

ソフトウェアと並列計算機 – 研究リソースの全学提供

宮下 夏苗

情報社会基盤研究センター

概要

近年、一般の PC がより安価に、手軽に手に入るようになってきた。タブレット端末も広く普及し、高性能化しており、大学院での教育、研究に利用する端末としてそれらを個人で所有する学生も増えている。情報社会基盤研究センター(以下情報センター)はこれを踏まえて全学に向け、個人用の PC、タブレット端末、学生用の共用端末などさまざまなデバイスからアクセスできるもの、個人で所有、利用することの難しいもの、高額なものなど、学生、研究者の教育研究に直接役立つ様々なサービスを統合的に構築、運用している。本稿では例としてこのうち2つのサービスを取り上げ、2014年の運用について述べる。

1 並列計算機サービス

研究分野によっては実験、データ取得に利用する計算リソースの性能が研究成果に大きく影響する。大量の計算を同時に行えればより多くのデータを取得し、実験精度を向上することができ、大容量のメモリを利用できれば、それだけ大規模な系を研究対象として扱うことができる。情報社会基盤研究センターは計算リソースを全学サービスとして集中管理、運用することで、全学の研究者、学生が個々の端末から自由にアクセス可能な潤沢な計算リソースを提供するとともに、研究者を機器管理、インストール、メンテナンスの手間から解放し、ジョブスケジューラによって個々の研究者の必要とする計算リソースを公平かつ効率的にサービスしている。

1.1 導入

2014年3月、情報社会基盤研究センターはこれまでの Appro PC クラスタ、NEC SX9 に代わる CX250 PC クラスタを導入した。

この導入に伴い、計算ジョブ投入の際のリソース配分の設計と動作チェックを行った。またユーザ向けサンプルスクリプトの作成、利用マニュアルとなる Web サイトを作成した。

また、ソフトウェアとしてはこれまでの PC クラスタでサポートしていた Materials Studio, Matlab, SX-9 でサポートしていた Gaussian09, GaussView をそれぞれ CX250 PC クラスタにインストール、設定、動作テストし、これらのサンプルスクリプトと Web マニュアルも準備した。パソコンとは違うクラスタ環境でどのようにインストール、設定するべきか、事前テストを行い手順を調査した。

1.2 メンテナンス/サポート

並列計算機はシステム自体一般的な PC とは違い、パーツも比較にならないほど多く、様々な技術を組み合わせられて構築されている。単純なパーツ故障であればアラートメール等で故障を即時認識できる場合も多いが、想定外のトラブルが起きる場合もある。ユーザからの問い合わせ、あるいは日常の監視状況から問題を確認し、サポートベンダーに情報を提供して解決を図っている。

問題の解決にシステムの停止等が必要な場合はベンダーとの日程調整とユーザアナウンスを行い、問題解消後に動作テストを行っている。

また、基本的なソフトウェアをインストールしたり、ライブラリ追加をサポートベンダーに依頼することもある。

2014年度は以下のようなソフトウェア/ライブラリなどをインストールし、利用法を調査した。

- Python with PC クラスタ
- OpenCV with PC クラスタ
- LAMMPS with Cray XC30
- Materials Studio with Altix UV1000

2 ソフトウェアサービス

フリーウェア、商用を問わず、適切なソフトウェア、ライブラリを利用することで研究の効率は大きく変わる。ある種の研究を進めるために、何らかのソフトウェアの使用がデファクトスタンダードとなっている場合もある。しかし、それを実際に計算機上で利用するためにそれなりの設定、コンパイルに関する知識を要するものもあれば、動作要件に制約のある場合もある。また、そもそもソフトウェア自体の購入、保守費用がかかる場合も多い。情報センターではこれらを統合的にサポートし、個々のソフトウェアについてフローティングライセンスの導入、共用機器へのインストール、設定など、どこからでも誰でも自由に利用できるソフトウェア環境の構築に取り組んでいる。

2.1 ライセンス

実際の利用率に基づいて、いくつかのソフトウェアについてはライセンスサーバを立て、ライセンスの管理と更新を行っている。各ソフトウェアについておおよそ年一回以上のライセンスの更新に基づいて、ライセンスサーバ上の各種ライセンスツール、およびライセンスファイルの更新を行い、実際にソフトウェアを動作させて更新の完了を確認している。集中管理しているソフトウェアとして、以下のようなものがある。

- Matlab, MaterialsStudio, Mathematica, SPSS, Intel コンパイラ, PGI コンパイラ etc.

2.2 利用率集計

年度末を目処に、ログをもとに MaterialsStudio, Mathematica などの利用率集計を行った。実際の利用率をもとに次年度のソフトウェア購入、保守継続について検討している。

2.3 インストール

共用 Linux 機、並列計算機を中心に、各種ソフトウェア、ライブラリのインストール方法を調査し、実際に作業を行った。

Linux 機では R と各種のオプションツール、ruby, python, hugs, hol, spin 他、ユーザの希望に基づいて利用環境の構築、コンパイル、インストールを行った。

並列計算機では前述の Materials Studio, Gaussian09, GaussView5, Matlab, OpenCV などのツール、ライブラリをインストール、並列計算機上で利用するための環境設定を行ったほか、ユーザからの希望のあった quantum-espresso などフリーソフトウェアのインストール方法を調査、ユーザに公開した。

3 講習会

ユーザが実際に並列計算機、また各種の便利なソフトウェアを利用してみる端緒とすべく、情報社会基盤研究センターは定期的に実機を用いた講習会を開催している。

昨年は以下の講習会を開催し、それぞれについて講習会の企画、ベンダーとの調整、テキストやサンプル

スクリプトのチェックおよび Web サイト作成, メールアナウンスによる広報を担当した.

[2014/Spring]並列計算機利用者オリエンテーション

[2014/Spring]Fujitsu CX250 初級者講習会 並列計算機で使うオープンソース(1)

[2014/Spring]SGI AltixUV1000 利用者講習会 並列計算機で使うオープンソース(2)

[2014/Spring]Cray XC30/MPI Programming 初級者講習会

[2014/Autumn]MATLAB Training Workshop

[2014/Autumn]Materials Studio Training Workshop

[2014/Autumn] Python @ CX250 Clustr User Workshop

[2014/Autumn]Cray XC30 講習会

3.1 企画

従来は MPI, OpenMP などの並列化プログラミング講習を主に開催してきたが, 近年ソフトウェア利用者が増え, 逆に自らプログラミングするユーザが減ってきていることから, これまでのプログラミング講習に代えてオープンソースのコンパイルをテーマにした講習会を 2 件企画した.

利用するオープンソースソフトウェアはコアユーザと実際に会話しながら決定し, ベンダーに講習会開催の協力を依頼した. 実際に必要なライブラリのインストール, 動産検証, サンプルプログラムの準備, 動作テストも行った.

また, ソフトウェアベンダーに依頼して, Materials Studio, Matlab など利用率の高いソフトウェアの操作方法, 利用方法をテーマとする講習会を企画した.

3.2 準備

講習会開催のために各講習の Web ページ, 参加登録フォームを用意し, 開催を知らせるメールアナウンスを行った. また, ソフトウェアライセンスを借りてライセンスサーバに適用する, 貸し出し用端末の利用方法やテキストダウンロード方法を記載した当日配布用の資料を作成するなど, 開催に向けた準備作業を行った.

4 まとめ

情報社会基盤研究センターが全学に向けたリソースの提供を行うにあたり, 以下の点が非常に重要となる.

1. 利用の簡単さ
2. アクセスのしやすさ

この視点から, 以下のような点を今後の課題と考える.

4.1 利用の簡単さ

講習会の開催, Web マニュアルの作成などユーザがリソースを利用するための導入にあたるコンテンツは適宜用意しているが, ユーザの利用方法は多岐にわたり, 動作させたいプログラムは様々, それに応じて適切なリソースの選定も必要となってくる. より多種のソフトウェア, 実験ツールについて自分自身がまず学び, それらに関する利用マニュアル, サンプルプログラムをより充実させたいと考える.

現時点で必要最低限のマニュアルは用意しているが, 利用するリソースサイズを変えた複数のサンプルを用意する, オプションの記載を充実させる, Web だけでなくノード内にテキストで読めるマニュアルを準備しておき, ターミナルからテキストでマニュアルを閲覧できるなど, 初学者から一般の利用者までユーザがわざわざ使い方やオプションを調べなくても利用できるようにしておきたい.

また, python, Java などこれまでサポートしてこなかったツールについてはマニュアルも用意できていない

状況にある。

これらのサポートも追加し、初めて計算機、ソフトウェアを利用するユーザが迷うことのないコンテンツを準備したい。

4.2 アクセスのしやすさ

ユーザが利用する端末は PC からタブレットまで様々である。どんな端末を使っても、自分がどこにいても、実験としての大規模計算や各種ソフトウェアの使用に不自由がないようにしたい。

Windows ソフトウェアの使用にあたっては、すでに当センターで運用している **Jaist Cloud Desktop** などがこの目的に対するひとつの成果と考えるが、大規模計算機については、ターミナルを開いて各々のログインノードに `ssh` ログインする、コマンドを叩いてジョブを投入し、データを取得するなどの煩雑さがある。

これに対して、ジョブの状況を確認する、実際にジョブ投入を行うなど、計算開始までのステップをテンプレートとして簡略化し、Web ブラウザやタブレットのアプリ経由で行えるようにするなど、改善を図りたい。

2015 年 9 月には、PC クラスタコンソーシアムが企画する第 1 回並列プログラミング講習会 in 金沢が執り行われる。また、金沢工大でも東大 FX10 が無料で利用できる **OpenFoam** 講習会が開催されることとなった。外部の大規模計算資源の利用が、研究活動においてより一般的になってきているということではないだろうか。

しかし、サイトの外にある資源を申請しての利用は一般的なユーザにとって必ずしも敷居の低いものではない。また、そもそも自分の研究に計算機の利用、ソフトウェアの利用を組み合わせることができると、ユーザ自身で考えつかない場合もある。

初心者もわかりやすく取得できるリソースがいつも身近にあることで、ユーザが自分のパソコンと自分で購入したソフトウェア以外の研究リソースに触れる、試せる機会を大きく広げることができる。もちろん、すでに利用しているユーザにとっては、研究、実験において必須のファクターとなっている場合が多い。逆にどんな優れたリソースがあっても、使い方が判らなければそれは絵に描いた餅にすぎない。

JAIST の計算リソースをより身近に、わかりやすく、使いやすく提供することが情報センターの業務のひとつである。