

# 主 論 文 要 旨

No.1

報告番号	甲乙第号	氏名	松谷健史
<p>主論文題目： IP(InternetProtocol)を用いた低遅延分散アーキテクチャ</p>			
<p>(内容の要旨)</p> <p>クラウドの普及と Ajax や WebSocket を利用した画面の自動更新に対応するサービスの普及によって、アプリケーションの高速化が求められるようになり、データセンタ内においてネットワークスループットの向上だけでなく、低遅延通信が要求されている。例えば、memcached(Facebook や Twitter で利用されている)などの分散メモリキャッシュサーバや、iSCSI などのストレージサービスでは、高速化のために低遅延通信が求められている。</p> <p>しかし、データセンタで使われている IP ネットワークは接続性に優れているが、InfiniBand や独自の専用ネットワークと比べて通信遅延が大きく、これらのサービスの処理性能を低下させている。低遅延通信には、RDMA が利用されるが、構造的に DMA リードによるオーバーヘッドが生じる。</p> <p>本研究では、DMA リードを廃することで、RDMA より通信遅延を少なくする IP-NUMA を提案する。IP-NUMA は、バークレイソケット API に代わり、ノード間の一部のメモリを NUMA アーキテクチャによって共有メモリ化し、データ送信時に DMA リードに代わって CPU ライトを使い、PC 上で生じる通信遅延を削減する手法である。IP-NUMANIC を用いたメモリ書き込みによる通信(PC+FPGA、Linux 上で構築)と、市販 NIC を用いたバークレイソケットによる通信を ping-pong プログラムを使って分析したところ、通信遅延を 90%削減した。片道通信遅延は、1.081<math>\mu</math>s で、RDMA を用いた場合と比べて、11%以上削減することができた。</p> <p>また、L3 スイッチの IP パケット転送に必要な処理と最小転送遅延をあきらかにした。FIBNIC は、パケット転送遅延を削減するため、本研究では全工程を最少ステージ数でパイプライン化し、パケットバッファを用いない IP パケット転送手法である。本手法を市販の FPGA ボードに実装したところ、41 万経路の IPv4 パケットの転送遅延が 976ns で一定であることが確認された。これは市販のギガビット L3 スイッチにて 64 バイトフレームを計測したときの 20~25%に相当する。</p> <p>本手法である IP-NUMA と FIBNIC を memcached に適用し、見積もりしたところ、ローカルの IP ネットワークにおいて、バークレイソケットと比べ、4 倍のトランザクション性能となった。コモディティネットワークを使用しながら、クラウドや HPC などのインターコネクトとして IP プロトコルの適用範囲が広がることが期待できる。</p> <p>キーワード： データセンタ,低遅延通信,RDMA、IP,memcached,FPGA.</p>			