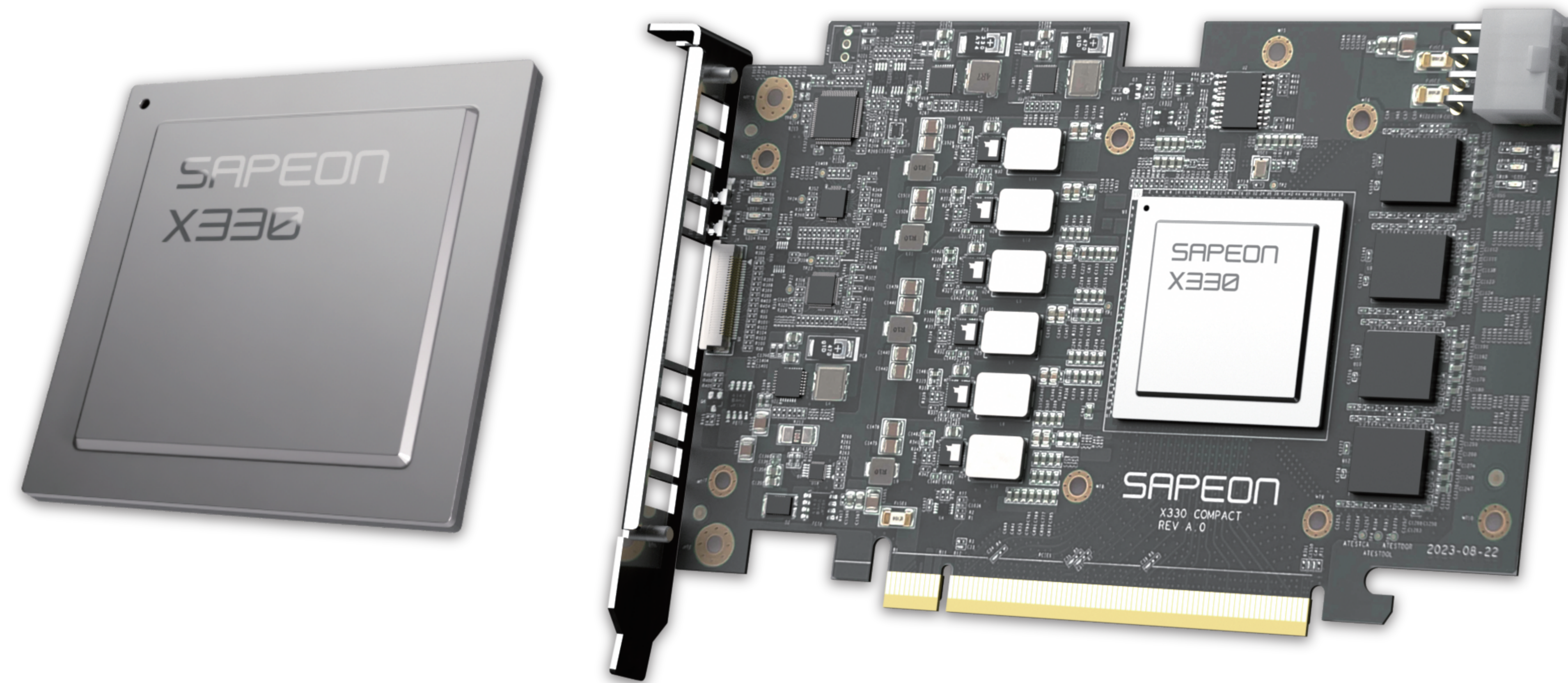


SAPEON Next Generation: X330

One tiny spark for the chip,
a quantum leap for AI inference



X330 Specifications



X330 Compact



X330 Prime

Precision		INT8, FP8, FP16	INT8, FP8, FP16
FP8 FP16 Performance		367 184 TFLOPS	734 368 TFLOPS
Memory	Type	GDDR6 X 8 (ECC)	GDDR6 X 16 (ECC)
	Capacity	16 GB	32 GB
	Bandwidth	256 GB/s	512 GB/s
Host interface		PCIe Gen5 16 Lane	PCIe Gen5 16 Lane
TDP		75~120 W	250 W
Form factor		FHHL Single	FHFL Single
Thermal solution		Passive	Passive

Industry leading performance MLPerf™ Benchmark

MLPerf™ benchmark results

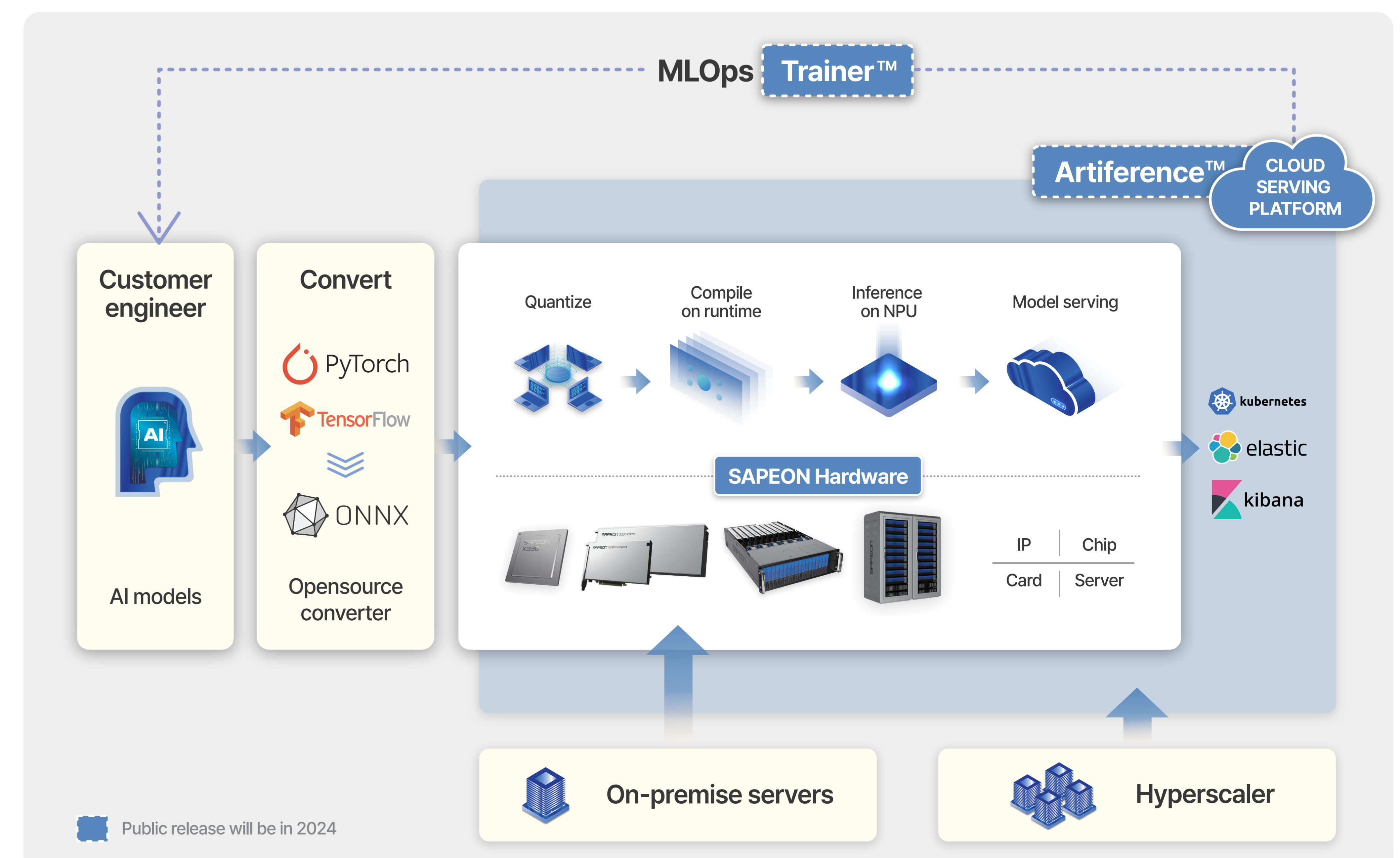
MLPerf™ benchmark v2.1 – Inference: datacenter



SAPEON estimation



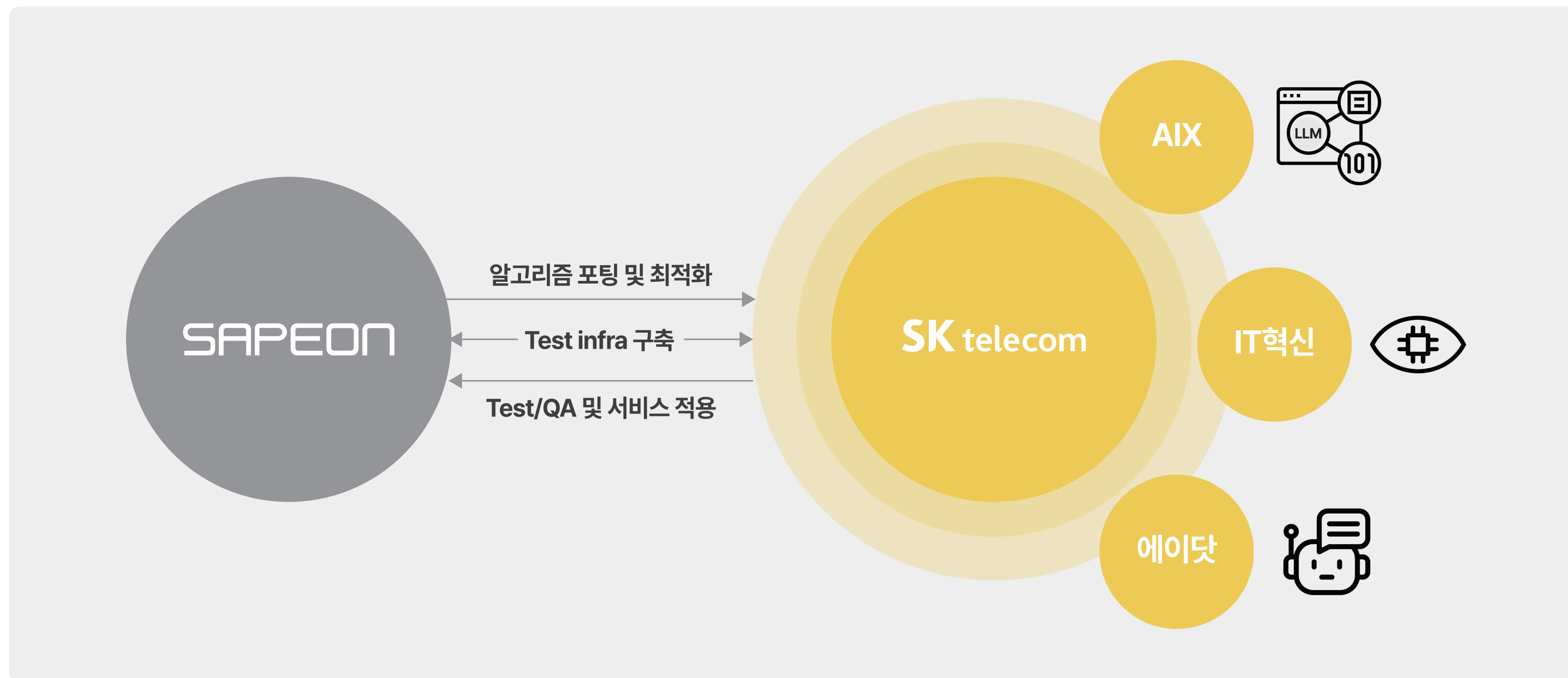
Zero-touch™ SAPEON AI full stack



SAPEON X SK telecom Collaboration

NPU Farm

SAPEON과 SKT는 NPU Farm을 통해 Test infra를 구축하고, 개발 협력을 통해 AI 혁신을 제공합니다.
SKT의 다양한 AI 조직이 SAPEON과 함께 AI 기술 기반의 혁신을 만들어 나가고 있습니다.

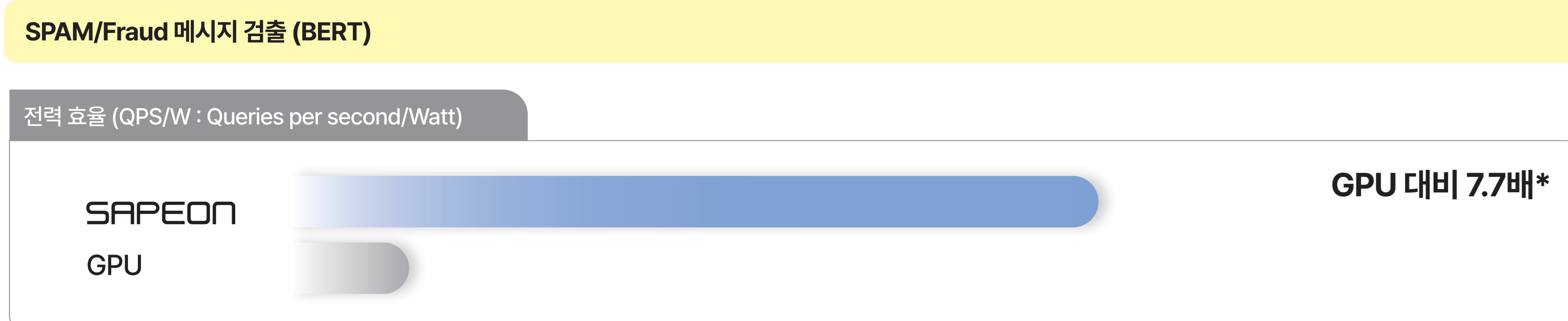
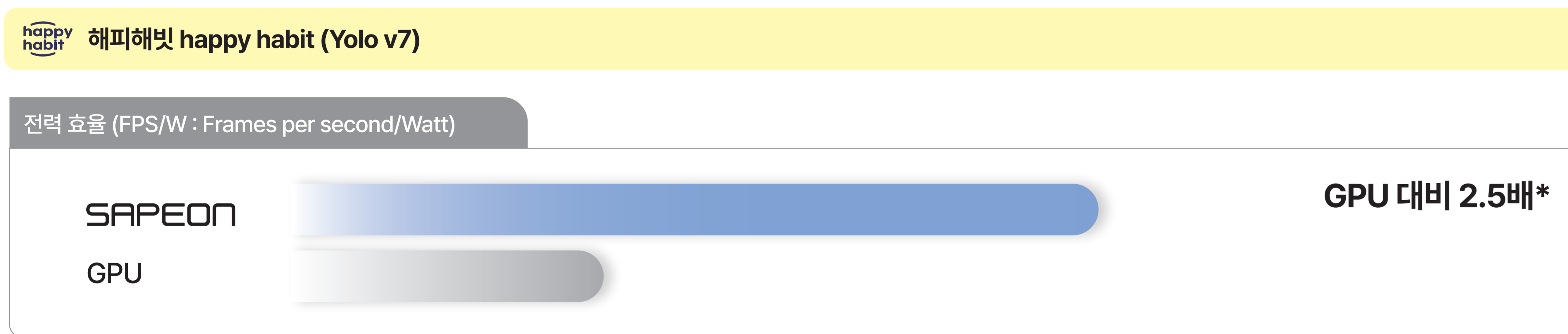


29 Servers
58 Cards

7.6 Peta OPS

Area	Application
Visual analysis	happy habit 해피해빗(컵 수거)
	AI CCTV
Language processing	X Caliber(동물병원 X-ray 분석)
	SPAM/Fraud 메시지 검출
Image enhancement	SUPERNOVA (Live converter)

성능비교

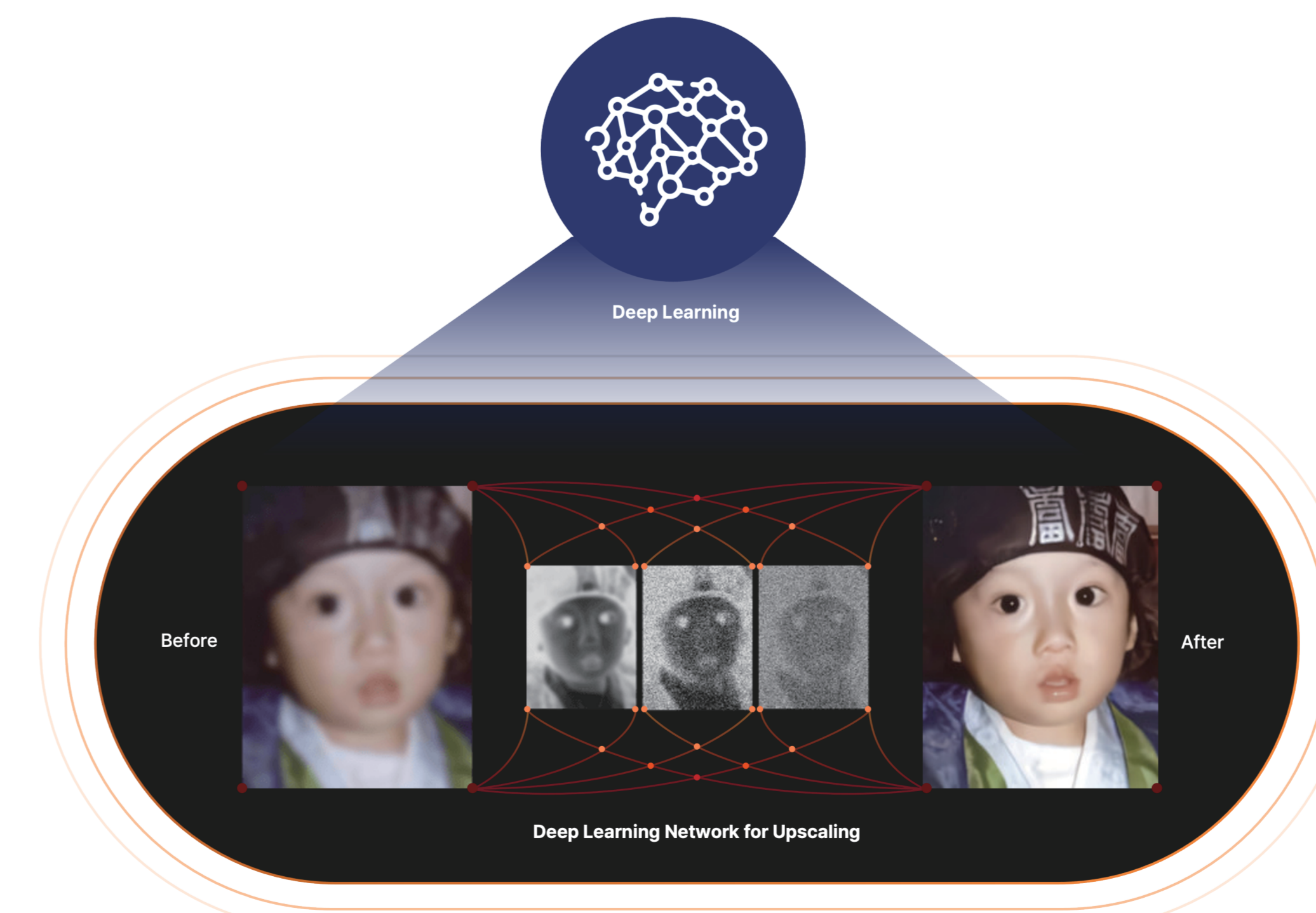


* INT8 기준, Tensor RT와 비교

SUPERNOVA

AI/Deep learning media technology for quality enhancement : SUPERNOVA

SUPERNOVA는 SK telecom의 AI 기술로, 딥러닝 기반 AI를 활용해 이미지, 영상, 오디오 등 다양한 미디어의 품질을 개선합니다.



Live converter | VoD transcoder powered by SAPEON

SAPEON x SKT SUPERNOVA 엔진을 기반으로 하는 VoD transcoder와 Live converter 개발을 완료하였으며, 개발과 동시에 지상파 방송사인 MBC에 VoD transcoder 공급을 완료하였습니다. SAPEON은 독보적인 데이터 처리속도를 구현하여 고품질 영상 서비스를 위한 핵심 기능을 빠르게 제공합니다.

