Welcome to my seminar

Let us:

• Work hard (efficiently)
• Have fun

Keep checking the class website

Register in the mailing list: 26-faloutsos
Modeling the Topology in Networks

- Measurement, modeling and generation of topologies
- Span different technologies: Internet, cellular, ad hoc and sensor networks
The Issues

- Questions we will try to answer:
  - How can I model the Internet topology?
  - What is a realistic topology to conduct simulations on?
  - How can I generate one?
  - How can I model the topology of a mobile network?
  - Are there any invariants in a mobile topology?
Some Unexpected Things

- We do not have a complete Internet topology
- We cannot model the business relationships
  - Customer - Provider ISPs
- We do not know how to generate a realistic Internet topology
- None uses a realistic mobility model
- We do not have tools to model mobility
Evaluation and Deadlines

- In class presentations (10%)
- Quiz (10%) End of November
- Participation (10%)
- A serious project (70%): (in pairs)
  - Project Proposal 10%, Delivered by THU, Oct 28 before or at the class.
  - Completed Project 60%, Delivered by MON, Dec 13.
Data Mining the Internet:
What we know, what we don’t and how we can learn more

Michalis Faloutsos, UC Riverside

Christos Faloutsos, CMU

Lectures based on a tutorial given at SIGCOMM 2002
Big Picture: Modeling the Internet

- Measure and model each component
  - Identify simple properties and patterns
- Model and simulate their interactions

Topology

Protocols

Traffic

Routing, Congestion Control
The Goal of Internet Modeling

- Find simple fundamental properties
- Understand why they appear and their effects

A real Internet instance

Power-law: Frequency of degree vs. degree
Claim: We Need The Right Tools

“This is just not effective… We need to get some chains”

The Far Side -- G. Larson
What This Lecture Is All About

What we do and don’t know about the Internet:

- Model the topology
- Analyze traffic and end-to-end behavior
- Examine effect of protocols traffic and topology
Oversimplified Tutorial Overview

We observe a mental switch in modeling

- Distributions: uniform ⌁ skewed, power-laws
- Processes: memoriless Poisson ⌁ long memory
- Behavior: smooth ⌁ bursty
What Do We Know

General background and basic concepts

Section I: Topology

Section II: Traffic and performance

Section III: The effect of protocols

Conclusions
Motivation

- We don’t know how to model the Internet
- We need realistic assumptions for simulations

Questions of interest
- Which topology should I use for my simulations?
- How should I generate background traffic?
- How can I recreate realistic packet loss?
- How can I detect abnormalities?
General Background

- Power-laws
- Fractals and Self-similarity
- Long Range Dependence
- Burstiness
What Is a Power-law?

Power-law is a formula:

\[ y = ax^c \]

where \( x, y \) variables and \( a, c \) constants

A power-law is a line in log-log scale:

\[ \log y = \log a + c \log x \]
Self-Similarity and Fractals

- **Objects of infinite detail**
- **Self-similar:**
  - A part is identical to the whole
- **Scale-free:**
  - Statistical properties are independent of scale of observation
- **Infinite detail:**
  - The closer I look, the more I see
- **Power-laws are intimately related to fractals**
Example: A Fractal Line

- Koch’s snowflake (dimension = 1.28)
- Repeat for ever:
  - Introduce a bump at every straight line
- Each side is identical to the initial line
- Infinite detail, infinite length
- More detail in part B
Long Range Dependence

- LRD captures the “memory” of the behavior
- It is quantified by a single scalar number
  - Hurst power-law exponent
- LRD appears in many aspects of networks
  - Traffic load, arrival times, delays, packet loss

Issues:
- How can we estimate the LRD
- How can we use LRD
The Definition of LRD

Given a signal $X_t$, the autocorrelation function $r(k)$ is

$$r(k) = E [(X_t - \bar{X})(X_{t+k} - \bar{X})] / \sigma^2$$

If $r(k)$ follows a power-law:

$$r(k) \sim k^{-\beta}$$

we say that the signal exhibits LRD.
The Intuition Behind LRD

- Capturing the “dependency” of the current measurement to previous values

- White Noise

- Brownian Noise

- Long Range Dependence
Fourier Transform

\[ x(t) = a_0 + \sum_{k=1}^{\infty} \left( a_k \cos(2\pi f_0 t) + b_k \sin(2\pi f_0 t) \right) \]

- Analyze a signal in the frequency domain
- Approximate a signal \( x(t) \) by sum of periodic signals
- Intuitively: think of the “equalizer in a stereo”
  - Decompose signal into frequencies

\( f_0 \) base frequency
\( a_k, b_k \) amplitude
The Time and Frequency Domains

A sinus wave corresponds to one frequency

Time

Frequency

★ A sinus wave corresponds to one frequency
Example: A Fourier Transform

- A signal with **four** different frequency components at four different time intervals...
Example: The Fourier Transform

Each peak corresponds to a frequency of a periodic component...
Part A.I: Topology

- General background and basic concepts
- Section I: Topology
- Section II: Traffic and performance
- Section III: The effect of protocols
- Conclusions
Motivation

- What is the topology I should use in my simulations?
- How can I generate a realistic topology?
- Can I define a hierarchy?
Why Is Topology Important?

“You can’t resolve the traffic jam problem of a city without looking at the street layout.”

- To conduct realistic simulations
- To interpret measured data
- To design and finetune protocols
Overview of Topology

- The topology is described by power-laws
  - Forget uniform distributions
- Growth of the network is super-linear
- It is compact and becomes denser with time
- The Internet looks like a jellyfish!
Part A.I. Topology: Roadmap

- Previous Models
- Power-laws of the Internet topology
- Time evolution
- Generating realistic topologies
- An Intuitive model: jellyfish
- Powerlaws in other communication networks
Real Internet Graphs

- Autonomous System (AS):
  - Individually administered network

- AS Level Topology: Each node is an AS

- Router level: each node is a router

- We focus on AS level graphs:
  - Routeviews – NLANR: archive
  - More complete data: using multiple data repositories
Previous Topological Models

- Models assume uniform distributions
  - All nodes have approximately the average degree
- Nodes uniformly distributed on a plane with edge probability decreasing with distance [Waxman]
- Hierarchical structure of simple graphs [Doar] [Zegura et al.]
The AS Topology exhibits Power-laws

- I. Degree of nodes vs. rank
- II. Frequency of degree (skip)
- III. Eigenvalues of adj. matrix
- IV. Pairs of nodes within $h$ hops

Accuracy: correlation coeff. $> 0.97$

Recently: power-laws for
- Distances
- Spanning Tree sizes
- Scaling of multicast trees
I. Power-law: rank exponent $R$

- The plot is a line in log-log scale
- Exponent = slope
  \[ R = -0.74 \]
- Rank: nodes in decreasing degree order
- Dec’98

[Faloutsos, Faloutsos and Faloutsos SIGCOMM’99]
I. Estimations Using the Rank Exponent $R$

Lemma:
Given the nodes $N$, and an estimate for the rank exponent $R$, we predict the edges $E$:

$$\square = \frac{1}{2(R + 1)} \cdot (1 \boxplus \frac{1}{N^{R+1}}) \cdot N$$
II. Powerlaw: Degree Exponent D

- Degree distribution of nodes: CCDF
- It holds even for the more complete graph: 99%
III. Eigenvalues

Let $A$ be the adjacency matrix of graph

The eigenvalue $\lambda$ is a real number s.t.:
- $A\mathbf{v} = \lambda \mathbf{v}$, where $\mathbf{v}$ some vector

Eigenvalues are strongly related to topological properties

More details in Part B
III. Power-law: Eigen Exponent $E$

Find the eigenvalues of the adjacency matrix

Eigenvalues in decreasing order (first 100)

Exponent = slope

$E = -0.48$

May 2001
Surprising Result!

- Exponent E is half of exponent D
- Theorem: Given a graph with relatively large degrees $d_i$ then with high probability:
  - Eigenvalue $\lambda_i = \frac{1}{d_i}$, where $i$ rank of decreasing order
- Thus, if we compare the slope of the plot the eigenvalues and the degrees:
  - $\log \lambda_i = 0.5 \log d_i$

[Fabrikant, Koutsoupias, Papadimitriou in STOC’01]
[Mihail Papadimitriou Random 02]
Powerlaws are here to stay
Degree distribution slope is invariant
Network becomes denser
The rich get richer phenomenon
The Number of ASes in Time

The number of AS doubled in two years
Growth slows down!
Degree Distribution Did Not Change!

❄️ Slope is practically constant for over 3 years
The Topology Becomes Denser!

- 6 hops reach approximately 98% of the network!
- Denser: 6 hops reach more nodes

Recall six degrees of separation
The Rich Get Richer

- The increase of the degree versus the initial degree
- New connections prefer “highly connected nodes”
The Origin of Powerlaws

- **Preferential attachment of nodes** [Barabasi Rekka]
- **Self Organizing Criticality** [Bak]:
  - The “steady state” of complex systems
- **Highly Optimized Tolerance** [Doyle Carlsson]:
  - Considering an element of design
- **Heuristically Optimized Tolerance** [Fabrikant et al]:
  - Optimizing with local constraints
Powerlaw Graph Generators

 предпочитательное присоединение, накопительный рост:
• Добавление новых узлов, предпочитающих ребра к узлам с высоким степенью
• Линейная предпочтительность: \( p_i = \frac{d_i}{\sum_k d_k} \) \cite{Barabasi et al}  
• Вариации к линейной предпочтительности \cite{Bu Towsley}

Природа мощної
• Установите каждый узел с степенью от желаемой степени распределения
• Соедините узлы их неподключенными ребрами
Heuristically Optimized Tolerance:

- Distribute nodes in Euclidean plane
- Add edges to minimize: $D_i + a C_i$
  - $D_i$: Path length from everybody else
  - $C_i$: Cost of building edge ($f()$ of Euclidean distance)
- Intuition: optimize hop-distance subject to local constraints
- Initial distribution of nodes does not affect result
- [Fabrikant, Koutsoupias, Papadimiitrou in STOC’01]
An Intuitive Model for the Internet

Can I develop a simple model of the AS Internet topology that I can draw by hand?
Can I identify a sense of hierarchy in the network?

Focus: Autonomous Systems topology
The Internet Topology as a Jellyfish

- **Core**: High-degree nodes form a clique
- **Each Layer**: adjacent nodes of previous layer
- **Importance decreases as we move away from core**
- **1-degree nodes hanging**

[Tauro et al. Global Internet 2001]
Developing An Intuitive Model

We need an anchor and a compass

Anchor:
- We need a starting point in the network

Compass:
- We want to classify nodes according to importance
Defining the Importance of a Node

- **Metrics for topologically importance**
  - **Degree:** number of adjacent nodes
  - **Eccentricity:** the maximum distance of a node to any other node
  - **Effective:** distance to 90%

- **Significance:** Significant nodes are near:
  1. many nodes
  2. significant nodes
Significance of a Node

- The significance of a node is the sum of the significance of its neighbors.
- The iterative procedure converges:
  - At each round, total significance is normalized to 1
- Surprise! This is equivalent to:
  - The eigenvector of the max eigenvalue of the adjacency matrix [Kleinberg]
- Relative Significance: Normalize to sum up to N
  - Relative Significance = 1, fair share of significance
Observation 1: Significant Nodes are in the “Center”

Significance vs. Eccentricity

- They are correlated
- Intuitively, significant nodes are in the middle of the network
Observation 2: One-Degree Nodes Are Scattered Everywhere

The distribution of 1-degree nodes follows a power-law.

- Important node connect with unimportant nodes

Order of decreasing degree
Observation 3: The Internet “Premise”: One Robust Connected Network

- Robust to random, sensitive to focused failures
- The network stays as one connected component
Observation 4: The Number of Alternate Paths Between Two Nodes

Number of paths

The Failure of the Donut Model

Path Length

- All alternate paths go through the same direction
- No shortcuts or loop-grounds
Defining a Hierarchy Recursively

Define the core:
- Maximal clique of highest degree node

Define the Layers:
- All nodes adjacent to previous layer

Define the Shells:
- A layer without its one-degree nodes
The Hierarchy: The Model Respects the Node Importance

- The importance of nodes decreases as we move away from the core.
- The effective eccentricity decreases by one in each layer (see paper for details).
The Evolution of the Jellyfish

- The jellyfish lives on!
- Percentage of node in each class in time
- The structure of the jellyfish has not changed much in the last three years
Why Is The Jellyfish a Good Model?

It’s cute, in addition…
The Jellyfish Captures Many Properties

- The network is compact:
  - 99% of pairs of nodes are within 6 hops
- There exists a highly connected center
  - Clique of high degree nodes
- There exists a loose hierarchy:
  - Nodes far from the center are less important
- One-degree nodes are scattered everywhere
- The network has the tendency to be one large connected component
And It Looks Like A Jellyfish...

- Independent Observation
- Router Level Topology
- Produced by CAIDA
Powerlaws In Other Networks

- Powerlaws appear in several other settings
- Graph of www pages:
- Peer-to-peer networks:
The WWW Page Topology

Distribution of in-degree and out-degree of a page

Diameter of the web: 19 clicks

[Albert, Barabasi, Huberman, Adamic, Lawrence, Giles, Rajagopalan et al]
The Size of Web Sites

CCDF of the web sites according to size

[Huberman Adamic]
The Peer-to-Peer Topology

- CCDF distribution: Frequency versus degree
- Number of adjacent peers follows a power-law

[Jovanovic+]

(a) Gnutella snapshot from Dec. 28, 2000 ($\rho = 0.94$)
Summary of Topology

- The topology is described by power-laws
  - Forget uniform distributions
- Growth is slowing down (?)
- It is compact and becomes denser with time
- The Internet looks like a jellyfish!
What We Still Don’t Know

❖ Need comprehensive set of metrics
  • Validate generators
  • Assess realism of graphs
❖ How topology affects
  • Simulations
  • Traffic
  • End-to-end Performance
❖ How to use new understanding for protocol design
# Table Overview

<table>
<thead>
<tr>
<th>Know</th>
<th>Don’t Know</th>
<th>How to learn more</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Topology</strong></td>
<td>Powerlaws, jellyfish</td>
<td>Growth pattern</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Compare, effect of topology</td>
</tr>
<tr>
<td><strong>Link</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>End-2-end</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Traffic Matrix</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
End of Topology Section
The World Wide Web is a Bow-Tie

- Captures several properties [Tomkins et al]
- The components are of comparable size
The Accuracy-Intuition Space Of Models

More tools…
- Self-similarity
- Power-laws
- Wavelets
- Eigenvalues

...less intuition
- Something a human can picture

Is it a real conflict?
Why Do We Need an Intuitive Model?

- Human mind is simple
- Visualizable: creates a mental picture
- Memorable: captures the main properties
- Maximizes \textit{information/effort ratio}
- Makes you think
Part A. What We Know

- General background and basic concepts
- Section I: Topology
- Section II: Traffic and performance
- Section III: The effect of protocols
- Conclusions
Questions of Interest

- How should I generate background traffic?
- How can I recreate realistic packet loss?
- How can I model end-to-end delay?
- How can I detect abnormalities?
- What is the flow matrix like?
Significance

- Need realistic assumptions for traffic
- Model the performance an application sees
- Fine-tune end-to-end protocols
  - TCP, RTP, playback buffer, real-time applications
Overview Of This Section

- Long Range Dependence describes many dynamic phenomena
  - Forget memoriless and Poisson processes
  - Link traffic is LRD
  - Packet loss and round-trip delay exhibit LRD

- Estimating LRD is tricky:
  - Common Pitfalls
  - Step Towards a systematic approach
Previous Models For Traffic

- **Fundamental assumption: Memoriless**
  - Only your current state affects your next state
- **Poisson arrivals**
- **Systems modeled by Markov processes**
- **Advantage: easier to study analytically**
- **Problem: nature is not like this**
Statistical Behavior of Link Traffic

- Aggregate behavior: Poisson becomes smooth
- Measured traffic is always bursty
  - Similar properties

© M. and C. Faloutsos
The Link Load is Self-Similar

Normalized Variance

Scale

💥 Intuition: it has large variance in many scales of observation  [Lelland et al 93, 94]
A Generator of Self-Similar Traffic

Many ON-OFF sources
- Times are heavy-tailed distributed
  - Non-zero probability of long intervals

Yields:
- Long Range Dependence

[Lelland+, Paxson+, Willinger+, Taqqu+, Riedi+]
Why Is Traffic Self-Similar?

Nature works non uniformly
- Applications/users are bursty
- File sizes and requests are skewed [Crovela et al]
- Effect of topology and TCP [Feldman+]
- Not all flows are equal [Sarvotham Riedi et al]
  - A few flows dominate a link (“Alpha flows”)
Real Web traces

Distributions are skewed  [Crovela et al]
Link Traffic and Dominant Flows

- The dominant flows are responsible for bursts
- The other flows exhibit long range dependence

Riedi Baraniuk+, INCITE project, Rice U.

Overall traffic = 1 Strongest connection + Residual traffic
Part A.II. Traffic: Roadmap

- Background
- Link traffic
- End-to-end performance
- Traffic Matrix
**End-to-end Performance Metrics**

- How the application sees the network

- **End-to-end (e2e) refers to**
  - One way
  - Round trip

- **Metrics**
  - Packet loss
  - Delay: one way or Round-Trip-Time (RTT)
  - Delay jitter: inter-arrival time
Significance of End-to-End Metrics

• **Round-trip-time (RTT):**
  - TCP estimates RTT to set time-out for packet retransmission

• **Delay jitter:**
  - Multimedia (RTP) uses jitter to tune playback-buffer

• **Packet loss:**
  - Direct effect on TCP sending rate
  - Define error recovery techniques in multimedia
Identifying Long Range Dependence

- Quantified by Hurst powerlaw exponent: \( H \)
  - When \( 0.5 < H < 1 \), we have LRD

- There are several methods to “estimate” it

- BUT, estimating LRD is not straightforward!
  - Many estimators, which often conflict
  - No ultimate generator for calibration
  - No systematic approach
LRD: Coping in Unknown Territory

1. How accurate are the LRD generators?
2. How accurate are the estimators?
3. How conclusive are the estimators?
4. How can I look for LRD in real data?
   - Missing data, “noise”, indecision
Our Approach To Understand LRD

- Develop a library of behaviors of known data
  - Compare with results of known behavior
- Three series of tests for the estimators:
  1. Evaluating the accuracy of the estimators
     - Synthetic Fractional Gaussian Noise (FGN)
  2. Deceiving the estimators with non-LRD data
     - Periodicity, Noise, Trend
  3. Applying the estimators on real data
     - Characterizing delay and packet loss

[Karagiannis+ GI 02]
1. Accuracy: Synthetic LRD Data

- Large difference in values!
- The Whittle and Periodogram are most accurate
- The rest can be significantly inaccurate!

Fractional Gaussian Noise Paxson’s Generator
2. Robustness: Deceiving the Estimators

- **Periodicity fools many estimators**
  - The Whittle, the Periodogram, the R/S and the Abry-Veitch falsely report LRD in series constructed by cosine functions and noise.

- **White noise affects the accuracy**

- **Trend also deceives estimators**
  - Whittle and Periodogram falsely report LRD
3. Analyzing Real Data

• Every 50msec send packet 400b
• From:
  • UCR
  • Cable modem, commercial ISP
• To:
  • Australia, Un. Of LaTrobe
  • CMU
  • Greece, Aristotelian Un. Of Thessaloniki
• Packet Loss and Round Trip Time (RTT)
The REALITI Measurement Tool

Enable us to control

- Sending rate, packet size and type
- Four time-stamps (at server too)

By M. Samidi, UCR
R. Venkataswaran, Tata Consulting Services
UCR-Australia: Loss/sec

1 day

1 hour
In More Detail...

Zoom in more

Zoom in even more
UCR-Australia: Loss is LRD

<table>
<thead>
<tr>
<th></th>
<th>R/S</th>
<th>Agg. Variance</th>
<th>Residuals</th>
<th>Periogram</th>
<th>Whittle</th>
<th>AV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.86</td>
<td>0.89</td>
<td>0.89</td>
<td>0.69</td>
<td>0.66</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>(99%)</td>
<td>(97%)</td>
<td>(97%)</td>
<td>(0.65-0.66)</td>
<td>(0.75-0.76)</td>
<td></td>
</tr>
</tbody>
</table>

All estimators detect LRD: 0.5 < H < 1
But not the same value of Hurst: 0.66 – 0.89
This is as close as it gets
Analyzing Delay: RTT

- Measured round trip time: UCR-CMU
- Initial signal does not exhibit LRD
- What do we do next?
A Closer Look at RTT

🌟 Is there a pattern?
The Measured Data Is Periodic

- There is periodicity throughout the dataset

Short-Time Fourier Transform Frequency Spectrum
The Periodicity Hides the LRD!

**Variance Method**

Measured
(periodicity)

Estimated: 0.55 and 0.68

**RS-plot Method**

Without
periodicity
Practical Lessons

Be cautious when you deal LRD

- LRD estimation and method must be reported
- LRD may exist even if all estimators do not agree
- There is no “consistent-winner” estimator
  - We need to consult all of them
  - If all find Hurst, then most likely LRD
- Estimation can be thrown off by
  - Noise, trend and periodicity
- Look at the plot
Towards a Systematic Approach

Goal: characterize the signal
Pre-process: clean data
Decompose signal
Characterize each component separately
Use all estimators
Compare results with those of known signals
The SELFYS Tool

- A platform for development and reference
  - Java-based
  - Modular
  - Free [developed by Thomas Karagiannis, UCR]

- Given a trace
  - Cleans data
  - Wavelet and Fourier analysis
  - Runs all LRD estimators

http://www.cs.ucr.edu/~michalis/PROJECTS/NMS/NMS.html
Part A.II. Traffic: Roadmap

- Background
- Link traffic
- End-to-end performance
- Traffic Matrix
Question of Interest

- Where are the sources and the receivers?
- Who communicates with whom?
- Can I identify clusters of users?
- How are the multicast members distributed?
Why Can’t We Measure Traffic Matrix?

- It is an open ended question
- It is affected by many parameters
- It is application dependent
- Caching obscures things more
Location of Web-Server Clients

- **A success story** [Krishnamurthy Wang 00]
- **Question**: Where are my clients?
- **Motivation**:
  - Install caches appropriately
  - Identify customer base and target advertising
- **Complication**:
  - Using first 3 bytes of IP addresses does not work!
Network-Aware Clustering

Cluster requests using routing data
- Get BGP routing tables
- Look up client IP address
- Find longest match between address and database
- Cluster together clients with same match

[ Krishnamurthy Wang 00]
Routing Database
101.23.54.9 /8
101.112.1.1 /16
101.112.21.16 /28
101.112.21.31
101.112.21.17
Experiments

- The method works well
- Experiments on wide range of Web servers
- Results
  - > 99% clients can be grouped into clusters
  - ~ 90% sampled clusters passed the validation tests
# The Clustering Data

<table>
<thead>
<tr>
<th>Log</th>
<th>Description</th>
<th>Date</th>
<th>Duration (days)</th>
<th># requests</th>
<th># clients</th>
<th># clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apache</td>
<td>Apache site</td>
<td>10/1/99-11/18/99</td>
<td>49</td>
<td>3,461,361</td>
<td>51,536</td>
<td>35,563</td>
</tr>
<tr>
<td>Ew3</td>
<td>AT&amp;T content hosting site</td>
<td>7/1/99-7/31/99</td>
<td>31</td>
<td>1,199,276</td>
<td>21,519</td>
<td>7,754</td>
</tr>
<tr>
<td>Nagano</td>
<td>1998 Winter Olympic Game</td>
<td>2/13/98</td>
<td>1</td>
<td>11,665,713</td>
<td>59,582</td>
<td>9,853</td>
</tr>
<tr>
<td>Sun</td>
<td>Sun Microsystems site</td>
<td>9/30/97-10/9/97</td>
<td>9</td>
<td>13,871,352</td>
<td>219,528</td>
<td>33,468</td>
</tr>
</tbody>
</table>

**Millions of requests, tens of thousands of clients, 1:2 to 1:6 clustering**

© M. and C. Faloutsos  
Advanced Networks
Distributions of Web Clients Are Skewed!

Most clusters have a small number of hosts; 95% of clusters contain 100 or less hosts.

Most clusters issue a small number of requests; 90% of clusters issued 1000 or less requests.
The Inter-Domain Traffic Matrix

- Inter-AS communication [Fang Peterson Globecom 00]
- Collected data Jan 1999:
  - vBNS: educational institutions
  - MCI: Mae-West
Distribution of Data Flow

9% of AS pairs is responsible for 86.7% of packets
Experience Suggests Skewness

- Skewed distributions of senders and destinations
  - In space and in time
- Skewed distributions of traffic intensity
- Correlations: Groups of common interest
  - I.e. gnutella destinations are probably sources of quake video games and likely to be active in the night
Some Open Questions

Traffic Matrix:
- Distribution of traffic among sources and receivers
- Models to generate realistic traffic matrices
- Temporal and spatial properties of traffic

Multicast members:
- Location of members
- Join and leave behavior
- Is multicast state aggregatable?
# Table Overview

<table>
<thead>
<tr>
<th></th>
<th>Know</th>
<th>Don’t Know</th>
<th>How to learn more</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Topology</strong></td>
<td>Powerlaws, jellyfish</td>
<td>Growth pattern, Compare graphs</td>
<td></td>
</tr>
<tr>
<td><strong>Link</strong></td>
<td>LRD, ON/OFF sources</td>
<td>Effect of topology and protocols</td>
<td></td>
</tr>
<tr>
<td><strong>End-2-end</strong></td>
<td>LRD loss and RTT</td>
<td>Troubleshoot, cluster and predict</td>
<td></td>
</tr>
<tr>
<td><strong>Traffic Matrix</strong></td>
<td>Skewness of location</td>
<td>Comprehensive model, troubleshoot</td>
<td></td>
</tr>
</tbody>
</table>
Part A: What We Know

- General background and basic concepts
- Section I: Topology
- Section II: Traffic and performance
- Section III: The effect of protocols
- Conclusions
Motivation

We want to know how protocols affect

- Traffic
- Performance
- Stability

Dominant protocols:

- BGP: routing protocol (our focus)
- TCP: end-to-end flow control
Part A. III: The Effect of Protocols

- Some background
- BGP and topology
- BGP and routing
- BGP and routing robustness
  - The attack of the worms
- BGP and scalability
Questions of Interest

- How does BGP affect routing?
- Will BGP scale?
- How does the BGP table grow?
- How robust is BGP?
- How does errors propagate?
What Is BGP?

- **Border Gateway Protocol, BGP version 4**
- **The de-facto inter-domain routing protocol**
  - Uses TCP to communicate
  - Distance Vector style: neighbor exchanges

**BGP was developed to achieve:**
- Flexible policy implementation
- Scalability via route aggregation given CIDR
BGP Modeling Brings New Issues

- Business policy is introduced in routing
- Manual and configurations errors
- Routing: paths are “inflated” due to policy
- Topology is modeled by a directed graph
  - Provider → Customer
- Convergence and stability become an issue

BGP is a hot research topic
How A BGP Network Looks Like

- Each AS has designated BGP routers
- BGP routers of an AS communicate internally with another protocol (IGP)

Recall: Autonomous System = Independent network
Routing Updates

- BGP routers advertise to each other:
  - IP prefixes and the related path

- Three steps:
  - Receive and filter an advertisement
  - Change your table, if necessary
  - Forward change selectively

- If a neighbor does not respond:
  - Invalidate all related paths (remember this)
IP Addresses and Prefixes

IPv4 addresses have 32 bits: 4 octets of bits
• 128.32.101.5 is an IP address (32 bits)

An IP prefix is a group of IP addresses
• 128.32.0.0/16 is a prefix of the first 16 bits
  • = 128.32.0.0 – 128.32.255.255 (2^16 addresses)
• 128.32.4.0/24 is a longer prefix 24 bits

Routing: find the longest match:
• IP prefix in table that matches most bits of the address
What Does a Routing Table Look Like?

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Origin AS</th>
<th>AS Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>128.32.0.0/16</td>
<td>123</td>
<td>14 56 123</td>
</tr>
<tr>
<td></td>
<td>123</td>
<td>34 101 203 123</td>
</tr>
<tr>
<td>128.32.101.0/24</td>
<td>15</td>
<td>50 50 15</td>
</tr>
</tbody>
</table>

- Origin AS “owns” the address
- Routing tables can have peculiarities and errors
Part A. III: The Effect of Protocols

- Some background
- BGP and topology
- BGP and routing
- BGP and routing robustness
  - The attack of the worms
- BGP and scalability
Customer – Provider: customer pays and is always right

Peer to Peer: Exchange traffic only between their customers

Sibling-Sibling: Exchange traffic at will
The BGP Logical Graph

A directed jellyfish! [Ge et al ITCom 01]
- Peers within a layer
- Higher layer are providers of lower layer
- More layers than the undirected jellyfish
Determining The Logical Graph

- The business relationships are critical
- How can I find the relationships?
  1. Infer relationships from routing tables
  2. IRR database: manually maintained – error prone
Two Inference Algorithms

1. Inference algorithm [Gao00]
   • Exploit the up-down path property
   • in a path, assume highest degree node as peak

2. Inference using multiple observation points [Subramanian et al 02]
   • Use multiple points of observation to improve results

Accuracy:
   • Fairly good but needs further investigation
Part A. III: The Effect of Protocols

- Some background
- BGP and topology
- BGP and routing
- BGP and routing robustness
  - The attack of the worms
- BGP and scalability
How BGP Policy Restricts Routing

Routing rules:
- Provider accept everything
- Peer only if it is for its customers

Path Properties:
- Up then down
- No up-down-up, at most 1 peer-peer steps
Policy Increases The Path Length

25-20% of paths are inflated by at least one hop

- Compared to the path on the undirected graph

[Siganos et al 02]
It’s Money That Matters…

- Sender pays up path
- Receiver pays down path
- Based on static and statistical agreements

Sender               Receiver
Policies And Routing Asymmetry

- A Provider exports traffic as soon as possible
- But a Provider will carry traffic for its customer
- Did anyone say traffic is asymmetric?
BGP Path-Length Asymmetry

- Consider only AS path-length
- Asymmetry: 46% of pairs differ by at least one AS hop!

[Siganos 01]
Part A. III: The Effect of Protocols

- Some background
- BGP and topology
- BGP and routing
- BGP and routing robustness
  - The attack of the worms
- BGP and scalability
Robustness: Path Updates Frequency

- Send updates for path no sooner than 30 sec
- Why? For stability and overhead reduction
- Side-effects: Convergence takes longer
- What is the right interval?
  - Recent studies say that 30s is too long
- Path Dampening:
  - Ignore frequently changing paths
  
  [Nicol, Premore, Griffin, Cowie, Oglieski, Feldman+]
Analyzing Update Messages

Number of prefix announcements in 30 sec intervals

September 18:
Notice over 20-fold exponential growth
returning back to baseline after 4 days!

By Renesys

# prefix announcements per 30 seconds
[Cowie Oglieski 01]
Initial Observations

_updates show daily and weekly periodicity

_There is no evidence of BGP disturbance on:_
  - The Baltimore tunnel train 18 July that destroyed Internet lines
  - The Sept 11 terrorist attack

_There are some spikes at:_
  - 19 July 2001
  - 18-22 September 2001
Prefix announcements by peer
RIPE NCC, September 10 - 22, 15-min intervals

September 18:
Long-tail wave of routing instabilities in BGP message streams from major Internet providers

By Renesys
The NIMDA Worm

By Renesys
The Attack of The Worm

September 18 BGP event correlates in time with Nimda worm attack

Smaller events: leakage of reserved AS numbers

But, how could the worm affect the routers?
How Did The Worm Affect BGP?

The Worm “Ate” the Router CPU Time!

Busy = non responsive

By Renesys
Another Opinion

- Observed correlation may have been an artifact of the measurement infrastructure [Wang et al IMW02]
- Monitoring links where multi-hop = more vulnerable than real BGP links
The Scope of AS Instability

Instability is contained locally (Good News)
Summary of BGP Instability

- Globally correlated BGP instability is not uncommon
- Some causes are well understood (misconfiguration, bad path announcements)
- Some others are less well understood, and more worrisome:
  - Worms, indirectly attack router CPU
Part A. III: The Effect of Protocols

- Some background
- BGP and topology
- BGP and routing
- BGP and routing robustness
  - The attack of the worms
- BGP and scalability
BGP Table Growth: The Prediction

Worst Case
Continued Exponential Growth
150,000 entries by January 2002

Best Case
Elimination of all extraneous routing entries
75,000 entries by January 2002

By G. Huston
The BGP Table Growth: The Truth

- Growth flattened out in 2001
- Why?
  - Better management
  - More aggregation of IP prefixes
  - Dot-com crash?

Time

By G. Huston

Advanced Networks

© M. and C. Faloutsos
Routing-Table Size Variation

Active BGP entries vs Time
Larger ASes have significantly larger tables

By G. Huston

© M. and C. Faloutsos
Advanced Networks
Some Open Questions

- Is there a pattern in BGP updates?
- How do floods of updates propagate?
  - Correlations and cascading phenomena
- How secure and robust is BGP?
  - Cyber-terrorism
- Can I predict BGP scaling and growth?
Practical BGP-Related Questions

❖ How can we handle massive data (100 Gb)?
❖ How can I identify correlations between BGP tables?
❖ Can we detect automatically pathologies?
   • Periodicities or unexpected bursts
Conclusions

We have seen major steps in Internet modeling
- Self-similarity and LRD to describe traffic and performance
- Power-laws to describe the topology

But still, we can not model a lot of things
- Spatio-temporal correlations
- Interest and group behavior
- Anomaly detection

Challenges:
- Massive multidimensional data
- Time – space correlations
- Case dependent phenomena
Can Data-Mining Help?

- Capture patterns and invariants
- Compare and cluster behaviors
- Detect: Identify irregular patterns
- Troubleshoot: correlate problem with cause
- Predict behavior
At Last, The End Of Part A

_for list of bibliography and good sites:
www.cs.ucr.edu/~michalis/tutorial/tutorial.html