IETF P2P Infrastructure Workshop Presentation

Richard Woundy
Senior Vice President, Software & Applications
Office of the CTO

Jason Livingood
Executive Director, Internet Systems Engineering
National Engineering & Technical Operations

May 28, 2008

Comcast

Happy to Be Here Today

- Thanks to the IETF, MIT, and others for helping to organize this important and timely workshop.
- Also appreciate the attendance and participation of application developers.
  - Apps providers are recognizing the need to be more network friendly.
  - Networks are recognizing the need to be more apps friendly, as technology advances over time and user needs and desires evolve.
  - We need to work as a community, via organizations like the IETF, to make the Internet work the best it can and continue the commercial success of the Internet
  - Basically this is the way the Internet has evolved for years including early FTP, web, and many other Internet experiences…

Update on Our Commitments to the Internet Community

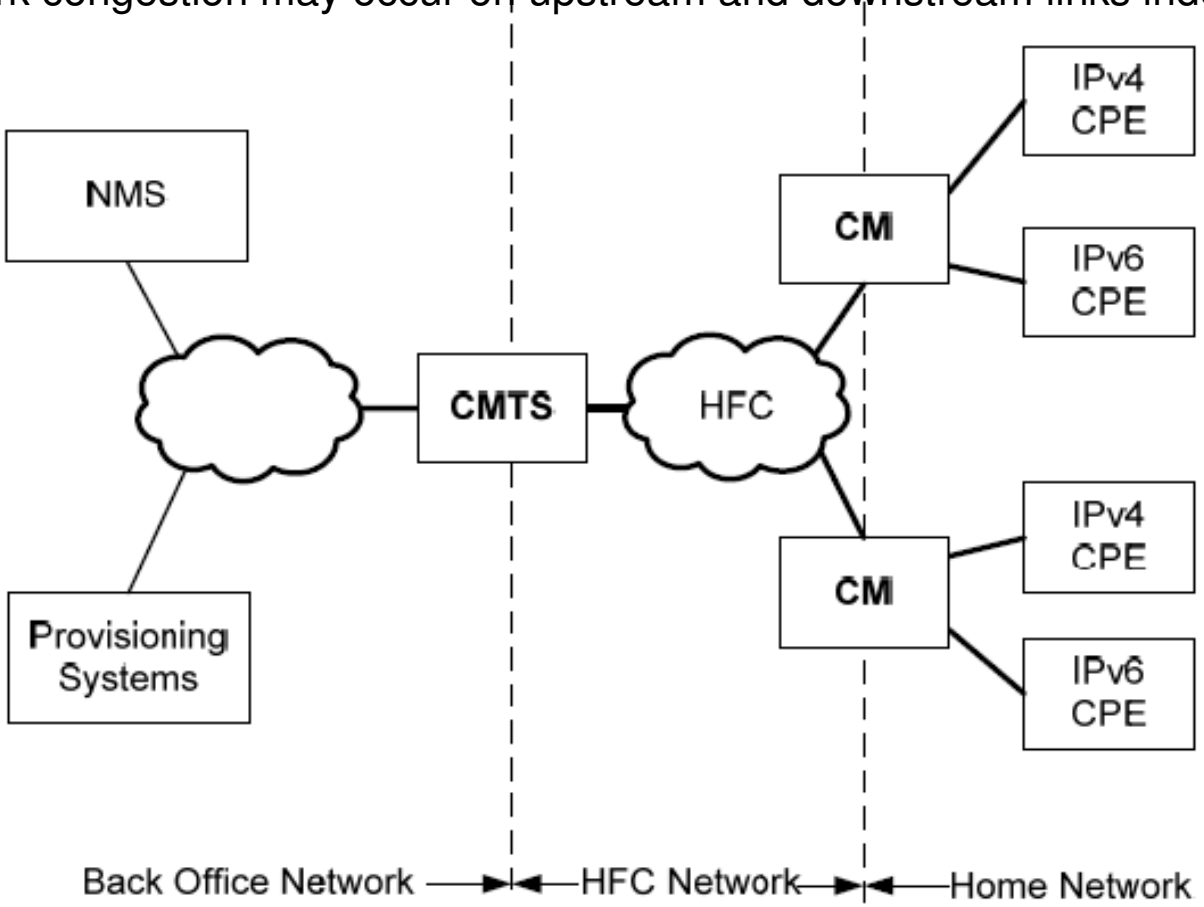
- Implement a protocol-agnostic network management technique instead of specifically targeting certain P2P applications by EOY.
- Increased Focus Points: transparency, disclosure, openness, fairness.
- Continue on our plan to roll-out DOCSIS 3.0 and more than double upstream capacity in several key markets by EOY
- *But let us begin with an overview of our network and try to help define the area of contention to some extent…*

Background for P2P Specific Network Management

- Back in 2005-2006, there were growing complaints related to delay sensitive traffic (particularly VoIP)
  - Public complaints submitted about Vonage service over Comcast's network
    - http://www.vonage-forum.com/ftopic11377.html, "Comcast vs. Vonage"
    - http://blogs.zdnet.com/ip-telephony/?p=961&tag=rbxccnbzd1, "Vonage over Comcast user: here's my proof there are problems"
    - Quote from http://www.csmonitor.com/2006/0315/p14s01-stct.html: "In one online forum, Vonage customers shared suspicions that cable company Comcast is degrading the quality of their Vonage phone calls. (Comcast is rolling out a digital phone service.) Comcast and Vonage Holding Corp. have denied that any such problem exists."
    - Note: an unrelated issue was congestion between two transit providers in the Comcast and Vonage path
  - At the same time, significant increases in P2P traffic, especially in upstream
    - Sandvine currently estimates that P2P represents 72% of upstream traffic
- Most ISPs started managing P2P traffic about 2-3 years ago
  - Comcast started deployment of P2P protocol-specific network management in 2006
    - Note: Comcast was not an early adopter of this technology
    - http://ubmtechnology.mediaroom.com/index.php?s=43&item=1474#, "P2P Bandwidth Hogs Drive Network Operators to Deploy Deep Packet Inspection, New Report Finds"
  - Customer complaints about delay sensitive traffic dropped as a result
    - Last official communication from Vonage to Comcast was back in March 2006

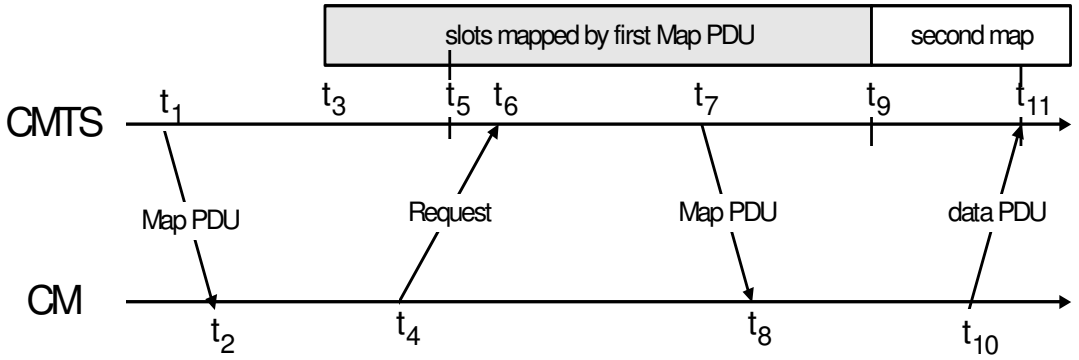
Overview of DOCSIS Network Architecture

- Focus is on the DOCSIS network, which is CMTS ←→ CM
  - A single CMTS node serves thousands of homes
  - A single CMTS node has multiple DOCSIS domains
  - Each DOCSIS domain could experience congestion independently of other domains
  - Network congestion may occur on upstream and downstream links independently.



Comcast

Upstream Data Transmission and Congestion



Simplified upstream data transmission process:

1. CMTS transmits MAP PDU to downstream CMs (t1)
2. CM scans MAP for request opportunity (t2)
3. CM sends Request PDU to CMTS during request opportunity (t4)
4. CMTS transmits MAP PDU that includes data grant (t7)
5. CM transmits data PDU according to MAP data grant (t8, t10)
6. CM may 'piggyback' next request in data PDU (t10)

How congestion manifests itself in the DOCSIS network

- Like all networks, typically results from offered instantaneous load exceeding available capacity
- If many CMs attempt to send simultaneous upstream requests, there may be collisions which result in upstream transmission delays

Note: how "TCP flow fairness" applies to the DOCSIS upstream is an open research question

Constraints for the ISP

- ISP must be responsive to dissimilar customer application demands
  - Customer care call volume is an obvious indicator of customer dissatisfaction, as well as its own support cost concern
  - Mix of desired customer applications tends to vary according to demographics, e.g., higher P2P usage in college environments
  - Interactive applications (VoIP, web, streaming video) tend to have much stronger diurnal consumption patterns than bulk file distribution (P2P)
- ISP must balance multiple external concerns
  - Internet community, government regulators, traffic sources & sinks, and sustainable business models, etc.
- Network capacity increases are not instantaneous
  - DOCSIS bandwidth augmentation usually requires fiber node splits and CMTS port allocations; it sometimes requires new fiber runs, additional CMTS blades and chassis, and occasionally the allocation of additional RF spectrum
  - Additional access network capacity can be consumed quickly

Necessary Elements for a Long-Term Solution

- Provide best possible network experience for broadest set of customers
  - Minimize or eliminate cross-customer service quality impacts
  - Reduce customer care calls
- Enable customers to control their own network experience
  - Inform customers of application bandwidth usage and network reaction
  - Perhaps enable customer-centric prioritization of application bandwidth usage
- Enable continued Internet evolution
  - Avoid ‘cat and mouse game’: detection and mitigation of specific protocols
  - Enable transparency of network operation for current/future applications
- Support a reasonable network capacity upgrade schedule
  - Support growth in number of customers
  - Support growth in per-customer average and peak bandwidth
  - Avoid uneconomic capacity upgrades that benefit only 5% of heavy usage customers

Comcast Engineering Activities

We break our engineering activities down into several key areas:

- Congestion Management Improvements
  - This is where a new network management technique comes into play.
- P2P Optimizations
  - Near-Term:
    - Tracker optimizations (optimizing & localizing P2P flows)
    - Caching
    - P2P client optimizations
  - Longer-Term:
    - Signaling from the Network to Clients
    - Signaling from Clients to the Network
    - Improved methods for allocating network resources
- DOCSIS 3.0 Technology
  - Part of normal technical evolution
  - Provides additional capacity but does not eliminate the need for congestion management

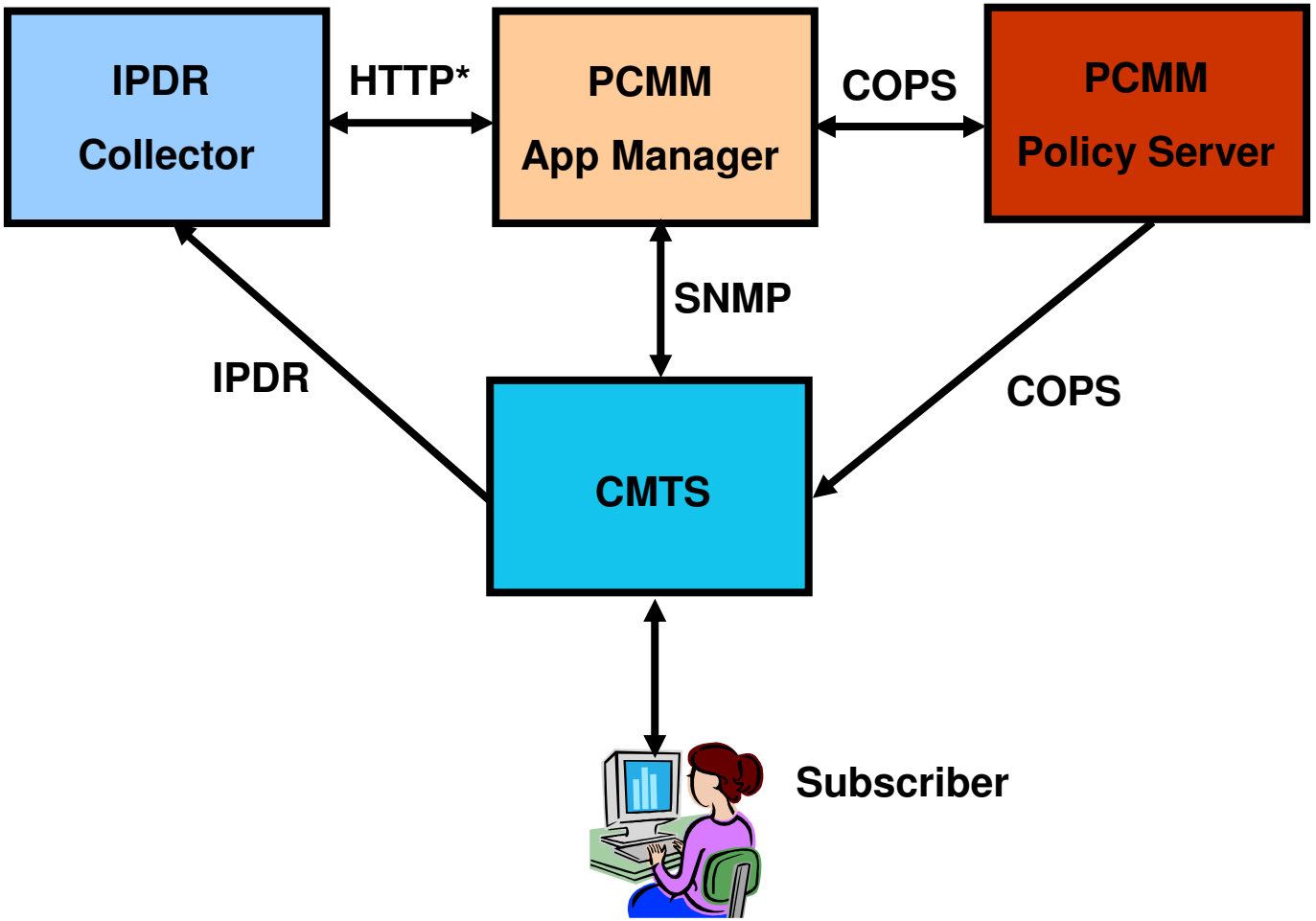
DOCSIS 3.0 and Upstream Capacity

- DOCSIS 3.0 is on track for ~20% of our network in 2008, additional markets in 2009 - 2010.
  - Increases broadband speeds offered to customers.
- Comcast committed earlier this year to double upstream capacity in several key markets by the end of the year
  - A plan to increase speeds will be announced by the company around the beginning of June.
  - A sub with 384Kbps upstream will go to 1Mbps
  - A sub with 768Kbps upstream will go to 2Mbps
- Putting this in context:
  - Normal course of business capacity augmentations & speed upgrades.
  - Widely understood that you can not build out of a peak network congestion problem which is largely the result of client software designed to maximize bulk bandwidth consumption.

A Protocol Agnostic Approach to Network Management

- Commitment: implement the new method by EOY, and provide useful consumer disclosure.
- Trials beginning in June:
  - Beginning in 2 markets, with a 3rd expected shortly thereafter.
  - Evaluating slight vendor and technical implementation variations.
  - Examining the effects of changes in markets with varying congestion characteristics.
  - Does not use P2P protocols or applications in use by a subscriber to make network management decisions.
  - Will make a determination on how best to implement by mid-summer, and will roll out this change to our entire network by the end of the year.

General Architecture for Network Management



* Interface has not been standardized. HTTP is one choice

A Protocol Agnostic Approach: Trial Details

- Before trial, all CHSI traffic considered “best effort”
- During trial, default traffic is re-classified as “priority.”
- As times of peak congestion approach, users who have exceeded certain usage thresholds have their traffic marking changed from priority to best efforts
  - Thus, when congestion subsequently occurs, subs with shorter-duration and burstier traffic patterns should be unaffected.
  - However, subs with longer-duration, bulk usage patterns may be affected via this best-effort QoS mechanism.
  - This does not reset connections.
  - Will attempt to protect real-time applications, where users would otherwise perceive delays/degradation.
  - Many of these real-time apps are competing over-the-top services, such as VoIP services, but may also be video conferencing, gaming, etc.
  - Protecting real-time application experiences is extremely important to users.

A Protocol Agnostic Approach: Trial Details

- IPDR is used to collect usage information from the CMTS.
- As peak congestion on a CMTS US or DS port approaches (the “Near Congestion State”), then users who exceed a certain threshold for a certain amount of time (entering a “Long Duration Bulk Consumption State”) will have traffic QoS markings changed from priority to best efforts.
- For a CMTS’s US or DS ports to enter the Near Congestion State, a Port Utilization Threshold must be exceeded for a specific period of time (the “Port Utilization Duration”).
  - The Port Utilization Threshold on the CMTS, measured as a %, will be varied during the trial.
  - The Port Utilization Duration on the CMTS, measured in minutes, will be varied during the trial.
- Once the CMTS is in a Near Congestion State, we will search for subscribers in a Long Duration Bulk Consumption State as candidates for temporary re-marking of their data flows.
- In order for a subscriber to temporarily enter a Long Duration Bulk Consumption State, they must consume a certain amount of their provisioned speed for a specific length of time.
  - The User Consumption Threshold, measured as a % of the provisioned US or DS bandwidth, will be varied during the trial.
  - The User Consumption Duration, measure in minutes, will be varied during the trial.

A Protocol Agnostic Approach: Trial Details

This describes the high-level mechanics of the technique. The exact values used will be varied during the course of the trial.

- Trial defaults, to be varied during the trial period:
  - *Upstream* Port Utilization Threshold ≥ AA%
  - *Downstream* Port Utilization Threshold ≥ BB%
  - *Upstream* Port Utilization Duration ≥ CC minutes
  - *Downstream* Port Utilization Duration ≥ DD minutes
  - *Upstream* User Consumption Threshold ≥ EE% of provisioned speed
  - *Downstream* User Consumption Threshold ≥ FF% of provisioned speed
  - *Upstream* User Consumption Duration ≥ GG minutes
  - *Downstream* User Consumption Duration ≥ HH minutes

Protocol-Specific Methods Remain Valuable

- There are clearly cases where protocol-specific methods remain valuable:
  - Where the ISP plays a role in the E2E application (ex: email using comcast.net, DNS).
  - Where illegal, illegitimate, or damaging uses of the network occur.
    - Spam
    - Bots
    - DoS attacks
    - Illegal content
    - Et cetera