On Fair Bandwidth Allocation in Connection-less Networks

B. Behsaz, P. Gburzynski, M. MacGregor
... is beautiful. It has survived the massive assault by ATM and is doing fine! This should be a lesson:
Connectionless ...

... is beautiful. It has survived the massive assault by ATM and is doing fine! This should be a lesson:

Don't fix what ain't broke!
... is beautiful. It has survived the massive assault by ATM and is doing fine! This should be a lesson:

Don't fix what ain't broke!

Many attempts at fixing the Internet address the general issue of QoS
Connectionless ...

... is beautiful. It has survived the massive assault by ATM and is doing fine! This should be a lesson:

Don't fix what ain't broke!

Many attempts at fixing the Internet address the general issue of QoS

Not all QoS “problems” are real problems: those that will go away when you add (global) bandwidth are not.
... is beautiful. It has survived the massive assault by ATM and is doing fine! This should be a lesson:

Don't fix what ain't broke!

Many attempts at fixing the Internet address the general issue of QoS

Not all QoS “problems” are real problems: those that will go away when you add (global) bandwidth are not.

The real ones can be redefined as fairness problems: do we know what fairness means?
What is a flow?

A TCP session? We know how to handle those.
What is a flow?

A TCP session? We know how to handle those.
Fairness among flows

What is a flow?

A TCP session? We know how to handle those.
What is a flow?

A TCP session? We know how to handle those.

But they are not synonymous with users.
Fairness among flows

What is a flow?

A TCP session? We know how to handle those.

But they are not synonymous with users ...

... and there is UDP traffic, too
Fairness among flows

What is a flow?

A TCP session? We know how to handle those.

But they are not synonymous with users ...

... and there is UDP traffic, too

... and IP streams like (Host, Port) (Host, Port) are not synonymous with users, either
What is a flow?

A TCP session? We know how to handle those.

But they are not synonymous with users ...

... and there is UDP traffic, too

... and IP streams like (Host, Port) (Host, Port) are not synonymous with users, either

In order to begin talking about fairness, we need some identification of the entity (user) being a meaningful recipient of fair/unfair treatment.
Note that host-based ...

... approaches will not do ...
Note that host-based ... 

... approaches will not do ... because of the possible collusion scenarios (like in BitTorrent):
A flux is a **ranked** set of flows grouped together by a router as a **unit of treatment**.
This involves a dissemination protocol akin to OSPF, BGP, RSVP ...
Suppose everything is equally ranked, also $R_1$ equally ranks the three hosts; then, at $R_1$ ...

- Each flux of $H_1$ will receive the rank $\frac{1}{(3 \times 3)} = \frac{1}{9}$
- Each flux of $H_2$ will receive the rank $\frac{1}{(3 \times 2)} = \frac{1}{6}$
- Each flux of $H_3$ will receive the rank $\frac{1}{(3 \times 4)} = \frac{1}{12}$
Generally ...

Hosts may diversify the ranks of their fluxes, according to local policies/importance/relevance.

The routers may diversify the ranks of their links.
Generally ...

Hosts may diversify the ranks of their fluxes, according to local policies/importance/relevance.

The routers may diversify the ranks of their links.

The situation is dynamic, not only because the configuration of fluxes changes at hosts ...
Generally ...

Hosts may diversify the ranks of their fluxes, according to local policies/importance/relevance

The routers may diversify the ranks of their links

The situation is dynamic, not only because the configuration of fluxes changes at hosts ...

The disseminated ranks must account for the current fraction of each flux propagated over the respective link
The formula:

\[ \rho(G^k_u) = \frac{\rho(F^j_i) \cdot \gamma_u(F^j_i, G^k_u)}{\sum_u \rho(F^j_i) \cdot \gamma_u(F^j_i, G^k_u)} \]

\( \{ F^j_i \} \) is the collection of all fluxes arriving at the router, i.e., j-th flux from neighbor (link) i.
The formula:

\[
\rho(G_u^k) = \frac{\rho(F_i^j) \cdot \gamma_u(F_i^j, G_u^k)}{\sum_{\psi_u} \rho(F_i^j) \cdot \gamma_u(F_i^j, G_u^k)}
\]

The (input) rank of the flux adjusted by the router's link (neighbor) rank
The formula:

\[
\rho(G_u^k) = \frac{\rho(F_i^j) \cdot \gamma_u(F_i^j, G_u^k)}{\sum_{F_i^j \in \Psi_u} \rho(F_i^j) \cdot \gamma_u(F_i^j, G_u^k)}
\]

\(G_u^k\) is the collection of all outgoing fluxes, i.e., k-th flux going to neighbor (link) u.
The formula:

\[
\rho(G^k_u) = \frac{\rho(F^j_i) \cdot \gamma_u(F^j_i, G^k_u)}{\sum_{F^j_i \in \Psi_u} \rho(F^j_i) \cdot \gamma_u(F^j_i, G^k_u)}
\]

The normalized fraction of flux \( F^j_i \) contributing to flux \( G^k_u \); for a given \( F^j_i \):

\[
\sum_{u=1}^{m} \gamma_u(F^j_i, G^k_u) = 1
\]
The formula:

\[ \rho(G^k_u) = \frac{\rho(F^j_i) \cdot \gamma_u(F^j_i, G^k_u)}{\sum_{F^j_i \in \Psi_u} \rho(F^j_i) \cdot \gamma_u(F^j_i, G^k_u)} \]

Set of all (input) fluxes contributing anything at all to link \( u \)
The formula:

\[
\rho(G^k_u) = \frac{\rho(F^j_i) \cdot \gamma_u(F^j_i, G^k_u)}{\sum_{F^j_i \in \Psi_u} \rho(F^j_i) \cdot \gamma_u(F^j_i, G^k_u)}
\]

Propagated rank of \textit{k}-th flux output on link \textit{u}
Maintaining fairness

We look at a single output link $u$ and all the fluxes feeding into it

\[ \lambda_i^j \] - the (monitored) arrival rate of flux $F_i^j$

\[ x_u^k = y_u(F_i^j, G_u^k) \times \lambda_i^j \] - the offered rate of $F_i^j$ on $u$

\[ B_u \] - the bandwidth of $u$

We only have a problem when:

\[ \sum_{k=1}^{n_u} x_u^k > B_u \]
Identifying "unfair" fluxes

\[
C = \{1, ..., n_u\}; \\
\]
\[
s = n_u; \\
\]
\[
B^\text{cont}_u = B_u; \\
\]
\[
\text{while } (\exists k \in C | x^k_u < B_u \cdot \rho(G^k_u)) \{ \\
\]
\[
C = C - \{k\}; \\
\]
\[
B^\text{cont}_u = B^\text{cont}_u - x^k_u; \\
\]
\[
s = s - 1; \\
\}
\]
Identifying “unfair” fluxes

\[
\begin{align*}
C &= \{1, \ldots, n_u\}; \\
\quad &s = n_u; \\
B_u^{cont} &= B_u; \\
\text{while } (\exists k \in C | x_u^k < B_u \cdot \rho(G_u^k)) \{ \\
\quad &C = C - \{k\}; \\
\quad &B_u^{cont} = B_u^{cont} - x_u^k; \\
\quad &s = s - 1; \\
\} 
\end{align*}
\]
Identifying “unfair” fluxes

\[ C = \{1, \ldots, n_u\}; \]
\[ s = n_u; \]
\[ B_u^{\text{cont}} \equiv B_u; \]
while (\exists k \in C \mid x_u^k < B_u \cdot \rho(G_u^k)) \{ \]
\[ C = C - \{k\}; \]
\[ B_u^{\text{cont}} = B_u^{\text{cont}} - x_u^k; \]
\[ s = s - 1; \]
\}
Identifying “unfair” fluxes

\[ C = \{1, \ldots, n_u\}; \]
\[ s = n_u; \]
\[ B_u^{\text{cont}} = B_u; \]
\[ \text{while } (\exists k \in C | x_u^k < B_u \cdot \rho(G_u^k)) \{ \]
\[ C = C - \{k\}; \]
\[ B_u^{\text{cont}} = B_u^{\text{cont}} - x_u^k; \]
\[ s = s - 1; \]
\[ \} \]

fair share
The problem

We are given a bunch of fluxes (maybe just one), offering some load $x_i$, the fair share being $b_i$, $b_i < x_i$.

The question is “how much to actually allocate”?

Our allocation should carry a certain universal message amounting to this:

The more you are going to exceed your fair share, the less of an actual bandwidth you are going to get.

Consequently, everybody's objective will be to navigate towards the fair share.
FMB functions:

Fair Maximum Bandwidth: $M(x_i, b_i)$ is a function with these properties:

\[
\frac{x_i}{b_i} = \frac{x_j}{b_j} \quad \Rightarrow \quad \frac{M(x_i, b_i)}{b_i} = \frac{M(x_j, b_j)}{b_j}
\]

\[
\frac{x_i}{b_i} > \frac{x_j}{b_j} \quad \Rightarrow \quad \frac{M(x_i, b_i)}{b_i} \leq \frac{M(x_j, b_j)}{b_j}
\]

Example:

\[
M(x, b) = x \times \left( \frac{b}{x} \right)^\alpha
\]

for $\alpha > 1$
## Test fluxes

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rational (RA)</strong></td>
<td>[b \geq x \Rightarrow x = x + 1]</td>
</tr>
<tr>
<td></td>
<td>[(b &lt; x) \land (x/b \leq V_{RA}) \Rightarrow x = x - 1]</td>
</tr>
<tr>
<td></td>
<td>[(b &lt; x) \land (x/b &gt; V_{RA}) \Rightarrow x = x/2]</td>
</tr>
<tr>
<td><strong>Near-rational (NR)</strong></td>
<td>[b \geq x \Rightarrow x = x + 1]</td>
</tr>
<tr>
<td></td>
<td>[b &lt; x \Rightarrow x = x/2]</td>
</tr>
<tr>
<td><strong>Greedy (GR)</strong></td>
<td>[b \geq x \Rightarrow x = d]</td>
</tr>
<tr>
<td></td>
<td>[b &lt; x \Rightarrow x = x + d - b]</td>
</tr>
<tr>
<td><strong>Near-greedy (NG)</strong></td>
<td>[(b &lt; x) \land (x/b \leq V_{NG}) \Rightarrow x = x + d - b]</td>
</tr>
<tr>
<td></td>
<td>[(b &lt; x) \land (x/b &gt; V_{NG}) \Rightarrow x = x/2]</td>
</tr>
</tbody>
</table>
Offered loads
Offered loads

<table>
<thead>
<tr>
<th>Flux</th>
<th>Src</th>
<th>Type</th>
<th>Dmnd</th>
<th>Inc Rank</th>
<th>Lnk1 Fract</th>
<th>Lnk2 Fract</th>
<th>Lnk3 Fract</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>RA</td>
<td>3000</td>
<td>0.3</td>
<td>-</td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>GR</td>
<td>4000</td>
<td>0.5</td>
<td>-</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>NG</td>
<td>3000</td>
<td>0.2</td>
<td>-</td>
<td>1.0</td>
<td>0.0</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>NR</td>
<td>5000</td>
<td>0.5</td>
<td>1.0</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>NG</td>
<td>5000</td>
<td>0.5</td>
<td>0.0</td>
<td>-</td>
<td>1.0</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>RA</td>
<td>2500</td>
<td>0.4</td>
<td>0.0</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>NR</td>
<td>2500</td>
<td>0.3</td>
<td>1.0</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>GR</td>
<td>2500</td>
<td>0.1</td>
<td>1.0</td>
<td>0.0</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>NG</td>
<td>2500</td>
<td>0.2</td>
<td>0.0</td>
<td>1.0</td>
<td>-</td>
</tr>
</tbody>
</table>
Treatment
Final comments

The issue of ranks ... It makes little sense to keep them fixed. They should be used to maximize the utility function.

The scheme doesn't have to be introduced everywhere at once (non-compliant hosts/routers can be treated as single-fluxed with some default rank).

Note: we essentially get all the benefits of the connection-oriented paradigm with none of its flaws!! E.g., an important session can be treated as a single (high-ranked) flux.