

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 V

西川武志^{†1}

概要: これまで FOCUS スパコンの Intel Xeon 搭載システムは A (Westmere-EP), D, E (Ivy-Bridge), F, H (Broadwell-EP) の 3 つのアーキテクチャ、5 つの CPU 性能、メモリ容量、ネットワーク・ストレージ機器構成の異なるシステムが存在する状況で課金インセンティブを適切に設定することで特定のシステムに大きく利用が偏ることが無いような運用を実現していることを報告した。今回は個別の利用者のジョブ利用傾向について詳細な報告、並列度の推移、計算資源利用量、システム間の遷移などについて報告する。

キーワード: インセンティブ設計、計算センター運用、運用統計、並列度向上

1. はじめに

これまでの報告[1-4]では計算科学振興財団 (FOCUS) が運用する産業界向けエントリースーパーコンピュータシステム「FOCUS スパコン」に関して 1 ジョブで複数のノードを確保して使うとノード課金時間単価を割引くという並列課金インセンティブを毎年度変化させて設定することでジョブ実行時の並列度を増加させることに成功してきた。

前回の報告では図 1 に示すように利用回数が 2 回以下と 6 回以上の利用課題において課金インセンティブの効果が小さいことが明らかとなった。リピート回数が少ないと並列ノード数も少なくリピート回数が増えるにつれて平均ノード数が増大すると予想し、5 回までのリピート数の利用課題はそのとおりであったが 6 回以上のリピート数の利用課題は並列課金インセンティブ以外の動機で FOCUS スパコンシステムを利用していることがヒアリングで後日判明した。

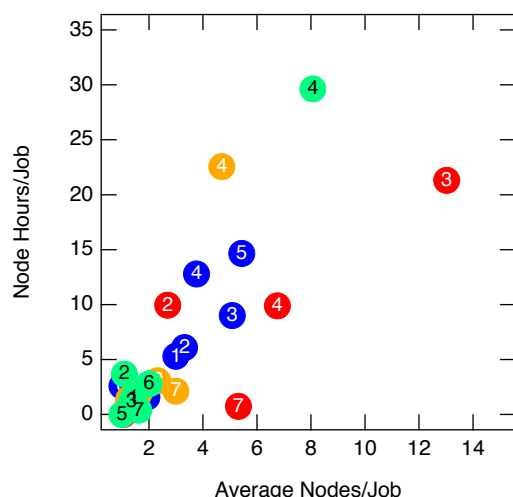


図 1 2017 年度 (2 月末まで) の A (赤), D (青), F (橙), H (緑) システムの全課題のジョブあたり平均ノード数に対するジョブあたり利用ノード時間。プロットの数値は 2011 年度から 2017 年度までの課題毎のリピート回数。

またこれまでの報告では利用課題には複数の利用者が含まれており、例えば、利用課題として 7 回リピートしていても個別の利用者には 1 回しか使っていないものも 7 回リピートの課題のジョブとして前回の報告では算入していた。

本稿では個々の利用者のリピート回数に着目して年度毎の詳細な利用推移を調査し、課金インセンティブが個々の利用者にどのような利用傾向をもたらしているかを報告する。

2. FOCUS スパコンシステム

FOCUS スパコンシステムの概要についてはこれまでの報告[3,4]に A, D, E, F, H システムについて概要を述べているが、2018 年度に F システム, V システムを増設したこともあり、あらためて A, D, F, H, V の各システムの基本仕様を述べる。

2.1 FOCUS スパコン A, D, F, H, V システム概要

FOCUS スパコン A, D, F, H 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-linfiniband 接続)

CPU : Xeon L5640 (Westmere-EP) 2.26 GHz 6 コア×2
108GFLOPS, RAM : 48GB, HDD : 500GB

(2) D システム (80 ノード)

高並列化環境 (56Gbps FDR-linfiniband 接続)

CPU : Xeon E5-2670 v2 (Ivy-Bridge) 2.5 GHz 10 コア×2
400GFLOPS, RAM : 64GB, HDD : 6000GB

(3) F システム (60+2 ノード)

高並列化環境 (56Gbps FDR-linfiniband 接続)

CPU : Xeon E5-2698 v4 (Broadwell) 2.2 GHz 20 コア×2
1152GFLOPS, RAM : 128GB, HDD : 6000GB

2 ノードには PCI 版 NVIDIA Tesla P100 をそれぞれ 1 基搭載

(4) H システム (136 ノード)

高密度高並列化環境 (34 ノード/3U シャーシ, シャーシ間 40Gbps Ethernet ×16 シャーシ内ノード間 10Gbps Ethernet ×2 接続)

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

CPU : Xeon D-1541 (Broadwell) 2.1 GHz 8 コア × 1
205GFLOPS, RAM : 64GB, SSD : 512GB

(5) V システム (62 ノード)

ベクトルエンジン環境 (56Gbps FDR-Infiniband 接続)

CPU : Xeon Gold 6148 (Skylake) 2.4 GHz 20 コア × 1
1024GFLOPS, RAM : 96GB, HDD : 240GB

PCI-Express 接続で NEC SX-Aurora TSUBASA Type 10B (周波数 1.4GHz 8 コア 2.15TFLOPS, メモリ帯域 1.22TB/s, HBM2 メモリ 48GB) をノードあたり 1 基搭載

このように FOCUS スパコンでは Westmere-EP, Ivy-Bridge, Broadwell, Skylake と 4 世代の Xeon が混在し, アクセラレータとして NVIDIA Tesla V100, NEC SX-Aurora TSUBASA も存在している. ノード性能と課金単価は 100 円/ノード時間の A システム 108GFLOPS から 500 円/ノード時間の F システム汎用 CPU1,152GFLOPS, F システムの 1,152GFLOPS に NVIDIA Tesla V100 の 4.7TFLOPS, V システムの汎用 CPU1024GFLOPS に NEC SX-Aurora TSUBASA 2.15TFLOPS と汎用性能 10.7 倍,アクセラレータを含んだ合計性能では 54 倍, 課金単価で 5 倍もの開きが存在する.

このため利用者がどのシステムを利用すれば良いか, コストパフォーマンスが自分が利用するアプリケーションでどうなのかについて FOCUS ではシステムの理論性能のみならず複数の実アプリケーション (Advance/ParallelWave, ABINIT-MP, ANSYS Fluent, iconCFD, NAMD, OpenFOAM, STAR-CCM+) の性能についてベンチマーク結果を開示している[5].

2.2 FOCUS スパコン課金制度とインセンティブ

FOCUS スパコンの課金制度には大きく二つあり従量制と予約制があるのは繰り返し述べてきた[6]. 従量制は常設キューを利用したノード数と経過時間の積にノード時間単価を乗じて利用課金する. 予約制は事前に 1 日単位、月単位、年度単位でノード数を確保し, 1 日単位の予約は従量制同一の時間単価で課金インセンティブを与えている.

図 2 に 2015 年度から 2017 年度 A, D, F, H システムの課金インセンティブ (従量制時間単価) を, 図 3 に 2015 年度から 2017 年度のコストパフォーマンス (GFLOPS/円) を示す.

これまでの課金インセンティブの設計方針は以下の通りである.

- 前年度に対し翌年度は同じノード数であれば課金単価が上昇させる.
- 利用を移行させたい移行元のシステムと移行先のシステムのコストパフォーマンスを 2 倍程度の比に設定する.
- 利用が特に進みにくいアクセラレータ Xeon Phi, Tesla P100, SX-Aurora 等の付加価値は設定しない. す

なわちアクセラレータ非搭載のノードの汎用 CPU と同じ単価にする.

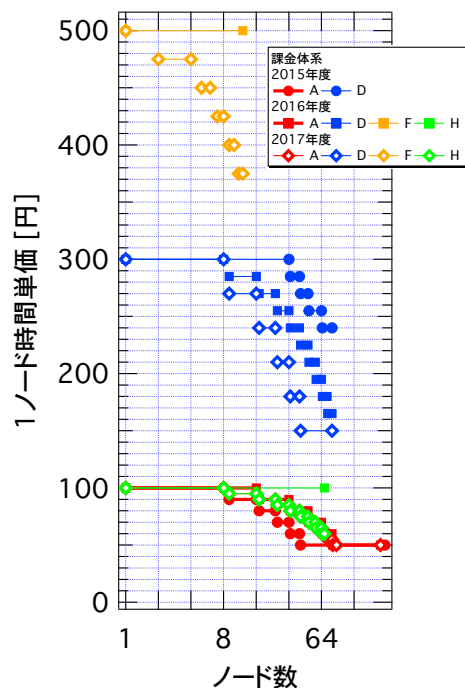


図 2 2015 年度から 2017 年度 A, D, F, H システムの課金インセンティブ (従量制時間単価)

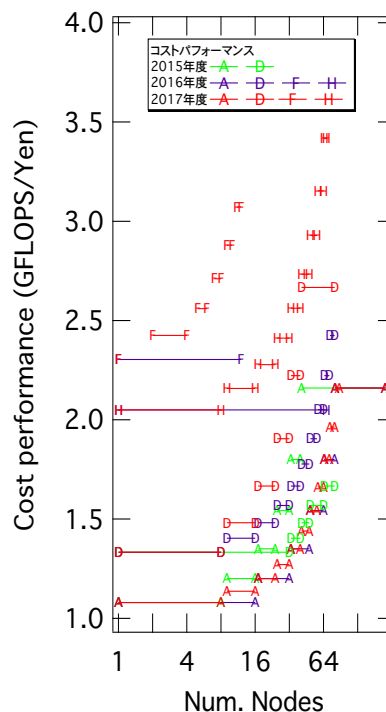


図 3 2015 年度から 2017 年度のコストパフォーマンス (GFLOPS/円)

3. 課金インセンティブの効果

3.1 システムの利用移行効果

表1に2011年度から2017年度における各システムで利用されたノード時間を、表2に各システムで利用された計算資源量(TF・h)を、表3同割合を示す。AよりもD、DよりもH、HよりもFのコストパフォーマンスが良くなるように課金インセンティブを設定した結果、ノード時間がD、F、Hに移行、特にAからDより多く移行したことが読み取れる。表1からノード時間としてはHがFの5倍強であるが、表2、3から計算資源量で見るとF、Hは同等の量となっていることがわかる。

表1 各システムで利用されたノード時間
(2011年度から2017年度)

ノード時間	A	D	F	H	合計
FY2011	4,118				4,118
FY2012	9,547				9,547
FY2013	12,918				12,918
FY2014	12,747	3,628			17,061
FY2015	13,067	5,558			19,207
FY2016	14,750	5,405	751	3,481	26,006
FY2017	5,450	3,463	905	5,186	16,152

表2 各システムで利用された計算資源量(TF・h)
(2011年度から2017年度)

TF・h	A	D	F	H	合計
FY2011	445				445
FY2012	1,031				1,031
FY2013	1,395				1,395
FY2014	1,377	1,451			3,102
FY2015	1,411	2,223			3,867
FY2016	1,593	2,162	865	714	5,981
FY2017	589	1,385	1,043	1,063	4,539

表3 各システムで利用された計算資源量割合
(2011年度から2017年度)

TF・h割合	A	D	F	H
FY2011	100%			
FY2012	100%			
FY2013	100%			
FY2014	44%	47%		
FY2015	36%	57%		
FY2016	27%	36%	14%	12%
FY2017	13%	31%	23%	23%

3.2 並列度向上効果

一度に利用するノード数が増えると課金単価を下げる、並列化速度向上率が悪化する傾向よりも課金単価が下がるようにすればコストパフォーマンスがシリアル実行や1ノード実行の場合より並列実行の場合が良くなるように課金イ

ンセンティブを設定している。結果、FOCUSスパコンシステム利用を繰り返すと並列度向上のインセンティブを繰り返し受けリピート回数が多くなるにつれて平均ノード数が増大すると期待する。

しかしながら課題毎に調べた結果、図1に示すように利用回数が2回以下と6回以上の利用課題において課金インセンティブの効果が小さいことが明らかとなった。

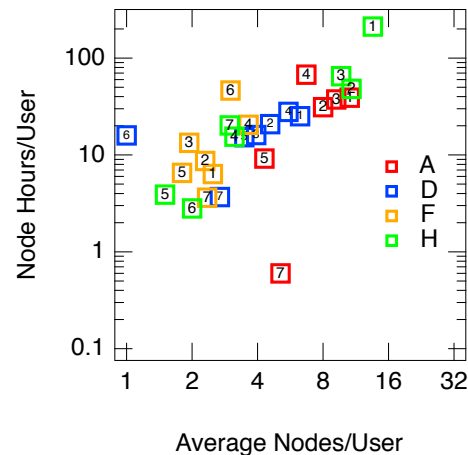


図4 利用者毎のリピート回数に応じた2017年度の平均ノード数と利用ノード時間の分布

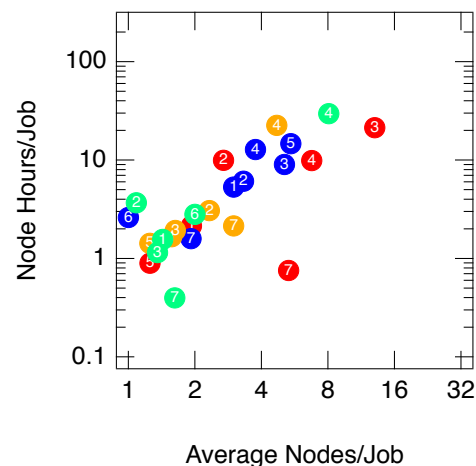


図5 課題毎のリピート回数に応じた2017年度の平均ノード数と利用ノード時間の分布

そこで今回は2017年度の平均ノード数と利用ノード時間の分布がどのようなになっているかリピート回数に応じて分類し利用課題に含まれる利用者毎でまとめたものを図4に、課題毎でまとめたものを図5に示す。

結局、利用者毎でまとめたものであってもリピート回数が5回以上では課金インセンティブの効果が並列度向上には現れていないことがわかった。6回以上のリピート数の利用課題および利用者は並列課金インセンティブ以外の動機、組織内計算需要の外部資源利用という目的でFOCUSスパコンシステムを利用していることがヒアリングで確

認できた。

4. 課金インセンティブの修正

並列度向上に対しては課金インセンティブの効果が小さかったが、システム利用移行に関しては大きく効き過ぎた感が否めない。表1、表2から2016年度から2017年にかけてAシステムの利用が急激に減少し約1/3になっている。これは課金インセンティブが効き過ぎたので2018年度の課金体系は図6に示すように修正し並列度向上も図った。

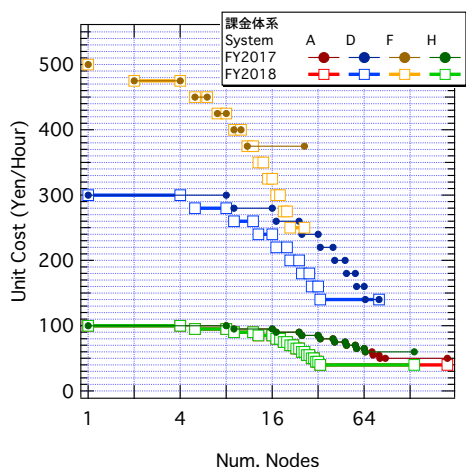


図6 2017年度と2018年度A, D, F, Hシステムの課金インセンティブ (従量制時間単価)

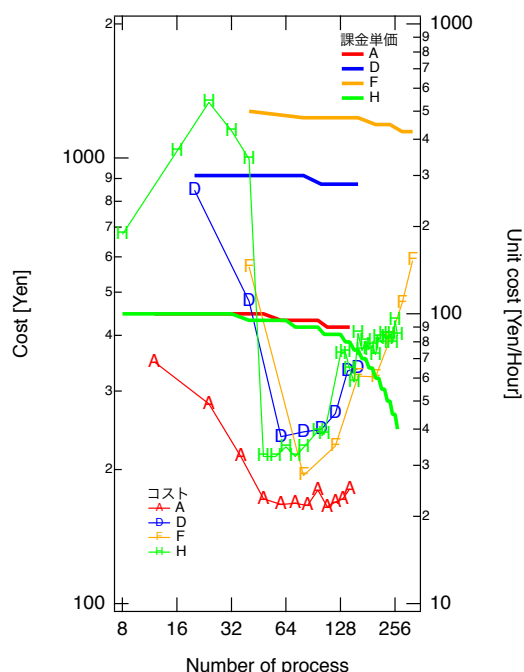


図7 2018年度の課金インセンティブの下での分子動力学法の同一ステップのシミュレーションを実行した時のコスト (Cost 円, 左軸) と並列数 (Numer of process) と課金単価 (Unit cost 円/時間, 右軸)

図7に2018年度の課金インセンティブの下での分子動力学法の同一ステップのシミュレーションを実行した時のコスト (Cost 円, 左軸) と並列数 (Numer of process) と課金単価 (Unit cost 円/時間, 右軸) でどのように変化するかを示す。AシステムのコストがD, F, Hシステムを利用するより常時低くなるようになっている。この修正により狙った通りにAシステムの利用がD, F, Hシステムから戻ってくるかどうかに関しては2018年度の利用について今後報告したい。

5. まとめ

本稿ではA, D, F, Hの複数のシステムの存在下で適切な課金インセンティブを設計することでシステム利用移行を促せることが明らかとなった。

一方で利用者の並列度向上に関しては2017年度の課金インセンティブの効果が小さいことが前回の報告と同様、明らかとなった。6回以上のリピート数の利用者は過去7年度繰り返してきた並列課金インセンティブ変更による並列度上昇の働きかけに反応しないため、あらたな課金インセンティブを検討する必要がある。

謝辞 FOCUS スーパーコンピュータシステムの運用や利用者の開拓に尽力されている計算科学振興財団の同僚と利用してくださっている利用者各位に、謹んで感謝の意を表する。

参考文献

- [1] 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, 研究報告ハイパフォーマンスコンピューティング (HPC), 2015-HPC-149(2),1-4 (2015-06-19).
- [2] 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 II, 西川 武志, 研究報告ハイパフォーマンスコンピューティング (HPC), 2016-HPC-157(10),1-5 (2016-12-14).
- [3] 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 III, 西川 武志, 研究報告ハイパフォーマンスコンピューティング (HPC), 2017-HPC-161(3),1-5 (2017-09-12).
- [4] 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 IV, 西川 武志, 研究報告ハイパフォーマンスコンピューティング (HPC), 2018-HPC-163(17),1-5 (2018-02-21).
- [5] FOCUS スパコンベンチマーク
<http://www.j-focus.jp/benchmark/>
- [6] FOCUS スーパーコンピュータシステム利用料金詳細,
<http://www.j-focus.or.jp/focus/fee.html>