บันทึกการวิจัย · พฤศจิกายน 2026

การประมวลผลคือฉันทามติ

AI × คริปโต × ชิป ในปี 2026 เริ่มจาก Pearl

Pearl · งานที่มีประโยชน์ · DePIN และเอเจนต์ · เศรษฐศาสตร์ ASIC · โหนดการผลิต · งบแบบเป็นขั้น

ทัน ชว่าย

tanshuai.com · i@tanshuai.com



ข้อสรุป

ทฤษฎีบทใช้ได้ แต่แรงจูงใจใช้ไม่ได้

01 Pearl ใช้ได้ในทางวิศวกรรม แต่ยังไม่เป็นงานที่มีประโยชน์

การคูณเมทริกซ์เลขจำนวนเต็มราว 90 EH/s รักษาความปลอดภัยของเชน แต่มีเพียงราว 1% ของบล็อกที่มีงาน AI จริง

02 อุปสงค์ที่จ่ายเงินจริงอยู่ที่การอนุมาน ข้อมูล และความเป็นส่วนตัว

Bittensor มีรายได้ที่ตรวจสอบได้ราว 70 ล้านดอลลาร์ต่อปี Grass 17 ล้านดอลลาร์ ในครึ่งแรกของปี 2026 TEE ใช้งานจริงแล้ว การชำระเงินของเอเจนต์ราว 5,000–11,000 ดอลลาร์ต่อเดือน

03 ASIC คุ้มทุนเมื่อกองรายได้ใหญ่ กติกามั่นคง และมีตลาดที่สองเท่านั้น

เพดานกำไรคือ R × (1 − 1/k) วันนี้ยังไม่มีเครือข่ายใดมีครบทั้งสามข้อ บทวิเคราะห์

≈90 EH/s

Pearl ณ 13 พฤศจิกายน 2026: เท่ากับปริมาณงานเลขจำนวนเต็มของ H100 ราว 122,000 ใบ

≈1%

ของบล็อก Pearl ที่มีงาน AI จริง (Blockworks)

4–11×

การปล่อยเหรียญของ Bittensor เทียบกับรายได้ที่ตรวจสอบได้

AI × คริปโต 6 กลุ่ม: อุปสงค์ที่จ่ายเงินจริงอยู่ที่ไหน

เชนงานที่มีประโยชน์ 01

จริงแต่เล็ก

≈$70M ต่อปี

รายได้จากภายนอกที่ตรวจสอบได้ของ Bittensor

สัญญาณ: การปล่อยเหรียญ 4–11× ของรายได้; Pearl มีบล็อก AI ≈1%

กลไกใช้ได้ ประโยชน์ต้องมีผู้จ่าย

การประมวลผลแบบกระจายศูนย์ 02

จริงแต่เล็ก

$2.35M ต่อปี

ค่าธรรมเนียมออนเชนของ Render สูงสุดที่ตรวจสอบได้

สัญญาณ: มูลค่าตลาด ≈360× ของค่าธรรมเนียม

ส่วนใหญ่ขายต่อการอนุมานทั่วไป

การเทรนและการอนุมาน 03

กำไรไปที่หุ้น

$100M+ ต่อปี

Prime Intellect บริษัทเอกชนบน GPU แบบรวมศูนย์

สัญญาณ: เมนเน็ตการเทรนของ Gensyn ยังไม่เปิด

ความก้าวหน้าจริง กำไรไปที่หุ้น

AI ที่ตรวจสอบได้ 04

ระยะแรก

398

ผู้ใช้แบบจ่ายเงินของคลาวด์ TEE ของ Phala ปลายปี 2025

สัญญาณ: ไม่มีการเปิดเผย zkML แบบจ่ายเงิน

TEE ใช้งานจริงแล้ว zkML ยังระยะแรก

การชำระเงินของเอเจนต์ 05

ยังไม่มีอุปสงค์

$5–11K ต่อเดือน

การชำระ x402 ที่เอเจนต์เป็นผู้เริ่มจริง

สัญญาณ: ≈200M ธุรกรรม ส่วนใหญ่เป็นการส่งสัญญาณ

รางมาก่อน อุปสงค์ยังไม่มา

เครือข่ายข้อมูล 06

มีผู้ซื้อจริง

$17M

รายได้ข้อมูลของ Grass ครึ่งแรกของปี 2026 (รายงานเอง)

สัญญาณ: มูลค่าตลาด ≈10× ของรายได้ต่อปี

หนึ่งในไม่กี่กลุ่มที่มีผู้ซื้อจริง

เงินอุดหนุนเป็นเชื้อเพลิงสำหรับการเริ่มต้น ไม่ใช่โมเดลธุรกิจ คริปโตทำหน้าที่กระจายและชำระราคา ส่วนกำไรตกอยู่กับหุ้นของบริษัท อย่างน้อยก็ในปี 2026

การคูณสะสมหนึ่งครั้งคือแฮชหนึ่งครั้ง

จากทฤษฎีบทสู่เมนเน็ต

Komargodski และ Weinstein (เมษายน 2025): การพิสูจน์งานที่มีประโยชน์สำหรับผลคูณเมทริกซ์ใดๆ ด้วยโอเวอร์เฮด 1 + o(1) เมนเน็ตของ Pearl เริ่มเมื่อ 27 เมษายน 2026

ทำงานอย่างไร

นักขุดคูณเมทริกซ์ที่เติมสัญญาณรบกวน แฮชตัวสะสมของแต่ละไทล์เป็นสลากลุ้นรางวัล และพิสูจน์ไทล์ที่ชนะด้วยการพิสูจน์ Plonky2 ขนาดไม่ถึง 60 KB

หน่วยวัด

1 H คือการคูณสะสม INT7 × INT7 → INT32 ที่นับเครดิตหนึ่งครั้ง: งบความปลอดภัยของบล็อกเชนที่คิดราคาเป็นปริมาณงานเทนเซอร์

Pearl ณ 13 พฤศจิกายน 2026

89.77 EH/s

แฮชเรตของเครือข่าย (ฟูลโหนดอิสระ)

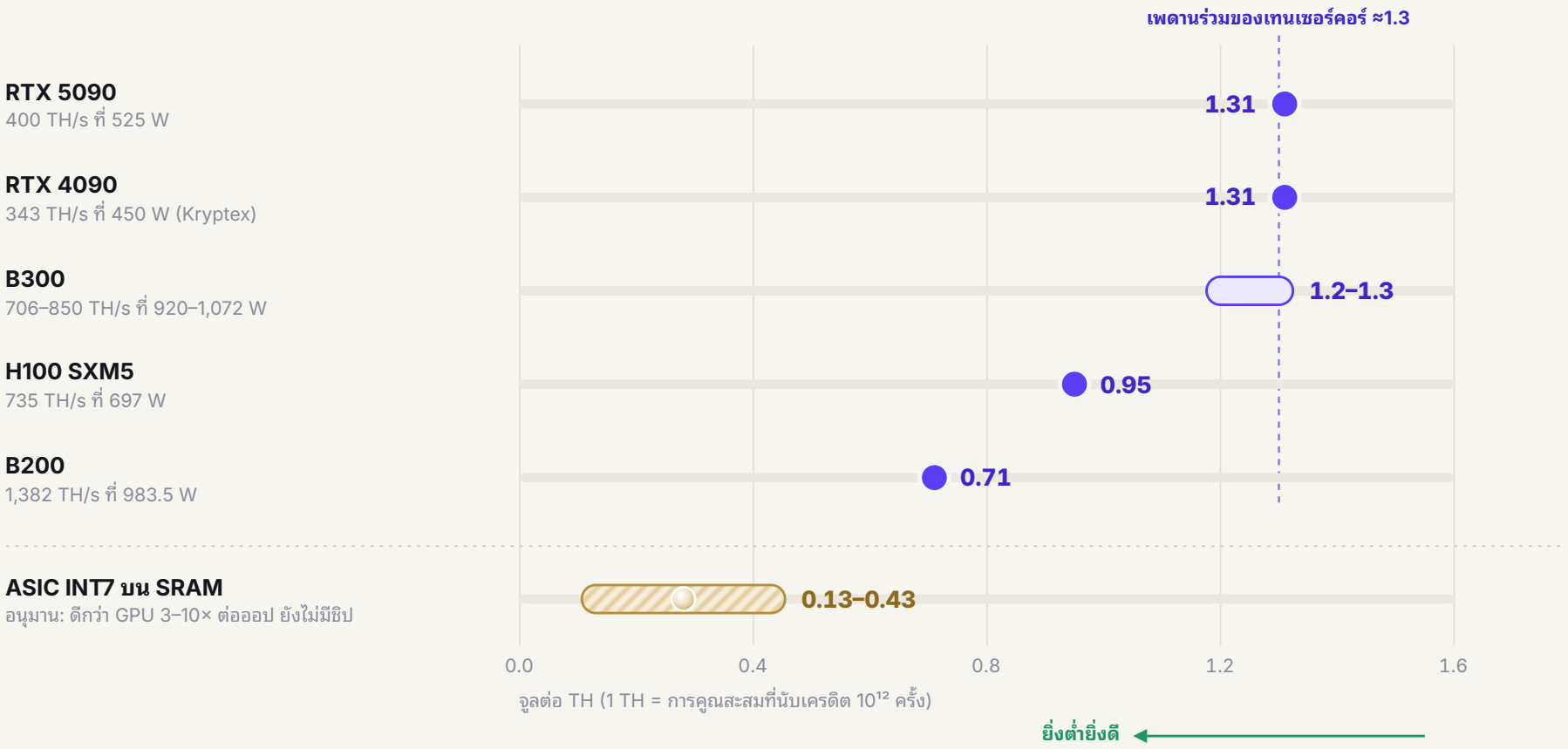
≈$500M

มูลค่าตลาด; แบบ fully diluted ราว 2.8 พันล้านดอลลาร์

+78%

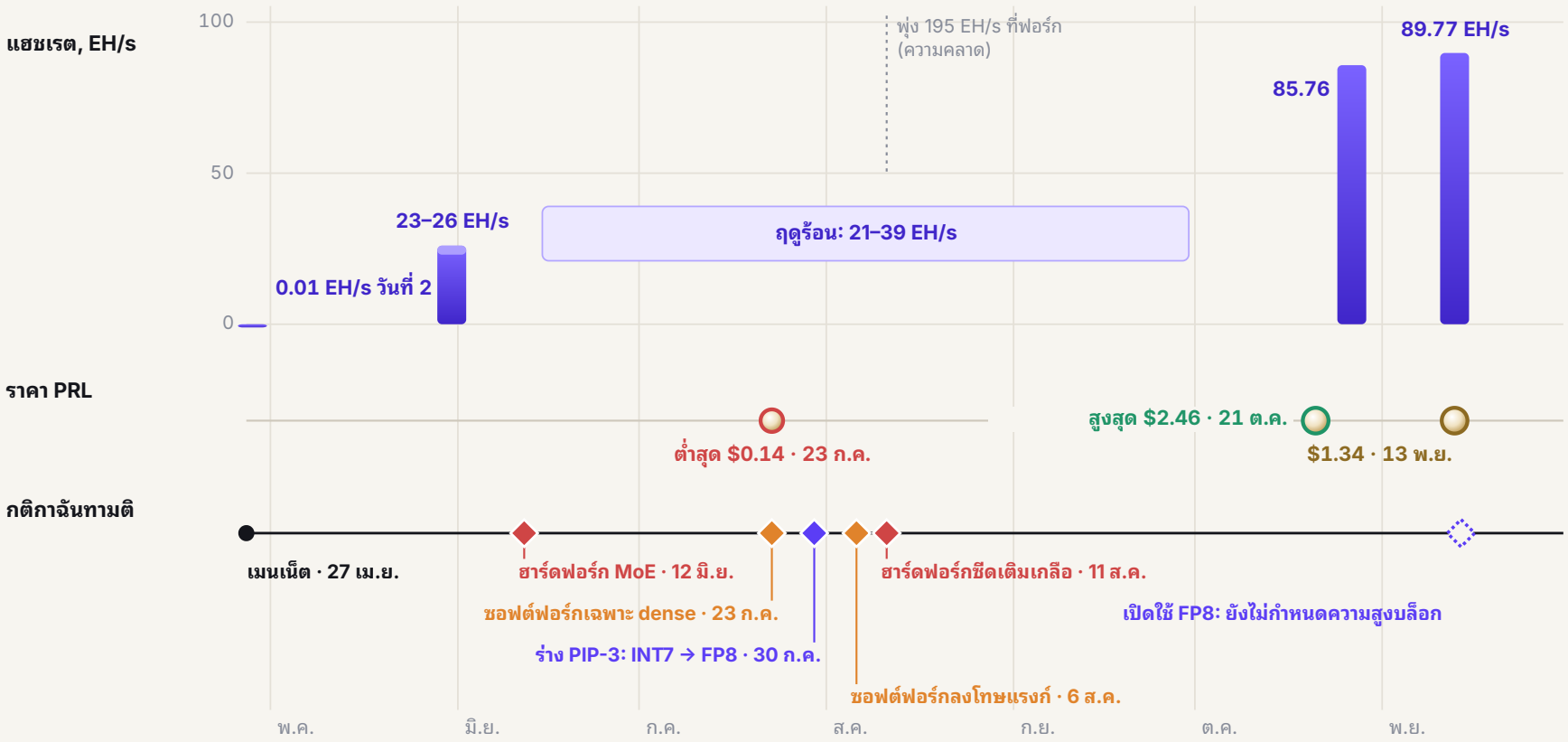
อุปทานใหม่ใน 12 เดือนข้างหน้า

บนงานเลขจำนวนเต็มของ Pearl GPU ชนเพดานพลังงานเดียวกัน



การ์ดเกมและ B300 มาบรรจบที่ราว 1.3 J/TH แม้แต่ B200 ก็ทำได้เพียง 0.71 ชิปที่ทำแค่คูณเลขจำนวนเต็ม 7 บิตจาก SRAM อาจทำได้ดีกว่า 3–10 เท่า บทวิเคราะห์

แฮชเรตตามราคา กติกาเปลี่ยนไม่หยุด



เหรียญราว 120 ล้านเหรียญ หรือ 32% ของอุปทานปัจจุบัน ถูกขุดในหกวันแรก RTX 5090 ตอนนี้ทำรายได้ราว 5–6 ดอลลาร์ต่อวัน การเช่าการ์ดมาขุดอย่างเดียวขาดทุน

ทฤษฎีบทใช้ได้ แต่แรงจูงใจใช้ไม่ได้

งานที่มีประโยชน์มีต้นทุนเท่าไร

ฝั่งการอนุมาน: โอเวอร์เฮดน้อย

Pearl รายงานว่า Llama 70B บน H200 สี่ใบใช้เวลาเพิ่ม 5.08% การทดสอบอิสระวัดได้ว่าปริมาณงานลดลง 9.2% บน GPU Hopper หนึ่งใบ

งานที่มีประโยชน์ได้ผลตอบแทนเท่าไร

ฝั่งการขุด: แฮชเรตน้อยเกินไป

การ์ดที่ให้บริการอนุมานขุดได้ราว 27% ของเครื่องขุดเฉพาะทางในช่วง prefill, 5.4% กับงานแชท และราว 1% เมื่อ decode เป็นหลัก

นักขุดที่มีเหตุผลจึงคูณเมทริกซ์สุ่ม ประโยชน์เป็นปัญหาเชิงพาณิชย์นอกเชน: ต้องมีคนจ่ายเงินซื้อการอนุมาน

สี่ฟอร์กในสองเดือน และต่อไปคือ FP8

12 มิถุนายน
ฮาร์ดฟอร์ก mixture-of-experts (บล็อก 71,935)

23 กรกฎาคม
ซอฟต์ฟอร์กเฉพาะ dense (บล็อก 91,630)

6 สิงหาคม
ซอฟต์ฟอร์กลงโทษแรงก์: บางการตั้งค่าเสียแฮชเรตที่มีผล 50–87.5%

11 สิงหาคม
ฮาร์ดฟอร์กซีดเติมเกลือ: เครื่องขุดเก่าผลิตได้แต่แชร์ที่ไม่ถูกต้อง

ยังไม่กำหนด
PIP-3: INT7 เป็น FP8 รับเฉพาะ Hopper และ Blackwell

2 ใน 3

มีประโยชน์ต่อ AI ตรวจสอบได้ทุกบิต เป็นกลางต่อฮาร์ดแวร์: รูปแบบตัวเลขได้อย่างมากสองข้อ

13 บิต

ที่การสะสม FP8 ของ Hopper เก็บไว้ พร้อมการตัดทิ้ง: ทำซ้ำข้ามรุ่นไม่ได้

≈82%

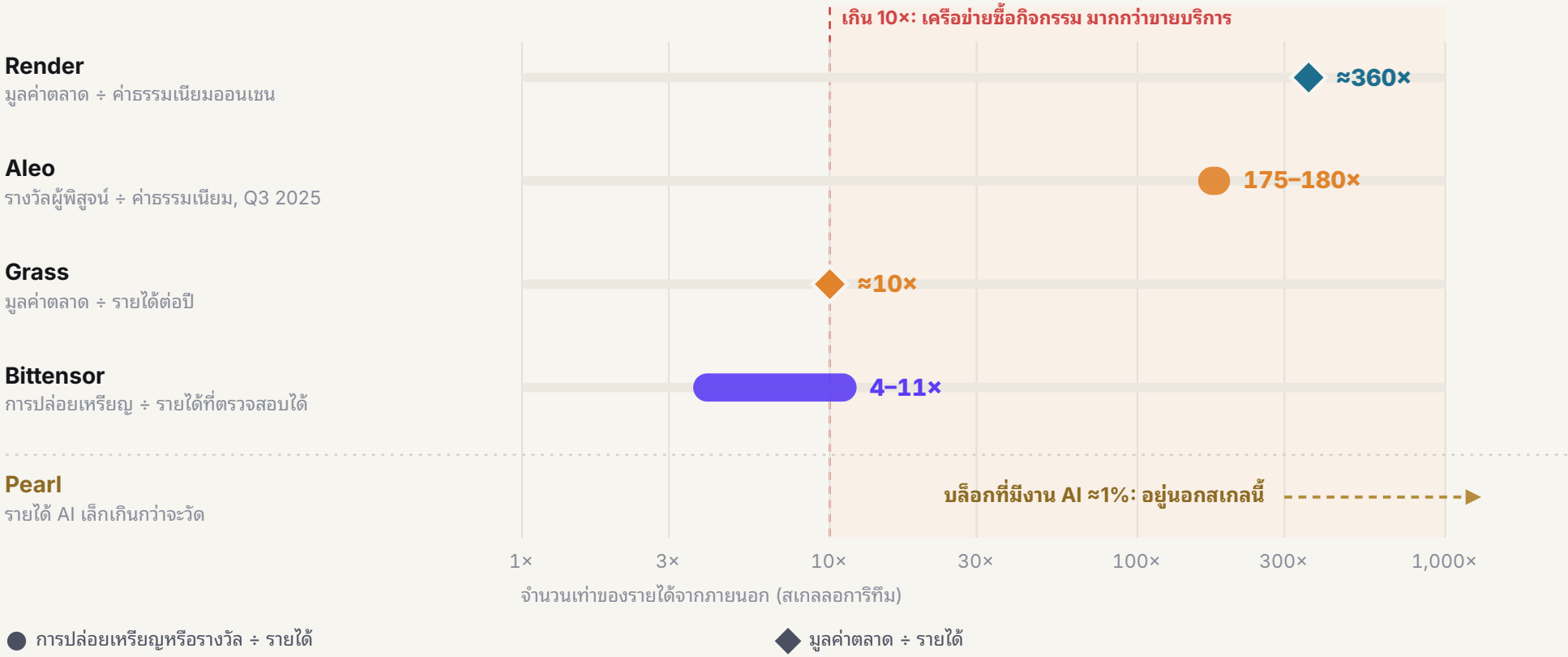
ของแฮชเรตอยู่บนการ์ดผู้บริโภคที่ PIP-3 จะไม่รับ (แบบจำลองความเชื่อมั่นต่ำ)

ใครคุมฟังก์ชันงาน คนนั้นกำหนดชะตาของนักลงทุนฮาร์ดแวร์ บทวิเคราะห์

ทั้งเมทริกซ์และ zero knowledge ต่างหยุด ASIC ไม่ได้

เครือข่าย	เกิดอะไรขึ้น	บทเรียน
Kaspa	ผลคูณเมทริกซ์ 64 × 64 คั่นระหว่างแฮช; มี ASIC หลัง 16 เดือน ประสิทธิภาพราว 1,000× ของ GPU; เครื่องที่ดีที่สุดตอนนี้ได้ 0.86 ดอลลาร์ต่อวัน	เมทริกซ์ไม่ได้มีประโยชน์และไม่กัน ASIC
Aleo	ปริศนา SNARK สังเคราะห์; มีเครื่องเฉพาะทางหลังห้าเดือน; รางวัลผู้พิสูจน์ 175–180× ของค่าธรรมเนียม	งานสังเคราะห์อยู่ได้ด้วยเงินอุดหนุน
Qubic	อ้างว่าเทรน AI; พลังประมวลผลของมันขุด Monero (อ้าง 51%) และตอนนี้ขุด Dogecoin	PoW อเนกประสงค์ถูกยึดจากภายนอกได้
Bittensor	ให้คะแนนถ่วงด้วยสเตก; รายได้ที่ตรวจสอบได้ราว 70 ล้านดอลลาร์ต่อปี เทียบการปล่อยเหรียญกว่า 300 ล้านดอลลาร์	รายเดียวที่มีรายได้จากภายนอกในระดับใหญ่
Monero, Sia, Ethereum	ฟอร์กทำให้ ASIC ใช้การไม่ได้; หลัง Merge นักขุดไม่ถึง 20% หางานใหม่ได้	ฮาร์ดแวร์ที่มีตลาดที่สองอยู่รอด

ตัวคูณเงินอุดหนุนคือการตรวจสุขภาพที่ซื่อตรงที่สุด



เมื่อเกินราว 10 เท่าของรายได้จากภายนอก เครือข่ายกำลังซื้อกิจกรรม มากกว่าขายบริการ รายได้ AI ของ Pearl เล็กเกินกว่าจะวางบนสเกลได้

อุปสงค์จริงแต่เล็ก และมูลค่ากำลังย้ายไปที่หุ้น

- **ตลาดการประมวลผล**

  ค่าธรรมเนียมออนเชนของ Render ราว 2.35 ล้านดอลลาร์ต่อปี; H100 ของ Akash ที่ 2.73 ดอลลาร์ต่อชั่วโมงถูกกว่า AWS และ Google แต่ไม่ถูกกว่า Vast.ai หรือ RunPod

- **การเทรน**

  เทรนโมเดล 8–40 พันล้านพารามิเตอร์ผ่านอินเทอร์เน็ต; INTELLECT-3 กลับไปใช้ H200 512 ใบบน InfiniBand

- **AI ที่ตรวจสอบได้**

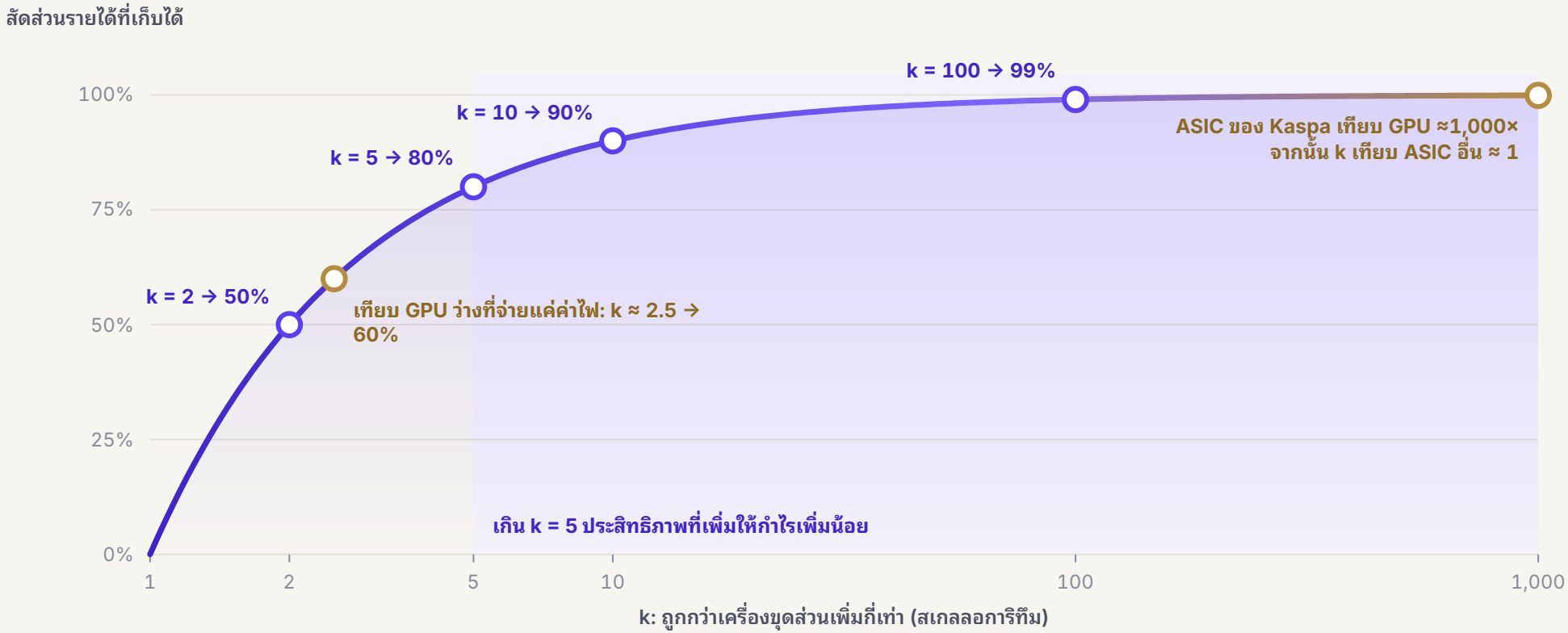
  TEE เพิ่มภาระไม่ถึง 7% และใช้งานจริงแล้ว; TEE.fail เจาะ attestation ได้ด้วย interposer ราคา 1,000 ดอลลาร์; zkML พิสูจน์ GPT-2 ได้ 174 โทเคนต่อนาที

- **เอเจนต์และข้อมูล**

  กระแสเงิน x402 ที่เอเจนต์เริ่มเองราว 5,000–11,000 ดอลลาร์ต่อเดือน; Grass มีรายได้ข้อมูล 17 ล้านดอลลาร์ในครึ่งแรกของปี 2026

> **บริษัทเอกชนสี่รายบนโครงสร้างพื้นฐานแบบรวมศูนย์มีมูลค่าราว 8.5 พันล้านดอลลาร์ ใกล้เคียงกับโทเคน AI ทั้งหมดในชั้นโปรโตคอลและชั้นแอปพลิเคชันรวมกัน**

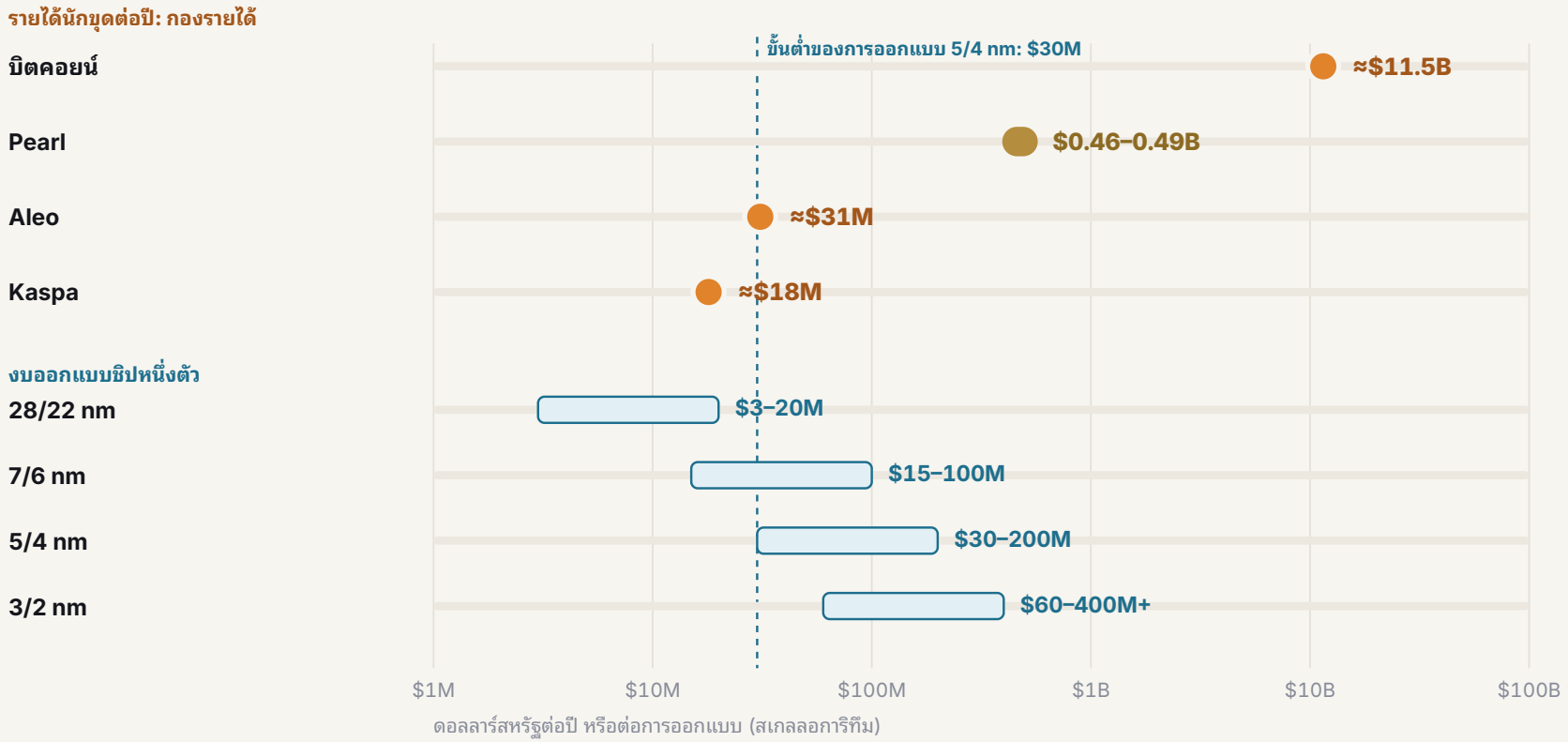
ทำไมประสิทธิภาพเลิกให้ผลหลัง k = 5



เพดานกำไร ≈ R × (1 − 1/k) เมื่อ ASIC กลายเป็นเครื่องขุดส่วนเพิ่ม k จะเริ่มนับใหม่เทียบกับ ASIC อื่น; เมื่อเทียบกับ GPU ว่างที่จ่ายแค่ค่าไฟ k อยู่ที่ราว 2.5

ประมาณการ

กองรายได้ส่วนใหญ่เล็กกว่างบออกแบบชิปหนึ่งตัว



แฮชเรตคืนทุน h* = NRE ÷ [R × (1 − 1/k) × T × p] สำหรับ Pearl ที่ NRE 20 ล้านดอลลาร์และ k = 5: ราว 4.9 EH/s เป็นเวลาหนึ่งปีถ้ากติกาคงอยู่ และ 9.8 EH/s ถ้าโอกาสที่กติกาคงอยู่เป็นครึ่งต่อครึ่ง ประมาณการ

ชิปห้าประเภท ห้าคำตอบ



ชิปขุด PoUW แบบ Pearl

ยังไม่ใช่ตอนนี้

k: ไม่มี HBM หรือหน่วยทั่วไป: 3–10× ต่อออป (อนุมาน)

ตลาดที่สอง: การอนุมาน AI ถ้าเป็น GEMM INT8/FP8 ทั่วไป

เดิมพันกับชิปที่เน้นการอนุมานและขุดได้ ไม่ใช่ชิป INT7 เฉพาะทาง



ชิปขุดบิตคอยน์ SHA-256

อย่าเข้า

k: ≤1 เมื่อเทียบผู้เล่นเดิม

ตลาดที่สอง: แทบไม่มี

เฉพาะเมื่อมีไฟฟ้าของตัวเอง รูปแบบเฉพาะกลุ่ม หรือขายเป็น IP



ชิปเทรน AI

ทำ IP ไม่ทำชิป

k: 1.1–1.8× ที่ความแม่นยำเท่ากัน

ตลาดที่สอง: —

แพ็กเกจและ HBM เป็นตัวตัดสิน; ทำ IP ชิปเล็ต และบริการ



ชิปอนุมาน AI

สร้างได้

k: แลกความเป็นสากล: อ้าง 10–20×

ตลาดที่สอง: นี่คือตลาดหลัก

เซ็นสัญญาลูกค้าก่อน แล้วค่อยเลือกการผูกขาด



ตัวเร่งการพิสูจน์ ZK

รอและป้องกันความเสี่ยง

k: อ้าง 10–100× ของ GPU

ตลาดที่สอง: จำกัดนอก ZK

เอนจิน MSM และ NTT ที่โปรแกรมได้; เริ่มจาก FPGA



ตัวเร่ง FHE

วางตำแหน่งแต่เนิ่น

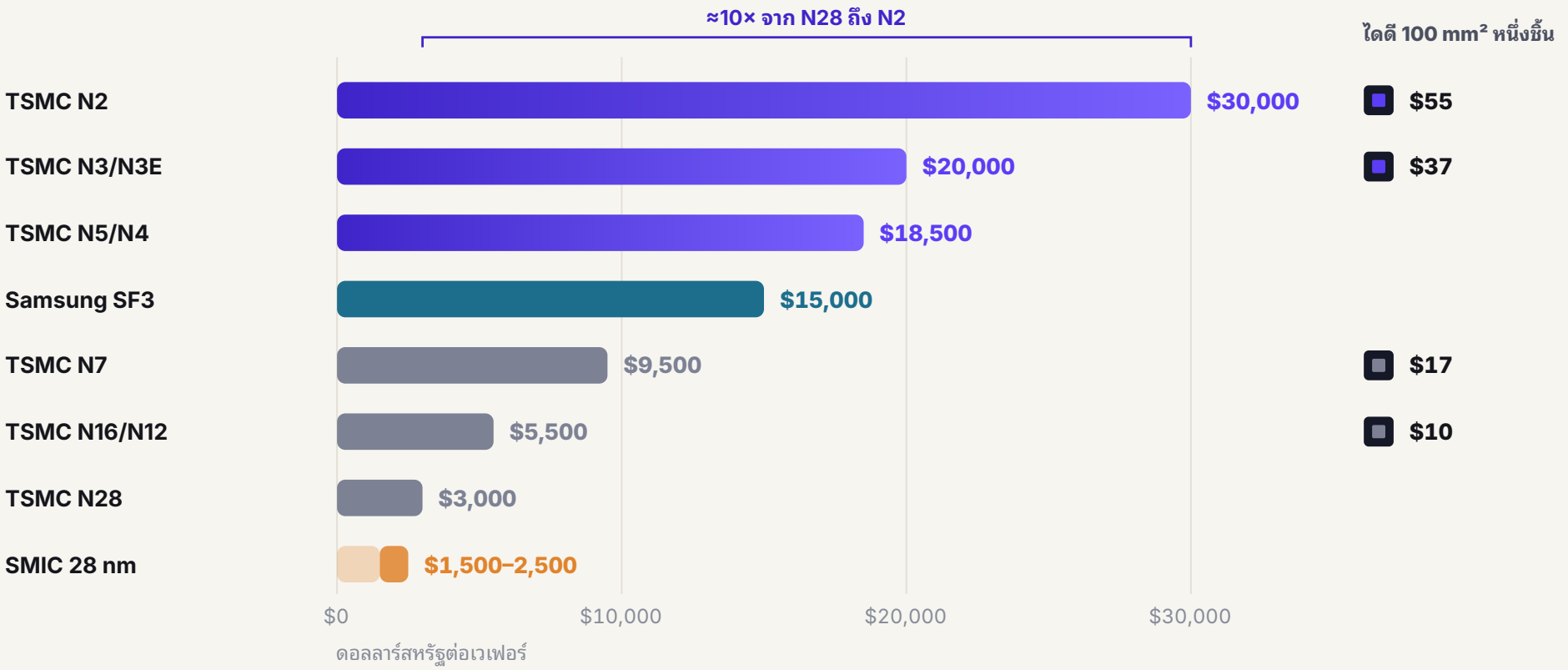
k: 1,074–5,547× ของ Xeon (Heracles)

ตลาดที่สอง: การประมวลผลแบบเข้ารหัส AI ส่วนตัว

หน้าต่างโอกาส 2027–2029; ตรวจสอบด้วยสินทรัพย์เบา

วันนี้ไม่มีเครือข่ายใดมีกองรายได้ใหญ่ กติกามั่นคง และ k มากกว่า 1 พร้อมกัน: k ของบิตคอยน์เป็นของผู้เล่นเดิม กติกาของ Pearl มั่นคงน้อยที่สุด กองรายได้ของ Kaspa และ Aleo เล็กเกินไป

เวเฟอร์โหนดล้ำสมัยแพงกว่าโหนดรุ่นเก่าสิบเท่า



ให้ต้นทุนตลอดอายุที่ใหญ่ที่สุดเป็นตัวเลือก: พลังงานสำหรับชิปประมวลผลที่ทำงานตลอดเวลา แพ็กเกจและ HBM สำหรับชิปเทรน (63% ของต้นทุน H100) การเลือก SRAM หรือ HBM สำหรับชิปอนุมาน

โหนดสำหรับชิปแต่ละแบบและทีมแต่ละแบบ

ชิป	ทีมเล็ก	บริษัทที่ขยายขนาดแล้ว	เส้นทางจีนแผ่นดินใหญ่
ชิปขุด PoUW	MPW 28 หรือ 16/12 nm; ผลิตจำนวนมากครั้งแรกที่ N7/N6	N3/N3E หรือ Samsung SF3	SMIC N+1/N+2 ใช้ได้ แพงกว่าราว 40–50% ต่อได
ชิปขุด SHA-256	อย่าไล่ตามโหนดแนวหน้า	ระดับ 3 nm	ยากจะเทียบประสิทธิภาพของโหนดล้ำสมัย
ชิปเทรน AI	IP ชิปเล็ต บริการออกแบบ	N3P แล้วตามด้วย N2/A16 พร้อม CoWoS และ HBM	SMIC N+2/N+3 กับ HBM ในประเทศ
ชิปอนุมาน AI	เน้น SRAM ที่ N6/N7 หรือ N4/N5/SF4	N3/N2 กับ HBM หรือชิปเล็ต SRAM	SMIC N+2 กับ LPDDR หรือ GDDR
ตัวเร่ง ZK และ FHE	เริ่มจาก FPGA หรือ MPW 16/12 nm	N5/N3; FHE กับ HBM	ZK ทำได้; FHE ถูกจำกัดด้วย HBM

งบแบบเป็นขั้น

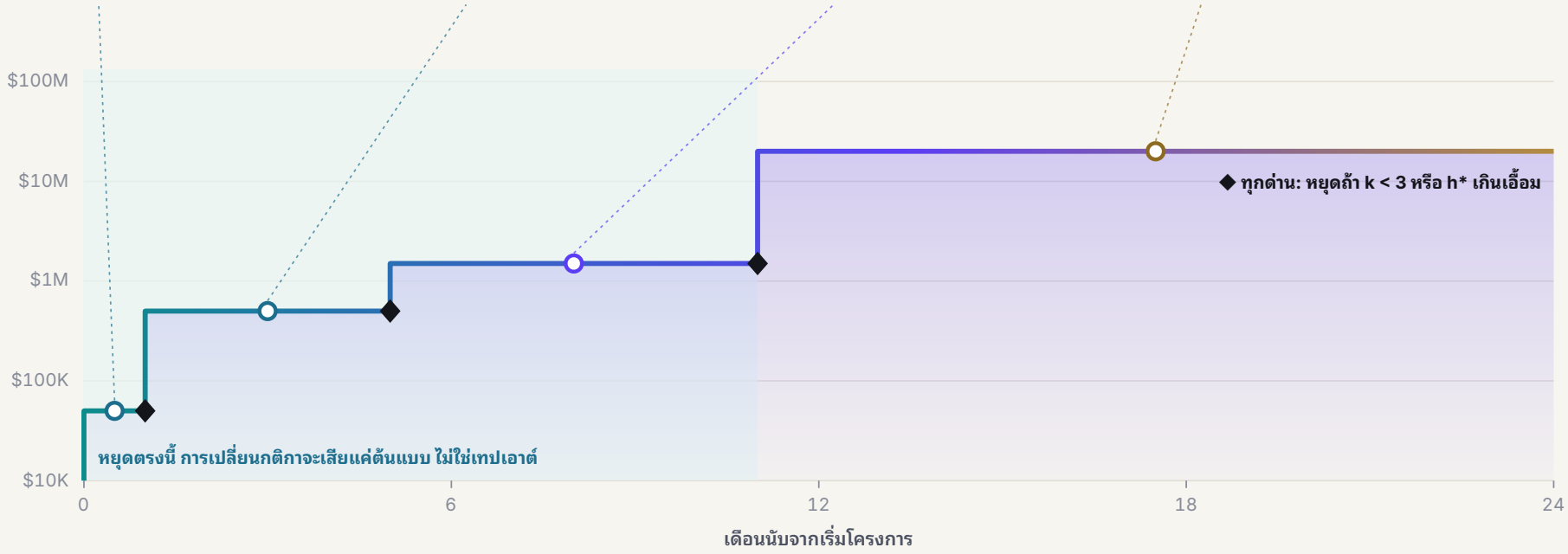
ซื้อหลักฐานก่อนซื้อหน้ากาก

ด่าน 0 · แบบจำลอง
≈$50K · หลักสัปดาห์
J/op และ $/op จากเลขคณิต SRAM และ I/O

ด่าน 1 · FPGA
≈$0.5M · หลักเดือน
งานจริงแบบ end to end: แชร์หรือโทเคน

ด่าน 2 · MPW
≈$1.5M · เดือนที่ 11
ซิลิคอนที่วัดได้จริงแทนที่แบบจำลอง

ด่าน 3 · หน้ากากและล็อตแรก
≈$20M · เดือนที่ 24
หน้ากาก N7 ≈$3.5M ในงานออกแบบ $15–100M



แบบจำลอง FPGA และ MPW พาไปถึงซิลิคอนที่วัดผลแล้วด้วยเงินราว 1.5 ล้านดอลลาร์ในสิบเอ็ดเดือน ส่วนชุดหน้ากากและล็อตแรกแพงกว่านั้นกว่าสิบเท่า ให้หยุดก่อนซื้อหน้ากากจนกว่าจะได้ผู้ซื้อหรือเวอร์ชันโปรโตคอลที่ตายตัว

คริปโตเป็นลูกค้ารายแรก AI เป็นตลาดที่สอง

ชิปสั่งทำ

- อุปสงค์มาจาก AI อย่างท่วมท้น: Broadcom มีเซมิคอนดักเตอร์ AI ราว 58 พันล้านดอลลาร์ในปีงบประมาณ 2026
- ทักษะชิปขุดใช้ต่อได้: อาร์เรย์แรงดันต่ำ การจ่ายไฟระดับกิโลวัตต์ เทปเอาต์ที่รวดเร็ว
- ชนะด้วยช่องทางสู่ตลาดที่เป็นอิสระ ไม่ใช่ช่องทางการขุด

การพัฒนา AI

- ชั่วโมงทำงานเป็นสินค้าโภคภัณฑ์: 25–49 ดอลลาร์ต่อชั่วโมงบน Clutch
- มูลค่าอยู่ที่รอยต่อ: เคอร์เนลสำหรับชิปสั่งทำ การขุดภายในบริการอนุมาน
- การอนุมานที่ตรวจสอบได้: TEE บวกการสุ่มตรวจด้วย ZK

ระบบ AI × คริปโต

- การตรวจสอบเป็นชั้น: TEE หรือการสุ่มตัวอย่าง แล้วจึงการโต้แย้งหรือการสุ่มตรวจด้วย ZK
- อุปสรรคคือการกระจายและสภาพคล่อง ไม่ใช่ต้นทุน
- เซ็นสัญญาผู้ซื้อพลังประมวลผลก่อนออกโทเคน

สามเส้นทาง แต่ละทางเริ่มด้วยก้าวเล็กๆ

เส้นทาง	ก้าวแรก
A. ตัวเร่งที่เน้นการอนุมานและขุดได้ด้วย	ศึกษา J/op และ $/op สี่สัปดาห์ของอาร์เรย์ INT8/FP8 บน SRAM เทียบ RTX 5090 และ H100 จากนั้น FPGA หรือ MPW บนโหนดรุ่นเก่า
B. บริการซิลิคอนสั่งทำ: คริปโตก่อน AI ทีหลัง	เปลี่ยนแบบจำลอง NRE และจุดคุ้มทุนของรายงานนี้เป็นการศึกษาความเป็นไปได้แบบเก็บเงิน มีหมุดหมายและล็อกเวอร์ชันโปรโตคอล
C. เครือข่ายที่เกิดมาเพื่อ AI โดยมีผู้ซื้อก่อนโทเคน	บันทึกการออกแบบว่าด้วยการตรวจสอบเป็นชั้นและเลขคณิตมาตรฐานที่เป็นกลางต่อผู้ขาย เขียนร่วมกับผู้ซื้อการอนุมานที่ระบุชื่อได้หนึ่งราย

ไม่เกินสามเส้นทาง และทดสอบแต่ละทางก่อนลงเงิน บทวิเคราะห์

สิ่งที่ต้องจับตา และคำถามที่ควรถาม

สัญญาณที่ต้องจับตา

เมื่อใดก็ได้	ความสูงบล็อกที่เปิดใช้ FP8 ของ Pearl และเชนจะแยกหรือไม่
รายเดือน	สัดส่วนบล็อกของ Pearl ที่มีงาน AI
รายไตรมาส	รายได้ที่ตรวจสอบได้ของ Bittensor เทียบการปล่อยเหรียญ; กระแส x402 ที่เอเจนต์เริ่มเอง
มกราคม 2027	TSMC ขึ้นราคา 2 nm และ 3 nm อีก 3–6% ตามที่มีรายงาน
Q1 2027	FP4 ใน Pearl ตามที่รายงานจาก AMA ของผู้ก่อตั้ง
เมษายน 2028	การลดรางวัลครึ่งหนึ่งครั้งถัดไปของบิตคอยน์

คำถามที่ควรถาม

01 ใครคุมฟังก์ชันงาน?

02 ตัวคูณเงินอุดหนุนเป็นเท่าไร?

03 ด่านถัดไปมีต้นทุนเท่าไร และอะไรจะทำให้คุณหยุด?

04 แฮชเรตคืนทุนเป็นเท่าไร เมื่อรวม p แล้ว?

05 ตลาดที่สองของชิปคืออะไร?

06 ลูกค้าที่จ่ายเงินรายแรกคือใคร ระบุชื่อ?

นี่คือด้านที่เราสำรวจร่วมกันได้ ผ่านการแลกเปลี่ยนและความร่วมมือทางเทคนิค: ชิปสั่งทำ การพัฒนา AI และระบบ AI × คริปโต

i@tanshuai.com · tanshuai.com