

九州大学情報基盤研究開発センターの PCクラスタ導入

九州大学情報基盤研究開発センター
天野浩文

1

本発表では...

- 本センターとして始めて導入したPCクラスタについて
 - 導入の背景
 - 新スーパーコンピュータシステムの概要
 - PCクラスタ導入に至る経緯
 - PCクラスタでのサービス
 - PCクラスタに寄せられる期待

2

センターの使命

- 1970年 旧大型計算機センターとしてスタート
 - 全国共同利用施設として、全国の大学等の研究者に計算サービスを提供
- 2000年 情報基盤センターへ改組
 - 学内情報基盤の整備・運用も担当
- 2007年 情報基盤研究開発センターへ
 - 情報統括本部の発足に合わせて
- 全国共同利用の計算サービスは今なお堅持

3

システムの変遷

- 旧大型計算機センター発足当初
 - いわゆるメインフレーム
- ベクトル型スーパーコンピュータの追加
 - 後に、ベクトル並列型スーパーコンピュータへ
- 2つのシステムを独立に更新
 - メインフレームは、大規模共有メモリを有するスカラ並列型計算機へ
- 今回の更新対象
 - ベクトル並列型スパコン

4

調達方法の改革

- 旧メインフレームと旧スパコンの予算規模
 - 4倍以上の開きがあるのを平衡させる
- 更新サイクル・時期の調整
 - 2つのシステムを、交互に等間隔で更新
 - なるべく利用者への少ない時期に更新
- 今回の調達では、予算規模の大きいほうのシステム(旧スパコン)を、2つに分けて
 - 次回の更新から、2年に一度、ほぼ同規模のシステムが最新機種へ
 - クラスタは、「新スパコン」の一部として導入

5

新システムの概要(1)

- バックエンドサーバ
 - FUJITSU PRIMEQUEST580
 - 13.1 TFLOPS・メモリ4TB
- バックエンドサーバB
 - FUJITSU PRIMERGY RX200S3 クラスタ
 - 18.4 TFLOPS・メモリ3TB
- フロントエンド/ファイルサーバ
 - FUJITSU PRIMEQUEST580
- ディスクアレイ
 - 250TB(データ) + 250TB(バックアップ)

6

新システムの概要(2)

- バックエンドサーバA
- フロントエンド/ファイルサーバ
- バックエンドサーバB



FUJITSU PRIMEQUEST 580



FUJITSU PRIMEERGY RX200S3 クラスタ



7

PCクラスタの詳細(1)

- 演算ノード: 富士通(株) PRIMERGY RX200S3
 - Intel Xeon 3.0GHz (デュアルコア) × 2 (=4コア)
 - メモリ8GB
- 相互結合網: InfiniBand 4xDDR × 1 / ノード
- 総ノード数: 192ノード × 2セット
 - 384プロセッサ (768コア) × 2セット
- 理論演算性能の総和: 18.4 TFLOPS
- 主記憶容量の総和: 3 TB

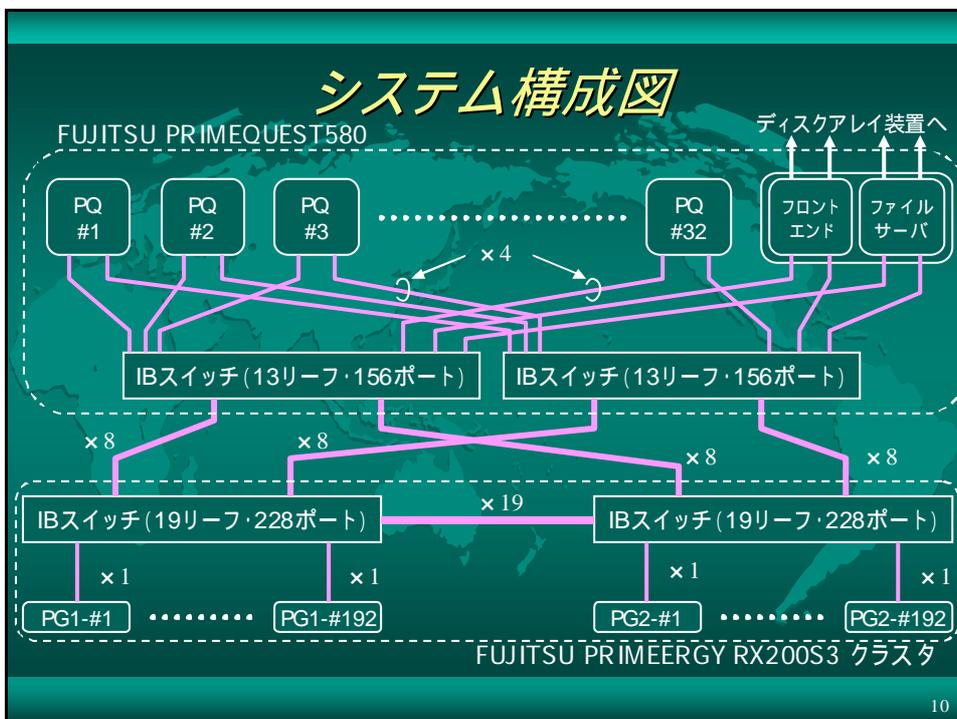
8

PCクラスタの詳細(2)

- OS: Red Hat Enterprise Linux WS (v.4)
- ファイルシステム: Parallelnavi SRFS for Linux (V2.0L10)
- バッチジョブ管理システム:
Parallelnavi NQS for Linux (V2.0L10)
PBS Professional 8.0
- 数値計算ライブラリ:
Parallelnavi Language Package for Linux
- アプリケーション: GAMESS, CHARM, AMBER

9

システム構成図



10

フォトギャラリー(1)



11

フォトギャラリー(2)



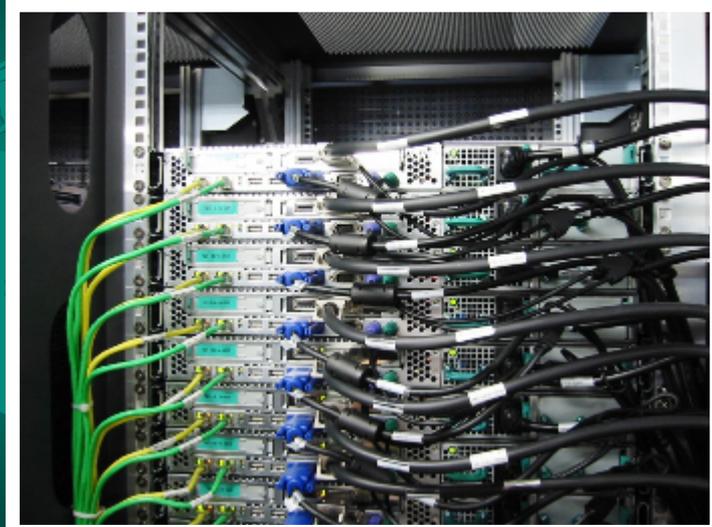
12

フォトギャラリー(3)



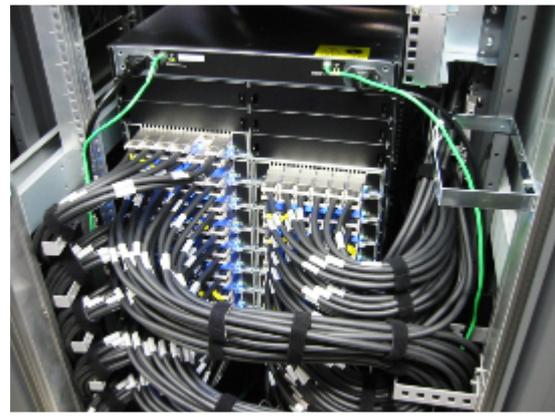
13

フォトギャラリー(4)



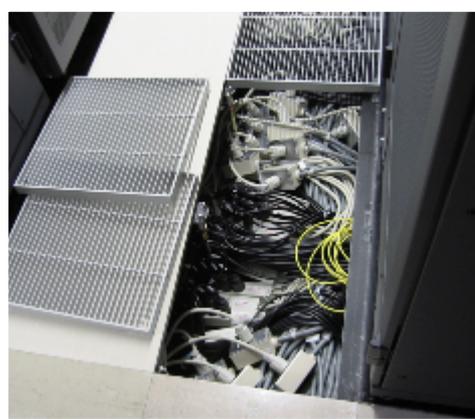
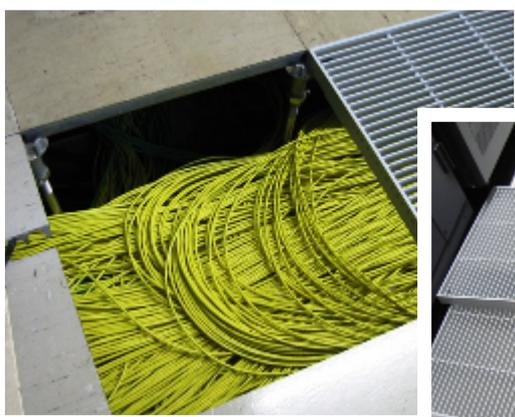
14

フォトギャラリー(5)



15

フォトギャラリー(6)



16

ケーブルの総延長

- InfiniBand
 - PRIMEQUEST用: 約 2,220 m
 - PRIMERGY用: 約 3,224 m
 - スイッチ間: 約 159 m
 - 総計: 約 5,603 m
- LAN
 - PRIMEQUEST用: 約 3,000 m
 - PRIMERGY用: 約 22,600 m
 - その他: 約 1,400 m(スイッチ, ディスクアレイ等)
 - 総計: 約 27,000 m
- Fibre Channel
 - 約 1,920 m

17

新システムの選択肢

- 旧機種はベクトル並列機だったので...
 - ベクトル並列型? もしくは
 - 自動ベクトル化の代わりにスレッド並列化を用いる, 大規模共有メモリ型スカラー並列機?
- 利用者には手の届かないような計算機を提供することを最重要視するなら:
 - PCクラスタは新システムの選択肢に入らない, とする考え方も?

18

クラスタ導入に至る経緯(1)

- 今回の調達から、リース料金を抑えて電力料金支払いの余力を残す
 - 電力料金の心配なく大規模なシステムを導入したい
 - 利用者の拡大を目指す
- ところが...
ベクトル並列型・大規模共有メモリ型だけでは:
 - われわれの予算規模では、大規模なシステムには手が届かない
 - より安価な大規模計算機資源の追加が必要

19

クラスタ導入に至る経緯(2)

- 利用者数減少の長期傾向続く
 - PCの普及
 - データベース検索サービスの縮小・撤退
 - PCクラスタに乗り換える利用者
- 新規利用者の候補として
 - 自前のPCクラスタを抱えて(苦労して)いる人々
- より安価で魅力的なPCクラスタを提供するには
 - やはり大規模な資源の追加が必要

20

クラスタでのサービス(1)

- 占有タイプ
 - バッチジョブが処理されるバックエンドサーバのノード数を保証
 - フロントエンドは他の利用者と共有
 - 年間の利用負担金は前払いで一定
 - 追加料金発生せず
 - 64コア・メモリ89.6GBで年額80,000円から
- 現在, 全体の半分以上を占有タイプに割り当て

21

クラスタでのサービス(2)

- 共有タイプ
 - 従来と同様に, 他の利用者とバックエンドサーバを共有
 - 占有するほどではない人向け
 - 128コア・メモリ179.2GBのジョブを投入する「権利」の価格: 10,000円

22

今後の展望(1)

- 民間からの利用に向けて
 - 『ペタコン』の全国共同利用・民間供用
 - 本センターには、利用者を「送り出す」任務
 - 将来的には、民間利用者も本格的に受け入れる必要あり
 - 『先端研究施設共用イノベーション創出事業【産業戦略利用】』への参加
 - 他の6つの全国共同利用センターとともに
 - <http://kyoyo.itc.u-tokyo.ac.jp/>
 - 本格供用に向けた経験を積むことを期待

23

今後の展望(2)

- グリッド技術の利用
 - グリッドコンピューティングにより、国レベルでの計算機資源の透過的な運用を
 - 7つの全国共同利用センター・他の同様のサービスを提供するセンターは、この体制への参加が求められる
- 本センターでも
 - クラスタにNAREGIミドルウェアを導入し運用するための検討を開始

24

まとめ

- 新スーパーコンピュータシステムの概要
- 本センターとして始めて導入したPCクラスタ
 - － 導入に至った経緯
 - － サービスの概要
 - － クラスタに寄せられる期待