

Internet Service Differentiation using Transport Options: the case for policy-aware congestion control

Panos Gevros
University of Cambridge
panos.gevros@cl.cam.ac.uk

RIPQoS Workshop, Karlsruhe, Germany
27 August 2003



**UNIVERSITY OF
CAMBRIDGE**

Computer Laboratory

Hypothesis

- many networks are **well provisioned** and therefore lightly utilised (high capacity links, delays ≤ 10 ths ms, loss rates rarely exceeding 0%)
- **Observation** : the stability of the network (and the service quality) depends more on long-term **provisioning** decisions and the **interconnection** practices of the service providers than on the **micro-management of transmission rates** at short time-scales
- **Hypothesis** : if the network is *not* congested, users can easily **transmit more aggressively** (than the current standards prescribe) without posing any threat the network or disrupting others
- in many cases **more optimistic (or informed) adaptive strategy** can lead to considerable performance improvements (make better use of the available capacity)
- Open issues :
 - how do we know in **which** networks and **when** this should be used
 - how do we know that users are more aggressive **only** when they should be and **not always** (regardless of remote endpoint location, or network load) ...



Transport Options (1) – the idea

- Think of the **part of a transport protocol** used for the adaptive control of the transmission rate (window) (‘ **congestion control** ’) as a **service** provided by the host OS.
- in practice, it turns out that this part of the OS has **economic significance** (influences bandwidth sharing at bottlenecks at arbitrary places inside the network)
- ..as such it could be a **service offered by the ISP**.

Example: the ‘ownership’ of the IP address (static or dynamic)

Transport Options (1) – the idea

- more than one, adaptive **transmission control behaviours** are available by the OS (called **Transport Options**)
- each behaviour creates a (notional) **service class** (1-to-1 relationship)
- different behaviours (classes) have different **performance expectations**
- there are **no quantitative specifications** in the service description, so differentiation between classes is **relative**
- provider offers service contracts (SLAs) which give the **right to use** a particular class
- provider should be able to **monitor** or **enforce** terms specified in the contract
- the actual **payoff to the user** (in performance terms) from using a particular Transport Option at a certain point in time is **uncertain** (although relative differentiation in the payoffs from different classes should be consistent).

Transport Options (2) - purpose

performance enhancement (in networks where more aggressive transmission behaviour most likely will not cause congestion problems)

Example 1: “relax” congestion control rules for all TCP connections within a corporate **intranet**.

Example 2: In **FreeBSD** TCP used a **very large initial cwnd** for all connections with destinations on the same LAN

```
if ( in_localaddr(inp->inp_faddr) )
    tp->snd_cwnd = mss*ss_fltsz_local ;
else
    tp->snd_cwnd = mss*ss_fltsz ;
```

Transport Options (2) - purpose

commercial exploitation : introduces a new type of service in the contracts (SLAs) offered by ISPs to their customers

Example: **lease IP addresses** to customers (DHCP), have different tariffs for *static* and *dynamic* allocations.

Analogy :

create a **menu** of available transport protocols or transmission control behaviours (which apply to a transport protocol e.g. TCP)

lease Transmission Control Behaviours which apply to all communications to/from certain network addresses as specified in a contract.

easy to offer SLAs which combine Transport Options **usage based charging** (volume)



UNIVERSITY OF
CAMBRIDGE

Computer Laboratory

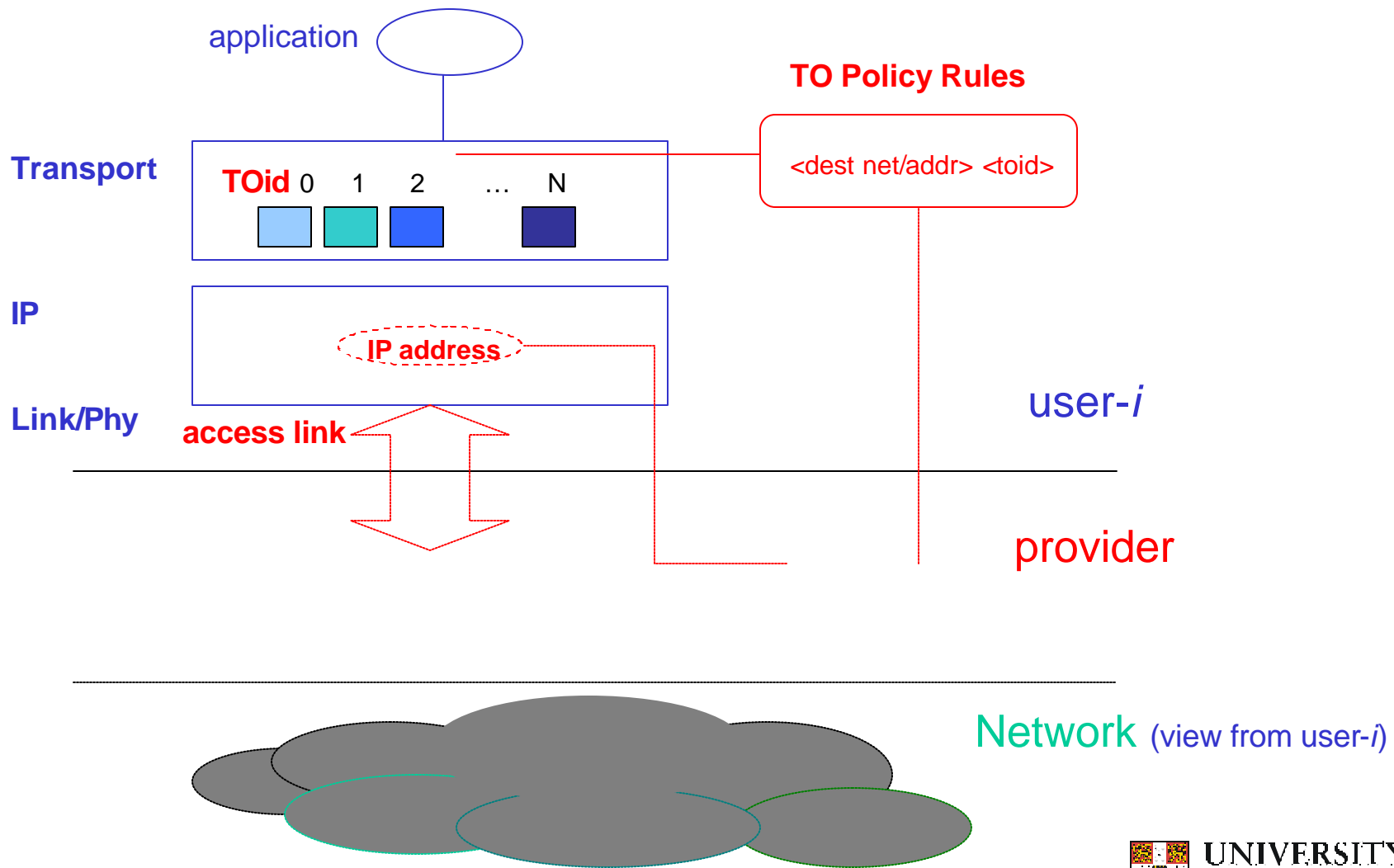
Transport Options (3) - implementation

- each Transport Option is assigned a unique **identifier (TOid)**
- TOid **semantics** are well-defined and compatible across hosts (e.g. TOid =0 corresponds to the default or reference transport)
- TOids are **ordered** :
 - minimum TOid corresponds to the **default** behaviour (e.g. 0 to standard TCP)
 - maximum corresponding to the **most “rewarding”** (in performance expectation terms)
 - the convention is : **greater TOids are preferred**
- use of appropriate TOid **negotiated at connection setup phase** (with TCP **3-way handshake**, similar to MSS negotiation)
- preconfigured **policy rules** at the connection end-points determine which TOid will be used, settle for the minimum of the two

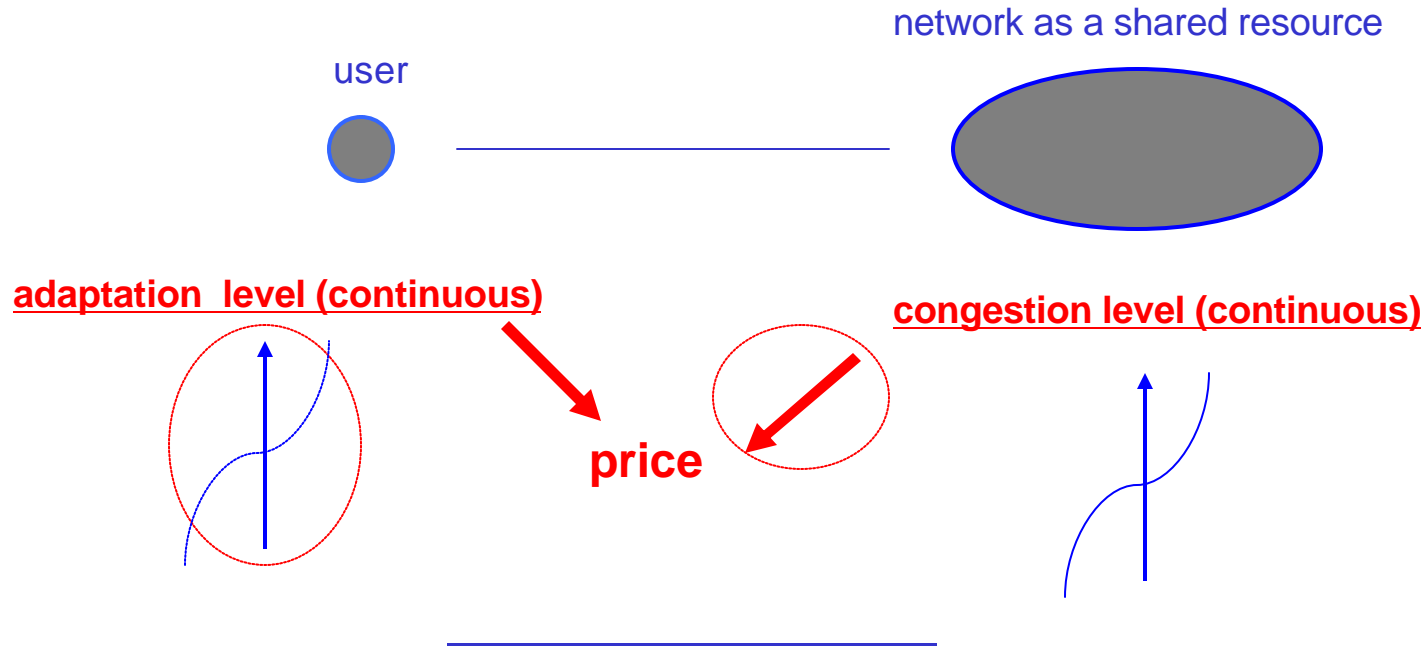
Transport Options (3) - implementation

- preconfigured **policy rules** at the connection end-points determine which TOid will be used, settle for the minimum of the two
- **Policy rules** :: table entries in a **configuration file**, with two parts
 - <condition>** **<action>**
 - <condition>** : **location** attribute e.g. Autonomous System number, network number, or IP address of the remote endpoint.
 - <action>** : **maximum allowed TOid** (range [0.. max])
- Use maximum TOid permitted for the **initialisation** of the connection's TCP control block structure (`cwnd` initialisation and update functions).
- Policy rules will be **explicitly stated in the SLA** and will normally be **static** i.e. will not change during the lifetime of the contract (short-term or long-term SLAs).

Transport Options (3) - implementation

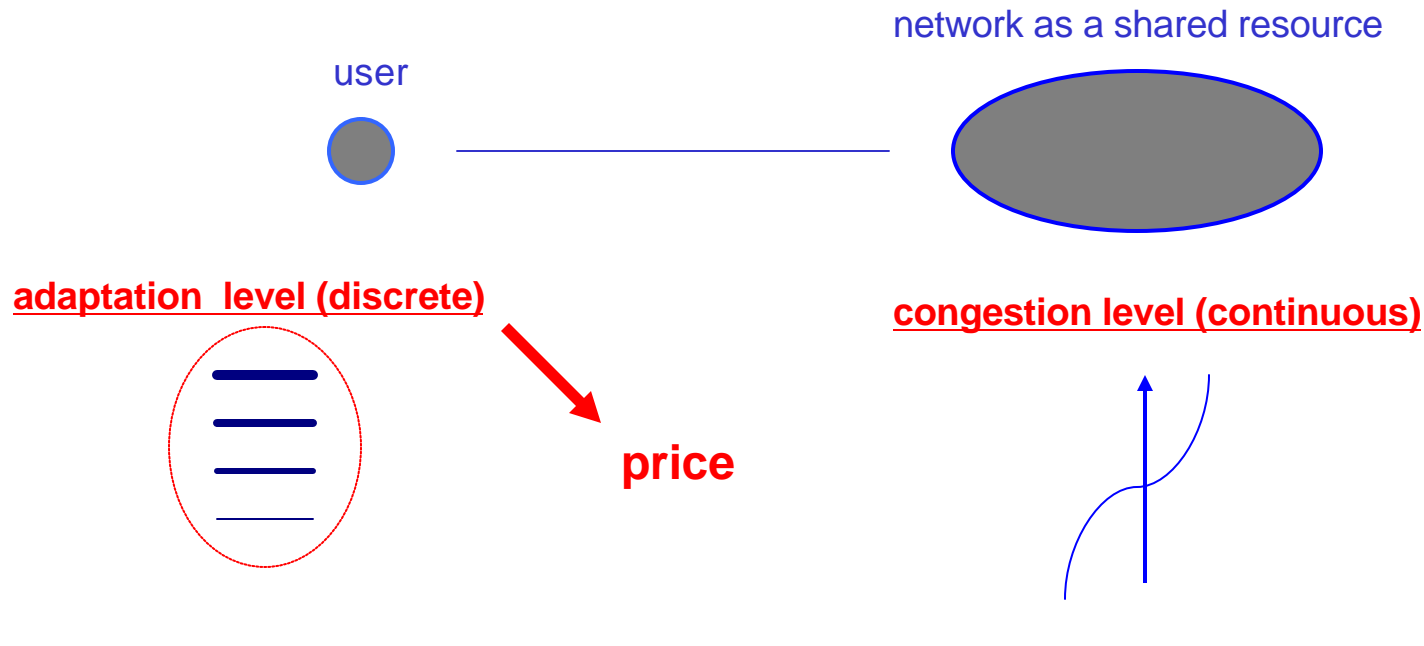


Congestion Pricing vs. Transport Options



special feedback required for conveying the congestion level (i.e. congestion marks -- packet loss is **not** sufficient indication)

Congestion Pricing vs. Transport Options



no special feedback required for conveying information about the congestion level (packet loss is sufficient indication)

Transport Options (4) – evaluation

- benefits to **providers**
 - (claim to be) ‘**deployable**’
 - built in long-term SLAs -> so **easy to manage**,
 - ‘**packaged**’ in a form in which users can buy
 - allow **flexibility in pricing** schemes (e.g. offer subsets of rules or TOid ranges)
 - approach is more **general** (I.e not confined to data transport)
- benefits to the **user**
 - user is **not involved** (TOs take effect automatically as soon as the appropriate configuration file is installed)
 - user can **verify** the difference
 - low level **details are hidden**

Transport Options (4) – evaluation

- downside

technically **difficult** to write contracts on **adaptive behaviour** compared to contracts on **traffic characteristics** (rate, packet size, burst size etc.)

easy to monitor traffic, but adaptive behaviour has to be **inferred** (by monitoring and traffic or analysis) by other means

extends provider control to user's 'private space' : **Big Brother concerns**

Implications for **provider interconnection** : not yet been explored ...

no quantitative guarantees (is this a problem?)

Summary (...)

- powerful abstraction: **pushes the policy enforcement boundary higher** ..uses the endpoints but delegates *responsibility* from the endpoint to the provider (ISP: globally recognizable/accountable entity),
- provider becomes reference point for **trust** – price paid: the provider should now police *not* the traffic but the **behaviour** of the endpoints (customers)
- new issues in **traffic exchange between ISPs** : open ...
- scalable way to differentiate at random bottlenecks **far beyond the provider's domain** (more important with broadband access which eliminates first/last hop bottleneck).



UNIVERSITY OF
CAMBRIDGE

Computer Laboratory