

# 筑波大学 計算科学研究センター スーパーコンピュータ COMA (PACS-IX) について

平成 26 年 2 月 筑波大学計算科学研究センター

## 1. はじめに

筑波大学計算科学研究センターは平成 26 年 4 月より、メニーコアプロセッサ技術を駆使した最先端スーパーコンピュータ COMA (PACS-IX)の稼働を開始します。同システムは最先端のメニーコアプロセッサである Intel Xeon Phi プロセッサ (通称名 MIC: Many Integrated Core) を演算加速装置部として搭載した、先進的大規模 PC クラスタです。また、同システムは筑波大学で 30 年以上に渡って研究開発を続けている PACS/PAX シリーズの第 9 世代に当たり、マシン名 COMA の他にシリーズコード PACS-IX が与えられています。通常の利用では COMA としてご参照下さい。

## 2. システム概要

COMA は、汎用 CPU 部として Intel Xeon E5-2670v2 (2.5GHz)を 2 基、さらに演算加速装置部として Intel Xeon Phi 7110P (61 core)を 2 基搭載した高性能計算ノードを合計 393 台結合した超並列 PC クラスタです。汎用 CPU 部のみを用いた一般的な OpenMP+MPI のようなプログラミングも可能であり、また先進的ユーザは Intel Xeon Phi を用いた演算オフローディングや汎用 CPU 部との平行利用によるヘテロジニアスプログラミング、あるいは Intel Xeon Phi のみを単独 Linux 環境で使用する native mode 利用も可能です。実際には、これらの運用を同時に行うために、それぞれのジョブクラス用の専用キューにジョブを投げることによって使い分けます。

相互結合網は InfiniBand FDR (ピーク性能 7GB/s)をフルバイセクションバンド幅を持つ Fat Tree ネットワークで、バッチジョブシステムによって割り当てられたあらゆる計算ノード間で均質な並列処理が実行可能です。また、Lustre+RAID6 による大規模高性能共有ファイルシステム (ユーザ空間約 1.5PB) が提供され、全ての計算ノードから自由に参照可能です。

COMA 上では以下の 3 つのクラスで計算ノードの並列実行をサポートします。

### (1) 汎用 CPU パーティション

各計算ノードの汎用 CPU (20 core) のうち、16 core を割り当て、それらを任意のノード数の MPI 通信で利用する。ノード内では OpenMP、ノード間では MPI によるハイブリッ

ド並列が可能。状況により、MIC パーティションのジョブと計算ノードを共有する（ジョブミックス実行）。

(2) MIC パーティション

各計算ノードの 2 台の MIC と汎用 CPU 部の 4 core を割り当て、それらを任意のノード数の MPI 通信で利用する。汎用 CPU core からは Intel compiler が提供する offload 機能を用い、適宜演算の一部を MIC 上でオフロード実行可能で、ノード間では MPI による並列化を行う。状況により、MIC パーティションのジョブと計算ノードを共有する（ジョブミックス実行）。

(3) 混合パーティション

各計算ノード上の全ての計算リソース（汎用 CPU 部 20 core+MIC 2 基）を占有可能、それらを任意のノード数の MPI 通信で利用する。MIC の native mode での利用や汎用 CPU とのヘテロジニアス計算等、先進的な利用が可能。本パーティションではジョブ実行中、計算ノードをノード単位で占有するため、他のジョブとのノード共有は発生しない。

COMA の主な諸元は以下の通りです。

Computation node	CPU	Intel E5-2670v2 (IvyBridge-EP) 2.5GHz x2
	# of cores	20 (10 cores / CPU)
	MIC	Intel Xeon Phi 7110P 61 core x2
	Main memory	64 GB (DDR3 1866MHz x 8 channel, 119.4GB/s)
	MIC memory	16 GB (8GB/MIC, 352GB/sMIC)
	Peak performance	400 GFLOPS (CPU) + 2147 GFLOPS (MIC)
	Network HCA	InfiniBand FDR
	Peak network b/w	7 GB/s /direction
Number of nodes		393
Interconnection configuration		Fat-Tree with full bisection b/w
Peak performance		1.001 PFLOPS (CPU: 157 TFLOPS, MIC:844 TFLOPS)
Network bisection b/w		2.75 TB/s
Shared file system		Lustre file system
File system capacity		1.5 PB (user space)
Operating System		Red Hat Enterprise Linux (login node), CentOS Linux (compute node)
Job Scheduler		SLURM

Programming Language	Fortran/C/C++ with Intel Composer/XE, OpenMP, MPI
----------------------	---

### 3. 運用

COMA は平成 26 年 3 月末に導入を完了し、同年 4 月 15 日より運用開始の予定です。システムは筑波大学計算科学研究センター学際共同利用、HPCI、計算基礎科学連携拠点学際共同利用、大規模有償一般利用等に供されます。