


Guide to Operational Resilience


eBook


Thank you for downloading this Elastic eBook. Carahsoft is the Master Government Aggregator and Distributor for Elastic Open Source solutions available via GSA, NASPO, NJSBA, and other contract vehicles.


To learn how to take the next step toward acquiring Elastic's solutions, please check out the following resources and information:


 For additional resources:
carah.io/ElasticResources

 For upcoming events:
carah.io/ElasticEvents

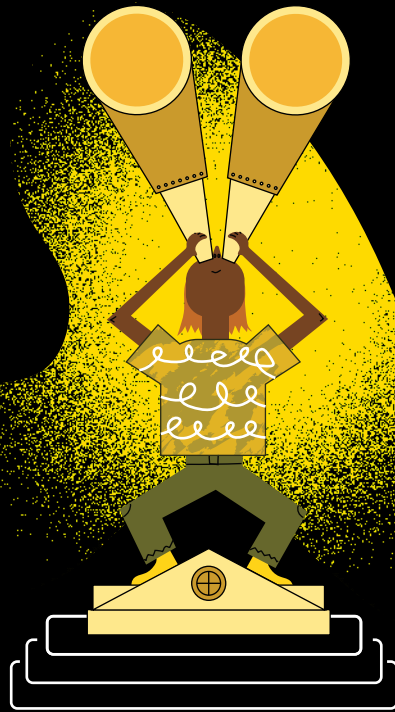
 For additional Elastic solutions:
carah.io/ElasticSolutions

 For additional Open Source solutions:
carah.io/OpenSourceSolutions

 To set up a meeting:
Elastic@carahsoft.com
877-742-8468

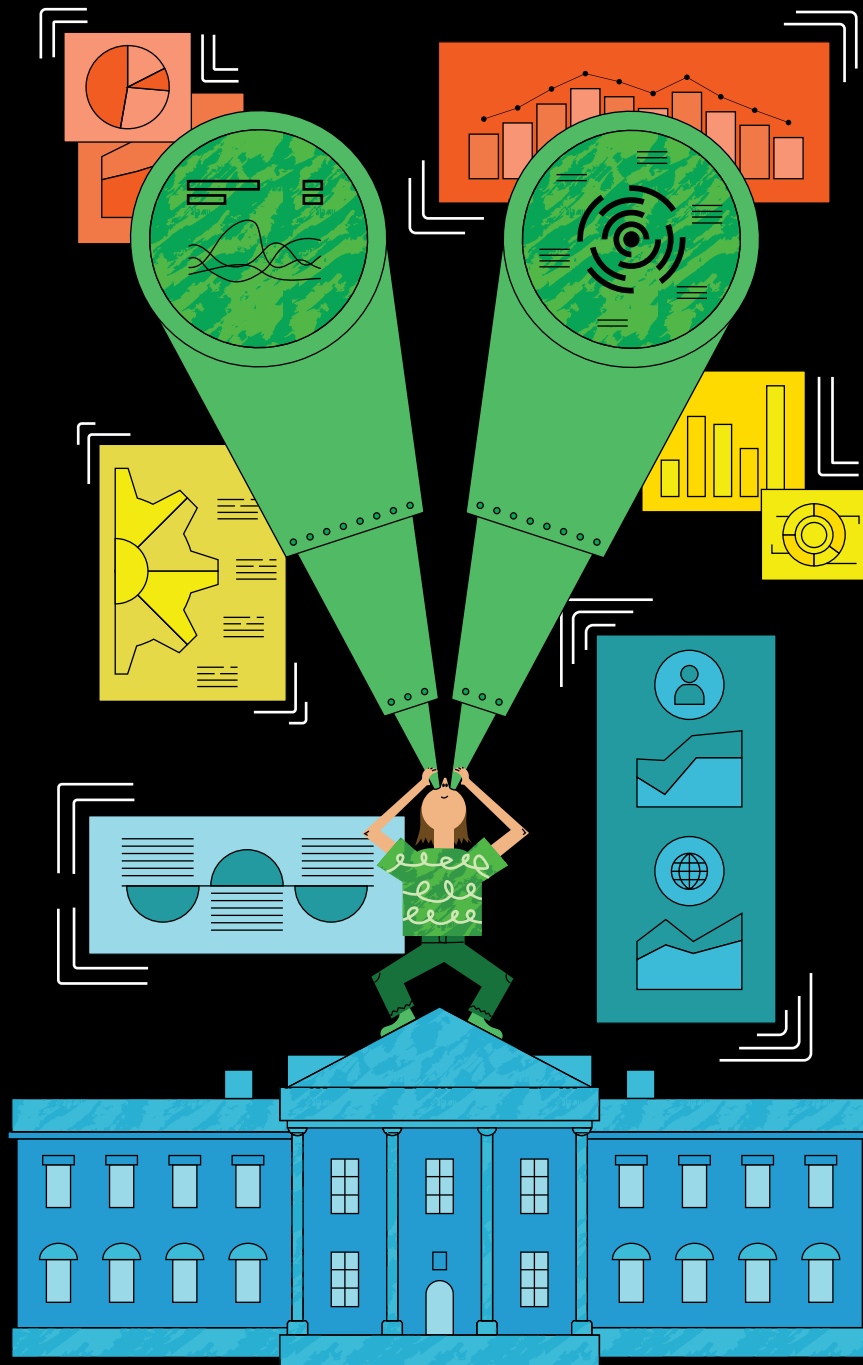
 To purchase, check out the contract vehicles available for procurement:
carah.io/ElasticContracts

For more information, contact Carahsoft or our reseller partners:
Elastic@carahsoft.com | 877-742-8468



How to Strengthen Your Operational Resilience





Here's the key to keeping systems from crashing and disrupting your services: Identify potential problems before they become big problems. And the key to that may not seem exciting, but it's as simple as good log management.

Across the network, systems automatically capture, or log, data on how things are running. But many agencies don't take full advantage of that data. The Biden administration wants to change that.

With its Executive Memorandum M-21-31, the White House laid out a maturity model for event log management, effectively making logging, log retention and log management a cornerstone of IT systems management in government.

Now that most agencies have become proficient at logging, they can make the simple next steps to full end-to-end observability, which lets you use those logs to improve individual app performance.

"Observability helps to fix those problems faster," said Nathan Stacey, Senior Manager for Solution Architects at Elastic, whose data security and observability solutions help teams gain full visibility and protect their valuable data. "The more observability that an agency has, the quicker they can remediate those problems."

Technology is not infallible, and improved observability offers a hedge against prolonged downtimes. Application managers have to manage more complex architectures than in the past, with many moving parts. This requires deeper visibility, beyond just logging, to ensure that when an issue arises, they know exactly what is happening and where, and what the operational impact may be. They need the ability to turn information into insight.

With full-stack observability, IT leaders and application owners are empowered to take a mission-driven approach to systems management. Expanded observability can help ensure not only that systems are kept running, but that the IT team's investigative and remediation efforts are targeted in ways that most effectively support continued mission outcomes.

Observability: An FAQ

What is observability, beyond log analysis?

"In the modern sense, observability is the ability to find the needle in the haystack: to not only gather the needed data on system operations, but to combine it and make it visible in a way that empowers the end user," Stacey said.

Observability is not just monitoring specific IT components but how they work together for the health of the application/mission. At a high level, this is done by looking at four key metrics of that mission: latency, errors, traffic and saturation.

What data informs observability?

Full-stack observability will be informed by telemetry data that is mainly — but not limited to — traces, metrics and logs. Traces (application performance monitoring, or APM) show the activity, the path that requests take through an application. While these are the most complex to retrieve sometimes, they are often the first stop when debugging issues or outages.

Metrics can be counters that increment each time something happens (i.e., when a page loads), they can be accumulators (bytes sent and received) or they can be aggregated or calculated over a period of time (system load).

Since log data may include metrics, agencies that have vigorous logging processes in place have in effect already started down this road.

"All this data is what allows you to assess the network: Are you getting as much traffic as normal? Are you getting packet loss?" Stacey said. "That visibility, in turn, allows you to keep the mission on track."

Why do workflow context and correlation matter?

When troubleshooting an application, the details of a problem rarely show up in the first place you look. The data is often like an iceberg: What can be seen is the warning, and the key details necessary to resolve the concern are hidden under the water. Insight derives from workflow context and correlation, bringing all datasets of that iceberg into a single place, drawing on enriched metadata and its relevant metrics — that is, looking at enriched metadata to jump directly to the relevant metrics.

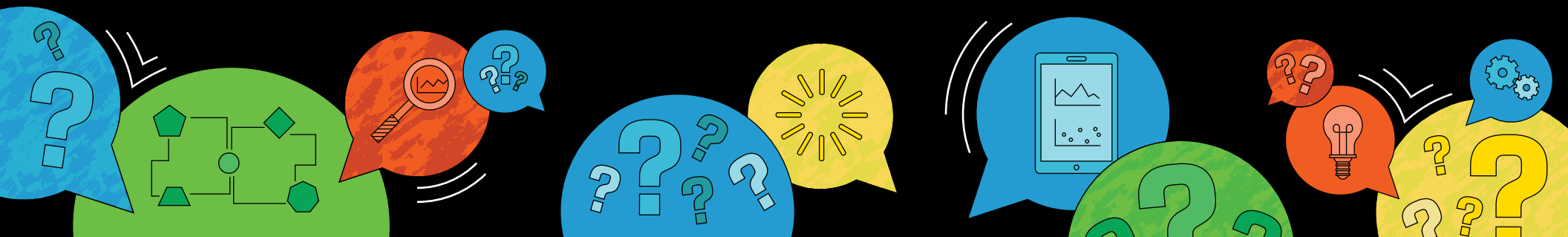
"You get data from the network, from the server, from the application — those are the three key datasets," Stacey said. "All that data will be formatted differently, so you need a common schema to establish that context and correlation. It's the difference between knowing you have a problem, and knowing you have a problem here."

What is the role of AIOps and machine learning?

To identify and remediate abnormal processes, IT teams have to wade through mountains of data. "You might be looking at 10,000 different fields, and there are three that are different than normal," Stacey said. "Artificial intelligence and machine learning enable you to focus on just those three."

This approach, known as AIOps, "then operationalizes the workflow so that things are automatically being rectified," said Tristan Ahmadi, principal solutions architect at Elastic.

"If bottlenecks are detected, if outages are identified, if latency is detected — anything that might be impacting a system or a network — processes kick in automatically, whether that is restarting this server or rerunning these processes in order to rectify the situation," Ahmadi said.





The Playbook

How to Put Observability to Work in Support of Mission Outcomes

With greater observability, federal agencies can ensure uninterrupted operation of mission-critical applications. With a methodical, step-by-step approach, they can achieve greater visibility and more actionable insights across their IT ecosystems.



Start with logging.

The recent [OMB memo](#) establishes the importance of logs in support of digital resilience. It lays out the basic logging categories and minimum logging data that most agencies already have implemented or are working to put in place. Here are three examples of how agencies are responding:

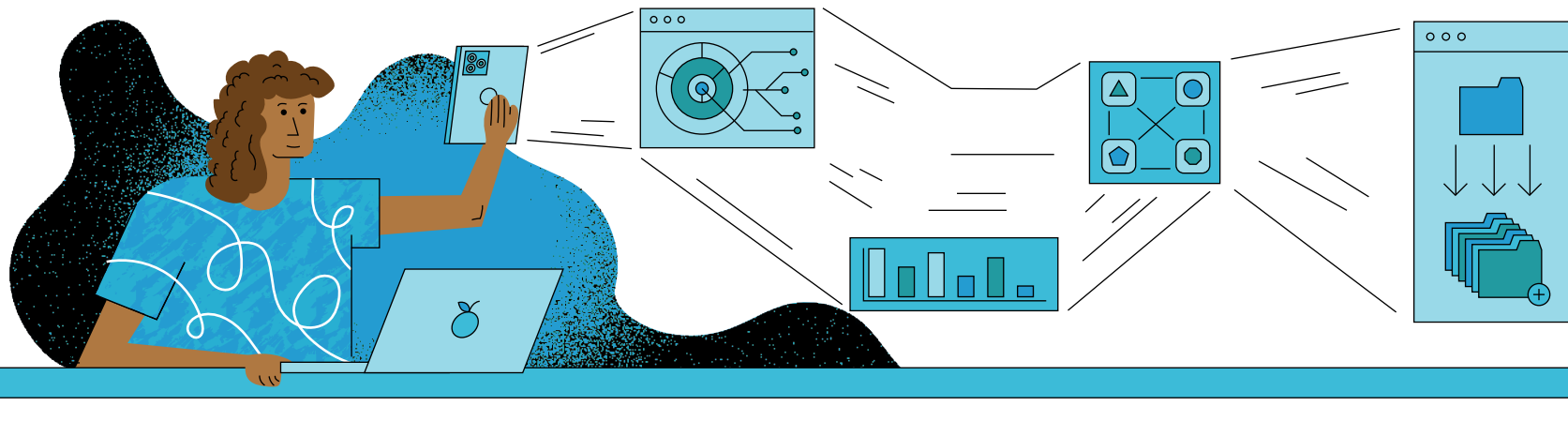
- The Treasury Department, for instance, reports it [is meeting the OMB guidance](#), known as M-21-31, by creating and monitoring traps for detecting data-stream disruption; by sharing the logs with CISA, the Federal Bureau of Investigation and others; and providing storage and retention for log data.
- The Secret Service likewise has been [making progress toward zero trust](#) under the guidance.
- The Justice Department is using the guidance to steer its [forthcoming budget request](#).

When applications crash, this logging information is by nature forensic: It tells you what happened, not what is happening. To get the most value out of such information, agencies need to implement tools and processes that enable fast access to historical data, no matter how old it is or where it's stored, via a single query.

One approach is to consider log data in a series of tiers, a collection of nodes with the same data role that typically share the same hardware profile. The newer data is viewed more than older data and thus should be easier to access. "Hot" tiers are used to bring in the logs and view them for the first few days. The logs can then transfer to "cold" tiers with a smaller cache ratio to the logs, queries will be slower but fewer queries are needed for older data.

To speed response to system failures, agencies can look to leverage solutions that include a "frozen tier" — nodes that store partially mounted indices of searchable snapshots. Such ready access to historical data "allows you to better understand the history of how a machine performs, what high disk utilization looks like beyond a certain time," Ahmadi said.

In tracking down the root cause of a problem, "it's valuable to have the historic data all available holistically, all in one spot," Stacey said. With tools like frozen tiering, "you don't have to silo the data anymore. You have the ability to bring the silos together in a single place."



Move on to app performance.

With robust logging capabilities in place, agencies can begin to more effectively leverage the data that informs quick remediation and keeps the mission applications on track.

In this phase, agencies will look at commonly available observability metrics, specifically CPU, memory, storage and network flow data. These logs can be retrieved in many different ways depending on the organization's permission policies. Third-party (product/cloud/vendor/application) specific datasets are also grabbed in this stage and are done by the specification of the vendor. All of these logs are tied together with APM to see the actual performance of the application.

"APM is a complex animal," Stacey said. "This is very valuable data, but it takes more time to set up and maintain."

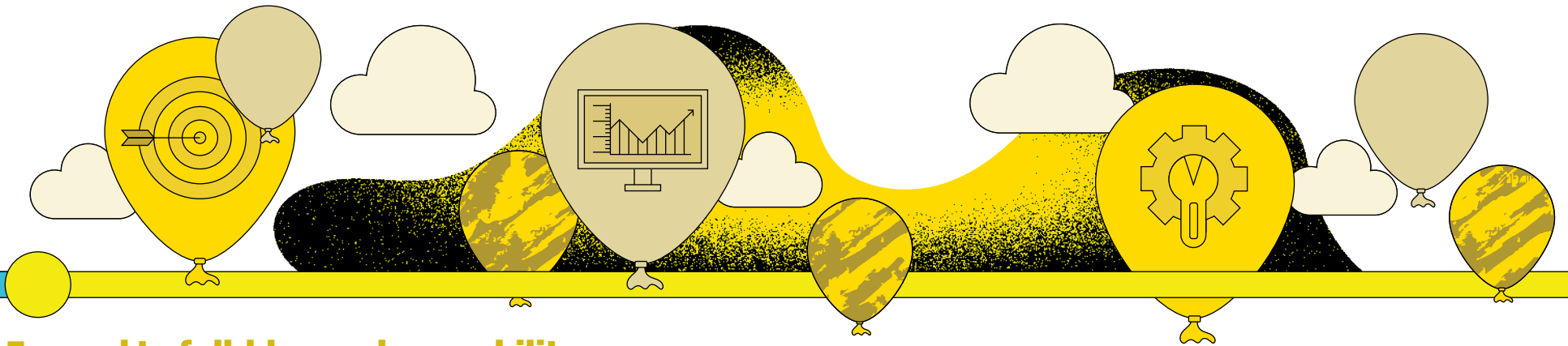
Given the volume of data and the complexity surrounding its use, agencies looking to prevent downtime need a solution that can help them ingest and utilize the full gamut of available metrics.

To make best use of that data and keep services up and running, IT leaders in this phase should be working in close collaboration with mission leaders. "You can start by logging everything, everywhere, but then you need to start talking to the application owners and figuring out what their needs are," Stacey said.

Some may need 90% availability, while others require 99%. "You can work with the app owner to establish what they need from an observability perspective. This is vital to help decipher if enough and the right metrics are being monitored. A 90% availability app needs significantly less observability than a 99% availability app. For cost efficiency, your monitoring effort should match the monitoring needs of the app owner," Stacey said.

With those targets in mind, metrics can help drive mission outcomes. "You can start to think about how that system is performing, not only point in time, but also historically," Ahmadi said. "You can see when disk utilization is getting to a point where you need to start thinking about scaling it, adding additional disk space or capacity."

For all these metrics, IT leaders will be guided by the four "golden signals" of monitoring: latency, traffic, errors and saturation. By understanding how long it takes to service a request, how much demand is being placed on the system, the rate of failed requests and how fully utilized the service is, it's possible to leverage observability to more directly support the application owners' varied mission needs.



Expand to full-blown observability.

Having tackled observability through the application lens, it's time to turn attention to full-stack observability. To minimize downtime, agencies need the ability to generate insights throughout the IT ecosystem and across all cloud environments — public, hybrid, on premises, and multi-cloud.

While not every agency will want to achieve full-blown observability immediately, most will at least be moving in this direction. A unified solution can support this effort, driving cost savings and tool consolidation.

A unified observability platform offers a single means with which to understand operational and mission data, along with context and correlation across all telemetries. It will deliver comprehensive visibility across cloud environments, including service dependency mapping and insights into cloud-native technologies. It will connect all of this together via a schema making sure all systems are talking in the same language so that it all connects.

Once it is all connected, it is important to select the high-level metrics that matter the most for the application. Oftentimes this is "user response time." This metric is the length of time after a user clicks a button for everything to finish. As in, everything in the backend completes and the results of the button click show up. With this, metric issues are less about a system that is down and more about whether the user is happy. When the user is not happy, we know that a downed system actually matters.

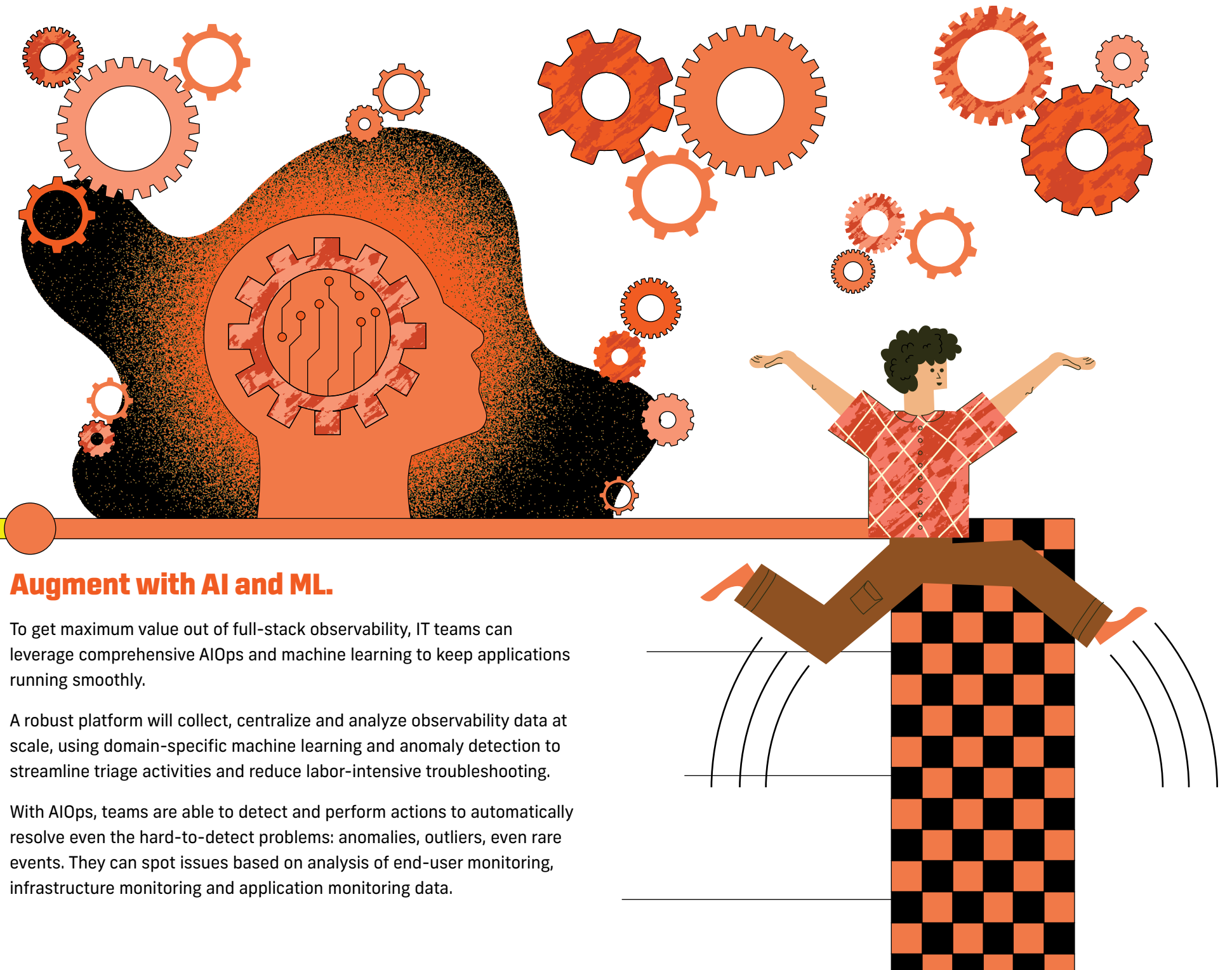
With this approach, agencies can take advantage of automated analytics for anomaly detection as well as for incident detection and response. A comprehensive platform also empowers proactive improvement, driving optimization with code and pipeline visibility along with measurable end-user experience metrics.

"With comprehensive visibility, I'm actually starting to include all of those additional data sources that previously were dropping on the floor," Ahmadi said. APM, for instance, becomes more meaningful, while automated operations drive improved mean time to response, or MTTR.

"With full-scale observability, you can take that full end-to-end user experience and optimize it in a way that improves the service to the team that's accessing the system," Ahmadi said.

All this supports the big-picture goals of system monitoring. The aim here is not just to know what happened: Logging alone could tell us that, in theory. The end goal of full-scale observability is to empower IT teams to take effective action.

With greater visibility, paired with automation, it becomes possible to solve problems more quickly. This, in turn, gives application owners the ability to continue doing the business of government, meeting mission capably and consistently.

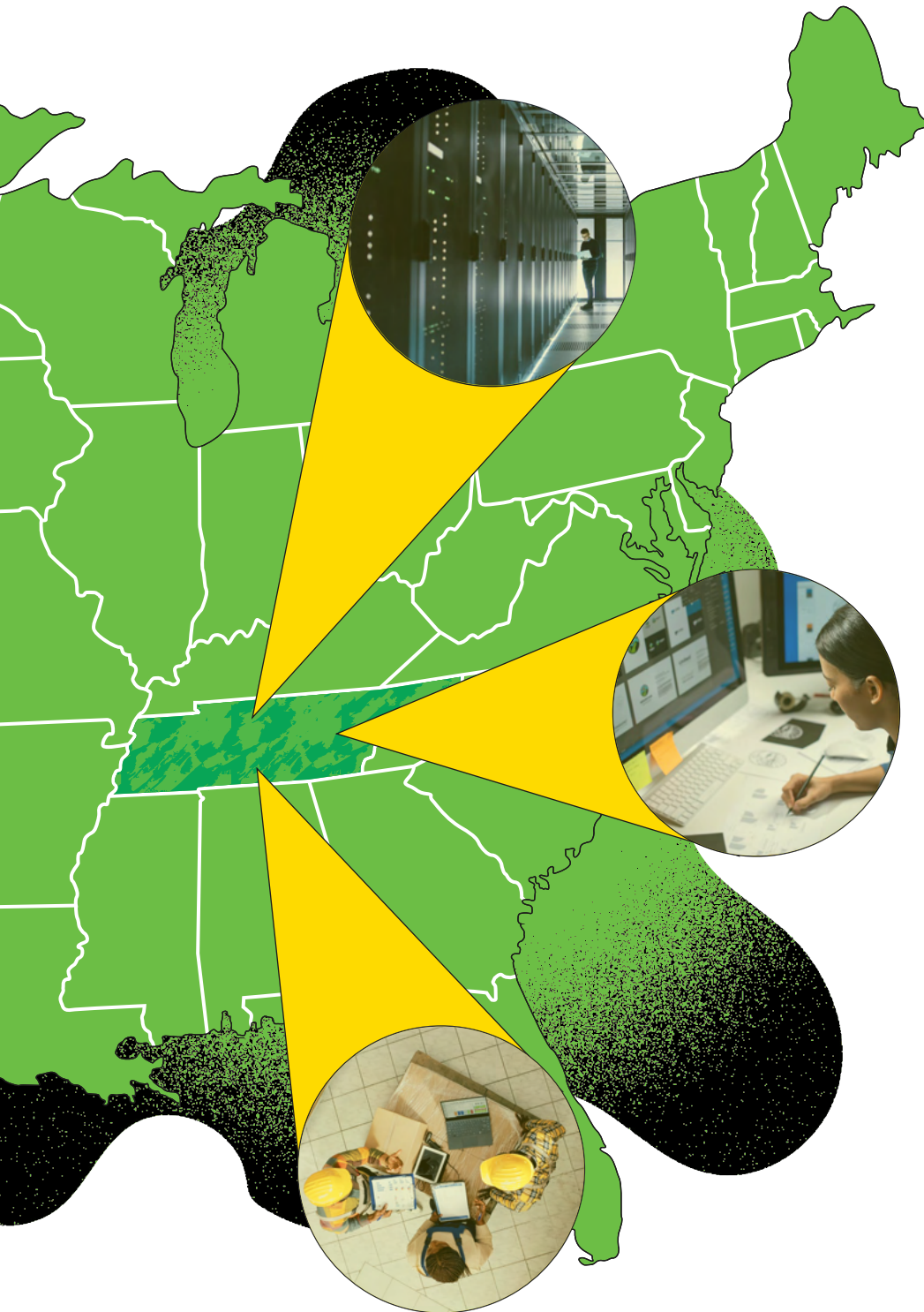


Augment with AI and ML.

To get maximum value out of full-stack observability, IT teams can leverage comprehensive AIOps and machine learning to keep applications running smoothly.

A robust platform will collect, centralize and analyze observability data at scale, using domain-specific machine learning and anomaly detection to streamline triage activities and reduce labor-intensive troubleshooting.

With AIOps, teams are able to detect and perform actions to automatically resolve even the hard-to-detect problems: anomalies, outliers, even rare events. They can spot issues based on analysis of end-user monitoring, infrastructure monitoring and application monitoring data.



Case Study:

How Oakridge National Laboratory Improved Supercomputing Resilience

In the hills of Tennessee, there's a top-secret government facility where scientists once raced to unlock the secrets of atomic energy. Today, Oak Ridge National Laboratory (ORNL) has a much broader mandate, studying everything from biological and environmental systems to clean energy.

The facility's supercomputing program enables researchers to push the limits of computational power in service of scientific advancement. ORNL's supercomputers can perform more than 200 quadrillion calculations per second.

Speed and performance are critical for ORNL teams, as is observability. "Reducing downtime is incredibly important for them," Ahmadi said. "They want to drive that down as low as possible, to speed up response or remediation when the issue occurs."

Working with computer speeds that are measured in petaFLOPS on a daily basis, the Oak Ridge team knows a thing or two about scale. But it has to be able to do something with all the incredible amounts of data being generated, beyond just saving it.

That's where Elastic comes in.

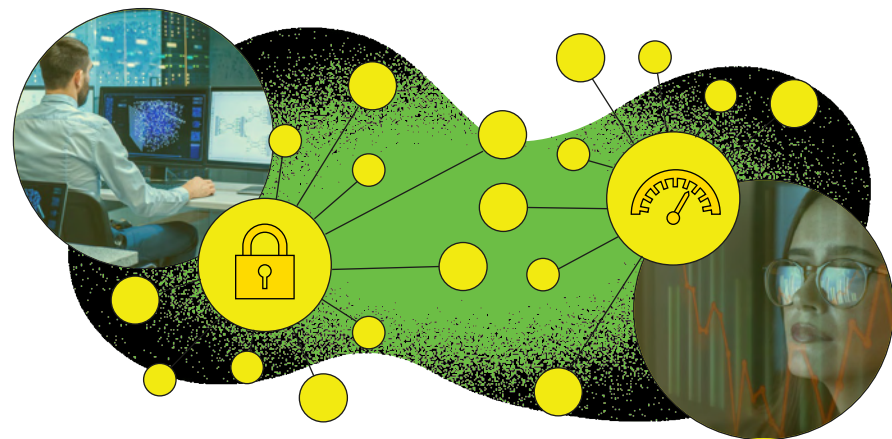
Building Toward Resilience

The Oak Ridge team uses Elastic as both a data store and an analytics engine. The Elastic implementation is key to the stability and performance of the supercomputer center. At the same time, Elastic's tools have helped ORNL achieve a higher level of observability in support of resilient systems operations.

ORNL's cybersecurity team also uses Elastic for security information and event management (SIEM). But it wasn't always that way: ORNL transitioned from another vendor to Elastic to improve its ability to manage roughly 20,000 endpoints through log monitoring and anomaly detection at scale.

For six years, ORNL's cybersecurity team used the other vendor as its SIEM. But ORNL's needs evolved, requiring it to ingest more and more data and run queries on massive indices of over 30 billion documents. Speed was another barrier. A key goal at ORNL is to facilitate research — yet some searches on its old clusters were taking as long as 15 minutes, diverting valuable time away from data analysis.

With the switch to Elastic, ORNL researchers were no longer limited in the amount of data they could ingest, and searches that once took minutes were down to just seconds. And with Elastic, ORNL was able to deploy a SIEM that increased speed and security.



Getting Ahead of Problems

Here's how ORNL has implemented Elastic to improve visibility and get ahead of problems:

- The lab's production architecture runs 25 Elasticsearch nodes, all within Docker, across 25 virtual machines. Its system ingests over 2 billion documents each day (roughly 1.5 terabytes (TB) of data) and maintains 180 days' worth of data (over 300 billion documents) across 10 hot nodes and seven warm nodes. It runs three machine learning nodes, three master nodes and two coordination nodes, with a total disk usage of about 120 TB of data.
- ORNL has a third cluster devoted to research and testing. This cluster, which ingests about 1.5 TB of data per day, features six Elasticsearch nodes on physical servers. Its final cluster, a single node, is used to monitor the other three clusters.
- Using Elastic's data visualization capabilities, the ORNL team is able to see at a glance what's going on with its users in general and focus on specific users as needed.

With this information, if a machine is known to be compromised, team members can quickly identify any other machines that have interacted with the infected device in order to contain and resolve the issue. The team is able to create dynamic, infographic-style dashboards to provide management with high-level views of activity at the laboratory.

With improved observability, ORNL researchers are empowered to develop new technologies for energy, medicine and materials, knowing that if issues should arise in the computing environment, they will be swiftly identified and remediated. "And by implementing a greater degree of observability with more automated means, we have freed up a lot of people's time and effort," Ahmadi said.



Find the mission answers that matter.

Elastic helps federal agencies find the answers that matter by transforming massive amounts of information into mission-critical insights through an AI-powered data platform.

Deployable on cloud or on-prem, Elastic accelerates time to insight, monitors infrastructure for anomalies, and protects highly targeted systems and data.

[Learn more](#)



Leveraging AI to Speed Incident Detection and Response

An interview with Nathan Stacey, Senior Manager, Solution Architects, at Elastic

With greater observability, federal agencies can gain actionable insight into the performance of their IT systems and applications. By adding AIOps, they can supercharge their ability to apply intelligent incident detection and response in order to speed remediation and keep the mission on track, said Nathan Stacey, Senior Manager, Solution Architects, at Elastic.

GovLoop: Can you briefly define AIOps as it relates to observability?

Stacey: Observability is the collection and coordination of something's health via data. AIOps is an always-improving OODA (observe, orient, decide, act) loop built from that observability. Just like turning the lights on in a room, the more data (light) and speed to search that data, the better the decision.

How do AIOps help streamline operations?

App owners have astronomically more levers than ever to pull for better performance, cost, improvements and reconfigurations. AIOps maximizes the value of these levers by having them done automatically.

When ML is understandable to the average user, then these levers can be custom-automated for better performance, lower cost and operational improvements.

And how does AIOps drive detection of hard-to-find problems and speed resolution?

Observability automatically combines different log fields together so that when there is a problem, we know what logs matter. AIOps looks at all of those logs, sometimes millions of logs, and shows which one has anomalies during that exact problem. It enables you to focus on what was different during the problem.

Where observability helps us see the needle in the haystack, AIOps removes the hay and hands us only the needles.

An all-in-one tactical package allows full AIOps in any location. Having the same AIOps workflows at the desk, in the field, or at central operations, ensures that teams can collaborate and troubleshoot in real time.

Each person needs unique access and data control to work together to resolve issues quickly. Elastic provides this role-based access, letting the data owner share data with anyone they want, no matter how close that person is to headquarters or the field.

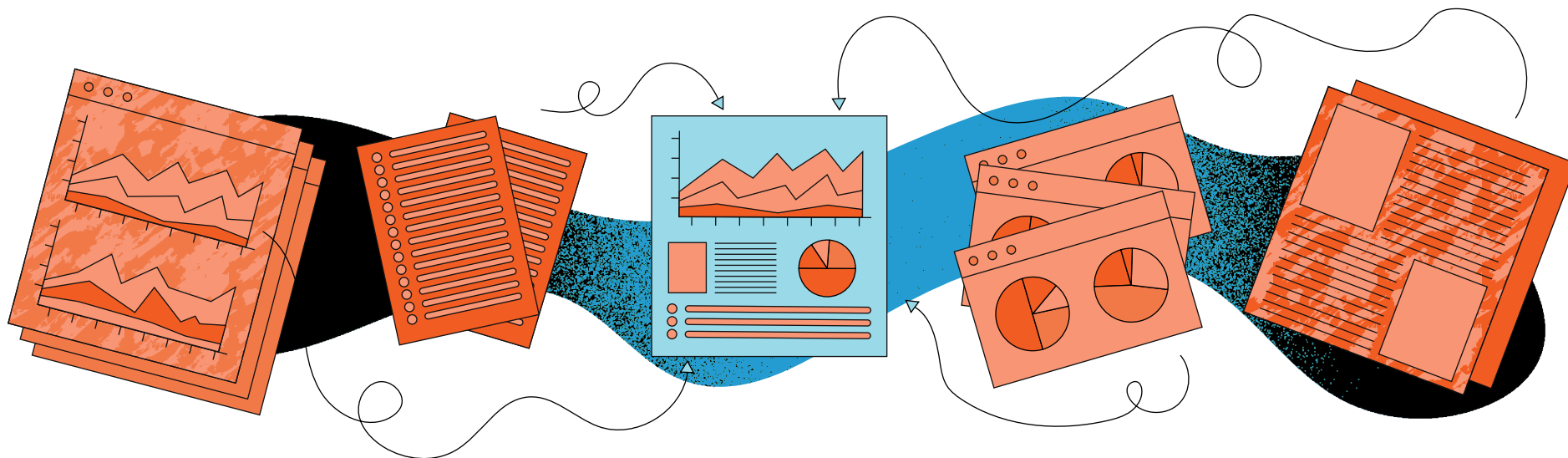
AIOps also enables you to integrate alerts. Why's that important?

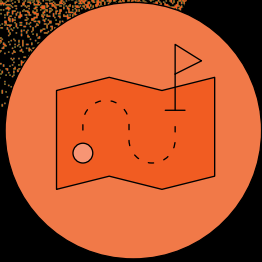
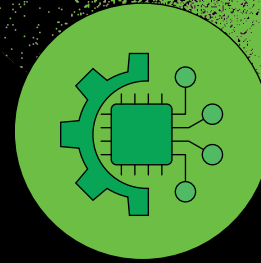
Imagine letting the field share the exact findings to headquarters that they want to share, when they have the network to share it, no matter the location. This facilitates fast communication and decision-making, and saves time versus manually looking for data anomalies.

Summing up, let's connect the dots back to observability. How do these capabilities support greater observability, and what kind of outcomes will agencies see as a result?

Elastic is built for the common user trying to improve their mission. From AIOps to tactical usages in far and remote locations, its setup and usage are easy to learn and configure for your needs.

Elastic's speed and scale also allow for limitless usage of observability data. Elastic's flexibility allows limitless mission needs to matter inside the IT world, while enabling the common mission user to bring as much IT performance as possible to the mission.





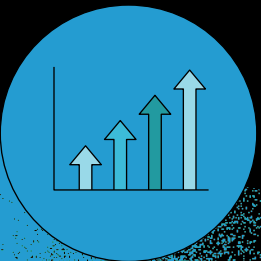
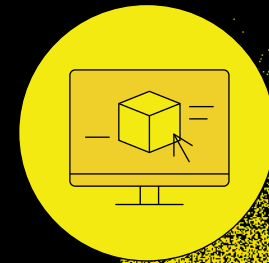
Conclusion

Spurred on by White House guidance, many federal agencies have made significant headway in their efforts to implement logging as a way to support resilience in IT systems. In order to keep missions on track, they now need to build on that strong start and go even deeper into observability.

To that end, government can leverage solutions that offer the ability to converge metrics, logs and traces, thus gaining unified visibility and actionable insights. They can look to tools that support a unified data strategy as a means to avoid siloed tool sprawl.

To ensure the systems' resilience needed to get the work of government done, agencies can embrace solutions that incorporate AIOps, taking advantage of intelligent automations and domain-specific ML rules to boost productivity. And they can adopt cloud monitoring solutions, with an eye toward gaining real-time insights into complex hybrid and multi-cloud environments.

By implementing a modernized approach to observability, IT teams can troubleshoot and remediate issues faster, ensuring that agencies are able to meet mission continuously and effectively.



Thank you to Elastic for their support of this valuable resource for public sector professionals.



About Elastic

Elastic enables public sector organizations to transform massive amounts of information into actionable, mission-critical insights through a unified, scalable, and flexible data platform built on the power of AI. By democratizing data access and findability, the Elasticsearch platform provides a foundation for real-time situational awareness, cybersecurity, application and infrastructure monitoring, and more.

To learn more about Elastic for public sector, please visit: elastic.co/industries/public-sector/

About GovLoop

GovLoop's mission is to inspire public sector professionals by serving as the knowledge network for government. GovLoop connects more than 300,000 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 the public sector.

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

www.govloop.com | [@GovLoop](https://twitter.com/GovLoop)