Content

Authors:

Bela Virag
Partner, Telecommunication, Information, Media & Electronics, Vienna
virag.bela@adlittle.com

Andrea Faggiano
Partner Telecommunication, Information, Media & Electronics, Dubai
faggiano.andrea@adlittle.com

Jonathan Rowan
Partner Telecommunication, Information, Media & Electronics, London
rowan.jonathan@adlittle.com

Agron Lasku
Principal , Telecommunication, Information, Media & Electronics, Stockholm
lasku.agron@adlittle.com

Executive summary

Consumers and businesses around the world increasingly demand virtually instantaneous digital experiences in whatever they do. Many industries have already embraced the related opportunities in serving their customers instantaneously and seamlessly, and in reengineering their business models and digital production methods accordingly. The more companies digitize – whatever that may mean in each individual context – the more instant digital interactions matter. These expectations are the key driver behind industrial digitization.

Looking back, as digitalization gained steam over the past decade it also drove a revolution in computing. Cloud computing began to spread, ways to leverage massive parallel computing were invented, storage technologies significantly accelerated, new file systems and new ways to store and retrieve data were invented, and so forth.

However, the products and services telecommunication operators provide to enterprises have not evolved at the same pace. Despite a few facts:

- Virtually all traffic (consumer and business) connects users to enterprises
- All networks terminate with computers or servers of one form or another
- Any acceleration in computing automatically lays incremental demands on networks, too
- Finally, we expect many new, high-performance use cases emerge in the next 5 years

Are today's enterprise networks fit for the task?

The good news is that telecommunication service providers continue to increase bandwidth in their access networks for both consumers and enterprises. Thus, we expect consumed bandwidths and transported data volumes to 5-10 fold over the next 5 years.

But when interviewing the infrastructure responsible executives of enterprise customers across industries, we learn that when they think about their future of enterprise networking, they do not think about bandwidth increases. To them this is by far not anymore the most important aspect when buying a corporate network. Automate-ability, availability, security, scalability all rank higher than performance.

To enterprises, automation is no longer a hype, but founded in functional and security requirements. Automation must span all infrastructure aspects: computing, storage, security and networking. Those operators believing that SD-WAN is merely a cost-saving technology are off the mark: it is about enabling enterprise infrastructure to support agile development approaches spanning multiple cloud environments and providing significantly enhanced cyber security capabilities, next to improved manageability for their corporate customers' operators.

Enterprise architects want to manage their networks via their own applications – both network oriented applications, e.g. for internal IT departments as well as non-network oriented applications. Some will even issue design policies making their applications network aware.

We believe operators must rethink their approach to enterprise networking and offer a comprehensive and integrated portfolio of services meeting the entire CIO-infrastructure agenda: computing, storage, security and networking. Or be superseded out by others who do.

1. Enterprise networks are at the heart of digitization

We have observed that while the various aspects of computing and security are evolving at blazing speeds, networks aren't. This is not to say they don't deliver more throughput – on the contrary. But most networks have not evolved in their capabilities. In this report we will describe what these capabilities could be.

Enterprise infrastructure executives have told us in interviews that many operators, with all of their ICT growth ambitions, do not yet have a comprehensive answer to the evolution of enterprise networking. If this is true, operators are risking existing revenue streams and leaving money on the table for enhanced services in the networking space. This misalignment between demand and supply leaves many enterprises' no option but to satisfy their networking needs elsewhere. Enterprises demanding – beyond bandwidth – automation, security, availability and overarching performance increasingly need to revert to "do-it-yourself" (DIY) models, stitching together solutions from non-telecom-operator networking companies.

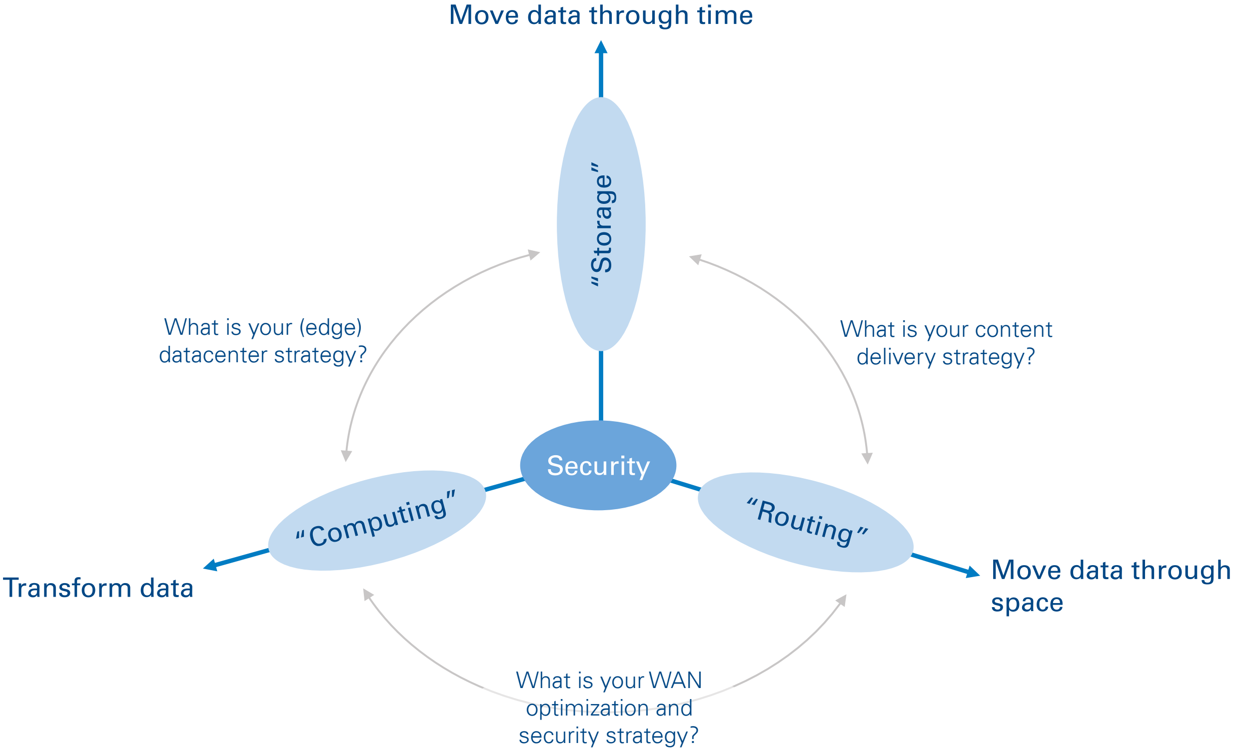
> Are telecom operators missing out on advanced enterprise networking opportunities?

Before diving into the topic deeper, we need to establish two thoughts:

1) Virtually all traffic (consumer and business) connects users to enterprises. The share of peer to peer traffic is generally low

Take sending an email as an example: this is traffic running between users and a mail-server operator such as Microsoft, Google, ISPs, telecommunication providers, et al. Another

Figure 1: Three things that can be done to data



Source: Arthur D. Little, expanding on John Leddy, Comcast "the smarter network" example could be file sharing (e.g., a video) – which requires companies such as Dropbox, Amazon, Facebook and Google to support it. Very little traffic actually runs between users. This is in stark contrast to what telecommunication operators would normally do: transmission of voice (which is peer to peer), sending of an SMS (in contrast to WhatsApp) or torrent downloads (in contrast to other file-sharing platforms). But overall, we estimate that more than 90% of all data traffic is not in a consumer-to-consumer context.

2) Enterprises need to satisfy the demand for an instantaneous end-user/device experience.

They must manage their applications and infrastructures highly efficiently and from an end-to-end perspective. Network connectivity is one part of this, and so are cloud computing and data storage.

Essentially, there are three things enterprises do to data[1]:

- They move data through space: which is called routing.
- They move data through time: which is called storage.
- They transform data: which is called computing

Of course, all of this needs to happen securely.

In this report we will outline how the increase in industrial digitization has already led to new forms of computing and is pushing the limits of enterprise networking.

1 Source: Arthur D. Little expanding on John Leddy, Comcast, "the smarter network"

2. Network bandwidth and data volumes will increase five- to tenfold by 2022

How it all started: Consumers enjoy broadband services (on ever smarter devices), becoming faster and faster. We predict that, whatever the baseline and whatever the access technology, speeds and volumes in access networks will increase five- to tenfold by 2022.

- Copper access speeds are being accelerated via vectoring, G.now and eventually G.Fast to deliver 10 times the current performance.
- Passive optical-fiber networks are progressing into a 4 times 10Gbps symmetrical performance with a 1:256 overbooking factor, delivering multi-Gbps speeds to end customers (and there is a world outside of PON, too).
- Docsis 3.1 (cable networks) delivers 10 Gbps in the downstream and 2 Gbps in the upstream, and will progress towards higher manageability and even more bandwidth, with full duplex and 40 Gbps in the shared segment already agreed in the standard.
- Mobile technologies will evolve to 4.5G (300 Mbps), 4.5G Pro (800 Mbps), 4.9G (1 Gbps) and eventually 5G. The expected migration of these services by appreciating consumers will lead to an increase in average connection speeds by 5 to 10 times in the next five years. And we expect mobile to play a greater role in fixed access than it has so far.

Cisco, in its visual networking index[2], projects a fourfold traffic increase by 2022. But this projection excludes a rise in internet gaming, virtual & augmented reality, artificial intelligence, security-as-a-service, immersive video, video surveillance, robotics and other use cases that are hard to predict. We believe the total volume transported may even grow five- to tenfold in the next five years.

The increase in access speeds and traffic volumes is a good indication that customers expect (and are willing to spend money on) a more instantaneous experience. This drives the emergence of new, high-performance, low-latency use cases.

2 http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html

3. New, high-performance, low-latency use cases begin to emerge

Enterprises of all types – national or global, privately or publicly held, whatever industry – drive digitization in order to improve customer experience, innovate, disrupt or lower cost:

From our casework and dozens of interviews conducted especially on this topic with industry executives responsible for their networks, we have seen new demands emerge: Broader availability of digital services (towards the customer as well as internally) drives an increase in networking needs:

- All kinds of video and imagery applications: video sharing, video calling, video games, video surveillance and visual sensors are a large traffic and bandwidth drivers for corporate networks.
- Migration of data from own maintained storage and servers to the cloud, leading to migration of traffic volume from intra-enterprise to the wide-area-networks (WAN).
- Big data-driven offerings and decision-making to optimize marketing, sales, logistics and production drive low-latency[3] needs.
- Real-time steering of self-controlling applications and remote control of machines, robots, vehicles, trains, planes, etc., drive up the number of connections as well as demand for low latency.
- Situational awareness for robotics, mining, security forces or emergency response teams, physical plant security – especially if remote, etc. – drive the need to keep networks agile.
- Natural language processing for entertainment purposes, smart homes, customer support, contact or incident categorization, customer experience measurements, human-machine interfaces, etc., drive the need for network-embedded computing.
- AI-supported applications for data cleansing, real-time performance improvement, scheduling, etc., drive all of the above.

Figure 2: Why companies digitize

Companies digitize to...

...improve CEX	...innovate	...disrupt	...lower cost
By engaging via simpler channels and more open with customers	By opening own capabilities and assets up to 3rd parties to leverage	By leveraging own capabilities and assets to change other industries	By increasing staff productivity and machine utilization

Source: Arthur D. Little

3 Low latency in this context refers to the time passed between when a request is submitted and an a response is generated (end to end)

4. Enterprises evaluate network vendors along multiple dimensions

When enterprises rethink their networks in light of increased demand, they no longer only consider bandwidth, but review all six dimensions: [availability, security, manageability, automate-ability, elasticity and bandwidth]. *reader to match figure 3.*

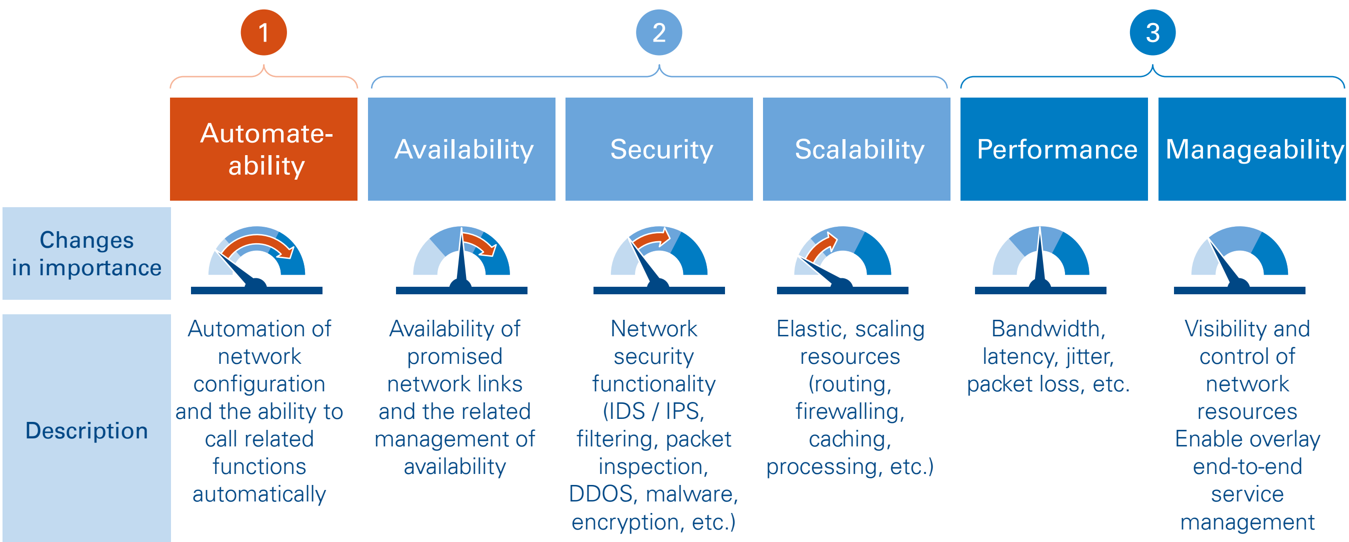
From the interviews we have held with executives responsible for the networking infrastructure of enterprises and from our case work we have learned that, while telecommunication operators often tout bandwidth – and they typically refer to access bandwidth – enterprises have a much more holistic view of how they define network performance. We have captured enterprise-networking performance expectations into the following dimensions:

The figure illustrates how executives think of networking qualities and how they expect the importance of each of the listed items to evolve in the near term. We have indicated this trend with the red arrows on the dials. "Automate-ability" is growing fastest in importance, followed by availability, security and scalability, and trailed by performance and manageability considerations.

1 The need to automate networks results directly from the need to more granularly and consistently manage network quality and networking resources.

The sheer volume of network ports enterprises manage in the most varied domains (e.g., offices, production facilities, clouds and data centers, etc.), combined with the increased need to segment the network[4] for service quality and security, leads to the need to automate network administration. Such automation must include own networking assets (e.g., in the LAN[5], WAN and datacenter(DC)-networks), as well as networking assets provided by various cloud service providers and networks from telecommunication network operators. This either calls for standardization and APIs, or for technology that can interface

Figure 3: Aspects enterprises evaluate when procuring networks



Source: Arthur D. Little

4 Segmenting a network refers to concept of logically splitting different traffic types within a network and treating each type differently. A single link may carry public WiFi, surveillance, employee internet, office collaboration, voice and corporate database traffic. Network administrators strive to treat each of these services differently across their networks.

5 LAN = Local Area Network. The network enterprises we use inside the offices or factories.

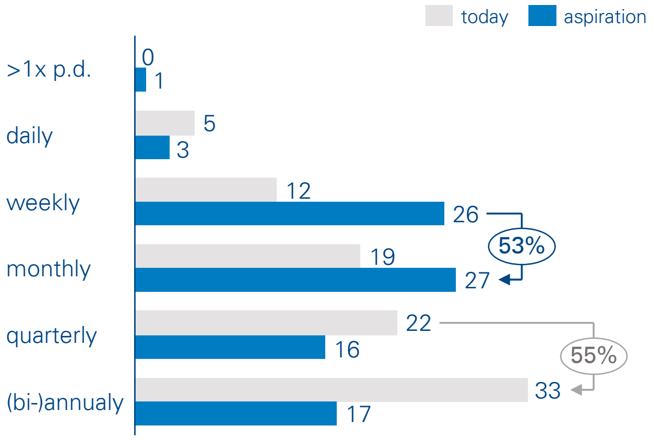
with the different standards and vendors. In most cases, though, operators do not offer such solutions. Many don't even offer automation solutions for their own networks!

Customers we have interviewed manage tens or hundreds of thousands of ports with dozens of services. They do so across multiple clouds and data centers, often in international settings. To deliver services securely and at predefined quality of service (e.g., database access, office IT), each service needs to be performance managed throughout the network and across all nodes. This effort exceeds the ability of many if managed manually.

But there is also a second key driver for network automation: the need for an agile infrastructure to empower agile enterprises. Companies have recognized that in order to remain competitive, they have to accelerate their software deployment cycles. Today, more than 50% of all companies deploy software quarterly at most. However, going forward, more than 50% aspire to deploy software either weekly or monthly.

As code is written and tested, it is moved from a development environment into a test environment, an integration environment and eventually a production environment. Each environment requires adjustments in access policies, firewalls, DNSs[6], routing policies, response simulation methods, test automation, deployment automation, etc., in order for software to function properly.

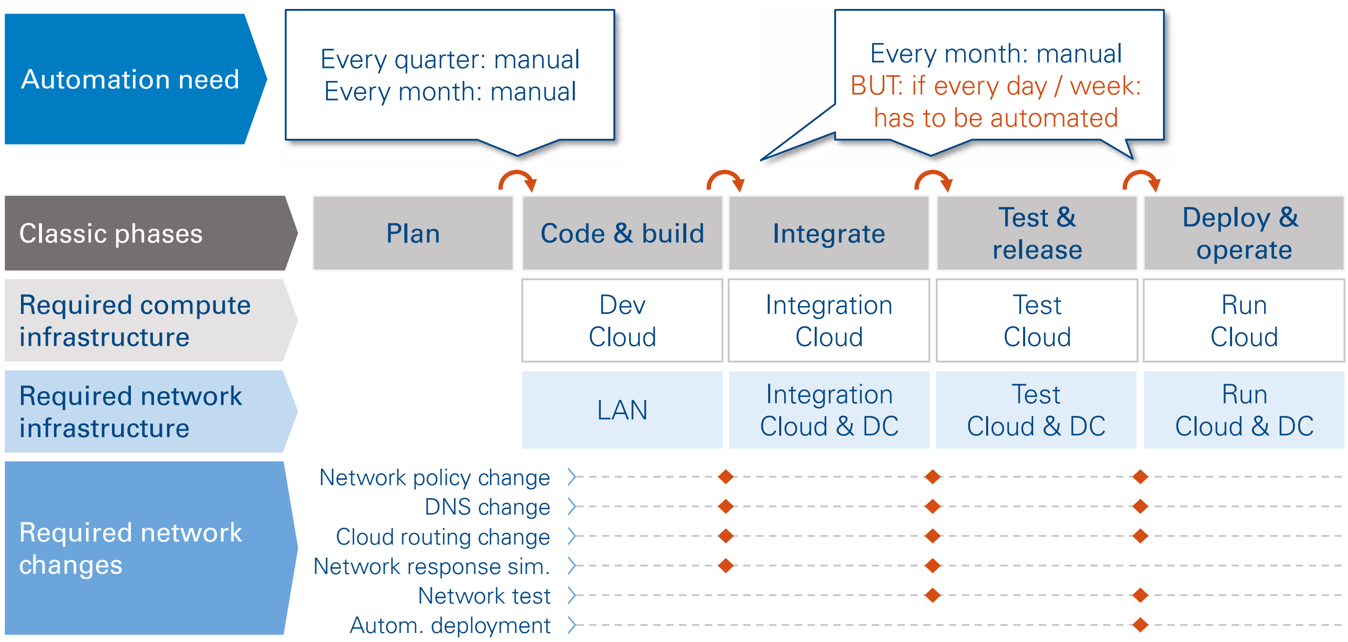
Figure 4: Deployment frequency



Source: Arthur D. Little based on RightScale's 2017 State of the Cloud

Therefore, the process of software development and deployment across distributed environments (e.g. multiple clouds and data centers), as well as the related reconciliation of the upstream environments on a daily, weekly or monthly basis, requires a high degree of automation. Deploying software more frequently means the maintenance of the development, integration and test environments requires more effort – to an extent that only automation is really feasible (unless security is jeopardized).

Figure 5: Aspects enterprises evaluate when procuring networks



Source: Arthur D. Little

6 DNS = Domain Name Servers. These are machines that maintain a directory of addresses and machine-addresses within a network.

2 After automation, considerations on availability, security and scalability gain importance, too. Security, especially, is an unsolved question for many enterprises – leading them to be ready to move beyond a DIY approach.

While availability is paramount to any infrastructure, it seems that corporates have understood the fundamentals of availability management well: agreeing on service levels, insisting on path redundancy and managing supplier performance. This seems to be sufficient in most cases – if operators deliver as agreed, rather than accepting the commercial penalty.

Despite major operators such as Vodafone and Deutsche Telekom having launched dedicated cyber-security teams, many enterprises, so far, rely on self-made approaches to cyber security. However, our discussions have shown that they are considering adjusting this approach to include security-as-a-service. This is mostly driven by the fact that enterprises have recognized an increase in threat complexity, breadth and need for massive infrastructure to fend off attacks, and they simply do not have the skill or budget to enact the related security demand.

So far, the most sourced services relate to infrastructure-heavy parts of cyber defense (e.g., DDoS[7], DNS protection, etc.). However, providing security is not an isolated matter in a highly dynamic network. And it is definitely not isolated from computing. Firewalling, identity management, federation and single-sign-on solutions, fraud and malware protection, data-loss prevention, intrusion detection and prevention, audits and forensics, etc., are all essential and need to work hand in hand with computing and networking[8].

Any security concept needs to be adjusted to the "network domain" it attempts to secure. Because enterprise networks span multiple domains, such as data centers, offices, cloud-service providers, suppliers, POSs, partners, campuses and customer front ends, some traffic may need to be diverted from their direct links to security providers before pursuing the original destination. Security-as-a-service providers are developing diverse solutions for security in such environments.

Eventually, we believe many enterprises will focus on their core competences and source security services from specialist vendors offering security-as-a-service functionality.

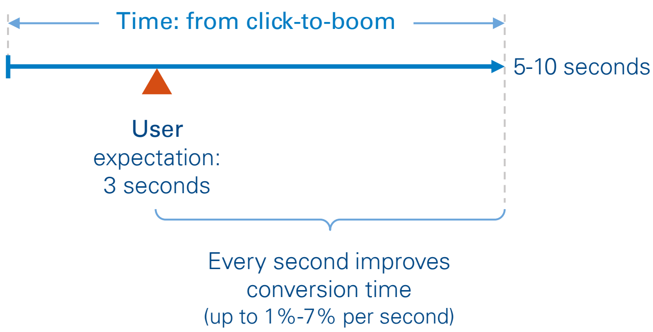
3 Surprisingly, the importance of network performance and manageability are not expected to increase much.

We were quite surprised to find that network performance improvements ranked third after network automation. Network performance includes the topics of bandwidth, latency, jitter, packet loss, and similar. Contrary to our initial expectations, our interview partners often stated that, to them, improving bandwidth was mostly a question of negotiation. While it seems that this is good news, it is also driving prices down.

The one topic in the "performance" category that caught more interest was the importance of managing and decreasing latency.

Enterprises reflect on this in two contexts: According to WebpageFX[9] and other sources, experiments show that users expect websites to load in less than three seconds or faster. At the same time the top e-commerce sites take 5-10 seconds to load. Multiple tests by Amazon, Google, Walmart, Tagman, Shopzilla and others have confirmed that increasing page load times by 1 second leads to a 1%-7% increase in sales.

Figure 6: User expectation: website load time



Source: Arthur D. Little

Latency, as enterprises look at it, is the time it takes for a client request to receive a response. This time interval is impacted by two main factors:

a) the time it takes a request to travel through the network to the server and back, and

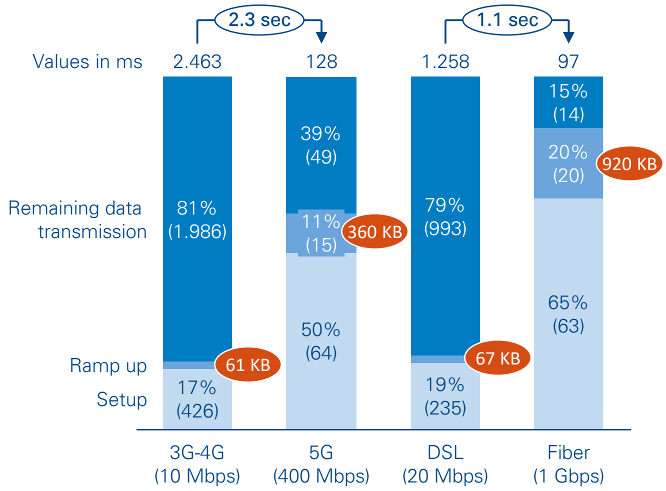
b) the time it takes the servers to compute an appropriate response

7 DDoS = Distributed Denial of Service. The attacker will send more requests than the defender can process. This causes the defender's systems to stall. In doing so, often, attackers use multiple devices to run the attack (sometimes deploying tens or hundreds of thousands of hijacked machines to attack in a concerted manner, causing up to 1terrabit per second – or 1 million gigabit per second – attacks.)

8 An example of how security concepts are intertwined with computing include functions needing to make networks change their security domains in run-time to satisfy regulatory constraints. Another is corporate sandboxes needing to be made available to outside developers, including access to cloud-services, etc., or – as illustrated above – software progressing through the development cycle into deployment.

9 www.webpagefx.com

Ad a) We have analyzed how delay builds up in great detail[10] in the various access technologies. To illustrate our findings, we assume a small, simple website of 2.5 MB in size with limited intelligence. As expected, the faster the access medium, the faster the site delivery, as you can see in the figure below.



Source: Arthur D. Little

Even though 5G is much faster than 3G or 4G, and fiber is much faster than DSL, the most relevant observation is that the increased maximum bandwidth does not linearly impact end-to-end performance. In faster media the relatively slow setup and ramp up phases increase in relevance – they make up 2/3s of the overall time. And by the time the maximum bandwidth has been reached, the network will have carried more than 50% of the payload already. Therefore, carrying small payloads on TCP[11]-oriented networks becomes inefficient quickly. The more requests and connections are being established, the less the maximum throughput matters.

In our example, the website will load in 2.5 seconds – excluding server-processing time. But what if a single element is not readily available on that server? This could be an external advertisement server, a system to suggest the "next-best offer", or a user authentication system which accesses centrally stored user-profile data. It will slow the entire process down – sometimes considerably.

But what if we are not talking about websites but enterprise IT? Let's think of ERP systems, industrial control processes, IoT processes, cash registers in supermarkets, tracking devices in logistics, measurements in predictive maintenance, robotics, manufacturing more generally, and the like: in these situations, the data exchanged is often very small and very frequent, with very low latency demand. Since bandwidth will not combat delay in the case of small amounts of data, different approaches are needed. This could include:

- WAN optimization
- static routes
- embedded/cached DNS lookups
- locally cached content
- no TLS/SSL handshakes
- moving servers closer to where they are needed
- etc.

Here again, operators seldom offer above mentioned services in their portfolios – let alone integrate them with their cloud portfolios.

> Will network operators add computing intelligence to their networks and expose network configuration to their customers?

But what if the above solutions don't help (e.g. in manufacturing processes, real-time monitoring and similar applications)? In these cases, enterprise IT departments typically keep servers on site, e.g., in small, local data centers/data rooms. This increases complexity and risk in regards to cyber security, increases the operational effort to maintain the infrastructure and foregoes scale efficiencies to be had in datacenters or the cloud.

Telecom operators could overcome these issues by deploying clouds in which the computing can be done in the network, close to the customer's sites and the networking configuration can be established by the client in an automated fashion. This would enable operators to process the entire transaction locally, while not having to manage the infrastructure.

Ad b) The second driver to latency is the time it takes computing environments to produce desired results.

Beyond what we discuss on computing and storage technologies in the appendix, there is one specific point impacting delay: the way software is written today.

10 Considering a technically relatively comprehensive perspective on the impact of TCP, DNS, TLS/SSL handshakes, frame alignment, scheduling, buffering, processing, and eventually the impact of slow-starting IP networks.

11 TCP = "Transmission Connection Protocol". It is part of the family of internet-protocols. The protocol is used in most internet applications, such as web-surfing, e-mail, etc. and manages the connection between client and server.

By far not all software developers consider the effect of network performance on application performance. Beyond, they often use development frameworks[12], both self-made and generally available, which sometimes obfuscates how servers and networks interact.

Let's use an example: suppose the task is to load a list of products and the related price for each product. How this task is coded has a great impact on overall performance. If the code is written so that the first query will load the products and the following queries will load the related price for each item, in a setup of 10.000 products this will mean executing 10.001 queries. While there may be other ways to code this, sometimes developers simply don't know how to and don't worry too much about it or the architectural setup doesn't allow them to do so more efficiently.

The real issue is that developers working in the development environment will not experience any meaningful delay, but once pushed into the production environment requests like this may even take minutes to execute.

In our interviews we asked if enterprises were beginning to include the concept of "network-aware software" in their design policies, and the answer was a sweeping: "No"! The contrary is the case: software is written in environments which are highly "local", with all software either on the same computer (e.g., the developer's), or at least in the same LAN. During the later stages of software integration and testing, response automation techniques are employed which simulate behavior of remote resources, only without delay. As a result, software is seldom developed "network aware".

However, many executives responsible for network infrastructure are not satisfied that software developers do not consider the network when coding.

What needs to happen for software to become location aware?

1. Developers need to consider network topologies and data localization prior to designing their software.
2. Architects need to embrace the idea that geographically redundant data and functions can - and maybe should be - maintained in real time. This means that the concept of having a middleware- or enterprise-bus-architecture needs to evolve beyond the idea of keeping data and functions as singular as possible and include the idea to keep data and functions maintained across the entire computational footprint redundantly and updated in real-time.
3. Finally, in more advanced settings, networks need to be addressable by software. Software will need to be able to assess how to optimally reach data and where to deploy computational requests optimally, and possibly how to do so in multiple locations simultaneously.

12 Framework in this context are tools which simplify the software development process by offering simple ways to invoke and manage complex processes. An example could be a simple database request. The function used by the developer would simply formulate the database command. The underlying framework would actually open a connection to the database, log on, execute the command and manage the response. While the software developer would use a single or a few simple lines of code, all the handling of issues that may occur to the request are being managed by an underlying framework. Issues may include failure of the database to respond, error in the execution of the command, no results found, interruptions in the connection, request queuing, etc.

5. Networking and computing will eventually be integrated

We already know that networking is a critical part of computing performance. Based on the following drivers, we predict that networks will become even more integrated into corporate computing environments:

- The increasing need for instantaneity
- Data and computing spreading across multiple environments
- The increasing need for an integrated approach to security
- The need for parallel processing and parallel storage, which require network automation and application-driven handling of network traffic

At the beginning of this paper, we stated that enterprises do three things to their data: transport it, store it and compute – in a secure way. Given that networks are between any of the aforementioned actions, we believe networks need to offer functionality for all three of these areas. Below is a first set of services that might be offered by networks in the future.

Traditional functions which networks can provide include:

- Traffic routing and path computing
- Firewalling
- Policy control
- Traffic and function spawning
- Error messaging and performance diagnostics
- Service-oriented quality management
- Path computing

Storage functions could include:

- Caching and CDN
- Application acceleration
- Database services such as NoSQL and Cassandra
- WAN optimization services

Computing functions could include:

- Application acceleration
- Microservices (Docker, Containers, Hadoop clusters, etc.)
- Transcoding
- Video or data analytics (e.g., security-camera image analytics)
- Real-time CEX support
- Status and control functions (e.g., for IoT applications)

However, this can only be a starting point for the discussion 'What functionalities the future enterprise network should have?'.

6. Enterprises revert to a DIY approach to networking

In the absence of market-available solutions, 15 years ago, the best known internet giants have begun to build their own networking solutions. In 1999 Google already realized there was no network provider out there that could meet its needs, and started building its own networking capability. Today its data centers operate at a mind-boggling 1 petabit per second – 1 million gigabits. Amazon and Facebook's followed suit.

Today, large enterprises, too, are formulating requirements which telecommunication network operators cannot respond to in a commercially-sensible way. Thus, they, too, revert to DIY approaches of sourcing individual parts from software and network equipment providers and assembling them into their own tailor-made solutions.

> Can it be that network operators have lost the edge in networking to software companies?

We are seeing many industrial players, in their attempts to digitize and cater to the new demands and competitive pressures, contemplating the following questions:

1. How can they build integrated network and IT architectures?
2. How can they avoid being locked in to single network service providers?
3. How can they design networks to leverage peering, instead of international leased lines?
4. How can they apply/unify/automate security if computing environments span multiple clouds?
5. How can they manage network-caused application latency?

Enterprise customers have begun asking their network service providers for answers to these questions – or at least for roadmaps. Sadly, more often than not, they found that telecom operators have not advanced their own networking technologies sufficiently. All the while "new entrants" such as cloud-network providers or SD-WAN equipment vendors provide better answers.

It may be that telecommunication operators are overwhelmed with reformatting their networks. But if they don't catch up on the six performance dimensions shown in Figure 3, they invite hardware vendors, software players, systems integrators and cloud-native players to develop solutions for their clients – on the back of the network operator's assets and services. Operators would essentially leave the space wide open – maybe even leave their enterprise customers no choice but to look to other software and hardware vendors to support more advanced enterprise networking. Players such as Cisco (after acquiring Meraki, Viptela, etc.), Riverbed, Masergy, Aryaka and many others are pushing their OTT solutions into the market.

Maybe a curiosity on the side: Both Google and Amazon have begun to offer data-transfer appliances to support customers' migration into their respective cloud environments. Amazon's "Snowball" solution is available in 50–100 TB sizes, while Google's "Transfer Appliance" is available up to 480TB for 1,800 USD plus shipping (500 USD). Their argument for "snail mail" is simple: transferring 10PB of data (typical enterprise volume) over a 100Mbps or 1 Gbps link takes three to 27 years.

Source: Arthur D. Little and company websites

7. Operators need to rethink their networking portfolio

Many operators are seeking growth in the B2B segment by providing ICT services. The B2B segment often contributes 20–35% of overall turnover, and accounts for 50% or more of growth expectations. Corporate data networks today typically amount to 4–7% of overall turnover – and are thus overlooked in the strategic plans. This has two consequences:

1. Networking revenue streams remain under pressure.
2. Growth in other ICT areas does not compensate declines in networking.

Let us be very clear: seeking growth in the ICT segment is a positive aspiration and a natural choice. Some operators become very successful in this space. However, we have also seen many operators try to fish in foreign ponds while ignoring the many ICT opportunities that may come from a more comprehensive networking, computing and security portfolio.

Not prioritizing the risk of revenue decline in the corporate networking segment is a dangerous mistake to make which may amount to the loss of 20% of B2B turnover and nullify growth expectations.

We expect software-oriented networking demands to emerge from customers – which, in turn, will require network operators to implement software-defined networking. Technologies such as segment routing will be more widespread and need to be exposed for use by enterprises. Beyond this, operators will need to install and expose path computing capabilities, as well as full, end-to-end support of IPv6, to facilitate efficient routing.

We expect operators to develop commercial and technical strategies for their (edge-) data centers, balancing resources between CPU, storage, GPU[13] and memory. Many operators are likely overwhelmed with even anticipating what will be needed, and deploying these resources following a "build-it-and-they-will-come" model may be of little financial attractiveness. Some operators may resort to having established computing companies – the likes of Google – deploy micro data-center environments for them. AT&T'S CORD initiative may be a good example at hand.

We even see systems integrators like WiPro, IBM and others develop networking capabilities - possibly even exploring the development of services based on their own 5G infrastructures. And we see internet giants enter this space. In the fields of the IoT, virtual reality, artificial intelligence, machine learning and other (presently hyped) areas, telecommunication operators may fail to grab incremental value if they fail to provide related competence and capabilities in their networks.

Operators need to prepare for, set up and initiate key client engagements to understand infrastructural requirements in a future networked world. Operators will need to redesign their networking portfolios to reflect the shift in importance of the six network-quality criteria:

- Automate their network operations and offer related capabilities (network deployment and configuration automation) to enterprise customers to dock onto – including the introduction of the necessary asset, inventory and capacity management capabilities.
- Develop or partner for an integrated solution that spans network and cloud security domains.
- Provide a broader set of network functions in all functional domains (networking, storing and processing).
- Be agnostic to who provides the underlying access circuits (probably the most painful one), as customers operate their solutions in multi-country environments and like sourcing from multiple access-network providers.
- Getting the above done likely requires a high level of standardization across domains. Operators need to forfeit their proprietary methods and processes and join forces to develop a unified, cross-domain standard.

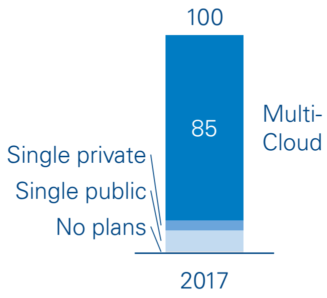
Doing so will force operators to invest into renewing their own networks, too, or acquiring related competencies from outside. In the first case, this may lead to a significant improvement in production cost (due to the increased internal efficiency). In the second case, they may actually develop their own "over-the-top" approaches to provide to the described customer demand of network automation. In both cases, operators need to develop commercial, operational and technical strategies for the future of enterprise networking – and possibly the future of networking as a whole.

13 GPU = graphical processing unit. The microchip that is designed to be particularly performant when handling graphical (2D and 3D) or otherwise highly parallelized computational tasks.

Appendix: How changes in computing and storage technology impact enterprise networking

Elaboration on changes in computing and storage technology and the impact on enterprise networking

a) Computing: To deliver enhanced computing performance, enterprises are increasingly utilizing modern computing methods in multiple cloud-computing environments.



Source: Arthur D. Little based on Right-Scale's 2017 State of the Cloud

RightScale[1] shows that 85 percent of companies employ multiple cloud environments. The most used services today are relational database-as-a-service applications. The most planned services are containers[2]-as-a-service.

Cloud infrastructure-based computing is comparatively more efficient than other methods of providing computing. Studies by Google and others have shown that 10 clusters of 1,000 servers are about 30 percent less efficient than one cluster of 10,000 servers. Thus, the larger the infrastructural pool enterprises use, the more efficient they will be in delivering computing results. Since cloud computing infrastructure is being provided on an IaaS[3], PaaS[4] or even FaaS[5] basis, enterprises can choose how to scale their computing needs in their own data centers, as well as in ever-larger cloud environments, exactly to the size they need – at any 100-millisecond interval!

This means companies can tune their environments to whatever performance requirements they may have, and still be exceptionally efficient without overprovisioning any kind of infrastructure. This ability, however, assumes that the data required for computation, and the application performing the computation, are available in real time – and geographically redundant.

b) Storage: Storage and retrieval technologies are rapidly evolving in terms of both hardware and software – they are not the problem. The problem is that data, the application and the processors are not in the same physical location.

Data-storage hardware is essentially evolving from spinning disks to memory chips. However, memory chips themselves are being connected using ever faster bus technologies, progressing from IDE[6] to flash drives, and finally to NVMe[7]. During this progression, storage latency has been reduced by a factor of 1.000 to about 10µs for NVMe drives over HDDs[8]. Thus, 100 I/Os today fit into 1ms. This makes NVMe only about 200 times slower than RAM[9], and consequently it does not, by far, pose the same delay issue as old-fashioned HDDs, which were 20.000 times slower than RAM. The real slowdown is in getting the data to the processor and back. If applications and data are physically where the computing resources are, storage should no longer be the bottleneck.

1 RightScale 2017 State of the Cloud Report
2 Containers in this context refer to fully encapsulated sets of functions that can be called from entitled applications and that will deliver responses. In contrast to virtual machines, they do not require an operating system. Due to containers, former monolithic software stacks are being broken into their individual functions, reducing and clarifying the dependencies within software code and easing maintenance and scale
3 Infrastructure-as-a service: the providing of a fully managed computing infrastructure
4 Platform-as-a service: This refers to the providing of execution environments (such as Java, various databases, etc.). Administrators are no longer concerned with hardware or operating system, but with the performance of the provided platform
5 Function-as-a service: This refers to individual functions, aka software code, to be provided as a service. In this case, administrators do neither worry about hardware, operating system or platform availability: they pay for the execution of software code
6 IDE = Integrated Drive Electronics, an interface to connect hard drives to mainboards
7 Non Volatile Memory Express. An interface which links SSDs (=Solid State Disks) with a PCI express interface. This allows NVMes to perform about 10x faster than HDDs with 1000x lower latency
8 HDD = Hard Disk Drive

However, we believe there are reasons data may actually be moving further away from where it is being computed:

1. Parallel processing of vast and exponentially growing amounts of data leads to greater distances between the data and where it can be processed.
2. "Containerization" and the use of multiple clouds lead to software calling functions which may be sitting in distant environments.

Modern file systems allow for fast retrieval of massive amounts of data. Enterprises generating and using vast amounts of data use systems such as the Hadoop Distributed File System and no-SQL databases such as Cassandra to store, sort, map and reduce large amounts of data – both structured and unstructured. These systems allow for highly scalable and widely distributed storage and retrieval – indeed, they are accelerated by the fact that they are distributed.

However, our observation is that, as storage arrays grow, they seem to be moving further away from processors. Having formerly been in-server, as volumes grew, storage arrays were moved to in-rack systems, later to in-data center arrays, and finally to the cloud. Since system performance depends on the slowest element of the system, retrieving even small elements required for computation may slow down the end-to-end processing.

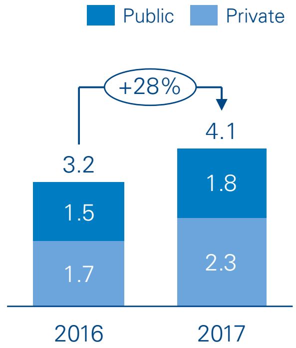
As computing power is being parallelized and distributed, and data storage is being distributed in various environments, the issue of distance between data and CPU[10] increases.

The above-quoted study from RightScale states that companies, on average, use 4.1 cloud environments: 1.8 public and 2.3 private clouds.

Given that software is decreasingly a monolithic piece of code, but increasingly containerized into hundreds or thousands of functions, it becomes increasingly less likely that both data and function are in the same physical place.

While storage arrays, file systems, operating systems and even software have adapted to delays in getting data, networks and security haven't.

We have tried to size the problem of intra-data-center or intra-cloud traffic (assuming that each cloud sits in its own data center), and found that traffic running inside a data center is at least 10–25 times larger than the traffic running into that data center. Put differently, companies require 25 times the networking speed inside a data center as in the access. If we assume that 10 percent of data needs to be transported from one location to another, this would potentially more than double the WAN traffic.



Source: Arthur D. Little based on Right-Scale's 2017 State of the Cloud

Clearly, not every enterprise is already facing this issue, but if access speeds accelerate and computing performance requirements lead to parallel, multi-cloud computing in containers, a number of new questions arise:

- Where to store data and code?
- How often to store the same data and software code in geographically distributed locations?
- How to ensure that data and code are kept synchronous across the entire corporate footprint?
- How to transport both data and functions to avoid having the network become a bottleneck in delivering on customer expectations for instantaneous business interactions?
- How to secure the data and functions as these transit multiple computing environments?

9 RAM = Random Access Memory. It is the main memory in a computer, generally referred to as "memory". There are other, faster types of memory in the 1-3 levels of CPU caching and on various other parts of computers (e.g. graphics cards, controllers, etc.), all of which are faster, but typically bound to specialty applications

10 CPU = central processing unit. It is the microchip inside a computer which hosts the operating system and executing the software running on it. It also does most of the calculations needed

Contacts

If you would like more information or to arrange an informal discussion on the issues raised here and how they affect your business, please contact:

Bela Virag
Karim Taga
virag.bela@adlittle.com

Belgium
Gregory Pankert
pankert.gregory@adlittle.com

China
Russell Pell
pell.russell@adlittle.com

Czech Republic
Dean Brabec
brabec.dean@adlittle.com

France
Julien Duvaud-Schelnast
duvaud-schelnast.julien@adlittle.com

Germany
Michael Opitz
opitz.michael@adlittle.com

India
Srini Srinivasan
srinivasan.srini@adlittle.com

Italy
Giancarlo Agresti
agresti.giancarlo@adlittle.com

Japan
Shinichi Akayama
akayama.shinichi@adlittle.com

Korea
Hoonjin Hwang
hwang.hoonjin@adlittle.com

Latin America
Guillem Casahuga
casahuga.guillem@adlittle.com

Middle East
Sander Koch
koch.sander@adlittle.com

The Netherlands
Martijn Eikelenboom
eikelenboom.martijn@adlittle.com

Norway
Diego MacKee
mackee.diego@adlittle.com

Sweden
Agron Lasku
lasku.agron@adlittle.com

Singapore
Yuma Ito
ito.yuma@adlittle.com

Spain
Jesus Portal
portal.jesus@adlittle.com

Switzerland
Clemens Schwaiger
schwaiger.clemens@adlittle.com

Turkey
Coskun Baban
baban.coskun@adlittle.com

UK
Jonathan Rowan
rowan.jonathan@adlittle.com

USA
Sean McDevitt
mcdevitt.sean@adlittle.com

The future of enterprise networking – How telecom operators need to accelerate to defend their enterprise networking competence

Arthur D. Little

Arthur D. Little has been at the forefront of innovation since 1886. We are an acknowledged thought leader in linking strategy, innovation and transformation in technology-intensive and converging industries. We navigate our clients through changing business ecosystems to uncover new growth opportunities. We enable our clients to build innovation capabilities and transform their organizations.

Our consultants have strong practical industry experience combined with excellent knowledge of key trends and dynamics. Arthur D. Little is present in the most important business centers around the world. We are proud to serve most of the Fortune 1000 companies, in addition to other leading firms and public sector organizations.

For further information, please visit **www.adl.com**.

www.adl.com/FutureOfEnterpriseNetworks