www.pipelinepub.com
Volume 22, Issue 10

How AI-Native Architecture is Unlocking a New Economics for Telecom

By: Pavan Thalak

Telecom operators have stopped debating whether AI belongs in their business. In NVIDIA's fourth annual State of AI in Telecommunications survey, 90 percent of operators reported that AI is already helping them grow revenue and reduce costs, and 89 percent said they plan to raise their AI budgets this year, up from 65 percent the year before. More than a third expect those budgets to grow by more than 10 percent. The capital is committed, and the conviction behind it is hard to miss.

What's far less settled is whether that spending translates to the business results operators were promised. Plenty of pilots look impressive in a controlled demo, then stall on the way to production. The distance between an AI budget and an AI outcome has become the defining question of this moment, and it has little to do with how much an operator spends.

The industry settled the ‘why AI?’ question, but then got stuck on ‘how’

The appetite for AI is genuine and broad. NVIDIA's data shows that 60 percent of operators now use or are assessing generative AI, up from 49 percent in 2024, and 65 percent report that AI already drives automation across their networks.

Network automation has overtaken customer experience as the top investment priority, with half of operators naming autonomous networks as their best-performing use case for return on investment. Looking further ahead, 77 percent expect AI-native networks to arrive before 6G, and a growing number of operators have started to describe themselves less as carriers of traffic and more as infrastructure companies for intelligence.

The momentum reaches the newest tools, too. Telecom posted the highest adoption of agentic AI of any industry in Nvidia's research, at 48 percent, and 89 percent of operators flagged open-source models as important to their strategy. Productivity gains are nearly universal, with 26 percent of operators reporting major improvements. The technology works, and operators know it works.

Readiness hasn't kept pace with that ambition

TM Forum's research into generative AI maturity tells the other half of the story, and the numbers are sobering. Only 25 percent of operators feel equipped to use advanced techniques like fine-tuning and retrieval-augmented generation, and just 16 percent feel confident using those methods to manage cost and prove return. About a third have a senior executive dedicated to AI strategy. Only 14 percent run more than 10 generative AI use cases in production, and roughly a third report a mature pipeline of work behind the ones they've already shipped.

The most common reason initiatives stall is accuracy, which tends to break down once a proof of concept meets the messiness of live operations. TM Forum reports that operators keep trying to attach generative AI to systems that were never designed for it, rather than building intelligence as a native part of how the business runs. That one architectural choice explains far more of the variance in results than the size of any budget does.

Architecture decides whether AI compounds or stalls

When AI sits on top of fragmented legacy systems, every use case turns into its own integration project. A customer-service bot might connect to one set of records while a fraud model connects to another, and neither shares context or learns from one another. What an operator ends up with is a portfolio of disconnected pilots that each deliver a thin slice of value while the cost of holding them together keeps climbing. The pilots that succeeded during the demo never compound into anything larger, because the architecture underneath them was never built to let them.

Embedding AI at the core of a unified stack helps solve this problem. The same intelligence layer that resolves a customer's billing question can recognize that the customer is a strong candidate for an upgrade without handing off to a separate system. Value compounds because customer experience, monetization and operations all draw on one model of the customer instead of three competing versions. A gain in one area feeds the next rather than stopping at a system boundary.

AI-native architecture creates value across three critical domains

The impact of AI-native architecture extends well beyond network operations. Operators investing in the core architecture are creating advantages across three areas: consumer experience, developer efficiency and operational excellence.

For consumers, AI-native systems remove friction from interactions by allowing subscribers to communicate in natural language rather than navigating disconnected applications and processes. The architecture translates customer intent into system actions, connecting previously isolated systems so requests can be resolved seamlessly. The result is a simpler experience that helps subscribers discover services and outcomes that may not have been visible through traditional workflows. For developers, a unified AI-native foundation improves efficiency across the software development lifecycle. Teams spend less time building and maintaining integrations between siloed systems and more time delivering customer-facing innovation. This advancement reduces complexity, accelerates feature releases and allows operators to respond faster to changing market demands.

Operationally, AI-native architecture helps teams identify, triage and resolve issues more effectively. By creating a shared intelligence layer across the organization, operators can proactively detect problems, automate troubleshooting and prevent disruptions before they impact customers. Operations can move from reactive management toward continuous optimization.

Consider what that looks like inside a single interaction.

A subscriber messages about an unexpected charge, and a layered system routes the question to a support bot that reads the billing record, resolves the charge and ends the conversation there.

A native system resolves the same charge, recognizes from the subscriber's usage that they keep hitting a data ceiling, offers a plan that fits the pattern and activates it on confirmation, all inside one exchange.

The first interaction closed a ticket. The second protected revenue, reduced the odds of churn and improved the subscriber's experience in the same motion.

The results from native deployments make the contrast more concrete. In work I've been close to, an AI concierge built into the core of the stack resolves 85 percent of customer queries without human escalation, handling tasks that range from network diagnostics to billing rather than the narrow scripts that earlier chatbots ran.

Operators running that kind of unified system have seen average revenue per user climb 22 percent and churn fall 9 percent, driven by personalized offers that the same system both recommends and fulfills, from what we've seen.

Those outcomes line up with the use cases Nvidia's respondents ranked highest for return, namely: autonomous networks, customer service and internal process optimization. The native approach produces them together, from one foundation, instead of one isolated pilot at a time.

The maturity path is open, but the window is closing

Maturity tends to move along a recognizable path, and knowing where you sit on it matters more than how quickly you push forward. Operators early on the path usually have a handful of disconnected pilots and no single owner for AI strategy, which is the profile TM Forum found across much of the industry. Their best first move is to pick one domain, customer care or network operations, and build it on an architecture that can carry the next use case, too.

Operators further along already have working use cases and can feel the cost of the boundaries between them. Their work is consolidation, pulling many narrow tools toward a shared intelligence layer that earlier investments can connect into. The operators setting the pace treat AI as the operating model of the business rather than a set of features bolted to the edge of it.

Underneath all of it sits a practical point about how to get there. Few operators should try to solve the architecture problem alone. The platform partners who've built native, full-stack systems have absorbed years of integration work that an individual operator would otherwise repeat from scratch. Buying into architecture that someone else has solved frees an operator's own engineers to focus on the problems specific to their market.

The timing matters more than it might appear

As connectivity becomes increasingly commoditized, operators that move first reach revenue categories that have little to do with selling data by the gigabyte. Embedded financial services offer the clearest near-term example.

A subscriber base that trusts an operator with monthly payments is a natural audience for wallets, cards and cross-border transfers, and operators that own a native platform can add those services without standing up a bank's worth of infrastructure behind them. Personalized commerce and anticipatory care follow the same logic, each drawing on the customer model the operator already

runs.

The economics of telecom are being rewritten

Whether an operator builds intelligence into the foundation or attaches it afterward will set the cost structure and the revenue ceiling it operates under for the next decade. That decision is being made this year, in budget cycles and vendor choices that rarely announce themselves as architectural. An operator can spend heavily and still end up with a portfolio of pilots that never compound, or it can put the same money into a foundation where each new capability makes the ones before it more valuable.

The first outcome leaves the business defending margins that thin a little more every year. The second turns the same subscriber base into the engine for revenue that connectivity alone can't reach. Operators who treat this as an architecture question will own the economics the rest of the industry spends the next decade trying to catch.