# Bandwidth-Sharing in Overloaded Networks

Sem Borst (Bell Labs & TU/e),

Regina Egorova (CWI & TU/e), Bert Zwart (Georgia Tech)

CISS 2008 – Special Track
on Optimization of Communication Networks

Princeton, March 19–21, 2008

# Introduction

Bandwidth-sharing networks as considered by Massoulié & Roberts provide natural modeling framework for describing dynamic flow-level interaction among elastic data transfers

Bandwidth-sharing commonly governed by utility optimization principle, e.g., so-called $\alpha$-fair strategies

Two views

- extension of single-node processor-sharing model to network setting with possibly several bottleneck links

- extension of network utility maximization problem to dynamic configuration of flows with finite random sizes

## Introduction (cont'd)

Under Markovian assumptions, Bonald & Massoulié proved that $\alpha$-fair bandwidth-sharing strategies achieve stability provided no individual link is overloaded, extending results of De Veciana, Lee & Konstantopoulos

Various further stability results obtained by several authors,

- Bramson (max-min fairness, general flow sizes, renewal arrivals)
- Massoulié (proportional fairness, phase-type flow sizes)
- Gromoll & Williams (fluid limit, linear and tree topologies, general flow sizes)
- Chiang, Shah & Tang (fluid limit, general flow sizes with bounded support)

# Introduction (cont'd)

Talk focuses on $\alpha$-fair bandwidth-sharing networks where load on one or several of links exceeds capacity

With adequate provisioning, network should not experience overload, or even approach overload, in normal operating conditions

However, even with proper dimensioning, actual traffic volume may strongly fluctuate and exhibit transient surges

Furthermore, understanding of overload behavior plays crucial role in analyzing performance in terms of long transfer delays or low flow throughputs as caused by large queue build-ups

## Introduction (cont'd)

As alluded to above, $\alpha$-fair bandwidth-sharing networks bear resemblance with single-server processor-sharing system

Overload behavior of single-server processor-sharing system first analyzed by Jean-Marie & Robert (1994)

Analysis extended by Altman et al. (2004) to discriminatory processor-sharing discipline, which corresponds to single-node network with weighted $\alpha$-fair strategy

Puha et al. (2006) studied single-server overloaded processor-sharing system in terms of measure-valued processes

# Introduction (cont'd)

Two key distinctions that arise in network scenario

- Rate received by class is no longer constant, but depends on number of flows of all classes in intricate fashion

- Network may show non-work-conserving behavior due to fact that congestion at other links may prevent link from utilizing its full capacity

## Introduction (cont'd)

These two features not only render flow-level performance largely intractable, but also complicate analysis of overload behavior

For example, even on underloaded links, workloads may grow because of non-work-conserving behavior

In addition, exact nature of growth patterns of various classes is far from clear, and may even potentially involve oscillatory effects

In order to characterize growth dynamics, we examine fluid limit, which emerges when number of flows of various classes is scaled in both space and time

# Model description

## Bandwidth-sharing network

- **Links labeled by** $j = 1, \ldots, J$

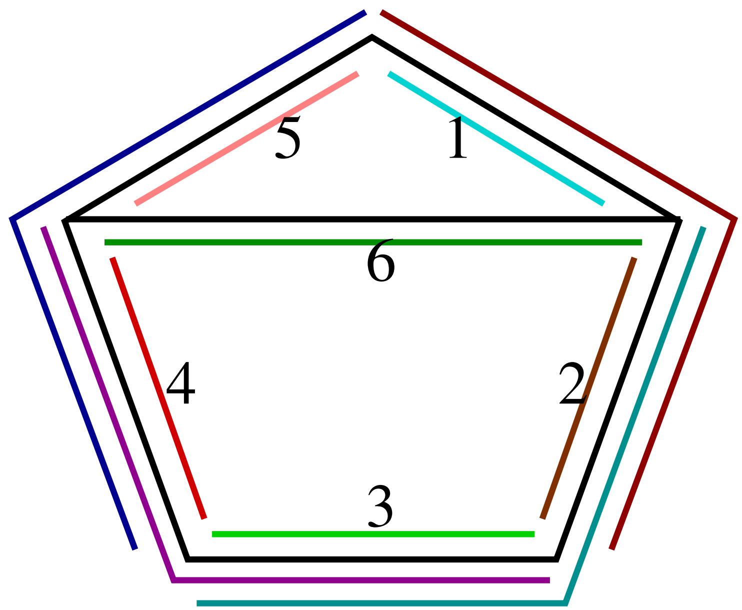
- $C = (C_1, \ldots, C_J)$ **is vector of link capacities**

- **Traffic classes indexed by** $i = 1, \ldots, I$

- **Each class is characterized by route, i.e., subset of** $\{1, \ldots, J\}$**, which represents links traversed by traffic from that class**

- $A$ **is** $J \times I$ **incidence matrix such that** $A_{ji} = 1$ **if link** $j$ **belongs to route of class** $i$**, and** $A_{ji} = 0$ **otherwise**

## Model description (cont'd)




J = 6  I = 10

## Model description (cont'd)

Traffic of various classes consists of elastic data transfers

Duration of flow depends on its size and simultaneous service rate it receives on all links along its route

Class-$i$ flows arrive as renewal process with rate $\lambda_i$ i.e., mean interarrival time is $1/\lambda_i$, and have generally distributed sizes $B_i$ with mean $1/\mu_i$

$\rho = (\rho_1, \ldots, \rho_I)$ with $\rho_i := \lambda_i/\mu_i$ is vector of traffic intensities

## Rate allocation

$\Lambda(z) = (\Lambda_1(z), \ldots, \Lambda_I(z))$ is vector of rates allocated to various classes based on population $z = (z_1, \ldots, z_I)$ of active flows

Any rate allocation vector $\Lambda$ must satisfy capacity constraints $A\Lambda \leq C$

Within each class, rate is fairly shared among all competing flows, i.e., if $z_i > 0$, then each class-$i$ flow receives service at rate $\Lambda_i(z)/z_i$

Thus, class-$i$ flow that is continuously active throughout time interval $[s, t]$, receives cumulative amount of service

$$S_i(s,t) = \int_s^t \frac{\Lambda_i(Z(u))}{Z_i(u)} du,$$

with $Z(t) = (Z_1(t), \ldots, Z_I(t))$ representing population of active flows at time $t$

# Rate allocation (cont'd)

**Weighted $\alpha$-fair strategy is parameterized by fairness coefficient $\alpha \in (0, \infty)$ and weight vector $w = (w_1, \ldots, w_I) \in \mathrm{R}_I^+$**

**For given population $z = (z_1, \ldots, z_I) \neq (0, \ldots, 0)$ of active flows, weighted $\alpha$-fair rate allocation $\Lambda(z)$ is uniquely determined by solution to optimization problem:**

$$(P) \quad \begin{array}{ll} \text{maximize} & G_z(\Lambda) \\ \\ \text{subject to} & A\Lambda \leq C,\ \Lambda \geq 0, \end{array}$$

**where**

$$G_z(\Lambda) = \begin{cases} \sum_{i=1}^{I} w_i\, z_i^{\alpha}\, \frac{\Lambda_i^{1-\alpha}}{1-\alpha}, & \alpha \in (0, \infty)/\{1\}, \\ \\ \sum_{i=1}^{I} w_i\, z_i\, \log \Lambda_i, & \alpha = 1 \end{cases}$$

## Rate allocation (cont'd)

**Family of $\alpha$-fair bandwidth-sharing strategies includes several common fairness concepts as special cases**

**In particular, case $\alpha = 1$ and limiting cases $\alpha \to 0$ and $\alpha \to \infty$ correspond to rate allocation that is proportional fair, achieves maximum throughput, and is max-min fair, respectively**

## Stability condition

Under Markovian assumptions, Bonald & Massoulie (2001) proved that $\alpha$-fair strategies achieve stability provided no individual link is overloaded, i.e., $A\rho < C$

Talk focuses on overload scenario where above condition is violated for at least one of the links

While total number of flows must grow in such a scenario, exact nature of growth patterns of various classes is not so evident

We will characterize growth dynamics in terms of fluid limit

# Fluid limit

**Fluid limit emerges from sequence of processes** $(Z(rt)/r, t \geq 0)$ **scaled in both space and time**

**Nonnegative continuous function** $(z(t), t \geq 0)$ **is said to be fluid-model solution if it satisfies**

$$z_i(t) = \lambda_i \int_0^t \mathbf{P}(B_i > S_i(s,t))ds, \tag{1}$$

**where** $S_i(s,t) = \int_s^t \frac{R_i(z(u))}{z_i(u)} du$ **is cumulative amount of service received by class-**$i$ **flow during time interval** $[s,t]$

$R_i(z)$ **represents slightly modified version of rate allocation functions which may be interpreted as service rates at fluid scale:** $R_i(z) \equiv \Lambda_i(z)$ **if** $z_i > 0$**, where** $\Lambda_i(\cdot)$ **is solution of optimization problem (P), and** $R_i(z) \equiv \rho_i$ **if** $z_i = 0$

**Above distinction reflects fact that on fluid scale,** $z_i = 0$ **requires that class** $i$ **receives service at rate** $\rho_i$**, rather than 0**

## Main result

### Theorem

**Fluid-model equation has linear solution** $(z_1(t), \ldots, z_I(t)) = (m_1, \ldots m_I)t$**, where vector** $(m_1, \ldots, m_I)$ **forms unique solution to fixed-point equation**

$$R_i(m) = \rho_i \mathbf{E}\left[\mathrm{e}^{-\frac{m_i}{R_i(m)} B_i^*}\right], \quad i = 1, \ldots, I, \tag{2}$$

**with** $B_i^*$ **representing residual flow size (i.e. random variable with density** $\mathrm{P}(B_i > x)/\mathbf{E}[B_i]$ **and** $\mathbf{E}\left[\mathrm{e}^{-sB_i^*}\right] = \mu_i(1 - \mathbf{E}\left[\mathrm{e}^{-sB_i}\right])/s$**)**

**Holds for arbitrary network topologies and arbitrary flow size distributions**

## Exponential flow sizes

**In case of exponential flow sizes, Equation (2) reduces to**

$$m_i = \lambda_i - \mu_i R_i(m), \qquad i = 1, \ldots, I, \tag{3}$$

**which makes sense, since $\mu_i R_i(m)$ is indeed departure rate of class-$i$ flows**

**This is also consistent with convention $R_i(m) = \rho_i$ when $m_i = 0$**

## Single-link scenario

**In single-link scenario, i.e., $J = 1$, it reduces to known results for single-server processor-sharing type systems**

**In particular, in single-class case, i.e., $I = 1$, we have, dropping class index, $R(m) = 1$, and Equation (2) specializes to**

$$1 = \rho\, \mathbf{E}[\mathrm{e}^{-mB^*}],$$

**which corresponds to result of Jean-Marie & Robert (1994)**

**In multi-class case, we have $R_i(m) = w_i m_i / \sum_{k=1}^{I} w_k m_k$, and Equation (2) takes form**

$$\frac{w_i m_i}{\sum_{k=1}^{I} w_k m_k} = \rho_i\, \mathbf{E}\left[\mathrm{e}^{-w_i^{-1} \sum_{k=1}^{I} w_k m_k B_i^*}\right], \qquad i = 1, \ldots, I,$$

**which agrees with result of Altman et al. (2004) for overloaded discriminatory processor-sharing queues**

## Heuristic interpretation

**Suppose that** $\frac{Z(r)}{r} \to m$ **a.s. as** $r \to \infty$ **for some vector** $m = (m_1, \ldots, m_I)$

**Then, for large** $t$**, class-**$i$ **flow will receive service at rate of approximately** $\frac{R_i(mt)}{m_i t}$

**Let** $a_i^n$ **be arrival epoch of** $n$**-th class-**$i$ **flow**

**Then size** $B_i$ **of that flow and its sojourn time** $T_i$ **may be related as:**

$$B_i = \int_{a_i^n}^{a_i^n + T_i} \frac{R_i(mu)}{m_i u} du$$

**Since** $R_i(mu) = R_i(m)$**, it follows that**

$$\frac{m_i}{R_i(m)} B_i = \int_{a_i^n}^{a_i^n + T_i} \frac{1}{u} du = \log(a_i^n + T_i) - \log a_i^n$$

# Heuristic interpretation (cont'd)

**Taking exponents on both sides, we obtain**

$$T_i = a_i^n \left( \mathrm{e}^{\frac{m_i}{R_i(m)} B_i} - 1 \right)$$

**Number of active class-$i$ flows at time $t$ may then be expressed as**

$$Z_i(t) = \#\{n : a_i^n + T_i^n \geq t, a_i^n \leq t\} = \#\{n : t \geq a_i^n \geq t\mathrm{e}^{-\frac{m_i}{R_i(m)} B_i}\}$$

**Because $a_i^n \approx n/\lambda_i$ for large $n$, we have**

$$Z_i(t) = \#\{n : t \geq n/\lambda_i \geq t\mathrm{e}^{-\frac{m_i}{R_i(m)} B_i}\} \approx \lambda_i t \left( \mathbf{E} \left[ 1 - \mathrm{e}^{-\frac{m_i}{R_i(m)} B_i} \right] \right)$$

**Dividing both sides by $t$ and letting $t$ tend to infinity, we deduce**

$$m_i = \lambda_i \left( 1 - \mathbf{E} \left[ \mathrm{e}^{-\frac{m_i}{R_i(m)} B_i} \right] \right),$$

**which is equivalent to Equation (2)**

# Uniqueness of fixed-point solution

Crucial role is played by related optimization problem

To formulate this optimization problem, rewrite fixed-point Equation (2) in equivalent form

$$\frac{m_i}{R_i(m)} = \beta_i^{-1}\left(\frac{R_i(m)}{\rho_i}\right), \qquad i = 1, \ldots, I, \tag{4}$$

where $\beta_i^{-1}(\cdot)$ is inverse of Laplace-Stieltjes Transform (LST) $\beta_i(y) = \mathbf{E}\left[\mathrm{e}^{-yB_i^*}\right]$, $\beta_i^{-1}(\beta_i(y)) = y$

Observe that right-hand side only depends on $m$ through $R_i(m)$

Now introduce function $H_i(x)$, with derivative $H_i'(x) = \left(\beta_i^{-1}\left(\frac{x}{\rho_i}\right)\right)^\alpha$

Since $\beta_i(x)$ is strictly decreasing in $x$, its inverse is strictly decreasing in $x$ as well

Consequently, $H_i(x)$ is strictly concave in $x$

## Uniqueness of fixed-point solution (cont'd)

**Now, consider optimization problem**

$$(Q) \quad \begin{array}{ll} \text{maximize} & H(R) = \sum_{i=1}^{I} w_i H_i(R_i) \\ \text{subject to} & AR \le C,\ R \ge 0 \end{array}$$

**This optimization problem is strictly concave, and hence has unique solution** $R = (R_1, \ldots, R_I)$

# Uniqueness of fixed-point solution (cont'd)

**Proposition**

**Rate allocation vector $R(m) = (R_1(m), \ldots, R_I(m))$ satisfying Equation (4) is unique solution of optimization problem (Q)**

**Proof sketch:**

**First consider Karush-Kuhn-Tucker (KKT) necessary conditions for problem (P)**

**If $\Lambda(z)$ is optimal solution to problem (P), then there exist constants $p_j(z) \geq 0$ such that**

$$w_i \left( \frac{z_i}{\Lambda_i(z)} \right)^{\alpha} = \sum_{j=1}^{J} p_j(z) A_{ji}, \quad \text{if } z_i > 0, \tag{5}$$

$$p_j(z) \left( \sum_{i=1}^{I} A_{ji} \Lambda_i(z) - C_j \right) = 0, \quad j = 1, \ldots, J \tag{6}$$

**Now consider KKT sufficient conditions for problem (Q)**

**Feasible solution** $R^* = (R_1^*, \ldots, R_I^*)$ **is global optimum, if there exist constants** $p_j^* \geq 0$ **such that**

$$w_i \left( \beta_i^{-1} \left( \frac{R_i^*}{\rho_i} \right) \right)^{\alpha} = \sum_{j=1}^{J} p_j^* A_{ji}, \qquad i = 1, \ldots, I, \tag{7}$$

$$p_j^* \left( \sum_{i=1}^{I} A_{ji} R_i^* - C_j \right) = 0, \qquad j = 1, \ldots, J \tag{8}$$

**By definition,** $R_i(m) \equiv \Lambda_i(m)$ **when** $m_i > 0$

**It then follows from KKT necessary conditions (5)–(6) for problem (P) that there exist constants** $p_j(m) \geq 0$ **such that**

$$w_i \left( \frac{m_i}{R_i(m)} \right)^{\alpha} = \sum_{j=1}^{J} p_j(m) A_{ji}, \qquad \textbf{if } m_i > 0, \tag{9}$$

$$p_j(m) \left( \sum_{i: m_i > 0} A_{ji} R_i(m) - C_j \right) = 0, \qquad j = 1, \ldots, J \tag{10}$$

**Now recall that** $R_i(m) \equiv \rho_i$ **when** $m_i = 0$

**Observing that** $\sum\limits_{i=1}^{I} A_{ji} R_i(m) \leq C_j$**,** $j = 1, \ldots, J$**, and writing** $\sum\limits_{i=1}^{I} A_{ji} R_i(m) = \sum_{i:m_i>0} A_{ji} R_i(m) + \sum_{i:m_i=0} A_{ji}\rho_i$**, it may be inferred that if** $m_i = 0$ **then for any** $j$ **with** $A_{ji} = 1$**, we have strict inequality**

$$\sum_{i:m_i>0} A_{ji} R_i(m) < C_j, \qquad (11)$$

**and thus Equation (10) forces** $p_j(m) = 0$

**We deduce**

$$\sum_{j=1}^{J} p_j(m) A_{ji} = 0, \quad \textbf{if } m_i = 0, \qquad (12)$$

**and**

$$p_j(m) \left( \sum_{i=1}^{I} A_{ji} R_i(m) - C_j \right) = 0, \quad j = 1, \ldots, J \qquad (13)$$

**Fact that $R(m)$ satisfies fixed-point Equation (4) implies $\beta_i^{-1}\left(\frac{R_i(m)}{\rho_i}\right) = \frac{m_i}{R_i(m)}$ when $m_i > 0$**

**Also, $\beta_i^{-1}\left(\frac{R_i(m)}{\rho_i}\right) = 0$ when $m_i = 0$**

**Substituting latter equalities in Equations (9) and (12), respectively, we obtain**

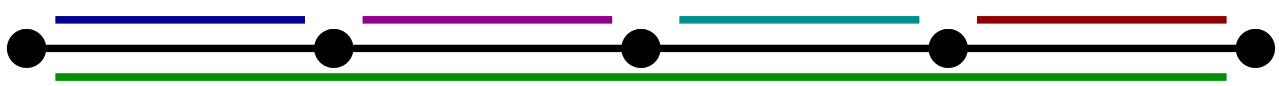
$$w_i \left(\beta_i^{-1}\left(\frac{R_i(m)}{\rho_i}\right)\right)^{\alpha} = \sum_{j=1}^{J} p_j(m) A_{ji}, \quad i = 1, \ldots, I \qquad (14)$$

**Equations (13)–(14) show that $R(m)$ satisfies KKT sufficient conditions (7)–(8) for problem (Q), and hence is global optimum**

## Linear networks

**Consider linear network with links $1, \ldots, L$, each of unit capacity, offered traffic from classes $0, 1, \ldots, L$**

**Class-$j$ flows require service from link $j$ only, $j = 1, \ldots, L$, while class-0 flows demand capacity on all links**



**Linear networks provide useful model for traffic that traverses several links and experiences bandwidth competition from independent cross-traffic**

## Linear networks (cont'd)

**Assume that load on at least one of links exceeds capacity, i.e.,** $\max_{j=1,\ldots,L} \rho_j > 1 - \rho_0$

**Proportional Fair bandwidth-sharing strategy, i.e.,** $G_z(\Lambda) = \sum_{i=0}^{L} z_i \log(\Lambda_i)$

**Capacity constraints take form** $\Lambda_0 + \Lambda_j \leq 1$**,** $j = 1, \ldots, L$

# Asymptotic growth rates

**There exist nonnegative Lagrange multipliers $p_j$ associated with various links so that $m_i$ and $R_i$ together with $p_j$ satisfy**

$$\begin{cases} m_0 = R_0 \sum_{j=1}^{L} p_j, \\ m_j = R_j p_j, \quad j = 1, \ldots, L, \\ p_j(R_0 + R_j - 1) = 0, \quad j = 1, \ldots, L, \end{cases}$$

**in conjunction with $\frac{m_i}{R_i} = \beta_i^{-1}\left(\frac{R_i}{\rho_i}\right)$, yielding**

$$\begin{cases} \beta_0^{-1}\left(\frac{R_0}{\rho_0}\right) = \sum_{j=1}^{L} p_j, \\ \beta_j^{-1}\left(\frac{R_j}{\rho_j}\right) = p_j, \quad j = 1, \ldots, L, \\ p_j(R_0 + R_j - 1) = 0, \quad j = 1, \ldots, L \end{cases} \tag{15}$$

**In total above system provides $2L+1$ equations for $2L+1$ unknown variables $R_i$, $i = 0, \ldots, L$, and $p_j$, $j = 1, \ldots, L$**

## Asymptotic growth rates (cont'd)

In order to solve above equations, consider nonempty subset $\mathcal{J}_+ := \{j : p_j > 0\}$ of links with strictly positive Lagrange multipliers

Observe that $p_j = 0$ means $m_j = 0$: growth rates of classes 0 and $j \in \mathcal{J}_+$ thus correspond to those in scenario with classes $j \notin \mathcal{J}_+$ as well as links $j \notin \mathcal{J}_+$ removed

In particular, when $\mathcal{J}_+ = \{j_+\}$, growth rates of classes 0 and $j_+$ are identical to those in single-node processor-sharing system with classes 0 and $j_+$

## Asymptotic growth rates (cont'd)

**For compactness, denote** $n = \sum_{j \in \mathcal{J}_+} p_j$

**Then solution to system of Equations (15) may be represented as**

$$R_0 = \frac{m_0}{n}; \qquad R_j = 1 - \frac{m_0}{n} \;\; \textbf{if } j \in \mathcal{J}_+; \qquad R_j = \rho_j \;\; \textbf{if } j \notin \mathcal{J}_+;$$

$$p_j = \frac{m_j}{R_j}, \;\; j = 1, \ldots, J$$

**Summing last equality over** $j \in \mathcal{J}_+$**, it follows that** $n = m_0 + \sum_{j \in \mathcal{J}_+} m_j$

## Asymptotic growth rates (cont'd)

**What remains is to determine subset $\mathcal{J}_+$ in terms of system parameters**

**Note that $j \in \mathcal{J}_+$ implies $R_0 + R_j = 1$, and thus necessitates $\rho_0 + \rho_j \geq 1$**

**However, latter inequality is not sufficient for $j \in \mathcal{J}_+$, since it is possible that $m_j = 0$ when other classes at other links sufficiently throttle service rate of class 0**

## Asymptotic growth rates (cont'd)

**In view of symmetry, may assume without loss of generality that links are indexed such that** $\rho_1 \leq \rho_2 \leq \ldots \leq \rho_L$

**Denote by** $\sigma_j$ **common service rate received by classes** $j, \ldots, L$ **in system with links** $j, \ldots, L$ **and classes 0 and** $j, \ldots, L$ **only,** $\sigma_j = 1 - \Lambda_0(m_j, \ldots, m_L) = 1 - \frac{m_0}{m_0 + \sum_{k=j}^{L} m_k}$

**Then subset** $\mathcal{J}_+$ **is of the form** $\{j_+, \ldots, L\}$**, with** $j_+ := \max\{j : \rho_{j-1} \leq \sigma_j\}$

## Interpretation

If $\rho_{j-1} \leq \sigma_j$, then competition from classes $j, \ldots, L$ alone is sufficient to throttle rate of class 0 to an extent that what remains available for classes $1, \ldots, j-1$ exceeds their respective loads, and hence $m_1 = \ldots = m_{j-1} = 0$

This occurs when loads of classes $1, \ldots, j-1$ are relatively low and loads of classes $j, \ldots, L$ are sufficiently high

Note that this may occur even when $\rho_0 + \rho_i > 1$ for some classes $i = 1, \ldots, j-1$

Although these classes rely on service from overloaded links, they remain stable thanks to much stronger competition at other higher-loaded links

In contrast, if $\rho_{j-1} > \sigma_j$, then competition from classes $j, \ldots, L$ alone is not sufficient to provide stability to class $j-1$, and hence $m_{j-1} > 0$

## Exponential flow sizes

**In order to determine growth rates explicitly, we need to specify flow size distributions of various classes that occur in set of fixed-point Equations (4)**

**In case of exponential flow sizes, growth rates of various classes may be represented in terms of single variable $n$ as**

$$m_0 = \lambda_0 - \mu_0 \frac{m_0}{n} = \frac{\lambda_0 n}{\mu_0 + n},$$

$$m_j = \lambda_j - \mu_j \frac{n - m_0}{n} = \lambda_j - \mu_j \left(1 - \frac{\lambda_0}{\mu_0 + n}\right), \qquad j \in \mathcal{J}_+$$

# Exponential flow sizes (cont'd)

**Summing above equations results in**

$$n = \frac{\lambda_0 n}{\mu_0 + n} + \sum_{j \in \mathcal{J}_+} \left( \lambda_j - \mu_j \left( 1 - \frac{\lambda_0}{\mu_0 + n} \right) \right),$$

**which yields quadratic equation**

$$n^2 + \nu n + \kappa = 0, \tag{16}$$

**with**

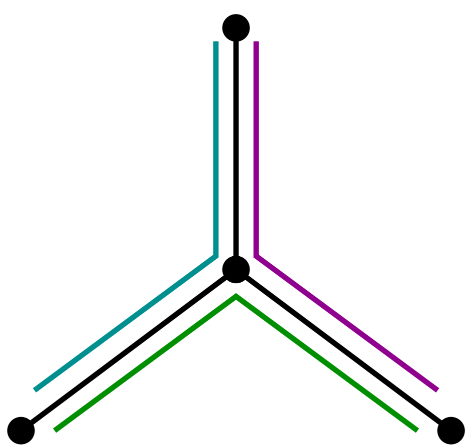
$$\nu := \mu_0 + \sum_{j \in \mathcal{J}_+} \mu_j - \lambda_0 - \sum_{j \in \mathcal{J}_+} \lambda_j,$$

$$\kappa := \sum_{j \in \mathcal{J}_+} \mu_j (\mu_0 - \lambda_0) - \mu_0 \sum_{j \in \mathcal{J}_+} \lambda_j$$

## Star networks

Consider star network composed of $L$ links, each of unit capacity, offered traffic from $L(L-1)/2$ classes labeled as $\{i,j\}$, $i,j = 1,\ldots,L$, $i \neq j$

Route of class $\{i,j\}$ simply consists of two links $i$ and $j$



Star networks offer convenient model for scenarios where core is over-provisioned and congestion predominantly occurs at edge with comparatively low-capacity access links

## Star networks (cont'd)

**Assume that load on at least one of links exceeds capacity, i.e.,** $\max_{j=1,\dots,L} \sigma_j > 1$**, with** $\sigma_j := \sum_{k \neq j} \rho_{\{j,k\}}$

**Proportional Fair bandwidth-sharing strategy, i.e.,** $G_z(\Lambda) = \sum_{j \neq k} z_{\{j,k\}} \log(\Lambda_{\{j,k\}})$

**Capacity constraints take form** $\sum_{k \neq j} \Lambda_{\{j,k\}} \leq 1$**,** $j = 1, \dots, L$

# Asymptotic growth rates

**Lagrange multipliers $p_j$ associated with various links and $R_{\{j,k\}}$ satisfy**

$$\begin{cases} \beta_{\{j,k\}}^{-1}\left(\frac{R_{\{j,k\}}}{\rho_{\{j,k\}}}\right) = R_{\{j,k\}}(p_j + p_k), & j \neq k, \\ p_j(\sum_{k\neq j} R_{\{j,k\}} - 1) = 0, & j = 1, \ldots, L \end{cases}$$

**In total above system provides $L(L+1)/2$ equations for $L(L+1)/2$ unknown variables $R_{\{j,k\}}$, $j \neq k$, and $p_j$, $j = 1, \ldots, L$**

**In case of exponential flow sizes, set of fixed-point equations takes explicit form in Equation (3), yielding**

$$\begin{cases} \lambda_{\{j,k\}} - \mu_{\{j,k\}} R_{\{j,k\}} = R_{\{j,k\}}(p_j + p_k), & j \neq k, \\ p_j(\sum_{k\neq j} R_{\{j,k\}} - 1) = 0, & j = 1, \ldots, L \end{cases}$$

**As before, consider subset of links with strictly positive Lagrange multipliers in order to solve above equations**

## Three-link network

**As an illustrative example, consider triangular network with three links and three classes**

**Distinguish three scenarios, (I), (II) and (III), depending on whether one, two or all three of Lagrange multipliers are strictly positive, respectively**

**With minor abuse of notation, we define $m_i := m_{\{1,2,3\}\setminus\{i\}}$, $R_i := R_{\{1,2,3\}\setminus\{i\}}$, and $\rho_i := \rho_{\{1,2,3\}\setminus\{i\}}$**

**Above equations may then be rewritten as**

$$\begin{cases} m_i = R_i(p_j + p_k), \quad \{i,j,k\} = \{1,2,3\}, \\ p_j(R_i + R_k - 1) = 0, \quad \{i,j,k\} = \{1,2,3\} \end{cases}$$

**For compactness, denote** $m := m_1 + m_2 + m_3$**, and** $n = m_i + m_j$

**Solutions in above three scenarios may be then represented as**

$$(I) \qquad \left(R_i, R_j, R_k\right) = \left(\tfrac{m_i}{n}, \tfrac{m_j}{n}, \rho_k\right),$$

$$p_i = p_j = 0, \quad p_k = n,$$

$$(II) \qquad \left(R_i, R_j, R_k\right) = \left(\frac{m_i}{m}, \frac{m_j + m_k}{m}, \frac{m_j + m_k}{m}\right),$$

$$p_i = 0, \quad p_j = \frac{m_j}{m_j + m_k} m, \quad p_k = \frac{m_k}{m_j + m_k} m,$$

$$(III) \qquad R_1 = R_2 = R_3 = \tfrac{1}{2},$$

$$p_i = \textstyle\sum_{j \neq i} m_j - m_i > 0, \quad i = 1, 2, 3$$

**Reflect trichotomy in behavior of triangular network**

**In case (I) network behaves as single-node processor-sharing system with classes** $i$ **and** $j$ **only**

Conditions for case (I) to occur in terms of system parameters also coincide with corresponding ones in linear network with $|\mathcal{J}_+| = 1$

Case (II) corresponds to case of linear network with $|\mathcal{J}_+| = 2$

Conditions for this case to arise subsume corresponding ones in linear network, but include additional condition that loads of three classes should be slightly unbalanced

If latter condition is violated, i.e., loads of three classes are nearly equal, then case (III) arises, which has no counterpart in linear network

In this case, each of three classes behaves as in isolated processor-sharing system with capacity $\frac{1}{2}$

## Extensions

- User impatience
- Non-zero initial states