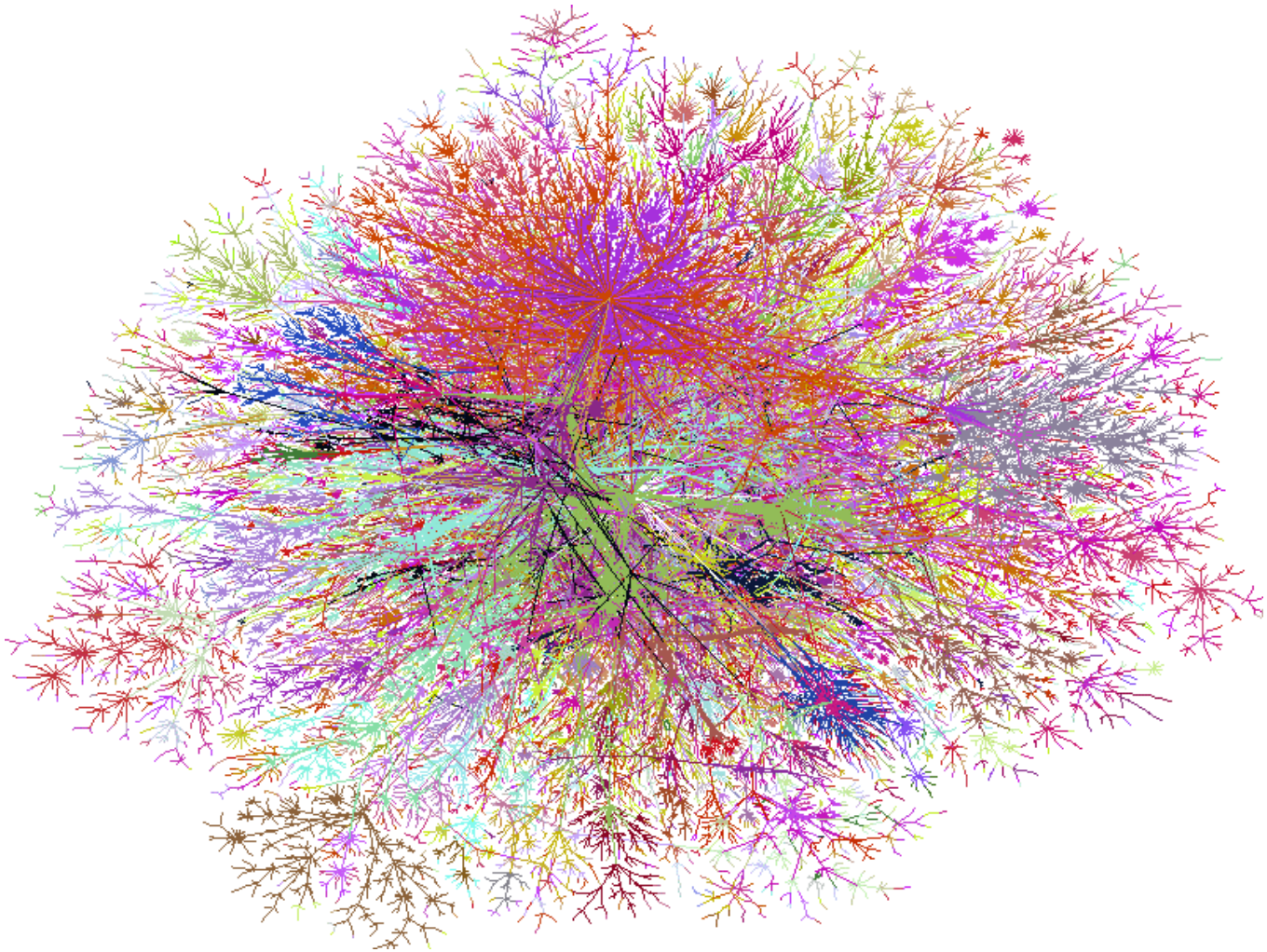
インターネット計測とデータ解析 第1回

長 健二朗

2010年9月29日

はじめに

世界中にはり巡らされたインターネットの全体像とは？



lumeta internet mapping http://www.lumeta.com

http://www.cheswick.com/ches/map/

はじめに (つづき)

世界中にはり巡らされたインターネットの全体像とは？

- 誰も把握できていない
- でも、誰もが知りたい

本授業のテーマ

- いろいろな切口からインターネットの実態を考える
  - 容易に計測できないものをどう計るか
  - 大量データからいかに情報を抽出する

このようなアプローチの仕方は今後の情報社会でますます重要となってくる

自己紹介

長 健二朗 (Kenjiro Cho)

- 肩書
  - 株式会社インターネットイニシアティブ 技術研究所 副所長
  - 慶應義塾大学環境情報学部 特別招聘教授
  - 北陸先端科学技術大学院大学 客員教授
  - WIDE プロジェクト ボードメンバー
- 経歴
  - 1984 年神戸大学電子工学科卒業。同年キヤノン (株) 入社
    - ハードウェア設計から始め、OS 屋に
  - 1993 年コーネル大学コンピュータサイエンス学科修士修了
    - コンピュータサイエンス、分散システムを勉強
  - 1996 年 (株) ソニーコンピュータサイエンス研究所入社
    - 本格的にインターネット研究 (QoS 通信、計測) を開始
  - 2001 年慶應義塾大学より博士号 (政策・メディア) 取得
  - 2004 年より (株) インターネットイニシアティブ勤務
- 専門分野
  - インターネットのトラフィック計測と解析
  - QoS 通信
  - オペレーティングシステムのネットワーク機能

インターネット計測とデータ解析

インターネット計測とデータ解析
(Internet measurement and data analysis)

- 担当教員: 長 健二朗 ⟨kjc@sfc.keio.ac.jp⟩
- TA: 空閑 洋平 ⟨sora@sfc.wide.ad.jp⟩
- URL: http://web.sfc.keio.ac.jp/~kjc/classes/sfc2010f-measurement/
- 教材・参考文献: 講義資料をオンライン配布
- 提出課題・成績評価の方法: 2 回の課題提出と学期末レポート提出

科目概要

いまや社会基盤となったインターネットの現状や挙動を把握し、今後を予想することは、技術面のみならず投資判断や政策決定にとっても重要な課題である。

しかし、大規模複雑システムであるインターネットを把握することは難しい。インターネット全体を網羅する大規模な計測は現実的でない一方で、従来のサンプリング手法も適用できない場合が多い。 さらに、技術的、社会的、経済的、法的にも多くの制約があり、その中で問題を解決する必要がある。

本授業は、インターネットの計測技術と大規模データ解析の概要について学び、情報社会で必須となる大量情報から新たな知識獲得をするための基礎能力を身につける。

主題と目的／授業の手法など

インターネット計測とデータ解析手法について学習し、ネットワーク技術と大規模データ処理の総合的な知識と理解を得る。具体的な応用例について、そこでの問題と制約、その工学的な解決手法を学び、同時に、その背後にあるネットワーク技術、数学、統計、アルゴリズムとそれらの関連を理解する。本授業は、システム系科目と解析系科目を関連づけて統合理解する科目である。

授業計画 (1/3)

- 第 1 回 イントロダクション (9/29)
  - ネットワーク計測とインターネット計測
  - ネットワーク管理ツール
  - 計測ツール
- 第 2 回 インターネットのサイズを計る (10/6)
  - ユーザ数、ホスト数
  - ウェブページ数
  - 精度 誤差 有効数字
- 第 3 回 インターネットの構造を計る (10/13)
  - インターネットアーキテクチャ
  - ネットワーク階層
  - 経路制御
  - トポロジー
  - グラフ理論
- 第 4 回 インターネットの速度を計る (10/20)
  - 速度計測
  - 利用可能帯域の推測
  - 平均 標準偏差
  - 線形回帰
  - **課題 1**

授業計画 (2/3)

- 第 5 回 インターネットの特徴量を計る (10/27)
  - 遅延、パケットロス、ジッタ
  - フロー計測
  - 相関と多変量解析
  - グラフによる可視化
- 第 6 回 インターネットの多様性と複雑さを計る (11/10)
  - ロングテールとさまざまな分布
  - サンプリング
  - 統計解析 (ヒストグラム、期待値と大数の法則、検定と信頼区間)
- 第 7 回 インターネットの時間変化を計る (11/17)
  - インターネットと時刻
  - 時系列解析
  - **課題 2**
- 第 8 回 インターネットの挙動を計る (12/1 休講、12/11 に振替予定)
  - トラフィック量
  - 経路情報
  - インターネット計測とプライバシー

授業計画 (3/3)

- 第 9 回 インターネットの異常や問題を計る (12/8)
  - 異常検出
  - スパム判定
  - ベイズ理論
- 第 10 回 データの記録とログ解析 (12/15)
  - データフォーマット
  - ログ解析手法
- 第 11 回 データマイニング (12/22)
  - パターン抽出
  - クラス分類
  - クラスタリング
- 第 12 回 スケールする計測と解析 (1/12)
  - 分散並列処理
  - クラウド技術
- 第 13 回 まとめ (1/19)
  - **最終レポートについて**

ネットワーク計測とインターネット計測

- ネットワーク計測
  - 比較的限定されたネットワークにおける計測
  - ある時点のスナップショット
- インターネット計測
  - 大規模分散開放系であるインターネットにおける計測
    - 大規模分散系
    - オープンシステム (常に変化し続ける)

インターネットの計測 – 掴みどころのないものを測る

- インターネットにおける一般的な測定データの必要性
  - 例えば、一般的なパケットサイズ分布など
- インターネットは開いた系で、つねに変化、発展、拡大
  - 中心も代表点もなく、測る場所や時間によって違う姿が観測される
  - インターネットの一般性を求める：掴みどころのないものを測る
- 現実にインターネットを運用、プロトコルや機器を開発
  - その時点で最善の一般性を模索、将来予想し、常に見直す努力
- 技術面だけでなく、社会的、政策的、経済的な影響も考慮が必要

計測の重要性

計測はすべての技術の基礎

- ネットワークにおいては、見えないネットワークを見ようとする試み
- 運用、設計、実装、研究のすべてで必要

- しかし、インターネットの商用化、利用の拡大で難しくなってきた現状
  - トラフィック情報などは事業者の企業機密で開示されない
  - プライバシー情報の漏洩リスク

計測、データ解析の目的

- 運用面
  - トラブルシューティング
  - 性能向上、信頼性向上のチューニング
  - 利用状況の把握、レポート
  - 回線容量や使用機器の中長期計画、コスト評価
- 工学面 (ソフトウェア、ハードウェア、プロトコル設計と実装)
  - 設計上のトレードオフ (バッファサイズとコスト)
  - 動作の検証
  - 予想外の現象の観測 (複雑な挙動)
- 研究面 (理論化、モデル化、新規発見)
  - ネットワークの挙動の特徴
  - モデル化 (web サービスの挙動など)
  - 複雑なシステムの挙動
    - 豊富なデータとツール
- 政策、投資計画等へのインプット

ネットワークのデータや挙動の特徴

- バラツキが大きく、偏った分布を持つ
  - パケットスイッチングの短時間にバースト的に転送する構造
  - 利用の偏り: 少数の利用者が大半のトラフィックを占めるなど
- さまざなな異常が日常的に発生
  - ソフトウェアのバグ、設定ミス、仕様の不整合、事故、メインテナンス
- さまざまな機能の相互干渉
  - 輻輳制御の例: イーサネットの衝突回避、パケットキューイング、TCP の輻輳制御、回線容量設計
- トラフィックやサービスの集約
  - 無数の要素の相互作用の結果、全体としてみれば個別要素の総和以上の独立な振舞い

計測には複合的なスキルが要求される

- 目的は運用や工学や研究
  - いずれにしても全ての視点が欠かせない
    - 動作環境に関する知識
    - 計測ツールに関する知識
    - ないものは自作する必要
- 成果は現状の把握、発見、新しい知見
  - 必ずしも研究的な新しさにこだわる必要はない
  - 事実の把握、可視化、特に長期的な解析は重要な貢献
- しかし具体的な目的を持つ事は重要
  - 実際に存在する問題を解決する
  - 何を把握する必要があるか考える

インターネット計測が難しい理由

- 大量、多様、変化するデータを扱う
- オープンな分散システムの複雑な挙動
  - 中心もなければ典型もない
  - さまざまな要因が複雑に絡む
- 動的変化
  - 適応的で障害に強いメカニズム
- さまざまな異常が日常的に発生

- いまだに体系的な理解に至っていない
  - いい教科書もない

大量データ

- インターネットの他に例をみない規模性と成長
- 解析能力を遥かに越えたデータ量
  - データサイズを小さくする必要
    - フィルタリング
    - 集約
    - サンプリング
  - 多変量の変数削減
- しかし時として詳細情報も重要
  - 大きな変化は往々にしてごく一部が引き起こす
  - 大局を見ながら、詳細にも気をくばる

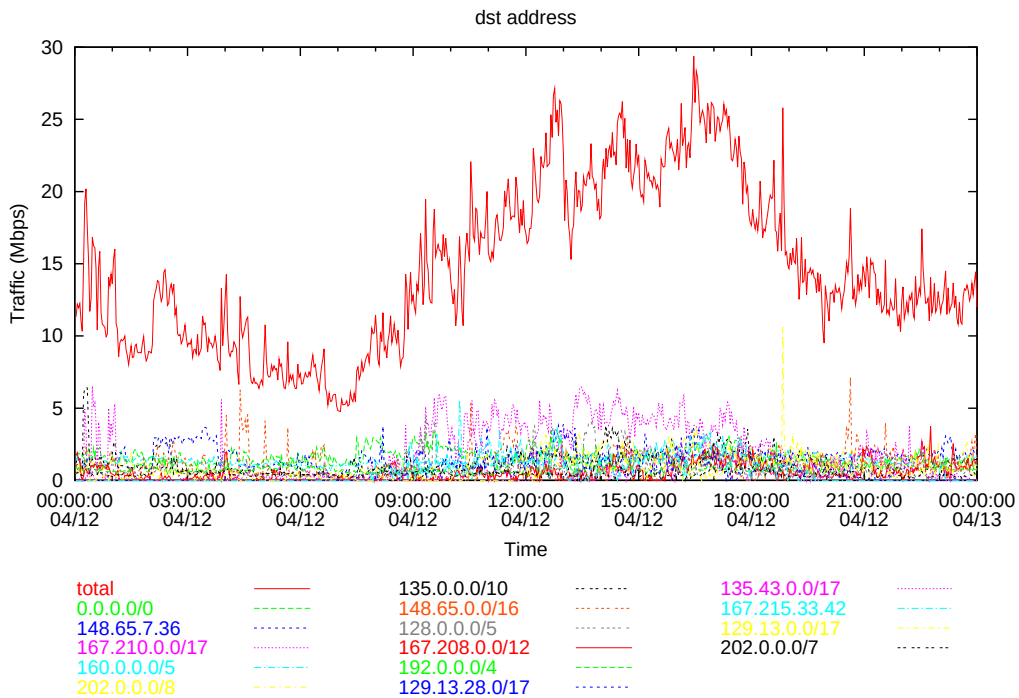
データの多様性

- 観測する場所によって異なる挙動が見える
  - 国、地域、時間
  - 企業と大学と家庭、バックボーンとアクセスネットワーク

典型的なネットワークは存在しない

時間とともに変化するデータ

- 時間帯や曜日による変化
- 長期的トレンド
  - 90 年代の web や 2000 年代の P2P ファイル共有、SNS で利用形態が大きく変化
- 将来予測は難しい



インターネット計測の制約

- 多くの問題がネットワーク境界で発生
  - 組織間協調が必要だが簡単ではない
- 測定そのものが測定対象に影響を与える
- 運用者の理解と協力が不可欠
  - 運用の現状を理解して実情にあった測定方法を工夫する必要
- 測定にはあまりコストをかけられない実情
  - 最新ルータを汎用 PC で測定する測定精度の限界
- データの解析とプライバシー、企業機密
  - 外部の研究者がデータ利用する障壁
  - 第三者が解析に使える汎用のデータを蓄積し公開する努力

まとめ

インターネットの計測とデータ解析

- 計測はすべての技術の基礎
- 掴みどころのないものを捉えようとする試み
- 技術面だけでなく、社会的、政策的、経済的な側面にも配慮

本授業のテーマ

- インターネットの計測とデータ解析を題材に
- 計測できないものをどう計るか
- 大量データからいかに情報を抽出する

ネットワーク管理ツール

一般的なネットワーク管理ツール

ネットワークの管理用ツール (もともと計測ツールではない)

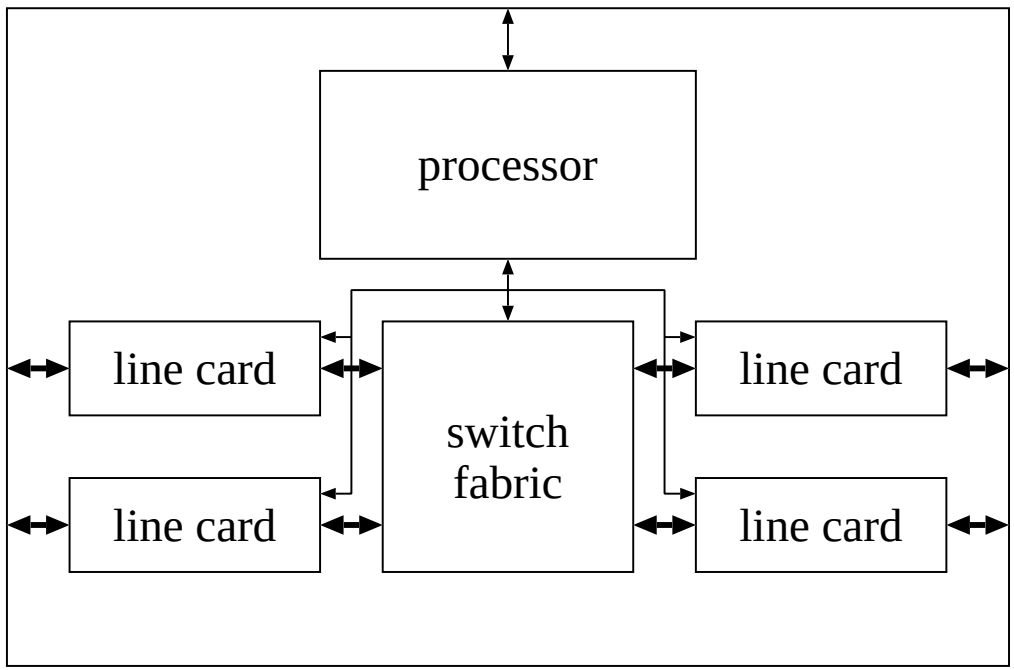
- ping
  - 到達性、ラウンドトリップタイム
- traceroute
  - 経路観測
- tcpdump
  - パケットキャプチャリング
- SNMP
  - ルータの状態把握

ping

- ネットワーク到達性確認ツール
- ICMP-echo request/reply
- 制約
  - 到達性がある $\neq$ ネットワークの正常動作
  - ICMP は遅延計測に適さない場合がある

ルータのアーキテクチャ

- fast path: ハードウェアサポート
- slow path: ソフトウェア処理
  - ICMP パケットは通常スローパスで処理



ping sample output

```
% ping -c 10 www.ait.ac.th
PING www.ait.ac.th (202.183.214.46): 56 data bytes
64 bytes from 202.183.214.46: icmp_seq=0 ttl=114 time=112.601 ms
64 bytes from 202.183.214.46: icmp_seq=1 ttl=114 time=106.730 ms
64 bytes from 202.183.214.46: icmp_seq=2 ttl=114 time=106.173 ms
64 bytes from 202.183.214.46: icmp_seq=3 ttl=114 time=111.704 ms
64 bytes from 202.183.214.46: icmp_seq=4 ttl=114 time=112.412 ms
64 bytes from 202.183.214.46: icmp_seq=5 ttl=114 time=114.603 ms
64 bytes from 202.183.214.46: icmp_seq=6 ttl=114 time=111.755 ms
64 bytes from 202.183.214.46: icmp_seq=7 ttl=114 time=115.273 ms
64 bytes from 202.183.214.46: icmp_seq=8 ttl=114 time=106.525 ms
64 bytes from 202.183.214.46: icmp_seq=9 ttl=114 time=111.562 ms

--- www.ait.ac.th ping statistics ---
10 packets transmitted, 10 packets received, 0% packet loss
round-trip min/avg/max/stddev = 106.173/110.934/115.273/3.142 ms
```

traceroute

- IP パケットのループ検出のための TTL (time-to-live) を利用
  - ルータはパケット転送時に TTL を 1 減らす
  - TTL が 0 になると ICMP TIME EXCEEDED を送信者に返す
- 制約
  - 経路は時間とともに変化する可能性
  - 非対称な経路も存在する
    - 行きのパスしか分からない
  - 通常ルータはインターフェイス毎に IP アドレスを持つ
    - IP アドレスだけでは同一ルータか判定できない

traceroute sample output

```
% traceroute www.ait.ac.th
traceroute to www.ait.ac.th (202.183.214.46), 64 hops max, 40 byte packets
 1  202.214.86.129 (202.214.86.129)  0.687 ms  0.668 ms  0.730 ms
 2  jc-gw0.IIJ.Net (202.232.0.237)  0.482 ms  0.390 ms  0.348 ms
 3  tky001ix07.IIJ.Net (210.130.143.233)  0.861 ms  0.872 ms  0.729 ms
 4  tky001bb00.IIJ.Net (210.130.130.76)  10.107 ms  1.026 ms  0.855 ms
 5  tky001ix04.IIJ.Net (210.130.143.53)  1.111 ms  1.012 ms  0.980 ms
 6  202.232.8.142 (202.232.8.142)  1.237 ms  1.214 ms  1.120 ms
 7  ge-1-1-0.toknf-cr2.ix.singtel.com (203.208.172.209)  1.338 ms  1.501 ms
 1.480 ms
 8  p6-13.sngtp-cr2.ix.singtel.com (203.208.173.93)  93.195 ms  203.208.172.
229 (203.208.172.229)  88.617 ms  87.929 ms
 9  203.208.182.238 (203.208.182.238)  90.294 ms  88.232 ms  203.208.182.234
(203.208.182.234)  91.660 ms
10  203.208.147.134 (203.208.147.134)  103.933 ms  104.249 ms  103.986 ms
11  210.1.45.241 (210.1.45.241)  103.847 ms  110.924 ms  110.163 ms
12  st1-6-bkk.csloxinfo.net (203.146.14.54)  131.134 ms  129.452 ms  111.408
 ms
13  st1-6-bkk.csloxinfo.net (203.146.14.54)  106.039 ms  105.078 ms  105.196
 ms
14  202.183.160.121 (202.183.160.121)  111.240 ms  123.606 ms  112.153 ms
15  * * *
16  * * *
17  * * *
```

tcpdump

- パケットキャプチャリングのためのツール
  - パケットの先頭 N バイトを記録
- 柔軟なフィルタリング機能
  - 例: 特定のホストからの TCP SYN パケット
- 詳細な解析が可能

- 制約
  - データサイズが大きい
  - 高速ネットワークでは技術的に困難

tcpdump sample output

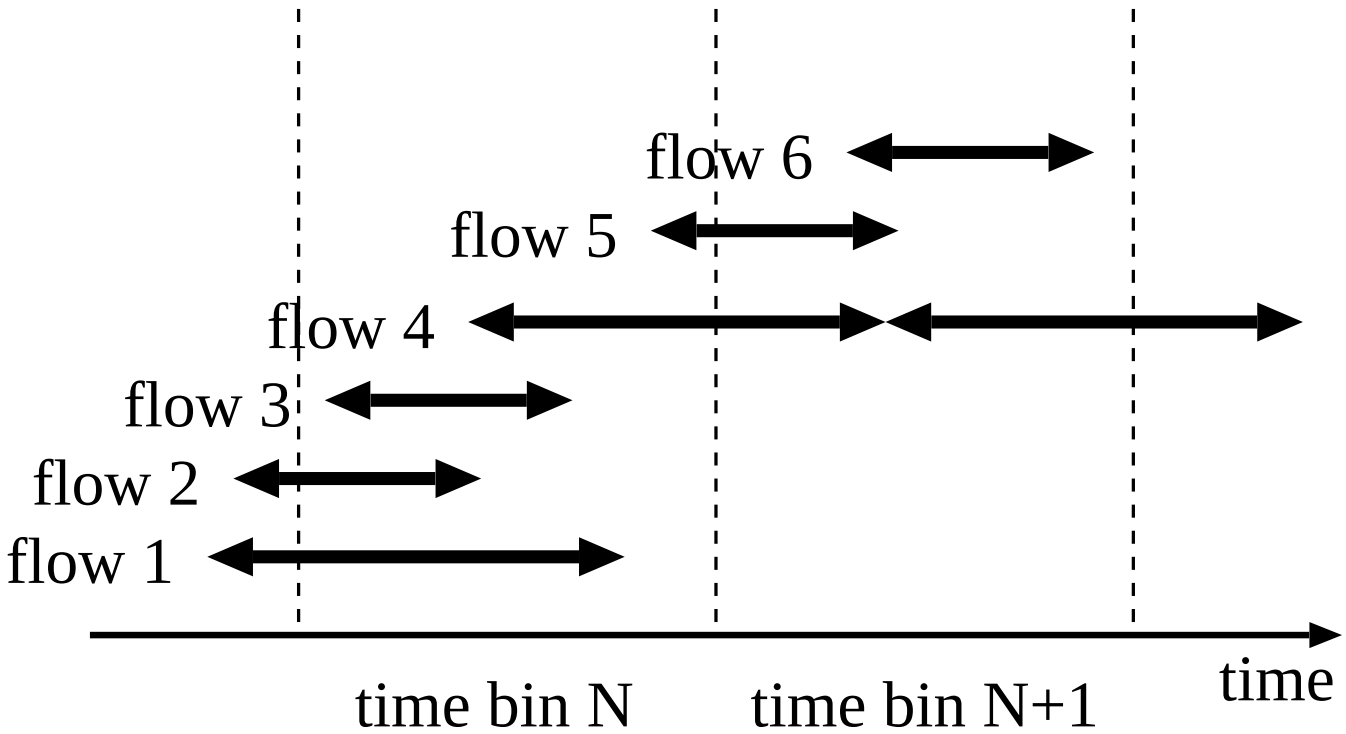
```
18:45:29.767497 IP 202.214.86.132.50052 > 202.210.220.18.80: \
  S 3304970307:3304970307(0) win 65535 <mss 1460,nop,nop,sackOK,nop, \
  wscale 1,nop,nop,timestamp 710778973 0>
18:45:29.770038 IP 202.210.220.18.80 > 202.214.86.132.50052: \
  S 3129218301:3129218301(0) ack 3304970308 win 65535 <mss 1460,nop, \
  ywscale 1,nop,nop,timestamp 2523776361 710778973,nop,nop,sackOK>
18:45:29.770090 IP 202.214.86.132.50052 > 202.210.220.18.80: \
  . ack 1 win 33304 <nop,nop,timestamp 710778973 2523776361>
18:45:29.787084 IP 202.214.86.132.50052 > 202.210.220.18.80: \
  P 1:521(520) ack 1 win 33304 <nop,nop,timestamp 710778975 2523776361>
18:45:29.791392 IP 202.210.220.18.80 > 202.214.86.132.50052: \
  P 1:222(221) ack 521 win 33304 <nop,nop,timestamp 2523776363 710778975>
18:45:29.887024 IP 202.214.86.132.50052 > 202.210.220.18.80: \
  . ack 222 win 33304 <nop,nop,timestamp 710778985 2523776363>
18:45:34.792726 IP 202.210.220.18.80 > 202.214.86.132.50052: \
  F 222:222(0) ack 521 win 33304 <nop,nop,timestamp 2523776864 710778985>
18:45:34.792763 IP 202.214.86.132.50052 > 202.210.220.18.80: \
  . ack 223 win 33304 <nop,nop,timestamp 710779475 2523776864>
18:45:42.528539 IP 202.214.86.132.50052 > 202.210.220.18.80: \
  F 521:521(0) ack 223 win 33304 <nop,nop,timestamp 710780249 2523776864>
18:45:42.531088 IP 202.210.220.18.80 > 202.214.86.132.50052: \
  . ack 522 win 33303 <nop,nop,timestamp 2523777637 710780249>
```

SNMP (Simple Network Management Protocol)

- SNMP
  - リモートから情報問い合わせ、情報の格納、トラップの設定
  - UDP の利用 (信頼性がない)
- 標準化されたトラフィック統計情報
  - ほとんどのルータ、スイッチ、ホスト OS に実装
  - 多くの管理ツールが存在
- MIB (Management Information Base)
  - SNMP オブジェクトの木構造データベース
    - 例: interfaces.ifTable.ifEntry.ifOutOctets
    - 標準 MIB とプライベート MIB
  - get, set, get-next to access MIB
- 制約
  - 標準化されている情報は限られている
  - オブジェクトのサポートを後から追加することは難しい
  - アクセスの効率が悪い

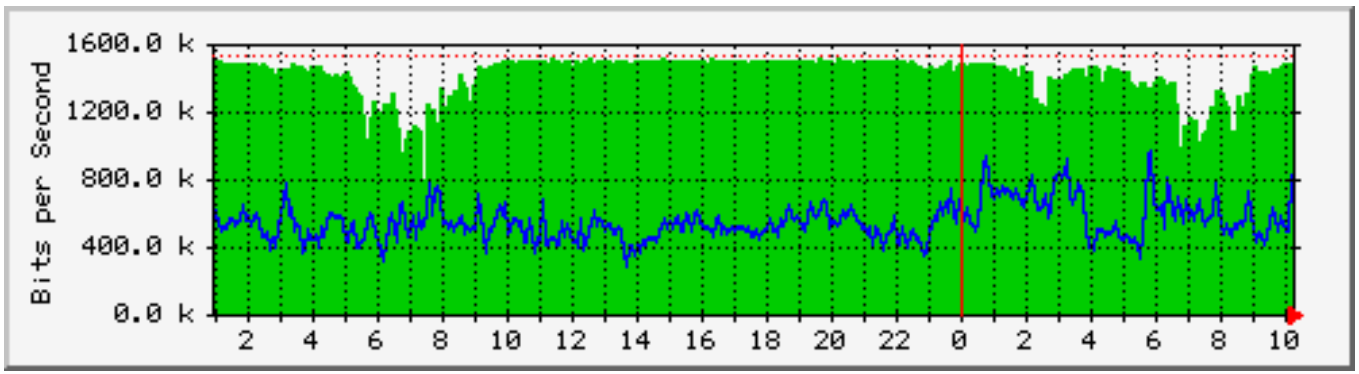
フロー計測

- SNMP によるインターフェイスカウンタ値による計測の限界
  - 総量は分かるが、それ以上の情報取得が困難
- フローベースの計測
  - 5 tupples (protocol, srcaddr, dstaddr, srcport, dstport), AS, etc
  - プロトコル: NetFlow、 sFlow、 IPFIX、 ...
- サンプリングによるデータ量削減も可能



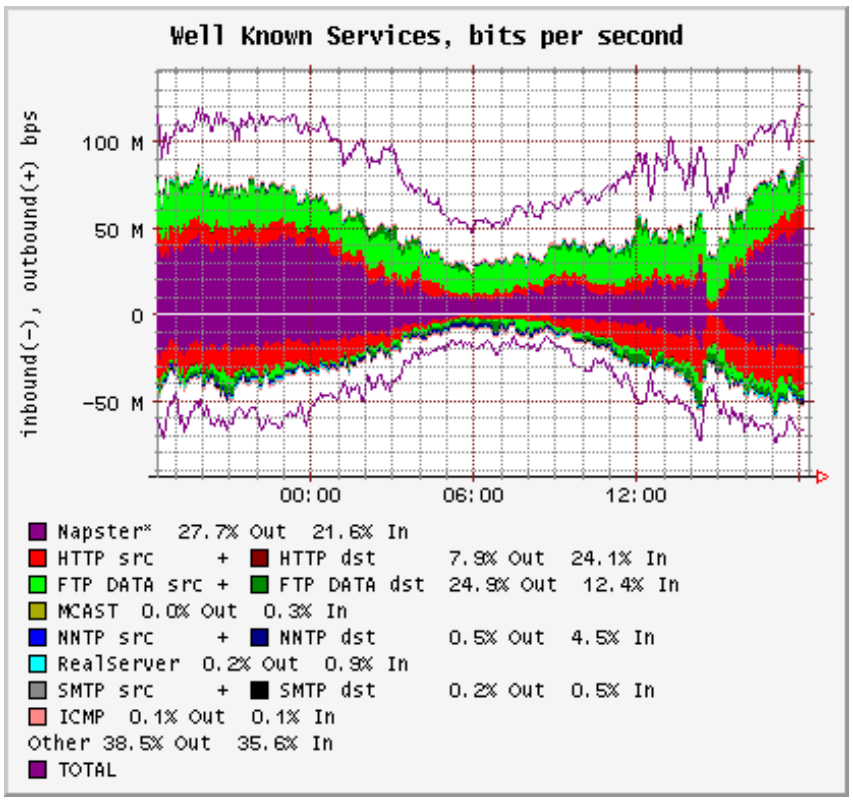
MRTG

- SNMP データをグラフ化するツール
- 古い時系列データを集約しデータ量を一定にする仕組み
  - daily, weekly, monthly, yearly
- inbound/outbound traffic
  - 他の時系列データにも利用可能



RRDtool

- RRDtool: MRTG の作者による後継ツール
  - 柔軟な設定が可能に
  - 任意の時系列データを扱えるよう工夫
- flowscan による RRDtool を使った NetFlow データのグラフ



from caida web site

まとめ

ネットワーク管理ツール

- もともと管理用ツールで計測ツールではない
- 多くの計測で利用されている
- 利用に際しては仕組みと制約を理解する必要がある

次回予定

第 2 回 インターネットのサイズを計る (10/6)

- ユーザ数、ホスト数を計る
- ウェブページ数を計る
- 精度 誤差 有効数字

参考文献

[1] Mark Crovella and Balachander Krishnamurthy. *Internet measurement: infrastructure, traffic, and applications*. Wiley, 2006.

[2] Antonio Nucci and Konstantina Papagiannaki. *Design, Measurement and Management of Large-Scale IP Networks: Bridging the Gap Between Theory and Practice* . Cambridge University Press, 2008.

[3] Pang-Ning Tan, Michael Steinbach and Vipin Kumar. *Introduction to Data Mining*. Addison Wesley, 2006.

[4] Raj Jain. *The art of computer systems performance analysis*. Wiley, 1991.

[5] 井上洋, 野澤昌弘. **例題で学ぶ統計的方法**. 創成社, 2010.

[6] 平岡和幸, 掘玄. **プログラミングのための確率統計**. オーム社, 2009.