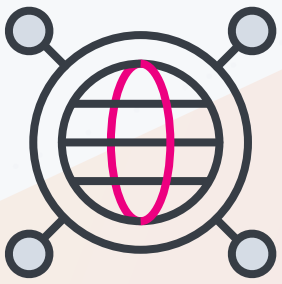
The Essential Guide to
Network Data

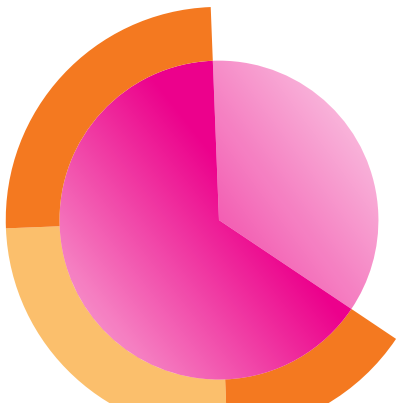


splunk>
turn data into doing™

Time-Series Data. Streaming Data. Dark Data.

It's no secret that data remains underused and undervalued in most organizations all over the world. Despite the constant talk of data-driven decisions, organizations of all sizes are still missing the mark on how to effectively capture and use the troves of data being generated every day, whether it comes from users, outside industry resources, or their own networked devices. In fact, most business and IT decision makers estimate that **55% of their data is dark data**, information you don't know you have, or can't fully tap.

This is a big missed opportunity. Important insights across IT, security and your organization lie hidden in this data. Data holds the definitive record of all activity and behavior of your customers and users, transactions, applications, servers, networks, mobile devices and more. Critical information on everything from configurations, APIs, message queues, diagnostic outputs, sensor data of industrial systems and more is all there — you just have to tap into it the right way.



With the right approach, data makes it simple to:

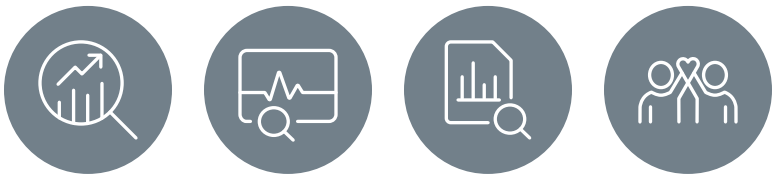
- Make better informed decisions about every part of your business.
- Run your operations more efficiently.
- Optimize user and customer experiences.
- Detect the fingerprints of fraud — or prevent it altogether.
- Uncover potential disasters before they happen.
- Find hidden trends that help your company leapfrog the competition.
- Make everyone who uses it look like a hero.
- … and so much more.

The challenge with leveraging the vast quantity of data that most companies collect is that it comes in a dizzying range of formats that traditional data monitoring and analysis tools aren't designed to handle. Many tools can't keep up with the varying data structures, sources or time scales. And it goes well beyond just machine data as well. But the upside to tapping into your data is tremendous, and this is where Splunk comes in.

With Splunk, you can bring data to every question, decision and action in your organization to create meaningful outcomes. Unlike any other platform, Splunk is truly able to take any data from any source and drive real action to benefit the business — from IT infrastructure and security monitoring to DevOps and application performance monitoring and management.

Data-to-Everything in Practice

Use data to:



The organizations that get the most value out of their data are those able to take disparate data types, enrich them and extract answers. But not knowing what data to ingest can stop businesses before they start.

Familiarizing yourself with general use cases in security, IT operations, business analytics, DevOps, the Internet of Things (IoT) and more — including the data types and sources involved — can get you on track right away.

Here's an example:

1. A customer's order didn't go through
2. The customer called support to resolve the issue
3. After too much time on hold, the customer gave up and tweeted a complaint about the company

What Does Machine Data Look Like?

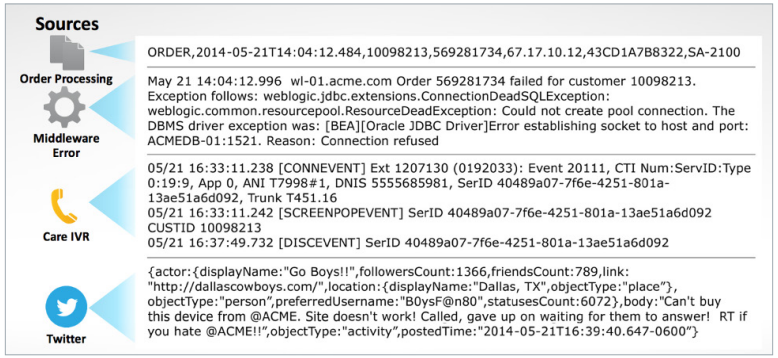


Figure 1: Data can come from any number of sources, and at first glance, can look like random text.

Machine Data Contains Critical Insights

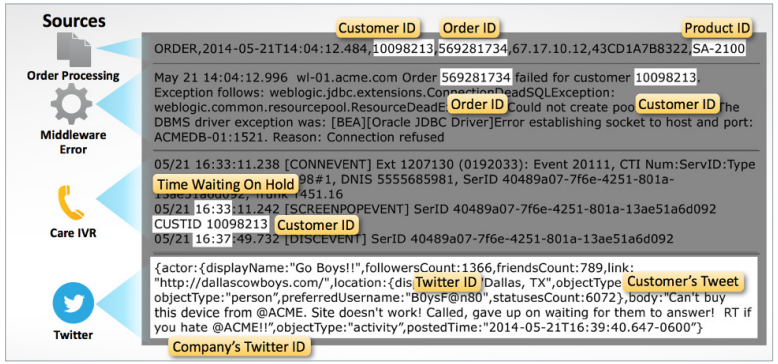


Figure 2: The value of data is hidden in this seemingly random text.

Machine Data Contains Critical Insights

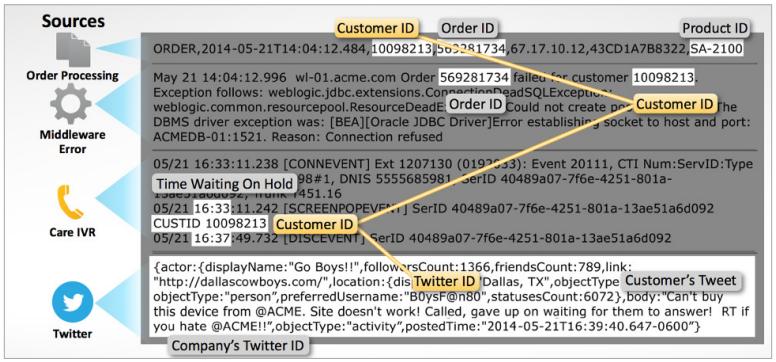


Figure 3: By correlating different types of data together, you can start to gain real insight into what's going on in your infrastructure, see security threats or even use the insights to drive better business decisions.

By taking all the data involved in the process — i.e. pulling information from order processing, middleware, interactive voice response systems and Twitter — an organization can get a full view of the customer experience problem.

Network Data

This book provides a high-level overview of the value you can get from the data on your network. This data can support a variety of use cases, including protecting corporate networks from attacks, providing visibility into network traffic, and helping to determine your network's role in the overall availability and performance of critical services.

While each organization's needs and data sources will vary by vendor, product and infrastructure, this book details where you should look for type of data and the value it can provide.

Many of the data sources listed in this book can support multiple use cases — this is a major part of what drives data's tremendous value.

Security and Compliance

IT Ops, App Delivery and DevOps

Table of Contents

Network Data

Deep Packet Inspection Data

Use Cases: Security and Compliance, IT Operations

Examples: Stream, PCAP, bro

Deep Packet Inspection Data (DPI) is a fundamental technique used by firewalls to inspect headers and the payload of network packets before passing them down the network subject to security rules. DPI provides information about the source and destination of the packet, the protocol, other IP and TCP/UDP header information and the actual data.

Use Cases

Security and Compliance: Packet Capture logs (PCAP) see everything traversing a network and are required to identify security attacks and incidents such as advanced persistent threats, data exfiltration, DDoS and malware. DPI also can be used to filter content subject to an organization's terms of service. PCAP data can also be used to provide and identify:

- DNS session analysis for malicious domain communications from each endpoint.
- Abnormal amounts of traffic or sessions.
- Abnormal amounts of domain and host communications.
- Known malicious traffic from a host.
- Expired SSL certification analysis.
- Abnormal host communications (internal and external).

IT Ops: Data on the network wire is authoritative and difficult to spoof (although encryption, steganography and advanced deception techniques can evade DPI). For example, DPI provides raw information of everything transmitted over a network, including things that aren't necessarily part of or difficult to extract from a log, such as database query results.

DHCP

Use Cases: Security and Compliance, IT Operations

Examples: DHCP Insight, Linux DHCP

DHCP is the network protocol most client devices use to associate themselves with an IP network. Implemented via a DHCP server, which could be standalone or embedded in a router or other network appliance, DHCP provides network clients with critical network parameters including IP address, subnet mask, network gateway, DNS servers, WINS or other name servers, time servers (NTP), a host and domain name and the address of other optional network services.

Use Cases

Security and Compliance: DHCP logs show exactly which systems are connecting to a network, their IP and MAC addresses, when they connect and for how long. This information is useful in establishing the state of a network when a security incident occurs and tracing an attacker's address back to a time of access and type of device by looking at the MAC ID and vendor identification string. The data can also be used to support user network access verification.

IT Ops: DHCP logs can be used when troubleshooting a client device that is having network problems, since it provides a definitive record of the device's primary IP parameters. The data may show that the DHCP server itself is at fault; for example, by not properly vending addresses, renewing IP leases or giving the same address to two separate devices.

DNS

Use Cases: Security and Compliance, IT Operations

Examples: BIND, PowerDNS, Unbound, Dnsmasq, Erl-DNS

The domain name system (DNS) is the internet's phone book, providing a mapping between system or network resource names and IP addresses. DNS has a hierarchical name space that typically includes three levels: a top-level domain (TLD) such as .com, .edu or .gov; a second-level domain such as "google" or "whitehouse;" and a system level such as "www" or "mail." DNS nameservers operate in this hierarchy either by acting as authoritative sources for particular domains, such as a company or government agency, or by acting as caching servers that store DNS query results for subsequent lookup by users in a specific location or organization; for example, a broadband provider caching addresses for its customers.

Use Cases

Security and Compliance: Security teams can use DNS logs to investigate client address requests such as correlating lookups with other activity, whether requests are made for inappropriate or otherwise suspicious sites and relative popularity of individual sites or domains. Since DNS servers are a frequent target of DDoS attacks, logs can reveal an unusually high number of requests from external sources. Likewise, since compromised DNS servers themselves are often used to initiate DDoS attacks against other sites, DNS logs can reveal whether an organization's servers have been compromised. DNS data can also provide detection of unknown domains, malicious domains and temporary domains.

IT Ops: DNS server logs provide operations teams with a record of traffic, the type of queries, how many are locally resolved either from an authoritative server or out of cache, and a picture of overall system health.

NETWORK DATA

Endpoint

Use Case: Security and Compliance

Examples: McAfee ePO, Symantec SEP

Endpoint security is used to protect corporate networks from inadvertent attacks by compromised devices using untrusted remote networks such as hotspots. By installing clients on laptops or other wireless and mobile devices, endpoint security software can monitor activity and provide security teams with warnings of devices attempting to spread malware or pose other threats.

In this context, endpoint refers to the security client software or agent installed on a client device that logs security-related activity from the client OS, login, logout, shutdown events and various applications such as the browser (Explorer, Edge), mail client (Outlook) and Office applications. Endpoints also log their configuration and various security parameters (certificates, local anti-malware signatures, etc.), all of which is useful in post-hoc forensic security incident analysis.

Use Cases

Security and Compliance: Endpoint data can be used for a variety of security uses, including identifying newly detected binaries, file hash, files in the filesystem, and registries. It can also help with identifying binary and hash registries that match threat intelligence, as well as unpatched operating systems and binaries, and to detect known malware.

Firewall

Use Cases: Security and Compliance, IT Operations

Examples: Palo Alto, Cisco, Check Point

Firewalls demarcate zones of different security policies. By controlling the flow of network traffic, firewalls act as gatekeepers collecting valuable data that might not be captured in other locations due to the firewall's unique position as the gatekeeper to network traffic. Firewalls also execute security policy and thus may break applications using unusual or unauthorized network protocols.

Use Cases

Security and Compliance: Firewall logs provide a detailed record of traffic between network segments, including source and destination IP addresses, ports and protocols, all of which are critical when investigating security incidents. The data may also reveal gaps in security policy that can be closed with tighter construction of firewall rules. Firewall data can help identify and detect:

- Lateral movement
- Command and Control traffic
- DDoS traffic
- Malicious domain traffic
- Unknown domain traffic
- Unknown locations traffic

IT Ops: When network applications are having communication problems, network security policies may be the culprit. Firewall data can provide visibility into which traffic is blocked and which traffic has passed through — helping identify if you have an app or network issue.

NETWORK DATA

FTP

Use Cases: Security and Compliance, IT Operations

Examples: OSSEC, Getwatchlist, UTBox, Security Onion, iSeries - AS400, Traffic Ray

FTP is one of the oldest and most rudimentary network protocols for copying data from one system to another. Before websites and HTTP, FTP was the best way to move large files across the internet. FTP is still used in organizations that need reliable, deterministic internet file transfer.

Use Cases

Security and Compliance: Analyzing FTP servers can help security teams identify when compromised credentials are used, when abnormal traffic is coming from different locations or at odd times, and when sensitive files and documents are being accessed.

IT Ops: FTP traffic logs record the key elements of a file transmission, including source (client) name and address and remote user name if the destination is password-protected. This and other data are crucial when troubleshooting FTP problems, regardless of the application.

Intrusion Detection/ Prevention

Use Case: Security and Compliance

Examples: Tipping Point, Juniper IDP, Netscreen Firewall, Juniper NSM IDP, Juniper NSM, Snort, McAfee IDS

IDS and IPS are complementary, parallel security systems that supplement firewalls — IDS by exposing successful network and server attacks that penetrate a firewall, and IPS by providing more advanced defenses against sophisticated attacks. IDS is typically placed at the network edge, just inside a perimeter firewall, although some organizations also put a system outside the firewall to provide greater intelligence about all attacks. Likewise, IPS is typically placed at the network perimeter, although it also may be used in layers at other points inside the network or on individual servers. IPS usually works by dropping packets, resetting network connections and blacklisting specific IP addresses or ranges.

Use Cases

Security and Compliance: IDS logs provide security teams detailed records of attacks including the type, source, destination and port(s) used that provide an overall attack signature. Special signatures may trigger alarms or other mitigating actions. IPS provide the same set of attack signature data, but also may include a threat analysis of bad network packets and detection of lateral movement. This data can also detect command and control traffic, DDoS traffic, and malicious or unknown domain traffic.

Load Balancer

Use Cases: IT Operations

Examples: Local Traffic Manager, Cisco Load Balancer, Citrix, Kemp Technologies, Radware AppDirector OnDemand

Load balancers allocate external network traffic bound for a particular server or application across multiple redundant instances. There are two categories of load balancer: local, in which all resources in a load-balanced pool are on the same subnet; and global or distributed, where the resource pool is spread across multiple sites. Load balancers use several user-selectable algorithms to allocate traffic including:

- Round robin (systems get an equal number of connections allocated sequentially).
- Weighted round robin (where the load is assigned according to the percentage weight assigned each system in a pool).
- Least connections (where new connections go to the system with the fewest number of existing clients).
- Weighted least connections (where the connection handling capacity of each system is taken into account when determining the least busy system for new connections).
- Random (connections are randomly assigned to each member of a pool).

Use Cases

IT Ops: Load balancer logs provide operations teams with a record of overall traffic to systems or particular applications and provide indicators of each system's traffic-handling capacity and health, along with the status and health of the load balancer itself.

Network Access Control (NAC)

Use Case: Security and Compliance

Examples: Aruba ClearPass, Cisco ACS

Network access or admission control is a form of client/endpoint security that uses a locally installed software agent to pre-authorize connections to a protected network. NAC screens client devices for contamination by known malware and adherence to security policies such as running an approved OS with the most recent patches. Clients failing NAC screens are rerouted to an isolated quarantine network until any detected problems are corrected.

Use Cases

Security and Compliance: NAC software collects data about the connecting clients such as an inventory of installed client software, compliance with security policies, OS and application patch versions, accessibility by remote access clients and user access to protected networks. NAC logs provide security teams with a detailed profile of a client's state and activity. It can provide details into unauthorized device connections and be used to correlate users/IP to a physical network location.

NETWORK DATA

Network Protocols

Use Cases: Security and Compliance, IT Operations

Examples: HTTP, Cisco NetFlow, Ntop, Flow-tools, FlowScan, EHNT, BPFT

Network protocols describe the structure of data that flows through networks. In most cases, network ports are assigned to specific protocols for both security and performance reasons. Some protocols operate at a lower level of the computing stack and are used to direct packet routing, such as TCP, UDP or IP. Other protocols, such as HTTP, HTTPS and TNS describe how packets are structured for applications — such as web services, databases and a wide range of client-based applications. By capturing, decrypting and analyzing network protocol data, you can better understand the kinds of applications, their usage, performance and even payload (content of the data) of applications. Since this data can be gathered directly from a network tap, or with specialized software, it provides a perspective on applications and how they interoperate that may not be otherwise available.

Use Cases

Security and Compliance: Network protocols are an important source for identifying advanced persistent threats, analyzing traffic flows for unusual activity and identifying potential data exfiltration. Aggregating and analyzing flow records also can show anomalous traffic patterns and flow destinations that are indicative of a breach, such as an APT phoning home to a command and control server for instructions, additional malware code, or copying large amounts of data to an attacker's system. The data can also be used to detect traffic related to DDoS, malicious domains, and unknown domains or locations.

IT Ops: Network protocol traffic analysis can help determine the network's role in overall availability and performance of critical services. Application traffic can be monitored for usage, performance, availability and can provide visibility into specific user data. For applications that cannot be instrumented on the servers, network traffic may be the only way to acquire performance data.

NETWORK DATA

Network Routers

Use Cases: Security and Compliance, IT Operations

Examples: Routers from Cisco, Juniper, Linksys, Arista, Extreme Networks, Avaya

If switches are network intersections, then routers are the signal lights and traffic cops — the devices responsible for ensuring that traffic goes to the right network segment. Unlike switches that operate at Layer 2, routers work at Layer 3, directing traffic based on TCP/IP address and protocol (port number). Routers are responsible for particular Layer 3 address spaces and manage traffic using information in routing tables and configured policies. Routers exchange information and update their forwarding tables using dynamic routing protocols.

Use Cases

Security and Compliance: Routers collect the same sort of traffic logs and statistics as switches; thus, their data is equally valuable to security teams as a source for flagging advanced persistent threats, analyzing traffic flows for unusual activity and identifying potential data exfiltration. As a wire-level data source, router statistics are almost impossible to spoof and thus a critical source of security data. Router data can also be used to detect configuration changes, and error or failure alerts correlating with security indicators.

IT Ops: Network engineers use router logs and statistics to monitor traffic flow and ensure that traffic is being correctly forwarded between network segments. Data from routing protocol updates can show whether your routers are appropriately exchanging route tables with other locations, that external traffic can reach you, and that internal traffic is correctly forwarded to external routers.

Network Switches

Use Cases: Security and Compliance, IT Operations

Examples: Ethernet Switch, Virtual Switches

Switches are network intersections, places where packets move from one network segment to another. In their purest form, switches work within a particular IP subnet and can't route Layer 3 packets to another network. Modern data center designs typically use a two-tier switch hierarchy: top-of-rack (ToR) switches connecting servers and storage arrays at the edge and aggregation or spine switches connecting to the network core. Although ethernet switches are far more widespread, some organizations also use fiber channels or infiniband for storage area networks or HPC interconnects, each of which has its own type of switch.

Use Cases

Security and Compliance: Switch data, often captured as NetFlow records, is a critical data source for flagging advanced persistent threats, analyzing traffic flows for unusual activity and identifying potential data exfiltration. As a wire-level data source, switch statistics are almost impossible to spoof and thus a crucial source of security data. This data can also be used to correlate users or IP addresses to a physical network location.

IT Ops: Operations teams use switch logs to see the state of traffic flow, such as source and destination, class of service and causes of congestion. Logs also can show traffic statistics in the aggregate, by port and by client, and whether particular ports are congested, failing or down.

Proxies

Use Cases: Security and Compliance, IT Operations

Examples: Blue Coat, Fortinet, Juniper IDP, Netscreen Firewall, Palo Alto Networks, Palo Alto Networks config, Palo Alto Networks system, Palo Alto Networks threat, Palo Alto Networks traffic, nginx

Network proxies are used in several ways in IT infrastructure: as web application accelerators and intelligent traffic direction, application-level firewalls, and content filters. By acting as a transparent 'bump-in-the-wire' intermediary, proxies see the entire Layer 7 network protocol stack, which allows them to implement application-specific traffic management and security policies.

Use Cases

Security and Compliance: Security teams are interested in proxies as application-layer firewalls. Here, proxy records can identify details about specific content traversing network control points including file names, types, source and destination, and metadata about the requesting client such as OS signature, application and username/ID (depending on the proxy implementation). The data can also be used to help detect command and control traffic, malicious domain traffic and unknown domain traffic.

Web proxies and some next generation firewalls may act in a transparent or explicit mode communicating with HTTP(S) servers on behalf of a client. Using a number of related technologies, the request and response can be inspected and permitted, or blocked, based on user role, site or resource category or attack indicator. Data logged in the events can potentially be used in detective correlation.

IT Ops: Operations teams often use proxies embedded in an application delivery controller (ADC), a more advanced, Layer 7-aware version of a load balancer. In this context, proxy logs can provide information about incoming requests and traffic distribution among available resources.

VoIP

Use Cases: Security and Compliance, IT Operations

Examples: Asterisk CDR, Asterisk event, Asterisk messages

Voice over IP protocol refers to several methods for transmitting real-time audio and video information over an IP-based data network. Unlike traditional phone systems using dedicated, point-to-point circuits, VoIP applications use packet-based networks to carry real-time audio streams that are interspersed with other ethernet data traffic. Since TCP packets may be delivered out of order due to data loss and retransmission, VoIP includes features to buffer and reassemble a stream. Similarly, VoIP packets are usually tagged with quality of service (QoS) headers to prioritize their delivery through the network.

Use Cases

Security and Compliance: VoIP deployments may expose organizations to potential security threats, and analyzing VoIP logs can help identify and prevent these exploits.

IT Ops: VoIP logs provide troubleshooting and usage data similar to that of other network applications. Details include source, destination, time and duration of calls, call quality metrics (e.g., packet loss, latency, audio fidelity/bit rate) and any error conditions. Integrating VoIP source/destination records with an employee database such as AD or LDAP and a DHCP database allows linking call records to actual people and IP addresses to physical locations; information that can assist in troubleshooting and billing.

SNMP

Use Cases: Security and Compliance, IT Operations

Examples: LogicMonitor, ManageEngine, Spiceworks, Ruckus Idera, Ipswitch

The simple network management protocol (SNMP) is one of the oldest, most flexible and broadly adopted IP protocols used for managing or monitoring networking devices, servers and virtual appliances. This includes network devices such as routers and switches, as well as non-networking equipment such as server hardware or disk arrays.

SNMP supports two different methods of obtaining data.

- **SNMP Traps** are essentially alerts, set to send an alert on a state change, critical threshold, hardware failure, and more. Traps are initiated by the SNMP device, and the trap is sent to an SNMP collector.
- **SNMP Polling** is an interactive query/response approach. Unlike traps, polling is initiated by the SNMP collector in the form of a request for certain, or all, SNMP data available on the SNMP device.

Although many now provide vendor-specific APIs for remote management and data collection, SNMP is still valuable in troubleshooting due to its ubiquity (nearly every device supports it) and inherently centralized design (a single instance of SNMP management software can collect data from every device on an internal network, even across route domains).

Use Cases

Security and Compliance: SNMP traps and alerts from network devices can help security teams identify abnormal activity over the network. SNMP Polling helps a security analyst to see the data transmission rates for a network-connected device that is suspected of malicious activity.

The data can also help identify abnormal amounts of traffic to a certain site or domain, an abnormal amount of specific SNMP traps from a certain host, and an abnormal number of unique SNMP traps from hosts compared to normal profiles.

IT Ops: SNMP data can provide current information about performance, configuration and current state. This allows the monitoring of the "normal" state of the environment, which is vital when using a service-level approach to monitoring the health of any environment. This could include current speed of all of the ports on a switch, the number of bytes sent (per port or in aggregate) through a router, the CPU temperature of a server, and any other information made available by the vendor per the SNMP MIBs for that device.

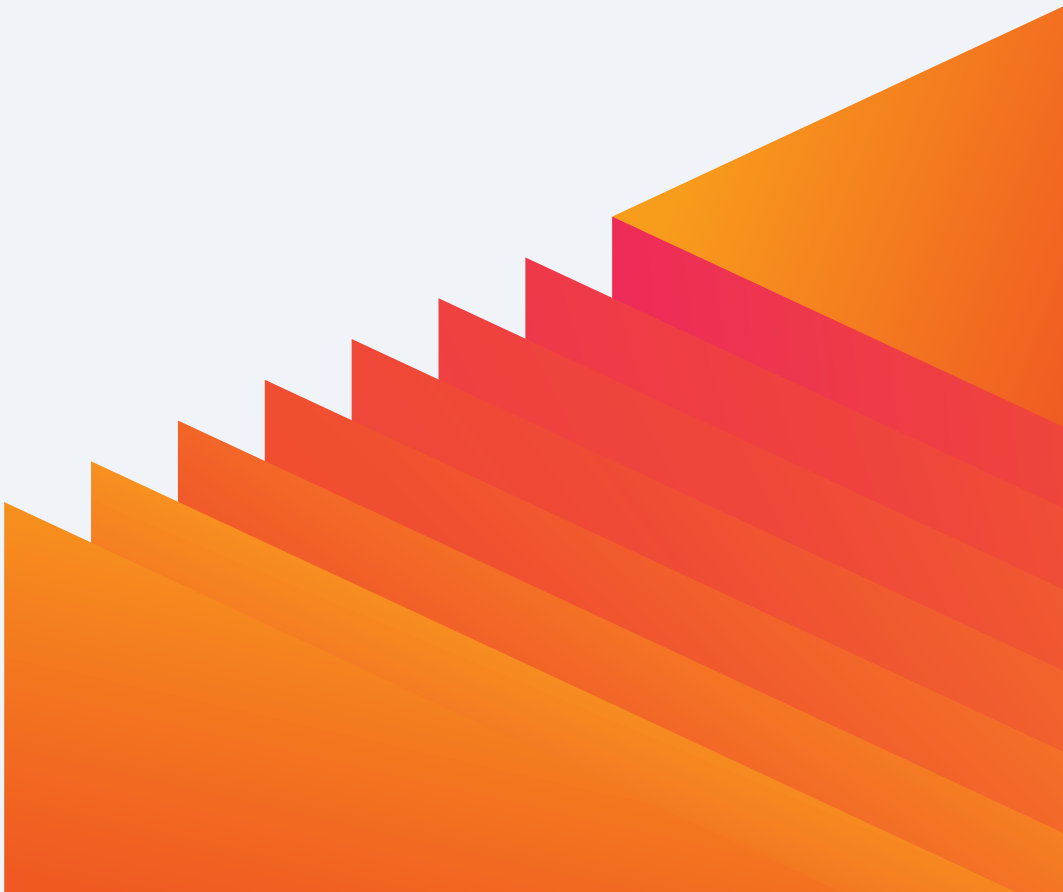
Many environments rely on SNMP traps for alerting when a critical state is reached (e.g., CPU temperature is critical) or when a failure occurs (e.g., RAID disk failure). SNMP traps are not only sent by devices to monitoring systems, in some environments SNMP traps are the de-facto method for multiple monitoring and alerting systems to aggregate errors to a single console.

NETWORK DATA

About **Splunk.**

Splunk turns data into doing with the Data-to-Everything™ Platform. Splunk technology is designed to investigate, monitor, analyze and act on data at any scale. Join millions of passionate users by trying Splunk for free.

Free Trial



splunk>

20-13476-SPLK-Essential-Guide-to-Data-Network-Data-105