

반도체(Neutral)

DeepSeek: AI에는 호재, Nvidia에는 악재

Issue Comment

[반도체] 송명섭 2122-9207 mssong@imfnsec.com [IT RA] 손우성 2122-9785 useong@imfnsec.com

DeepSeek가 불러온 충격

DeepSeek의 등장이 AI 산업 전반에 걸쳐 엄청난 충격을 불러오고 있다. 이는 대규모 최신 GPU 구매만이 AI 개발을 위한 유일한 방법이라는 그간의 개발 방식에 변화를 가져올 수 있기 때문이다.

군비 경쟁에 비유되는 현재 AI 산업에서의 경쟁 구조는 Nvidia의 초고가, 최신 GPU를 대량 구매해 전력을 더 소모하더라도 더 강력한 컴퓨팅 파워를 갖추는 방향으로 진행되어 왔다. 그러나 DeepSeek는 하드웨어 강화가 아닌 알고리즘, 최적화 기술, 오픈 소스 등 소프트웨어 기술로 여타 최신 LLM (GPT-4o, GPT-o1 등) 서비스와 동등하거나 그 이상의 성능을 내는 것으로 나타났다.

벤치마크 테스트 결과에 따르면 DeepSeek는 coding, mathematical reasoning, task-specific optimization 등 보다 전문적이고 특화된 분야에서 여타 최신 LLM 서비스보다 앞선 성능을 보이는 것으로 드러났다. 개발자에 따르면 비용 측면에서도 GPT-4o의 1백만 토큰 (LLM 모델이 텍스트를 이해하고 생성하는 기본 단위)당 Input Cost는 \$2.5, Output Cost는 \$10.0인 반면, DeepSeek-V3는 Input Cost가 \$0.14, Output Cost가 \$0.28에 불과해 DeepSeek-V3가 GPT-4o 대비 29.8배 저렴하다.

반도체 채용 측면에서 놀라운 점은 DeepSeek가 Nvidia의 A100 또는 H800 등 저성능 구형 GPU로 초고가, 고성능 AI GPU를 사용한 미국의 LLM 서비스들에 필적하는 성능을 구현했다는 점이다.

물론 그간 미국의 제재에도 불구하고 Nvidia의 H100, H200 등 고성능 GPU 들이 중국으로 불법 반입되었다는 의심이 존재해왔고 또한 시장 일부에서는 DeepSeek R-1의 개발 비용이 Meta LLAMA 3.1의 5억 달러에 비해 557만 달러에 불과하다는 주장에 대해 의구심을 가지고 있는 것으로 보인다.

그리고 일부 보고서에 따르면 DeepSeek R-1은 개발자가 명시한 10K의 A100 대신 50K의 Hopper GPU를 사용했을 수 있다는 의심도 존재한다. 따라서 개발자 발표 내용의 진위를 파악하기 위해 시간이 좀더 필요한 것으로 판단된다. 물론 개발자의 발표 내용이 거짓으로 밝혀진다면 AI 개발 방향은 이전과 같은 하드웨어에 대한 대규모 투자 방식으로 되돌아 갈 것이다.

그러나 만약 DeepSeek의 성공 모델이 사실이라면 이제부터 AI 혁신은 얼마나 지출하는지에 의해서가 아니고, 얼마나 효율적으로 개발하느냐에 따라 승패가 결정될 가능성이 높아질 것이다. 물론 앞으로도 로보틱스, 자율 주행, AGI 등 고용량 데이터 처리를 위해 고성능 GPU는 여전히 필요할 것이나, 빅테크 업체들은 대규모 지출보다는 더 큰 비중으로 효율적인 투자와 소프트웨어 개발에 관심을 기울이게 될 가능성이 높다.

AI 산업 전반에는 호재. Nvidia에는 악재

저비용, 고효율 DeepSeek의 성공은 AI 산업 전반에 긍정적인 영향을 줄 것으로 판단된다. 이는 사명과 달리 극히 폐쇄적인 OpenAI와 다르게 DeepSeek는 시장 모두에 개방된 Open Source이기 때문이다. 그간 비용 문제로 AI 혁신에 배제되어온 수많은 주체들이 AI 개발에 적극적으로 참여하게 될 경우 장기적으로는 AI 시장 규모가 도리어 빠르게 확장될 수 있다.

단 반도체 산업에는 DeepSeek의 성공이 위기이자 기회가 될 것으로 보이며 특히 Nvidia에는 부정적인 영향을 미칠 가능성이 높은 것으로 판단된다. 현재까지 Nvidia의 고성장은 끊임없이 신규 고성능 GPU를 출시하고 이를 AI 개발 업체들이 경쟁적으로 대량 구매해왔음에 따른 것이다.

그러나 구형 저성능 GPU로 구현된 DeepSeek의 성공이 Nvidia의 성장 모델에 의문을 제기하는 상황이 되었다. 동사의 초고가 GPU 판매량과 70%에 달하는 이익률에 하락 압력이 발생할 가능성이 높아졌기 때문이다. 향후 고객들은 무조건적인 초고가 AI GPU 구매보다는 효율성과 비용을 감안하여 동사 GPU를 구매하게 될 것으로 보인다.

물론 AI 산업의 성장 속도가 빨라짐에 따라 중저가 반도체를 포함한 전반적인 AI 반도체 시장 규모가 커질 수 있으므로, 이는 반도체 업계 전체에 기회가 될 수도 있다. 단 이는 이미 고성능 AI 반도체 시장을 독점하며 성장해온 Nvidia에게는 긍정적인 방향이 아닌 것으로 판단된다.(계속)

일부 보도에 따르면 DeepSeek에는 메모리 반도체를 ChatGPT보다 75% 덜 사용했다고 한다. 이러한 추세가 일반화될 경우 GPU 내 DRAM 채용량 정체 또는 감소로 향후 HBM 판매에 부정적인 영향을 줄 가능성이 있다. DRAM 업체들에게는 HBM 대비 저용량, 저가인 GDDR 등 DRAM이 AI 용으로 얼마나 빨리 성장해서 HBM의 성장세 둔화를 상쇄할 것인지가 관건이 될 수 있다. Nvidia의 최신 GPU에 고용량, 고성능 HBM을 공급하며 동반 성장해온 DRAM 업체들에게도 시장의 성격이 바뀔을 의미하므로 당분간은 쉽지 않은 환경이 조성될 수 있다는 의미이다.

그리고 만약 미국 정부가 이번 사건을 계기로 Nvidia의 중국 전용 GPU (H800, H20 등)의 판매를 향후 모두 금지하게 된다면, Nvidia뿐 아니라 동 GPU에 HBM을 공급해온 한국 DRAM 업체에도 부정적인 영향이 발생할 전망이다.

표1. ChatGPT와 DeepSeek 간 비교

Feature	ChatGPT	DeepSeek
Parameters	1.8 trillion	671 billion
Training Cost	\$100 million	\$5.5 million
Training Efficiency	FP16 precision	FP8 precision
Open-Source	No	Yes
Input Cost	\$2.50/million tokens	\$0.14/million tokens
Output Cost	\$10.00/million tokens	\$0.28/million tokens
Best For	General-purpose tasks	Coding, math, specialized tasks

자료: iM증권 리서치본부

표2. Nvidia와 AMD의 AI GPU Roadmap과 채용 HBM

기업	AI GPU	2024E				2025E				2026E			
		1Q24	2Q24	3Q24	4Q24	1Q25	2Q25	3Q25	4Q25	1Q26	2Q26	3Q26	4Q26
Nvidia	H200	HBM3E 8H 144GB (24GB*6), TSMC 4N											
	B200	HBM3E 8H 192 (24GB*8), TSMC 4NP											
	GB200	HBM3E 8H 384 (2 x 24GB*8), TSMC 4NP											
	B300	HBM3E 12H 288GB (36GB*8), TSMC 4NP											
	B300A	HBM3E 12H 144GB (36GB*4), TSMC 4NP											
	Rubin	HBM4 8H/12H											
AMD	MI300X	HBM3 8H 192GB (24GB*8), TSMC 5/6nm											
	MI325X	HBM3E 8H 256 (32GB*8), TSMC 5/6nm											
	MI355X	HBM3E 12H 288GB (36GB*8), TSMC 3nm											
	MI400	TBD											

자료: Trendforce, iM증권 리서치본부

Compliance notice

당 보고서 공표일 기준으로 해당 기업과 관련하여,

- 회사는 해당 종목을 1%이상 보유하고 있지 않습니다.
- 금융투자분석사와 그 배우자는 해당 기업의 주식을 보유하고 있지 않습니다.
- 당 보고서는 기관투자자 및 제 3자에게 E-mail등을 통하여 사전에 배포된 사실이 없습니다.
- 회사는 6개월간 해당 기업의 유가증권 발행과 관련 주관사로 참여하지 않았습니다.
- 당 보고서에 게재된 내용들은 본인의 의견을 정확하게 반영하고 있으며, 외부의 부당한 압력이나 간섭 없이 작성되었음을 확인합니다.

본 분석자료는 투자자의 증권투자를 돕기 위한 참고자료이며, 따라서, 본 자료에 의한 투자자의 투자결과에 대해 어떠한 목적의 증빙자료로도 사용될 수 없으며, 어떠한 경우에도 작성자 및 당사의 허가 없이 전재, 복사 또는 대여될 수 없습니다. 무단전재 등으로 인한 분쟁발생시 법적 책임이 있음을 주지하시기 바랍니다.

[투자의견]

종목추천 투자등급

종목투자의견은 향후 12개월간 추천일 증가대비 해당종목의 예상 목표수익률을 의미함.

- Buy(매수): 추천일 증가대비 +15% 이상
- Hold(보유): 추천일 증가대비 -15% ~ 15% 내외 등락
- Sell(매도): 추천일 증가대비 -15% 이상

산업추천 투자등급

시가총액기준 산업별 시장비중대비 보유비중의 변화를 추천하는 것임

- Overweight(비중확대)
- Neutral(중립)
- Underweight(비중축소)

[투자등급 비율 : 2024-12-31 기준)

매수	중립(보유)	매도
92.4%	6.9%	0.7%