Game Theory-Based Nonlinear Bandwidth Pricing for Congestion Control in Cloud Networks

Abouzar Ghavami[§], Zhuozhao Li[*] and **Haiying Shen**[*]

[§]Department of Electrical and Computer Engineering, University of Maine, USA
[*]Department of Computer Science, University of Virginia, USA

UNIVERSITY of VIRGINIA

Outline

- Introduction
- System model
- Problem formulation and pricing policy design
- Performance Evaluation
- Conclusion and remark

Introduction

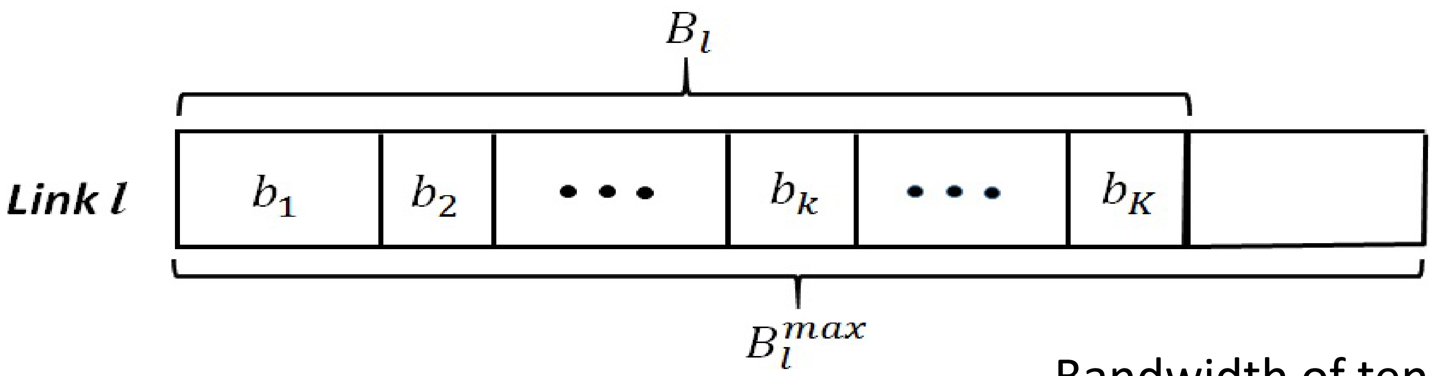
- Cloud computing
  - multiplexes computation, storage and network resources among different tenants
- Network resource
  - shared in a best-effort manner
  - avoiding congestion is very important to ensure QoS
- Charge in a flat rate
  - do not help to provide congestion control effectively

Contributions

- Maximize the social welfare
  - maximize the total satisfaction of the tenants
  - minimize the congestion over the link
- Propose a Nonlinear Bandwidth Pricing Model
- Propose a Best Response Bandwidth Update of Tenants process to converge to the socially optimal bandwidth schedule
- Conduct simulation study and real implementation to verify the performance of our proposed mechanism

System Model

Hose model



Bandwidth of tenant k

Total bandwidth

$$B_l = \sum_{k=1}^{K} b_k \leq \eta B^{max}$$

Maximum capacity of the link

Safety factor

System Model

Social welfare of tenants = Total satisfaction – Total congestion cost.

$$P(\mathbf{b}) = \sum_{k=1}^{K} \mathcal{U}_k(b_k) - \mathcal{V}(\sum_{k=K}^{S} b_k)$$

Social welfare

Bandwidth schedule of tenants

Satisfaction of tenant k

Bandwidth of tenant k

Congestion cost function

s.t. $\quad B_l - \eta_l B^{max} \leq 0.$

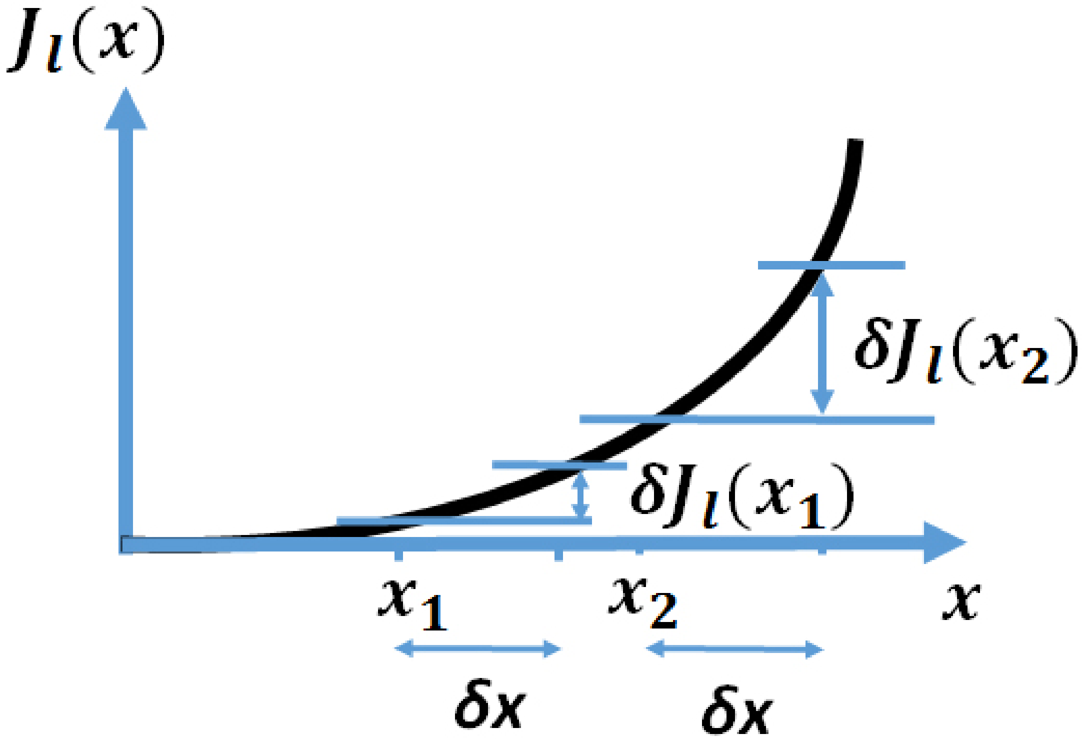
Goal: Find a bandwidth schedule that maximizes social welfare.

System Model

- Nonlinear pricing policy
  - Higher unit price when the current bandwidth usage is higher
- Disincentivizes tenants to use bandwidth when the bandwidth usage is higher
- Bandwidth payment (congestion cost) is based on a strictly convex function of the bandwidth usage

System Model

- Increment of the bandwidth pricing function using the derivative of the congestion cost function
- Congestion cost function is higher at the higher bandwidth usage



System Model

- Transformation of social welfare by adding constraints as penalty functions:

$$S(\mathbf{b}) = \sum_{k=1}^{K} \mathcal{U}_k(b_k) - \mathcal{V}(\sum_{k=1}^{K} b_k) - \mathcal{C}\left(\sum_{k=1}^{K} b_\mathrm{k} - \eta B^{max}\right)$$

Social welfare

Bandwidth schedule of tenants

Utility of tenant k

Congestion cost function

Overload penalty function

Bandwidth of tenant k

- Definiton: A bandwidth schedule is ***socially optimal*** if it maximizes the social welfare of the tenants.

Proposed Pricing Policy

- The cloud provider can minimize the congestion cost and avoid overloading by transferring cost to the tenants.
- Total cost of tenants = congestion cost + overloading cost

$$\mathcal{Z}(B_l) = \mathcal{V}(B_l) + \mathcal{C}(B_l - \eta B^{max})$$

Total cost of tenants

Total bandwidth

Total congestion cost of tenants

Overload cost

Proposed Pricing Policy

- Total cost a tenant ***k*** imposing on one core switch = total cost of all tenants – (total cost of all tenants excluding tenant ***k***)

$$\mathcal{Y}_k\,(\mathbf{b}_{-k},\mathbf{b}_k) - \mathcal{Y}_k(\mathbf{b}_{-k},\mathbf{0})$$
$$= \mathcal{Z}\underbrace{\left(\sum_{j\in\mathcal{K}-\{k\}} b_j\ + b_k\right)}_{\text{Total cost of all tenants}} - \mathcal{Z}\underbrace{\left(\sum_{j\in\mathcal{K}-\{k\}} b_j\right)}_{\text{Total cost of tenants excluding tenant } \boldsymbol{k}}$$

Bandwidth schedule of other tenants excluding tenant ***k***

- **Proposed bandwidth payment function** of a tenant = total cost it imposes on the link

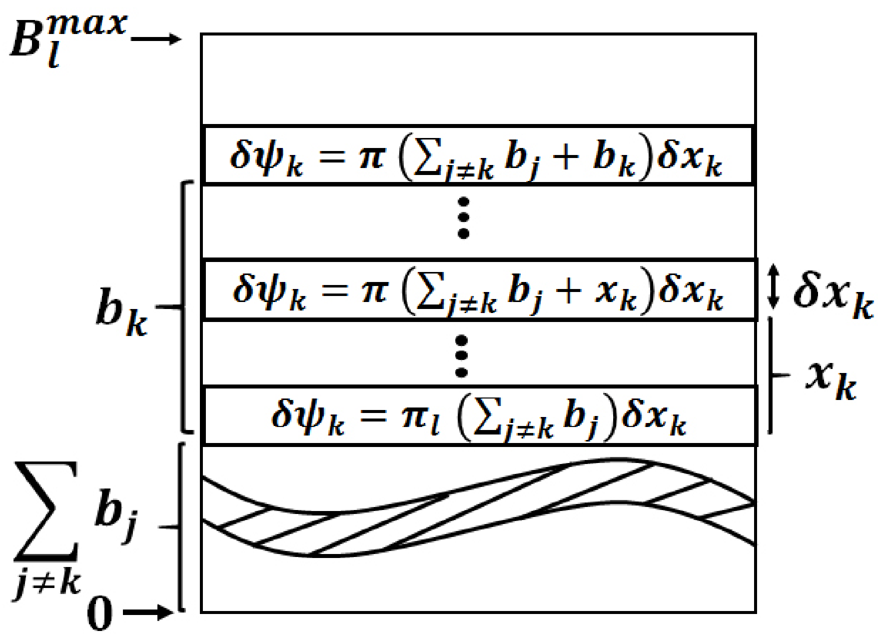
$$\psi_k(\mathbf{b}_{-k},\mathbf{b}_k) = \mathcal{Y}_k(\mathbf{b}_{-k},\mathbf{b}_k) - \mathcal{Y}_k(\mathbf{b}_{-k},\mathbf{0})$$

Proposed Pricing Policy

- **Proposed bandwidth payment function** of a tenant = total cost it imposes on the link

$$\psi_k(\mathbf{b}_{-k}, \mathbf{b}_k) = \mathcal{Y}_k(\mathbf{b}_{-k}, \mathbf{b}_k) - \mathcal{Y}_k(\mathbf{b}_{-k}, \mathbf{0})$$

Increments of the bandwidth pricing function for tenant k using the derivative of the congestion cost function, π.



Proposed Pricing Policy

- The utility of tenant ***k***

$$\mathcal{F}_k(\mathbf{b}_{-k}, \mathbf{b}_k) = \mathcal{U}_k(b_k) - \psi_k(\mathbf{b}_{-k}, \mathbf{b}_k)$$

Utility of tenant ***k***

Satisfaction function of tenant ***k***

Payment function of tenant ***k***

- Each tenant attempts to maximize its own utility

UNIVERSITY of VIRGINIA

Strategic Game

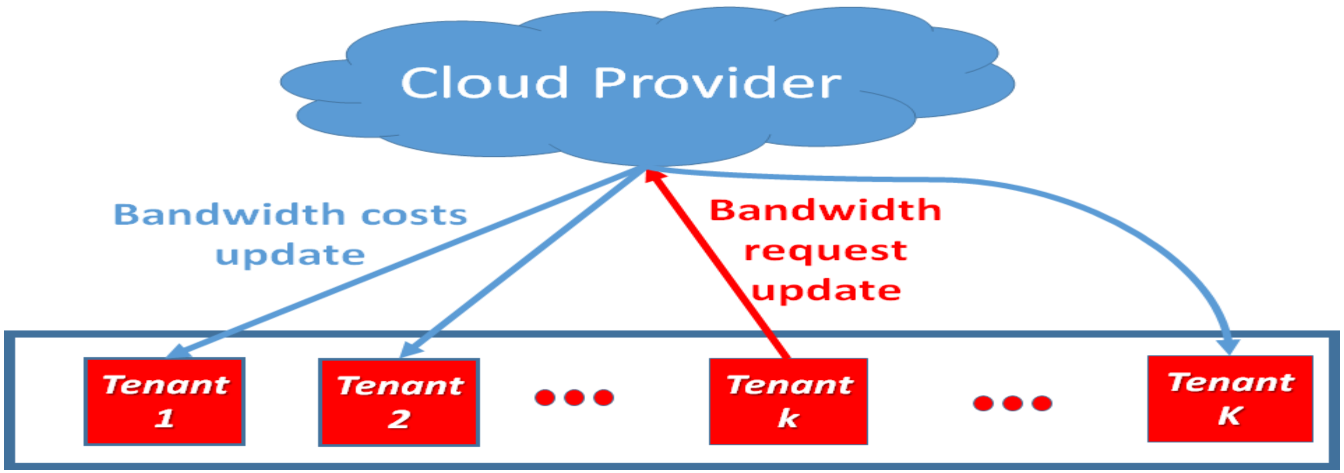
- The bandwidth allocation can be formulated as a strategic game.
- Tenants attempt to maximize their individual profits, i.e., utility function.
- We show that the Nash equilibrium exists for this strategic game.
- We also prove that the Nash equilibrium is equal to the unique socially optimal bandwidth schedule.

Strategic Game

- Theorem 1. The socially optimal bandwidth schedule is a Nash equilibrium, and vice versa.
- Lemma 1. The socially optimal bandwidth schedule is unique.
- Theorem 2. The Nash equilibrium bandwidth schedule is unique.

Best Response Update Process

- The tenants do not reveal their private information such as utility function to the cloud provider.
- It results in decentralized bandwidth schedule optimization.
- At each step only one tenant (randomly chosen) updates its bandwidth schedule: $$\mathbf{b}_k^{m+1} = arg \max_{\mathbf{b}_k \in \mathcal{B}_k} \mathcal{F}_k(b_{-k}^m, b_k)$$



Theorem: The best response update process converges to a socially optimal bandwidth schedule.

Best Response Update Process

- Lemma 2. At each update step, the optimal bandwidth allocation of the updating tenant is unique.

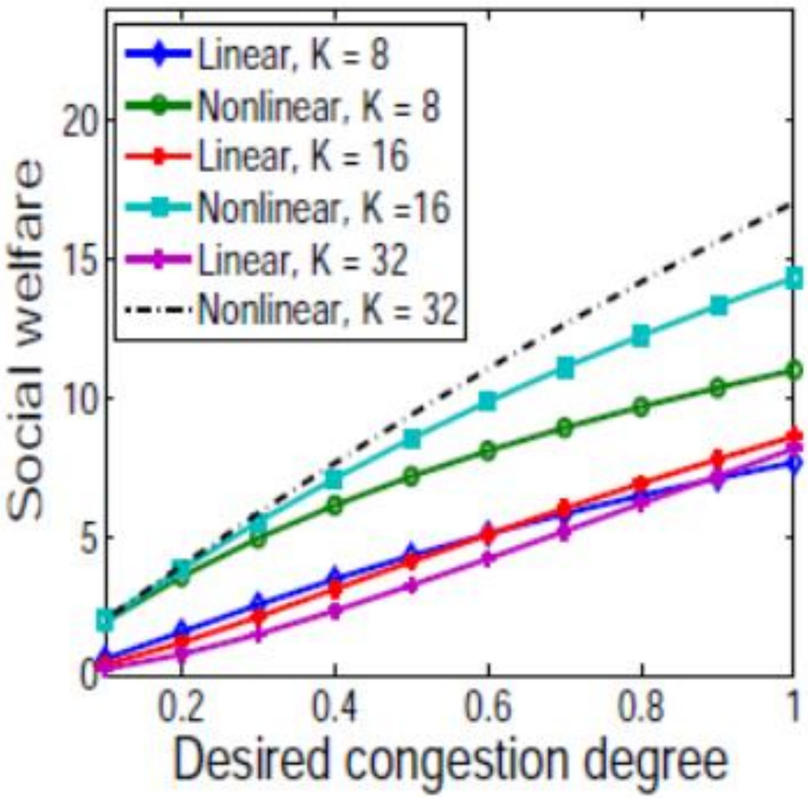
- Theorem 3. The asynchronous update process converges to the socially optimal bandwidth schedule.

Performance Evaluation

- Bandwidth pricing policy
  - K tenants (K varies in the experiment).
  - 2 VMs connected with each other on the same link l.
  - data sent from one VM to another VMs through link l.
  - $B_l^{max} = 40Gbps$
  - two cases: homogeneous tenants satisfaction function $\mathcal{U}_k(b_k) = log(1 + b_k)$ and heterogeneous tenants $\mathcal{U}_k(b_k) = log\left(1 + (1 + (\frac{k}{K})b_k)\right)$
  - overload cost function $\mathcal{C} = \gamma(x)^2$

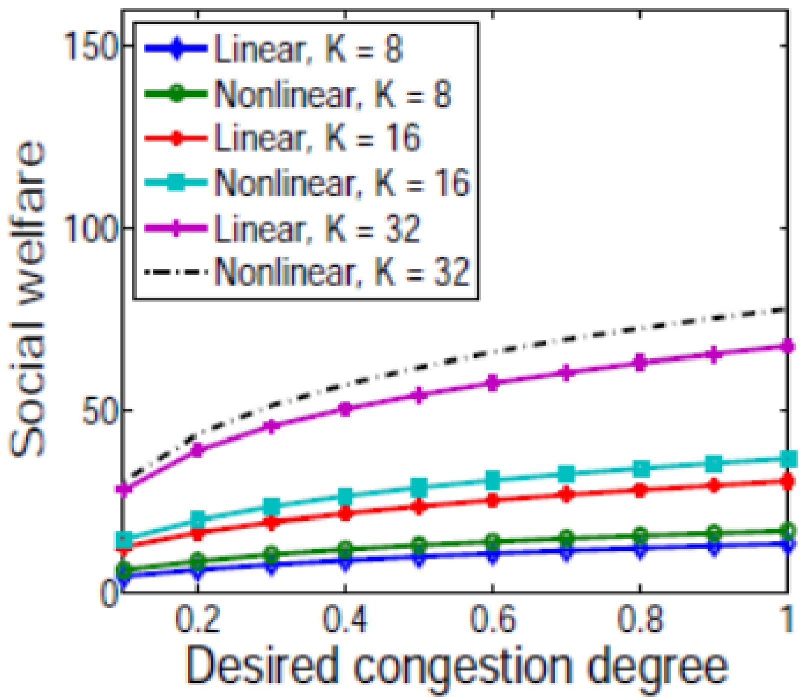
- Comparison
  - Linear congestion cost function, $\mathcal{V}_1(x) = \alpha(\frac{x}{B_l^{max}})$
  - Nonlinear congestion cost function, $\mathcal{V}_2(x) = \beta\left(\frac{x}{B_l^{max}}\right)^2$
  - The cloud provider determines the cost factors, α and β, to achieve a specified congestion degree on link l.

Performance Evaluation



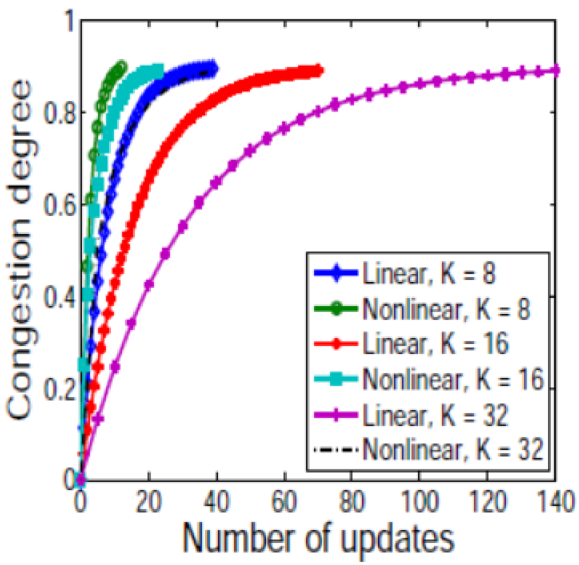
- Social welfare versus desired congestion degree results for homogeneous tenants
- Desired congestion degree is defined as desired bandwidth usage on the link divided by the link capacity.
- # of tenants K varies from 8 to 32.
- Observations: the social welfare is higher with nonlinear pricing model than with linear pricing model

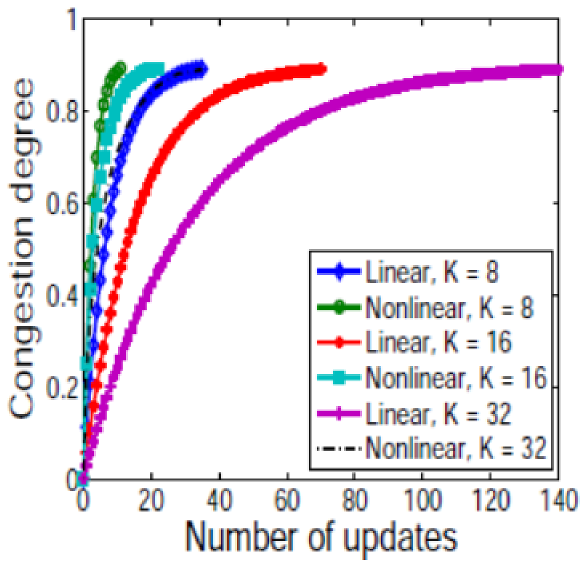
Performance Evaluation



Social welfare versus desired congestion degree results for heterogeneous tenants

Performance Evaluation





Desired congestion degree versus number of updates for homogeneous tenants

Desired congestion degree versus number of updates for heterogeneous tenants

Both homogeneous and heterogeneous converge very fast!!!

Conclusion

- We proposed nonlinear bandwidth pricing for congestion control in a communication link in the cloud networks.
- We defined the social welfare of the tenants as the total satisfaction functions of the tenants minus the congestion cost over the link. We formulated the utility optimization of the tenants as a strategic game.
- We showed different properties of the Nash equilibrium of the designated game. We implemented a decentralized asynchronous-based best response strategy to find the Nash equilibrium bandwidth schedule.
- In the future, we will consider the scenario that the tenants share multiple links in the network.

Thank you!

Questions & Comments?

Dr. Haiying Shen

hs6ms@Virginia.edu

Associate Professor

Pervasive Communication Laboratory

University of Virginia

Strategic Game

- The bandwidth allocation can be formulated as a strategic game, $< \mathcal{K}, \mathbf{b}_k, \mathcal{F}_k >$.
- Tenants attempt to maximize their individual profits, i.e., utility function.
- We show that the Nash equilibrium exists for this strategic game.
- We also prove that the Nash equilibrium is equal to the unique socially optimal bandwidth schedule.