

## 2016 年度大規模計算論授業概要

### 2016 High Performance Computing Class Overview

松岡 聡 Satoshi Matsuoka

2016 年 9 月 23 日(金) Friday, Sep 23, 2016

本年度は、ハイパフォーマンスコンピューティング(HPC)、並列処理、クラスタ計算、クラウド、の分野において、エクサスケールに向かうスパコンや大規模 IDC にとって非常に重要な「ビッグデータ・AI の HPC 的手法による性能モデル化・大幅な性能やスケールの向上」の学術的な成果の近年の論文をサーベイする。

This year's class will deal with "Immense acceleration of Big Data and AI through convergence with HPC"

- 基本的に、各自がハイパフォーマンスコンピューティング(HPC)、並列処理、クラスタ計算、IDC、ビッグデータ、クラウド、の分野において、「ビッグデータ・AI の HPC 的手法による性能モデル化・大幅な性能やスケールの向上」にまつわる特定のテーマの英語の論文を 3 編選ぶ。そのうち最低一本は当方から指定できる。それらを発表の 2 週間前に、下記のメイリングリストに出典とアブストラクトを mail する。すると、松岡はそのうち一本を選択するので、それを自分の回に発表する。  
A student is designated per class to present a paper regarding the theme "immense acceleration of Big Data and AI through convergence with HPC". The designated student will pick 3 papers (written in English), and email their abstracts 2 classes prior to the assigned class. At least one of the papers may be selected by the instructor. We will pick one to present out of the 3, and the specified paper will be presented in detail in front of the class.
- ビッグデータと HPC の融合が近年様々な形で模索されている。これは HPC にまつわるデータ処理をビッグデータ技術を用いて大規模化・広帯域化・高速化するだけでなく、ビッグデータに固有な処理を HPC 技術を用いて高速化・大規模化する課題もある。  
Convergence of HPC and Big Data are being explored in various ways. This includes acceleration of large-scale data in HPC, but also utilizing HPC technologies to accelerate big data processing such as graph processing and deep learning.
  - テーマの例 Example Themes:
    - ◇ Architectures for accelerating the handling of extremely large-scale data in HPC and AI.
    - ◇ Performance Modeling of Big Data and AI processing such as Deep Learning
    - ◇ Large Scale Parallel frameworks for Pre- and Post- processing of large-scale simulations using Big Data and AI methods
    - ◇ Utilization of new non-volatile storage and memory for large-scale Big Data and AI scaling
    - ◇ Acceleration of Large-scale Graph Algorithms on Supercomputing Platforms
    - ◇ Accelerating Deep Learning using GPUs and other HPC accelerators
    - ◇ Large-scale parallel filesystems and their algorithms
    - ◇ Utilization of Many-core processors such as GPUs and Xeon Phi for big data and AI
    - ◇ HPC-class networks for Big Data and AI.
  - We will not deal with small-scale "big data", pure HPC, or purely theoretical big data issues such as advanced theory of deep learning.
- The students should search for the appropriate papers in conferences and journals that deal with HPC and big data such as ACM/IEEE Supercomputing, ACM ICS,

ISC (International Supercomputing Conference), IEEE Big Data, ACM HPDC, IEEE IPDPS ([www.ipdps.org](http://www.ipdps.org)), ICS, Cluster Computing, CCGrid([www.ccgird.org](http://www.ccgird.org)) USENIX, IEEE Transactions on Big Data, as well as Big Data and Deep Learning conferences such as NIPS and ICML. They are easily searchable using Google Scholar and ACM/IEEE digital library portals.

- Such conferences often have associated workshops that specifically deal with convergence issues, so search widely.
- Papers that are too commercial or deal with non-innovative technologies, just stating the state-of-practice, should be avoided.
- The paper must have been published in the last 6 years, e.g., including 2011 and after. Also, papers that have been selected prior by other students cannot be selected. This implies that you should present as early as possible. You cannot present your own paper or a paper from your group.
- The selected abstracts and papers should be discussed and announced in the mailing list as indicated below.
- Although one paper is to be presented in a principal fashion, all three selected papers should be read and introduced as related topics. References to all papers should be indicated.
- Each class will involve presentation by one speaker; there will be many questions that will be asked by the instructor, the TA, and the student audience. Questions will be asked to audiences as well from the instructor and the TA, so full attention should be paid during the class.
- The grades will be marked in the following fashion
  - Presentation: 50%
  - Report: 50%
  - For each class, if a student voluntarily asks a question or states a technical opinion, additional 5% credit will be added
- Presentation will be done online using Powerpoint. A PDF version should be prepared and sent to the mailing list [hpc16@smg.is.titech.ac.jp](mailto:hpc16@smg.is.titech.ac.jp) A DAY PRIOR TO THE CLASS (Sunday) so that the student audience can refer to it during the presentation. The presentation will also be archived at the following page.
  - <http://matsu-www.is.titech.ac.jp/lecture/lecture-wiki/index.php?hpc2016>
- The report should summarize the general topic covering and including ALL THREE PAPERS regarding the state of the art in HPC and Big Data convergence