Address and traffic dynamics in a large enterprise network

Thomas Karagiannis, Richard Mortier
Microsoft Research, Cambridge

Enterprise networks

- Followed the Internet's design principles
  - Radically different requirements/features

- Characteristics
  - Hundreds to several thousands of hosts
  - Single co-operative administrative domain
    - Hosts are partially trusted to share network information
  - Significant management and bandwidth costs
    - E.g., WAN optimizers
  - Do not want net-neutrality!
    - Can estimate the value of each application's traffic

Network Management today…

- Enterprise networks
  - High complexity, costly, error-prone
  - 80% of IT budgets just for maintenance

- Challenges:
  - Application complexity constantly grows
  - Limited analysis of network characteristics
    - E.g., traffic dynamics
    - Access restrictions and data sensitivity

Profiling enterprise networks

Traffic dynamics

- Can we profile enterprise traffic by sampling (a few) hosts?
  - Functional role (e.g., client vs. server)
  - "Heavy" hitters

Address dynamics

- What are the mobility characteristics of hosts?
  - Stability of Address-Name-Subnet mappings
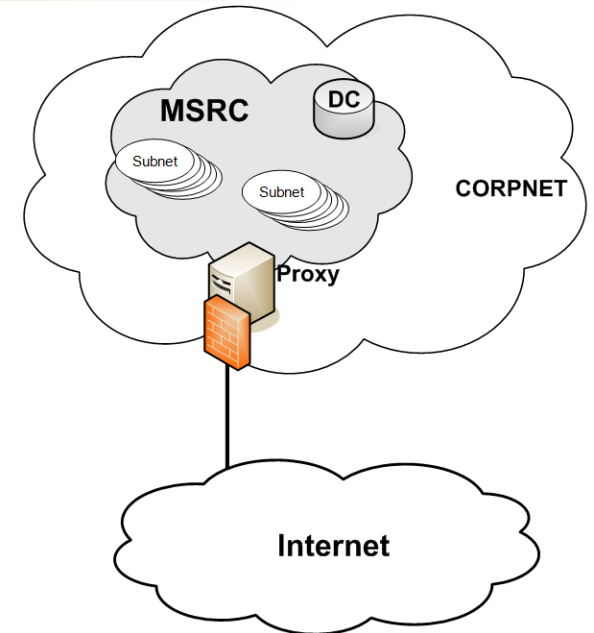 - Host mobility within the enterprise

Data traces

- Collected at the MSRC

- **Traffic**
  - August 2005, 3.5 weeks
  - 34K IPs (591 local), 13B packets, 12.5TB
- **Topology**
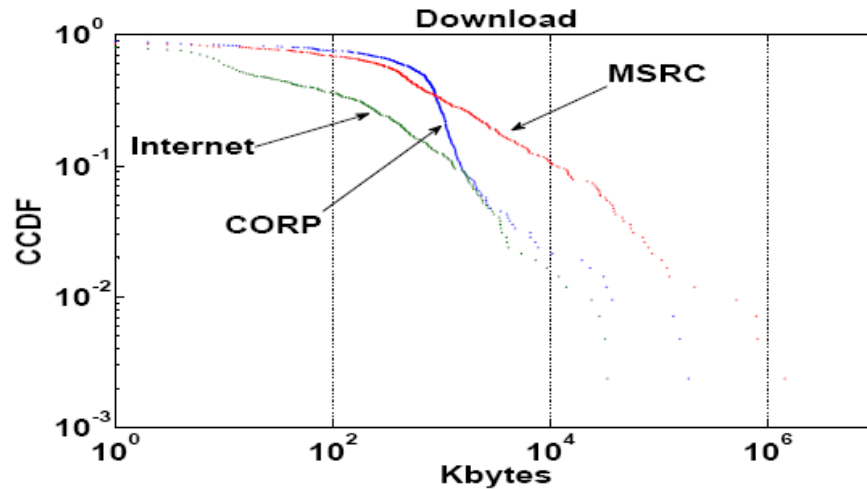 - OSPF, 3 years, stub and backbone
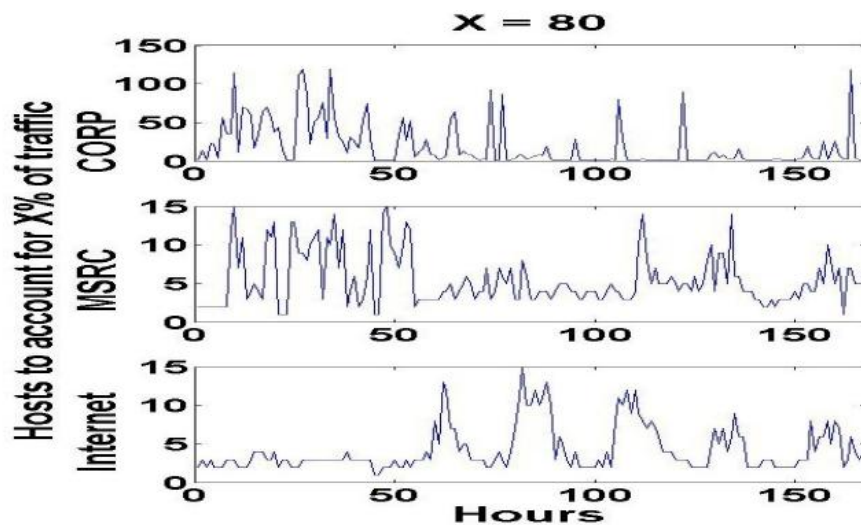
Profiling traffic dynamics

Expected to find:

- Functional role of hosts should be easy to detect from traffic contributions
  - Mostly client-server applications

- "Heavy" hosts should be stable over time
  - A small set of servers (mostly in DC)

Profiling traffic dynamics

- ◻ CCDFs of hourly averages
- ◻ Heavy-tailed distributions
  - ◻ Small-set of hosts dominates traffic

- ◻ Temporal & spatial variability
  - ◻ Heavy set varies over time
  - ◻ Unable to determine host functional role

Profiling traffic dynamics

Expected to find:

- Functional role of hosts should be easy to detect from traffic contributions
  - Mostly client-server applications

- "Heavy" hosts should be stable over time
  - A small set of servers (mostly in DC)

Data analysis:

- Traffic contributions cannot distinguish client vs. server hosts

- Significant variability!

Implication:

- Sampling hosts does not help!
- Connectivity appears to be a better metric (details in the paper & tech report)

Profiling address dynamics

How often should IPs be considered as unique identifiers?

- IP addresses map to several hosts and vice versa
  - E.g., DHCP, multi-homing, multi-machine services
- Examine the stability of address mappings
  - DNS packets & router configuration files
- Three types of mappings:
  - Name-address : Unique names per IP
  - Address-name : Unique IPs per name
  - Subnet-name   : Unique subnets per name

Profiling address dynamics
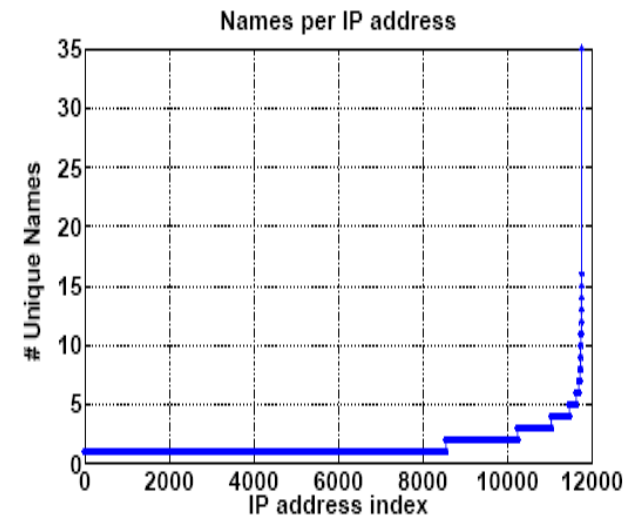
Findings:

- ## Name-address:
  - 73% of addresses map to a unique name
  - Addresses can map to 10s of names

- ## Address-name:
  - 63% of the names map to a single address
 - Multi-homing and clusters the main factor for multiple IPs per name

- ## Subnet-name:
  - 63% of the names map to a single subnet
  - 30% of the names map to two subnets
  - 4% due to travelling!

Names per IP address

(plot: # Unique Names vs IP address index)

Host mobility

- **Of general interest**
  - DTN settings   (e.g., Infocom 2006, Mobicom 2007, Infocom 2008)
  - Understanding human mobility   (e.g., Barabasi-Nature Jun08)

- **Examine "host" trips within the enterprise**
  - Extract subnet-name mappings
  - 9,269 names in 110 cities across 63 countries
  - Analyze location changes (trips) across enterprise sites
    - Residence time, return time

Host mobility
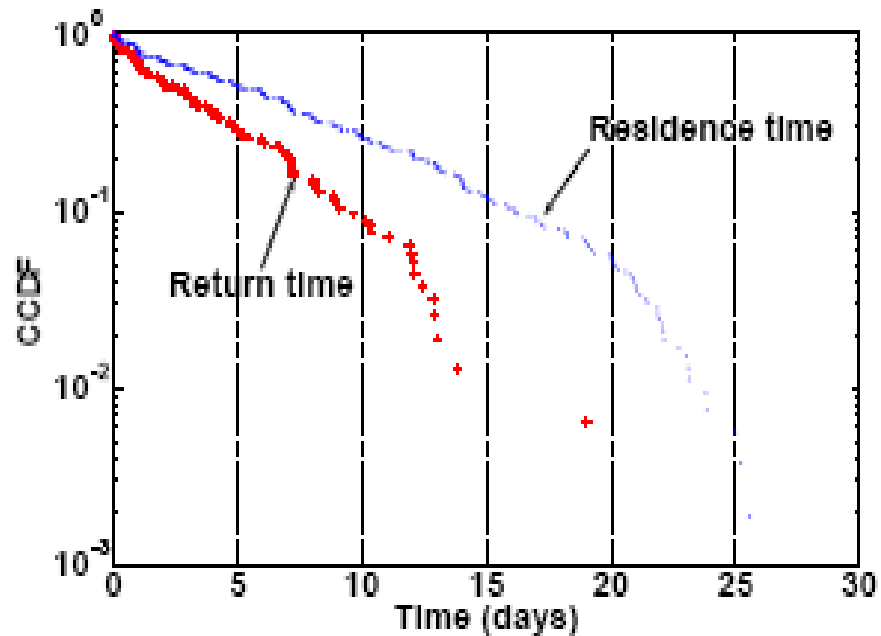
- Exponential distributions
- 38% of residence time is < 3 days
- Means
  - Residence time : 5.5 days
  - Return time : 3.8 days

Concluding remarks

Two perspectives of enterprise network dynamics

- **Traffic**
  - Sampling a few hosts is not straightforward
  - Engage hosts in network management (SIGCOMM 08)

- **Address**
  - Analyzing traces requires more than just packets
    - Only 2/3 of mappings are unique

Thank you!