Cache Capacity Allocation for BitTorrent-like Systems to Minimize Inter-ISP Traffic

Valentino Pacifici, Frank Lehrieder, György Dán

School of Electrical Engineering
KTH Royal Institute of Technology
Stockholm, Sweden

Institute of Computer Science
University of Würzburg
Würzburg, Germany

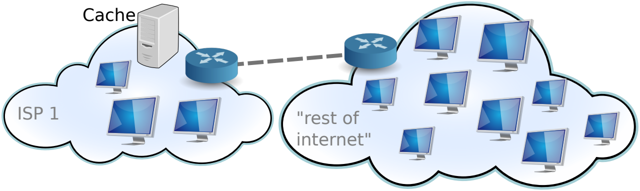
Orlando, March 29, 2012

P2P Traffic

- Up to 70 % of network traffic
- Source of Inter-ISP traffic ⇒ cost for low level ISPs

Decreasing Inter-ISP traffic

1. Locality awareness
2. P2P caching



P2P Caching

Cache resource management

1. Storage capacity ⇒ cache eviction (LRU,LFU,GDS,ARC,...)
2. Bandwidth ⇒ not actively managed (e.g. Web caches)

Should bandwidth be actively managed so as to minimize the amount of Inter-ISP traffic?

P2P Caching

Cache resource management

1. Storage capacity ⇒ cache eviction (LRU,LFU,GDS,ARC,...)
2. Bandwidth ⇒ not actively managed (e.g. Web caches)

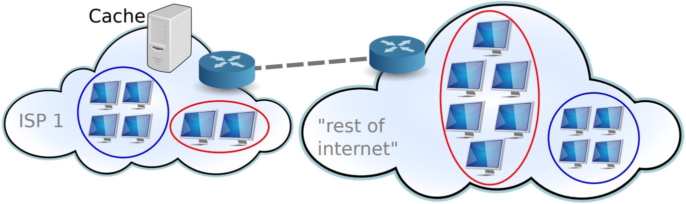
Should bandwidth be actively managed so as to minimize the amount of Inter-ISP traffic?

P2P Caching

Cache resource management

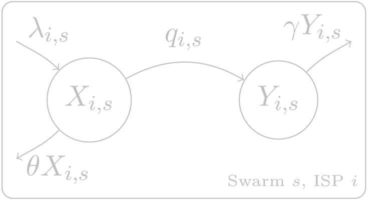
1. Storage capacity ⇒ cache eviction (LRU,LFU,GDS,ARC,...)
2. Bandwidth ⇒ not actively managed (e.g. Web caches)

Should bandwidth be actively managed so as to minimize the amount of Inter-ISP traffic?



System Model without Cache

- Set of ISPs $\mathcal{I} = \{1, \ldots, I\}$, Set of swarms $\mathcal{S} = \{1, \ldots, S\}$
- Markovian model of system dynamics
  - System state $Z_{i,s}(t) = (X_{i,s}(t), Y_{i,s}(t))$
  - Parameters $(\lambda_{i,s}, \theta, \gamma, \mu, \eta)$

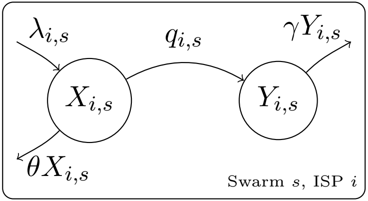


$$q_{i,s} = \underbrace{\frac{X_{i,s}}{X_s}\mu(\eta X_s + Y_s)}_{\text{available upload rate}}$$

- Incoming inter-ISP traffic rate $I_{i,s}(Z_s(t), .)$

System Model without Cache

- Set of ISPs $\mathcal{I} = \{1, \ldots, I\}$, Set of swarms $\mathcal{S} = \{1, \ldots, S\}$
- Markovian model of system dynamics
  - System state $Z_{i,s}(t) = (X_{i,s}(t), Y_{i,s}(t))$
  - Parameters $(\lambda_{i,s}, \theta, \gamma, \mu, \eta)$

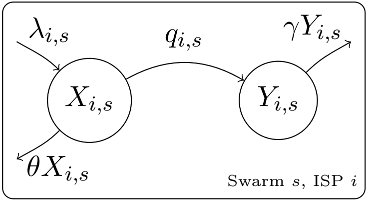


$$q_{i,s} = \underbrace{\frac{X_{i,s}}{X_s}\mu(\eta X_s + Y_s)}_{\text{available upload rate}}$$

- Incoming inter-ISP traffic rate $I_{i,s}(Z_s(t), .)$

System Model with Cache

- Set of ISPs $\mathcal{I} = \{1, \ldots, I\}$, Set of swarms $\mathcal{S} = \{1, \ldots, S\}$
- Markovian model of system dynamics
  - System state $Z_{i,s}(t) = (X_{i,s}(t), Y_{i,s}(t))$
  - Parameters $(\lambda_{i,s}, \theta, \gamma, \mu, \eta, \kappa_{i,s})$
- $K_i < \infty$ bandwidth capacity of cache in ISP i



$$q_{i,s} = \underbrace{\frac{X_{i,s}}{X_s}\mu(\eta X_s + Y_s) + \kappa_{i,s}}_{\text{available upload rate}}$$

- Incoming inter-ISP traffic rate $I_{i,s}(Z_s(t), \kappa_{i,s}(t))$

Cache Bandwidth Allocation Problem

- Cache bandwidth allocation of ISP i at time t

$$\kappa_i(t) = (\kappa_{i,1}(t), \ldots, \kappa_{i,S}(t))$$

- Defined by *policy* π: $\kappa_i(t) = \mathcal{F}^{\pi}\left(\left(Z(u)\right)_{u<t}, \left(\kappa_i(u)\right)_{u<t}\right)$

Cache Bandwidth Allocation Problem

- Cache bandwidth allocation of ISP i at time t

$$\kappa_i(t) = (\kappa_{i,1}(t), \ldots, \kappa_{i,S}(t))$$

- Defined by *policy* π: $\kappa_i(t) = \mathcal{F}^\pi\Big(\big(Z(u)\big)_{u<t}, \big(\kappa_i(u)\big)_{u<t}\Big)$

- Expected incoming inter-ISP traffic under allocation policy π

$$C_i^\pi(Z(0), T) = E_{Z(0)}^\pi\left[\int_0^T \sum_{s\in\mathcal{S}} I_{i,s}(Z_s(t), \kappa_{i,s}(t))dt\right]$$

Cache Bandwidth Allocation Problem

- Cache bandwidth allocation of ISP i at time t

$$\kappa_i(t) = (\kappa_{i,1}(t), \ldots, \kappa_{i,S}(t))$$

- Defined by *policy* π: $\kappa_i(t) = \mathcal{F}^\pi\left(\left(Z(u)\right)_{u<t}, \left(\kappa_i(u)\right)_{u<t}\right)$

- Expected incoming inter-ISP traffic under allocation policy π

$$C_i^\pi(Z(0), T) = E_{Z(0)}^\pi\left[\int_0^T \sum_{s\in\mathcal{S}} I_{i,s}(Z_s(t), \kappa_{i,s}(t))dt\right]$$

Find the optimal policy $\pi^* \in \Pi$ s.t.

$$\inf_{\pi\in\Pi} C_i^\pi(Z(0)) = \inf_{\pi} \limsup_{T\to\infty} \frac{1}{T} C_i^\pi(Z(0), T).$$

Cache Bandwidth Allocation Problem

- Cache bandwidth allocation of ISP i at time t

$$\kappa_i(t) = (\kappa_{i,1}(t), \dots, \kappa_{i,S}(t))$$

- Defined by *policy* π: $\kappa_i(t) = \mathcal{F}^\pi\left(\left(Z(u)\right)_{u<t}, \left(\kappa_i(u)\right)_{u<t}\right)$

- Expected incoming inter-ISP traffic under allocation policy π

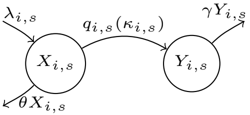
$$C_i^\pi(Z(0), T) = E_{Z(0)}^\pi \left[\int_0^T \sum_{s\in\mathcal{S}} I_{i,s}(Z_s(t), \kappa_{i,s}(t))dt\right]$$

Find the optimal policy $\pi^* \in \Pi$ s.t.

$$\inf_{\pi\in\Pi} C_i^\pi(Z(0)) = \inf_{\pi} \limsup_{T\to\infty} \frac{1}{T} C_i^\pi(Z(0), T).$$

Existence of Optimal Stationary Policy

- Markov Decision Process
  $< \mathcal{Z}, \mathcal{K}, \underline{\underline{Q}}(\kappa), I(z,\kappa) >$



Theorem

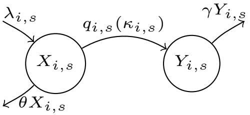
There exists an optimal stationary policy $\pi^* \in \Pi$ *that minimizes* $C_i^\pi(Z(0))$

The optimal policy π^*

- Stationary: $\kappa_i(t)$ is only a function of the system state $Z(t)$
- Calculation requires steady state probabilities
  - Prohibitive even for few ISPs and swarms
- Use of simple approximations to gain insight...

Existence of Optimal Stationary Policy

- Markov Decision Process
  $< \mathcal{Z}, \mathcal{K}, \underline{\underline{Q}}(\kappa), I(z,\kappa) >$



Theorem

There exists an optimal stationary policy $\pi^* \in \Pi$ *that minimizes* $C_i^{\pi}(Z(0))$

The optimal policy π^*

- Stationary: $\kappa_i(t)$ is only a function of the system state $Z(t)$
- Calculation requires steady state probabilities
  - Prohibitive even for few ISPs and swarms
- Use of simple approximations to gain insight...

One-Step Look Ahead (OLA)

- Minimize the incoming inter-ISP traffic rate given the system state

$$\kappa_i(t) = \arg\min_{\kappa_i \in \mathcal{K}_i} \sum_{s \in \mathcal{S}} I_{i,s}(Z_s(t), \kappa_{i,s})$$

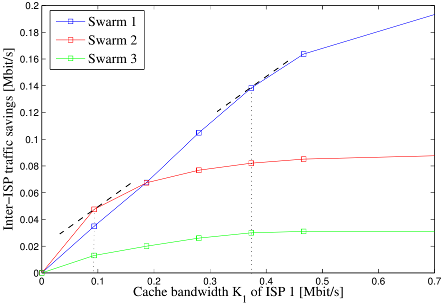
- Short term approximation → disregards system dynamics

One-Step Look Ahead (OLA)

- Minimize the incoming inter-ISP traffic rate given the system state

$$\kappa_i(t) = \arg\min_{\kappa_i \in \mathcal{K}_i} \sum_{s \in \mathcal{S}} I_{i,s}(Z_s(t), \kappa_{i,s})$$

- Short term approximation $\rightarrow$ disregards system dynamics



Optimal $\kappa_i(t)$ leads to equal marginal traffic saving for every swarm

$$\kappa_{i,s} > 0 \quad \Rightarrow \quad \frac{\partial I_{i,s}(z_s, \kappa_{i,s})}{\partial \kappa_{i,s}} = \zeta$$

$$\kappa_{i,s} = 0 \quad \Rightarrow \quad \frac{\partial_{-} I_{i,s}(z_s, \kappa_{i,s})}{\partial \kappa_{i,s}} \geq \zeta$$

Steady-State Optimal (SSO)

- Minimize the incoming inter-ISP traffic rate at steady state

$$\overline{\pi}^* = \arg\min_{\kappa_i \in \mathcal{K}_i} \sum_{s \in \mathcal{S}} \overline{I}_{i,s}(\kappa_{i,s})$$

- Long term approximation $\rightarrow$ non adaptive policy
- $\overline{I}_{i,s}(\kappa_{i,s}) = I\left(\overline{x}^{\overline{\pi}}_{i,s}(\kappa_{i,s}), \overline{y}^{\overline{\pi}}_{i,s}(\kappa_{i,s}), \kappa_{i,s}\right)$
- Based on fluid model [1] of cache impact on system state

$$\begin{aligned}
\overline{x}^{\overline{\pi}}_{i,s} &= \frac{\lambda_{i,s}}{\nu\left(1+\frac{\theta}{\nu}\right)} - \frac{\kappa_{i,s}}{\mu\eta\left(1+\frac{\theta}{\nu}\right)} - \Delta_i(\mathbf{x},\mathbf{y},\kappa) \\
\overline{y}^{\overline{\pi}}_{i,s} &= \frac{\lambda_{i,s}}{\gamma\left(1+\frac{\theta}{\nu}\right)} + \frac{\kappa_{i,s}\theta}{\mu\eta\gamma\left(1+\frac{\theta}{\nu}\right)} + \frac{\theta}{\gamma}\Delta_i(\mathbf{x},\mathbf{y},\kappa),
\end{aligned}$$

[1] *F. Lehrieder, G. Dán, T. Hoßfeld, S. Oechsner and V. Singeorzan "The Impact of Caching on BitTorrent-like Peer-to-peer Systems" in Proc. IEEE Int'l Conf. Peer-to-Peer Computing (P2P), Aug. 2010*

Steady-State Optimal (SSO)

- Minimize the incoming inter-ISP traffic rate at steady state

$$\overline{\pi}^* = \arg\min_{\kappa_i \in \mathcal{K}_i} \sum_{s \in \mathcal{S}} \overline{I}_{i,s}(\kappa_{i,s})$$

- Long term approximation $\rightarrow$ non adaptive policy
- $\overline{I}_{i,s}(\kappa_{i,s}) = I\left(\overline{x}^{\overline{\pi}}_{i,s}(\kappa_{i,s}), \overline{y}^{\overline{\pi}}_{i,s}(\kappa_{i,s}), \kappa_{i,s}\right)$
- Based on fluid model [1] of cache impact on system state

$$\begin{aligned} \overline{x}^{\overline{\pi}}_{i,s} &= \frac{\lambda_{i,s}}{\nu\left(1+\frac{\theta}{\nu}\right)} - \frac{\kappa_{i,s}}{\mu\eta\left(1+\frac{\theta}{\nu}\right)} - \Delta_i(\mathbf{x}, \mathbf{y}, \kappa) \\ \overline{y}^{\overline{\pi}}_{i,s} &= \frac{\lambda_{i,s}}{\gamma\left(1+\frac{\theta}{\nu}\right)} + \frac{\kappa_{i,s}\theta}{\mu\eta\gamma\left(1+\frac{\theta}{\nu}\right)} + \frac{\theta}{\gamma}\Delta_i(\mathbf{x}, \mathbf{y}, \kappa), \end{aligned}$$

[1] *F. Lehrieder, G. Dán, T. Hoßfeld, S. Oechsner and V. Singeorzan "The Impact of Caching on BitTorrent-like Peer-to-peer Systems" in Proc. IEEE Int'l Conf. Peer-to-Peer Computing (P2P), Aug. 2010*

Smallest-Ratio Priority (SRP)

- Approximation of SSO for small cache bandwidth
  - For two ISPs at steady state:

$$\overline{I}_1(\kappa_1) \approx \frac{\overline{x}_1}{\overline{x}_1 + \overline{x}_2}\mu(\eta\overline{x}_2 + \overline{y}_2)$$

  - $\frac{\partial \overline{I}_1(\kappa_1)}{\partial \kappa_1}\Big|_{\substack{\kappa_1=0 \\ \kappa_2=0}} < 0$ and decreases monotonically in $r = \frac{\lambda_2}{\lambda_1}$
- Swarms with lowest ratio $\frac{\lambda_i}{\sum_{j \neq i} \lambda_j}$ have highest priority
- Practical implementation $\hat{r}_{i,s} = \frac{x_{i,s}(t)}{\sum_{j \neq i} z_{j,s}(t)}$
- Adaptive policy

Smallest-Ratio Priority (SRP)

- Approximation of SSO for small cache bandwidth
  - For two ISPs at steady state:

$$\overline{I}_1(\kappa_1) \approx \frac{\overline{x}_1}{\overline{x}_1 + \overline{x}_2}\mu(\eta\overline{x}_2 + \overline{y}_2)$$

  - $\frac{\partial \overline{I}_1(\kappa_1)}{\partial \kappa_1}\Big|_{\substack{\kappa_1=0 \\ \kappa_2=0}} < 0$ and decreases monotonically in $r = \frac{\lambda_2}{\lambda_1}$
- Swarms with lowest ratio $\frac{\lambda_i}{\sum_{j \neq i} \lambda_j}$ have highest priority
- Practical implementation $\hat{r}_{i,s} = \frac{x_{i,s}(t)}{\sum_{j \neq i} z_{j,s}(t)}$
- Adaptive policy

Evaluation Methodology

- Model of the incoming inter-ISP traffic for OLA and SSO policies

Simulations

- Flow level simulation in the ProtoPeer framework
- 6.5 hours of simulated time, up to 12.000 BitTorrent peers

Evaluation Methodology

- Model of the incoming inter-ISP traffic for OLA and SSO policies

Simulations

- Flow level simulation in the ProtoPeer framework
- 6.5 hours of simulated time, up to 12.000 BitTorrent peers

Experiments

- 500 PlanetLab nodes running BitTorrent
  - BitTorrent Mainline client 4.4.0
  - 4 hours experiments, 1 hour of warm-up period
  - Up to 8400 peers distributed among 12 swarms
- Dedicated Linux computer running the P2P cache
  - Bandwidth allocation policies implemented in Linux kernel TC

Evaluation Methodology

- Model of the incoming inter-ISP traffic for OLA and SSO policies

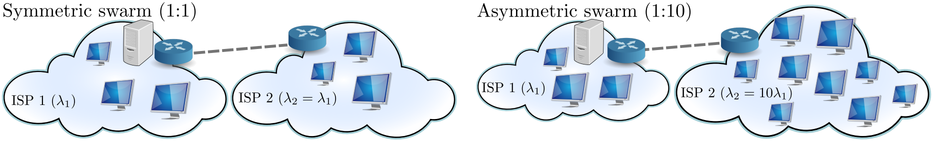
Simulations

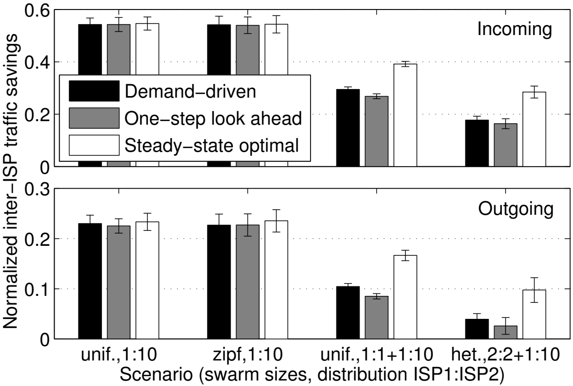
- Flow level simulation in the ProtoPeer framework
- 6.5 hours of simulated time, up to 12.000 BitTorrent peers

Experiments

- 500 PlanetLab nodes running BitTorrent
  - BitTorrent Mainline client 4.4.0
  - 4 hours experiments, 1 hour of warm-up period
  - Up to 8400 peers distributed among 12 swarms
- Dedicated Linux computer running the P2P cache
  - Bandwidth allocation policies implemented in Linux kernel TC

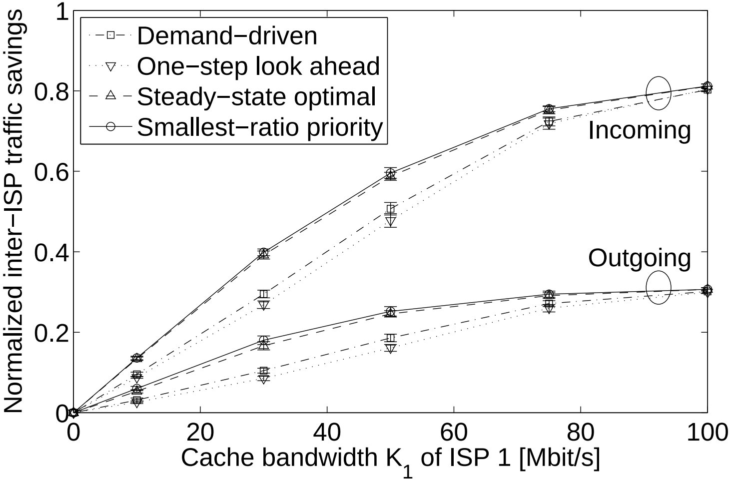
When Bandwidth Allocation Matters - Simulations



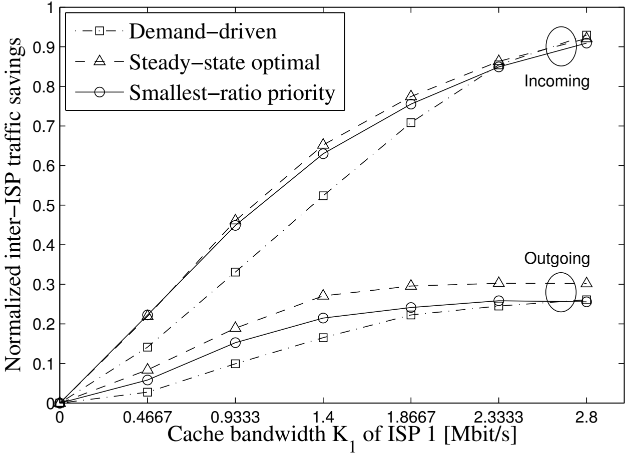


Bandwidth Allocation Policies Evaluation - Simulations

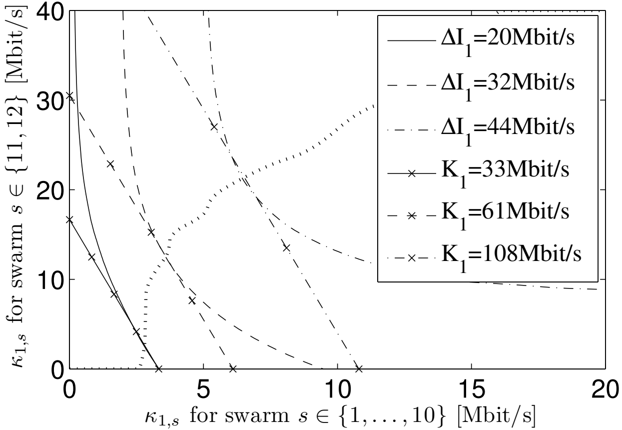
- Scenario *unif.,1:1+1:10*:



Validation - Experiments on PlanetLab



Indifference Map - Simulations



Conclusions

- Cache upload bandwidth allocation problem
- Existence of a stationary bandwidth allocation policy
- Various adaptive bandwidth allocation policies

Main observations

- Cache's impact on system dynamics is important
- Difference in swarms symmetry is the key
- Significant traffic savings possible
  - ~60% improvement in incoming inter-ISP traffic saving
  - ~250% improvement in outgoing inter-ISP traffic saving

Conclusions

- Cache upload bandwidth allocation problem
- Existence of a stationary bandwidth allocation policy
- Various adaptive bandwidth allocation policies

Main observations

- Cache's impact on system dynamics is important
- Difference in swarms symmetry is the key
- Significant traffic savings possible
  - ∼60% improvement in incoming inter-ISP traffic saving
  - ∼250% improvement in outgoing inter-ISP traffic saving

Cache Capacity Allocation for BitTorrent-like Systems to Minimize Inter-ISP Traffic

Valentino Pacifici, Frank Lehrieder, György Dán

School of Electrical Engineering
KTH Royal Institute of Technology
Stockholm, Sweden

Institute of Computer Science
University of Würzburg
Würzburg, Germany

Orlando, March 29, 2012