Network Distribution Capacity and Content-Pipe Divide

Mung Chiang
Electrical Engineering Department, Princeton

CISS Optimization of Networks
March 19, 2008

Overview

A talk with no results (except pointers to recent results), but questions

A talk with lazy slides too

- Content-pipe divide
- Three examples
- General interactions
- Network distribution capacity
- Special cases

Acknowledgement

- Collaboration: Rob Calderbank, Phil Chou, Prashanth Hande, Wenjie Jiang, Jin Li, Ying Li, Zhu Li, Shao Liu, Jennifer Rexford, Sudipta Sengupta, Rui Zhang-Shen
- Discussion: Keith Cambron, Gary Chan, Xing Jin, Raj Savoor, Steve Sposato
- Partners: AT&T, Microsoft, and Motorola

Content-Pipe Divide

Network usage now dominated by sharing and viewing content, especially video

- IPTV and triple play
- PPLive
- YouTube and Facebook
- Web 2.0

The third wave of Internet usage has arrived

Shake many basic assumptions in network design:

- Asymmetry between uplink and downlink data rates
- "Horizontal decomposition" into access-metro-core hierarchy
- "Vertical decomposition" into application layer and all layers below

Leads to Content-Pipe Divide

Content Side

Those who generate and distribute content

- Media companies who own video and music
- End-users who post video online
- Operators of content distribution network (CDN)
- Operators of peer-to-peer (P2P) sharing systems

- Seek the best way to distribute content, through technologies including multimedia signal processing as well as content caching, relaying, and sharing
- Take the network as just a means of transportation

Pipe Side

Those who design and operate the network

- Internet Service Providers (ISP)
- Equipment vendors
- Network management software vendors
- Municipalities and enterprises running their own networks

- Seek the best way to meet end-user requirements, through technologies including those that manage resources on each link, between links, and end-to-end
- Take the content as just bits to transport between given nodes in the network

“Divide” Example 1: Traffic Management

- ISP run traffic management protocols (TCP congestion control and intra-AS routing), assuming that the traffic matrix is fixed and can be accurately estimated

- On possibly different timescales, server selection by CDN or peer selection by P2P changes the traffic matrix by adapting in their own way to the user-perceived delay and throughput, as induced by ISP's traffic management

A feedback loop is present

Particularly challenging to ISP because, unlike the voice applications, the way videos are generated, shared, and viewed are quickly evolving through disruptive technologies and user-initiated protocols

"Divide" Example 2: Video Delivery

Generation and processing of multimedia signals have traditionally been designed in separation from the way the resulting packets are treated inside the network, e.g., shaping, marking, and dropping

Opportunities of jointly designing how video packets are coded and transported

- Dropping packets by frame-utility
- Assigning multiple streams of video packets coded differently for the same source on multiple paths

Content-Aware Networking?

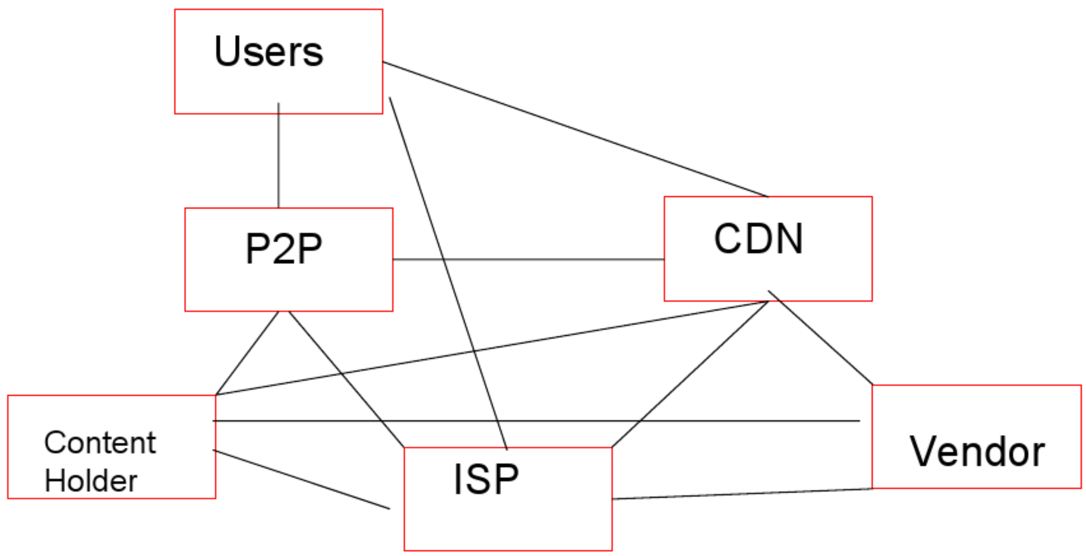
“Divide” Example 3: Net Neutrality

What kind of pricing structures by ISP over different content will be

- Efficient
- Fair
- Incentive-compatible

To all parties?

Question on Six Party Interactions



- Any pairwise interaction is interesting
- Triangle of interactions even more interesting
- So are "multipaths" in the interaction diagram

Special Case: ISP-P2P Interactions

	ISP no change	ISP changes
P2P no change	Current practice	
P2P changes		

- Collaboration III: Sharing control
- Collaboration II: Sharing information
- Collaboration I: Anticipate reaction by the other party

Bottleneck: Lack of unilaterally-actionable, incrementally-deployable, and backward-compatible strategies for cooperation

Question on Fundamental Limits: NDC

Various existing notions of capacity:

- Transportation of flows in a graph
- Largest rate subject to vanishingly small decoding error probability
- Largest set of arrivals subject to queue stability

Network Distribution Capacity (NDC) is a combination of these:

- Combinatorial problems of overlay graph construction ("distribution" in NDC)
- Communication network problems over various degrees of freedom ("network" in NDC)

Introduction of wireless components will bring further issues:

- Mobility, shared medium, time-varying links

NDC: Constants

Given

- A directed graph $G = (V, E)$

- A set of contents $D = (D_1, \ldots, D_N\}$, where each D_i consists of 3-tuples: the size of the content M_i, a set T_i of destinations $T_{it} \in V,\ t = 1, 2, \ldots, |T_i|$, who demand the content, and a set S_i of sources $S_{is} \in V,\ s = 1, 2, \ldots, |S_i|$, who can supply the content

- The set of sources of content can become larger after more nodes $v \in V$ obtain the content

- A node can be a source but not a destination (server), a destination but not a source (client), or both (peer), or neither (router)

- Extensions: content chunk availability and peer churn

NDC: Some of the Variables

• For each content i and destination T_{it}, a subset of S_i, denoted by S_{it}, that serve node T_{it}. For each source S_{is}, there is a set T_{is} of destinations served by it. These subsets (S_{it}, T_{is}) are variables, while the original source and destination sets (S_i, T_i) are constants

• Transmission rate of each node $x_v \geq 0$, and queue management policy in each router

• Routing matrix A, which in turn depends on the variable of load balancing matrix H and the constant of physical topology matrix W: $A = HW$, $H_{vp} \in [0, 1]$ is the fraction of traffic from node v on path p, and $W_{pl} \in \{0, 1\}$ is the boolean indicator of whether link $l \in E$ is on path p or not

• Link capacity c_l and node capacity c_v

• The variables are obviously constrained with each other

• Source and destination sets are changing, and the construction of content distribution topology over time is a design variable

NDC: Some of the Metrics

• When $\{M_i\}$ are infinite, what are the time-averaged throughputs R_{it} of content distribution for each of the receivers T_{it}?

• For finite $\{M_i\}$ arriving at the system according to some pattern, what are the completion times Q_{it} of content distribution for each of the receivers T_{it}?

Other possibilities:

• Utility function based on user-perceived video quality

• ISP cost functions

• Robustness metrics

NDC Models

Model so far is neither complete nor tractable

- Adding details to the formulations and taking asymptotic limits in appropriate dimensions will be necessary

- Even then, only special cases (holding some degrees of freedom as constants) over special (G, D) will be analytically tractable

NDC

What are the best achieavable R_{ij}, denoted as R^*_{ij}, and Q_{ij}, denoted as Q^*_{ij}?

- Can also consider stability capacity as the arrival of D that would keep Q finite over time
- For special graph G (a less desirable restriction) and special demands D (a more interesting restriction), closed form solutions may be possible
- In general, either asymptotic results in terms of the order of growth of R or decay of Q, or an efficient computation of R and Q are more likely
- Computing the answers to the question is NP-hard for sufficiently general G and D

Achieving NDC

What R_{ij} and Q_{ij} are achieved by the current practice of network providers and content distributors?

By "practice", we mean a set of architectures and protocols of controlling an entity's own variables and of interacting with other entities. For example,

• ISP runs OSPF at the timescale of hours, TCP Reno at the timescale of round trip time, and bandwidth allocation at the timescale of WDM wavelength assignment

• P2P runs BitTorrent with tit-for-tat and opportunistic unlocking for peer selection

• CDN runs locality-based, user-delay-minimization for content caching

• Their interaction is only based on each entity's measurements without any explicit message passing

Special Cases: P2P Streaming Capacity

Metrics:

- Throughput

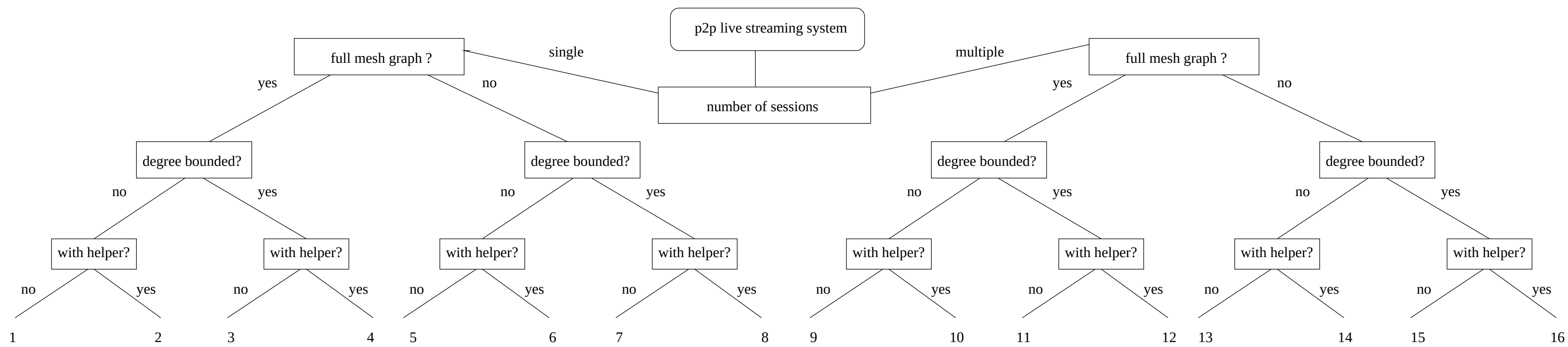
Degrees of freedom:

- Overlay topology and peering relationship: Tree (single or multi) or mesh or hybrid, Pull or push, Locality based, Clustered architectures for scalability
- Streaming rate (at application layer)

Streaming capacity (optimal, bound, achieving, partially achieving):

- Li, Chou, Zhang 2004 (solved 2 cases)
- Chen, Ponec, Sengupta, Chou, Li 2008 (solved 2 more cases)
- Liu, Zhang-Shen, Jiang, Rexford, Chiang 2008 (solved 2 more cases)
- Liu, Sengupta, Chiang, Li, Chou 2008 (solved 6 more cases)

Taxonomy


p2p live streaming system
number of sessions
single
multiple
full mesh graph ?
full mesh graph ?
yes
no
yes
no
degree bounded?
degree bounded?
degree bounded?
degree bounded?
no
yes
with helper?
no
yes
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

Extensions

Metrics:

- Delay (buffering latency, delay, jitter)
- Robustness (peer churn and breakdown)
- ISP-friendliness (pricing, congestion)

Degrees of freedom:

- Delivery schemes (P2P)
- Construction of caching nodes (ISP)
- Underlay traffic engineering (ISP)
- Underlay congestion control (ISP)

Tradeoff surfaces in 4-dimensional metric space by tuning many knobs?

More Extensions

	Technology	Neutrality policy
Content distribution alone	Slides 19-20	
Interact with transportation		
Interact with video processing		

Many under-explored areas in this emerging field

The Future of Networking Research

Understand content, its distribution and sharing over a network, and the interactions between content and pipes

Contacts

chiangm@princeton.edu
www.princeton.edu/∼chiangm