Xiaomi MiMo-V2.6-Pro 完全解剖

1兆パラメータのオープンウェイトMoEがもたらす革新と、本番導入のリアリティ

[1] オープン首位の性能
（AA Intelligence Index 46.32）

[2] 破壊的低コスト
（商用モデル比 1/30のAPI価格）

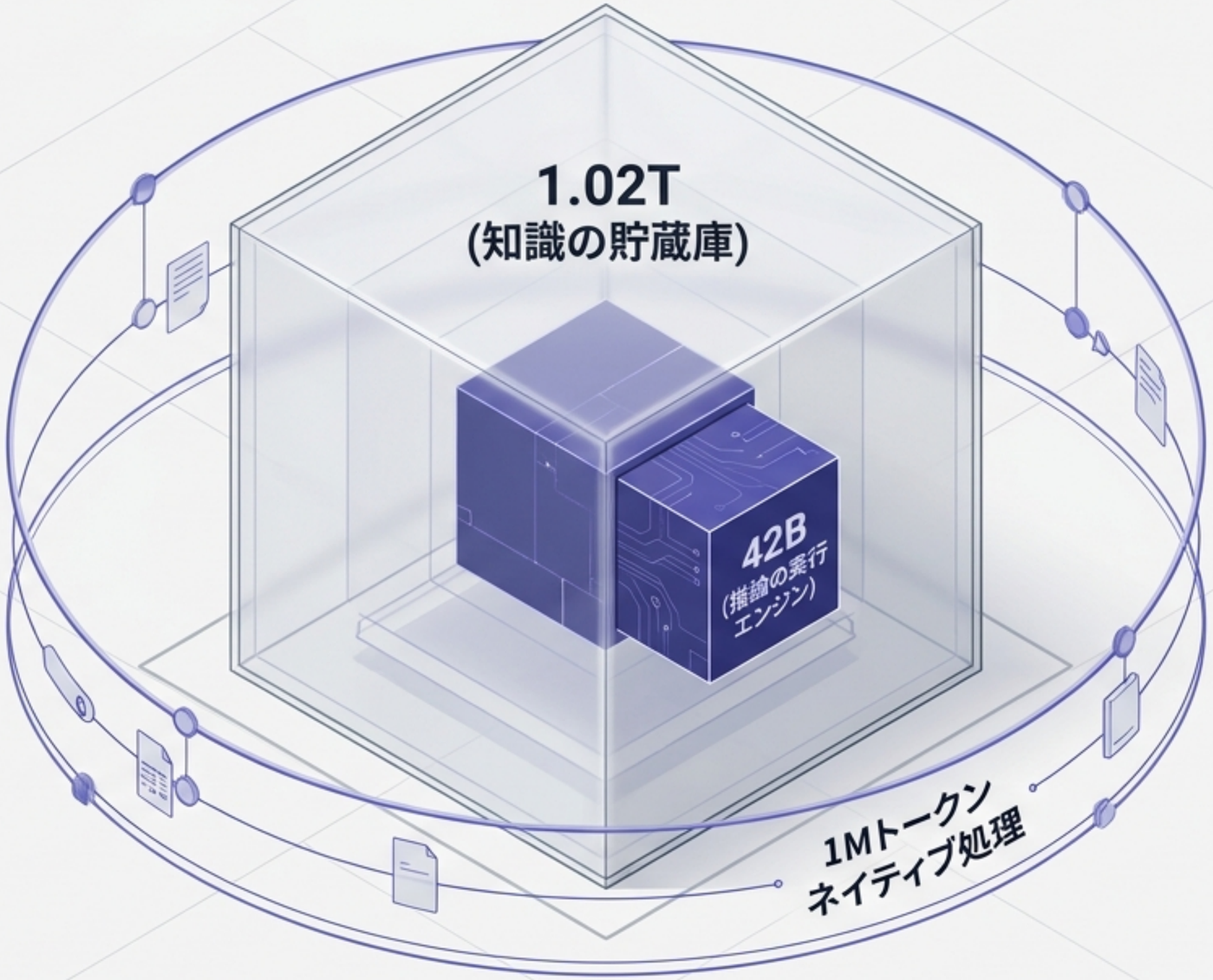
[3] 越えるべきインフラの壁
（TP16/EP16の必須要件）

全モダリティ (Omnimodal) 基底モデルの正体

総パラメータ
1.02T (1兆200億)

トークン活性化
42B (420億)

キーインサイト: 中間変換を挟まず、テキスト・画像・音声・動画を同一シーケンス空間で直接処理する統合型アーキテクチャ。巨大な知識を保持しつつ、推論時には軽量に動作するスパースMoEの基本特性を持つ。

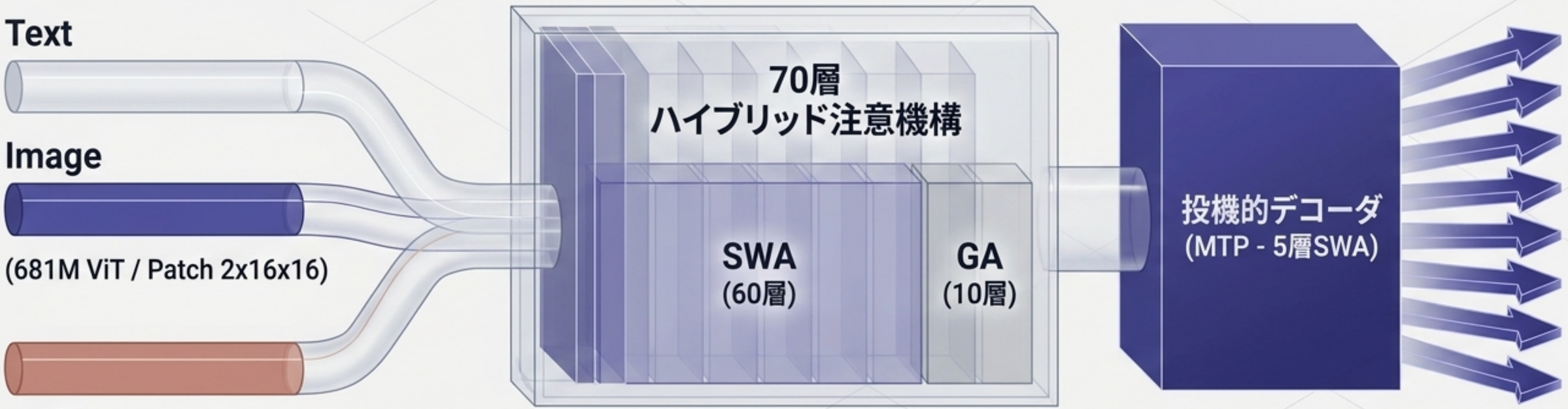


統合処理パイプラインと高速化メカニズム

直接統合：
翻訳層なしで直接トークン空間へマッピング

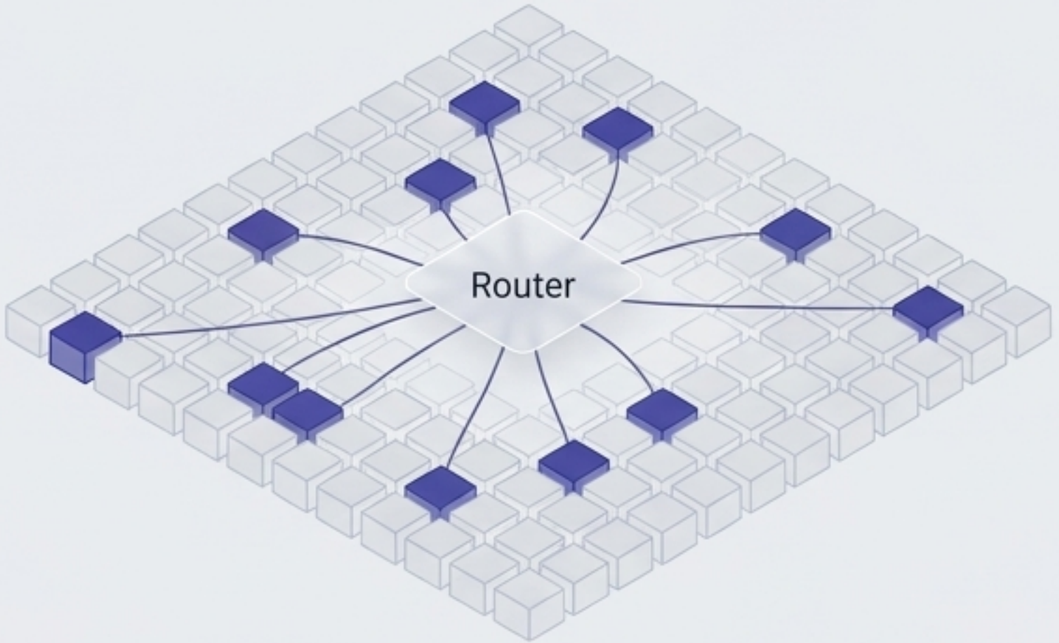
SWA/GAハイブリッド：
ローカル(128トークン)とグローバル注意機構の混成による低レイテンシ処理

MTPによるスループット向上：
1回のパスで後続7トークンを並列予測・検証し、MoE特有のメモリ帯域ボトルネックを緩和



384エキスパートの制御と「ルータ凍結」によるスケーリング

スパースMoE構造



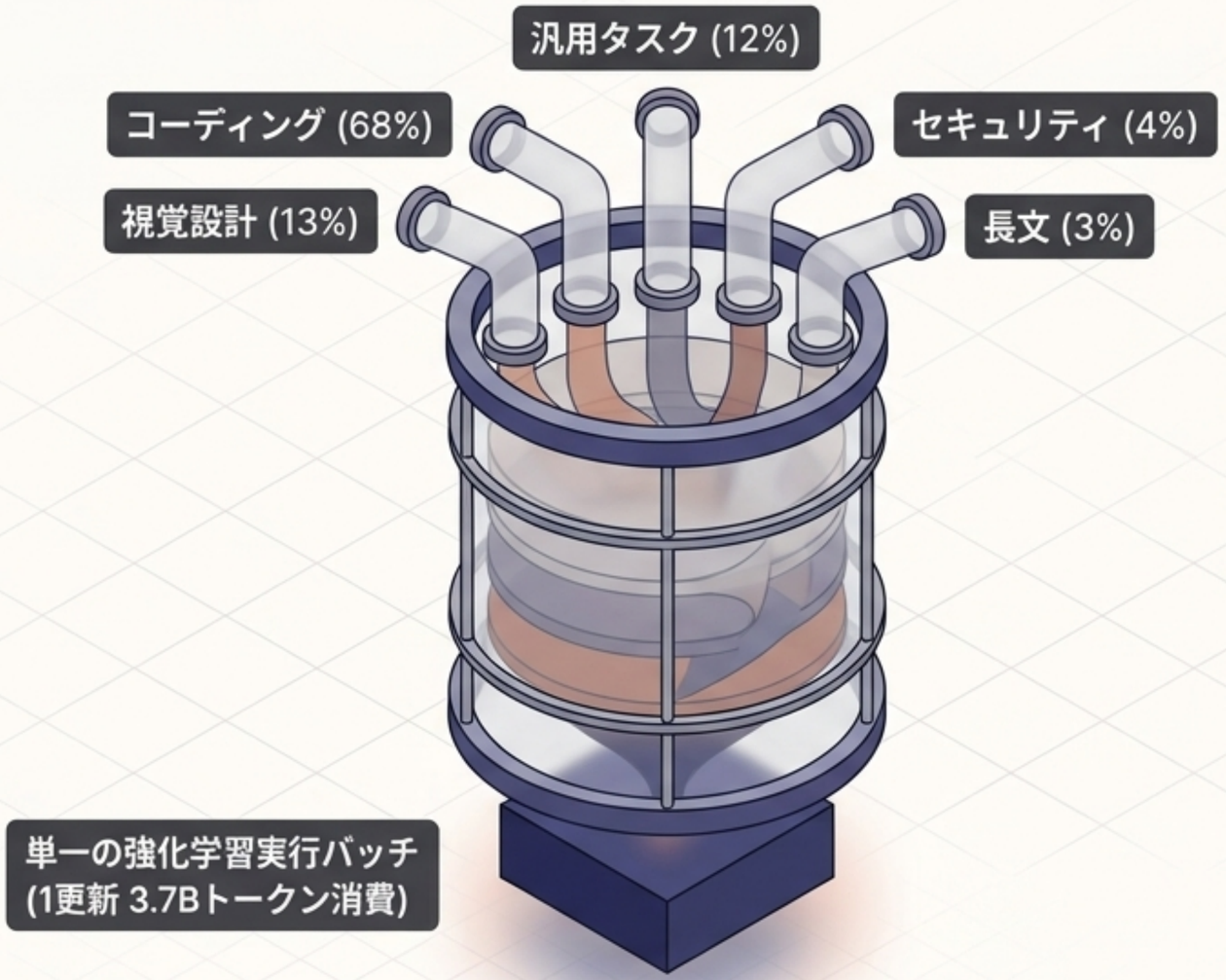
384個のエキスパートから、トークンごとに最適な8個を動的に選択（Top-8活性化）。

課題　　解決策

課題: RL中にMoEルータ重みを更新すると、一部のエキスパートに負荷が極端に集中し崩壊する。

解決策 (Router Freezing): 訓練全体を通じて「MoEルータの重みを完全凍結」することで、エキスパート内部の表現力を維持したまま、安全なバッチスケーリングを実現。

新RLパラダイム YORO: 単一混合実行の相互強化

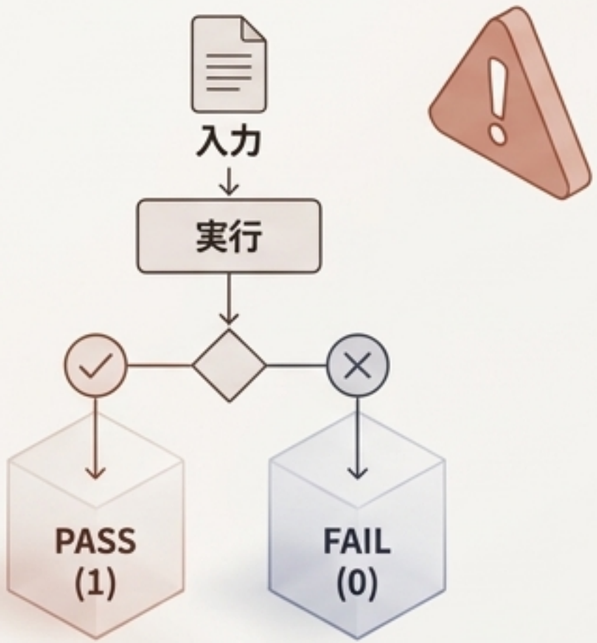


相互強化作用 (Cross-Domain Transfer)

従来のようにドメインごとに独立して学習させるのではなく、単一のRL工程に全領域を混在させる。これにより、コード生成で培われた「厳密な論理検証能力」が、サイバーセキュリティや視覚推論のタスクへと転移し、モデル全体の推論性能を劇的に底上げする。

評価エンジンGAR: 報酬ハッキングを無効化する相対評価

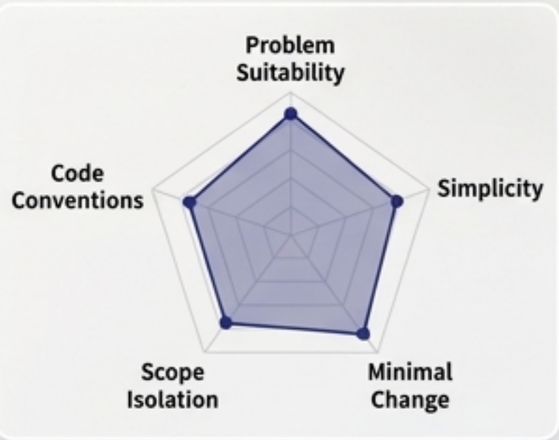
従来の二値評価 (0 or 1)



テストを通過させるためだけの「冗長で防衛的なハックコード」が正解となり、モデルの思考パスが劣化する。

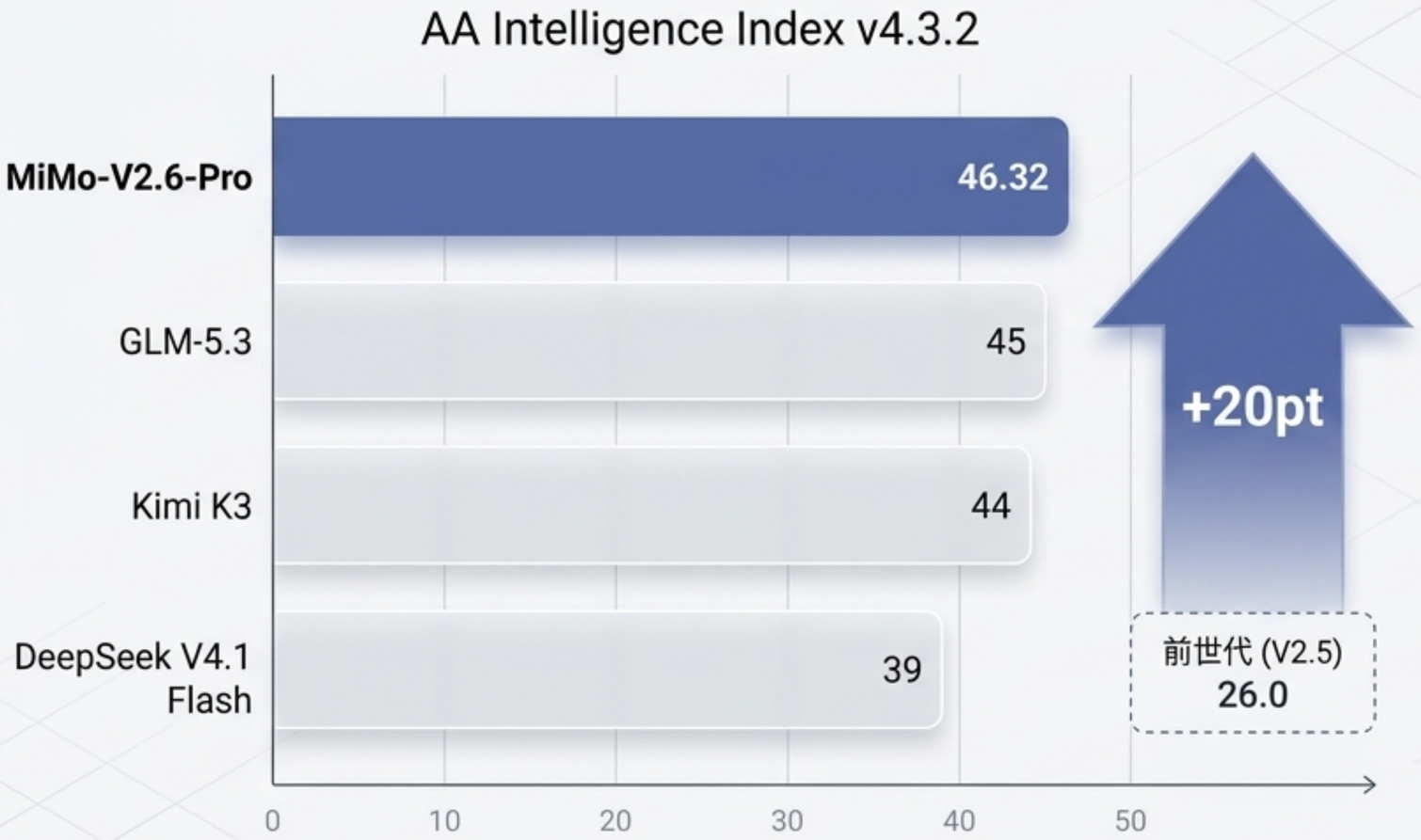
GAR (Groupwise Advantage Redistribution)

同一プロンプトに対する16個の回答を同一空間で相対評価。報酬ハッキングはスコアゼロとして排除。



- 問題適合性 (解法が直接的か)
- 実装の簡潔性 (過剰な防御の排除)
- 変更の最小性 (Diffの最小化)
- スコープ隔離性 (偶発的改変の検知)
- コーディング規約 (既存作法との整合)

オープンウェイト市場における圧倒的優位性



飛躍を牽引したエージェント性能

前世代からの劇的な進化 (+20pt) の中核は、ツール呼び出しと環境フィードバックを伴うエージェント系タスクにある。

- Terminal Bench 4.0: 1.5 → **34.9**
- AutomationBench: 16.0 → **53.1**

フロンティアモデルとの能力差（強みと現在地）

	MiMo-V2.6-Pro	Claude Opus 5	GPT-6 Astra
デスクトップ操作 (AutomationBench)	53.1 ↑	45.8	46.2
視覚コーディング (MiMo VisualCoding)	72.3 ↑	73.4	69.1
高度セキュリティ探索 (ExploitBench)	47.9 ↓	70.0	100.0 ↑

優位・匹敵領域 (Agent & Vision)

デスクトップ操作や視覚コーディングにおいて、商用最高峰のClaude Opus 5と同等以上のパフォーマンスを達成。エージェントとしての実務遂行能力は極めて高い。

ギャップ・限界領域 (Extreme Security)

高度な脆弱性探索・悪用 (Exploit) においては、GPT-6 Astraに圧倒的な大差をつけられている。長期推論探索の安定性や安全機構の最適化深度には明確な壁が存在する。

破壊的経済性とエコシステムの展開

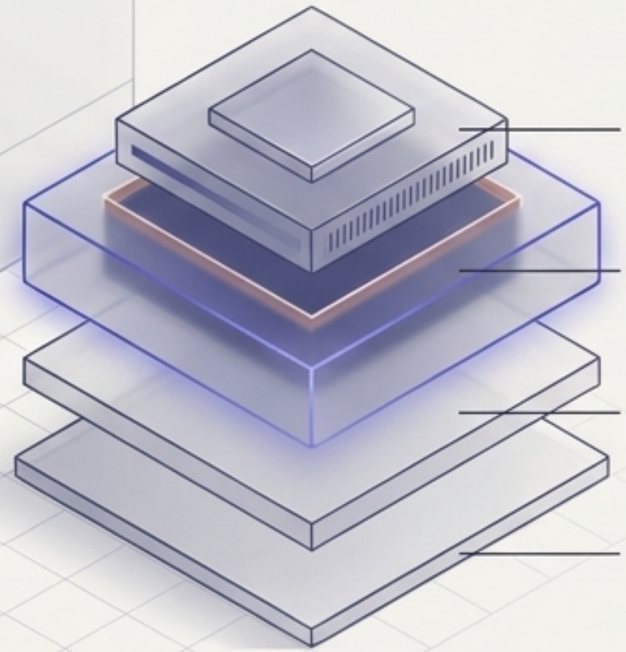
THE COST DISRUPTION

1/30

Kingy.ai 実測検証		
Claude Opus 5	タスク完遂	消費コスト: **$1.03**
MiMo-V2.6-Pro	タスク完遂	消費コスト: **$0.03**

API価格 (1M tokens): 入力 **$0.435** / 出力 **$0.87**。
キャッシュヒット時の加重平均コストは約 **$0.18** に圧縮され、商用モデル比で圧倒的なROIを実現。

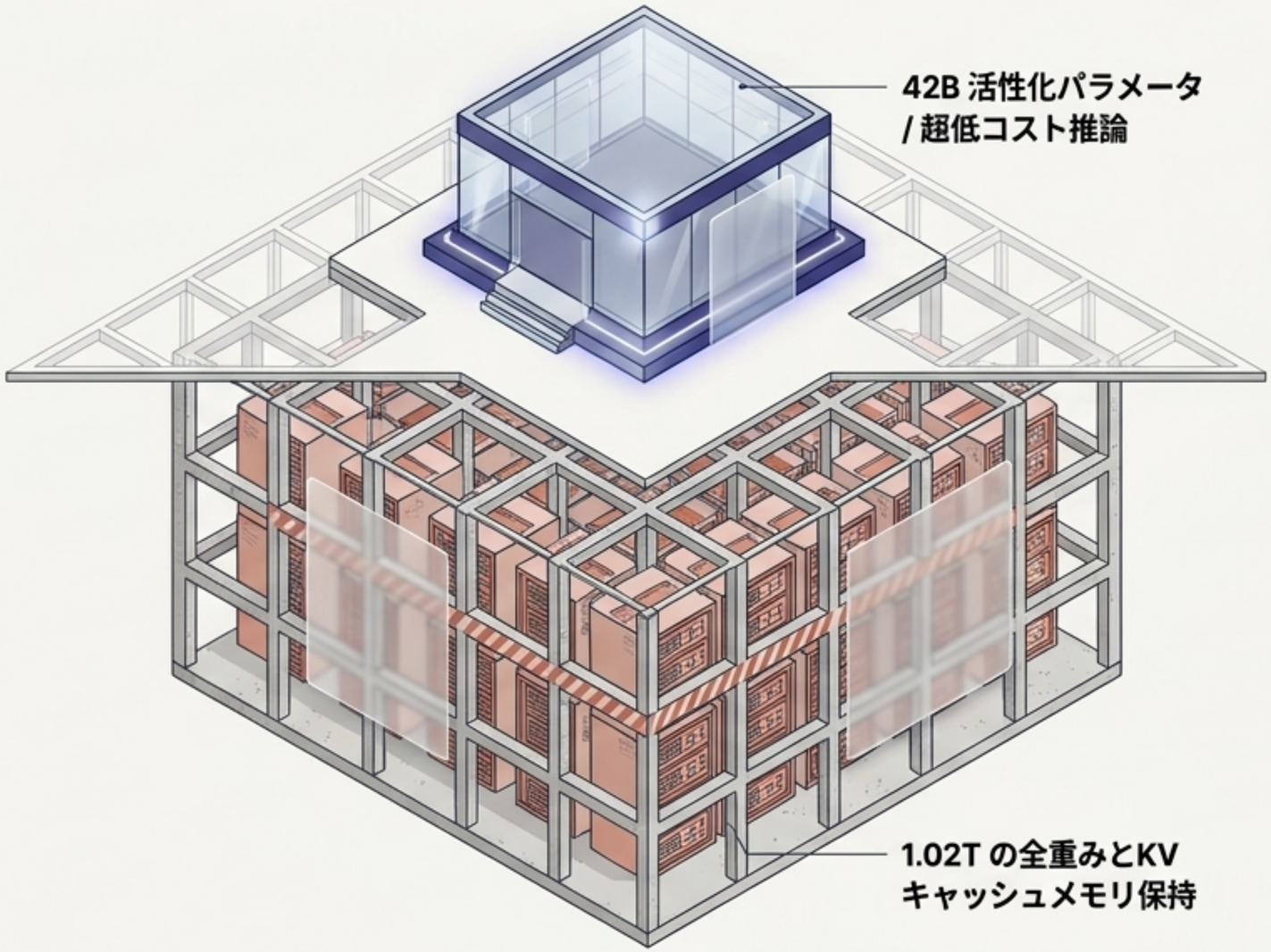
THE ECOSYSTEM ARCHITECTURE



- **UltraSpeed** (最大20倍の低遅延層)
- **Pro** (1.02T フルスペックMoE)
- **Flash** (軽量高効率版 15B)
- **Distill-Qwen-9B** (オープンRL研究用)

MiMo Desktop:
PC操作を自動化する専用クライアント

Reality Check 1: 本番導入を阻むインフラ要件の壁



マルチノードクラスタの必須化

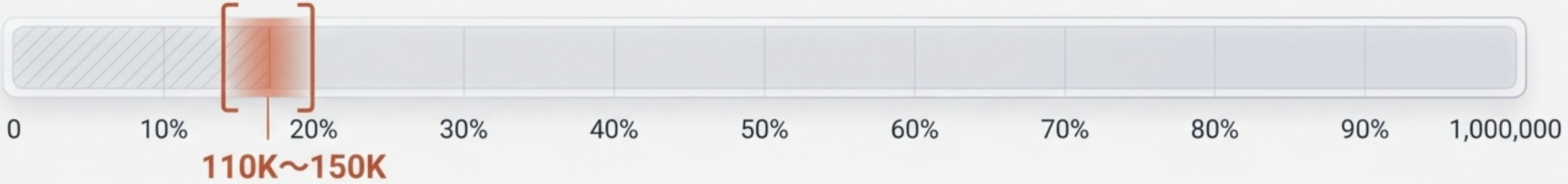
推論には1.02Tの全パラメータのメモリ常駐が不可欠。公式推奨は TP16 + EP16。

単一ノードやオンプレミスでの自前運用は極めて非現実的であり、クラウドインフラ依存が前提となる。

SPOF（単一障害点）リスク

現状、APIプロバイダが限られており、レートリミット (HTTP 429) やエンドポイント障害時の自動フェイルオーバー構築が困難。可用性設計が重大なボトルネックとなる。

Reality Check 2: 長文コンテキストの真実と「自己改善」の限界



コンテキスト長の実効性: 最大仕様は1Mトークンだが、実際のRL検証軌跡長は110K〜150Kに留まる。1M近傍での複雑な因果推論の信頼性は未検証。

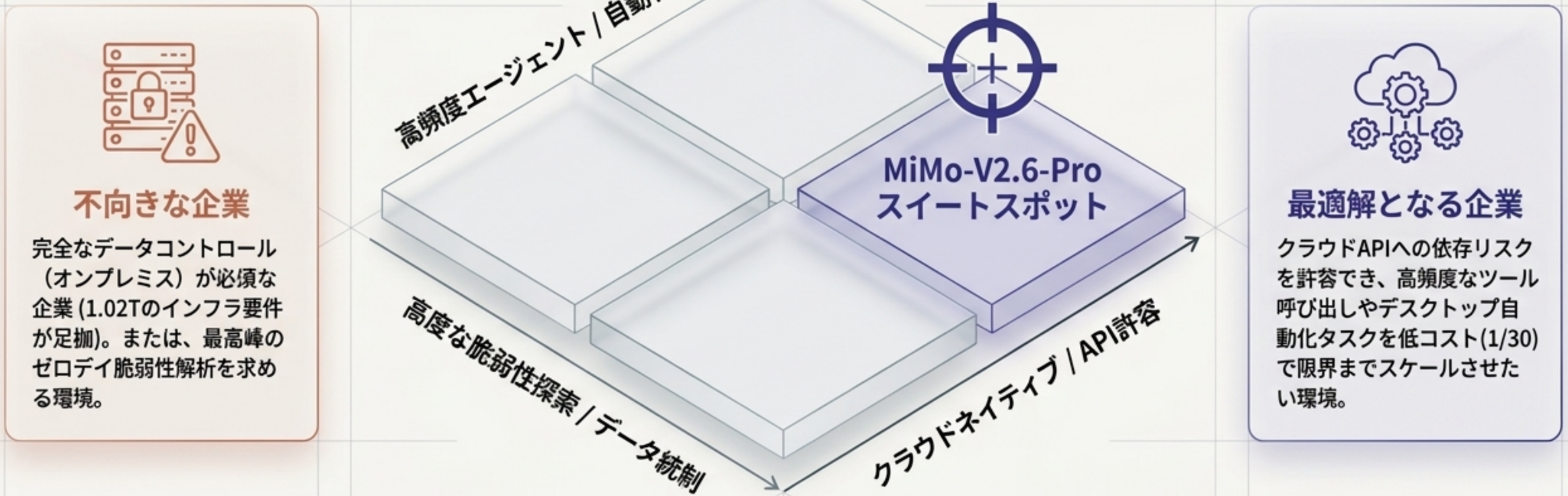
「自己改善（RSI）」への過大評価

モデルが完全に自律改変しているわけではない。環境選定、ルール策定、GARアルゴリズム調整などの**「外側のループ（Outer Loop）」**は人間の研究チームが制御する**『人間主導の自己改善支援ループ』**である。

ベンチマーク過剰適合の懸念

Epoch AIの監査により、評価の主軸であるDeepSWEにおいてテスト側の不備 (偽陰性) が**23問**発見された。これらが訓練時の内部最適化メトリクスとして機能してしまった過剰適合のリスクが指摘されている。

導入のための意思決定: 誰にとって「最強」か？



不向きな企業

完全なデータコントロール（オンプレミス）が必須な企業 (1.02Tのインフラ要件が足枷)。または、最高峰のゼロデイ脆弱性解析を求める環境。

最適解となる企業

クラウドAPIへの依存リスクを許容でき、高頻度なツール呼び出しやデスクトップ自動化タスクを低コスト(1/30)で限界までスケールさせたい環境。

アーキテクトの総括: 人間主導のRLが生んだ現在最高峰のエージェント実行エンジン。
ただし、インフラ設計はクラウドネイティブを大前提として描くべし。