splunk>

SOLUTIONS GUIDE

The Splunk Guide to Operational Intelligence

Turn Machine-Generated Data Into Real-Time Visibility, Insight and Intelligence

What is Splunk® Enterprise?

Splunk Enterprise is the leading platform for real-time Operational Intelligence. It's the easy, fast and secure way to search, analyze and visualize the massive streams of machine data generated by your IT systems and technology infrastructure—physical, virtual and in the cloud. Use Splunk Enterprise and your machine data to deliver new levels of visibility, insight and intelligence for IT and the business. Splunk Enterprise is available as software or via the Splunk Cloud™ software as a service (SaaS) offering.

The Machine Data Opportunity

All your IT applications, systems and technology infrastructure generate data every millisecond of every day. This machine data is one of the fastest growing and most complex areas of big data. It's also one of the most valuable, containing a definitive record of all user transactions, customer behavior, sensor activity, machine behavior, security threats, fraudulent activity and more.

Making use of this data, however, presents real challenges. Traditional data analysis, management and monitoring solutions are simply not engineered for this high volume, high velocity and highly diverse data.

Consider traditional information management systems, such as business intelligence and data warehouse tools. These systems are batch-oriented and designed for structured data with rigid schemas. IT management and security information and event management (SIEM) tools, on the other hand, provide a very narrow view of the underlying data and are hard-wired for specific data types and sources. They also don't provide historical context.

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 how to innovate and offer new services, as well as trends and customer behaviors.

The Splunk Approach

Splunk Enterprise is the first enterprise-class platform that collects and indexes any machine data—whether it's from physical, virtual or cloud environments. Splunk Enterprise can read data from virtually any source, such as network traffic or wire data, web servers, custom applications, application servers, hypervisors, GPS systems, stock market feeds, social media, sensors and preexisting structured databases. It gives you a real-time understanding of what's happening and deep analysis of what's happened across your IT systems and technology infrastructure, so you can make informed decisions.

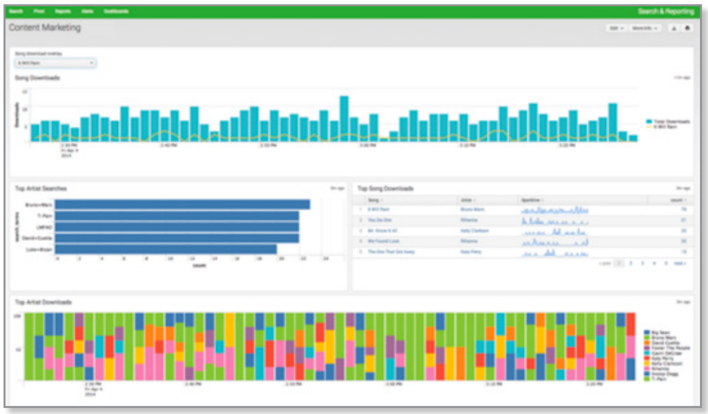
Splunk Enterprise has many critical uses across IT and the business:

Security, Compliance and Fraud: provide rapid incident response, real-time correlation and in-depth monitoring across data sources; conduct statistical analysis for advance pattern detection and threat defense

Infrastructure and Operations Management: proactively monitor across IT silos to ensure uptime; rapidly pinpoint and resolve problems; identify infrastructure service relationships, establish baselines and create analytics to report on SLAs or track SLAs of service providers

Application Delivery: accelerate and empower developers, QA and operations; provide end-to-end visibility across distributed applications and infrastructures; troubleshoot and isolate problems across application environments; measure service levels and application performance; gain insight on mobile app performance

Internet of Things: collect data from devices, control systems, sensors, SCADA, mobile and hand-held point of sale devices, and more to monitor operations, analyze usage, and integrate these insights into an end-to-end view of your business operations



Create powerful, interactive dashboards.

Business Analytics: gain visibility and intelligence on customers, services and transactions; gain end-to-end visibility into business processes; identify trends and patterns in real time; fully understand the impact of new product features on adoption and usage; gain a deeper understanding of user experience to prevent drop-offs and improve customer satisfaction, conversions and online revenues

Finding and fixing problems, following the trail of an attacker, reporting for compliance, analyzing service usage and customer behavior requires a complete view. Splunk Enterprise analyzes machine-generated data to arm network engineers, system administrators, security and compliance analysts, developers, support/service desk staff and business users alike with new levels of visibility, analysis and insight. This is called Operational Intelligence.

How is Splunk Enterprise Different?

Splunk Enterprise is different from previous approaches to managing, auditing, securing and gathering intelligence from IT systems and technology infrastructure. Here's how:

Immediate results without the risk. Users can download Splunk Enterprise for free, install it in minutes, use a simple wizard to onboard data and immediately get productive. No more armies of consultants or DBAs to make it work. Most users download and install Splunk Enterprise while they're under fire and the proof is immediate. A serious service problem or security incident can now be investigated in minutes, versus the hours or days it used to take.

Based on high-performance indexing and search technology. Every day over a billion people search and navigate web pages served all over the world. Search is flexible, intuitive and delivers immediate results. At its core, Splunk Enterprise has powerful indexing and search technology, bringing a whole new meaning to speed and responsiveness. Search billions of events in seconds and start seeing results immediately.

Leverage powerful analytics. Advanced features enable technical users to explore and interact with their machine data with a powerful user interface and Search Processing Language (SPL™) designed to search, correlate and visualize data. Business users can gain rapid insights using a simple drag-and-drop interface to analyze data without learning the search language. Pattern detection, instant pivot and an advanced field extractor makes it easy for everyone in your organization—including non-specialists—to turn machine data into powerful insights.

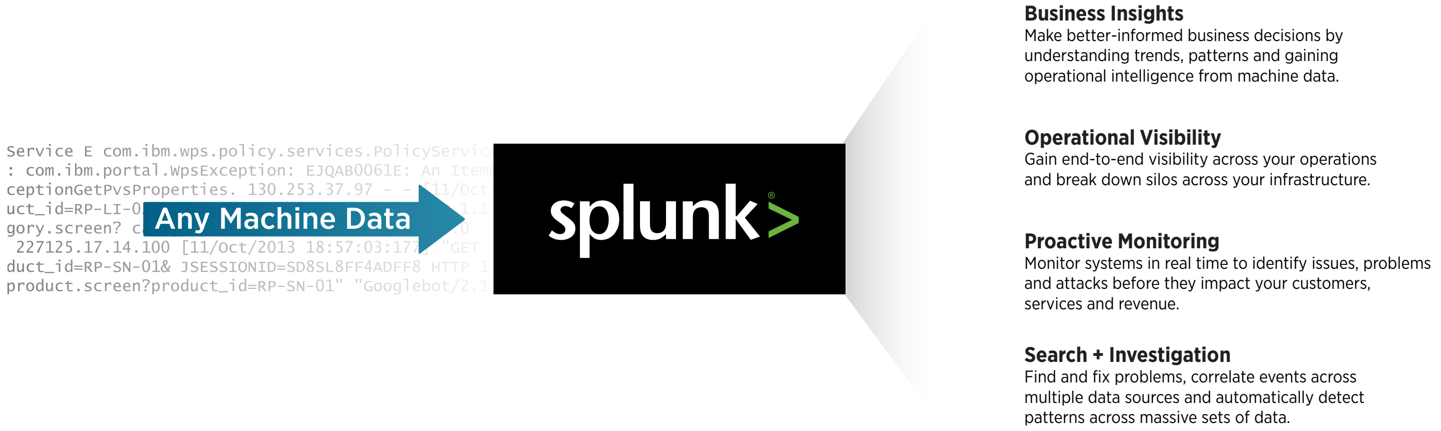
Collect and index any machine data. Machine data is high volume, high velocity, highly variable and incredibly diverse. It contains all time-stamped events generated by machine-to-machine and human-to-machine interactions. The traditional set of tools—system management, SIEM, CEP/ECA and log management—require weeks or months to develop and to configure custom connectors for each data source. Splunk Enterprise collects and indexes any machine data from virtually any source, format or location in real time. This includes data streaming from packaged and custom applications, app servers, web servers, databases, wire data from networks, virtual machines, telecoms equipment, operating systems, sensors and much more. There's no requirement to "understand" the data upfront. You can use a simple wizard to load data into Splunk Enterprise or deploy forwarders to reliably stream data from remote systems at scale. Splunk Enterprise immediately starts collecting and indexing, so you can start searching and analyzing.

Analyze real-time and historical data. Traditional IT systems force a decision between real-time monitoring and historical analysis. With Splunk Enterprise, you can search and analyze real-time streaming and terabytes of historical data from one place. This means you can identify and respond to patterns of behavior or activities of interest immediately. Most data management projects are designed to answer a pre-set list of questions and must fit into a brittle schema. Indexed data in Splunk Enterprise doesn't have these limitations because the schema is applied at the time of search—so users can immediately ask new questions while they search. Any interpretation of the data you need, such as extracting a field or tagging a subset of hosts, can be done easily—as you search.

Create custom dashboards and views. You need to make sense of huge volumes of machine data and satisfy the needs of the different users and groups in your organization. With Splunk Enterprise, you can quickly create custom dashboards that integrate multiple charts and views of your real-time data and access them from your desktop or mobile device. Personalize dashboards for different users in your organization—managers, business and security analysts, auditors, developers and operations teams. Users can edit dashboards using a simple drag-and-drop interface and use integrated charting controls to change chart types or views dynamically. Dashboard panels can be saved independently and be published to a panel library, so insights can be shared across teams.

Software that people want to use. It used to make sense to manage your IT infrastructure in silos. But with today's distributed, virtualized and cloud-based environments, this just doesn't work anymore. Splunk Enterprise breaks down the IT silos. Search, report, monitor and analyze all your data from every application, server and device—physical, virtual or in the cloud. Easily integrate with existing enterprise management, security and compliance tools, or view insights from Splunk Enterprise in common business applications like Salesforce or Microsoft SharePoint. For cloud deployments, Splunk Cloud is available* and includes all of the features of Splunk Enterprise as a cloud service. Splunk Enterprise AMIs are available for Amazon Web Services (AWS) environments. Finding and fixing problems, following the trail of an attacker, tracing transactions and gaining new insights from your operational data is suddenly orders of magnitude faster and easier.

Do more with Splunk Apps. Take advantage of hundreds of apps that run on top of the Splunk Enterprise platform. Apps deliver a targeted user experience for different roles and use cases. There are a growing number of apps, built by our community, partners and Splunk. These apps help you visualize data geographically or provide pre-defined compliance views for your mission-critical technologies such as VMware®, Microsoft Exchange®, Microsoft Active Directory®, Cisco® and Citrix. There are apps for different technologies such as Windows®, Linux®, Unix, virtualization, networking technologies and more. View and download apps and add-ons at Splunkbase.



```
Service E com.ibm.wps.policy.services.PolicyServic
: com.ibm.portal.WpsException: EJQAB0061E: An Item
ceptionGetPvsProperties. 130.253.37.97 - - [11/Oct
uct_id=RP-LI-0                                  1.1
gory.screen? c                                   .0
 227125.17.14.100 [11/Oct/2013 18:57:03:177] "GET
duct_id=RP-SN-01& JSESSIONID=SD8SL8FF4ADFF8 HTTP 1
product.screen?product_id=RP-SN-01" "Googlebot/2.1
```

Business Insights
Make better-informed business decisions by understanding trends, patterns and gaining operational intelligence from machine data.

Operational Visibility
Gain end-to-end visibility across your operations and break down silos across your infrastructure.

Proactive Monitoring
Monitor systems in real time to identify issues, problems and attacks before they impact your customers, services and revenue.

Search + Investigation
Find and fix problems, correlate events across multiple data sources and automatically detect patterns across massive sets of data.

Transform machine data into real-time operational intelligence.

Splunk Premium Solutions. Go beyond apps to apply real-time data intelligence to manage your security posture, IT operations, and more. These purpose-built solutions extend the Splunk platform to monitor your specific environment, offering rich dashboards and key performance indicator (KPI) tracking, investigative capabilities, workflows and more that enable you to achieve and maintain operational excellence. Splunk Premium Solutions include: Splunk Enterprise Security, Splunk User Behavior Analysis and Splunk IT Service Intelligence.

Build enterprise-scale big data projects. Splunk Enterprise scales to collect and index hundreds of terabytes of data per day, across multi-geography, multi-datacenter and cloud-based infrastructures. And because the insights from your data are mission critical, Splunk provides the resilience you need, even as you scale out your low-cost, distributed computing environment. Automatic load balancing optimizes workloads and response times and provides built-in failover support. Out-of-the-box reporting and analytics capabilities deliver rapid insights from your data, removing the need for data scientists or complex development timeframes. Multisite clustering provides high availability and disaster recovery to ensure that your machine data is available when you need it.

A platform for enterprise developers. Enable developers to integrate data and functionality from Splunk Enterprise into applications across the enterprise using software development kits (SDKs) for Java, JavaScript, C#, Python, PHP and Ruby. Developers can also build Splunk Apps with custom dashboards, flexible UI components and custom data visualizations using common development languages such as JavaScript and Python.

Delivering the Key Capabilities for Operational Intelligence

- Universal collection and indexing of machine data, from virtually any source
- Powerful search processing language to search and analyze real-time and historical data
- Automatic detection of interesting patterns in your data
- Real-time monitoring for patterns and thresholds, trigger alerts when specific conditions arise
- Powerful reporting and analysis
- Custom dashboards and views for different users and roles
- Resilience and scale on commodity hardware
- Granular role-based security and access controls
- Support for multitenancy and flexible, distributed deployments
- Connectivity with other data stores includes scalable, real-time integration with relational databases and bi-directional connectivity with Hadoop and NoSQL data stores.
- Robust, flexible platform for developing enterprise apps

Universal Indexing

Individual components in your infrastructure generate hundreds of events per second. A datacenter can log many terabytes of data per day. You'll probably begin wondering how you're going to access all this data in all the different formats and locations. Splunk Enterprise offers a variety of flexible input methods and doesn't need custom connectors for specific data formats. You can immediately index logs, clickstream data, configurations, traps and alerts, messages, scripts, performance data and statistics from your applications, servers, mainframes and network devices—physical, virtual and in the cloud.

Flexible data input. Collect and index data from just about any source imaginable, such as network traffic, web servers, custom applications, application servers, hypervisors, GPS systems, sensors, stock market feeds, social media and preexisting structured databases. No matter how you get the data, or what format it's in, it's indexed the same way—without any specific parsers or connectors to write or maintain. Getting data in is fast and easy—just point Splunk Enterprise at your data and an intuitive user interface guides you through the rest. And for real-time DevOps and IoT data collection, Splunk Enterprise offers a high-capacity developer-standard HTTP/JSON API and SDKs.

Forwards data from remote systems. Splunk Enterprise forwarders can be deployed in situations where the data you need isn't available over the network or visible to the server where Splunk software is installed. Splunk Enterprise forwarders deliver reliable, secure, real-time universal data collection for tens of thousands of sources. Monitor local application log files, clickstream data, the output of status commands, performance metrics from virtual or non-virtual sources, or watch the file system for configuration, permissions and attribute changes. Forwarders are lightweight and can be deployed quickly, at no additional cost.

Real-time indexing. IT teams depend on up-to-date information for troubleshooting, security incident investigations, compliance reporting and other valuable tasks. Splunk Enterprise continually indexes machine data in real time—your logs, configuration data, change events, the output of diagnostic commands, data from APIs and message queues, even logs from your custom applications.

Captures everything. Store both raw data and the rich index in an efficient, compressed, redundant, file-system-based data store, with optional data signing and auditing to prove data integrity.

No rigid schemas. Splunk Enterprise has no predefined schema. Solutions that rely on brittle schemas have limited flexibility to answer new questions and break when data formats change. Any interpretation you need to do on the data, such as extracting a common field, or tagging a subset of hosts, is done at search time.

Automates chronology. All this streaming data means extracting and normalizing timestamps is very important. Splunk software automatically determines the time of any event—even with the most atypical or non-traditional formats. Data missing timestamps can be handled by inferring timestamps based on context.

Search and Investigation

Splunk Enterprise lets users search and navigate their data from one place.

Search and investigate anything. Freeform search supports intuitive Boolean, nested, quoted string and wildcard searches familiar to anyone comfortable on the web. This allows users to quickly iterate and refine their searches without knowing anything about specific data formats.

Powerful search processing language. The Splunk Search Processing Language (SPL) is a query language that provides a powerful means to operate on your data. It supports five different types of correlation (time, transactions, sub-searches, lookups, joins) and over 135 analytical commands. You can also conduct deep analysis and event pattern detection for spotting patterns or new opportunities in your data.

Real-time search. Searching real-time streaming data and indexed historical data from the same interface is best-in-class. With Splunk Enterprise you can analyze behavior and activity in real time and see the historical context.

Time-range search. Given the large volume and repetitive nature of machine data, users often start by narrowing their search to a specific time range. With the focus on when events happen, Splunk Enterprise lets users combine time and term searches. This ability to search across every tier of your infrastructure for errors and configuration changes in the seconds before a system failure occurs is incredibly fast and powerful.

Event Pattern Detection. Machine data can vary widely across your infrastructure—the data from your storage systems may not look like the data from your applications. Splunk Enterprise automatically detects meaningful patterns in machine data, regardless of data source or type. It then enables users to zoom in and out using a visual timeline so they can identify trends, spikes and drill down into the results.

Transaction search. Sending an email, placing an order on a website or connecting a VOIP call will create a number of events across different IT components. Often you'll want to search for these collections of events that are all part of the same transaction. For example, you can find all the sendmail events with the same user-ID, between a login and a logout, that occur within ten minutes.

Splunk Enterprise lets you correlate events by finding common characteristics and then saving that search as a transaction, so you can find the same type of transactions again for different search parameters.

Sub-searches. Take the results of one search and use them in another to create if/then conditions. Using a sub-search allows users to see the results of a search only if a set of other conditions are met (or not). Security event management systems operate on this premise. For example, you may only be interested in viewing one event if the threshold for another event is met in a given time period.

Lookups. Used to enhance, enrich, validate or add context to data collected in Splunk Enterprise. Correlating intrusion detection data (IDS) with data from an asset management system can reduce IDS false-positives. For example, an attack based on a Windows OS vulnerability seen by an IDS can be correlated with data from an asset being attacked within the AIX OS.

Joins. Support for "SQL-like" inner and outer joins are similar in concept to "joins" in an SQL database. Inner and outer joins are supported. "Join" as part of a search string can link one data set to another based one or more common fields. Two completely different datasets can be linked together based on a user name or event ID field presenting the results in a single view.

Interactive results. Compared to command line scripts and tools, an interactive interface dramatically improves the user's experience and the speed with which tasks can be accomplished. Zoom in and out on a timeline of results to quickly reveal trends, spikes and anomalies. Dynamically drilldown in dashboards anywhere in a chart to the raw events or define custom views and eliminate noise to get to the needle in the haystack. Whether you're troubleshooting a customer problem or investigating a security alert, you'll get to the answer quickly rather than taking many hours or days.

Add Knowledge

Adding machine data to Splunk Enterprise is possible with the native or custom input framework. Splunk Enterprise automatically discovers knowledge from your data and lets users add their own, unlocking your data's full potential. Knowledge about events, fields, transactions, patterns and statistics can be added to your data. You can identify, name and tag this data as well. Go from finding all events with a particular username, to instantly getting statistics on specific user activities. You can also correlate and name transactions that span multiple data sources. Splunk Enterprise marries the flexibility of freeform search with the power of working with your data, in a way you've never experienced before.

Map knowledge at search time. Avoid the problems caused by traditional approaches by mapping knowledge to data at search time, rather than attempting to normalize the data into a brittle database schema up front. And there's no more need for the complex management of custom parsers and connectors. Easily enrich your machine data with information from external asset management databases, configuration management systems and user directories. Now you have a flexible way to manage your data, so as it changes, you don't have to.

Work smarter. Splunk Enterprise lets every user add their own knowledge as they go. As you're saving searches and identifying different types of fields, events and transactions, you make the system smarter for everyone else. And that knowledge doesn't walk out the door when someone leaves.

Monitor and Alert

Rather than using search to simply react to ad hoc incidents or problems, you want to be proactive. Gain flexible alerting capabilities that improve your monitoring coverage. Because Splunk Enterprise works across your entire IT infrastructure, it's the most flexible monitoring solution in your arsenal.

Turn searches into real-time alerts. Searches can be saved and scheduled for continual monitoring and can trigger alerts via email or RSS. You can even kick off a script to take remedial actions, send an SNMP trap to your system management console or generate a service desk ticket. Scheduling alerts is a great way to complete the investigation of a problem or security incident by proactively looking for similar occurrences in the future.

Correlate complex events. Splunk Enterprise lets you correlate complex events from multiple data sources across your IT infrastructure so you can monitor more meaningful events. For example, you can track a series of related events as a single **transaction to measure duration or status.**

Monitor for specific conditions. Alerts can be based on a variety of threshold and trend-based conditions and to any level of granularity. The search language goes beyond simple Boolean searches into fielded searches, statistical searches and sub-searches. You can correlate on anything you want and alert on complex patterns such as abandoned shopping carts, brute force attacks and fraud scenarios.

Add context to alerts. Alerts can be embedded with machine data context, thereby reducing mean-time-to-resolution (MTTR).

Report and Analyze

If you've ever wanted to generate a report on-the-fly from hard-to-understand machine data, you'll love Splunk Enterprise. The Splunk Enterprise platform is capable of generating reports on an immense amount of data at lightning fast speeds. With built-in acceleration technologies, you have access to key data for a specified time window to make business-critical decisions. With instant pivot, you can create powerful, information-rich reports from any search, without any knowledge of search commands. You can schedule delivery of any report via PDF and share it with management, business users or other IT stakeholders.

Report on search results. Easily build advanced graphs, charts and sparklines from search results and visualize important trends, see highs and lows, summarize top values and report on the most and least frequent types of conditions. The simplicity of analyzing massive amounts of data will amaze you (and your boss). For example, a report can show the total bytes sent by IP address from firewall activity events; a table showing bytes per protocol per IP address; or a chart illustrating firewall traffic by hour for a specific employee's laptop. Virtually any field can be used as reporting criteria. And remember, because fields are identified as you search, you can specify new fields without re-indexing your data.

Analyze correlated events. Splunk Enterprise supports five types of correlation.

- Time-based correlations, to identify relationships based on time, proximity or distance
- Transaction-based correlations, to trace transactions that span multiple silos, systems and data sources so you can report on and analyze important activities
- Sub-searches, to take the results of one search and use them in another
- Lookups, to correlate data with external data sources outside of Splunk software, including relational databases
- Joins, to support SQL-like inner and outer joins

Instant pivot. Pivot directly on any search and allow any user to discover relationships in the data and build reports, all without having to learn SPL.

Plays well with others. Now your entire organization can leverage the value of machine data. Reports can be saved and shared with management or other colleagues in secure, read-only formats, such as PDF and even integrated into dashboards. Dashboard panels can be built and shared through a shareable library, allowing them to be added to any dashboard.

Custom Dashboards and Views

Make more sense of the huge volumes of data at your disposal. Create custom dashboards and views for different types of users, technical and non-technical. Integrate reports, search results and even data from external applications. Edit dashboards using a simple drag-and-drop interface; integrated charting controls mean you can change chart types on-the-fly. Doing this all through the Splunk UI means that you can empower business users to do the same.

Real-time, interactive dashboards. Dashboards integrate multiple charts, views and reports of live and historical data to satisfy the needs of different users. You can add workflows enabling users to click through to another dashboard, form, view or external website. Quickly build and personalize dashboards for management, business or security analysts, auditors, developers and operations teams.

Mashups with other apps. Create mashups with other web-based apps, such as Tivoli, SAP, security consoles and more, to provide a seamless view across silos.

Prebuilt panels. Quickly create dashboards using prebuilt panels that are shareable and integrate multiple charts and views of your data. Use geospatial mapping to visualize and communicate results using geographic or custom-defined areas. Leverage single value displays for "at-a-glance" indicators and relevant contextual data for war-room displays and management discussions.

Dashboards wherever you are. Charts and timelines in Splunk Enterprise don't use Flash, which means dashboards can be viewed and edited on tablets, smartphones and non-Flash browsers.

A Platform for Apps and Developers

Now that you're indexing and making use of all your machine data, you can make use of apps that let you do even more.

Innovate on your own. Easily create apps that deliver a targeted user experience for different roles and use cases. The Splunk Web Framework provides the ability to develop and package apps using modern development languages. Deliver a user experience tailored to a specific use case or augment existing vendor technologies.

Share and download apps. You can share and reuse apps within your organization and the rest of the Splunk community. There are a growing number of apps available on Splunkbase, built by our community, partners and Splunk. You can find apps that monitor a wide-variety of data sources or that enrich the experience and functionality of any task.

Simple management. Once Splunk Enterprise is installed, you can apply role-based access controls and deploy apps with a tailored user experience across the organization, extending the value of your data to different users.

Extendable platform. The Splunk platform makes is easy to customize and extend the power of Splunk Enterprise. Developers can debug and troubleshoot applications during development and test cycles or integrate data from Splunk Enterprise into custom applications. The Splunk Enterprise platform has built-in SDKs for JavaScript and JSON with additional downloadable SDKs for Java, Python and PHP. Developers can also build Splunk Apps with custom dashboards, flexible UI components and custom data visualizations using common development languages such as JavaScript and Python.

Enterprise-Scale Big Data

With Splunk Enterprise you can scale your installation from a single commodity Windows, Linux or Unix server, to the largest most complex multigeography, multidatacenter or cloud-based infrastructures indexing hundreds of terabytes of data per day. The Splunk Enterprise architecture is distributed and scales vertically with CPU power and horizontally across commodity servers to unlimited data volumes. You'll find a wide range of options to access data, store it, search it and route it to other systems.

Easy installation. A self-contained software package with no dependencies on third-party programs makes Splunk Enterprise easy-to-install and get running. It works on all major operating systems and hardware platforms. And because Splunk Enterprise is software, it can operate across physical or virtual infrastructures rather than requiring dedicated hardware, power and datacenter space.

Analyzes big data. Your datacenter generates more machine data than you probably ever imagined. A single production server can generate hundreds of megabytes of data a day. Firewalls and web servers can each generate many times that amount. In fact, machine data is one of the fastest, most complex segments of big data.

This volume of data is also subject to retention requirements ranging from a few days for incident response, to months and years for compliance. When considering performance and comparing approaches to collect, index and analyze and visualize your machine data here are some things to look for and consider:

Indexing throughput. Events-per-second (EPS) is a common throughput measurement, but consider that event sizes can vary from a few hundred bytes to a megabyte or more. EPS ratings are usually calculated at whatever size is optimal for one specific vendor's appliance or solution. Splunk Enterprise indexes every byte in your data, without the need for custom parsers or connectors. If the vendor is unable or unwilling to quote you EPS figures based on this criteria, move on and find someone who will.

Search speed. Searches of any type should return results in seconds, not minutes or hours. Based on a distributed computing framework, Splunk Enterprise automatically converts searches into a parallel program, providing the ability to quickly retrieve and analyze massive data sets. A single commodity server will support searching of billions of events in seconds.

Storage efficiency. Measured as a percentage of the original data stream size, storage efficiency determines the amount of storage capacity you'll need to retain your data and the associated indexes. A good solution will require 25% to 50% of the original data size to retain your data and a useful set of indexes. Beware of solutions that claim 10% or less of the original data size. That indicates just the storage of compressed data and no indexing.

Archiving. Eventually you may decide to tier the storage of your data. Tiered storage can provide lower cost and better redundancy. Archiving data based on disk utilization or age will come in handy for building a multi-tiered data store. Make sure your solution lets you set up an archiving policy based on datastore size or age and restore your archives on demand. Splunk Enterprise supports the option of archiving (including subsequent search) to Hadoop clusters.

Linear scalability. You can scale Splunk Enterprise horizontally and vertically by simply adding more computing power. You can run a distributed configuration on different physical servers, a combination of virtual and non-virtual servers, or on a large multicore, multiprocessor machine. Balance workloads by configuring multiple indexers and search engines across your configuration. Search head clustering enables additional concurrent searches and reduces TCO by eliminating the need for NFS storage requirements.

Availability. The availability and integrity of data are foundational elements for an enterprise. The data is mission critical and needs to be available at all times. Gain greater protection against data loss while maintaining data integrity. The high availability architecture of Splunk software delivers built-in resilience, so the right data is available when you need it.

Distributed and hybrid search. Often it won't be feasible to physically centralize all your data in one place. You will likely need to search across multiple Splunk Enterprise deployment sites and data stores in different technology or geographic silos. Hybrid search allows you to search across independent Splunk Enterprise, Splunk Cloud and Hunk installations.

Centralized management. Since access to machine data analytics is mission critical, the distributed management console allows you monitor your Splunk deployment from a single location. Monitor the health and performance metrics of your large-scale clusters, and receive alerts based on pre-defined conditions.

Integration. If you're like most IT shops, you've made significant investments in other management tools, monitoring tools and analysis tools. Wouldn't it be great if you could integrate Splunk software with all of them? Imagine launching in-context searches from your network management console, sending Splunk alerts to your system management console or automating trouble ticket creation when unusual activity occurs. Splunk Enterprise provides multiple integration points; a robust, documented API; and a simple-to-use Custom Alert Actions function.

Security

You'll need to keep your machine data secure, especially as you realize what a valuable information asset you have. Splunk Enterprise provides secure data handling, access controls, auditability, assurance of data integrity and integration with enterprise single sign-on solutions.

Secure data access and transport. Machine data can be sensitive. Splunk Enterprise supports advanced anonymization to mask confidential data from results. Private consumer, healthcare or corporate information also requires secure access, transport and storage. Encrypted access to data streams, using protocols such as TCP/SSL is a must-have feature for ensuring data security. User access should also be secured using protocols such as HTTPS or SSH for command-line access.

Granular access control. Of course you also need the ability to control the actions users can take and what data, tools and dashboards they can access. You don't necessarily want to allow the application development team access to your IDS scans, alerts and firewall logs. Splunk is a flexible, role-based system that lets you build your own roles to map to your organization's policies for different classes of users.

In some environments, like multitenant services, you may need to physically control access to data. The ability to route select data to distinct Splunk installations will let you physically separate data in different data stores. You'll also want to integrate with LDAP and Active Directory and map groups to different roles.

Single sign-on. If you're using access controls internally and have organizational access control policies, you'll want to make sure you can integrate your Splunk Enterprise solution with your authentication system, whether it's LDAP, Active Directory, e-Directory, SAML or another authentication system.

Audit capability. Once you have your access controls set-up, you need to monitor who's doing what. Splunk logs administrative and user activities so you can audit who's accessing what data and when.

Data integrity. You'll also need to ensure the integrity of your data. How do you know the search results or report you're viewing is based on data that hasn't been tampered with? With Splunk software, individual events can be signed and streams of events block signed. Splunk also provides message integrity measures that prove no one has inserted or deleted events from the original stream.

Hardened deployment. Keeping an audit trail and signing events is worthless if the server running Splunk Enterprise can be compromised. Be sure your vendor provides hardening guidelines.

ROI and Splunk

Splunk Enterprise customers typically achieve an ROI measured in weeks or months, sometimes even before Splunk Enterprise has been deployed into production. Splunk Enterprise users can troubleshoot application problems and investigate security incidents in minutes instead of hours or days, dramatically improve service levels, reduce outages and deliver compliance reporting at a lower cost. This visibility, typically unavailable prior to Splunk software, delivers organizations a fast ROI, new productivity and powerful insights. Here are a few examples:

- A leading provider of healthcare management solutions avoided a $100K SLA penalty—found during the Splunk evaluation phase. This same customer achieved an annual ROI of over $700,000.
- One of the world's largest business publishers replaced their old server monitoring software with Splunk Enterprise and other open source software. This eliminated maintenance fees and reduced operations costs by $1.6 million/year.
- A major communications manufacturer avoided a $1.5M software license upgrade for their existing SIEM, reassigned five full-time analysts to other duties ($600,000/year) and now monitors new data sources to identify previously unknown attacks.
- The world's largest B2B poker provider, hosting 25 of the industry's top brands and up to 45,000 concurrent players at peak hours, reduced downtime by 30% and quantified an annual savings of $1.9 MM (16x ROI in the 1st year).
- One of the world's largest online travel sites demonstrated an annual ROI over $14 million. This ROI was a combination of tools consolidation, retired licenses, outage avoidance and troubleshooting efficiencies gained using Splunk Enterprise.

Free Download

Download Splunk for free. You'll get a Splunk Enterprise 6 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 sales@splunk.com.

Seeking a best-in-class solution for managing your machine data? Here’s what to look for:

1	Index Any Machine Data
a	Indexes any machine data generated by applications, servers or network devices including logs, wire data, clickstream data, configurations, messages, traps and alerts, sensors, GPS, RFID, metrics and performance data without custom parsers or connectors for specific formats (includes virtual and non-virtual environments).
	Data can be loaded and indexed easily and intuitively. UIs and wizards are available to guide the process.
b	Flexible real-time and on-demand access to data from files, network ports and databases and custom APIs and interfaces.
	Captures wire data containing network communication across layers 3-7.
	Listens to TCP and UDP network ports to receive syslog, syslog-ng and other network inputs.
	Consumes archive files.
	Offers an HTTP/JSON API, SDK and endpoint for real-time event collection from apps or IoT devices.
	Captures new events in live log files in real time.
	Monitors files for changes.
	Queries database tables via DBI.
	Monitors Windows events remotely via WMI.
	Natively accesses the Windows event API.
	Monitors the Windows registry for changes.
	Connects to OPSEC LEA and other key security event protocols.
	Subscribes to message queues such as JMS.
	Captures the output of Unix/Linux system status commands like ps, top and vmstat.
	Remotely copies files via scp, rsync, ftp and sftp.
	Extensible via scripted inputs to capture the output of new status commands, connect to new event APIs and subscribe to different kinds of message queues.
c	Universally indexes data in virtually any format without custom parsers or connectors for specific data formats.
	Identifies events in single line, multi-line and complex XML structures.
	Recognizes and normalizes timestamps. Handles bad or missing timestamps through contextual inference.
	Captures and indexes the structure of each event.
	Tracks and indexes the host and source of each event.
	Classifies source formats dynamically.
d	Densely indexes every term in the original data.

e	Retains original, unaltered machine data.
f	Builds an unstructured index on disk without schema.
g	Supports forwarding and receiving of data from remote hosts for load balancing, failover and distributed deployments.

2	Search, Investigate, Explore
a	Search events across components in multiple formats at once.
b	Search live and historical data from the same interface and automatically backfill historical data for real-time windowed searches.
c	Deliver rapid data analysis through field extraction that adds context and meaning to machine data.
d	Fast results from searches on terms instead of queries optimized for specific fields/columns in a persistent schema.
e	Free form ad hoc search on any term in the original events with support for Booleans, nesting, quoted strings and wildcards.
f	Precise searches using fields identified within the data at search time. Supports multiple schema views into the same data without redundant storage or re-indexing.
g	Type-ahead suggestions to make it easy to discover what to search.
h	Navigate to related events and refine searches by clicking on fields or terms within the search results.
i	Search by time across multiple data formats.
j	Automatically detects anomalies and patterns across massive sets of machine data to discover meaningful outliers and patterns.
k	Visualize trends and navigate results using interactive time-based charts, histograms, sparklines and summaries.
l	Search for transactions across different data sources and components.
m	Persist searches as event and transaction types and search, filter and summarize by event and transaction type.
n	Discover fields, event types and transactions interactively at search time.
o	Save searches in reports, dashboards or views to simplify routine search scenarios.
p	Browser based, interactive AJAX user interface. No plug-ins required.
q	Optional scriptable CLI interface for both real-time and historical search.
r	Build data models that define objects, attributes and relationships in the underlying machine data.
s	Accelerate any data model with consistent authoritative views of machine data to drive high performance analytics.

3	Add Knowledge
a	Enable the system and the user to automatically add semantic meaning to machine data.
b	Automatically discovers knowledge from the machine data, such as timestamp, name/value pairs, headers, etc.
c	Let users add additional knowledge about the events, fields, transactions and patterns in their machine data.
d	Assign tags to field values to help search groups of events with related field values more efficiently.
e	Identify and classify transactions by correlating events across multiple data sources.
f	Save searches that return interesting results by either saving the search string (to run the search later) or the search results (to review the results later).
g	Share and promote saved searches, saved reports and event types with other authorized users.
h	Define a custom input capability and reuse other inputs; ensure that all inputs are available for use in the management interface.

4	Monitor and Alert
a	Run time-based search on a schedule and set alerting conditions based on thresholds and deltas in the number and distribution of results.
b	Trigger alert actions to personnel or other applications via email, RSS, SNMP, scripts or pre-built integrations.
c	Take automated corrective or follow-on actions via scripted alerts.
d	Embed sophisticated correlation rules in alerts via sub-searches.
e	Add context about the event that triggered the alert.

5	Report and Analyze
a	Build summary reports based on the results of any search interactively by clicking on available fields and statistics.
b	Create reports using fields and schemas identified at search time. Supports multiple schema views into the same data without redundant storage or re-indexing.
c	Supports sophisticated statistical and summary analysis by pipelining advanced search commands together in a single search.
d	Leverage report acceleration features to efficiently report on the very large volumes of data, e.g., long-term trends.
e	Accelerate reports by maintaining summaries that are up-to-date, scalable and used by other eligible searches.
f	View report results in tabular form; as interactive line, bar, pie, scatterplot and heat map charts. Visualize data using geospatial visualizations, single-value displays.
g	Pivot or drill down into any field, term or search without requiring knowledge of an advanced language.
h	Click through to another dashboard, form, view, or external website, carrying forward any relevant context.
i	Cache the results of scheduled reports for reuse.
j	Create real-time reports based on live streaming data sources.
k	Generate PDF versions of reports either on-demand or on a scheduled basis.
l	Schedule searches or report for automated delivery via email or RSS.

6	Create Custom Dashboards and Views
a	Create and edit dashboards that combine searches, reports, charts and tables using a visual dashboard editor.
b	Share pre-built panels to quickly build dashboards that integrate multiple charts and views.
c	Build sophisticated dashboards with entirely custom user interfaces and rich visualizations, including mashups with other applications and data from external sources.
d	Provide pre-packaged dashboards depicting key information and user activity—such as admin activity, search activity, index activity and inputs activity.
e	Leverage report acceleration features to efficiently report on the very large volumes of data, e.g., long-term trends.
f	Expand or restrict the role-based read and write permissions for a dashboard.
g	Create composite dashboards based on live and historical data sources. Deploy dashboards to devices and web browsers that do not support Flash.
h	Provide the ability to build complex reports, tables and visualization without a query language.
i	Provide integrated mapping software with geo-IP location.
j	Integrate with common third-party applications.

7	Build and Deploy Apps
a	Provide the ability to build and deploy apps on top of the machine data platform for specific use cases.
b	Package custom dashboards and configurations ranging from scripts, knowledge objects and back-end settings as apps.
c	Easily browse and dynamically switch between apps running on the Splunk platform by using an app launcher interface. Instantly see all installed apps on instance that the user has permissions to see.
d	Provide a powerful framework to support the creation of robust apps at all levels.
e	Expand or restrict the role-based read and write permissions to the app.

8	Developer Platform and Integration
a	Provide APIs to enable the quick integration with other applications, IT management tools and systems.
b	Minimum interface requirements should include, command-line interface, DBI, data routing, documented SDKs, REST API, scripted alerts, scripted inputs.
c	Leverage common development languages like Python and JavaScript for custom development.

9	Scale and Deploy
a	A self-contained software package with no dependencies on third-party programs. It runs on premises, in the cloud or in virtualized server and storage environments.
b	Native packages (rpm, deb, pkg, dmg, msi, etc.) and archive format distributions (.tgz, .zip, .tar.Z) are available for most widely deployed operating systems including Linux, Windows, Solaris, HP-UX, AIX, Free BSD and Mac OSX.
c	Servers work together support both centralized and decentralized models for machine data management across the organization.
d	Provides real-time centralization of machine data from production servers with reliable data transport over TCP.
e	Distributed architecture to support highly available configurations with integrated resilience, failover and load balancing.
f	Policy-based data routing among servers and to third-party systems.
g	Linear scaling to terabytes per day via distributed search and data balancing based on the MapReduce technique.
h	Single view across silos via distributed search.
i	Maintains a complete, signed audit trail of administrative actions and search history.
j	Monitors its own configurations for unauthorized change.
k	Centrally monitor the health and performance of deployments and large-scale clusters.
l	Centralized, policy-based configuration management across servers in a distributed deployment.

m	REST API enables quick integration with other IT management tools and systems.
n	Tunable indexing levels can be set for different sources or events.
o	Extremely fast search speed, delivers results fast across billions of events.
p	Highly efficient compressed storage—12-48% of the original data size typical for syslog depending on indexing level.
q	Data store uses local or network storage and is compatible with incremental file system backup utilities.
r	Index is segregated by time to support extended retention times without impact to search performance.
s	Configurable archiving and data retirement policy by age or size.
t	Archive and restore compressed or fully indexed data on demand. Facilitates maintaining oldest data using lower cost nearline storage for extended retention times.
u	Integrated use of MapReduce to enable scaling of real-time and historic search functions across commodity hardware.
v	Supports the option of archiving (including subsequent search) to Hadoop clusters.

10	Secure
a	Flexible roles for controlled user and API access. Supports granular data access and capabilities by role. Enables restricted access to specific data sources, data types, time periods, specific views, reports or dashboards.
b	Authentication and authorization integration with Active Directory, eDirectory and other LDAP-compliant implementations.
c	Integration to SAML and enterprise single sign-on solutions enabling pass-through authentication of third-party credentials.
d	Real-time remote indexing of data to minimize the opportunity for alteration of audit trails on compromised hosts.
e	Secure data stream access and distributed functionality via SSL/TCP or HTTP/JSON. Secure user access via HTTPS.
f	Block-signs events to demonstrate data integrity.
g	Maintains a complete, signed audit trail of administrative actions and search history.
h	Monitors its own configurations for unauthorized change.
i	Validates integrity of indexed data on demand to ensure security and compliance.