KEK 新スーパーコンピュータシステム

松古 栄夫 （計算科学センター）

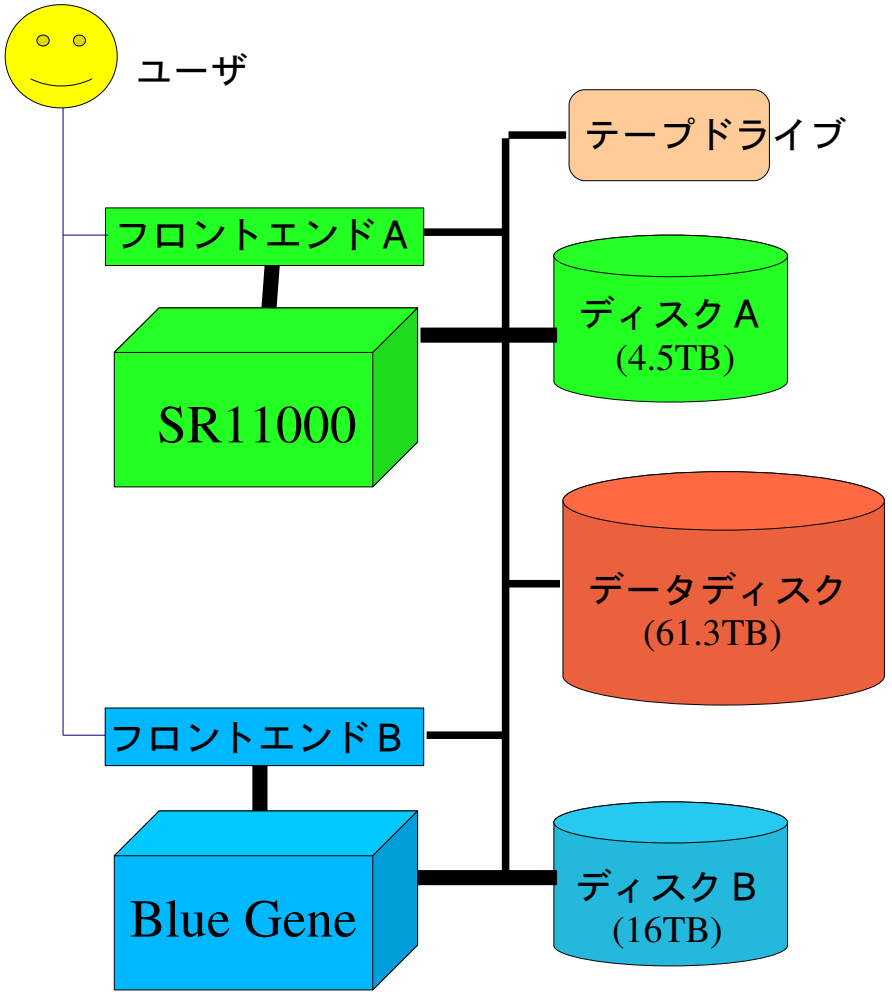
High Energy Accelerator Research Organization (KEK)

目次

- システム概要
  - 全体構成
  - システムA (Hitachi SR11000)
  - システムB (IBM Blue Gene)
  - 周辺装置（ディスク/テープ/WWWサーバ）
- 大型シミュレーション研究
  - テスト運用期間
  - 本運用期間
- プログラム開発支援
- スケジュール

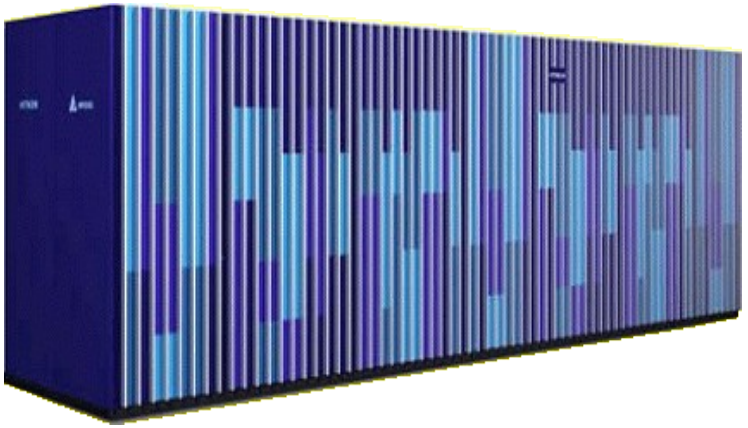
システム概要/システム全体構成

- システムA: Hitachi SR11000
  - 16 node, peak 2.15TFlops
  - 汎用計算サーバ
  - ディスクA: 4.5TB
- システムB: IBM Blue Gene
  - 10 rack (10240 node), 57.3TFlops
  - 高並列型計算に特化
  - ディスクB: 16TB
- データ保管用ディスク: 61.3TB
- バックアップ用テープドライブ



システム概要/システムA

- 計算サーバ：Hitachi SR11000 (16 node, peak 2.15 TFlops)
  - プロセッサ: POWER5+ (2.1GHz, 8.4 Gflops, 36MB L3 cache)
  - 1ノード=16CPU, SMP並列 (134.4 GFlops), 24GBメモリ
    - 自動並列化, MPI並列も可能
  - ノード間: クロスバスイッチ
    - MPI並列
- 高速アクセスディスクA: 4.5 TB
- フロントエンドA
  - ログイン(ssh2のみ)、コンパイル、キュー投入など
  - OS: AIX Unix



システム概要/システムA

システムAのキュー構成（テスト運用期間）

キュークラス	ノード数	制限時間	備考
q1s	1	5 min	テスト用
q1m	1	30 min	
q1a,q1b	1	2 h	
q1c,q1d	1	4 h	
q1e,q1f	1	8 h	
q4a,q4b	4	4 h	ノード間MPI並列

- 1ノードあたりの使用可能メモリ: 24GB
- LoadLeveler によるキュー管理 (NQSと同様の機能)
- ジョブ実行開始、終了(性能情報)をメールで通知
- プロファイラによるチューニング支援

システム概要/システムB

- 計算サーバ: IBM Blue Gene (10 rack, 57.3TFlops)
  - 1 rack = 1024 nodes = 2048 CPU
  - プロセッサ: PowerPC440 (700MHz, 2.8 Gflops)
    - Double FPU: 複素数演算を高速に処理
  - 1 ノード= 2CPU, 4MB L3 cache (共有)
    - 1ノードあたりのメモリ: 512MB
    - Virtual node mode: ２つのCPUが独立に計算
    - Coprocessor mode: 片方のCPUが通信を担当
  - ノード間ネットワーク
    - midplane(rack/2): 3次元トーラス(8x8x8)
    - global tree

(1 Feb 2006)

システム概要/システムＢ

- 高速アクセス用ディスクＢ: 16TB
- フロントエンドＢ (3台)
  - ログイン(ssh2のみ)、コンパイル、キュー投入など
  - OS: SuSE Linux

システム概要/システムＢ

システムＢのキュー構成（テスト運用期間）

キュークラス	ノード数	制限時間	ネットワーク	備考
q32s	32	10 min	Mesh	テスト用
q32a-c	32	4 h	Mesh	
q128a-c	128	8 h	Mesh	
q512a-i	512	12 h	Torus	
q1024a-e	1024	12 h	Torus	

- 1ノードあたりのメモリ: 512MB
- LoadLeveler によるキュー管理 (NQSと同様の機能)
- ジョブ実行開始、終了をメールで通知
- ジョブの性能情報を一日一回通知 (mail or file)
- プロファイラによるチューニング支援

システム概要/周辺装置

- データ保管用ディスク (61.3 TB)
  - システムＡ、Ｂからアクセス可能
- バックアップ用テープ装置: SONY PetaSite
  - システムＡフロントエンドから操作
  - 最大780TB
  - SAITテープ 1巻 500GB(非圧縮時)--1.3TB(最大圧縮時)
- ウェブサーバ
  - URL: http://scwww.kek.jp/
  - システム関連情報を提供

システム概要/Benchmark の結果

格子QCD, Wilson solver

- システムＡ
  - 24^3 x 48 on 16 node: ～28% (0.61 TFlops)
- システムＢ
  - 24^3 x 48 on half-rack: ～29% (0.82 TFlops)

今日の第2マシン室

２００６年２月７日　１３時頃

大型シミュレーション研究

テスト運用期間 (2006年3月1日--8月末)

- 審査会を経ずに利用可
  - システムA: 1ノードキュー 500時間まで
  - システムB: 32ノード、128ノードそれぞれ300時間まで
  - プログラムが一定の基準を超えれば、上記キューの追加や他のキューも利用が可能 (追加申請)
  - テスト運用終了時にディスクのクリアの可能性あり
- 申請 (http://ohgata-s.kek.jp/ 参照)
  - 申請書、メンバー表、ユーザ登録用書類
  - 締切: 2月17日(金) –3/1より利用の場合 / その後も随時受け付け
- 講習会 3月6,7日

大型シミュレーション研究

本運用期間 (平成18年度: 2006年9月--年度末)

- 7月中旬申請締切、8月に審査会
- 各キューの利用時間、ディスク資源を配分

プログラム開発支援

- チューニングマニュアル、各種マニュアル – Web で配布
- 講習会 -- 3/6,7 に開催 (申込は下記ウェブサイト参照)
- プログラミング相談
  - メール: scconsult-a@sc.kek.jp, scconsult-b@sc.kek.jp
- 一部ベンチマークプログラムの公開 (予定)
- お知らせ
  - システム関連: http://scwww.kek.jp/
  - 大型シミュレーション関連: http://ohgata-s.kek.jp/

スケジュール

- 2月17日 テスト運用期間申請締切
- 2月17日 新システム講習会申込締切
- 3月1日 運用開始
- 3月6、7日 新システム講習会
  - 6日(13:00-17:30): システム概要、Aシステム
  - 7日(9:00-12:30): Bシステム
- 6月 平成18年度本運用期間の申請アナウンス
- 7月中旬 本運用申請締切
- 8月 本運用審査会
- 9月初旬 本運用開始