

環オホーツク情報処理システム概要

先端技術支援室 小野数也

環オホーツク観測研究センターからの依頼により、環オホーツク情報処理システム(以下、情報システム)のネットワーク構成、Web サーバインストールなどを行った。図1に構成を示す。

1. ネットワーク概要

ゲートウェイひとつにグローバル IP を設定して、内部ネットワークにプライベート IP を設定してしまうと、一度ゲートウェイにログインしなくては各マシンにログインできない。また、低温研内にそれぞれのマシンを置き、グローバル IP を設定すると、そのマシンごとの管理をしなくてはならなくなり、監視が行き届きにくくなる。その

ためセキュリティと利便性を考えて、本情報システムでは、低温研内にゲートウェイサーバを設置し、その中に環オホーツク独自のネットワーク(サブドメイン)を構築した。そして、各サーバにグローバル IP を与え、直接ログインを可能にした。

図1でのゲートウェイはあくまでもパケットフィルタリング用のマシンなので、踏み台としてのログインは必要ない。本設定では iptables を使い必要の無いポートを閉じて、外部からは ssh、scp の使用のみを許可し、内部からはそれに加え ntp、lpd、smtp などのポートを許可している。

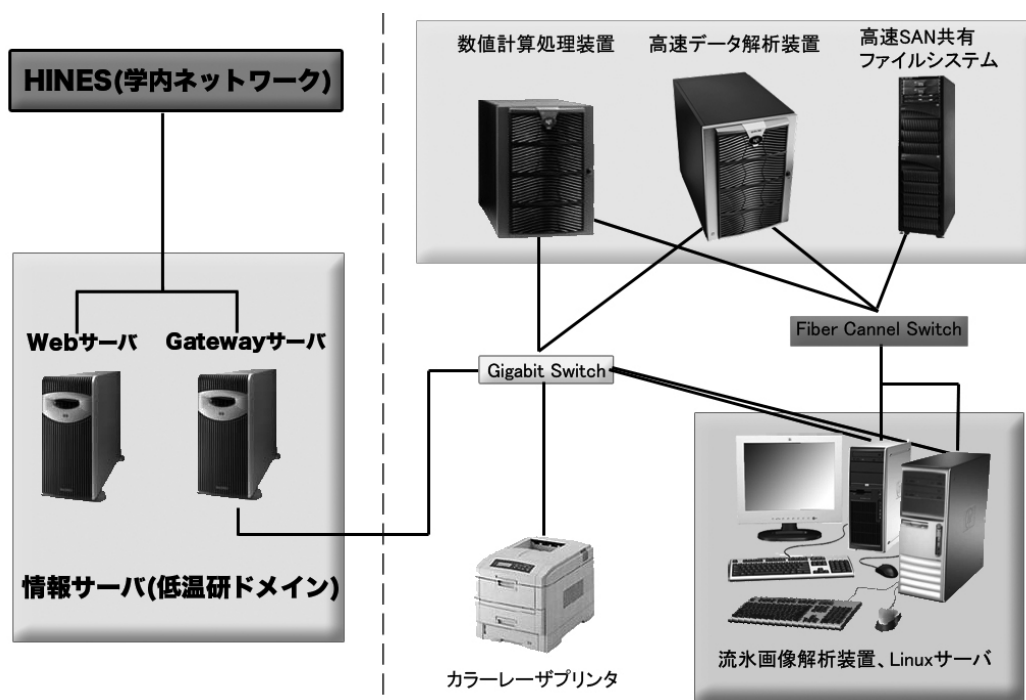


図1：情報処理システム構成図、破線より右側が環オホーツクドメインのネットワーク群である。

2. ハードウェア構成

ハードウェアに関してはセキュリティのこともあり詳しくは表記しないが、SGI 社製のものに加え、hp 社製の製品で主に構成されている。ホストやNISドメインなどの名前には環オホーツクに関連して、「オホーツク海を取り巻く都市などの名前」を採用した。

これらのマシンをファイバチャネルでつなぎ、ひとつのハードディスクを全てのマシンで共有できるようにした。したがって、システム内にあるサーバ間のデータのやり取りは必要ない。あるマシンで並列計算して出力した結果をコピーするという手間を省略して、Windows や IA32 マシンなどの別マシンで解析が可能である。

共有するハードディスクは Raid5+ホットスペアで構成されており、利用可能な容量は約 10TB である。もしハードディスクに不良ブロックなどのエラーが発生した場合は、本体からは警告音を発し、さらにメールで通知するようになってある。

3. パケットフィルタリング

本情報システムではゲートウェイサーバに iptables を用いたパケットフィルタリングを行い、セキュリティを強化している(図2)。ポリシーとしては、すべて drop(捨てる)する。Reject(拒否)とは違い、外部からの反応が何も無いのであたかも起動していないかのように装うことができる。その基本ポリシーに ssh などの必要な通信サービスを許可した。後に示すバッチジョブとのからみでメール通知するため情報システムからは SMTP などを許可した。ほかにも ntp や DNS などもセキュリティを守りつつ許可した。

低温研内(133.87.216.0/22)から本情報システム(133.87.xxx.0/28)に接続しようとしたとき、本来違うネットワーク群にあると予想されるため情報が上位ルータである低温研ゲートウェイを通過しそうに思われるが、133.87.xxx.0/28 に接続しようとしたものは上位ルータの設定を変更することにより、本情報システムのゲートウェイサーバに投げる。低温研外から接続した場合にも上位ルータなどの設定により問題なくログインが可能である。

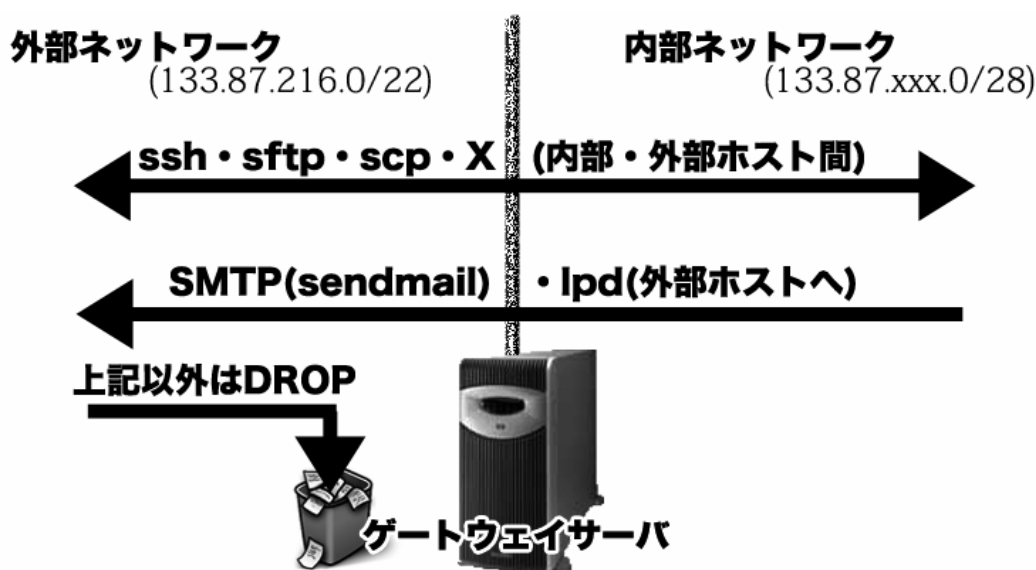


図2: Gateway サーバ プロトコル許可概要

4. バッチジョブ(OpenPBS)

PBS (ポータブルバッチシステム)とは、ネットワーク接続された複数の PC をひとつのリソース群として、そのリソースで有効に計算を行うために開発されたシステムのことである。使用するユーザは、ジョブを実行するために必要な CPU 数、メモリ容量、CPU 時間に応じて、バッチシステムのキューと呼ばれる待ち行列にジョブを投入する。概念図を図 3 に示す。

本情報システム内には二台 OpenPBS を用いたバッチシステムを採用している。利用者が計算するときにはバッチシステムを利用するが、それを使用せず直接計算させようとすると 30 分で計算がストップするように設定した (ulimit -m で)。具体的な方法として、/etc/profile.d/以下にファイルを設置し、利用者に対してインタラクティブな計算の制限をかけた。また本システムで利用できるプログラム言語は f77、f90、f95、c、c++などである。また、性能解析ツールなども入っている。

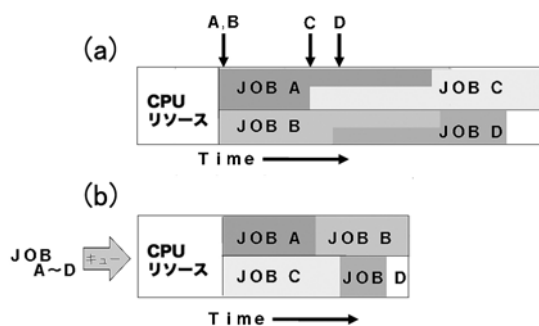


図3: Job 投入方法の概念図。(a) 複数のユーザがインタラクティブにジョブを投入した場合。(b) PBS を用い、キューにジョブを投入した場合。システムの利用状況を見て、自動的にジョブの実行が行われるため、同じ CPU リソースで効率よく計算が行われる。

各ユーザが投入したジョブが正常終了、異常終了したときには、sendmail を使いユーザにメールで通知することも可能である。メールを低温研メールサーバなどに送信させるため、ゲートウェイは内側からの smtp ポートの使用を許可している。

5. 環オホーツク Web サーバ

環オホーツクの Web サーバは観覧用ページのコンテンツを載せるだけではなく、保有するデータ公開や共有にも用いる予定である (<http://www.woc.lowtem.hokudai.ac.jp>)。

海洋短波レーダのホームページでは準リアルタイムで宗谷沖、紋別沖の海洋表層の流速の分布図をアップロードしている(概要は <http://www.lowtem.hokudai.ac.jp/tech/ats/> を参照のこと)。また、紋別沖に設置してある流氷・大気観測レーダに関しても、同様の技術を用いた準リアルタイム更新を目指す。

また、これから環オホーツク観測研究センターが保有してあるデータの公開用のアプリケーションとして、LAS(Live Access Server)を使用する。LAS は IPRC (International Pacific Research Center) と共同でデータの共有を Web 上で行い、さらにデータの観覧に関しては、ブラウザなどを用いて簡単に出来るようにしたものである (<http://apdrc.soest.hawaii.edu/> を参照)。また LAS の日本語化も検討中である。

6. 謝辞

本情報システムのゲートウェイの設定などは住商情報システム株式会社に、サーバ全般(OpenPBS など)の設定に関しては SGI 株式会社に協力していただいた。LAS は IPRC の Shen, Yingshuo 氏に設定していただいた。ここに深く感謝します。