

## 第1部

### 特集1 AIとWIDE

浅井 大史

---

---

#### 第1章 はじめに

---

---

生成AIの登場により、さまざまな領域でのAI技術の応用が始まっている。特に、大規模言語モデル(LLM: Large Language Model)は、自然言語のインターフェイスを持つため、これまでデジタル化の進まなかった領域までその応用が期待され始めている。

WIDE Projectでは、2024年7月29日から31日にかけて開催した夏のボード主催研究会において、AI技術の最新動向について議論をした。また、秋の研究会のボードブレナリにおいて、AIとWIDEと題して夏のボード主催研究会の内容を報告するとともに、WIDEメンバーと今後WIDEで取り組む研究課題について議論した。

本稿では、夏のボード主催研究会の報告を中心に、WIDEで取り組むべき研究課題について報告する。

---

---

#### 第2章 夏のボード主催研究会の報告

---

---

2024年7月29日から31日の3日間にわたり夏のボード主催研究会を開催した。研究会のアジェンダは、以下のとおりであった。

- Introduction to Foundation Models of AI
- Operations and Management of AI Cloud (guest speaker's talk & discussion)
- Data Center Networking (guest speaker's talk & discussion)
- The Internet and Data Governance
- Discussion on AI and Internet as a social infrastructure

---

#### 2.1 Introduction to Foundation Models of AI

---

まず、最初のIntroduction to Foundation Models of AIでは、AI技術に関する基礎知識・前提の共有が行われた。中でも注目すべきこととして、AI基盤モデル(Foundation Model)の急速な発展を説明するスケーリング則[1]についての情報共有があった。スケーリング則では、モデルサイズ(パラメータ数)、学習に利用したデータセットのサイズ(トークン数)、学習に要した計算資源量(FLOPS-days)の3つの要素に従い、AIモデル性能が向上する(スケールする)という法則である。昨今の大規模なGPU基盤の整備とインターネット・Web上の多量のデータがこれらを支えている。また、Transformer[2]の登場によりGPUの計算資源を効率的に利用することができるようになったこともスケーリング則を支える要因のひとつである。

本研究会では、大規模学習に加えて、追加学習や推論についての動向についても報告があった。特に推論については、モデルの大規模化に対応するために、パラメータの量子化に4ビット等の低精度浮動小数点や整数が用いられ始めている。これにより、モデルの利用するメモリ量を減らすことができ、大規模モデルを分散して格納するのに必要なGPU数を減らすことができる。また、一般にGPUは低精度浮動小数点や整数の方が高い演算性能を有する(例えば、単精度浮動小数点の演算性能は、倍精度浮動小数点の演算性能の倍程度となることが一般である。)ことや、データサイズの削減によりメモリ帯域幅の利用量も削減されるため、推論のスループットを高めることができる。低精度浮動小数点や整数を用いた場合のモデルの性能については課題があるが、性能を下げないような研究が進められており、低精度浮動小数点の利用は今後も広がることが予想されている。

このような低精度浮動小数点や整数の利用が広がってきているものの、モデルの大規模化も同時並行的に進んでいるため、AI計算基盤においてスケーラビリティがますます重要となっている。

## 2.2 Operations and Management of AI Cloud

Operations and Management of AI Cloudのセッションでは、ソフトバンクの川上氏をゲストスピーカーとして招待し、AIクラウドの設計から運用までの課題について議論した。大規模学習に用いるAIクラウド(AIデータセンター)の議論が中心であったが、今後AI技術がインターネットサービスの一部として普及していく中で、地域分散アーキテクチャのあり方も変わってくる可能性についての指摘もあり、継続的に集中と分散アーキテクチャに関する研究を進める必要があることを確認した。特に、AI計算基盤は学習・推論のどちらにおいても、大量の電力を消費するため、ネットワークだけでなく、電力・冷却を同時に検討する必要がある。

また、大規模学習や追加学習においては、データを保管するストレージについても必要不可欠な要素であり、ストレージとAI計算基盤間の接続も重要である。

## 2.3 Data Center Networking

Data Center Networkingのセッションでは、Arrcusの海老澤氏をゲストスピーカーとして招き、AIデータセンターで利用されるネットワーク技術について、プロトコルからハードウェアまでを網羅的に議論した。まずは、大規模学習で用いられる分散計算技術として、Data Parallelism、Tensor Parallelism、Pipeline Parallelismの3種類の技術およびこれらの分散計算技術で用いられる集合通信(Collective Communication)についての解説があった。また、この集合通信をEthernet上で行う技術として、RDMA over Converged Ethernet version 2 (RoCEv2)の解説があった。図1にRoCEとRoCEv2の packets フォーマットを示す。ここに示したとおり、RoCEはEthernet上に実装されていたが、RoCEv2ではIPヘッダとUDPヘッダのオーバーヘッドは増えるものの、スケーラビリティや運用性が向上するため、UDP上にInfiniBandプロトコルをカプセル化する形で実装されるようになった。

RoCEv2は、通常はノード間でパケットロスのない lossless ネットワークを想定しており、ロスレスを実現するための技術としてPriority Flow Control (PFC)が使われている。図2にPFCの概念図を示す。送信側および受信側

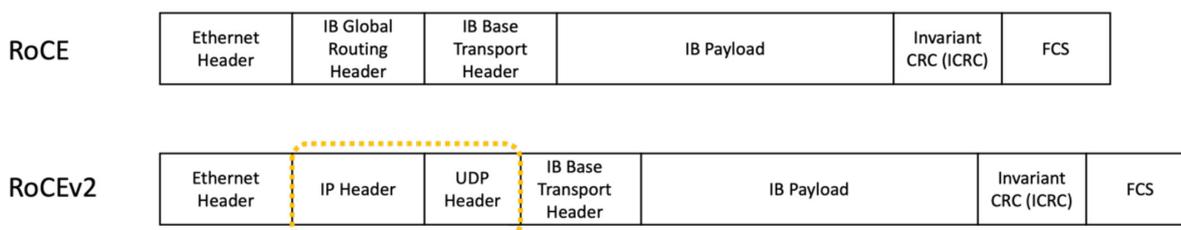


図1 RoCEとRoCEv2のパケットフォーマット

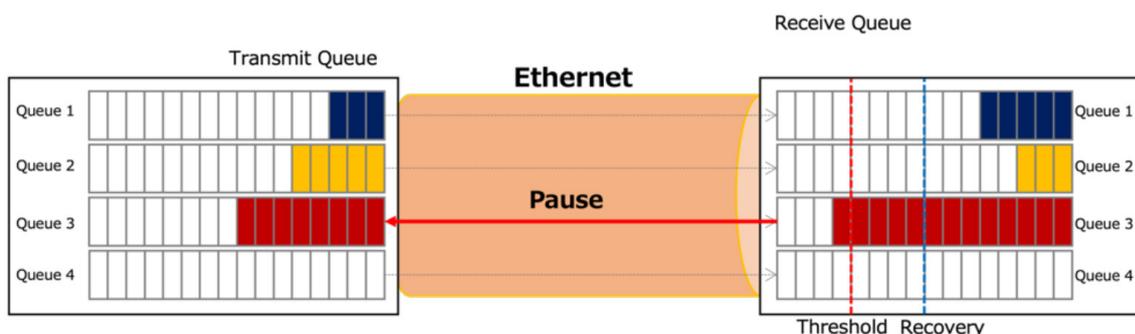


図2 Priority Flow Control

のネットワーク機器でPriorityごとにキューを持っており、受信側のキューが一定の量を超えた場合に送信停止を指示するパケットを送信することで、バッファが溢れてパケットロスが発生することを防ぐ仕組みである。なお、Priorityという用語ではあるものの優先度を表してはならず、キューの番号と扱ってよい。これに加えて、100Gbpsやそれを超えるような超広帯域の輻輳制御としてExplicit Congestion Notification (ECN)を組み合わせたData Center Quantized Congestion Notification (DCQCN)が採用されている。Losslessネットワークを実現するためには、スイッチのバッファ管理が重要となる。一方で、PFCを用いてlosslessネットワークを実現するには十分なメモリが必要であるが、パケットの伝搬遅延と帯域幅遅延積からデータセンター内であっても高帯域ネットワークで多ポートのスイッチにおいては十分なメモリを搭載できないこともあることから、ポート間でのメモリ共有によるオーバーサブスクリプションが利用されており、パラメータのチューニングが重要となる旨の説明があった。これらの技術自体は、従来のインターネットQoSでも議論されており、技術的に新しいものではないが、多ポートかつ高帯域ネットワークでの応用においてその重要性をあらためて確認することとなった。

また、大規模学習の集合通信において最も利用されているAll-Reduceの処理に最適化したネットワークトポロジーの研究として、Rail Optimized Topologyの紹介があった。従来のSpine-LeafやCLOSトポロジーとは大きく異なり、All-Reduceにおいて最大の実効帯域を得るためにはノード内のGPUやNVLink等のスケールアップネットワーク、PCIeスイッチのトポロジーを考慮した設計が必要となる。そのため、マルチテナンシーやさまざまなワークロード・分散学習において、Rail Optimized Topologyが最適化どうかは詳細に検討し、評価する必要がある。

さらに、RoCEv2で課題となるスケーラビリティやオーバーサブスクリプション、ロードバランスの課題を解決するためにEthernetをベースとしてHigh Performance Computing (HPC)やAI・機械学習向けのインターコネクタ技術の標準化を目指しているUltra Ethernet Consortium (UEC)について紹介があった。UECのスコープには、In-Network Computing (集合通信の計算をネットワーク機

器で行う機能)なども含まれており、トラフィックエンジニアリングだけでなく、計算リソースの管理も必要になってくることが予想される。UECについては、現段階はメンバー企業内のみで議論が行われているが、今後は仕様が公開され、対応製品が登場することが予想されるため、運用性や相互接続性を含め、継続的に動向を追う必要がある。

## **2.4 The Internet and Governance**

The Internet and Data Governanceのセッションでは、本報告書で別途説明したIAB Workshop on AI Controlの開催案内やEUのAI Act等の法規制を中心に、AI時代にインターネット基盤で考えるべきガバナンスや技術について議論した。AI技術が急速に発展し、利用が進む中で、これらについては結論を出すことが難しいが、自律分散システムとしてのインターネットから得られた知見をもとに検討を続けることとした。

## **2.5 Discussion on AI and Internet as a social infrastructure**

Discussion on AI and Internet as a social infrastructureのセッションでは、直前のセッションのインターネットとガバナンスを含め、AIがデジタル基盤に取り組みられる中で、インターネットとAI基盤が社会基盤として担う役割と課題について議論した。特に、インターネットを取り巻くさまざまな研究開発に関係してきたWIDE Projectがその知見を活かして取り組むべき課題について議論した。

まずは、大きく分類してインターネットのためのAIとAIのためのインターネットという2つの視点で議論した。しかし、その2つの分類で基盤を考えるのは困難であったため、WIDE Projectとして取り組む課題や戦略に関連するトピック・キーワードを出し合った。CopyrightやIntellectual property、robots.txtなど、AI学習に関係するキーワードだけでなく、AccountabilityやExplainability、Responsibilityなど、AI技術の運用に関するものや、NeutralityやInteroperabilityなど、インターネットが持続可能な技術であるための特徴など、合計51個のキーワードが出された。これ以外に、AI計算基盤の要素技術であるメモリ技術やネットワーク技術、冷却技術などのキーワードも挙がった。

これらについて、本研究会では時間の制約で具体的な議論ができなかったため、短期的なもの、中長期的なもの2つに分類して、WIDEとしてAIに関する戦略を検討することとした。

---

### 第3章 WIDE合宿Board Plenary報告

---

秋のWIDE合宿のBoard Plenaryでは、夏のボード主催研究会の報告をするとともに、Discussion on AI and Internet as a social infrastructureのセッションで挙げたキーワードをもとに、WIDE Projectの取り組むべき課題について議論した。しかし、AI技術は日進月歩であり、その応用領域も急速に広がっていることから、robots.txtのような特定の課題はあり、それらについて取り組む必要はあるものの、WIDE Projectとして中長期的な戦略を定めることが困難であることがわかった。そのことから、実際にAI技術に触れながら、また運用を通じて知見を集積させ、課題解決に取り組むことという方向性が示された。その方針にもとづき、WIDE Projectのメンバーが関わるAI関係のプロジェクトや研究テーマについて情報共有を行った。今後もこれらのプロジェクトや研究について、研究会等でその情報共有を継続し、課題発見や解決に取り組むことで、AIを含めたデジタル社会基盤の研究開発に取り組む予定である。

---

### 第4章 まとめ

---

本稿では、夏のボード主催研究会を中心にWIDE ProjectのAI技術・基盤への戦略に関する議論について報告した。急速に発展するAI技術に対しては戦略と言える結論には至らなかったが、WIDE Projectの強みである運用を通じた課題発見と解決を行うための土台となるプロジェクトにはWIDE Projectメンバーが取り組んでおり、研究会等におけるそれらのプロジェクト間での交流やWIDE Projectメンバーのプロジェクト参加により、WIDE ProjectとしてAI基盤に貢献していく方針である。