計算科学振興財団(FOCUS)
スパコン産業利用セミナー2017 〜他社事例に学ぶ企業のシミュレーション技術活用〜
2017年7月25日@神戸商工会議所会館

OpenFOAMによる流体解析ベンチマークテスト

FOCUS・クラウド・スパコンでのチャネルおよびボックスファン流れ解析

今野　雅
オープンCAE学会V&V委員会
東京大学情報基盤センター客員研究員
株式会社OCAEL

Copyright © 2017 OCAEL All rights reserved.

オープンCAE学会共通OpenFOAMベンチマーク

- **大学のスーパーコンピュータ**
  - ✓ 近年は産業利用など教育・公共機関以外でも利用可能なシステム有り
  - ✓ 通常，課題審査が必要．通常1ヶ月～1年単位での課金
- **クラウドサービス**
  - ✓ 誰でも課題の審査無く利用可能
  - ✓ 使った分だけ課金(分または時間単位)
- **産業界専用のスーパーコンピュータFOCUS**
  - ✓ 法人の場合，課題の審査無く利用可能
  - ✓ 1円単位で使った分だけ課金．ただし1年毎にアカウント発行料も必要

- **オープンCAE学会V&V委員会で，チャネル流れによる共通OpenFOAMベンチマークを作成し，解析速度，並列化効率，対費用効果を比較した**
- **測定結果は学会のGitHubで公開している**

Copyright © 2017 OCAEL All rights reserved.

産応協ボックスファンベンチマーク

- **オープンCAE学会共通ベンチマーク問題のチャネル流れは，ベンチマークが容易に行えるよう，以下の特徴を重視して選択した**
  - ✓ 解析形状が矩形で，格子が構造格子かつ等分割のため，格子数の変更が容易
  - ✓ メッシュ生成に時間を要しない
  - ✓ 乱流モデルを使用せず，圧力と速度のみ解くので，「圧力線形ソルバの解析時間が支配的」という非圧縮性流体解析の特性を素直に示す
- **一方，チャネル流れは産業界で解析される流れ場とは乖離している欠点もある**

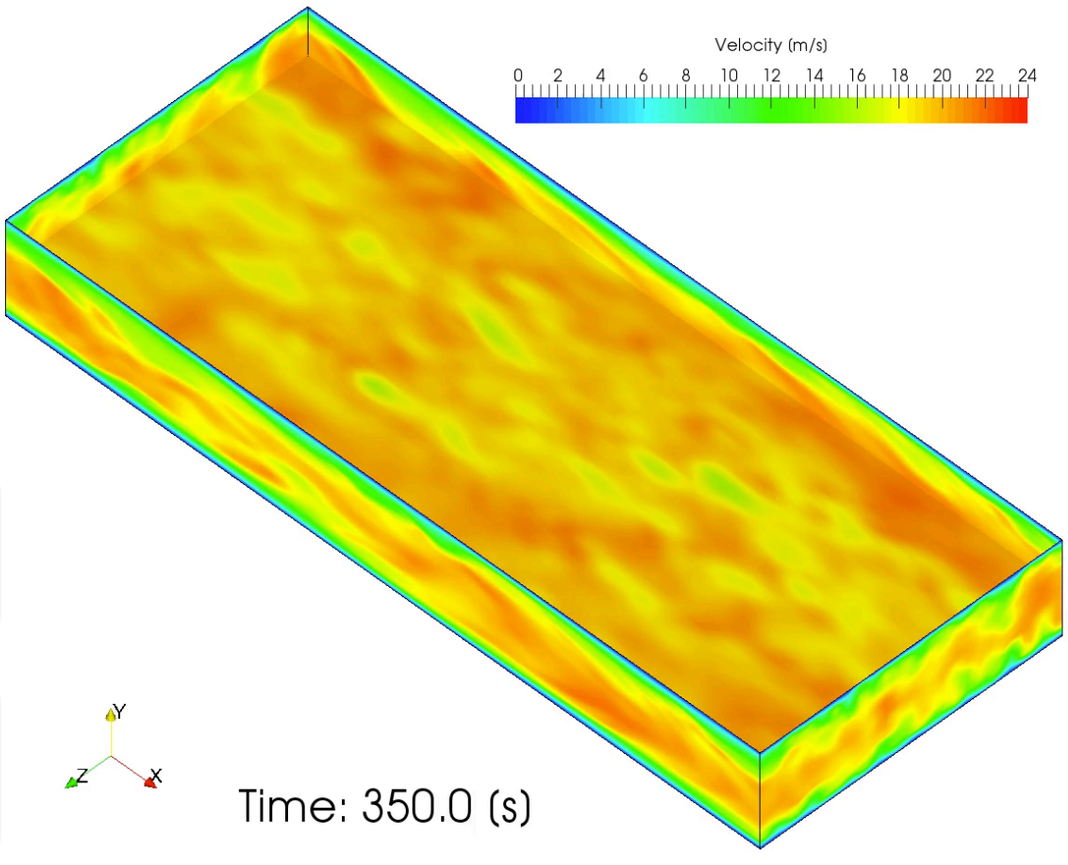
▼

- **スーパーコンピューティング技術産業応用協議会(通称：産応協)のHPCものづくりWSが作成した共通ベンチマークであるボックスファンは，乱流モデル，複雑形状，動的格子を扱う，より実務的なベンチマークである**
- **今回，産応協から共通メッシュや実験値を，FOCUSから計算機資源のご提供を受け，ボックスファンベンチマークをOpenFOAMで解析することにより，FOCUSと大学スパコンの解析速度，並列化効率，対費用効果を比較した**

Copyright © 2017 OCAEL All rights reserved.

オープンCAE学会チャネル流ベンチマーク

チャネル流れ ($Re_\tau = 110$)
格子数約3M
(Githubに24Mの結果も有り)



解析条件

$L_x \times L_y \times L_z = 5\pi \times 2 \times 2\pi$

$Re_\tau = u_\tau \, \delta / \mu = 110$ [-]

ここで

L_x, L_y, L_z: 各方向のチャネル幅 [m]

u_τ：壁面摩擦速度 [m/s]

δ ：チャネル半幅 [m] ($=L_y / 2$)

μ ：動粘性係数 [m^2/s^2]

主流方向(x): 一定の圧力勾配

主流方向(x), スパン方向(z): 周期境界

ソルバ: OpenFOAM-2.3.0, pimpleFoam

(GPUシステムではRapidCFD)

乱流モデル: 無し(laminar)

速度線型ソルバ: BiCG (前処理DILU)

圧力線型ソルバ: PCG (前処理DIC)

領域分割手法：scotch(周期境界面は同領域)

- **2〜51ステップのCPU時間(Execution time)から1時間あたりのステップ数を算出**

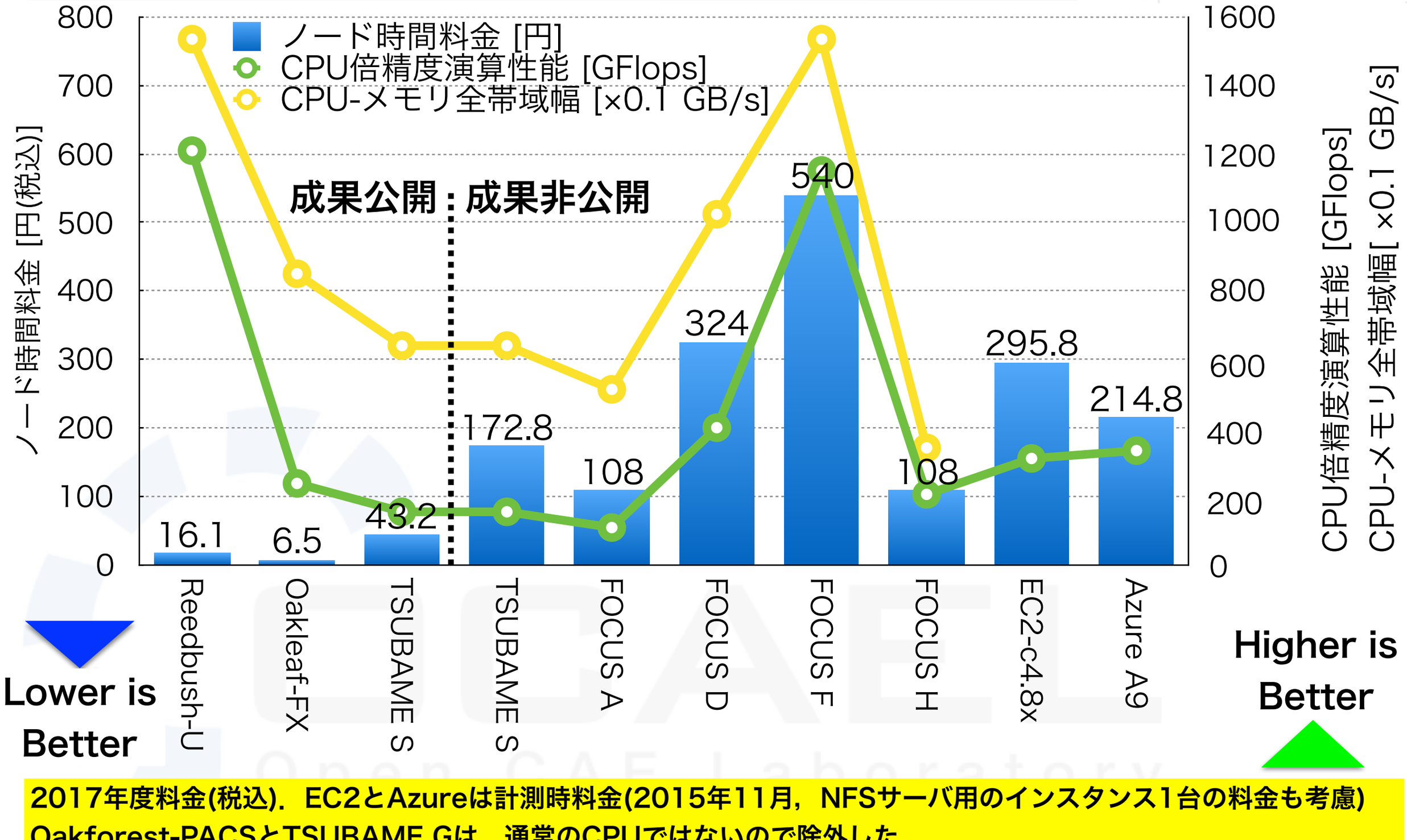
Copyright © 2017 OCAEL All rights reserved.

チャネル流ベンチマーク計測システム

機関	システム (略称)	CPU [GPU] (周波数[GHz])	CPU数 (コア)	倍精度性能 [GFlops]	メモリ[GiB] (帯域幅[GB/s])	インターコネクト (帯域幅[Gbps])
JCAH PC	Oakforest-PACS (OFP)	Intel Xeon Phi 7250, Knights Landing(1.4)	1(68)	3046	96(115.2), MCDRAM 16(490)	Intel Omni-Path (100)
東京大学	Reedbush-U (RBU)	Intel Xeon E5-2695 v4 (2.1-3.3)	2(36)	1210	256 (76.8×2)	Infiniband EDR(100)
	Oakleaf-FX FX (FX)	Fujitsu SPARC64 IXfx (1.848)	1(16)	237	32 (85)	Tofu(40)×双方向 ×10(4方向同時通信)
東京工業大学	TSUBAME 2.5 S	Intel Xeon E5-2670 (2.93-3.2)	2(12)	154	54 (32×2)	Infiniband QDR (40)×2
	TSUBAME 2.5 G	[nVIDIA Tesla K20X] (0.732)	3GPU	1310×3	6×3 (150×3)	
FOCUS	A	Xeon L5640(2.26)	2(12)	108	48 (25.6×2)	Infiniband QDR(40)
	D	E5-2670 v2(2.5)	2(20)	400	64 (51.2×2)	Infiniband FDR(56)
	F	E5-2698 v4(2.2)	2(40)	1152	128 (76.8×2)	
	H	D-154(2.1)	1(8)	205	64 (34.1)	10GbE(10)×2 or 4
Amazon	EC2 c4.8xlarge	Intel Xeon E5-2666 v3(2.9)	2(18)	310	60 (不明)	10GbE(10)
Microsoft	Azure A9	Intel Xeon E5-2670(2.6)	2(16)	333	112 (不明)	Infiniband QDR(40)

Copyright © 2017 OCAEL All rights reserved.

ノード時間料金・CPU演算性能・メモリ帯域幅



2017年度料金(税込)．EC2とAzureは計測時料金(2015年11月，NFSサーバ用のインスタンス1台の料金も考慮)
Oakforest-PACSとTSUBAME Gは，通常のCPUではないので除外した．

Copyright © 2017 OCAEL All rights reserved.

ノード時間料金の算出条件

システム	成果公開	ノード時間料金[円](※1)	ノード時間料金の算出条件
Oakforest-PACS	公開	10.7	最も高価となるグループコース(8ノード，企業)，利用期間1ヶ月間
Reedbush-U		16.1	最も高価となるグループコース(4ノード，企業)，利用期間1ヶ月間
Oakleaf-FX		6.5	最も高価となるグループコース(12ノード，企業)，利用期間1ヶ月間
TSUBAME S		43.2	成果非公開は成果公開の4倍の料金．GはSの1/2の料金．従量利用，最大計算時間：1時間，優先度：標準，実行時間ごとの係数：1の場合．学内・共同研究利用(ノード時間料金10円，40円)は除外
TSUBAME G		21.6	
TSUBAME S	非公開	172.8	
TSUBAME G		86.4	
FOCUS A		108	多ノード割引きを考慮
FOCUS D		324	
FOCUS F		540	
FOCUS H		108	
EC2-c4.8xlarge		295.8	計測時：2015年11月15日．リージョン：東京．NFSサーバ: 132.8円/h(c3.4xlarge)(※2)
Azure A9		214.8	計測時：2015年11月25～26日．リージョン：West US．NFSサーバ: 33.3円/h(D3)(※2)
(※1)税込．計測時明記以外は2017年度料金 (※2)NFSサーバ用のインスタンス1台の料金も考慮			

Copyright © 2017 OCAEL All rights reserved.

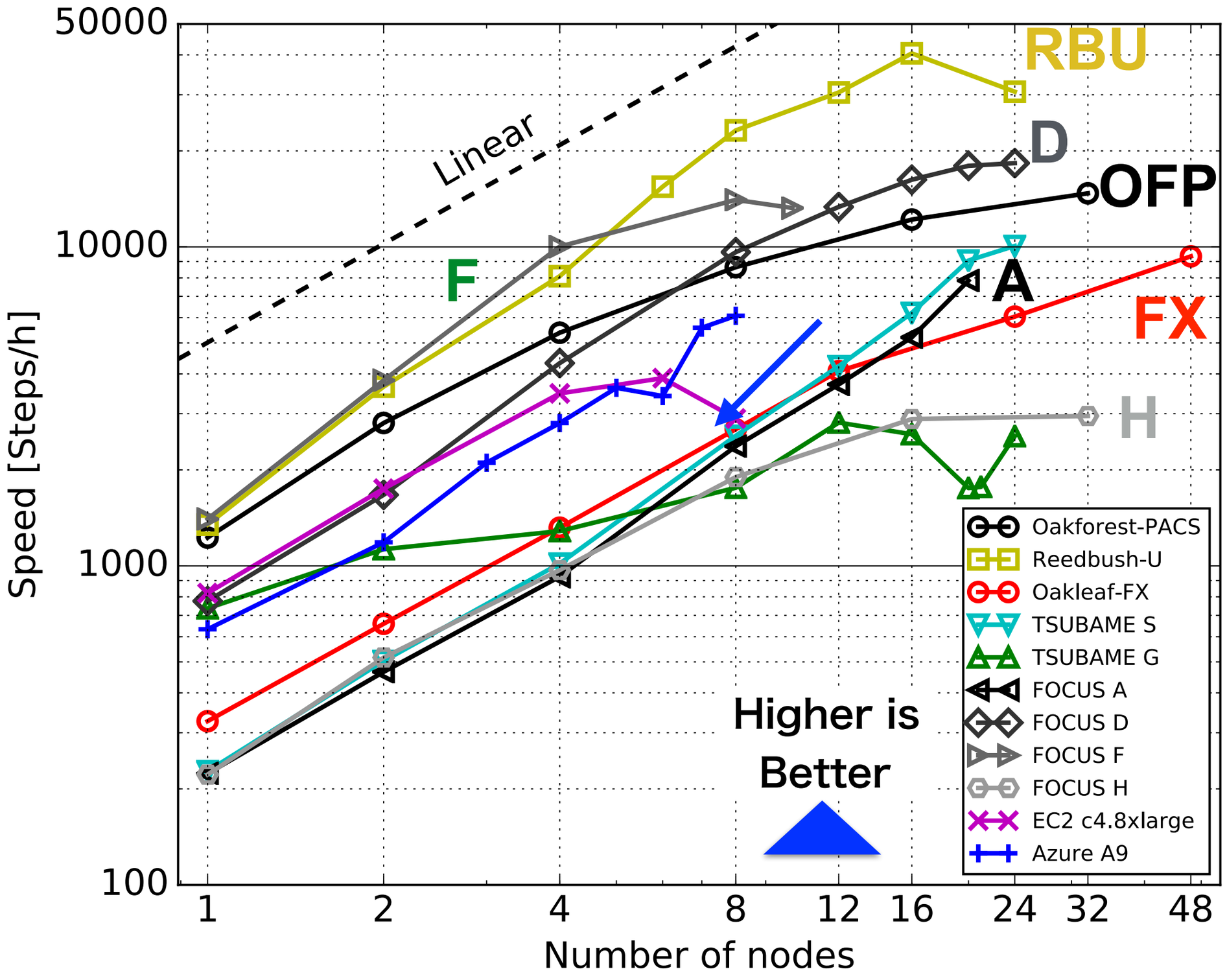
使用OpenFOAMバージョン・コンパイラ・MPI

システム	バージョン	コンパイラ(※1)	MPI
Oakforest-PACS (Flat, libhbm使用)	v1612+	Icc 2017.1 (KNL向け最適化)	Intel MPI 2017.1(※3)
Reedbush-U	2.3.0	Gcc-4.8.5	OpenMPI 1.8.3
Oakleaf-FX		FCC GM-1.2.1-09	FJMPI GM-1.2.1-09
TSUBAME S		Gcc-4.8.4	OpenMPI 1.6.5(※4)
TSUBAME G(GPU)	RapidCFD (※2)	nvcc (cuda-6.5)	OpenMPI 1.8.4(※5)
FOCUS A, D, F, H	2.3.0	Gcc 4.8.3	OpenMPI 1.6.5(※4)
EC2 c4.8xlarge		Gcc 4.8.5	OpenMPI 1.8.5(※6)
Azure A9		Gcc 4.8.3	Intel MPI 5.1.1(※7)

※1) 最適化フラグ: -O3，OakforestPACSのみ-O3 -DvectorMachine -xmic-avx512 ※2) rev: d3733257dee5fb9999b918f5c26a1493cebb603c ※3) unset KMP_AFFINIY;mpirun -env LD_PRELOAD libhbm.so -env HBM_SIZE 100 -env HBM_THRESHOLD 16 -env MPI_BUFFER_SIZE 1000000 -env I_MPI_PIN_PROCESSOR_EXCLUDE_LIST 0,1,68,69,136,137,204,205 -env KMP_HW_SUBSET 1T -env I_MPI_PIN_DOMAIN 4 ※4) mpirun -bind-to-core -mca btl openib,sm,self ※5) mpirun -bind-to core -mca btl openib,sm,self ※6) mpirun -bind-to core ※7) Azure A9のLinux OSでは，MPI通信にRDMAを使うためにはIntel MPIが必要

Copyright © 2017 OCAEL All rights reserved.

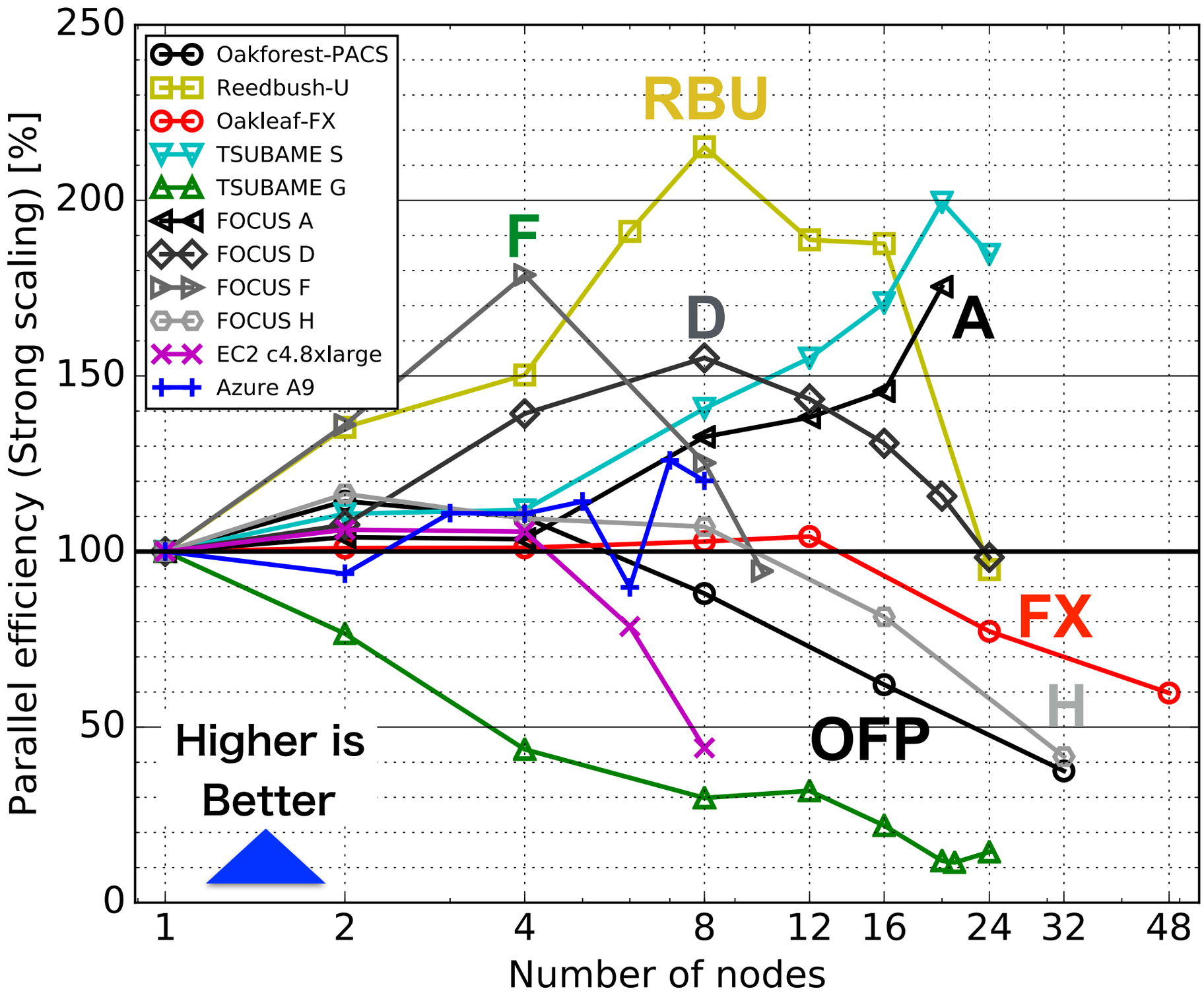
各システムの解析速度比較



- 最高速
  - ✓ 4ノードまで: FOCUS F
  - ✓ 8ノード以上: RBU
- 飽和ノード数
  - ✓ A: 不明
  - ✓ D: 不明
  - ✓ F: 8
  - ✓ H: 32
  - ✓ FX: 不明
  - ✓ RBU: 16
  - ✓ OFP: 不明

Copyright © 2017 OCAEL All rights reserved.

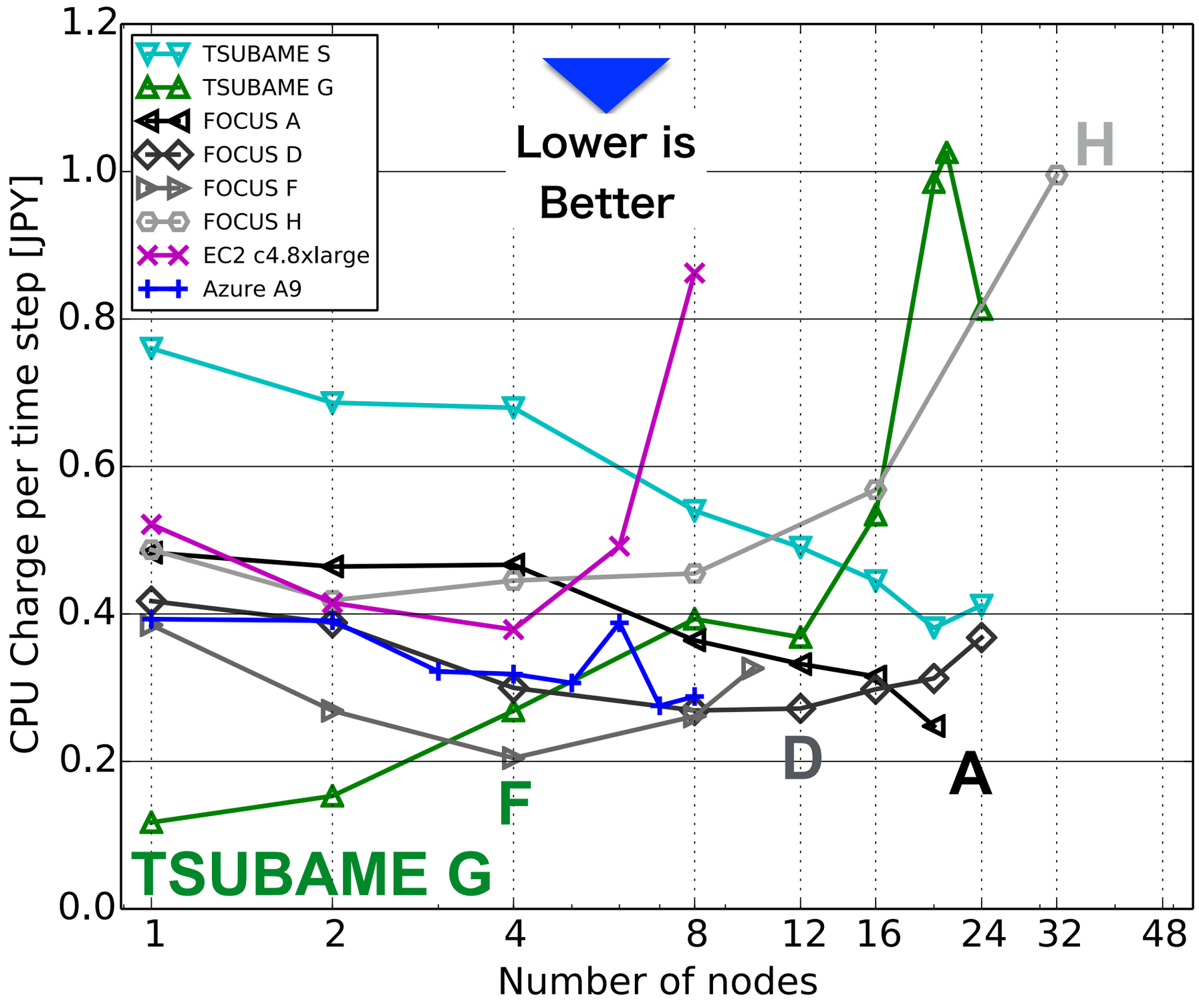
各システムのStrong scaling並列化効率比較



- Intel Xeon機 (RBU, FOCUS) はスーパリニア
- FXは12ノードまでほぼリニア
- OFPの並列化効率は悪い
- ピークのノード数
  - ✓ A 不明
  - ✓ D: 8
  - ✓ F: 4
  - ✓ H: 2
  - ✓ FX: 12
  - ✓ RBU: 8
  - ✓ OFP: 2

Copyright © 2017 OCAEL All rights reserved.

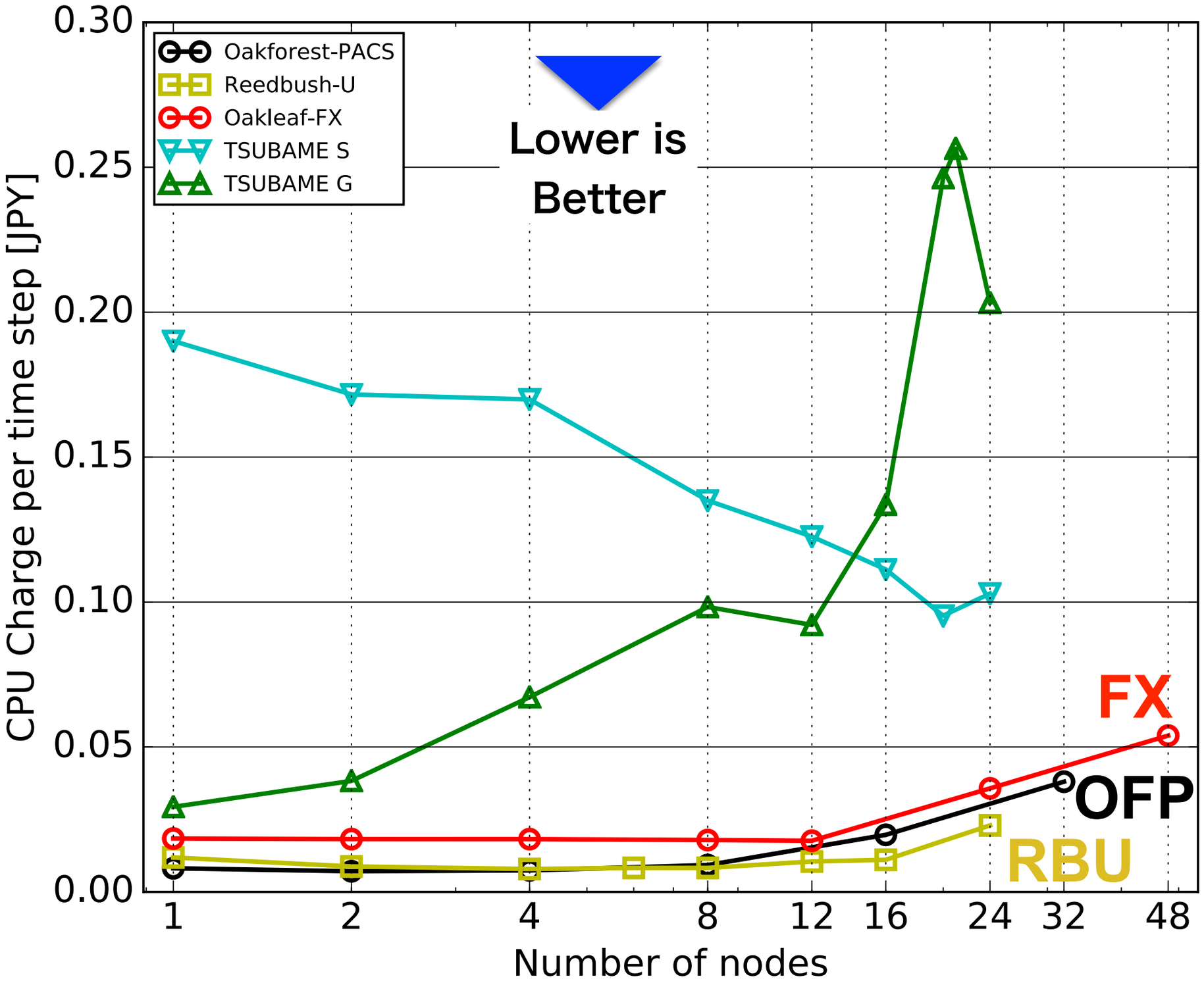
成果非公開型システムのステップ毎の課金比較



- 最安価
  - ✓ 1, 2ノード: TSUBAME G
  - ✓ 4, 8ノード: FOCUS F
  - ✓ 12~16, 24ノード：FOCUS D
  - ✓ 20ノード: FOCUS A
- 基本的に課金額は並列化効率に反比例するが，FOCUSでは多ノード割引率にも依存する.

Copyright © 2017 OCAEL All rights reserved.

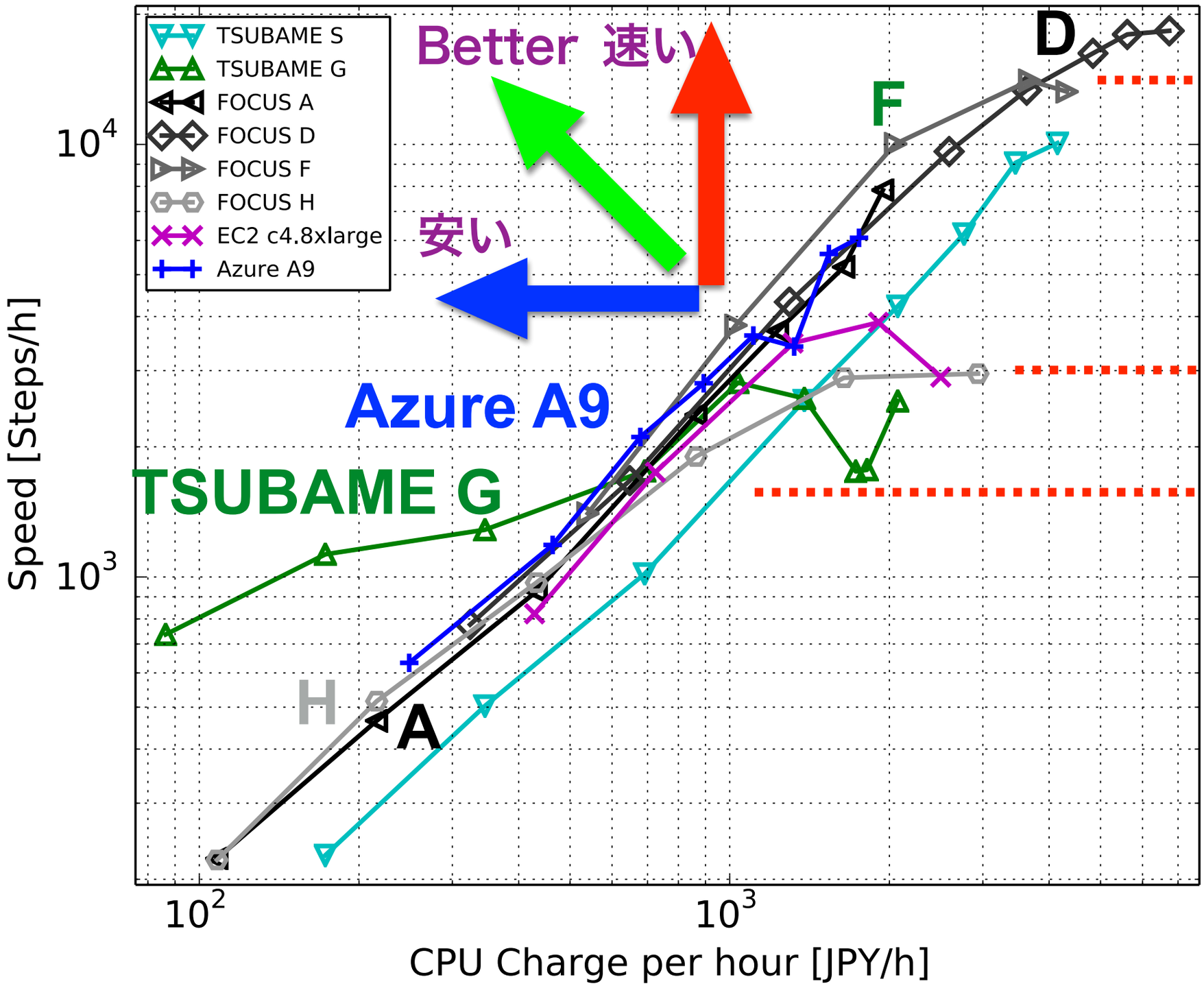
成果公開型システムのステップ毎の課金比較



- 最安価
  - ✓ 1～2ノード: OFP
  - ✓ 4～ノード：RBU
  - ✓ 次点FX

Copyright © 2017 OCAEL All rights reserved.

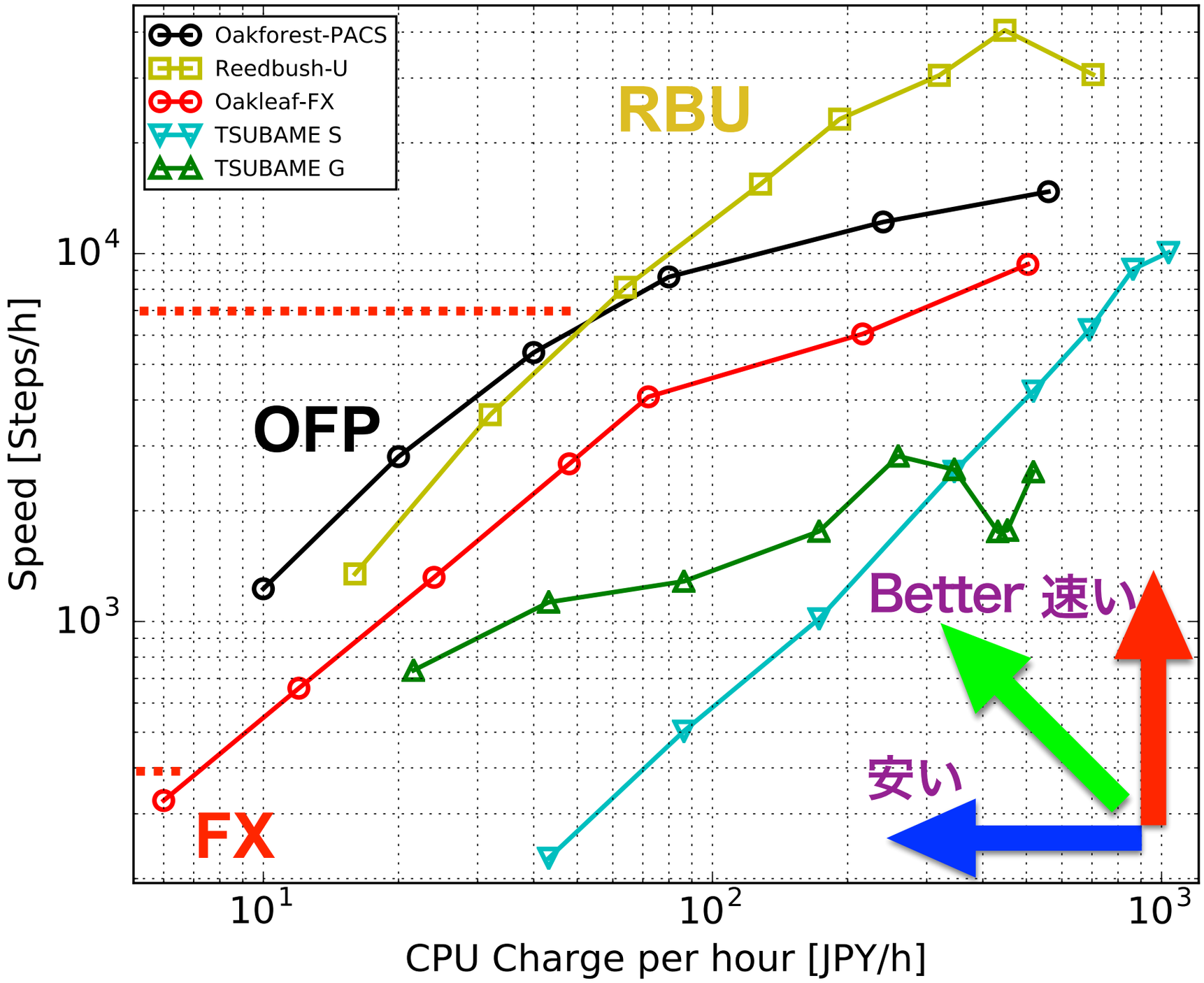
成果非公開型システムでの課金–解析速度曲線



ステップ数	最安価
～20K	FOCUS D
～15K	FOCUS F
～3K	Azure A9
～1.5K	TSUBAME G(GPU)

Copyright © 2017 OCAEL All rights reserved.

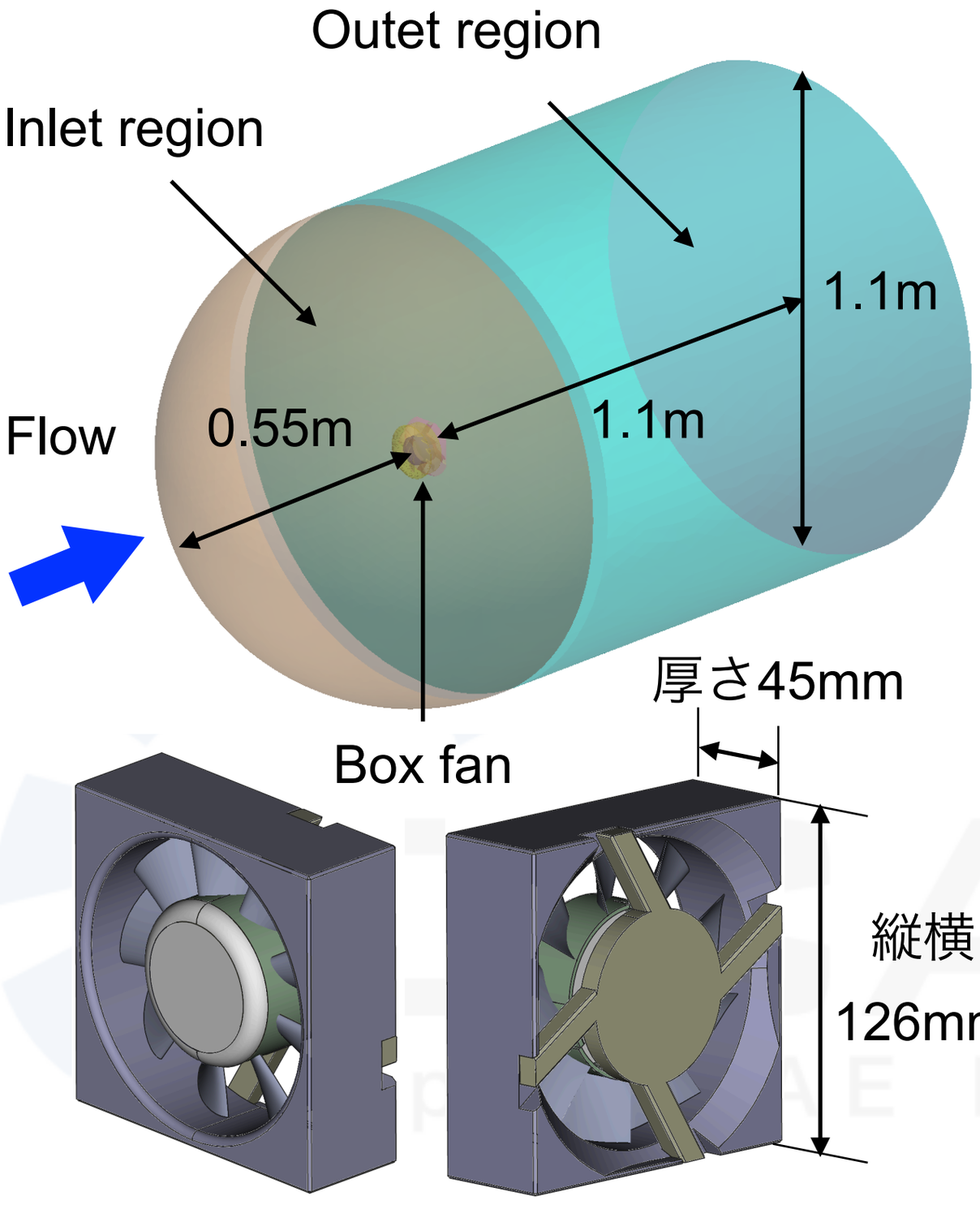
成果公開型システムでの課金–解析速度曲線



ステップ数	最安価
～40K	Reedbush-U
～7K	Oakforest-PACS
～40	Oakleaf-FX

Copyright © 2017 OCAEL All rights reserved.

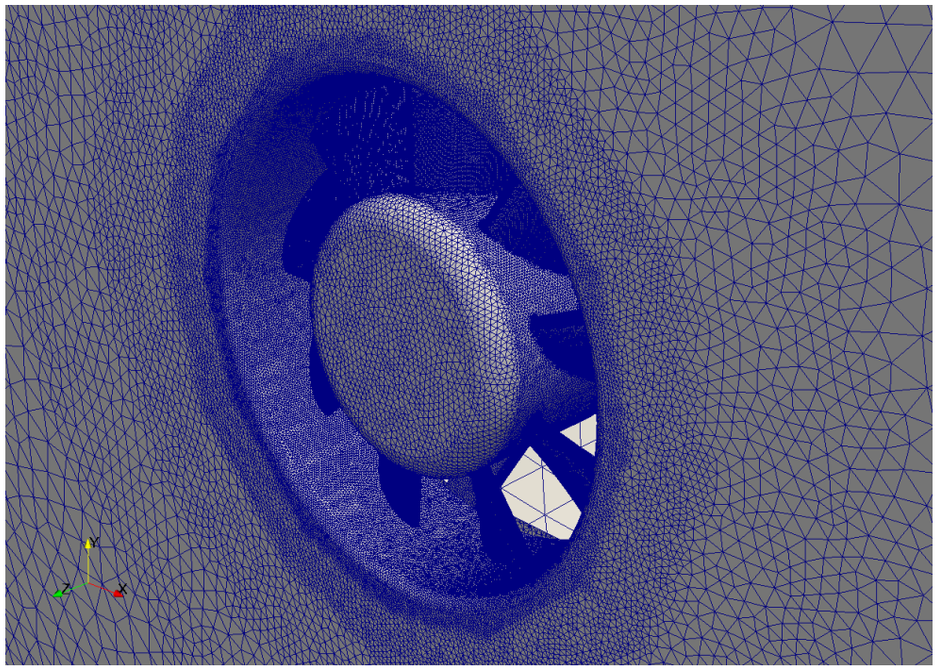
産応協HPCものづくりWSボックスファン



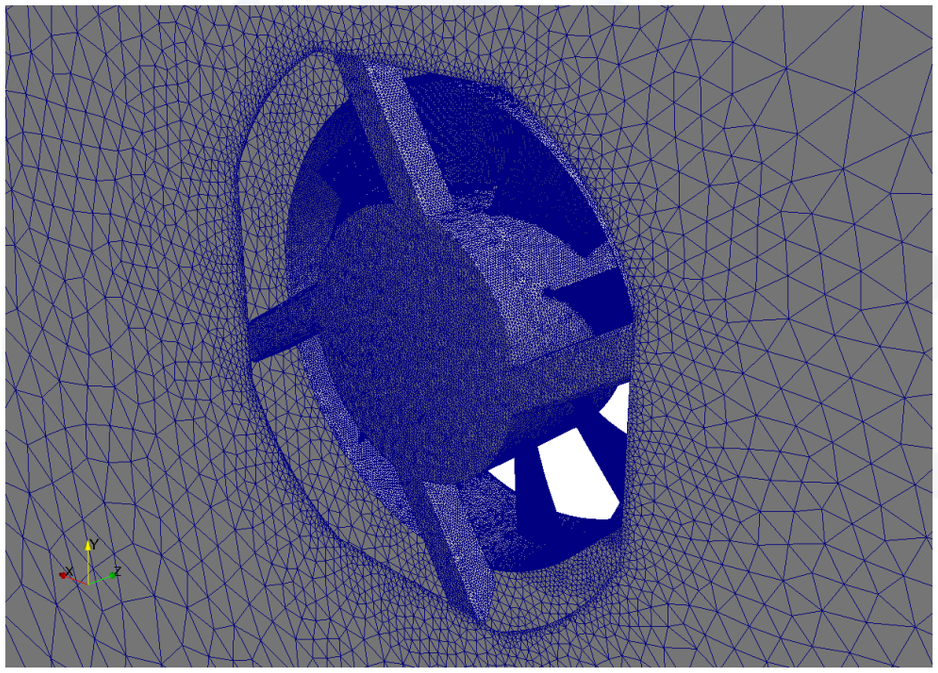
- 産応協HPCものづくりWS作成の共通ベンチマーク問題
- ボックスファン
  - ✓ 翼枚数：9 [枚]
  - ✓ 回転数：3000 [rpm]
  - ✓ プロペラ外径：111[mm]
- 2箇所で実験実施：流量特性測定(JIS B8330準拠)および騒音測定
- HPCものづくりWS(2017/6/25)
  - ✓ 結果発表機関：7
  - ✓ 使用CFDコード：CFX, Fluent, FrontFlow/Blue，Helyx, OpenFOAM，SCRYU/Tetra, StarCCM+

Copyright © 2017 OCAEL All rights reserved.

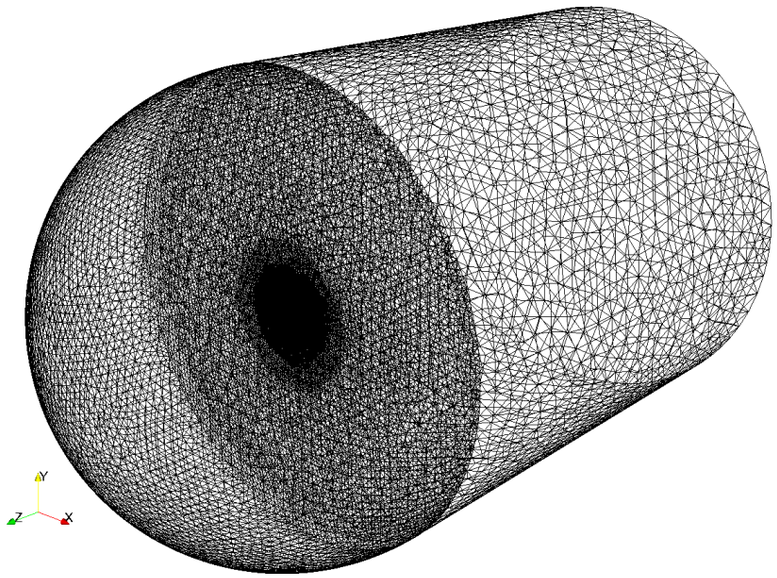
ボックスファンベンチマーク共通メッシュ



Inlet region

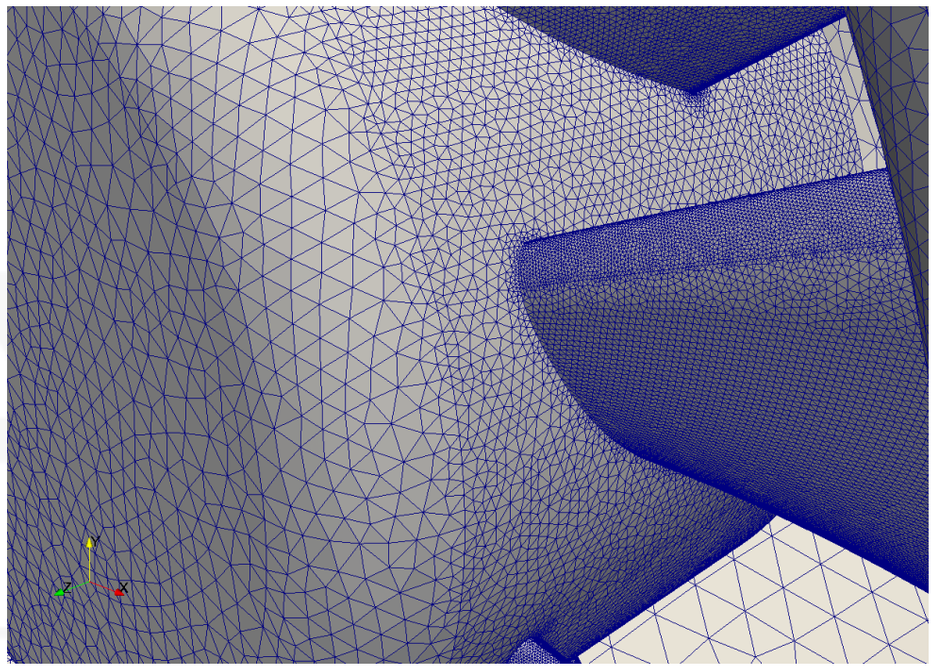


Outlet region



Mesh

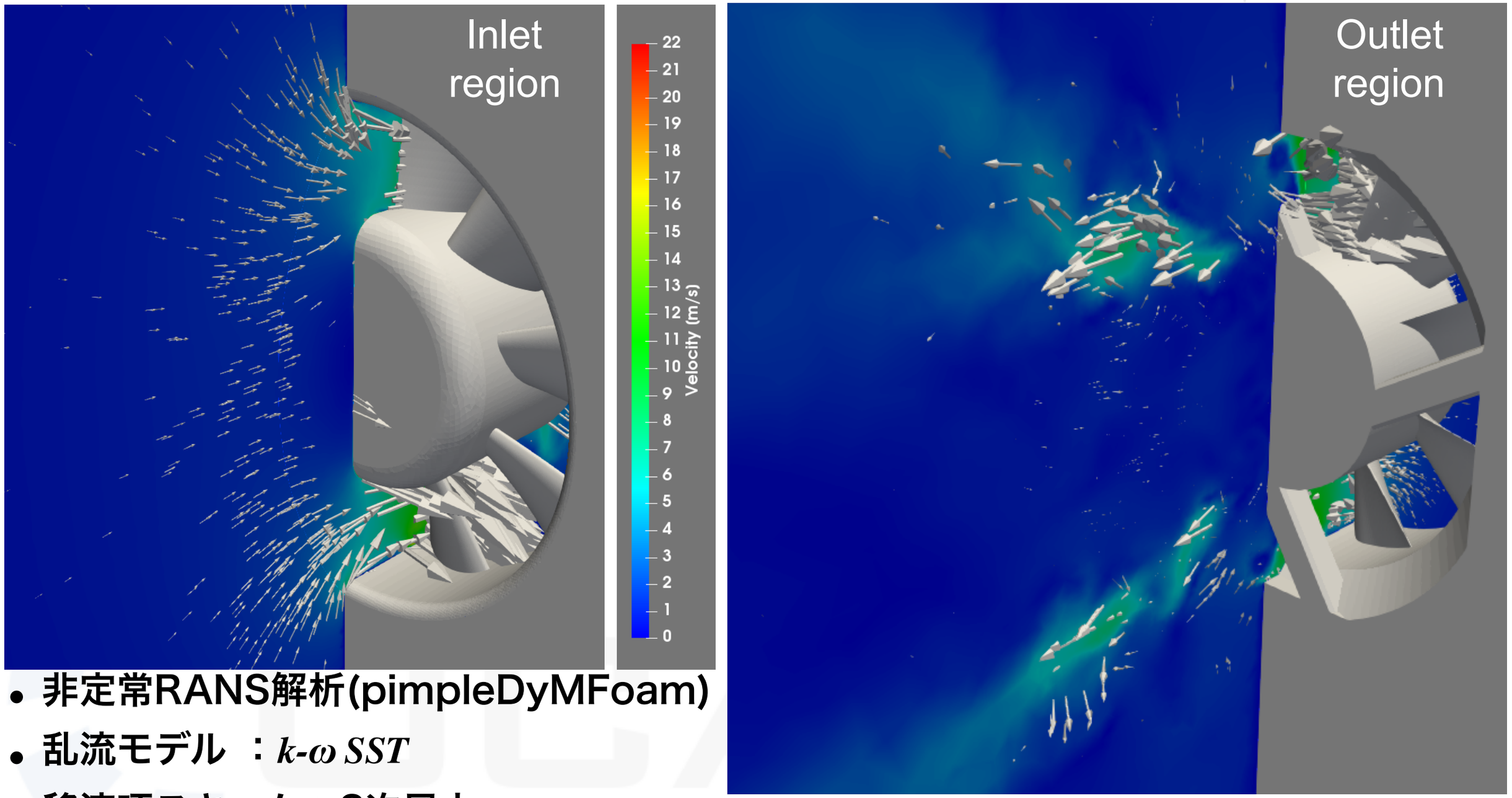
- 商用メッシャー使用
- 格子数：約1310万
- テトラ：約1050万
- プリズム：約260万
- プリズム第1層厚さ：約0.065mm



Blade and boss surface

Copyright © 2017 OCAEL All rights reserved.

ボックスファン速度分布瞬時値



- 非定常RANS解析(pimpleDyMFoam)
- 乱流モデル ：k-ω SST
- 移流項スキーム：2次風上
- 時間刻み：5×10^{-5}[s] (400step/rev)
- 時間積分：定常解から2秒後(100回転後)

流量2.186m^3/minでの静圧上昇の解析値：
実験値から約10%低い

Copyright © 2017 OCAEL All rights reserved.

ボックスファンベンチマーク条件・計測システム

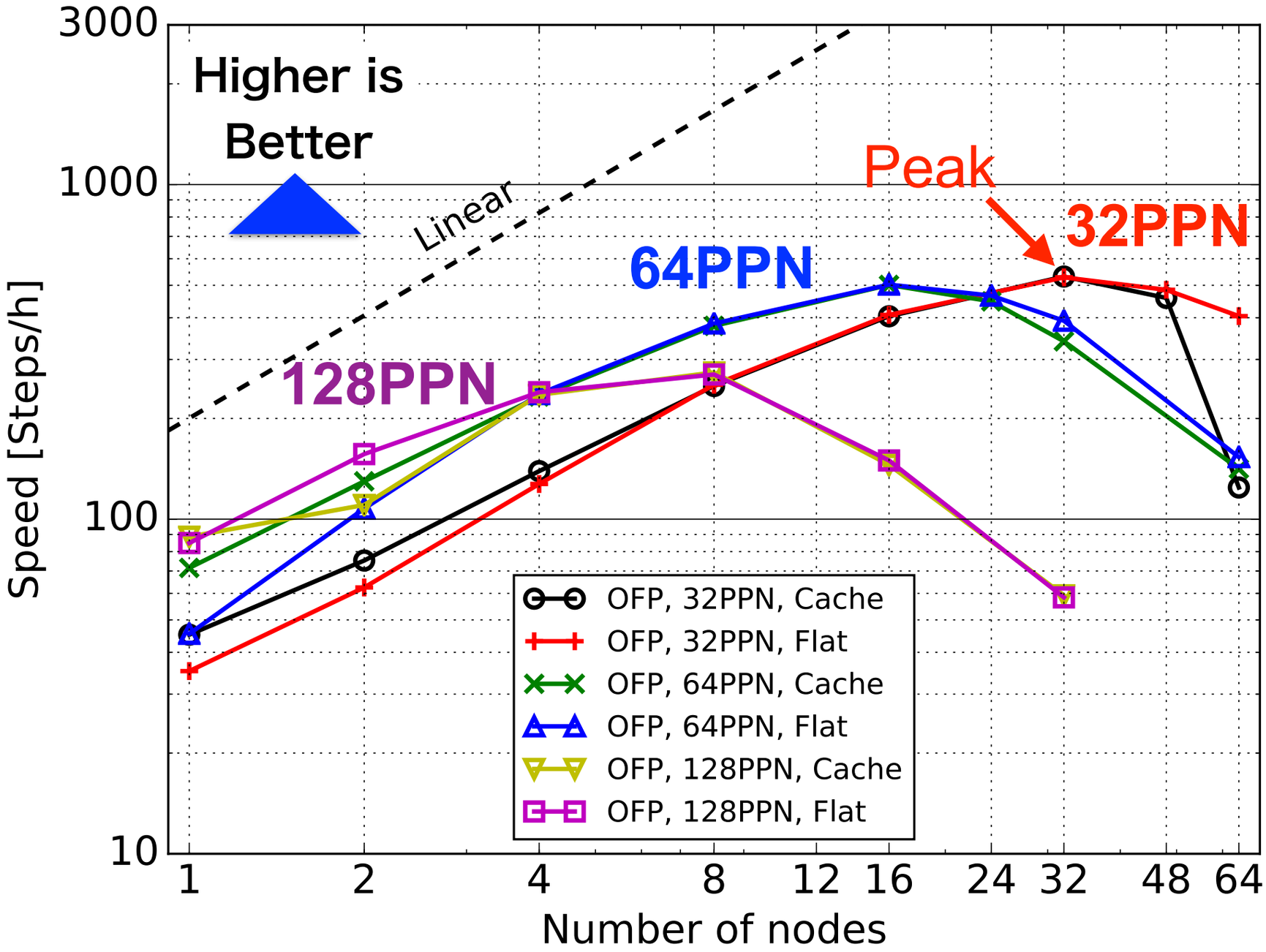
- 非定常RANS解析(OpenFOAM-v1612+，pimpleDyMFoam)
- 乱流モデル ：k-ω SST，移流項スキーム：1次風上
- 時間刻み：5×10^{-5}[s] (400step/rev)
- 2〜21ステップのCPU時間(Execution time)から1時間あたりのステップ数を算出
- 圧力線型ソルバ: PCG，前処理FDIC，許容残差10^{-7}
- 圧力以外の線型ソルバ: Gauss Seidel，許容残差10^{-5}
- 領域分割手法：scotch

機関	システム (略称)	CPU [GPU] (周波数[GHz])	CPU数 (コア)	倍精度性能 [GFlops]	メモリ[GiB] (帯域幅[GB/s])	インターコネクト (帯域幅[Gbps])
JCAHPC	Oakforest-PACS (OFP)	Intel Xeon Phi 7250, Knights Landing(1.4)	1(68)	3046	96(115.2), MCDRAM 16(490)	Intel Omni-Path (100)
東京大学	Reedbush-U (RBU)	Intel Xeon E5-2695 v4 (2.1-3.3)	2(36)	1210	256 (76.8×2)	Infiniband EDR(100)
FOCUS	A	Xeon L5640(2.26)	2(12)	108	48 (25.6×2)	Infiniband QDR(40)
	D	E5-2670 v2(2.5)	2(20)	400	64 (51.2×2)	Infiniband FDR(56)
	F	E5-2698 v4(2.2)	2(40)	1152	128 (76.8×2)	
	H	D-154(2.1)	1(8)	205	64 (34.1)	10GbE(10)×2 or 4

Copyright © 2017 OCAEL All rights reserved.

Oakforest-PACSの解析速度

コンパイラ：Icc 2017.4，KNL用最適化：-O3 -DvectorMachine -xmic-avx512
MPI：Intel MPI 2017.3，タイル0除外，Flatモードではlibhbm使用

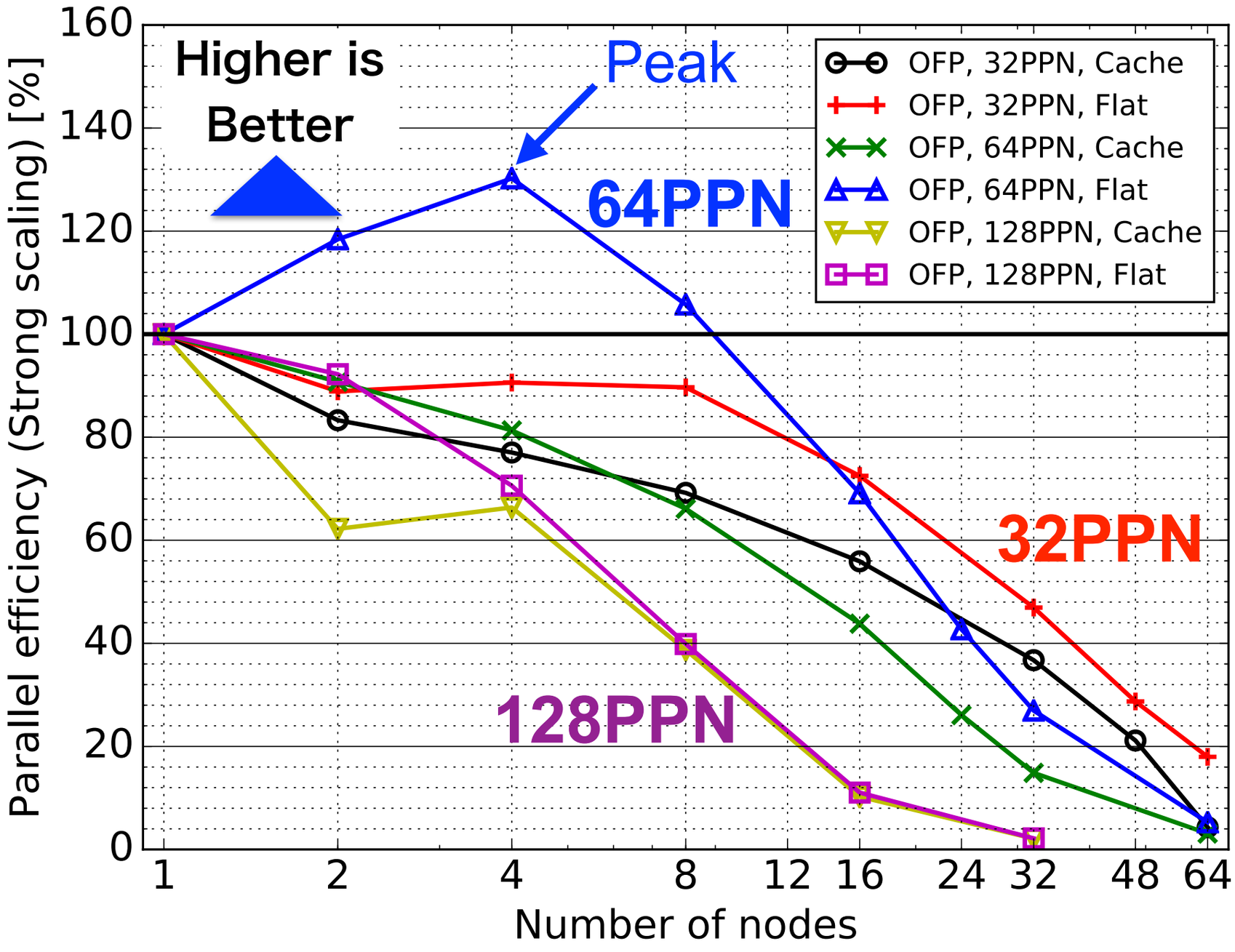


- 最適化オプションはKNL用がデフォルト設定(-O3)より僅かに速かった.
- 2ノード以下では，PPN(ノードあたりの使用プロセッサ数)がHyper thread利用の128が高速.
- 多ノードではCacheよりFlatのほうが高速

解析速度最大条件：
32PPN，32ノード(1024プロセッサ)，Flat

Copyright © 2017 OCAEL All rights reserved.

Oakforest-PACSのstrong scaling並列化効率



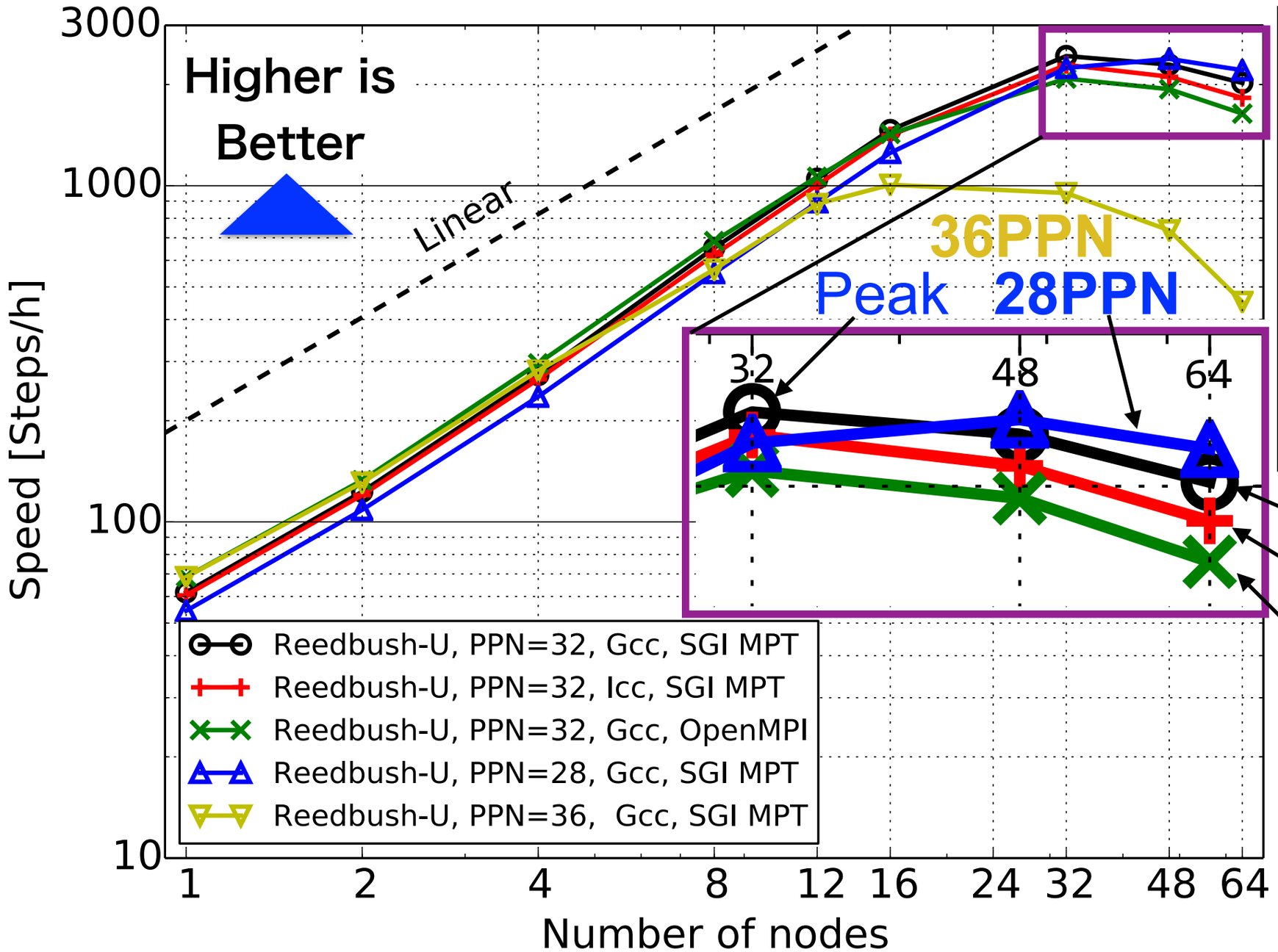
- Flat > Cache
- ～8ノード：
  64 > 32 >128PPN
- 16ノード～：
  32 > 64 >128PPN

並列化効率最大条件：
64PPN，4ノード(256プロセッサ)，Flat

Copyright © 2017 OCAEL All rights reserved.

Reedbush-Uの解析速度

コンパイラ：Gcc 4.8.5, Icc 2017.1，オプション：-O3(Intelは-xHost使用で悪化),
MPI：SGI MPT 2.16, OpenMPI 2.1.1 (Intel MPIは多ノードで動作せず)



- 多ノード，32PPN
  - ✓ Gcc > Icc
  - ✓ MPT > OpenMPI
- 4ノード以上：
  36PPN > 32PPN
- 48ノード以上：
  28PPN > 32PPN

Gcc, MPT

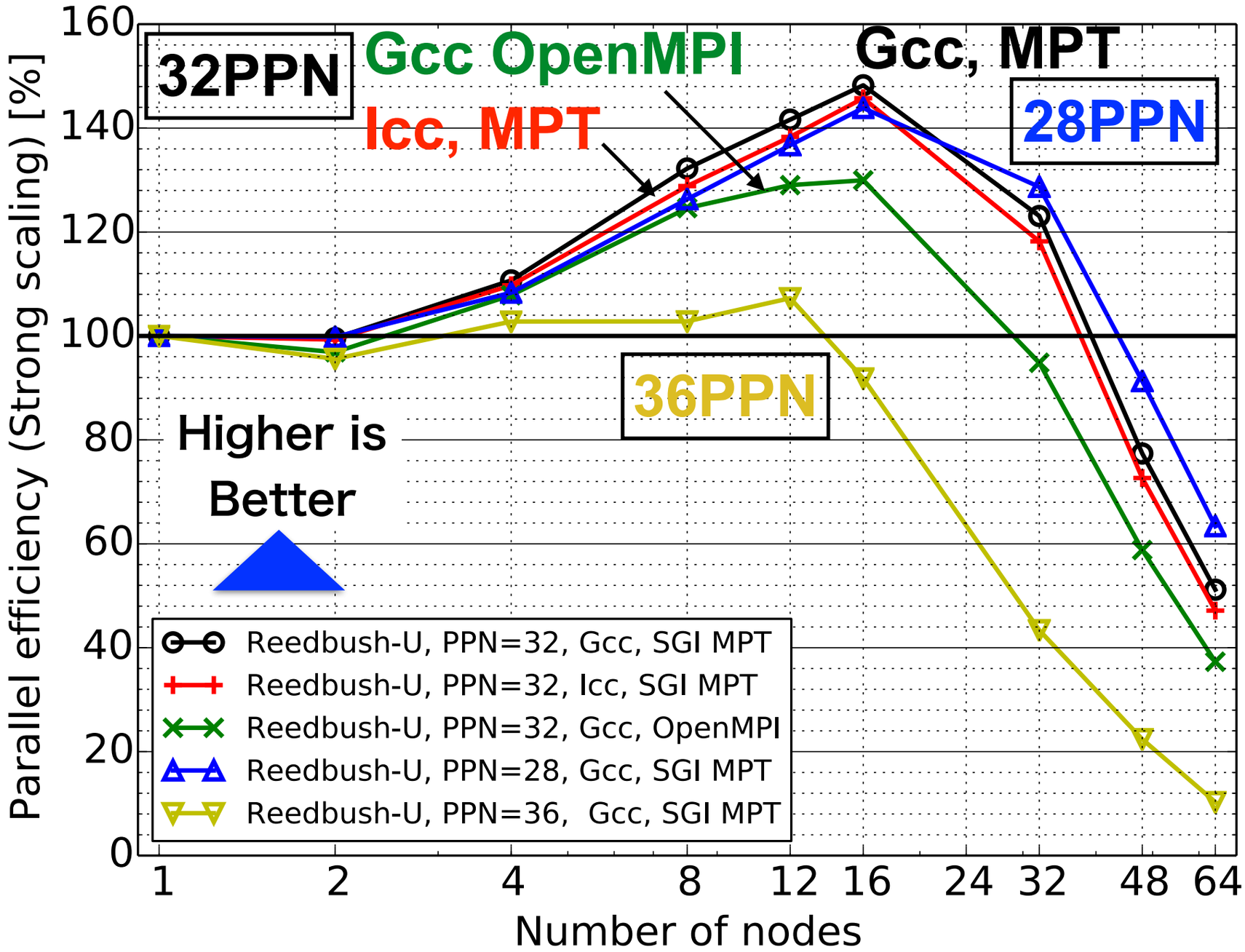
Icc, MPT **32PPN**

Gcc OpenMPI

解析速度最大条件:
32PPN，32ノード
(1024プロセッサ),
Gcc，SGI MPT

Copyright © 2017 OCAEL All rights reserved.

Reedbush-Uのstrong scaling並列化効率

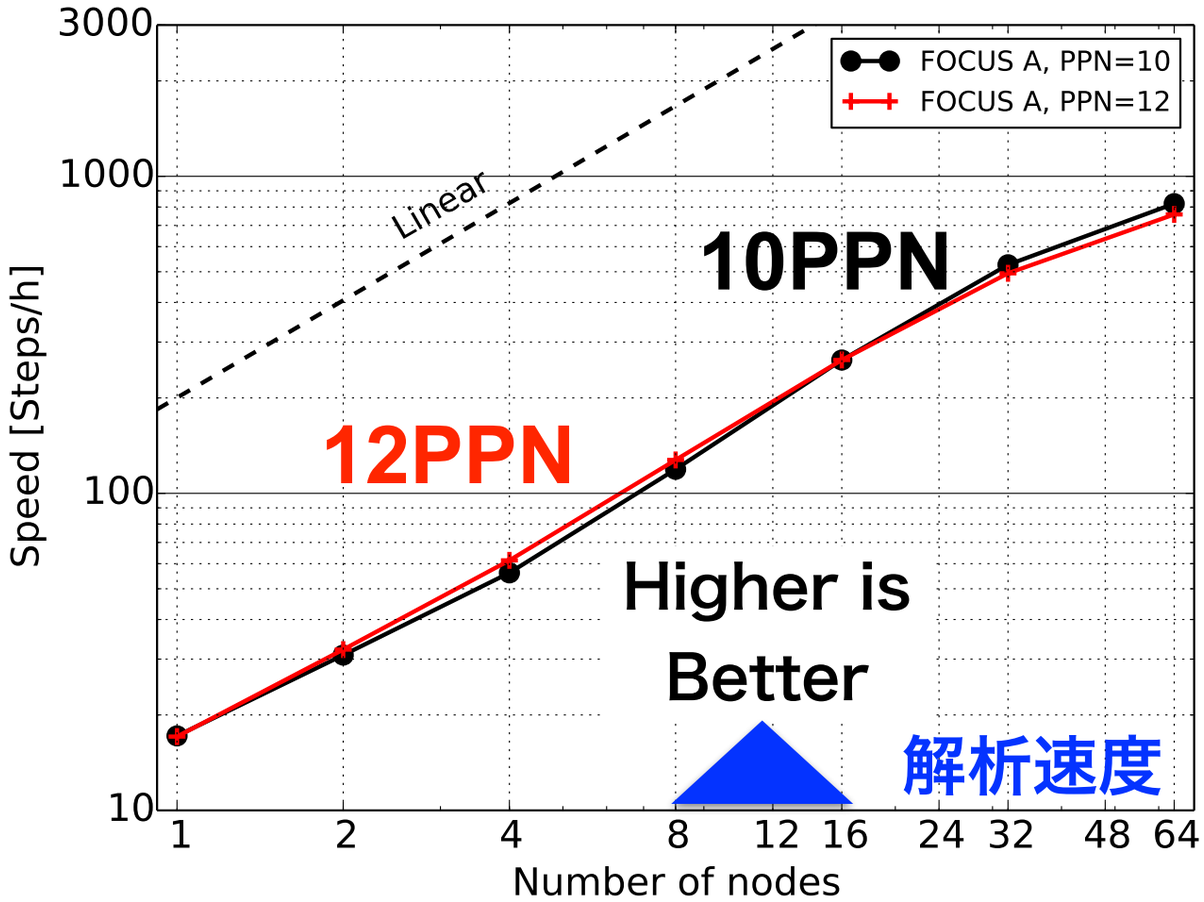


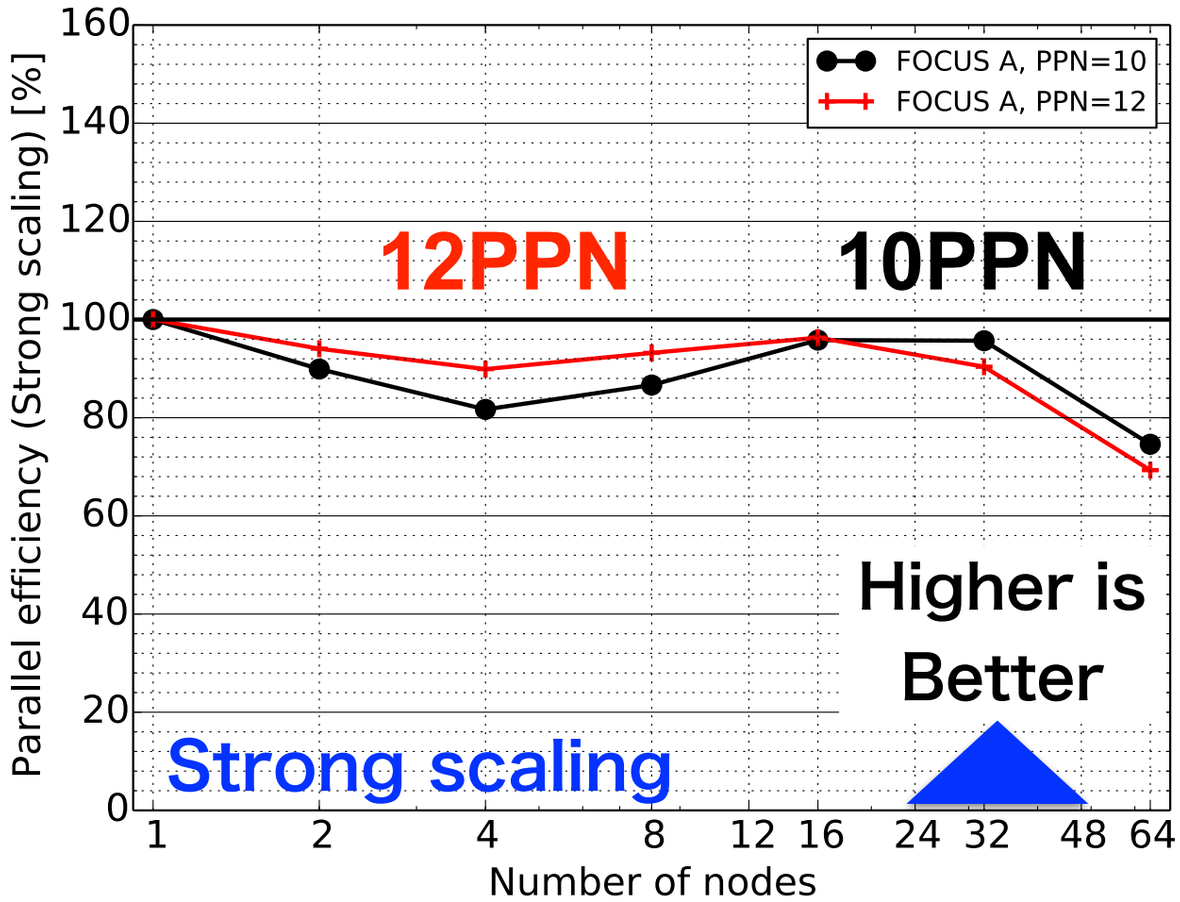
- 32PPN
  - ✓ 12ノード～：MPT > OpenMPI
  - ✓ ～32ノード：SGI MPTはスーパリニア
- 36PPN
  - ✓ ～12ノード：リニア
  - ✓ 32ノード～：悪化
- 28PPN
  - ✓ ～32ノード：スーパーリニア
  - ✓ 48ノード：リニア

Copyright © 2017 OCAEL All rights reserved.

FOCUS Aのベンチマーク結果

(FOCUS共通) コンパイラ：Gcc 4.8.5，オプション：-O3，MPI： OpenMPI 2.1.1





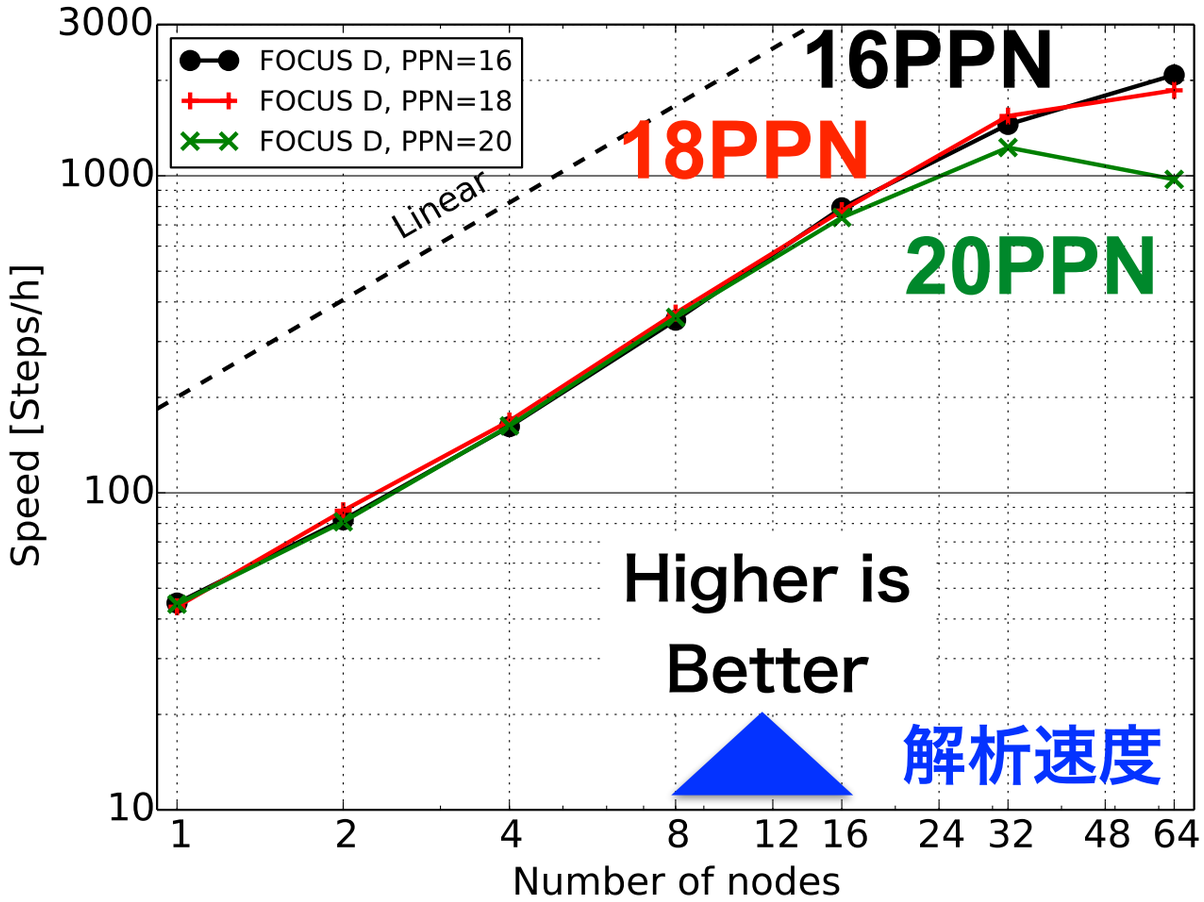
- 解析速度
  - ✓ 2～8ノード：12PPN > 10PPN
  - ✓ 32ノード～：10PPN > 12PPN
- Strong scaling： ～32ノードでほぼリニア

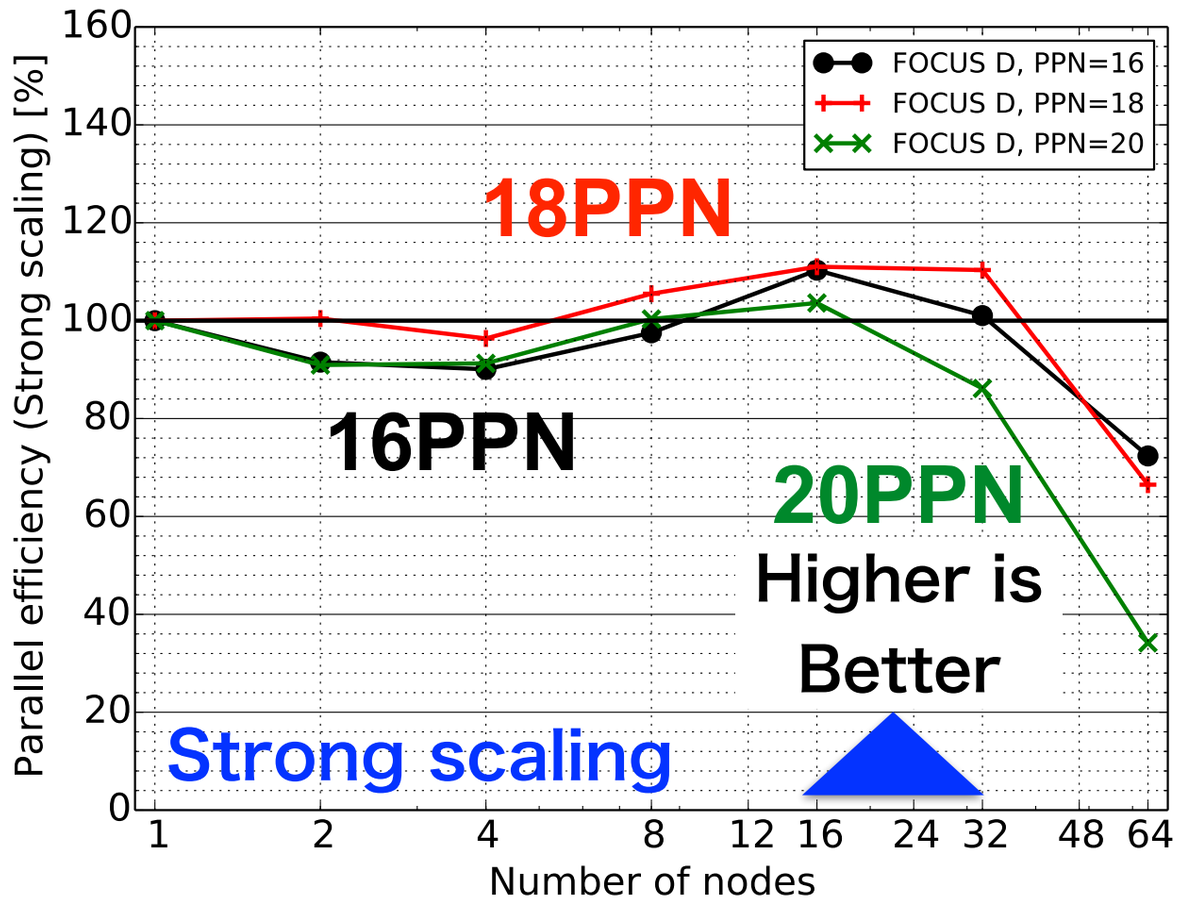
解析速度最大条件：10PPN，64ノード(640プロセッサ)

Copyright © 2017 OCAEL All rights reserved.

FOCUS Dのベンチマーク結果

(FOCUS共通) コンパイラ：Gcc 4.8.5，オプション：-O3，MPI： OpenMPI 2.1.1





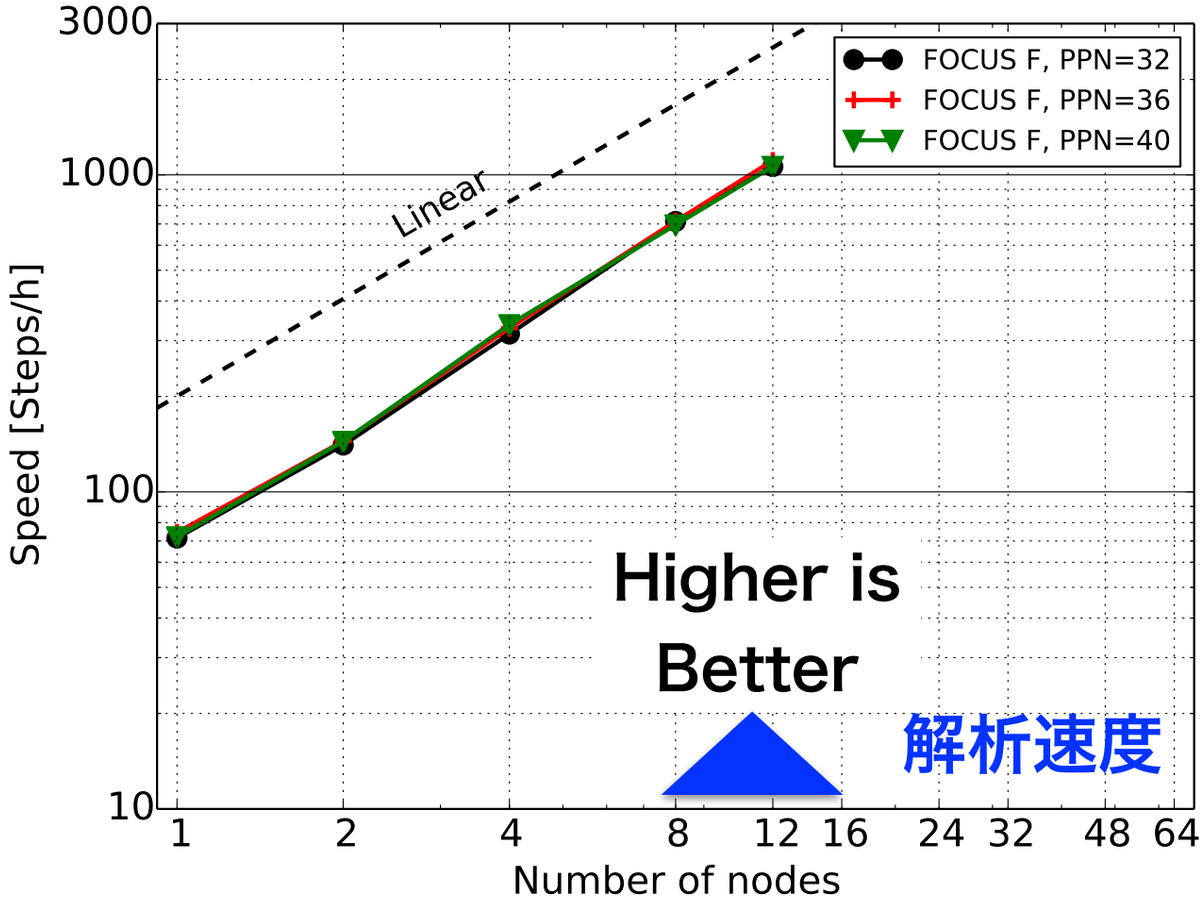
- 解析速度
  - ✓ ～16ノード 16PPN ≒ 18PPN ≒ 20PPN
  - ✓ 64ノード：16PPN > 18PPN > 20PPN
- Strong scaling：～32ノード 概ねリニア

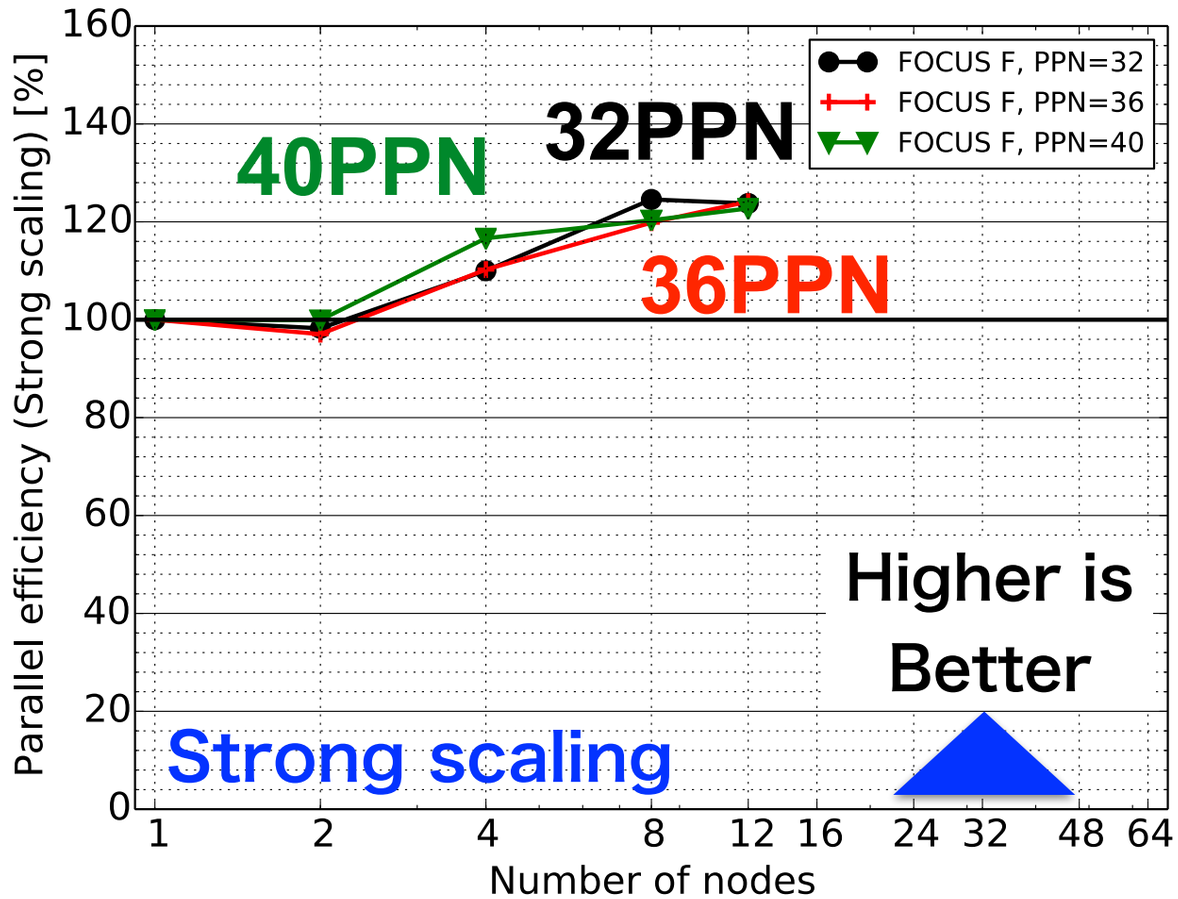
解析速度最大条件：16PPN，64ノード(1024プロセッサ)

Copyright © 2017 OCAEL All rights reserved.

FOCUS Fのベンチマーク結果

(FOCUS共通) コンパイラ：Gcc 4.8.5，オプション：-O3，MPI： OpenMPI 2.1.1





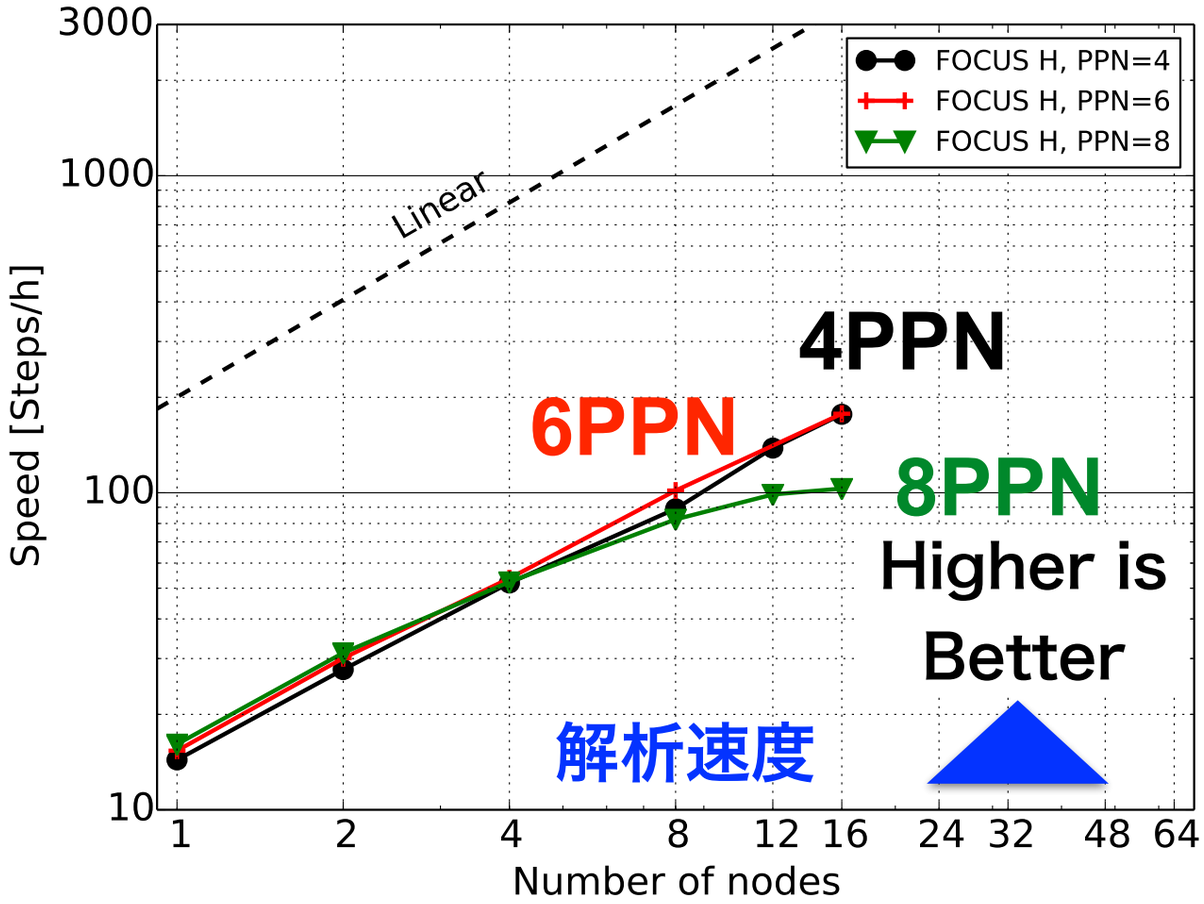
- 解析速度：全ノードで32PPN ≒ 36PPN ≒ 40PPN
- Strong scaling：4ノード～ スーパーリニア

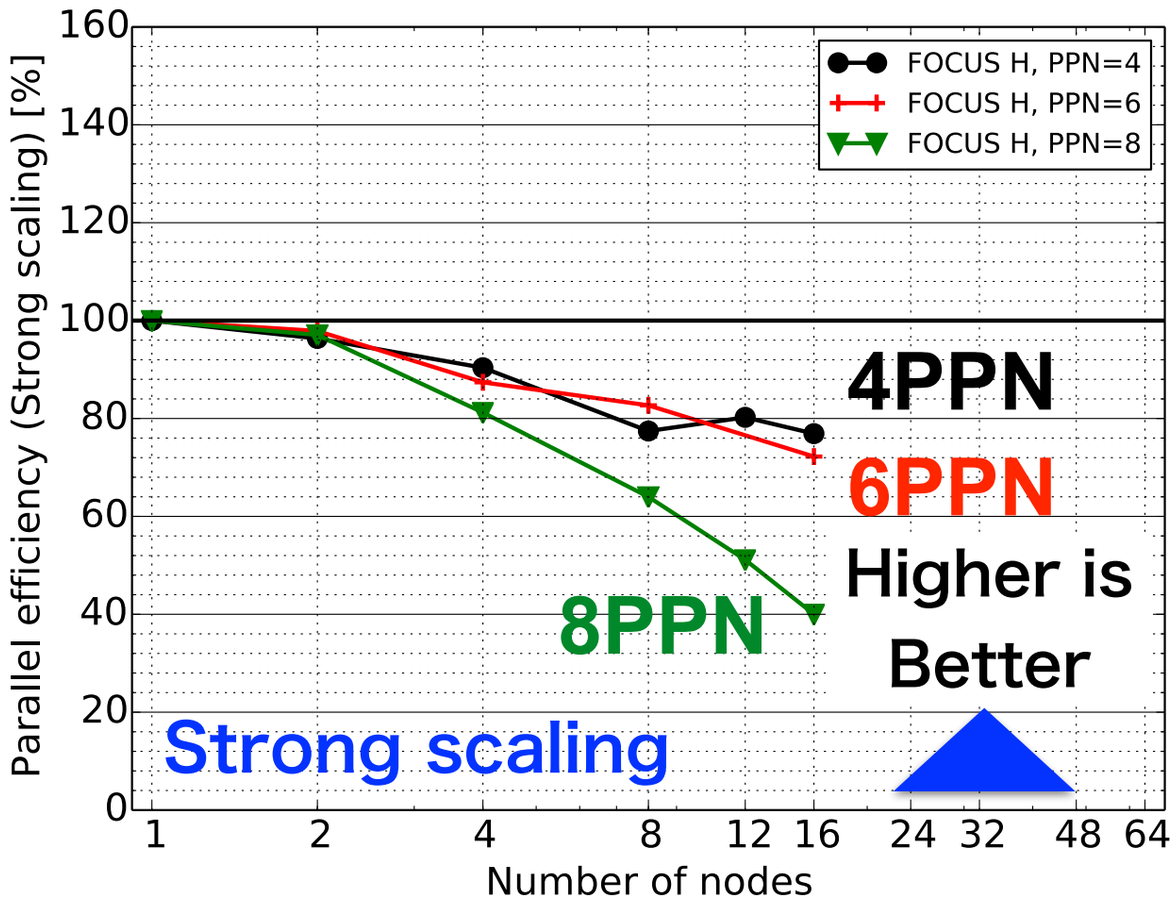
解析速度最大条件：32PPN，12ノード(384プロセッサ)

Copyright © 2017 OCAEL All rights reserved.

FOCUS Hのベンチマーク結果

(FOCUS共通) コンパイラ：Gcc 4.8.5，オプション：-O3，MPI： OpenMPI 2.1.1





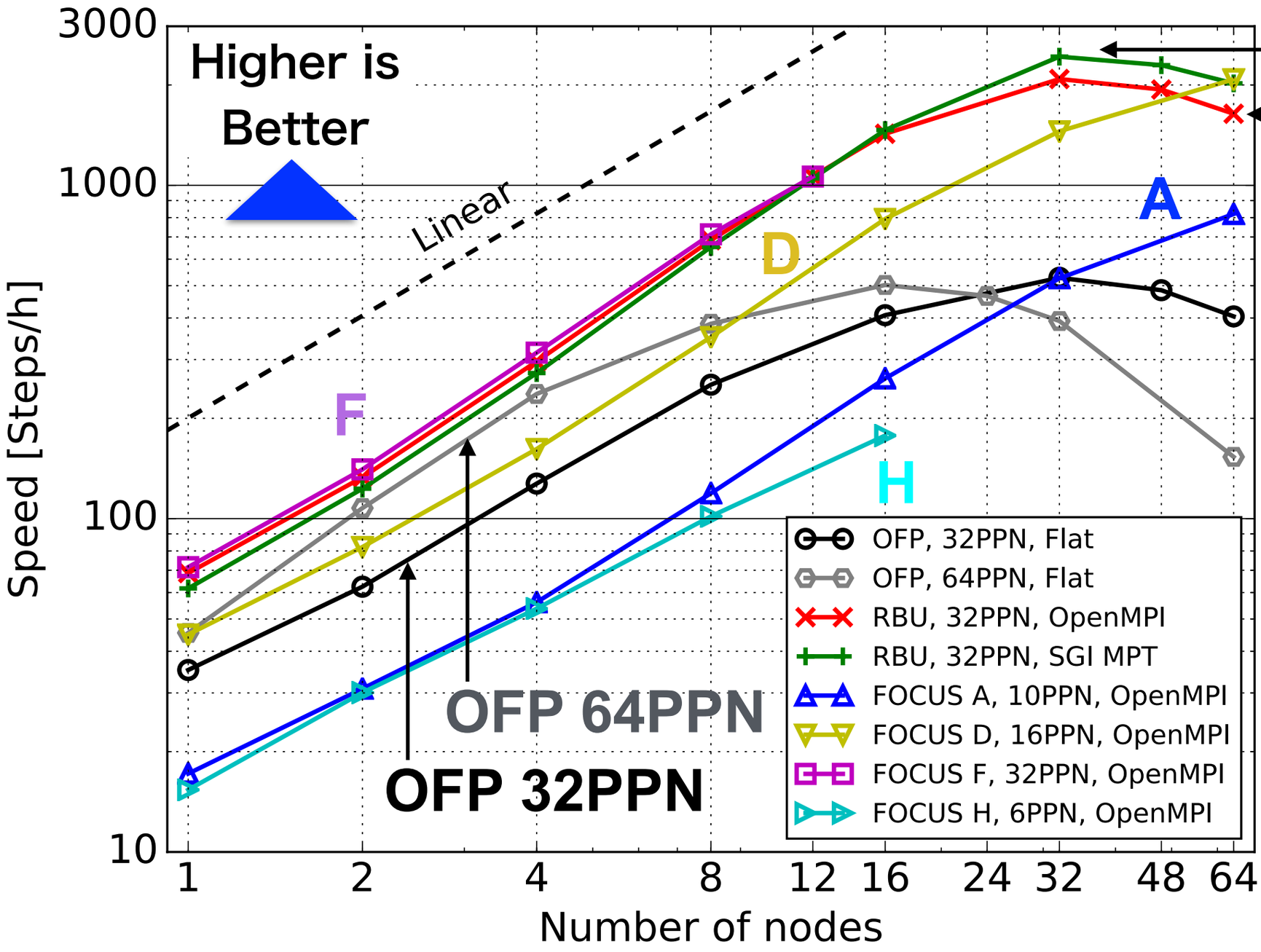
- 解析速度：4PPN ≒ 6PPN > 8PPN
- Strong scaling：
  - ✓ 4PPN ≒ 6PPN > 8PPN
  - ✓ インターコネクトが10GbEであるため，2ノード以上から除々に悪化する

解析速度最大条件：6PPN，16ノード(96プロセッサ)

Copyright © 2017 OCAEL All rights reserved.

各システムの解析速度比較

各システムでピーク性能が高い条件での解析速度を比較．OpenMPIを用いるFOCUSとの比較のため，Reedbush-UはMPIライブラリがMPT以外にOpenMPIも掲載



RBU MPT

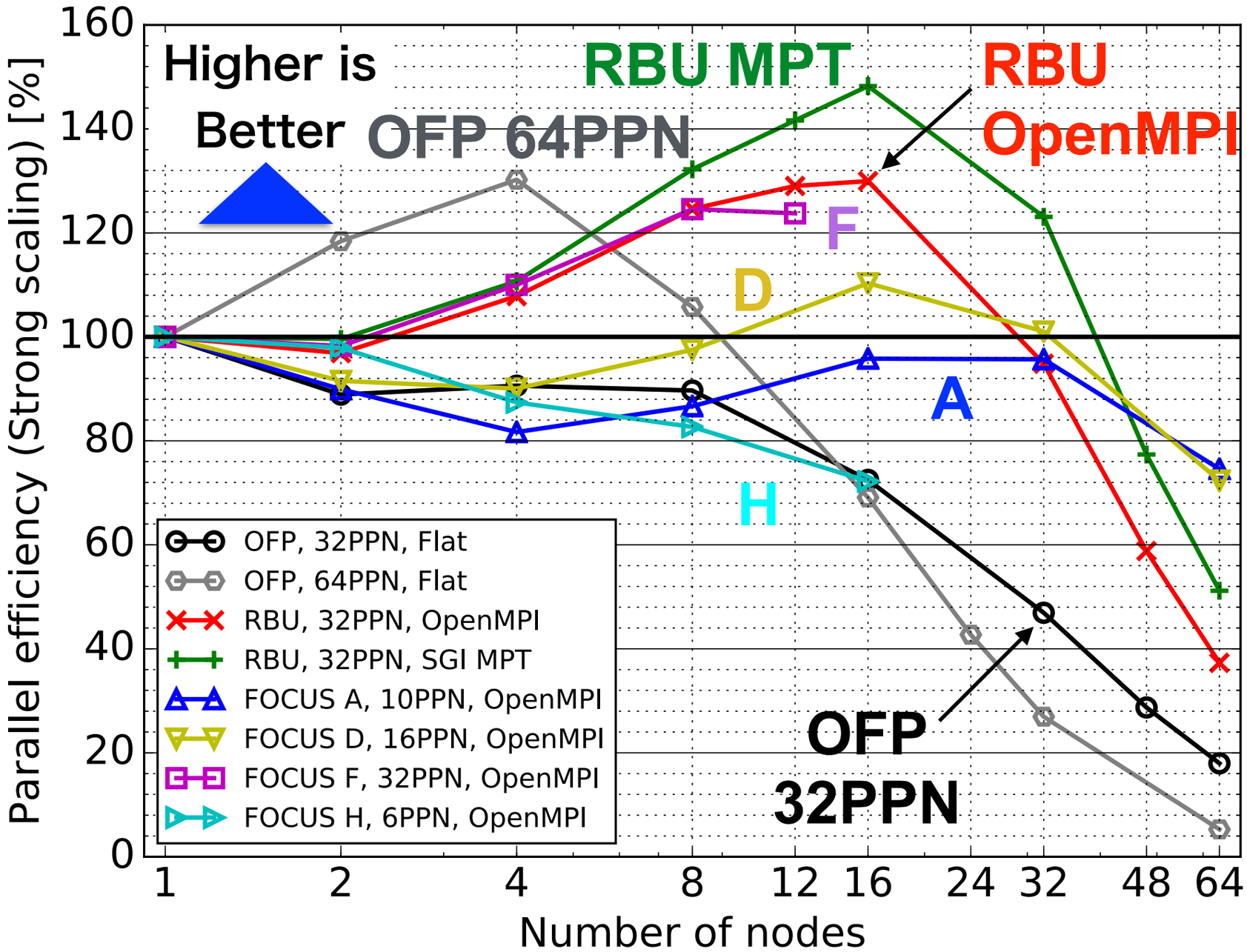
RBU OpenMPI

解析速度順

1. FOCUS F
2. Reedbush-U (～12ノード：OpenMPI, 16ノード～：MPT)
3. Oakforest-PACS, 64PPN(～8ノード)
4. FOCUS D
5. Oakforest-PACS, 32PPN(～32ノード)
6. FOCUS A
7. FOCUS H

Copyright © 2017 OCAEL All rights reserved.

各システムのstrong scaling並列化効率比較



スーパースケール

- Oakforest-PACS, 64PPN (～8ノード)
- Reedbush-U, MPT (～32ノード)
- Reedbush-U, OpenMPI(～16ノード)
- FOCUS F(～12ノード)

概ねスケール

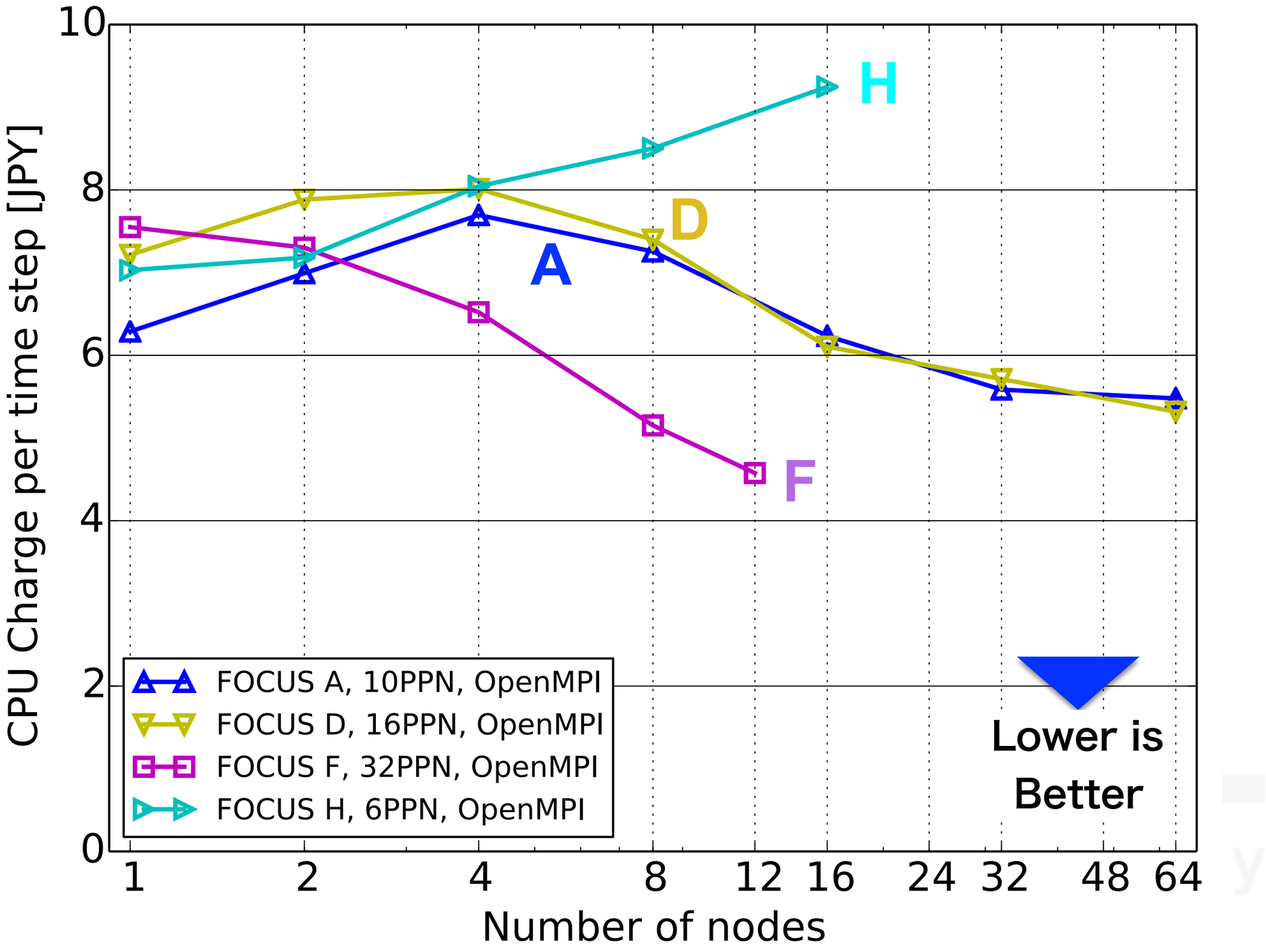
- FOCUS D(～32ノード)
- FOCUS A(～32ノード)

非スケール

- FOCUS H
- Oakforest-PACS, 32PPN

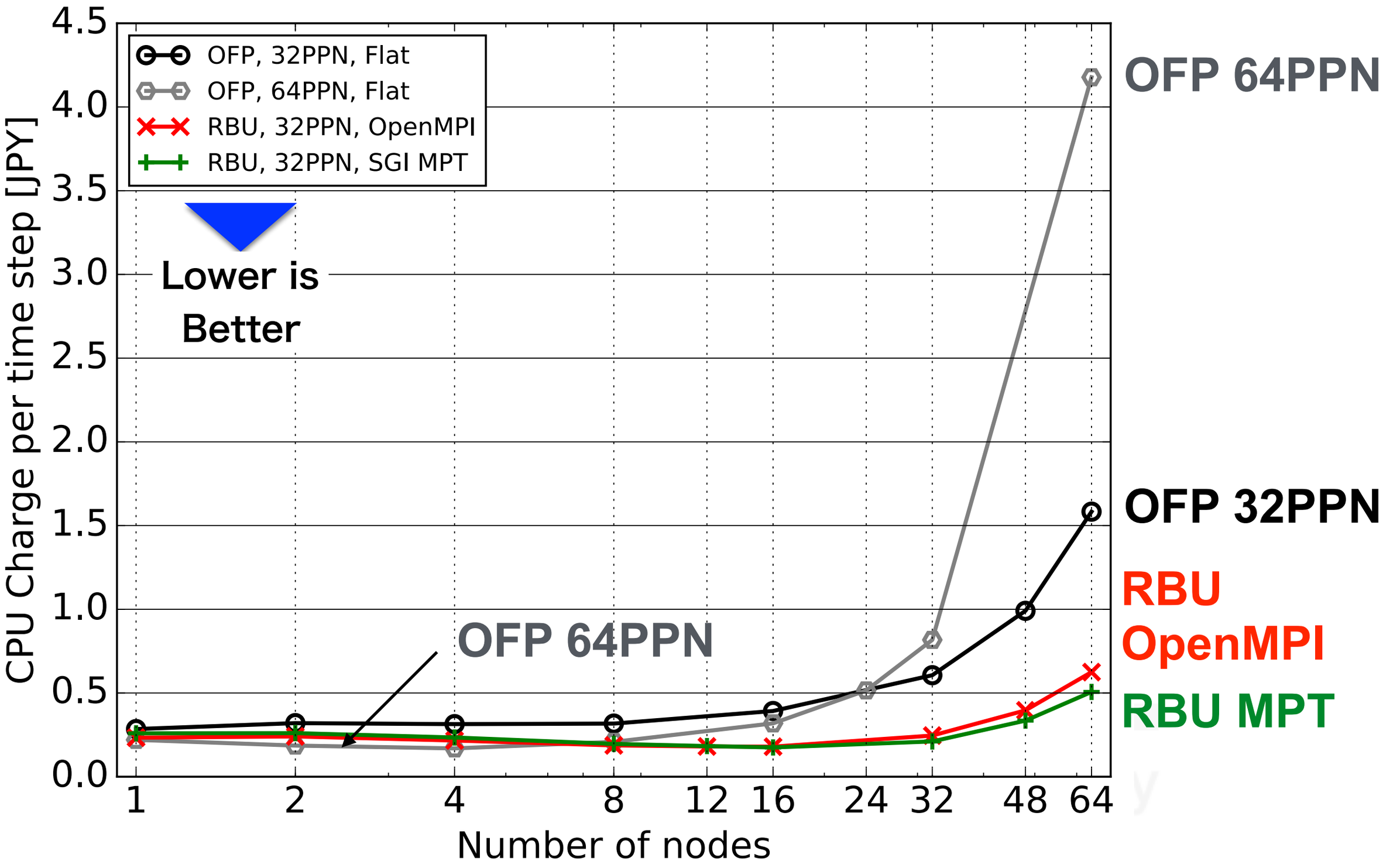
Copyright © 2017 OCAEL All rights reserved.

FOCUSのステップ毎の課金比較



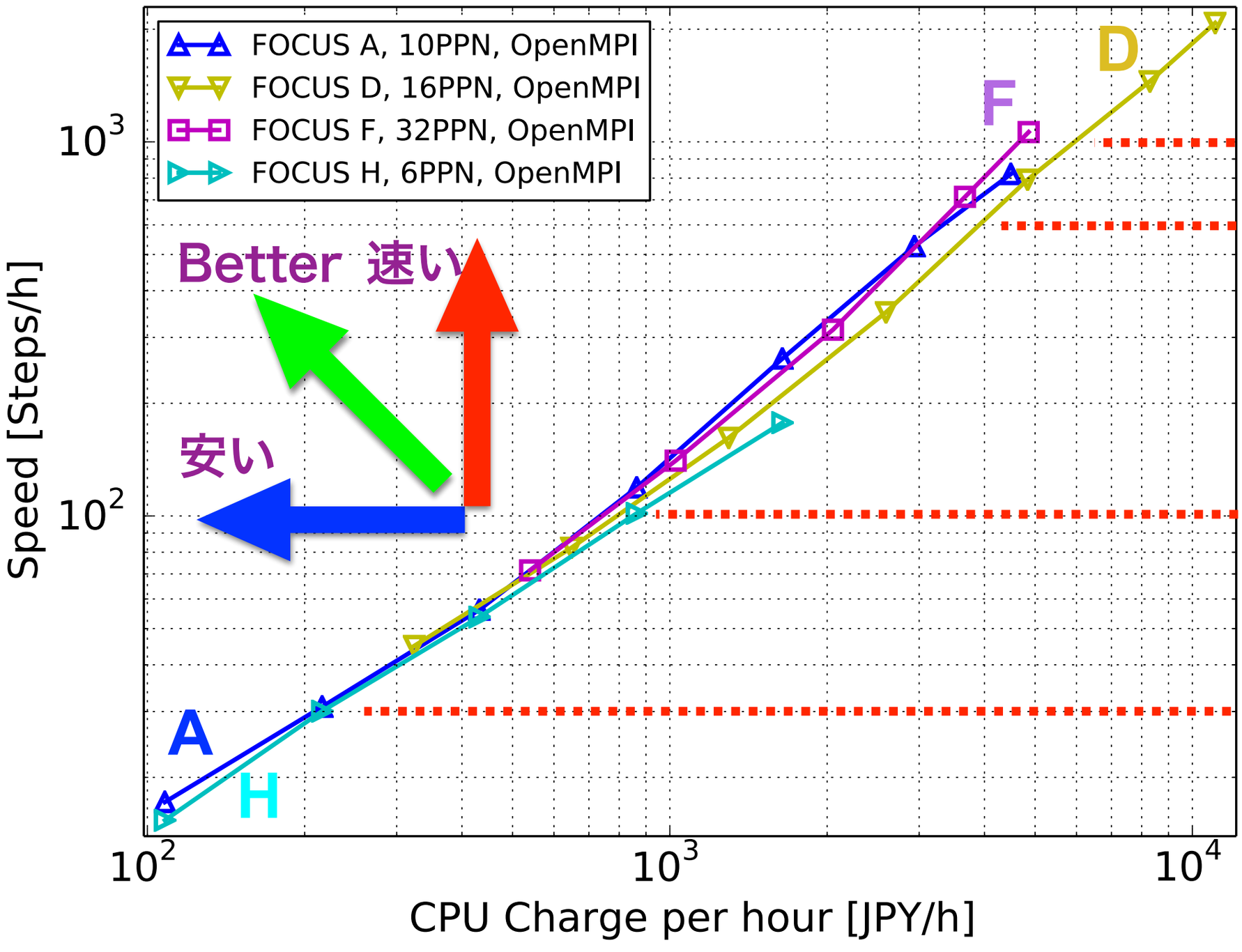
Copyright © 2017 OCAEL All rights reserved.

成果公開型システムのステップ毎の課金比較



Copyright © 2017 OCAEL All rights reserved.

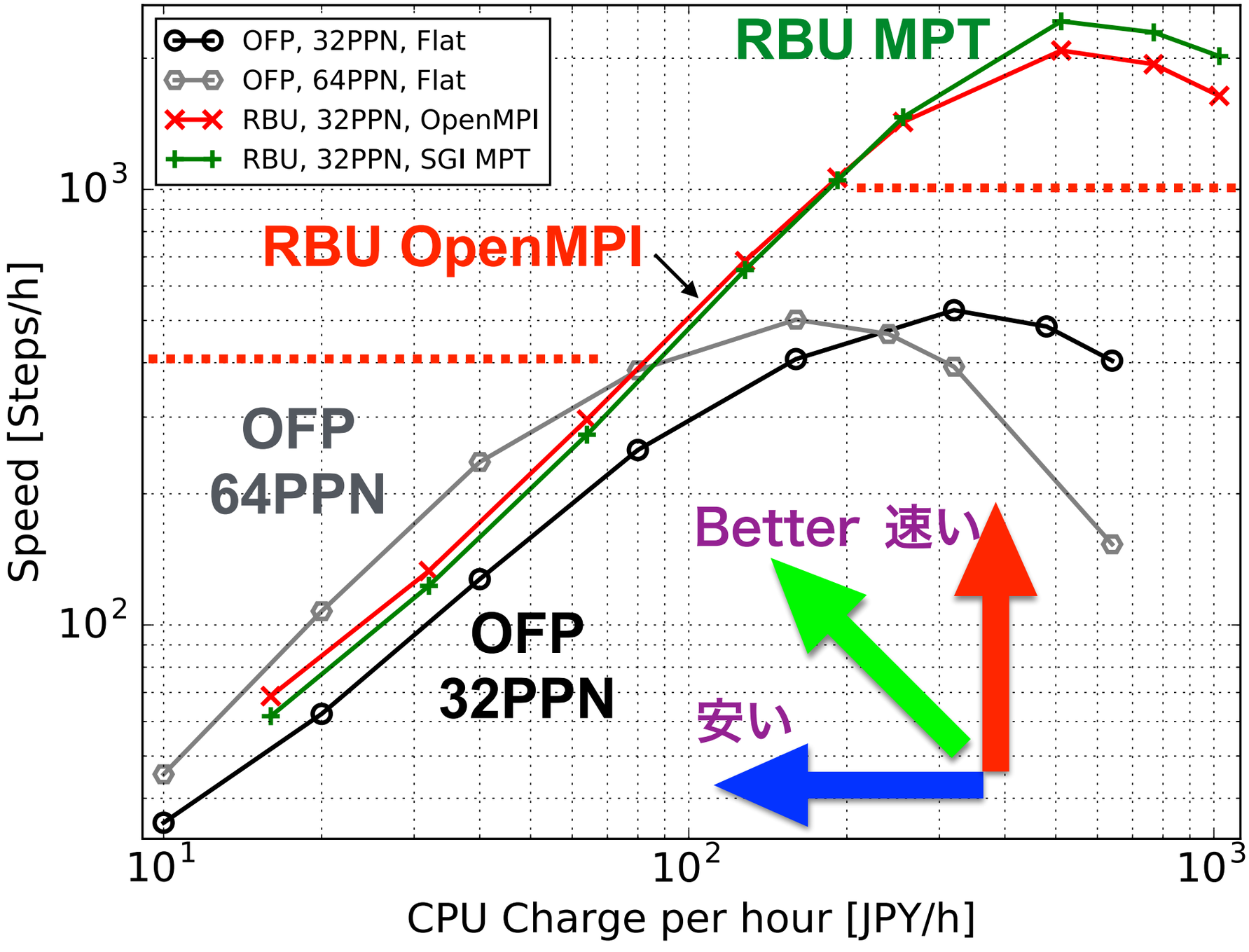
FOCUSの課金–解析速度曲線



ステップ数	最安価
1K～	FOCUS D
600～1K	FOCUS F
100～600	FOCUS A FOCUS F
30～100	ほぼ重なる
～30	FOCUS A

Copyright © 2017 OCAEL All rights reserved.

成果公開型システムの課金–解析速度曲線



ステップ数	最安価
1K～	Reedbush-U, MPT
～1K	Reedbush-U, OpenMPI
～400	Oakforest-PACS, 64PPN

Copyright © 2017 OCAEL All rights reserved.

まとめ

- FOCUSや大学のスパコンおよびクラウドにおいて，チャンネル流れおよびボックスファンのベンチマークテストを実行し，各システムでの解析速度や並列化効率，対費用効果を比較した.
- 課金ー解析速度の関係から，解析速度範囲に応じて最安価なシステムが分れる結果となった.
- FOCUSはシステムが多様だが，課金ー解析速度の関係がほぼ重なる課金システムなので，求める解析速度とキューの混雑状況に応じて，柔軟に投入キューを選択するのが望ましい.

謝辞 東京工業大学学術国際情報センター共同利用推進室の佐々木様から，OpenFOAMとRapidCFDの評価用として，TSUBAME 2.5の計算機リソースをご提供頂いた．日本Microsoft(計測当時)の佐々木様から，Microsoft Azure A9でのベンチマークの結果をご提供頂いた．電通国際情報サービス(計測当時)の住友様には，Amazon EC2のベンチマーク結果をご提供頂いた．青子守歌様には，RapidCFDのビルドについてご協力頂いた．スーパーコンピューティング技術産業応用協議会からボックスファンの共通メッシュや実験結果等についてご提供頂いた．FOCUSからボックスファンベンチマーク用に計算機資源を提供頂いた．ここに深く感謝する.

Copyright © 2017 OCAEL All rights reserved.