

## 20/20 Network Visibility

By: Scott St. John

Let's face it: humans are great, but we have our limitations. There is only so much we can do. I would even go so far as to say there is only so much that we *want* to do. And it's not our fault. Every day we are bombarded with data from our smartphones, homes, watches, and connected cars. I receive nearly 500 email messages during an average day. It took me nearly two hours just to configure my new car. Do we really want to deal with more data? The short answer is: no.



But the world today is inundated with data and there is no avoiding it. Fortunes have been built upon it. Economies depend on it. Your company's future may rely on mastering it. This isn't anything necessarily new. [We have been talking about Big Data for years](#) now and it shows no signs of slowing. In fact, [Statista predicts](#) the Big Data market volume will more than double, from \$40.8 to \$84 billion, in as few as six years. But if we lack the ability, capacity or desire, then what is our role in a world of data?

Historically, humans were the data interface. We were the integral component to read, interpret and take action upon data. But this was back when data was relatively small and simplistic. All you had to do was look at the numbers and make a call. When data instances became more complex to manage, we built simple systems to automate a particular set of actions based on a specific and static use case. That worked for a while and, in some instances, it may still work. But every new system, application, device, and service today produces more and more important data. In fact, the role of managing multiple systems has become nearly unmanageable, to the point where many service providers and enterprises have created thousands of lines of custom code to specifically ignore particular data sets in an attempt to make dealing with data more manageable. But this too is a problem.

## Death by data

*Pipeline* recently had the opportunity to interview Anand Thummalapalli, Head of Product Management at [gen-E](#), an expert and leading solution provider that helps service providers and enterprises tame their data, to discuss this growing issue.

"We have seen organizations intentionally disregard 90 to 95 percent of their network data," Thummalapalli told *Pipeline*. "They, and the systems they have in place in many cases, view the reduced data set as the entire picture and that can create serious issues."

By disregarding data, organizations are unable to correlate its relationship to present events. For service providers, this means they are unable to identify or predict network faults because of a lack of understanding of the relationship between disregarded data and the present data set – which can, and often does, lead to failures. Reducing the number of faults for the sake of fault reduction no longer works. All data needs to be rationalized in context to understand the current and future state of networks and systems. The root causes of issues need to be proactively addressed, which in turn reduces the number of both alarms *and* failures.

Managing so many hand-crafted rules has also become incredibly difficult to maintain. Changes to hardware and software make the old rules obsolete, and it takes teams of highly experienced and talented professionals countless hours to manage even a reduced data set. Their roles are consumed by determining what actions may be required to take based upon a limited view of what

is occurring, and they have limited – if any – time to consider which rules need to be changed, updated or replaced until failures occur.

“In one instance, we were working with a large, nationwide North American mobile operator whose network was getting very noisy after the installation of new LTE devices, and it was having a negative impact on their customers’ services,” commented Thummalapalli. “Because much of the data, called informs, had been previously disregarded by manual rules, the operator was unable to see that the noise was being created by the software running on the new LTE devices. By automatically analyzing the entire data set, we were able to quickly identify the issue and the operator was then able to work with the device manufacturer to address the root cause, which both reduced the number of alarms and improved the quality of service in the affected areas.”

Having only the ability and capacity to view and manage a limited set of data also means organizations are unable to incorporate additional data sources to make better, more informed and proactive decisions. For example, highly accurate weather data can be used to predict impacted areas so that rules can automatically be applied to those regions and enable network operators to focus on non-weather-related, critical alarms during storms. Travel and event data could be used to identify influx patterns to proactively increase capacity and service availability. Temperature data can be used in data centers to identify and circumvent routers that may be overheating and are prone to shutting down. In another instance, SDN networks report more health data well before a hard fault, and that data can serve as predictive to head off a major disruption. But many organizations are simply drowning in data and are struggling just to keep their heads above water, even with a reduced data set. They can’t even begin to think about adding more data to the mix. Yet they may no longer have a choice.

The momentum behind software-defined networking deepens the waters and causes additional waves of data disruption. Even if you could apply hand-crafted rules in such a dynamic landscape, it would take droves of hard-to-find, highly trained individuals with years of domain-specific experience to effectively manage it. Not only is this inefficient, it’s cost prohibitive, prone to human error, and still restricted by the limitations of human capacity.

“Hand-crafting rules no longer works in a software-defined networking environment,” commented Thummalapalli. “There is a much higher requirement for automation and pattern matching to understand what the data is telling you. It needs to happen in real time, and this goes beyond human speed and capacity. More and more of our customers, both large and small, are coming to us for those predictive, pattern-matching models.”

Perhaps at the pinnacle of software-defined networks is self-organizing networks (SON), and SON is a key enabler of next-generation networks. SON has already proven itself and is being used widely to automate many aspects of network management, performance and coverage. However, as 5G deployments continue to pick up momentum, SON becomes a mission-critical component. 5G networks rely on small cells due to the poor building penetration properties of 5G millimeter wavelength frequencies. Small cell deployments, particularly in hybrid network scenarios, will require a high degree of automation using SON that can leverage environmental and situational data from service assurance systems. In addition, many future 5G use cases will depend on zero latency and mission-critical connectivity where there is no room for fault, human error, or time for human intervention.

## Teaching machines

Artificial intelligence (AI) and machine learning provide a viable solution to contend with the complexity of software-defined networks and the massive amounts of data being generated by the proliferation of new system, application, network, and third-party information sources. But first, it must be taught.

“AIs and machine-learning systems are like children. They must be taught what to do, and told what not to do,” said Thummalapalli.

In a machine-learning environment, the human role becomes one of a teacher-parent to that of the

machine-child. How children and systems develop is largely dependent on a combination of environment, nurturing, and genetics. In the machine-learning world, this can be defined as the data sets, user-guided learning, and the pre-defined models, patterns, and actions with which the system has been designed.

Thummalapalli explains that gen-E takes a unique approach to this, not by breaking down or attempting to integrate data or system silos, but by [embracing the data](#) within each separately. To do this, the gen-E OpsCenter solution contains two integral components. OpsCenter Search consumes data from a variety of data sources to provide a complete, holistic view. It can also collect, compare and contextualize disregarded data and understand the difference between it and current data sets. Additionally, OpsCenter Search can create the foundation for analytics-driven service assurance. Now with a complete, contextual view of all data in a central platform, The OpsCenter Machine Learning component possesses the models to identify patterns and automate actions that can be taken to resolve known issues.

The third component and final component, the human role, shifts to nurture the platform so that it can become increasingly independent. The expertise grows within the system, not the human, lowering staffing requirements. Over time, an increasing number of issues can be resolved without error through automated pattern matching, which increases operational efficiency and improves quality of service. As a result, humans can then spend their time contemplating how to leverage additional data sources to improve network performance and enhance new service offerings.

## Listen to your networks

Networks are noisy, but it's important you listen to them. Big Data is only getting bigger and human speed and capacity has its limitations. The push to software-defined networking is making matters worse, as more network functions are hosted in datacenter environments. Hand-crafting can't and won't keep up. We have reached our limits and must adapt to continue to thrive in the data-driven world.

Excluding data also no longer works. We must embrace all data to gain a more perfect picture. We need systems that can learn *from us* so that they can work *for us*. Our focus needs to shift to raising systems that can grow, consume and rationalize all data from all data sources. These systems can become increasingly independent, automated and sophisticated over time – so that we can spend less time dealing with data and more time being human.