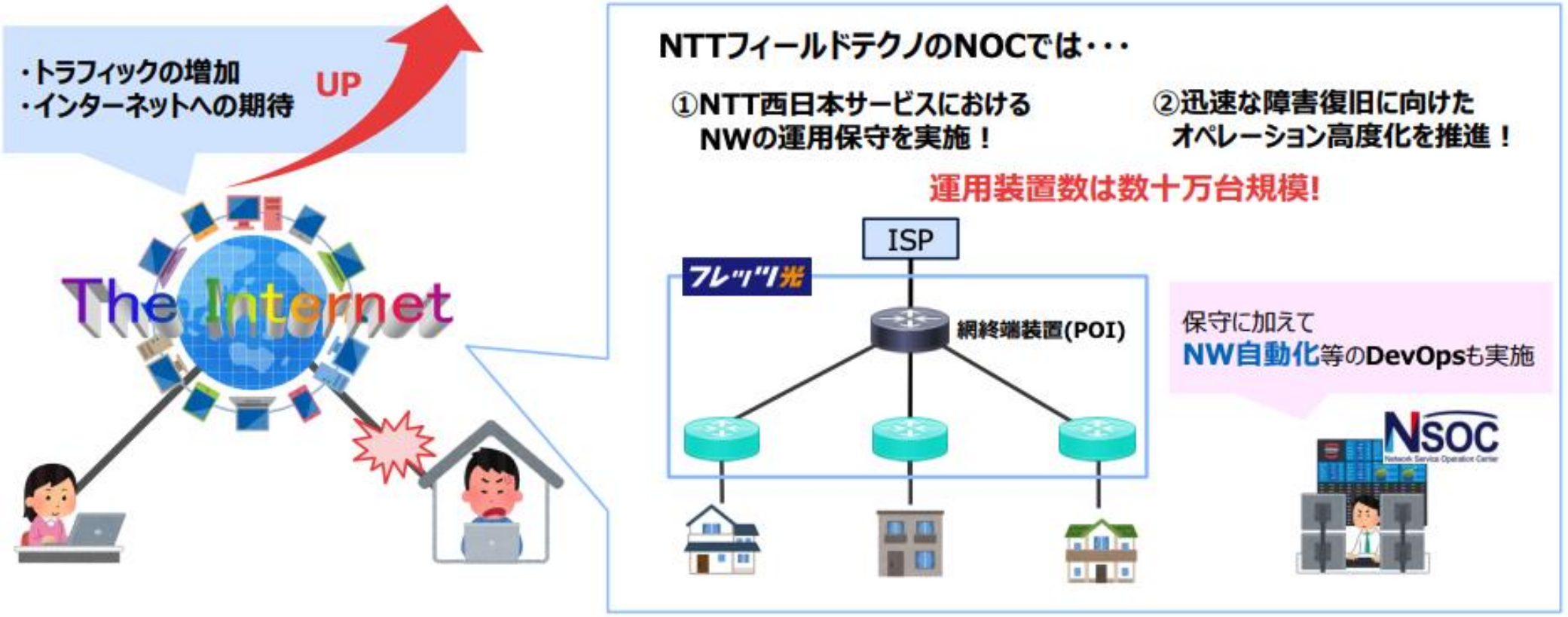
AIOpsのためのAgent Harness運用と課題について

株式会社NTTフィールドテクノ 佐藤 亮介 重松 史哉
NTT株式会社 田中 文貴

NTTフィールドテクノ業務紹介

- リモートワークの普及や大規模なネットワーク障害を通してNW保守の重要性が高まっており
- NTTフィールドテクノのNOCでは日々迅速な障害復旧に向けて努めています



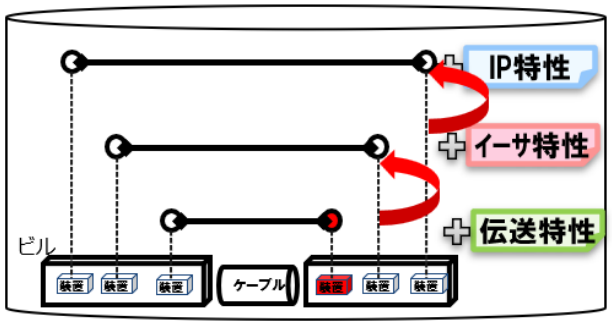
NTT アクセスサービスシステム研究所　業務紹介

- お客様とNTTビルを結ぶアクセスネットワークに関する研究・開発

研究内容の一例

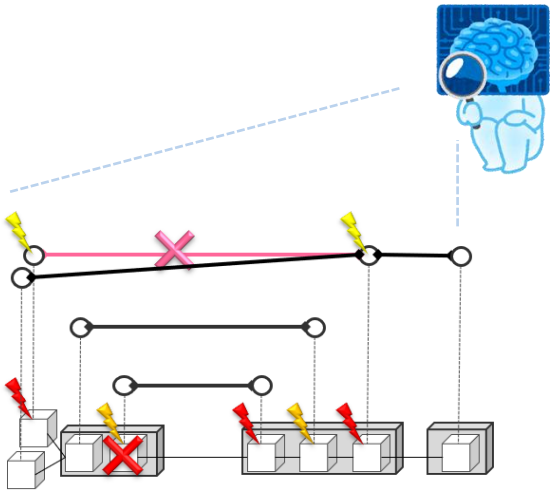
NWリソース管理

NW構成情報管理モデルによる
汎用的データ形式を利用した
NW情報の情報活用基盤

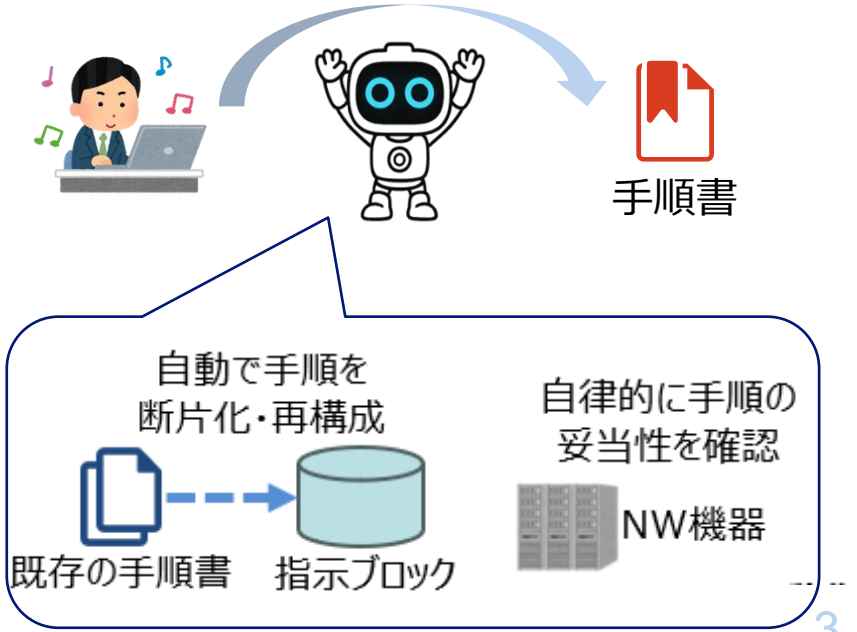


自律的故障箇所推定

マルチレイヤにおける
故障箇所推定

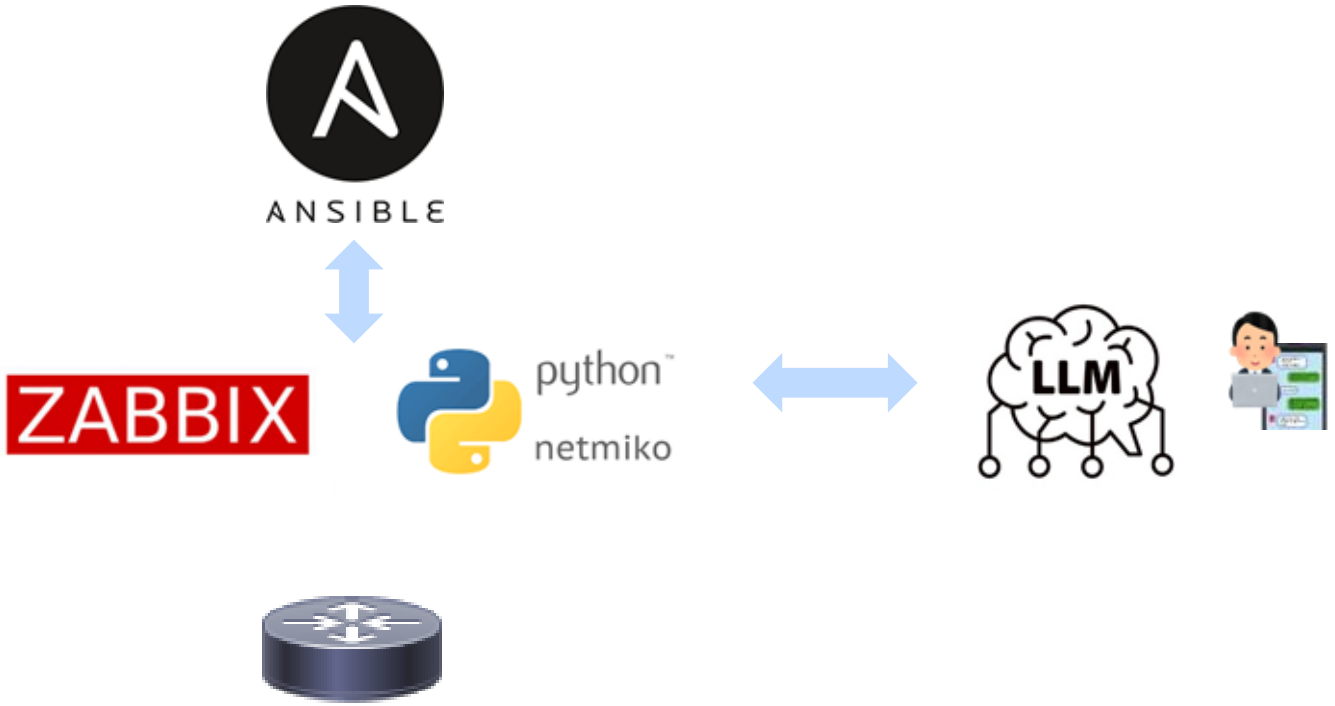


手順書補完・精緻化



取り組みの背景・目的

- 大規模かつ成長し続けるNWを限りあるリソースで保守するため, 様々なケースのNW自動化に取り組んでいます.
- LLM等AI技術に大きな期待をしており, 現在の**Network Automation(AIOps)をより高度に拡張しオペレータの業務を全面的に任せられるレベル**にすることを目的として検討を進めています.

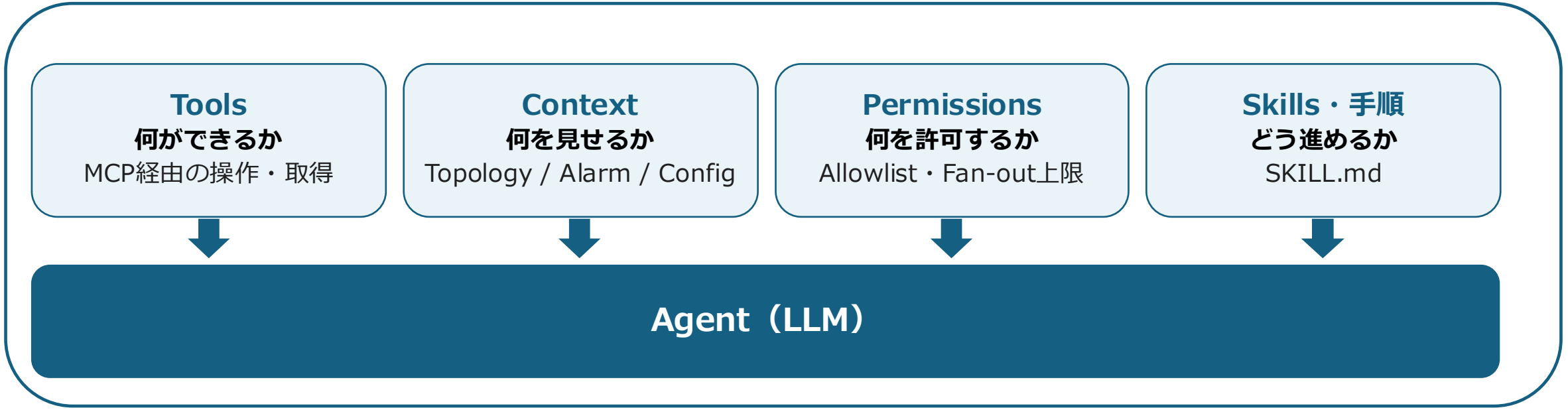


AIOpsに関連する技術動向と発表概要

- Agentic AIOpsへの期待が高まり、Agent SkillsやMCP AppsなどAgentの能力と所掌範囲を拡張する技術が急速に発展している
- プロンプトチューニングなど個別要素の改善だけでは精度・維持管理性の面で厳しくなっており、**Agentに何を与え、何を判断させ、誰と協働させるか**という実行環境全体の設計（Agent Harness）が重要になっている
- NW運用におけるAI活用（AIOps）に適したハーネスの事例を共有し、知見・課題を議論させていただきたい

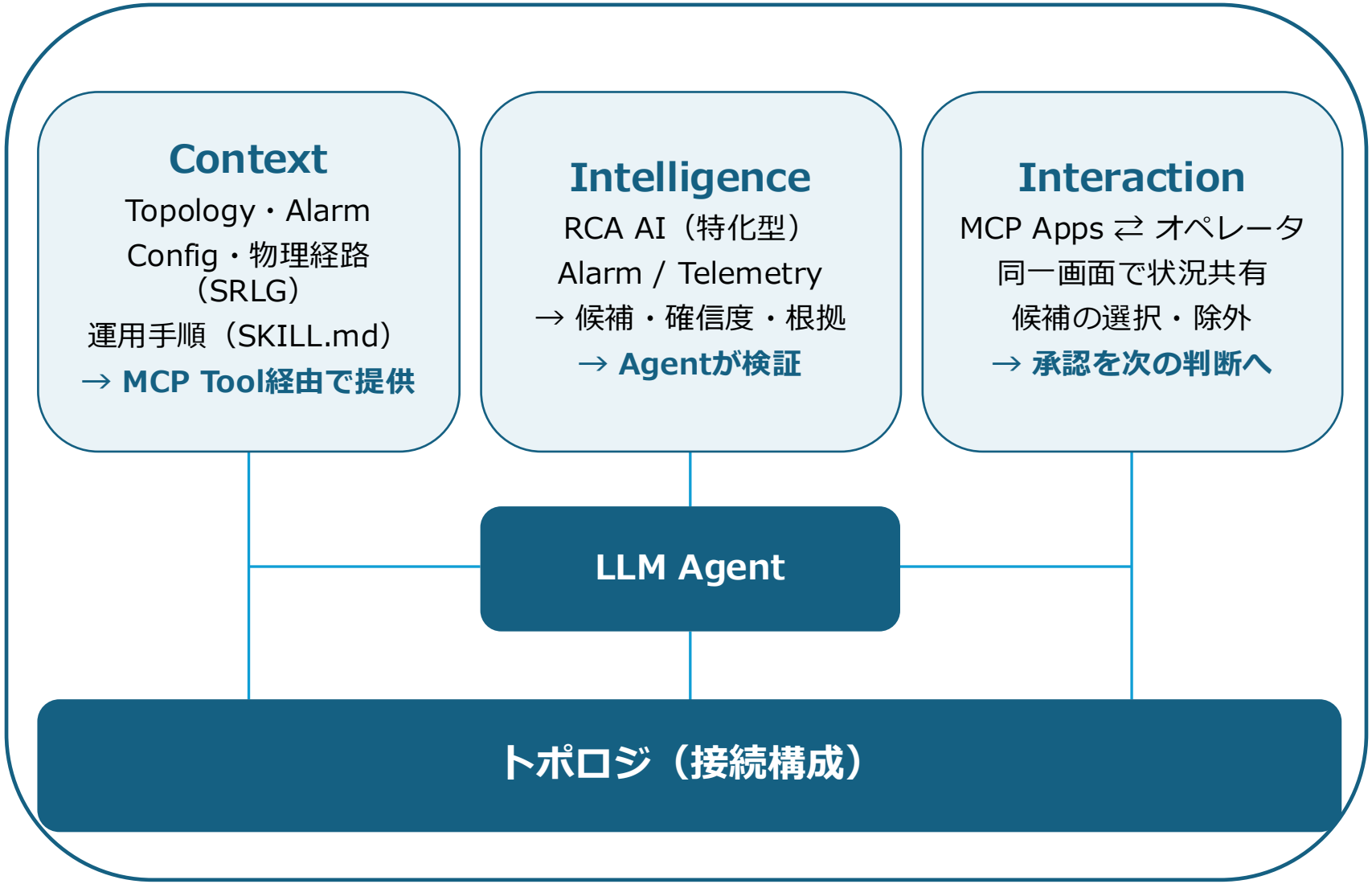
Agent Harnessとは

- Anthropicなどが提唱している、**Agentを動かす実行環境**（Tool・Context管理・権限制御）
- 何をAgentに見せるか・何を判断させるか・何を許可するかを規定するもの
- 同じAgentでも、**Harnessの設計次第で精度や信頼性は大きく変わる**



AIOpsにおけるHarness構成要素

- NW運用の判断は、接続構成（トポロジ）が基盤
- トポロジを1本の軸に、Context・Intelligence・Interactionの3本柱を組む

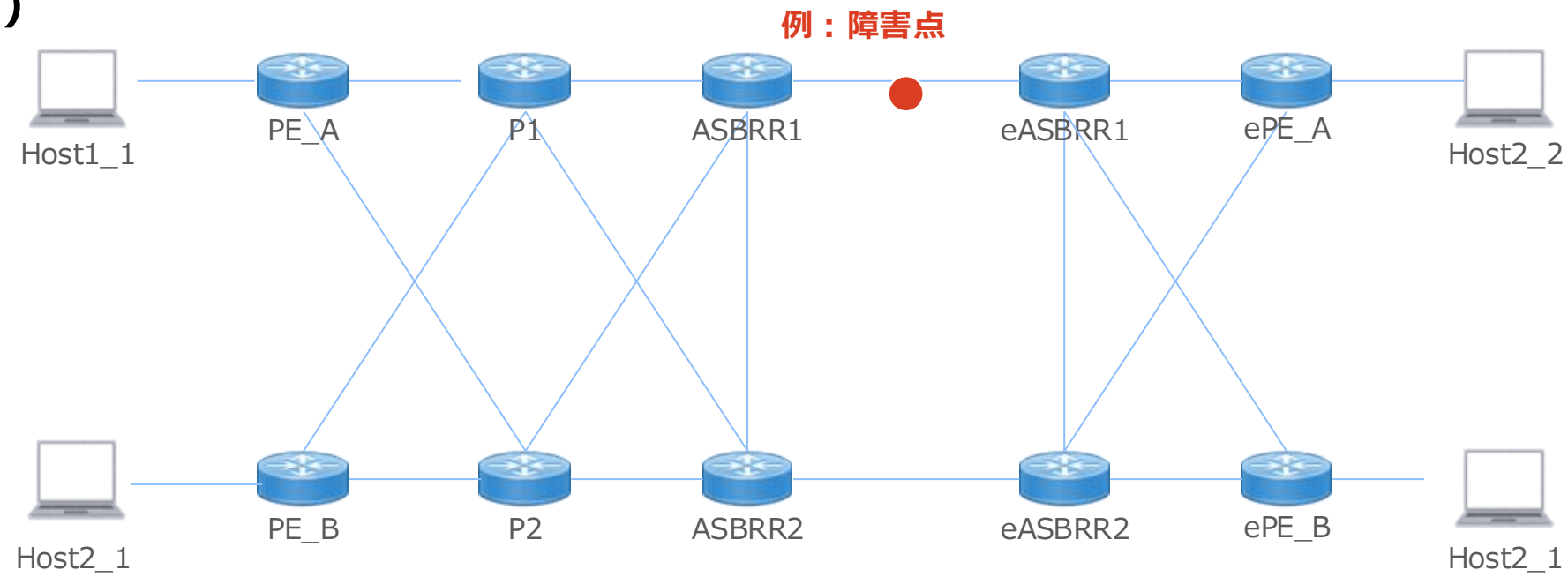


Context：何を与えるか

- Agentが参照する情報は、Topology・Alarm・Config・cable-routes.csv（物理経路）・SKILL.md（運用手順）
  - いずれもMCP Tool経由で取得させ、Agentに直接ファイルを触らせない
- 論理トポロジだけでは見えない共通障害点が、**物理経路情報**を突き合わせることで見えてくる

Intelligence：RCA AIへの推論委譲

- 全ての推論をLLMに任せず、NW障害に特化したRCA AIへ一部の推論を委譲することで、判断制度やトークン消費効率を向上させる
- RCA AIはAlarm・Topologyなどから原因候補・確信度・根拠を提示し、LLM Agentがそれを検証・裏付けする
- RCA AIの推定は答えとしてそのまま採用せず、**検証すべき仮説として扱う（Agentがその後裏取りをする）**



① 観測データ
Alarm／Topology

→

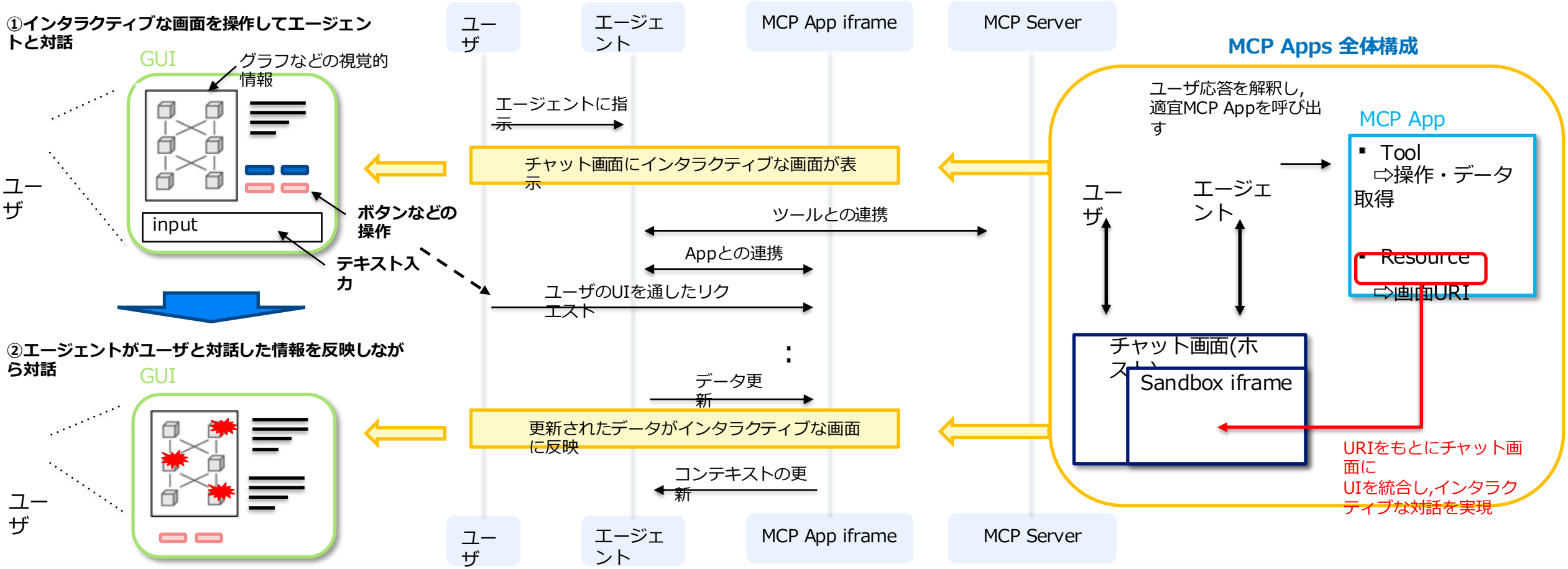
② RCA AI が推定
原因候補・確信度・根拠を提示

→

③ LLM Agent が検証
仮説として裏付け → 結論

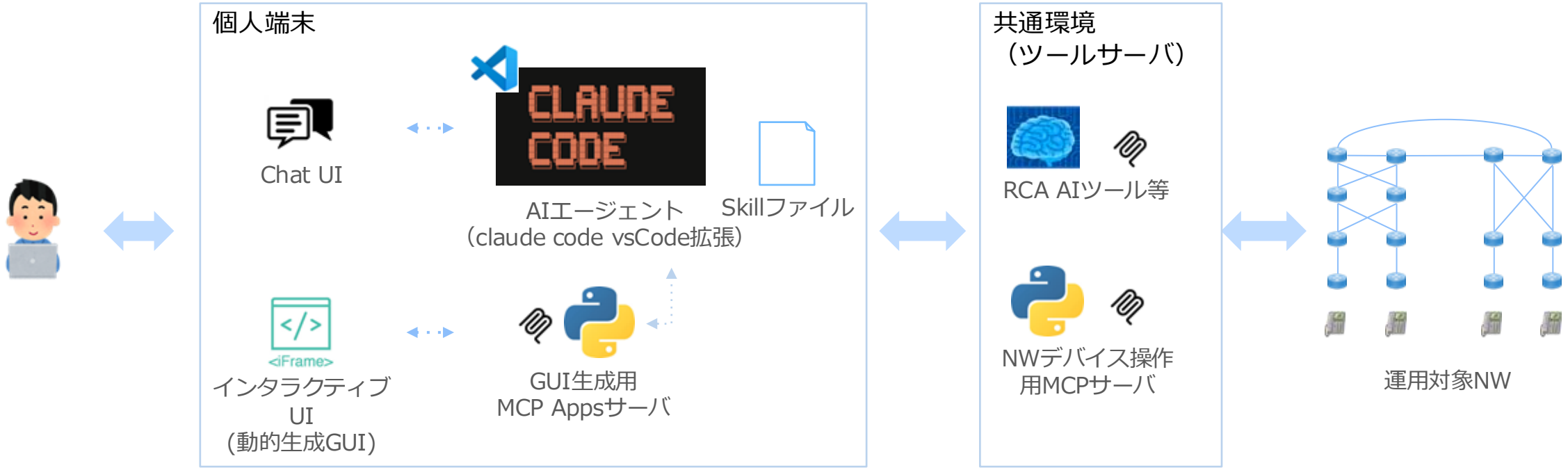
Interaction：ChatからOperational Workspaceへ

- Topology・Alarm時系列・RCA AIの候補・Agentの調査根拠を**同じ画面で共有し**
  オペレータのGUI操作（候補の選択・除外、追加調査や対応案の承認）をAgentの次の判断へ反映
- Chatに打ち込むだけの利用から、同じ状況を見て判断する**Operational Workspace**へ



ハーネスの実装構成

- 個人端末にAgent（Claude Code）・Skillファイル・MCP Appsサーバを配置
- 共通環境にNW操作用MCPサーバとRCA AI等のツールを配置
- Skillは個人・チームのオペレーション作法に合わせて最適化。MCP Appsのコードはセッション中にAgentが修正できるため、UI仕様も状況に応じて最適化できる



比較する3構成と評価軸

構成	Agentに与えるもの	観測された傾向
LLM-only	NW操作Tool のみ	調査手順が定まらず発散しやすい
LLM + Skill	＋ 障害切り分け手順（SKILL.md）	手順に沿って正解に到達。Tool Call数は多い
LLM + Skill + RCA AI	＋ RCA AIの原因候補・確信度・根拠	case4は大幅に効率化／case9は誤診（過信）

評価軸：精度 ／ Tool Call数 ／ Token ／ 処理時間 ／ 説明の妥当性
シナリオ：case4 = 既知パターンの単一障害／case9 = 未知パターンの複合障害

実証：Permissionsはpromptでは強制されない

- SKILL.mdの文面任せでは実際には強制されておらず、**任意コマンドを無検証で実行できる実装ギャップ**が見つかった
- 是正：Allowlist化／設定変更コマンドの分離／大規模NW向けのFan-out上限を実装
  - いずれもprompt文面ではなく、harnessの機構として強制する形に変更した
- 示唆：**権限はpromptで頼むものではなく、harnessで担保するもの**

実証：Agent Skillsは「渡すだけ」では機能しない

- 既存のfailure-simスキルは、**明示的なトリガー文言がないと自発的に呼ばれなかった**
- 一般的な障害申告の文面でも自発起動する新しいSkillを作って改善した
  - ただし起動しても誤診した回があり、「起動する」ことと「正しく診断する」ことは別問題
- 示唆：Skillは置くだけでなく、**description設計と起動条件までharness側で作り込む**必要がある

実証：case4の効果とcase9の過信リスク

case4：既知パターンの単一障害

- **RCA AIありの構成が正解に到達**
- Tool Call数・処理時間を大幅に削減
- 既知パターンでは推論委譲の効率化効果がはっきり出た

→ ハーネスとして「効く」ことを確認

case9：未知パターンの複合障害

- **RCA AIが既知パターンに引っ張られて誤診**
- Agentも誤った候補に沿って調査を進めた
- 専門AIへの推論委譲は過信リスクと表裏一体

→ 委譲先の出力を疑う仕組みが必要

示唆：推論委譲は「効率化」と「過信」の両面を持つ。RCA AIの出力は答えではなく仮説として扱い、Agent側で反証させる手順までharnessに組み込む必要がある

まとめ

- NW運用向けのAgent Harnessとは、**トポロジという単一の構造を軸にContext・Intelligence・Interactionを組む設計問題**である
- Skillも権限も「渡す／書く」だけでは機能せず、harness側の機構として作り込む必要があった
- 専門AIへの推論委譲は効率化に効くが、既知パターンへの引き込みという過信リスクと表裏一体