splunk>

TECH BRIEF

# The Splunk Light Difference

Use the Power of Splunk's Proven Technology to Put Out IT Fires

## Splunk® Light Fights IT Fires

Splunk Light is a comprehensive solution for small IT environments to automate log search and analysis. Splunk Light speeds tactical troubleshooting by gathering all of your log data into one place in real time, and providing a powerful search and analytical language for real-time analysis. The net result is the ability to quickly and proactively analyze problems and take immediate action—all without having to manually gather, organize and sift through gigabytes of data. This paper will provide a complete overview of the capabilities you need to give your small IT team the power to fight, and even prevent, IT fires.

### Delivering the Key Capabilities for Log Search and Analysis

- Universal collection and indexing of machine data, from virtually any source
- Powerful search processing language to query and analyze real-time and historical data
- Customizable dashboards and views
- Powerful reporting and analysis
- Real-time monitoring for patterns and thresholds, trigger alerts on important conditions

### The Splunk Approach

#### Universal Data Collection & Indexing

The first step in preventing IT fires is to get a handle on all of your data. Given the many sources of machine data in your organization, you'll probably begin wondering how you're going to gather and make sense of all of these different formats. Splunk Light offers the following native features to make this possible.

**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, and social media. 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 Light at your data and an intuitive user interface guides you through the rest.

**Forwards data from remote systems.** Splunk 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 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 Light 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.

**No rigid schemas.** Splunk Light 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.** Streaming data means extracting and normalizing timestamps is 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.

#### Monitor and Alert

After you've collected and indexed all of your data, you can start using it to move from reactive to proactive. Rather than using search to simply react to ad hoc incidents or problems, Splunk Light integrates monitoring and alerting with these features.

**Correlate complex events.** Splunk Light enables you to 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.

**Continuously monitor for specific conditions.** Alerts can be based on a variety of thresholds and trend-based conditions, and to any level of granularity. Your alerts can go 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 server or network performance degradation, 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).

**Turn searches into real-time alerts.** Searches can be saved and scheduled for continual monitoring and can trigger alerts via email or RSS.

**Have alerts take action.** Alerts can be set to run scripts that take remedial actions, send an SNMP trap to your system management console or generate a service desk ticket.

### Report and Analyze

Once you've set up your alerts, you may want to get regular updates on key parts of your operations. Splunk Light is capable of rapidly generating reports on an immense amount of data. 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. 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 Light 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 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
- **Joins,** to support SQL-like inner and outer joins

**Easy Report Sharing.** 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

Sometimes you just want to create a set of screens— or even a single screen—to give you the big picture view of your IT operations. Splunk Light allows you to create custom dashboards and views.

**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.

**Drag-and-drop interface.** Edit dashboards using a simple drag-and-drop interface; integrated charting controls mean you can change chart types on-the-fly.

**Prebuilt panels.** Quickly create dashboards using prebuilt panels that are shareable and integrate multiple charts and views of your data.

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

### Search and Investigation

Having all of your operational data in one place is useless if you can't use it to solve problems and understand your systems. Splunk Light enables users to 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 and analytical 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 Light 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 a focus on when events happen, Splunk Light enables users to 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 Light 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 Light enables you to correlate events by finding common characteristics, save that search as a transaction, and find the same type of transactions again for different search parameters.

**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 can get to the answer quickly, rather than taking many hours or days.

### Security

You'll need to keep your machine data secure, especially as you realize what a valuable information asset you have. Splunk Light 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 Light 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.

**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 nobody has inserted or deleted events from the original stream.

**Hardened deployment.** Keeping an audit trail and signing events doesn't help if the server running Splunk Light can be compromised. Be sure your vendor provides hardening guidelines.

## Which Splunk is Right for You?

Splunk Light and Splunk Enterprise are both powerful solutions, but you want to make sure you choose the one that's right for you. Splunk Light is ideal for log search and analysis for small IT environments with no more than 5 users and Splunk software operating on a single server.

For an enterprise-grade environment, take a look at Splunk Enterprise. It offers the functionality, scalability and security to meet the needs of any company. With pre-packaged apps for IT operations, security, business analytics, and Internet of Things, Splunk Enterprise delivers value across your organization. Regardless of which you chose, you will have an option of a cloud service or installable software.

See below for a selected feature comparison or click here to learn more about Splunk Enterprise.

| Features | Splunk Light | Splunk Enterprise |
|---|---|---|
| Maximum Daily Indexing Volume | 20GB | Unlimited |
| Maximum Users | 5 | Unlimited |
| Universal Data Collection/ Indexing | • | • |
| Data Collection Add-Ons | • | • |
| Monitoring and Alerting | • | • |
| Dashboards and Reports | • | • |
| Search and Analysis | • | • |
| Automatic Data Enrichment | • | • |
| Anomaly Detection | • | • |
| Data Models and Pivot | | • |
| Packaged Apps | | • |
| Scalability | Single Server | Unlimited |
| High Availability | | • |
| Disaster Recovery | | • |
| Clustering | | • |
| Distributed Search | | • |
| Performance Acceleration | | • |
| Access Control | User and Admin only | Granular and Customizable |
| Single Sign-On/LDAP | | • |
| Developer Environment | | Full access to APIs and SDKs |
| Support | Standard | Enterprise/Global |

**Start Putting Out IT Fires—Try Splunk Light**

You have IT problems that Splunk Light can solve—so what are you waiting for? Try Splunk Light as a cloud service trial or free download. For those ready to buy: purchase now with a credit card or through a reseller.

Splunk Light is available in the U.S., Canada, Europe, Middle East, Africa, Australia and New Zealand.

250 Brannan St., San Francisco, CA 94107 info@splunk.com | sales@splunk.com 866-438-7758 | 415-848-8400 splunkbase.splunk.com





Item # TB-Splunk-Light-109

## Appendix – Selection of key Splunk Light features for log search and analysis

| 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-ing and other network inputs. |
| | Consumes archive files. |
| | 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** | Zoom in and out on a timeline of results to quickly reveal trends, spikes and anomalies |
| **k** | Automatically detects patterns across massive sets of machine data to discover meaningful patterns. |
| **l** | Visualize trends and navigate results using interactive time-based charts, histograms, sparklines and summaries. |
| **m** | Search for transactions across different data sources and components. |
| **n** | Persist searches as event and transaction types and search, filter and summarize by event and transaction type. |
| **o** | Discover fields, event types and transactions interactively at search time. |
| **p** | Save searches in reports, dashboards or views to simplify routine search scenarios. |
| **q** | Browser based, interactive AJAX user interface. No plug-ins required. |

| 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 alerts via email, RSS, SNMP or scripts. |
| **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** | View report results in tabular form; as interactive line, bar, pie, scatterplot and heat map charts. |
| **e** | Create real-time reports based on live streaming data sources. |
| **f** | Generate PDF versions of reports either on-demand or on a scheduled basis. |
| **g** | 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** | Easily build dashboards with rich visualizations. |
| **d** | Create composite dashboards based on live and historical data sources. |
| **e** | Provide integrated mapping software with geo-IP location. |