Living on the Edge: Singtel, Microsoft and NVIDIA Dial Up AI Over 5G

Singtel rolls out multi-access edge computing over 5G in Singapore and elsewhere in partnership with Microsoft Azure, backed by NVIDIA accelerated computing.

Author: Joao Gomes

For telcos around the world, one of the biggest challenges to upgrading networks has always been the question, “If you build it, will they come?”

Asia’s leading telco, Singtel , believes the key to helping customers innovate with AI across industries — for everything from traffic and video analytics to conversational AI avatars powered by large language models (LLMs) — is to offer multi-access edge compute services on its high-speed, ultra-low-latency 5G network.

Multi-access edge computing, or MEC, moves the computing of traffic and services from a centralized cloud to the edge of the network, where it’s closer to the customer. Doing so reduces network latency and lowers costs through sharing of network resources.

Singtel is collaborating with Microsoft and NVIDIA to combine AI and 5G, so enterprises can boost their innovation and productivity. Using NVIDIA’s full-stack accelerated computing platform optimized for Microsoft Azure Public MEC, the telco is creating solutions that enable customers to leverage AI video analytics for multiple use cases and to deploy 5G conversational avatars powered by LLMs.

Singtel has been rolling out enterprise 5G and MEC across ports, airports, manufacturing facilities and other locations. In addition to running low-latency applications at the edge using Singtel’s 5G network, the solution has the potential to transform operations in sectors such as public safety, urban planning, healthcare, banking, civil service, transportation and logistics. It also offers high security for public sector customers and better performance for end users, enabling new intelligent edge scenarios.

Customers can use these capabilities through Microsoft Azure, only paying for the amount of compute and storage they use for the duration in which they use it. This replicates the cloud consumption model at the network edge and lets users save on additional operational overhead.

Singtel is working with video analytics software-makers participating in NVIDIA Inception , a free program that offers startups go-to-market support, expertise and technology. These ISVs will be able to use the NVIDIA Jetson Orin module for edge AI and robotics in conjunction with Microsoft MEC to identify traffic flows at airports and other high-population areas, retail video analytics and other use cases.

Singtel and NVIDIA are also showcasing their technology and solutions, including a real-time LLM-powered avatar developed by system integrator Quantiphi and based on NVIDIA Omniverse digital twin technology, at a May 11 launch event in Singapore. The avatar, built with NVIDIA Riva speech AI and the NeMo Megatron transformer model, enables people to interact in natural language on any topic of interest. Businesses can deploy these avatars anywhere over 5G.

Using Singtel’s high-speed, low-latency 5G — combined with NVIDIA AI accelerated infrastructure and capabilities — enterprises can explore use cases on everything from computer vision and mixed reality to autonomous guided vehicles.

Singtel plans to expand these new capabilities beyond Singapore to other countries and affiliated telcos, as well. This collaboration will help redefine what’s possible through the powerful combination of compute and next-generation networks, unlocking new operational efficiencies, revenue streams and

customer experiences.

Original URL: https://blogs.nvidia.com/blog/2023/05/11/singtel-microsoft-ai-over-5g/