FCP: A Flexible Transport Framework for Accommodating Diversity

Dongsu Han (KAIST)
Robert Grandl†, Aditya Akella†, Srinivasan Seshan*
† University of Wisconsin-Madison
*Carnegie Mellon University
Evolution of Transport Protocols

- Reliability/Loss recovery
- In-order delivery
- Flow control
- Congestion control

Transport
Network (IP)
Evolution of Transport Protocols

Transport

Quality of Experience (QoE)
Evolu;on of Transport Protocols

Support for diversity:
Multi-homing, multiple paths, multiple sub-streams, multimedia streaming, Web applications, Data transfer with a deadline
Evolution of Transport Protocols

- Reliability/Loss recovery
- In-order delivery
- Flow control
- Congestion control

Transport
Evolution of Transport Protocols

Resource Allocation

Congestion control

Coordination

Fairness and efficiency
Problems in Supporting Diversity in Congestion Control

Cannot ensure coexistence.
1. End-point flexibility: Purely end-point based
2. Coexistence: Invariant for fairness
   (TCP friendliness)
End-point based vs. Router-Assisted

High Flexibility, Diversity

End-point based [TCP]

Can we achieve the best of both worlds?

Router-Assisted [XCP, RCP]

Feedback on network’s state

High Efficiency
Our Approach

1. Decouple coexistence issues (fairness and efficiency) from end-point control

2. Introduce generic abstractions for resource allocation.

Network's state (feedback)

End-point's future behavior (feed-forward)
Decoupling for Flexibility

Sender can distribute its budget to its flows.

Flow $i$’s budget $w_i$, subject to $W \geq \Sigma w_i$

Flow $i$’s rate:

$$R_i = \frac{\text{budget/price}}{= w_i}{P_i} (\text{Byte/sec})$$

Flow $i$’s price: $P_i (\$/\text{Byte})$
Decoupling for Flexibility

1. Flexibility at the end-points in how its budget is used.

2. Flexibility in network price generation.

Invariant for fairness

Flow i’s budget $w_i$, subject to $W \geq \Sigma w_i$

Budget $W$ ($$/sec$$)
• Feedback: “congestion price” reflecting the “cost” of sending data across the link [Kelly]
• Feedback: “congestion price” reflecting the “cost” of sending data across the link [Kelly]

Sender updates the rate.

\[ Rate = \frac{\text{budget}}{\text{price}} \]

\[
\text{Price feedback} = p_3 + p_2 + p_1
\]

This implements proportional fairness [Kelly].
Router updates the price, upon packet reception.

Price ($/byte) = f(Average recent load)

Averaging window = 2 x AvgRTT
Router updates the price, upon packet reception.

Instantaneous Incoming budget ($)

= Price ($/byte) x Bytes Received (bytes)

Price ($/byte) = \frac{\text{Average Incoming Budget}}{\text{Remaining Link Capacity}}

Averaging window = 2 \times \text{AvgRTT}
Feedback: Pricing

Router stores recent history of price, $price(t)$

Instantaneous
Incoming budget ($$)

\[ price(t - rtt) \cdot size \cdot dt \]

New Price
Round Trip Time (rtt)

Incoming budget
\[ l(t) = \int price(t - rtt) \cdot bytes(t) \, dt \]

Time

Price ($$/byte) = \frac{\text{Average Incoming Budget}}{\text{Remaining Link Capacity}}
Feed-forward: Preloading

Sender

Rate = \(1/p\) Bps

Price: \(p\$\)/byte

Rate = \(10/p\) Bps

1 ($/sec) \rightarrow 10 ($/sec)
Feed-forward: Preloading

Sender

Rate = 1/p

1 ($/sec) \rightarrow 10 ($/sec)

Rate = 10/p'

Incoming budget ($)

(1+preload) x price(t-rtt) x bytes

Price goes up (from p to p')

Congestion Header

Preload=9

Round Trip Time

Price
Evaluation

1. Basic performance of FCP

Budget W ($/sec)

2. End-point budget allocation.

Flow 1

PRICE

Flow 2

3. Flexibility in network price generation.
Fast Convergence/Accurate Feedback

RTT = \{25\text{ms}, 50\text{ms}, 125\text{ms}, 250\text{ms}, 625\text{ms}\}
Fast Convergence/Accurate Feedback

\[ \text{Ideal} \]

\( \text{RTT} = \{25\text{ms}, 50\text{ms}, 125\text{ms}, 250\text{ms}, 625\text{ms}\} \)
Fast Convergence/Accurate Feedback

RTT = \{25\text{ms}, 50\text{ms}, 125\text{ms}, 250\text{ms}, 625\text{ms}\}
Fast Convergence/Accurate Feedback

RTT = \{25\text{ms}, 50\text{ms}, 125\text{ms}, 250\text{ms}, 625\text{ms}\}
Fast Convergence/Accurate Feedback

RTT = \{25\text{ms}, 50\text{ms}, 125\text{ms}, 250\text{ms}, 625\text{ms}\}
Fast Convergence/Accurate Feedback

Ideal

RTT = \{25\text{ms}, 50\text{ms}, 125\text{ms}, 250\text{ms}, 625\text{ms}\}
Fast Convergence/Accurate Feedback

Graph showing the comparison between Ideal and FCP sending rates over time.
Fast Convergence/Accurate Feedback

Sending Rate (Mbps)

Time

Ideal

RCP

Overloaded when new flows arrive

FCP
Fast Convergence/Accurate Feedback

![Graphs showing comparison of Ideal, RCP, FCP, and XCP sending rates over time.](image)

- **Ideal**: Constant sending rate without any overload issues.
- **RCP**: Overloaded when new flows arrive.
- **FCP**: Dynamic adjusting to new flows but with some fluctuation.
- **XCP**: Fairness AIMD approach showing a balanced response to new flows.

Sending Rate (Mbps) vs. Time
Faster Convergence/Accurate Feedback

FCP’s converges faster and the feedback is more accurate.

More results in the paper:
- Fairness, efficiency
- Local stability
- Effectiveness of preloading
- Overhead of price calculation
End-point Flexibility: Topology

Sender0 (S0)
Budget 1

Sender1 (S1)
Budget 2

Sender2 (S2)
Budget 1

Round-trip latency: 12 ms

250 Mbps link

200 Mbps links

50 Mbps link

Top flow

Btm flow

Receiver0

Receiver1
Budget allocation is up to end-points

![Diagram showing the relationship between budget allocation and S1's throughput.](chart.png)
Budget allocation is up to end-points

<table>
<thead>
<tr>
<th>Equal-budget</th>
<th>Equal-throughput</th>
<th>Max-throughput</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>Top flow</td>
<td>Bottom flow</td>
</tr>
<tr>
<td>S1’s Throughput (Mbps)</td>
<td>Budget of S1’s top flow ($/sec)</td>
<td></td>
</tr>
</tbody>
</table>
End-point Diversity: Background flows

Non-bottleneck Link capacity: 200Mbps

Flow0

50Mbps

100Mbps

Flow2 (Background)

Flow1

Flow3

Link latency: 3ms
Diversity in Network Pricing

• FCP can also support diverse behaviors in network price generation.

• Examples (in the paper)
  – Deadline support [D³ SIGCOMM 11]
  – Aggregate resource allocation in a multi-tenant data-center
  – Stable bandwidth allocation for streaming
  – Multicast congestion control
Deadline flows preload to specify the desired rate

Price ($/byte) = \frac{\text{Average Incoming Budget}}{\text{Remaining Link Capacity}}
Differential Pricing for Deadline Support

- Constant Pricing for Deadline Flows
- Variable Pricing for BE flows
Differential Pricing for Deadline Support

- Constant Pricing for Deadline Flows
- Variable Pricing for BE flows

Price increase

Best effort (BE)

Dead flows

Rejected/BE flows

Sending Rate (Mbps)

Time (sec)
Conclusion

• FCP accommodates diverse behaviors in resource allocation while utilizing explicit feedback.
• FCP maximizes end-point’s flexibility by simplifying the mechanism of coexistence.
• FCP’s explicit feed-back and feed-forward provides a generic interface for efficient resource allocation.