

学術研究ネットワークにおける ネットワーク計測

池田 貴俊

NICT/KDDI

Ikeda 'at' kddnet.ad.jp

コミュニケーションクオリティ研究会
第7回QoSワークショップ

2009/11/20

目次

- 学術研究ネットワーク
- ネットワーク性能計測
- perfSONAR
- perfSONARの展開
- まとめ

学術研究ネットワーク

- テストベッドネットワーク
 - 先進的技術の実験・検証用
 - ネットワーク研究
 - センサーネットワーク
 - Future Internet
 - アプリケーション開発
 - 高精細映像伝送
 - E-Science アプリケーション
- 学術情報インフラ
 - 学術研究及び教育活動のインフラ
 - 科学技術振興
 - 物理学研究
 - 宇宙科学
 - 教育
 - e-Learning
 - 知的情報の共有



ネットワーク性能計測

多い問い合わせ

① あるホストとの通信で速度がでないんだけど、ネットワークに異常はないですか？

② アラームもなく、特にネットワーク機器での障害は発生していません。

③ でも、100Mbpsで接続しているのに10Mbpsしか出ないんですけど、

④ トラフィックを確認しましても、特に輻輳はございません。



研究者



運用者

⑤ そうですか、でも、何かあると思いますが。

⑥ ...

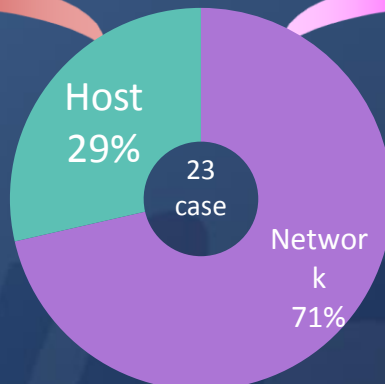
「絶対ネットワークがおかしい」

「ネットワークの異常は検知してない。ホスト側だろう」

End-to-End性能問題の要因

ホスト

- ハードウェアの性能
- TCP ウィンドウサイズ
- NICバッファサイズ
- NIC 速度とデュープレックス
- アプリケーション



*Problems inquired
at APAN Tokyo XP NOC

ネットワーク

- ネットワークの品質劣化
- ネットワークの輻輳
- 機器の高負荷or障害
- 予期しない経路制御

ネットワーク運用者の役割

- ネットワーク性能劣化を検知する
- ネットワーク内の要因を探し出し、解決する
- ネットワーク内に要因がないことを証明する

➡ ネットワークの性能計測が必要

一般的な計測

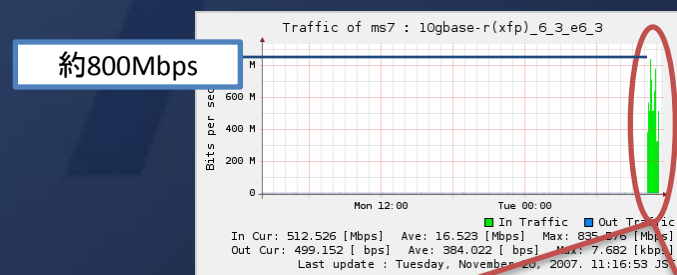
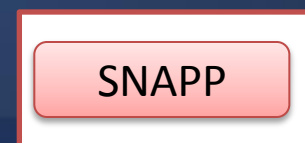
計測メトリック	計測手法	ツール
ネットワーク到達性	Ping	HP OpenView NNM, Nagios
RTT(往復遅延)	Ping	Smokeping
機器MIB情報 (IFラヒック, CPU, メモリ 使用率等)	SNMPによりMIB各種情報を 取得し, グラフ作成	MRTG, RRDtool
パケットフロー	Netflow	Peakflow

高時間分解能トラヒックグラフ

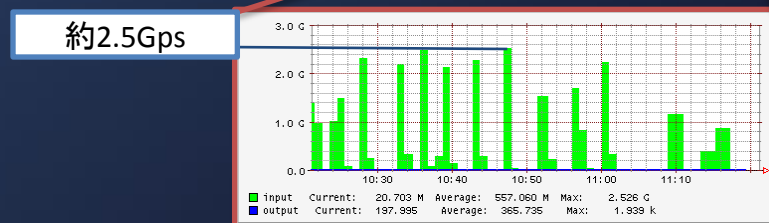
- 高時間分解能(10秒毎)でMIBデータを取得
- RRDtoolでグラフ表示
- SNAPP
 - RRDtoolのフロントエンド
 - 米国インディアナ大 GlobalNOCにて開発
 - <http://sourceforge.net/projects/snapp/>
- APAN-JP NOCデータ
 - <http://nms2.jp.apan.net/snapp/>

【利点】

- 5分間隔でデータ取得するツールと比較して、より細かいトラヒックパターン、ピーク値を表示可能。
- トラヒック発生後、10秒後にはグラフで表示されるため、準リアルタイムでトラヒックを確認可能



Cricket(5分毎)



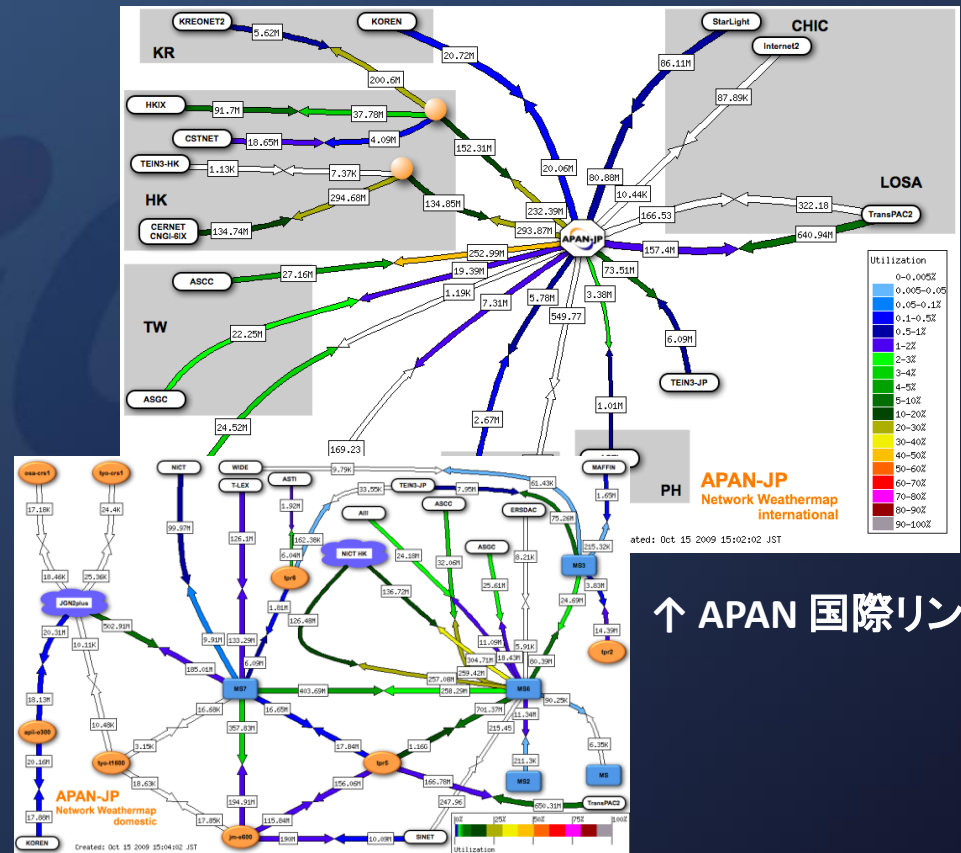
SNAPP(10秒毎)

Traffic Weather MAP

- リンク利用状況をMAP上に表示
- RRDデータを利用
- APAN-JP weather MAP
 - <http://monitor.jp.apan.net/weathermap/>

【利点】

- 複数リンクの現在のリンク利用状況が一目で把握
- トラフィックの流入、流出が一目瞭然
- デモ、イベント時のトラフィック監視に便利



↑ APAN 国際リンク

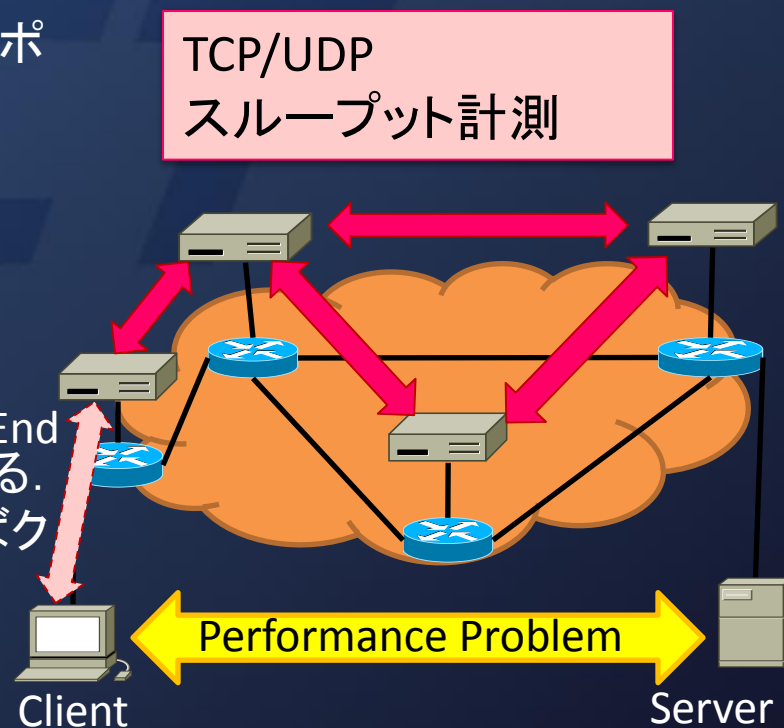
↑ APAN内リンク

スループット計測

- TCP, UDPのスループットを定常およびオンデマンドで計測
- Iperf + BWCTL
 - Iperf - ネットワークスループット性能計測
 - <http://www.dast.nlanr.net/Projects/Iperf/>
 - BWCTL - Iperfの実行スケジューリングとポリシー管理
 - <http://e2epi.internet2.edu/bwctl/>

【利点】

- 実際にトラフィックを流し、計測することでユーザ環境に近い計測が可能
- 各主要ノードに配置することで、End-to-End性能問題発生時の切り分けが容易となる。
- クライアントにIperf/BWCTLを導入すればクライアントとネットワーク間で計測



ルーティング計測

- Compath Tool

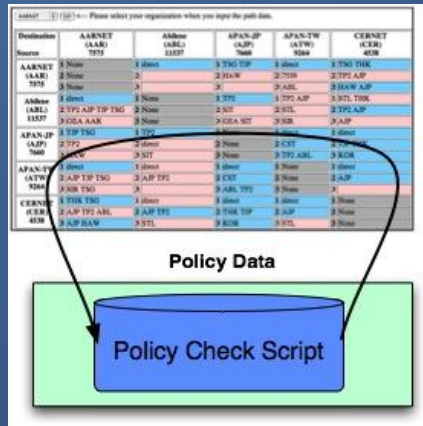
学術研究ネットワークにおいて、非効率、ポリシーに反する経路を自動検知し、それを表示するツール

- 機能

- ルートビューサーバ(Quagga)で、世界の学術研究ネットワークとBGPピアし、経路情報を収集
 - 現在27ピア
 - 各学術研究ネットワークのルーティングポリシーを収集
 - ルーティングポリシーと実際のルーティングが異なる経路を検知
 - 明らかに非効率な経路の検知
 - 大陸間ルーティング (Asia -> Internet2 -> Asia)
 - 検知した結果をWEBに表示

- <http://compath.jp.apan.net/>

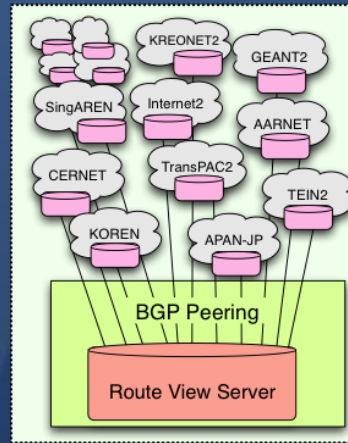
ルーティング計測(続き)



Compare the route with policy

Prefix	AS PATH	Choice AS PATH
40.250.0.0/24	7660 23855 7610 4249 I	*[7660_](24287_)(24490_)(23855_)*
58.83.128.0/18	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*
58.83.128.0/19	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*
58.83.128.0/20	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*
58.83.128.0/21	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*
58.83.136.0/21	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*
58.83.144.0/20	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*
58.83.144.0/21	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*
58.83.152.0/21	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*
58.83.159.0/24	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*
58.83.160.0/19	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*
58.83.160.0/20	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*
58.83.160.0/21	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*
58.83.168.0/21	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*
58.83.176.0/20	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*
58.83.176.0/21	7660 24287 24489 4538 9406 7497 9308 I	*[7660_](4538_)*

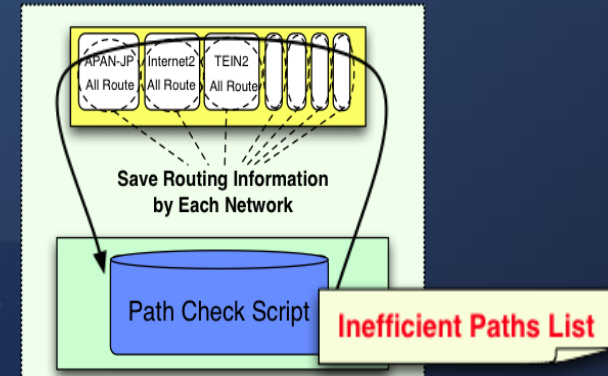
Unexpected routing



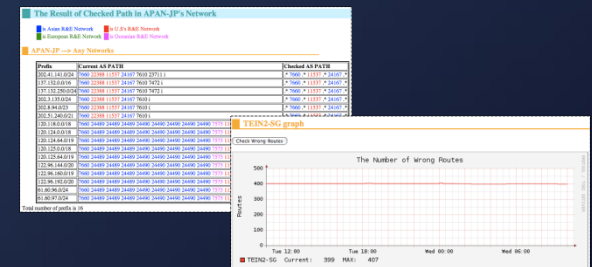
R&E Route view server

BGP Table of APAN-JP_tpr6					
BGP table version is 0, local router ID is 203.181.248.35					
Status codes: s - suppressed, d - dampened, h - history, * - valid, > best, i - Internal,					
Origin codes: i - IGP, e - EGP, ? - Incomplete					
Network	Next Hop	Metric	LocPrf	Weight	Path
6.1.0.0/16	203.181.248.144	0	0	0	7660 22388 11537 668 I
6.2.0.0/22	203.181.248.144	0	0	0	7660 22388 11537 668 I
6.3.0.0/18	203.181.248.144	0	0	0	7660 22388 11537 668 I
6.4.0.0/16	203.181.248.144	0	0	0	7660 22388 11537 668 I
6.5.0.0/19	203.181.248.144	0	0	0	7660 22388 11537 668 I
6.6.0.0/16	203.181.248.144	0	0	0	7660 22388 11537 668 I
6.7.0.0/16	203.181.248.144	0	0	0	7660 22388 11537 668 I
6.8.0.0/20	203.181.248.144	0	0	0	7660 22388 11537 668 I
6.9.0.0/20	203.181.248.144	0	0	0	7660 22388 11537 668 I
6.10.0.0/15	203.181.248.144	0	0	0	7660 22388 11537 668 I
6.14.0.0/15	203.181.248.144	0	0	0	7660 22388 11537 668 I
8.6.244.0/23	203.181.248.144	0	0	0	7660 22388 11537 11096 6356 I
8.6.245.0/24	203.181.248.144	0	0	0	7660 22388 11537 11096 I
9.4.0.0/16	203.181.248.144	0	0	0	7660 22388 11537 20965 559 I
12.5.48.0/20	203.181.248.144	0	0	0	7660 22388 11537 10578 1742 I
12.6.208.0/20	203.181.248.144	0	0	0	7660 22388 11537 10578 1742 I

Current routing information



Check the routes



inefficient routing

perfSONAR

既存の計測システム

• ドメイン毎の計測システム

• 異なる計測ツールの採用

- 同計測メトリックに対する計測手法
- 保存フォーマット

• 異なるユーザインターフェース

- データ公開手法 (WEB, 専用アプリ)
- グラフフォーマット

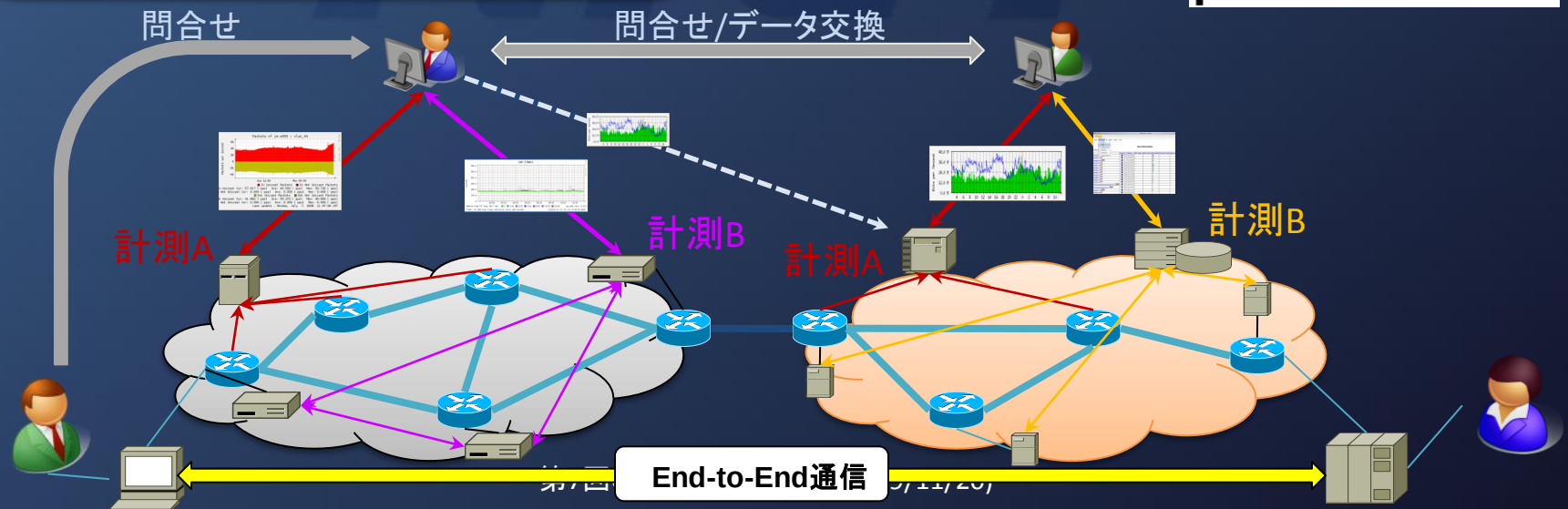
• 異なる計測ポリシー

- データ公開ポリシー

マルチドメインへの対応に課題

- ドメイン間の問合せ/データ交換
- 計測データの比較
- ドメイン間での性能計測実施

perfSONAR



Overview

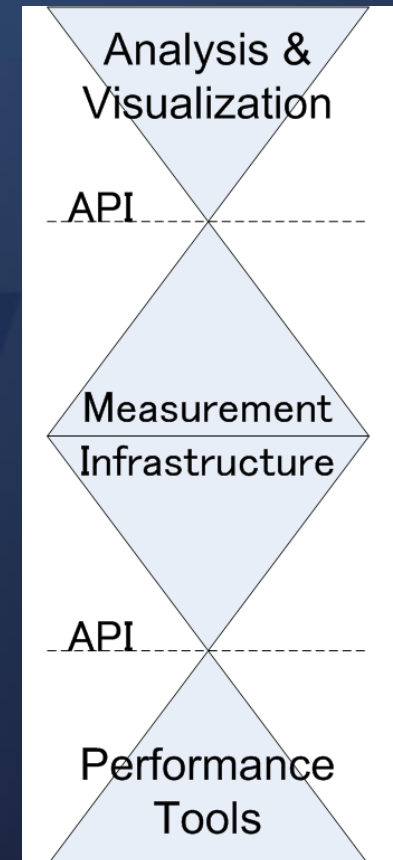
- perfSONARとは？
 - マルチドメイン対応計測ミドルウェア
 - ドメイン間でのNW計測連携のための枠組み。
 - 実計測システムと可視化アプリとの間に位置
 - 実計測システムの差を吸収
- マルチドメインネットワークで
 - ネットワーク性能計測の実施
 - 計測データの公開、交換

例えば、これを利用して、

- マルチドメインネットワークのステータス監視
- 複数メトリックのネットワークデータを自分好みに表示

Overview (Cont.)

- アーキテクチャ
 - SOA (service-oriented architecture)
 - Webサービスベース SOAP API
 - 分散型
 - 統合:
 - Network measurement tools and data archives
 - Data manipulation
 - 標準化:
 - OGF Working Group
 - Data : NM-WG (Network Measurements Working Group)
 - Protocol : NMC (Network Measurement & Control Working Group)
 - Network Topology : Network Markup Language WG
- プロジェクト
 - ESnet, GEANT, Internet2, RNP等がジョイント
 - 現在の参加者
 - <http://www.perfsonar.net/partners.html>



第7回QoS'ワークショップ (2009/11/20)



Services

	サービス	サービス内容
情報系	Lookup Service (LS)	各種perfSONARサービスの登録と検索
	Topology Service(TopS)	ネットワークトポロジーの公開
	Authentication Service(AS)	サービス利用の認証, 承認
計測系	Measurement Point Service (MP)	計測の実行とデータの公開
	Measurement Archive Service (MA)	保存計測データの公開
	Transformation Service (TS)	計測データの変換、加工
	Resource Protector Service (RP)	計測リソースの管理(競合の仲裁)

Implementation

		perfSONAR MDM v3.1	perfSONAR-PS 3.1
Developer		GEANT2 and EU folks	Internet2 and U.S. folks
Language		JAVA	Perl
Information	LS	Lookup Service(gLS, hLS)	Lookup Service (gLS, hLS)
	TopS	-	Topology Service
	AS	Authentication Service	-
Measurement	MP	SSH Telnet MP Command Line BWCT MP	pingER
	MA	RRD MA SQL MA	perfSONAR-BUOY SNMP MA Status Circuit Status
	TS	-	-
	RP	-	-

perfSONARサービス動作例

例1) インターフェース使用状況

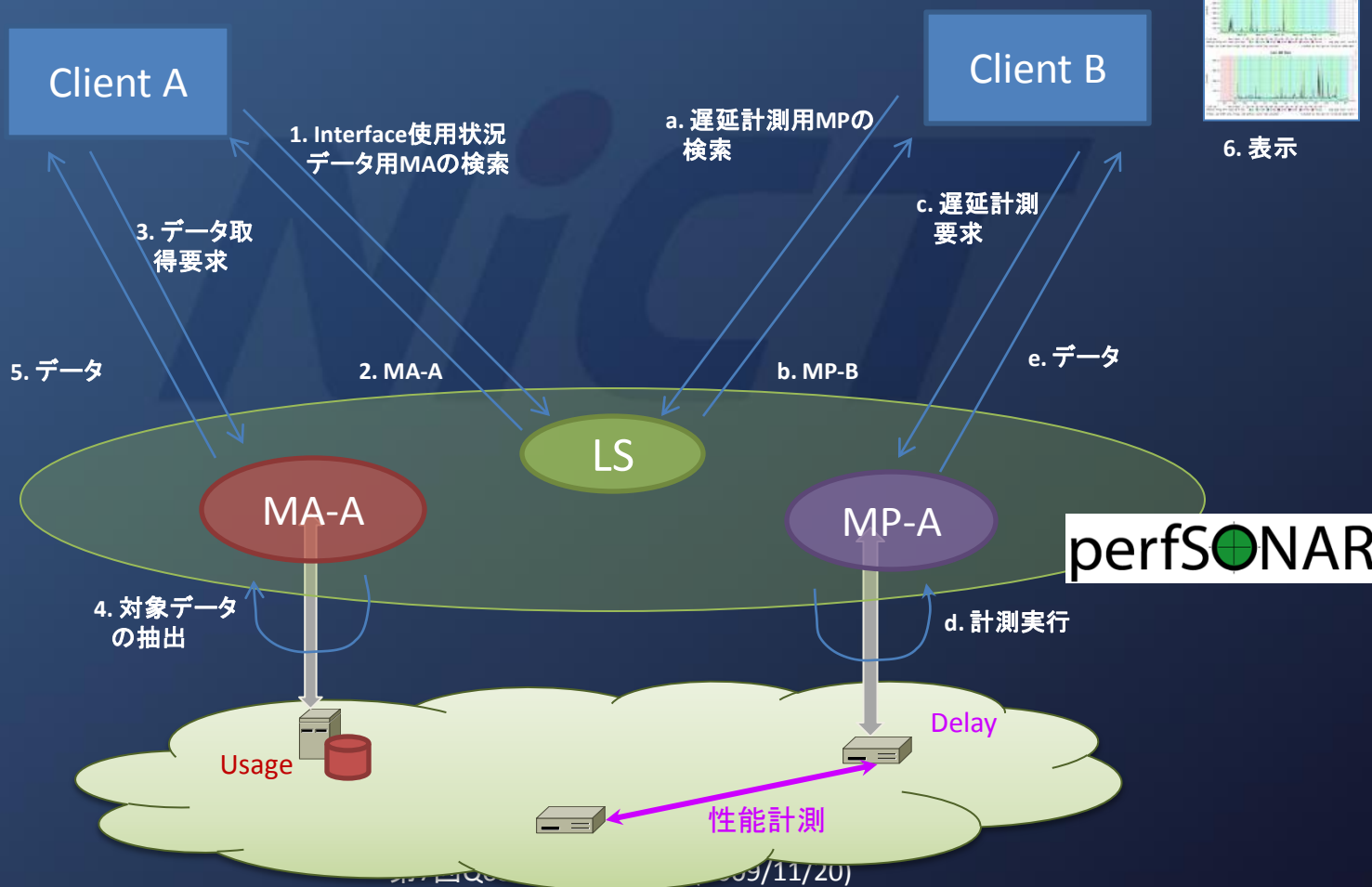


6. 表示

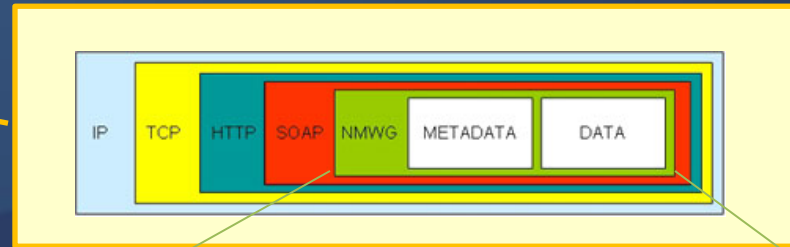
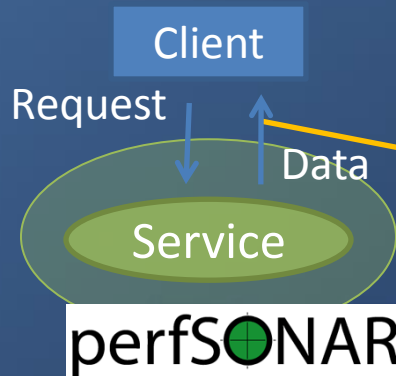
例2) 遅延計測



6. 表示



SOAP Message



Request

```
<nmwg:message id="msg1"
  type="SetupDataRequest"
  xmlns:netutil="http://ggf.org/ns/nmwg/characteristic/utilisation/2.0/"
  xmlns:nmwg="http://ggf.org/ns/nmwg/base/2.0/"
  xmlns:nmwgt="http://ggf.org/ns/nmwg/topology/2.0/"
  xmlns:select="http://ggf.org/ns/nmwg/ops/select/2.0/">
  <nmwg:metadata id="tpr5-1">
    <netutil:subject
      xmlns:netutil="http://ggf.org/ns/nmwg/characteristic/utilisation/2.0/" id="test">
        <nmwggt:interface xmlns:nmwgt="http://ggf.org/ns/nmwg/topology/2.0/">
          <nmwggt:ifAddress type="ipv4">192.168.0.1</nmwggt:ifAddress>
          <nmwggt:direction>in</nmwggt:direction>
          <nmwggt:hostname>hostname</nmwggt:hostname>
          <nmwggt:ifName>so-1/1/0.0</nmwggt:ifName>
        </nmwggt:interface>
      </netutil:subject>
      <nmwg:eventType http://ggf.org/ns/nmwg/characteristic/utilization/2.0
    </nmwg:eventType>
  </nmwg:metadata>

  <nmwg:data id="data1" metadataIdRef="tpr5-2"/>
</nmwg:message>
```

Data

Data

```
<nmwg:message xmlns:nmwg="http://ggf.org/ns/nmwg/base/2.0/" messageIdRef="msg1"
  id="message.5619252" type="SetupDataResponse"><nmwg:metadata
  xmlns:nmwg="http://ggf.org/ns/nmwg/base/2.0/" id="metadata.9571913" metadataIdRef="tpr5-2">
    <netutil:subject xmlns:netutil="http://ggf.org/ns/nmwg/characteristic/utilization/2.0/" id="tpr5-s-in-netutil-
    29">
      <nmwggt:interface xmlns:nmwgt="http://ggf.org/ns/nmwg/topology/2.0/">
        <nmwggt:ifAddress type="ipv4">192.168.0.1</nmwggt:ifAddress>
        <nmwggt:hostname>hostname</nmwggt:hostname>
        <nmwggt:ifName>so-1/1/0.0</nmwggt:ifName>
        <nmwggt:ifDescription>Description</nmwggt:ifDescription>
        <nmwggt:ifIndex>29</nmwggt:ifIndex>
        <nmwggt:direction>in</nmwggt:direction>
        <nmwggt:capacity>10000000000</nmwggt:capacity>
      </nmwggt:interface>
    </netutil:subject>
    <nmwg:eventType>http://ggf.org/ns/nmwg/characteristic/utilization/2.0</nmwg:eventType>
    <nmwg:eventType>http://ggf.org/ns/nmwg/tools/snmp/2.0</nmwg:eventType>
  </nmwg:metadata>

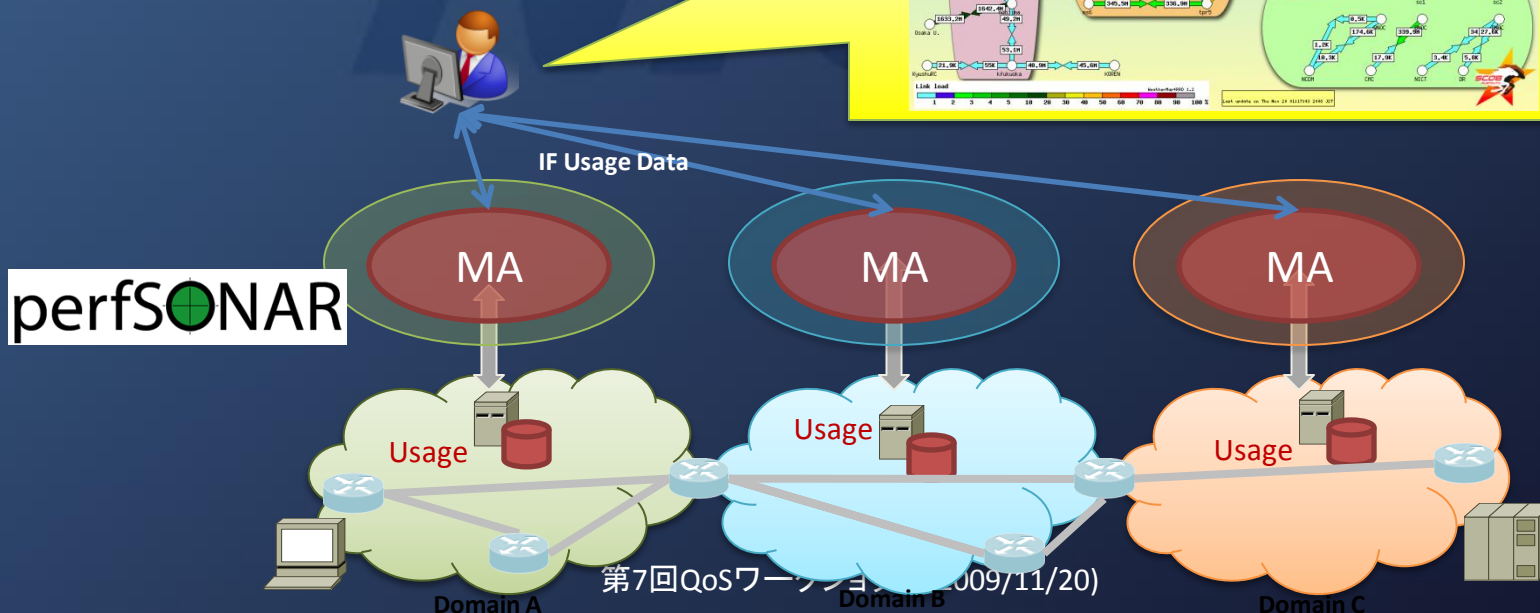
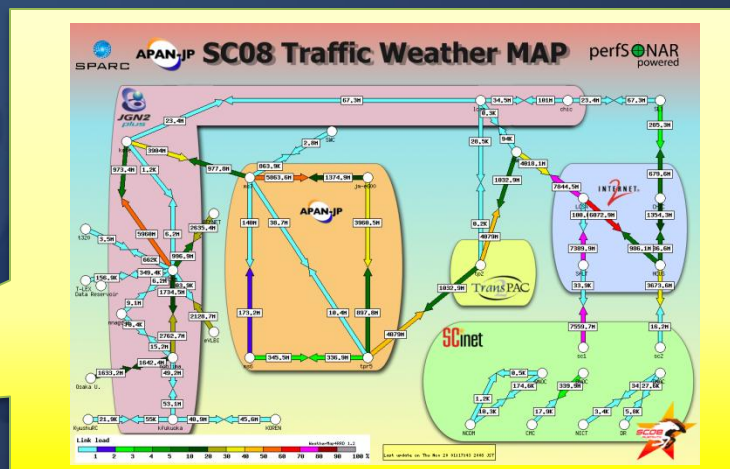
  <nmwg:data metadataIdRef="metadata.9571913" id="data.6546994">
    <nmwg:datum timeType="unix" value="1.9313008150e+07" valueUnits="Bps" timeValue="1222307360"/>
    <nmwg:datum timeType="unix" value="1.9043827930e+07" valueUnits="Bps" timeValue="1222307380"/>
    <nmwg:datum timeType="unix" value="1.7634486450e+07" valueUnits="Bps" timeValue="1222307400"/>
    <nmwg:datum timeType="unix" value="1.6736562060e+07" valueUnits="Bps" timeValue="1222307420"/>
  </nmwg:data></nmwg:message>
```

Data

利用例1

• マルチドメインTraffic WeatherMAP

- インターフェース利用状況をEnd-to-Endで確認
- ボトルネック箇所の特定, パケットロス箇所の特定などが容易



利用例2

- e2e Link Monitoring

- End-to-EndでL2パスの監視確認
- 障害箇所の特定が容易

Status of E2E Link CERN-FERMI-LHCOPN-002

Time of State Aggregation: 2008-09-24, 09:19:14 MEST (Cycle time: 60 s)

Operational State: Up

Administrative State: Normal Oper.

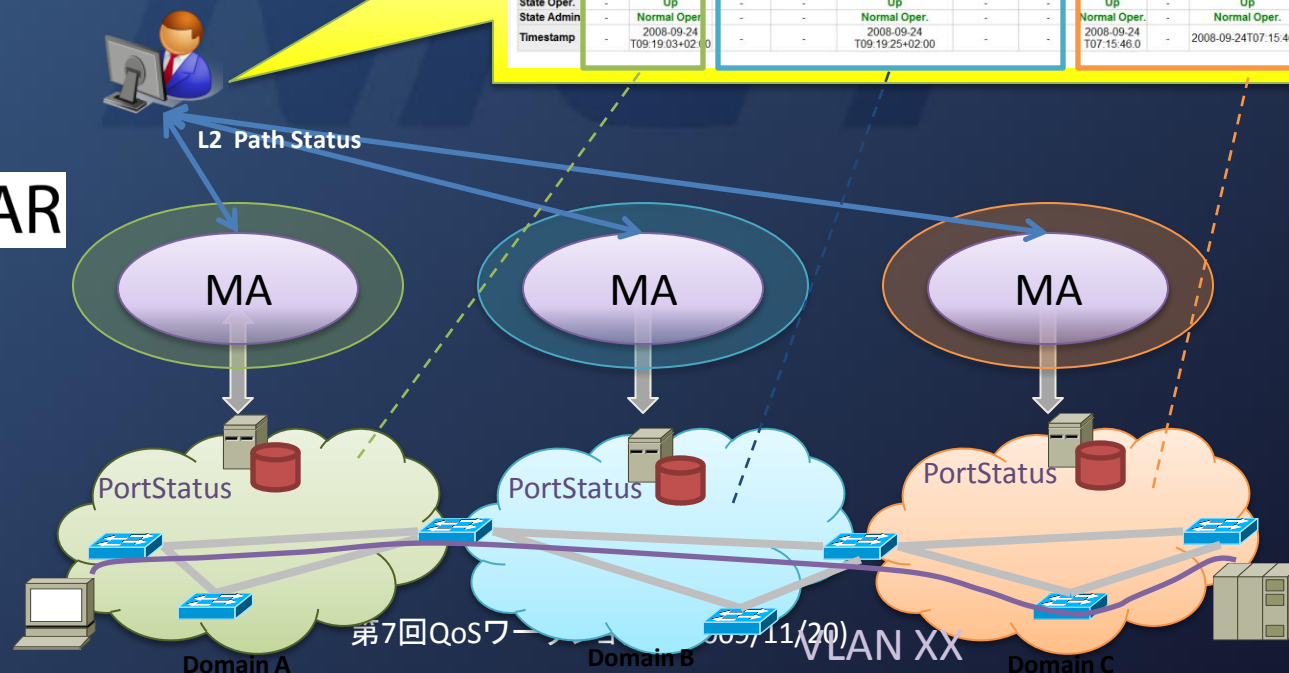
Error: E2E Link is not contiguous (End Point missing or gap found)

Warning: Operational state is not known for all involved links

Warning: Administrative state is not known for all involved links

Domain	CERN	USLHCNET	ESNET
Link Structure	EP	DP	DP
Type	EndPoint	Domain Link	Domain Link
Local Name	CERN-T0	USLHCNET-GEN	ESNET-AOFA1
State Oper.	Up	Up	Up
State Admin	Normal Oper.	Normal Oper.	Normal Oper.
Timestamp	2008-09-24 09:19:03+02:00	2008-09-24 09:19:25+02:00	2008-09-24 07:15:46.0

perfSONAR

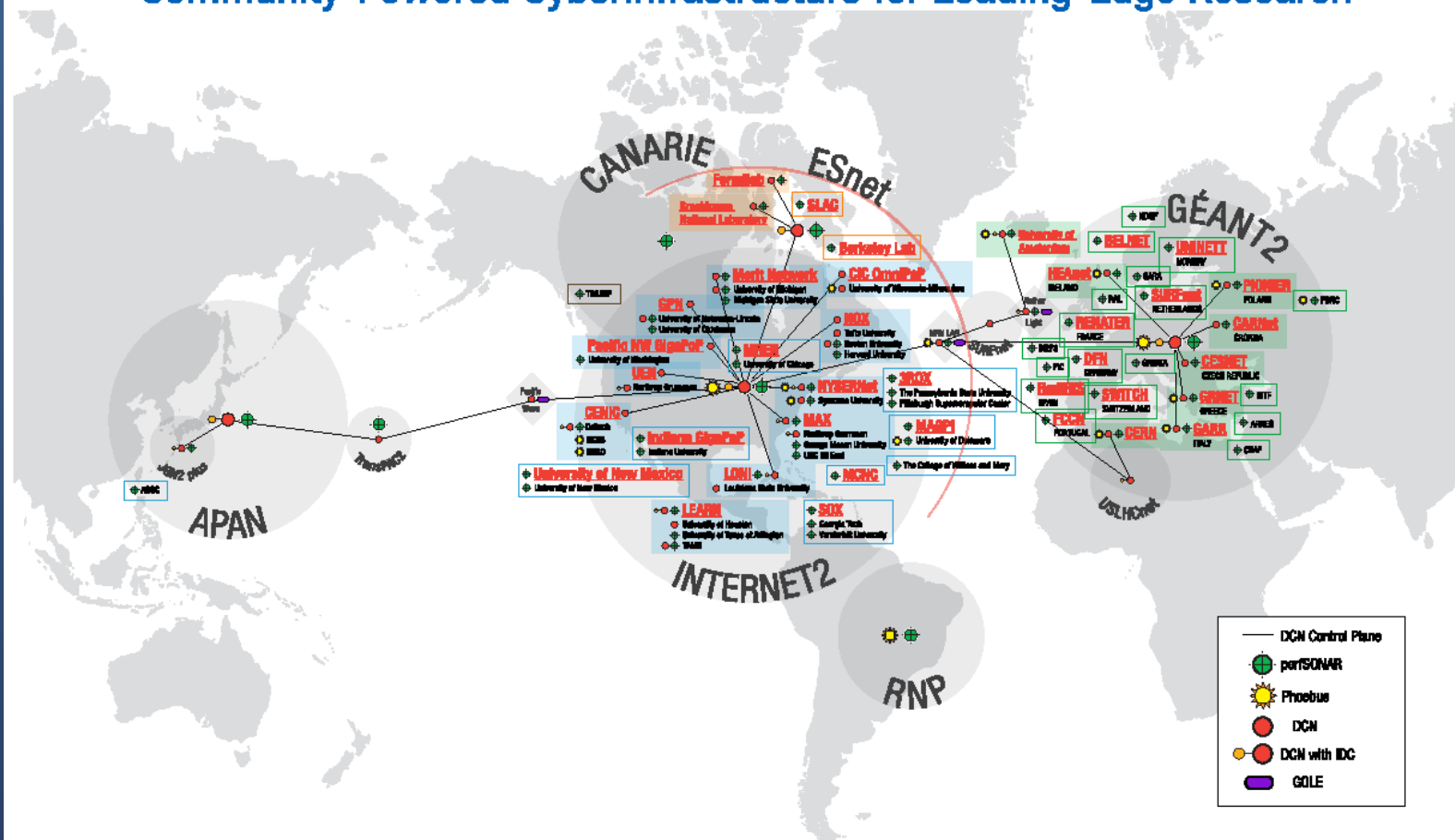


perfSONARの展開

展開マップ

2009/1現在

Community-Powered Cyberinfrastructure for Leading-Edge Research



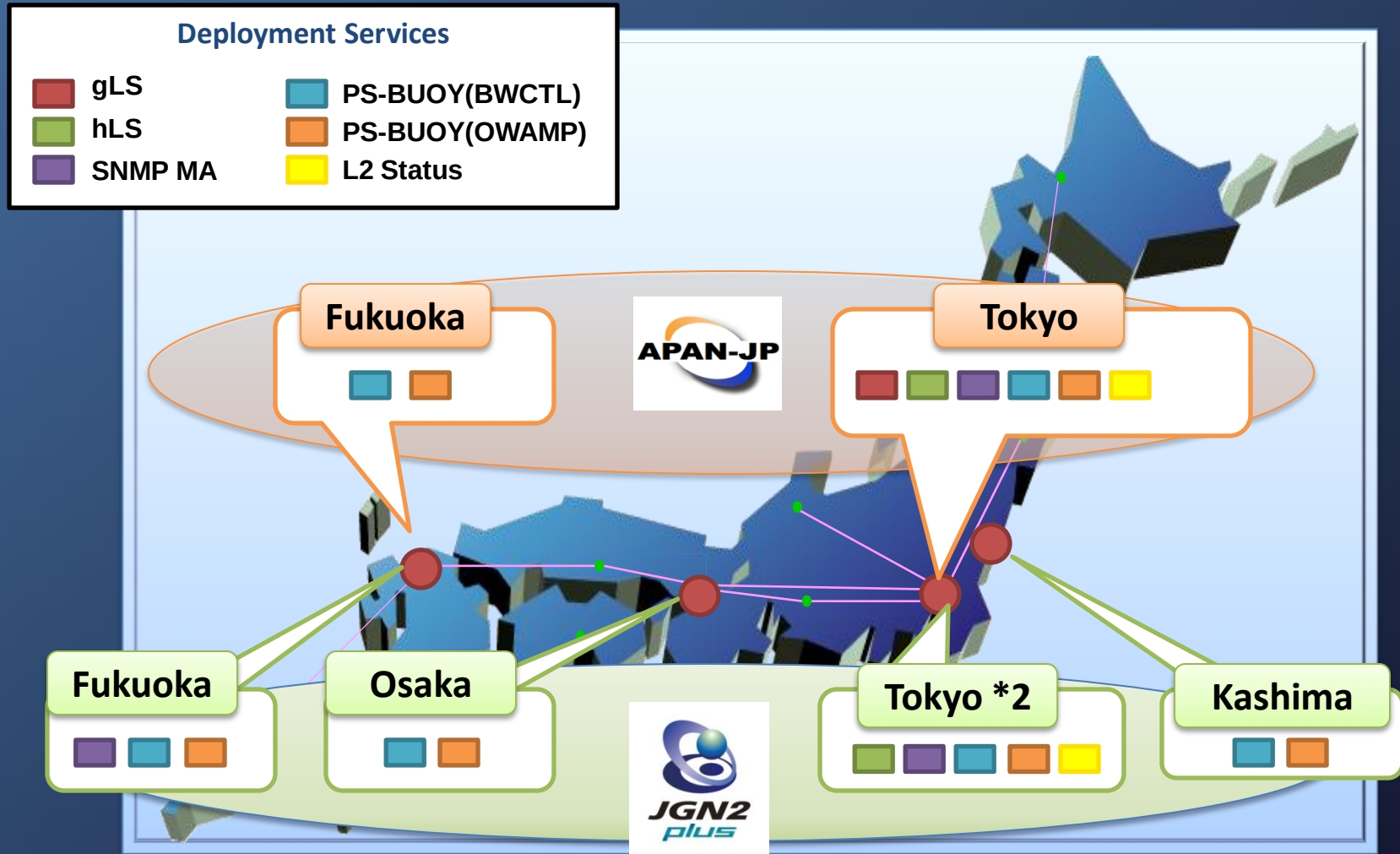
サービスリスト

- グローバルサービスリスト
 - <http://dc211.internet2.edu/cgi-bin/perfSONAR/serviceList.cgi>

Service Name	Service Type	Address	Description
MWT2_IU OWAMP Server	owamp	tcp://iua2-net1.in.edu861	OWAMP Server at MWT2_IU in Indianapolis, Indiana, USA
Internet2 OWAMP Server	owamp	tcp://iua2-net1.wash.net.internet2.edu861	OWAMP Server at Internet2 in Washington, D.C., USA
Internet2 OWAMP Server	owamp	tcp://iua2-net1.newy32aon.net.internet2.edu861	OWAMP Server at Internet2 in New York, NY, USA
Internet2 OWAMP Server	owamp	tcp://edu-1.nnu-net1.kans.net.internet2.edu861	OWAMP Server at Internet2 in Kansas City, MO, USA
Australia-ATLAS OWAMP Server	owamp	tcp://icant-cleveland10g.atlas.unimelb.edu.au861	OWAMP Server at Australia-ATLAS in Melbourne, Australia
Internet2 OWAMP Server	owamp	tcp://iua2-net1.loa.net.internet2.edu861	OWAMP Server at Internet2 in Los Angeles, CA, USA

Service Name	Service Type	Address	Description
Internet2 OWAMP Server	owamp	tcp://iua2-net1.hous.net.internet2.edu861	OWAMP Server at Internet2 in Houston, TX, USA
Internet2 OWAMP Server	owamp	tcp://iua2-net1.salt.net.internet2.edu861	OWAMP Server at Internet2 in Salt Lake City, UT, USA
Internet2 OWAMP Server	owamp	tcp://iua2-net1.atla.net.internet2.edu861	OWAMP Server at Internet2 in Atlanta, GA, USA
Australia-ATLAS OWAMP Server	owamp	tcp://iua2-net1.unimelb.edu.au861	OWAMP Server at Australia-ATLAS in Melbourne, Australia
Internet2 OWAMP Server	owamp	tcp://iua2-net1.chic.net.internet2.edu861	OWAMP Server at Internet2 in Chicago, IL, USA
OWAMP Server	owamp	tcp://polycom141.machac.org861	OWAMP Server
OWAMP Server	owamp	tcp://perfsonar.nerc.gov861	OWAMP Server
The College of William and Mary OWAMP Server	owamp	tcp://128.239.20.10.861	OWAMP Server at The College of William and Mary in Williamsburg, VA, USA
Vanderbilt University OWAMP Server	owamp	tcp://pmd2.accre.vanderbilt.edu861	OWAMP Server at Vanderbilt University in Nashville, TN, US
UW-Madison OWAMP Server	owamp	tcp://atlas-perfsonar1.chc.wisc.edu861	OWAMP Server at UW-Madison in Madison, Wisconsin, US
University of Chicago OWAMP Server	owamp	tcp://uct2-net1.chicago.edu861	OWAMP Server at University of Chicago in Chicago, Illinois, USA
T2_US_Nebraska OWAMP Server	owamp	tcp://129.93.239.182.861	OWAMP Server at T2_US_Nebraska in Lincoln, NE, USA
Indiana Purdue Fort Wayne OWAMP Server	owamp	tcp://ingtoolkit-w19.igfw.edu861	OWAMP Server at Indiana Purdue Fort Wayne in Fort Wayne, IN USA
Ocala Electric Utility OWAMP Server	owamp	tcp://216.255.240.2.861	OWAMP Server at Ocala Electric Utility in Ocala, FL, USA
PNNL OWAMP Server	owamp	tcp://192.101.102.27.861	OWAMP Server at PNNL in Richland, WA
UTA-SWT2 OWAMP Server	owamp	tcp://utsmoon1.atlas-rwt2.org861	OWAMP Server at UTA-SWT2 in Arlington, Texas, USA
Thomas Jefferson National Accelerator Facility OWAMP Server	owamp	tcp://jab4.jlab.org861	OWAMP Server at Thomas Jefferson National Accelerator Facility in Newport News, Virginia
CUHK OWAMP Server	owamp	tcp://fbl.crc.cuhk.edu.hk861	OWAMP Server at CUHK in Hong Kong
SFU - Surrey Campus OWAMP Server	owamp	tcp://142.58.190.12.861	OWAMP Server at SFU - Surrey Campus in Surrey, BC, Canada

日本での展開計画



JGN2plus プラットフォームサービス ー 計測サービス ー

- サービス

サービス	提供データ	perfSONAR-PS	段階
Lookup	各サービスの登録と検索	Lookup	提供可
MA	Interface (IF) 利用状況	SNMP MA	
MP&MA	スループット	perfSONAR-BUOY (BWCTL)	
MP&MA	片方向遅延	perfSONAR-BUOY (OWAMP *)	
Topology	ネットワークポロジ	Topology	来年を予定
MA	L2パス	Circuit Status MA Status MA	

* OWAMP – One-Way Active Measurement Protocol

- ノード

- ー NICT大手町, 大手町AP, NICT鹿島, 大阪堂島, 福岡
- ー タイ、シンガポール、香港(検討中)

JGN2plus プラットフォームサービス

－ 課題と今後の予定 －

- プロモーション
 - － 各種会合における発表
 - － デモ、実験の実施
- 利用者支援
 - － 利用マニュアル、利用例の作成
 - － 利用者初期導入支援
- 他ドメインとの連携
 - － 国内R&Eネットワークへの導入支援
 - － 国外R&Eネットワークと連携した計測環境構築

まとめ

まとめ

- 学術研究ネットワーク
 - テストベッドネットワーク
 - 学術情報インフラ
- End-to-End性能問題
 - 7割はネットワークに要因
 - 調査、解決には、ネットワーク性能計測が必要
- ネットワーク性能計測
 - 高時間分解能トラヒックグラフ
 - Traffic Weather MAP
 - スループット計測
 - ルーティング計測
- perfSONAR
 - インタードメインでの計測、データ交換の枠組み
 - 世界の学術研究ネットワークが展開
 - JGN2plusプラットフォームサービス

参考資料

- JGN2plus
 - <http://www.jgn.nict.go.jp/>
- SINET3
 - <http://www.sinet.ad.jp/>
- APAN
 - <http://www.apan.net/>
- APAN-JP NOC
 - <http://www.jp.apan.net/noc/>
- Indiana University Global NOC
 - <http://globalnoc.iu.edu/>
- Internet2 performance
 - <http://www.internet2.edu/performance/>
- perfSONAR
 - www.perfsonar.net
- Internet2 perfSONAR-PS
 - <http://www.internet2.edu/performance/pS-PS/>
- APAN perfSONAR
 - <http://www.jp.apan.net/noc/perfSONAR/>

ご静聴ありがとうございました！