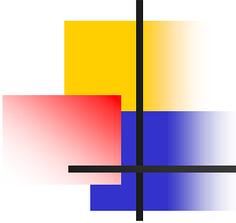


SCore と eScience の install と 設定について

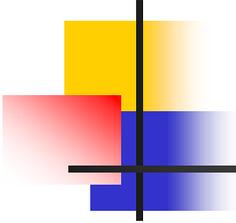
理化学研究所
亀山豊久



はじめに

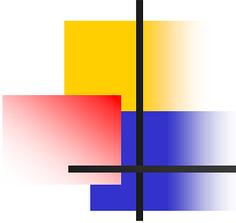
eScience プロジェクトの成果物は SCore パッケージに統合して配布することになった。SCore をインストールすることによって eScience のすべての成果物が容易に使用できるようになる。

ここでは、SCore を使用して eScience の成果物をインストール及び設定する方法について説明する。



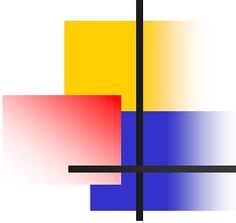
SCore の概要

- PC クラスタのための高性能並列プログラミング環境
 - 通信ライブラリ PMX
 - クラスタオペレーティングシステム SCore-D
 - SCore-D 上で動く mpi (mpich2/SCore)



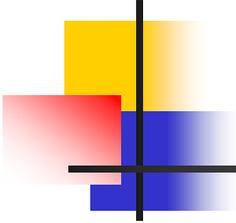
SCore の build システム smake

- SCore を build するコマンド
- 異なる環境でソースファイル、インストールディレクトリとも同一の directory を共有できる
- make の wrapper



SCore のコンパイルドライバ (1)

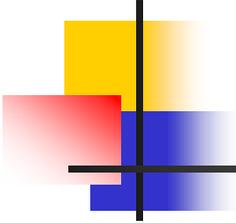
- SCore プログラムをコンパイルするためのプログラム
 - コンパイルドライバと設定ファイル、設定ファイル確認ユーティリティからなる
 - scorecc, scorec++, scoref77, scoref90
 - score_compiler_list



SCore のコンパイルドライバ (2)

- コンパイラの指定が可能
 - コンパイラによるオプションの違いをある程度吸収
- “環境” によって必要なオプションを設定
 - PMX 環境なら PMX のプログラムに必要なライブラリなどを設定
- コンパイラ、環境は環境変数によって指定可能

eScience 成果物の SCore への 取り込み



- eScience プロジェクトの成果物を既存の SCore プログラムと一緒に動くようにする
 - 複数の環境でディレクトリを共有できる
 - top ディレクトリからは smake で make できるようにする
 - オリジナルの build 環境とは異なる場合がある
 - MPI プログラムは MPICH2/SCore で動くようにする

SCore 環境での Xcrypt の使い方

- Xcrypt 関連のコマンドがパスに含まれている

```
$ xcrypt file.xcr
```

- 実際は wrapper 経由で起動
 - 環境変数 XCRJOBSCHED でスケジューラを指定できる。default は sh
 - その他の環境変数 XCRYPT, PERL5PATH の設定は不要

SCore 環境での Xabclib の使い方 (1)

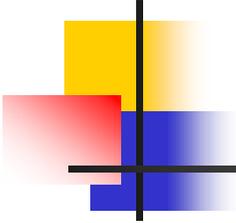
- Xabclib は scoref90 に統合

```
$ scoref90 -compiler compilr -env xabclib  
program.f
```

- インクルードファイルパスやライブラリの指定、
openmp の指定は scoref90 が行う

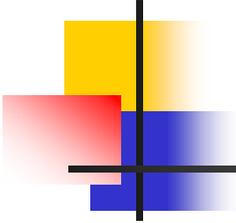
SCore 環境での Xabclib の使い方 (2)

- 環境変数 SCORE_COMPILER_ENV を xabclib に設定することで `-env xabclib` を省略可能
- 環境変数 SCORE_BUILD_COMPILER を設定することで省略時のコンパイラを指定可能。default は gnu
- 必要な環境設定を行うと以下で OK
`$ scoref90 program.f`



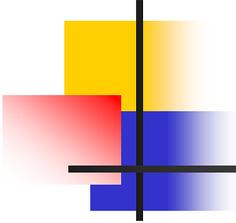
SCore のパッケージ (1/2)

- SCore は独立してインストールできるパッケージに分かれている
- ソース上はひとつのパッケージはひとつの directory の下に置かれている
 - すべてのソースは score7-src の下に置かれていて、パッケージはその下のサブディレクトリになる
 - SCore の基本部分は例外



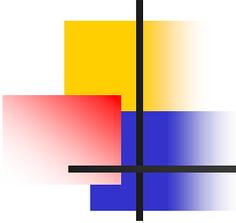
SCore のパッケージ (2/2)

- source ファイルはパッケージごとに tar ファイルでリリース
- rpm は tar ファイルをもとに作成
- コンパイル,インストールはまとめて行える
 - top directory からの make
 - bininstall コマンド



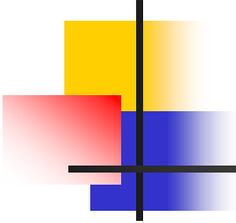
パッケージの内容 (1/5)

- SCore
 - PMX, SCore-D
 - サポートツール、全体的なドキュメント、起動スクリプトなども含む
 - rpm は kernel モジュールのみを comp として分割
 - 残りのパッケージに必要



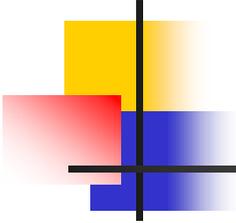
パッケージの内容 (2/5)

- MPI
 - mpilib (score7-src/mpi/lib)
 - MPI を動かすための共有 library, wrapper
 - mpiimpl (score7-src/mpi/impl)
 - MPI 実装
 - 現在は mpich2 のみ
 - mpilib が必要



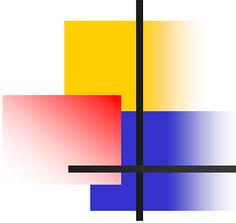
パッケージの内容 (3/5)

- eScience
 - (XcalableMP)
 - まだ統合していない
 - MPI-Adapter (score7-src/eScience/Xruntime/mpi-adapters)
 - mpiimpl が必要
 - mpich_score と centos 5 に付属している MVAPICH2, OpenMPI 相互間の adapter を default で生成



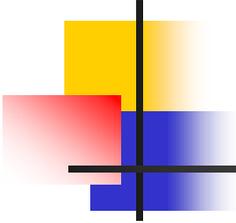
パッケージの内容 (4/5)

- eScience の続き
 - (Catwalk)
 - SCore に含まれる
 - Xcrypt (score7-src/eScience/xcrypt)
 - Xabclib (score7-src/eScience/xabclib)
 - libioc (score7-src/eScience/Xruntime/libioc)
 - stg (score7-src/eScience/Xruntime/stg)
 - mpiimpl が必要



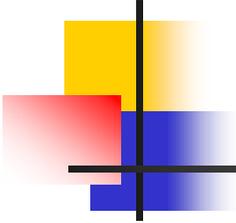
パッケージの内容 (5/5)

- benchmark (score7-src/benchmark)
 - 現在は Intel MPI Benchmark のみ
- test (score7-src/test)
 - SCore 及び mpi のテスト用 script
 - 主として SCore 開発者向け
 - rpm は存在しない



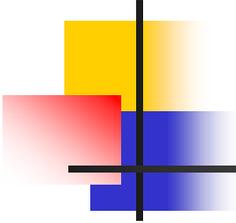
ソフトウェアの入手先

- SCore
<http://www.pccluster.org/>
- XcalableMP
<http://www.xcalablemp.org/>
- Xcrypt
<http://super.para.media.kyoto-u.ac.jp/~tasuku/xcrypt>
- Xabclib
<http://www.abc-lib.org/Xabclib/>



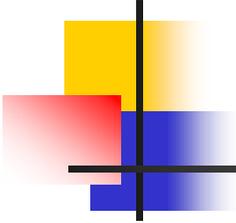
SCore のインストール

- DVD
 - OS のインストールから行う
- rpm
 - 既存のクラスタにインストールする
- ソース
 - root 権限を持たない環境
 - rpm が提供されていない環境/コマンド
 - i386, papion, PGI コンパイラなど



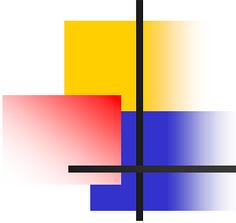
DVD からのインストール (1/2)

- インストール DVD をダウンロード
- DVD を作成
- DVD から boot して score を入力
- 通常の centos のインストール
 - SCore の実行に必要なパッケージがインストールされる



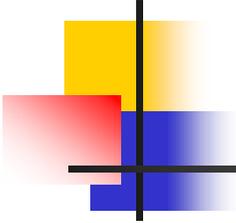
DVD からのインストール (2/2)

- firstboot で firewall なし, selinux 無効を選択
- DVD を mount して rpm install



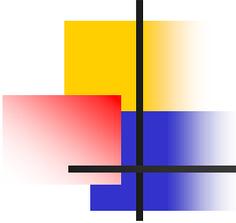
rpm からのインストール

- DVD もしくは rpm 一式をダウンロード
- bininstall を実行
 - # sh ./bininstall -server
 - # sh ./bininstall -comp
- サーバと計算ホストとの共有のときは -comp を実行



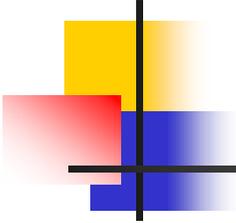
ソースからのインストール (1/5)

- DVD からもしくはソース tar ファイル一式をダウンロード
- tar ファイルを展開
- score7-src で
 - # ./configure
 - # make
 - make install は不要
 - パッケージごとに install するため



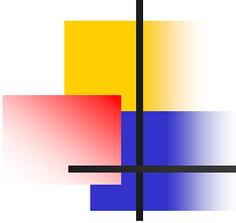
ソースからのインストール (2/5)

- `configure` のオプション
 - `--prefix=/foo`
 - インストール先を `/foo` に
 - `--scorehome=/foo`
 - インストール先を `/foo/score_version` に
 - バージョンを上げてても設定ファイルを維持する



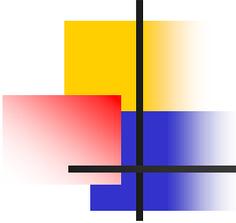
ソースからのインストール (3/5)

- `configure` のオプション (続き)
 - `--mx=/usr/local/mx`
 - MX のインストールディレクトリを指定
 - `/opt/mx` にインストールしている場合は検出する
 - `--ofed=/usr/local/ofed`
 - OFED のインストールディレクトリを指定
 - `/usr` にインストールしている場合は検出する



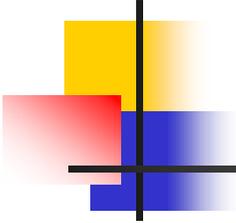
ソースからのインストール (4/5)

- `configure` のオプション (続き)
 - `--glut=/usr/local/glut`
 - `glut` のインストールディレクトリを指定 (scope)
 - `/usr` にインストールしている場合は自動検出
 - `--papi=/usr/local/papi`
 - `papi` のインストールディレクトリを指定 (papion)



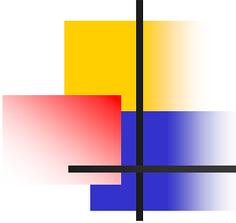
ソースからのインストール (5/5)

- configure のオプション (続き)
 - `--compilers=intel,fujitsu,pgi,pathscale`
 - オプションのコンパイラを指定
 - `mpi, xanclib`などを個別に作成する場合は
`/opt/score/etc/compilers/site`を編集して該当パッケージだけをインストールすればよい



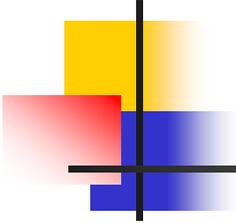
パッケージのインストール (1/4)

- rpmなどで一度 SCore を install した後でそれぞれのパッケージをソースからインストールすることも可能
 - コンパイラや対応 MPI の種類を増やすことができる
 - mpi や libioc など時間がかかり、更新する必要のないパッケージのコンパイルを避けることができる



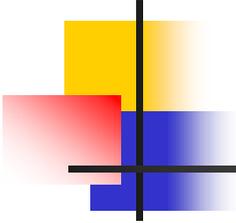
パッケージのインストール (2/4)

- そのパッケージの tar ファイルのみを展開
- ソースの置いてあるディレクトリに cd
- smake を実行
 - # cd score7-src/ *eScience/xabclib*
 - # smake
 - # smake install



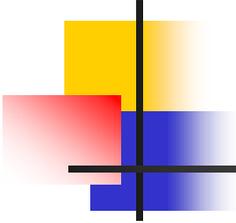
パッケージのインストール (3/4)

- mpiimpl, xabclib
 - 使用可能なすべてのコンパイラでライブラリを作成する
 - /opt/score/etc/compilers/site.x86_64-rhel5-linux2_6 を編集してインストール
 - script compiler=path...
mpif90 pgi=pgf90 intel gnu
scoref90 pgi=pgif90 intel gnu



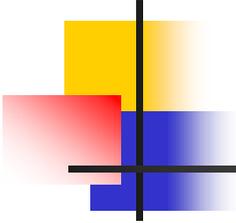
パッケージのインストール (4/4)

- mpi-adapter
 - score7-src/eScience/Xruntime/mpi-
adapters/mpi-local.x86-64rhe/5-
linux2_6.conf を書くことで対象 mpi を追加で
きる
 - mpiname type install_root
mpich mpich /usr/local/mpich2
 - type は mpich, mpich_score, ompi, hpmpi



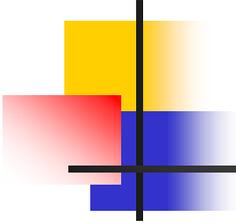
SCore の設定順序

- まず rsh-all, scout を使えるようにする
- NFS, PATH などの実行するための基本的な設定
- PMX の設定及び確認
- daemon, SCore データベース
- 環境変数による default の設定



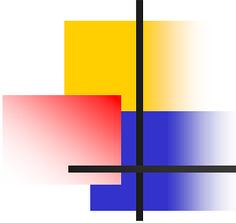
rsh-all, scout の設定 (1)

- SCORE_RSH を設定
 - 指定が無い場合は ssh
- scout の場合、計算ホストから別の計算ホストへ ssh/rsh を行う



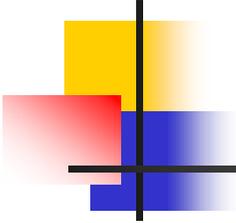
rsh-all, scout の設定 (2)

- rsh を使う場合
 - kerberos を使っていないと余分なメッセージが...
 - /usr/bin/rsh に設定しておいたほうが良い
 - /etc/hosts.equiv \$HOME/.rhosts
 - root の場合は /etc/securetty に rsh を追加
- SSH を使う場合
 - ssh-agent, null passphrase
 - ssh-copy-id で公開鍵転送



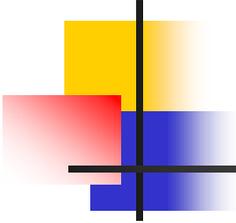
NFS の設定

- 現在の SCore7 ではプログラムの計算ホストの転送は行わない
 - /home などコマンドを置く場所は NFS にしたほうが望ましい



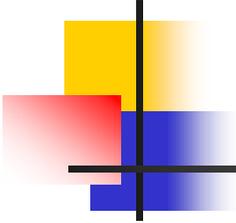
実行環境の設定 (1/2)

- PATH, MANPATH の設定
 - rpm の場合は `/etc/profile.d/score*` で設定
 - source で root 権限があるならば,
`# /opt/score/install/setup -score_server`
 - root になれない場合は以下を参照
`/install_root/install/profile/score.{sh,csh}`
`/install_root/install/profile/scorempi.{sh,csh}`



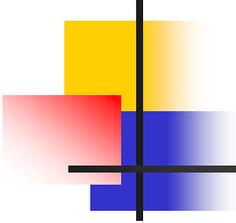
実行環境の設定 (2/2)

- mpi-selector
 - 複数の mpi がインストールされている場合、どれを使用するか個人が選べるようにするしかけ
 - mpi-selector-menu で選択できる
 - 設定したあと一度 logout する必要がある
 - mpi-selector がインストールされている場合、`/opt/score/install/setup -score_server` で設定



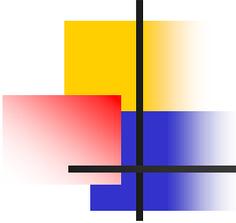
PMXの設定 (1/4)

- PMX で ethernet を使用するには kernel module をロードする必要がある
 - PMX/sctp sctp module (kernel 標準)
 - PMX/Ethernet pmx_ethernet_loader, pmx_ethernet (SCore)
 - PMX/Etherhxb pmx_ethernet_loader, pmx_etherhxbmem, pmx_etherhxb, patch をあてた NIC driver (SCore)



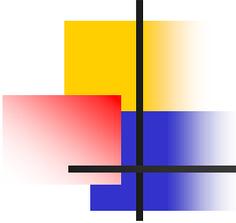
PMXの設定 (2/4)

- kernel module のロードは root 権限が必要
 - /opt/score/install/setup -comp で設定
 - /etc/init.d/sctp
 - /etc/init.d/pmx_ethernet
- PMX/Etherhxb で必要な NIC driver は SCore で用意していないものがある
 - /opt/score/doc/HowtoAdd-PMX-EtherHXB-Driver.txt 参照



PMX の設定 (3/4)

- sctp モジュールがロードされているかの判定
 - /proc/net/sctp/ があるかどうか?
- PMX/Ethernet, PMX/Etherhxb が使用可能かの判定
 - /proc/pmx/ether_loader を見る
 - PMX/Etherhxb の場合は使用可能な interface もチェック



PMXの設定 (4/4)

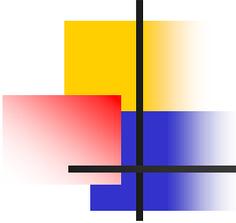
- 動作確認 pmxtest

- scout 上で動かす

- \$ scout -g comp0+comp1

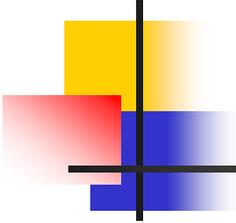
- \$ pmxtest -network *ethernet*

- \$ pmxtest -composite 2 -network *ethernet*



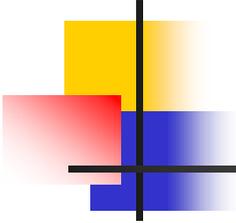
daemon の設定

- daemon は無くても一通りは動く
 - oomadjserv
 - oomkiller のときに SCore のジョブを優先的に落とす
 - scoreboard
 - SCore database が無くても動く
 - 環境変数 SCBDSERV でファイルを指定可能
 - msgbserv
 - scout のジョブが複数動くことを抑制
- /opt/score/install/setup -score_server で設定



score database の必要性

- SCore7 ではなくても動く
 - scout など host 名や host 名が書かれたファイルを指定できる
- 細かい指定などでほしいときがある
 - PMX/Ethernet で一部 interface 名が違うとき
 - group を使用したいとき
 - ファイルで代用可能
 - msgbserv を使用したいとき



score database の設定

- network レコードと host レコード
 - network レコードは PMX に関する記述
 - host レコードは host に関する記述
- host レコードは scbdrec で生成

```
# rsh-all -q -s -P -g comp_hosts ¥  
scbdrec > /opt/score/etc/scorehosts.db
```

 - msgbserv, group などは追加する必要がある