

超並列メニーコアクラスター Oakforest-PACS について

1. はじめに

Oakforest-PACS は筑波大学計算科学研究センター及び東京大学情報基盤センターが共同運営する「最先端共同 HPC 基盤施設」(JCAHPC: Joint Center for Advanced High Performance Computing) が管理運用する超並列メニーコアクラスターである。同システムは 2016 年 12 月から行われている試験運用期間を経て、2017 年 4 月から公式運用が開始される。各計算ノードに最先端の Intel Xeon Phi 7250P (メニーコア型プロセッサ) を搭載しており、極めて高い演算性能及びメモリバンド幅を提供する。なお、Oakforest-PACS は簡単のため OFP と略されることもある。

2. Oakforest-PACS システムの概要

Oakforest-PACS のハードウェア諸元について以下にまとめる。

項目	仕様
ベースクラスター部計算ノード構成	
CPU	Intel Xeon Phi 7250P (68 core) x 1 台
CPU コア数	68
CPU ピーク性能	3 TFLOPS
CPU メモリ	超高速メモリ (MCDRAM) 16GB 高速メモリ (DDR4) 96GB
並列処理ネットワーク	
ネットワークシステム	Intel Omni Path Architecture
ノード当たりバンド幅	100Gbps
トポロジ	Fat-Tree (full bisection bandwidth)
全体システム構成	
総ノード数	8208
ファイルシステム	26PB のハードディスクドライブベースの Lustre file system、及び 960TB の SSD ベースのバーストバッファ (高速ファイルキャッシュ)
ファイルシステムへのアクセス	全ノード上のプロセスから任意にアクセス可能、ただしバーストバッファの利用についてはステージングを行う
ログインノード	8 台

Oakforest-PACS へのアクセスはログインノードへのログインを通じて行う。フロントエンドノードは `ofp.jcahpc.jp` の代表ホスト名が与えられており、負荷に応じて自動的に 8 台のうち 1 台にログインサービスが振り分けられる。ジョブ投入及びファイルサーバへのアクセス機能については 8 台のログインノードは均等である。外部からのファイル転送はログインノードに対する `scp` コマンドを通じてのアクセスのみが許される。

計算ノードに対するジョブ実行は富士通 PJS によるバッチキュー・システムを通じて行う。各計算ノードへの直接ログインは許されていない。計算に必要なファイル及びジョブスクリプト群を予めファイルサーバに置き、ジョブを投入する。ファイルサーバはフロントエンドノード及び全計算ノードから透過にアクセス可能である。計算結果は各計算ノードによってファイルサーバに出力されるようプログラムする。

Oakforest-PACS のファイルシステムは全計算ノード及びログインノードから均質にアクセス可能な大規模共有クラスタファイルシステムであり、一般的なハードディスクドライブを用いた 26PB の Lustre ファイルシステムと、特に高速なファイルアクセスを要求するアプリケーションを想定した超高速ファイルキャッシュシステム（バーストバッファ、DDN IME) 960TB が備えられている。後者は前者のキャッシュファイルとなっているが、アクセスするにはジョブスケジューラと連動したファイルのステージイン及びステージアウト処理が必要となる。

Oakforest-PACS の運用状況等については JCAHPC 内の下記 URL を参照のこと。

http://jcahpc.jp/ofp/ofp_intro.html

Oakforest-PACS システムの構成及び利用方法の詳細については、「Oakforest-PACS 利用の手引き」（プロジェクト採択後にアクセス方法を通知）を参照のこと。