Six steps to an effective network observability strategy

IBM

Overview

Whether you're looking to improve your application and service delivery, consolidate existing performance monitoring tools and responsibilities, or justify the impact of a new technology deployment, the following six steps can help you create the fundamental building blocks of an effective network observability strategy:

1. Collect
2. Learn
3. Alert
4. Report
5. Analyze
6. Share

Many existing network performance management tools can draw you a pretty graph. However, what's required is pulling together a comprehensive network observability strategy that looks at what data you're gathering and how that data will be treated during each of these six steps.

Breaking down your strategy into these components will make it easier to understand, articulate and reach consensus on the network observability requirements for your business. This white paper also highlights the questions you should ask network observability vendors to ensure they support your strategy.

1. Collect

Any network observability strategy starts with data collection. If you can't monitor it, you can't manage it. It sounds simple enough, but there are many obstacles that prevent proper data collection and leave you with visibility gaps. To prevent this, look for a network observability system that supports the following data collection criteria:

Comprehensive data collection
Your network performance monitoring system should be data agnostic. It should be able to ingest time series data, regardless of source. This includes support for collecting data using standards such as SNMP, NetFlow, IP SLA, WMI, JMX, NBAR and more. But your system should also support where the market is headed. Next-generation networks, both in the wide area network (WAN) and the data center, have begun to move away from supporting standard protocols such as SNMP, in favor of vendor-specific application programming interfaces (APIs).

High-frequency polling
Traditional five-minute polling cycles are insufficient in many instances. A spike in traffic that lasts only a couple seconds represents less than 1% of a five-minute polling cycle. The anomaly will be completely flattened and undetectable when averaged out over that time span. Yet this brief spike can disrupt business transactions, video communications and other latency sensitive applications. Many leading customers today gather data at least every minute. Be sure your network observability strategy enables you to poll for performance data at the frequency your business requires. Don't get trapped into a strategy that limits what you collect and how often based on the polling limits of an inferior network observability system.

Raw data retention
Granular data collection is only useful when you can maintain that data for a sufficient time frame. Some monitoring solutions may average and consolidate historical data over time because of storage limitations, which results in less understanding of historical events. It also weakens your ability to forecast future capacity needs with accuracy. Many organizations seek a network observability system that maintains a year of as-polled data, but this shouldn't require an investment in extra storage capacity.

Massive scale
We live in the age of big data. Applications, systems and network devices produce massive volumes of performance data. Your network observability system must scale with your data collection needs. Inability to scale your solution forces you to make tough decisions about what you will and won't monitor. This creates a visibility gap. You never know what might go wrong in your environment. Therefore, the best strategy is to take a broad approach to data collection supported by a highly scalable network observability system.

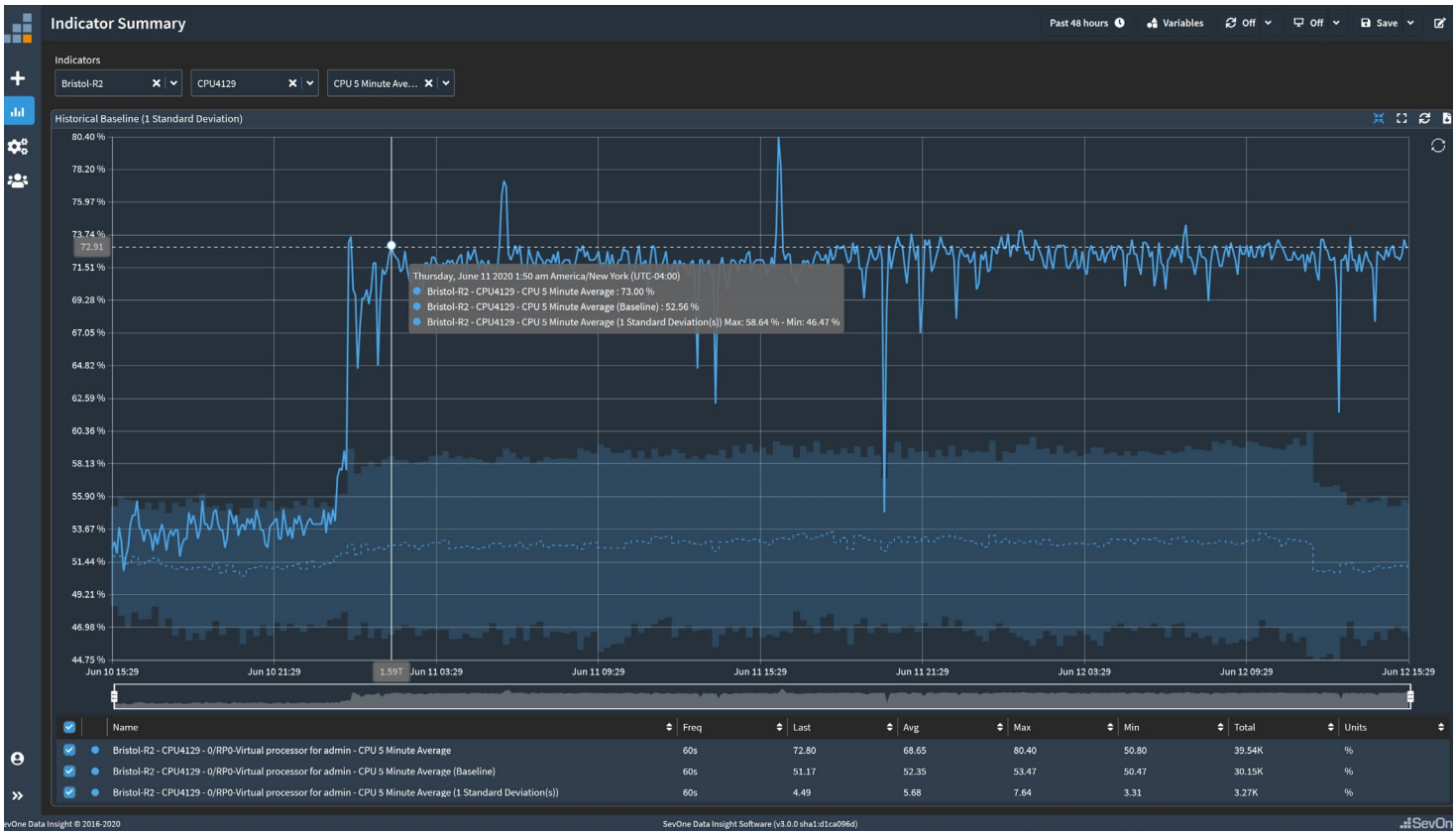


Figure 1. Performance baselines help you understand how your real-time performance compares to historical norms. They also serve as the foundation for a more effective alerting methodology.

2. Learn

Once you've collected the broadest set of performance data at the required granularity, it's time to use advanced machine learning to establish a baseline for "normal" performance. Your network and infrastructure performance monitoring platform should do this automatically for every metric you collect.

Machine learning–driven baselines provide a historical reference point for every 15-minute time frame, every day of the week. With baselines, you can compare real-time network performance to historical norms. You can also view capacity trends and deviations that cause performance-impacting events. Virtualized data centers with rapid elasticity make baseline technology that much more important. Users may provision and retire resources commensurate with demand and without human intervention. This rapid fluctuation in your environment presents monitoring challenges. It's imperative to understand what normal looks like at any given moment. Baselines then become your basis for a more effective alerting method.

3. Alert

You may employ two types of alerts: those based on static thresholds and those based on deviation from machine learning–driven baselines.

Static thresholds are useful in cases such as wanting to know when a central processing unit (CPU) exceeds 95% utilization for a period of 15 minutes or more. Some performance monitoring tools may only allow you to set an upper-level threshold. You should look for a solution that allows you to specify a lower-level range as well. For example, you'd want to know when the voltage on an uninterruptible power supply (UPS) falls outside a specified range, or when the temperature of a device exceeds a high or low manufacturer recommendation.

But you may not always understand what an acceptable range is for the performance of a specific device or metric. Often, administrators guess at threshold values. This results in a significant increase in noise from false positive alerts.

In an environment that generates a lot of noise, it's more effective to alert when the performance of any metric deviates from historical norms. For example, if your company always runs a backup procedure at 3:00 AM, you don't want a daily alert about high bandwidth usage. But you would want an alert when an unexpected spike occurs during working hours due to a unique user-initiated action. You should be able to specify how many standard deviations you consider acceptable for any metric. This requires an understanding of baseline historical performance for all metrics monitored.

This method provides a more reliable predictor of service-impacting events. It answers the most critical question: "What is happening in my environment right now that is unique that I need to know about?"

4. Report

Collecting network and infrastructure performance data is hard. Reporting on it should be the easy part. Yet many vendors deliver canned reports that simply reveal most utilizedinterfaces, highest packet loss and other key metrics. Though these reports are helpful, they may not allow for the level of manipulation required. Troubleshooting the more challenging performance scenarios demands extreme report flexibility.

When developing your network observability strategy, verify that you can answer "yes" to all of the following questions:

- Can I experience immediate value from out-of-the-box, Day 1 reporting, with a series of auto-populating and fully editable reports of customer metric and flow data for common network performance reporting needs?
- Can I ask application type questions of my network data?
- Can I easily edit any of the out-of-the-box reports, or create new reports from scratch—in both dark and light mode—and then reuse them for different regions or offices, saving hours of report building?
- Can I reuse and share any report to enable consistent insights across multisite operations teams?
- Can I quickly pivot and easily visualize related metric, flow and alert data by chaining visualizations together—for example, does selecting data in one visualization drive the content of the rest of the dashboard?
- Can I create troubleshooting workflows by linking multiple reports together—and then share those workflows across teams?

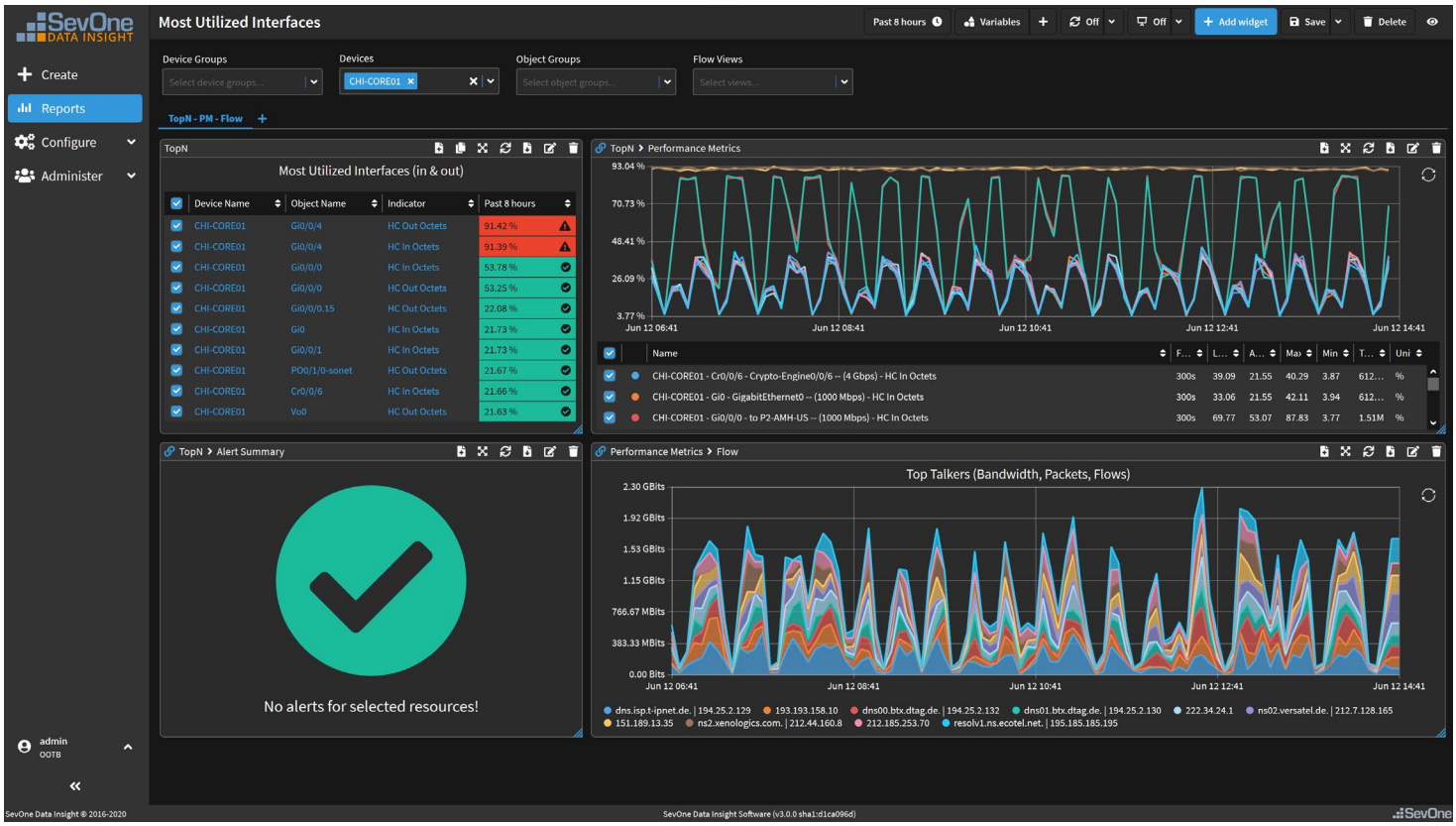


Figure 2. Place a high value on report flexibility because it's the best way to ensure you provide actionable insight to your teams. You should be able to combine any performance metrics in a single dashboard as needed. This may include status maps, alert notifications, TopN reports, NetFlow charts and other data sources.

Finally, you need to understand how increases to the number of devices and the amount of metric and flow data you're collecting impact the speed of your reporting platform. Reports that fail at providing near real-time information are unacceptable. Performance monitoring solutions that rely on a centralized database architecture suffer significant degradation to reporting speed as your monitored domain expands. The centralized reporting database becomes a bottleneck, unable to process report requests with speed. That's why reports sometimes take minutes or hours rather than seconds. It's best to maintain information in a distributed fashion and have the system query the data when needed.

5. Analyze

Data analysis and visualization may take a number of unique forms. At the end of the day, you're looking for actionable insight that allows you to:

Detect and avoid performance events before they impact users or customers

Fine-tune your infrastructure to make the most of current resources

Make more informed decisions about the impact the infrastructure has on your business

Shift to proactive
Proactive analysis and troubleshooting mean you have a better chance of avoiding service-impacting events. For example, in Step 3, we reviewed the value of receiving alerts based on performance deviation as opposed to static thresholds. But what if you could receive a notification when a key WAN link has hit a bursting threshold multiple times over an hour? This type of automated analysis helps you stay a step ahead of your users. Look for ways to make the shift from reactive troubleshooting to proactive analysis that helps you avoid performance events in the first place.

Understand correlations
Many organizations struggle with performance analysis because their platform fails to provide dashboards and reports that can present disparate data sources in a single view. This makes correlation problematic. You end up referencing multiple screens to understand causation. Plan a strategy that allows you to automatically pivot from traditional performance metrics—such as SNMP or IP SLA—to insightful flow data.

Forecast capacity
Capacity teams expect data analysis to reveal which resources are near exhaustion and how much time they have to complete a resource upgrade. Prepare to answer these questions:

– How much bandwidth does your business need?
– How close to maximum utilization are your servers?
– Which network interfaces will be most used 30 days from now?

The ultimate tool for WAN and data center capacity teams is a report that reveals the number of days until specific resources reach a user-defined threshold. For example, capacity planners would like to see all resources that will exceed 80% utilization in 30 days or less. This gives them the ability to plan upgrades based on an understanding of the lapsed time required to complete the upgrade.

Remember, when it comes to capacity planning, volume isn't everything. You must also comprehend the composition of that traffic. Understanding the type of activity taking place can make a big difference in investment plans and monetization strategy. Capacity planning is a numbers game. But the best projection models take into account the value of different types of traffic.

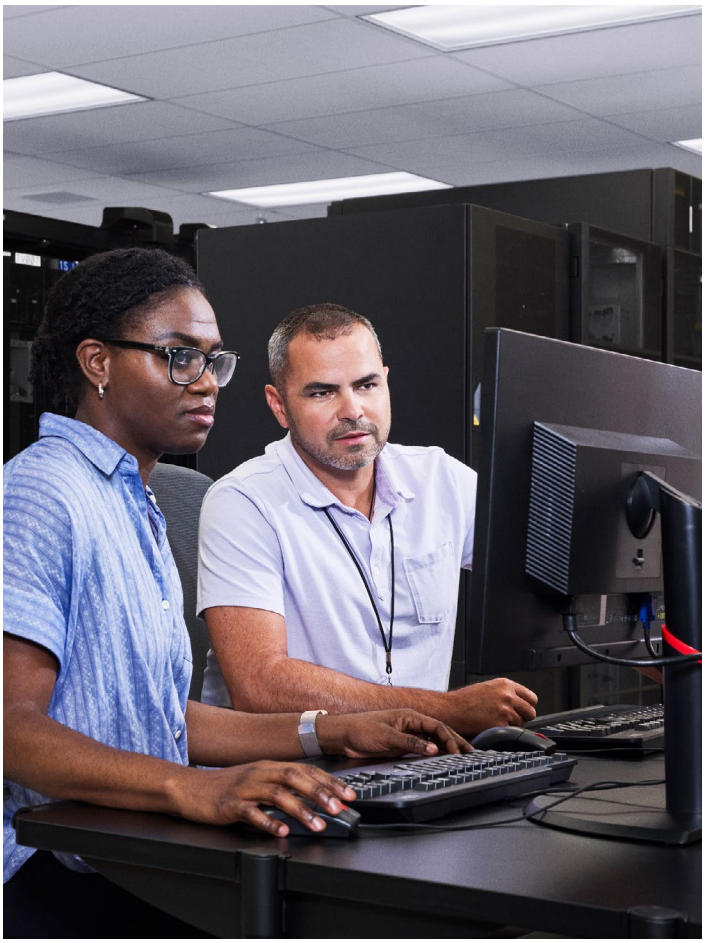
6. Share

You now have a strategy to collect, baseline, set alerts on, report on, and analyze your network performance data. But who can benefit from your insight? Your audience needs may vary:

– Customers through a secure self-service portal
– Coworkers with a .pdf report
– Capacity teams with a .csv export
– Management and executives using real-time dashboards

Providing data for the sake of data is not a productive strategy. The most important lesson to keep in mind when it comes to sharing performance data is to know your audience. A chief information officer doesn't want to get bogged down in granular metrics about a particular link's capacity or some other device status. They're much more interested in a service-level or even market-level view of performance. For example, "What percent of time are we within SLA compliance?" or "How do our multiple carriers impact the performance of our top applications?" Often, a dashboard featuring a simple status map with red, yellow and green indicators for current service health suffices.

Sharing information also means sharing data with other platforms, such as fault or configuration management solutions. It should be just as easy to export data as it is to ingest it. Does your network observability system use an API for communicating with other platforms as part of your maintenance plan, or is that an additional expense?



Summary

The journey from raw performance monitoring data to actionable insight about the health of your network is a critical path. The more you can refine the process to eliminate wasted time and energy, the more successful you'll be. We often think of this cycle as mean time to repair (MTTR), but it can be more accurately stated as mean time to action (MTTA). Ultimately, you're looking to make decisions on the best possible data in the shortest amount of time. This requires a clear understanding of the steps in the process and how you can shave critical moments off each stage.

Breaking down your performance monitoring process into six fundamental components provides clarity around your strategy. It allows you to answer questions such as:

- How much and what types of data should I prepare to collect?
- Do I understand what normal performance looks like in my environment?
- How can I see when something is happening right now that I need to know about?
- Can I understand exactly what happened at any time in the past year?
- Can I comprehend the future capacity needs of my organization?
- Who needs this information and how can I make it actionable for them?

By understanding the core requirements of a monitoring strategy, you also arm yourself with the knowledge to make an informed buying decision when evaluating network observability vendors.

Why IBM?

IBM® SevOne® provides a single source of truth to help assure network performance across multivendor, enterprise, communication and managed services provider (MSP) networks.

Learn more about IBM SevOne and how it can help your organization monitor and manage the performance of both your existing and next-gen network and infrastructure resources more effectively.

IBM Corporation
New Orchard Road
Armonk, NY 10504

Produced in the United States of America
May 2024