



Scaling AI education for hundreds of students with Backend.AI at Kookmin University

Discover how Kookmin University built a scalable
GPU-based learning platform with Backend.AI



Scaling AI education for hundreds of students with Backend.AI at Kookmin University

“When we started with just three GPU servers, Backend.AI enabled 80+ students to run modeling exercises simultaneously by container-level GPU virtualization, all without dedicated administrative support.”

KMU
Professor, AI, Big Data, and Convergence Management
Yoonho Cho

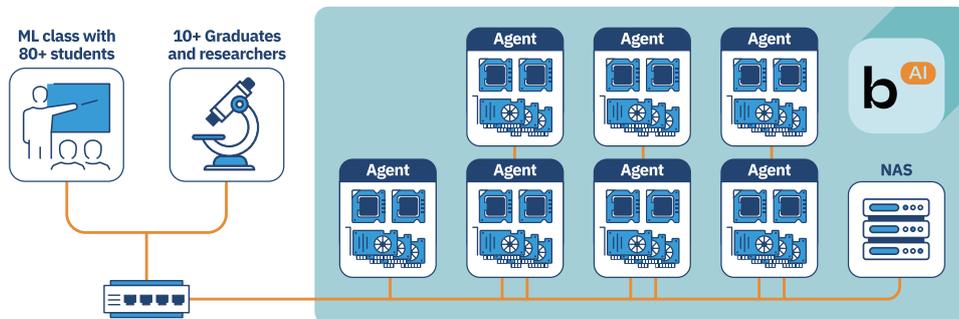
As one of the leading educational institutions, Kookmin University (KMU) is one of the leading institutions in Korea. KMU required powerful GPU resources to support hands-on machine learning courses for a large number of students. The growing demand for AI and data science education led the university to seek an efficient solution that could provide simultaneous access to GPU resources for hundreds of students ranging from MBA professionals to undergraduates and researchers in the business school.

Limited GPUs, but wide demand

The university faced the challenge of limited GPU availability, limiting concurrent usage. Supporting over 80 MBA students and hundreds of undergraduates simultaneously with only 16 GPUs required a scalable infrastructure capable of handling multiple instances without resource contention. For a business school environment, securing experts with specialized skills was particularly challenging. The university therefore required an infrastructure with two key features: flexible resource policies capable of separating usage across different groups, and an intuitive administrative interface that could be operated efficiently without the need for dedicated on-site technical staff.

Backend.AI enables virtualization of GPUs into hundreds of instances

To address these challenges, KMU implemented Backend.AI, which virtualized 16 nodes of GPUs into hundreds of instances, enabling graduate and undergraduate students to engage in machine learning labs concurrently. Due to Backend.AI's demonstrated high performance and utilization, the university expanded GPU infrastructure progressively over 3 years, increasing the total count to 57.



With Backend.AI, two teaching assistants without CS backgrounds efficiently managed resources across multiple servers. The platform enforced clear policies separating students and researchers and supported dynamic workload allocation by academic schedule. During breaks, it shifted to corporate AI training, maximizing GPU use and efficiency.