

## NSⅢにおけるネットワークの設計と実装

大川博文 (航空宇宙技術研究所)

### Design and Implementation of the Network System in NSⅢ

by

Hirofumi OHKAWA (NAL)

#### ABSTRACT

In 2002, NAL introduces the NS III, which has 4GB bi-directional crossbar network for interconnection among nodes, 4 striped GSN network (500MB) for high-speed connection between the computing system and the visualization system, and 1Gbps Ethernet network for user communication. In this report, we describe the requirements needed about the network composition mounted in NSIII and show the designed apparatus composition.

#### 1. はじめに

大規模な計算サーバのネットワークを構築する上で重要な点は、各計算ノード間および各計算ノードと I/O ノード間のデータ伝送のボトルネックとならないことである。それに加え、航空宇宙技術研究所が 2002 年に導入した第 3 期数値シミュレータシステム(以下、NSⅢ : Numerical Simulator III と呼ぶ。)においては、現有可視化システムとのシステム連携、外部ユーザの利用経路といったネットワークについても、今後の利用形態の方向性を鑑みると重要である。

本稿では、NSⅢにおいて実装するネットワーク構成について必要とされる要件及び設計方針を述べるとともに、実装している機器構成について示す。また後半において、NSⅢの外部利用に関しての実装構成と、その利用サービスについて述べる。

#### 2. ネットワークシステム要件

NSⅢにおけるネットワークを設計するにあたり、次の 5 つに大別する各ネットワークの要件を設定した。

##### 1)内部ネットワーク

各計算ノードと I/O ノードとを高速に結合する。NWT での採用実績を踏まえ、単段クロスバにより構成し、1 パスでのピーク転送性能が、NWT でピーク性能の約 10 倍となる、入出力それぞれ 4 GB/秒以上を有する。

##### 2)可視化連携ネットワーク

可視化システムとのリアルタイム可視化を可能とする、500MB/秒以上のデータ転送性能を有する。

##### 3)運用ネットワーク

ユーザ利用を目的とする。Gigabit Ether スイッチによる所内 LAN との接続とともに、外部ユーザとのセキュアな接続を可能とする。

##### 4)管理ネットワーク

運用アプリケーションの通信を目的とする。

##### 5)制御ネットワーク

SCF (System Control Facility) 制御網専用ネットワークとして、システムの異常検出およびシステム制御を目的とする。

#### 3. NSⅢのネットワーク構成

NSⅢの全体構成の概要を図 1 に示す。また、ネットワークシステム要件に対する、使用媒体および通信のピーク性能を表 1 に示す。

表 1 ネットワーク種別

ネットワーク	用途	使用媒体	ピーク性能
内部ネットワーク	ノード間 高速接続用	クロスバ	4GB/s IN/OUT 同時
可視化ネットワーク	可視化システム 接続用	GSN	800MB/s×4 ストライピング*
運用ネットワーク	ユーザ利用	1000BASE-SX 100BASE-TX	1Gbps 100Mbps
管理ネットワーク	運用アプリ 通信用	100BASE-TX	100Mbps
制御ネットワーク	SCF 制御網	10BASE-T	10Mbps

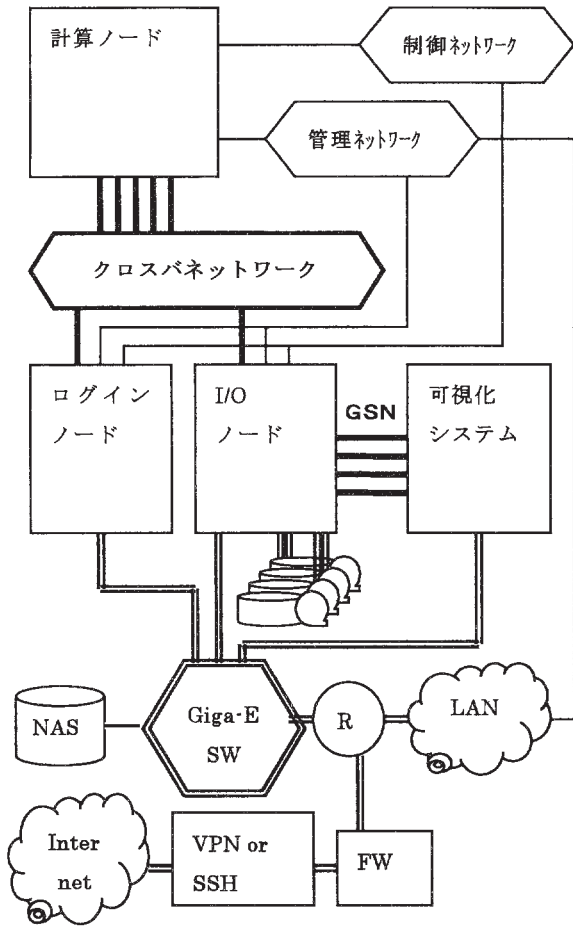


図1 NSIIIネットワーク構成概念図

#### 4. ネットワークシステムの実装

NSIIIを構成する筐体は、各ノードで必要とするネットワークインターフェイスを有するように分割される。筐体のタイプにより保有するネットワークインターフェイスは異なる。表2に筐体のタイプごとに有するネットワークインターフェイスとその数を示し、表3に各ノードの必要とするネットワークインターフェイスの一覧を示す。

表2 筐体の有するネットワークインターフェイス

筐体タイプ	IF種別	IF数
Type A	DTU	4
	100BASE-TX	16
Type B	DTU	4
	100BASE-TX	16
Type C	DTU	2
	1000BASE-SX	4
	100BASE-TX	16
Type D	DTU	2
	GSN	4
	1000BASE-SX	4
	100BASE-TX	16

表3 ノードの有するネットワークインターフェイス

ノードタイプ	IF種別	対応ネットワーク
計算ノード	DTU	内部ネット
	100BASE-TX	管理ネット
I/Oノード	DTU	内部ネット
	1000BASE-SX	運用ネット
	100BASE-TX	管理ネット
IOノード	DTU	内部ネット
	1000BASE-SX	運用ネット
	GSN	可視化ネット
	100BASE-TX	管理ネット

#### 4.1. 内部ネットワーク

内部ネットワークは、NSIIIにおける各ノードのそれぞれをクロスバにて結合する。各ノードは、少なくとも1つのDTU (Data Transfer Unit) をクロスバのインターフェイスとして持ち、計算ノード間のデータ通信や、IOノードを介しての大規模ストレージとのデータ転送を行う。

#### 4.2. 可視化ネットワーク

可視化ネットワークは、IOノードと可視化サーバの間をGSN (GigaByte System Network) を4本使用して結合する。ストレージからのデータはIOノード内でバッファに蓄積された後、ストライピング処理を受けて4本のGSNを利用してSTプロトコルで通信を行う。

GSNのピーク性能としては1パスあたり800MB/secを有しているが、実際にはGSNデバイスドライバおよび、STプロトコルドライバのオーバーヘッドがあり、単体ではリアルタイム可視化の要件である500MB/secを満たさない。そこで、ストライピング処理を行うことで、実行性能として要件を満たした。

#### 4.3. 運用ネットワーク

運用ネットワークは冗長構成をとっており、2系統のGiga Ether Switch に対してそれぞれの要素が接続されている。

運用ネットワークでは、ログインノードおよび可視化サーバを、Giga Ether Switch に対して1000BASE-SXを用いて接続している。また、NSIIIユーザのホームディレクトリが配置されるNAS (Network Attached Storage) もGiga Ether Switch に接続され、小規模ファイル等はこのストレージで管理する。このGiga Ether Switch は所内バックボーンスイッチとも1000BASE-SXで接続され、LANからの高速な利用を可能にしている。

運用ネットワークの外部との接続に関しては、生活線とはファイアーウォールを独立させ、LANの外部接続ポリシーとは別に自由に設定、運用ができるようにしてい

る。外部との接続はインターネット、専用線、ISDN 回線等の利用を想定しているが、その通信において暗号化を行うために、SSH サーバおよび VPN ルータを接続している。

#### 4.4. 管理ネットワーク

管理ネットワークは、NSIIIの管理用のネットワークであり、運用管理に用いる管理 PC と各ノード間のアプリケーション通信に使用される。各ノードは完全に冗長構成をとった形で、2 系統の HUB に対して 100BASE-TX で接続している

#### 4.5. 制御ネットワーク

制御ネットワークはハードウェアの異常検出、自動制御を行う SCF (System Control Facility) のための専用ネットワークとして用いられ、筐体単位で接続をしている。制御ネットワークも同様に冗長構成をとっている。

### 5. NSIIIシステムの外部利用

#### 5.1. 外部接続の概要

NWTでの外部接続方法は、トークンカードにより生成されたワンタイムパスワードを利用してファイヤーウォールにログインし、そこから LAN 内にある NWT に対して telnet を行う形であった。

この場合、ファイヤーウォールへのログインパスワードは、トークンカードのサービスによって作成される使い捨ての共通鍵であるため、外部から悪意を持って侵入するといったことについては困難であったが、途中経路における通信は平文で行われるため、例えば盗聴といった問題については無防備であった (図2)。

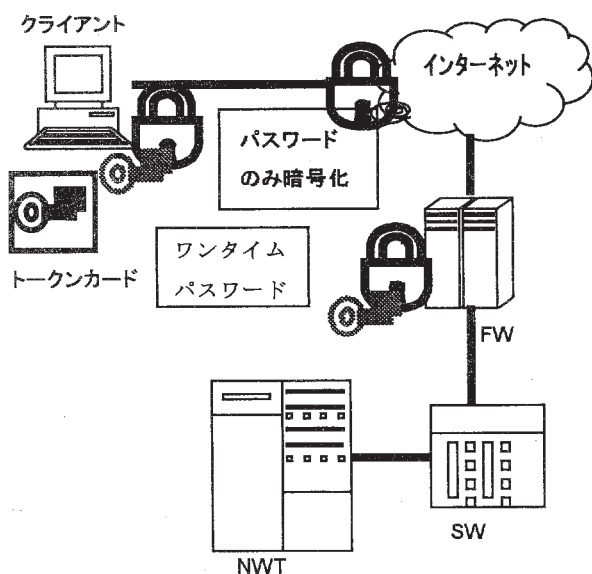


図2 従来の外部接続

それに対して、NSIIIでは外部との通信路において、公開鍵暗号化 (PKI) 方式による暗号化を行うことで、通信経路での盗聴の問題を排している。証明書に関しては、NAL 内の CA 局による認証を行い、同じく NAL 内で生成したユーザ秘密鍵と併せて CD-R 媒体に書き込んで外部利用ユーザに郵送する。ユーザは証明書、鍵をクライアントにインストールして SSH サーバにログインし、暗号路を確立した後で、NSIIIとの通信を行うことになる (図3)。

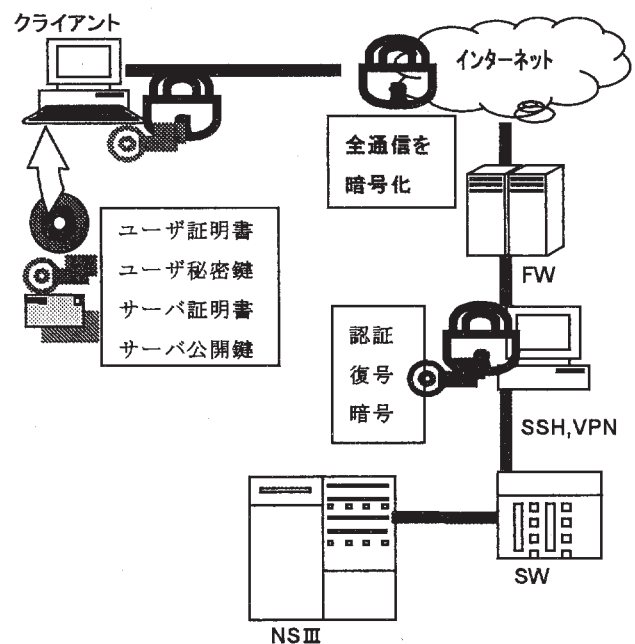


図3 NSIIIへの外部接続

#### 5.2. 外部接続の構成

暗号路確立のためのアプリケーションとして、VPN クライアント、SSH.COM 社製クライアント、OpenSSH、そして Windows 用の端末ソフト Putty を想定している。それぞれ証明書、鍵の取り込みなどをクライアントソフトに設定し、ファイヤーウォールを経由して VPN ルータ、もしくは SSH サーバに接続を行う。その際、認証局に対して証明書が失効していないかどうかの確認がなされ、暗号路が確立される。暗号路確立後は、クライアントから NSIIIのログインノードに任意の telnet クライアントでログインする。NSIIIとのファイル転送に関しては、現在のところ、sep を想定している。

また、セキュアノードに関しては、平文となる以降からは物理的に直接つなぐことで、他者の通信とは分離し、よりセキュアな環境にしている。

クライアントからの NSIIIの操作としては、上記以外に https を用いての Web アクセスを想定しており、JOB 投入、結果の確認などの操作が、ブラウザ上からも可能である。

5.3. 申請手続き

外部から NSIII を利用するために必要な申請書類は表 4 の通りである。

表 4 申請書類

職員	業務システム利用申請書 (職員用) —NS システム用—	L3C487R2
職員外	業務システム利用申請書 (職員外用) —NS システム用—	L3C488R2
	誓約書	L3C48BR1

申請書フォーマットについては、NAL イン트라ネットおよび、NSIII ホームページからダウンロードすることが可能である。

申請者には、航技研との契約 (共同研究等) の航技研側受入担当責任者を通して、NSIII システムの受付窓口にて申請書を提出していただく。連絡を受けて、管理者はユーザ秘密鍵の生成およびユーザ証明書を発行し、CA 証明書、サーバ公開鍵、サーバ証明書とともに CD-R 媒体に記録する。NSIII ユーザ情報の追記された申請書の控えと、鍵、証明書の記録された CD-R は受付窓口より受入担当責任者に戻され、受入担当者からユーザに郵送される。

6. NSIII で提供するサービス

図 5 に NSIII で提供するサービスと、ファイヤーウォールなどで通過を許可するプロトコルについて示す。

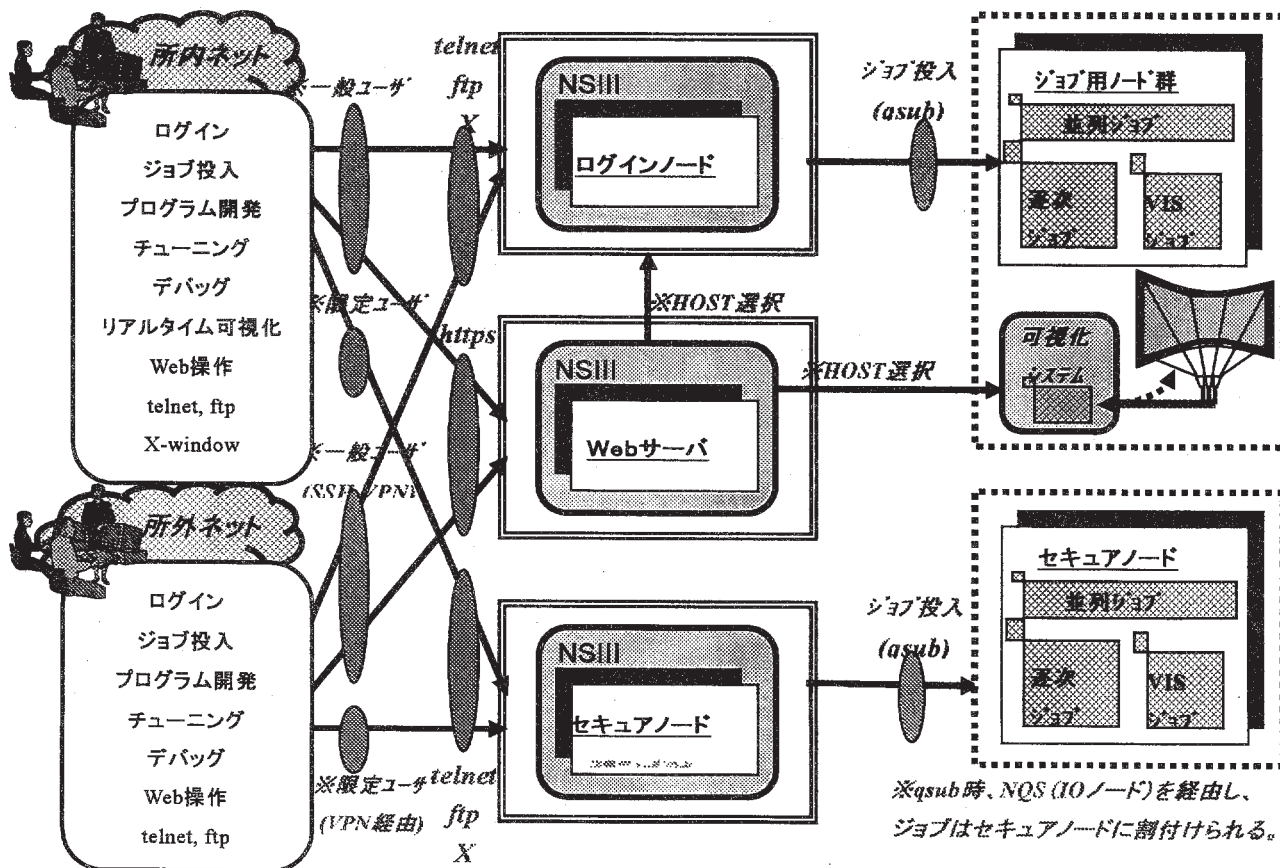


図 5 NSIII運用サービス

LAN および、外部から接続する一般ユーザは、ログインノードにログインして、NSIII へのジョブ投入、プログラム開発、デバッグなどの操作を行うことができる。

よりセキュアな利用を目的とする限定ユーザは、一般ユーザの利用するノードとは切り離された、セキュアノードに対して、直接ログインする。

プロトコルとしては、telnet, ftp (scp), http, https, X-protocol を想定している。LAN からの NSIII への接続に関しては、特に制限を加えていないが、NSIII 側からの LAN への通信に関しては、原則禁止する。また、NSIII と外部ユーザ間の通信については、原則として暗号路が確立された特定の通信についてのみ許可する。

7. おわりに

本稿では NSIII で実装するネットワーク構成を示した。また、NSIII のサービスについて、ネットワーク利用の観点から述べた。

今後は、実装したネットワーク構成における実性能値の測定と、NSIII のサービス提供の質を低下させない確実なネットワーク運用が必要となる。