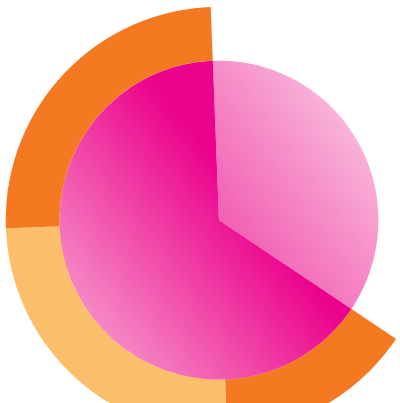
The Essential Guide to
Infrastructure Data

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:

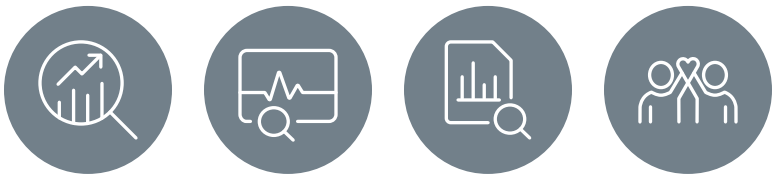
- Make better informed decisions about every part of your business.
- Run your operations more efficiently.
- Optimize user and customer experiences.
- Detect the fingerprints of fraud — or prevent it altogether.
- Uncover potential disasters before they happen.
- Find hidden trends that help your company 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?

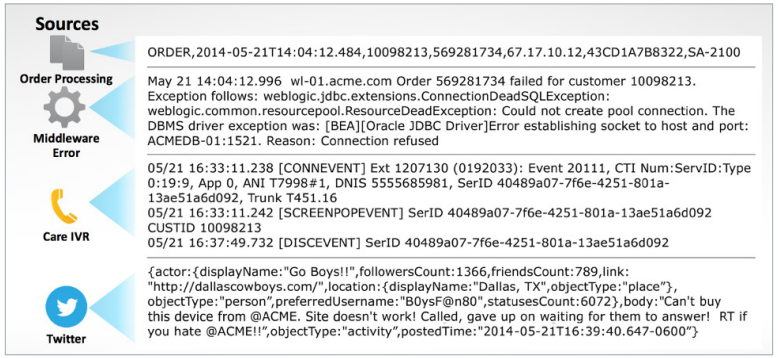


Figure 1: Data can come from any number of sources, and at first glance, can look like random text.

Machine Data Contains Critical Insights

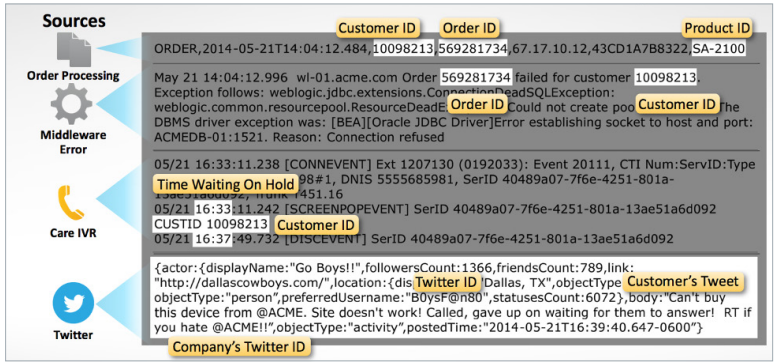


Figure 2: The value of data is hidden in this seemingly random text.

Machine Data Contains Critical Insights

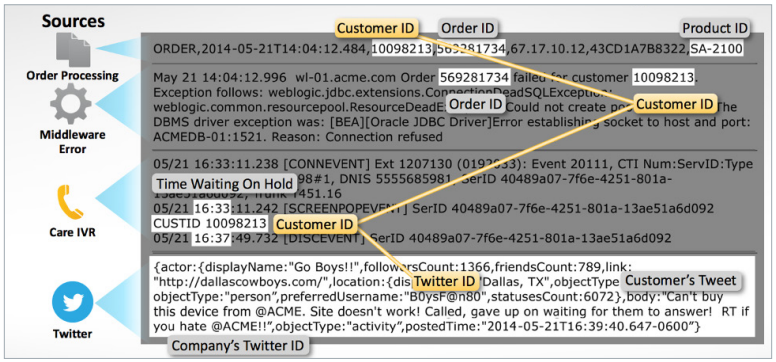


Figure 3: By correlating different types of data together, you can start to gain real insight into what's going on in your infrastructure, see security threats or even use the insights to drive better business decisions.

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.

Infrastructure Data

This book provides a high-level overview of the value you can get from the data created by your virtual and physical infrastructure as part of normal operations. This data can support a variety of use cases, ranging from monitoring your cloud deployments to identifying breach attempts and plugging vulnerabilities.

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 machine data and the value it can provide to IT, security, IoT and business analytics use cases.

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

Security and Compliance

IT Ops, App Delivery and DevOps

Internet of Things

Business Analytics

Table of Contents

Virtual Infrastructure Data 6

- AWS Services 6
- Google Cloud Platform (GCP) 7
- Microsoft Azure 7
- Pivotal Cloud Foundry (PCF) 8
- VMware Server Logs, Configuration Data and Performance Metrics 9

Physical Infrastructure Data 10

- Backup 10
- Environmental Sensors 11
- Industrial Control Systems (ICS) 11
- Mainframe 12
- Medical Devices 12
- Metric Line Protocols 13
- Patch Logs 14
- Physical Card Readers 14
- Point-of-Sale Systems (POS) 15
- RFID/NFC/BLE 16
- Sensor Data 17
- Server Logs 18
- Smart Meters 18
- Storage 19
- Telephony 19
- Transportation 20
- Wearables 20

Virtual Infrastructure Data

AWS Services

Use Cases: Security and Compliance, IT Operations

Examples: CloudTrail, CloudWatch, Config, S3

AWS is the largest and most widely used public cloud infrastructure, providing on-demand compute, storage, database, big data and application services with consumption-based pricing. AWS can be used to replace traditional enterprise virtual server infrastructure in which software runs on individual virtual machines (VM) or to host cloud-native applications built from a collection of AWS services. AWS includes a host of service management, automation, security, network and monitoring services used to deploy, scale, decommission, audit and administer one's AWS environment, subscriptions and hosted applications.

Use Cases

Security and Compliance: Security data from AWS services includes login and logout events and attempts, API calls and logs from network and web application firewalls.

IT Ops: AWS services provide similar types of system and service data as traditional IT infrastructure, much of which is consolidated by the CloudWatch service. These include service monitoring, alarms and dashboards for metrics, logs and events generated by other AWS resources and applications. Typical events and measures include when instances are instantiated and decommissioned, CPU usage, network traffic and storage consumption.

Google Cloud Platform (GCP)

Use Cases: Security and Compliance, IT Operations

Examples: Stackdriver

GCP is a popular and widely used public cloud infrastructure, providing on-demand compute, storage, database, big data and application services with consumption-based pricing. GCP can be used to replace traditional enterprise virtual server infrastructure in which software runs on individual VMs, or to host cloud-native applications built from a collection of GCP services. GCP includes a host of service management, automation, security, network and monitoring services used to deploy scale, decommission, audit and administer one's GCP environment, subscriptions and hosted applications.

Use Cases

Security and Compliance: Security data from GCP services includes login and logout events and attempts, API calls and logs from network and web application firewalls.

IT Ops: GCP services provide similar types of system and service data as traditional IT infrastructure, much of which is consolidated by Stackdriver. These include service monitoring, alarms and dashboards for metrics, logs and events generated by other GCP resources and applications. Typical events and measures include when instances are instantiated and decommissioned, CPU usage, network traffic and storage consumption.

Microsoft Azure

Use Cases: Security and Compliance, IT Operations

Examples: WADLogs, WADEventLogs, WADPerformanceCounter, WADDiagnostInfrastructure

Azure is a popular and widely used public cloud infrastructure, providing on-demand compute, storage, database, big data and application services with consumption-based pricing. Azure can be used to replace traditional enterprise virtual server infrastructure in which software runs on individual VMs, or to host cloud-native applications built from a collection of Azure services. Azure includes a host of service management, automation, security, network and monitoring services used to deploy, scale, decommission, audit and administer one's Azure environment, subscriptions and hosted applications.

Use Cases

Security and Compliance: Security teams can use Azure service logs to audit and attest to compliance with established policies. Log data also is invaluable for incident forensic analysis, such as identifying unauthorized access attempts from access logs, tracking resources and configuration change events and identifying vulnerabilities in hosts or firewalls.

IT Ops: Azure services provide detailed metrics and logs for monitoring one's infrastructure across the entire technology stack, VMs, containers, storage and application services. The data is useful in maintaining application delivery quality and service levels, measuring user behavior, resource utilization and for capacity planning and cost management.

VIRTUAL INFRASTRUCTURE DATA

Pivotal Cloud Foundry (PCF)

Use Cases: IT Operations and DevOps

Examples: Loggregator, PCF Healthwatch

Pivotal Cloud Foundry is a platform-as-a-service (PaaS) built on top of Cloud Foundry, an open source cloud computing platform that allows developers to easily deploy, operate and scale cloud-native applications. Enterprises can manage the entire application lifecycle, from packaging to deployment to execution, as Cloud Foundry supports many cloud frameworks and application languages. With PCF, the installation and administration of cloud-native applications is simplified with capabilities around infrastructure management and provisioning, OS patching, container orchestration, security and more.

Use Cases

IT Ops and DevOps: Operations teams can use PCF metrics, much of which is consolidated via the Loggregator Firehose to gain insights into deployment health, capacity needs and application health before end users are impacted by degraded performance. Since PCF allows DevOps to run their applications on any cloud rapidly and to scale on demand, PCF data is critical for teams to get the end-to-end visibility into the entire lifecycle and visibility between each individual component. When it comes to operating PCF deployments at scale, understanding performance relies on dependencies among the various layers within the app, container and larger architecture.

VIRTUAL INFRASTRUCTURE DATA

VMware Server Logs, Configuration Data and Performance Metrics

Use Cases: Security and Compliance, IT Operations

Examples: vCenter, ESXi

VMware vSphere ESXi is the most commonly used enterprise server virtualization platform. The VMware management platform, whether one of the vSphere products or standalone hypervisor, produce a variety of data and fall into four main categories:

- **vCenter Logs:** vCenter is the "control center" of a vSphere environment. The vCenter logs show information including: who is logging in to make changes, which individuals made changes and authentication failures.
- **ESXi Logs:** Every vSphere environment includes one or more ESXi hypervisors; these are the systems that host the virtual machines. ESXi logs contain information that is useful when troubleshooting hardware and configuration issues.
- **Inventory Information:** the vCenter environment tracks configuration about a number of configuration items including: hypervisors, virtual machines, datastores, clusters and more. This includes the configuration of each item, and how a given item relates to any other. This information is not represented in the log files from either the vCenter or ESXi servers. This information can be viewed using the vSphere client or by using vSphere APIs to pull this information. In both cases this information is pulled from the vCenter servers.
- **Performance Information:** for each configuration item, the vCenter server tracks a number of performance metrics about that item. Datastore latency, virtual or physical CPU utilization, and over 100 other metrics fall into this category. As with the inventory information, this information is not present in the log files and must be viewed through the vSphere client or polled through the vSphere API.

Use Cases

Security and Compliance: The uncoupled nature of virtual resources and underlying physical hardware can cause complex challenges during incident investigations, capacity analyses, change tracking and security reporting. One common security use case for VMware data comes from the vCenter logs, which audit the activity of individuals using the vSphere interface to re-assign user permissions within the VMware environment.

IT Ops: Operations teams can use VMware data to measure the health of the overall hypervisor environment and underlying guest operating systems. Admins can use this data for capacity planning, and for troubleshooting of ongoing performance issues, such as datastore latency issues.

This data also records hardware resource usage that can be used to optimize VM deployments across a server pool to maximize resource consumption without having workloads overwhelm any given server.

VIRTUAL INFRASTRUCTURE DATA

Physical Infrastructure Data

Backup

Use Case: IT Operations

Despite the use of data replication to mirror systems, databases and file stores, data backup remains an essential IT function by providing for long-term, archival storage of valuable information, much of which has legal and regulatory requirements regarding its preservation. Backups also can be used to store multiple versions of system images and data, allowing organizations to reverse changes, accidental deletions or corrupted data quickly, restoring the system or database to a known good state. Backup software can use different types of storage media depending on the likelihood of needing the data: external disks or virtual tape libraries for active data and tape, optical disks or a cloud service for long-term storage.

Use Cases

IT Ops: Backup systems routinely log activity and system conditions, recording information such as job history, error conditions, backup target and a detailed manifest of copied files or volumes. This data allows operations teams to monitor the health of backup systems, software and jobs; triggers alerts in the case of errors; and assists in debugging backup failures. It also allows teams to locate where specific data may be stored, when a recovery is required.

Environmental Sensors

Use Cases: Internet of Things, Business Analytics

Examples: Bosch Sensortec, Mouser Electronics, Raritan, Schneider Electric, TSI, Vaisala

Environmental sensors provide data on barometric air pressure, humidity, ambient air temperature and air quality. They are applied in everything from combating pollution and detecting gasses to keeping data centers from overheating.

Use Cases

Internet of Things: Environmental sensors are a class of smart meters that have been optimized to monitor the environment. In some instances, such as a data center, the information provided by these sensors is used to automatically alter temperature setting and heat flow.

Business Analytics: Environmental sensor data collect can be used in retail applications capable of answering predictive questions, such as "what impact inclement weather might have on foot traffic in a mall?"

Industrial Control Systems (ICS)

Use Cases: Security and Compliance, Internet of Things, Business Analytics

Examples: ABB, Emerson Electric, GE, Hitachi, Honeywell, Rockwell Automation, Siemens, Toshiba

Within the context of a manufacturing environment, industrial control systems make use of programmable logic controllers to both acquire data and execute supervisory functions. Much of the process automation employed in a manufacturing facility is enabled by the industrial control systems.

Use Cases

Security and Compliance: Industrial control systems play a critical role in delivering services to industry and municipalities across the world. These systems live on top of traditional IT infrastructure and — while typically separate from enterprise IT — digital transformation is driving organizations to provide connectivity to these systems, increasing exposure to attacks. These systems tend to be unmanned from a security perspective. Regardless of how ICS might get attacked or infected, data from ICS devices can provide visibility and can be used to analyze and identify malicious activity and potential threats. This visibility enables companies to measure impact and risk, and associate them with business processes.

Internet of Things: Machine data from ICS can be used to gain real-time visibility into the uptime and availability of critical assets. This enables companies to detect an issue, perform root cause analysis and take preventive action to prevent certain events from happening in the future. Companies are also leveraging machine data from ICS systems to secure these mission-critical assets.

Business Analytics: Organizations can apply machine learning algorithms against the machine data created by industrial control systems to increase productivity, uptime and availability. ICS data can also drive visibility into complex manufacturing processes, helping identify bottlenecks and remove inefficiencies.

PHYSICAL INFRASTRUCTURE DATA

Mainframe

Use Cases: IT Operations

Mainframes are the original business computer: large, centralized systems housing multiple processors, system memory (RAM) and I/O controllers. Despite their 60-year legacy, mainframes still are widely used for mission-critical applications, particularly transaction processing. Although they usually run a proprietary OS, mainframes also can be virtualized to run Unix and Linux or, with add-on processor cards, Windows Server. Mainframes are valued for their bulletproof reliability and security, using highly redundant hardware and resilient, stringently tested software. As such, they appeal to organizations wanting to consolidate workloads onto a small number of systems and that need the added reliability and versatility.

Use Cases

IT Ops: Like other servers, mainframes measure and log numerous system parameters that show their current status, configuration and overall health. Since most mainframe subsystems are redundant, system logs also show non-disruptive hardware failures or anomalous behavior that is predictive of an impending failure. Due to their use for critical applications, mainframes often record application performance data such as memory usage, I/O and transaction throughput, processor utilization and network activity.

Medical Devices

Use Cases: Internet of Things, Business Analytics

Examples: Abbott Laboratories, Apple, Baxter, Boston Scientific, GE, Siemens, St. Jude Medical

Everything from intensive care units to wearable devices generates multiple types of machine data. In fact, just about every aspect of patient care inside and out of a hospital setting can be instrumented. While the primary goal is to save lives, a crucial secondary goal is to reduce healthcare costs by reducing both the number of potential visits to a hospital as well as the length of stay.

Use Cases

Internet of Things: Most devices inside a hospital are connected to local monitoring applications. But it's possible to monitor patient care remotely using sensors that communicate with either a wearable device or some other system for monitoring patients in their homes.

Business Analytics: Machine data also makes it simpler for medical professionals to analyze both patient and anonymous data across a broader range of geographically distributed regions — for example, to see how certain diseases are affecting a group of people more than another.

Metric Line Protocols

Use Cases: IT Operations, Application Delivery, Internet of Things

Examples: collectd, statsd

Metrics are measurements generated by a process running on a system that provide a regular data point around a given metric, such as CPU utilization. Metrics data sources generate measurements on regular intervals and generally consist of:

- Timestamp
- Metric Name
- Measurement (a data point)
- Dimensions (that often describe the host, kind of instance, or other attributes that you might want to filter or sort metrics on)

Metrics are typically generated by a daemon (or process) that runs on a server (OS), container, application. Each data measurement is delivered by a network protocol, such as UDP or HTTP, to a server that indexes and analyzes that information.

Metrics are particularly useful for monitoring. For example, a heart monitor that regularly checks a patient's pulse, metrics provide insight into trends or problems that affect the performance and availability of infrastructure and application. However, a heart monitor won't tell you why a patient has a sudden issue with their heart rate - you need other means to quickly identify the cause and stabilize the patient. It's the same with machine data. When combined with other data sources, usually logs, you gain insight into both what's going on, and why it's happening.

Examples of Metric Line Protocols

collectd: Collectd is a protocol that involves an agent running on a server that is configured to measure specific attributes and transmit that information to a defined destination. Collectd is an extensible measurement engine, so you can collect a wide range of data. Currently, collectd is most often used for core infrastructure monitoring insights, such as getting insight on the workload, memory usage, I/O, and storage of servers and other infrastructure components. Collectd is part of the open source community, and you can learn much more about collectd by visiting **http://collectd.org**.

statsd: is a network daemon that runs on node.js. It has gained popularity with windows administrators, application performance experts and others. Statsd provides some capabilities that allow for metrics to be delivered in batch, and while it uses the less dependable UDP network method, many administrators like how easy it is to deploy. Much like collectd, statsd is focused on collecting metrics, mostly involving the usage and performance of applications and application components, and sending them via the network to a tool that can collect and analyze that information.

Use Cases

IT Ops and Application Delivery: Metrics Line Protocols provides usage, performance and availability data across operating systems, storage devices, applications and other components of IT infrastructure. Metrics are particularly useful for the monitoring portion of IT Operations and Application Delivery, where trends can help identify where there is a problem. Once trends and thresholds illustrate performance issues, other data sources are often correlated to determine the root cause of the problem.

Internet of Things: As devices become more intelligent, more metrics based telemetry will be on board. Metrics line protocols represent an efficient way for these devices to report their status and performance.

PHYSICAL INFRASTRUCTURE DATA

Patch Logs

Use Cases: Security and Compliance, IT Operations

Keeping operating systems and applications updated with the latest bug fixes and security patches is an essential task that can prevent unplanned downtime, random application crashes and security breaches. Although commercial apps and operating systems often have embedded patching software, some organizations use independent patch management software to consolidate patch management and ensure the consistent application of patches across their software fleet and to build patch jobs for custom, internal applications.

Patch management software keeps a patch inventory using a database of available updates and can match these against an organization's installed software. Other features include patch scheduling, post-install testing and validation and documentation of required system configurations and patching procedures.

Use Cases

Security and Compliance: Security teams can use patch logs to monitor system updates and determine which assets could be at risk, due to failed or out-of-date patches.

IT Ops: Operations teams use patch logs to verify the timely and correct application of scheduled patches, identify unpatched systems and applications, and alert to errors in the patching process. Correlating errors to patch logs can indicate when an error is due to a patch.

Physical Card Readers

Use Case: Security and Compliance

Most organizations use automated systems to secure physical access to facilities. Historically, these have been simple magnetic strips affixed to employee badges; however, locations with stringent security requirements may use some form of biometric reader or digital key. Regardless of the technology, the systems compare an individual's identity with a database and activate doors when the user is authorized to enter a particular location. As digital systems, badge readers record information such as user ID, date and time of entry and perhaps a photo for each access attempt.

Use Cases

Security and Compliance: For IT security teams, the data from card readers provide the same sort of access information for physical locations as a network firewall log. The data can be used to detect attempted breaches and be correlated to system and network logs to identify potential insider threats and provide overall situational awareness. It can also be used to detect access at unusual times and locations or for unusual durations.

Point-of-Sale Systems (POS)

Use Cases: Security and Compliance, Internet of Things, Business Analytics

Examples: IBM, LightSpeed, NCR, Revel Systems, Square, Toshiba, Vend

Point-of-sale systems are most often associated with transactions generated at a retail outlet. However, thanks to the rise of mobile POS solutions, many of these systems are starting to be deployed in temporary locations, such as a community fair or a high school event.

The typical POS system incorporates a cash register based on a PC or embedded system, monitor, receipt printer, display, barcode scanner, and debit/credit card reader. Machine data generated by POS systems provides organizations with real-time insight into everything from what's sold, to the amount of cash being generated per transaction, to what payment methods are being used.

Use Cases

Security and Compliance: POS systems are typically used for financial transactions and are often targeted since they contain account, payment and financial information. Because the POS transaction information is highly sought after for its value to attackers, and the POS can be used as an entry point to the network, it's critical to protect these systems. Furthermore, POS systems are usually unmanned, run an underlying operating system, and versioning/monitoring typically fall outside of IT's purview — adding additional complexity to their security. Visibility and analysis of POS systems and data can provide insights that are critical to protecting financial information, detecting fraud and securing vulnerabilities.

Internet of Things: Historically, POS systems were either not connected or managed on a dedicated private network. Thanks to the rise of the IoT, these systems are being connected directly to cloud platforms that make remotely administering these devices from a central location much simpler. There's no longer a need to dispatch IT personnel to manually update each system. This is critical because a POS failure can result in longer lines that inconvenience customers and potentially lead to lost revenue. A negative customer experience can easily translate to customers opting to shop somewhere else in a retail industry that is intensely competitive.

Business Analytics: POS systems contain information about what's sold, how it's paid for, as well as the pace at which it's being sold. Organizations can use this data to monitor revenue in real time, which can feed into how to better market 1:1 against customers, track product placement and sales in a store, or detect potentially fraudulent transactions in real time. This type of real time Big Data analysis can have a profound impact on customers cross- and up-sell opportunities. POS data also delivers visibility into customer experience such as which coupons are most popular or the combinations of products that are selling together. When enriched with geolocation data, it can also drive valuable insights into location-based analytics.

PHYSICAL INFRASTRUCTURE DATA

RFID/NFC/BLE

Use Cases: Internet of Things, Business Analytics

Examples: Alien Technology, BluVision, CheckPoint Systems, Gimbal, MonsoonRF, Radius Networks, STMicroelectronics, TAGSYS RFID, ThingMagic

The two primary wireless methods organizations use today to keep track of objects and interact with customers in retail stores involve two distinct types of wireless communications technologies. The better known is radio-frequency optimization (RFID), which involves the use of tags capable of storing information such as product information or what goods might be loaded in a shipping container.

At the same time, organizations are adopting Bluetooth Low Energy (BLE) wireless connectivity solutions that can broadcast signals to other devices. BLE is used most widely in beacons that are employed, for example, to inform shoppers of new sales in retail stores on their smartphones or update fans on events that might be occurring during a sporting event.

Use Cases

Internet of Things: RFID is arguably one of the first instances of an IoT application. Deployed in place of traditional barcode readers, RFID tags are used in everything from shipping to keeping track of farm animals. IoT deployments make it possible to capture RFID data in a way that makes it simpler to track events involving anything that has an attached RFID tag. Data insights from RFID can help improve overall supply chain, order processing and inventory management.

BLE, meanwhile, is used to engage customers more directly as they move about a specific location, which in turn creates data that can be used to optimize the customer experience.

Business Analytics: Whether it's inventory tracked using RFID tags or customers and employees moving around specific locations, new classes of analytics applications are using the data generated by these devices to serve up actionable business insights in near real time. Retailers can leverage this data for several use cases, such as making sure that inventory is located as close as possible to the locations where customers are most likely to want to purchase.

PHYSICAL INFRASTRUCTURE DATA

Sensor Data

Use Cases: Security and Compliance, IT Operations, Internet of Things, Business Analytics

Examples: Binary and numeric values including switch state, temperature, pressure, frequency, flow, from MQTT, AMQP and CoAP brokers, HTTP event collector

Industrial equipment, sensors and other devices often have embedded processors and networking that allows them to record and transmit a vast array of information about operating conditions. Regardless of device, their data provides unprecedented detail about performance parameters and anomalies that can indicate larger problems — for example, a device ready to fail or issues with another system. Aggregating and correlating data from multiple devices and subsystems provides a complete picture of equipment, system, factory or building performance.

Use Cases

Security and Compliance: Sensor data can help protect mission-critical assets and industrial systems against cybersecurity threats by providing visibility into system performance or set points that could put machines or people at risk. Data can also be used to satisfy compliance reporting requirements.

IT Ops: Some of the most important parameters for operations teams to monitor are environmental conditions such as temperature, humidity, airflow and voltage regulation in a data center. Similar readings are available from individual servers and network equipment that, when correlated, can highlight problems in the facility or equipment ready to fail.

Additional Use Cases

Preventative Maintenance and Asset Lifecycle Management: Sensor data can provide insights into asset deployment, utilization and resource consumption. Operational data can also be used to proactively approach long-term asset management, maintenance and performance.

Monitoring and Diagnostics: Monitoring sensors can help ensure that equipment in the field operates as intended, for example, monitoring and tracking unplanned device or system downtime. The data can also be used to understand the cause of failure on a device to improve efficiency and availability, and to identify outliers and issues in device production or deployment.

PHYSICAL INFRASTRUCTURE DATA

Server Logs

Use Cases: Security and Compliance, IT Operations, Application Delivery

Server operating systems routinely record a variety of operational, security, error and debugging data such as system libraries loaded during boot, application processes open, network connections, file systems mounted and system memory usage. The level of detail is configurable by the system administrator; however, there are sufficient options to provide a complete picture of system activity throughout its lifetime. Depending on the subsystem, server logs are useful to system, network, storage and security teams.

Use Cases

Security and Compliance: Server logs include data from security subsystems such as the local firewall, login attempts and file access errors that security teams can use to identify breach attempts, track successful system penetrations and plug vulnerabilities. Monitoring server logs such as file access, authentication and application usage can help secure infrastructure components.

IT Ops and Application Delivery: Server logs provide a detailed record of overall system health, and forensic information about the exact time of errors and anomalous conditions that are invaluable in finding the root cause of system problems.

Smart Meters

Use Cases: Internet of Things, Business Analytics

Examples: ABB, GE, Google, eMeter, IBM, Itron, Schneider Electric, Siemens

Smart meters record consumption of energy, usage of water, or usage of natural gas so that the information can be continually processed and shared. Typically, smart meters allow for bi-directional communication in real time in a way that allows a gauge of some type to be adjusted.

Use Cases

Internet of Things: Smart meters are deployed across critical systems at large utilities companies, for example, power, gas and water utilities. These systems are the lifeblood of infrastructure and failure can lead to catastrophic outcomes. Real time monitoring of smart meters can help organizations better analyze failures remotely, by way of detecting remotely line down failures. Equally important is securing the devices from tampering that could lead to malicious attacks and breaches.

Energy companies and water utilities make extensive use of smart sensors to track everything from oil reserves to the quality of the water supply.

Business Analytics: A wide variety of industries are applying analytics to the data being collected by smart meters to optimize service. For example, an oil or gas company no longer needs to physically send a worker to a location to read a meter. The provider already knows how much fuel has been consumed and how much remains.

Smart meters in the future will be used in everything from modern traffic control systems to defense systems designed to protect critical infrastructure. Aggregating data from these smart meters can give utilities critical insights into the demand. Heavily regulated utilities are required to meet established SLA's during demand response events, and machine data from smart meters can drive visibility into how they are responding.

Storage

Use Case: IT Operations

Examples: EMC, Netapp, IBM, Amazon EBS

Data center storage is provisioned in two general ways: built into servers and shared using various network storage protocols, or via a dedicated storage array that consolidates capacity for use by multiple applications that access it using either a dedicated storage area network (SAN) or ethernet LAN file-sharing protocol. The activity of internal, server-based storage is typically recorded in system logs, however storage arrays have internal controllers/storage processors that run a storage-optimized OS and log a plethora of operating, error and usage data. Since many organizations have several such arrays, the logs often are consolidated by a storage management system that can report on the aggregate activity and capacity.

Use Cases

IT Ops: Shared storage logs record overall system health (both hardware and software), error conditions (such as a failed controller, network interface or disks) and usage (both capacity used per volume and file or volume accesses). Collectively, the information can alert operations teams to problems, the need for more capacity and performance bottlenecks.

Telephony

Use Cases: IT Operations

Examples: Cisco Unified Communications Manager, Shoretel, Twilio

Real-time business communications are no longer limited to voice calls provided by plain old telephone service (POTS); instead, voice, video, text messaging and web conferences are IP applications delivered over existing enterprise networks. Unlike traditional client-server or web applications, telephony and other communications applications have strict requirements on network quality of service, latency and packet loss, making service quality and reliability much more sensitive to network conditions and server responsiveness. Traditional POTS has conditioned people to expect immediate dial tone when picking up the phone and be intolerant of noise, echo or other problems that can plague IP telephony; as such, the systems and supporting infrastructure require careful monitoring and management to assure quality and reliability.

Use Cases

IT Ops: Like VoIP, telephony logs provide an overview of system health along with troubleshooting and usage data similar to that of other network applications. Details include source, destination, time and duration of voice/video calls, web conferences and text messages, call-quality metrics (e.g., packet loss, latency, audio fidelity/bit rate), error conditions and user attendance at web conferences. By integrating telephony records of source/destination address with an employee database such as AD or LDAP and a DHCP database, organizations can link call records to actual user IDs and IP addresses to physical locations; information that can assist in troubleshooting and billing. Logs also can reveal any network segments experiencing congestion or other performance problems that may indicate equipment problems or the need for an upgrade.

PHYSICAL INFRASTRUCTURE DATA

Transportation

Use Cases: Internet of Things, Business Analytics

Examples: Boeing, BMW, Ford, GE, General Motors, Daimler-Benz, John Deere, Volkswagen

Vehicles of all sizes and types generate massive amounts of machine data every day that can be used to gain real-time visibility into the health and performance of an asset, and to drive predictive maintenance applications. Armed with that data, an airplane or automobile manufacturer can follow a maintenance regime that is data driven rather than driven "by the book."

That information can then be used to improve availability and reliability, and extend the life cycle of a vehicle that has not been extensively used or, conversely, replace components that have seen extensive wear and tear sooner.

Use Cases

Internet of Things: Vehicle manufacturers are attaching sensors to every mechanical and electronic component they use. This allows companies to gain a unified view of assets to quickly identify and diagnose operational issues, and to monitor, track and avoid unplanned asset downtime. This helps to ensure that equipment is operating as intended. They can also detect anomalies and deviations from normal behavior to take corrective action — improving uptime, asset reliability and longevity.

Business Analytics: With access to machine data, vehicle manufacturers are applying analytics in ways that fundamentally changes their business models. Instead of selling a vehicle, manufacturers increasingly prefer to lease vehicles based on actual usage. The longer that vehicle can be used between repairs, the more profitable that leasing service becomes. The key to providing this type of service economically is advanced analytics, which are applied to all the aggregate data that's collected.

Wearables

Use Cases: Internet of Things, Business Analytics

Examples: ARM, Intel, Lenovo, Microsoft, Samsung

From smartwatches that double as fitness aids to medical devices that enable physicians to remotely monitor vital statistics, wearable devices have proven they are here to stay. Wearable devices are one of the most recognizable parts of the Internet of Things.

Use Cases

Internet of Things: Beyond merely syncing with smartphones, the latest generation of smartwatches is taking advantage of geo-positioning systems and application programming interfaces to give device owners an optimal application experience that includes both their location and often time of day.

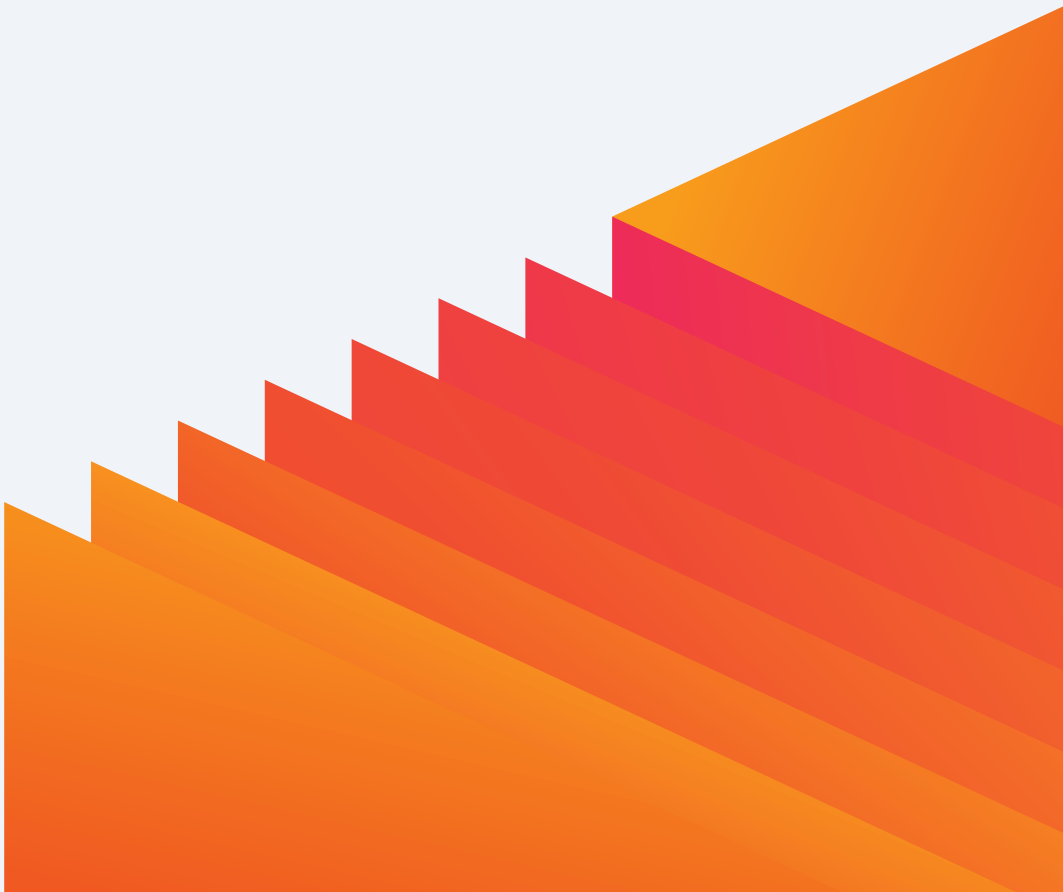
Going forward, there soon will be whole new classes of wearable devices taking advantage of everything from virtual reality applications delivered via a headset to sensors embedded in the latest fashion.

Business Analytics: As more people become comfortable with sharing data via wearable devices, many are experiencing the power of analytics firsthand. Developers of applications optimized for wearables are making recommendations concerning everything from how to improve life expectancy to where to find a meal. Analytics from wearables can help improve user experience and drive product innovation. For example, product managers can understand how consumers are interacting with devices to build better features.

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-Infrastructure-Data-104