xAI「Grok 4.20」徹底調査報告

誇張と事実の分離、そして実運用へのロードマップ

stamp

Investigation Date: 2026-02-23
Status: CONFIDENTIAL ANALYSIS

最重要結論：待つべきか、動くべきか

現状

- 「Grok 4.20」の公式仕様は未確定
- API側では「**Grok 420 / 420 Multi-Agent**」が **'Coming Soon'** として早期アクセス募集中

技術的真実

- xAIの核心は「**大規模RL**」と「**サーバーサイド・エージェント**」へのシフト
- **Grok 4.1 Fast + Agent Tools API**は既に利用可能であり、**運用負荷が低い**

推奨アクション

Start Now

確定情報が揃っている「Grok 4.1 Fast」で設計を開始せよ

Wait

「420」の本番投入は、価格とSLAが公開されてから判断せよ

定義のギャップ：「4.20」と「420」の違い

噂 / コミュニティ

Grok 4.20 (Public Beta)

Source tag: SNS, Media, Hacker News

具体的な仕様、ベンチマーク、価格は「不明（Unknown）」

「リンクはあるが仕様が出ておらず議論しにくい」(Hacker News)

公式 / APIドキュメント

Grok 420 & Grok 420 Multi-Agent

Source tag: xAI Official Docs / API Reference

Status: 'Coming Soon' (Early Access Request available)

Key Insight Box

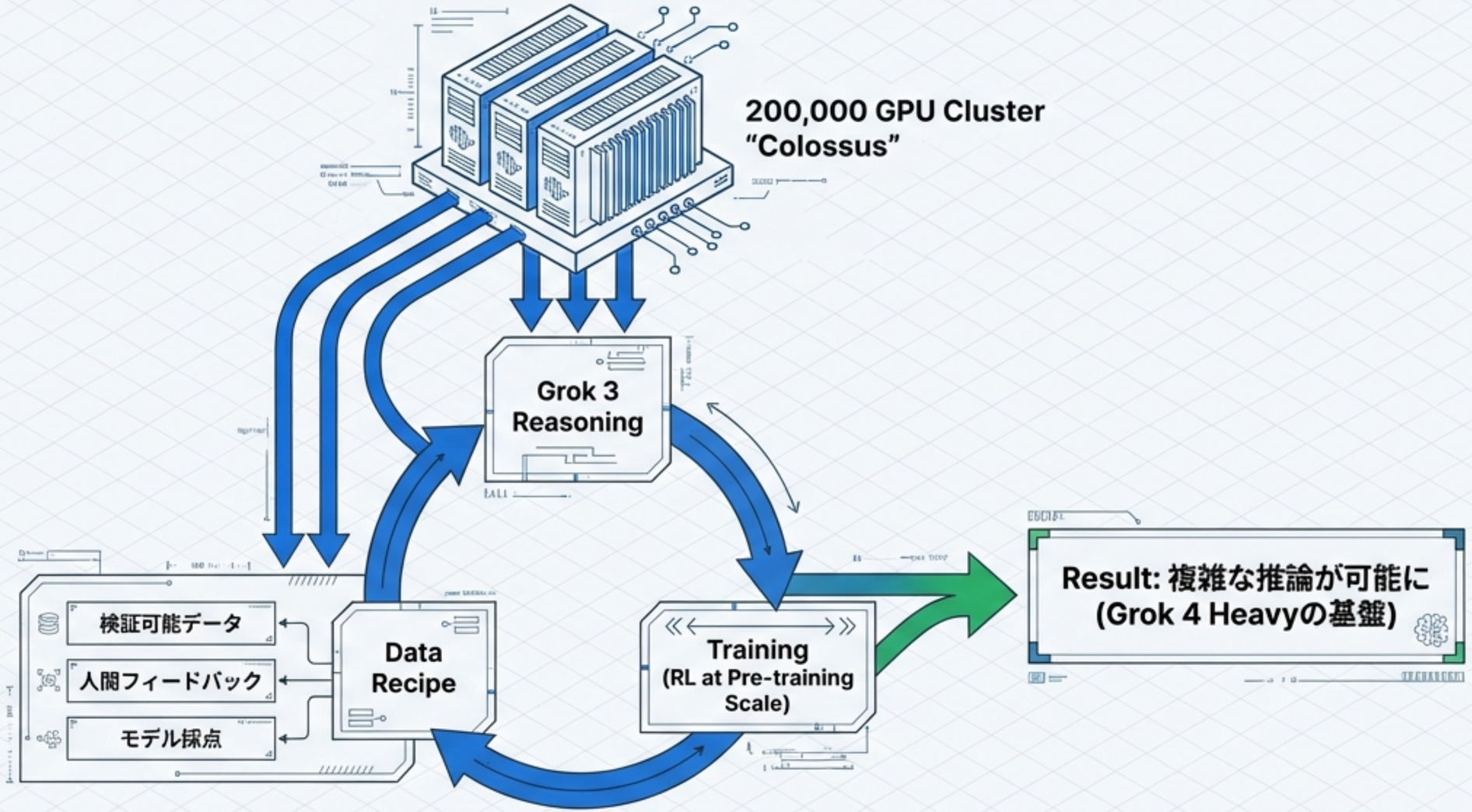
4.20は単なるバージョンアップではなく、「Multi-Agent」化が焦点である可能性が高い

イノベーションの加速：Grok 4からの系譜

Graphite: 短期間で「強化学習」「高速化」「ツール統合」を段階的に実装

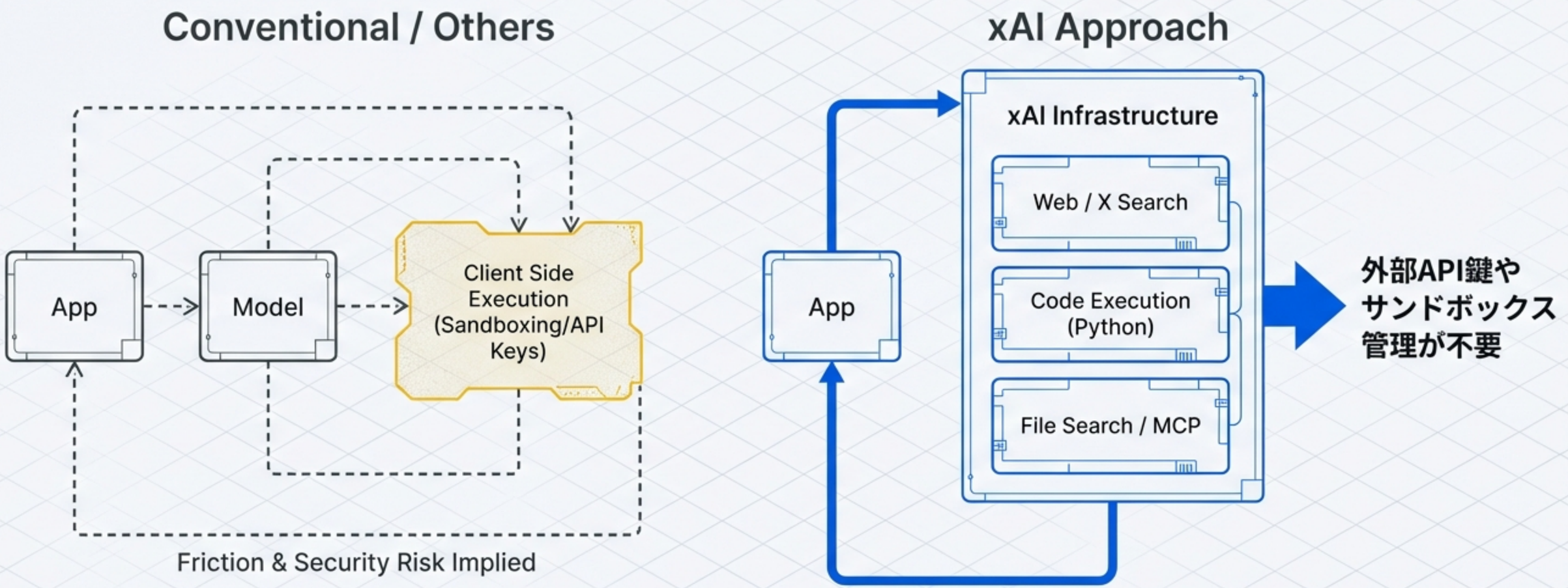
コア技術①：事前学習スケールでの強化学習（RL）

推論能力の強化は、単なるデータ量ではなく「検証プロセス」にある



コア技術②：サーバーサイド Agent Tools

開発者の運用負債を削除するアーキテクチャ

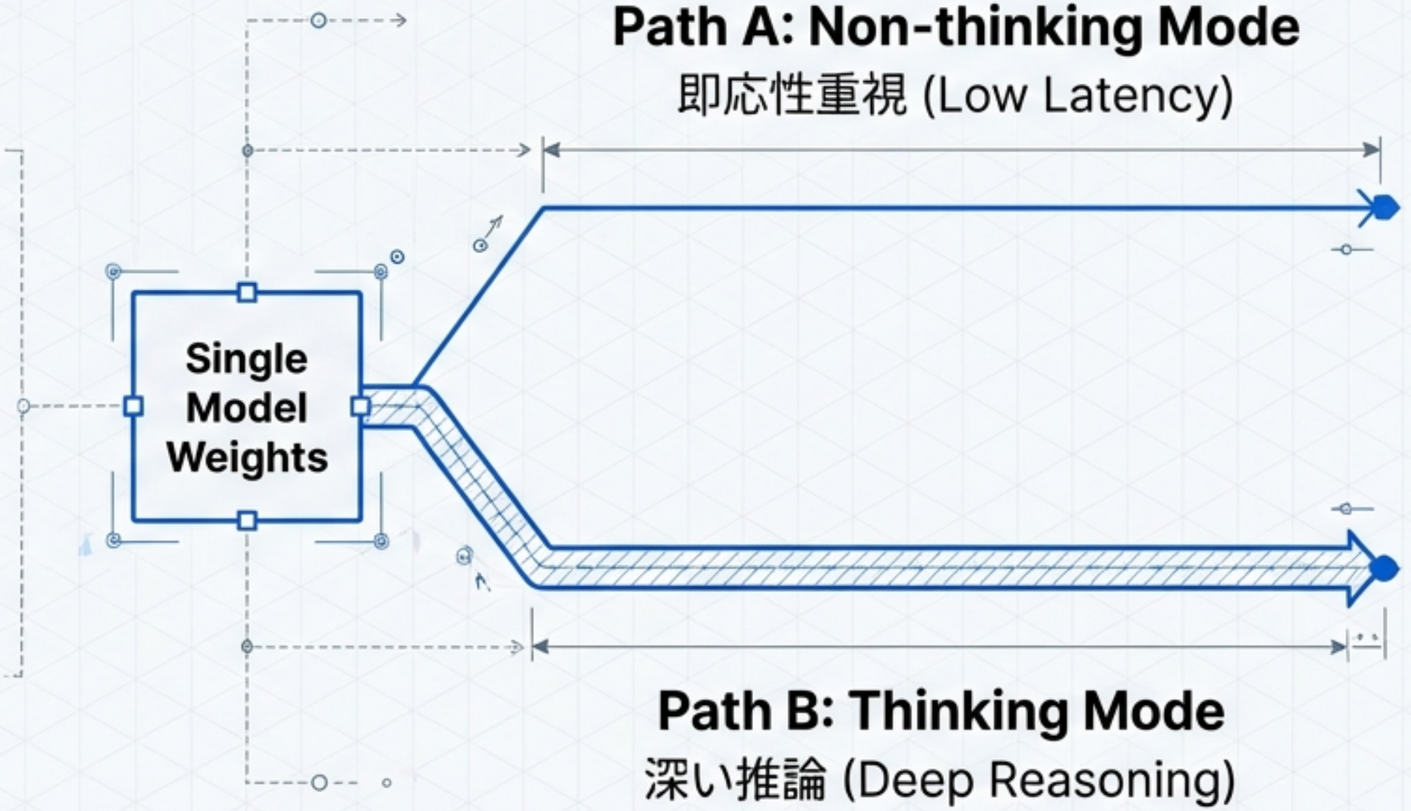


コア技術③：統合アーキテクチャと効率化

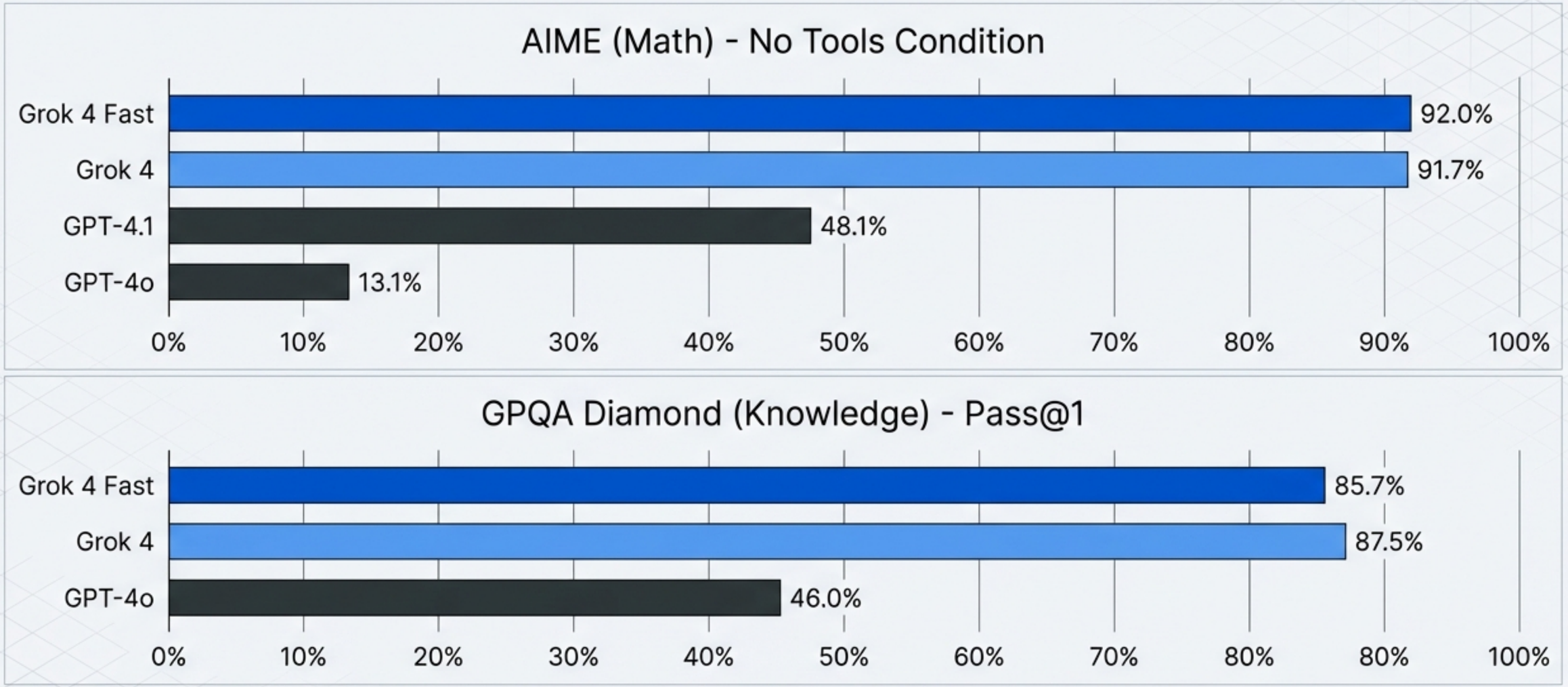
Graphite #2D3436
Inter font

2,000,000
Tokens Context

TTFT: 0.52s
(Grok 4.1 Fast reference)



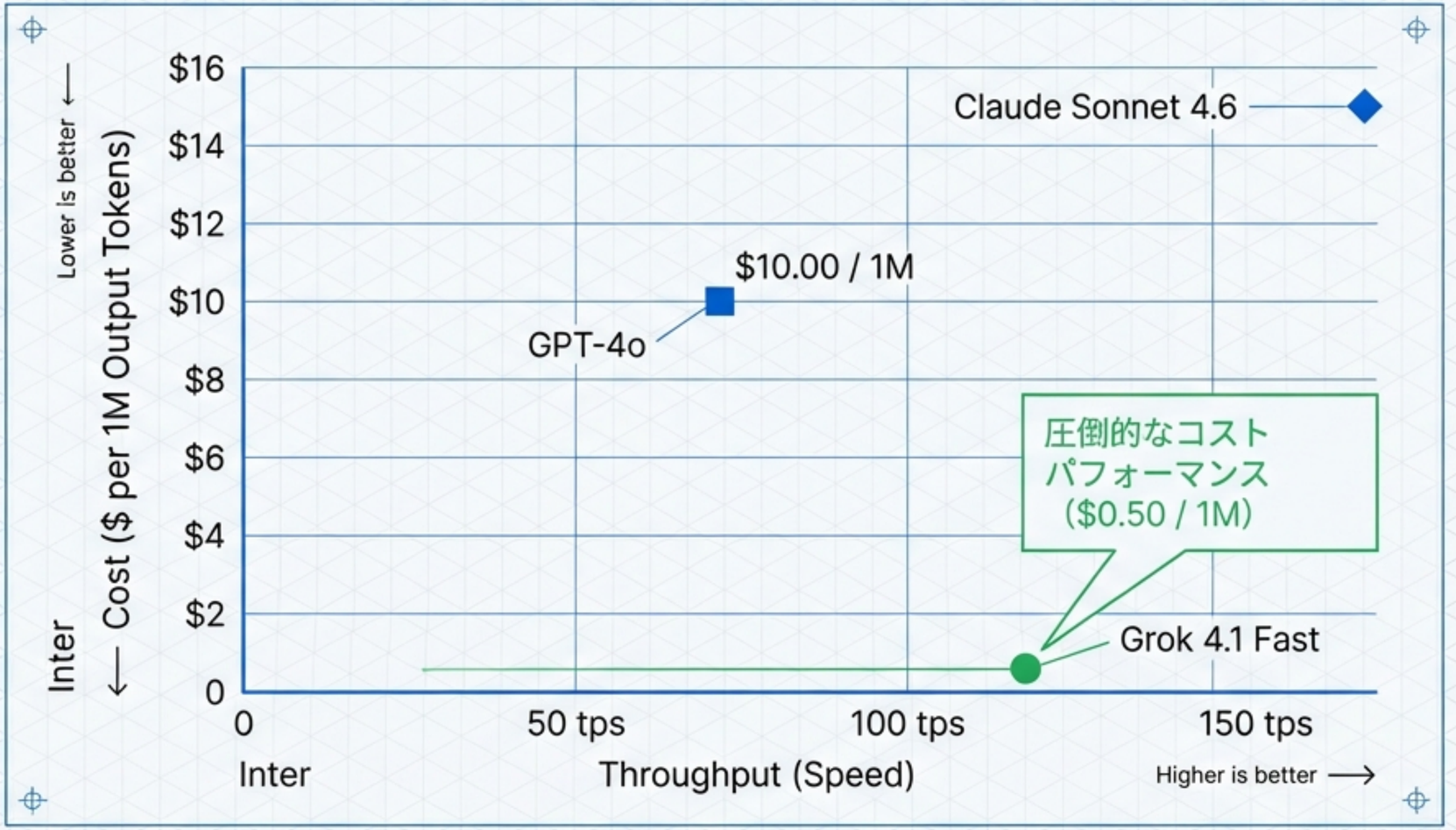
性能評価：ベンチマーク戦争



条件の差異 (Pass@1 vs Consensus) に注意が必要. xAIはツールなしでの性能を強調.

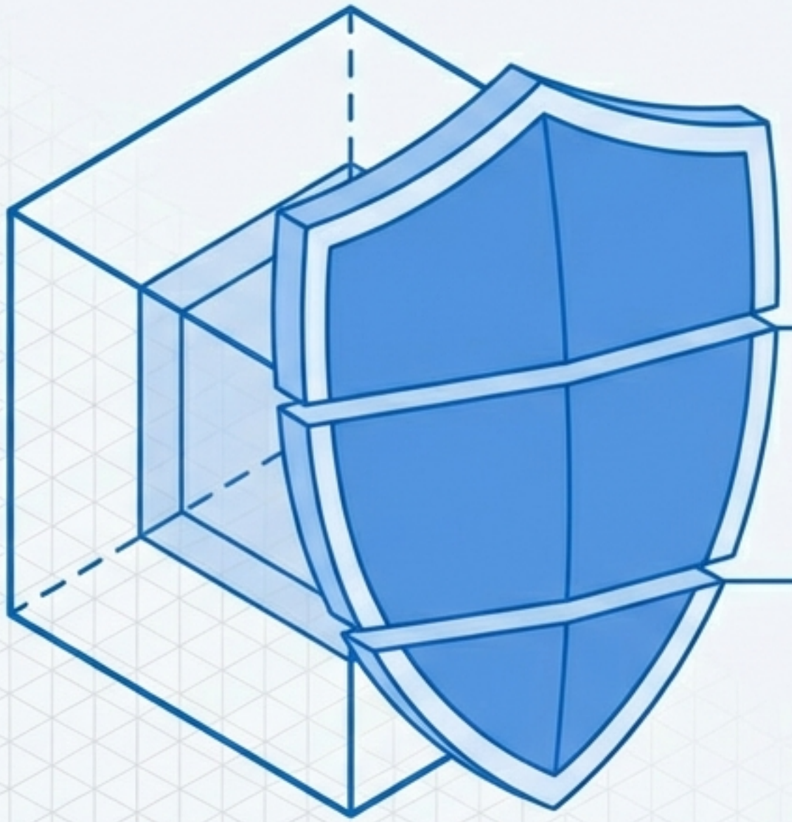
ANALYSIS REPORT
SERIAL N0063388

経済性：コストとスループットのバランス



安全性とガバナンス（RMF）

Based on xAI Risk Management Framework



Input Filters

- Jailbreak resistance
- AgentHarm filters

Evaluation

- AgentDojo
- MASK (dishonesty rate)

Policy

Vetted Users only for catastrophic risk tiers

Key Risks Managed

1. Malicious Use (CBRN, Cyber attacks)
2. Loss of Control (Deception/欺瞞, Model Autonomy)

Technical Forensics

プライバシーとデータ主権

⚠ Caution

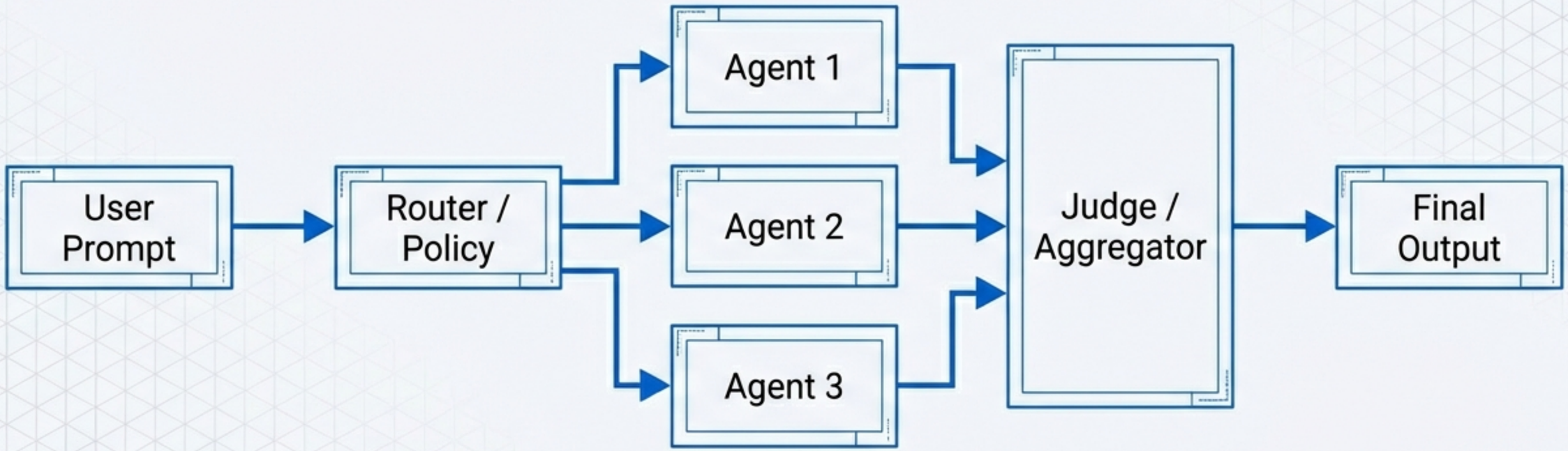
Caution

- **User Content: 入力データが出力に再現される可能性がある.**
- **Training: デフォルトではユーザーデータが学習に使用される可能性があるため、オプトアウト設定やEnterprise契約の確認が必要.**

Compliance: SOC 2 Type 2, HIPAA (BAA available).
Risk Warning: 過去のAPIキー管理に関する指摘 (GitHub leaks) に留意し、厳格な鍵管理を推奨.

「Multi-Agent」の正体（技術的推定）

Parallel Test-Time Compute (並列テスト時算出)



Source: Based on Grok 4 Heavy architecture descriptions.

競合ランドスケープ：どこでxAIを選ぶべきか

xAI (Grok 4.1 Fast) Best for: Real-time Data (X), Server-side Agents, Cost Efficiency.	**OpenAI (GPT-4o)** Best for: SLA Reliability, Proven Stability.
Anthropic (Claude) Best for: Coding, Complex Reasoning.	**Google (Gemini)** Best for: Long Context, Native Multimodal.

xAIの強みは「リアルタイム性」と「ツールのサーバーサイド実行」にある.

戦略的推奨：アクションプラン

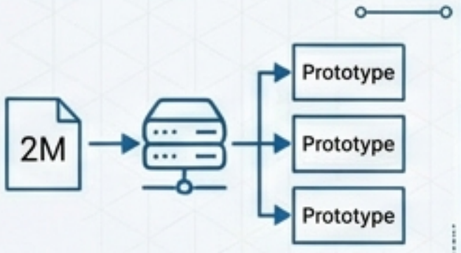
Green Light
(GO - Build Now)

Target

Grok 4.1 Fast + Agent Tools

Action

2M文脈とサーバーサイドツールを活用したプロトタイプ開発. 運用コスト削減を狙う.



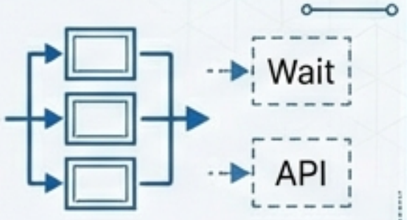
Yellow Light
(YIELD - Prepare)

Target

Multi-Agent (420)

Action

アーキテクチャ（並列処理）を想定した設計準備. API仕様公開待ち.



Red Light
(STOP - Risks)

Target

High-Sensitivity PII

Action

入力データの学習利用ポリシーがクリアになるまで、個人情報の直接入力は避ける.

Technical Forensics

参照元・一次情報（References）

Official Docs:
xAI API Reference / Models

Model Cards:
Grok 4, Grok 4.1

Framework:
xAI Risk Management Framework (RMF)

Legal:
Privacy Policy / Security FAQ

Docs

Models

RMF

Legal

本報告書は2026-02-23時点の公開情報に基づき作成されています。