บันทึกการวิจัย · ฉบับเผยแพร่สาธารณะ

2026 · 11



การพิสูจน์งานที่มีประโยชน์ · คริปโต × ซิลิคอน

การประมวลผลคือฉันทามติ

AI × คริปโต × ชิป ในปี 2026 เริ่มจากการพิสูจน์งานที่มีประโยชน์ของ Pearl

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

ข้อสรุป

Pearl พิสูจน์แล้วว่าการคูณเมทริกซ์ใช้รักษาความปลอดภัยของบล็อกเชนได้ แต่ยังไม่ได้พิสูจน์ว่างานนั้นมีประโยชน์ ในภาพรวมของ AI × คริปโต อุปสงค์ที่จ่ายเงินจริงอยู่ที่การอนุมาน ข้อมูล และความเป็นส่วนตัว และ ASIC จะคุ้มทุนก็ต่อเมื่อกองรายได้ใหญ่ กติกามั่นคง และชิปมีตลาดที่สอง

ทัน ชว่าย

นักเทคโนโลยีและผู้ก่อตั้ง · นักวิจัยและนักพัฒนา AI

tanshuai.com · i@tanshuai.com

ข้อมูล ณ วันที่ 13 พฤศจิกายน 2026

จำนวนเงินเป็นดอลลาร์สหรัฐ

สารบัญ

สารบัญ



◆ **บทสรุปผู้บริหาร** 04
ทฤษฎีบทใช้ได้ แต่แรงจูงใจใช้ไม่ได้

01 **ขอบเขตและวิธีการ** 07
ขอบเขต วิธีการ และการทดสอบสี่ข้อ

02 **Pearl · กลไก** 08
Pearl: การคูณเมทริกซ์ในฐานะฉันทามติ

03 **Pearl · เศรษฐศาสตร์** 10
Pearl: ขนาด การปล่อยเหรียญ และช่องว่างด้านประโยชน์

04 **Pearl · กติกา** 12
Pearl: ความเสี่ยงด้านกติกา ระบบตัวเลข และธรรมาภิบาล

05 **เชนงานที่มีประโยชน์** 13
เชนงานที่มีประโยชน์: ทั้งเมทริกซ์และ zero knowledge ต่างหยุด ASIC ไม่ได้

06 **ภาพรวมที่กว้างขึ้น** 16
ส่วนอื่นของ AI × คริปโต: อุปสงค์จริงแต่เล็ก และมูลค่ากำลังย้ายไปที่หุ้น

07 **เศรษฐศาสตร์ ASIC** 18
คำถามเรื่อง ASIC: หนึ่งสูตร สามไม้บรรทัด

08 **ชิปทีละประเภท** 21
ชิปห้าประเภท ห้าคำตอบ

09 **โหนดการผลิต** 24
โหนดการผลิต: ให้ต้นทุนตลอดอายุที่ใหญ่ที่สุดเป็นตัวเลือก

10 **งบแบบเป็นขั้น** 27
จากแบบจำลองสู่ซิลิคอน: งบแบบเป็นขั้นสำหรับชิปตัวแรก

11 **โอกาส** 29
โอกาสอยู่ที่ไหน และผมจะทำอะไร

12 **สัญญาณที่ต้องจับตา** 32
สิ่งที่ต้องจับตา พฤศจิกายน 2026 ถึง 2027

13 **คำถามที่ควรถาม** 33
คำถามที่คุณไม่ได้ถาม แต่ควรถาม

◆ **เกี่ยวกับผู้เขียน** 35
ทัน ชว่าย

◆ **แหล่งอ้างอิง** 36

ภาพประกอบ

ภาพประกอบ

ภาพที่ 01	AI × คริปโต 6 กลุ่ม: อุปสงค์ที่จ่ายเงินจริงอยู่ที่ไหน	06
ภาพที่ 02	บนงานเลขจำนวนเต็มของ Pearl GPU ชนเพดานพลังงานเดียวกัน	09
ภาพที่ 03	หกเดือนแรกของ Pearl: แฮชเรตตามราคา กติกาเปลี่ยนไม่หยุด	11
ภาพที่ 04	ตัวคูณเงินอุดหนุน: เครือข่ายจ่ายกี่เท่าของรายได้จากภายนอก	15
ภาพที่ 05	ทำไมประสิทธิภาพเลิกให้ผล: ASIC เก็บรายได้ได้ไม่เกิน 1 − 1/k	19
ภาพที่ 06	กองรายได้ส่วนใหญ่เล็กกว่างบออกแบบชิปหนึ่งตัว	20
ภาพที่ 07	ชิปห้าประเภท ห้าคำตอบ	24
ภาพที่ 08	เวเฟอร์โหนดล้ำสมัยแพงกว่าโหนดรุ่นเก่าสิบเท่า	25
ภาพที่ 09	ซื้อหลักฐานก่อนซื้อหน้ากาก: งบแบบเป็นขั้นสำหรับชิปตัวแรก	28

ป้ายกำกับในเนื้อหา

- ประมาณการ — การคำนวณหรือการอนุมานของผู้เขียนจากข้อมูลที่อ้างอิง
- ผู้ขายอ้าง — เผยแพร่โดยบริษัทที่ขายผลิตภัณฑ์นั้น ยังไม่มีการตรวจสอบโดยอิสระ
- บทวิเคราะห์ — การตีความหลักฐานของผู้เขียน

เอกสารนี้แปลจากฉบับภาษาอังกฤษ หากเนื้อหาไม่ตรงกัน ให้ยึดฉบับภาษาอังกฤษเป็นหลัก

ข้อมูล ณ วันที่ 13 พฤศจิกายน 2026 · จำนวนเงินเป็นดอลลาร์สหรัฐ

บทสรุปผู้บริหาร

ทฤษฎีบทใช้ได้ แต่แรงจูงใจใช้ไม่ได้

ข้อสรุป

Pearl แสดงให้เห็นว่า “การประมวลผลคือฉันทามติ” ใช้ได้ในทางวิศวกรรม แต่ยังไม่ได้แสดงว่ามันมีประโยชน์

≈90 EH/s
เครือข่าย Pearl ณ วันที่ 13 พฤศจิกายน 2026 โดย 1 H เท่ากับการคูณสะสม INT7 หนึ่งครั้ง

≈1%
ของบล็อก Pearl ที่มีงาน AI จริง (Blockworks, 28 ตุลาคม 2026)

≈$70M
ต่อปี คือรายได้จากภายนอกที่ตรวจสอบได้ของ Bittensor เทียบกับการปล่อยเหรียญกว่า 300 ล้านดอลลาร์

175–180×
รางวัลผู้พิสูจน์ของ Aleo เทียบกับค่าธรรมเนียมในไตรมาสที่ 3 ปี 2025

R × (1 – 1/k)
เพดานกำไรของ ASIC ที่ถูกกว่าเครื่องขุดส่วนเพิ่ม k เท่า ประมาณการ

≈$1.5M
เพื่อให้ได้ซิลิคอนที่วัดผลแล้วบน N7 ผ่านแบบจำลอง FPGA และ MPW ก่อนชุดหน้ากากราว 3.5 ล้านดอลลาร์ ประมาณการ

Pearl

ณ วันที่ 13 พฤศจิกายน 2026 เครือข่าย Pearl ทำงานที่ราว 90 EH/s ในโค้ดฉันทามติ “แฮช” หนึ่งครั้งคือการคูณสะสมเลขจำนวนเต็มหนึ่งครั้ง เครือข่ายจึงเท่ากับปริมาณงานเลขจำนวนเต็มของ H100 ราว 122,000 ใบ แต่มีเพียงราว 1% ของบล็อกที่มีงาน AI จริง และความร่วมมือด้านการอนุมานที่เปิดเผยต่อสาธารณะเพียงรายเดียวก็ไม่มีให้บริการบน serverless API ของผู้ให้บริการแล้ว สาเหตุเป็นเรื่องเศรษฐศาสตร์ ไม่ใช่วิทยาการเข้ารหัส: GPU ที่ขุดไปพร้อมกับให้บริการอนุมานได้แฮชเรตเพียงราว 1% ถึง 27% ของเครื่องขุดเฉพาะทาง นักขุดที่มีเหตุผลจึงคูณเมทริกซ์สุ่ม

อุปสงค์ที่จ่ายเงินจริง

ในภาพรวมของ AI × คริปโต อุปสงค์ที่จ่ายเงินจริงซึ่งตรวจสอบได้ปรากฏในสามที่: การขายการอนุมาน การขายข้อมูล และการขายความเป็นส่วนตัว รายได้จากภายนอกที่ตรวจสอบได้ของ Bittensor อยู่ที่ราว 70 ล้านดอลลาร์ต่อปี ราวสี่ในห้ามาจากซับเน็ตด้านการประมวลผลและโครงสร้างพื้นฐาน Grass รายงานรายได้ข้อมูล 17 ล้านดอลลาร์ในครึ่งแรกของปี 2026 การอนุมานแบบเป็นความลับในสภาพแวดล้อมการประมวลผลที่เชื่อถือได้ (TEE) ใช้งานจริงแล้ว ในทางกลับกัน การชำระเงินที่เอเจนต์ AI เป็นผู้เริ่มจริงอยู่ที่ราว 5,000–11,000 ดอลลาร์ต่อเดือน และ zkML กับการเทรนแบบกระจายศูนย์ยังไม่มีการเปิดเผยอุปสงค์ที่จ่ายเงิน

เงินอุดหนุนและมูลค่า

ตัวคูณเงินอุดหนุนคือการตรวจสุขภาพที่ซื่อตรงที่สุด รางวัลผู้พิสูจน์ของ Aleo อยู่ที่ราว 175–180 เท่าของค่าธรรมเนียม การปล่อยเหรียญของ Bittensor ราว 4–11 เท่าของรายได้ที่ตรวจสอบได้ มูลค่าตลาดของ Render ราว 360 เท่าของค่าธรรมเนียมออนเชน และรายได้นักขุดของ Pearl ราว 460 ล้านดอลลาร์ต่อปี ส่วนใหญ่เป็นเงินอุดหนุนบล็อก แทบไม่เกี่ยวกับอุปสงค์ AI ที่จ่ายเงิน เงินอุดหนุนเป็นเชื้อเพลิงสำหรับการเริ่มต้น ไม่ใช่โมเดลธุรกิจ ขณะเดียวกัน มูลค่ากำลังย้ายจากโทเคนไปที่หุ้น: บริษัทเอกชนสี่รายที่ใช้โครงสร้างพื้น

ฐานแบบรวมศูนย์และช่องทางกระจายแบบคริปโตมีมูลค่าราว 8.5 พันล้านดอลลาร์ ใกล้เคียงกับโทเคน AI ในชั้นโปรโตคอลและชั้นแอปพลิเคชันรวมกัน

ASIC

สำหรับชิป เพดานกำไรของ ASIC โดยคร่าวคือรายได้ต่อหน่วยการประมวลผลคูณ (1 − 1/k) โดย k คือจำนวนเท่าที่มันถูกกว่าฮาร์ดแวร์ส่วนเพิ่ม เพราะพจน์นี้เว้า ประสิทธิภาพที่เกิน k = 5–10 แทบไม่เพิ่มกำไร สิ่งที่ตัดสินผลคือขนาดของกองรายได้ กติกาอยู่ได้นานแค่ไหน และชิปมีตลาดที่สองหรือไม่ คำตอบของผมตามประเภทชิป: อย่าเทปเอาต์ชิป INT7 เฉพาะทางสำหรับ Pearl ตอนนี้ ให้เดิมพันกับตัวเร่งที่เน้นการอนุมานและขุดได้ด้วยแทน ผู้เล่นใหม่ควรอยู่ห่างจากตลาดเครื่องขุดบิตคอยน์เพื่อขายทั่วไป ทีมเล็กไม่ควรสร้างชิปเทรนแบบแยกเดี่ยว สร้างชิปอนุมานเมื่อเซ็นสัญญาลูกค้าแล้ว ป้องกันความเสี่ยงด้วยตัวเร่ง ZK และวางตำแหน่งแต่เนิ่นใน FHE

โหนดและงบประมาณ

โหนดการผลิตควรเลือกตามรายการที่ใหญ่ที่สุดในต้นทุนตลอดอายุของชิป: พลังงานสำหรับชิปที่ทำงานตลอดเวลา แพ็กเกจและหน่วยความจำสำหรับชิปเทรน และการตัดสินใจเลือก SRAM หรือ HBM สำหรับชิปอนุมาน และชิปตัวแรกควรซื้อเป็นขั้น ได้แก่ แบบจำลอง FPGA เวเฟอร์หลายโครงการ และจากนั้นจึงเป็นชุดหน้ากาก เพื่อให้การเปลี่ยนแปลงของโปรโตคอลเสียแค่ราคาของต้นแบบ ไม่ใช่ราคาของเทปเอาต์ โอกาสอยู่ที่จุดที่สาขาเหล่านี้ตัดกัน: คริปโตเป็นลูกค้ารายแรก AI เป็นตลาดที่สอง

คำถาม	ข้อวินิจฉัย	ตัวเลขสำคัญ (แหล่งที่มาอยู่ในแต่ละบท)
① **Pearl ใช้ได้หรือไม่ และงานของมันมีประโยชน์หรือไม่?**	ใช้ได้: บล็อกเชนที่รักษาความปลอดภัยด้วยการคูณเมทริกซ์ในระดับใหญ่ แต่ยังไม่มีประโยชน์: การขุดเกือบทั้งหมดทำบนเมทริกซ์สุ่ม	ราว 90 EH/s ณ 13 พฤศจิกายน 2026; 1 H = การคูณสะสม INT7 หนึ่งครั้ง; บล็อกที่มีงาน AI ราว 1%
② **อุปสงค์ที่จ่ายเงินจริงใน AI × คริปโตอยู่ที่ไหน?**	การอนุมาน ข้อมูล และความเป็นส่วนตัว การชำระเงินของเอเจนต์ zkML และการเทรนแบบกระจายศูนย์ยังอยู่ก่อนจะมีอุปสงค์	Bittensor มีรายได้ที่ตรวจสอบได้ราว 70 ล้านดอลลาร์ต่อปี; Grass 17 ล้านดอลลาร์ใน H1 2026; การชำระเงินของเอเจนต์ราว 5,000–11,000 ดอลลาร์ต่อเดือน
③ **ASIC คุ้มค่าจะสร้างหรือไม่?**	เฉพาะเมื่อกองรายได้ใหญ่ กติกามั่นคง และชิปมีตลาดที่สอง วันนี้ไม่มีเครือข่ายใดมีครบทั้งสามข้อ	เพดานกำไร ≈ R × (1 − 1/k); รายได้นักขุดบิตคอยน์ราว 11.5 พันล้านดอลลาร์ต่อปี Pearl ราว 0.46–0.49 พันล้านดอลลาร์ Kaspa ราว 18 ล้านดอลลาร์
④ **ทีมเล็กควรสร้างชิปแบบใด?**	ตัวเร่งที่เน้นการอนุมานและขุดได้ด้วย ชิปอนุมานสำหรับลูกค้าที่เซ็นสัญญาแล้ว งาน FHE ระยะแรก ไม่ใช่เครื่องขุดบิตคอยน์เพื่อขายทั่วไปหรือชิปเทรน	ASIC ที่ดีที่สุดของ Kaspa ได้ 0.86 ดอลลาร์ต่อวัน; งบออกแบบ 3/2 nm 60–400 ล้านดอลลาร์หรือมากกว่า
⑤ **โหนดการผลิตใด?**	ต้นทุนตลอดอายุที่ใหญ่ที่สุดเป็นตัวตัดสิน: พลังงาน แพ็กเกจ และหน่วยความจำ หรือสถาปัตยกรรมหน่วยความจำ	TSMC N2 ราว 30,000 ดอลลาร์ต่อเวเฟอร์ เทียบกับราว 3,000 ดอลลาร์สำหรับ N28; HBM และ CoWoS ราว 63% ของต้นทุน H100
⑥ **ชิปตัวแรกมีต้นทุนเท่าไร และควรซื้ออย่างไร?**	เป็นสี่ขั้น: แบบจำลอง FPGA เวเฟอร์หลายโครงการ แล้วจึงหน้ากาก ให้หยุดก่อนซื้อหน้ากากหากกติกาหรือผู้ซื้อยังไม่ลงตัว	สล็อต MPW เชิงพาณิชย์ 50,000–500,000 ดอลลาร์; หน้ากาก N7 ราว 3.5 ล้านดอลลาร์; งบออกแบบ 7/6 nm 15–100 ล้านดอลลาร์

ภาพที่ 01

AI × คริปโต 6 กลุ่ม: อุปสงค์ที่จ่ายเงินจริงอยู่ที่ไหน

อุปสงค์ที่จ่ายเงินจริงและตรวจสอบได้ เทียบกับสัญญาณด้านเงินอุดหนุนและมูลค่า พฤศจิกายน 2026

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

แหล่งที่มา: DefiLlama, SubConnect (ผ่าน hoge.gg), TRM Labs, Phala, Dealroom, SiliconANGLE, CoinGecko, Blockworks; การสังเคราะห์ของผู้เขียน

บทที่ 01



01 ขอบเขตและวิธีการ

ขอบเขต วิธีการ และการทดสอบสี่ข้อ

ทำไมต้องมีรายงานนี้ และทำไมต้องตอนนี้

ในเดือนเมษายน 2026 บล็อกเชนใหม่ชื่อ Pearl นำแนวคิดที่มีมานานไปใช้จริง: ให้นักขุดแข่งกันชิงบล็อกด้วยการคูณเมทริกซ์แบบที่ AI ใช้ แทนการแฮช ภายในหกเดือน เครือข่ายแตะราว 90 EH/s และโทเคนมีมูลค่าตลาดราว 500 ล้านดอลลาร์ ในช่วงเดียวกัน ตลาดการประมวลผล AI เครือข่ายการเทรนแบบกระจายศูนย์ รางการชำระเงินของเอเจนต์ และเครือข่ายข้อมูลระดมทุนได้หลายพันล้านดอลลาร์ และนักขุดบิตคอยน์เริ่มแปลงไซต์ไฟฟ้าของตนเป็นศูนย์ข้อมูล AI ผมอยากรู้ว่าสิ่งเหล่านี้รายใดมีลูกค้าจริง ซิลิคอนสั่งทำเหมาะกับส่วนใดบ้าง และทีมชิปใช้โหนดการผลิตใดและโรงหล่อใดได้จริง

การทดสอบสี่ข้อสำหรับโครงการ AI × คริปโตใดๆ

งานศึกษาด้านความปลอดภัยของ Gridcoin ในปี 2017 เสนอว่าการประมวลผลที่มีประโยชน์จะรักษาความปลอดภัยของบล็อกเชนได้ก็ต่อเมื่อแก้ยาก ตรวจสอบง่าย และให้รางวัลง่าย ผมเพิ่มการทดสอบข้อที่สี่ ทั้งสี่ข้อรวมกันตัดสินเนื้อหาส่วนใหญ่ที่ตามมา

01 **แก้ยาก** การประมวลผลอย่างซื่อตรงต้องเป็นวิธีที่ถูกที่สุดในการสร้างการพิสูจน์ที่ถูกต้อง ต้องไม่มีทางลัด

02 **ตรวจสอบง่าย** การตรวจงานต้องถูกกว่าการทำงานมาก มิฉะนั้นเครือข่ายต้องจ่ายสองครั้ง

03 **ให้รางวัลง่าย** โปรโตคอลต้องคำนวณรางวัลได้จากการพิสูจน์เพียงอย่างเดียว โดยไม่ต้องเชื่อใจบุคคลที่สาม

04 **มีคนจ่ายเงินซื้อผลลัพธ์** หากไม่มีลูกค้านอกเครือข่ายจ่ายเงินซื้อผลลัพธ์ งานที่ “มีประโยชน์” ก็เป็นเพียงเงินอุดหนุนที่มีเรื่องเล่าประกอบ

ปริศนาแฮชผ่านการทดสอบสามข้อแรกแต่ตกข้อที่สี่ งาน AI จริงผ่านข้อที่สี่แต่ตกข้อที่สอง: การพิสูจน์การประมวลผล AI แบบ zero knowledge มีต้นทุน 100 ถึง 10,000 เท่าของการประมวลผลเอง และมีการแสดงแล้วว่ากลไก proof-of-learning ถูกปลอมได้ในราคาถูก การคูณเมทริกซ์แบบเติมสัญญาณรบกวนของ Pearl เป็นสะพานแรกที่น่าเชื่อถือซึ่งผ่านการทดสอบด้านการตรวจสอบ ขณะที่ยังคงรูปแบบเดียวกับงานประมวลผล AI

ผมทำงานอย่างไร

นี่คือการวิจัยเอกสาร ทำแบบเดียวกับที่ผมจะทบทวนชิปหรือแผนธุรกิจ ผมอ่านโค้ด ข้อเสนอโปรโตคอล และไวต์เปเปอร์ของ Pearl ข้อมูลบนเชนจากฟูลโหนดอิสระ เอกสารที่บริษัทยื่นต่อหน่วยงานและประกาศผลประกอบการ งานวิจัยด้านวิทยาการเข้ารหัส สถาปัตยกรรมคอมพิวเตอร์ และการเรียนรู้ของเครื่อง ข้อมูลราคาจากโรงหล่อและผู้ขาย รวมถึงสื่อด้านการค้าและเทคโนโลยี ตัวเลขทุกตัวระบุวันที่ ตัวเลขที่ผมคำนวณเองจะระบุไว้และติดป้าย ประมาณการ ตัวเลขที่มีเพียงผู้ขายรายงานติดป้าย ผู้ขายอ้าง การตีความหลักฐานของผมติดป้าย บทวิเคราะห์ จำนวนเงินเป็นดอลลาร์สหรัฐ สถานะเป็นข้อมูล ณ วันที่ 13 พฤศจิกายน 2026

บทที่ 02

02 Pearl · กลไก
Pearl: การคูณเมทริกซ์ในฐานะฉันทามติ

แนวคิด จากทฤษฎีบทสู่เมนเน็ต

ในเดือนเมษายน 2025 Ilan Komargodski และ Omri Weinstein เผยแพร่งาน “Proofs of Useful Work from Arbitrary Matrix Multiplication” ทั้งสองสร้างการพิสูจน์งานที่มีประโยชน์สำหรับผลคูณเมทริกซ์ใดๆ ด้วยโอเวอร์เฮดเพียง 1 + o(1) เท่าของการคูณแบบตรงไปตรงมา ความปลอดภัยของมันอาศัยความยากในการแก้ชุดสมการเชิงเส้นแรงก์ต่ำแบบสุ่ม ซึ่งเป็นข้อคาดการณ์ ไม่ใช่สิ่งที่พิสูจน์แล้ว งานนี้ระบุความยากที่แท้จริงของสาขาไว้: การพิสูจน์งานที่มีประโยชน์ซึ่งมีโอเวอร์เฮดเป็นค่าคงที่นั้นทำได้ง่ายแต่ไร้ประโยชน์ ส่วนที่ยากคือการรักษาให้การประมวลผลอย่างซื่อตรงเป็นวิธีที่ถูกที่สุดในการชนะ เมื่อนักขุดเลือกข้อมูลนำเข้าเอง

เมนเน็ตของ Pearl เริ่มเมื่อวันที่ 27 เมษายน 2026 เวลา 09:00 UTC นักขุดคำนวณผลคูณที่เติมสัญญาณรบกวน (A + E)(B + F) โดยสัญญาณรบกวนแรงก์ต่ำได้มาจากคอมมิตเมนต์ BLAKE3 แล้วจึงลอกสัญญาณรบกวนออกเพื่อได้ A·B คืนมา สถานะของตัวสะสมของแต่ละไทล์ผลลัพธ์ถูกแฮชเป็นสลากลุ้นรางวัล ไทล์ที่ชนะจะถูกเปิดและพิสูจน์ด้วยการพิสูจน์แบบ zero knowledge ของ Plonky2 แบบเรียกซ้ำ ขนาดไม่ถึง 60 KB ภายในใบรับรองบล็อกที่จำกัดไว้ที่ 65 KB การตรวจสอบคือการพิสูจน์ขนาดเล็กที่มีขนาดคงที่ ไม่ขึ้นกับขนาดเมทริกซ์: ตรวจได้ถูกพอๆ กับปริศนาแฮช ขณะที่การประมวลผลเองอาจมีประโยชน์ได้

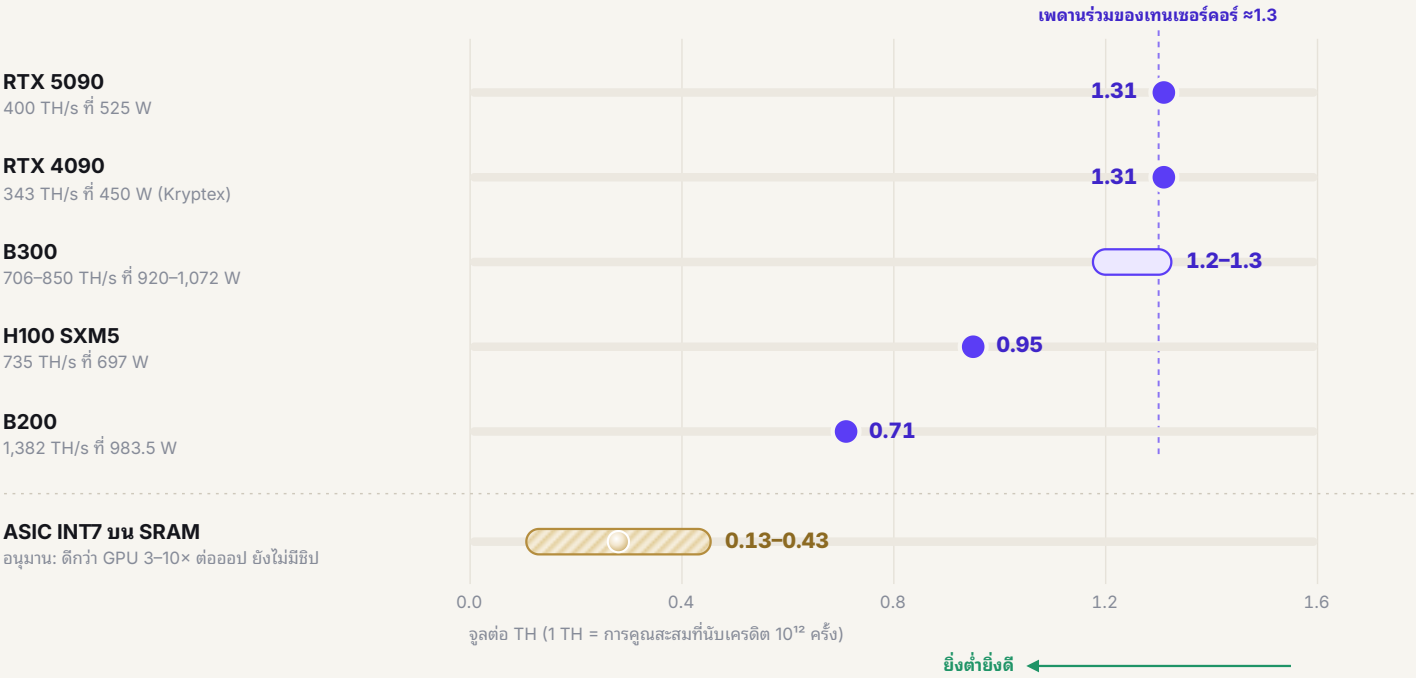
การคูณสะสมหนึ่งครั้งคือ “แฮช” หนึ่งครั้ง

สำหรับนักออกแบบชิป ข้อเท็จจริงที่สำคัญที่สุดคือหน่วยวัด โค้ดฉันทามติของ Pearl ปรับเป้าหมายของแต่ละความพยายามตามขนาดไทล์คูณความยาวผลคูณจุด ดังนั้น **1 H คือการคูณสะสม (MAC) INT7 × INT7 → INT32 ที่นับเครดิตหนึ่งครั้งพอดี** ตั้งแต่บล็อก 96,251 การลงโทษแรงก์กำหนดให้แรงก์ของสัญญาณรบกวนต้องไม่ต่ำกว่า 128 ที่แรงก์ขั้นต่ำ 1 H ยังคงเท่ากับ 1 MAC “EH/s” ของ Pearl จึงหมายถึงการคูณสะสมเลขจำนวนเต็ม 10^18 ครั้งต่อวินาทีตามตัวอักษร: **เป็นครั้งแรกที่งบความปลอดภัยของบล็อกเชนคิดราคาเป็นปริมาณงานเทนเซอร์ของตัวเร่ง AI** แฮชเรตของ GPU ที่วัดได้อยู่ที่ 60% ถึง 95% ของจุดสูงสุดการคูณสะสม INT8 แบบ dense ของการ์ดแต่ละใบ ซึ่งสอดคล้องกับการตีความนี้

ฮาร์ดแวร์	แฮชเรต Pearl	กำลังไฟ	จูลต่อการคูณสะสมหนึ่งล้านล้านครั้ง (J/TH)	แหล่งที่มา (วันที่)
RTX 5090	400 TH/s	525 W	1.31 ประมาณการ	Kryptex (12 พฤศจิกายน)
RTX 4090	ราว 305 TH/s (ช่วงจากชุมชน 280–330); Kryptex ระบุ 343 TH/s	450 W (Kryptex)	1.31 (ตัวเลขของ Kryptex) ประมาณการ	prlstats (3 ตุลาคม); Kryptex
H100 SXM5	735 TH/s	697 W	0.95 ประมาณการ	PeakMiner ผ่าน prlstats (19 กันยายน)
B200	1.382 PH/s	983.5 W	0.71 ประมาณการ	PeakMiner ผ่าน prlstats
B300	706–850 TH/s	920–1,072 W	1.2–1.3 ประมาณการ	prlstats; SimeonOnSecurity

ภาพที่ 02 **บนงานเลขจำนวนเต็มของ Pearl GPU ชนเพดานพลังงานเดียวกัน**

จูลต่อการคูณสะสม INT7 หนึ่งล้านล้านครั้ง (J/TH) ที่แฮชเรต Pearl ที่วัดได้ ยิ่งต่ำยิ่งดี



แหล่งที่มา: Kryptex (12 พฤศจิกายน 2026); ผลทดสอบ PeakMiner และ SRBMiner ผ่าน prlstats; SimeonOnSecurity; ช่วงของ ASIC เป็นการอนุมานของผู้เขียน ไม่ใช่การวัด

รายละเอียดสองข้อในตารางนี้สำคัญต่อนักออกแบบชิป ข้อแรก การ์ดผู้บริโภค RTX 4090 และ 5090 กับ B300 สำหรับศูนย์ข้อมูลมาบรรจบที่ราว 1.3 J/TH และแม้แต่การ์ดที่ดีที่สุดอย่าง B200 ก็ทำได้เพียงราว 0.71 J/TH GPU ศูนย์ข้อมูลราคาหลายหมื่นดอลลาร์ใช้พลังงานต่อการคูณสะสมเลขจำนวนเต็มเกือบเท่าการ์ดเกม ซึ่งชี้ไปที่เพดานร่วมของพลังงานเทนเซอร์คอร์ต่อ MAC แบบ INT8 และเพดานนั้นคือสิ่งที่ชิปเฉพาะทางอาจทลายได้ ข้อสอง B300 ช้ากว่า B200: NVIDIA ลดปริมาณงาน INT8 ลงอย่างมากใน Blackwell Ultra งานเลขจำนวนเต็มจึงต้องใช้เส้นทางที่ช้ากว่าบนชิปนั้น ซิลิคอน AI กระแสหลักกำลังเคลื่อนออกจาก INT8 ไปสู่ FP8 และ FP4 เมื่อรูปแบบตัวเลขที่ฉันทามติเลือกแยกทางกับทิศทางของฮาร์ดแวร์ โปรโตคอลก็ถูกบีบให้เปลี่ยนกติกา และนั่นคือรากของฟอร์ก FP8 และของข้อวินิจฉัยเรื่อง ASIC ของผมในตอนหลังของรายงานนี้

บทที่ 03

03

Pearl · เศรษฐศาสตร์

Pearl: ขนาด การปล่อยเหรียญ และช่องว่างด้านประโยชน์

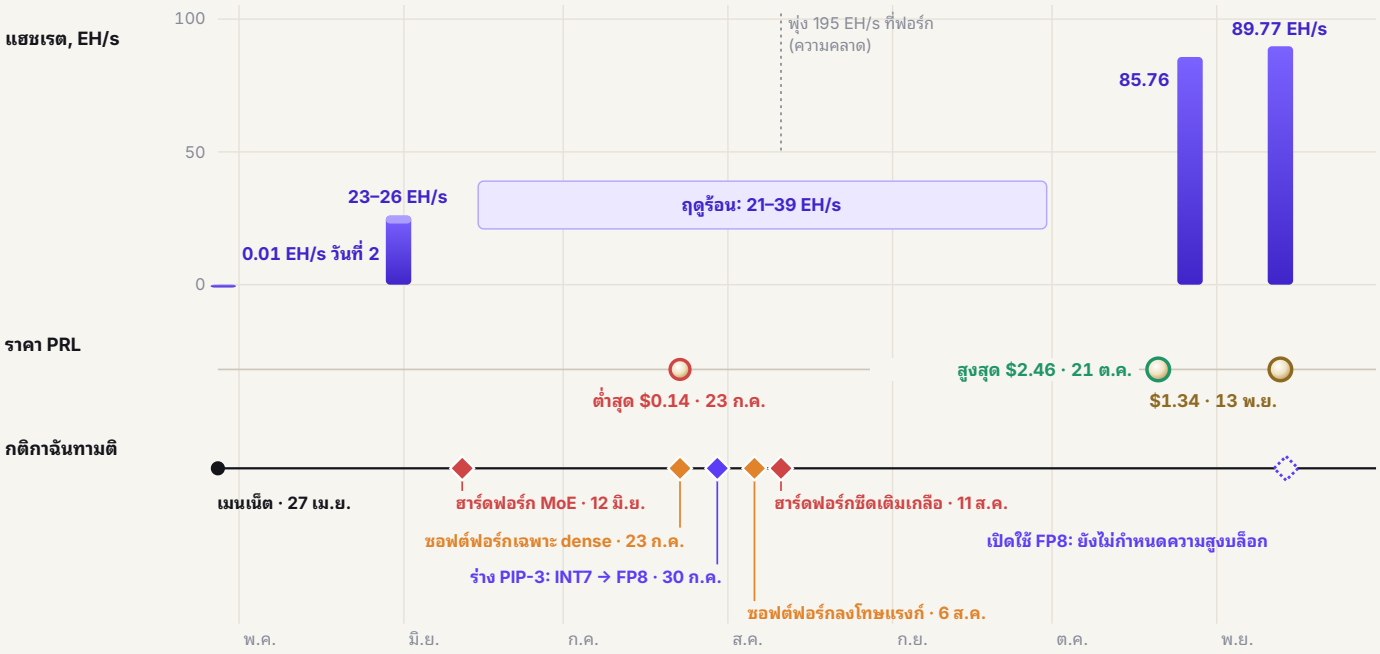
จาก 0.01 สู่ 90 EH/s ในหกเดือน

ตัวชี้วัด	ค่า (13 พฤศจิกายน 2026)	ฐานและแหล่งที่มา
แฮชเรตของเครือข่าย	ราว 90 EH/s (ค่าฐาน 89.77; ช่วงในวันเดียวกัน 88.20–93.22; สูงสุดระหว่างวันของ 12 พฤศจิกายน 110–123)	ฟูลโหนดอิสระ prlstats; Kryptex
ความยากและความสูงบล็อก	58,841,982 และ 142,530	prlstats
ฮาร์ดแวร์เทียบเท่า	H100 ราว 122,000 ใบ หรือ RTX 5090 ราว 225,000 ใบ	ที่ 735 และ 400 TH/s ต่อการ์ด ประมาณการ
ราคา มูลค่าตลาด และมูลค่าแบบ fully diluted	1.32–1.34 ดอลลาร์; ราว 500 ล้านดอลลาร์; ราว 2.8 พันล้านดอลลาร์ (อันดับ 101 บน CoinGecko)	CoinGecko
สูงสุดและต่ำสุดตลอดกาล	2.46 ดอลลาร์ (21 ตุลาคม) และ 0.1399 ดอลลาร์ (23 กรกฎาคม)	CoinGecko
เหรียญที่ขุดแล้ว	377.6 ล้าน PRL หรือ 17.98% ของเพดาน 2.1 พันล้าน	prlstats
เงินอุดหนุนบล็อกและการปล่อยเหรียญรายวัน	ราว 2,173 PRL และราว 0.97–1.03 ล้าน PRL	prlstats; Kryptex
อุปทานใหม่ใน 12 เดือนข้างหน้า	ราว 78% ของอุปทานปัจจุบัน	prlstats
รายได้นักขุด	ราว 1.25 ล้านดอลลาร์ต่อวัน (แบบจำลอง) ราว 460 ล้านดอลลาร์ต่อปี	prlstats
รายได้ต่อหน่วยแฮชเรต R	ราว 14–15 ดอลลาร์ต่อ PH/s ต่อวัน; RTX 5090 ได้ราว 5–6 ดอลลาร์ต่อวัน	ประมาณการ ; Kryptex
บล็อกที่มีงาน AI จริง	ราว 1% (ประมาณการ ณ 28 ตุลาคม)	Blockworks
การกระจุกตัวของพูล	Kryptex ขุดพบ 46.4% ของบล็อกใน 24 ชั่วโมงล่าสุด; สองพูลแรกราว 62%	prlstats
การกระจุกตัวของผู้ถือ	สัมประสิทธิ์จีนีของที่อยู่ 0.969; 100 ที่อยู่แรกถือ 53.0%	prlstats

แฮชเรตของ Pearl ใกล้เคียงกับฟังก์ชันที่ตามหลังราคา หนึ่งวันหลังเปิดตัว เครือข่ายทำงานที่ราว 0.01 EH/s ถึงปลายเดือนพฤษภาคมแตะ 23–26 EH/s ตลอดฤดูร้อน ราคาร่วงราว 86% จากจุดสูงสุดในเดือนพฤษภาคม และแฮชเรตเคลื่อนไหวในกรอบ 21–39 EH/s หลังราคาขึ้นราวสิบเท่าในเดือนกันยายนและตุลาคม แฮชเรตแตะ 85.76 EH/s เมื่อ 27 ตุลาคม และราว 90 EH/s ในเดือนพฤศจิกายน จุดพุ่งถึง 195 EH/s เมื่อ 11 สิงหาคมเกิดพร้อมกับฟอร์ก และควรอ่านว่าเป็นความคลาดของการประมาณค่า พลังงานเพิ่มตามไปด้วย: งานศึกษาหนึ่งประเมินไว้ราว 112 MW สำหรับทั้งเครือข่ายที่ 24 EH/s และแบบจำลองของ prlstats ประเมินว่าวันนี้อยู่ที่ราว 190–215 MW

ภาพที่ 03 **หกเดือนแรกของ Pearl: แฮชเรตตามราคา กติกาเปลี่ยนไม่หยุด**

แฮชเรตของเครือข่าย จุดราคาสามจุด และการเปลี่ยนฉันทามติทุกครั้ง 27 เมษายน ถึง 13 พฤศจิกายน 2026



แหล่งที่มา: ฟูลโหนดและชุดข้อมูลของ prlstats; Kryptex; CoinGecko; params.go และ PIP-3 ของ Pearl จุดพุ่งวันที่ 11 สิงหาคมเป็นความคลาดของการประมาณค่าที่จุดฟอร์ก

การปล่อยเหรียญขยายการสะท้อนกลับ

เพดานอุปทานคือ 2.1 พันล้านเหรียญ และไม่มีการลดรางวัลครึ่งหนึ่ง (halving): เงินอุดหนุนบล็อกลดลงอย่างราบรื่นตาม E(t) = S·H / ((t + H)(t + H − 1)) โดย H = 650,226 แต่การปล่อยเหรียญนับตามบล็อก ไม่ใช่ตามเวลา และความยากช่วงแรกเริ่มต่ำเกินไปและปรับตัวช้า **เหรียญราว 44.06 ล้านเหรียญถูกขุดในวันแรก และราว 120 ล้านเหรียญในหกวันแรก** ราว 32% ของอุปทานปัจจุบัน ประมาณการ บล็อกช่วงหลังซึ่งใช้เวลาราว 171–183 วินาที ยังเร็วกว่าเป้าหมาย 194 วินาที ผลคืออุปทานจะเพิ่มราว 78% ใน 12 เดือนข้างหน้า เป็นการเจือจางที่นักขุด ผู้ผลิตฮาร์ดแวร์ และผู้ถือเหรียญทุกคนต้องเผชิญ

เศรษฐศาสตร์ของนักขุดเป็นไปตามนั้น รายได้นักขุดทั้งเครือข่ายอยู่ที่ราว 1.25 ล้านดอลลาร์ต่อวัน (แบบจำลอง) การปล่อยเหรียญรายวันของ Kryptex คูณราคาปัจจุบันได้ราว 1.35 ล้านดอลลาร์ต่อวัน ดังนั้น **รายได้ต่อหน่วยแฮชเรต R ≈ 14–15 ดอลลาร์ต่อ PH/s ต่อวัน** ประมาณการ RTX 5090 ได้ราว 5–6 ดอลลาร์ต่อวัน H100 หรือ H200 ได้ราว 11 ดอลลาร์ ขณะที่ค่าเช่าแบบตามต้องการสาธารณะที่ถูกที่สุดอยู่ที่ราว 2.09 ดอลลาร์ต่อการ์ดต่อชั่วโมง หรือราว 50 ดอลลาร์ต่อวัน การเช่าการ์ดมาขุดอย่างเดียวจึงขาดทุนหนัก R แกว่งตัวกว้าง: ตามชุดข้อมูลของ prlstats อยู่ที่ราว 10 ดอลลาร์ในเดือนสิงหาคมและกันยายน และขึ้นไปราว 33 ดอลลาร์ชั่วครู่ในกลางเดือนตุลาคม ประมาณการ ภายในไม่กี่สัปดาห์หลังเปิดตัว รายได้รายวันโดยประมาณของ RTX 5090 ลดลงครึ่งหนึ่งจากราว 33.80 ดอลลาร์เหลือราว 17.19 ดอลลาร์ การเช่า GPU บนคลาวด์มาขุดคุ้มเฉพาะในสัปดาห์แรกๆ และช่วงราคาพุ่งในเดือนตุลาคม

ทฤษฎีบทใช้ได้ แต่แรงจูงใจใช้ไม่ได้

ปัญหาที่แท้จริงคืออุปสงค์ ไม่ใช่วิทยาการเข้ารหัส การเชื่อมต่อเชิงพาณิชย์ที่เปิดเผยต่อสาธารณะเพียงรายการเดียวของ Pearl คือ เอนด์พอยต์ Gemma-4-31B-it-pearl ของ Together AI ที่เปิดตัวเมื่อ 15 พฤษภาคม ในราคาต่ำกว่าปกติกว่า 25% โดยส่วนลดนั้นได้ทุนจากมูลค่าในอนาคตที่คาดหวังของการปล่อยเหรียญ ณ วันที่ 13 พฤศจิกายน หน้าโมเดลแสดงว่า **ไม่มีให้บริการบน serverless API ของ Together** ในเดือนมิถุนายน การสำรวจอินสแตนซ์การขุดที่ออนไลน์ 8,012 รายการ ราว 21% ของเครือ

ข่าย สรุปว่าเครือข่ายไม่ได้ผลิตงานประมวลผล AI ที่มีประโยชน์เลย (ผู้เขียนงานสำรวจเรียกหลักฐานนี้ว่าบ่งชี้ ไม่ใช่ข้อยุติ) และแสดงว่าพูลยอมรับแชร์ที่สร้างจากเมทริกซ์สุ่มล้วนๆ เมื่อ 28 ตุลาคม Blockworks ประเมินว่า **มีเพียงราว 1% ของบล็อกที่มีงาน AI จริง** และเขียนว่าหากไม่มีงานนั้น Pearl ก็คือบิตคอยน์ที่เล็กกว่า ปลอดภัยน้อยกว่า และเฟ้อเร็วกว่า

เหตุผลคำนวณได้ โอเวอร์เฮดฝั่งการอนุมานน้อยจริง: Pearl รายงานว่าเวลาแบบ end-to-end ของ Llama 70B บน H200 สี่ใบเพิ่มขึ้น 5.08% และการทดสอบอิสระวัดได้ว่าปริมาณงานลดลง 9.2% บน GPU Hopper หนึ่งใบ ฝั่งการขุดได้ผลตอบแทนน้อยเกินไป ปลั๊กอิน vLLM อย่างเป็นทางการขุดเฉพาะในแบตช์ prefill ที่มีอย่างน้อย 1,024 โทเคน และไม่ขุดในขั้น decode เลย การ์ดที่ให้บริการอนุมานได้แฮชเรตราว 27% ของเครื่องขุดเฉพาะทางเมื่อทราฟฟิกเป็น prefill ล้วน ราว 5.4% กับงานแชทที่เป็น decode 80% และเพียงราว 1% กับทราฟฟิกที่ decode เป็นหลัก หากไม่มีอุปสงค์การอนุมานที่จ่ายเงิน นักขุดที่มีเหตุผลจะคูณเมทริกซ์สุ่ม โปรโตคอลตรวจสอบว่าเมทริกซ์ที่คอมมิตไว้ถูกคูณอย่างซื่อตรง ไม่ได้ตรวจว่ามันมีประโยชน์หรือไม่ **ทฤษฎีบทใช้ได้ แต่แรงจูงใจใช้ไม่ได้** ประโยชน์ของ Pearl เป็นปัญหาเชิงพาณิชย์นอกเชน: ต้องมีคนจ่ายเงินซื้อการอนุมาน

บทที่ 04

04 Pearl · กติกา
Pearl: ความเสี่ยงด้านกติกา ระบบตัวเลข และธรรมาภิบาล

สี่ฟอร์กในสองเดือน

อีกด้านหนึ่งของ Pearl คือความเร็วที่กติกาเปลี่ยน ตั้งแต่เดือนมิถุนายนถึงสิงหาคม 2026 ทีมแล็บผลักดันการเปลี่ยนฉันทามติสี่ครั้งที่ความสูงบล็อกซึ่งเขียนตายตัวในโค้ด: ฮาร์ดฟอร์ก mixture-of-experts (บล็อก 71,935 ราว 12 มิถุนายน) ซอฟต์ฟอร์กเฉพาะ dense (91,630 ราว 23 กรกฎาคม) ซอฟต์ฟอร์กลงโทษแรงก์ (96,251 ราว 6 สิงหาคม) และฮาร์ดฟอร์กชีดเติมเกลือ (99,000 ราว 11 สิงหาคม) หลังฟอร์กชีดเติมเกลือ เครื่องขุดเก่าผลิตได้แต่แชร์ที่ไม่ถูกต้อง ส่วนการลงโทษแรงก์ทำให้แฮชเรตที่มีผลของการตั้งค่าแรงก์สูงลดลง 50% ถึง 87.5%

ขั้นต่อไปคือ PIP-3 ซึ่งยังเป็นร่าง: ย้ายการขุดจาก INT7 ไปเป็น FP8 × FP8 → FP32 และกำหนดให้ใช้โปรไฟล์อุปกรณ์ที่ตรงกันทุกบิต **รองรับเพียงสองสถาปัตยกรรม คือ Hopper และ Blackwell** เอกสารเรียกการพึ่งพาฮาร์ดแวร์ว่าเป็นความชั่วร้ายที่จำเป็นของโปรโตคอลแบบทศนิยมลอยตัว และเตือนว่าความคลาดเคลื่อนใดๆ ที่ยอมให้ในการตรวจสอบจะเชื้อเชิญการโจมตีแบบ grinding เมื่อ 12 พฤศจิกายน พารามิเตอร์เมนเน็ตของแบรนช์ fp8-scheme ยังไม่กำหนดความสูงบล็อกที่เปิดใช้ ภายใต้ร่างปัจจุบัน การ์ดซีรีส์ RTX 40, A100 และการ์ด AMD จะไม่ได้รับอนุญาต ขณะที่แบบจำลองความเชื่อมั่นต่ำของ prlstats ประเมินว่าราว 82% ของแฮชเรตมาจากการ์ดผู้บริโภค การดันฟอร์กให้ผ่านจึงเสี่ยงต่อการแยกเชนหรือการต่อต้านจากนักขุด

ไตรลักษณ์ของรูปแบบตัวเลข

Pearl เผชิญไตรลักษณ์: รูปแบบตัวเลขหนึ่งอาจมีประโยชน์ต่อ AI (FP8 หรือ FP4) ตรวจสอบได้ตรงกันทุกบิตข้ามฮาร์ดแวร์ และเป็นกลางต่อฮาร์ดแวร์ แต่ได้อย่างมากสองในสามข้อ แบบ INT7 ตรวจสอบได้และเป็นกลาง แต่แยกทางกับทิศทางของซิลิคอน AI FP8 พร้อมรายชื่ออนุญาตมีประโยชน์และตรวจสอบได้ แต่ผูกฉันทามติไว้กับผู้ขายรายเดียว FP8 พร้อมค่าความคลาดเคลื่อนที่ยอมให้มีประโยชน์และเป็นกลาง แต่เปิดประตูให้การโจมตีแบบ grinding เลขคณิตทศนิยมลอยตัวในเทนเซอร์คอร์โดยทั่วไปไม่ได้เป็นไปตามมาตรฐานเชิงตัวเลข และทำซ้ำข้ามรุ่นไม่ได้ ตัวอย่างเช่นบน Hopper การสะสม FP8 เก็บบิตเศษส่วนไว้เพียง 13 บิตและตัดทิ้งส่วนที่เหลือ จากมุมมองของนักออกแบบชิป “โปรไฟล์ Hopper” คือเลขคณิตที่กำหนดไว้ครบถ้วนซึ่ง ASIC อาจทำได้ตรงทุก

บิต และอาจมีประสิทธิภาพกว่า GPU ด้วย แต่การทำเช่นนั้นจะเป็นธุรกิจในพื้นที่สีเทาที่ขัดกับเจตนาที่โปรโตคอลประกาศไว้ และมีความเสี่ยงด้านธรรมาภิบาลสูงมาก บทวิเคราะห์

ธรรมาภิบาล

ธรรมาภิบาลกระจุกตัว ทีมแล็บเขียนข้อเสนอโปรโตคอลทุกฉบับ ฟอร์กเปิดใช้ที่ความสูงบล็อกที่เขียนตายตัวโดยไม่มีการส่งสัญญาณจากนักขุด ไม่มีการพัฒนาไคลเอนต์ตัวที่สอง และจนถึงต้นเดือนตุลาคมยังไม่มีการเผยแพร่ผลการตรวจสอบโดยบุคคลที่สามของโหนดหรือวงจร zero knowledge หนังสือพิมพ์ TheMarker ของอิสราเอลรายงานว่าทีมแล็บระดมทุนได้ราว 120–130 ล้านดอลลาร์ในรอบลับ (ทุนเรือนหุ้นราว 55 ล้านดอลลาร์ บวกกองทุนโครงสร้างพื้นฐาน GPU ราว 75 ล้านดอลลาร์) Pearl ยังไม่ได้ยืนยันผ่านช่องทางทางการ “ไม่มีการจัดสรรโทเคนให้นักลงทุน” และ “บริษัทที่มีผู้ถือหุ้นแบบร่วมลงทุน” อาจเป็นจริงทั้งคู่ แต่ความตึงเครียดยังคงอยู่: รายได้ AI อาจไหลไปที่หุ้นของบริษัท ขณะที่การเจือจางตกอยู่กับผู้ถือโทเคน

เมื่อมองรวมกัน Pearl คือการทดลอง “การประมวลผลคือฉันทามติ” ที่สะอาดที่สุดจนถึงตอนนี้: กลไกใหม่ ตรวจสอบได้บนเชน และมีหน่วยวัดที่จับคู่หนึ่งต่อหนึ่งกับฮาร์ดแวร์ AI แต่ประโยชน์ยังมาไม่ถึง และงบความปลอดภัย แฮชเรต และความกระตือรือร้นของนักขุดล้วนแขวนอยู่กับโทเคนที่อุปสงค์หลักยังไม่ก่อตัว และอุปทานจะเพิ่มราว 78% ใน 12 เดือน สัญญาณสามข้อจาก Blockworks ควรติดตาม: การอัปเกรดจะผ่านไปอย่างราบรื่นหรือไม่ สัดส่วนบล็อกที่มีงาน AI จะเพิ่มขึ้นอย่างมีนัยสำคัญหรือไม่ และจะมีผู้ให้บริการอนุมานรายใดส่งต่อรายได้จากการขุดให้ลูกค้าหรือไม่

บทที่ 05

เชนงานที่มีประโยชน์

05 เชนงานที่มีประโยชน์: ทั้งเมทริกซ์และ zero knowledge ต่างหยุด ASIC ไม่ได้

Kaspa: เมทริกซ์ไม่ได้มีประโยชน์และไม่กัน ASIC

Kaspa คือกรณีตัวอย่างตามตำรา kHeavyHash ของมันใส่ผลคูณเมทริกซ์กับเวกเตอร์ขนาด 64 × 64 บนค่า 4 บิตไว้ระหว่างแฮชสองครั้ง และทิ้งผลคูณนั้นไป ราว 16 เดือนหลังการเปิดตัวแบบ fair launch ในเดือนพฤศจิกายน 2021 IceRiver ประกาศ ASIC สำหรับ Kaspa รุ่นแรกเมื่อ 14 มีนาคม 2023 และถึงปลายปี 2023 การขุดด้วย GPU ก็จบลง เมื่อเทียบ RTX 4090 ที่ 3.8 GH/s และ 300 W กับ KS7 ของ Bitmain ที่ 77 J/TH ช่องว่างด้านประสิทธิภาพอยู่ที่ราว **1,000 เท่า** ประมาณการ หลังงานเลี้ยงก็ถึงช่วงบีบตัว: เมื่อ 13 พฤศจิกายน เครือข่ายทำงานที่ 358.9 PH/s ต่ำกว่าระดับ 1 EH/s ของเดือนสิงหาคม 2024 ราว 64% และ KS7 ที่ดีที่สุดได้เพียง **0.86 ดอลลาร์ต่อวัน** ที่ค่าไฟ 0.10 ดอลลาร์ต่อ kWh

Aleo: zero knowledge ก็หยุดฮาร์ดแวร์เฉพาะทางไม่ได้เช่นกัน

ปริศนาผู้พิสูจน์ของ Aleo ออกแบบมาเพื่อเลียนแบบงานสร้างการพิสูจน์ SNARK มันไม่ได้พิสูจน์ธุรกรรมของผู้ใช้ และโครงการเปิดรับ ASIC สำหรับ ZK อย่างเปิดเผย ราวห้าเดือนหลังเมนเน็ตในเดือนกันยายน 2024 ก็มีเครื่องเฉพาะทางวางขาย ในไตรมาสที่ 3 ปี 2025 ผู้พิสูจน์ได้รับ 31.2 ล้าน ALEO (ราว 7.8 ล้านดอลลาร์) ขณะที่เครือข่ายเก็บค่าธรรมเนียมได้เพียงราว 43,000–44,000 ดอลลาร์: **ตัวคูณเงินอุดหนุนราว 175–180** ราคาร่วงจากราว 9 ดอลลาร์ในช่วงซื้อขายแรกเหลือ 0.22 ดอลลาร์ ณ สิ้นไตรมาสนั้น และตอนนี้ ARC-46 กำหนดให้ต้องสเตก 100,000 ถึง 2.5 ล้าน ALEO ต่อคำตอบ ซึ่งเปลี่ยนการขุดให้กลายเป็นเกณฑ์ขั้นต่ำของการสเตก

Qubic และ Bittensor

Qubic ทำการตลาด “การพิสูจน์งานที่มีประโยชน์” ที่เทรนโครงข่ายประสาทเทียม Aigarth ของตน แต่ผลผลิตที่เห็นชัดที่สุดของพลังประมวลผลในปี 2025–2026 คือบล็อกของเชนอื่น แรกเริ่มมันขุด Monero เพิ่มส่วนแบ่งจากไม่ถึง 2% ในกลางเดือนพฤษภาคม 2025 เป็นกว่า 25% ณ ปลายเดือนกรกฎาคม และอ้างว่าได้ 51% ในเดือนสิงหาคม ซึ่งทำให้เกิดการจัดเรียงเชนใหม่หกบล็อก (อีกแหล่งหนึ่งว่า 18) เมื่อ 1 เมษายน 2026 มันเปลี่ยนไปขุด Dogecoin ด้วย ASIC และนำรายได้ไปซื้อคืนและเผาเหรียญ ไม่มีหลักฐานอิสระว่าการเทรน AI ของมันมีประโยชน์ และมูลค่าตลาดร่วงจากราว 280–300 ล้านดอลลาร์เหลือราว 50 ล้านดอลลาร์เมื่อ 13 พฤศจิกายน มันยังเผยให้เห็นผลกระทบภายนอกอย่างหนึ่ง: เชนที่จ่ายเงินให้พลังประมวลผลอเนกประสงค์จะสร้างแหล่งแฮชเรตขนาดใหญ่ที่เช่าได้และประสานกันได้ ซึ่งหันไปเล่นงานเชนที่เป็นมิตรกับ CPU ได้

Bittensor ไม่ใช่การพิสูจน์งาน แต่เป็นการให้คะแนนเชิงอัตวิสัยที่ถ่วงด้วยสเตก: ผู้ตรวจสอบให้คะแนนนักขุด และน้ำหนักถูกตัดที่ค่ามัธยฐานถ่วงด้วยสเตก มันเป็นเครือข่ายเดียวในกลุ่มนี้ที่มีรายได้จากภายนอกในระดับใหญ่ ดัชนี SubConnect ซึ่งนับเฉพาะลูกค้าภายนอกที่ตรวจสอบได้ แสดงตัวเลขราว **70 ล้านดอลลาร์ต่อปี** ในต้นเดือนพฤศจิกายน 2026 (ราว 35 ล้านดอลลาร์เมื่อ 1 กันยายน) โดย 78% ถึง 82% มาจากการประมวลผลและโครงสร้างพื้นฐาน แต่ที่ราคาต้นเดือนพฤศจิกายน การปล่อยเหรียญเกิน 300 ล้านดอลลาร์ต่อปี รายได้จึงครอบคลุมได้เพียง 9% ถึง 23% ประมาณการรายได้ต่างกันมากตามฐานที่ใช้ ตั้งแต่ราว 17 ล้านดอลลาร์ (Artemis) ถึง 172 ล้านดอลลาร์ต่อปีจากยอดรวมที่รายงานเอง ในเดือนเมษายน ซับเน็ต 67 จาก 128 แห่งเผาสิ่งจูงใจของตนเกิน 99% และเมื่อทีมชั้นนำอย่าง Covenant ลาออกเพราะเรื่องธรรมาภิบาล TAO ร่วง 24% ภายในไม่กี่ชั่วโมง

ประวัติศาสตร์ให้คำตอบเดียวกัน

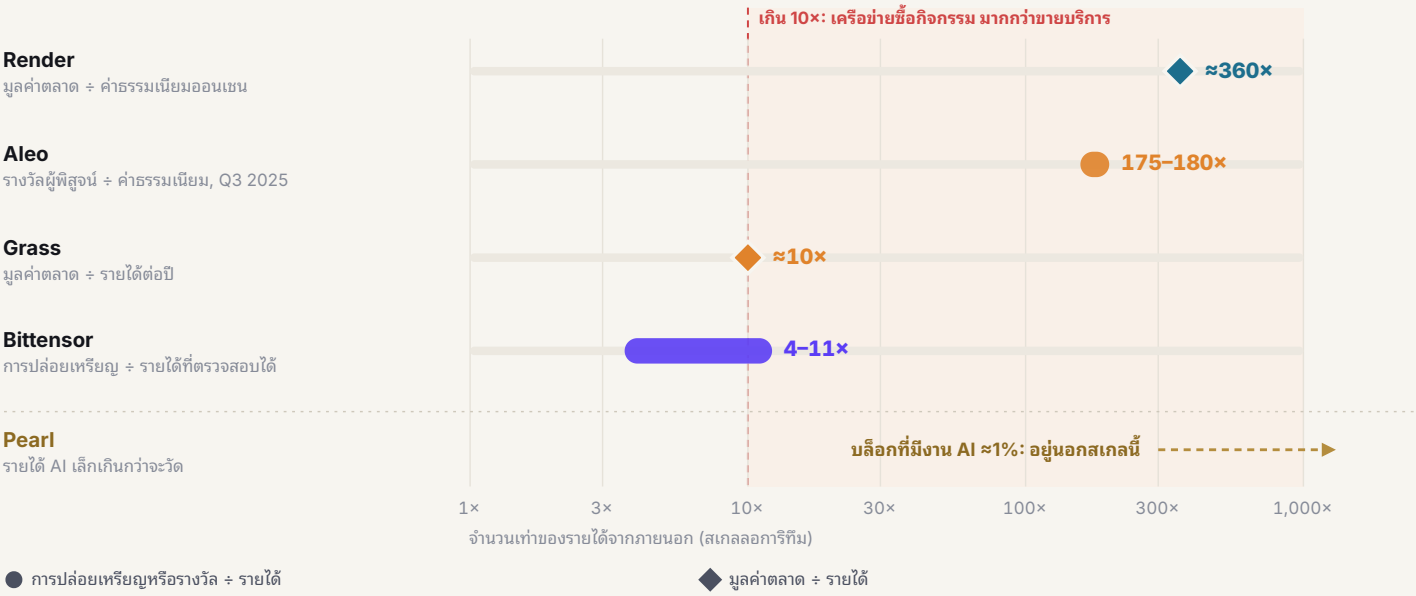
สายจำนวนเฉพาะของ Primecoin ตรวจสอบได้เร็ว แต่ไม่มีใครต้องการมัน Gridcoin พึ่งเว็บไซต์สถิติของบุคคลที่สามในการรับรองการมีส่วนร่วมใน BOINC และนักวิจัยปลอมบล็อกได้โดยไม่ต้องประมวลผลอะไรเลย การพิสูจน์การจัดเก็บของ Filecoin ใช้ได้ในทางเทคนิค แต่ในไตรมาสที่ 3 ปี 2025 อัตราการใช้งานอยู่ที่เพียง 36% และ 99.5% ของค่าธรรมเนียม 793,000 ดอลลาร์เป็นค่าปรับ โดยค่าธรรมเนียมพื้นฐานมีเพียง 4,100 ดอลลาร์

โครงการ	งานหรือฉันทามติ	มูลค่าตลาด (วันที่)	เส้นทางฮาร์ดแวร์	อุปสงค์ที่จ่ายเงินเทียบเงินอุดหนุน	ข้อวินิจฉัย
Pearl	การคูณเมทริกซ์แบบเติมสัญญาณรบกวน (MAC แบบ INT7 บวกใบรับรอง ZK)	ราว 500 ล้านดอลลาร์ (13 พฤศจิกายน)	GPU; ร่าง FP8 จะใส่ Hopper และ Blackwell ในรายชื่ออนุญาต	บล็อกที่มีงาน AI ราว 1%; รายได้นักขุดส่วนใหญ่เป็นเงินอุดหนุน	กลไกที่น่าเชื่อถือที่สุด แต่อุปสงค์ยังไม่ได้พิสูจน์
Kaspa	kHeavyHash (ผลคูณเมทริกซ์กับเวกเตอร์ 64 × 64 แบบ 4 บิตระหว่างแฮช)	ราว 840 ล้านดอลลาร์ (13 พฤศจิกายน)	ASIC ราว 16 เดือนหลังเปิดตัว ตอนนี้เป็น ASIC ทั้งหมด	ไม่อ้างว่ามีประโยชน์; รายได้นักขุดราว 18 ล้านดอลลาร์ต่อปี ประมาณการ	เมทริกซ์ไม่ได้มีประโยชน์และไม่กัน ASIC
Aleo	ปริศนา SNARK สังเคราะห์บวกการพิสูจน์ด้วยสเตก	ราว 118 ล้านดอลลาร์ (สิ้นไตรมาส 3 ปี 2025)	เครื่องเฉพาะทางหลังราวห้าเดือน	รางวัลผู้พิสูจน์ราว 175–180 เท่าของค่าธรรมเนียม	งานสังเคราะห์ ขับเคลื่อนด้วยเงินอุดหนุน
Qubic	“uPoW” (อ้างว่าเทรน Aigarth)	ราว 50 ล้านดอลลาร์ (13 พฤศจิกายน)	CPU และ GPU; ASIC สำหรับ Dogecoin ตั้งแต่เมษายน 2026	ไม่มีหลักฐานอิสระว่ามีประโยชน์; พลังประมวลผลใช้ขุดภายนอกและซื้อคืน	การรวมพลังประมวลผลที่ห่อด้วยเรื่องเล่า AI
Bittensor	ฉันทามติ Yuma (การให้คะแนนถ่วงด้วยสเตก)	ราว 2.55 พันล้านดอลลาร์ (11 พฤศจิกายน)	GPU อเนกประสงค์; ไม่มีเส้นทาง ASIC	รายได้ที่ตรวจสอบได้ราว 70 ล้านดอลลาร์ต่อปี เทียบการปล่อยเหรียญกว่า 300 ล้านดอลลาร์	รายเดียวที่มีรายได้จากภายนอกในระดับใหญ่ แต่ยังเป็นเงินอุดหนุน 4–11 เท่า

โครงการ	งานหรือฉันทามติ	มูลค่าตลาด (วันที่)	เส้นทางฮาร์ดแวร์	อุปสงค์ที่จ่ายเงินเทียบเงินอุดหนุน	ข้อวินิจฉัย
Filecoin (อ้างอิง)	การพิสูจน์การจำลองและการพิสูจน์ปริภูมิ-เวลา	ราว 1.5 พันล้านดอลลาร์ (สิ้นไตรมาส 3 ปี 2025)	—	อัตราการใช้งาน 36%; ค่าธรรมเนียมพื้นฐานรายไตรมาส 4,100 ดอลลาร์	มีประโยชน์ แต่อุปสงค์ไม่พอ

ภาพที่ 04 **ตัวคูณเงินอุดหนุน: เครือข่ายจ่ายกี่เท่าของรายได้จากภายนอก**

การปล่อยเหรียญหรือมูลค่าตลาด หารด้วยรายได้ที่จ่ายจากนอกเครือข่าย สเกลลอการิทึม



แหล่งที่มา: Blockworks (Aleo ไตรมาส 3 ปี 2025), DefiLlama (ค่าธรรมเนียม Render), CoinGecko, hoge.gg (Bittensor), Dealroom (Grass), Blockworks 0xResearch (Pearl); การคำนวณของผู้เขียน

บทเรียนสำหรับฮาร์ดแวร์: งานที่ตายตัวและสม่ำเสมอซึ่งจ่ายด้วยโทเคนที่มีสภาพคล่องจะมี ASIC ภายในราว 3 ถึง 16 เดือน บ่อยครั้งมาจากผู้ท้าชิงก่อนผู้เล่นเดิม ข้ออ้างว่า “มีประโยชน์” ยืนได้เฉพาะเมื่อมีลูกค้าภายนอกจ่ายเงินซื้อผลลัพธ์ และตัวการหลักที่ทำให้ฮาร์ดแวร์ค้างเติ่งคือเงินอุดหนุนที่สูงเมื่อเทียบกับรายได้ และการที่นักพัฒนาควบคุมฟังก์ชันงาน

บทที่ 06

ภาพรวมที่กว้างขึ้น

06 ส่วนอื่นของ AI × คริปโต: อุปสงค์จริง แต่เล็ก และมูลค่ากำลังย้ายไปที่หุ้น

ตลาดการประมวลผลแบบกระจายศูนย์

การประมวลผลแบบกระจายศูนย์มีลูกค้าที่จ่ายเงิน แต่อุปสงค์เล็ก กระจุกตัว และส่วนใหญ่เป็นการขายต่อการอนุมานของโมเดลเปิด ค่าธรรมเนียมที่ตรวจสอบได้บนเชนมีน้อย: DefiLlama แสดงค่าธรรมเนียมราว 2.35 ล้านดอลลาร์สำหรับ Render ในปีที่ผ่านมา และราว 680,000 ดอลลาร์สำหรับ Livepeer และทั้งกลุ่ม DePIN มีรายได้ออนเชนเพียงราว 72 ล้านดอลลาร์ในปี 2025 ตัวเลขที่รายงานเองสูงกว่าแต่นิยามหลวม: Aethir รายงานรายได้ “มากกว่า 127.8 ล้านดอลลาร์” สำหรับปีงบประมาณ 2025 โดยไม่ระบุว่าเป็นยอดเรียกเก็บรวมหรือส่วนที่โปรโตคอลได้ มูลค่าโทเคนกับรายได้แยกทางกันทั้งสองทิศ: มูลค่าตลาดของ Render ราว 360 เท่าของค่าธรรมเนียมออนเชน ขณะที่มูลค่าตลาดของ Aethir ราว 87.7 ล้านดอลลาร์ต่ำกว่ารายได้ที่รายงานเอง ประมาณการ

ข้ออ้างว่าถูกกว่า 70% ถึง 85% เป็นจริงเฉพาะเมื่อเทียบกับราคาแบบตามต้องการของไฮเปอร์สเกลเลอร์ เมื่อ 3 พฤศจิกายน ราคา H100 ถัวเฉลี่ยถ่วงน้ำหนักบน Akash อยู่ที่ราว **2.73 ดอลลาร์ต่อ GPU-ชั่วโมง** (ช่วง 2.04–3.16 ดอลลาร์) ต่ำกว่าราว 6.88 ดอลลาร์บน AWS และ 11.06 ดอลลาร์บน Google Cloud แต่สูงกว่าราว 1.73 ดอลลาร์บน Vast.ai, 1.99 ดอลลาร์บนคอมมูนิตี้คลาวด์ของ RunPod และราว 2.44 ดอลลาร์สำหรับการจองหนึ่งปี (ราคาบนแพลตฟอร์มอื่นอ่านเมื่อ 12 กันยายน; Akash ผู้เผยแพร่ตัวเลขเป็นผู้มีส่วนได้เสีย) อุปสงค์ส่วนใหญ่เป็นการอนุมานทั่วไปที่กระจายผ่าน OpenRouter ราว 1.7 พันล้านโทเคนต่อวันบน AkashML และกว่า 4 พันล้านบน io.net แข่งขันตรงกับนีโอคลาวด์ทุกราย เงินดอลลาร์ก้อนใหญ่ที่สุดของ “การประมวลผลแบบกระจายศูนย์” กำลังกลับไปรวมศูนย์: Axe Compute ซึ่งถือโทเคนของ Aethir เป็นสินทรัพย์คงคลัง เซ็นสัญญาห้าปีมูลค่า 1.5 พันล้านดอลลาร์เพื่อให้เช่าคลัสเตอร์ B300 จำนวน 9,200 ใบของตนเอง เทียบกับรายได้ไตรมาสแรกปี 2026 ที่ 35,000 ดอลลาร์

การเทรนและการอนุมานแบบกระจายศูนย์

การเทรนผ่านอินเทอร์เน็ตได้รับการสาธิตแล้วที่ 8 ถึง 40 พันล้านพารามิเตอร์ Pluralis-8B เทรนนาน 40 วันบน 330 โหนด ส่วนใหญ่เป็น GPU ผู้บริโภค และได้ประสิทธิภาพ 63% ของค่าฐาน H100 แบบรวมศูนย์ แม้งานวิจัยจะยอมรับว่าการเทรนนั้นไม่ได้กระจายศูนย์ และต้องใช้โหนดของ Pluralis เองกับหน่วยงานรับเข้า Psyche ของ Nous Research เทรนล่วงหน้า Consilience ขนาด 40 พันล้านพารามิเตอร์สำเร็จ ในระดับแนวหน้า งานกลับไปรวมศูนย์: INTELLECT-3 ของ Prime Intellect โมเดล mixture-of-experts ขนาด 106 พันล้านพารามิเตอร์ เทรนนานสองเดือนบน H200 จำนวน 512 ใบที่เชื่อมด้วย InfiniBand 400 Gbps ข้อจำกัดชัดเจน: โหนดต้องมีแบนด์วิดท์อย่างน้อย 200 Mbit/s และการบีบอัดการสื่อสารราว 100 เท่า และความเชื่อใจกับการตรวจสอบยังขึ้นกับบทบาทพิเศษของผู้ดำเนินการ

เงินไปอยู่กับบริษัทเอกชนที่ขายผลิตภัณฑ์แบบรวมศูนย์ Prime Intellect มีรายได้ต่อปีมากกว่า 100 ล้านดอลลาร์ก่อนระดมทุน 130 ล้านดอลลาร์ที่มูลค่า 1 พันล้านดอลลาร์ในเดือนกรกฎาคม Venice แอปอนุมานแบบส่วนตัว รายงานรายได้ต่อปีมากกว่า 70 ล้านดอลลาร์และมีกำไรตั้งแต่ไตรมาสแรกปี 2026 (ตัวเลขของบริษัท) ระดมทุน 65 ล้านดอลลาร์ที่มูลค่า 1 พันล้านดอลลาร์ และทำงานบน GPU ที่เช่ามา ฝั่งโทเคน เครือข่ายการเทรนหลักของ Gensyn ยัง “รอกำหนด” และโทเคนมีมูลค่าราว 17.2 ล้านดอลลาร์ โทเคนของ Ritual เลเยอร์ 1 ที่เกิดมาเพื่อ AI ร่วง 72% จากจุดสูงสุดหลังเมนเน็ต (มีเพียงพาดหัวข่าว)

AI ที่ตรวจสอบได้

การใช้งานจริงเลือกสภาพแวดล้อมการประมวลผลที่เชื่อถือได้ (TEE) การประมวลผลแบบเป็นความลับของ Hopper เพิ่มโอเวอร์เฮดไม่ถึง 7% สำหรับคิวรี LLM ทั่วไป และแทบไม่มีเลยสำหรับโมเดลใหญ่ โอเวอร์เฮดในการใช้งานจริงของ Phala อยู่ที่ราว 0.5% ถึง 5% แต่ ณ สิ้นปี 2025 มีผู้ใช้ 398 จาก 10,004 รายที่จ่ายเงิน รากของความเชื่อถือของ TEE คือผู้ขายชิป: TEE.fail ซึ่งเปิดเผยในเดือนตุลาคม 2025 ใช้ interposer บนบัส DDR5 ราคาไม่ถึง 1,000 ดอลลาร์ดึงคีย์ attestation ออกจากเครื่อง Intel TDX และ AMD SEV-SNP ที่อัปเดตครบถ้วน ซึ่งส่งผลต่อเนื่องให้การประมวลผลแบบเป็นความลับบน GPU ของ NVIDIA อ่อนแอลงด้วย และผู้ขายทั้งสองรายกล่าวว่าการโจมตีทางกายภาพอยู่นอกแบบจำลองภัยคุกคามของตน นั่นคือสภาพของ DePIN พอดี ที่ผู้ดำเนินการโหนดเข้าถึงเครื่องได้ทางกายภาพ zkML ตอนนี้พิสูจน์ LLM ได้ทั้งตัว: DeepProve ของ Lagrange พิสูจน์ GPT-2 ได้ราว 174 โทเคนต่อนาที และ Gemma 3 ได้ราว 86 ห่างจากการอนุมานในการใช้งานจริงหลายระดับขนาด และไม่มีการเปิดเผยการใช้งานแบบจ่ายเงิน

การชำระเงินของเอเจนต์และข้อมูล

x402 เป็นตัวอย่างเตือนใจที่ดีที่สุด TRM Labs นับได้ราว 198.9 ล้านธุรกรรม และ 52.7 ล้านดอลลาร์ที่ชำระผ่านผู้อำนวยความสะดวกที่รู้จักตั้งแต่เดือนพฤษภาคม 2025 โดย 99.6% เป็น USDC หลังตัดการจ่ายให้ตัวเอง กระแสแบบแบตช์ และผู้ขายที่มีผู้ซื้อไม่ถึงสิบราย ส่วนที่เอเจนต์เป็นผู้เริ่มจริงมีเพียง 0.6% ถึง 7.5% หรือ **ราว 5,000–11,000 ดอลลาร์ต่อเดือน** Artemis ประเมินว่ากิจกรรมกว่า 95% เป็นการส่งสัญญาณของโปรโตคอล มากกว่าการซื้อขาย ขณะเดียวกัน รางกำลังถูกทำให้เป็นมาตรฐานโดยผู้เล่นเดิม: x402 Foundation ก่อตั้งภายใต้ Linux Foundation เมื่อ 14 กรกฎาคม โดยมี Visa, Mastercard, Stripe, Google, AWS, Cloudflare และ Coinbase เป็นสมาชิกหลัก และ Tempo ซึ่งร่วมเขียน Machine Payments Protocol กับ Stripe เปิดเมนเน็ตในเดือนมีนาคมที่มูลค่า 5 พันล้านดอลลาร์ แพลตฟอร์มโทเคนเอเจนต์ยังทำงานด้วยตรรกะแบบลอนช์แพด: Virtuals รายงานเองว่ามี “aGDP” 479 ล้านดอลลาร์ และผู้ก่อตั้ง ElizaOS ประกาศว่าโทเคนของตนตายแล้วหลังการยอมความคดีแบบกลุ่ม มูลค่ามีแนวโน้มจะตกอยู่กับผู้ออกสเตเบิลคอยน์และผู้เล่นเดิมด้านการชำระเงินมากกว่า บทบาทเฉพาะตัวของคริปโตในที่นี้อาจเป็นตัวตนและความรับผิดชอบของเอเจนต์

ข้อมูลเป็นหนึ่งในไม่กี่กลุ่มย่อยที่มีผู้ซื้อจริงนอกเชน Grass รายงานรายได้ข้อมูล 17 ล้านดอลลาร์ในปี 2025 และอีก 17 ล้านดอลลาร์ในครึ่งแรกของปี 2026 ตั้งเป้า 75 ล้านดอลลาร์สำหรับทั้งปี และกล่าวว่ามีกำไร GRASS มีมูลค่าราว 328 ล้านดอลลาร์ ราวสิบเท่าของรายได้ครึ่งแรกเมื่อคิดเป็นรายปี ประมาณการ คู่เมืองของมันคือขนาดการเก็บข้อมูลผ่านแบนด์วิดท์ที่อยู่อาศัยแบบกระจาย ไม่ใช่ตัวคริปโตเอง Data DAO และตลาดข้อมูลยังไม่มีอุปสงค์ที่จ่ายเงินให้เห็น: Vana มีมูลค่าราว 21.5 ล้านดอลลาร์ และ Ocean ราว 27.6 ล้านดอลลาร์ และ Ocean อยู่ระหว่างคดีความกับ Fetch.ai การพิสูจน์ความเป็นบุคคลได้การเชื่อมต่อจริงมากขึ้นเมื่อเอเจนต์เพิ่มจำนวน: World กล่าวว่ายืนยันตัวตนคนไปแล้วเกือบ 18 ล้านคน และ Zoom, Docusign และ Tinder รองรับ World ID บางส่วน

ทุนลงคะแนนแล้ว

กลุ่ม	อุปสงค์ที่จ่ายเงินและตรวจสอบได้	สัญญาณด้านเงินอุดหนุนและมูลค่า	ข้อวินิจฉัย
เชนงานที่มีประโยชน์	Bittensor มีรายได้จากภายนอกที่ตรวจสอบได้ราว 70 ล้านดอลลาร์ต่อปี; Pearl มีบล็อกที่มีงาน AI ราว 1%	Aleo 175–180 เท่า; Bittensor 4–11 เท่า; Pearl ส่วนใหญ่เป็นเงินอุดหนุน	กลไกมีคุณค่า ประโยชน์ต้องมีลูกค้าที่จ่ายเงิน
ตลาดการประมวลผลแบบกระจายศูนย์	ค่าธรรมเนียมออนเชนราว 0.68–2.35 ล้านดอลลาร์ต่อปี (Livepeer, Render); รายงานเอง 4–128 ล้านดอลลาร์	มูลค่าตลาดของ Render ราว 360 เท่าของค่าธรรมเนียม	จริงแต่เล็ก ส่วนใหญ่เป็นการขายต่อการอนุมานทั่วไป
การเทรนและการอนุมานแบบกระจายศูนย์	บริษัทเอกชน: Prime Intellect กว่า 100 ล้านดอลลาร์ต่อปี Venice กว่า 70 ล้านดอลลาร์ (โครงสร้างพื้นฐานแบบรวมศูนย์)	เมนเน็ตการเทรนของ Gensyn ยังไม่เปิด; โทเคนของ Ritual −72% จากจุดสูงสุด	ความก้าวหน้าทางเทคนิคเป็นของจริง กำไรไปที่หุ้น

กลุ่ม	อุปสงค์ที่จ่ายเงินและตรวจสอบได้	สัญญาณด้านเงินอุดหนุนและมูลค่า	ข้อวินิจฉัย
AI ที่ตรวจสอบได้	TEE ใช้งานจริง (ผู้ใช้ที่จ่ายเงิน 398 รายที่ Phala); ไม่มีการเปิดเผย zkML แบบจ่ายเงิน	—	TEE มีประโยชน์ zkML ยังระยะแรก
การชำระเงินของเอเจนต์	กระแสที่เอเจนต์เริ่มเองราว 5,000–11,000 ดอลลาร์ต่อเดือน	ราว 200 ล้านธุรกรรม ส่วนใหญ่เป็นการส่งสัญญาณและการเก็งกำไร	รางมาก่อน อุปสงค์มาทีหลัง
เครือข่ายข้อมูล	Grass 17 ล้านดอลลาร์ใน H1 2026 (รายงานเอง)	GRASS ราว 10 เท่าของรายได้ต่อปี	หนึ่งในไม่กี่กลุ่มที่มีผู้ซื้อจริง

เมื่อวางหกกลุ่มเทียบกัน ข้อสรุปสอดคล้องกัน: **เงินอุดหนุนเป็นเชื้อเพลิงสำหรับการเริ่มต้น ไม่ใช่โมเดลธุรกิจ** เมื่อ 13 พฤศจิกายน หมวด AI ของ CoinGecko มีมูลค่ารวมราว 18.6 พันล้านดอลลาร์ แต่ราว 56% ของจำนวนนั้นเป็นเลเยอร์ 1 อเนกประสงค์อย่าง NEAR และ ICP บวกโครงการระดับฉันทามติอย่าง TAO, TAO ที่สเตกไว้ (อาจนับซ้ำกับ TAO) และ Pearl โทเคน AI ในชั้นโปรโตคอลและชั้นแอปพลิเคชันรวมกันราว 8 พันล้านดอลลาร์หรือน้อยกว่า ประมาณการ ซึ่งใกล้เคียงกับมูลค่ารวมของบริษัทเอกชนสี่ราย: Tempo (5 พันล้านดอลลาร์), Nous (ตามรายงาน 1.5 พันล้านดอลลาร์), Prime Intellect (1 พันล้านดอลลาร์) และ Venice (1 พันล้านดอลลาร์) เงินร่วมลงทุนในคริปโตติดอยู่ที่ราว 4.5–5 พันล้านดอลลาร์ต่อไตรมาส โดยส่วนที่เพิ่มขึ้นมาจากโครงสร้างพื้นฐานที่เกี่ยวกับ AI ซึ่งส่วนแบ่งของยอดที่เปิดเผยเพิ่มจาก 8.1% เป็น 18.2% ในไตรมาสที่ 3 ปี 2026 และ 84.7% ของโทเคน 118 ตัวที่เปิดตัวในปี 2025 ซื้อขายต่ำกว่าราคาเปิดตัว สรุปในประโยคเดียว: **คริปโตทำหน้าที่กระจายและชำระราคา ส่วนหุ้นของบริษัทเป็นผู้รับกำไร** อย่างน้อยก็ในปี 2026

บทที่ 07

07

เศรษฐศาสตร์ ASIC

คำถามเรื่อง ASIC: หนึ่งสูตร สามไม้บรรทัด

หนึ่งสูตร: ทำไมประสิทธิภาพเลิกให้ผลหลัง k = 5

ให้ R เป็นรายได้รายวันต่อหน่วยการประมวลผล (รายได้นักขุดของเครือข่ายหารด้วยแฮชเรตของเครือข่าย) c เป็นต้นทุนรายวันของการประมวลผลหนึ่งหน่วยบนฮาร์ดแวร์ส่วนเพิ่ม (บน Pearl คือ GPU) และ k เป็นจำนวนเท่าที่ ASIC ถูกกว่าฮาร์ดแวร์ส่วนเพิ่มนั้น (c_ASIC = c/k) ตราบใดที่ GPU ยังเป็นเครื่องขุดส่วนเพิ่ม การเข้าและออกอย่างเสรีจะดัน R เข้าหา c ดังนั้น ASIC จะได้ R − c/k ต่อหน่วยการประมวลผล:

เพดานกำไรของ ASIC ≈ R × (1 − 1/k)

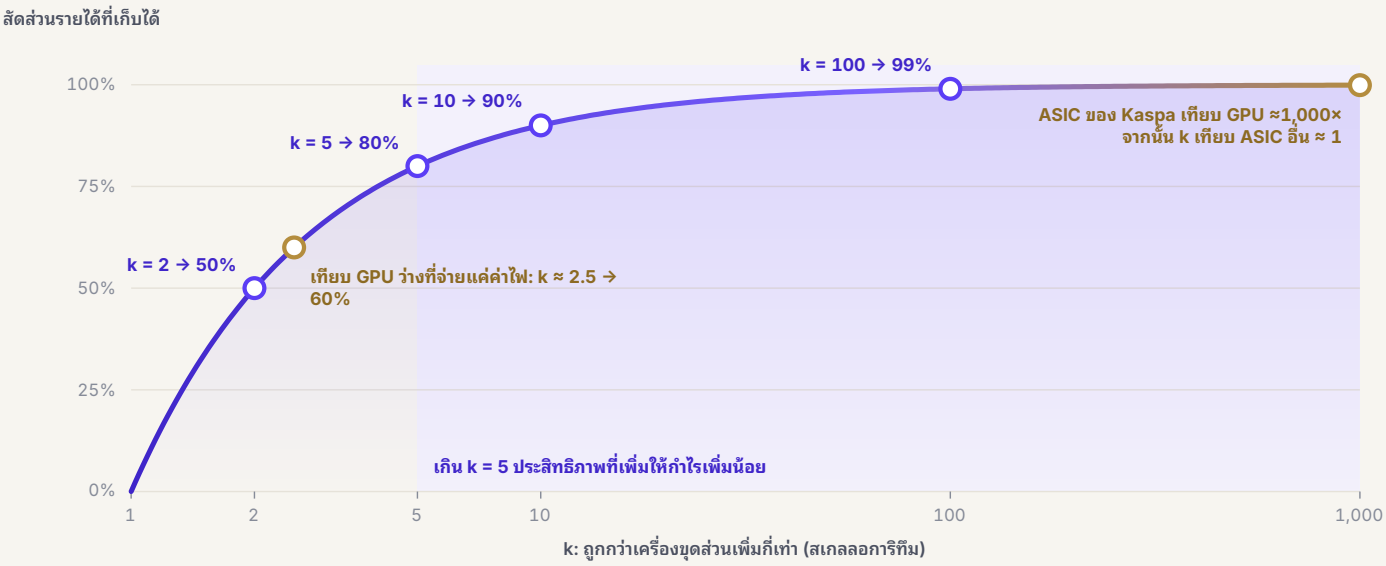
ในระดับทั้งเครือข่าย เพดานกำไรขั้นต้นต่อปีของฝูง ASIC โดยคร่าวคือรายได้นักขุดของเครือข่าย × (1 − 1/k) และเป็นเช่นนั้นจนกว่า ASIC จะเข้าแทนที่ GPU ในฐานะเครื่องขุดส่วนเพิ่มเท่านั้น

สูตรนี้มีผลสามข้อ ข้อแรก (1 − 1/k) เป็นฟังก์ชันเว้า: ที่ k = 2 ASIC เก็บรายได้ได้ 50% ที่ k = 5 ได้ 80% ที่ k = 10 ได้ 90% ที่ k = 100 ได้ 99% **เมื่อเกิน k = 5–10 ประสิทธิภาพที่เพิ่มแทบไม่เพิ่มกำไร** ขนาดของ R และของกองรายได้เป็นตัวตัดสิน ข้อสองจุดอ้างอิงเริ่มนับใหม่: เมื่อ ASIC เป็นเครื่องขุดส่วนเพิ่มแล้ว k ต้องวัดเทียบกับ ASIC อื่น ASIC ของ Kaspa มีความได้เปรียบด้านประสิทธิภาพเหนือ GPU ราว 1,000 เท่า ซึ่งซื้อได้เพียงลาภลอยช่วงสั้น วันนี้ KS7 ที่ดีที่สุดได้ 0.86 ดอลลาร์ต่อวัน ข้อสาม บนเชนที่เป็นมิตรกับ GPU อย่าง Pearl จุดอ้างอิงมักไม่ใช่ราคาซื้อ GPU แต่เป็น **ค่าไฟของ GPU ที่ว่างอยู่**: RTX 5090 ที่มีอยู่แล้ว ซึ่งกินไฟ 525 W ที่ค่าไฟ 0.10 ดอลลาร์ต่อ kWh มีค่าไฟราว 1.26 ดอลลาร์ต่อวัน และทำรายได้ราว 5–6 ดอลลาร์ ประมาณการ เมื่อ

เทียบกับ GPU ที่เป็นต้นทุนจมแล้ว แม้ ASIC ที่ใช้พลังงานหนึ่งในห้าก็มี k ที่มีผลจริงเพียงราว 2.5 หากค่าเสื่อมราคาเท่ากับค่าไฟ และมีเพดานกำไรราว 60% ของ R (สมมติฐานเพื่อการอธิบาย)

ภาพที่ 05 **ทำไมประสิทธิภาพเลิกให้ผล: ASIC เก็บรายได้ได้ไม่เกิน 1 − 1/k**

สัดส่วนรายได้ที่ ASIC เก็บได้เมื่อเทียบกับเครื่องขุดส่วนเพิ่ม ตามความได้เปรียบด้านต้นทุน k



แหล่งที่มา: แบบจำลองของผู้เขียน: เมื่อฮาร์ดแวร์ส่วนเพิ่มเป็นผู้กำหนดราคา รายได้ต่อหน่วยการประมวลผลจะลู่เข้าหาต้นทุนของมัน ตัวเลข Kaspa และ RTX 5090 มาจากรายงาน

สามไม้บรรทัด: กองรายได้ ขนาดคืนทุน และกติกา

ไม้บรรทัดอันแรกคือ **กองรายได้** เมื่อแปลงรางวัลบล็อกราว 2,893 BTC ในสัปดาห์แรกของเดือนพฤศจิกายน (ราว 222 ล้านดอลลาร์) เป็นรายปี นักขุดบิตคอยน์มีรายได้ราว 11.5 พันล้านดอลลาร์ต่อปี ประมาณการ นักขุดของ Pearl มีรายได้ราว 0.46–0.49 พันล้านดอลลาร์ (ที่ราคาปัจจุบัน) รางวัลผู้พิสูจน์ของ Aleo เมื่อแปลงจากไตรมาสที่ 3 ปี 2025 เป็นรายปีราว 31 ล้านดอลลาร์ และนักขุดของ Kaspa ที่รางวัลบล็อก 1.945 KAS และสิบบล็อกต่อวินาที ราว 18 ล้านดอลลาร์ ประมาณการ **รายได้นักขุดของ Kaspa ทั้งปีน้อยกว่าขั้นต่ำของงบออกแบบชิป 5/4 nm เพียงตัวเดียว (ราว 30 ล้านดอลลาร์)**

ไม้บรรทัดอันที่สองคือ **แฮชเรตที่ต้องใช้เพื่อคืนทุน NRE**:

h* = NRE ÷ [R × (1 − 1/k) × T × p]

โดย T คือช่วงเวลาที่ใช้งานได้เป็นวัน และ p คือความน่าจะเป็นที่กติกาจะไม่เปลี่ยนในช่วงนั้น สำหรับ Pearl ให้ R = 14 ดอลลาร์ต่อ PH/s ต่อวัน, k = 5, T = 365 วัน และ NRE 20 ล้านดอลลาร์ (ชิปอาร์เรย์ซ้ำบน N7/N6 อยู่ที่ขั้นต่ำของช่วง 5–30 ล้านดอลลาร์สำหรับโหนดระดับกลางจาก Silicon Analysts และ 15–100 ล้านดอลลาร์สำหรับ ASIC ประมวลผล 7/6 nm จาก AnySilicon เพราะชิปขุดใช้ IP น้อยและมีโครงสร้างซ้ำ เป็นสมมติฐานของผม) ที่ p = 1 ต้องใช้ราว **4.9 EH/s** ทำงานเต็มหนึ่งปี ราว 5% ของเครือข่าย ที่ p = 0.5 ต้องใช้ราว **9.8 EH/s** ราว 11% และถ้า R ลดกลับไปที่ระดับฤดูร้อนราว 10 ดอลลาร์ แม้ p = 1 ก็ต้องใช้ราว 6.9 EH/s ประมาณการ ขนาดนี้ไม่ได้เกินเอื้อม ปัญหาคือ p และ T

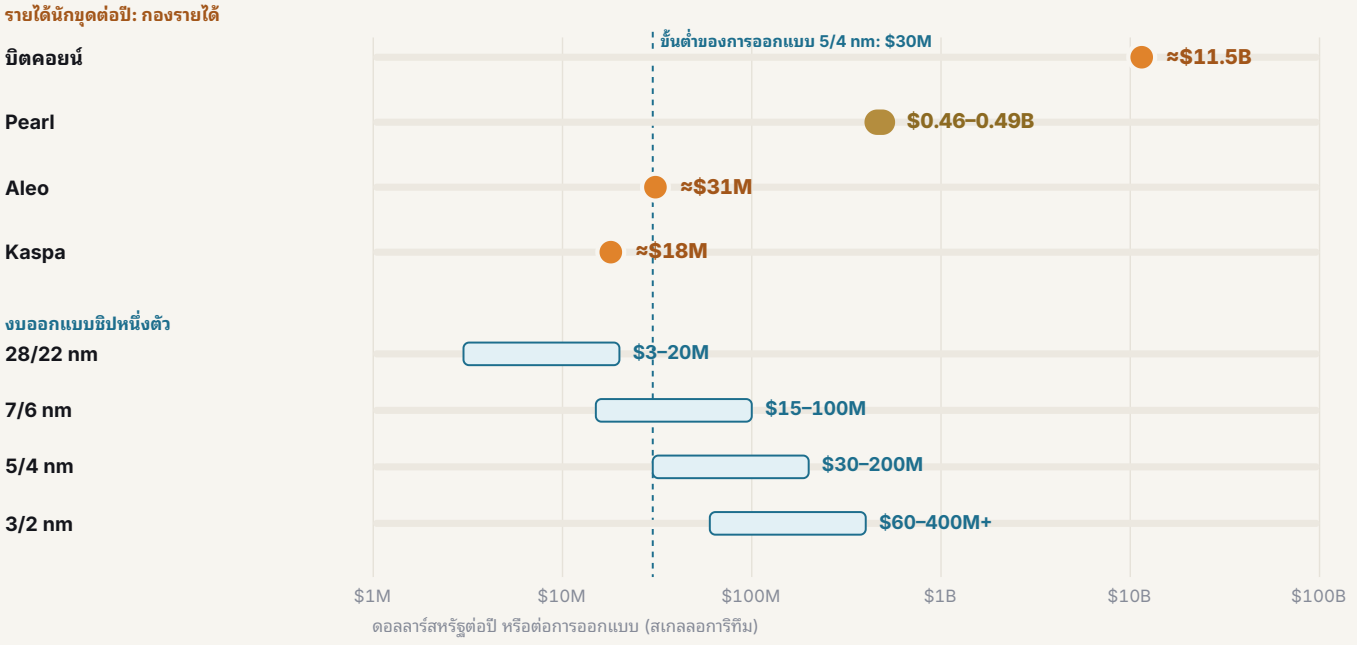
ไม้บรรทัดอันที่สามคือ **ความเสี่ยงด้านกติกา** และมักเป็นตัวชี้ขาด ฟอร์กต่อต้าน ASIC ของ Monero ในปี 2018 ทำให้ X3 ของ Bitmain ค้างเติ่ง และทำให้เครือข่ายเสียแฮชเรตไปเกือบครึ่งชั่วคราว ฟอร์กของ Sia ในปีเดียวกันทำให้เครื่องขุดของ Bitmain และ Innosilicon ใช้การไม่ได้ ขณะที่ราคามือสองของเครื่องขุด Obelisk ซึ่งเชื่อมโยงกับนักพัฒนา เพิ่มจาก 2,499 ดอลลาร์เป็นกว่า 5,000 ดอลลาร์ ราวหนึ่งเดือนเศษหลัง Merge ของ Ethereum อดีตนักขุด Ethereum ไม่ถึง 20% หางานใหม่ได้ **ใครคุม**

ฟังก์ชันงาน คนนั้นกำหนดชะตาของนักลงทุนฮาร์ดแวร์ และ Sia แสดงให้เห็นว่า “การต้านทาน ASIC” อาจกลายเป็นคูเมืองของนักพัฒนาเอง ฮาร์ดแวร์ที่มีตลาดที่สองอยู่รอด: หลัง Merge GPU ย้ายไปสู่ AI และคลาวด์ ขณะที่ ASIC อัลกอริทึมเดียวไปแออัดกันใน Ethereum Classic หน้าต่างโอกาสก็หดลงด้วย: ASIC ตัวแรกของ Kaspa ใช้เวลาราว 16 เดือน ของ Aleo ราวห้าเดือน และผู้ที่ไปถึงพรมแดนประสิทธิภาพก่อนมักเป็นผู้ท้าชิง มากกว่าผู้เล่นเดิม

เครือข่าย	รายได้นักขุดต่อปี (โดยประมาณ)	ฮาร์ดแวร์ส่วนเพิ่มหรือฮาร์ดแวร์อ้างอิง	ความมั่นคงของกติกา	k ที่ผู้เล่นใหม่ได้	ข้อวินิจฉัย
บิตคอยน์	11.5 พันล้านดอลลาร์	Bitmain S23 XP Hyd (8.9 J/TH, 25.5 ดอลลาร์ต่อ TH)	สูงมาก	≤1	กองรายได้ใหญ่ แต่ k เป็นของผู้เล่นเดิม
Pearl	0.46–0.49 พันล้านดอลลาร์ (ที่ราคาปัจจุบัน)	GPU ที่มีอยู่แล้วหรือว่างอยู่ (ต้นทุนส่วนเพิ่ม ≈ ค่าไฟ) และ GPU ศูนย์ข้อมูล	ต่ำมาก: สี่ฟอร์กในสองเดือน; รายชื่ออนุญาต FP8 รอดำเนินการ	อาจ 3–10 เท่าด้านประสิทธิภาพ บทวิเคราะห์ ; ต่ำกว่านั้นเมื่อเทียบ GPU ที่เป็นต้นทุนจม	กองรายได้ใหญ่พอ แต่ความเสี่ยงด้านกติกาคัดค้าน
Aleo	ราว 31 ล้านดอลลาร์ (ไตรมาส 3 ปี 2025 แปลงเป็นรายปี)	IceRiver AE2 และเครื่องเฉพาะทางอื่นๆ	ปานกลาง: เกณฑ์สเตกของ ARC-46	≈1	กองรายได้เล็ก และเกณฑ์กำลังสูงขึ้น
Kaspa	ราว 18 ล้านดอลลาร์	KS7 (77 J/TH)	สูง	≈1	รายได้ต่อปีต่ำกว่าขั้นต่ำของงบออกแบบ 5/4 nm หนึ่งตัว

ภาพที่ 06 **กองรายได้ส่วนใหญ่เล็กกว่างบออกแบบชิปหนึ่งตัว**

รายได้นักขุดต่อปีของสี่เครือข่าย เทียบกับต้นทุนออกแบบชิปหนึ่งตัวตามโหนดการผลิต สเกลลอการิทึม



แหล่งที่มา: Hashrate Index (บิตคอยน์ พฤศจิกายน 2026), prlstats และ Kryptex (Pearl), Blockworks (Aleo ไตรมาส 3 ปี 2025), minerstat (Kaspa), AnySilicon (งบออกแบบ 2026); ผู้เขียนแปลงเป็นรายปี

ช่อง “กองรายได้ใหญ่ กติกามั่นคง และ k > 1” ว่างเปล่าในวันนี้: บิตคอยน์มีกองรายได้และกติกาที่มั่นคง แต่ k เป็นของผู้เล่นเดิม กองรายได้ของ Pearl ใหญ่พอ แต่กติกามั่นคงน้อยที่สุด Kaspa และ Aleo มีกติกามั่นคงพอ แต่กองรายได้เล็กเกินไป

บทที่ 08

08 ชิปทีละประเภท
ชิปห้าประเภท ห้าคำตอบ

PoUW แบบ Pearl: ไม่ทำชิป INT7 เฉพาะทาง ให้เดิมพันกับเครื่องขุดที่ให้บริการอนุมานได้

เหตุผลสนับสนุน ASIC มีจริง วันนี้มีเพียงราว 1% ของบล็อกที่มีงาน AI ข้อโต้แย้งทางเศรษฐศาสตร์ในไวต์เปเปอร์ที่ว่า ASIC เฉพาะทางเสียความสามารถในการทำงานที่มีประโยชน์และจึงได้รายได้น้อยกว่า จึงอ่อนที่สุดในตอนนี้ ส่วนที่แพงที่สุดของ GPU คือ HBM และการแพ็กเกจ CoWoS คิดเป็นราว 63% ของต้นทุนการผลิต H100 โดยไดลอจิกคิดเป็นเพียงราว 9% ชิปที่ไม่ทำอะไรนอกจากการคูณสะสมเลขจำนวนเต็ม 7 บิต เก็บตัวถูกดำเนินการไว้ใน SRAM บนชิป และทำงานที่แรงดันต่ำ อาจทำได้ดีกว่า GPU ที่โหนดเดียวกัน 3 ถึง 10 เท่าในด้านพลังงานต่อออป และได้เปรียบมากกว่านั้นในด้านต้นทุนต่อออป [บทวิเคราะห์; ยังไม่มีการวัด] เหตุผลคัดค้านหนักกว่า ระยะเวลารอของโรงหล่อที่โหนดล้ำสมัยอาจเกิน 50 สัปดาห์ และทีมใหม่ต้องใช้ราว 18 ถึง 24 เดือนตั้งแต่เริ่มจนถึงการผลิตจำนวนมาก ประมาณการ กว่าชิปจะมาถึง Pearl อาจย้ายไปเป็น FP8 แล้วตามด้วย FP4 แล้วก็ได้ และเพราะ PIP-3 กำหนดเลขคณิตที่ยอมรับผ่านรายชื่ออุปกรณ์ที่อนุญาต ชิป INT7 จะหยุดใช้งานได้ทันทีเมื่อ PIP-3 เปิดใช้

ข้อสรุปของผมจึงเป็น: **อย่าเทปเอาต์ชิป INT7 เฉพาะทางสำหรับ Pearl ตอนนี้** มีเดิมพันที่ดีกว่าอยู่สองทาง ทางแรกคือตัวเร่งที่เน้นการอนุมานและขุดได้ด้วย: สร้างรอบ GEMM INT8/FP8 ทั่วไปและตัวถูกดำเนินการที่อยู่ใน SRAM เพื่อที่หากถูกปิดช่องทางการขุด มูลค่าจะไปตกในตลาดการอนุมานจริง ทางที่สองคือเข้าร่วมการถกเถียงเรื่องโปรโตคอล และผลักดันเลขคณิตมาตรฐานที่เป็นกลางต่อผู้ขายซึ่ง ASIC ทำได้ตรงทุกบิต พร้อมคำมั่นเรื่องความมั่นคงของอัลกอริทึม ซึ่งเป็นเงื่อนไขเบื้องต้นที่ช่อง “กองรายได้ใหญ่ กติกามั่นคง k > 1” จะมีอยู่ได้ ระหว่างนี้ สถาปัตยกรรมตรวจสอบได้บน FPGA หรือเวเฟอร์หลายโครงการ (MPW) บนโหนดรุ่นเก่า ซึ่งคุมความเสี่ยง NRE ไว้ในระดับหนึ่งล้านดอลลาร์ บทที่ 10 แปลงเรื่องนี้เป็นงบแบบเป็นขั้น

เครื่องขุดแบบบิตคอยน์: ผู้เล่นเดิมเป็นเจ้าของ k และไฟฟ้าคือสินทรัพย์ที่ขาดแคลน

ความเสี่ยงด้านกติกาของบิตคอยน์ต่ำที่สุด แต่การแข่งขันดัน k ของผู้เล่นใหม่ให้ต่ำกว่า 1 S23 XP Hyd ของ Bitmain ทำได้ 8.9 J/TH ขายที่ 25.5 ดอลลาร์ต่อ TH และส่งมอบตั้งแต่เดือนพฤศจิกายน ประสิทธิภาพของรุ่นเรือธงลดจาก 17.5 J/TH ของ S21 ในปี 2023 เหลือ 8.9 J/TH ราว 49% ในสามปี ผู้เล่นใหม่ที่ส่งมอบจำนวนมากในปี 2027–2028 ต้องทำได้ราว 8 J/TH หรือดีกว่าจึงจะยืนอยู่ได้ ประมาณการ ผู้ท้าชิงกำลังถอนตัว: Canaan รายงานรายได้ไตรมาสที่ 2 ปี 2026 ที่ 31.9 ล้านดอลลาร์ ขาดทุนขั้นต้น 29.3 ล้านดอลลาร์ และให้แนวโน้มไตรมาสที่ 3 เพียง 11–15 ล้านดอลลาร์ ยอดขายเครื่องขุดของ Bitdeer เองให้ผู้ซื้อภายนอกในไตรมาสที่ 3 ลดเหลือ 0.3 ล้านดอลลาร์ โดยอัตรากำไรขั้นต้นของกลุ่มอยู่ที่ 3.4% Auradine เปลี่ยนชื่อเป็น Velaura AI และหันไปทำชิป AI และ Core Scientific ยกเลิกคำสั่งซื้อเครื่อง 3 nm ราว 15 EH/s จากหน่วย Proto ของ Block โดยบันทึกผลขาดทุน 41.9 ล้านดอลลาร์

อุปสงค์ก็หดตัวเช่นกัน เมื่อ 9 พฤศจิกายน hashprice อยู่ที่ราว 31.89 ดอลลาร์ต่อ PH/s ต่อวัน เส้นราคาล่วงหน้าถึงเดือนเมษายน 2027 เฉลี่ยราว 32.7 ดอลลาร์ และค่าธรรมเนียมคิดเป็นเพียง 0.72% ของรางวัลบล็อก ที่ค่าไฟ 0.045 ดอลลาร์ต่อ kWh S23 XP Hyd คืนทุนในราว 3.1 ปี ประมาณการ ใกล้กับการลดรางวัลครึ่งหนึ่งครั้งถัดไปที่คาดว่าจะเกิดราวปี 2028 ที่สำคัญกว่าคือต้นทุนค่าเสียโอกาสของไฟฟ้า: สัญญาเช่าสำหรับ AI และการประมวลผลสมรรถนะสูงตอนนี้ได้ราว 1.67–1.9 ล้านดอลลาร์ต่อเมกะวัตต์ตามสัญญาต่อปี ซึ่งที่รายได้จากการดำเนินงานสุทธิ 80% ถึง 85% เท่ากับราว 1.5 ถึง 2.9 เท่าของกำไรขั้นต้นจากการขุดต่อเมกะวัตต์ (ราว 0.56–0.9 ล้านดอลลาร์) [ประมาณการ; ฐานเมกะวัตต์ของทั้งสองต่างกัน] Core Scientific จะหยุดขุดใน

ไตรมาสที่ 4 Bitfarms ซึ่งตอนนี้คือ Keel ปิดไซต์ขุดในสหรัฐฯ ทั้งหมด IREN มีรายได้ 200 ล้านดอลลาร์จากคลาวด์ AI และเพียง 28 ล้านดอลลาร์จากการขุดในไตรมาสเดียว **ไฟฟ้าและกำลังของศูนย์ข้อมูลคือสินทรัพย์ เครื่องขุดคือวัสดุสิ้นเปลือง**

การคำนวณ NRE ก็ไม่ปรานีเช่นกัน ประสิทธิภาพระดับแนวหน้าตอนนี้ต้องใช้กระบวนการระดับ 3 nm (ตามแหล่งข้อมูลทุติยภูมิ MicroBT ใช้ Samsung SF3 และชิปของ Auradine และ Proto เป็น 3 nm) หน้ากากมีต้นทุนราว 10–20 ล้านดอลลาร์ และ NRE ทั้งหมด 30–100 ล้านดอลลาร์หรือมากกว่า ที่ราคา 20–25 ดอลลาร์ต่อ TH และอัตรากำไรขั้นต้น 30% ทุก 1 EH/s ที่ขายได้ให้กำไรขั้นต้นราว 6–7.5 ล้านดอลลาร์ การครอบคลุมค่าวิจัยพัฒนาต่อปีในระดับของ Bitdeer ราว 110–160 ล้านดอลลาร์ จึงต้องขายให้ผู้ซื้อภายนอกราว 20–25 EH/s ต่อปี ราว 2% ของเครือข่าย ประมาณการ ข้อสรุป: **ทีมใหม่ไม่ควรเข้าตลาดเครื่องขุดบิตคอยน์เพื่อขายทั่วไป** ยังมีสามกรณีที่สมเหตุสมผล: การบูรณาการแนวตั้งกับไฟฟ้าราคาถูกของตนเอง รูปแบบเฉพาะกลุ่มเช่นเครื่องขุดสำหรับบ้านและเครื่องขุดที่ให้ความร้อน (Canaan ก็กำลังเปิดตัวผลิตภัณฑ์ให้ความร้อนสำหรับบ้าน) และการมีส่วนในแผนงานของผู้เล่นเดิมผ่าน IP และบริการออกแบบ

ชิปเทรน AI: ไม่ใช่สำหรับทีมเล็ก

สำหรับชิปเทรน k หาได้ยากที่สุด ที่รุ่นและความแม่นยำเดียวกัน Ironwood ของ Google นำ B300 ในประสิทธิภาพ FP8 ระดับชิปเพียงราว 1.1 ถึง 1.8 เท่า (กำลังไฟของ Ironwood เป็นการประมาณของผู้สื่อข่าว) ประมาณการ ตัวเลข 30 ถึง 80 เท่าที่มักอ้างจากงานวิจัย TPU v1 วัดเทียบกับ K80 ซึ่งไม่มีเทนเซอร์คอร์ คันโยกที่แท้จริงคือรูปแบบตัวเลข: การย้ายจาก FP8 เป็น FP4 มักให้ปริมาณงาน 2 ถึง 3 เท่าบนซิลิคอนเดิม และ NVIDIA เคลื่อนที่เร็วที่สุดในเรื่องนี้: อ้างว่า Rubin ทำการอนุมาน NVFP4 ได้ 50 PF และผลิตจำนวนมากในครึ่งหลังของปี 2026 เกณฑ์ด้านต้นทุนและอุปทานสูงกว่านั้นอีก: งบออกแบบ 3/2 nm อยู่ที่ 60–400 ล้านดอลลาร์หรือมากกว่า กำลังการผลิต CoWoS ในปี 2026 ราว 106,000–130,000 เวเฟอร์ต่อเดือน และมีรายงานว่า TSMC กำลังพิจารณาขึ้นราคาสูงสุด 20% ที่ราคาสัญญา 15.6 ดอลลาร์ต่อ GB สำหรับ HBM3E แค่ HBM ในตัวเร่งขนาด 288 GB ก็มีต้นทุนราว 4,500 ดอลลาร์ ประมาณการ ตลาดนี้เป็นของการพัฒนาร่วมกับไฮเปอร์สเกลเลอร์แบบ Broadcom (บทที่ 11) ข้อสรุป: **ทีมเล็กไม่ควรสร้างชิปเทรนแบบแยกเดี่ยว** ให้สร้าง IP ชิปเล็ต หรือบริการออกแบบแทน

ชิปอนุมาน AI: แลกความเป็นสากลกับ k และคุ้ม

การอนุมานคือที่ที่ ASIC มีพื้นที่มากที่สุด เพราะซื้อ k ได้ด้วยการยอมสละความเป็นสากล Etched อ้างว่าเซิร์ฟเวอร์แปดชิปของตนรัน Llama-70B ได้กว่า 500,000 โทเคนต่อวินาที ราว 20 เท่าของ H100 แปดใบ Taalas ฝังน้ำหนักโมเดลแบบตายตัวลงใน mask ROM และอ้างว่ากินไฟต่ำกว่า 10 เท่าและต้นทุนการผลิตต่ำกว่า 20 เท่า (ทั้งสองข้อ ผู้ขายอ้าง) สถาปัตยกรรมที่เน้น SRAM เลี่ยง HBM และ CoWoS ได้ และใช้โหนดที่จัดการได้ง่ายกว่า: Taalas ใช้ TSMC 6 nm และ Groq ย้ายจาก 14 nm ไปเป็น Samsung 4 nm ช่องทางกระจายแบบคริปโตไม่ช่วยชิปอนุมาน เพราะเครือข่ายการประมวลผลแบบกระจายศูนย์ทำงานบน NVIDIA เกือบทั้งหมด ทางเข้าคือเศรษฐศาสตร์การจัดซื้อของนีโอคลาวด์ และตลาดการอนุมานในประเทศขนาดใหญ่ของจีน ซึ่งศูนย์ข้อมูลที่รัฐให้ทุนต้องใช้ชิป AI ในประเทศ การนำเข้า H200 ถูกจำกัด และปักกิ่งต้องการให้บริษัทอินเทอร์เน็ตรายใหญ่รันการอนุมานบนชิปในประเทศ ข้อสรุป: **คุ้มค่าจะสร้าง** ตราบใดที่มีลูกค้าชัดเจน และระดับการผูกติดกับโมเดลและสถาปัตยกรรมเป็นการเลือกอย่างตั้งใจ มันยังเป็นตลาดที่สองที่ดีที่สุดสำหรับชิปขุด PoUW

ตัวเร่ง ZK และ FHE: ป้องกันความเสี่ยงด้วย ZK วางตำแหน่งแต่เนิ่นใน FHE

GPU ไล่ทันการพิสูจน์ ZK แบบเรียลไทม์แล้ว: SP1 Hypercube ของ Succinct พิสูจน์บล็อก Ethereum ได้ 99.7% ภายใน 12 วินาทีบน RTX 5090 จำนวน 16 ใบ Cysic อ้างว่าชิป ZK C1 ของตนมีประสิทธิภาพ 10 ถึง 100 เท่าของผู้พิสูจน์บน GPU และจะส่งมอบในปี 2026 แต่อุปสงค์ส่วนใหญ่มาจากรางวัลโทเคนของตนเอง ผู้ขายอ้าง ผมไม่พบหลักฐานสาธารณะว่า ASIC สำหรับปริศนาของ Aleo ถูกใช้ในการพิสูจน์ ZK จริง การพิสูจน์ ZK เป็นงานเลขจำนวนเต็มและกำหนดผลได้แน่นอน มีรูปแบบใกล้กับงานของชิปขุด แต่ระบบการพิสูจน์เปลี่ยนเร็ว (SP1, RISC Zero, Halo2, Plonky2 และอื่นๆ) ข้อสรุป: **รอและป้องกันความเสี่ยง** ด้วยเอนจิน MSM, NTT และแฮชที่โปรแกรมได้ เริ่มบน FPGA ก่อน แล้วจึงเป็น ASIC

FHE เป็นพื้นฐานการเข้ารหัสที่ผูกกับฮาร์ดแวร์มากที่สุดและมีศักยภาพสูงสุด Heracles ของ Intel บน Intel 3 ขนาด 197 mm², 176 W และ HBM3 48 GB รันการดำเนินการ FHE ได้เร็วกว่า Xeon 24 คอร์ 1,074 ถึง 5,547 เท่า แต่ Intel ไม่มีแผนเชิงพาณิชย์ และหัวหน้าโครงการลาออกในเดือนธันวาคม 2025 Niobium กำลังสร้างบน Samsung 8LPU โดยมี SemiFive เป็นพันธมิตรด้านการออกแบบ (สัญญาพัฒนาราว 6.86 ล้านดอลลาร์) และระดมทุนได้กว่า 23 ล้านดอลลาร์ Optalysys ซึ่งใช้แนวทางโฟโตนิกส์ ระดมทุนได้ 31 ล้านดอลลาร์ และกล่าวว่าชิปเชิงพาณิชย์เต็มรูปแบบยังห่างอีกสองถึงสามปี และงานออกแบบ FPGA แบบโอเพนซอร์สของ Zama ทำได้ราว 13,000 programmable bootstrap ต่อวินาทีที่ราว 200 W ซึ่งเป็นค่าฐานสาธารณะที่ ASIC ใดๆ ต้องชนะอย่างชัดเจน ข้อสรุป: **วางตำแหน่งแต่เนิ่นและตรวจสอบด้วยสินทรัพย์เบา** สำหรับหน้าต่างโอกาสราวปี 2027–2029 อุปสงค์มีแนวโน้มจะยึดอยู่กับ AI แบบเข้ารหัสสำหรับองค์กรและภาครัฐ มากกว่าสกุลเงินดิจิทัลเอง และ FHE ต้องใช้ HBM โครงสร้างต้นทุนจึงคล้ายชิปเทรน

ประเภทชิป	k มาจากไหน	อุปสงค์และกองรายได้	NRE และเกณฑ์	ความเสี่ยงด้านกติกาหรือการผูกติด	ตลาดที่สอง	ข้อวินิจฉัย
ชิปขุด PoUW แบบ Pearl	ตัด HBM, CoWoS และหน่วยอเนกประสงค์ออก: อาจ 3–10 เท่าใน J/op บทวิเคราะห์	รายได้นักขุดราว 0.46–0.49 พันล้านดอลลาร์ต่อปี ผันผวน	อาร์เรย์ซ้ำบน N7/N6 ราว 10–30 ล้านดอลลาร์ บทวิเคราะห์	สูงมาก: โปรโตคอลเป็นผู้กำหนดรูปแบบตัวเลขและรายชื่ออุปกรณ์ที่อนุญาต	การอนุมาน AI หากสร้างเป็น GEMM INT8/FP8 ทั่วไป	ไม่ทำชิป INT7 เฉพาะทาง สร้างแบบเน้นการอนุมานและขุดได้ หรือรอเลขคณิตมาตรฐาน
ชิปขุดบิตคอยน์ SHA-256	k ≤ 1 เมื่อเทียบ Bitmain	ราว 11.5 พันล้านดอลลาร์ต่อปี ตลาดเพื่อขายทั่วไปอยู่ในมือสองบริษัท	หน้ากากระดับ 3 nm 10–20 ล้านดอลลาร์; NRE ทั้งหมด 30–100 ล้านดอลลาร์หรือมากกว่า	ต่ำ (อัลกอริทึมมั่นคง); การแข่งขันและการลดรางวัลครึ่งหนึ่งเสี่ยงสูง	แทบไม่มี (อัลกอริทึมเดียว)	ผู้เล่นใหม่อยู่ห่างจากตลาดเพื่อขายทั่วไป ทำได้เฉพาะเมื่อมีไฟฟ้าเอง รูปแบบเฉพาะกลุ่ม หรือ IP
ชิปเทรน AI	systolic array รุ่นเดียวกันได้เพียง 1.1–1.8 เท่า รูปแบบตัวเลขคือคันโยก	ใหญ่ที่สุด (Broadcom มีรายได้เซมิคอนดักเตอร์ AI ราว 58 พันล้านดอลลาร์ในปีงบประมาณ 2026)	ออกแบบ 3/2 nm 60–400 ล้านดอลลาร์หรือมากกว่า; CoWoS และ HBM	จังหวะออกรุ่นรายปีของ NVIDIA	—	ทีมเล็กทำ IP ชิปเล็ต และบริการออกแบบ ไม่ทำชิปแยกเดี่ยว
ชิปอนุมาน AI	สละความเป็นสากล: อ้าง 10–20 เท่า ผู้ขายอ้าง	ใหญ่และกระจัดกระจาย; ตลาดในประเทศของจีนแผ่นดินใหญ่ใหญ่ในตัวเอง	สถาปัตยกรรม SRAM ที่ N6/N7–N4 จัดการได้	การผูกติดกับโมเดลและสถาปัตยกรรม	นี่คือตลาดหลัก	คุ้มค่าจะสร้าง: เซ็นสัญญาลูกค้าก่อนแล้วค่อยเลือกระดับการผูกติด
ตัวเร่ง ZK	อ้าง 10–100 เท่าของ GPU ผู้ขายอ้าง	เล็ก ขับเคลื่อนด้วยรางวัลโทเคนเป็นหลัก	N7/N6 แล้วตามด้วย N5; FPGA ก่อน	ระบบการพิสูจน์เปลี่ยนเร็ว	จำกัดนอก ZK	รอและป้องกันความเสี่ยง: เอนจิน MSM และ NTT ที่โปรแกรมได้
ตัวเร่ง FHE	1,074–5,547 เท่าของ Xeon (Heracles)	ยังไม่ก่อตัว; ยึดอยู่กับองค์กรและภาครัฐ	ต้องใช้ HBM; โครงสร้างต้นทุนใกล้ชิปเทรน	มาตรฐานและระบบนิเวศยังไม่ลงตัว	การประมวลผลแบบเข้ารหัส AI ส่วนตัว	วางตำแหน่งแต่เนิ่น ตรวจสอบด้วยสินทรัพย์เบา; หน้าต่างโอกาส 2027–2029

ภาพที่ 07 **ชิปห้าประเภท ห้าคำตอบ**

ความได้เปรียบด้านต้นทุน k มาจากไหน ซิลิคอนขายเข้าตลาดใดได้ถ้ากติกาเปลี่ยน และข้อสรุป



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

แหล่งที่มา: การสังเคราะห์ของผู้เขียนจากบทที่ 7 และ 8; ตัวเลขของผู้ขายติดป้ายไว้ในรายงาน

บทที่ 09



09

โหนดการผลิต

โหนดการผลิต: ให้ต้นทุนตลอดอายุที่ใหญ่ที่สุดเป็นตัวเลือก

บันไดต้นทุน: เวเฟอร์โหนดล้ำสมัยแพงกว่าโหนดรุ่นเก่าสิบเท่า

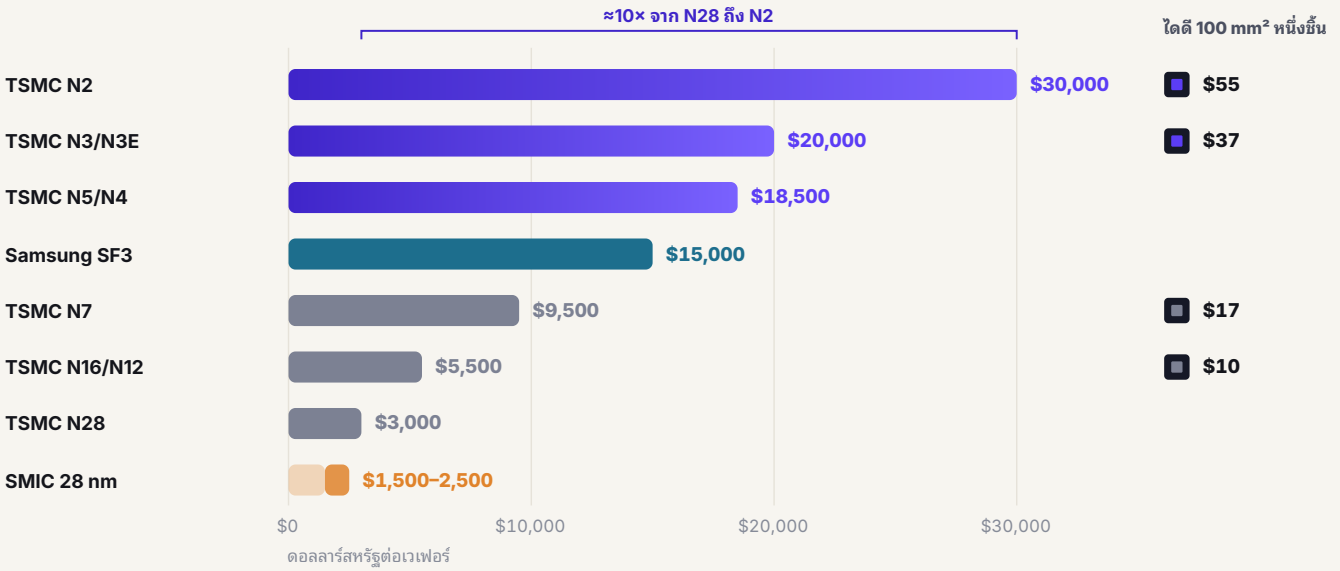
เริ่มจากบัญชี เวเฟอร์โหนดล้ำสมัยตอนนี้มีราคาราวสิบเท่าของโหนดรุ่นเก่า: TSMC N2 ราว 30,000 ดอลลาร์ N3 ราว 20,000 ดอลลาร์ N7 ราว 9,500 ดอลลาร์ และ N28 ราว 3,000 ดอลลาร์ Samsung SF3 ราว 15,000 ดอลลาร์ ต่ำกว่า N3 ราว 25% และมีรายงานว่าต้นทุนการผลิตระดับ 7 nm ของ SMIC สูงกว่าของ TSMC 40% ถึง 50% มีรายงานว่า TSMC จะขึ้นราคาอีก 3% ถึง 6% ตั้งแต่เดือนมกราคม 2027 โดยขึ้นมากที่สุดสำหรับ 2 nm และ 3 nm ต้นทุนหน้ากากและการออกแบบต่างกันยิ่งกว่านั้น และแหล่งข้อมูลไม่ตรงกัน ตารางจึงแสดงไว้ทั้งสองค่า

โหนด	ราคาเวเฟอร์ (ประมาณการ)	ชุดหน้ากาก (ค่าจากแบบจำลอง / ช่วงอ้างอิง)	งบออกแบบ (AnySilicon)	MPW (ต่อโครงการ)
TSMC N2	ราว 30,000 ดอลลาร์	—	3/2 nm: 60–400 ล้านดอลลาร์หรือมากกว่า	—
TSMC N3/N3E	ราว 20,000 ดอลลาร์ (18,000–27,000 ดอลลาร์)	15 ล้านดอลลาร์ / 10–20 ล้านดอลลาร์	เช่นเดียวกับข้างต้น	—

โหนด	ราคาเวเฟอร์ (ประมาณการ)	ชุดหน้ากาก (ค่าจากแบบจำลอง / ช่วงอ้างอิง)	งบออกแบบ (AnySilicon)	MPW (ต่อโครงการ)
Samsung SF3	ราว 15,000 ดอลลาร์ (13,000–17,000 ดอลลาร์)	12 ล้านดอลลาร์	—	—
TSMC N5/N4	ราว 18,500 ดอลลาร์ (16,000–21,300 ดอลลาร์ ความเชื่อมั่นต่ำ)	6.5 ล้านดอลลาร์ / 10–20 ล้านดอลลาร์	5/4 nm: 30–200 ล้านดอลลาร์	—
TSMC N7	ราว 9,500 ดอลลาร์ (8,100–10,900 ดอลลาร์)	3.5 ล้านดอลลาร์ / 5–10 ล้านดอลลาร์	7/6 nm: 15–100 ล้านดอลลาร์	เชิงพาณิชย์ 100,000–500,000 ดอลลาร์
TSMC N16/N12	ราว 5,500 ดอลลาร์	1.8 ล้านดอลลาร์	16/12 nm: 6–40 ล้านดอลลาร์	เชิงวิชาการราว 35,000 ดอลลาร์ต่อ mm²
TSMC N28	ราว 3,000 ดอลลาร์	1 ล้านดอลลาร์ / 1–3 ล้านดอลลาร์	28/22 nm: 3–20 ล้านดอลลาร์	เชิงพาณิชย์ 50,000–150,000 ดอลลาร์; เชิงวิชาการราว 12,000 ดอลลาร์ต่อ mm²
SMIC 28 nm	1,500–2,500 ดอลลาร์	—	—	—
SMIC ระดับ 7 nm (N+2/N+3)	สูงกว่า TSMC 40–50%	—	—	—

ภาพที่ 08 **เวเฟอร์โหนดล้ำสมัยแพงกว่าโหนดรุ่นเก่าสิบเท่า**

ราคาประมาณต่อเวเฟอร์ 300 mm ตามโหนด และต้นทุนไดดีหนึ่งชิ้นขนาด 100 mm² ที่ yield 85%



แหล่งที่มา: ราคาเวเฟอร์ของ Silicon Analysts (ประมาณการ N5 ความเชื่อมั่นต่ำ), Wccftech (SMIC); ต้นทุนไดเป็นการคำนวณของผู้เขียนโดยสมมติ yield

จากนั้นคือฟิสิกส์ การวัดคลาสสิกของ Horowitz แสดงว่าคำสั่งที่โปรแกรมได้หนึ่งคำสั่งใช้พลังงานราว 70 pJ ขณะที่ตัวเลขคณิตเองใช้เพียงไม่กี่พิโคจูล: **เลขคณิตถูก โอเวอร์เฮดแพง** ซึ่งเป็นเหตุผลที่ systolic array และชิปเฉพาะทางมีอยู่ การลดขนาดลายวงจรลงครึ่งหนึ่งทำให้พลังงานต่อออปลดลงราวครึ่งหนึ่ง แต่แรงดันไฟเลี้ยงหยุดนิ่งอยู่ราว 0.9–1.1 V เวเฟอร์ N2 แพงกว่า N3 50% ถึง 67% ดังนั้นหากความหนาแน่นไม่เพิ่มเกินราว 1.5 เท่า ต้นทุนต่อทรานซิสเตอร์จะเพิ่มขึ้นแทนที่จะลดลง ประมาณการ: โหนดล้ำสมัยต้องคุ้มทุนด้วยประสิทธิภาพและสมรรถนะ สำหรับไดขนาด 100 mm² ที่ yield 85% ต้นทุนของไดดีแต่ละชิ้นอยู่ที่ราว

55 ดอลลาร์ที่ N2, 37 ดอลลาร์ที่ N3, 29–34 ดอลลาร์ที่ N5, 17 ดอลลาร์ที่ N7 และ 10 ดอลลาร์ที่ N16 [ประมาณการ; yield เป็นค่าสมมติ]

การเลือกโหนดตามประเภทชิปและทีม

กฎของผมคือ: **ให้รายการที่ใหญ่ที่สุดในต้นทุนตลอดอายุของชิปเป็นตัวเลือกโหนด** สำหรับชิปประมวลผลที่ทำงานเต็มกำลังตลอดเวลา (การขุด ZK) ค่าไฟครองต้นทุนตลอดอายุ เมื่ออัลกอริทึมมั่นคงแล้วจึงควรย้ายไปที่ N5/N4 และ N3/SF3 ขณะที่อัลกอริทึมยังไม่ลงตัวหรือเงินทุนจำกัด ให้ตรวจสอบบน MPW ของโหนดรุ่นเก่าก่อน สำหรับชิปเทรน คอขวดคือการแพ็กเกจและหน่วยความจำ (HBM และ CoWoS คิดเป็นราว 63% ของรายการวัสดุของ H100) โหนดจึงเป็นเรื่องรอง สำหรับชิปอนุมาน การตัดสินใจแรกคือสถาปัตยกรรมหน่วยความจำ: การออกแบบที่เน้น SRAM อยู่ที่ N6/N7 ถึง N4 ได้ ขณะที่การออกแบบที่ใช้ HBM รับโครงสร้างต้นทุนของชิปเทรนมาด้วย FHE ถูกจำกัดด้วยแบนด์วิดท์หน่วยความจำและเลี่ยง HBM ไม่ได้

ประเภทชิป	สตาร์ตอัปหรือทีมเล็ก	บริษัทที่ขยายขนาดแล้ว	เส้นทางจีนแผ่นดินใหญ่	ประเด็นสำคัญ
ชิปขุด PoUW แบบคูณเมทริกซ์ (อยู่ใน SRAM ไม่มี HBM)	ทำต้นแบบบน MPW 28 nm หรือ 16/12 nm; ผลิตจำนวนมากครั้งแรกบน N7/N6; ไป N5/N4 เมื่ออัลกอริทึมและรายได้มั่นคง; ไดขนาด 50–150 mm² ในแพ็กเกจ flip-chip มาตรฐาน	N3/N3E หรือ Samsung SF3	SMIC N+1/N+2 ใช้ได้ (ไม่ต้องใช้ HBM หรือ CoWoS) โดยแพงกว่าราว 40–50% ต่อไดดี	พลังงานเป็นต้นทุนหลัก; กติกาของโปรโตคอลเป็นความเสี่ยงใหญ่ที่สุด
ชิปขุด SHA-256	อย่าไล่ตามโหนดแนวหน้า; รูปแบบเฉพาะกลุ่มใช้โหนดรุ่นเก่าได้	ระดับ 3 nm; ผู้เล่นเดิมย้ายจาก TSMC 5 nm ไป 2 nm และจาก Samsung SF3 ไป 2 nm (แหล่งข้อมูลทุติยภูมิ)	ยากจะเทียบประสิทธิภาพของโหนดล้ำสมัย	ราคาไฟฟ้าและวัฏจักรการลดรางวัลครึ่งหนึ่งเป็นตัวกำหนดอัตรากำไร
ชิปเทรน AI	ไม่ทำชิปแยกเดี่ยว: IP ชิปเล็ต บริการออกแบบ	N3P แล้วตามด้วย N2/A16; 2–8 ชิปเล็ต; จอง CoWoS และ HBM แต่เนิ่น	SMIC N+2/N+3 กับ HBM ในประเทศ (แบบ Huawei); FP8 ต่อชิปราว 0.2–0.5 เท่าของ NVIDIA ชดเชยด้วยขนาดของระบบ	การแพ็กเกจและ HBM เป็นคอขวด
ชิปอนุมาน AI	เน้น SRAM หรือฝังแบบตายตัว N6/N7 หรือ N4/N5/SF4; เฉพาะ INT8, FP8 และ FP4	N3/N2 กับ HBM (บริบทยาว mixture of experts) หรือ SRAM แบบหลายชิปเล็ต	SMIC N+2 กับ LPDDR หรือ GDDR; หน่วยความจำระดับ HBM คือจุดคอขวด	ระดับการผูกติดของสถาปัตยกรรมคือการตัดสินใจหลัก
ตัวเร่ง ZK	เริ่มบน FPGA หรือ MPW 16/12 nm; แล้วจึง N7/N6 และ N5	N5/N3	SMIC N+1/N+2 ใช้ได้	งานเลขจำนวนเต็มที่กำหนดผลได้แน่นอน; ตรวจสอบง่าย
ตัวเร่ง FHE	FPGA (เทียบกับค่าฐานโอเพนซอร์ส) หรือ MPW 16/12 nm; การผลิตจำนวนมากต้องใช้ HBM	N5/N3 กับ HBM	ถูกจำกัดด้วยการเข้าถึง HBM	แบนด์วิดท์หน่วยความจำ ไม่ใช่เลขคณิต เป็นตัวกำหนดความเร็ว

ขั้นสูงหรือรุ่นเก่า โรงหล่อสี่แบบ และชิปเล็ต

ทางเลือก	จุดแข็งหลัก	จุดอ่อนหลัก	เหมาะที่สุดสำหรับ
โหนดล้ำสมัย (N3/N2, SF3/SF2, Intel 18A)	แต่ละโหนดเต็มลดพลังงานต่อออปลงราวครึ่งหนึ่ง; SRAM หนาแน่น	เวเฟอร์ 3–10 เท่า หน้ากาก 5–15 เท่า และงบออกแบบ 10–20 เท่าของโหนดรุ่นเก่า; ระยะเวลารอเกิน 50 สัปดาห์; ต้นทุนต่อทรานซิสเตอร์ของ N2 อาจเพิ่มขึ้น	ชิปประมวลผลที่อัลกอริทึมมั่นคงและทำงานตลอดเวลา; ชิปเทรน

ทางเลือก	จุดแข็งหลัก	จุดอ่อนหลัก	เหมาะที่สุดสำหรับ
โหนดรุ่นเก่าและระดับกลาง (N7/N6, N16/12, N28)	หน้ากากในระดับล้านดอลลาร์ MPW ราคาถูก ทำซ้ำได้เร็ว มีโรงหล่อหลายราย (TSMC, UMC, GlobalFoundries, SMIC)	ประสิทธิภาพสู้ไม่ได้เมื่อทำงานเต็มกำลังตลอดเวลา	ต้นแบบ ตลาดเฉพาะกลุ่ม การอนุมานที่ขอบเครือข่าย
TSMC	yield และระบบนิเวศดีที่สุด	ราคาสูงที่สุด; กำลังการผลิต N2/N3 ตึงตัว; คิว CoWoS ยาว	ทีมที่ไล่ตามพรมแดนประสิทธิภาพและมีคำมั่นด้านปริมาณการสั่งซื้อ
Samsung Foundry	เวเฟอร์ถูกกว่าราว 20–25% (SF3 ราว 15,000 ดอลลาร์ เทียบกับราว 20,000 ดอลลาร์สำหรับ N3); ใช้โดย Groq รุ่นที่สองและรายอื่นๆ	ธุรกิจโรงหล่อขาดทุน ตั้งเป้าคุ้มทุนภายในสิ้นปี 2027	ชิปประมวลผลที่อ่อนไหวต่อต้นทุน
Intel Foundry	ชิปซับซ้อนอย่าง Heracles พิสูจน์แล้วบน Intel 3	ไม่มีข้อมูลราคาหรือลูกค้าภายนอกสำหรับ 18A ที่เปิดเผย; ธุรกิจโรงหล่อขาดทุน	ประเมินระบบนิเวศและการสนับสนุนเป็นรายกรณี
โรงหล่อในจีนแผ่นดินใหญ่ (SMIC และรายอื่น)	อยู่ใกล้ตลาดในประเทศขนาดใหญ่; เวเฟอร์ 28 nm ที่ 1,500–2,500 ดอลลาร์	กำลังการผลิตระดับ 7 nm ราว 45,000–60,000 เวเฟอร์ต่อเดือน ส่วนใหญ่ถูก Huawei ใช้ไป; yield ราวหนึ่งในสามของ TSMC; ต้นทุนสูงกว่า 40–50%; DUV เท่านั้น	ชิปอนุมาน ชิปขุด และชิป ZK สำหรับตลาดในประเทศ
ชิปเล็ต	ข้ามขีดจำกัดของ reticle และ yield; เป็นมาตรฐานสำหรับระบบ HBM (MI355X, Ironwood, Trainium3)	ต้นทุนแพ็กเกจสูงกว่าและพึ่งพา CoWoS; การทดสอบ known-good-die เพิ่มต้นทุน	ชิปเทรนและชิปอนุมานขนาดใหญ่; ไม่จำเป็นสำหรับชิปขุดหรือชิป ZK ขนาดเล็ก

บทที่ 10

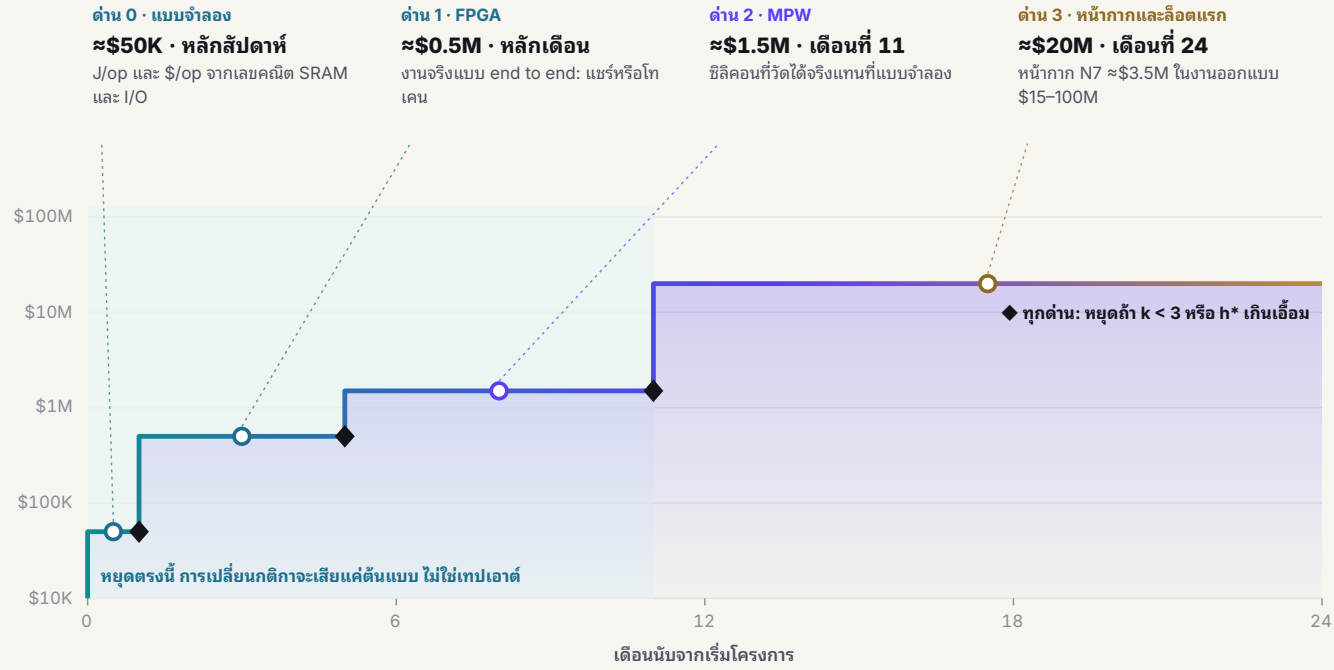
10 งบแบบเป็นขั้น
จากแบบจำลองสู่ซิลิคอน: งบแบบเป็นขั้นสำหรับชิปตัวแรก

ซื้อหลักฐานก่อนซื้อหน้ากาก

ชิปแบบ full-custom คือชุดของการเดิมพันที่แต่ละครั้งใหญ่กว่าครั้งก่อน และวินัยสำคัญคือการซื้อหลักฐานก่อนซื้อหน้ากาก ผมใช้สี่ด่าน บทวิเคราะห์ ด่าน 0 คือแบบจำลอง: การศึกษาสถาปัตยกรรมที่ประมาณ J/op และ $/op จากเลขคณิต SRAM และ I/O เทียบกับฮาร์ดแวร์อ้างอิงในบทที่ 7 ใช้เวลาหลักสัปดาห์และเงินหลักหมื่นดอลลาร์ ด่าน 1 คือต้นแบบ FPGA ที่รันงานจริงแบบ end to end ไม่ว่าจะเป็นแชร์การขุดหรือโทเคนการอนุมาน และพิสูจน์ดาต้าพาธกับการผสานเข้ากับโปรโตคอล ใช้เวลาไม่กี่เดือนและเงินหลักแสนดอลลาร์ ด่าน 2 คือเวเฟอร์หลายโครงการ: สล็อตเชิงพาณิชย์มีราคาราว 50,000–150,000 ดอลลาร์ที่ N28 และ 100,000–500,000 ดอลลาร์ที่ N7 และซิลิคอนที่วัดได้จริงเข้าแทนที่ค่าประมาณพลังงานของแบบจำลอง ด่าน 3 คือชุดหน้ากากเต็มชุดและล็อตการผลิตแรก: หน้ากากราว 1 ล้านดอลลาร์ที่ N28, 3.5 ล้านดอลลาร์ที่ N7 และ 15 ล้านดอลลาร์ที่ N3 ภายในงบออกแบบ 15–100 ล้านดอลลาร์ที่ 7/6 nm แต่ละด่านต้องมีตัวเลขที่สั่งหยุดโครงการ เช่น k ต่ำกว่า 3 เมื่อเทียบกับฮาร์ดแวร์ส่วนเพิ่ม หรือแฮชเรตคืนทุน h* ที่ช่วงเวลาใช้งานไปไม่ถึง

ภาพที่ 09 **ซื้อหลักฐานก่อนซื้อหน้ากาก: งบแบบเป็นขั้นสำหรับชิปตัวแรก**

ยอดใช้จ่ายสะสมเป็นดอลลาร์สหรัฐของโครงการ N7/N6 ผ่านสี่ด่าน สเกลลอการิทึม แต่ละด่านมีตัวเลขที่สั่งหยุดโครงการ



แหล่งที่มา: ราคาเวเฟอร์ MPW และหน้ากากตามบทที่ 9 (Silicon Analysts, AnySilicon, ค่าจากแบบจำลอง); เวลาและยอดใช้จ่ายของแต่ละด่านเป็นสมมติฐานเพื่อการอธิบายของผู้เขียน

บันไดนี้ชัน ในโครงการ N7/N6 สามด่านแรกรวมกันมีต้นทุนราว 1.5 ล้านดอลลาร์และใช้เวลาราวสิบเอ็ดเดือน ส่วนด่านที่สี่แพงกว่านั้นกว่าสิบเท่า ประมาณการ นั่นคือเหตุที่ลำดับสำคัญกว่าโหนด: หากโปรโตคอลเปลี่ยนเลขคณิต หรือผู้ซื้อถอนตัว ทีมที่หยุดไว้ที่ด่าน 2 จะเสียแค่ราคาของต้นแบบ ไม่ใช่ราคาของเทปเอาต์

ชิปแต่ละแบบมีต้นทุนเท่าไร และอะไรต้องเป็นจริง

โครงการ	ด่าน 2 (ต้นแบบ)	ด่าน 3 (โหนดและงบออกแบบ)	ไดดีขนาด 100 mm² หนึ่งชิ้น	อะไรต้องเป็นจริงจึงจะคืนทุน
ตัวเร่งที่เน้นการอนุมานและขุดได้	MPW 16/12 nm หรือ N7, 0.1–0.5 ล้านดอลลาร์	N6/N7 ราว 15–30 ล้านดอลลาร์ ประมาณการ	ราว 17 ดอลลาร์ที่ N7 ก่อนรวมแพ็กเกจและหน่วยความจำ	ลูกค้าการอนุมานที่เซ็นสัญญาแล้วแบกรับ NRE ส่วนใหญ่; รายได้จากการขุดเป็นส่วนเพิ่ม ไม่ใช่กรณีฐาน
เครื่องขุด INT7 เฉพาะทางสำหรับ Pearl (ไม่แนะนำในตอนนี้)	MPW 28 nm, 0.05–0.15 ล้านดอลลาร์	N7/N6 ราว 20 ล้านดอลลาร์	ราว 17 ดอลลาร์ที่ N7	ราว 4.9 EH/s ตลอดหนึ่งปีที่ p = 1 หรือ 9.8 EH/s ที่ p = 0.5 (บทที่ 7)
ตัวเร่ง ZK	FPGA แล้วจึง MPW 16/12 nm	N7/N6 แล้วจึง N5	ราว 10 ดอลลาร์ที่ N16, 17 ดอลลาร์ที่ N7	ระบบการพิสูจน์ลงตัวที่เคอร์เนล MSM และ NTT ซึ่งการออกแบบที่ตายตัวรองรับได้
ตัวเร่ง FHE	FPGA เทียบกับค่าฐานโอเพนซอร์ส	N5/N3 กับ HBM, 30 ล้านดอลลาร์หรือมากกว่า	ถูกครอบงำด้วย HBM และการแพ็กเกจ	มีผู้ซื้อ AI แบบเข้ารหัสในปริมาณมาก ซึ่งยังไม่มีอยู่

ประเด็นเชิงพาณิชย์สองข้อตามมา ข้อแรก ประสิทธิภาพที่ถูกที่สุดมาจากสถาปัตยกรรม: การเก็บตัวถูกดำเนินการไว้ใน SRAM และตัด HBM กับ CoWoS ออกขจัดต้นทุนส่วนใหญ่ของ GPU (ราว 63% ของรายการวัสดุของ H100) ไปก่อนจะเลือกโหนดใดๆ ข้อ

สอง เวลาคือต้นทุน: ทีมใหม่ต้องใช้ราว 18 ถึง 24 เดือนตั้งแต่เริ่มจนถึงการผลิตจำนวนมาก และระยะเวลารอที่โหนดล้ำสมัยเกิน 50 สัปดาห์ โปรโตคอลที่เปลี่ยนเลขคณิตทุกไม่กี่เดือนจึงอาจทำให้ชิปล้าสมัยก่อนที่ด่าน 3 จะคืนทุน

กฎสามข้อสำหรับทีมเล็ก

เมื่อมองรวมกัน ตัวเลขเหล่านี้ให้กฎสามข้อ บทวิเคราะห์ ข้อแรก **อยู่ที่ด่าน 1 หรือ 2 จนกว่าจะมีผู้ซื้อที่ระบุชื่อได้หรือเวอร์ชันโปรโตคอลที่ตรึงแล้ว** ทั้งราคาโทเคนและแฮชเรตไม่อาจใช้แทนสิ่งใดสิ่งหนึ่งในสองข้อนี้ได้ ข้อสอง **ซื้อ k ด้วยสถาปัตยกรรมก่อน แล้วจึงด้วยโหนด**: อาร์เรย์ที่อยู่ใน SRAM บน N7 อาจชนะชิปอเนกประสงค์บน N3 สำหรับงานเดียว ด้วยต้นทุนหน้ากากเพียงเศษเสี้ยว ข้อสาม **เขียนแฮชเรตหรือปริมาณที่ต้องใช้เพื่อคืนทุน พร้อมค่า p ไว้ก่อนเทปเอาต์** และตั้งราคาการจ่ายตามหมุดหมายโดยอิงเวอร์ชันโปรโตคอลที่ล็อกไว้ เพื่อให้ลูกค้า ไม่ใช่บริษัทออกแบบ เป็นผู้แบกรับความเสี่ยงด้านกติกา

บทที่ 11

11 โอกาส
โอกาสอยู่ที่ไหน และผมจะทำอะไร

ชิปสั่งทำ: อุปสงค์มาจาก AI อย่างท่วมท้น

ในปี 2026 อุปสงค์ชิปสั่งทำมาจาก AI อย่างท่วมท้น รายได้เซมิคอนดักเตอร์ AI ของ Broadcom ในไตรมาสที่ 3 ของปีงบประมาณ 2026 อยู่ที่ 16.7 พันล้านดอลลาร์ เพิ่มขึ้น 221% จากปีก่อน ราว 58 พันล้านดอลลาร์สำหรับทั้งปี โดย 73% มาจาก XPU สำหรับลูกค้าหกราย (จากบันทึกการประชุมผลประกอบการที่ถอดความอัตโนมัติ) รายได้ไตรมาสที่ 2 ของ Global Unichip เพิ่มขึ้น 128% จากปีก่อน โดยรายได้การผลิตแบบ turnkey เพิ่มขึ้น 188% และอัตรากำไรขั้นต้น 21.5% รายได้ไตรมาสที่ 1 ของ Alchip แยกตามโหนดเป็น 21% ที่ 3/2 nm, 63% ที่ 7/5 nm และ 16% ที่ 12 nm ขึ้นไป คำสั่งซื้อใหม่ปี 2026 ของ VeriSilicon ถึง 8.24 พันล้านหยวน (ราว 1.16 พันล้านดอลลาร์) ณ สิ้นเดือนเมษายน โดย 91.37% เป็นงานประมวลผล AI ตลาดมีสามชั้น: ชั้นบนสุดคือการพัฒนาร่วมหลายปีมูลค่าหลายพันล้าน ชั้นกลางคือโครงการตั้งแต่ 7 nm ถึง 2 nm ที่มี NRE ตั้งแต่ไม่กี่ล้านถึงกว่าร้อยล้านดอลลาร์ (5–30 ล้านดอลลาร์ที่โหนดระดับกลาง 30–100 ล้านดอลลาร์หรือมากกว่าที่โหนดล้ำสมัย) และหางยาวของ ASIC บนโหนดรุ่นเก่าที่มี NRE 0.5–5 ล้านดอลลาร์ ลูกค้าด้านคริปโตและการขุดเป็นกลุ่มเฉพาะที่เล็กและเป็นวัฏจักร แต่เป็น “ลูกค้ารายแรก” ที่เร็วที่สุดและไม่ต้องขออนุญาตใครที่สุดซึ่งจ่ายเงินซื้อพลังประมวลผลในวันนี้

ความสามารถของทีมชิปขุดที่ใช้ต่อได้ดีที่สุด ได้แก่ การทำงานแรงดันต่ำใกล้เกณฑ์ อาร์เรย์ประมวลผลที่ทำซ้ำจำนวนมหาศาล การจ่ายไฟและการระบายความร้อนระดับกิโลวัตต์ และเทปเอาต์บนโหนดล้ำสมัยที่รวดเร็ว คือสิ่งที่ชิปอนุมานและตัวเร่ง ZK และ FHE ต้องการพอดี ตลาดทุนตีราคาเรื่องนี้ไว้แล้ว: Velaura AI (เดิมชื่อ Auradine) ระดมทุนได้ 110 ล้านดอลลาร์ในเดือนสิงหาคมที่มูลค่าเกิน 1 พันล้านดอลลาร์ ตัวอย่างตรงข้ามคือหน่วยชิป AI ที่ขอบเครือข่ายของ Canaan ซึ่งมีรายได้ 0.9 ล้านดอลลาร์ในปีงบประมาณ 2024 ขณะที่ใช้ค่าใช้จ่ายในการดำเนินงาน 15% และถูกปิดในเดือนมิถุนายน 2025 ความแตกต่างอยู่ที่ช่องทางสู่ตลาดที่เป็นอิสระ ไม่ใช่การใช้ช่องทางการขุดซ้ำ

การพัฒนา AI: ชั่วโมงทำงานเป็นสินค้าโภคภัณฑ์ มูลค่าอยู่ที่รอยต่อ

ในระดับล่าง บริการพัฒนา AI กลายเป็นสินค้าโภคภัณฑ์: อัตราทั่วไปบน Clutch อยู่ที่ 25–49 ดอลลาร์ต่อชั่วโมง และ 50–99 ดอลลาร์ในสหรัฐอเมริกา แคนาดา และตลาดที่คล้ายกัน ทีมพัฒนาบล็อกเชนมีต้นทุนราว 15,000 ถึงกว่า 90,000 ดอลลาร์ต่อเดือน การขายชั่วโมงทำงานไม่มีคูเมือง มูลค่าอยู่ที่รอยต่อระหว่างสาขา: เคอร์เนลและคอมไพเลอร์สำหรับชิปสั่งทำ การขุดที่ฝังอยู่ใน

บริการอนุมาน (ปลั๊กอิน vLLM ของ Pearl เป็นตัวอย่างหนึ่ง) การอนุมานที่ตรวจสอบได้ซึ่งรวม TEE เข้ากับการสุ่มตรวจด้วย ZK และการชำระเงินกับตัวตนสำหรับเอเจนต์

ระบบ AI × คริปโต: เช็นสัญญาผู้ซื้อพลังประมวลผลก่อนออกโทเคน

สำหรับใครก็ตามที่ออกแบบเชนหรือเครือข่ายการประมวลผลที่เกิดมาเพื่อ AI ชุดเครื่องมือตรวจสอบของปี 2026 ก่อรูปแล้ว เรียงตามต้นทุนต่อหน่วยงานจากต่ำไปสูง: การสุ่มตัวอย่างเชิงสถิติ (proof of sampling ของ Hyperbolic) การมอบหมายแบบมีผู้ตัดสิน (Verde ของ Gensyn ซึ่งต้องการผู้ตรวจสอบที่ซื่อตรงเพียงรายเดียวแต่ต้องประมวลผลตรงกันทุกบิต) TEE (ความเร็วใกล้เคียงเนทีฟ แต่ต้องเชื่อใจผู้ขายชิป) การทำซ้ำเต็มรูปแบบ และ zkML ทั่วไป การออกแบบที่ใช้ได้จริงเป็นแบบชั้น: TEE หรือการสุ่มตัวอย่างเพื่อความหน่วงต่ำ เกมโต้แย้งหรือการสุ่มตรวจด้วย ZK เพื่อความเป็นที่สุด งานวิจัยเดือนธันวาคม 2025 จำลองการประมวลผลแบบเป็นความลับบน H100 ร่วมกับ fraud proof และการสุ่มตรวจด้วย ZK ได้ปริมาณงาน 99% ของแบบรวมศูนย์ และต้นทุนส่วนเพิ่ม 0.07 ดอลลาร์ต่อคิวรี ปรัชญาด้านฮาร์ดแวร์เป็นตัวกำหนดความปลอดภัยและการกระจาย: PoUW แบบ GPU-native ในช่วงแรกได้รับการคุ้มครองจาก GPU บนคลาวด์ที่เช่าได้ ต้นทุนการโจมตีจึงผูกกับราคาเช่า และ Qubic แสดงให้เห็นว่าการพิสูจน์งานบนฮาร์ดแวร์อเนกประสงค์ถูกยึดทั้งหมดได้ด้วยสิ่งจูงใจจากภายนอก

ต้นทุนและเวลาไม่ใช่อุปสรรค Pearl ไปจากงานวิจัย (เมษายน 2025) สู่เมนเน็ตในราว 12 เดือน โดยฟอร์กจาก btcd และมีฟูลโหนด วอลเล็ต ไลต์ไคลเอนต์ และโปรแกรมขุดพร้อมตั้งแต่เปิดตัว การตรวจสอบ DeFi ส่วนใหญ่มีค่าใช้จ่าย 25,000–100,000 ดอลลาร์ การตรวจสอบระดับองค์กรหรือหลายเชน 150,000–250,000 ดอลลาร์หรือมากกว่า วงจร ZK แพงกว่า 80% ถึง 120% และโครงการที่มีมูลค่าจริงตกอยู่ในความเสี่ยงตั้งงบด้านความปลอดภัย 150,000–500,000 ดอลลาร์ต่อปี อุปสรรคที่แท้จริงคือการกระจายและสภาพคล่อง ผู้ก่อตั้งรายงานว่า Binance ขอ 0% ถึง 3.5% และกล่าวหาว่าถึง 8% ของอุปทานโทเคนเพื่อการลิสต์ (Binance ปฏิเสธว่าไม่ได้หากำไรจากการลิสต์) ทั้งหกโครงการที่ Cointelegraph ทบทวนซึ่งใช้ข้อตกลงดูแลสภาพคล่องแบบ “ยืมโทเคนบวกออปชันซื้อ” ล้วนราคาร่วง และผู้ดูแลสภาพคล่องของ Movement ขายโทเคนราว 38 ล้านดอลลาร์ในวันลิสต์ ทางเลือกคือสภาพคล่องบนเชน อย่าง Hyperliquid ซึ่งมีมูลค่าตลาดเกิน 10 พันล้านดอลลาร์โดยไม่ได้ลิสต์บน Binance และตลาด perpetual ตาม HIP-3 มี open interest เกิน 3.2 พันล้านดอลลาร์ในเดือนมิถุนายน และการขายต่อสาธารณะผ่านกระดานเทรดรายใหญ่ อย่างที่ Monad ระดมทุนได้ 269 ล้านดอลลาร์ผ่าน Coinbase มูลค่าแบบ fully diluted ที่สูงพร้อมสัดส่วนหมุนเวียนต่ำเป็นตัวทำนายที่ชัดที่สุดของราคาร่วงหลังเปิดตัวในปี 2025–2026

การออกแบบโทเคนเป็นส่วนหนึ่งของการออกแบบระบบ การปล่อยเหรียญควรตามหลังอุปสงค์ที่จ่ายเงิน ไม่ใช่นำหน้า: เครือข่ายในรายงานนี้ที่มีตัวคูณเงินอุดหนุนดีที่สุดผูกรางวัลไว้กับงานที่ผู้ซื้อภายนอกจ่ายเงิน ส่วนเครือข่ายที่แย่ที่สุดจ่ายให้แฮชเรตหรือสเตกโดยไม่มีอะไรรองรับ กฎที่ใช้ได้จริงคือประกาศตัวคูณเงินอุดหนุนที่เครือข่ายตั้งเป้าไว้สำหรับปีที่หนึ่ง สอง และสามก่อนเปิดตัว และลดการปล่อยเหรียญเมื่อพลาดเป้า บทวิเคราะห์

ทิศทาง	ใครซื้อ และความเต็มใจจ่าย	หน้าต่างโอกาส	ความสามารถหลัก	ความเสี่ยงหลัก
ตัวเร่งที่เน้นการอนุมานและขุดได้	นีโอคลาวด์ การติดตั้งแบบส่วนตัว เครือข่าย PoUW	2027–2028 (ขึ้นกับฟอร์ก FP8 และ FP4)	GEMM แบบ INT8, FP8 และ FP4; สถาปัตยกรรม SRAM; การผสานกับโปรโตคอล	การรับเข้าของโปรโตคอล การผูกติดของสถาปัตยกรรม
ชิปอนุมานแบบ SRAM หรือฝังแบบตายตัว	นีโอคลาวด์ ศูนย์ข้อมูลในจีนแผ่นดินใหญ่ ขอบเครือข่าย	ตอนนี้	คอมไพเลอร์และการปรับโมเดล; การออกแบบ N6–N4	โมเดลเปลี่ยนเร็วเกินไป
ชิปขุด (เฉพาะกลุ่มหรือบูรณาการแนวตั้ง)	นักขุดที่มีไฟฟ้าเอง; เครื่องขุดสำหรับบ้านและเครื่องขุดที่ให้ความร้อน	ตอนนี้ ภายใต้ขอบเขตของการลดรางวัลครึ่งหนึ่งและ hashprice	อาร์เรย์แรงดันต่ำ ประสิทธิภาพ ห่วงโซ่อุปทานการผลิตจำนวนมาก	สงครามราคาจากผู้เล่นเดิม ราคาไฟฟ้า

ทิศทาง	ใครซื้อ และความเต็มใจจ่าย	หน้าต่างโอกาส	ความสามารถหลัก	ความเสี่ยงหลัก
การเร่งการพิสูจน์ ZK	เครือข่ายผู้พิสูจน์ rollup	รอจนระบบการพิสูจน์ลู่เข้าหากัน	MSM และ NTT ที่โปรแกรมได้	อุปสงค์ขึ้นกับรางวัลโทเคน
การเร่ง FHE	AI แบบเข้ารหัสสำหรับองค์กรและภาครัฐ; เชนด้านความเป็นส่วนตัว	2027–2029	NTT ระบบ HBM วิศวกรรมการเข้ารหัส	อุปสงค์ยังไม่ก่อตัว; วงจรการขายยาว
การเสริมความแข็งแกร่ง TEE และการอนุมานที่ตรวจสอบได้	เครือข่ายการอนุมานแบบเป็นความลับ DePIN	ตอนนี้	การเข้ารหัสหน่วยความจำพร้อมความถูกต้อง รากของ attestation ที่เป็นอิสระ การสุ่มตรวจด้วย ZK	การพึ่งพาระบบนิเวศของผู้ขาย
เชนหรือเครือข่ายการประมวลผลที่เกิดมาเพื่อ AI	ผู้ซื้อการอนุมาน นักขุด นักพัฒนา	ตอนนี้	การออกแบบการตรวจสอบ การออกแบบการปล่อยเหรียญ การกระจาย	ราคาร่วงหลังเปิดตัว สภาพคล่องของโทเคน

สามเส้นทางที่ผมจะเลือก แต่ละทางมีก้าวแรก

หากต้องเลือก ผมจะเลือกไม่เกินสามเส้นทาง แต่ละทางมีก้าวแรกเล็กๆ ที่ทดสอบก่อนลงเงิน บทวิเคราะห์

- **เส้นทาง A: ตัวเร่งที่เน้นการอนุมานและขุดได้ด้วย** ก้าวแรก: ศึกษาสถาปัตยกรรมสี่สัปดาห์ เพื่อสร้างแบบจำลอง J/op และ $/op ของอาร์เรย์ GEMM INT8/FP8 ที่อยู่ใน SRAM เทียบกับ RTX 5090 และ H100 ทั้งบนงานของ Pearl และการอนุมาน transformer ตามด้วยการตรวจสอบบน FPGA หรือ MPW ของโหนดรุ่นเก่าที่มี NRE ในระดับหนึ่งล้านดอลลาร์
- **เส้นทาง B: บริการซิลิคอนสั่งทำที่มีคริปโตเป็นลูกค้ารายแรกและ AI เป็นตลาดที่สอง** ก้าวแรก: เปลี่ยนแบบจำลอง NRE และจุดคุ้มทุนของรายงานนี้เป็นการศึกษาความเป็นไปได้แบบเก็บเงินสำหรับโครงการชิปประมวลผล โดยจ่ายตามหมุดหมายและตรวจรับเทียบกับเวอร์ชันโปรโตคอลที่ล็อกไว้
- **เส้นทาง C: เครือข่ายที่เกิดมาเพื่อ AI ซึ่งเซ็นสัญญากับผู้ซื้อพลังประมวลผลก่อนออกโทเคน** ก้าวแรก: บันทึกการออกแบบว่าด้วยการตรวจสอบเป็นชั้น (TEE บวกการสุ่มตัวอย่างบวกการสุ่มตรวจด้วย ZK) และเลขคณิตมาตรฐานที่เป็นกลางต่อผู้ขาย เขียนร่วมกับผู้ซื้อการอนุมานที่ระบุชื่อได้หนึ่งรายก่อนออกแบบโทเคนใดๆ

บทที่ 12

12 สัญญาณที่ต้องจับตา
สิ่งที่ต้องจับตา พฤศจิกายน 2026 ถึง 2027

เมื่อใด	สิ่งที่ต้องจับตา	ทำไมจึงสำคัญ
เมื่อใดก็ได้	ความสูงบล็อกที่เปิดใช้ FP8 ของ Pearl	ตัดสินว่าฮาร์ดแวร์เลขจำนวนเต็มและการ์ดที่ไม่ใช่ NVIDIA จะยังได้รับอนุญาตหรือไม่ และเชนจะแยกหรือไม่
รายเดือน	สัดส่วนบล็อกของ Pearl ที่มีงาน AI	การทดสอบที่ดีที่สุดเพียงข้อเดียวว่า “มีประโยชน์” กำลังกลายเป็นจริงหรือไม่
รายเดือน	ผู้ให้บริการอนุมานส่งต่อรายได้จากการขุดให้ลูกค้าหรือไม่	อุปสงค์ที่จ่ายเงินมาบรรจบกับการปล่อยเหรียญ: แบบจำลองที่ Pearl ต้องการ
รายไตรมาส	รายได้ที่ตรวจสอบได้ของ Bittensor เทียบกับการปล่อยเหรียญ	เครือข่ายงานที่มีประโยชน์ที่ใหญ่ที่สุดกำลังปิดช่องว่างเงินอุดหนุนหรือไม่
รายไตรมาส	สัดส่วนกระแส x402 ที่เอเจนต์เป็นผู้เริ่มจริง	การชำระเงินของเอเจนต์จะเปลี่ยนจากการส่งสัญญาณไปสู่การค้าหรือไม่
มกราคม 2027	TSMC ขึ้นราคา 3% ถึง 6% ตามที่มีรายงาน โดยขึ้นมากที่สุดสำหรับ 2 nm และ 3 nm	ต้นทุนของเทปเอาต์บนโหนดล้ำสมัย
ไตรมาสแรกของปี 2027	FP4 ใน Pearl ตามที่รายงานจาก AMA ของผู้ก่อตั้ง	การเลื่อนงานไปสู่รูปแบบตัวเลขใหม่ล่าสุดของ NVIDIA อีกขั้น
ปลายปี 2027	เป้าหมายคุ้มทุนของ Samsung Foundry	แหล่งผลิตรายที่สองที่ถูกกว่าจะยังแข็งแรงอยู่หรือไม่
เมษายน 2028	การลดรางวัลครึ่งหนึ่งครั้งถัดไปของบิตคอยน์	กองรายได้ของเครื่องขุด SHA-256 ลดลงครึ่งหนึ่ง

บทที่ 13

13 คำถามที่ควรถาม
คำถามที่คุณไม่ได้ถาม แต่ควรถาม

คำถามด้านล่างคือคำถามที่ผมจะถามก่อนลงเงินหรือซิลิคอนในโครงการ AI × คริปโตใดๆ

ใครคุมฟังก์ชันงาน? หากทีมเล็กๆ เปลี่ยนเลขคณิตได้ที่ความสูงบล็อกที่เขียนตายตัว อายุของฮาร์ดแวร์คุณก็ขึ้นกับการตัดสินใจของพวกเขา ขอคำมั่นเรื่องความมั่นคงเป็นลายลักษณ์อักษร หรือออกแบบฮาร์ดแวร์ที่มีตลาดที่สอง

ตัวคูณเงินอุดหนุนเป็นเท่าไร? หารการปล่อยเหรียญ (หรือมูลค่าตลาด) ด้วยรายได้ที่จ่ายจากนอกเครือข่าย ตัวคูณที่เกิน 10 หมายความว่าเครือข่ายกำลังซื้อกิจกรรม ไม่ได้ขายบริการ

ด่านถัดไปมีต้นทุนเท่าไร และตัวเลขใดจะทำให้คุณหยุด? เวเฟอร์หลายโครงการมีราคาหลักหมื่นหรือหลักแสนดอลลาร์ ชุดหน้ากากมีราคาหลักล้าน กำหนดตัวเลขที่จะหยุดไว้ก่อนใช้เงิน

แฮชเรตคืนทุนเป็นเท่าไร เมื่อรวม p แล้ว? หาร NRE ด้วย R × (1 − 1/k) × T × p หากคำตอบเป็นสัดส่วนใหญ่ของเครือข่าย ชิปนั้นก็คือการเดิมพันกับกติกา ไม่ใช่กับซิลิคอน

ตลาดที่สองของชิปคืออะไร? หากโปรโตคอลฟอร์กหรือโทเคนร่วง ซิลิคอนนี้จะขายเข้าตลาดใดได้อีก GPU รอดจาก Merge ของ Ethereum เพราะมีตลาดที่สอง ส่วน ASIC อัลกอริทึมเดียวไม่รอด

มูลค่าไปตกที่โทเคนหรือที่หุ้นของบริษัท? ในปี 2026 บริษัท AI × คริปโตที่ทำกำไรได้ส่วนใหญ่เป็นบริษัทเอกชนที่ใช้ช่องทางกระจายแบบคริปโต ตัดสินใจให้ได้ว่าคุณกำลังสร้างแบบใดก่อนระดมทุน

ลูกค้าที่จ่ายเงินรายแรกคือใคร ระบุชื่อได้หรือไม่? กรณีที่ยั่งยืนทุกกรณีในรายงานนี้ ตั้งแต่ Grass ไปจนถึงซับเน็ตด้านการประมวลผลของ Bittensor เริ่มต้นจากผู้ซื้อนอกเครือข่ายที่ระบุชื่อได้

นี่คือด้านที่ผมอยากสำรวจร่วมกับผู้อื่น ผ่านการแลกเปลี่ยนและความร่วมมือทางเทคนิค: ชิปสั่งทำสำหรับการขุด การเทรนและการอนุมาน AI และตัวเร่ง การพัฒนา AI และระบบ AI × คริปโตตั้งแต่เชน โหนด ไปจนถึงชิป ติดต่อผมได้ที่ i@tanshuai.com

ข้อจำกัดของรายงานนี้

นี่คือการวิจัยเอกสารจากเอกสารสาธารณะจนถึงวันที่ 13 พฤศจิกายน 2026 สำหรับฉบับนี้ ผมไม่ได้สัมภาษณ์แหล่งข้อมูลปฐมภูมิและไม่ได้วัดผลอย่างอิสระ ตัวเลขสำคัญหลายตัวเป็นข้ออ้างของบริษัทหรือผู้ขาย เช่น รายได้ของ Aethir รายได้ของ Grass สมรรถนะของ Etched และ Taalas และประสิทธิภาพของ Cysic ซึ่งติดป้ายไว้ตามนั้น ตัวชี้วัดบนเชนของ Pearl มาจากฟูลโหนดอิสระและจากพูล และเปลี่ยนแปลงรายชั่วโมง ผมใช้ค่าที่อ่านได้ราว 07:13 UTC ของวันที่ 13 พฤศจิกายน และให้ช่วงค่าในวันเดียวกัน ต้นทุนเวเฟอร์ หน้ากาก และการออกแบบเป็นประมาณการของบุคคลที่สามและค่าจากแบบจำลองที่ไม่ตรงกัน ผมจึงแสดง

เป็นช่วง การคำนวณจุดคุ้มทุนของ ASIC และงบแบบเป็นขั้นใช้สมมติฐานเพื่อการอธิบาย (k, T, p, NRE และต้นทุนของแต่ละด่าน) ซึ่งผู้อ่านควรแทนด้วยค่าของตนเอง

เกี่ยวกับผู้เขียน



เกี่ยวกับผู้เขียน

ทัน ชว่าย เป็นนักเทคโนโลยีและผู้ก่อตั้ง มีประสบการณ์กว่ายี่สิบปีในงานด้านชิป ฮาร์ดแวร์ และซอฟต์แวร์ระบบ และได้รับสิทธิบัตร 20 ฉบับ เขามีประสบการณ์หลายปีในการวิจัยพัฒนาและการบริหารชิปแบบ fabless รวมถึงการออกแบบดาต้าพาธ ASIC แบบ full-custom สำหรับชิปขุดสกุลเงินดิจิทัลและชิปกล้องมัลติมีเดีย ซึ่งเป็นเลนส์ของรายงานนี้: พลังงานต่อออป ต้นทุนต่อเวเฟอร์ สิ่งที่โปรโตคอลยอมให้ตรึงไว้ในซิลิคอน และใครเป็นผู้จ่าย งานของเขาพาไปทั้งซิลิคอนแวลลีย์ นิวยอร์ก ชิคาโก ไต้หวัน เขตเกรตเตอร์เบย์แอเรีย (เซินเจิ้นและฮ่องกง) เวียดนาม และมาเลเซีย ปัจจุบันเขาทำงานเป็นนักวิจัยและนักพัฒนา AI

พูดคุยเชิงเทคนิค: i@tanshuai.com · tanshuai.com

หมายเหตุและข้อจำกัดความรับผิดชอบ

รายงานนี้เป็นการวิจัยเอกสารอย่างอิสระจากเอกสารสาธารณะที่มีอยู่จนถึงวันที่ 13 พฤศจิกายน 2026 จัดทำเพื่อให้ข้อมูลและการเรียนรู้เท่านั้น และไม่ใช่คำแนะนำด้านการลงทุนหรือวิศวกรรม รายงานนี้พิจารณาเฉพาะด้านเทคโนโลยีและธุรกิจ ราคา ตัวชี้วัดบนเชน และกติกาของโปรโตคอลเปลี่ยนแปลงได้ โปรดตรวจสอบสิ่งที่คุณพึ่งพากับแหล่งข้อมูลปฐมภูมิ ตัวเลขที่ติดป้าย ผู้ขายอ้าง ประมาณการ หรือ บทวิเคราะห์ ควรอ่านตามนั้น ราคาสินทรัพย์คริปโตผันผวนสูง

© 2026 ทัน ชว่าย อนุญาตให้แบ่งปันรายงานนี้โดยไม่ดัดแปลงพร้อมระบุแหล่งที่มา



แหล่งอ้างอิง

เข้าถึงทั้งหมดระหว่างวันที่ 10 ถึง 13 พฤศจิกายน 2026 จัดกลุ่มตามบท

Pearl: โปรโตคอล โค้ด และข้อมูลบนเชน

- Komargodski and Weinstein, "Proofs of Useful Work from Arbitrary Matrix Multiplication", arXiv 2504.09971: https://arxiv.org/abs/2504.09971
- Pearl Research Labs, INT whitepaper: https://pearlresearch.ai/research/int-whitepaper
- Pearl source code: genesis block: https://github.com/pearl-research-labs/pearl/blob/main/node/chaincfg/genesis.go ; consensus parameters and fork heights: https://github.com/pearl-research-labs/pearl/blob/main/node/chaincfg/params.go ; hashrate accounting: https://github.com/pearl-research-labs/pearl/blob/main/zk-pow/src/api/sanity_checks.rs ; salted-seed fork upgrade guide: https://github.com/pearl-research-labs/pearl/blob/main/docs/salted-seed-fork-upgrade-guide.md ; fp8-scheme branch parameters: https://github.com/pearl-research-labs/pearl/blob/fp8-scheme/node/chaincfg/params.go
- Pearl Improvement Proposals: https://github.com/pearl-research-labs/pips ; PIP-3, floating-point mining: https://github.com/pearl-research-labs/pips/blob/main/PIPS/pip-0003.md
- prlstats (ฟูลโหนดอิสระ): network: https://prlstats.com/api/network ; series: https://prlstats.com/api/series ; home and miner income model: https://prlstats.com/api/home ; mined supply: https://prlstats.com/api/mined ; pools: https://prlstats.com/api/pools ; holders: https://prlstats.com/api/holders ; hardware and overhead assumptions: https://prlstats.com/api/assumptions
- Kryptex Pool: Pearl network: https://pool.kryptex.com/prl/about-coin ; hashrate: https://pool.kryptex.com/prl/hashrate ; RTX 5090: https://pool.kryptex.com/de/device/gpu/nvidia/nvidia-geforce-rtx-5090 ; RTX 4090: https://pool.kryptex.com/device/gpu/nvidia/nvidia-geforce-rtx-4090 ; "Pearl FP8 fork": https://pool.kryptex.com/articles/pearl-fp8-fork-en
- CoinGecko, Pearl: https://api.coingecko.com/api/v3/coins/pearl-2 ; https://www.coingecko.com/en/coins/pearl-research ; price history: https://api.coingecko.com/api/v3/coins/pearl-2/market_chart?vs_currency=usd&days=365&interval=daily ; "Perle" ticker search: https://api.coingecko.com/api/v3/search?query=perle
- SimeonOnSecurity, "Mining Pearl (PRL) in 2026": https://simeononsecurity.com/articles/pearl-prl-mining-guide-2026/
- Hashrate Index, "Pearl (PRL): Inside the AI-Compute Cryptocurrency", June 2, 2026: https://hashrateindex.com/blog/pearl-prl-ai-compute-cryptocurrency/
- Tom's Hardware, "New AI compute cryptocurrency Pearl sparks a GPU mining rush, but profitability is sliding": https://www.tomshardware.com/tech-industry/cryptomining/new-ai-compute-cryptocurrency-pearl-sparks-a-gpu-mining-rush-but-profitability-is-sliding
- Together AI, partnership with Pearl Research Labs: https://together.ai/blog/together-ai-partners-with-pearl-research-labs ; model page: https://together.ai/models/gemma-4-31b-it-pearl
- Basu, census of Pearl mining instances, arXiv 2606.04819: https://arxiv.org/html/2606.04819v2
- Blockworks, 0xResearch newsletter, October 28, 2026: https://blockworks.com/newsletter/0xresearch/issue/post_daf0d74a-2bf4-41e4-86d3-2a1275475a24
- Datawallet, "Pearl Explained": https://www.datawallet.com/crypto/pearl-explained
- Coingabbar on TheMarker's report of Pearl Research Labs' funding: https://www.coingabbar.com/en/pearl-research-labs-funding-130m-stealth-round-themarker

เชนงานที่มีประโยชน์และประวัติศาสตร์

- Grothe et al., Gridcoin security analysis, USENIX WOOT 2017: https://www.usenix.org/system/files/conference/woot17/woot17-paper-grothe.pdf
- Proof-of-learning spoofing, arXiv 2208.03567: https://arxiv.org/abs/2208.03567
- hoge.gg, verification and trust in decentralized inference: https://hoge.gg/decentralized-inference-2026-verification-trust/ ; Bittensor revenue and treasury: https://hoge.gg/bittensor-quantum-wallet-revenue-tao-treasury-2026/
- Kaspa: kHeavyHash matrix code: https://github.com/kaspanet/rusty-kaspa/blob/master/consensus/pow/src/matrix.rs ; launch: https://coinranking.com/coin/V8GxkwWow+kaspa-kas/about ; IceRiver KS1/KS2 announcement: https://www.iceriver.io/2023/03/14/1677644580/ ; GPU-to-ASIC history: https://captainmining.com/blogs/crypto-news/the-evolution-of-kaspa-mining-from-gpus-to-asic-powerhouses ; GPU hashrates: https://coincalculators.io/algorithms/kheavyhash ; KS7: https://www.asicminervalue.com/miners/bitmain/antminer-ks7-40th ; network: https://minerstat.com/coin/kas ; 1 EH/s in 2024: https://ourcryptotalk.com/news/kaspa-reaches-1-ehs-in-hashrate ; Crescendo: https://ourcryptotalk.com/news/kaspa-crescendo-live
- Aleo: mainnet FAQ: https://aleo.org/post/aleo-mainnet-faq/ ; decentralized proving: https://aleo.org/post/decentralized-proving-advantages/ ; ASICs: https://whattomine.com/coins/432-aleo-zksnark/asics ; Blockworks, "State of Aleo Q3 2025": https://app.blockworks.com/report/state-of-aleo-q3-2025 ; ARC-46: https://chaincatcher.com/en/article/2144326
- Qubic: The Block on the Monero reorganization: https://www.theblock.co/post/366535/monero-faces-chain-reorganization-fears-after-qubic-says-it-controls-51-of-hashrate ; Halborn: https://www.halborn.com/blog/post/explained-the-monero-51-percent-attack-august-2025 ; Dogecoin mining: https://hackernoon.com/qubic-is-now-mining-dogecoin-on-its-ai-compute-infrastructure ; CoinGecko: https://www.coingecko.com/en/coins/qubic

- Bittensor consensus documentation: https://www.bittensor.com/llms.mdx/docs/internals/consensus/content.md ; FalconX, "State of Bittensor": https://www.falconx.io/newsroom/state-of-bittensor-subnet-adoption-trends-network-mechanics-and-covenants-departure
- Primecoin: https://en.wikipedia.org/wiki/Primecoin ; Blockworks, "State of Filecoin Q3 2025": https://app.blockworks.com/report/state-of-filecoin-q3-2025
- Monero and ASIC-resistance forks: https://lbry.com/news/forking-for-asic-resistance ; Sia fork: https://hotfix.coindesk.com/markets/2018/11/01/sia-completes-hard-fork-as-bricked-miners-form-rival-blockchain ; Obelisk prices: https://www.cryptoglobe.com/latest/2018/11/obelisk-hardware-skyrockets-in-selling-price-after-sia-hard-fork/ ; miners after the Merge: https://hotfix.coindesk.com/business/2022/11/11/ex-ethereum-miners-token-hop-to-stay-alive-after-the-merge

กลุ่มต่างๆ ของ AI × คริปโต

- DefiLlama: Render: https://defillama.com/protocol/render-network-bme ; Livepeer: https://defillama.com/protocol/livepeer
- Cointelegraph on Messari's DePIN revenue: https://cointelegraph.com/news/depin-revenues-grow-despite-token-losses
- Aethir, 2025 wrap-up: https://aethir.com/blog-posts/aethirs-2025-wrap-up-decentralized-gpu-cloud-milestones
- CoinGecko, AI category: https://www.coingecko.com/en/categories/artificial-intelligence
- Akash, H100 rental prices: https://akash.network/the-bid/h100-rental-price-2026-cost-per-hour/ ; Q1 2026 report: https://akash.network/blog/akash-network-q1-2026-report/
- io.net, $8 million deal and daily tokens: https://lifestyle.kbew98country.com/story/271004/io-net-to-burn-up-to-12m-tokens-as-network-closes-8m-deal-and-hits-4-billion-daily-ai-tokens/
- DCD, Axe Compute contract: https://www.datacenterdynamics.com/en/news/axe-compute-secures-15bn-compute-contract-with-unnamed-customer/
- Pluralis-8B: https://www.alphaxiv.org/abs/2607.13332.md ; Nous Research, Psyche: https://nousresearch.com/the-next-phase-of-psyche ; Implicator on INTELLECT-3: https://www.implicator.ai/prime-intellects-intellect-3-open-source-ambition-meets-centralized-reality/
- SiliconANGLE, Prime Intellect funding: https://siliconangle.com/2026/07/08/prime-intellect-raises-130m-1b-valuation-ai-training-platform/ ; The Block, Venice: https://www.theblock.co/post/406934 ; RuntimeWire, Nous funding: https://runtimewire.com/article/nous-research-75-million-round-1-5-billion-valuation
- ownyourmind, Gensyn: https://ownyourmind.ai/projects/gensyn/ ; Phemex, Ritual (มีเพียงพาดหัวข่าว): https://phemex.com/news/article/ritual-launches-ai-layer-1-mainnet-and-ritual-token-price-dives-72-from-high-101259
- Confidential computing overhead, arXiv 2409.03992: https://arxiv.org/abs/2409.03992v2 ; Phala 2025 report: https://phala.com/posts/phala-2025-report ; TEE.fail: https://thehackernews.com/2025/10/new-teefail-side-channel-attack.html ; DeepProve, IACR ePrint 2026/1112: https://eprint.iacr.org/2026/1112
- TRM Labs, measuring AI agent payments: https://www.trmlabs.com/trm-tech-blog/whos-actually-paying-measuring-ai-agent-payments-onchain ; Forkast on the x402 Foundation: https://forkast.news/the-x402-foundation-activated-a-27-year-old-internet-code-200-million-transactions-later-the-real-volume-is-still-tiny/ ; The Block, Tempo mainnet: https://www.theblock.co/amp/post/394131/tempo-mainnet-goes-live-with-machine-payments-protocol-for-agents ; BlockEden on Virtuals' aGDP: https://blockeden.xyz/blog/2026/04/21/agdp-agent-gdp-virtuals-protocol-ai-blockchain-valuation-primitive-tvl-displacement/ ; The Block on ElizaOS: https://theblock.co/post/410774/eliza-labs-native-token-dead
- Dealroom, Grass revenue: https://dealroom.co/news/157794-grass-backed-by-multicoin-projects-75m-in-2026-data-revenue/ ; CryptoSlate, Fetch.ai and Ocean: https://cryptoslate.com/cryptos-flagship-ai-pact-fracture-fetch-sues-ocean-over-263m-fet-community-sales/ ; Fortune on World: https://fortune.com/2026/06/01/internet-according-to-sam-altman-seeing-orb/
- Galaxy, crypto venture capital: https://www.galaxy.com/insights/research/crypto-blockchain-venture-capital-q3 ; Crypto Briefing (มีเพียงพาดหัวข่าว): https://cryptobriefing.com/crypto-vc-investment-q3-2026-galaxy/ ; Blockhead on Tiger Research: https://www.blockhead.co/2026/10/14/bitcoin-rose-43-in-q3-crypto-venture-funding-did-not-follow/ ; ForkLog on 2025 token launches: https://forklog.com/en/85-of-tokens-launched-in-2025-fell-below-initial-prices/

เศรษฐศาสตร์ ASIC และประเภทชิป

- Hashrate Index, roundup of November 9, 2026: https://hashrateindex.com/blog/hashrate-index-roundup-november-9-2026/
- AnySilicon, ASIC design cost by node in 2026: https://anysilicon.com/news/how-much-does-asic-design-cost-in-2026-real-price-ranges-by-process-node/ ; Silicon Analysts, fabless tape-out cost guide: https://siliconanalysts.com/analysis/fabless-startup-tapeout-cost-guide ; market data: https://siliconanalysts.com/market-data
- MillionMiner, Bitmain S23 XP Hyd: https://millionminer.com/news/bitmain-wdms-2026-s23-xp-hyd-launch
- Khattak and Mikaitis, FP8 accumulation in tensor cores, arXiv 2512.07004: https://arxiv.org/html/2512.07004v2
- Canaan, second-quarter 2026 results: https://finviz.com/news/389337/canaan-inc-reports-unaudited-second-quarter-2026-financial-results ; Bitdeer, third-quarter 2026 results: https://greenstocknews.com/news/nasdaq/btdr/bitdeer-reports-unaudited-third-quarter-2026-results-record-revenue-of-us-320-7-million-adjusted-ebitda-up-173-to-us-69-6-million
- TheEnergyMag: Velaura AI: https://theenergymag.com/news/market-news/velaura-ai-series-a-bitcoin-pivot ; Core Scientific and Block: https://theenergymag.com/news/2026-07-29/core-scientific-exit-block-bitcoin-deal-ai ; AI lease economics: https://theenergymag.com/news/2026-08-06/miner-weekly-ai-pumpability-out-fuel
- Blockworks, Cipher and Fluidstack: https://blockworks.co/news/cipher-mining-inks-3b-ai-hosting-deal-with-fluidstack ; Cointelegraph, Core Scientific: https://cointelegraph.com/news/core-scientific-to-quit-bitcoin-mining-by-year-end-as-ai-business-surges ; ForkLog, Bitfarms: https://forklog.com/en/former-bitfarms-shuts-down-all-mining-sites-in-the-us/ ; IREN, Q1 FY27 results: https://globenewswire.com/news-release/2026/11/05/3397020/0/en/iren-reports-q1-fy27-results.html
- Spark, bitcoin ASIC market concentration: https://www.spark.money/research/bitcoin-asic-manufacturer-market-concentration

- Google, TPU7x documentation: https://docs.cloud.google.com/tpu/docs/tpu7x?hl=es ; The Register, Ironwood: https://www.theregister.com/2025/04/10/googles_7thgen_ironwood_tpus_debut/ ; SemiAnalysis, B300: https://inferencex.semianalysis.com/chips/b300 ; TPU v1 paper, arXiv 1704.04760: https://arxiv.org/abs/1704.04760 ; Glenn Lockwood, MI355X: https://glennklockwood.com/garden/processors/mi355x
- Tom's Hardware, Vera Rubin: https://www.tomshardware.com/pc-components/gpus/nvidia-launches-vera-rubin-nvl72-ai-supercomputer-at-ces-promises-up-to-5x-greater-inference-performance-and-10x-lower-cost-per-token-than-blackwell-coming-2h-2026 ; TrendForce, CoWoS pricing: https://trendforce.com/news/2026/10/27/news-tsmc-reportedly-eyes-20-cowos-price-hike-skips-equipment-price-renegotiations-amid-ai-demand-surge
- Menon Lab, AI inference accelerators compared: https://themenonlab.blog/blog/ai-inference-accelerators-compared ; Introl, AI accelerators beyond GPUs: https://introl.com/blog/ai-accelerators-beyond-gpus-tpu-trainium-gaudi-cerebras-2025
- StratNews (Reuters), domestic AI chips in state-funded data centers: https://stratnewsglobal.com/team-sng/china-orders-state-funded-data-centres-to-use-domestic-ai-chips/ ; TNW, H200 imports: https://thenextweb.com/news/nvidia-h200-chips-hong-kong-china-import-approval
- Succinct, real-time proving on 16 GPUs: https://blog.succinct.xyz/real-time-proving-16-gpus/ ; Cysic ZK ASIC products: https://cysic-docs.pages.dev/hardware-products/zk-asic-products
- Tom's Hardware, Intel Heracles: https://www.tomshardware.com/tech-industry/cyber-security/intels-heracles-chip-computes-fully-encrypted-data-without-decrypting-it-chip-is-1-074-to-5-547-times-faster-than-a-24-core-intel-xeon-in-fhe-math-operations ; IEEE Spectrum: https://spectrum.ieee.org/fhe-intel ; Evertiq, Niobium and SemiFive: https://evertiq.com/design/2026-02-20-semifive-secures-design-win-with-niobium-for-fhe-accelerator ; The Quantum Insider, Niobium funding: https://thequantuminsider.com/2025/12/03/niobium-23m-fhe-funding/ ; Optica OPN, Optalysys: https://www.optica-opn.org/home/industry/2026/february/optalysys_raises_us_$31_million_to_push_photonics_computing_toward_cloud_infrastructure/ ; Zama, HPU on FPGA: https://www.zama.org/post/announcing-hpu-on-fpga-the-first-open-source-hardware-accelerator-for-fhe

โหนดการผลิต

- Silicon Analysts, wafer pricing: https://siliconanalysts.com/data/wafer-pricing
- Tom's Hardware, TSMC A16 pricing rumours: https://www.tomshardware.com/tech-industry/semiconductors/tsmc-could-charge-up-to-usd45-000-for-1-6nm-wafers-rumors-allege-a-50-percent-increase-in-pricing-over-prior-gen-wafers ; Wccftech, SMIC pricing: https://wccftech.com/smic-5nm-chip-prices-up-to-50-percent-more-expensive/amp/
- EUROPRACTICE, TSMC MPW schedule 2026: https://europractice-ic.com/wp-content/uploads/2025/11/TSMC-Europractice-MPW-schedule-2026.pdf
- Horowitz, "Computing's Energy Problem", ISSCC 2014: https://www.Gwern.net/doc/cs/hardware/2014-horowitz-2.pdf
- Tom's Hardware, Huawei Ascend roadmap: https://www.tomshardware.com/tech-industry/artificial-intelligence/huawei-ascend-npu-roadmap-examined-company-targets-4-zettaflops-fp4-performance-by-2028-amid-manufacturing-constraints ; The Substrate, China's AI chip supply chain: https://www.the-substrate.net/p/where-chinas-ai-chip-supply-chain ; Tom's Hardware, Trainium3: https://www.tomshardware.com/tech-industry/artificial-intelligence/amazon-launches-trainium3-ai-accelerator-competing-directly-against-blackwell-ultra-in-fp8-performance-new-trn3-gen2-ultraserver-takes-vertical-scaling-notes-from-nvidias-playbook

โอกาสและเศรษฐศาสตร์การเปิดตัว

- Webull (Benzinga), Broadcom earnings call: https://www.webull.com/news/15511095937917952 ; Quartr, Global Unichip Q2 2026: https://quartr.com/events/global-unichip-corp-3443-q2-2026_F9×6NJQY ; Taiwan News, Alchip: https://www.taiwannews.com.tw/news/6370946 ; Design & Reuse, VeriSilicon: https://www.design-reuse.com/news/202530457-surge-in-custom-chip-business-propels-verisilicon-s-q1-revenue-to-double-with-new-orders-skyrocketing-to-8-2-billion-yuan-increased-r-d-investment-leads-to-wider-net-loss-earnings-insight/ ; The Block, Canaan's AI unit: https://www.theblock.co/post/359250/crypto-miner-canaan-to-shutter-non-core-ai-semiconductor-business-unit-amid-strategic-realignment
- Clutch, AI development pricing: https://clutch.co/developers/artificial-intelligence/pricing ; PixelPlex, blockchain development cost: https://pixelplex.io/blog/blockchain-development-cost/
- Hyperbolic, proof of sampling: https://hyperbolic.ai/blog/deep-dive-into-hyperbolic-proof-of-sampling ; Gensyn, Verde: https://www.gensyn.ai/research/verde-a-verification-system-for-machine-learning-over-untrusted-nodes ; layered verification, arXiv 2512.20176: https://arxiv.org/abs/2512.20176v1
- Sherlock, audit pricing reference for 2026: https://sherlock.xyz/post/smart-contract-audit-pricing-a-market-reference-for-2026
- Blockworks, Binance listing standards: https://blockworks.com/news/binance-listing-standards-scrutiny ; Cointelegraph, market-maker deals: https://cointelegraph.com/magazine/market-maker-deals-quietly-killing-crypto-projects ; Blockworks, Movement: https://blockworks.co/news/movement-rushi-manche-suspension-market-maker-web3port ; Nansen, HIP-3: https://nansen.ai/post/what-is-hip-3-hyperliquid ; The Block, Monad: https://www.theblock.co/post/380094/monad-mainnet-launches

การพิสูจน์งานที่มีประโยชน์ · คริปโต × ชิลิคอน

การประมวลผลคือฉันทามติ

AI × คริปโต × ชิป ในปี 2026 เริ่มจากการพิสูจน์งานที่มีประโยชน์ของ Pearl

ตัวเลขสำคัญทุกตัวในรายงานนี้ระบุแหล่งที่มาและวันที่ ตัวเลขที่ผู้เขียนคำนวณเองจะติดป้ายว่าเป็นประมาณการ ตัวเลขที่มีเพียงผู้ขายรายงานจะติดป้ายว่าเป็นข้ออ้างของผู้ขาย รายงานนี้พิจารณาเฉพาะด้านเทคโนโลยีและธุรกิจ

ทัน ชว่าย

นักเทคโนโลยีและผู้ก่อตั้ง · นักวิจัยและนักพัฒนา AI

tanshuai.com · i@tanshuai.com

ฉบับเผยแพร่สาธารณะ · ข้อมูล ณ วันที่ 13 พฤศจิกายน 2026

