APRIL 2024

Embracing the AI-native Network Transformation for Communications Service Providers

Jim Frey, Principal Analyst

Abstract: As AI influence spreads across every economic sector, the time for communications service providers (CSPs) to adopt an AI-native stance is now. Those embracing AI technologies will be best positioned to reduce risk and improve agility and efficiency, remaining vitally competitive throughout a critical time in the industry. Google Cloud and Intel are bringing essential enabling technology and sustainability to the table to help providers achieve this important transformation in their networks.

Introduction

Transformation: It's a word and mandate that organizations of all shapes and sizes must embrace. Organizations must transform, evolve, and adapt, or they will fall behind and become obsolete. The CSP sector has already undergone many major transformations over the past few decades, and there will likely be more ahead. First there was deregulation, followed by digital voice, IP telephony, and the explosive growth of internet services. This heralded a true shift to digital technologies, with a decoupling of hardware from applications and services that has since dominated the provider landscape.

Traffic continues to grow, and customer experience continues to be top of mind for CSPs to ensure market share and service differentiation. To keep pace, retain customers, grow average revenue per user, and protect operating margins, providers have been forced to reexamine all aspects of their business and figure out how to embrace and apply cloud-native techniques and approaches that have become pervasive in today's modern economy. Cloud-native network transformation drove the shift from vertically integrated stacks to disaggregated platforms and practices to deliver 5G Core and Cloud RAN hybrid infrastructures, and today it means taking the next step on the transformative journey: toward AI-native.

What is meant by *AI-native*? It helps to first consider an essential enabling step, which is transforming to cloud-native architectures. For those providers on the journey to cloud-native, it was necessary to step away from building highly specialized, proprietary service architectures that were heavily engineered, expensive to operate and maintain, and painfully slow to evolve. Cloud-native technologies opened the door to greater agility and reduced costs in developing services and applications based on software microservices, container compute architectures, CI/CD, and autoscaling and flexible placement of compute capacity, which is particularly important for edge network, data processing, and AI workloads. It is this type of distributed network that will enable CSPs to meet requirements to more easily satisfy customer experience, data regulations, and sovereignty requirements.

AI-native is the next natural step, building on and enabled by cloud-native architectures to insert, embrace, and leverage the power of AI technologies across virtually every aspect of providers' business planning and operations. The payoff will be, as with other transformational evolutions, greater operational flexibility and control of profitability while maintaining a sharp competitive edge.

Key to getting there efficiently and effectively will be the selection of supporting technologies that enable the transformation to AI-native architectures. Google Cloud and Intel have been working together to provide solutions that are specifically designed with CSPs in mind and are optimized for network workloads and for helping providers at each stage of this essential transition.
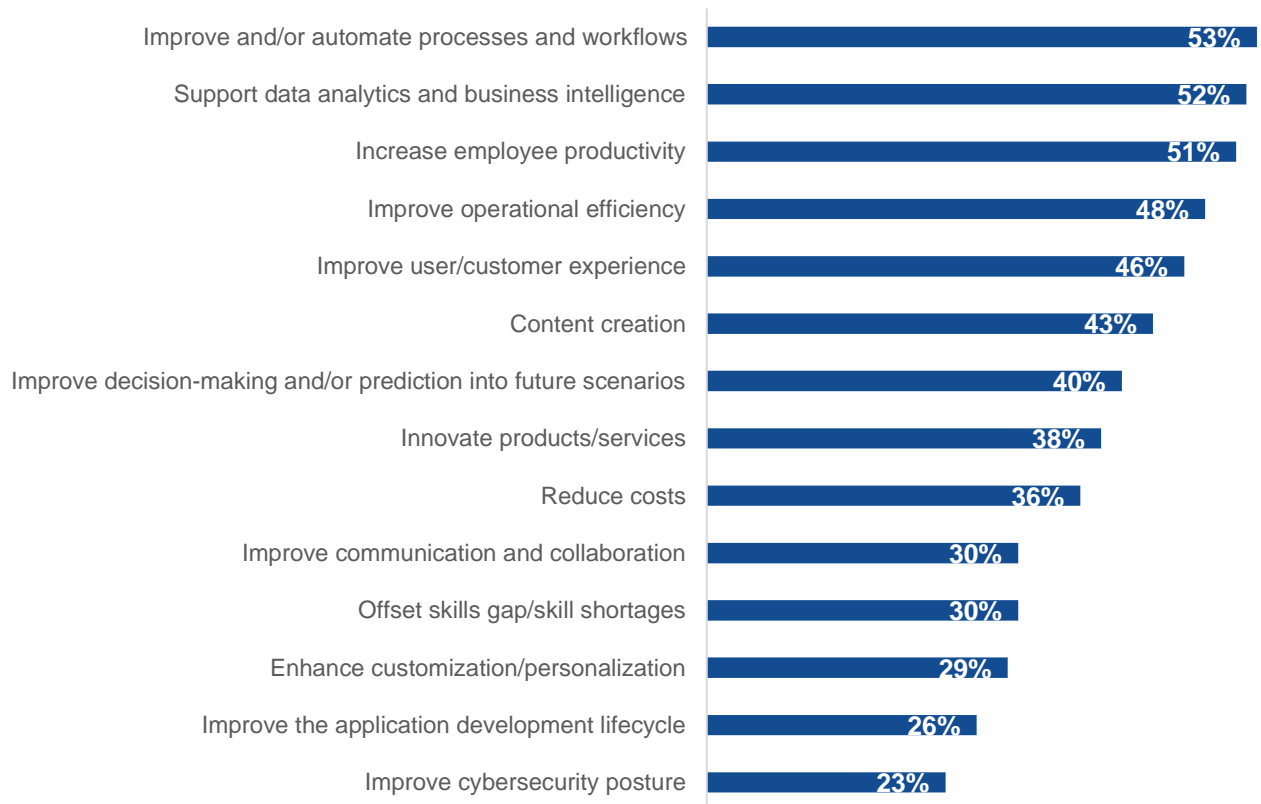
The Need for AI-native Transformation

Few would argue against AI's transformative potential for businesses across most sectors of the economy. AI is not new by any means—its roots go back to the 1950s—but until recently, most AI applications have been highly specialized and have required substantial effort to adapt to business uses. But innovations around large language models, generative AI (GenAI), and AI as a Service have changed that, broadening the range of use cases and accelerating the path to operational and cost efficiencies.

What makes AI perhaps so compelling is its range of applications and benefits. Even when considering GenAI alone, research from TechTarget's Enterprise Strategy Group has found a broad range of business benefits experienced by those deploying the technology (see Figure 1).[1]

Figure 1. Business Benefits of Using GenAI

Broadly speaking, what would you consider to be the primary benefits of using generative AI in your organization? (Percent of respondents, N=670, multiple responses accepted)

Benefit	Percent
Improve and/or automate processes and workflows	53%
Support data analytics and business intelligence	52%
Increase employee productivity	51%
Improve operational efficiency	48%
Improve user/customer experience	46%
Content creation	43%
Improve decision-making and/or prediction into future scenarios	40%
Innovate products/services	38%
Reduce costs	36%
Improve communication and collaboration	30%
Offset skills gap/skill shortages	30%
Enhance customization/personalization	29%
Improve the application development lifecycle	26%
Improve cybersecurity posture	23%

Source: Enterprise Strategy Group, a division of TechTarget, Inc.

---

[1] Source: Enterprise Strategy Group Research Report, *Beyond the GenAI Hype: Real-world Investments, Use Cases, and Concerns*, August 2023.

These benefits support several specific objectives relevant to providers. For instance, regarding operations, GenAI technologies are showing positives around improving workflow efficiency (including automations), improving communication and collaboration, and offsetting skill shortages. Important from a subscriber experience perspective, AI's ability to recognize patterns indicating pending service degradations or outages within monitoring data, recommend corrective actions to operations staff, and accurately deploy approved changes could finally make fully proactive service assurance a viable and reachable reality.

On the service product management front, GenAI has demonstrated utility for data analytics and business intelligence to better support decision-making and tune future scenario planning. GenAI can be purposed with reviewing existing service usage and adoption by customer tier and geography and suggest new offerings to improve competitiveness, reduce churn, and increase service revenues. When applied to customer care, AI assistants can improve self-help results, reducing call center demand and improving customer satisfaction.

Fortunately, AI is rapidly maturing and is already finding adoption for use cases such as network planning and operations. According to recent Enterprise Strategy Group research, 83% of IT organizations have been using AI to support their networking environment for two years or more.[2] And when those networking pros choose their AI technology, the most common source (cited by 57% of respondents) was CSP.[3] This means that there is a path today for making the journey to AI-native architectures by working with the same technology suppliers that made cloud-native transformation possible.

Google Cloud and Intel's Approach to AI-native Transformation for CSPs

As a long-standing operator of a global service network, Google Cloud knows what's important when it comes to the provider business and is making the transition to AI-native strategies based upon its own experiences and needs. Together with Intel, Google Cloud is optimizing and tuning the hardware, software, and automation capabilities into AI-native architectures.

Google Cloud first helped CSPs with their cloud-native transformation journey, bringing disaggregated cloud platforms for deployment of agile networks with built-in autonomous operations, and then delivered the same architecture to the on-premises edge as Google Distributed Cloud (GDC). The GDC, powered by Intel 3rd generation Xeon processors, delivered carrier-grade capabilities to run 5G Core Networks on premises and offloading control plane functions to the public cloud (GKE) to deliver a truly scalable hybrid cloud-native architecture. Innovating in virtualized radio access networks (vRANs), the journey continued with vRAN DU functionality using the Intel ACC100 vRAN accelerator and now natively using Intel 4th generation Xeon Sapphire Rapids EE processors.

Intel 4th Generation Xeon processors come with built-in accelerators that are powering the next generation of GDC products. Intel® Advanced Matrix Extensions (Intel® AMX) for AI in each core match deep learning training requirements for numerous applications that traditionally necessitated offloading to a separate GPU. The Intel® Advanced Vector Extensions 512 (Intel® AVX-512) instructions can accelerate the performance for vector-processing workloads, bringing AI inferencing to the edge without the overhead of additional hardware accelerator offload. This range of capabilities fuels the AI-native transformation, including:

- AI Anywhere
  - Consistently bring the best of Google Cloud data and AI services from the cloud to the on-premises edge using GDC.
  - Develop, build, train and execute AI models on this infrastructure to deliver AI anywhere, including validating models like Gemma, Llama 2, and more, meaning no lock-in.

---

[2] Source: Enterprise Strategy Group Complete Survey Results, *The Role of AIOps in Network Infrastructure Operations*, April 2024.
[3] Ibid.

- Operations and automation
  - ○ AI will process diverse data sets to provide real-time actionable insights and recommendations.
  - ○ Automation enables fast and accurate deployment of focused, closed-loop corrective or proactive measures into the control fabric, including Google Telecom Network Automation, which is based on open standards Nephio from the Linux Foundation.
- Real-time intelligence
  - ○ Deliver inference on the edge to optimize 5G Core, vRAN, and other workloads.
  - ○ Enable security applications to eliminate security blind spots by ingesting, normalizing, analyzing, and searching all security telemetry in real time to improve security posture.
- Multimodal Experience
  - ○ Build a set of LLM-based network assistants that transform how CSPs interact with their systems and networks.
  - ○ Bring image, speech, and text LLM capabilities together to drive richer experiences in employee productivity, field operations, issue isolation, and more.
- Sustainability
  - ○ As they continue to innovate in AI, Google Cloud and Intel are committed to sustainability.
  - ○ With Intel 4th Generation Xeon processors, significant improvements in performance, reduction in power consumption, reduction in CO_2 emissions, and optimized footprints are some of the key advancements.
  - ○ Strong collaboration with network equipment vendors to optimize their 5G and RAN workloads is also driving improved efficiencies.

Employing technologies and practices such as these puts operators in position to assume an AI-native stance, gaining business advantages and a competitive edge via outcomes such as:

- Improved operational agility through acceleration of work processes and new service introductions.
- Better operating margins via cost control and reduced operational risk.
- Higher subscriber retention thanks to curated, customized experiences.

Conclusion

Business change is a must, and in today's hypercompetitive communications services marketplace, providers must embrace the latest innovations to remain in the game. Making the journey to AI-native technologies and strategies is one of those changes that will be essential to help organizations survive and even thrive. For further information on how Google Cloud and Intel have joined forces, learn more here.

About Enterprise Strategy Group
TechTarget's Enterprise Strategy Group provides focused and actionable market intelligence, demand-side research, analyst advisory services, GTM strategy guidance, solution validations, and custom content supporting enterprise technology buying and selling.

contact@esg-global.com
www.esg-global.com