

FOCUS スーパーコンピュータシステムにおける 並列課金インセンティブの効果

西川武志^{†1}

スーパーコンピュータ「京」の産業利用推進を目的として計算科学振興財団（FOCUS）が運営している FOCUS スーパーコンピュータシステムでは並列度向上のためジョブの並列度が大きくなると課金単価を割引くというインセンティブを設けている。2011 年度の運用初年度の結果から利用者にはノード内並列に留まる階層と積極的にノード間並列を行う二つの階層に明確に分かれていることが判明し、2012 年度から 2014 年度までの三年度は 16 ノードを境界に二つの階層を意識した並列度向上を目指す課金インセンティブを毎年度更新してきた。その結果、明確に課金インセンティブによる並列度向上が見られ効果があったことを報告する。

Effects of parallelism charge incentive on the FOCUS supercomputer system

TAKESHI NISHIKAWA^{†1}

The FOCUS supercomputer system is to encourage the utilization of the K computer, has provided the incentive discounting unit price parallelism of the job for increasing the parallelism. The operation results of the FOCUS super computer system in the 2011 fiscal year that was the first year for operation showed that users are divided clearly into two hierarchy to carry out a program. One is running a program in parallel inside one node, the other is running program in parallel among nodes. The billing incentive from fiscal 2012 to fiscal 2014 has been updated every year to aim of increasing parallelism of a job was conscious of two of the hierarchy in the 16-nodes border. As a result, we report that the change of incentive was affected improvement of the degree of parallelism.

1. はじめに

公益財団法人計算科学振興財団は（Foundation for Computational Science: FOCUS）次世代スーパーコンピュータ（現：スーパーコンピュータ「京」）を活用した研究開発や産業利用を推進するとともに、広く普及啓発を行うことにより、計算科学分野の振興と産業経済の発展に寄与するため兵庫県、神戸市、神戸商工会議所が出資して平成 21 年 1 月 22 日に設立された。平成 21 年の設立以降、地元の主要大学や経済界との密接な連携・協力のもと、産業利用の推進や、普及・啓発に係る事業を着実に実施し平成 26 年度からは内閣府認可の公益財団法人となった。

平成 22 年度までの 3 年度の活動では、企業コンサルティングにより潜在的な HPC 利用需要を掘り起こすとともに、企業のシミュレーション技術の高度化を支援してきた。実践スクールや研究会などを通じて、企業技術者の HPC 技術の習得・向上に貢献し、コミュニティ形成に向けた基盤の構築を行ってきた。平成 23 年度からは神戸市ポートアイランドの理化学研究所計算科学研究機構（京コンピュータ施設）に隣接して建設された設置された計算科学センタービル（図-1）の 1・2 階に設置された高度計算科学研究支援センターを活動のインフラとして、産業界向けエントリ-スーパーコンピュータシステム「FOCUS スパコン（ABC

システム：合計理論ピーク性能 26.8TFLOPS)」、実習室、セミナー室、端末理容室、展示コーナー、貸研究室等を活用し、毎月の各種セミナー、FOCUS スパコン利用サービス等、全国を活動範囲にスーパーコンピュータ京の産業利用推進活動を始めた。平成 24 年 9 月 28 日からは「京」ネットワークと光ファイバーで直結された HPCI アクセスポイント神戸が高度計算科学研究支援センターで共用開始された。平成 26 年 1 月からは FOCUS スパコンシステムに DE システム（合計理論ピーク性能:245.3TFLOPS、汎用 CPU51.2TFLOPS、コプロセッサ 194.1TFLOPS）を追加し、4 月からは 272TFLOPS のシステムとして一体化運用を開始した。

FOCUS スパコンは、産業界利用者を利用対象者とするため、CPU に Intel x86-64 アーキテクチャ、OS に Linux を採用し、Intel x86-64 アーキテクチャ向けに提供されている商用アプリケーションの円滑な利用環境提供を行っている。同時に「京」向けに開発チューニングされた各種アプリケーションの Intel x86-64 アーキテクチャ版を事前にインストールし、1 つの計算ジョブで数万並列以上の利用が望ましい「京」と数十並列以下の計算ジョブが多い産業利用とのギャップを埋めるための環境整備を行っている。FOCUS スパコンシステムは「京」の利用に向けて、いわば自動車教習所のような役割を果たし「京」利用の支援や推進を図っている。

^{†1} (公財)計算科学振興財団
Foundation for Computational Science

今回、産業向けの FOCUS スーパーコンピュータシステムのジョブ統計の一部を示すが、大学の大型計算センターのジョブ統計が公表されること[1]は以前からあっても産業界の利用ジョブ統計の公表というのはこれまでほとんど存在しなかったため 2011 年度の運用開始当初は手探り状態であった。2011 年度の FOCUS スパコン運用初年度の結果は 8 ノード (96 並列) 以内が計算資源利用量の 47% を占め、1 ノード内並列 (12 並列以下) に留まる利用が 15% にも達した。一方で 32 ノード超 (384 並列超) の利用も 11% と決して少なくないものであった。ノード内並列に留まる階層と積極的にノード間並列を行う二つの階層に明確に分かれていることが判明したため、翌 2012 年度以降は 16~32 ノードを境界に二つの階層を意識した並列度向上を目指す課金インセンティブを毎年度更新した。本報告では 2011 年度から 2014 年度までの四年度の課金インセンティブ設定の効果、複数年度継続利用者において明確に課金インセンティブによる並列度向上が見られ効果があったこと、一方で商用ソフトウェアを主に利用する利用者では課金インセンティブよりソフトウェアライセンス利用料によって並列度向上が抑制されていることも明らかとなったことを報告する。

2. FOCUS スパコンシステム

FOCUS スパコンシステムの概要と利用概況について述べる。

2.1 FOCUS スパコン概要

FOCUS スパコンシステムの概要は次の通りである。

(1) A システム (224 ノード)

高並列化環境 (40Gbps QDR-Infiniband 接続)

CPU : Xeon L5640 2.26 GHz 6 コア×2 108GFLOPS, RAM : 48GB, HDD : 500GB

(2) B システム (2 ノード)

プリポスト環境 (40Gbps QDR-Infiniband 接続)

CPU : Xeon X7520 1.86 GHz 4 コア×4 119GFLOPS, RAM : 512GB, HDD : 1200GB

(3) C システム (22 ノード)

ノード内並列実行環境 (Gigabit Ethernet 接続)

CPU : Xeon L5640 2.26 GHz 6 コア×2 108GFLOPS, RAM : 48GB, HDD : 500GB

(4) D システム (80 ノード)

高並列化環境 (56Gbps FDR-Infiniband 接続)

CPU : Xeon E5-2670 2.5 GHz 10 コア×2 400GFLOPS, RAM : 64GB, HDD : 6000GB

(5) E システム (48 ノード)

大容量メモリ、高並列化・コプロセッサ利用環境 (56Gbps FDR-Infiniband 接続)

CPU : Xeon E5-2670 2.5 GHz 10 コア×2 400GFLOPS, RAM : 128GB, HDD : 2000GB, コプロセッサ Xeon Phi

P5110 60 コア×4 4044GFLOPS, ノード内合計理論演算性能 4444GFLOPS

(6) G システム (4 ノード)

デバッグ環境 (10Gbps Ethernet 接続)

CPU : Xeon E5-2640 2.5 GHz 6 コア×2 240GFLOPS, RAM : 64GB, HDD : 1000GB, コプロセッサ Xeon Phi P5110 60 コア×1 1011GFLOPS, ノード内合計理論演算性能 1251GFLOPS

(7) NAS システム (ホーム領域 360TB)

ホーム領域 : 180TB NAS×2 利用課題毎の独立した専用ファイルシステムを提供

NAS 構成 (10Gbps Ethernet×2 接続)

CPU : Xeon E5-2609 2.4GHz 4 コア×1, RAM : 4GB, HDD : 3TB 7200rpm NL-SAS×60 RAID6 (8 Data + 2 Parity) 構成 実行容量 144TB

(8) 大容量分散高速ストレージシステム (ホーム兼ワーク領域 1640TB)

ホーム兼ワーク領域 : 物理容量 1640TB, 実効容量 1250TB Lustre ファイルシステム, 利用課題全体で共通のファイルシステムを提供

56Gbps FDR-Infiniband 接続

SFA12K-20 + SS8460 5 Enclosure Model 1 セット

ExaScaler OST : 4TB 7200rpm SAS ×410 RAID6 (8d+2p)

ExaScaler MDS : (Active/Active) ×2

ExaScaler MDT EF3015 : 450GB 15Krpm SAS×12 RAID10 (5+5)

ExaScaler OSS : (Active/Active) ×4

以上より OST×40, 10 OST/OSS 構成を取っている。

ストレージサブシステム全体実効スループットは 14GB/s 以上のカタログ値で実測は 11GB/s

(9) ソフトウェアシステム

学会 OS : Linux (Red Hat/Cent OS 6 系)

開発環境 : GCC, MPICH2, OpenMPI, インテル® Parallel Studio XE 2015 Cluster Edition for Linux

ジョブ管理システム : 京のジョブ管理システムとコマンドラインオプション互換性をもたせたラッパーを作成した SLURM

従量課金アプリケーション : Gaussian 09, Parallel CONFLEX, MIZUHO Biostation

自社開発や商用ソフトウェアが利用課題毎に独立したファイルシステムに自由にインストール可能

詳細は <http://www.j-focus.or.jp/> を参照していただきたい。

2.2 FOCUS スパコン課金制度

図 1 に FOCUS スパコン A システムの課金単価を示す。

2011 年度はスーパーコンピュータ「京」が 12 ノード 96 並列で Tofu ネットワークの 1 単位として構成され 1 ラックは 8 単位 96 ノード 768 並列から構成されていること[2]から FOCUS スパコン A システムは 1 ノード 12 コアからなり

16 ノードで 192 並列 32 ノードで 384 並列とそれぞれスーパーコンピュータ「京」の最小単位の 2 倍，1 ラックの半分を並列割引の境界と設定した。

2012 年度は，2011 年度の課金インセンティブが得られる並列度が利用者の並列度に見合わなかったことを反省しより細かくかつ，8 ノード 96 並列すなわちスーパーコンピュータ「京」の Tofu ネットワーク 1 単位分の並列度が向上するたびに基準から 10% ずつ割引くというシンプルかつスーパーコンピュータ「京」をより強く意識した課金区分とした。

2011年度		2013年度		2014年度	
ノード数	課金単価 [JPY]	ノード数	課金単価 [JPY]	ノード数	課金単価 [JPY]
1<n<=16	100	n=1	100	n=1	100
17<n<=32	80	n=2	97.5	2<n<=4	95
33<n	60	3<n<=4	95	5<n<=8	90
		5<n<=6	92.5	9<n<=16	85
		7<n<=8	90	17<n<=32	80
		9<n<=12	88	33<n<=48	75
		13<n<=16	86	49<n<=64	70
		17<n<=20	84	65<n<=80	65
		21<n<=24	82	81<n<=96	60
		25<n<=32	80	97<n<=112	55
		45<n<=64	70	113<n	50
		65<n<=84	60		
		85<n	50		

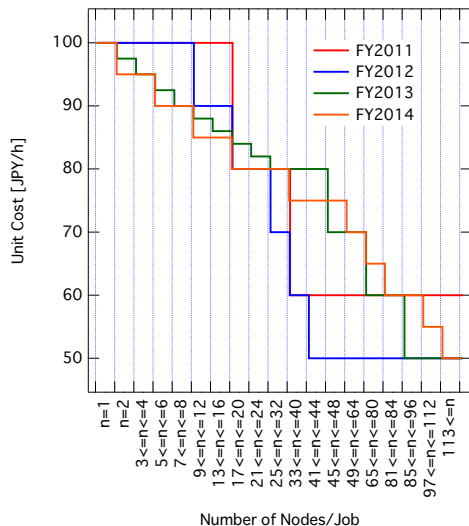


図 1 A システムの課金単価
Figure 1 Unit cost of the A system.

2013 年度は，スーパーコンピュータ「京」を意識した 8 ノード 96 並列を意識しつつも，より強く 1 ノード内並列からノード間並列へのステップアップを意識した。2 ノードから 16 ノード以下までの少しでもノード間並列度を上げれば課金単価を割引く一方で 16 ノードを超えると前年度と同じ並列度に停留すれば課金単価が上昇するように設定した。既に 192 並列を超えるような利用者にはさらなる並列化向上のため負のインセンティブを導入した。

2014 年度も 16 ノードを境界に低並列利用者層には少しでも並列度を上げれば課金単価を割引き、高並列利用者層には前年度と同程度に止まれば課金単価上昇の基本方針は維持しつつも，5%刻みに統一して課金体系をシンプルにし

た。

2.3 FOCUS スパコン利用概況

図 2 に FOCUS スパコンの利用状況を示す。2011 年度は 58 法人 64 課題，利用率 68%，2012 年度は 98 法人 108 課題，利用率 61%，2013 年度は 129 法人 148 課題，利用率 68%，2014 年度は 143 法人 168 課題，利用率 44%であった。平成 26 年 1 月からの D, E システムの増強により平成 26 年度は計算資源に余裕ができ利用者は 12 月から翌年 3 月までの繁忙期を除き，大規模並列ジョブをほとんど待ち時間無く実行できる環境が提供できるようになった。

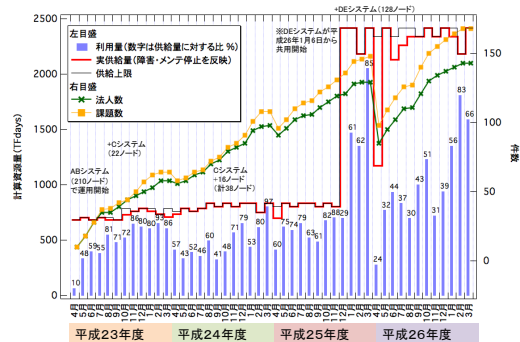


図 2 法人数，課題数，利用量
Figure 2 Number of corporations, number of projects and development of usage.

3. 課金インセンティブの効果

FOCUS スパコン A システムにおける課金インセンティブの効果を前年度と比較して 2012 年度対 2011 年度、2013 年度対 2012 年度、2014 年度対 2013 年度と前年度と比較して述べる。

3.1 2012 年度対 2011 年度

図 3 に 2011 年度と 2012 年度のジョブあたりの並列数区分と計算資源利用率を示す。並列数区分は 2012 年度を基準とした。図 3 の 41 ノード以上の区分が明らかに 2011 年度より増加しており課金インセンティブの効果が現れている。

3.2 2013 年度対 2012 年度

図 4 に 2012 年度と 2013 年度のジョブあたりの並列数区分と計算資源利用率を示す。図 4 の 12 ノード以下と 41 ノード以上の区分で明らかに 2012 年度より並列度が高い利用率が増加しており課金インセンティブの効果が現れている。特に 2012 年度に 41 ノード以上で課金単価が半額になり 43 ノードで 512 並列の利用に集中していた利用が 2013 年度では課金インセンティブの効果で 45 ノード以上 64 ノード以下と 85 ノード以上の区分に分散し並列度が向上している。

3.3 2014 年度対 2013 年度

図 5 に 2013 年度と 2014 年度のジョブあたりの並列数区分と計算資源利用率を示す。

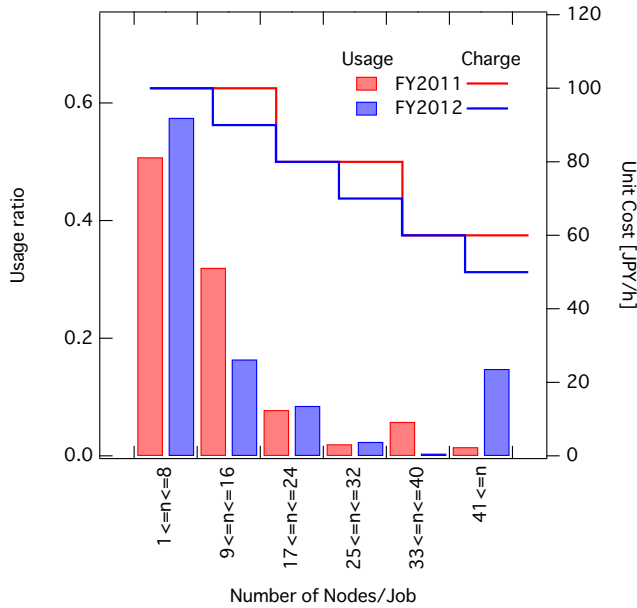


図 3 2011 年度と 2012 年度のジョブあたりの並列数区分と計算資源利用率

Figure 3 Computational resource usage and number of nodes per job in FY2011 and FY2012.

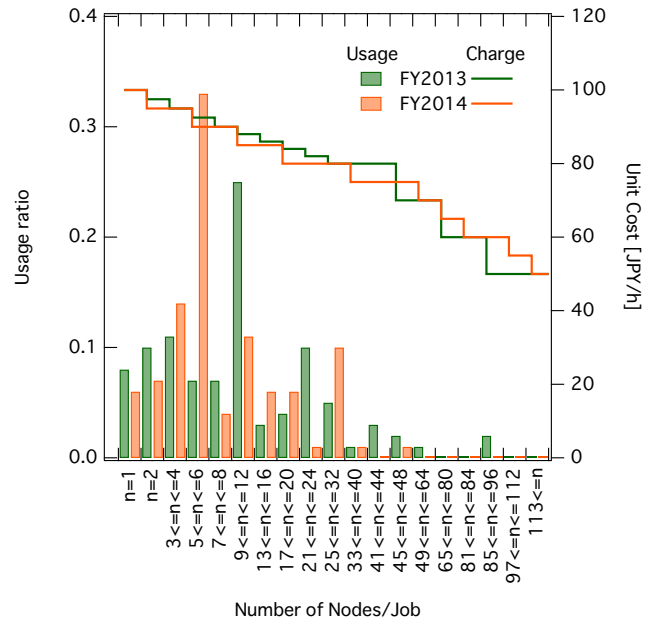


図 5 2013 年度と 2014 年度のジョブあたりの並列数区分と計算資源利用率

Figure 5 Computational resource usage and number of nodes per job in FY2013 and FY2014.

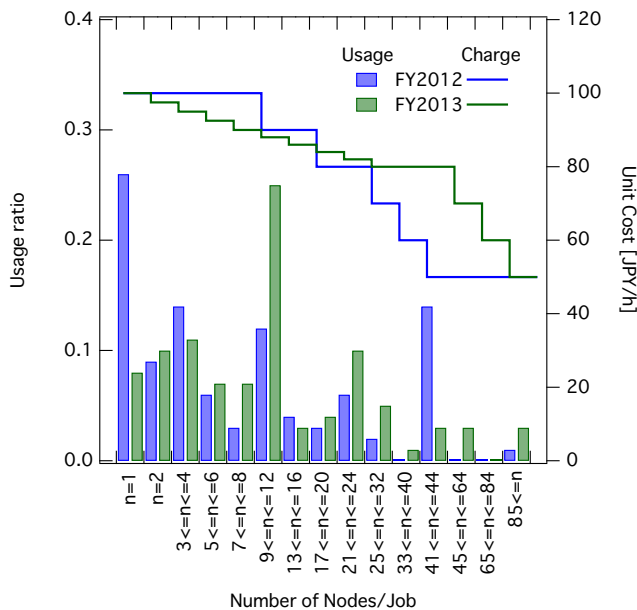


図 4 2012 年度と 2013 年度のジョブあたりの並列数区分と計算資源利用率

Figure 4 Computational resource usage and number of nodes per job in FY2012 and FY2013.

図 5 の 3 ノード以上 32 ノード以下の区分で明らかに 2013 年度より並列度が高い利用が増加している。一方 49 ノード以上の利用が 2014 年度は 2013 年度と比較するまでも無くほぼ消失している。ノード数が増えることによる課金単価割引よりも並列化効率の悪化による速度向上の停滞と課金単価上昇が相まって利用が消失したと考えられる。

4. おわりに

FOCUS スーパーコンピュータシステムにおける並列度向上のため低並列領域ではジョブの並列度が大きくなると課金単価を低い並列数から割引を適用し、高い並列度を利用している課題には前年度と同じ程度の並列度に止まれば値上げにあるという二面性を持つ課金インセンティブを設けることで継続的に並列度の向上に成功してきた。しかしながら 2014 年度には負の課金インセンティブ(ペナルティ)を課す領域では利用実績が消失するなど失敗もあった。FOCUS スーパーコンピュータシステムは運用が始まってから日が浅いため、課金インセンティブの他にも解決されていない問題点が多々ある。これらを様々な方々の御協力を仰ぎつつ、少しでも使いやすくするための改良を加えていくつもりである。運用に関する要望や意見を、是非 request@j-focus.or.jp までお寄せいただきたい。

謝辞 FOCUS スーパーコンピュータシステムの運用に際して富士通株式会社、株式会社アルグラフィックスおよび計算科学振興財団の業務・運用グループ各位に、謹んで感謝の意を表する。

参考文献

- 1) 石田晴久,高橋延匡,斎藤五郎:東京大学大型計算センターにおけるジョブ統計,情報処理,Vol.13,No.10,pp.714-719 (1972).
- 2) 黒川原佳,庄司文由:スーパーコンピュータ「京」システム概要,情報処理,Vol.53,No.8,pp.759-766 (2012).

正誤表

訂正箇所	誤	正
p.1 左 6 行	平成 21 年	平成 20 年
p.1 左 7 行	平成 21 年	平成 20 年
p.1 右 2 行	端末理容室	端末利用室