

How Edge Computing Is Transforming IT Infrastructure

New technologies such as IoT and cloud architectures are driving computing to the edge. Companies must embrace this trend in order to survive.

The definition of computing infrastructure is changing. While large traditional data centers have been the mainstay of information technology for the past 60 years, we're seeing a perfect storm today where mobility, economics, and technology are all converging to fundamentally redefine the IT challenge as well as the IT solution. In a nutshell, most everything we learned as IT professionals about leveraging economies of scale as it pertains to the delivery of IT in the corporate world is being turned on its side, and instead being viewed from the users' perspective.

This change in perspective is forcing the entire IT function to be recast. The organizations that navigate this transformation most effectively will be rewarded, while those that don't will be left in the dust.

Since IBM introduced the mainframe in the early 1950s, the common assumption is that "real computing" happens in real data centers: Data centers that take the form of concrete and steel, massive power feeds, and rows and rows of racks filled with servers, switches, and storage. Bigger has been the mantra for the past six decades, and still today when average IT professionals think about what actually makes up their world of computing, the cloud, and all the major web-scale properties, they still visualize small numbers of these massive structures delivering the required computing. In reality, the trend is quite different.

Sure, we've seen a distribution of some IT technology to the edge for years. At any retail or financial institution, there's a significant amount of IT technology in each store or branch. That said, when you focus on the function being delivered by those distributed edge deployments, what you'll find is that technology is mostly aggregating communications for processing in a centralized data center. Every one of those branches has been outfitted with user stations and networking equipment that essentially route all transactions to a centralized data center or in some cases the public cloud, which has also been historically comprised of massive data centers. While those organizations have successfully distributed their networking access function, the actual processing of transactions is still done in a centralized core.

All this is changing today. New technologies such as the internet of things, modern cloud architectures with micro zones, and carrier 5G deployments are driving distributed infrastructure designs to include all the needed computing capabilities to do real work at the edge. These new designs don't just distribute the network function, but also include the processing and storage required to perform work and process transactions completely.

Moreover, while it has been fairly common for a remote site to go offline from time to time, new business strategies are demanding that the same level of performance enjoyed in centralized data centers be recreated in these distributed sites. That means definable and defendable risk and resilience models across the entire structure, regardless of where each portion is physically. In some extreme cases, the large centralized data centers themselves are vanishing altogether, in favor of the deployment of dozens or hundreds of these fully functional miniature data processing centers or micro-zones of computing.

The result? Companies that embrace this truly distributed computing approach, along with all the business fundamentals at play are quickly transforming. All of those small cabinets you see installed in the back of a retail store or bank branch, down the hall in a campus environment, or even at the base of a cellular tower are becoming full micro-zones of computing, which have been amalgamated into the bigger corporate processing

function. This is a significant change to the topologies, structures, economics, and strategies the past two generations of IT professionals are used to.

Make no mistake, distributed or "edge" computing is here to stay. The various technologies required to aggregate instances of computing have matured, and the deployments of such technologies are well past the pilot stages. In many cases, real computing is already happening across an organizations' entire footprint, and the number of instances of computing that these organizations must actively manage and maintain can easily number into the hundreds and thousands. Capacity must be planned using modern models that aggregate these distributed functions as the ability for this capacity to be shared across the entire structure becomes an increasingly critical IT strategy.

Every company today needs to make sure that its distributed computing function is treated as its corporate computing function, with the same level of mission-critical importance that the large monolithic structures have been in years past. How well businesses manage the transformational challenge to build truly distributed computing Infrastructure will make or break their business.

Mark Harris is senior vice president of strategy at [Uptime Institute](#). He has been driving data center innovation for more than 25 years in a wide range of market segments including facilities, computing, storage, security, and networks.