리서치 노트 · 2026년 11월

합의가 된 연산

2026년의 AI × 크립토 × 칩: Pearl에서 출발하다

Pearl · 유용한 작업 · DePIN과 에이전트 · ASIC 경제학 · 공정 노드 · 단계별 예산

탄 슈아이

tanshuai.com · i@tanshuai.com



정리는 성립하지만, 인센티브는 성립하지 않는다

01 Pearl은 공학으로는 작동하지만, 아직 유용한 작업은 아니다

약 90 EH/s의 정수 행렬 곱셈이 체인을 보호하지만, 실제 AI 작업을 담은 블록은 약 1%뿐이다.

02 유료 수요는 추론, 데이터, 프라이버시에 있다

Bittensor는 검증 가능한 수익이 연 약 7,000만 달러, Grass는 2026년 상반기 1,700만 달러, TEE는 실서비스 중; 에이전트 결제는 월 약 5,000–11,000달러.

03 ASIC은 큰 보상 풀, 안정된 규칙, 두 번째 시장이 있어야 수익을 낸다

이익 상한은 R × (1 − 1/k)이다. 지금 세 가지를 모두 갖춘 네트워크는 없다 분석.

≈90 EH/s

2026년 11월 13일의 Pearl: H100 약 122,000장 분량의 정수 처리량

≈1%

실제 AI 작업을 담은 Pearl 블록(Blockworks)

4–11×

검증 가능한 수익 대비 Bittensor의 발행량

여섯 부문의 AI × 크립토: 유료 수요는 어디에 있나

유용한 작업 체인 01

실재하나 작음

연 ≈$70M

Bittensor의 검증 가능한 외부 수익

신호: 발행량은 수익의 4–11×; Pearl AI 블록 ≈1%

메커니즘은 작동; 유용성에는 지불자가 필요

탈중앙화 컴퓨팅 02

실재하나 작음

연 $2.35M

Render의 온체인 수수료, 검증 가능한 최대치

신호: 시가총액 ≈ 수수료의 360×

대부분 범용 추론의 재판매

학습과 추론 03

이익은 지분으로

연 $100M+

중앙화 GPU를 쓰는 지분 회사 Prime Intellect

신호: Gensyn 학습 메인넷 미가동

진전은 실재; 이익은 지분으로

검증 가능한 AI 04

초기

398

Phala TEE 클라우드 유료 사용자, 2025년 말

신호: 공개된 유료 zkML 없음

TEE는 실서비스 중; zkML은 초기

에이전트 결제 05

수요 이전

월 $5–11K

에이전트가 실제로 개시한 x402 결제

신호: ≈2억 건 거래, 대부분 신호용

레일이 먼저, 수요는 아직

데이터 네트워크 06

실제 구매자

$17M

Grass 데이터 수익, 2026년 상반기(자체 보고)

신호: 시가총액 ≈ 연환산 수익의 10×

실제 구매자가 있는 드문 부문

보조금은 콜드 스타트의 연료이지 사업 모델이 아니다. 크립토는 유통과 결제를 맡고, 이익은 지분이 가져간다. 적어도 2026년에는 그렇다.

곱셈-누산 한 번이 해시 하나다

정리에서 메인넷까지

Komargodski와 Weinstein(2025년 4월): 모든 행렬 곱에 대해 1 + o(1) 오버헤드의 유용한 작업 증명. Pearl 메인넷은 2026년 4월 27일 시작됐다.

작동 방식

채굴자는 노이즈를 넣은 행렬을 곱하고, 각 타일의 누산기를 해시해 복권으로 삼으며, 당첨 타일을 60 KB 미만의 Plonky2 증명으로 입증한다.

단위

1 H는 인정된 INT7 × INT7 → INT32 곱셈-누산 1회: 텐서 처리량으로 매긴 블록체인 보안 예산이다.

2026년 11월 13일의 Pearl

89.77 EH/s

네트워크 해시레이트(독립 풀 노드)

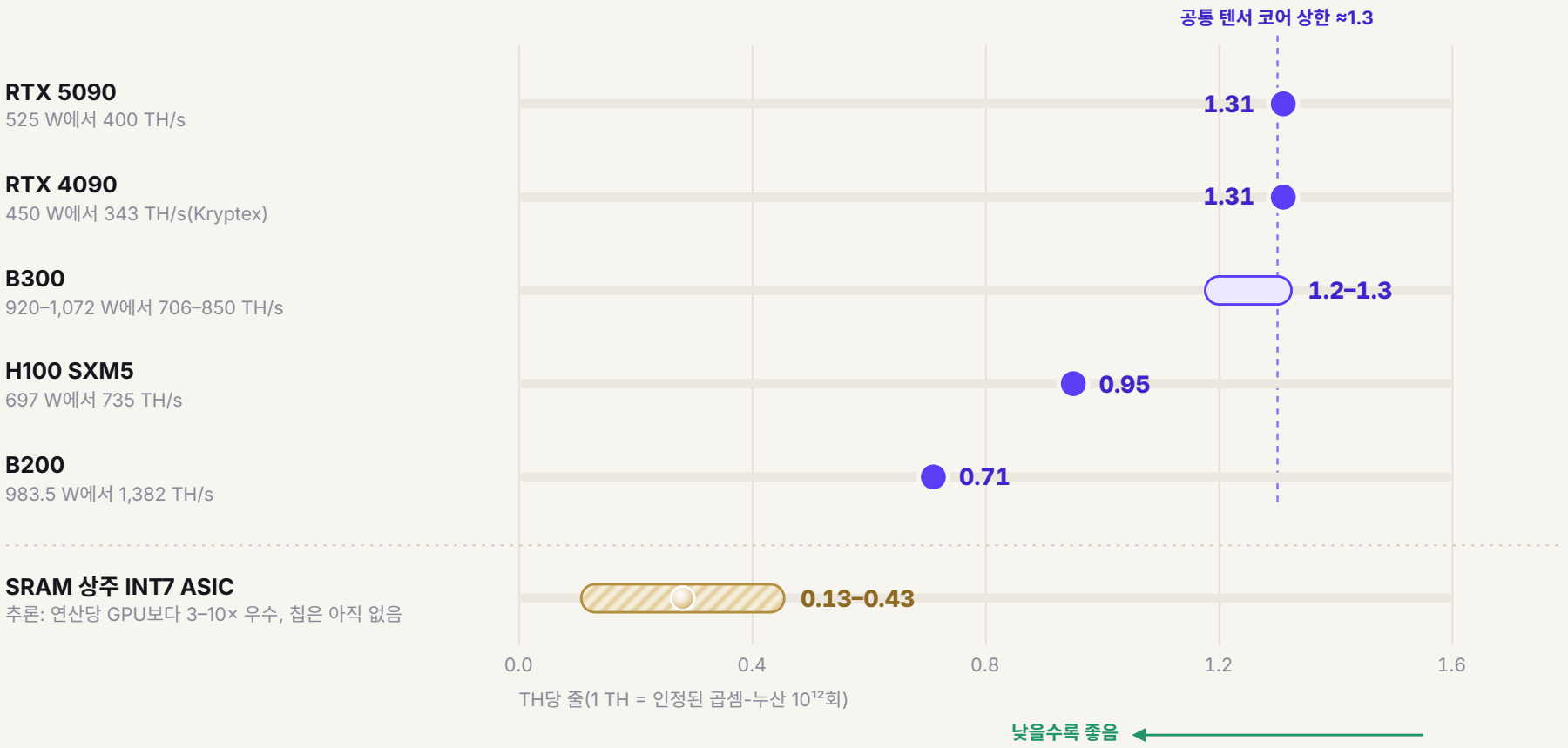
≈$500M

시가총액; 완전 희석 기준 약 28억 달러

+78%

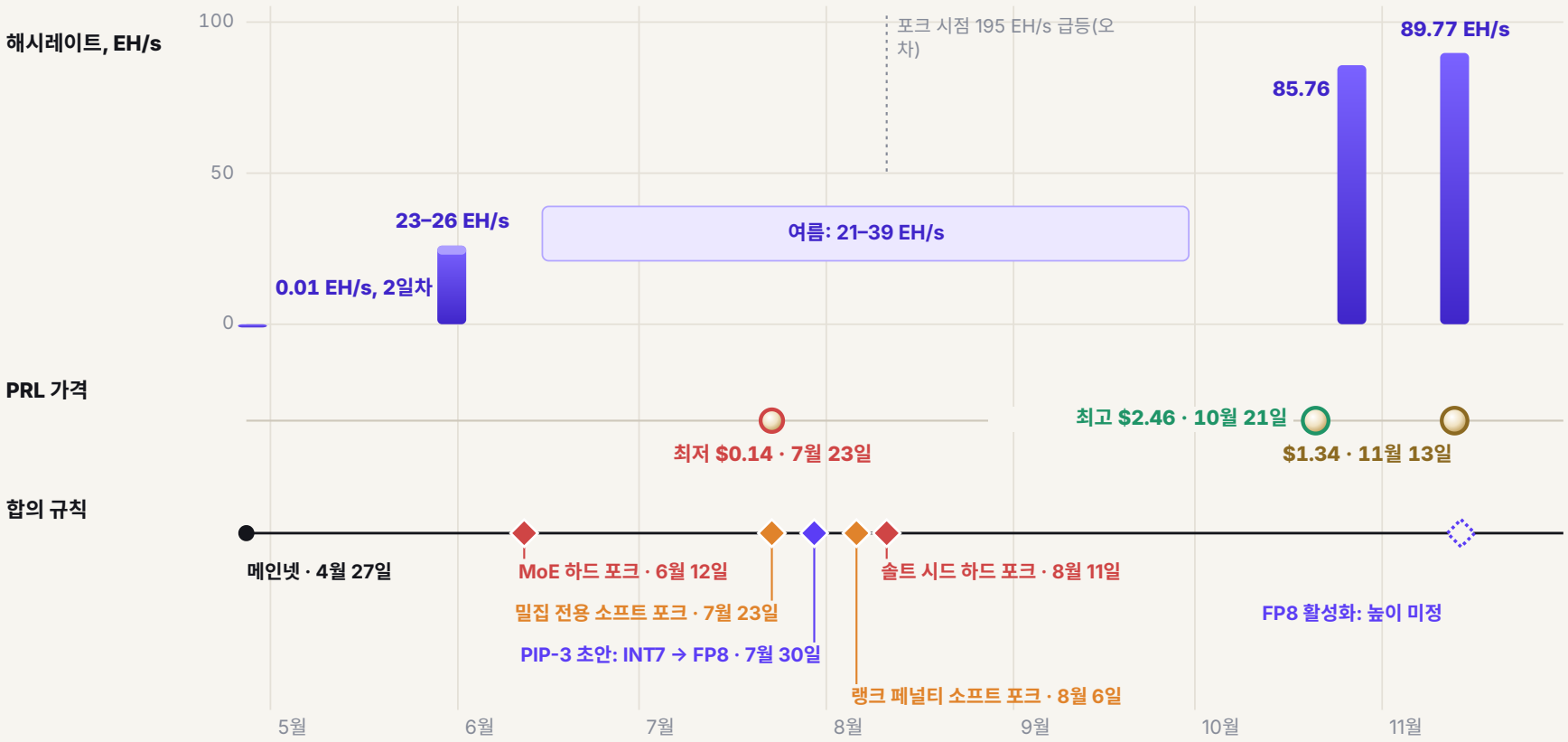
향후 12개월 신규 공급

Pearl의 정수 연산에서 GPU는 같은 에너지 상한에 걸린다



게이밍 카드와 B300은 약 1.3 J/TH에서 만나고, B200도 0.71에 그친다. SRAM에서 7비트 정수 곱셈만 하는 칩이라면 3–10배 나을 수 있다 분석.

해시레이트는 가격을 따르고, 규칙은 계속 바뀐다



현재 공급량의 32%인 약 1억 2,000만 개가 첫 6일 만에 채굴됐다. RTX 5090은 지금 하루 약 5–6달러를 벌며, 채굴만을 위해 카드를 빌리면 손해다.

정리는 성립하지만, 인센티브는 성립하지 않는다

유용한 작업의 비용

추론 쪽: 작은 오버헤드

Pearl은 H200 네 장에서 Llama 70B의 소요 시간이 5.08% 늘었다고 보고했다. 독립 시험에서는 Hopper GPU 한 장의 처리량이 9.2% 줄었다.

유용한 작업의 수입

채굴 쪽: 너무 적은 해시레이트

추론을 서비스하는 카드는 전용 채굴기 대비 프리필에서 약 27%, 채팅에서 5.4%, 디코딩이 지배하면 약 1%만 채굴한다.

그래서 합리적인 채굴자는 무작위 행렬을 곱한다. 유용성은 오프체인의 상업 문제다. 누군가 추론 비용을 내야 한다.

두 달 만에 포크 네 번, 다음은 FP8

6월 12일
전문가 혼합(MoE) 하드 포크 (블록 71,935)

7월 23일
밀집 전용 소프트 포크(블록 91,630)

8월 6일
랭크 페널티 소프트 포크: 일부 구성은 유효 해시레이트의 50–87.5%를 잃는다

8월 11일
솔트 시드 하드 포크: 구형 채굴기는 무효 셰어만 생산한다

미정
PIP-3: INT7에서 FP8로, Hopper와 Blackwell만 허용

3 중 2

AI에 유용, 비트 단위 검증 가능, 하드웨어 중립: 수치 형식은 최대 둘만 얻는다

13비트

Hopper의 FP8 누산이 절삭과 함께 유지하는 비트: 세대 간 재현 불가

≈82%

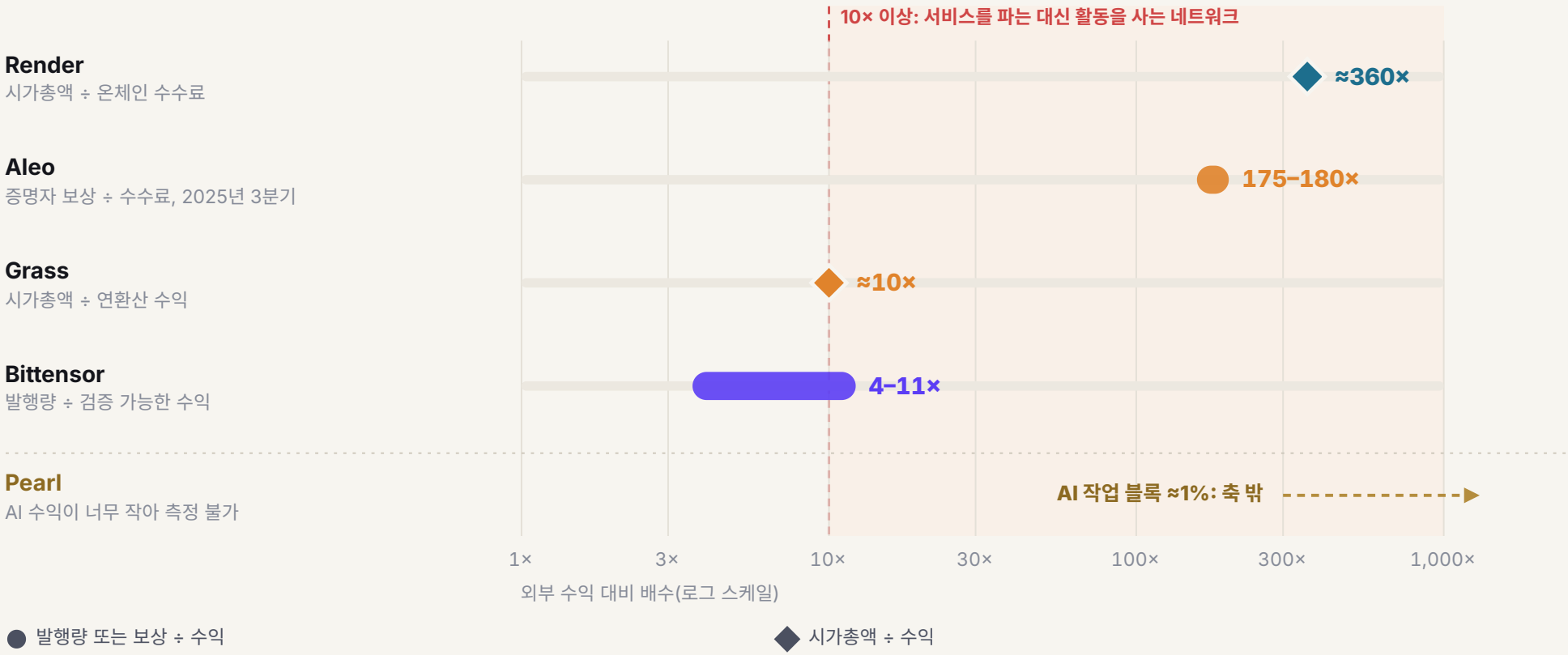
PIP-3가 허용하지 않을 소비자용 카드의 해시레이트 비중(저신뢰 모델)

작업 함수를 통제하는 쪽이 하드웨어 투자자의 운명을 통제한다 분석.

행렬도, 영지식도 ASIC을 막지 못한다

네트워크	일어난 일	교훈
Kaspa	해시 사이의 64 × 64 행렬 곱; 16개월 뒤 ASIC 등장, GPU의 약 1,000배 효율; 최고 기종도 지금 하루 0.86달러	행렬은 유용하지도, ASIC 내성도 없다
Aleo	합성 SNARK 퍼즐; 5개월 뒤 전용 장비; 증명자 보상은 수수료의 175–180배	합성 작업은 보조금으로 돌아간다
Qubic	AI 학습을 내세우지만 연산은 Monero(51% 주장), 지금은 Dogecoin을 채굴	범용 PoW는 외부에서 장악될 수 있다
Bittensor	지분 가중 채점; 검증 가능한 수익 연 약 7,000만 달러 대 발행량 3억 달러 이상	규모 있는 외부 수익을 가진 유일한 사례
Monero, Sia, Ethereum	포크로 ASIC이 무용지물이 됐고, Merge 후 새 일을 찾은 채굴자는 20% 미만	두 번째 시장이 있는 하드웨어가 살아남는다

보조금 배수는 가장 정직한 건강 검진이다



외부 수익의 약 10배를 넘으면 네트워크는 서비스를 파는 것이 아니라 활동을 사고 있다. Pearl의 AI 수익은 너무 작아 축에 올릴 수 없다.

실재하지만 작은 수요, 그리고 지분으로 옮겨가는 가치

- **컴퓨팅 시장**

  Render의 온체인 수수료는 연 약 235만 달러; Akash의 H100 시간당 2.73달러는 AWS와 Google보다 싸지만 Vast.ai나 RunPod보다는 비싸다.

- **학습**

  인터넷으로 80억–400억 파라미터를 학습; INTELLECT-3는 InfiniBand로 연결된 H200 512장으로 돌아갔다.

- **검증 가능한 AI**

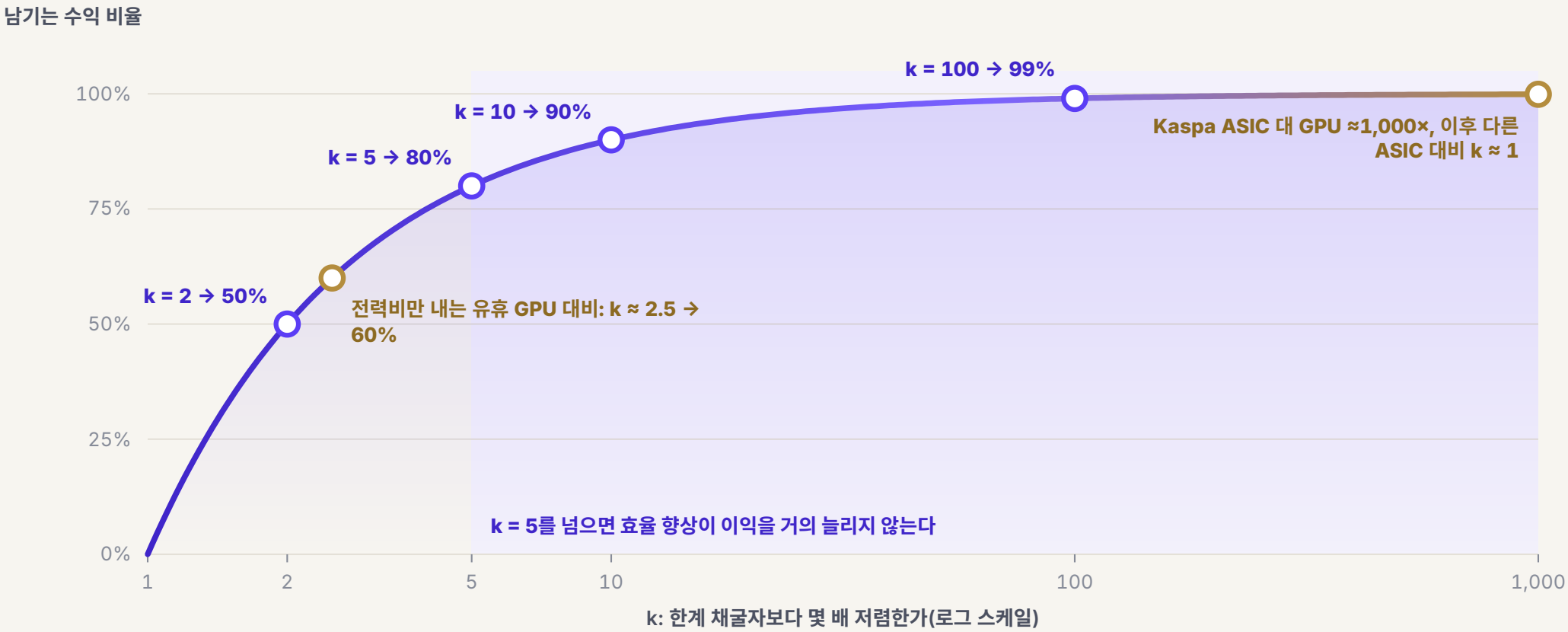
  TEE는 오버헤드 7% 미만으로 실서비스 중; TEE.fail은 1,000달러짜리 인터포저로 증명을 깼다; zkML은 GPT-2를 분당 174토큰으로 증명한다.

- **에이전트와 데이터**

  에이전트가 개시한 x402 흐름은 월 약 5,000–11,000달러; Grass는 2026년 상반기 데이터 수익 1,700만 달러.

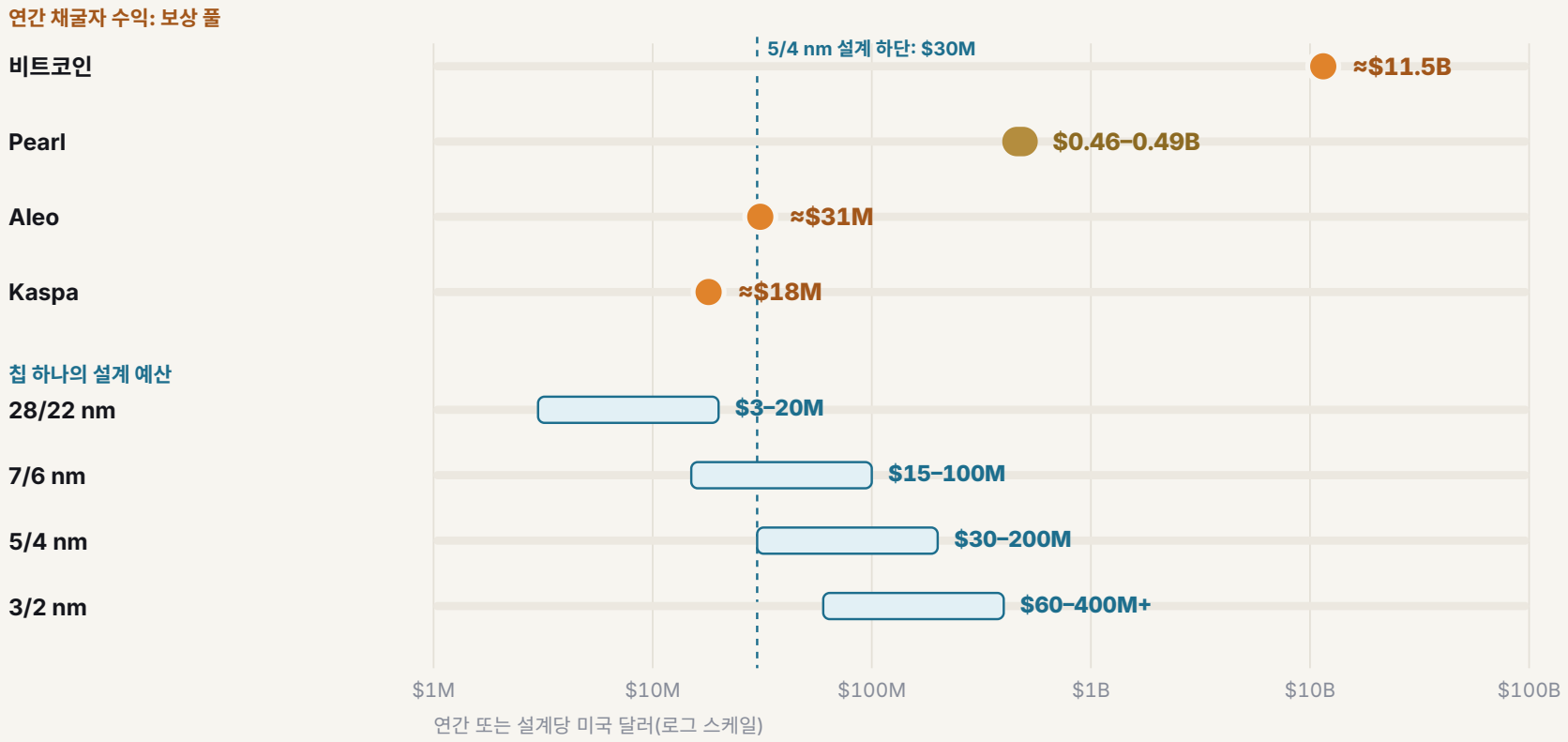
> **중앙화 인프라 위의 지분 기업 네 곳의 가치는 약 85억 달러로, 프로토콜·애플리케이션 계층 AI 토큰 전부와 비슷하다.**

k = 5를 넘으면 효율이 더는 돈이 되지 않는 이유



이익 상한 ≈ R × (1 − 1/k). ASIC이 한계 채굴자가 되면 k는 다른 ASIC 기준으로 재설정된다. 전력비만 내는 유휴 GPU 대비 k는 약 2.5다 추정.

대부분의 보상 풀은 칩 하나의 설계 예산보다 작다



회수 해시레이트 h* = NRE ÷ [R × (1 − 1/k) × T × p]. NRE 2,000만 달러, k = 5인 Pearl이라면 규칙이 유지될 때 1년간 약 4.9 EH/s, 유지 확률이 반반이면 9.8 EH/s 추정.

칩별 분석

다섯 가지 칩, 다섯 가지 답



Pearl형 PoUW 채굴 칩

지금은 아님

k: HBM·범용 유닛 없음: 연산당 3–10×(추론)

두 번째 시장: 범용 INT8/FP8 GEMM이면 AI 추론

전용 INT7 칩이 아니라 채굴도 되는 추론 우선 칩에 걸라



비트코인 SHA-256 채굴 칩

관망

k: 기존 강자 대비 ≤1

두 번째 시장: 거의 없음

자체 전력, 틈새 폼팩터, 또는 IP로만



AI 학습 칩

칩 대신 IP

k: 같은 정밀도에서 1.1–1.8×

두 번째 시장: —

패키징과 HBM이 좌우; IP·칩렛·서비스를 만들라



AI 추론 칩

개발

k: 범용성 포기: 10–20× 주장

두 번째 시장: 이것이 주력 시장

고객과 먼저 계약하고 종속 수준을 정하라



ZK 증명 가속기

대기·헤지

k: GPU의 10–100× 주장

두 번째 시장: ZK 밖에서는 제한적

프로그래머블 MSM·NTT 엔진; FPGA부터



FHE 가속기

조기 포지션

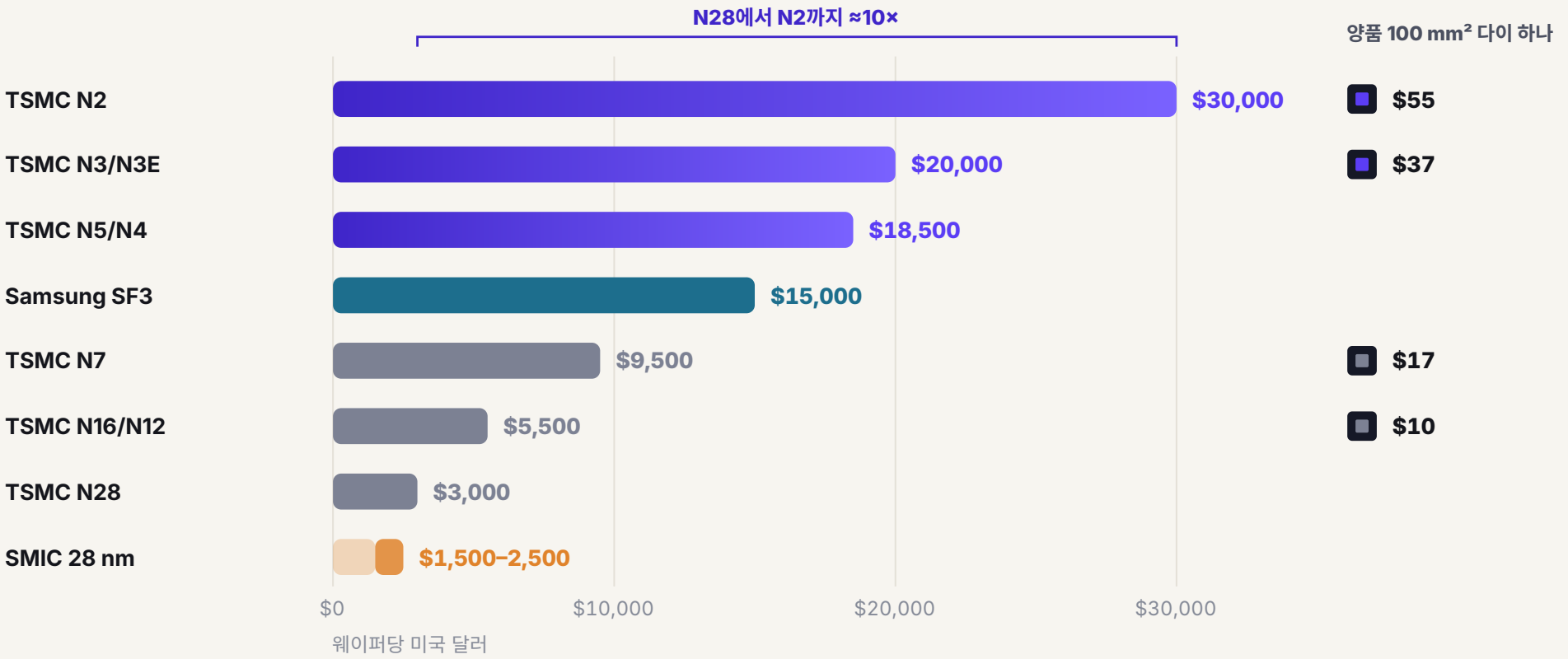
k: Xeon의 1,074–5,547×(Heracles)

두 번째 시장: 암호화 연산, 프라이빗 AI

2027–2029년 기회; 가벼운 자산으로 검증

큰 보상 풀, 안정된 규칙, 1보다 큰 k를 함께 갖춘 네트워크는 지금 없다. 비트코인의 k는 기존 강자의 것이고, Pearl의 규칙은 가장 불안정하며, Kaspa와 Aleo의 보상 풀은 너무 작다.

최첨단 웨이퍼는 성숙 공정의 열 배다



> 수명 주기 비용 중 가장 큰 항목이 선택하게 하라: 24시간 도는 연산 칩은 에너지, 학습 칩은 패키징과 HBM(H100 비용의 63%), 추론 칩은 SRAM이냐 HBM이냐.

공정 노드

칩마다, 팀마다 맞는 노드

칩	소규모 팀	규모를 갖춘 회사	중국 본토 경로
PoUW 채굴 칩	28 또는 16/12 nm MPW; 첫 양산 N7/N6	N3/N3E 또는 Samsung SF3	SMIC N+1/N+2 가능, 다이당 약 40–50% 더 비쌈
SHA-256 채굴 칩	최첨단을 쫓지 말 것	3 nm급	최첨단 노드 효율을 따라잡기 어려움
AI 학습 칩	IP, 칩렛, 설계 서비스	N3P 다음 N2/A16, CoWoS와 HBM	SMIC N+2/N+3와 국산 HBM
AI 추론 칩	N6/N7 또는 N4/N5/SF4의 SRAM 중심 설계	HBM을 단 N3/N2, 또는 SRAM 칩렛	SMIC N+2와 LPDDR 또는 GDDR
ZK·FHE 가속기	FPGA 또는 16/12 nm MPW부터	N5/N3; FHE는 HBM과 함께	ZK는 가능; FHE는 HBM이 제약

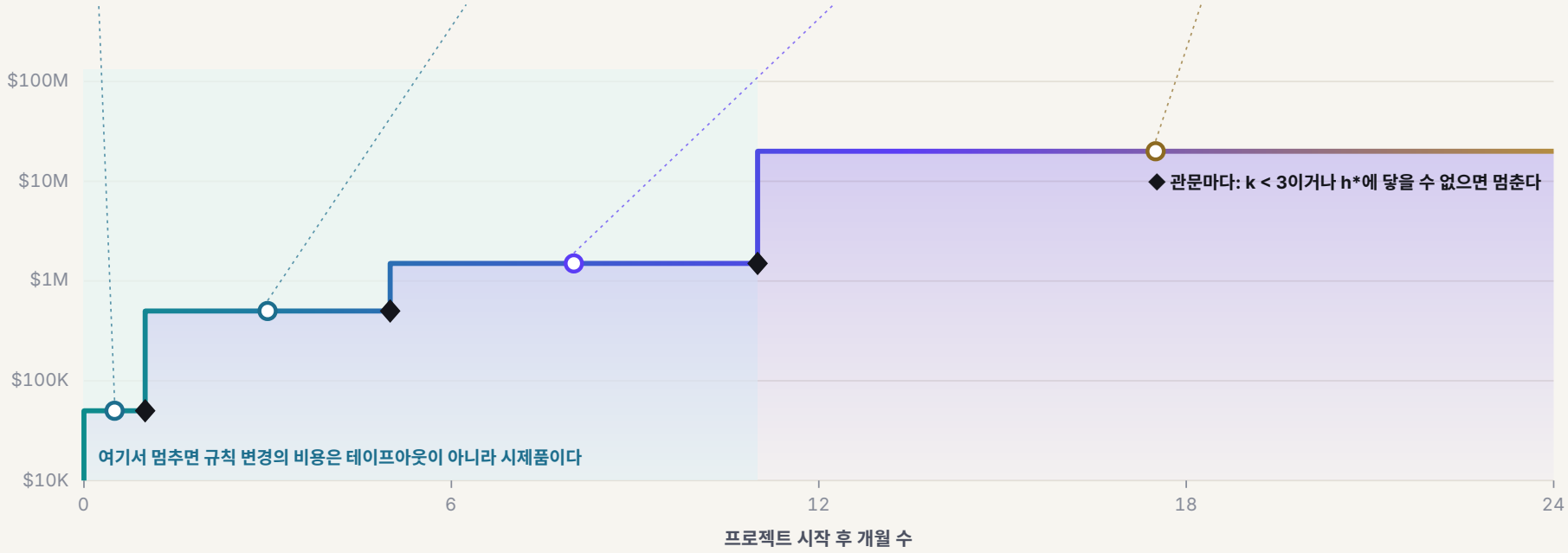
마스크보다 증거를 먼저 사라

관문 0 · 모델
≈$50K · 수 주
산술, SRAM, I/O로 J/op와 $/op 추정

관문 1 · FPGA
≈$0.5M · 수개월
실제 워크로드를 끝까지: 셰어 또는 토큰

관문 2 · MPW
≈$1.5M · 11개월차
실측 실리콘이 모델을 대체

관문 3 · 마스크와 첫 로트
≈$20M · 24개월차
$15–100M 설계 안의 N7 마스크 ≈$3.5M



모델, FPGA, MPW로 11개월에 약 150만 달러면 실측 실리콘에 이른다. 마스크 세트와 첫 로트는 그 열 배 이상이 든다. 구매자나 프로토콜 버전이 확정될 때까지 마스크 전에 멈춰라.

크립토는 첫 고객, AI는 두 번째 시장

맞춤형 칩

- 수요는 압도적으로 AI: Broadcom의 2026 회계연도 AI 반도체 약 580억 달러
- 채굴 칩 역량은 옮겨간다: 저전압 어레이, 킬로와트급 전력, 빠른 테이프아웃
- 채굴 유통망이 아니라 독립적인 시장 진입 경로로 승부하라

AI 개발

- 시간은 범용 상품: Clutch에서 시간당 25–49달러
- 가치는 연결부에 있다: 맞춤형 칩용 커널, 추론 서빙 안의 채굴
- 검증 가능한 추론: TEE와 ZK 표본 검사

AI × 크립토 시스템

- 계층형 검증: TEE 또는 샘플링, 그다음 분쟁 또는 ZK 표본 검사
- 장벽은 비용이 아니라 유통과 유동성
- 토큰보다 연산 구매자와 먼저 계약하라

내가 할 일

세 갈래 길, 각각 작은 첫걸음

경로	첫걸음
A. 채굴도 되는 추론 우선 가속기	SRAM 상주 INT8/FP8 어레이의 J/op와 $/op를 RTX 5090·H100과 비교하는 4주 연구, 그다음 FPGA 또는 성숙 노드 MPW
B. 맞춤형 실리콘 서비스: 크립토 먼저, AI 다음	이 보고서의 NRE·손익분기 모델을 마일스톤과 고정 프로토콜 버전을 둔 유료 타당성 조사로 바꾼다
C. 토큰보다 구매자가 먼저인 AI 네이티브 네트워크	계층형 검증과 업체 중립적 표준 산술에 관한 설계 노트를, 이름 있는 추론 구매자 한 곳과 함께 쓴다

> **길은 최대 세 갈래, 모두 돈을 걸기 전에 시험한다** 분석.

지켜볼 것, 그리고 던져야 할 질문

이정표

수시	Pearl의 FP8 활성화 높이, 그리고 체인 분열 여부
매월	AI 작업을 담은 Pearl 블록 비중
분기별	Bittensor의 검증 가능한 수익 대 발행량; 에이전트가 개시한 x402 흐름
2027년 1월	보도된 TSMC의 2 nm·3 nm 가격 3–6% 인상
2027년 1분기	창업자 AMA에서 보도된 Pearl의 FP4 도입
2028년 4월	비트코인의 다음 반감기

던져야 할 질문

01 작업 함수는 누가 통제하는가?

02 보조금 배수는 얼마인가?

03 다음 관문의 비용은 얼마이고, 무엇이 멈추게 하는가?

04 p를 반영한 회수 해시레이트는 얼마인가?

05 칩의 두 번째 시장은 무엇인가?

06 첫 유료 고객은 구체적으로 누구인가?

맞춤형 칩, AI 개발, AI × 크립토 시스템은 기술 교류와 기술 협력을 통해 함께 탐구할 수 있는 분야다. i@tanshuai.com · tanshuai.com