

SCore 7 の概要

PC Cluster Consortium 開発部会
Allinea Software
堀 敦史

2007年9月14日

第7回 PC クラスタシンポジウム

1/11

SCore 7 で変わるもの

- ・ 新しいネットワークへの対応
 - Myrinet 10G
 - InfiniBand
- ・ マルチコアでの性能向上
- ・ MPICH-2 のサポート

2007年9月14日

第7回 PC クラスタシンポジウム

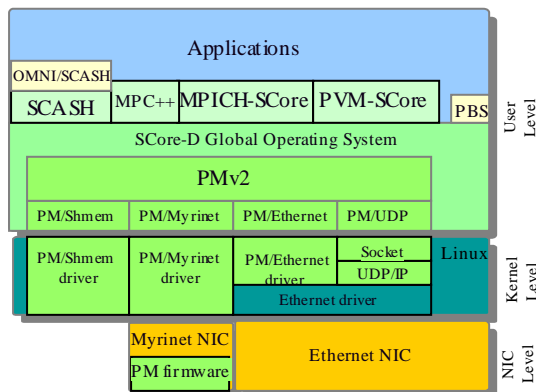
2/11

SCore 7 で変わらないもの

- Ethernet, Shmem のサポート
- ギャングスケジューリング
- チェックポイント(パリティ)
- 並列ジョブ

⇒外観上(使い勝手)ほとんど変化なし
⇒しかし、内部はほぼスクラップ&ビルド

SCore 6 の根本的問題点



- SCoreパッケージ
 - 密接に関係
 - 分割不可
 - 独自API
 - 全て独自開発
- 問題点
 - 外部通信ライブラリの取込困難
 - 新たな通信ハードへの対応はかなりの労力が必要

SCore 6 の問題点

- SCore はクラスタオペレーティングシステム
 - SCore の開発がクラスタ用 OS の研究から
 - 特殊 (当初はユニーク) な設計思想、API
 - ユーザレベル通信、かつ OS 機能を実現する
- PMv2 の性能上の問題点
 - マルチコアでの性能低下が顕著
 - PMv2 の再設計が必要

2007年9月14日

第7回 PC クラスタシンポジウム

5/11

PMv2 の問題点

- プロセス間で共有されるエンドポイント
 - SCore-D ではネットワークは共有資源
 - SCore-D がエンドポイントを生成し、ユーザプロセスに与える
 - = > MX や Infiniband ではエンドポイントはプロセス間共有が許されていない！！
 - = > かなりのハックが必要
 - = > 開発が長期になる (約1年！)

2007年9月14日

第7回 PC クラスタシンポジウム

6/11

SCore 7 チャレンジ

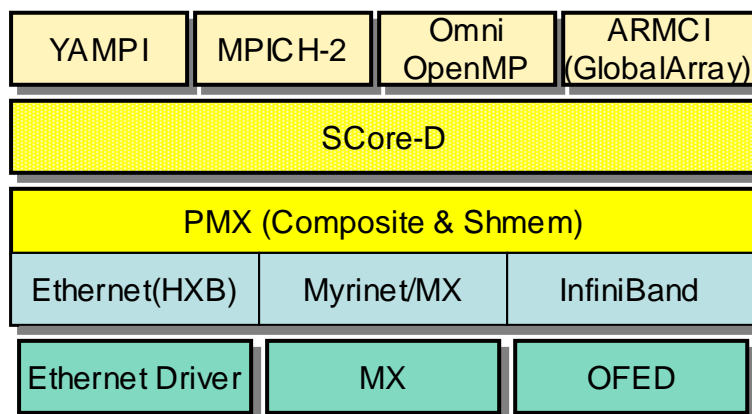
- 新しい SCore 通信ライブラリ *PMX*
 - 第3者開発の通信ライブラリと親和性を高める
 - MX - Myrinet
 - OpenFabrics - InfiniBand (iWarp ?)
 - マルチコアでの性能向上
 - PMv2 の利点は残す
 - マルチネットワークのサポート
 - 複合ネットワークのサポート
 - ギャングスケジューリングやチェックポイントも可能

2007年9月14日

第7回 PC クラスタシンポジウム

7/11

SCore 7 ソフトウェア階層



ARMCI: Aggregated Remote Memory Copy Interface
OFED: OpenFabrics Enterprise Distribution

2007年9月14日

第7回 PC クラスタシンポジウム

8/11

SCore 7 : PMX

- 高速化
 - Composite と Shmem の一体化
 - ロック機構の見直し(マルチコアの性能向上)
- 外部通信ライブラリの取込み
 - Myrinet(10G): MX 通信ライブラリ
 - InfiniBand: OFED
 - 他のネットワークへの対応が容易になる
- マルチネットワーク、複合ネットワーク機能は従来通り

2007年9月14日

第7回 PC クラスタシンポジウム

9/11

SCore 7 : SCore-D

- オペレーティングシステム的な色合いを薄く
 - MPICH-2 のプロセスマネージャに近い
 - PMX の管理はしない(資源ではない)
 - プロセスの生成と管理に徹する
 - 標準入出力の管理
 - 並列ジョブ
 - ギャングスケジューリング
 - チェックポイントの起動
 - デバッガとのインターフェイス、等々

2007年9月14日

第7回 PC クラスタシンポジウム

10/11

SCore 7 開発状況(11月末)

- PMX (SC07@Reno でデモ)
 - API 設計ほぼ終了
 - Composite(Shmem) 動作確認済み
 - Ethernet 基本動作確認済み
 - Ethernet-HXB 開発中
 - MX (Myrinet) 基本動作確認済み
 - InfiniBand 基本動作確認済み
- MPICH-2/SCore
 - 基本動作確認済み

2007年9月14日

第7回 PC クラスタシンポジウム

11/11

リリーススケジュール

2007/12 MPICH2 までの動作確認

今後

- 仕様細部の詰め
- ゼロコピー通信の実装
- 性能チューニング
- チェックポイント
- ギャングスケジューリング

2008/03 SCore 7 リリース

2007年9月14日

第7回 PC クラスタシンポジウム

12/11