to monitor and troubleshoot their end-to-end asset delivery infrastructure. Before Splunk, there were critical business metrics they couldn't get from their traditional web analytics solutions. They expanded their deployment of Splunk and now measure program popularity, views by device, reconcile royalty payments for digital rights, measure abandonment rates and more.

"Only Splunk easily gives us the business reports about our web-based digital assets that we need."

Sondra Russel, Online Metrics Analyst

Pegasus Solutions

Pegasus Solutions, a major power behind the travel and hospitality industry, caters to hundreds of thousands of hotels, websites and travel agencies and processes 4-5 billion transactions per month.

Pegasus uses Splunk to gather real-time insights from their operational data. Results from using Splunk include reduced escalations and troubleshooting, accelerated response to customer inquiries and unparalleled insights on the health of their system and business.

"Splunk scales to give us real-time monitoring as well as deep historical trend analysis across 50+ systems and 2 billion transactions a month. It is amazingly flexible—we get deep, detailed information and high-level health metrics—all from the same set of data."

Peter Elhke, Principal Systems Engineer, Pegasus Solutions

MetroPCS

MetroPCS, the fifth largest facilities-based wireless carrier in the United States with nearly 10 million subscribers, deployed Splunk to provide better intelligence on wireless usage, revenue per user and optimize their partner carrier charges. Since their customers were charged a flat rate fee per month, determining the lowest-cost route for calls, translated directly to higher margins and profits.

"With an implementation time of a few weeks, Splunk surpassed all ROI expectations and has delivered millions of dollars in savings."

Gregg Woodcock, Manager Corporate Engineering, MetroPCS

Free Download

Download Splunk for free. You'll get a Splunk Enterprise license for 60 days and you can index up to 500 megabytes of data per day. After 60 days, or anytime before then, you can convert to a perpetual Free license or purchase an Enterprise license by contacting apac_sales@splunk.com.

CONTACT SPLUNK

APAC Headquarters

Unit 601, 6/F., Miramar Tower,
132 Nathan Road, Tsim Sha Tsui,
Kowloon, Hong Kong

Phone: +852 3975 4000
Fax: +852 3975 4001

apac_sales@splunk.com

Australia and New Zealand

Level 27, 101 Collins Street
Melbourne VIC 3000, Australia

Phone: +61 03.9221.6206
Fax: +61 03.9221.6208

splunkanz@splunk.com

Japan

Level 21 Shiodome, Shibarikyu Building,
1-2-3 Kaigan Minato-ku,
Tokyo 105-0022 Japan

Phone: +81 (03) 5403 6500

splunkjp@splunk.com

People Republic of China

Level 19 Tower E2, Oriental Plaza 1,
East Chang An Avenue,
Dong Cheng District,
Beijing 100738 China

Phone: +86 (010) 8520 0536
Fax: +86 (010) 8520 0220

splunkchina@splunk.com

South East Asia and India

Level 18 Republic Plaza II,
9 Raffles Place,
Singapore 048619

Phone: +65 6823.6864
Fax: +65 6823.6866

splunksea@splunk.com

Taiwan

6F, No 6, Sec 4, Hsinyi Road,
Da-an District, Taipei City
106, Taiwan

Phone: +886 (02) 5551 1266
Fax: +886 (02) 2703 6007

splunktw@splunk.com

splunk® > listen to your data™

www.splunk.com