

災害からの早期復旧を目指した レジリエント情報通信ネットワーク

2013年12月17日

国立情報学研究所(NII)

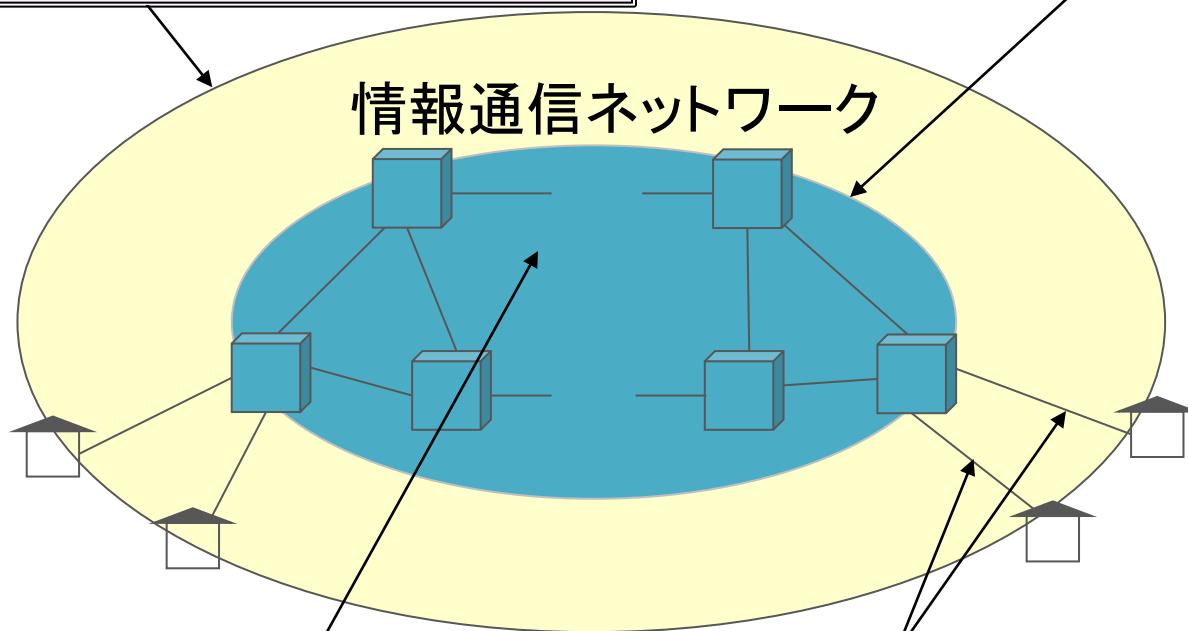
山田茂樹

(shigeki@nii.ac.jp)

講演内容

1. バックボーンネットワーク
とアクセスネットワーク

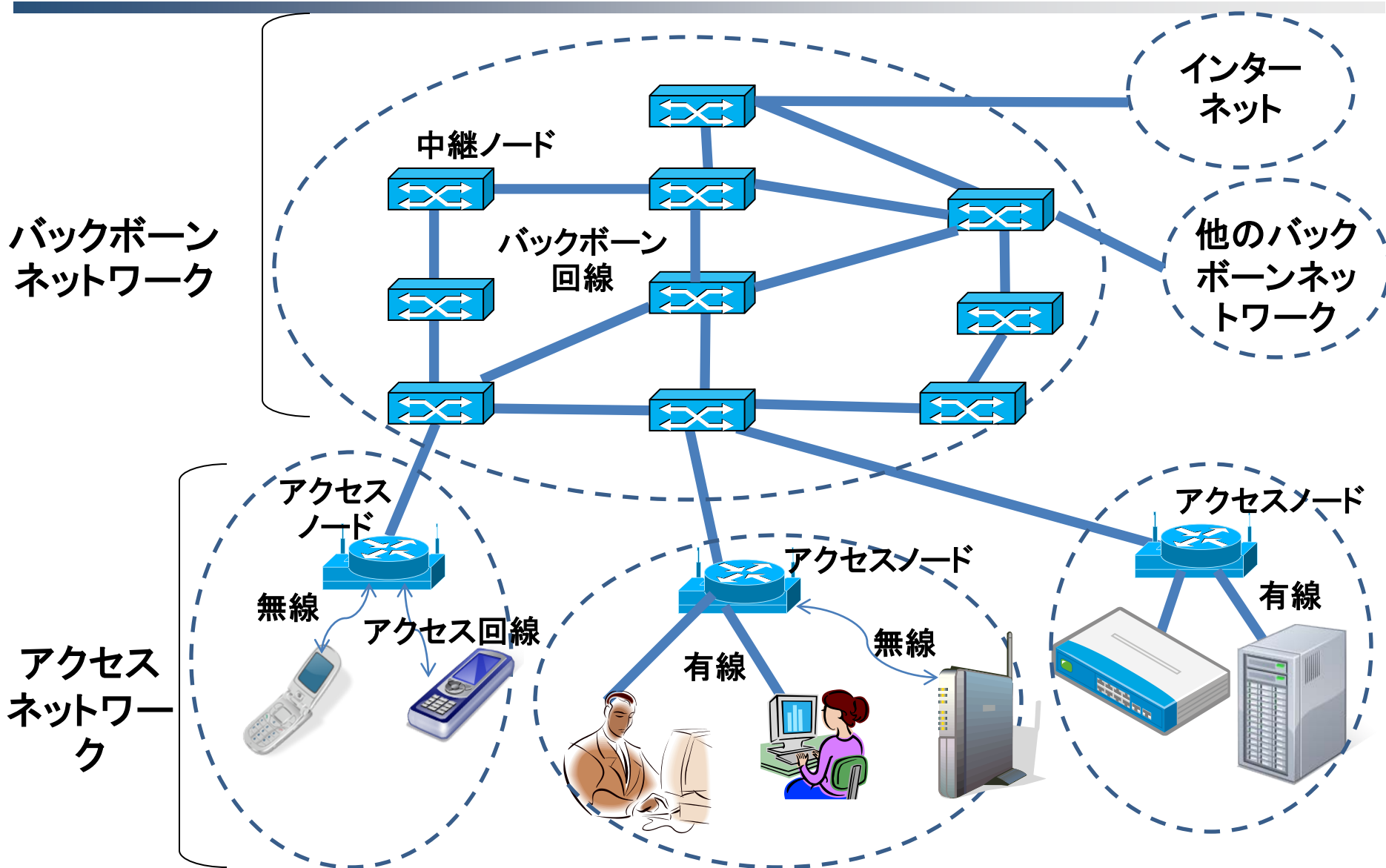
2. 学術情報ネットワーク
SINETの役割



3. レジリエント・バック
ボーンネットワーク

4. レジリエント・アクセス
ネットワーク

バックボーンネットワークとアクセスネットワーク(1)



バックボーンネットワークとアクセスネットワーク(2)

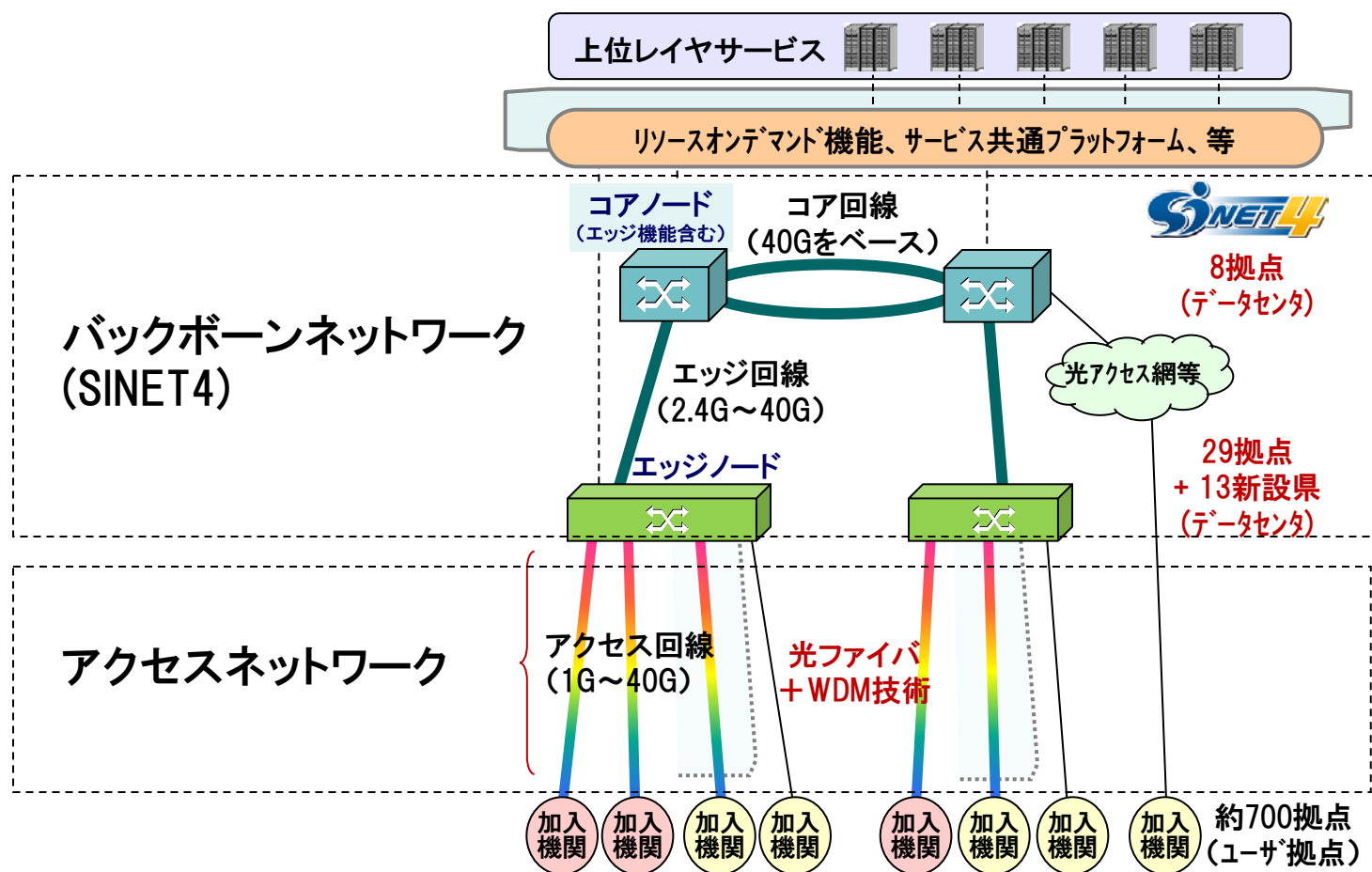
• バックボーンネットワーク

- 大規模な通信ネットワークにおいて、集線装置間や拠点間、あるいは事業者間、国家間などを結ぶ大容量の通信回線網
- 光ファイバによる有線回線が採用されている場合が多い。
- 通常、各ISP (Internet Service Provider)等が構築。
- ISP間の相互接続ポイント、インターネットへの接続ポイントを有する。
- 広域エリアをカバーし、予備回線や迂回ルート等、高信頼化が図られている。
- その結果、大規模災害でもバックボーンネットワークの一部は生き残ることができる。
- NIIが構築運用する学術情報ネットワークSINETも基本的にはバックボーンネットワークの一つ。

• アクセスネットワーク

- バックボーンから個々の機器・端末あるいは利用者・加入者などを結ぶネットワーク
- 銅線(メタル・ケーブル)や光ファイバの有線系ネットワークと、無線LANやFWA(fixed wireless access)など無線系ネットワークがある。
- 小地域をカバーし、代替ルート等を持たない／少ない場合が多い。
- 最も単純な構成は、予備回線を持たない単一アクセス回線

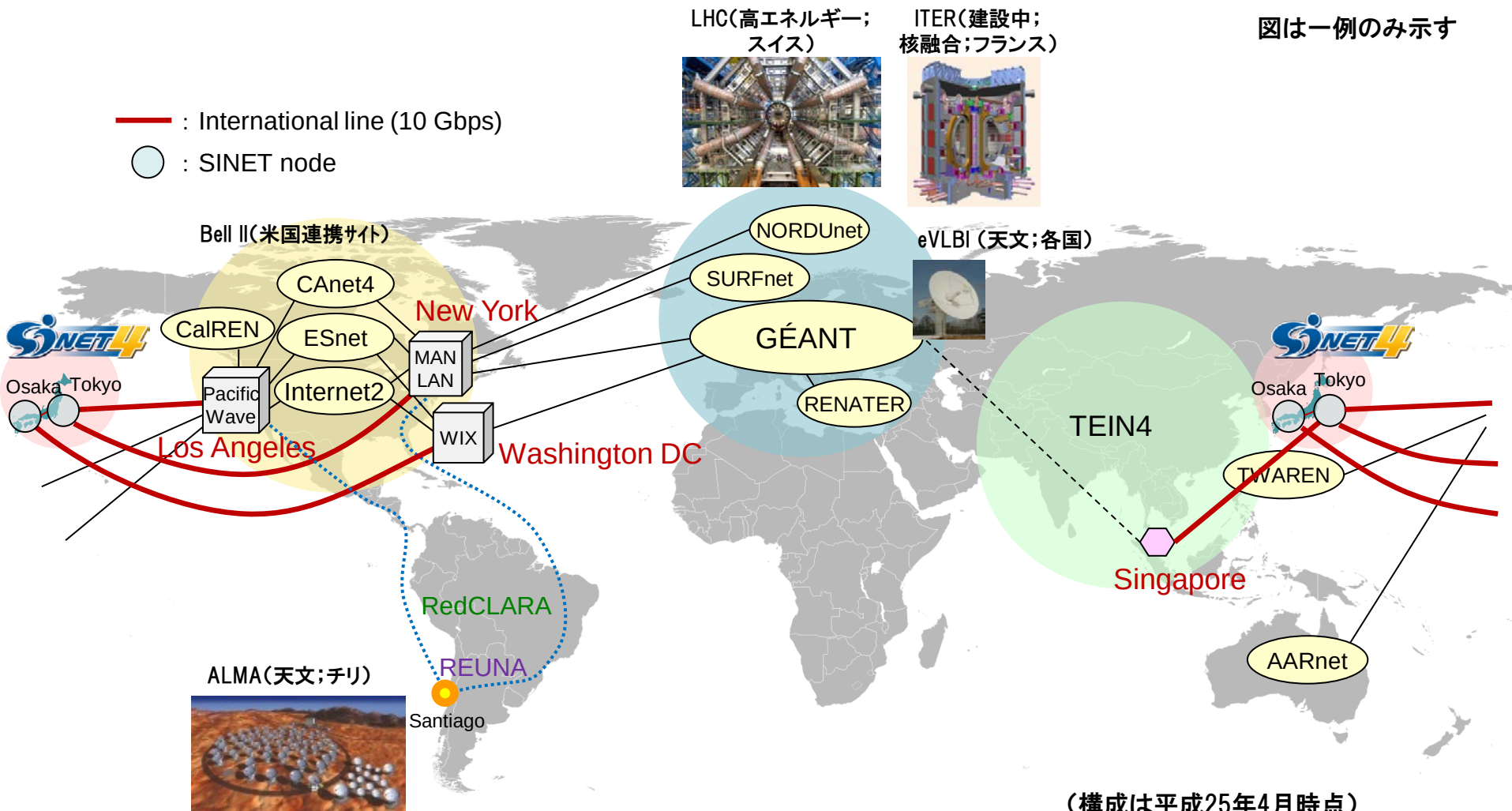
- ◆ NWの高速化: ネットワーク構成の見直しや光ファイバ+WDM技術などにより経済的に高速化
- ◆ エッジ高安定化: エッジノード・コアノードを全てデータセンタへ設置
- ◆ 格差の解消: ノード未整備県の解消、非ノード校のアクセス系の経済的高速化(共同調達)
- ◆ サービスの多様化: SINET3のアーキテクチャを継承し、リソースオンデマンド機能等を強化・拡張
- ◆ 上位レイヤ展開: 上位レイヤサービスを支援するインタフェースやサービス共通プラットフォームを整備



-
- :エッジ・コア回線(40Gbps)
 ■ :エッジ・コア回線(10Gbps)
 ■ :エッジ回線 (2.4Gbps)
 ■ :アクセス回線 (10~40Gbps)
- :コアノード(データセンタ)
 ■ :エッジノード(データセンタ)
 ■ :エッジノード(新設県;データセンタ)

国際連携による共同研究の支援

- ◆ 国際的な先端共同研究プロジェクトを、SINETの日米回線(合計帯域30Gbps)やアジア回線(10Gbps)を用いて、国際パートナー(Internet2, GÉANT, TEIN等)と連携して支援

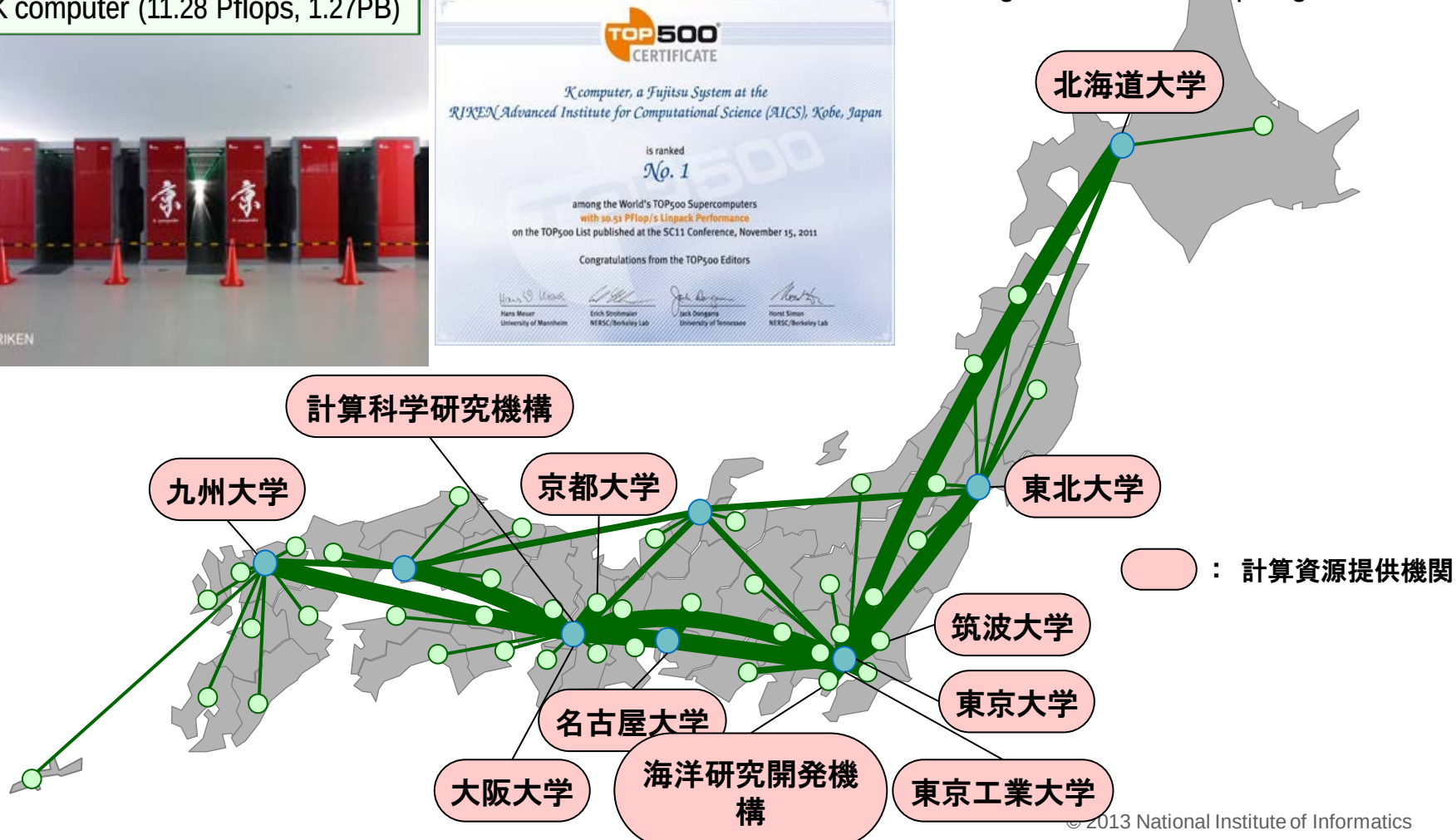


- ◆ 計算科学研究機構や情報基盤センターなどのスパコンやストレージをSINET4を用いて共同利用
- ◆ 認証基盤の整備についてもNIIが中心的な役割を担当
- ◆ 平成24年9月末より本格運用開始

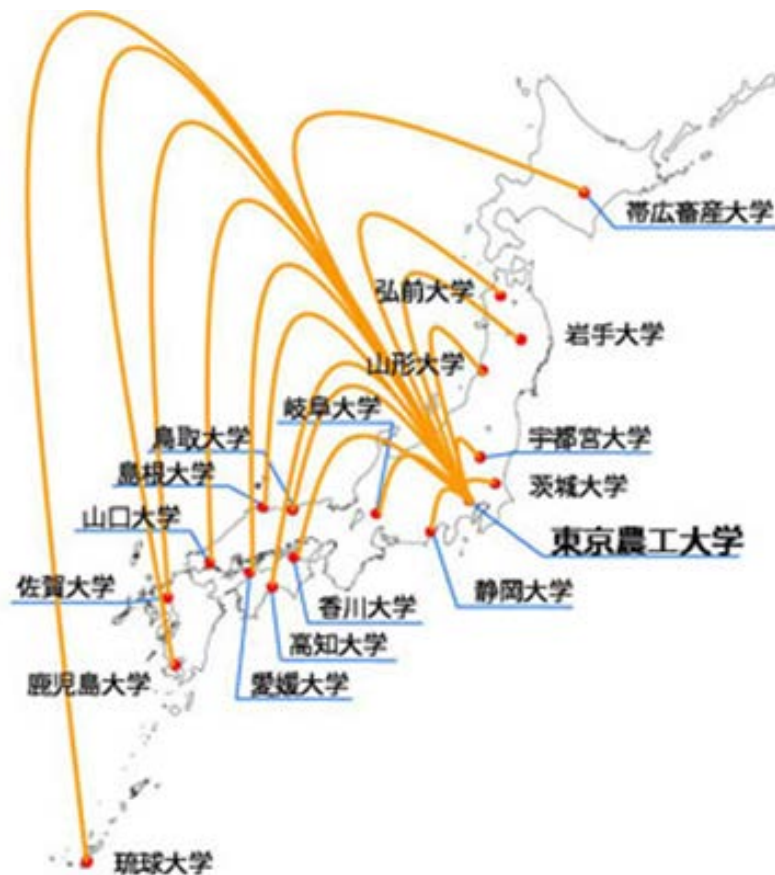
K computer (11.28 Pflops, 1.27PB)



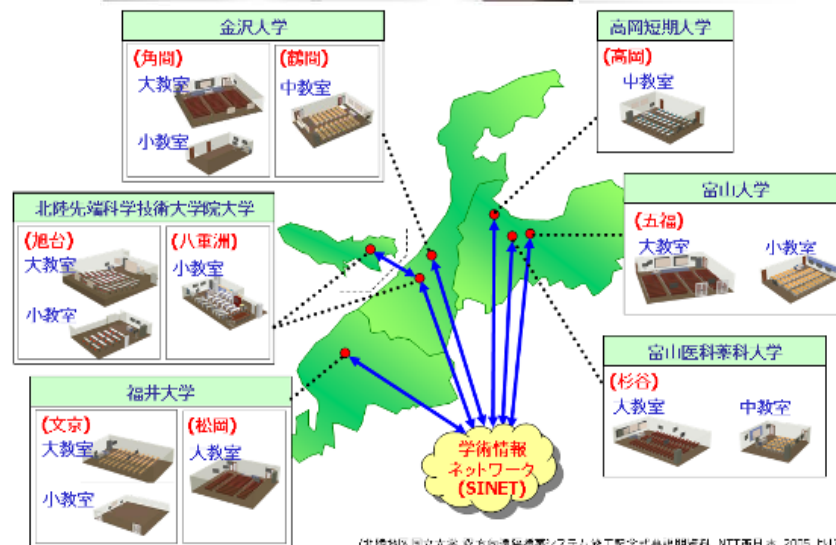
HPCI: High Performance Computing Infrastructure



- ◆ 安定した遠隔授業の実施を実現し、また、単位互換制度の推進の一役も担っている。
- ◆ 例えば、全国18の国立大学にまたがる連合農学研究科を結ぶ遠隔講義、北陸地区の大学間での双方向遠隔授業、琉球大学等での海外大学との遠隔講義、などの実施を支援



東京農工大を基点とした多地点遠隔講義



北陸地区での遠隔授業

SINET利用例 — 高エネルギー研究

- ◆ 小林・益川理論の検証を目的としたBelle実験において、KEKのBelle測定器から出される膨大なデータを連携大学に転送あるいは大学から直接データにアクセスし、並行解析を実施
- ◆ スイスの大型ハドロン衝突型加速器（LHC）における高エネルギー陽子衝突の反応を記録するATLAS測定器から、国際回線を通じて約4Gbpsの超大容量のデータ転送を実施中

提供： 高エネルギー加速器研究機構

『CP対称性の破れ』

膨大な実験データを各大学においてリアルタイムに並行解析

KEK筑波実験室
Belle測定器

400GB~1TB/日
のデータが発生

400GB~1TB/日

400GB~1TB/日

SINET L3VPN

大阪大学
理学部

名古屋大学
理学部

東京大学
理学部

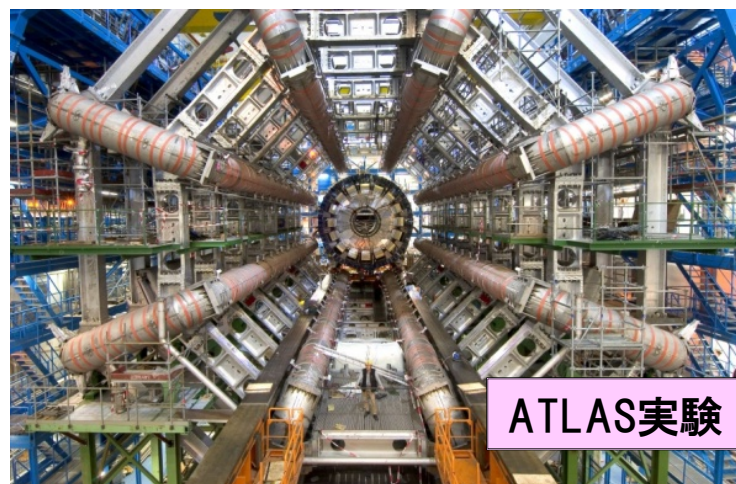
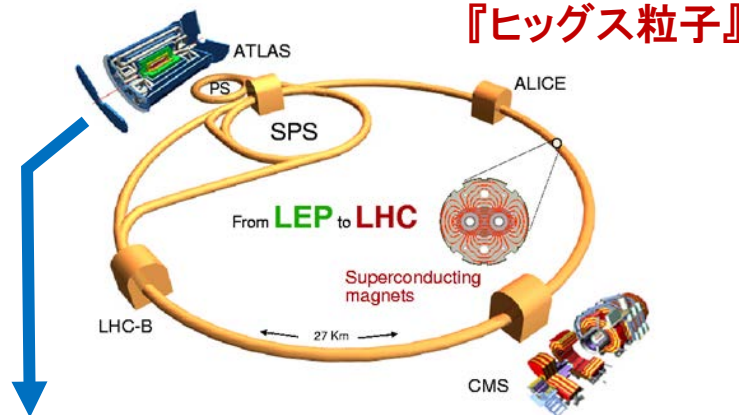
東京工業大学
理学部

東北大学
理学部

Belle実験

Belle IIが2015年
より運用開始

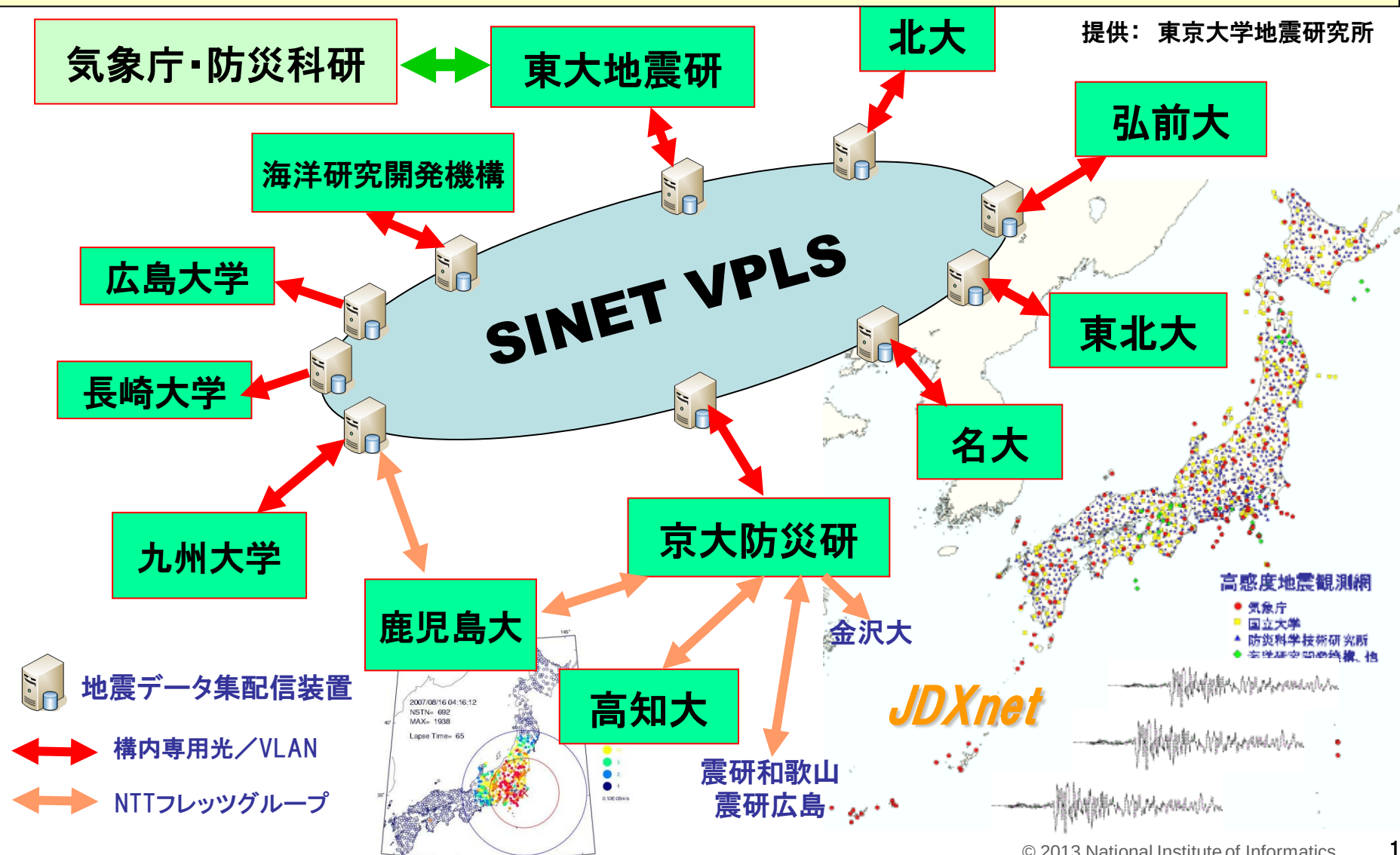
『ヒッグス粒子』



ATLAS実験

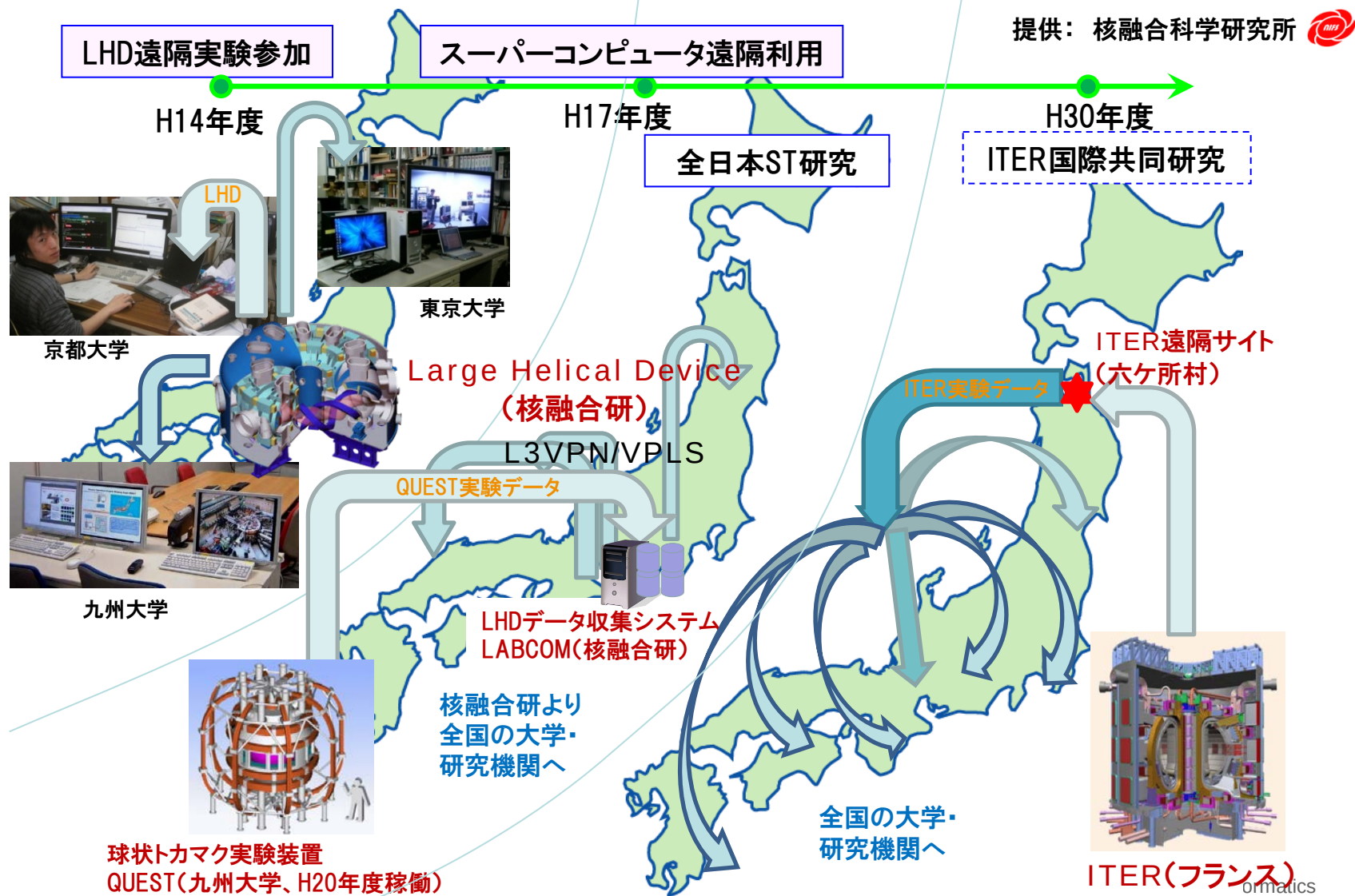
SINET利用例 — 地震研究

- ◆ 全国各地の地震観測データを高優先機能とマルチキャスト機能を用いて各拠点に安定的に配信し、最先端の地震研究を支援（約1300の観測点の地震データが流通）。クラウド機能の活用を検討中。



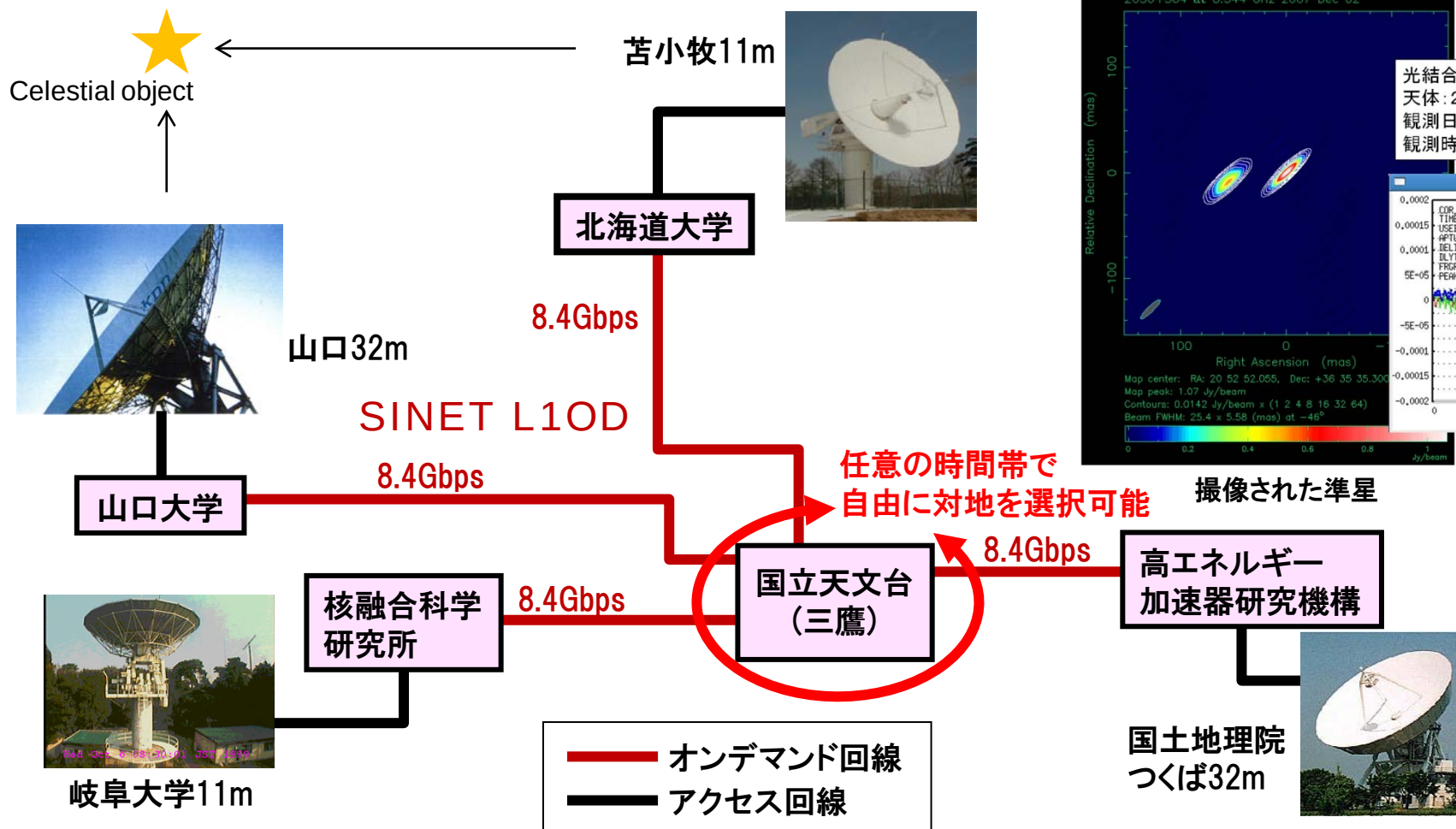
SINET利用例 — 核融合研究

- ◆ 大型ヘリカル実験装置(LHD)や新実験装置(QUEST)からのデータをVPNを用いてセキュアに転送
- ◆ ITER遠隔サイト(六ヶ所村)がオープンし、EUファンドのスパコンが設置済みで利用開始



- ◆ ユーザ要望によりNIIで独自開発したL1オンデマンド機能を用いて、指定した日時だけ、任意の電波望遠鏡と国立天文台間を接続し、大容量の観測データを転送
- ◆ SINET4では、帯域8.4Gbps/電波望遠鏡(最大値)で観測を実施

提供: 国立天文台

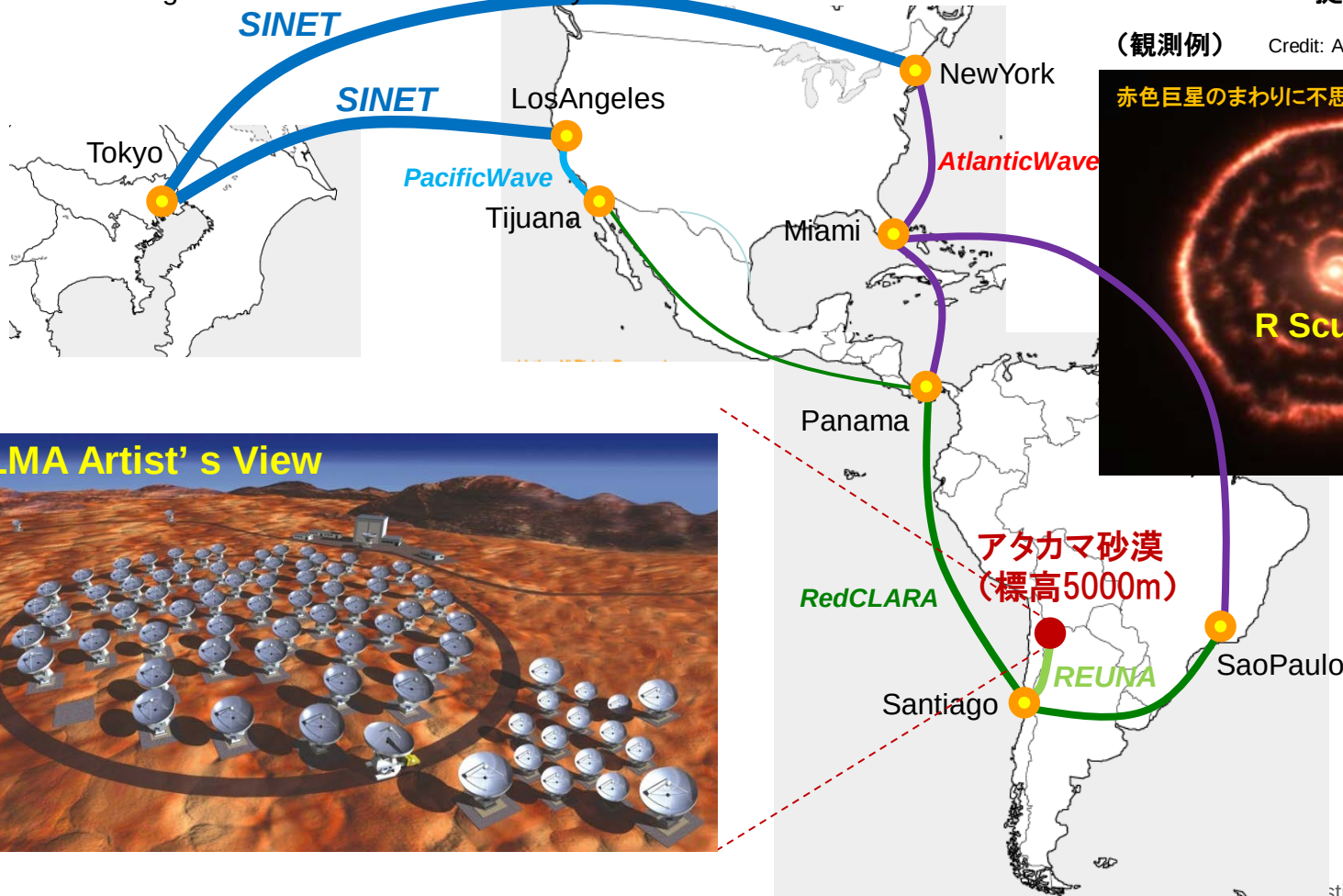


SINET利用例 — ALMA

- ◆ 東アジア(日本が主導)・北米・ヨーロッパ・チリの協力で、チリのアタカマ砂漠に66台の高精度アンテナを設置し一つの超性能な電波望遠鏡を形成。16台による初期運用が2011年9月30日から開始
- ◆ 観測データは、日米欧の3つのアルマ地域センタに転送され、ミラーアーカイブ。東アジア地域センタ(国立天文台)へは、REUNA, RedCLARA, Internet2, SINETを介して観測データを転送。

ALMA: Atacama Large Millimeter/submillimeter Array

提供: 国立天文台



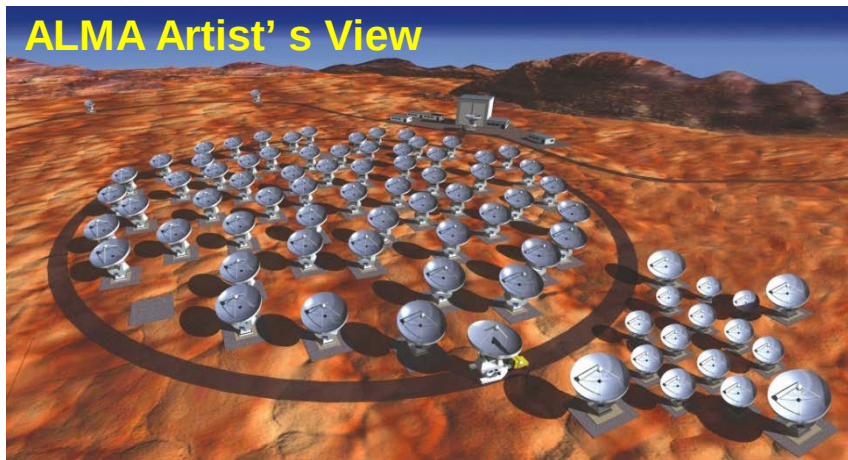
(観測例)

Credit: ALMA (ESO/NAOJ/NRAO)

赤色巨星のまわりに不思議な渦巻き構造を発見

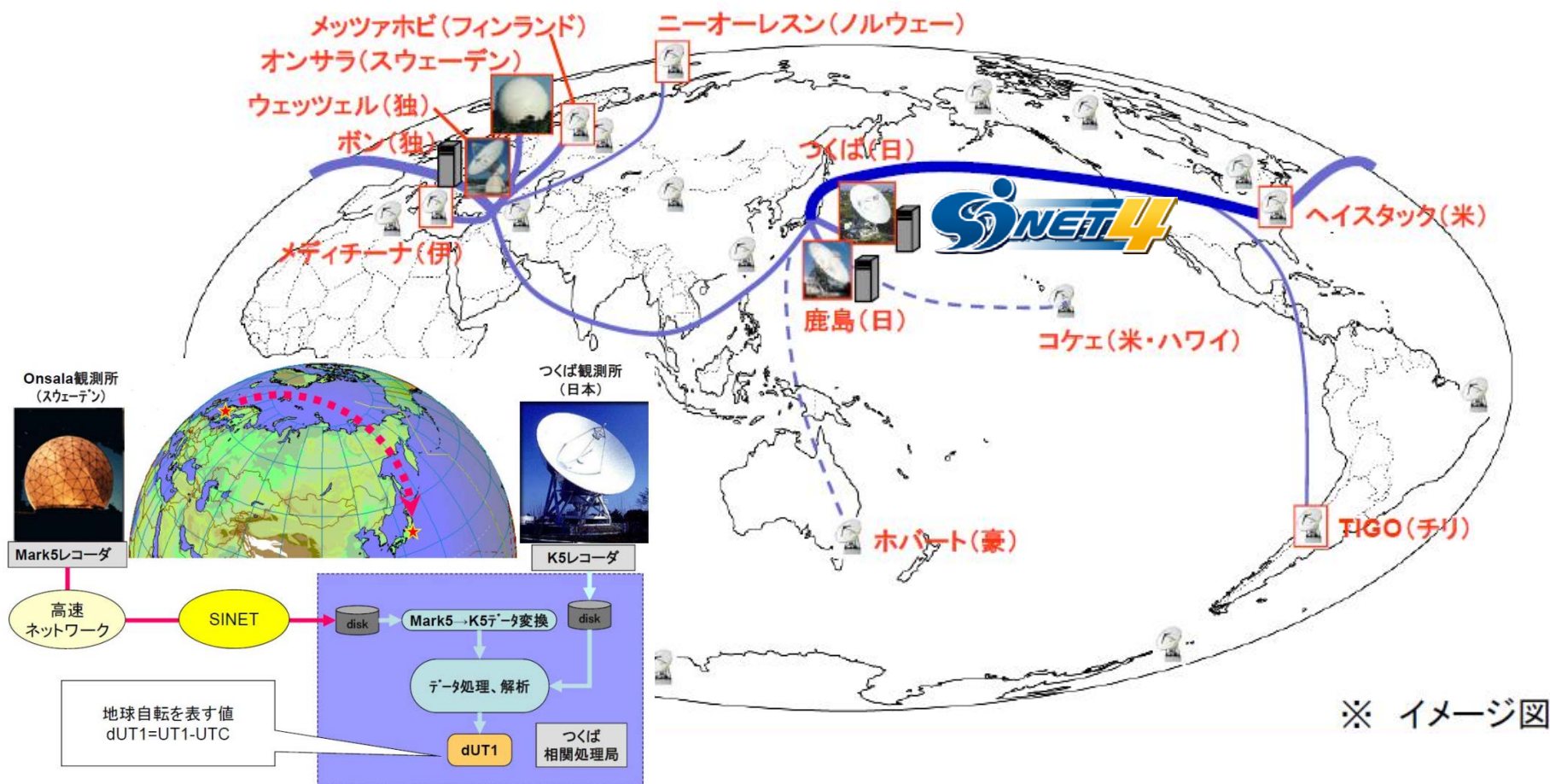


ALMA Artist' s View



- ◆ 国際回線を通じてヨーロッパ各国や米国等との間で電波望遠鏡の観測データを送受信し、地球上の経緯度の基準の決定やプレート運動などの地殻変動検出、地球の自転の振る舞いや天球上での電波星の位置の調査を実施
- ◆ 現在400Mbps程度の帯域で観測を実施

提供： 国土地理院

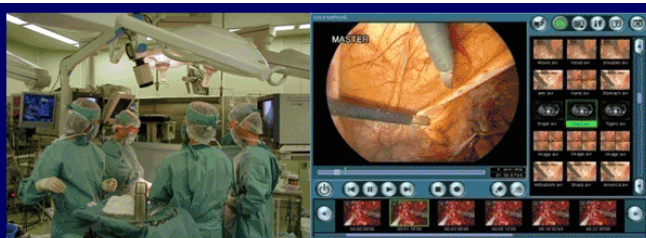


- ◆ ハイビジョン映像(帯域は約30Mbps)を用いて、**内視鏡を用いたライブ手術やカンファレンス**などを、アジア各国(26か国・126機関)と実施
- ◆ 今後は、北米・中南米・ヨーロッパ・アフリカとの連携や遠隔診断への応用を進める予定

提供:九州大学・清水周次先生

手術室

手術画像



- 腹腔鏡下胃切除術
- 暗号ソフト



日本側:九州大学

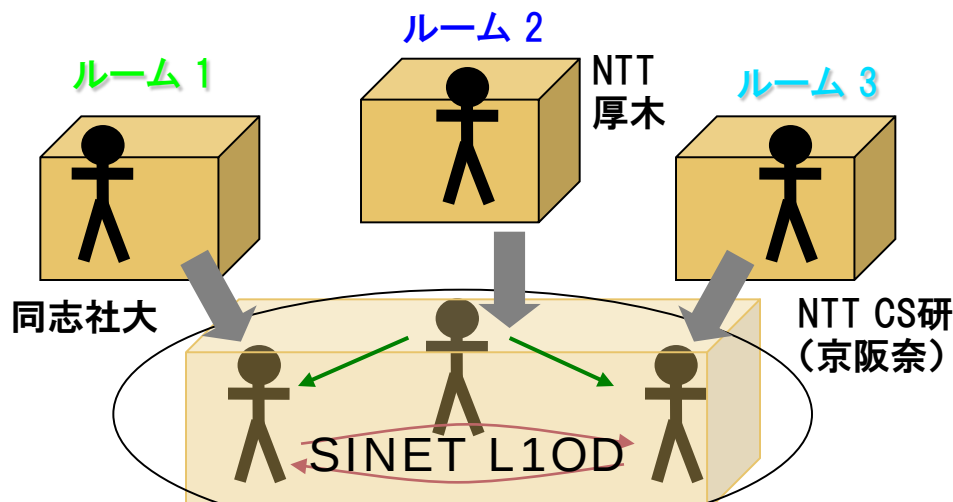


韓国側:国立ガンセンター

◆ 広帯域の通信環境を用いて、高臨場感や超臨場感を与えるメディア装置の開発を実施

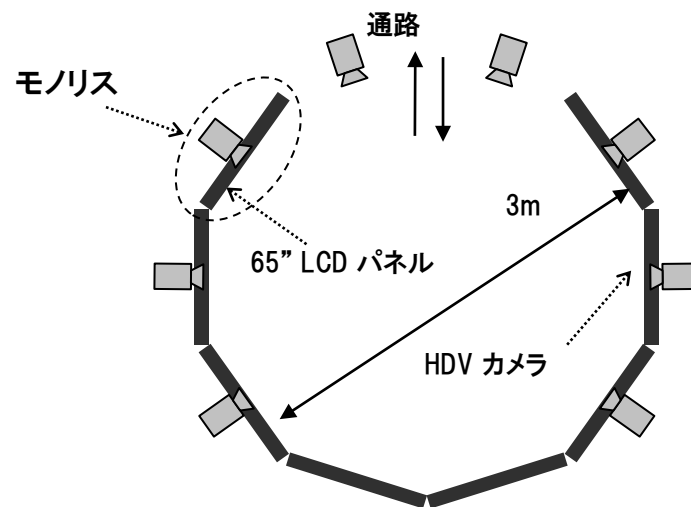
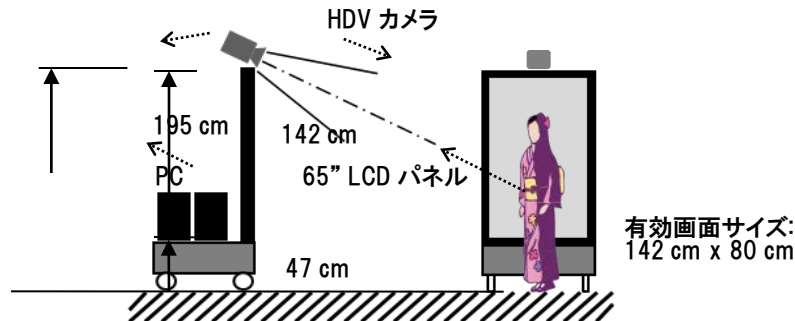
- ・ 例：地理的に離れた人々が同じ部屋にいる感覚（同室感）を共有するビデオシステム(t-Room)の開発
帯域300Mbps(8台分のHD Video、音声、制御信号)で通信。

提供：同志社大学



“モノリス” モジュール:

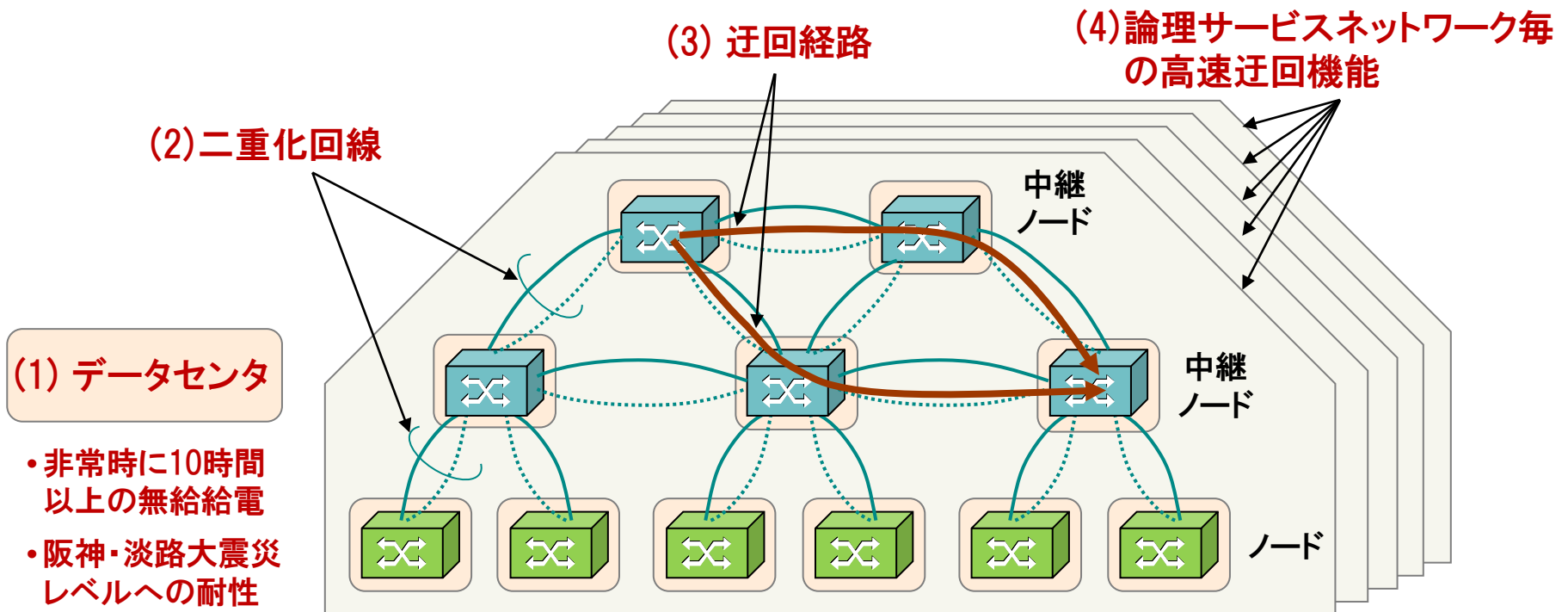
側面図 (左)および 正面図(右)



SINET4の高信頼化設計

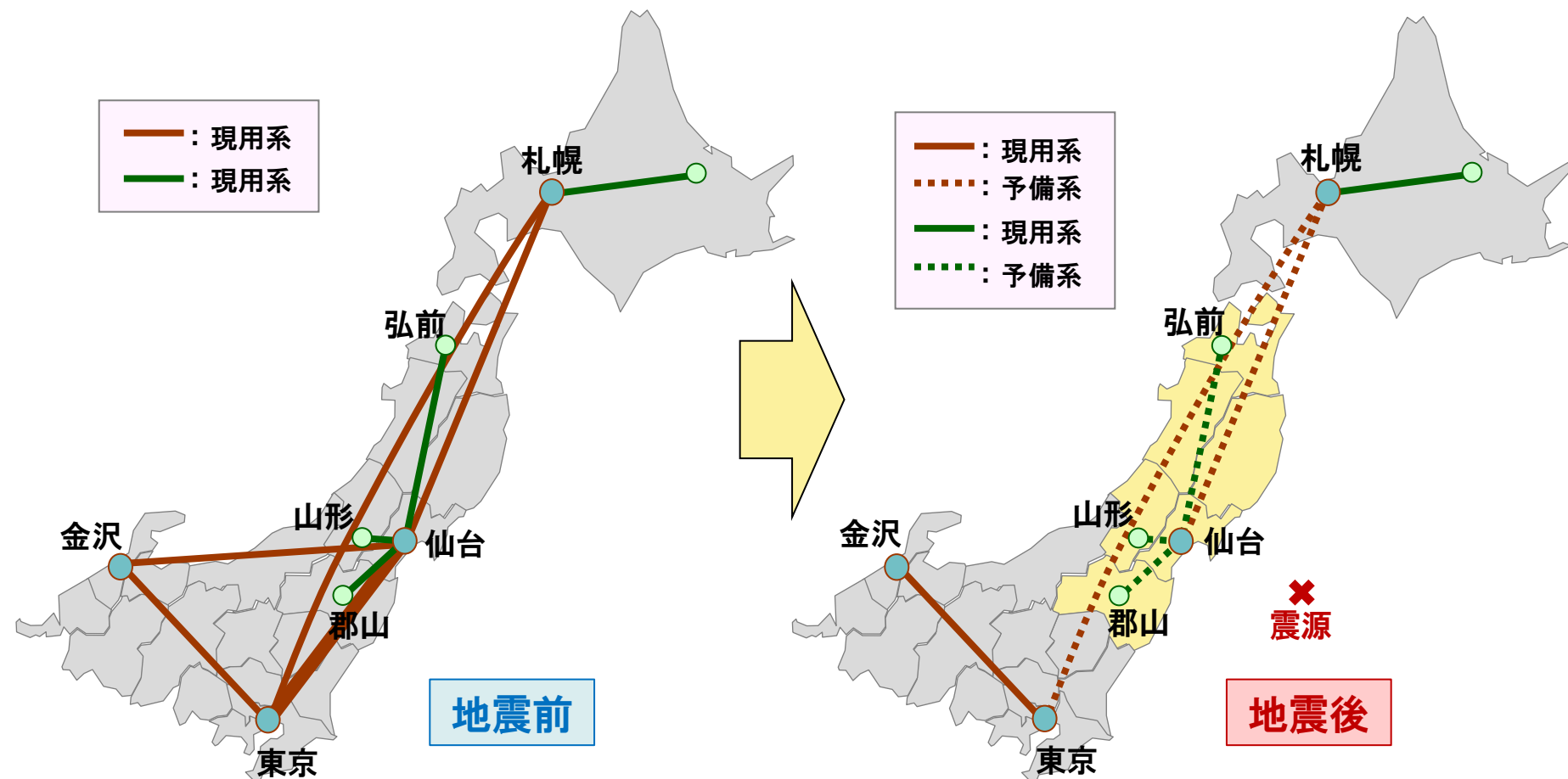
◆ SINET4では、以下の設計・構築ポリシーにより、高信頼化を実現

- (1) 全ノードの**データセンターへの設置**(安定した給電と耐震性を確保)
- (2) 全回線の**異経路二重化**(現用系と予備系)
- (3) 中継ノード間での**迂回経路**確保
- (4) 論理サービスネットワーク(インターネット, L3VPN, L2VPN, L2OD, L10D)毎の**高速迂回機能**実装



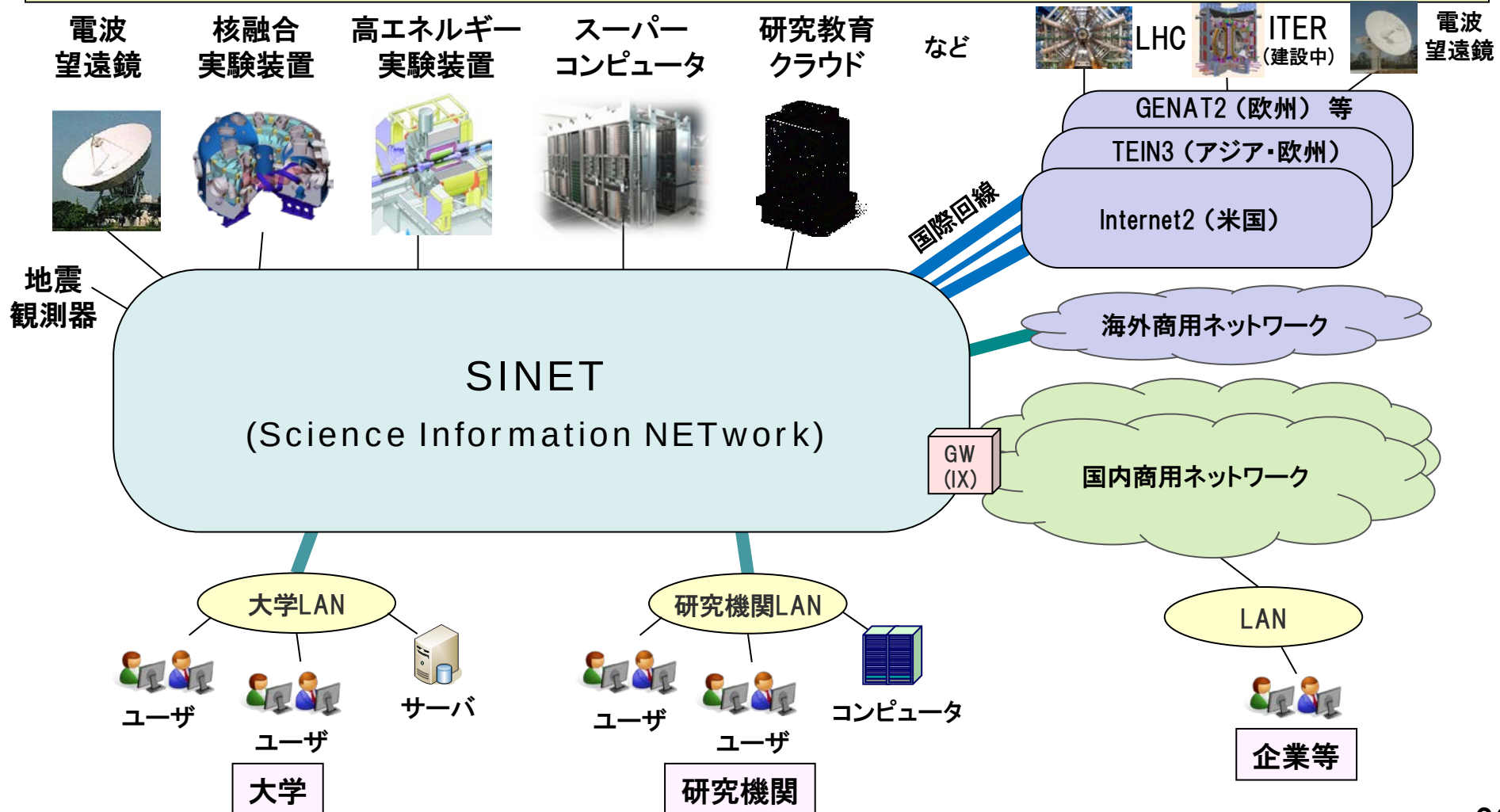
SINET4の被災状況(回線)

- ◆ 異経路二重化回線と迂回経路のおかげで、全エリアで通信経路を確保
 - ・ 7回線の現用系が切断。仙台－東京間、仙台－金沢間は予備系も切断(＝回線断)
 - ・ しかしながら、すべてのエリアで迂回経路を確保でき、東北エリアは孤立を回避
- ◆ サービス毎の迂回機能は、確保できた通信経路へ即座にトラフィックを迂回



SINET4の役割と将来

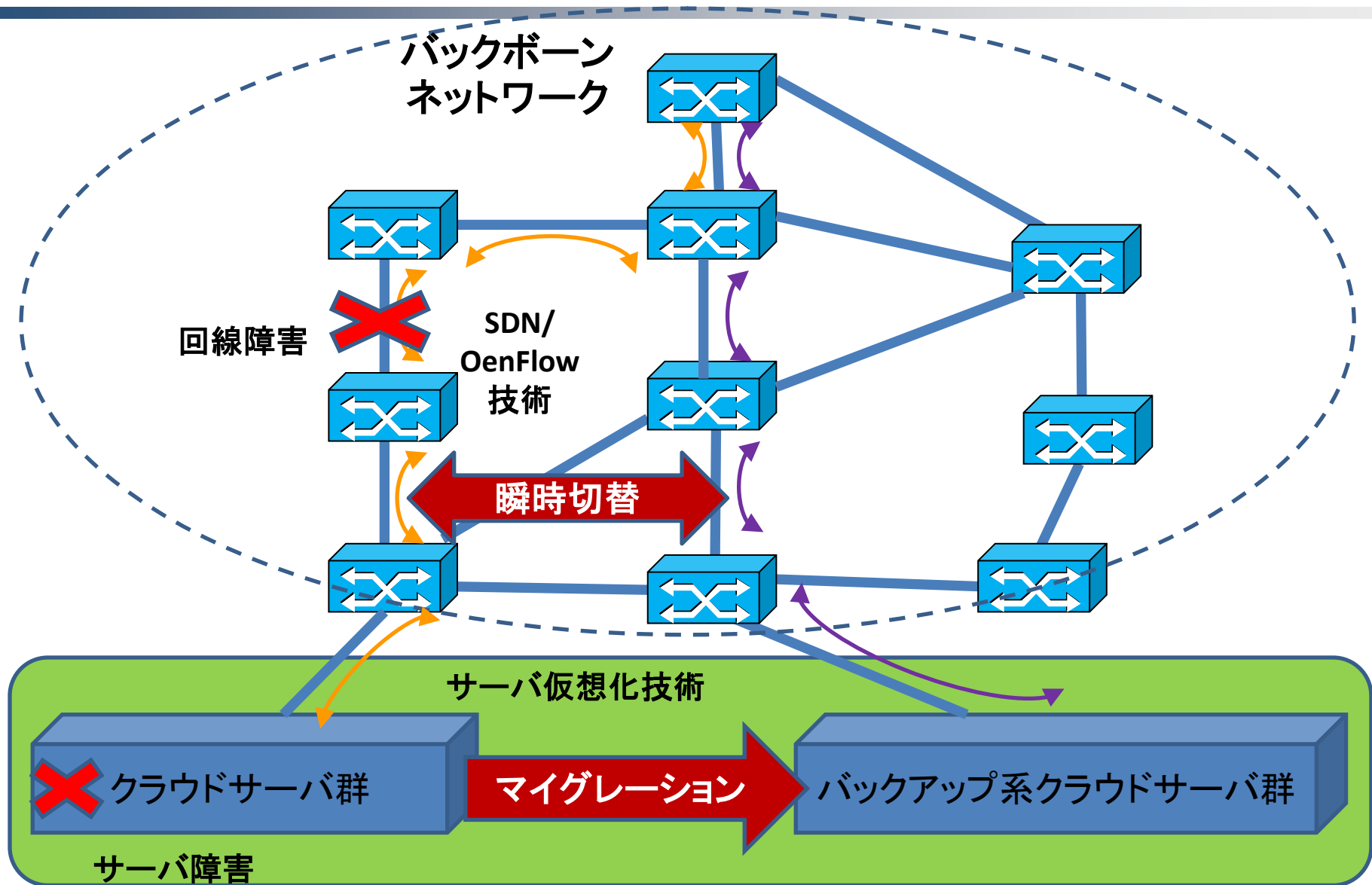
- ◆ 文科省からの予算で日本の700以上の大学、研究機関等を接続する情報通信ネットワーク
- ◆ 学術研究・教育の生産性や質の向上、新しい未来価値や知的ブレークスルーの創出を支援
- ◆ 2016年度に次期SINET運用開始予定。引き続き、SINETへの御支援と御協力をお願いします



レジリエントネットワークの必要性

- 大災害や社会危機などのクライシスには、復元力(レジリエンス)の強い社会・生活空間を構築するため、根本から見直した新たなアプローチが重要。
- 日本学術振興会「クライシスに強い社会・生活空間創成研究プロジェクト」先導的研究開発委員会で議論:レジリエントネットワークは、重要な研究開発課題の一つ
- レジリエンスの（我々の）定義
 - 絶対ダウンさせないという考え方でなく、ダウンしても、**復元力(レジリエンス)を持って早期に回復**するという考え方
- バックボーンネットワークのレジリエンス
 - 障害時に、ダウンしている回線／ノードから機能している回線／ノードに高速に切り替える技術が基本(**ネットワークの部分修復**)
 - SINET4では固定バックアップルートへの切り替えである程度のレジリエンス化は実現。しかし、ユーザに**サービスレベルの無瞬断**を提供するには、更なるレジリエンスの向上が必要。
- アクセスネットワークのレジリエンス
 - 障害時、全ダウンの可能性。修復よりも、新たに別なアクセスネットワークを迅速に構築する技術が重要(**別ネットワークの早期構築**)。
 - 災害時にも、ユーザに平常時から**使い慣れたネットワーク環境**を提供することが望ましい。

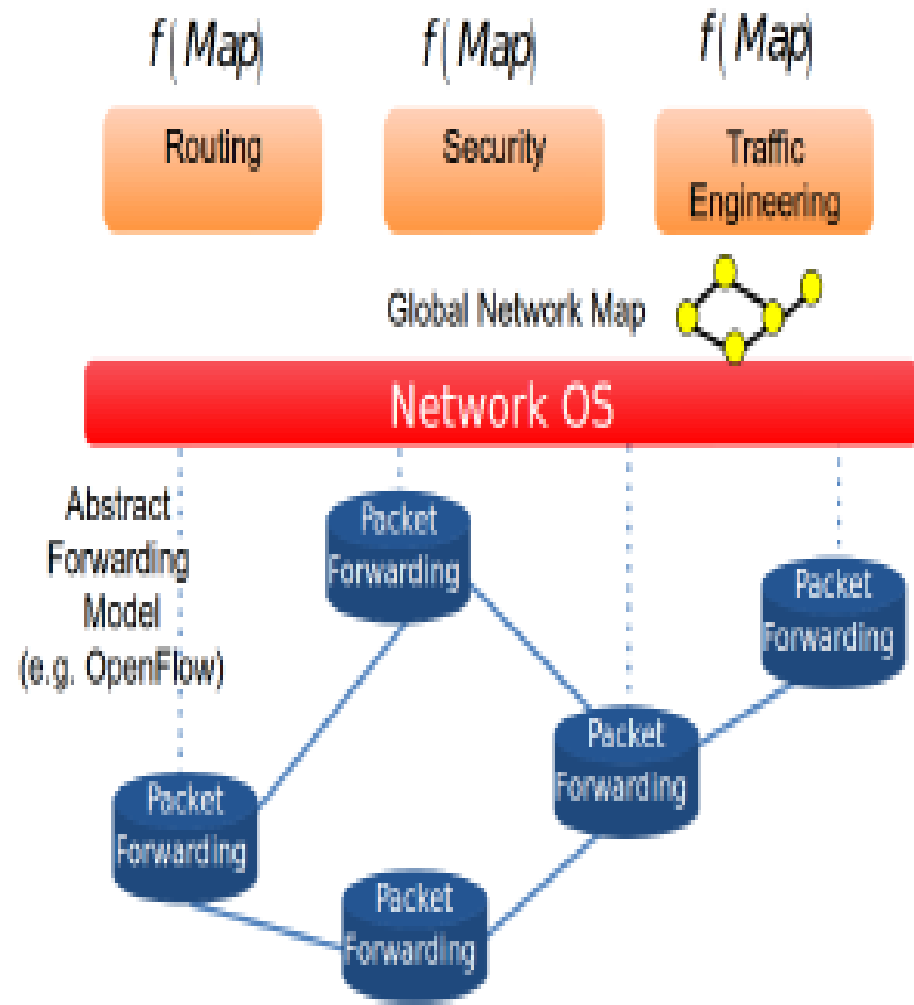
レジリエント・バックボーンネットワークの研究



レジリエントバックボーンネットワークの実現法

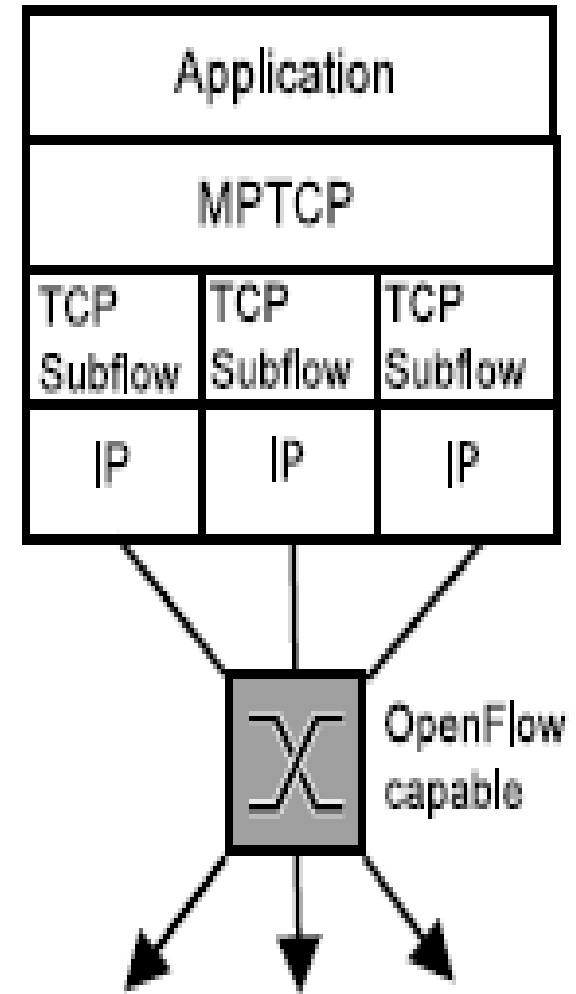
- SDN (Software Defined Networking)/OpenFlow技術の活用
- SDN:集中化したプログラム(Controller)によってネットワーク全体の制御を行うアーキテクチャ
- OpenFlow: Controllerとネットワーク機器の間のプロトコルとインタフェース
- レジリエンス向上手法: 通常は回線をマルチパスで負荷分散、回線障害時にはOpenFlowで正常パスの回線に瞬時切り替える。
- リンクレイヤとトランスポートレイヤの2種類の方法で実現
- 最終的目標はサーバ/コンピュータシステムと連動してサービスレベルでの無瞬断化

SDN Architecture

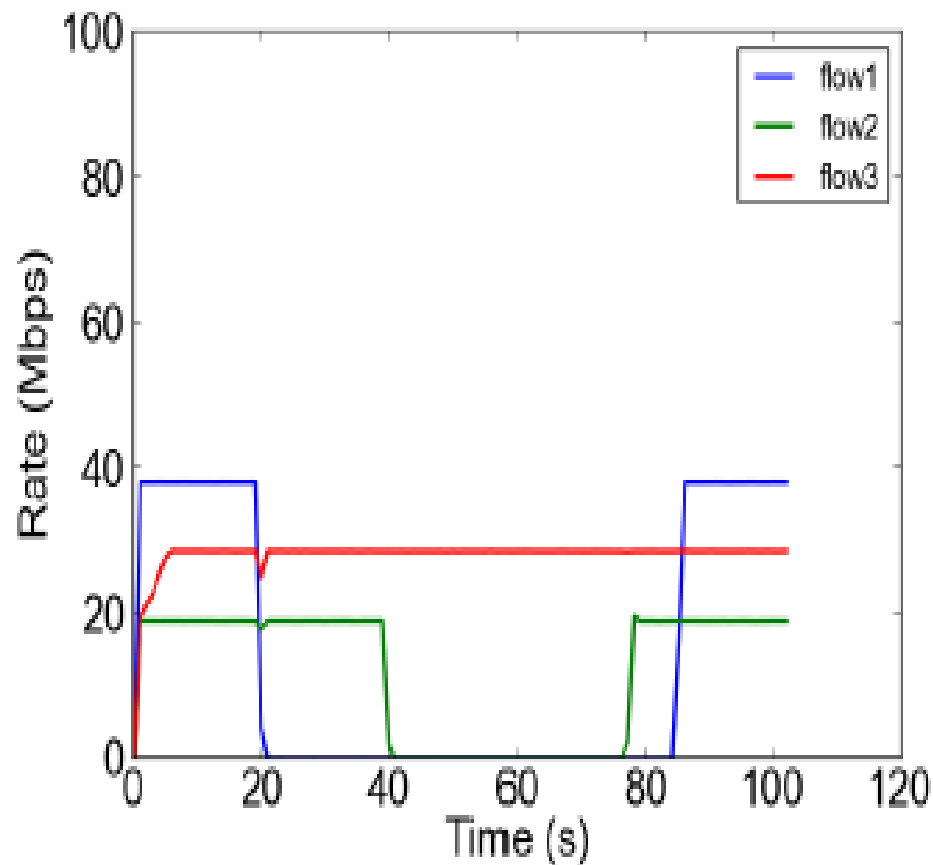
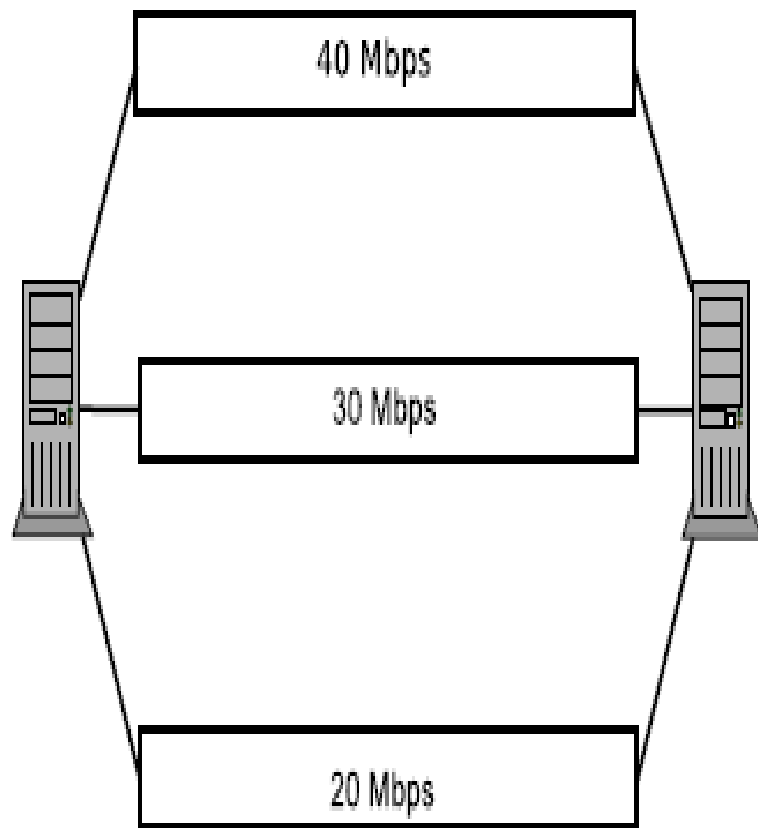


トランスポートレイヤでの実現:MPTCP (1)

- MPTCP: Multi Path TCP (Transmission Control Protocol)
- IETF (Internet Engineering Task force)で2013年1月に標準化
- 目的はリソースの最大限の活用とパスの冗長化。IPv4, IPv6対応
- 従来のアプリケーションは変更せずに使用可能: アプリケーションのデータフローが複数のサブフローに分割される
- 各サブフローは一本のTCPフローと類似の動作: エンドツエンドでパス管理、輻輳制御、負荷分散
- 我々のアプローチ: 複数パスへの分散とパス間切り替え制御をOpenFlowで行う

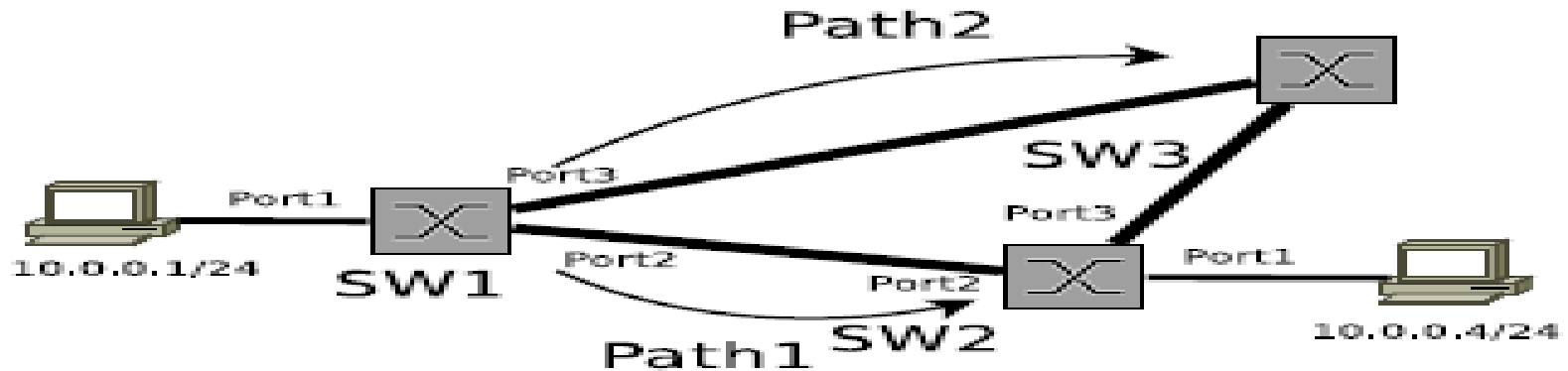


トランスポートレイヤでの実現:MPTCP (2)



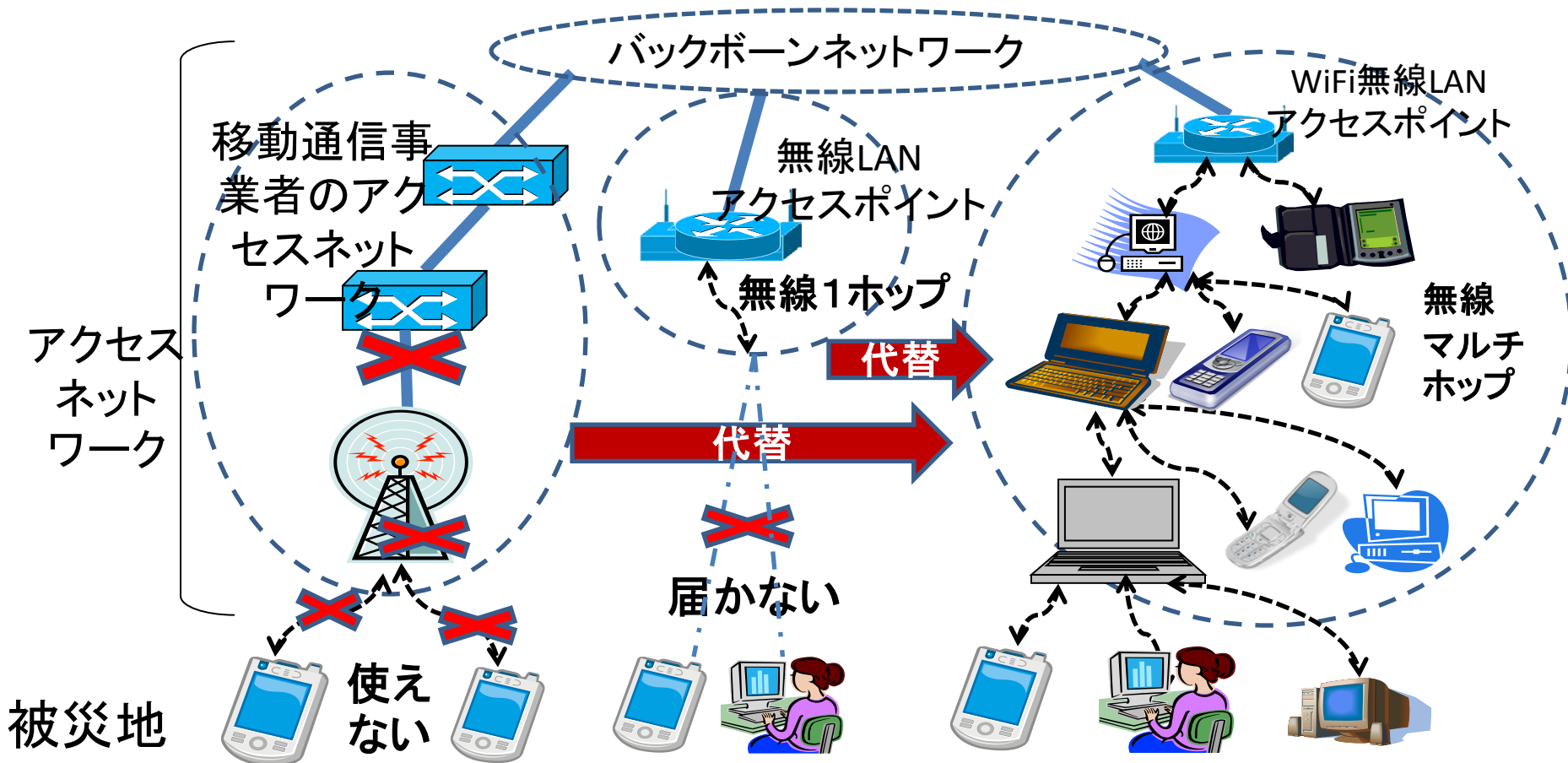
- パスの冗長化: いずれかのパスがダウンしても他のパスには影響しない。

リンクレイヤでの実現: Group Table



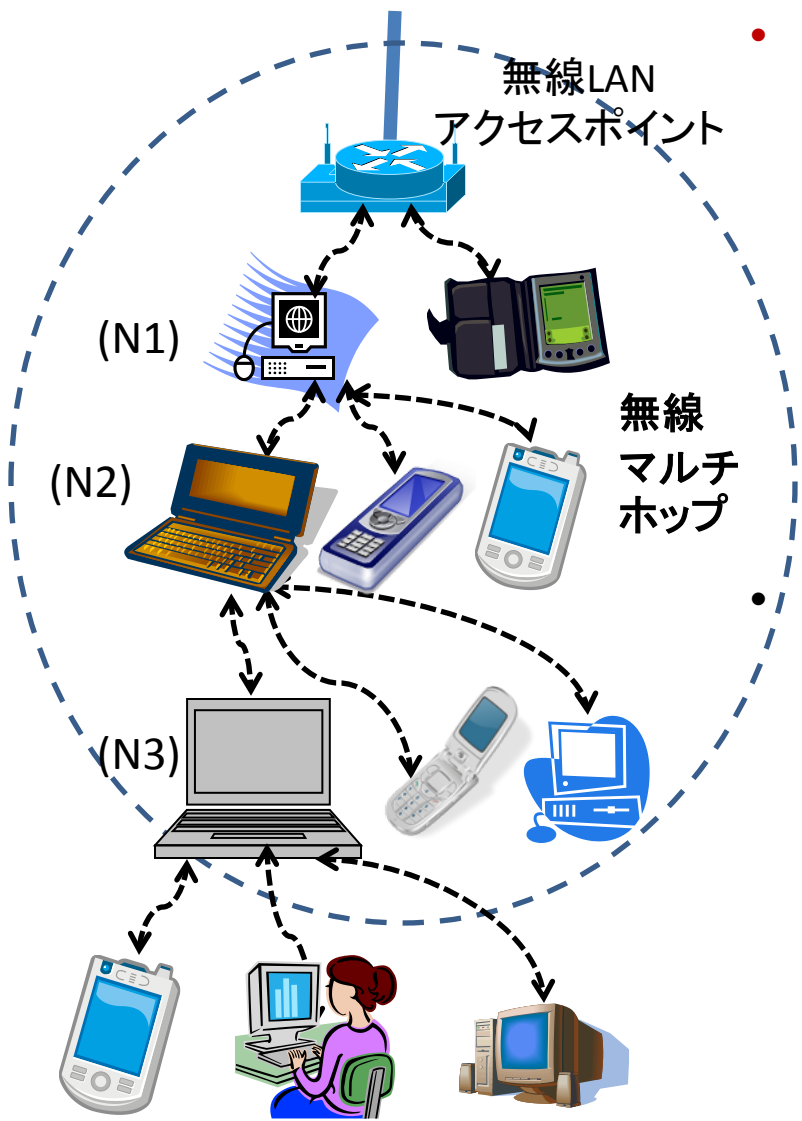
- OpenFlow 仕様の **Group Table** 機能 (複数のポートをグループ化して処理を定義) を活用。
- **Fast Failover Group** 機能
 - スイッチポート障害時にグループ内の別なポートにパケットを送出する機能
 - バックアップパス(Path2)用のポート(SW1-Port3)をSW1のFast Failover Groupに所属させる
 - 正常時: 入力トラフィックは正常パスPath1 (SW1のPort1→Port2)を通過
 - Path1障害時にはFast Failover Group機能により、入力トラフィックがバックアップパスPath2(SW1のPort1→Port3)に迂回される: 実証的に確認済み
 - バックアップパスは通常時、不使用なので、ネットワーク使用効率が悪い
- **Select Group** 機能
 - 通常は複数のパスでトラフィックを運ぶ。障害時、障害パスのトラフィックを他のパスに迂回。
 - Path1とPath2の同時並行使用(負荷分散)しながら、Path1障害時に、他の使用中回線(Path2)に迂回させる。Fast Failover Group方式より更に高速に迂回できることを実証的に確認済み

レジリエントアクセスネットワークの研究



孤立した被災地から、インターネットアクセスができるようにする。
被災者が平常時に使用のPC、スマートフォン等を用いて平常時に使用
のアプリケーション(Web メール、Web 閲覧、SNS等)がそのまま使える
ようにする。

レジリエントアクセスネットワークの実現方法



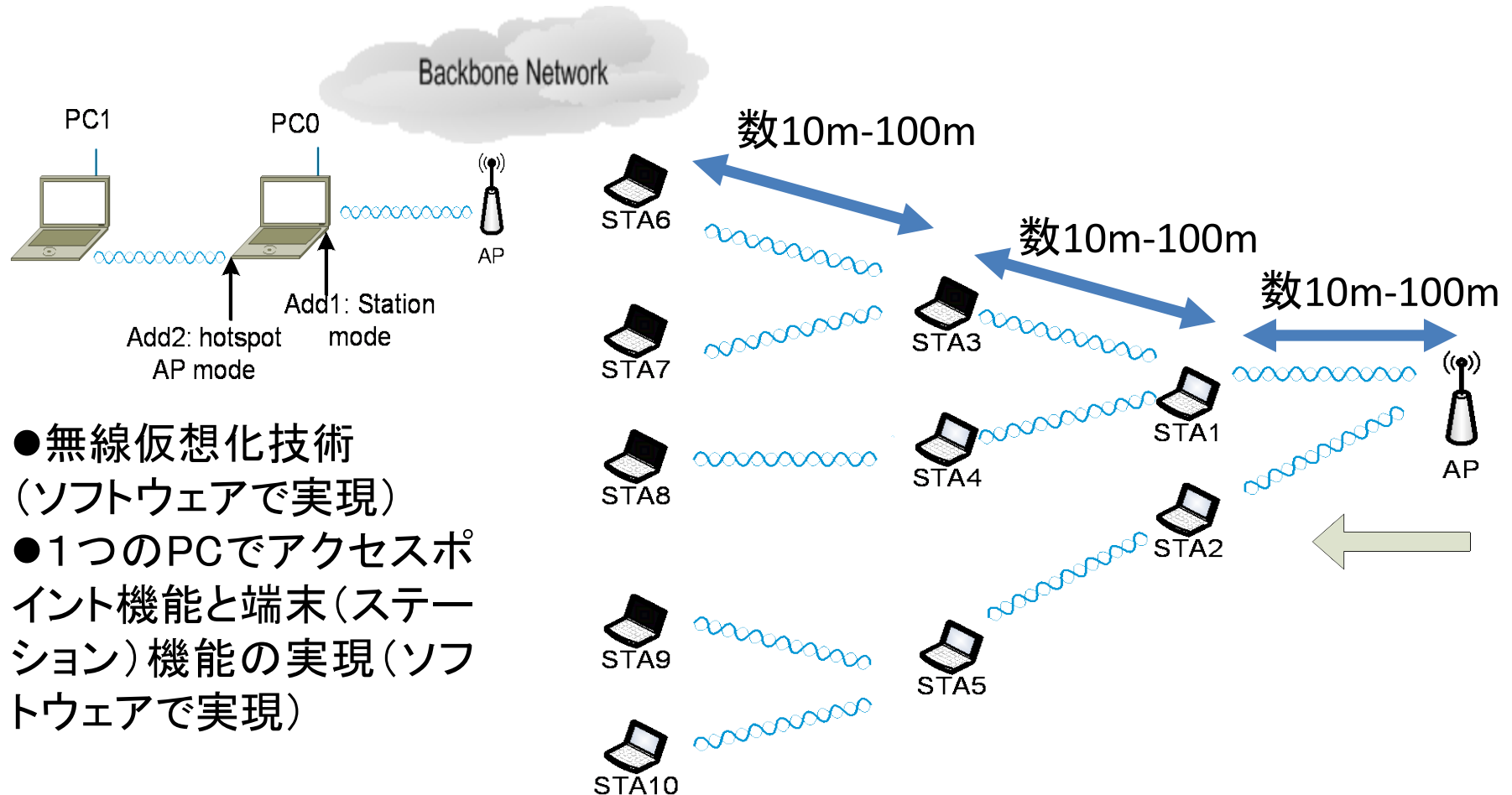
- マルチホップ仮想化

- 距離制限のあるシングルホップのWiFi(無線LAN)アクセスポイントをマルチホップ接続に拡張し、物理距離を拡大
- 近隣のWiFiアクセスポイントまでのネットワークをPC等を使って迅速に構築し、インターネットアクセス
- マルチホップをシングルホップと同程度の操作性で実現できるようにする

- 実現技術

- 無線仮想化技術: 1つの無線インタフェースカードを時分割で複数の無線リンクとして使えるようにする仮想化技術
- ソフトウェアによる無線LANアクセスポイント機能(親機) + 端末(子機)機能: PCやスマートフォンに本ソフトウェアを搭載
- 木構造ネットワーク: 各端末がアクセスポイントとなって周辺の複数の端末を收容して木構造ネットワークを構成

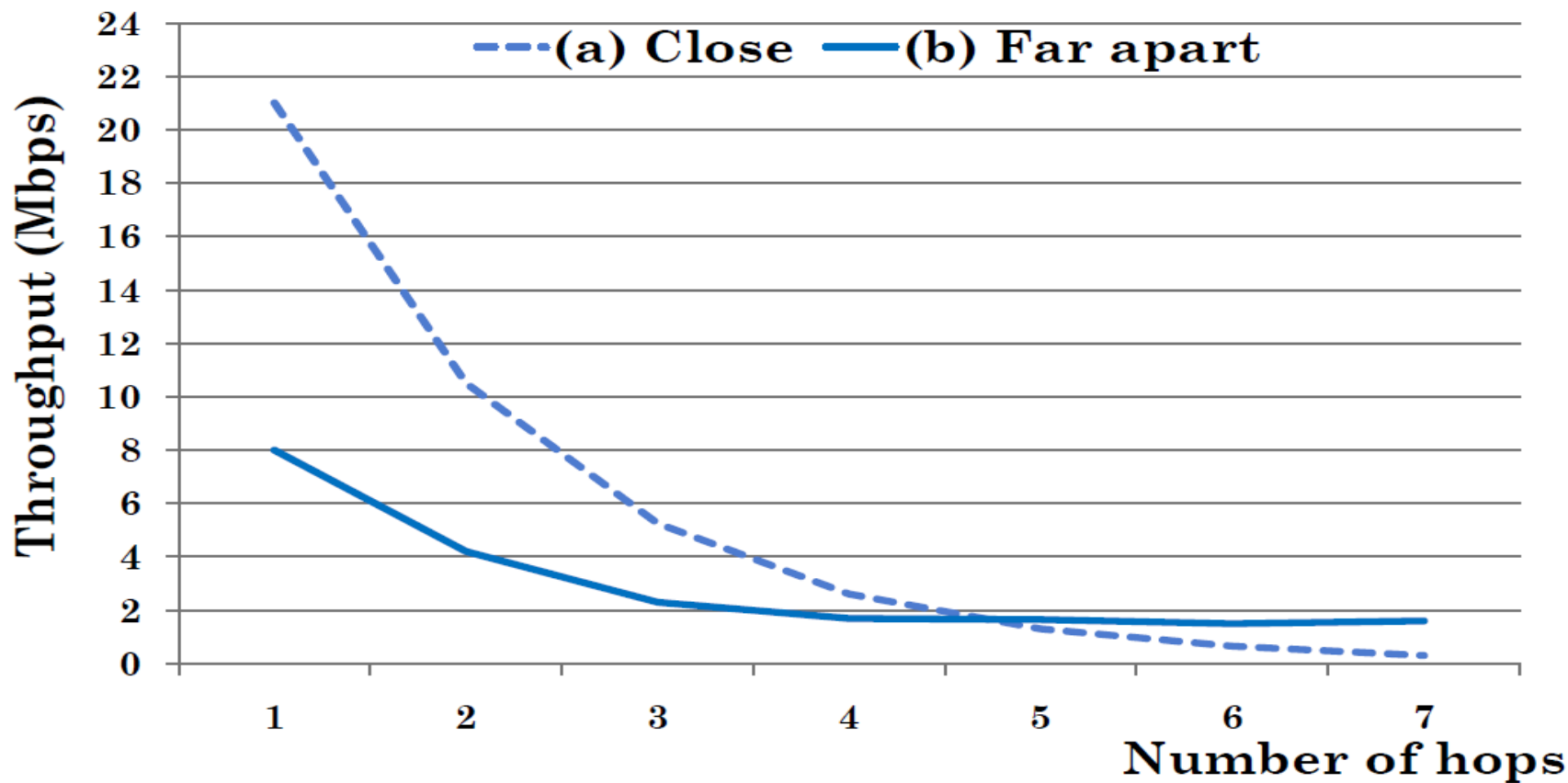
木構造アクセスネットワーク



- 無線仮想化技術
(ソフトウェアで実現)
- 1つのPCでアクセスポイント機能と端末(ステーション)機能の実現(ソフトウェアで実現)

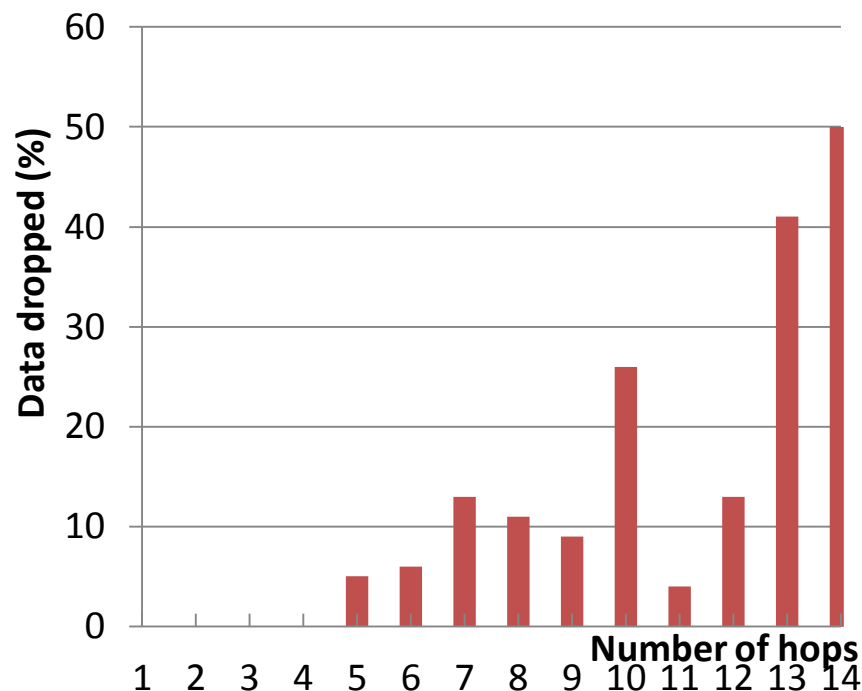
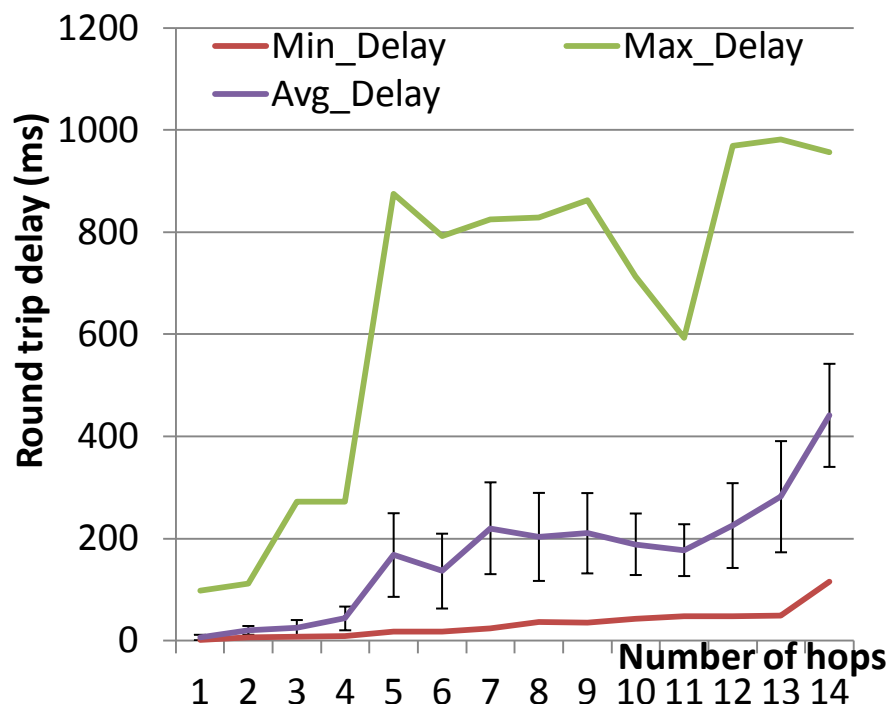
- 木構造アクセスネットワーク
(簡単なルーチングをソフトウェアで実現)

予備実験結果(1)



- ホップ数が増えるほど、木構造末端(葉の位置)のノードの転送性能が大きく低下.

予備実験結果(2)(岩手県立大学と協力)



- 以下の条件でインターネットアクセスが可能であることを実証
 - ◆PC間50m間隔で14ホップ (アクセスネットワーク延長距離：約700m)まで
 - ◆PC間100m間隔で10ホップまで(アクセスネットワーク延長距離：約1000m)まで

まとめ

- レジリエントネットワークの研究：災害でネットワークの一部がダウンしても、**復元力(レジリエンス)**を持って早期再構築を行い、サービス／アプリケーションレベルで、できる限り停止させない
- **レジリエントバックボーンネットワーク**
 - SDN/OpenFlow技術の活用
 - 通常は回線をマルチパスで負荷分散、回線障害時に正常パスの回線に瞬時に切り替える方式を検討中
- **レジリエントアクセスネットワーク**
 - マルチホップ仮想化：WiFiアクセスポイントをマルチホップ接続
 - 被災地で近隣のアクセスポイントまでのネットワークをPC等を使って迅速に構築し、インターネットアクセス
 - 本機能を実現するソフトウェアを研究開発中。**フリーウェアとして公開予定。**
- ネットワーク仮想化技術とサーバ仮想化技術(クラウド)を合わせて、**エンドツエンド**でレジリエントネットワークの実現を目指す。