

登録機関の利用者支援について ～ これまでの経験を踏まえて ～

平成21年6月

(財)高度情報科学技術研究機構

1. 利用者支援とは

■ 従来の利用者支援とは

- ・ 高度なスーパーコンピュータに関わる利用技術、環境を支援
計算資源の提供、配分、プログラム相談、講習、広報等が主。
アプリに係わる技術問題 ⇒ 利用者の自助努力で解決が基本

■ 次世代スパコンの利用者支援は従来に比べ何が違うか

- 世界最先端の新科学技術(発見、発明、新産業)を拓くことに利用
⇒ 新しい理論をベースにした複雑、大規模なシミュレーション問題を扱う例が多いと予想
- 並列規模が拡大(約10~20倍)し、世界最大級の超並列環境になる。
⇒ 利用には高度な計算科学技術が必要と予想

次世代スパコン



従来スパコン

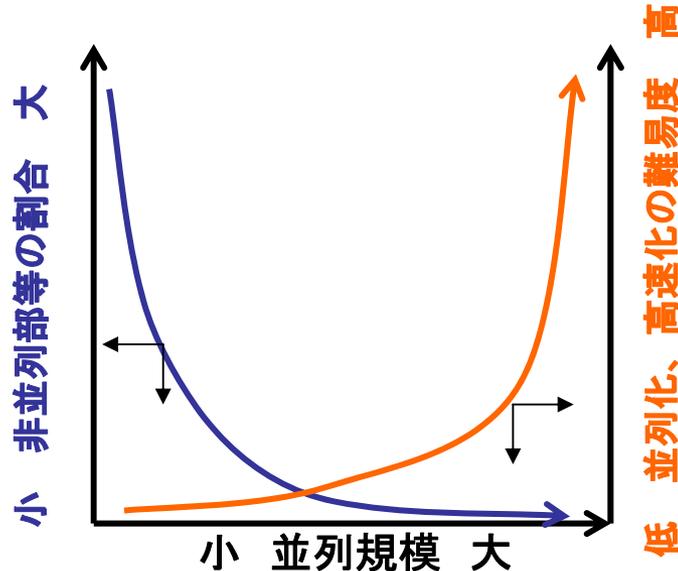


“豊富な経験と
技術が必要...”

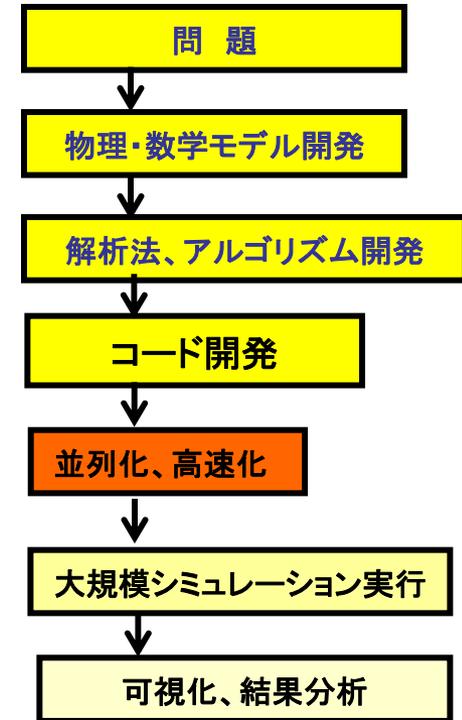
□ 具体的に、次世代スパコンの利用者支援には

- ① 幅広い先端科学技術アプリの並列化、超並列化、高速化、可視化などの技術支援
- ② 超並列に向けた新モデル、新アルゴリズム開発などへの支援にも幅広く対応する必要があるだろう。

⇒ 利用者のシミュレーション工程に積極的に関わり、
計算科学技術問題の解決を支援する。



超並列では非並列部、通信の割合を極力下げる
⇒ 並列モデル、アルゴリズムの工夫 ⇒ 難易度が高い



従来とは異なる利用者支援のあり方が望まれる。

⇒ この観点を踏まえた、利用支援の「巧い仕組み」設計が必要

2. 利用支援の仕組みの設計要件



■ 要員について

幅広い経験と能力を有した人材が必要。

- (1) 大型スパコンを利用したシミュレーションを経験、その能力を有する。
- (2) 大規模計算科学技術コード開発、その並列化、高速化の経験を有する。
- (3) 科学技術分野で研究、開発を経験し、その能力を有する。

■ 機能について

- (1) 利用支援項目の重点化
- (2) 戦略、一般課題の広範囲な分野のアプリに対応できること
- (3) 支援が有効に機能するマンパワーがあること
- (4) 関係する機関、研究者との一体的な連携がとれること
- (5) 人材・技術を蓄積、伝播、継承できること



3. 要件の具体化

3. 1 要員に求められる経験と能力(例)

■ 全般的に

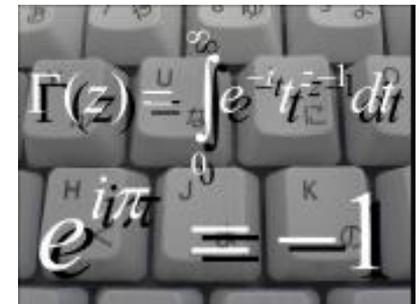
- 科学技術分野の素養、専門分野で実社会経験があればなお可。
- 並列計算機、ソフトウェア、ICTの知識と技能
- アプリの背景にある理論、算法等を理解する物理、化学、数学能力
- 語学(英語)能力(ある程度の)・・・海外専門家との討論、交流能力

■ アプリ経験及び知識など

- 領域・・・ナノ、ライフ、工学(製造技術)、航空、原子力、防災、基礎科学、地球科学、宇宙・天文、計算機科学・・・
- 算法など・・・FDM(差分法)、FEM(有限要素法)、FVM(有限体積法)、スペクトル法、MD(分子動力学)、MO(分子軌道法)、DFT(密度汎関数法)、DEM(離散要素法)、モンテカルロ法、データ同化法、・・・
- カーネルなど・・・疎行列、密行列、構造格子、非構造格子、スペクトル、N体問題、モンテカルロ法・・・

■ 並列計算機の経験と知識

- いろいろな機種 of 並列計算機を利用した経験が望ましい。
・・・例えば、国産機、米国機、スカラー機、ベクトル機、PCクラスターなど
- 大規模シミュレーションの経験が望ましい
- 多くの言語系の対応できることが望ましいが・・・FORTRAN、C言語など、
- 並列ライブラリ、MPI、OpenMPなどの経験が望ましい
- 可視化処理の経験、知識



これら条件を満たせる人材は少ないのが現状・・・

3.2 利用支援組織の機能

■ 利用支援項目の重点化

共用法で示す利用者支援業務のうち、

- ① 課題選定業務の支援業務・・・選定委員会への支援
- ② 研究実施相談業務 ⇒ **この業務を重点化**

■ 戦略、一般課題などの広範囲のアプリに対応できること

- ① グループ制をとる ⇒ 一人で多様なアプリ調整に対応することは無理！
異なる得意分野を持つ専門家同士が補完しあい問題解決に当たる方式
- ② 複数グループで平行して多くの課題に対応：
例；ナノ、バイオなどその分野に得意な専門者が
リーダーとなって、グループを組む。

③ グループの構成員

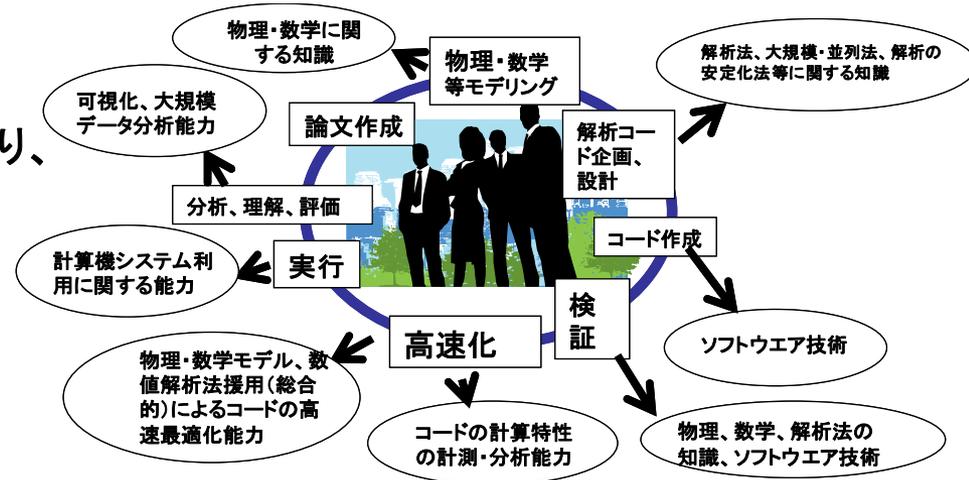
・「計算科学支援員」・・・グループリーダー層。
各分野アプリの並列化、高度化などに経験があり、
スパコン特性も把握している。

アプリ高度化支援、講習・指導ができる。

・「ソフト技術支援員」：

プログラミング、システムエンジニアリング、
可視化などに経験と技能を有する。

利用者の問題に相談・技術対応ができる。



また、利用者環境を支援するものとして

- ・ ネットワーク管理者(計算機接続、データ転送管理等)、
- ・ 情報安全管理者(不正アクセス、情報漏洩防止管理の保全を担当する要員)、が必要。

■ 支援が有効に機能するマンパワーがあること

● 利用支援対応のイメージ

相談課題の難易度に応じて以下のように対応

- A アドバイスレベル： 比較的容易な並列化の場合
→ 研究実施者の努力を支援
- B 並列化支援レベル： 並列化が比較的難しい場合
→ 研究実施者と共同作業、または作業分担
- C 課題解決レベル： 新並列化手法などを開発する場合
→ 研究実施者と共同的に作業

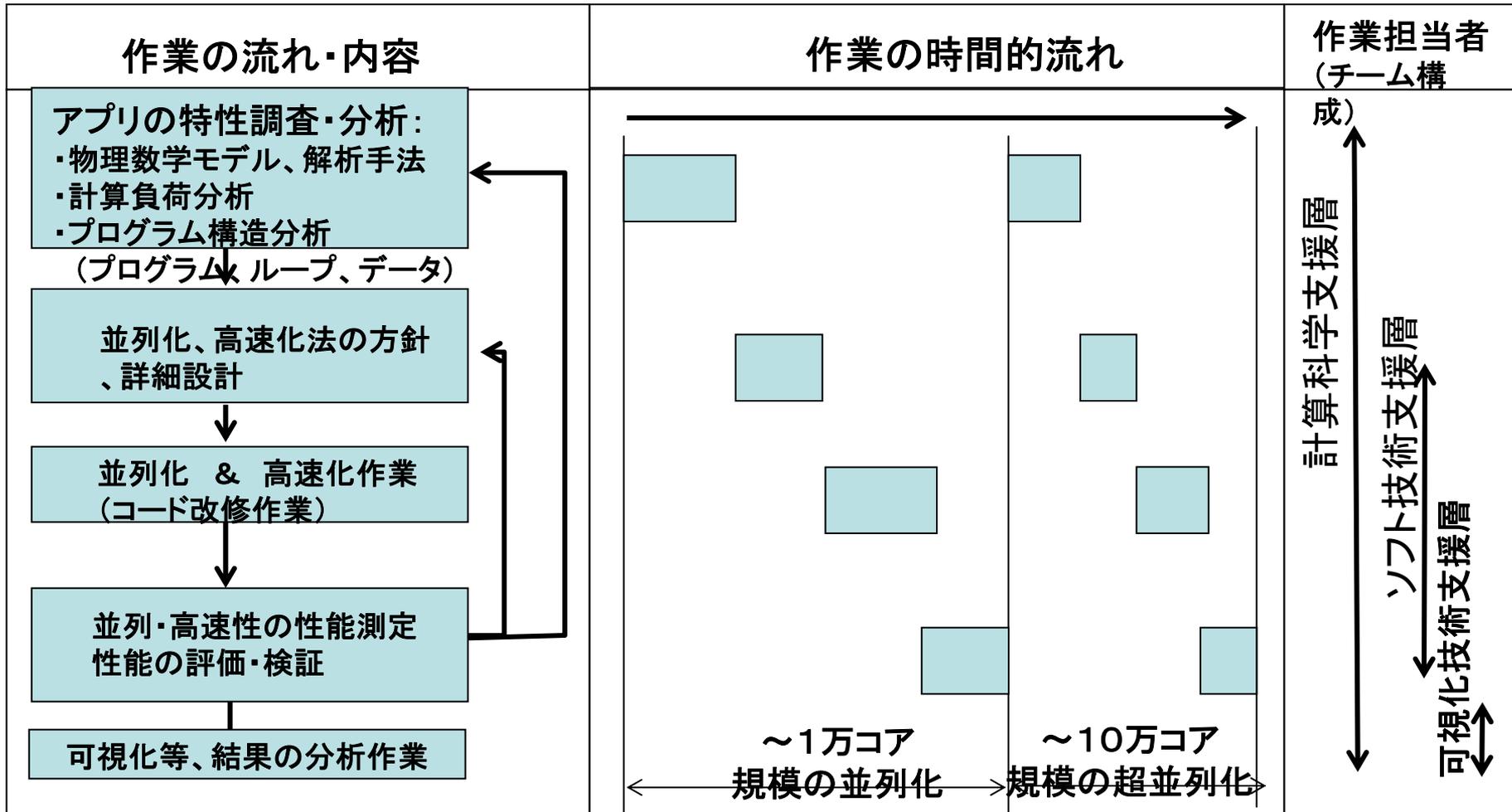
⇒ Bの対応が多いものと想定。

● 効果的な人員数、期間、対応アプリ数を試算する。 (条件)

- アプリのプログラム規模(行数)：
- 並列化調整の程度
 - … 逐次処理から並列化(数万並列)、さらに超並列化(約10万並列以上)
 - … 並列処理から超並列化



逐次処理プログラムからの超並列化



* 既並列コード⇒超並列化: 既並列コードの特性調査、超並列化検討などを必要とするが、逐次処理から並列化する場合の約2/3の期間で対応可と予想。

支援能力の試算(概数として)
 一つのプログラムを超並列化するときに必要な人数、期間
 コード規模(行数):平均で約5千から5万行と設定

並列化プロセス	並列粒度	期間(月)	人数(人/月)
逐次 ⇒ 並列、超並列	～1万規模	6～12ヶ月	4人/月
	～10万規模	9～18ヶ月	
並列 ⇒ 超並列 (約1万規模から)	～10万コア規模	6～12ヶ月	

戦略分野【約5分野】を想定し、

- ・ 4名/チーム (経験から適切と思われる構成人数)
 - ・ 1チーム/分野 × 5分野 × 4人/チーム … 20人程度で処理
- ① 逐次処理コードを、超並列化(約10万コア)にする場合
 (約2課題/年/チーム)×5チーム/組織 … 10 課題/年 程度の対応
- ② 並列コードを超並列化する場合、
 (2～3課題/年/チーム)×5チーム/組織 … 10～15 課題/年 程度の対応

■ 関係する機関、研究者と密接な連携がとれること

- 運用体と密接な連携が不可欠
 - …支援の柔軟性、機動性を確保するために
(組織、担当間の)「壁」をなくすことが重要。
- 研究実施者との密接な意思疎通を図ること
- 専門家の創意工夫を活かせる柔軟な関係
- 目標を共有すること。(世界一、世界初、新物質発見…)



■ 人材・技能の蓄積、伝播、継承が図れること

- 並列化、高速化等の知識、技術経験の蓄積
- 技術ノート、記録、論文などの保存、公開、講習
- 並列化、高速化などの作業に学生、若手研究者、企業技術者などヘインターンの機会を提供。
(例:新しい産業技術分野として興隆を期待)



まとめ

次世代スパコンの利用者支援について

- ◆ 次世代スパコンの目的、特性を活かすには、研究実施者のアプリ群を適切に**超並列化するなどの**計算科学技術面の支援が重要となる。
- ◆ その支援には**有能な人材**を集め、**効果的に機能する**ような組織作りが必要である。
- ◆ 試算すれば、約**20人前後の要員**が必要。約10～15課題／年程度の対応が可能か。
- ◆ 運用体と利用支援とは、**目標などを共有**し、密接な連携を図ることが望ましい。
- ◆ 次々世代計算科学へむけ、利用支援の人材、技術等の**蓄積、伝播、継承**を図る仕組みも必要。

参 考：これまでの主な経験

- (財)高度情報科学技術研究機構(RIST)の目的
高度計算科学技術の推進と普及(中立、公益)
- RISTの主な経験
 - 高性能大型並列計算機の利用研究
 - ・ 主に地球シミュレータ、東大、原研等の並列計算機など
 - 大学、研究機関の研究者との並列コード開発
 - ・ 科学技術庁 科学技術振興調整費「高精度の地球変動予測のための並列ソフトウェアに関する研究」(H10~H14)
 - 大学、研究機関の研究者との大規模シミュレーション研究
 - ・ 文部科学省ITプログラム「戦略的基盤ソフトウェアの開発」、「戦略的革新シミュレーションソフトウェアの研究開発」、「イノベーション創出の基盤となるシミュレーションソフトウェアの研究開発」
 - ・ ES共同研究利用枠: ナノ材料、DDS, 超伝導シミュレーション研究会
 - 民間企業研究者との共同利用研究
 - ・ ES産業利用枠: 新奇材料開発
 - 研究法人からの委託研究
 - ・ 熱流体、構造、量子力学、核融合などのコードの並列化、高速化、高度化
 - 大型計算機システムの運用支援、遠隔ネットワーク運用支援 など