

VLBI@home ~分散型ソフトウェア関連システムの開発

竹内央、近藤哲朗、小山泰弘、中島潤一、木村守孝 (鹿島宇宙通信研究センター/CRL)
高島和宏 (国土地理院)

概要

近年の汎用 PC の性能向上により、従来のハードウェア関連器の代替として、PC によるソフトウェア関連器の使用が現実的になって来た。通常の用途で使用されている PC の余剰リソースを有効に活用すれば、安価に高速分散処理が実現できる。そこで SETI@home 等で採用されている server/client モデルによるスクリーンサーバ型の分散処理を用いて VLBI 関連処理を行うシステムの開発を行った。

1. はじめに

表 1 に CRL 鹿島グループ内でプリンタサーバや mail,プログラミング等の一般的用途で使用されている 5 台の PC について、平日の 5 日間にわたり CPU の使用率をモニタした結果を示す。高々10%程度のリソースしか使用していない事がわかる。この結果からも分かる通り、多数の PC が備わる研究機関等には、大きなコンピュータリソースが眠っていると言える。この余剰リソースを有効に活用する例として SETI@home 等で採用されてい

	個人 PC1	個人 PC2	個人 PC3 (昼のみ)	共用 PC1	共用 PC2
平均使用率	3.90%	10.19%	11.57 %	9.47%	7.09%

表 1 : PC の CPU 使用率モニタ(2003/12/1 ~ 12/5)結果

る server/client 型の分散処理がある。Server/client 型の分散処理は、大量のデータをこま切れにして多数の client PC に送り、各 PC でデータに対して何らかの演算処理をさせ、その結果を server に返すという 3 つの機構を有する事で特徴付けられる。この方式の分散処理が有効に働くための必要条件の一つとして処理すべきデータの転送時間が各 client でそのデータを処理するのに必要な時間に比べ短い事が挙げられる。現在 CRL において開発されているソフトウェア関連器は、XF 型 1 bit 32 ラグの関連処理を 4Mbps で行う (Pentium3 1GHz 相当の PC を使用) [近藤他 本集録]ため、10Mbps,100Mbps 程度の回線速度を持つ環境では、VLBI 関連処理に対し server/client 型分散処理が有効になり得る事がわかる。そこで、VLBI 関連処理を行う分散処理システムとして SETI@home 型の方式を採用しシステムの開発を行った。

2. システム概要

図 1 にシステムの構成図を示す。システム全体を管理する 1 台のコントロールサーバと受信した VLBI 生データを送信するために各観測局に配置された FTP server、及び多数のクライアントによってシステムが構成される。スクリーンサーバ型の client は、起動する

とまずコントロールサーバに対して処理するデータの所在を XML 通信により問い合わせる。コントロールサーバは問い合わせに応じて、処理すべき基線に対応する 2 局分の FTP サーバの URL と、処理すべきファイル名を client に返す。client は指定されたアドレスから処理すべきファイルをダウンロードし、相関処理を行う。最後に相関処理結果を返すと

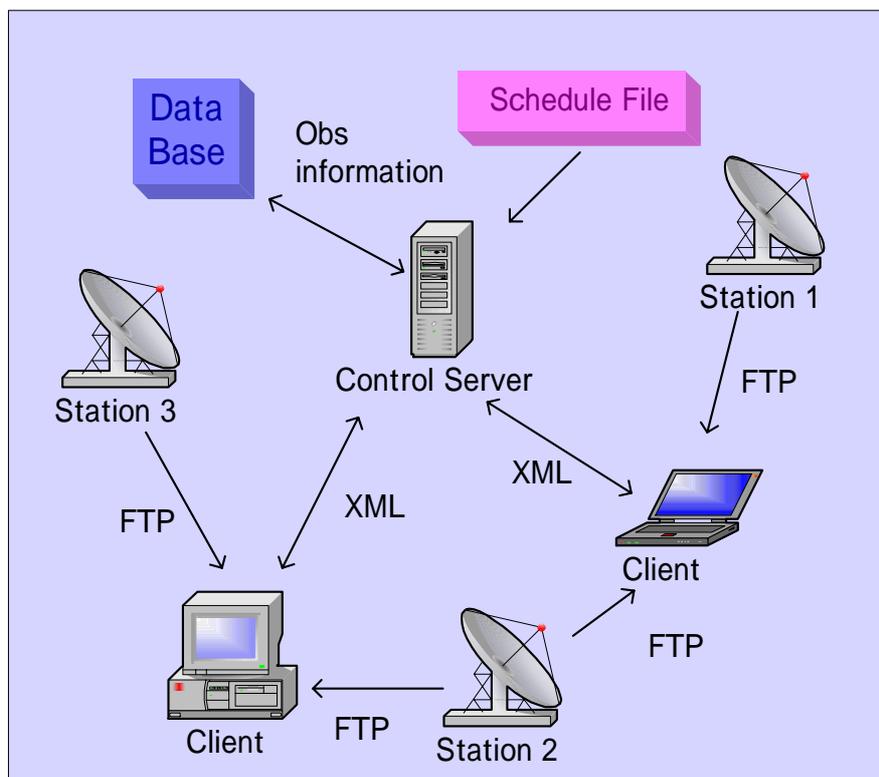


図 1 : VLBI@home のシステム構成図

共に、データ転送及び相関処理に要した時間を報告する。図 2 に、コントロールサーバの概念図を示す。処理すべき obs の情報を待ち行列(図 2 上)に保持し、client から接続があるたびに順番に要求を渡していく。Client に渡された obs 情報は相関処理終了待ち FIFO(図 2 下)に移され、処理終了の報告があるまで保持される。相関処理待ち FIFO が空の時に client から接続があった時は相関処理終了待ち FIFO の先頭の情報を client に渡す事により、いつまでも処理が終わらない obs が出る事を防いでいる。コントロールサーバ内のデータ構造を LIST 構造により構成する事により最小限のオーバーヘッドで必要な機能が実現される。この事は、将来 client 数が莫大な数になった時に大変重要になる。次にスクリーンセーバ型の client プログラムの仕様について示す。相関処理を行う相関部は XF 型のソフトウェア相関プログラム cor.exe [近藤他 本集録]を外部実行する事により実現している。各 VLBI 局の FTP サーバとの間のデータ送受信部は、プログラム内部の FTP クライアント機能により実現している。コネクション毎に独立したスレッドを立ててデータ通信を行うため、複数任意個のファイルを同時にダウンロード可能である。また FTP 伝送が途中で中断され

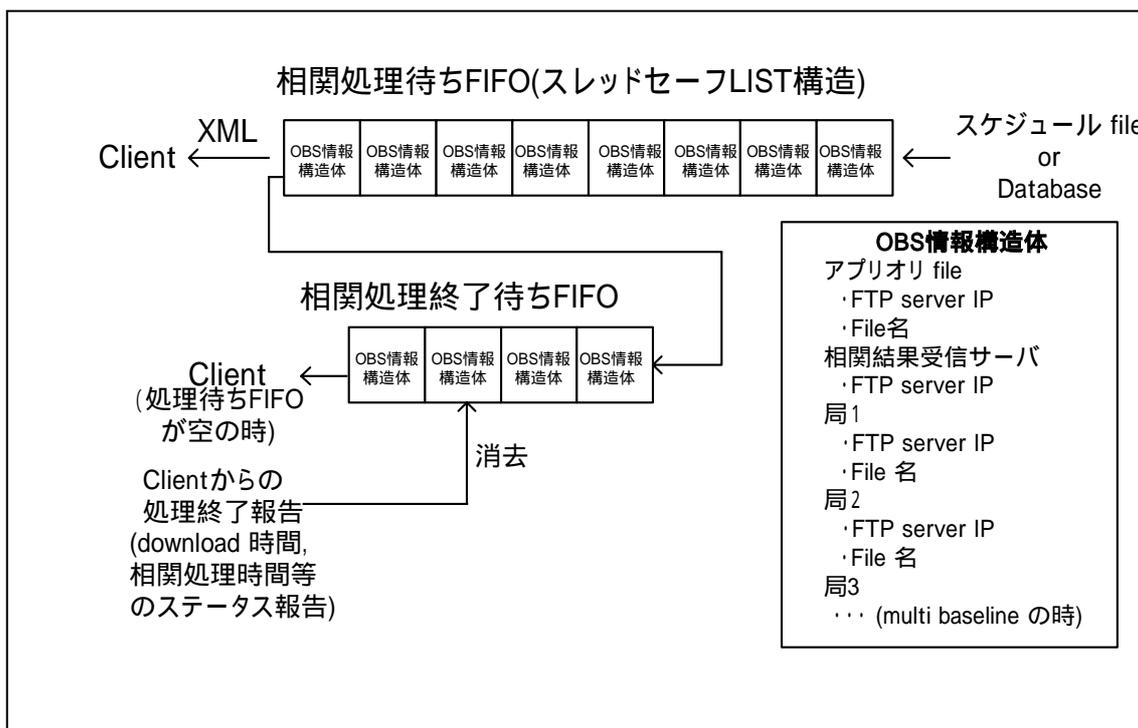


図 2：コントロールサーバの模式図

た際には、再開時に再度中断箇所からダウンロードをやり直す resume 機能に対応しており、処理が中断した時のオーバーヘッドを最小にしている。また FTP 送受信機能を内部に持つため、受信したデータを HDD に書き出す事なくメモリに直接取り込む事が可能である。将来、相関処理部の速度が 100Mbps 程度になった時は HDD 書き込みがボトルネックになるが、メモリ内に保持したデータを直接相関処理する事によりボトルネックを回避できる。コントロールサーバとの間のコマンド通信は XML socket により実現している。コマンド通信部では処理すべきファイルの URL を受信し、相関処理時間、ダウンロード速度等の情報を送信する。

3. システムのボトルネックと改善策

本システムのボトルネックになりうる要素として、クライアント処理速度、ネットワーク回線速度、ファイルサーバの送出速度の 3 つの要素が挙げられる。ボトルネックを改善しシステムの向上を進めるためには、3 つのうちどの要素がシステムのボトルネックになっているか常に把握できるようにする必要がある。Client から報告される相関処理時間、ダウンロード時間、及び、FTP サーバのデータ転送速度の履歴を常に記録しデータベース化する事によりこの目的を達している。現在の典型的なクライアント処理速度は 1CPU の PC1 台あたり 2~12Mbps である。クライアントの処理速度がシステムのボトルネックである場合はクライアント数増加、相関プログラム高速化が必要になる。ネットワーク回線速度の典型値は 10M, 100M, 1 ~ 数 Gbps / 回線 である。ダウンロード時間が処理時間に比べて極端

```
File download started.  
Downloading R1970007.dat from byakko.cr.lgo.jp: 13991936/20000040bytes: 414.52 KB/s  
Downloading H1970007.dat from seiryuu.cr.lgo.jp: 14548992/20000040bytes: 434.71 KB/s
```



図3：Client プログラムのスクリーンショット。相関処理に必要な2局分のデータを同時にダウンロードしている時の図。

に遅いクライアントが存在するとネットワーク環境がシステム全体のボトルネックになり得る。ネットワーク環境を高速化・最適化するか、若しくはそのようなクライアントに対しては単位データ量あたりの処理時間が長いような目的(ラグ数の長い相関処理, 飛翔体探査用の相関処理等)を割り当てる事によりボトルネックを改善できる。ファイルサーバのデータ送出速度は HDD のアクセス速度(ランダムアクセス数百 Mbps, シーケンシャルアクセス 1Gbps 程度)で制限を受ける。ファイルサーバ送出速度がボトルネックになる場合は、FTP サーバを複数ミラー化する、ラムディスク化する(リアルタイム処理の場合)等の方法によりボトルネックの改善が可能である。

4. まとめ

本システムを用いる事により、十分なクライアント数が確保できれば、現在技術的に可能なネットワーク速度 10Gbps 程度を上限として、いくらでも相関処理速度を高める事が可能である。今後、数ヶ月程度で 10~20 台の PC による 128Mbps リアルタイム観測を行い、その後、数百台の PC による 2~3 Gbps のリアルタイム観測を実現したい。