



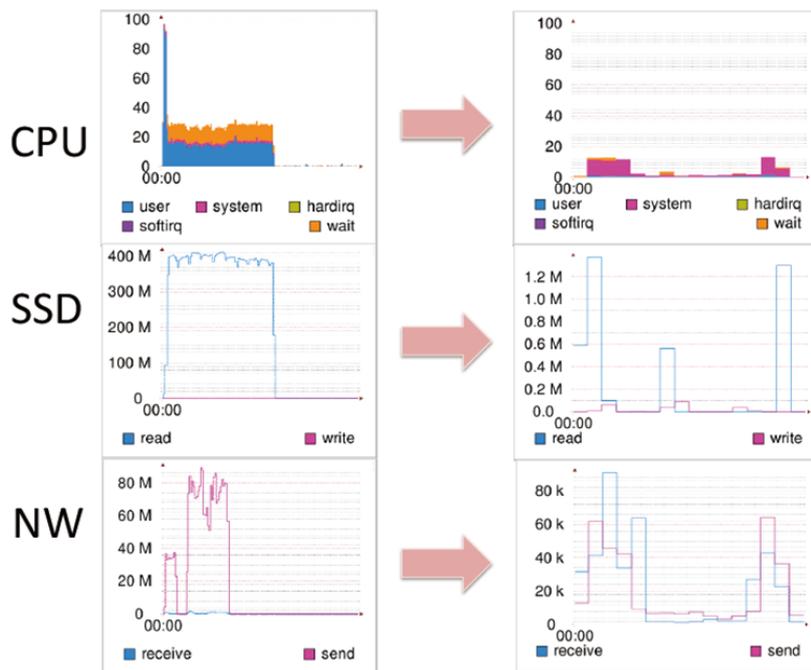
テーマ名	ビッグデータの高速処理を小型かつ安価な構成で達成する密結合型ハードウェアと並列処理を組合せた計算機システム
組織名	国立大学法人 電気通信大学 情報システム学研究所 吉永 務 教授
技術分野	IT

## 概要

ビッグデータの高速処理を実現するために、ストレージ、光通信ネットワーク、FPGA、SSD等を密接に結合させたハードウェアと高効率の並列処理を組合せ、小型かつ安価なシステム構成でありながら、Hadoop、Impalaなど他の分散ソフトウェアと比べて10倍以上の高速化を実現する計算機システムを開発しています。ビッグデータの高速処理を希望する企業、ビッグデータ処理基盤を開発している企業、データベース処理について課題を抱えている企業など、連携に意欲がある企業を歓迎いたします。

## 簡略図

ビッグデータの高速処理を小型かつ安価な構成で達成する密結合型ハードウェアと並列処理を組合せた計算機システム



Impala(137.6sec)

Aval(13.5sec)

## 計算機資源の大幅な効率化

小型、安価なシステムでありながらHadoop、Impalaなどの他の分散ソフトウェアよりも10倍以上の高速処理を実現



## 背景

世の中で流通するデータ量が膨大に増え、ビッグデータ化が進展しています。そして、ビッグデータを統計的に分析して特徴を抽出し、ビジネスへ生かすデータサイエンティストの活躍の場も増えています。

このような傾向につれ、ビッグデータを高速に処理するデータベース処理基盤が普及しています。Hadoop、Impalaなどは代表的な手法ですが、ビッグデータ化につれて更に高速にデータを処理するニーズが高まっています。しかしそのためにデータベース処理基盤が大型化、高コスト化してしまえばビジネスメリットが薄れてしまいます。

本研究では、ストレージ、光通信ネットワーク、FPGA、SSD等を密接に結合させたハードウェアと高効率の並列処理アルゴリズムを組合せ、小型かつ安価なシステム構成でありながら、Hadoop、Impalaなど他の分散ソフトウェアと比べて10倍以上の高速化を実現します。

ビッグデータの高速処理を希望する企業、本システムの製品化に興味がある企業、ビッグデータ処理基盤を開発しており、本研究の成果を活用したい企業、データベース処理について課題を抱えている企業など、連携に意欲がある企業を歓迎いたします。

## 技術内容

ハードウェア構成は以下の通りです。

□システム内部の通信ネットワークは、「高速インターコネクト」として、

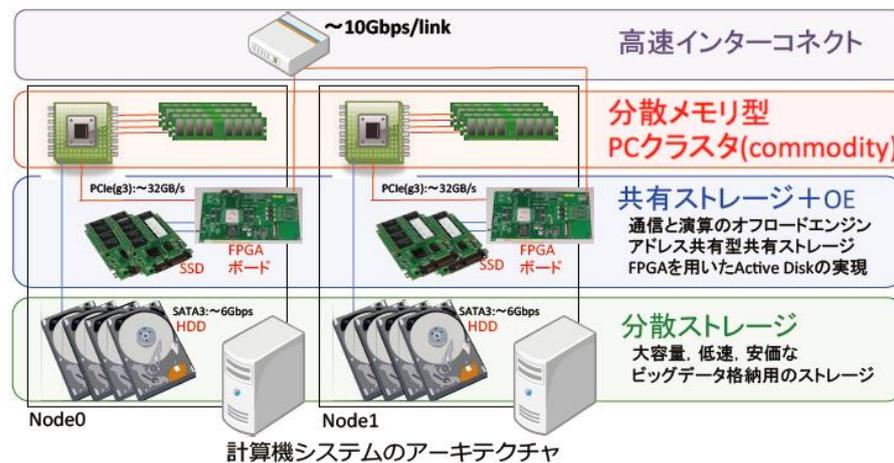
10Gbpsまでの通信速度を持つ光通信を使用しています。

□ビッグデータを複数の計算機で効率良く並列処理するためのハードウェアとアルゴリズムを搭載した「分散メモリ型PCクラスタ」を開発しています。

□データの中継と一次保存を行うための「共有ストレージ+OE」としてFPGAボードとSSDが組み込まれています。

本システムではここで通信と演算の負荷を徹底的に効率化していることが特徴です。

□データを保管するための「分散ストレージ」は大容量かつ低速、安価な一般のストレージです。

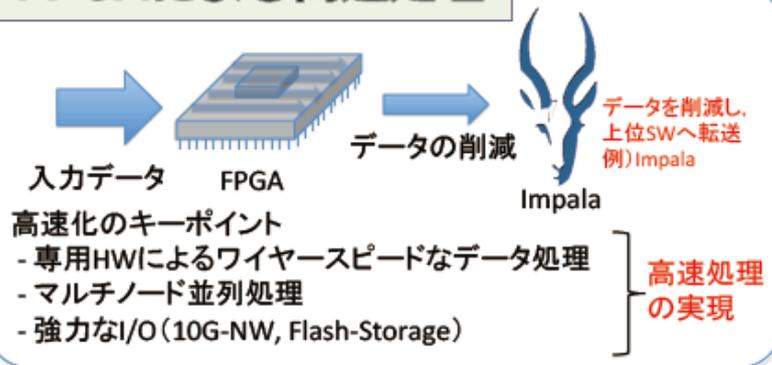




前記のとおり、FPGA ボードを用いた高速処理に特徴があります。

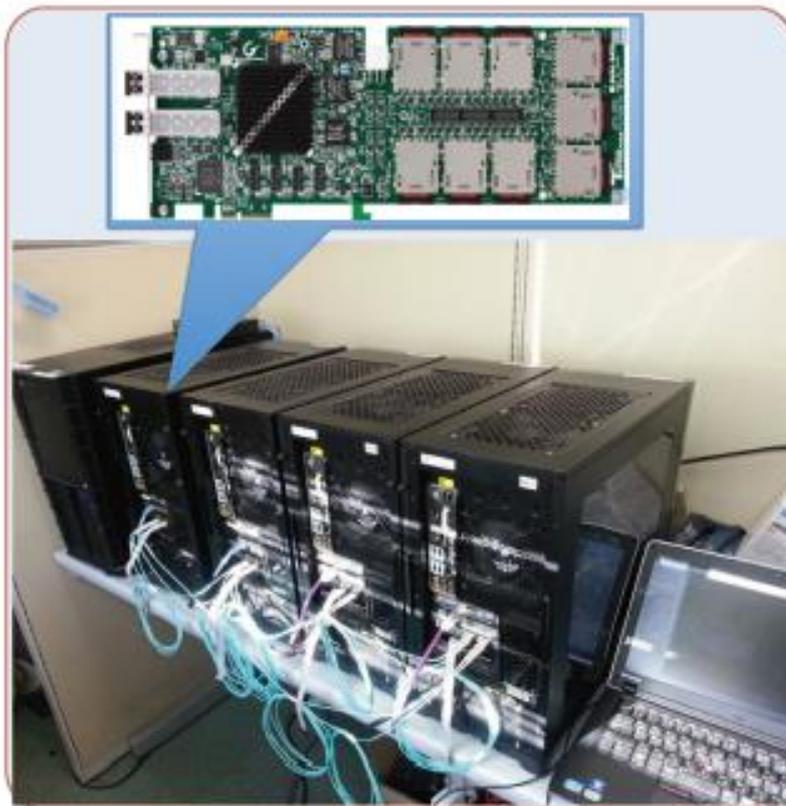
1. ビッグデータ処理に特化した専用のハードウェアにより、ワイヤースピード（理論上の最高通信速度）によるデータ処理
2. データを複数の計算機で効率良く並列処理する機能（マルチノード並列処理）、
3. 光通信ネットワーク（10G-NW）に対応し、中間メモリであるフラッシュストレージ（SSD）を効果的に活用するインタフェース（I/O）機能を持っています。

## FPGAによる高速処理



システムは下記の通り小型です。

FPGA ボードが組み込まれています。





### 技術・ノウハウの強み(新規性、優位性、有用性)

下記が性能評価結果です。

4words では、本システム (Aval) では 9.15 秒であり、Hadoop (738 秒) の約 80 倍、Impala (136 秒) の約 15 倍です。

256word では、本システム (Aval) では 13.5 秒であり、Hadoop (5798 秒) の約 430 倍、Impala (137.6 秒) の約 10 倍の処理速度です。

### Simple Word Countingの性能評価



### 分散ソフトウェアと比べて 10倍以上の高速処理

計算機システムの効率低下は主に下記の「ネットワーク」「待機時間」によって起こります。それに対する本システムの特徴は以下の通りです。

#### 1. ネットワーク

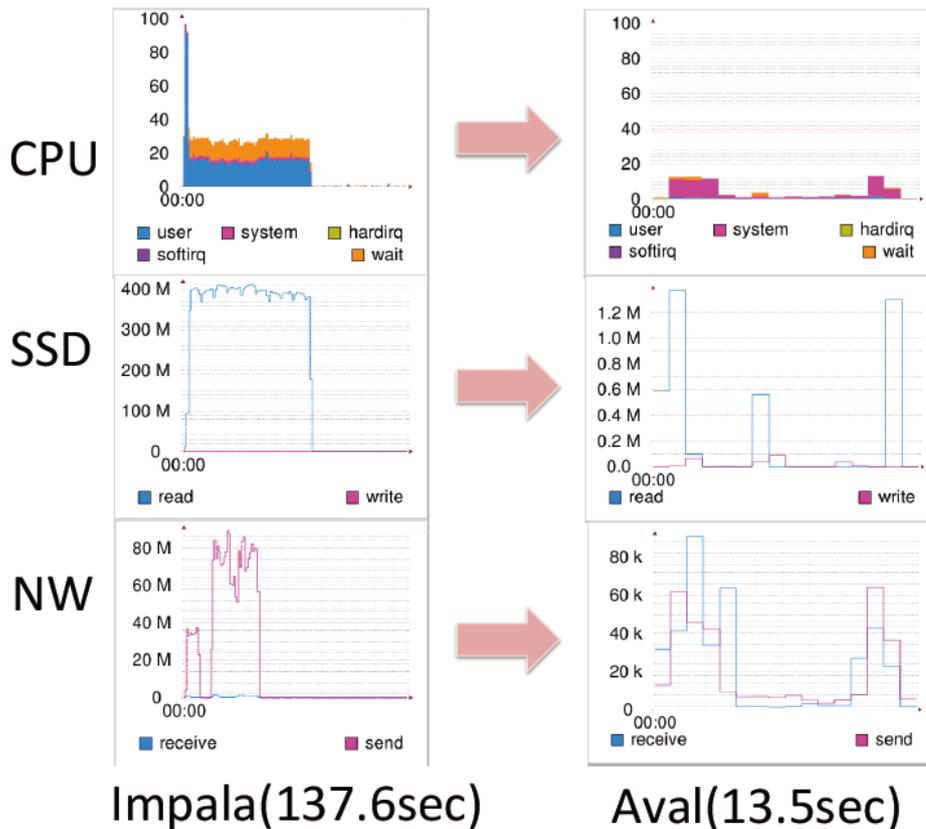
- 計算機システム内では短距離であっても大量のデータが行き来します。ネットワークが遅い場合には遅延となり効率低下に繋がります。
- 本システムではネットワーク内のデータ伝送を光通信で行い、インタフェース部も光通信へ対応させることにより、高速化しています。

#### 2. 待機時間

- 中継点でデータを一時的に保存するメモリには、保存可能なデータ容量に限りがあり、そのデータ容量を超えるデータ量が届いても保存できず、待機時間が発生し、遅延による効率低下に繋がります。
- 本システムでは専用の FPGA と接続した SSD と、効率的な並列処理アルゴリズムを搭載することにより、高効率化しています。



下記のとおり、システムを構成する各コンポーネントである CPU、SSD、ネットワーク等の計算機資源を大幅に効率化しています。



## 計算機資源の大幅な効率化

### 連携企業のイメージ

1. ビッグデータの高速処理を希望する企業
2. 本システムの製品化に興味がある企業
3. ビッグデータ処理基盤を開発しており、本研究の成果を活用したい企業
4. データベース処理について課題を抱えている企業

など、連携に意欲がある企業を歓迎いたします。

データベース、並列処理、クラウドコンピューティング等のキーワードに関連して様々な研究を行っておりますので、上記に限らず技術指導や社内人材の育成も含めてお気軽にご相談ください。

### 技術・ノウハウの活用シーン(イメージ)

ビッグデータを高速に処理する必要がある場面で活用可能です。

特に、ビッグデータを高速に処理したいが、システムを小型で安価にしたい、とのニーズに最適です。



### 技術・ノウハウの活用の流れ

本研究にご興味があればお気軽にお問合せください。  
システムのデモ紹介や技術の詳細なご説明など対応させていただきます。

### 専門用語の解説

#### 【ビッグデータ】

大容量のデジタルデータを指します。また、非定型でかつリアルタイム性が高いデータを扱うケースが多いことも特徴です。近年、ビッグデータを高速かつ簡単に分析できる技術が登場し、ビッグデータを活用すれば、これまで予想できなかった新たなパターンやルールを発見できることが明らかとなったことから注目されています。

#### 【Hadoop、Impala】

Hadoop は大規模データの分散処理を支える Java ソフトウェアフレームワークであり、フリーソフトウェアとして配布されています。アプリケーションが数千ノードおよびペタバイト級のデータを処理することを可能としています。

Impala は大規模並列処理（MPP）エンジンであり、Hadoop のデータに対して SQL を用いてクエリを実行できるソフトウェアです。

#### 【並列処理】

コンピュータにおいて複数のプロセッサで1つのタスクを動作させることです。並列コンピューティングや並列計算とも呼ばれます。問題を解く過程はより小さなタスクに分割できることが多い、という事実を利用して処理効率の向上を図る手法です。

大規模なタスクを複数の計算機で並列処理するためには、各計算機に仕事を最適に割り当てるスケジューリングが非常に重要であり、割り当ての効率が悪いと遅延が発生してしまいます。

吉永教授は並列コンピューティングについて長年研究しており、そのノウハウを本システムでは生かしています。

#### 【密結合】

密結合とは、細分化された個々のコンポーネント同士が密接に結びついている状態を指します。逆に、コンポーネント間の結びつきが弱いシステムを疎結合と呼びます。

密結合状態のシステムでは、緊密で高速な動作が可能となる利点がありますが、システム内で固有のインターフェースに基づいて接続されているため、一方が他方を容易に取り替えられない弱点もあります。

本システムでは、ビッグデータのデータ処理に目的を特化し、システム全体の光通信ネットワーク化や専用 FPGA 等を使用した密結合ハードウェアと並列処理を組み合わせることにより、高速処理・小型・安価な計算機システムを実現しています。



**【SSD（フラッシュストレージ）】**

記憶媒体としてフラッシュメモリを用いるドライブ装置であり、ハードディスクドライブ(HDD)と同じ接続インタフェースを備え、ハードディスクの代替として利用できます。SSDはハードディスクのようにディスクを持たないため、読み取り装置(ヘッド)をディスク上で移動させる時間(シークタイム)や、目的のデータがヘッド位置まで回転してくるまでの待ち時間(サーチタイム)がなく、高速に読み書きできます。

また、HDDを備えたストレージに比べて、消費電力量が少ないことです。フラッシュストレージを使えば、家庭用の大型テレビ1台分の消費電力量で、20テラバイト(TB)のデータを保存できるという試算もあります。

**お問い合わせ先**

下記から御問合せください。

[http://www.open-innovation-portal.com/university/it/bigdata\\_command.html](http://www.open-innovation-portal.com/university/it/bigdata_command.html)