the modern enterprise data fabric: helping the public sector transform digitally

splunk>

industry perspective

executive summary

Governments at all levels have been adopting digital technologies in large numbers over the past few decades to fulfill their missions – be it protection of the country, welfare of the citizen or even to ensure efficacies across their agencies

But nothing could have prepared them for the new digital transformation that is underway. While promising to deliver unprecedented benefits, including meeting rising citizen expectations, real-time situational awareness and agency efficiencies, this transformation led by innovative technologies is fundamentally changing how an agency functions.

An outcome of this transformation is the amount of data that is being created, which is extraordinary by any measure. To put it in some perspective, numbers released by IBM found that an equivalent of 90 percent of all the data in the entire world, and in fact in all of human history, has been generated in the past two years. As an example, a connected airplane will generate 40 terabytes of data during a single flight, while an automated factory could produce up to a petabyte every day. And this is happening across all sectors, with a Gartner report revealing 50 percent of CEOs expressing confidence that their industries will be digitally transformed by 2020.

Organizations that can manage this digital transformation and the massive amount of data it generates will thrive. Those that can't will have a difficult time surviving these changes, risking failed missions and tarnished reputations.

In this industry perspective, GovLoop and Splunk, provider of a leading software platform for real-time Operational Intelligence, have partnered to discuss what digital transformation means for the public sector, how the data deluge is adding to the complexity and how a segment of big data, called machine data, is offering opportunities to ride the digital transformation wave seamlessly and effectively.

digital transformation in the public sector

In the public sector, profits are not the driving factor; citizen welfare and protection are.

And digital transformation is critical to these efforts. Budgets are tight and agencies are constantly driven to do more with less. Finding ways to efficiently interact with the public as well as ways to optimize complex processes or improve workflows can go a long way to better serve citizens, improve public safety, handle emergencies and tackle a whole lot of other issues.

Data can help in this arena. And in many ways, the public sector is generating even more data than before. Data in government can be generated by almost anything, like a police officer's body camera, an agency's social media account, website servers, firewalls, traditional desktop computers, VoIP phone systems and the millions of mobile devices on the job as part of a BYOD or official mobility program.

And that does not even consider the Internet of Things (IoT), which are being used for everything from checking the health of street lights to monitoring how many times the courthouse door is opened every day. Those tiny little sensors only record one or two data points, but can clog almost any collection or analysis system by their sheer numbers, which only continue to grow.

While the data growth is immense as are its velocity and variability, harnessing and analyzing this data can improve efficiencies, enhance security and safety, deliver faster services, increase citizen satisfaction and fortify agency reputation.

machine data helps navigate the digital transformation storm

Machine-generated data is one of the fastest-growing and most complex areas of big data.

While this transformation can seem daunting and bring its own challenges, the technologies powering this journey leave a digital footprint - the digital exhaust of all activities initiated and performed by users, systems, applications and devices. They help connect the dots between the various activities to get a fuller, more holistic picture of what is happening across the agency. This is what is called machine data and if harnessed properly, can help ride this transformation wave and solve a myriad of challenges, in real-time, like never before. In fact, machine data is the fastest growing and most complex segment of big data.

Going beyond log files, machine data includes configurations, data from APIs, message queues, call detail records, sensor data and more. In a modern network, machine data is generated from everything - user transactions, customer behavior, system behavior, security events, fraudulent activity and more. And all those IoT sensors embedded in everything from bridges to cars to HVAC systems add their own machine data to this mix.

As ubiquitous as machine data is, and as massive as those data sets quickly become, they are also the most valuable pieces of information in an organization. If properly collected and analyzed, they can create endless opportunities to improve network efficiency, ensure operational performance, offer real-time situation awareness and transform customer interactions. And it can be used to spot trouble, either from failing equipment or outside manipulation – but only if properly used.

Properly collecting that data and deciding what is important is no easy task. Every nugget is potentially valuable. With machine data, getting value is not like finding a needle in a haystack; it's more like finding a needle in a stack of needles.

In previous papers on this topic, GovLoop and Splunk have examined how using the Splunk platform can help in two major areas within the public sector: security and IT operations. For security, Splunk can help turn machine data into Operational Intelligence, providing new levels of visibility, information and insight. Using analytics and intelligence drawn from machine data, public-sector cybersecurity teams can gain operational visibility into the layers of their environment and turn data from across silos of operations into cogent, actionable information that allows agencies to respond to incidents faster, hunt down malware, and get proactive with their security posture.

In the same fashion, the Splunk platform can transform IT operations by gathering data from all IT layers, correlating and synchronizing them by time and providing powerful keyword searching so that analysts can connect the dots to identify problems, track end-to-end performance, ensure system uptime, and reduce mean-time-to resolution.

To take advantage of these opportunities and mitigate risks, you need to capitalize on the machine data that is at the core of your agency's work.

You need the ability to collect, analyze, share and provide access to data across your organization to enable mission-critical use cases in real time.

You need an enterprise machine data fabric.

weaving the enterprise data machine fabric

Splunk helps agencies overcome the inherent challenges of machine data by delivering a platform that can collect, correlate, analyze, share and provide access to data across the enterprise.

Edward Tufte, a renowned statistician and professor emeritus of political science, statistics and computer science at Yale University, once remarked that the clutter and confusion associated with information are not attributes of the data – they're design shortcomings. The millions of computers, servers, programs, devices and apps being used in every modern network, as well as the billions of IoT devices that also now contribute to that clutter, were not designed to operate within a single system, or even to speak the same language.

Splunk helps agencies overcome the inherent challenges of machine data by delivering a platform that can collect, correlate, analyze, share and provide access to data across the enterprise. This weaves an Enterprise Machine Data Fabric (EMDF) wherein the same data that's collected can be used to solve a variety of mission-critical use cases across the agency in real-time.

With an EMDF, data is correlated as you stitch it together based on the questions you ask, the story you want to create or insights you want to mine. Because the data remains in its native format and does not have to be modified to fit into a pre-defined database, operators, analysts, managers and executives can ask any questions they want, essentially letting different people ask different questions of the same data to pursue their initiatives and mission. It can thus be weaved into every business function to help agencies solve any type of problem, and provides strong ROI – if you can use the same data and use the same platform to solve multiple problems, then the value is increased.

Fundamentally, the Enterprise Machine Data Fabric:

Can process data in any format from any source and in any ingest mode - batch, streaming, real time, etc.

Ingests, enriches, correlates and analyzes this data for real-time use

Extends to multiple, heterogeneous data stores cutting across silos of operations

Makes it easy to customize with a service-oriented model, accessible developer tools and well defined and long-lived APIs

Leverages a third-party ecosystem of developers building analytics apps to extending investments

Extends to world-class technologies (including open source and ecosystem) for single-pane-of-glass analytics across the enterprise

the future is now

Searching massive volumes of data can be a humanly prohibitive endeavor and still be a reactive proposition in many cases.

If all that the EMDF accomplished was to join diverse types of machine data into a common, searchable format, it would still be a powerful tool to enable teams to address all types of mission-critical challenges in near real-time. But just searching massive volumes of data can be a humanly prohibitive endeavor and still be a reactive proposition in many cases. On top of enabling the enterprise machine data fabric for real-time Operational Intelligence, the Splunk platform extends machine learning by adding outlier and anomaly detection, adaptive thresholding and predictive analytics capabilities using packaged and custom models. As a core capability of the Splunk platform, machine learning lets you operationalize your machine data

Splunk's machine-learning capabilities can help agencies discover potential problems based on the interactions of the millions of pieces of data stored within the unified EMDF. Machine learning can analyze vast amounts of data much faster, helping separate the signal from the noise, focusing the cognitive requirements of inquisition and decision making to humans. It can use anomaly detection, behavior, baseline samples, modeling and other techniques to point out connections and cause-and-effect relationships that would not be obvious or even detectable without it. As the platform learns over time, it gets even more accurate, becoming a unique and valuable expert that is laser-focused on the specific problem at hand.

real-world examples of the EMDF at work

U.S. Immigration Services:

The United States' Department of Citizenship and Immigration Services (USCIS) is a Splunk user. The agency has been able to leverage the platform to solve a number of mission critical challenges. The security team uses Splunk to conduct kill chain analysis – while previously they did not have all the data to conduct such an analysis they are now able to do this much faster with end-to-end visibility. The agency also uses Splunk for IT troubleshooting, ensuring that systems that underpin their mission critical services are up. Splunk also provides them with a complete view of assets on their network helping them with inventory, their state and visibility into items that they did not even know existed.

Denver Water:

Denver Water promotes the efficient use of water to 1.3 million people in the city. Given this scope, supporting the infrastructure and monitoring applications was a challenge. A deluge of machine data from logs and databases often overwhelmed IT administrators, hampering efforts to pinpoint problems when users notified the help desk. Aside from measuring and regulating water quality to its customers, with Splunk Enterprise, Denver water is now being proactive to technical errors and help desk tickets, the utility's IT team now can quickly see failure trends and address them ahead of time. In addition to troubleshooting, the same machine data is helping the utility get a better picture of how their applications are being used by employees.

Sacramento Sheriff's Department:

The Sacramento County Sheriff's Department is responsible for law enforcement services for unincorporated areas of Sacramento County, CA, and several incorporated cities within the county. The department implemented Intelligence-Led Policing (ILP), a strategy that depends on high-quality data analysis, and needed a way to harness data that existed in disparate, siloed systems. Since deploying Splunk Enterprise, the Sheriff's Department has seen benefits, including: Integrated and visualized crime, management and corrections data; enhanced accountability, helping to reduce crime report backlog; and supported proactive policing based on big data analytics.

conclusion

No organization can tame the data maelstrom that digital transformation is generating without fundamentally changing the way they interact with information. And public-sector agencies, with their diverse array of devices producing mountains of data, are especially vulnerable.

The solution is to take that data and weave it into a common fabric so that different people who are asking different questions can all get answers from the same data.

The Splunk platform extending powerful EMDF and machine learning capabilities can help turn opportunities for digital transformation into an incredible advantage and a powerful tool to help agencies better serve the citizens who rely on them for services, information and support.

about splunk

Splunk Inc. provides the leading platform for operational intelligence. Splunk® software searches, monitors, analyzes and visualizes machine-generated big data from websites, applications, servers, networks, sensors and mobile devices. More than 11,000 organizations use Splunk software to deepen business and customer understanding, mitigate cybersecurity risk, improve service performance and reduce costs.

about govloop

GovLoop's mission is to "connect government to improve government." We aim to inspire public-sector professionals by serving as the knowledge network for government. GovLoop connects more than 250,000 members, fostering cross-government collaboration, solving common problems and advancing government careers. GovLoop is headquartered in Washington, D.C., with a team of dedicated professionals who share a commitment to connect and improve government.

For more information about this report, please reach out to info@govloop.com.

www.govloop.com | @GovLoop

govloop

1152 15th St. NW, Suite 800
Washington, DC 20005

Phone: (202) 407-7421 | Fax: (202) 407-7501

www.govloop.com
@GovLoop