



T2K Open Supercomputer Alliance

T2K オープンスパコン 設計思想とアーキテクチャ

主演：中島 浩（京都大学）


**共演：佐藤 三久（筑波大学）
朴 泰祐（筑波大学）
石川 裕（東京大学）**



T2K Open Supercomputer Alliance

目次

- T2K オープンスパコン・アライアンス
 - 「同盟」メンバーと目的
- T2K オープンスパコン仕様
 - 何がどう「オープン」なのか？
 - 共通仕様が定めるものと定めのないもの
- T2K オープンスパコン
 - ノードアーキテクチャ／サイトごとの構成
 - FAQ
 - 過去&未来




 T2K オープンスパコン・アライアンス
同盟メンバー


 **京都大学 (Kyoto U.)**
学術情報メディアセンター


 **筑波大学 (U. Tsukuba)**
計算科学研究センター

 **東京大学 (U. Tokyo)**
情報基盤センター






T2K Open Supercomputer Alliance

 T2K オープンスパコン・アライアンス
同盟の目的

- 旧来型のスパコン調達から...
 - 受動的&ベンダー主導
 - 製品市場からの選択型
- 新たな調達スキームに転換し...
 - 能動的&大学主導
 - 技術市場からの創造型
- 高性能計算の新たなソリューション提供
 - 最先端技術に基づく設計
 - 幅広い大学ユーザへの対応

T2K Open Supercomputer Alliance

**T2K オープンスパコン仕様
何がどうオープンなのか？**

アーキテクチャが open

- コモティティ部品ベース
e.g. x86, IB/Myri-10G
- 現在の技術市場から入手可能な最先端・高性能デバイスによる構成
- HPC 専用の特殊な部品は排除

ソフトウェアが open

- オープンソースの標準 OS/ミドルウェア
e.g. Linux, MPI, Globus
- オープンソースの HPC ミドルウェア/ツール/ライブラリ等の実装基盤

ユーザニーズに open

- 計算科学分野の伝統的な FP ユーザだけでなく...
- PC クラスタを使っている INT ユーザも誘引
 - 検索/マイニング
 - 自然言語処理
 - 遺伝子情報処理

T2K Open Supercomputer Alliance

**T2K オープンスパコン仕様
何が共通仕様なのか？**

- **共通の要求要件**
 - **ハードウェア**
 - 16以上の x86 コア+容量 32GB 以上 & 合計転送性能 40GB/sec 以上のメモリからなる共有メモリノード
 - 合計転送性能 5GB/sec 以上の偶数本ノード間リンク群
 - 250GB 以上の RAID-1 ノードディスク(optional) & IPMI2.0
 - **ソフトウェア**
 - Red Hat または SuSE Linux
 - Fortran, C, C++ (OpenMP 対応 & 自動並列化)
 - Java (JITコンパイラ付)
 - MPI : バンド幅 4GB/sec 以上 & 往復遅延 8.5- μ sec 以下
 - BLAS, LAPACK, ScaLAPACK
 - **ベンチマーク (要求性能値は個別)**
 - SPEC CPU2006, SPEC OMP2001, HPC Challenge (部分)
 - 自作のメモリ性能, MPI性能, ストレージ性能ベンチマーク


T2K Open Supercomputer Alliance

**T2K オープンスパコン仕様
何が共通仕様ではないのか？**

■ サイト固有仕様

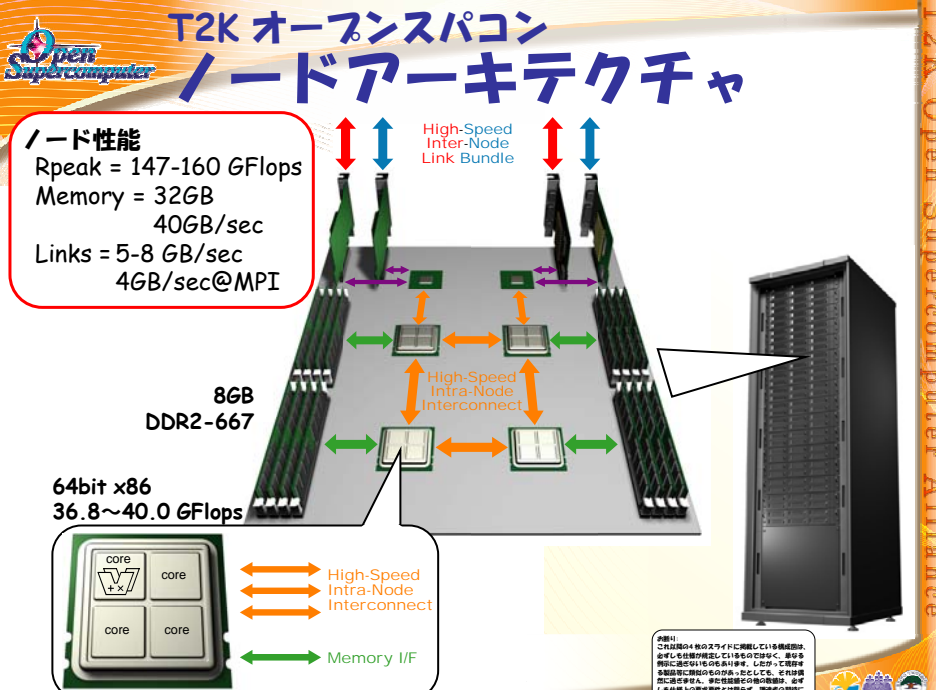
- **ハードウェア**
 - コア/ソケット/ノードのピーク演算性能
 - 総ノード数/総ピーク演算性能
 - ノード間結合網の構成方式とピークバンド幅
 - ストレージ容量/転送性能
 - 消費電力
- **ソフトウェア**
 - ジョブスケジューラ, 耐故障機能, システム管理
 - 商用ライブラリ/アプリケーション
- **ベンチマーク**
 - 共通ベンチマークの要求性能
 - サイト固有の応用ベンチマーク

T2K Open Supercomputer Alliance



**T2K オープンスパコン
ノードアーキテクチャ**

ノード性能
 Rpeak = 147-160 GFlops
 Memory = 32GB
 40GB/sec
 Links = 5-8 GB/sec
 4GB/sec@MPI



8GB DDR2-667

**64bit x86
36.8~40.0 GFlops**


High-Speed Inter-Node Link Bundle

High-Speed Intra-Node Interconnect

Memory I/F

※留意:
 この図は4コアのスライドに構成している構成図例。
 全ては各製メーカーにより異なる可能性があります。また、
 構成により異なる場合があります。したがって、実際の
 構成は各メーカーの仕様を必ずご確認ください。また、
 本図はあくまで参考図であり、最終的な構成は
 各メーカーの仕様を必ずご確認ください。必ず
 して仕様上の留意事項を必ずご確認ください。製造者の説明に
 必ずして必ずご確認ください。

T2K Open Supercomputer Alliance



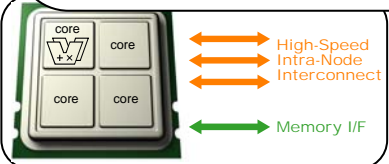
**T2K オープンスパコン
ノードアーキテクチャ**

ノード性能
Rpeak
Memory
Link

↑ ↑ High-Speed Inter-Node Link Bundle ↑ ↑

お断り:
これ以降の4枚のスライドに掲載している構成図は、必ずしも仕様が規定しているものではなく、単なる例示に過ぎないものもあります。したがって現存する製品等に類似のものがあっても、それは偶然に過ぎません。また性能値その他の数値は、必ずしも仕様上の要求要件とは限らず、調達者の期待に基づく値である場合もあります。

6
36



High-Speed Intra-Node Interconnect

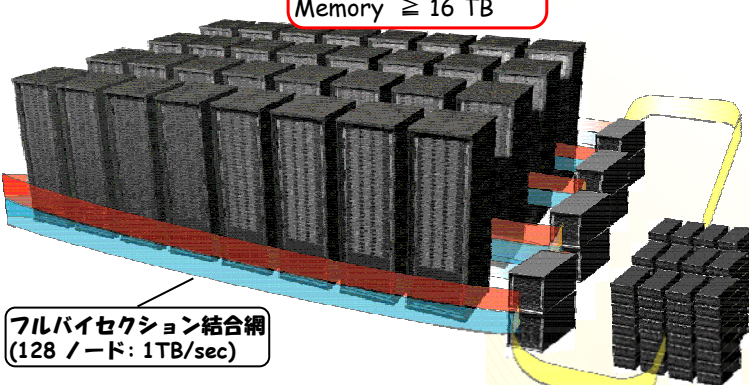
Memory I/F

※留意:
これは図の4枚のスライドに掲載している構成図は、必ずしも仕様規定しているものではなく、単なる例示に過ぎないものもあります。したがって現存する製品等に類似のものがあっても、それは偶然に過ぎません。また性能値その他の数値は、必ずしも仕様上の要求要件とは限らず、調達者の期待に基づく値である場合もあります。

T2K Open Supercomputer Alliance

**T2K オープンスパコン
筑波大学の構成**

ノード数 ≥ 512
Rpeak ≥ 75 TFlops
Memory ≥ 16 TB

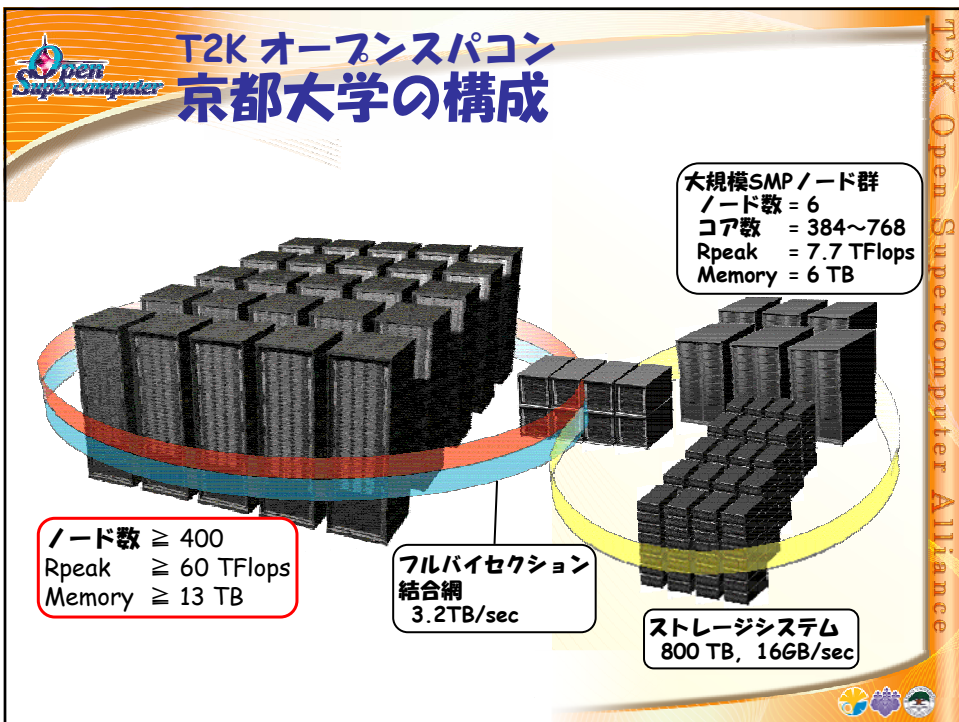
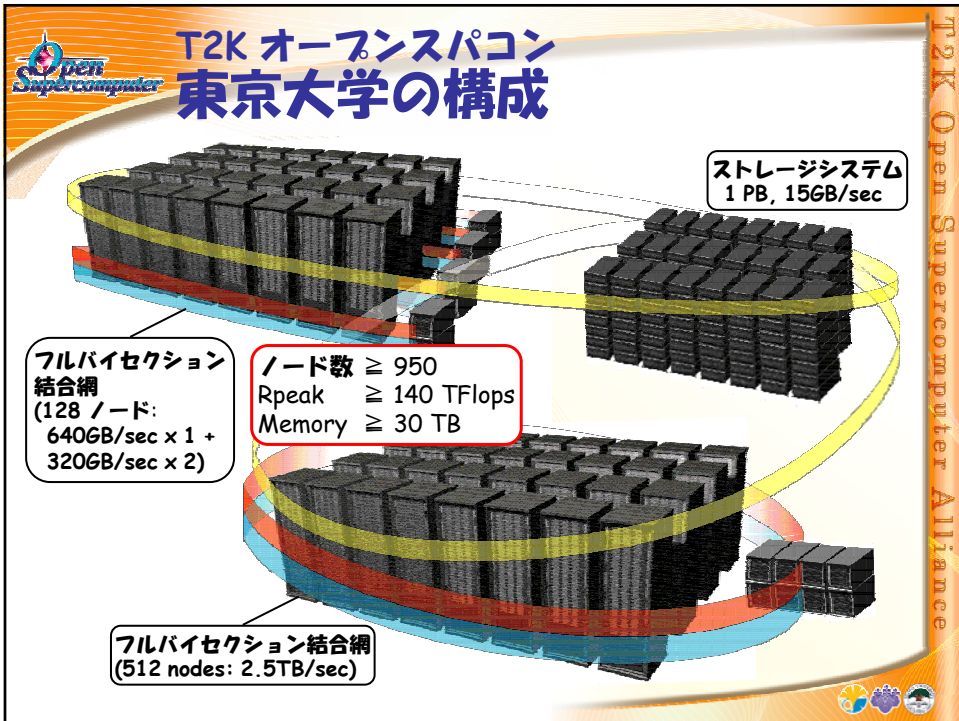


フルバイセクション結合網
(128 ノード: 1TB/sec)

1/8~1/1 x フルバイセクション
結合網: 0.5~4TB/sec

ストレージシステム
400 TB, 10GB/sec

T2K Open Supercomputer Alliance





T2K オープンスパコン TOP500 番付予想

	筑波大	東大	京大
想定 Rpeak (TFlops)	75.0	140.0	61.0
想定 Rmax (Rpeak x 0.7)	52.5	98.0	42.7
Nov'07 順位	20	8	25
Jun'08 想定順位 (Jun'07 → Nov'07)	30	14	36



T2K オープンスパコン FAQ

Q1: 共同調達なの？

A1: 違います。3大学で独立した調達ですので、異なるベンダーが落札するかもしれません。

Q2: Tsubame とか Ranger@TACC の類似品では？

A2: 全く違います。たとえば T2K のノード間結合網は、Tsubame や Ranger の 3~4 倍の転送バンド幅を有しています。

Q3: 10PFlops マシン@理研の先駆けなの？

A3: さあ。それなりに繋がるといいですけどね。

Q4: 調達完了後は同盟は解体するの？

A4: いいえ T2Kでの研究開発協力体制を続けます。





T2K オープンスパコン 昨日と明日とずっと先

T2K Open Supercomputer Alliance

- **昨日まで**
 - 06/05: T2K 最初の打合せ
 - 06/09: 記者会見&公開シンポジウム
 - 06/10: 資料招請&調達手続開始
 - 07/03: 仕様書案公表
 - 07/10: 最終仕様書公表&入札手続開始
 - 07/11: 応札締切
- **明日から**
 - 07/12: 開札 (on Christmas Morning !!)
 - 08/06: 納入&運用開始
- **ずっと先**
 - T2K マシンのリフレッシュ
京大=2012, 筑波大=2013, 東大=2014
 - T2K の研究開発協力はまだまだ続く...

