

반도체 (Overweight)

Rubin Ultra와 HBM4E의 변화?

스펙 변화의 가능성과 향후 주목할 요소

Rubin Ultra와 HBM4E의 스펙 변화 가능성

NVIDIA의 CY27 차세대 AI 가속기 Rubin Ultra는 당초 GPU 패키지당 HBM4E 1TB (1,024GB)의 용량을 목표로 발표되었으나, 두 가지 제조 병목 (패키징 레티클 한계, HBM 적층 수율)으로 인해 384GB 축소안이 거론되고 있는 것으로 보인다. 다만 이는 현재 NVIDIA와 메모리 업체 간 협의가 진행 중인 사안으로, 공식 확정된 스펙은 아닌 것으로 판단된다. 축소안의 정황 증거로는 GTC 2026에서 발표된 Rubin Ultra 프로토타입의 형태가 정사각형인 것이 2 레티클 구조임을 시사한다는 점, 듀얼 히트 스프레더 설계가 고볼륨 양산 단계에서 워피지 (휘어짐) 요건을 맞추지 못해 싱글 스프레더로 변경될 것이라는 보도가 있었다는 점 등이다.

스펙 변화의 내용은 GPU당 HBM 탑재량 = (컴퓨터 레티클 수) × (레티클당 HBM 큐브 4개) × (큐브당 DRAM 용량)의 구조에서, 원안은 $4 \times 4 \times 64GB = 1,024GB$ (1TB)인데 반해 축소안에서는 $2 \times 4 \times 48GB = 384GB$ 로 바뀐다는 것이다 (<표1> 참조). 즉 GPU를 구성하는 레티클의 개수가 원안의 4개에서 2개로, HBM 큐브 당 코어 다이 적층 수가 원안의 16단에서 12단으로 축소된다는 것이 핵심이다. 단 코어 DRAM 다이는 두 구성 모두 32Gb (4GB)칩으로 동일하다.

스펙 변화 가능성의 원인

기술적 측면에서 첫번째로 거론되는 스펙 변화 가능성의 원인은 현재 최대 가능 면적이 8,150mm²인 CoWoS-L 패키지의 한계이다. 원안대로 4개의 레티클을 사용할 경우, 4개 레티클, 16개 HBM, 2개의 I/O와 오버헤드 면적까지 모두 고려하면 총 6,750mm²의 면적에 달하게 된다. 반면 축소안대로 2개의 레티클을 사용할 경우, 2개 레티클, 8개 HBM, 2개의 I/O 면적, 오버헤드 면적을 감안하면 총 3,800mm²의 면적이 예상된다.

4-die의 6,750mm²는 최대 인터포저 면적인 8,150mm²보다 작으므로 산술적으로 제조가 가능하다. 그러나 인터포저의 절대 크기가 커질수록 워피지 현상이 대각선 길이의 제곱에 가깝게 비선형적으로 증가하게 된다. 따라서 결과적으로 모서리에 응력이 집중되며 범프 피로도 상승, 접합 수율 저하, 박리 리스크가 높아지는 문제가 발생할 가능성이 높다. 8,150mm²의 인터포저 면적은 물리적 최대치이지 양산 가능 구간은 아니란 의미이다. 만약 NVIDIA가 Rubin Ultra의 컴퓨터 다이를 원안의 4개가 아닌 현재 Rubin까지 채택 중인 2개로 되돌리면, HBM 큐브도 원안의 16개에서 8개로 자동 감소하게 된다.

두번째 원인으로 거론 중인 것은 16단 HBM4E의 적층 수율이다. 16단은 micro bump의 두께 및 경제성의 한계로 하이브리드 본딩 또는 초고난도 MR-MUF가 필요하다. JEDEC이 설정한 775μm 높이의 제한 내로 16층을 적층하기 위해서는 웨이퍼를 약 30μm까지 박막화할 필요가 있다. 물론 16단 HBM4는 MR-MUF 기반으로 2H26에 양산 계획이 잡혀 있다. 다만 1C 나노 코어 DRAM과 16Gbps의 입출력 속도를 가지는 16단 HBM4E는 열, 신호 마진이 더 타이트해 CY27 양산 성공이 어려울 수 있다. 따라서 이 경우 CY27 양산 HBM4E의 주력은 48GB (32Gb x 12단)이고, 64GB (32Gb x 16단)는 CY27 말 이후 시장 진입이 유력하게 된다.

(다음 페이지에서 계속)

[반도체]
송명섭
2122-9207
mssong@imfnsec.com

[IT RA]
박정하
2122-9195
jhpark@imfnsec.com

적층 (16단)은 메모리 3사가, 레티클/인터포저는 TSMC (CoWoS-L → CoPoS)가 풀어야 할 과제이며, 완전한 1TB HBM 용량의 Rubin Ultra (16큐브 × 64GB)는 메모리 16단 수율과 TSMC의 CoPoS 패키징이 동시에 성숙해야 가능할 전망이다.

향후 HBM4E 관련 관전 포인트

향후 약 1년간의 HBM4E 관련 핵심 관전 포인트는 '16단 HBM4E 샘플을 누가 먼저 제출하고, 누가 먼저 인증을 통과하느냐'인 것으로 판단된다. 단 16단 후공정에서 삼성전자와 SK하이닉스의 공정 노선이 갈릴 가능성이 높다는 점이 중요하다 (<표2> 참조).

16단 인증은 삼성전자의 하이브리드 신공정 vs SK하이닉스의 검증된 MR-MUF 연장의 대결이 될 가능성이 높은 것으로 보인다. 단기 (16단) 적으로는 안정성에서 SK하이닉스의 MR-MUF가 유리할 수 있고, 중장기 (20단 이상) 적으로는 삼성전자의 하이브리드 본딩 선행 투자가 빛을 볼 수 있는 상반된 리스크 구도인 것으로 평가된다.

일단 현재까지는 지난 5월 29일에 삼성전자가 세계 최초로 12단 48GB HBM4E 샘플을 출하한 바 있으며, SK하이닉스는 당초 2H26로 예정되었던 동 제품 샘플 출하 계획을 6월, 7월 경으로 앞당긴 상황인 것으로 보인다 (<표3> 참조).

표1. Rubin Ultra에서의 HBM4E 용량 원안과 축소안 비교

HBM4E	GPU당 Reticle 수	Reticle 당 HBM Cube 수	GPU당 HBM Cube 수	Core Die (GB)	Stack (단)	Cube당 Density (GB)	GPU당 HBM Density (GB)
원안	4	4	16	4	16	64	1,024
축소안	2	4	8	4	12	48	384

자료: Nvidia, Trendforce, Wccftech

표2. HBM 업체별 HBM4E 16단 본딩 방식 전망

업체	16단 공정	리스크
삼성전자	Hybrid Bonding (Cu-Cu 직접 접합) 16단부터 적용 전망	신공정이므로 초기 수율 리스크 존재 성공 시 두께, 열 마진 우위 성공 시 20단 이후에도 경쟁력 우위 가능성
SK하이닉스	Advanced MR-MUF를 16단까지 유지하고 Hybrid Bonding은 20단 이후 도입할 듯	검증된 공정의 연장이므로 수율 리스크 낮음 단 MR-MUF로 16단 두께는 물리적 한계 근접
마이크론	1y 공정, base die TSMC 외주	HBM4E 양산 2027 목표 커스텀 경쟁 상대적 후발

자료: 각 사, iM증권

표3. HBM 업체별 HBM4E 진행 상황

업체	샘플/양산 일정	기술 구성
삼성전자	5/29 세계 최초 12단 48GB HBM4E 샘플 출하 (NVIDIA 등)	1c DRAM + 자체 4nm base die 최대 16Gbps, 스택당 4TB/s
SK하이닉스	6월~7월 샘플 출하 추진 당초 하반기 계획에서 앞당김	1c DRAM base die TSMC 3nm 검토
마이크론	1y 공정, base die TSMC 외주	1y 공정 EUV 도입 base die TSMC 외주

자료: 각 사, iM증권

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]

<p>종목추천 투자등급</p> <p>종목투자의견은 향후 12개월간 추천일 증가대비 해당종목의 예상 목표수익률을 의미함.</p> <p>· Buy(매수): 추천일 증가대비 +15% 이상</p> <p>· Hold(보유): 추천일 증가대비 -15% ~ 15% 내외 등락</p> <p>· Sell(매도): 추천일 증가대비 -15% 이상</p>	<p>산업추천 투자등급</p> <p>시가총액기준 산업별 시장비중대비 보유비중의 변화를 추천하는 것임</p> <p>· Overweight(비중확대)</p> <p>· Neutral(중립)</p> <p>· Underweight(비중축소)</p>
--	---

[투자등급 비율 2026-03-31 기준]

매수	중립(보유)	매도
88.2%	11.8%	-