

IBKS NVIDIA GTC 참관기

# 전쟁과 유가의 무풍지대

나는 여기서 AI의 미래를 보았다.

IT/반도체 김운호

02) 6915-5656

unokim88@ibks.com







IBK투자증권에서 반도체 담당하는 김운호입니다.

아직 미국 일정이 남아서 보고서는 산호세에서 작성하고 있습니다. 이런저런 제약 요건은 있지만 빨리 마무리하는 것이 최선이라고 생각해서 현지에서 마무리했습니다.

제목을 “전쟁과 유가의 무풍지대 - 나는 여기서 AI의 미래를 보고 간다”로 정했습니다. 제법 거창한 제목입니다. AI의 미래는 확정된 것은 아닐 것이고 누구라도 정확하게 아는 것도 아니겠지만 지금 여기에 있는 만큼 그리 생각하고 있습니다. 그리고 확실히 이곳은 전쟁도 유가도 불가침의 영역으로 보였습니다.

NVIDIA GTC는 제가 다녀온 CES, MWC, COMPUTEX, IFA와는 결을 달리 합니다. 근본적인 차이점은 한 회사가 주관한다는 점입니다. 그러다 보니 모든 행사의 근간에는 NVIDIA가 중심이 되어 있습니다. GPU 중심으로 성장한 회사가 앞으로 다가올 AI 시대의 변화를 준비하면서 솔루션을 제시한 행사로 생각합니다.

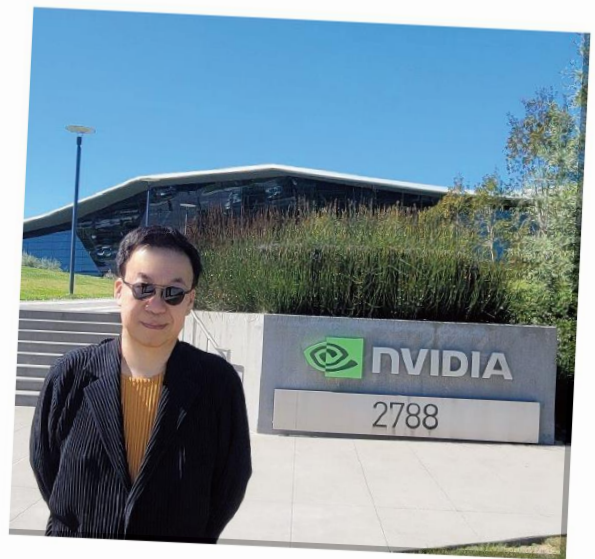
이번 NVIDIA GTC에서 제가 중점적으로 체크한 내용은 다음 3가지입니다.

1. GPU(NVIDIA)의 영향력 강화
2. Open Model의 역할과 중요성
3. Agentic AI 시대 개막

GTC 참관 전후 제가 갖고 있던 NVIDIA에 대한 시각이 조금 바뀌었습니다. 가상화폐, ChatGPT로 인해 우연히 대박난 회사로 생각하다가 여기 와서 보니 NVIDIA는 AI 생태계를 조성하고 솔루션을 제공하는 선구자 역할을 하고 있는 것으로 보입니다.

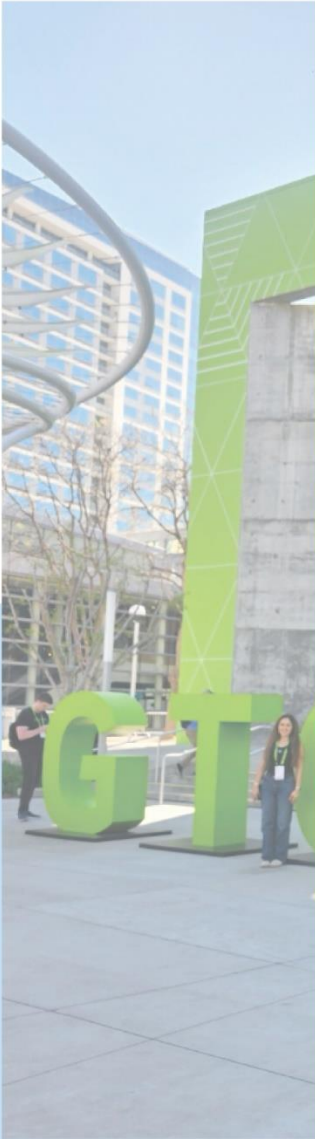
그리고 운이 좋게도 현지 벤처캐피탈 관계자를 만날 기회가 있었습니다. 이와 관련된 내용도 보고서에 있습니다. 피상적으로 알았던 AI 관련 내용을 현실적으로 이해할 수 있는 계기가 되었습니다. High End GPU의 수요, AI 관련 투자에 대한 이모저모, 쿼텀 관련 VC에서 보는 시각, Physical AI에 대한 접근 방법, 그리고 AI가 우리에게 미치는 영향력까지 정리 요약할 수 있었습니다.

지난 보고서에 이어 이번 보고서를 현지에서 작성하다 보니 이래 저래 부족한 점이 많습디다만 제가 느낀 점에 대해서는 가감없이 반영된 것으로 생각합니다.



# CONTENTS

개발자들의 축제 NVIDIA GTC .....	5
NVIDIA GTC 개요.....	5
행사 구성.....	6
Jensen Huang의 Vision(Keynote).....	9
Forum 이모저모 .....	13
The Era of GPU Data Processing .....	14
Inside the NVIDIA Inference Platform and Ecosystem.....	16
Open Models : Where we are?.....	18
Agentic AI 101.....	20
어느 실리콘밸리 벤처투자가의 이야기 .....	21
전시장 풍경 – NVIDIA H/W 신제품, Liquid Cooling, Memory.....	23
NVIDIA Vera Rubin System.....	23
Liquid Cooling System.....	31
Memory: HBM, SOCMAM2, SSD.....	32



## 개발자들의 축제 NVIDIA GTC

### NVIDIA GTC 개요

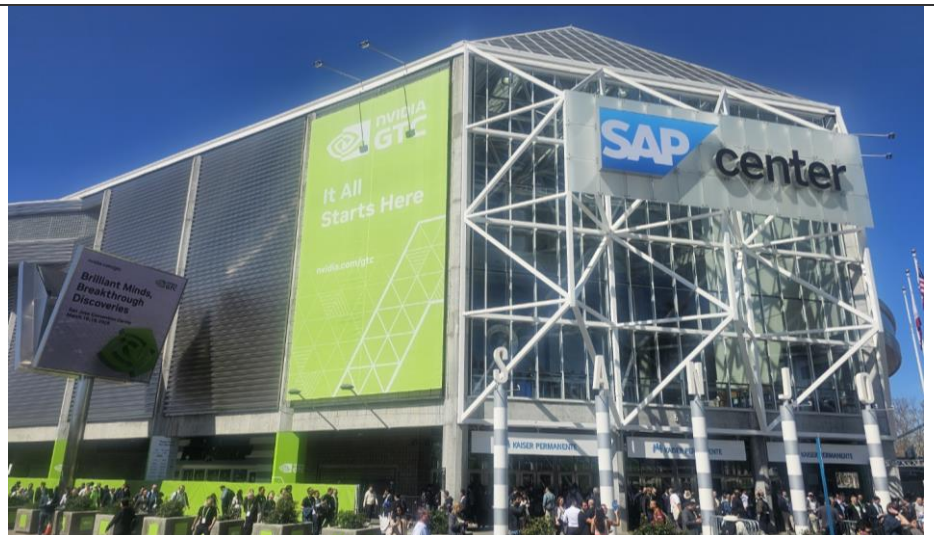
NVIDIA 주관 AI 중심  
개발자 컨퍼런스

NVIDIA GTC(GPU Technology Conference) 2026은 3월 16~19일(현지시간) 캘리포니아주 산호세 맥에너리 컨벤션센터에서 진행됐다. GTC는 2009년 산호세 페어몬트 호텔에서 처음 시작된 AI 및 가속컴퓨팅 중심 개발자 컨퍼런스이다. 초창기 GTC는 작은 전시회 분위기에 첫 행사 참석자도 약 1,500명 수준에 그쳤으나 현재는 190개 이상 국가에서 30,000명 이상이 참여하는 대형 행사로 성장했다.

AI 발달에 따른  
논의 범위 확대

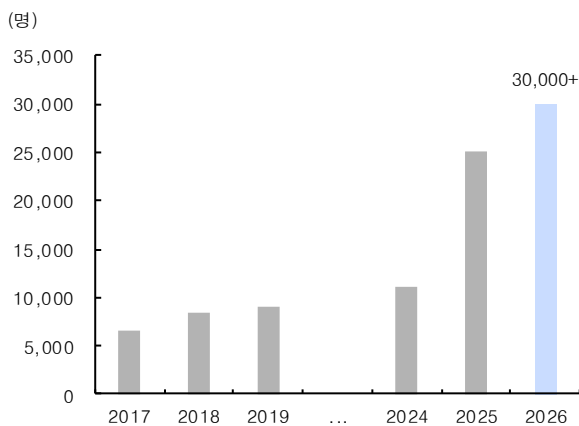
행사의 주제도 다양해졌다. 초기에는 GPU와 게임 분야에 대한 논의가 주를 이뤘지만, 생성형 AI 확산과 함께 GPU 기반 연산 수요가 늘어나면서 Robotics, Automotive, 양자컴퓨팅 등으로 논의 범위가 넓어진 모습이다.

그림 1. GTC 2026 Keynote 행사가 열리는 SAP Center 전경



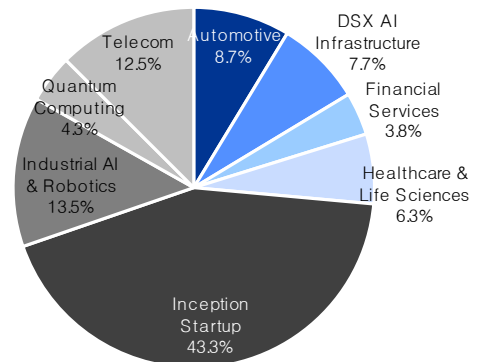
자료: IBK투자증권

그림 2. 연도별 GTC 오프라인 참석자 수



자료: 언론보도, IBK투자증권  
주: 2020~2023년 행사는 온라인 개최

그림 3. GTC 2026 업종별 전시 기업 비중



자료: NVIDIA, IBK투자증권

## 행사 구성

### 14개 주요 주제

GTC 2026은 NVIDIA CEO인 Jensen Huang의 Keynote를 비롯해 총 700개 이상의 세션으로 행사가 진행됐다. 세션은 1) Agentic AI / Generative AI, 2) AR / VR, 3) Computer Vision / Video Analytics, 4) Content Creation / Rendering, 5) Data Center / Cloud, 6) Data Science, 7) Developer Tools and Techniques, 8) Edge Computing, 9) MLOps / AIOps, 10) Networking and Communications, 11) Quantum Computing, 12) Robotics, 13) Simulation / Modeling / Design, 14) Trustworthy AI / Cybersecurity 등 14개 분야를 중심으로 편성했다.

### Keynote: 다양한 산업을 아우르는 NVIDIA 플랫폼

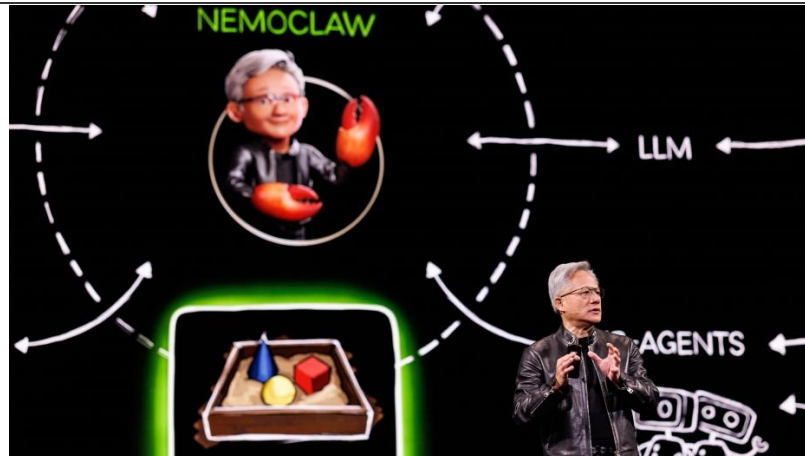
NVIDIA의 창립자이자 CEO Jensen Huang의 Keynote는 3월 16일 진행됐다. “가속 컴퓨팅과 AI의 미래, 그리고 이러한 기술들이 어떻게 전 산업에 걸쳐 차세대 컴퓨팅 시대를 만들어갈 것인가”가 주제였다. Jensen Huang 대표는 CUDA-X 라이브러리를 활용해 다양한 산업의 AI를 지원할 수 있음을 강조했다. Vera Rubin 아키텍처, OpenClaw용 NVIDIA NemoClaw 그리고 Newton 엔진과 NVIDIA Omniverse 기반 시뮬레이션이 적용된 디즈니의 올라프 로봇도 Keynote에서 공개됐다.

표 1. GTC 2026 세션 주요 주제

Key Topics	내용
Agentic AI	기업 및 오픈소스 AI 혁신
AR / VR	XR 및 공간 컴퓨팅 분야의 발전이 workflow와 몰입형 경험을 혁신적으로 변화
Computer Vision / Video Analytics	컴퓨터 비전 및 비디오 분석 분야의 최첨단 기술과 최신 발전 사항
Content Creation / Rendering	차세대 생성형 AI 및 렌더링 기술
Data Center / Cloud	인공지능 시대에 확장 가능한 인프라 구축
Data Science	NVIDIA CUDA-X 라이브러리 활용 가속 인프라 기반 데이터 파이프라인
Developer Tools and Techniques	NVIDIA CUDA 및 CUDA-X 모델, 라이브러리 등 개발자 도구
Edge Computing	강력한 AI 컴퓨터를 통한 차세대 엣지 제품 구현
MLOps / AIOps	생성형 AI를 운영 환경에 배포하고 확장하기 위한 MLOps 전략 수립 및 구현
Networking and Communications	AI Factory를 지원하는 고대역폭 네트워킹 및 가속 컴퓨팅
Quantum Computing	양자 컴퓨팅의 미래
Robotics	인공지능 기반 로봇의 개발, 시뮬레이션 및 배포를 위한 end-to-end 솔루션
Simulation/Modeling / Design	NVIDIA 기술이 설계 및 생산 모든 단계에서 디지털 트윈 개발 혁신
Trustworthy AI / Cybersecurity	사이버 보안에서 혁명을 일으키는 Agentic AI

자료: NVIDIA, IBK투자증권

그림 4. Keynote 발언하는 Jensen Huang 대표



자료: NVIDIA, IBK투자증권

그림 5. Jensen Huang 대표와 디즈니 올라프 로봇



자료: NVIDIA, IBK투자증권

다양한 불거리의 AI 축제

GTC 2026은 세션과 전시뿐 아니라 행사장 내외에서도 AI를 활용한 다양한 불거리를 제공했다. 컨벤션센터와 전시장에서는 음식 및 기념품을 배송하는 로봇을 확인할 수 있었고 행사장 외부 GTC Park에서는 참가자들이 직접 개인용 Agent를 만들 수 있는 Build-a-Claw 이벤트가 개최됐다.

전문가뿐 아니라 일반인도 참여 가능한 체험형 전시

행사장 내부에는 Jensen Huang 대표를 본뜬 홀로그램 AI 에이전트 “Toy Jensen(TJ)”이 배치됐다. TJ는 LiveX.ai 플랫폼과 NVIDIA Nemotron Open Model, NVIDIA NIM 마이크로서비스 등을 기반으로 구현됐으며, 참가자들의 질문에 답하고 사진 촬영에도 응하는 체험형 전시로 운영됐다. 이번 프로그램들은 기술 중심 컨퍼런스를 넘어, 비전문가 참가자들도 생성형 AI를 실제 활용하고 경험하도록 구성됐다.

그림 6. Toy Jensen과 Jensen Huang 대표



자료: NVIDIA, IBK투자증권

## Jensen Huang의 Vision (Keynote)

NVIDIA GTC 2026은 2개의 캐치프레이즈를 사용하고 있다. 하나는 지난 해에도 내세웠던 It all starts here이고, 다른 하나는 The next generation of AI begins이다.

차세대 AI / Agentic AI 시대가 오고 있다.

많은 개발자, 창업자들이 여기에 참가하고 여기에 참가했던 엔지니어가 창업자가 되고 있어서 It all starts here(여기서 모든 것이 시작된다)를 자신 있게 캐치프레이즈로 사용하는 것으로 보인다. 두번째 The next generation of AI begins(AI의 다음 세대가 시작되었다)는 한 마디로는 Agentic AI 시대로 볼 수 있다. GTC 2026에서 강조한 GPU의 역할 강화, AI Factory, Open Model, Agentic AI, Physical AI 등의 주제와 일맥 상통하고 있다는 인상이다.

AI가 발전할수록 Computing수요가 더 커진다.

Training AI 이후 Inference AI 시장이 도래하면 High End GPU 시장은 점차 사그라들 것으로 인식되었으나 Inference AI에서도 High End GPU의 역할은 더욱 커지고 있다. 가속기로서의 역할이 아니라 점차 연산에도 관여하고 있고 그 의존도가 높아지고 있다. Hyperscaler들의 모델 개발과 학습에 필요한 것뿐만 아니라 AI 모델의 Inference를 지원하는 플랫폼 기업들의 수요도 아직은 가파르게 성장하고 있는 국면이기 때문이다.

AI Factory: 단순한 H/W가 아닌 시스템이 필요해진다.

AI Factory는 2026년에 처음 등장한 말은 아니지만 Vera Rubin 시스템으로 좀 더 강화되는 인상이다. Computing을 넘어선 네트워킹 / 메모리 / 패키징 / 전력 / 냉각 / 추론 SW까지 아울러서 AI Factory라고 언급한다.

AI의 현재 버전은 Agentic AI

Generative AI가 등장했을 때는 모두가 Generative AI 엔진을 확보하는 데 집중했지만, 이후 더 높은 성능의 AI에 대한 요구가 커지면서 AI는 Reasoning AI로 진화하게 되었다. 단순 검색 이상의 결과를 원하는 수요자를 대응하기 위한 모델이다. 그리고 지금 등장하기 시작한 것이 Agentic AI이다. Agentic AI는 생각에서 그치는게 아닌 계획하고 실행까지 가능하게 하는 솔루션을 의미한다.

Open Model 확장. 이는 결국 NVIDIA의 미래 먹거리

가장 강조한 섹션 중에 하나는 Open Models이다. 이 세션은 Jensen Huang이 직접 사회를 봤는데 LangChain, Cursor, Perplexity, Mistral과 같은 Open Model 회사들과 토론하며 Open Model의 방향성을 제시했다.

이는 NVIDIA가 Open Model의 확산을 주도하고 있고 이를 통해서 Computing 시장과 네트워크 기능의 확산을 유도하는 것을 의미한다. NVIDIA의 미래 먹거리와도 연결 고리를 갖고 있는 것으로 판단한다.

광 네트워크, LPU를 도입해서 Agentic AI 지원

Inference를 대비하기 위해서 GPU를 지원하는 LPU를 도입했고, Latency 문제를 물리적으로 해결하기 위해서 CPO(Co Packaged Optics)를 시스템에 채택했다.

## Key Note를 중심으로 요약하면

### 1. NVIDIA는 Platform 기업이다

단순 GPU 메이커가 아닌 AI 시장 형성에 필요한 모든 서비스를 제공

GPU를 만드는 Chip Maker이지만 AI의 핵심인 CUDA를 세상에 알렸고, 이를 바탕으로 한 AI 엔진들이 시장을 주도한 영향도 있고, AI Enterprise, NIM(NVIDIA Inference Microservices), Mission Control 등의 소프트웨어도 NVIDIA의 칩과 연동해서 제품 라인을 구축하고 있다.

Open Model을 강조하고 NVIDIA가 제시하는 솔루션으로 창업 문턱을 낮추는 것도 결국은 Platform 기업으로 입지를 강화하는 수단으로 판단한다.

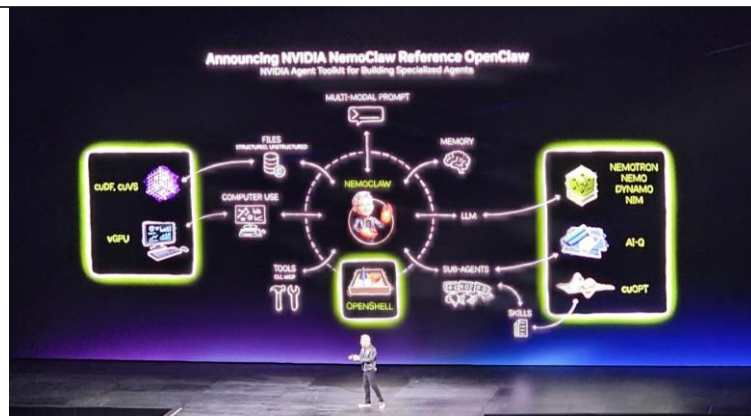
### 2. AI Chip 시장 규모는 예상보다 빠르게 성장한다.

추론 시장에서도 Computing은 핵심 자원

Jensen Huang은 2025년 GTC에서 AI Chip 시장 규모를 \$0.5T로 전망했던 것을 2026년에는 \$1T이 넘을 것으로 상향 조정했다. 연산 관련된 Chip 수요가 예상보다 훨씬 더 빠르게 성장한다는 것을 방증하는 지표로 판단한다. 2026년 회계연도 NVIDIA 매출액은 \$215.9B이다.

또 다른 내용은 AI 모델이 Agent로 진화하면서 토큰 소비량이 급격히 증가한다는 내용이다. ChatGPT에 비해서 Claude code는 토큰을 10,000배 더 필요로 한다. 그 중간 단계에 추론형 O1 Reasoning Model은 Chat GPT 대비 10배 더 사용하고, Claude code는 O1 reasoning model 대비 100배 더 많이 사용하고 있기 때문이다.

그림 7. NVIDIA의 NemoClaw



자료: NVIDIA, IBK투자증권

### 3. 신제품 소개(Vera Rubin System + Groq3, Feynman)

AI 성장을 위한 새로운 동력이 Vera Rubin 시스템

2025년부터 예견했던 Vera Rubin 시스템을 공개했다. 10년 전 대비 연산 능력이 4천만배 더 커진 규모이다. 이전 모델인 Blackwell 대비 연산속도가 빨라졌지만 HBM4를 사용하고 용량은 288GB로 동일하다.

여기에 강력한 CPU인 Vera와 함께 시스템을 구성하고 있다. Blackwell은 Grace CPU와 시스템을 구성했다. 시장 규모는 이전 시스템 대비 \$150B 이상 커질 것으로 전망한다.

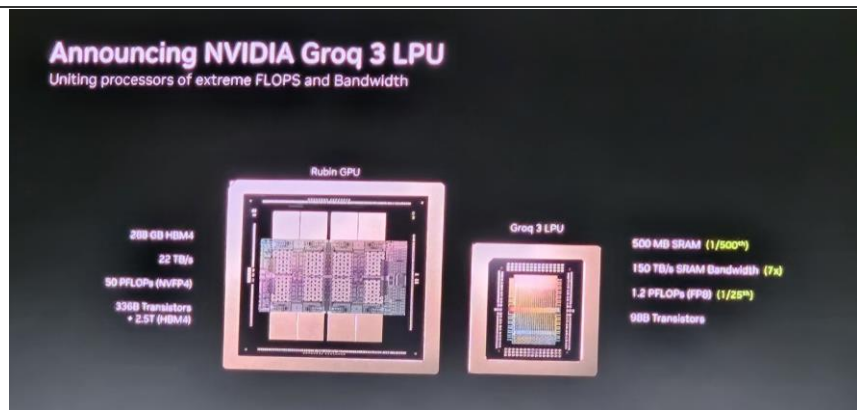
신규로 추가된 Chip은 LPU Groq3이다. SRAM 500MB를 사용한다. SRAM은 DRAM 대비 고가이고 빠른 속도가 특징이다. 용량 측면에서는 DRAM이 훨씬 앞서서 주로 Cache Memory로 사용된다. Groq3가 포함된 Vera Rubin + LPX 시스템은 시장 규모가 \$300B 수준일 것으로 보인다.

표 2. 1GW AI 공장 기준 Hopper, Vera Rubin 성능 비교

	X86 + Hopper	Vera Rubin
GPU 수	600K	300K
AI FLOPS	1.2 ZFLOPS	16 ZFLOPS
AI-to-AI Scale-up	7.2 TB/s	260 TB/s
Memory BW-per-Domain (Groq SRAM)	2 EB/s	100 EB/s
Tokens per Second	2M	700M

자료: NVIDIA, IBK투자증권

그림 8. NVIDIA Rubin GPU와 Groq3 LPU



자료: NVIDIA, IBK투자증권

#### 4. NemoClaw

Open Model을 지원하는 새로운 솔루션이 Nemo Claw

NemoClaw는 NVIDIA가 만든 OpenClaw용 보안·운영 레이어이다. NVIDIA의 Open Model인 Nemotron에서 Nemo를 인용했다. NVIDIA 공식 설명에 따르면 NemoClaw는 NVIDIA Agent Toolkit과 OpenShell을 써서 OpenClaw를 단일 명령으로 설치·보호하고 정책 기반 보안, 네트워크 가드레일, 프라이버시 제어 기능을 추가한 제품이다.

OpenClaw는 오픈소스 AI 에이전트 자체 / 생태계를 언급하는데 항상 동작하고 있고 스스로 진화하는 에이전트(Claw)를 업무에 쓰는 흐름 자체이다. OpenClaw를 오픈소스 AI Agent system으로 볼 수 있다.

OpenClaw를 최적화할 수 있는 모델이 NVIDIA의 NemoClaw이고 이는 Open Model 확산의 핵심 톨로 NVIDIA가 제공하는 솔루션이다.

#### 5. 우주 데이터센터 / 자율 주행 / Physical AI

자율 주행은 코 앞에 와 있고, Physical AI도 가상 시뮬레이션으로 빠르게 도입 목표

향후 전략으로는 우주 데이터센터 / 자율 주행 / Physical AI에 대해서 소개했다. 인공지능 인프라 영역을 우주까지 확장하겠다는 계획을 발표했고 이미 Vera Rubin Space-1이라는 우주 전용 컴퓨터를 개발 중에 있다.

자율주행 분야에서는 현대, 닛산, BYD 등 글로벌 완성차 업체들이 Robotaxi-Ready 플랫폼에 합류했다고 발표했다. 앞으로 빠르게 사업 영역을 확장할 계획이다. 벤츠 CLA 자율주행 영상도 시연했다.

Physical AI 분야에서는 Groot와 Omniverse 플랫폼으로 로봇이 현실 세계에 적응하는 속도를 혁신적으로 단축할 것이라 발표했다.

알고리즘부터 칩, 시스템까지 수직적으로 통합해 성능을 올리되, 이를 모든 클라우드와 On-premise 환경에서 적용할 수 있도록 수평적으로 개방하는 전략을 고수할 계획이다.

## Forum 이모저모

NVIDIA GTC는 개발자 중심으로 14가지 기술과 20개 산업을 대상

NVIDIA GTC는 다른 행사와는 많이 다른 양상이다. Forum 주제가 기술 섹터가 14개, 산업 섹터가 20개이다. 이와 관련된 세미나 수가 1,000개를 상회한다. 기술 개발 워크샵이 9개가 있고, Lab은 60개를 넘게 진행되었다. 전시장도 있었지만 다른 전시회에 비해서는 규모가 적은 편이었다.

NVIDIA가 등장하는 유일한 전시장

COPUTEX 같은 전시장에 가면 AI 관련 하드웨어가 많은 메이커들의 제품을 중심으로 공개되어 있지만 정작 NVIDIA 부스는 없다. NVIDIA의 신제품이 GTC에서 주로 공개된다는 특징이 있어서 많은 개발자들이 Vera Rubin, 서버, 하드웨어에 많은 관심을 가진 것도 다른 전시회와 차별점이다.

관심 내용은 AI 성장과 이에 따른 하드웨어, 새로운 모델, 산업 발전 방향 이해에 초점

관심 갖고 참석했던 세션은 AI 진화 방향, GPU의 미래, Open Model, Agentic AI이고, 전반적인 맥락은 NVIDIA GTC Pre-game에서 언급한 Accelerated computing beyond AI, Open Models, Agentic systems로 요약할 수 있다. 이것은 AI의 진화 방향이기도 하지만, NVIDIA가 제공하는 AI 모델의 솔루션으로 볼 수 있다.

그래서 언급할 세미나는 다음과 같다.

- 1) GPU Processing 시대
- 2) Platform기업 NVIDIA
- 3) Open Model의 현재 위치와 갈 방향
- 4) Agentic AI 시대

## The Era of GPU Data Processing

GPU의 기능은  
모델 학습을 뛰어 넘어  
데이터 인프라로 확장 중

GPU가 이제 AI 학습용만이 아니라, 데이터 처리 전반의 기본 엔진이 되고 있다는 것을 의미한다. 예전 CPU의 역할까지 하고 있다는 의미인데 SQL(Structured Query Language) / DataFrame 분석 → 벡터·시맨틱 검색 → ML(Machine Learning) → 의사결정 최적화까지 하나의 흐름으로 GPU 위에서 처리하는 시대가 되었다는 의미이다.

NVIDIA가 성장 논리를 “GPU = 모델 학습”에서 “GPU = 데이터 인프라 전체”로 넓히고 있는 것이다.

표 3. Data Processing 단계별 사례

단계	역할	예시
SQL	데이터를 조회/집계	지역별 매출 분석
Semantic Search	의미 기반으로 정보 탐색	비슷한 기술 문서 찾기
Vector Search	semantic search의 구현 방식	임베딩 유사도 검색
Decision Optimization	최적 행동 계산	물류 경로, 생산 계획 최적화

자료: NVIDIA, IBK투자증권

세션은 5가지로 요약할 수 있다.

### 1) 데이터 처리의 중심이 CPU에서 GPU로 이동

데이터 처리 전반에서  
GPU 역할이 확대

GPU Processing 범위를 단순 SQL 가속이 아닌, SQL / DataFrames, Vector search, ML, Decision Optimization까지 확대한다는 의미이다. 즉 데이터 파이프라인 전체를 GPU native로 재구성하자는 것이고 GPU가 과거 CPU의 역할을 하게 되었다는 것을 얘기하고 있다.

참고로 ML은 과거에 CPU가 담당했던 영역이다.

### 2) 왜 SQL이 다시 중요해졌나

데이터의 출발점은 SQL

이 세션의 부제는 From SQL to Search and Back Again이다. 이것은 한때 AI / 검색 쪽으로 무게가 이동했지만 결국 기업 데이터의 출발점은 여전히 정형 데이터와 SQL이라는 것이어서 NVIDIA는 GPU native 엔진으로 Massive Lakehouse-scale Dataset에서도 인터랙티브 분석이 가능하다고 설명한다.

### 3) 검색은 이제 키워드 검색이 아니라 시맨틱 / 벡터 검색

GPU의 기능은 모델 학습을 뛰어 넘어 데이터 인프라로 확장 중

검색이 단순 텍스트 매칭이 아니라, 임베딩 기반의 의미 검색으로 바뀌었고, 이걸 대규모로 실시간 처리하려면 GPU가 중요하다는 의미이다.

### 4) ML과 최적화까지 한 플랫폼으로 묶으려는 의도

GPU의 기능은 모델 학습을 뛰어 넘어 데이터 인프라로 확장 중

분석과 검색만 말하는 게 아니라 cuOpt로 ML and Decision Optimization까지 포함해 GPU에서 문제를 빠르게 해결하고자 한다. 여기서 Decision Optimization은 물류, 공급망, 스케줄링, 금융 같은 문제를 푸는 영역이다. ML과 마찬가지로 과거에는 CPU에서 최적화 문제를 해결했었다. 하지만 변수가 많아지면 CPU의 연산 속도가 느려지는 단점이 있었다.

GPU는 “데이터를 보고 끝”이 아니라, 데이터 분석 → 예측 → 최적 의사결정까지 지원이 가능하다는 의미로 해석된다.

### 5) 진짜 메시지: AI 시대엔 데이터 엔진도 GPU native여야 한다

GPU의 기능은 모델 학습을 뛰어 넘어 데이터 인프라로 확장 중

결국 이 세션의 본론은, 생성형 AI와 Agentic AI가 커질수록 데이터 준비, 검색, 피처 처리, 검색 증강, 운영 최적화가 전부 병목이 되기 때문에, 모델 추론만 가속해서는 부족하다는 것이다. 그래서 NVIDIA는 GPU-native data platforms라는 표현을 쓰면서, 2026년 이후 데이터 플랫폼 아키텍처도 GPU 중심으로 가는 것이 바람직하다는 주장이다.

## Inside the NVIDIA Inference Platform and Ecosystem

이번 세션은 발표자가 Ian Buck인데, CUDA의 아버지로 불린다. 모델을 만든 연구자가 모델에서 Platform으로 시대가 바뀌고 있다고 주제발표를 한 세션이다.

Platform기업으로서  
NVIDIA  
이를 가능하게 하는  
기술에 대한 설명

내용은 “좋은 모델 하나”보다 “그 모델과 에이전트를 대규모로 실제 서비스하는 추론 플랫폼 전체”가 더 중요해졌다는 점을 강조하는 세션으로 요약할 수 있다. 궁극적으로는 NVIDIA도 이를 향한 행보가 진행 중이라는 얘기이기도 하다. Jensen Huang의 Keynote에서 큰 그림으로 발표한 내용을 보완해서

1. 그걸 가능하게 하는 기술이 뭔지
2. 실제로 어떤 변화가 생기는지
3. 어떤 산업이 기회를 얻는지에 관한 내용이다.

가능하게 하는 기술은 GPU, HBM, NVLink, Networking, Storage, Dynamo, NIM, Open Model stack을 말한다. 이 기술을 통해서 추론 속도 개선, 비용 절감, latency 감소, throughput 증가, 더 많은 AI Agent 운영 가능, 데이터센터 구조 변화 효과를 얻는다. 기회를 얻게 되는 산업은 클라우드, 제조업, 금융, 통신, 자율주행, 로봇틱스, 반도체 공급망 등이다.

세션 내용은

### 1) Inference가 새로운 중심축

GTC 핵심 키워드  
Inference

GTC 2026에서 NVIDIA는 회사의 초점을 **학습보다 추론(inference)** 쪽으로 강하게 옮겨 놓았고, 이 세션도 **추론 플랫폼과 생태계**를 해설하는 자리였다. 학습도 Pre-training 보다는 Post-training을 훨씬 더 강조하는 모습이 여러 세션에서 등장했다. 참고로 Jensen Huang이 Keynote에서 “inference”를 36번 언급했다고 한다.

### 2) 플랫폼 관점: 칩 하나가 아니라 시스템

GPU 보다는 추론시스템이  
중요

이 세션 제목이 “Chip”이 아니라 **Platform and Ecosystem**인 이유가 핵심이다. 포인트는 “GPU 성능” 하나보다 **GPU + CPU + Networking + Storage + Software**가 합쳐진 추론 시스템이 중요하다는 것이다.

### 3) 소프트웨어 축: Dynamo

Dynamo가 추론 운영  
시스템. 이를 중심으로  
Platform 기반 마련

이 플랫폼의 중심 소프트웨어가 **Dynamo**이다. NVIDIA는 GTC 2026에서 **Dynamo 1.0**을 **AI Factories용 Inference Operating System**으로 발표했고, 특히 **Generative AI와 Agentic Inference at Scale**을 위한 오픈소스 기반 운영층으로 설명하고 있다. 이를 바탕으로 NVIDIA는 Platform 기업으로 도약을 모색한다.

#### 4) Agentic AI가 추론 플랫폼을 바꾼다

Agentic AI는 기존의 플랫폼을 새롭게 구성. 요구하는 데이터가 급증하기 때문

이번 GTC 전체에서 핵심 키워드는 **Agentic AI**이다. Keynote에서 Agentic AI가 45번 언급됐고 Open Model, NemoClaw, Agent 프레임워크를 GTC 행사 모든 부문에서 강조하고 있다. Agent는 훨씬 더 많은 추론, 더 긴 컨텍스트, 더 많은 데이터 접근을 요구하기 때문에 Reasoning 단계보다 더 중요한 추론의 핵심 성장 동력이다.

#### 5) 그 결과 필요한 하드웨어 / 시스템 변화

LPU와 KV Cache 강화를 위한 하드웨어 도입

기존 시스템을 바꾸어야 하는 이유인데 이번 GTC에서 NVIDIA는

1. Groq3 LPU / LPX로 저지연 decode 가속을 보강했고
2. BlueField-4 STX로 Agentic AI에서 병목이 되는 Storage-KV Cache-데이터 이동 문제를 해결했다.

세션 흐름은 “추론이 중요해졌다 → Agent가 그 수요를 폭증시킨다 → 그래서 Chip만이 아니라 Decode, Memory, Storage, Networking까지 바뀌어야 한다”로 요약된다.

#### 6) Open Model / 생태계까지 포함한 Inference Ecosystem

앞서 언급한 모든 것들을 아우른 생태 시스템

마지막으로 “Ecosystem”이라는 단어의 의미는 NVIDIA가 이번 GTC에서 **Open Model, Open Agent, Open Source 추론 스택**을 강조한 배경과 일맥 상통한다. NVIDIA가 단순 하드웨어 업체가 아니라 **종합 AI 플랫폼 제공자**로 자신을 포지셔닝하고 있다는 의미로 해석할 수 있다.

## Open Models: Where we are?

Open Model이  
GTC 2026의 핵심 키워드

GTC 2026에서 Open Model은 주요 행사장에 빠짐없이 등장하는 키워드이다. Open Model은 보통 AI 모델의 가중치나 사용 방식이 비교적 개방된 모델을 의미한다. 즉 API(Application Programming Interface)로만 쓰는 완전 폐쇄형보다, 개발자가 직접 가져다 쓰고 배포하고 튜닝하기 쉬운 모델을 일컫는다.

Open Model은  
모델의 개방성을 의미

다음과 같이 이해하면 편하다.

- Closed model: 남이 운영하는 모델을 빌려 씀(ex. 날씨 앱, 지도 앱)
- Open Model: 내가 직접 가져와서 돌리거나 수정하기 쉬움(NVIDIA의 Nemotron, Mistral, Perplexity, Meta의 Llama 등)

Jensen Huang이 주관한  
세션의 무게

이번 세션에서는 다양한 Open Model 관련 기업들을 2번의 세션으로 나누어서 Jensen Huang이 토론을 주관했다. 그만큼 Open Model에 관심과 공을 들인 인상이다. 그렇다면 NVIDIA는 왜 Open Model을 강조하고, 중요하게 생각하고 있을까?

이유는

1) Agentic AI 확산을 위한 진입장벽을 낮추기 위해서

Agentic AI 확산의  
도화선으로 인식

GTC 2026 공식 자료를 보면 NVIDIA는 Nemotron Open Models를 전면에 두고, 이걸 Agentic AI systems 구축용이라고 직접 설명하고 있다. Agent Toolkit도 Open Models, Open Agents, Open Skills, Open Runtimes 조합으로 소개한다. 즉 폐쇄형 API만 쓰는 구조보다, 개발자가 직접 가져다 튜닝하고 연결할 수 있는 생태계를 키우려는 의도이다. 이를 통해서 보다 많은 Agentic AI Model이 만들어 질 수 있다고 보고 있기 때문이다.

2) 모델보다 플랫폼이 더 중요해졌기 때문

Platform이 보다 많은  
Token을 만든다

NVIDIA는 이제 “최고의 모델 회사”라기보다 AI Factory 플랫폼 회사가 되려고 한다. 그래서 Open Model 자체로 수익을 창출하는 것보다, Open Model이 많이 작동함으로써 GPU, 추론 SW, 클라우드, 네트워킹 사용량을 키우는 게 더 의미가 있어진 상황이다.

이를 보여주는 게 Dynamo 1.0인데, NVIDIA는 이걸 Open Source Foundation For Inference At Scale이라고 했고, vLLM·SGLang·LangChain 같은 오픈 생태계와의 native 통합을 강조하고 있다.

3) 개발자 Lock-in보다 “생태계 표준 자리잡기”를 노리기 때문

생태계를 장악함으로써  
CUDA 효과 노림수

Open Model은 개발자 입장에서 채택 장벽이 낮다. 그래서 모델을 받아서 로컬 / On-premises / 클라우드 어디서든 구동할 수 있고, 자기 데이터로 특화시키기도 쉬운 구조이다. NVIDIA는 이런 흐름을 끌어오기 위해 Nemotron Coalition까지 만들어서 여러 AI 연구기관과 함께 차세대 Open Frontier Model을 만들 계획이다. 즉 “우리만의 닫힌 모델”보다, 업계가 널리 쓰는 기본 레이어를 잡으려는 의도가 강해 보인다.

4) Physical AI·로보틱스는 오픈 접근이 특히 중요해서

Physical AI와 유사한 특성

GTC 2026 메인 페이지 자체가 Open Models, Libraries, Simulation Frameworks가 차세대 Factory·Robot·Autonomous Vehicle 개발을 돕는다고 강조하고 있다. 로보틱스 / 자율주행 / 시뮬레이션은 특정 API 하나 호출해서 끝나는 구조가 아니라, 데이터 생성·튜닝·시뮬레이션·배포까지 길게 이어지기 때문에 개방형 모델과 툴체인이 훨씬 잘 맞는다.

5) 오픈이 오히려 추론 수요를 키우기 때문

결과적으로는 추론 수요 강화. 이는 NVIDIA 생태계 확장에 유리

Open Model 확산은 NVIDIA에 불리한 게 아니라 오히려 유리하게 작용할 것으로 기대한다. 모델이 많이 공개되고 채택될수록 더 많은 기업이 자체 서비스, 에이전트, 산업용 AI를 만들고, 그 결과 추론 토큰량이 늘어날 것으로 예상하기 때문이다.

NVIDIA는 바로 이 추론 확장을 위해 Dynamo 같은 오픈소스 Inference Stack을 강조하고, Rubin GPU에서 성능 향상과 토큰 비용 절감을 강조했다. 이는 “모델 라이선스 수익”보다 인프라 사용량 확대를 노린 전략으로 해석된다.

표 4. Open Model 확산 효과

단계	무슨 일이 벌어지나	수혜 / 영향
Open Model 확산	더 많은 기업이 자체 AI 구축	개발자 / 기업 채택 증가
자체 구축 증가	추론 서비스 수 증가	GPU 사용량 증가
Agentic AI 결합	호출 수, 토큰 수, 반복 추론 증가	inference CAPEX 확대
대형 모델 운영	메모리 대역폭 병목 심화	HBM 수요 증가
고성능 AI 서버 출하	GPU + HBM 통합 패키징 확대	CoWoS 수요 증가

자료: IBK투자증권

## Agentic AI 101

이번 세션은 말 그대로 Agentic AI 개론이다. 이 세션은 Agentic AI가 무엇인지, 지난 1년간 어떻게 바뀌었는지, Frontier Models와 Open Models가 어떻게 새로운 종류의 Agents를 가능하게 하는지, 그리고 오늘 당장 어떻게 시작할 수 있는지에 대해서 언급하고 있다. 5가지 항목인데

### 1) Agentic AI란 무엇인가

Agentic AI는 능동적으로 수행하는 AI로 정의

이 세션의 출발점은 AI를 “수동 응답기”에서 “능동 수행자”로 인식하는 것이다. 즉 그냥 질문에 답하는 모델이 아니라, 목표를 받고 계획하고, 필요한 도구를 써서, 여러 단계를 거쳐 작업을 수행하는 시스템을 Agentic AI라고 정의한다.

### 2) 지난 1년간 뭐가 달라졌나

Chatbot에서 Agent로 1년새 진화

2025~2026년 사이의 변화 포인트는 AI가 단순 ChatBot에서 실행 가능한 Agent로 진화하고 있다는 점이다. 이 변화는 GTC 2026 전반의 큰 테마와도 맞물려 있었고, NVIDIA도 2026년 들어 Agentic AI Inflection Point(변곡점)라는 표현을 공식적으로 쓰고 있다.

### 3) Frontier Model과 Open Model이 왜 중요하나

구동하는 모델은 Frontier Model, Open Model 모두 대안

이 세션은 Frontier Models와 Open Models를 둘 다 언급하고 있다. 뜻은, 최고 성능의 대형 모델과 개방형 모델 모두가 Agent 구축에 쓰일 수 있고, Agent 생태계는 둘 중 하나만으로 돌아가는 게 아니라는 것을 의미한다. 즉 Closed API만의 세계도 아니고, Open Model만의 세계도 아니라, 둘 다 Agent의 재료가 되는 구조라는 설명이다.

### 4) Agent를 이루는 핵심 구성요소

모델 하나로는 의미가 없고 도구, 상태, 제어 로직 모두가 포함되어야

Agentic AI의 구성요소는 모델 하나가 아니라, 모델 + 도구 + 상태 + 제어 로직 모두 포함해서 Agent가 된다. 모델은 판단·생성 담당, 도구(tool use)는 검색, DB, API, 계산기 등의 역할을 하고, 메모리 / 상태(state)는 중간 결과를 유지한다. 계획(planning)은 단계를 분해하고, 실행 루프를 통해서 결과를 보고 다음 행동을 수정하는 단계이다.

### 5) 어디에 쓰이나

업무형 AI

Agent는 단순 질의응답이 아니라 실제 workflow 자동화에 쓰이는 개념이고, “설명형 AI”보다 업무형 AI에 가깝다는 점이 핵심적인 설명이다.

## 어느 실리콘밸리 벤처투자자의 이야기

우연치 않게 출장 일정이 길어지면서 성사가 된 실리콘밸리 내 벤처투자자와 미팅이 있었다. 실리콘밸리 투자자들이 보는 AI에 대한 시각, GPU 수요, 실제 성사되고 있는 투자의 실체, 퀀텀 컴퓨팅과 Physical AI 등에 관한 의견을 나누었다.

### 1) AI는 어떤 의미로 해석되는가

산업혁명 수준.  
AI가 이제는 생각을  
대체하는 시대

AI 관련 최대의 고민거리이거나 생각할 문제인데 대답은 생각을 대체한다는 것이다. 지금까지 생각할 수 있는 어떤 것도 세상에 등장한 적이 없었고, 가장 비근하게 세상에 크게 영향을 미친 것이 인터넷인데 이와는 결이 다르다고 판단하고 있다.

일반론이지만 인터넷은 새로운 산업을 만들기는 했으나, 전 산업에 미치는 영향이 제한적인 것으로 판단한다. 그에 비해 AI는 전 산업으로 파고 들어 생산성을 높이는 역할을 할 것으로 보고 있다.

### 2) GPU 누가 쓰고 있는가

Inference Training 전문  
회사의 등장은 GPU의  
새로운 수요처

Hyperscaler, Enterprise, Sovereign 등으로 알려져 있지만 새로운 플랫폼 기업들의 수요가 생각보다는 크게 자리매김 하고 있는 것으로 들린다. 대표적인 회사가 Together AI인데 여러 모델들이 필요한 추론 작업을 대행해준다. 매출이 매년 10배씩 증가하고 있다. 이들이 가장 필요한 부품이 NVIDIA의 GPU이다.

### 3) 투자 대상으로 어떤 회사들이 있는가

실체가 있기도 없기도  
하지만 모금 실패 사례는  
아직 없다

다소 내부적인 문제일 수도 있지만 실리콘밸리 투자 사회 전반적인 상황을 대상으로 했다. Together AI처럼 확실한 비즈니스 모델, 실적이 뒷받침되는 회사도 있고, 회사는 있지만 아직 매출이 발생하지 않은 회사들도 수억 달러의 투자자금을 유치하고 있다고 한다.

AI에 대한 기대감이 아직은 투자업계에서 유효한 상황으로 볼 수 있다. 다만, 아직 실체도 실적도 없고 모델 아이디어만 가지고 투자를 받을 수 있다는 점은 AI 열풍에 거품이 낀 것은 아닌가 하는 우려도 할 수 있는 상황으로 해석된다.

4) 퀀텀 컴퓨팅에 대한 시각

퀀텀 관련 투자는 없다

GTC 2026에서도 언급된 테마이고 한 세션을 듣기도 해서 무척 궁금한 부분이었는데 대답은 다소 냉정했다. 퀀텀 컴퓨팅 관련 펀딩 사례도 많지 않고 아직은 투자자들이 크게 관심을 가지는 섹터는 아니라고 한다.

GTC 세션에서 들은 내용도 아직은 검증하는 단계라고 결론을 내리고 있었다.

5) Physical AI에 대한 시각

확산될 여지가 크다.  
Humanoid 중심

국내 여러 그룹들이 활발히 활동하는 것에 대해서는 긍정적인 시각이었다. 현대그룹이 가장 빠르게 움직인 것에 비해 삼성그룹의 대처가 늦은 점은 아쉬운 대목이라 평가한다.

Humanoid에 대해서는 대량 생산을 위한 선택으로 보기도 하고, 굳이 Humanoid이어야 하는 지에 대해서는 논쟁 중이었다.

6) Neocloud에 대한 시각

Neocloud는 Niche market

AI 투자에서 한 축을 담당하고 있지만 궁극적으로는 Hyperscaler들에 의해 잠식될 것이라는 의견과 그럼에도 높은 가격 경쟁력으로 생존할 것이라는 의견이 아직은 맞서고 있다는 인상이다.

## 전시장 풍경 - NVIDIA H/W 신제품, Liquid Cooling, Memory

### NVIDIA Vera Rubin System

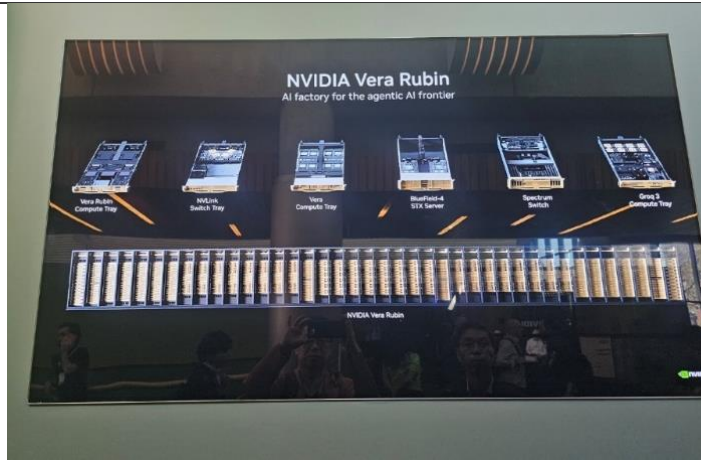
GTC 주인공은 NVIDIA

CES, MWC, COMPUTEX 모두가 AI 관련 전시하는 행사지만 NVIDIA 부스가 있던 곳은 없었고 COMPUTEX에만 Ecosystem을 전시한 정도였다. GTC는 NVIDIA가 주최하는 행사여서 가장 큰 부스는 NVIDIA였다. CES 2025는 LG전자, MWC는 Huawei, COMPUTEX는 Gigabyte가 주인공들이었다.

Vera Rubin 시스템은 아래와 같이 구성된다.

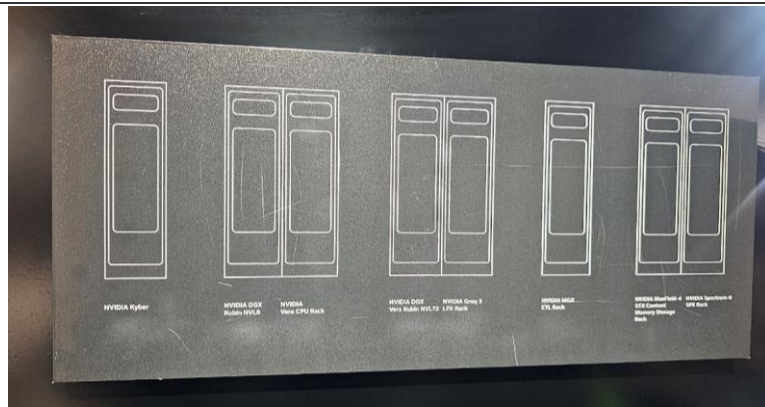
1. Vera Rubin Compute Tray: Vera CPU 2개와 Rubin GPU 4개(NVL72에서 18개)
2. NVLink Switch Tray: Rack 내부 여러 Compute Tray에 GPU들을 하나의 초고속 NVLink fabric으로 묶는 스위치 전용 모듈(NVL72에서 9개)
3. Vera Compute Tray: GPU 연산이 아니라, AI Factory 인프라, Orchestration, Data Processing, HPC, CPU-dense services 지원. Unit 당 최대 4개.
4. BlueField-4 STX Server: STX는 긴 컨텍스트 추론이나 Agentic AI에서 생기는 KV Cache, Context Memory, 고속 스토리지 계층을 처리하기 위한 저장 인프라 설계이다. BlueField-4 STX Rack / Server는 이 STX 아키텍처를 실제로 구현한 Context Memory Storage 인프라 노드다. 즉, 대형 추론 시스템에서 GPU 메모리만으로 감당하기 어려운 긴 컨텍스트를 위해 별도의 고대역폭 저장 계층을 만드는 장비다.
5. Spectrum Switch: Rack과 Rack 사이를 연결하는 Scale-out 네트워크 스위치이다. Spectrum-X는 단순 스위치 1개 이름이라기보다, NVIDIA의 AI용 Ethernet 플랫폼을 말한다. Spectrum-6 SPX Networking Rack이 전체 POD를 하나의 슈퍼컴퓨터처럼 연결하는 역할을 한다. 이 Rack은 Spectrum-X Ethernet 또는 Quantum-X800 InfiniBand 스위치로 구성될 수 있고, Rack-to-Rack 저지연·고대역폭 연결을 담당한다.
6. Groq3 Compute Tray: LPX Rack 안에 들어가는 1U 추론 전용 Tray 1개이며, LPU 8개 + Host Processor + Fabric Logic으로 구성된 모듈이다. LPX에는 32개 Tray가 내장된다. NVIDIA Groq3 LPX = Vera Rubin 옆에 붙는, 실시간 응답용 Inference 전용 Rack이다

그림 9. Vera Rubin NVL72 System



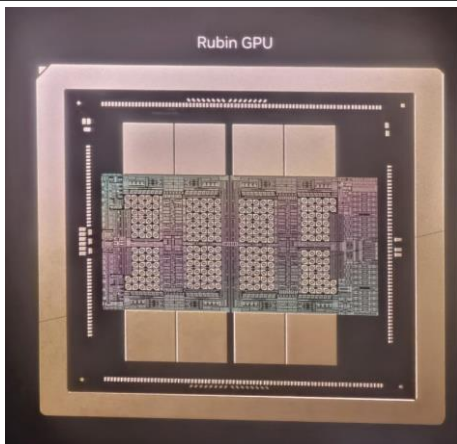
자료: NVIDIA, IBK투자증권

그림 10. Vera Rubin System 구성 Rack



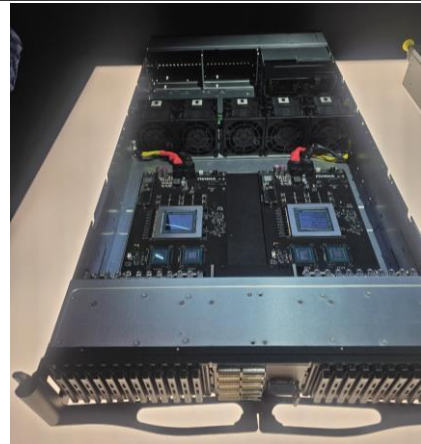
자료: NVIDIA, IBK투자증권

그림 11. Rubin GPU



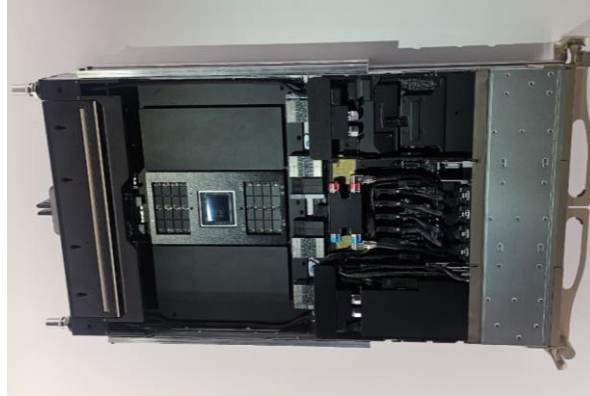
자료: NVIDIA, IBK투자증권

그림 12. NVIDIA Context Memory Storage



자료: NVIDIA, IBK투자증권

그림 13. Vera CPU Tray



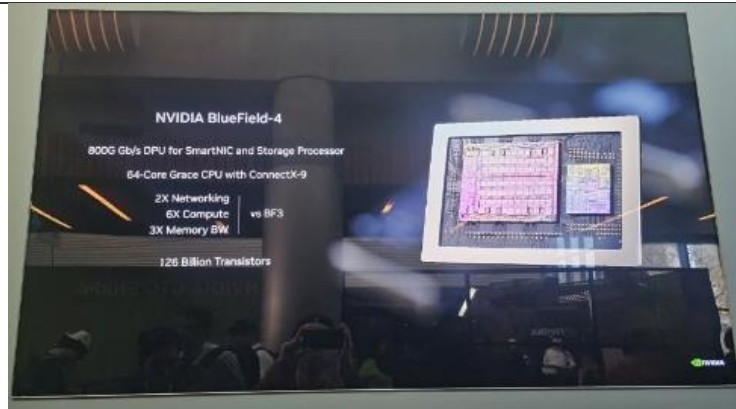
자료: NVIDIA, IBK투자증권

그림 14. Rubin GPU Tray



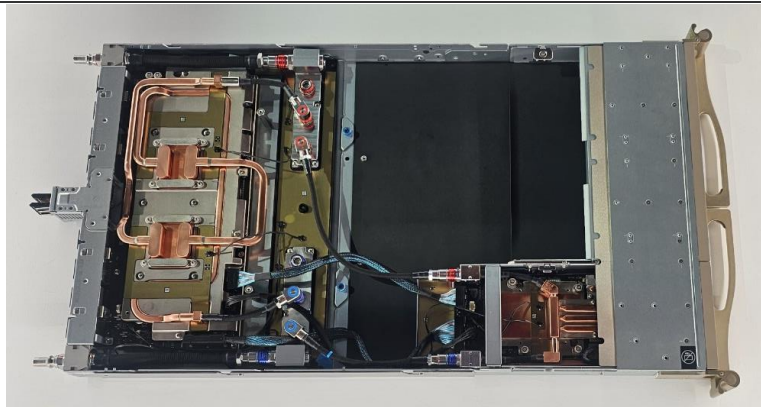
자료: NVIDIA, IBK투자증권

그림 15. NVIDIA BlueField-4



자료: NVIDIA, IBK투자증권

그림 16. Vera Rubin Switch Tray

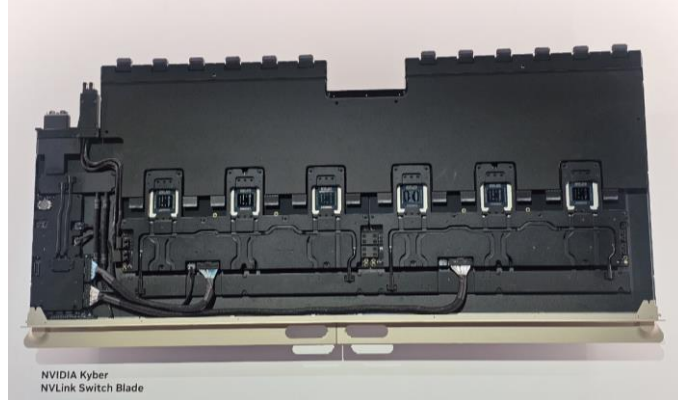


자료: NVIDIA, IBK투자증권

Kyber는 NVIDIA의 차세대 Rack 스케일(scale-up) 시스템용 내부 인터커넥트 / Rack 아키텍처다. GPU Blade와 NVLink Switch Blade를 더 높은 밀도로 묶고, Rack 안에서 더 큰 NVLink 도메인과 더 효율적인 전력·배치 구조를 구현하려는 플랫폼이다.

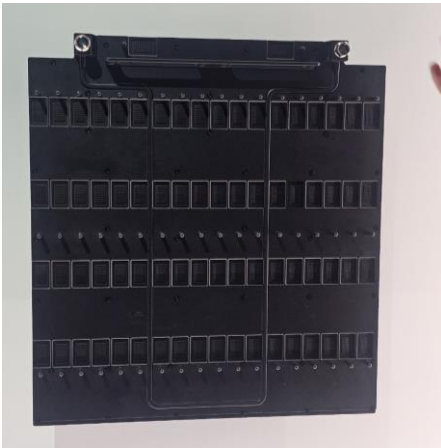
Kyber를 통해 Blade를 책처럼 세워 넣는 구조를 사용해 새시당 최대 18개 Compute Blade를 담고, 뒤쪽의 Cabless Midplane으로 NVLink Swtich Blade를 연결한다.

그림 17. NVIDIA Kyber NVLink Switch Module



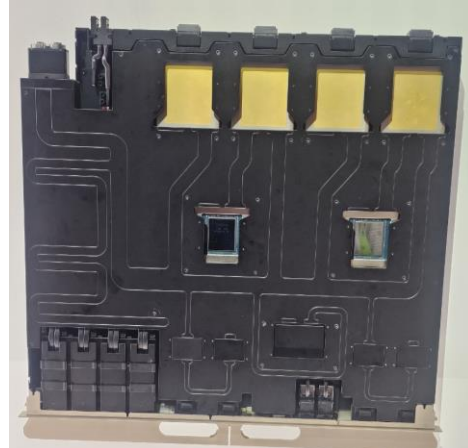
자료: NVIDIA, IBK투자증권

그림 18. NVIDIA Kyber Midplane



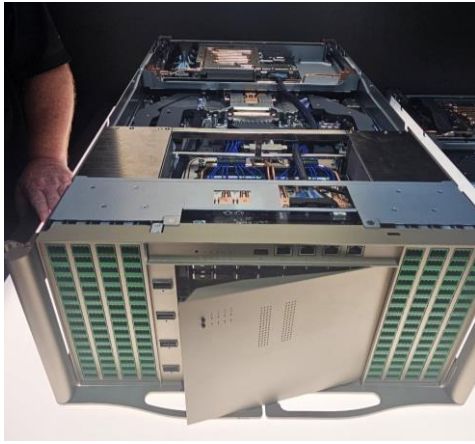
자료: NVIDIA, IBK투자증권

그림 19. NVIDIA Kyber Compute Blade



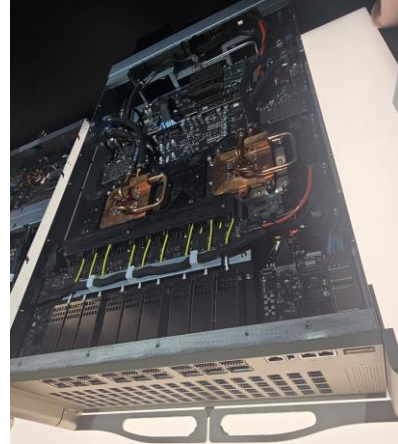
자료: NVIDIA, IBK투자증권

그림 20. NVIDIA Spectrum-X



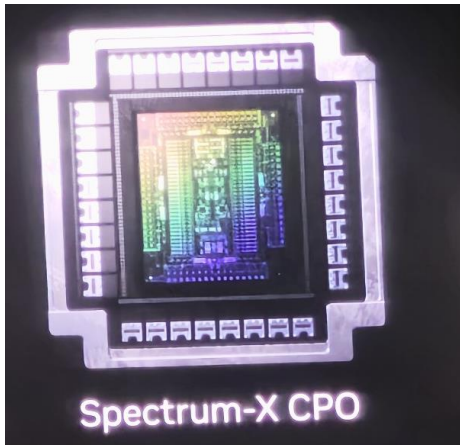
자료: NVIDIA, IBK투자증권

그림 21. NVIDIA Quantum-X



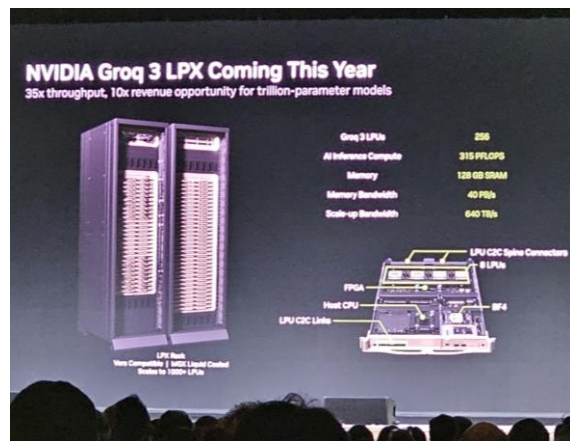
자료: NVIDIA, IBK투자증권

그림 22. Spectrum-X CPO



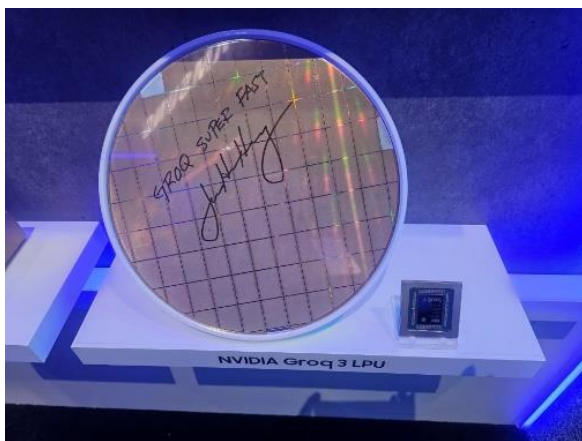
자료: NVIDIA, IBK투자증권

그림 23. NVIDIA Groq3 LPX



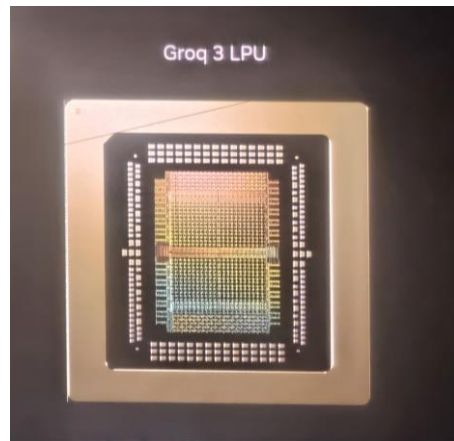
자료: NVIDIA, IBK투자증권

그림 24. 삼성전자의 Groq3 Wafer



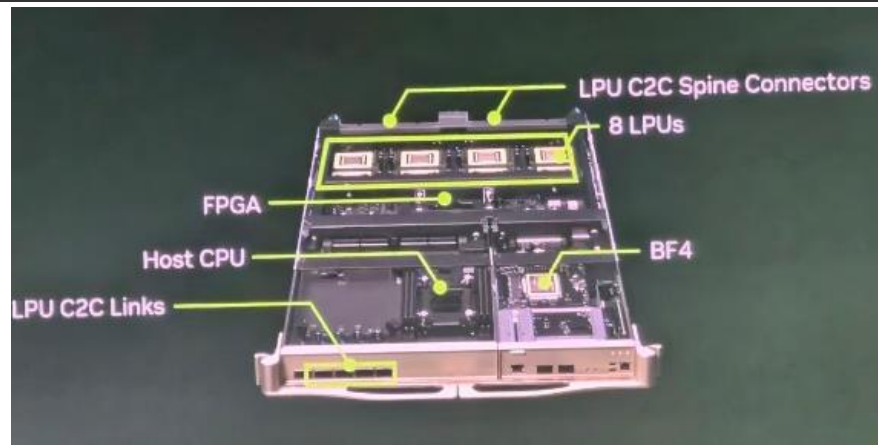
자료: 삼성전자, IBK투자증권

그림 25. 삼성전자의 Groq3 Chip



자료: 삼성전자, IBK투자증권

그림 26. NAND LPU Tray



자료: NVIDIA, IBK투자증권

그림 27. NVIDIA ConnectX-9



자료 NVIDIA, IBK투자증권

그림 28. NVIDIA AI Factory MGX Ecosystem



자료: NVIDIA, IBK투자증권

그림 29. Rack ODM



자료: NVIDIA, IBK투자증권

## Liquid Cooling System

그림 30. NAND Switch Board Liquid Cooling



자료: NVIDIA, IBK투자증권

그림 31. Vertiv Cooling System



자료: Vertiv, IBK투자증권

그림 32. Vertiv Cooling



자료: Vertiv, IBK투자증권

그림 33. ASUS Liquid Cooling 구조



자료: ASUS, IBK투자증권

그림 34. HPE Liquid Cooling



자료: HPE, IBK투자증권

## Memory: HBM, SOCAMM2, SSD

HBM4는 마무리 국면,  
HBM4E가 새로운 격전지

GTC 메모리 핵심은 HBM4이다. 메모리 3사가 모두 전시했으나 차별화 포인트는 없어 보였다. 공급 시점, 공급 물량과 관련된 노이즈는 있지만 대충 마무리 되어 가는 그림이다. 3사 모두 계획 수준의 물량은 공급할 것으로 보인다.

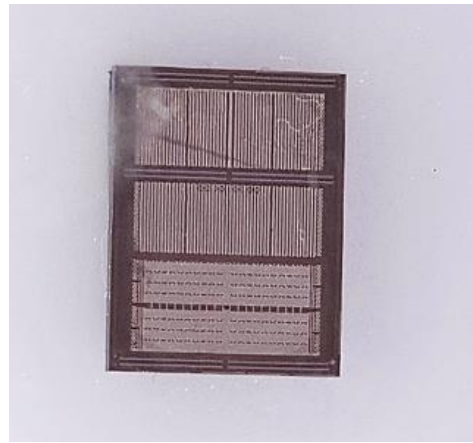
삼성전자는 HBM4 Base Die Wafer와 HBM4E Core Die Wafer를 공개한 점이 차별화 되었고, 메모리는 아니지만 Groq3 웨이퍼도 전시했다.

그림 35. 삼성전자 HBM4 Base Die Wafer



자료: 삼성전자, IBK투자증권

그림 36. 삼성전자 HBM4 Chip



자료: 삼성전자, IBK투자증권

그림 37. 삼성전자 HBM4E Core Die Wafer



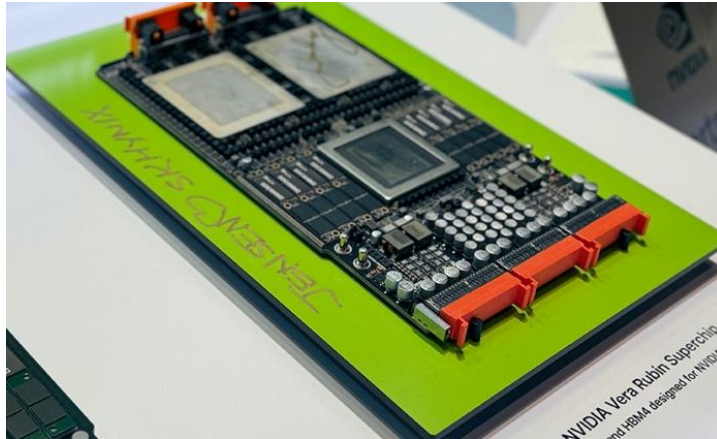
자료: 삼성전자, IBK투자증권

SOCAMM2도 매력적  
시장. 다만 Vera CPU  
물량에 유동적

SOCAMM2도 메모리 3사가 모두 전시하였는데 Vera CPU를 지원하는 메모리이다. 메모리는 LPDDR5를 사용하고, Vera CPU를 지원하기 위해서 특별히 패키징을 하는데 이를 SOCAMM2라고 부른다. 이는 Vera CPU 전용이어서 다른 CPU에는 사용하지 않는다.

Vera CPU에 최대 2TB SOCAMM이 필요하다. Vera Rubin NVL72에는 72TB가 필요한 셈이다.

그림 38. NVIDIA Vera Rubin Board



자료: NVIDIA, IBK투자증권

그림 39. SK하이닉스 SOCAMM2

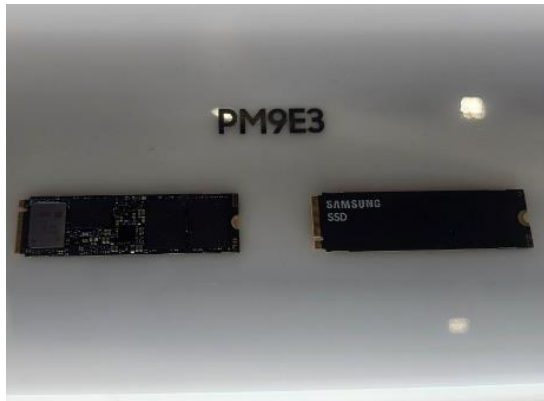


자료: SK하이닉스, IBK투자증권

NAND 주력은 제각각

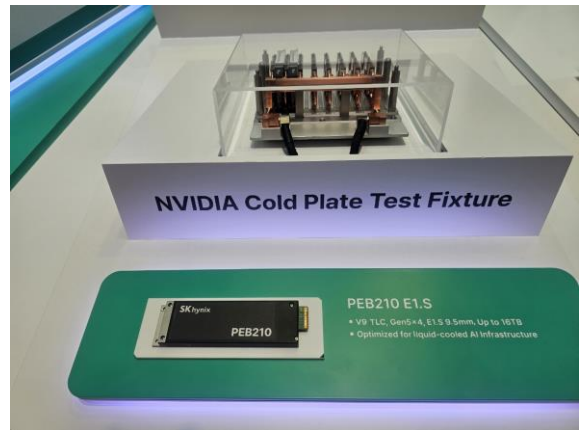
NAND도 다양한 제품으로 경쟁하고 있다. 삼성전자는 DGX로 공급하고, SK하이닉스는 Liquid Cooling AI 관련 제품을 공개했다. Micron은 BlueField 최적화 제품을 공개했다.

그림 40. 삼성전자의 DGX용 SSD



자료: 삼성전자, IBK투자증권

그림 41. SK하이닉스 Liquid cooled AI 최적화 SSD



자료: SK하이닉스, IBK투자증권

그림 42. Micron의 BlueField용 SSD



자료: Micron, IBK투자증권

### Compliance Notice

동 자료에 게재된 내용들은 외부의 압력이나 부당한 간섭없이 본인의 의견을 정확하게 반영하여 작성되었음을 확인합니다.  
동 자료는 기관투자자 또는 제3자에게 사전 제공한 사실이 없습니다.  
동 자료는 조사분석자료 작성에 참여한 외부인(계열회사 및 그 임직원등)이 없습니다.  
조사분석 담당자 및 배우자는 해당종목과 재산적 이해관계가 없습니다.  
동 자료에 언급된 종목의 지분을 1%이상 보유하고 있지 않습니다.  
당사는 상기 명시한 사항 외 고지해야 하는 특별한 이해관계가 없습니다.



## IBKS Research Center

성명	직급	담당업종	전화	이메일
용대인	전무(부문장)	총괄	6915-5400	daeinyong@ibks.com
이승훈	상무대우(본부장)	SI/인터넷/게임	6915-5680	dozed@ibks.com

### 투자분석부

변준호	연구위원	Strategy	6915-5670	ymaezono@ibks.com
정용택	수석 Economist	Economy	6915-5701	ytjeong0815@ibks.com
김인식	연구위원	자산배분/ETF	6915-5472	kds4539@ibks.com
정형주	연구위원	채권/크레딧	6915-5654	hj.jeong@ibks.com

### 기간산업분석부

이동욱	연구위원	에너지/소재	6915-5671	treestump@ibks.com
남성현	연구위원	유통·식자재/지주	6915-5672	rockrole@ibks.com
김유혁	연구위원	미디어/엔터/레저	6915-5673	yuhyuk.kim@ibks.com
이현욱	연구원	자동차/2차전지	6915-5659	hwle1125@ibks.com
오지훈	연구원	조선/기계	6915-5662	jihoonoh@ibks.com

### 혁신기업분석부

김운호	연구위원	IT/반도체	6915-5656	unokim88@ibks.com
김태현	연구위원	음식료/유틸리티/통신	6915-5658	kith0923@ibks.com
조경진	연구위원	해외주식	6815-5464	ckjins@ibks.com
조정현	연구원	건설/부동산	6915-5660	controlh@ibks.com

### 코스닥리서치센터

이건재	연구위원	소재·부품·장비/스몰캡	6915-5676	geonjaelee83@ibks.com
정이수	연구위원	제약/바이오	6915-5677	ysjeong306@ibks.com
강민구	연구원	IT/디스플레이/미드·스몰캡	6915-5473	kmg@ibks.com

## “국민과 중소기업에 필요한 참 좋은 IBK투자증권”



IBK기업은행 금융그룹

서울특별시 영등포구 여의도동 국제금융로 6길 11  
대표번호 02-6915-5000  
고객지원부 1588-0030, 1544-0050

IBKS Family Office	02) 536-4070	IBK WM센터 대구	053) 752-3535
영업부	02) 6915-2626	IBK WM센터 광주	062) 382-6611
강남센터	02) 2051-5858	IBK WM센터 일산	031) 904-3450
강남역 금융센터	02) 532-0210	IBK WM센터 판교	031) 724-2630
분당센터	031) 705-3600	IBK WM센터 평촌	031) 476-1020
IBK WM센터 강남센트럴	02) 556-4999	IBK WM센터 천안	041) 569-8130
IBK WM센터 목동	02) 2062-3002	IBK WM센터 부산	051) 741-8810
IBK WM센터 도곡	02) 2057-9300	IBK WM센터 창원	055) 282-1650
IBK WM센터 한남동	02) 796-8500	IBK WM센터 울산	052) 271-3050
IBK WM센터 중계동	02) 948-0270	IBK WM센터 시화공단	031) 498-7900
IBK WM센터 반포자이	02) 3481-6900	IBK WM센터 남동산단	032) 822-6200
IBK WM센터 동부이촌동	02) 798-1030		