R&Dフェーズ｜研究　**実用化時期**｜FY28-

Mehr Licht!
More Photonics for Intelligence

#A05

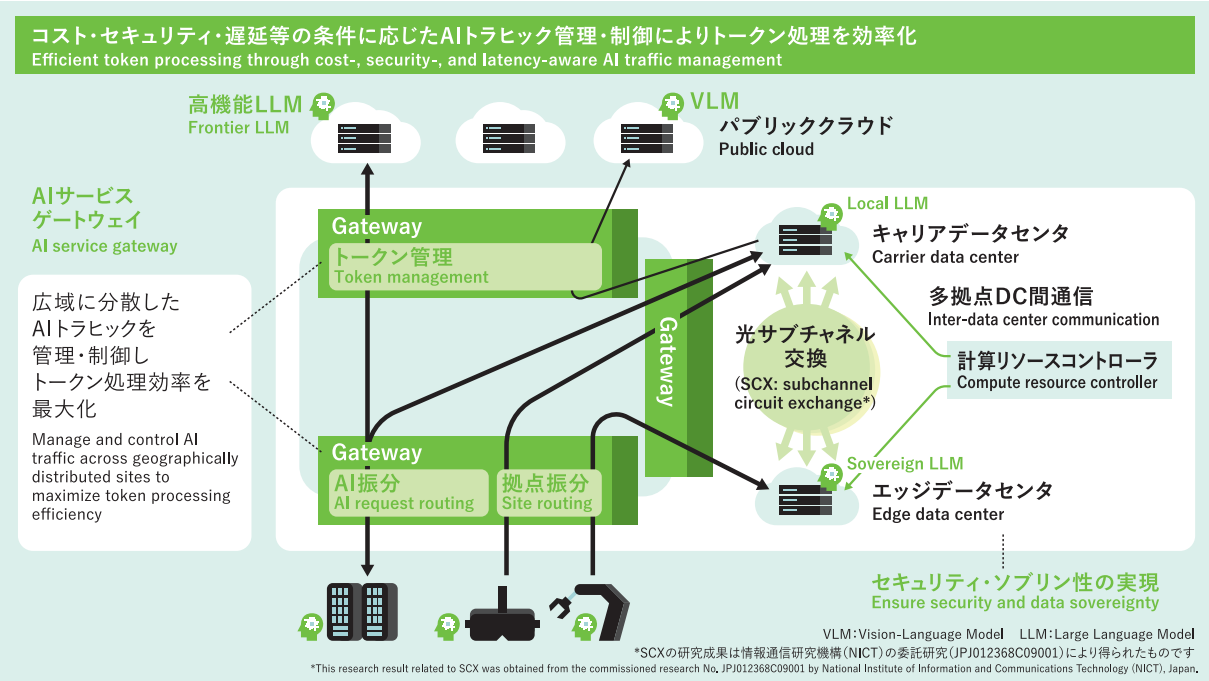
#AI・光量子基盤#DX・AX（デジタル変革）

AIトラヒック制御で効率化する分散インフラ

AIトラヒック制御によって、トークン処理効率を最大化します

技術課題・ゴール

AI需要の急増に伴い、計算資源・電力の有効活用が課題となっています。トークン効率を最大化する事により、計算資源や電力を効率化しつつ、応答速度などのユーザ体感を向上します。



技術ポイント

要素技術

・分散した計算資源をリソース稼働状況に応じて使い分け、計算資源の利用効率を最大化
・分散した計算資源を効率的・柔軟に接続する多地点間確定通信

市中技術差異点

・従来の市中技術と比較して、分散した計算資源上でのLLMトークン処理性能を約40%、タスク完遂率を約20%向上
・光波長パスのリソースを論理分割して制御することで、多地点間の柔軟な接続と確定通信品質、通信リソースの高効率利用を両立

適用ビジネス/ユースケース

情報技術

AIエージェントを活用するソフトウェア開発において、GPUの使用率に応じて処理をふりわけることにより、開発品質を維持しつつ、トークン処理性能を向上します。

出展組織

NTT株式会社 ネットワークサービスシステム研究所
株式会社NTTドコモ
NTTドコモビジネス株式会社

社外連携先

F5 Networks.Inc
KDDI株式会社（国プロ NICT受託研究 課題90）
楽天モバイル株式会社（国プロ NICT受託研究 課題90）
日本電気株式会社（国プロ NICT受託研究 課題90）
1FINITY株式会社（国プロ NICT受託研究 課題90）

お問い合わせ：ネットワークアーキテクチャプロジェクト