

3.1.2 ED-Net 対応システムの整備

目 次

(1) 業務の内容

- (a) 業務題目
- (b) 担当者
 - (c) 業務の目的
 - (d) 5 ヶ年の年次実施計画
 - (e) 平成 15 年度業務目的

(2) 平成 15 年度の成果

- (a) 業務の要約
- (b) 業務の実施方法
 - 1) 既存システムを用いた実大実験データ配信・公開システムの整備
 - 2) 実験データ配信・公開システムの性能検証
- (c) 業務の成果
 - 1) 既存システムを用いた実大実験データ配信・公開システムの整備
 - 2) 実験データ配信・公開システムの性能検証
- (d) 結論ならびに今後の課題
- (e) 引用文献
- (f) 成果の論文発表・口頭発表等
- (g) 特許出願，ソフトウェア開発，仕様・標準等の策定

(3) 平成 16 年度業務計画案

(1) 業務の内容

(a) 業務題目 ED-Net 対応システムの整備

(b) 担当者

所 属	役 職	氏 名	メールアドレス
独立行政法人 防災科学技術研究所 防災基盤科学技術研究部門 特定プロジェクトセンター 実大三次元震動破壊実験施設 利用プロジェクト	統括主任研究員	佐藤 一雄	ksato@bosai.go.jp
	プロジェクトディレクター	佐藤 正義	m.sato@bosai.go.jp
	プロジェクトサブディレクター	井上 貴仁	dinoue@bosai.go.jp

(c) 業務の目的

国内外の研究機関を高速ネットワークで結び、E - ディフェンスの実験研究成果を世界に発信し、実験データ、研究成果の幅広い利用促進を図るとともに、相互のデータ流通や共同実験の効率的実施等を目標とした「ED-Net (E-Defense Network)対応システム」を構築する。

(d) 5 ヶ年の年次実施計画

1) 平成14年度：

システムの全体構想を取り纏める。

実大実験実験データ配信・公開システムの基本設計およびソフトウェアの整備を行う。

2) 平成15年度：

基本設計に基づく実大実験データ送受信システムを、既存施設(つくば)を利用して構築する。

構築したシステムを用いてシステム試験を行い、システムの高速度・安全性を検証する。

3) 平成16年度：

つくばで確認したシステムを基に三木におけるシステムを具体化する。

実験と解析を遠隔地同士でリアルタイムに協調実施するための仕組みを構築する。

4) 平成17年度：

具体化したシステムで、つくばと三木間の試験運用を行い、システムの問題点の抽出と改良および通信速度の高速化を図る。

遠隔協調実験システムにより実験・解析を実施する。

5) 平成18年度：

システムの本格運用を開始し、共同研究者との送受信および実験情報のHP提供を行う。また、実際の運用での問題点抽出と改良を行う。

(e) 平成15年度業務目的

- 1) つくば Wide Area Network (以降つくば WAN と称す)¹⁾を用い、防災科学技術研究所(以降、防災科研と称す)と E-ディフェンスを仮想した機関(農林水産省農林水産技術会議事務局筑波事務所、以降農水研と称す)の機器・装置を整備し、実大実験データ配信・公開システムを構築する。
- 2) 構築したシステムを用いて実大実験データ配信・公開システムに必要な広域データ共有及び広域アプリケーション共有の性能検証を行う。

(2) 平成15年度の成果

(a) 業務の要約

平成15年度は次の各項目を実施した。

- 1) つくば WAN を用いた実大実験データ配信・公開システムを構築するため、防災科研と農水研のファイバ・チャンネル網の装置・機器を整備した。
- 2) 構築したシステムを用いて実大実験データ配信・公開システムに必要な広域データ共有及び広域アプリケーション共有の性能検証を行うため、試験計画を立案・実施した。

(b) 業務の実施方法

1) 既存システムを用いた実大実験データ配信・公開システムの整備

つくば地区の高速ネットワーク網であるつくば WAN を用い、防災科研と E-ディフェンスを仮想した機関(農水研)との間に、実大実験データ配信・公開システムを構築するため、ファイバ・チャンネル網の機器・装置を整備する。

2) 実験データ配信・公開システムの性能検証

既存システムを用いて整備する実大実験データ配信・公開システムの性能を検証するため、広域データ共有及び広域アプリケーション共有試験を実施する。

(c) 業務の成果

1) 既存システムを用いた実大実験データ配信・公開システムの整備

前年度、国内外の共同研究者との迅速な実験データの交換を行うための広域データ共有システムとして、SAN (Storage Area Network) /CXFS(Clustered eXtended File System) ファイルシステム技術が有効であることを検証した。しかし、この検証試験は防災科研内で行われており、高速ネットワークに対応する ED-Net 環境における広域データ共有の性能評価を行うと共に、実運用への適用性を検討する必要がある。

そのため、本年度は高速なネットワークであるつくば WAN に実験用として既に設定されているファイバ・チャンネル網を利用し、防災科研と検証に必要な距離の離れた拠点間(仮想三木(E-ディフェンス))において、SAN/CXFS ファイルシステム技術の広域データ共有への適用性を検証する。

想定するもう一方の拠点としては、防災科研に設置されている SGI 社製 Origin3800 と

同様の装置を保持し、防災科研からの距離が検証に必要な距離以上離れていること、の条件を満たすことが必要であり、今回はつくば WAN に参画している機関の中でその条件を満たす農水研を検証試験の相手とした。

防災科研に設置されている SGI 社製 Origin3800 において、SAN/CXFS ファイルシステム技術を用いた広域データ共有を可能にするため、既設 2 ファイバ単方向切替リング対応光分岐挿入装置（日本電信電話株式会社製）に必要なインタフェース等を設置した。すなわち、既設の 2 ファイバ単方向切替リング対応光分岐挿入装置のトランスポンダ機能部にファイバチャネル(2.12GBd)パッケージを 1 回線挿入した。

また、防災科研と農水研に、ネットワークを整備するため、既存のつくば WAN 上に設定されているファイバ・チャンネル網をこの 2 拠点間に設定し直し、必要な装置・機器を移転・設置した。

図 1 に構築された防災科研・仮想 E - ディフェンス間の実大実験データ配信・公開システムを示す。

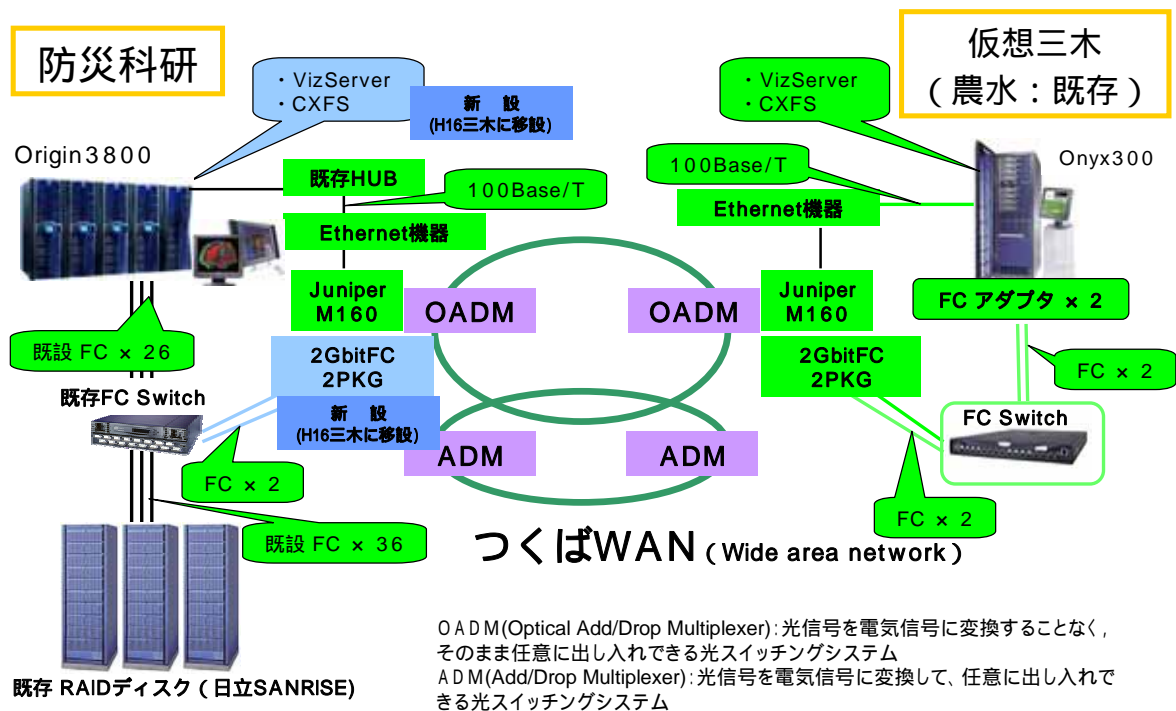


図 1 既存システムを用いた実大実験データ配信・公開システム

2) 実験データ配信・公開システムの性能検証

既存システムを用いて整備した実大実験データ配信・公開システムの性能を検証するため、広域データ共有及び広域アプリケーション共有試験を実施する。

a) 広域データ共有試験

）試験構成

使用する機器として、防災科学技術研究所側は CXFS サーバとして現有の Origin3800 128CPU (ホスト名: o3k)、CXFS クライアントとして現有の Origin3800 256CPU (ホスト名: o3kb)、磁気ディスク装置は SANRISE1100 を用いる。つくば WAN 越しの農水研の CXFS クライアントとして Onyx300 (ホスト名: ayame) を用いる。

今回用いる Origin3800 及び Onyx300 については、現有機器をできるだけ使い、最小限の構成変更と機器追加で対応できるように設計している。また、ファイバ・チャネル (以下、FC と記述) については、現有の機器を用いるため、Origin3800 側は全て 1Gbit 仕様の機器となる。Onyx300 側は、将来の拡張性などを勘案して、2Gbit 対応の FC で構成されている。従って、防災科研側ではつくば WAN に対応するため既設の 2 ファイバ単方向切替リング対応光分岐挿入装置のトランスポンダ機能部にファイバチャネル(2.12GBd)パッケージを 1 回線挿入したが、Origin3800 との接続では機器が保有する 1Gbit の速度が上限となる。

本構成が有する冗長性と性能は以下の通りである。

- ・ 磁気ディスク装置(既存)には RAID5 を使い、また、キャッシュのバックアップなどの機能があり、冗長性を有する。
- ・ 1つの RAID 構成(4+1)に対して、1つのコントローラを使い、高いスループットを得ている。
- ・ Origin3800 側のスイッチは 4 台 (既存) を使い、いずれのスイッチが障害となってもデータアクセスを継続できる冗長性を有する。

防災科研及び農水研の現行構成を変更し、つくば WAN への接続したときの構成を図 2 に示す。現行の磁気ディスク装置の構成変更をできるだけ少なくつくば WAN へ接続を行うため、図中の点線で示す「Silkworm2800 - o3k」間の FC を矢印で示すように「Silkworm2800 - Juniper M160」間に移設・接続した。また、Juniper M160 側は 2G-FC PKG に FC を直接接続した。これによりつくば WAN 越しにある農水研の Onyx300 より防災科研の SANRISE1100 のラック 1 及びラック 2 の磁気ディスクがアクセス可能となる。なお、現行の o3k については、ラック 1 及びラック 2 の各磁気ディスクへのバンド幅が若干小さくなるが、フロントエンドシステムであることから運用への影響は少ない。

）試験方法

試験では、防災科研内で CXFS サーバ、CXFS クライアントを FC ケーブルで直結した構成でディスク I/O 性能測定を行い基礎データ (Local と表示) とした。一方、OADM 装置を接続し、つくば WAN を経由した場合の CXFS クライアント (農水研) におけるディスク I/O 性能 (Remote と表示) を測定し、基礎データと比較することにより、OADM 機器を使用したことによる性能劣化の検証を実施した。

性能測定ソフトウェアとしては、UNIX にて使用できるディスク I/O 測定ツールの lmdo および IRIX にて提供されているディスク I/O 測定ツールの diskperf を使用した。

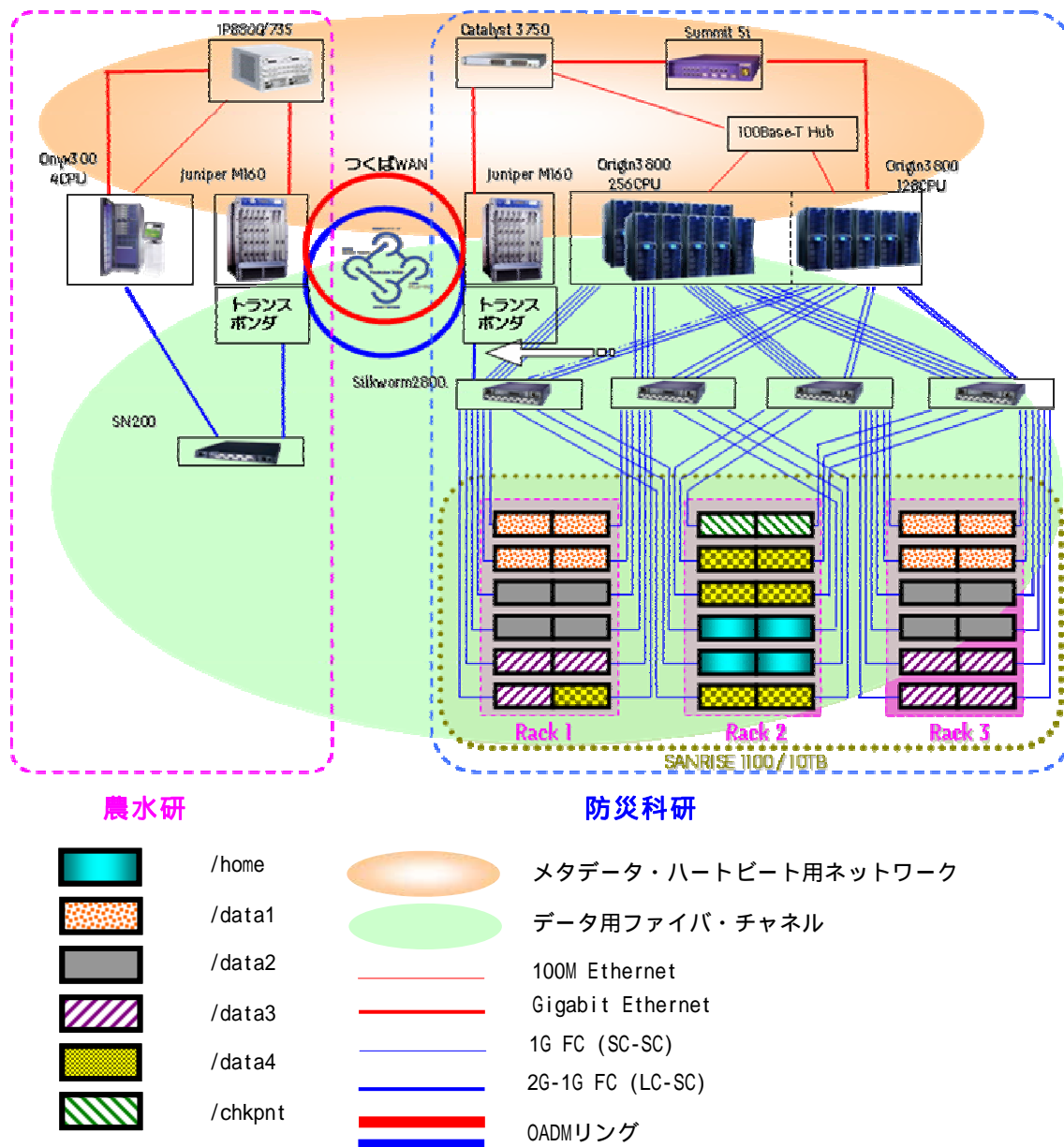


図2 機器構成図

lmddd コマンド

lmddd コマンドは dd コマンドの改良版で、データのコピーに要する時間を測定し、性能データを出力する。任意のブロックサイズやダイレクト I/O、コピーするファイルサイズなどを指定することができ、今回使用したパラメータは下記の通りである。

ブロックサイズ 128kB, 256kB, 512kB, 1024kB, 4096kB

移動ファイルサイズ 1.0, 2.0, 3.0 GB

ダイレクト I/O Yes

diskperf コマンド

diskperf コマンドは IRIX のディスクのパフォーマンスを測定するツールで、ディスクアクセスについて次の 3 パターン実施した。

fwd_wt フォワード書き込み

fwd_rd フォワード読み込み

bwd_wt バックワード書き込み

bwd_rd バックワード読み込み

rnd_wt ランダム書き込み

rnd_rd ランダム読み込み

また、使用したコマンドパラメータは、# diskperf -D -W -c1g /disk2/1gdata である。

）試験結果

Imdd による Direct I/O の性能評価の試験結果及び diskperf による I/O 性能評価試験の一例としてランダム読み書き性能の試験結果を図 4 に示す。つくば WAN 環境下でのディスク I/O 性能は、ローカル環境での性能より低下しているが、ディスク共有では読み書き性能は 30MB/sec (0.24Gbps) の性能がみられ、昨年度想定したデータの転送量 (表 1 参照) と同等の帯域が確保できることが判った。なお、ブロックサイズやリクエストサイズが大きくなるとつくば WAN 環境下でのディスク I/O 性能は頭打ち状態となっている。この原因となっているつくば WAN 側の制限を解消することにより、ディスク I/O 性能は向上すると考えられる。

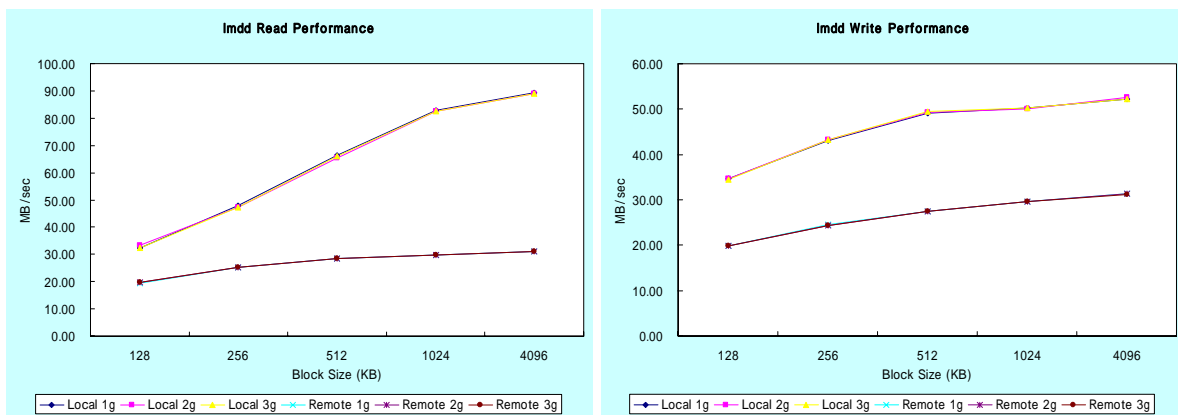


図 3 Imdd による Direct I/O の性能評価

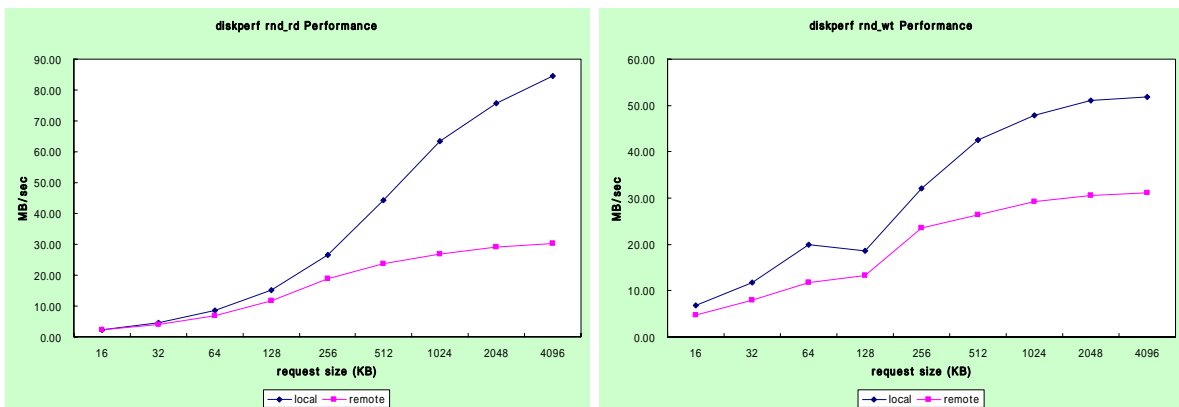


図 4 diskperf による I/O 性能評価 (ランダム読み書き性能)

表 1 ネットワークに必要なとされる帯域幅

作業内容	転送に要する時間	転送データ量	全データを転送するための容量	必要な帯域幅
実験中の共同研究者 (国内外) へのデータ転送 (主要な実験及び画像データ)	至急 (5分)	実験データ : 500Mbit 画像データ : 75Gbit	0.26Gbps	1Gbps
ホームページからの実験状況の画像配信	リアルタイム	180Mbit/秒 (30万画素、25フレーム/秒)	0.18Gbps	1Gbps

帯域決定には通信負荷を考慮し余裕を見込んでいる

b) 広域アプリケーション共有試験

）試験構成

防災科研側 Origin3800 及び農水研側 Onyx300 を下記のような構成で接続した。IP はつくば WAN グリッド実験網を用い、Origin3800 の映像を農水側の Onyx300 にて投影する形で検証した。アントとして Onyx300 を想定する。(図 5 参照)

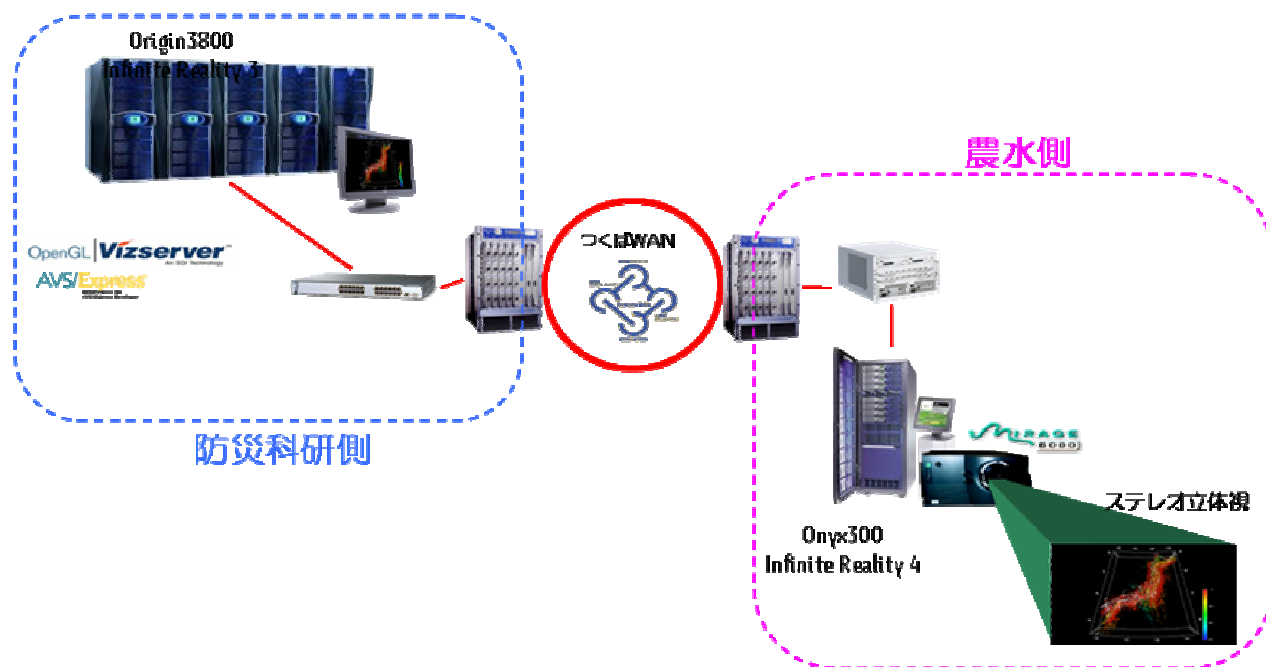


図 5 試験構成

）試験方法

ア) TTCP による帯域性能測定

防災科研 Origin3800 - 農水研 Onyx300 間における帯域測定を TTCP(Test-TCP)で実施した。TTCP は、2 台のコンピュータ上で TCP もしくは UDP プロトコルを使って通信路を開き、一方的にパケットを送信して、そのパフォーマンス(単位時間あたりに送信/受信できた bytes 数)を表示するコマンドである。

イ) VizServer の機能検証

防災科研側 Origin3800 にインストールされている VizServer をサーバとして使用し、OpenGL GUI アプリケーションである AVS Express 6.2 を使用した場合に、つくば WAN を経由して農水研側 Onyx300 において VizServer の下記に挙げる機能の検証を実施した。

シングルユーザ機能、立体視機能、コラボレーション機能、

コラボレーション + 立体視機能

ウ) VizServer の機能測定

防災科研側 Origin3800 の VizServer をサーバとして使用し、防災科研側の OpenGL GUI アプリケーションである AVS Express Developer 6.2 の描画測定を実施した。

）試験結果

ア) TTCP による帯域性能測定

防災科研 (NIED) Origin3800 - 農水研(MAFF) Onyx300 間における帯域測定結果を表 2 に示す。表から分かるとおり、防災科研より農水研への帯域は、47.8MB/sec、農水研から防災科研への帯域は平均 55.1MB/sec あることが分かった。

表 2 TTCP による帯域性能測定

1回目

送信側	受信側	NIED	MAFF
NIED	MAFF	47.8	47.6
MAFF	NIED	56.7	56.7

2回目

送信側	受信側	NIED	MAFF
NIED	MAFF	47.9	47.8
MAFF	NIED	53.4	53.5

(単位は MB/sec)

イ) VizServer の機能検証

農水研側 Onyx300 において VizServer のシングルユーザ機能、立体視機能、コラボレーション機能及びコラボレーション + 立体視機能を全て検証できた。

ウ) VizServer の機能測定

コラボレーション機能を使用しない場合と使用した場合で、描画画像の圧縮度をパラメータとして測定したフレームレート及び使用帯域を表 3 に示す。表 2 に示すつくば WAN グリッド実験網の帯域が 47.8MB/sec と十分なため、描画性能は圧縮度にまったく依存しないことも分かった。この結果は、つくば WAN グリッド実験網の帯域において、VizServer によるアプリケーション共有機能は非圧縮でリアルタイムにコラボレーション環境を実現すること示している。

表 3 VizServer の機能測定結果

圧縮度	single		collaboration	
	フレームレート [fps]	帯域 [MB/s]	フレームレート [fps]	帯域 [MB/s]
非圧縮	44.5	48	35.5	47
1:1	48.4	25	58.9	17
4:1	49.6	4.8	59	3
8:1	50.1	2.6	59.2	1.6
16:1	50	1.4	59.2	0.9
32:1	50.3	0.74	59.4	0.45

以上の結果、SAN CXFS 機能及び VizServer 機能を組み合わせることにより、リアルタイムに視覚化された情報を遠隔地間で確認しながら研究を遂行することができ、かつ必要に応じて可視化結果を保存できるような環境が実現できると考えられる。

(d) 結論ならびに今後の課題

- 1) 国内外の共同研究者との迅速な実験データの交換ができる実大実験データ配信・公開システムの試験システムを、既存の高速ネットワークであるつくば WAN を用い、防災科研と農水研との間で整備・構築した。
- 2) 構築・整備したシステムを用い、広域データ共有及び広域アプリケーション共有の性能を検証した結果、運用に耐える構成であることを確認した。
- 3) つくばと三木の間で広域ディスク及びアプリケーション共有を行う場合、その距離を考えると Gigabit ネットワークのインフラが整備されていることが望ましい。

(e) 引用文献

- 1) つくば WAN, <http://tsukuba-wan.mexttci.go.jp/>

(f) 成果の論文発表・口頭発表等

著者	題名	発表先	発表年月日
なし			

(g) 特許出願，ソフトウェア開発，仕様・標準等の策定

1) 特許出願

なし

2) ソフトウェア開発

なし

3) 仕様・標準等の策定

なし

(3) 平成 16 年度業務計画案

(a) つくばの既存システムを用い性能を確認したシステムを基に、E-ディフェンスにおけるシステムを具体化する。

(b) 大型震動台実験のような多数の研究者の協力を必要とする実験において、実験場所にいない研究者でもネットワークを通じて容易に実験に参画することを可能とすることを目的として、実験と解析を遠隔地同士でリアルタイムに協調実施するための仕組みを構築する。