NETSCOUT’s “Visibility Without Borders” vision is centered on the belief that digital transformation and virtualization erase the borders across elements and layers that exist in today’s networks and bring end-to-end visibility into how the networks work and perform. Tearing down borders that isolate network components frees operators from the constraints of location, but it does not abstract network performance from location. On the contrary, networks without borders unlock the power of location.

Just as the internet takes down borders across cultures and nations, but also supports highly localized content and services, removing the borders in our communication networks allows operators to extract the value of location in ways that are not possible in today’s networks, where function is still tied to a fixed location within the network architecture, and traffic is treated as a uniform stream of bits transmitted across the network.

As borders come down, operators not only gain (and need) visibility, they also gain (and need) flexibility. In a virtualized network, they get to choose what goes where. Which functions should be kept in a centralized location or in the cloud? Which ones should instead be moved towards the edge? And where is the suitable edge – the cell site, the basement of an enterprise, the central office, or a metropolitan data center? How distributed should the network be? And how should different traffic flows, services and content types be managed within such distributed networks? Which bits should be transmitted first?

In the age of 5G, networks become dynamic, agile and self-optimizing, and performance increasingly depends on real-time resource allocation and network topology – which in a virtualized network translate into the location of function.

And location does not only impact performance. It impacts the cost of deploying and running the network, the type of services and the quality of service the network can support, and the revenue streams it can command.
Latency is a prime example of this. By deploying computing resources closer to the edge and using network slicing to keep the latency low for specific types of traffic or services, operators have to modify their networks topology, but they can also generate new revenues from new services that depend on latency, such as online gaming or some IoT enterprise applications.

Edge computing and network slicing are the main technologies that give location its new prominence. They operate orthogonally: edge computing horizontally from the center to the periphery of the network; network slicing vertically with parallel channels that cross the network. Their intersection magnifies the power of location in optimizing the use of network resources. Not all traffic is created equal, and edge computing and network slicing are designed to manage the diversity in traffic requirements, within the capabilities of the deployed wireless infrastructure, and extract the highest value from the network.

But the adoption of edge computing and network slicing is only the first step in extracting value from the choice of location. Even more importantly, however, operators have to decide how to implement them to fully benefit from latency – as well as the higher capacity, reliability and security – that 5G promises. There is no unique answer to the what-goes-where questions we asked earlier. Each operator will have to find its own answers and, because this is all new territory, the whole wireless ecosystem has to learn – vendors included – how to use the data available, but still largely underused, to extract more value from their networks.

That begs the question of where the value of the network comes from. Traditional metrics, such as throughout or dropped calls, are no longer sufficient to capture network value. To maximize network value, operators have to optimize network performance for specific outcomes and strategic goals.

What should an operator optimize? What cost-benefit tradeoffs is it willing to make? In the latency example, which applications should have guaranteed low latency? And which ones are ok on best efforts? How is the operator going to balance the requirements of different traffic flows cost-effectively? How much is it willing to pay in additional cost and effort to lower the latency on some applications? How much should it expect to save from running some traffic as best-efforts?

Operators need to answer these two sets of questions – what goes where and what to optimize – as they decide how to deploy edge computing and network slicing in commercial deployments. They need visibility across the network to guide through this process, make the right decisions, and continue to refine their networks capabilities.

Because edge computing and network slicing add two dimensions (horizontal and vertical) that interact with each other, they also increase the complexity of the optimization process and the amount of data to be processed. And to get to the right answers for their specific network, services, demand and strategy, operators are moving to a more powerful but also more intensive approach to understand their networks and use what they learn in a continuous optimization process:

- Collect reliable, detailed, location-aware, real-time data on network performance at the application or service level
- Develop the capabilities to access the data as needed (e.g., performance data across vendors)
- Drill down network data at the layer level, at the network slice level, and at the micro service level and relate it to the quality of experience and performance for different users or devices and services
- Identify the relevant data (e.g., anomaly detection, user experience) and ignore the rest
- Analyze, monitor and troubleshoot the network in real-time, with a high spatial granularity
- Generate responses to address issues and optimize the network topology and the real-time resource allocation
- Automate the process, and repeat to continue to improve network performance

This is a challenging transformation that may end up giving operators data that is too detailed to lead to issue resolution or learning, creating duplication and fragmentation if the optimization process is done separately for different functions, or more generally creating excessive complexity. Visibility may end up obfuscating the workings of the network rather than exposing them.

To avoid falling into this predicament, operators have to establish a robust and reliable optimization process that allows them to go deeper to get a better understanding of the network when needed, but without adding unmanageable overhead. The combination of learning (AI and machine learning) and automation will help operators manage the additional complexity that technologies like edge computing and network slicing bring and make it possible to extract the value of location, and enable new ways to operate and profit from the network.

For sure, we are still taking the first steps in this direction and the move to a distributed, location-aware and function-aware network requires time, effort and commitment. But it is also an opportunity that operators cannot afford sit out if they want to make their 5G networks shine.