

申請書の書き方

木戸 善之

2021年11月10日

そもそも公募型利用制度とは

- 概要より

大規模計算機システムを活用する研究開発の育成・高度化支援の観点から、今後の発展が見込まれる**萌芽的な研究課題**や、本センターの**大規模計算機システム**を最大限活用することで**成果が見込まれる研究課題**を公募しております。

本来であれば、全部の課題を通したい！
大規模計算機を必要としている研究者にお届けしたい！！

- でも計算機資源は限られている．．



- 大規模計算機が必要な研究？
- 必要としている計算機資源量は妥当？

課題申請書

- 課題申請書は公募利用制度のWebページを参照



http://www.hpc.cmc.osaka-u.ac.jp/wp-content/uploads/2021/10/J_2022_Application-for-Proposal-based-Use-of-HPC.doc

課題申請書の構造

- 代表者の情報
- 課題のタイトル
- **1. 計算計画の概要**
 - 計算内容
- **2. 研究課題の概要**
- **3. 計算計画**
- 4. 利用者リスト
- 5. 外国人リスト
- 6. JHPCN/HPCIの応募状況
- 7. JHPCN構成機関が実施する公募型利用制度への応募状況
- 8. 前年度の公募型利用制度採択課題との関係

大阪大学 サイバーメディアセンター
大規模計算機システム 公募型利用制度申請書

※本様式は、本センター 大規模計算機システム担当 system@cmc.osaka-u.ac.jp までお送りください。

代表者所属			
代表者氏名		代表者職名	
年齢/性別	年齢 歳 (2022 年 4 月 1 日時点) / 男・女		
電話連絡先			
E-mail アドレス			
住所	〒		
指導教員	(代表者職名が大学院生の場合に記入) 氏名 : E-mail アドレス :		
研究課題名			

計算計画の概要

- 科研費などの研究計画と大きく違うところ
- 計算機をどれだけ使えば扱う問題が解けるのかを見積もる必要がある
- 本来、定量的に計算量を見積もる
- ノード時間：ノードを使用する時間積
 - 24ノード時間
= 24計算ノードを1時間使用
= 1計算ノードを24時間使用

1. 計算計画の概要			
申請研究課題で利用したい計算機システムについて記載ください。			
高性能計算・データ分析基盤システムの資源量		OCTOPUS の資源量	
汎用 CPU ノード	ノード時間	汎用 CPU ノード	ノード時間
GPU ノード	ノード時間	GPU ノード	ノード時間
ベクトルノード	ノード時間	Xeon Phi ノード	ノード時間
ストレージ容量	TB	大容量主記憶搭載ノード	ノード時間
		ストレージ容量	TB
計算内容 (例 ○○○○○○:□□□□法による△△△計算を行うため XXX ノード時間の計算を YYY 回実行する。)			
(3. の計算計画に記載した内容を要約して記入してください。)			

ノード時間の算出方法（考え方）

- 例えば OCTOPUS, 8ノードで20時間かかっていたプログラム「foo」があった場合
 - fooの処理は160ノード時間（OCTOPUS）必要
- ではSQUIDではどうか？
 - SQUIDでは動かしてないからわからないよ！
 - 単体論理性能ではOCTOPUSの約8倍
 - SQUIDでの見積りノード時間 = $160 / 8 = 20$ ノード時間

あくまでも見積なので厳密さは問わないけど...

	CPU	単体の理論性能	1ノードあたりのCPUの数
SQUID	Intel Xeon Platinum 8368 Icelake / 2.40 GHz 38コア	約 3,981.2 GFLOPS	2基
OCTOPUS	Intel Xeon Gold 6126 Skylake / 2.6 GHz 12コア	約 499.2 GFLOPS	2基

計算計画の概要の書き方

- 研究の背景，目的を端的に記述し，提案手法である計算での計算機の必要性および必要なノード時間（見積り）を記入.
- **ノード時間の見積りは現在稼働している計算機より算出．以下の項目を根拠とする**
 - FLOPS
 - 並列数に依存する場合，ノード数，コア数

研究課題の概要

- 学術的背景、目的・意義、研究計画、萌芽的研究である理由(若手・女性研究者萌芽支援)あるいは期待される成果(大規模HPC利用枠)を，それぞれの位置づけがわかるように記述してください
- とあるものの，一番書いてほしいのは当該研究で「**なぜ大規模計算機が必要なのか**」
- 通常の計算機では終わらない理由（現在のデータセットが〇〇TBで，今回の試行では〇〇PBとなり，研究室内のクラスターでは現実的な時間では終了しないことが予測される．．等）

計算計画

- 利用を想定するCPU数, メモリ量, 計算時間, および, 必要なディスク容量などについて研究計画と関連づけて記入してください。
- 実際に利用したいノード時間を書いてもらえればよいが...
- 計算機資源は有限であるため**提供資源量からおおよそ1課題につくノード時間を目安に要求**

		若手 (3-5課題)	HPC (1-2課題)	AI (1-3課題)
SQUID	CPU	2,600,000	1,300,000	700,000
	GPU	70,000	40,000	20,000
	ベクトル	60,000	30,000	15,000
OCTOPUS	CPU	210,000	210,000	100,000
	GPU	30,000	30,000	20,000
	Xeon Phi	40,000	40,000	20,000
	大容量メモリ	2,000	2,000	1,000

ノード時間

悪い例

メタゲノム解析において、様々な生物種の類似性を示唆するためにはタンパク質の解析において構造予測をする必要がある。複数の微生物において網羅的・大規模に行うため、本研究では計算資源（SQUID）を2,600,000ノード時間必要とする。

数的根拠があげられていないためNG

大腸菌の全タンパク質は4000と言われており、1タンパク質-1低分子化合物の解析では600GFLOPの**計算機1ノードにて5～10時間ほどかかる**。試行回数が必要となり20回の試行を各タンパク質-1低分子化合物の解析を行うとして...

と必要なノード時間の数的根拠を記述するとよい

若手研究での提供計算機資源のノード時間の上限値になっている。
これではこの研究を通すと他の研究に割り当てる資源がなくなる。なので
若手研究の場合、3－5件の枠であるため1件あたり500,000 - 800,000
ノード時間の要求でないと難しい

まとめ

- 計算計画の概要，研究課題の概要，計算計画の書き方（考え方）
 - 要求するノード時間の数的根拠を示す
 - 研究でなぜ大規模計算機が必要なのかを示す
 - 要求するノード時間は，有限なノード時間上限を目安にしていきたい．．