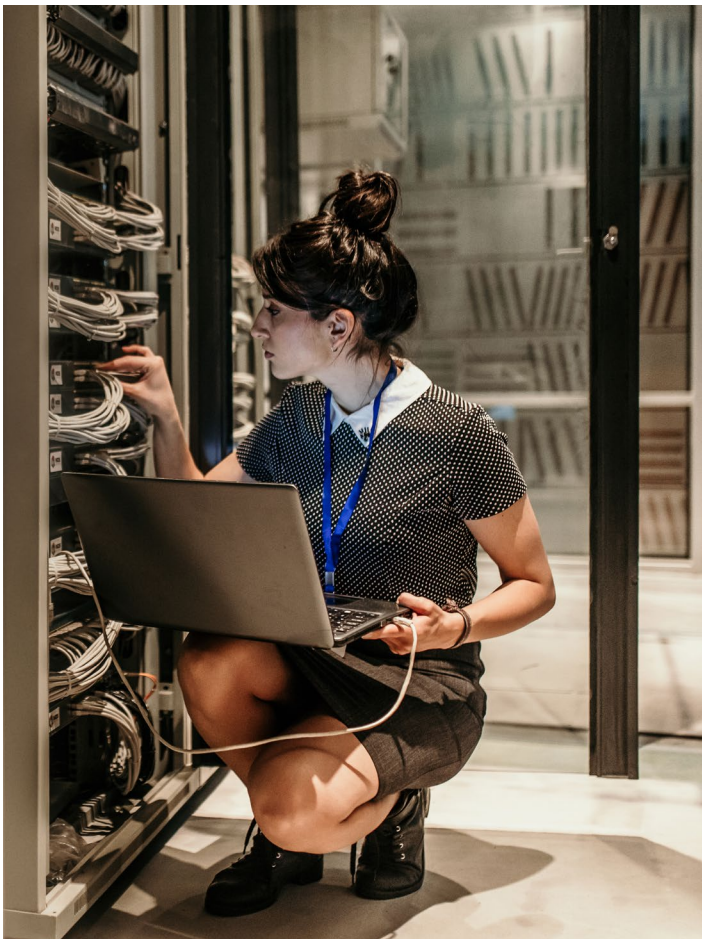
3 ways to improve network capacity forecasts

IBM

Today's organizations are under constant pressure to increase the capacity of their networks and other infrastructures to keep up with user demand. Whether a large enterprise, managed service provider (MSP) or communications service provider (CSP), they're expected to keep costs down so they can effectively compete in the market. Given these opposing requirements, they need to make investments in areas that truly warrant upgrades now as opposed to later. To consistently make those decisions correctly, organizations must be adept at a crucial discipline: capacity planning.

Capacity planning involves determining when the demand for a given resource—bandwidth, central processing unit (CPU), disk space, memory and so on—will outweigh the capacity to deliver it. Done correctly, capacity planning makes it possible to determine which infrastructure upgrades will deliver a return in network efficiency, or in the case of MSPs and CSPs, increased customer business and reduce churn. Make a mistake and they may find themselves overprovisioned or not delivering capacity when it's needed most.



Current state of capacity planning

Most network operations teams conduct capacity planning by collecting data from various systems and pouring it into spreadsheets on a weekly or monthly basis. This labor-intensive method is no longer fast enough to keep up with the rapid changes that occur as more and more networks include everything from wireless to cloud-based IT services.

If teams have a modern network observability system, it's already collecting detailed data on all crucial components of the infrastructure. In this paper, we'll outline three ways to take advantage of that data to dramatically improve capacity planning capabilities.

First, let's look at how most teams currently tackle the job. In a typical scenario, they're importing metrics from all the important systems and devices. These metrics may show peak usage on wide area network (WAN) and local area network (LAN) links, percentage of used capacity on a storage array and other relevant statistics, depending on the system in question.

Next, the teams import the metrics into spreadsheets and perform calculations to try to determine which of them merit attention. That's a difficult task because the metrics typically aren't correlated with one another. For example, they can't immediately tell if a spike in demand on one WAN link had something to do with a failure on another.

It's also a labor-intensive process, given there are often thousands of systems, devices and variables to monitor, including bandwidth, CPU utilization, storage, memory and so on. What's more, it's hard to determine whether teams are using the most appropriate metrics and calculations. And should they decide they want to rerun a given calculation, it may be difficult because the raw data is often gone once it's extracted from the source device or system.

It can be tough to extract data from certain devices that don't support standard management protocols such as Simple Network Management Protocol (SNMP). And it's difficult to ensure that each device and system is accounted for, given the rapid pace of change in most environments.

In the end, teams can't be sure they're making the most appropriate capacity planning decisions. That calls into question whether they're investing in the areas that need it most and will deliver the best return.

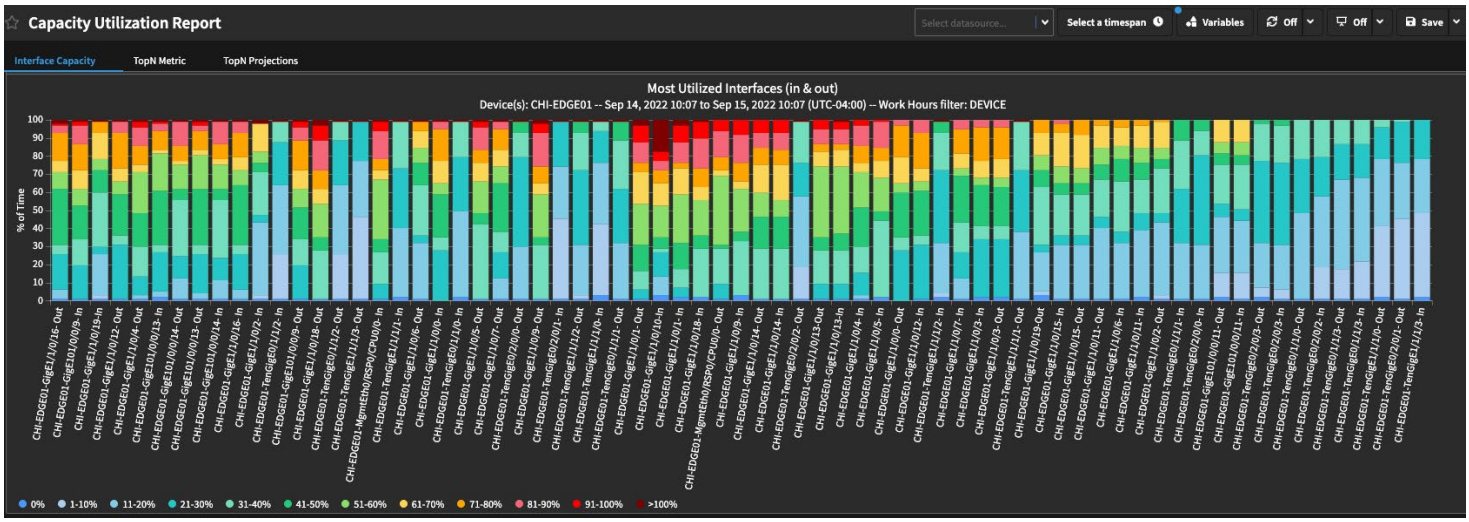


Figure 1. Using the IBM® SevOne® widget development kit, a power user built this widget to visualize the amount of time interface utilization has been in a specific capacity range throughout a 24-hour period.

1. Go beyond TopN reports.

Now let's discuss the ways modern network observability systems can help enterprises and service providers deliver more accurate capacity planning reports. Some platforms aren't being used to their full potential and can dig far deeper into performance metrics. For example, it's common to rely on TopN reports that show which resources are used the most, such as servers, WAN trunks or applications.

Such reports are useful to a point, but they do come with some big caveats. For instance, they represent an average of use over a period of time, which isn't the most useful approach for capacity management. A TopN report may show one WAN link as being most utilized, but it's not necessarily the one that's suffering repeated usage peaks over time.

A TopN report may show a WAN link averaged 90% of its capacity. But a closer look with a network observability system system can highlight when the link burst over capacity and for how long. If the peak lasted only a second, it may not feel like an immediate concern. But if it went on for many minutes or longer, that's a different story. And time of day matters too: If the spike lasted a long time during the normal peak period, then it's an area that's ripe for an upgrade.

With the right network observability system, teams can get to that level of detail using things such as percentage reports. For example, they might establish a threshold for when capacity is considered too high, maybe 80% or 90%. They can then run reports that show how often a given system reached these thresholds. What's more, TopN reports only tell a part of the story. What's not included is information that's crucial for capacity planning, such as:

- When do peaks occur?
- How often do the peaks occur?
- How long did each peak last?

2. Report on groups of resources.

Enterprises and service providers can run reports to identify when they're consistently exceeding a threshold and for how long, so they can quickly home in on problem areas. Consider the difference such detailed data can make to a network services provider. Although a TopN report would show areas that experienced heavy use, what the provider is really looking for are the ones that had heavy use for extended periods during prime network use hours. Those are the areas that would demand immediate attention.

In some instances, it's not a single WAN trunk or server that's causing concern but how a whole group of resources is performing as a whole. Examples may include groups of interfaces to a given application, a server farm or an entire customer site. Maybe a data center hosting provider has 5 or 6 lines coming into its facility from a large branch or customer, all attached to a load balancer. The provider will likely be more concerned with the performance of the lines as a whole than with any individual line. With a modern network observability system, the organization can create a group object that treats the lines as one, apply a key performance indicator (KPI) that defines what proper performance should look like, and continually monitor to ensure the load balancer is distributing traffic effectively.

In large service provider networks and enterprises, it would be cumbersome to create such groups manually. Therefore, it's important that the network observability system be able to put devices and systems into groups using application programming interfaces (APIs) or service management systems. This makes it possible to automate the grouping process.

A wireless service provider, for example, may use this capability to keep track of the capacity at various points in its network. Perhaps it groups the base stations and backhaul lines that aggregate traffic according to region. The provider may want to place highly stringent KPIs on the core or backbone of the network because it affects so many customers, while allowing somewhat higher traffic concentrations on the downstream links.

In short, a network observability system would allow teams to put a premium on those backhaul connections that serve the most customers so that they never run out of capacity. What's more, they could get alerts in real time if capacity was threatened—for instance, if a power outage occurred in an area where everyone was using cellular devices.

Similarly, an internet service provider may want to group the lines that pertain to its various classes of service—one for real-time traffic such as voice and video and other groups for less crucial and best-effort traffic. A network observability system would enable it to look at the sum of each type of traffic and ensure each is within its defined performance parameters. Both examples illustrate the type of capacity planning that's beyond the reach of even a weekly spreadsheet report.

Need to incorporate user activity and wireless data?

If you're a CSP and want to tackle capacity planning, it's good to know not just how much capacity is being used but what it's being used for.

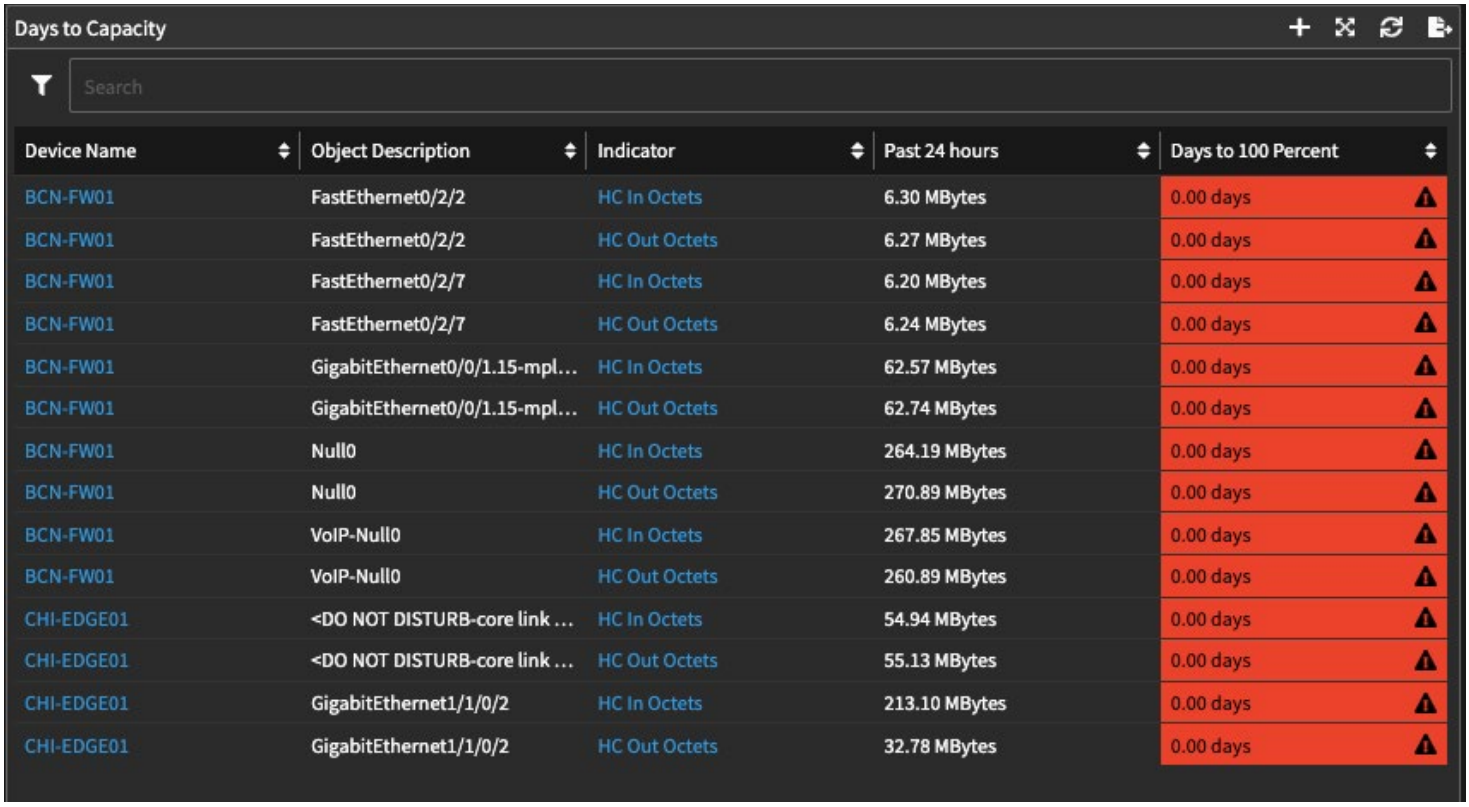
The right network observability system can help on both fronts. To determine what capacity is being used for, NetFlow provides deep insight into IP traffic. Analyzing NetFlow makes it possible to determine where packets are going and coming from, and enable the tracking of application usage.

CSPs can use this data to better plan capacity by type of traffic. For example, a wireless carrier may want to track usage of 5G traffic on its network to determine when it makes sense to add more 5G capacity. Such information can also factor into how companies charge for different services. When a cable company sees rapid growth of a particular feature, it may want to charge more for it. Similarly, if other services aren't catching on, the company could run a promotion to spark interest.

In some instances, companies may also want to pull in data from portions of their infrastructure that don't support typical IT protocols such as SNMP. A modern network observability system should be able to interface with your existing log analysis systems so you can visually correlate log data and performance metrics. It should also be able to normalize the data so it can be compared to other data to find trends that help in capacity planning.

A cable company, for example, may have to incorporate data from several devices to determine the capacity of its network in any given area. These include cable modems and routers in customer homes, cable modem termination systems in the access layer, and headend solutions in hub sites.

With a modern network observability system that can import data from all these devices, the company can accurately determine how many customers it can add, considering their service mix, before capacity runs out. At the same time, it can ensure that as customers are added, service performance doesn't degrade—causing disgruntled customers, cancellations and lost revenue.



Days to Capacity

Search

Device Name	Object Description	Indicator	Past 24 hours	Days to 100 Percent
BCN-FW01	FastEthernet0/2/2	HC In Octets	6.30 MBytes	0.00 days
BCN-FW01	FastEthernet0/2/2	HC Out Octets	6.27 MBytes	0.00 days
BCN-FW01	FastEthernet0/2/7	HC In Octets	6.20 MBytes	0.00 days
BCN-FW01	FastEthernet0/2/7	HC Out Octets	6.24 MBytes	0.00 days
BCN-FW01	GigabitEthernet0/0/1.15-mpl…	HC In Octets	62.57 MBytes	0.00 days
BCN-FW01	GigabitEthernet0/0/1.15-mpl…	HC Out Octets	62.74 MBytes	0.00 days
BCN-FW01	Null0	HC In Octets	264.19 MBytes	0.00 days
BCN-FW01	Null0	HC Out Octets	270.89 MBytes	0.00 days
BCN-FW01	VoIP-Null0	HC In Octets	267.85 MBytes	0.00 days
BCN-FW01	VoIP-Null0	HC Out Octets	260.89 MBytes	0.00 days
CHI-EDGE01	<DO NOT DISTURB-core link …	HC In Octets	54.94 MBytes	0.00 days
CHI-EDGE01	<DO NOT DISTURB-core link …	HC Out Octets	55.13 MBytes	0.00 days
CHI-EDGE01	GigabitEthernet1/1/0/2	HC In Octets	213.10 MBytes	0.00 days
CHI-EDGE01	GigabitEthernet1/1/0/2	HC Out Octets	32.78 MBytes	0.00 days

Figure 2. A “days to capacity” report is the ultimate tool for WAN and data center capacity planners, giving them the ability to plan for upgrades based on an understanding of the lapsed time required to complete the upgrade. In the report above, all of the interfaces shown for the two devices have capacity issues that need to be addressed.

3. Run “days to capacity” reports.

Finally, modern NPM systems can run predictive reports. Such reports show how many days until a given resource runs out of capacity or reaches a predefined limit.

It’s a powerful capability that greatly simplifies capacity planning. If it will take, say, two months to upgrade the capacity of a given resource, organizations may want to know three months prior to when that resource will be at maximum capacity. For example, a wireless provider can’t wait until it hits 90% capacity before it makes the call to add more routers to the core or construct extra cell sites.

Such predictive reporting brings a just-in-time infrastructure planning capability. As a result, network operations teams aren’t scrambling to increase capacity after they’ve run out. At the same time, they’re not so far ahead of demand that they’re wasting capacity and capital.

Conclusion

The days of trying to do capacity planning by manually importing data into spreadsheets are coming to an end. This type of process can't sustain the rapid changes required to keep up with customer demands and market forces. What's more, a spreadsheet-based system makes it too difficult to correlate relevant variables to get the best, most actionable information.

A modern NPM system offers many benefits that can bring a new level of sophistication to capacity planning efforts.

A platform with open APIs makes it possible to gather data from a wide range of external devices, not just those that support traditional IT management protocols such as SNMP. It also enables the automation of gathering and reporting on data, which means it's no longer a gating factor. Results can be produced in real time and reports can be created anytime—not just on a weekly or monthly schedule.

This type of NPM system will deliver more accurate capacity planning information, determining exactly where the most pressing problem areas are—and which ones can wait until later.

It also enables teams to retain raw data for extended periods of time—often a year or more. That's critical for making accurate capacity forecasts, because it's not a good idea to base projections on inaccurate usage data that's averaged over time, such as hourly usage data rolled up to daily usage views after 30 days. And it's an important consideration if the need arises to rerun certain reports or to run them using different variables.

What's more, the right NPM system provides better institutional knowledge of seasonal events. For example, when planning for big annual events, enterprises and providers can look at data from the same or similar events in the prior year. With this data in hand, a wireless service provider and cable company can plan for major events such as the World Cup, Super Bowl or Olympics, or even the annual influx of college students to certain cities and towns. With sound data capacity planning data in hand, they can be confident customers will get the performance they expect. And an enterprise can plan for the capacity needed for a big service upgrade.

Finally, when evaluating the latest NPM systems, it's important to make sure users won't encounter bottlenecks that give them a less than satisfactory service experience—and that money isn't being spent on unnecessary infrastructure upgrades.

Why IBM?

IBM SevOne Network Performance Management (NPM) provides a single source of truth to help assure network performance across multivendor, enterprise, communication and MSP networks.

Learn more about SevOne NPM 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