

システム構成の見直し案について

2009年5月25日
理化学研究所
次世代スーパーコンピュータ開発実施本部

構成見直し(新システム構成)のポイント

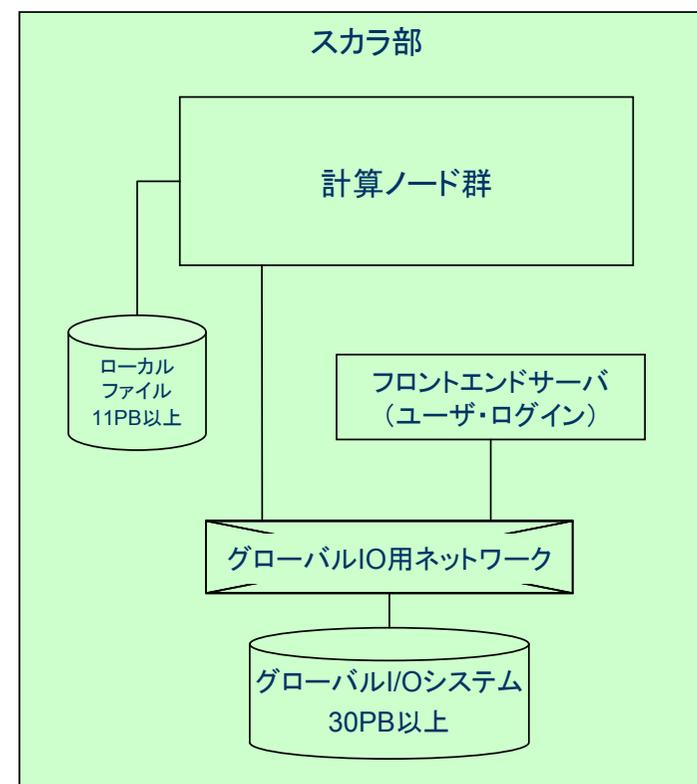
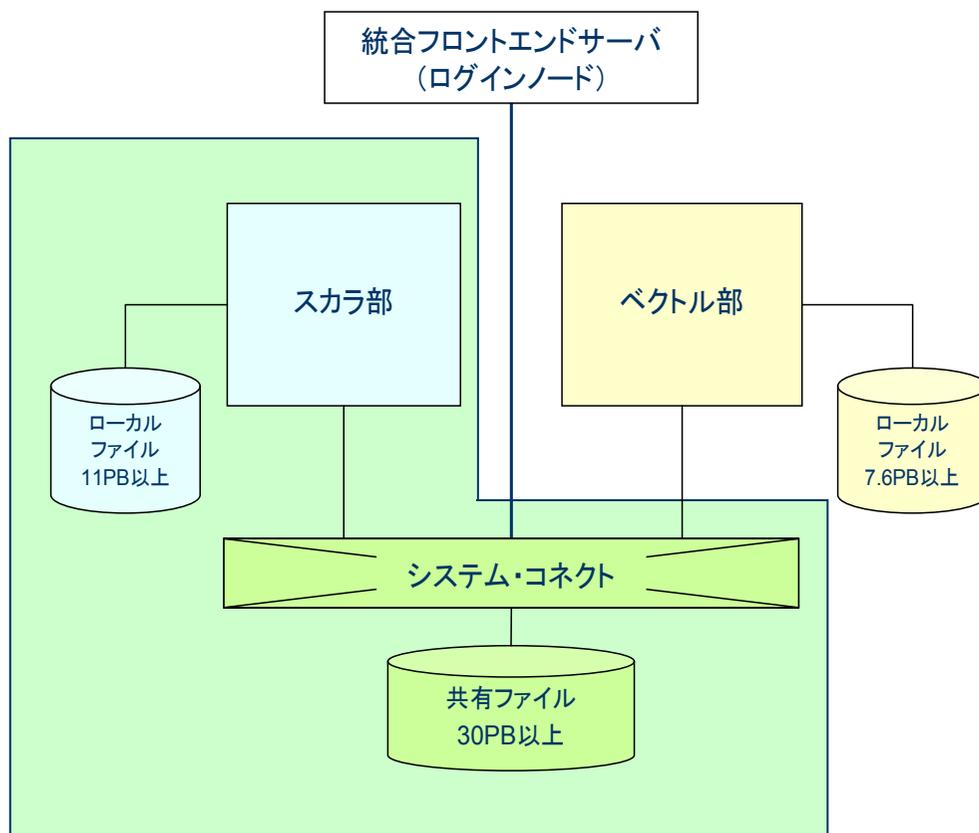
- スカラ部の単独構成により、世界最速への挑戦を継続
 - これまでの詳細設計の成果を活かしつつ、製作能力を増強し、可能な限り整備スケジュールを加速。
 - 平成23年(2011年)11月までにLINPACK性能 10PFLOPSを達成(6ヶ月前倒し)
 - その時点でのTOP500において世界最速を目指す。
 - 世界最高クラスのアプリケーション実行効率(ピーク性能に対するアプリケーション実行性能比)を目指す。
- 高信頼性システムの実現により、優れた成果を早期に創出
 - 世界最大のメモリ空間を確保し、世界最大規模のアプリケーション実行環境をユーザに提供
 - 高信頼性のプロセッサ及びネットワークによる早期安定稼動を実現
 - 出来る限り早期に新システムの運用を開始し、世界的な研究教育拠点の形成に向け、人材の育成、優れた成果の創出に貢献

新システム構成案

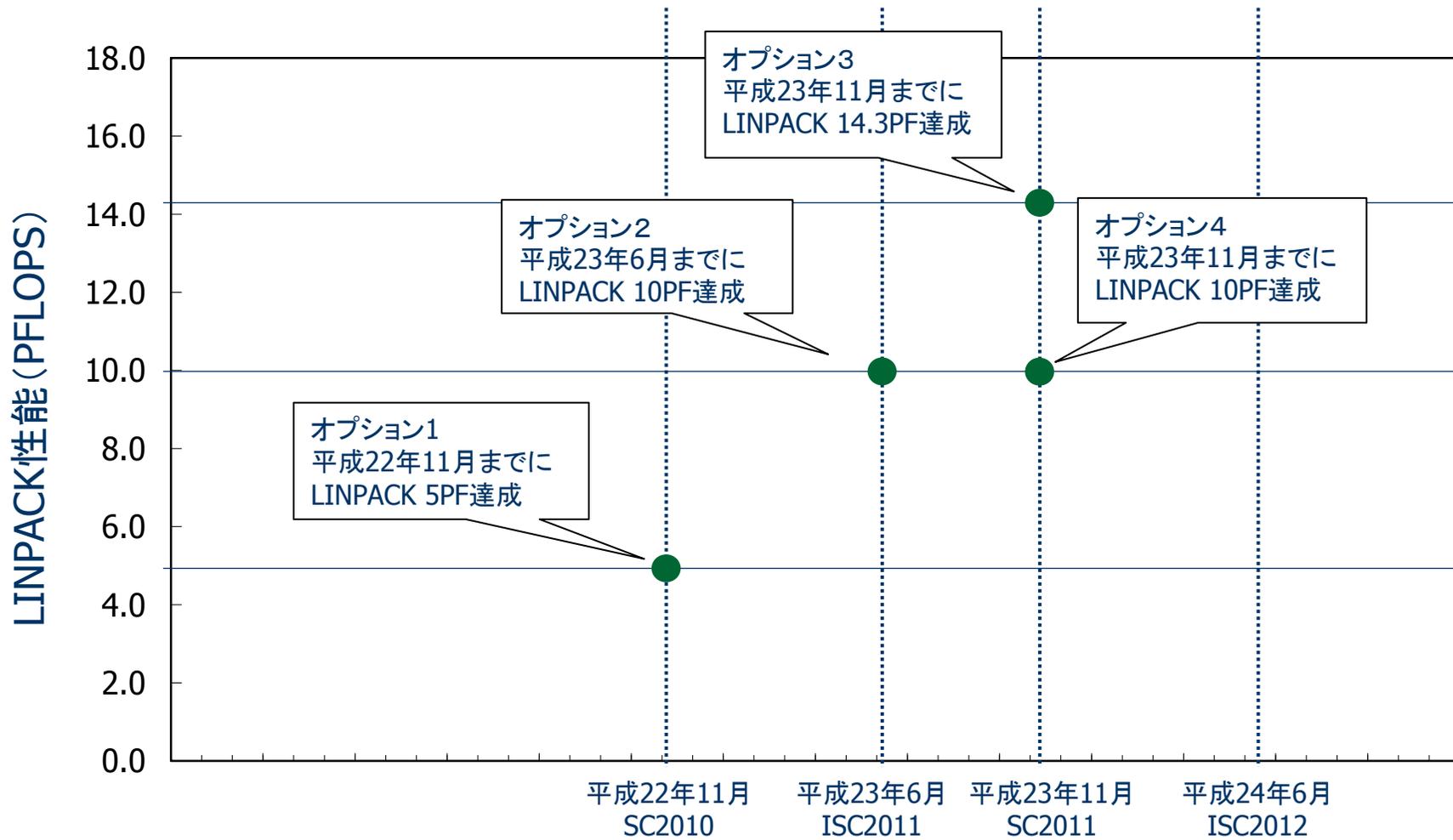
従来の構成
(スカラ部とベクトル部から成る
複合システム)



新システム構成
(スカラ部による単独構成)



検討対象としたオプション



各オプションの搬入据付スケジュール

	2010年												2011年												2012年					
	平成22年度												平成23年度												平成24年度					
	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6			
現計画	[Redacted]																													
オプション1	[Redacted]																													
オプション2	[Redacted]																													
オプション3	[Redacted]																													
オプション4	[Redacted]																													

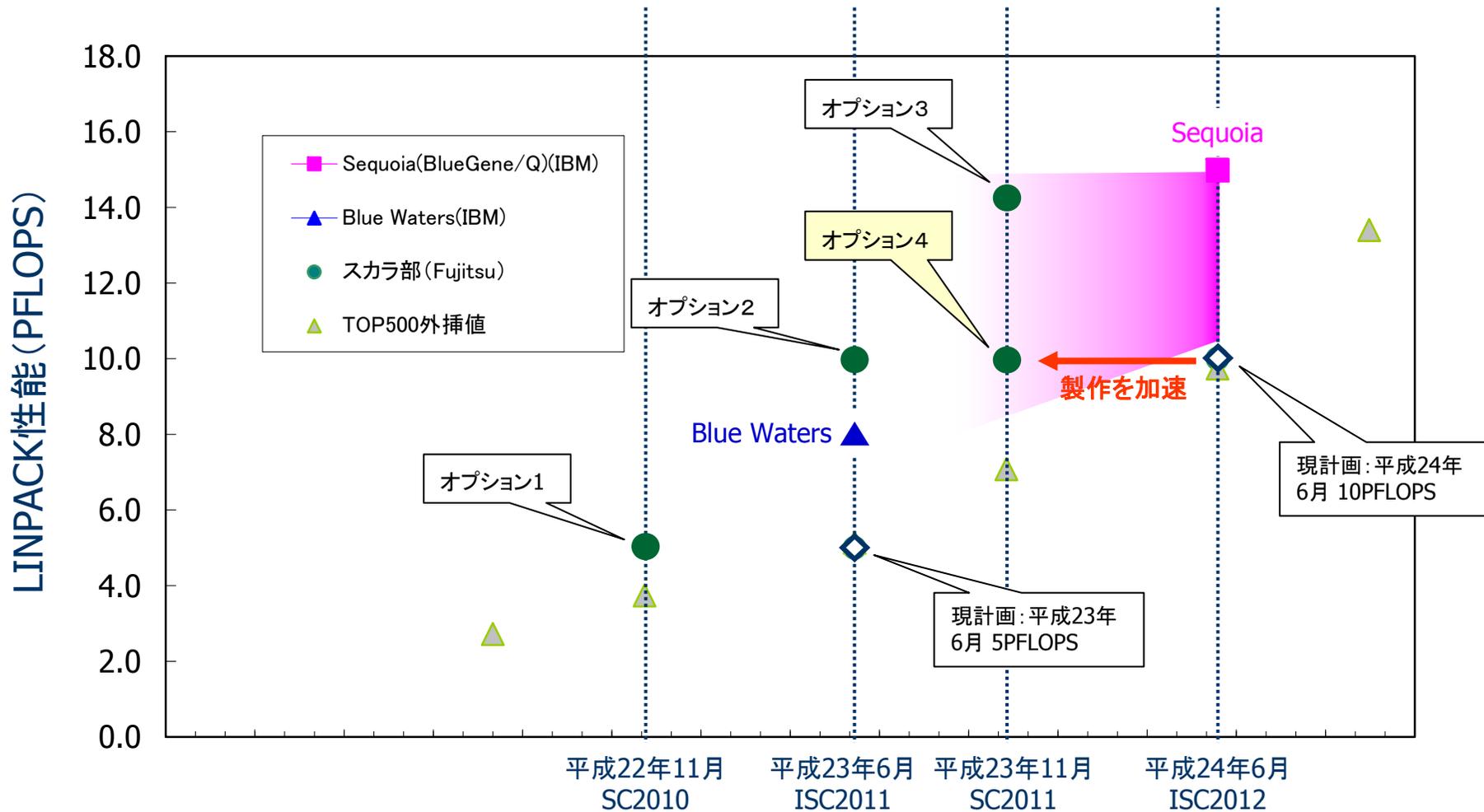
各オプションの検討結果

オプション1	オプション2	オプション3	オプション4
平成22年(2010年)11月, LINPACK 5PFLOPS規模のシステム	平成23年(2011年)6月, LINPACK 10PFLOPSのシステム	平成23年(2011年)11月に LINPACK 14.3PFLOPSのシステム	平成23年(2011年)11月に LINPACK 10PFLOPSのシステム

各オプションの比較

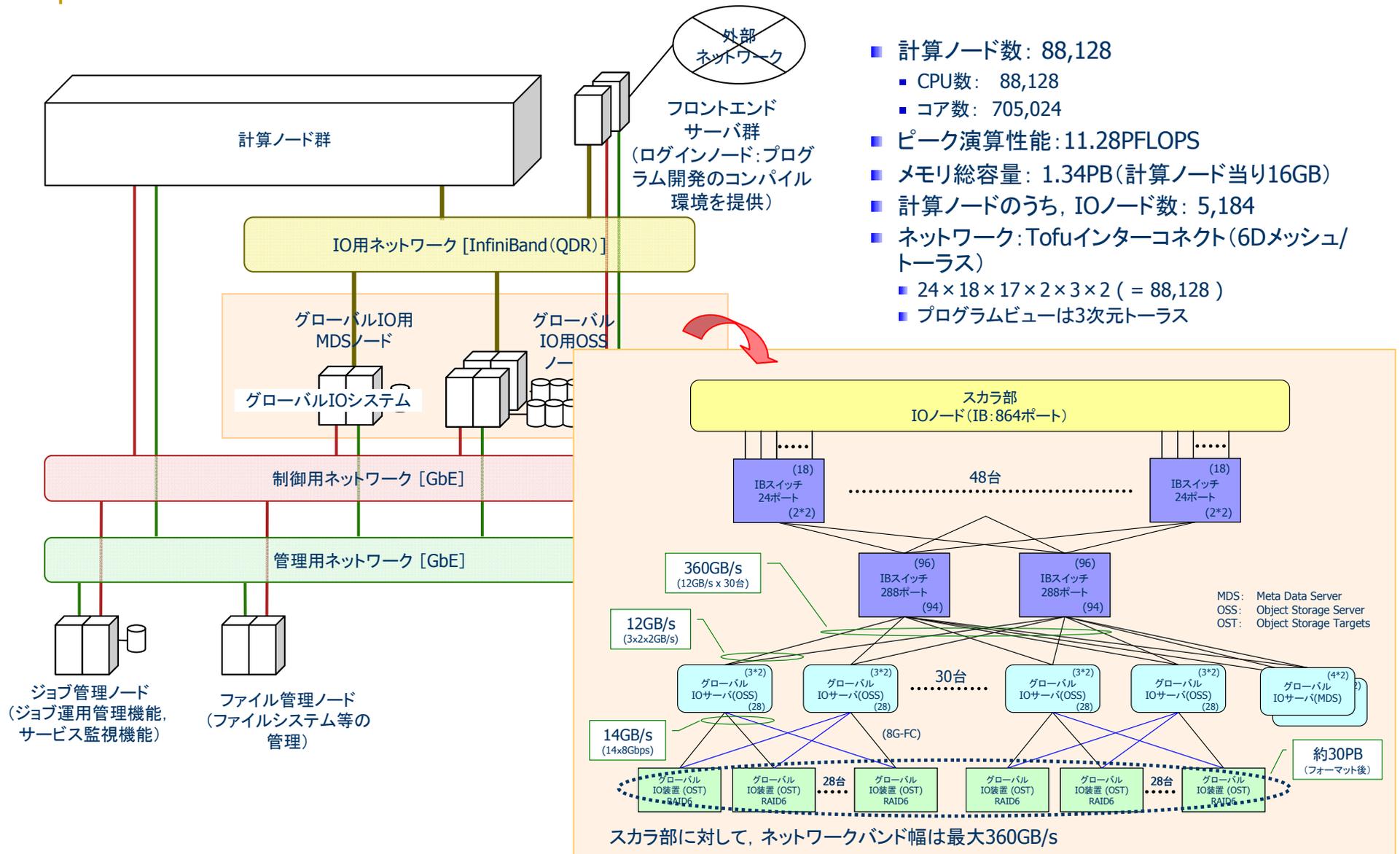
		オプション1	オプション2	オプション3	オプション4	現計画	
技術的なリスク		極めて大 (非現実的)	極めて大 (非現実的)	極めて大 (非現実的)	小		
目標LINPACK性能 (達成時期)		5PF (平成22年11月)	10PF (平成23年6月)	14+PF (平成23年11月)	10PF (平成23年11月)	10PF (平成24年6月)	
競争システム	システム名	—	Blue Waters(IBM)	Blue Waters, Sequoia(IBM)	Blue Waters, Sequoia(IBM)	Sequoia	
	推定LINPACK性能	3.8PF	8.2PF	8.2PF—15PF	8.2PF—15PF	15PF	
目標達成度	TOP500 ランキング第一位	◎	◎	○	△	×	
	達成時期:平成23年6月	○	○	×	×	×	
	HPC Award	Global HPL	○	○	○	△	×
		G-RandomAccess	△	×	×	×	×
		Global FFT	△	×	×	×	×
EP stream per system		○	○	○	○	×	
資金計画へのインパクト (見込み)	スカラ部増減の規模				約100億円		
	当初資金計画額 (1,154億円)との関係				当初資金計画の範囲内で可能		
	光熱水料の増減 (平成24年度以降)				スカラ部: 0億円 ベクトル部: 約6.8億円減		
共用開始時期		—	—	—	平成24年7月	平成24年9月	
建屋へのインパクト		—	—	—	計算機棟 スカラ部設置フロアは現設計どおり	—	

各オプションと競合システムとの比較



※Sequoiaを早期整備するとの情報によれば、平成23年11月におけるLINPACK性能は10-15PFLOPSの可能性もある。

新システム構成



- 計算ノード数: 88,128
 - CPU数: 88,128
 - コア数: 705,024
- ピーク演算性能: 11.28PFLOPS
- メモリ総容量: 1.34PB (計算ノード当り16GB)
- 計算ノードのうち、IOノード数: 5,184
- ネットワーク: Tofuインターコネクト (6Dメッシュ/トラス)
 - $24 \times 18 \times 17 \times 2 \times 3 \times 2$ (= 88,128)
 - プログラムビューは3次元トラス

新システム構成の目標性能

- 平成23年(2011年)11月までに, LINPACK性能 10PFLOPSを達成.
- LINPACK 10PFLOPSを達成した時点(平成23年11月)で, HPC Challenge Award 4項目において, 以下の性能を達成.
 - Global HPL 10PFLOPS
 - Global RandomAccess 200GUPS
 - Global FFT 180TFLOPS
 - EP Stream (triad) per system 3800TB/s

新システムとBlue Waters, Sequoiaとの比較

第1位

第2位

	新システム	Blue Waters	Sequoia
特徴	<ul style="list-style-type: none"> 高信頼性, 耐障害性に優れたシステム 多様なアプリケーションに対する高い実行効率 	<ul style="list-style-type: none"> 単一プロセッサの性能が最も高いが, 性能あたりのメモリ容量が少ない. 大域的な通信性能が高いシステム 	<ul style="list-style-type: none"> 性能あたりの低消費電力が低いシステム メモリ系が弱く, 実行効率が低い(性能あたりのメモリ容量及びメモリ帯域が小さい)
CPU数	88,128	38,912	98,304
システム性能 (ピーク性能)	11.3PFLOPS	10.06PFLOPS	20PFLOPS
総メモリ帯域	5.4PB/s	5.0PB/s	4.0PB/s
性能あたりのメモリ容量 (総メモリ容量)	0.125B/FLOPS(1.34PB)	0.055B/FLOPS(0.556PB)	0.08B/FLOPS (1.6PB)
総ネットワーク帯域	2.8PB/s	1.37PB/s	3.0PB/s
バイセクションバンド幅	49TB/s	171TB/s*	60TB/s
ディスク容量	41.2PB	23.3PB	50PB (他のシステムと共有)
IO性能	1.33TB/s	2.02TB/s	0.512TB/s

※ ノード構成から仮定したネットワークポロジ及び総ネットワーク帯域から推定
(<http://www.it.utah.edu/leadership/committees/papers/IBMinEducation.ppt>)

新システムの整備スケジュール

	2009年			2010年								2011年								2012年																				
	平成21年度				平成22年度								平成23年度								平成24年度																			
	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
【スカラ部】 現計画																																								
見直し案																																								
評価用システム																																								
現地システム																																								
ソフトウェア ・OS、ファイルシステム ・コンパイラ、ライブラリ等																																								

LSI製作, システム設置スケジュール

	2009年												2010年												2011年		
				平成21年度									平成22年度														
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3
スカラ部(ユニットA)																											
CPU (45nm)																											
ICC (65nm)																											
小規模システム (工場内)																											
大規模展開																											

開発状況(スカラ部)

- LSI開発

- CPU

- [Redacted]

- ICC(インタコネク用LSI)

- [Redacted]

- [Redacted]

- 45nm半導体プロセス

- [Redacted]

- 実装関連

- [Redacted]

- システムソフトウェア

- [Redacted]
 - [Redacted]

【参考】米国の最新整備状況

プロジェクト		2009	2010	2011	2012	2013
開発計画	次世代スパコン 広範な分野の研究基盤を提供し、我が国の競争力強化等に資する			5.6PF(スカラ部) 1PF(ベクトル部)	11.3PF(スカラ部) 3.1PF(ベクトル部)	
	国防省(DOD) -HPCS計画※- ※High Productivity Computing System計画 国家安全保障及び産業界のために寄与する経済性・運用性に優れた次世代スパコンの開発を目指す		計画が加速	Cascade(Cray)・PERCS(IBM) 2PF(4PF超の拡張性) から10PFへと計画が加速 2PF→10PF (実行性能)		
	米国科学財団(NSF) -Cyber Infrastructure計画- コンピュータ、データ及びネットワーク技術を活用した科学・工学研究の為の環境を整備する			Blue Waters 10PF(IBM)	※HPQS計画で完成したマシン(PERCS)をイリノイ大学に設置	
	エネルギー省(DOE) -NLCF計画※- ※National Leadership Computing Facility計画 大規模かつ高度なプロジェクトにフォーカスし、科学分野のブレークスルー及び持続的技術革新を創出する				Cascade 20PF(Cray) BlueGene/Q 20PF(IBM)	※開発状況により2011-2013の間で導入(オークリッジ国立研究所) ※開発状況により2011-2013の間で導入(アルゴンヌ国立研究所)
	エネルギー省(DOE) -ASC計画※- ※Advanced Simulation and Computing計画 核兵器のパフォーマンスを分析する等、ターゲットを絞って世界最速を目指す			2009/4/29の会議で発表	10~20PF(IBM) Sequoia 20PF(IBM) ※ローレンス・リバモア国立研究所に設置	2009/2/3 DOEプレス発表 ¹⁾
	米国航空宇宙局(NASA) -Pleiades計画- 航空・宇宙分野のモデリングとシミュレーションを行う		Pleiades 1PF(SGI)			Pleiades 10PF(SGI)

(注) 数字はピーク性能

※上記点線は世界スパコンランキングが実施される6月、11月を表す

1) <http://nnsa.energy.gov/2277.htm>