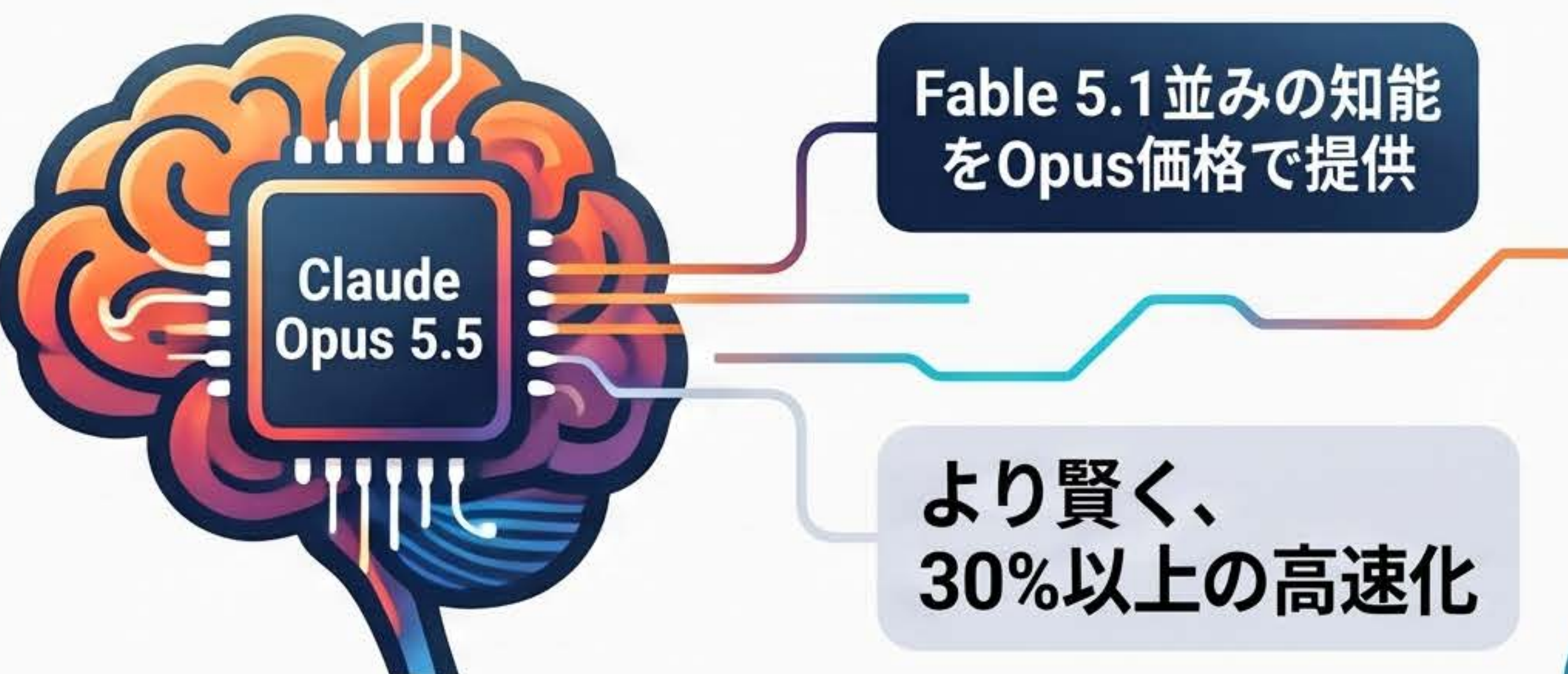


次世代フラグシップ「Claude Opus 5.5」：Fable級の知能を圧倒的なコスト効率で

Claude Opus 5.5の主要な性能向上、劇的なコスト削減、および導入時に必須となる仕様変更

スペックの飛躍と驚異的なコストパフォーマンス



キャッシュ読み取り 単価を60%削減

100万トークンあたり\$0.20、エージェント適用のコストを劇的に抑制



100万トークンの 巨大コンテキストを標準化

長文利用の追加料金を撤廃。大規模リポジトリの移行や監査が1日で完了可能に

前世代モデル(Opus 5)と比較した際の、API利用料の明確な引き下げ

料金項目 (100万トークンあたり)	Claude Opus 5.5	Claude Opus 5	変化率
通常入力 (Base Input)	\$4.00	\$5.00	-20.0%
通常出力 (Base Output)	\$20.00	\$25.00	-20.0%
キャッシュ読み取り (Cache Read)	\$0.20	\$0.50	-60.0%

実運用・移行時に不可欠な 3つの重要ポイント



思考(Thinking)プロセスの常時義務化
無効化は不可(400エラー)。推論の深さは新設の`effort`パラメータで制御



`effort`設定による「コスト逆転」に注意
⚠️ 設定「max」では思考が深まり、総コストがOpus 5を20%上回る場合がある。通常は「medium」を推奨



特定分野での「自動ルーティング」発生
Opus 4.8 / Opus 5

サイバー攻撃や生物学タスクでは、安全のためOpus 4.8やOpus 5へ自動で切り替わる