

SCore 7 の概要

PC Cluster Consortium 開発部会
Alinea Software
堀 敦史

2007年9月14日

PC クラスタワークショップ in 九州

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

2007年9月14日

PC クラスタワークショップ in 九州

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

⇒実質的にほとんど変化なし

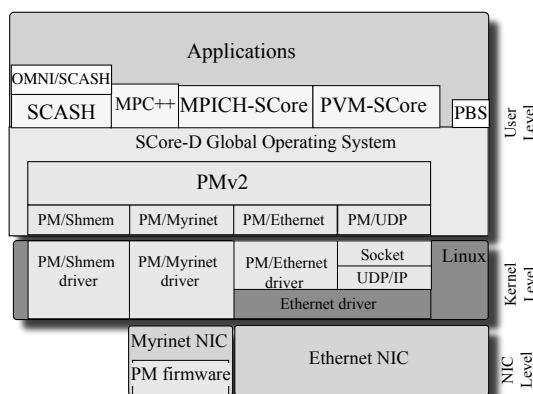
⇒しかし、内部的にはスクラップ&ビルド

2007年9月14日

PC クラスタワークショップ in 九州

3/11

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

2007年9月14日

PC クラスタワークショップ in 九州

4/11



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

PC クラスタワークショップ in 九州

5/11



SCore 7 技術的チャレンジ

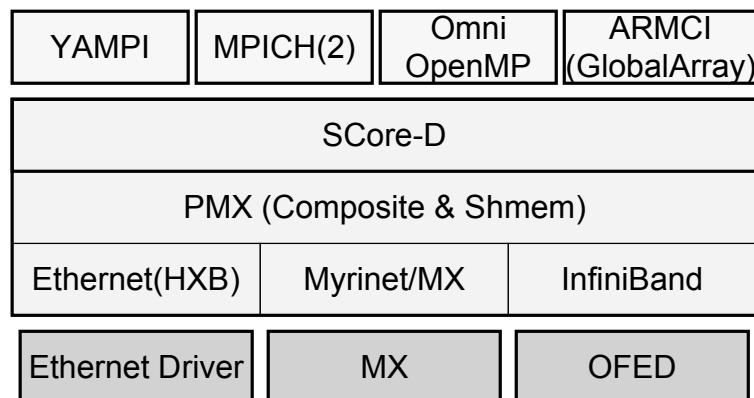
- 新しい SCore 通信ライブラリ *PMX*
 - 第3者開発の通信ライブラリと親和性を高める
 - MX - Myrinet
 - OpenFabrics - InfiniBand
 - マルチコアでの性能向上
 - PMv2 の利点は残す
 - マルチネットワークのサポート
 - 複合ネットワークのサポート
 - ギャングスケジューリングやチェックポイントも可能

2007年9月14日

PC クラスタワークショップ in 九州

6/11

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

PC クラスタワークショップ in 九州

7/11

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

PC クラスタワークショップ in 九州

8/11



SCore 7 : SCore-D

- オペレーティングシステムの的な色合いを薄く
 - MPICH-2 のプロセスマネージャに近い
 - PMX の管理はしない
 - 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

2007年9月14日

PC クラスタワークショップ in 九州

9/11



SCore 7 開発状況

- PMX
 - API 設計ほぼ終了
 - PMX/Composite(Shmem) 動作確認済み
 - PMX/Ethernet デバッグ中
 - PMX/MX 基本動作確認済み
 - PMX/InfiniBand コーディング開始
- SCore-D
 - コーディング開始

2007年9月14日

PC クラスタワークショップ in 九州

10/11



リリーススケジュール

- 2007/11 基本機能
ギャングスケジューリングなし
- 2007/12 ベータリリース
チェックポイントなし
- 2008/03 SCore 7 リリース

2007年9月14日

PC クラスタワークショップ in 九州

11/11