„Backbone“ Measurement

Anja Feldmann
Deutsche Telekom Laboratories
TU-Berlin, Germany

The current Internet
The "future networks"

- Social phenomena:
  - Cyperspace
  - Changing/redefining communication
    - Human to human, human to computer, ….
  - Service:
    - Web, email, news, SMS, telephony, P2P
- A physical entity
- An infrastructure
- An object to be studied
- The foundation of someone's business

The "Internet": challenges

- Explosive growth and moving target
- Heterogeneity any which way you look
- New requirements: security, reliability, …
- Complex user behavior and traffic dynamics
- Designed as open, cooperating, static system
- Highly interacting systems
  - Temporal: between users, hosts and networks
  - Spatial: among different components
  - Vertical: across different networking layers

Scenarios to be addressed

Example 1:

- Situation: network connectivity fails
- Presumed action: call system administrator
- Effect: no phone call possible
- Why: telephone service via VoIP

Example 2:

- Situation: remote computer/router fails
- Presumed action: connect to device via control net.
- Effect: no connection
- Why: no reliable backup network

The "network": advantages

- Highly engineered structure
  - Well specified and documented
- Unique measurement capabilities
  - In theory unlimited access to data :-)
- Exploiting available data
  - Learn from the existing artifact
  - Use invariants not details
  - Consider emerging phenomena
  - Use network wide data
  - Take advantage of structural models

Challenges

The total is more than the sum of its pieces

- Specify and manage
  services rather than components

- Address the gap in understanding between
  individual pieces and the overall

Internet: a complex layered SWS

- Physical connectivity: Links
- Point-to-point connectivity: NIC, switches
  - distributed hardware, protocols - local management
- End-to-end connectivity: Routers
  - Forwarding, addressing, routing
  - Distributed hardware, protocols, software, management by Internet Service Providers (ISPs)
- Process-to-process connectivity: TCP, UDP
  - De-/multiplexing, reliability, congestion control, …
- Applications: Web, P2P, …
  - Users
  - Distributed, independent, autonomous, …

Tools

- Instrumentation and analysis
  - Integrate measurements into the design process
  - Collect data at a variety of different locations/levels
  - Find invariants and correlate various datasets
- Simulation
  - Build a mirror world for "what if" studies
  - Verify explanations
- Test-Labs
  - Incorporate variability
  - Provide an friendly/unfriendly environment

Measurement analysis

- Interesting when combining multiple datasets
  - Same data different locations
    - Root cause analysis of BGP events
    - Internet topology derivation
    - …
  - Different data sets
    - Intra-domain traffic matrix
    - Inter-domain traffic matrix
    - …

Measurement capabilities packet monitors

- Available data:
  - All protocol information
  - All content
- Possible analysis:
  - Application performance
  - User behavior (search enginee comparisons)
  - Application usage (P2P usage)
  - Abuse detection (intrusion detection system)
- Disadvantages:
  - Data flood
  - Data aggregation
  - Needle in a haystack
  - Only captures on network information (no device info)
  - Usually needs fixed installations

Measurement capabilities packet monitors (2.)

- Deployment scenarios:
  - Needs cooperation of the network operator
  - Limited number
  - Specialized hardware/software
  - Data collection / aggregation infrastructure
- Challenges
  - Data integrity
  - Incomplete data
  - User privacy & network security
  - Data correlation
  - Data privacy vs. data sharing
  - Data filtering
  - Data collection across network confederations

Measurement capabilities flow statistics

- Available data:
  - Summary information about traffic flows
- Possible analysis:
  - (Application performance)
  - User behavior
  - Application usage (P2P usage)
  - Abuse detection (intrusion detection system)
- Disadvantages:
  - Coarser grain information
  - Data flood
  - Data aggregation
  - Needle in a haystack
  - Only captures on network information (no device info)
  - Usually needs to be configured on network devices

Measurement capabilities flow statistics (2.)

- Deployment scenarios:
  - Needs cooperation of the network operator
  - Larger number
  - Specialized hardware/software
  - Data collection/aggregation infrastructure
- Challenges
  - Lack of detail
  - Data integrity
  - Incomplete Data
  - Data correlation
  - Data privacy vs. data sharing
  - Data collection across network confederations

Measurement capabilities
SNMP statistics

- Available data:
  - Summary information about devices
- Possible analysis:
  - (User behavior)
  - Anomaly detection (intrusion detection system)
- Disadvantages:
  - Very coarse grain information
  - (Data flood)
  - Data aggregation
  - Needle in a haystack
  - Only captures on network device information

Measurement capabilities
SNMP statistics (2.)

- Deployment scenarios:
  - Needs cooperation of the network operator
  - Data collection/aggregation infrastructure
- Challenges
  - Hard to see all
  - Data integrity
  - Level of detail – no connection to application
  - Data correlation
  - Data privacy vs. data sharing
  - Data collection across network confederations

Measurement capabilities routing information

- Available data:
  - Summary information about devices
- Possible analysis:
  - Network dynamics
  - Anomaly detection
  - Root cause analysis
- Disadvantages:
  - Very coarse grain information
  - (Data flood)
  - Data aggregation
  - Needle in a haystack
  - Only captures on network device information

Measurement capabilities routing information (2.)

- Deployment scenarios:
  - Collected anyhow by the network operator (currently basis for network management)
  - Data collection/aggregation infrastructure
- Challenges
  - Lack of detail
  - Data integrity
  - Data correlation
  - (Data privacy vs. data sharing)
  - (Data collection across network confederations)

Measurements: other

- Application protocol data
- Server related data
- Access networks
- Mobile networks
- Adhoc networks
- Sensor networks
- ...

Measurements:

- Challenges:
  - Scalability
    - how to reduce the amont of data to be analysed
  - Data flood
    - what to measure when the purpose is unclear
    - expect the unexpected
  - Validation
    - how to verify any inference
- Importance:
  - Enables network management
  - Enables debugging
  - Accountability
  - Verifies presumed assumptions

Internet dynamics: time scale

- Years: introduction of new protocols, e.g. IPv6
- Months: dimensioning a new circuit
- Weeks, days: different # of users responsible for weekly/daily cycle of traffic load
  different application mix
- Hours: variability of traffic volume
- Seconds: retransmissions
- Subseconds: round trip times

Internet control: time scale

- Years: IETF
- Months: network planning
- Weeks: network engineering
- Days: traffic engineering
- Hours: routing changes
- Seconds: TCP

Yet: user demand influences network performance but is also influenced by network performance

Clean slate design & measurements

Build in measurement capabilities at all levels

- Protocol
- Network device
- Network management
- Simplify data correlation
- Tackle scalability problem
- Enable data sharing
- Enable data validation
- Enable network debugging
- Enable network service debugging
- ….