Transport Options

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

TAPAS Project Meeting, Bologna, Italy
25 September 2003

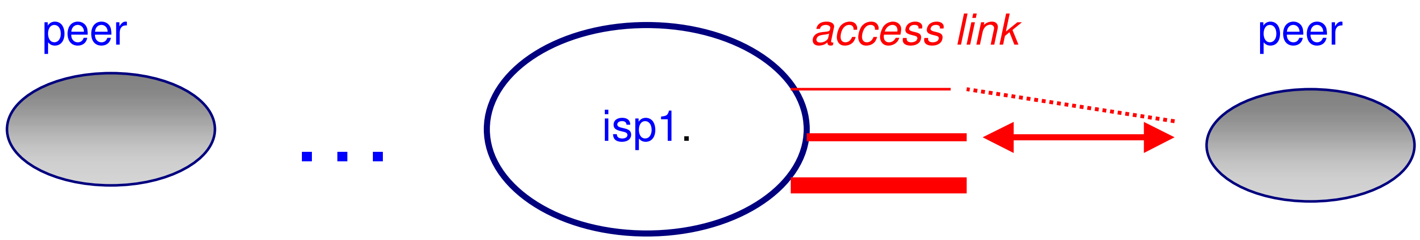
UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Outline

- Motivation : overview of end-to-end performance limitations
  access provider (first/last hop , ISP), network, endpoint, (protocol)
- Transport Options
  idea, purpose, implementation, deployment, evaluation
- Practical Example
  server-side, client-side, (enforcement)
- Inter-provider issues

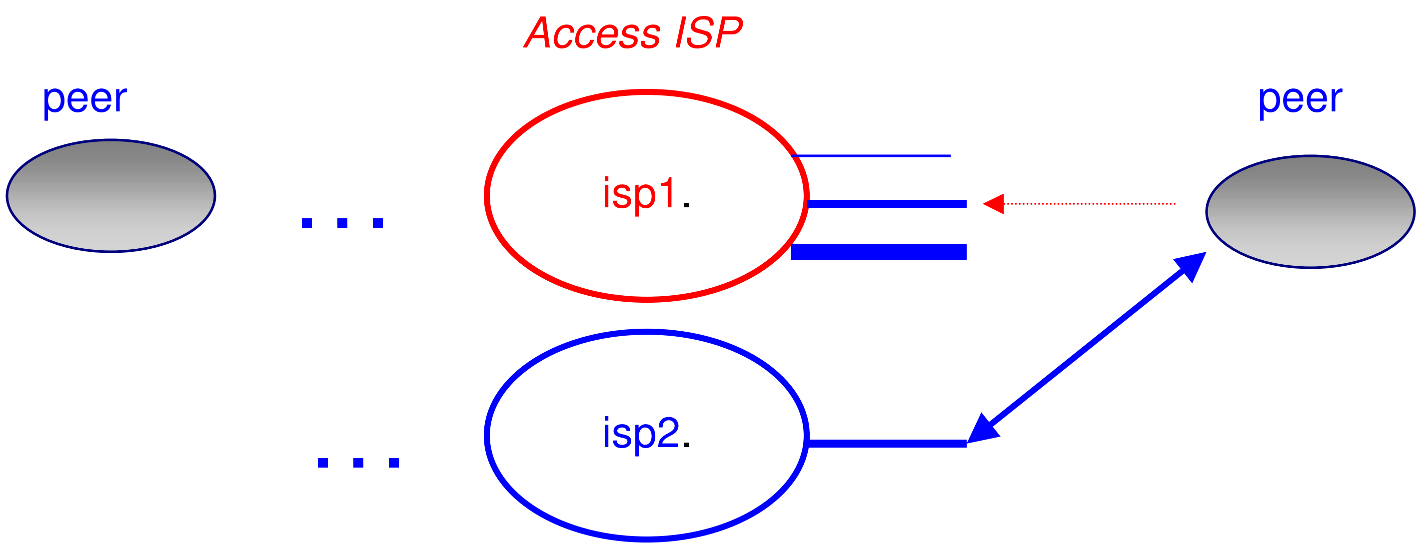
Performance limitations (1)

- bottleneck is the access link (first or last mile problem) – solution: capacity upgrade
- easy to verify - traffic direction irrelevant (either client or server : *peers*)



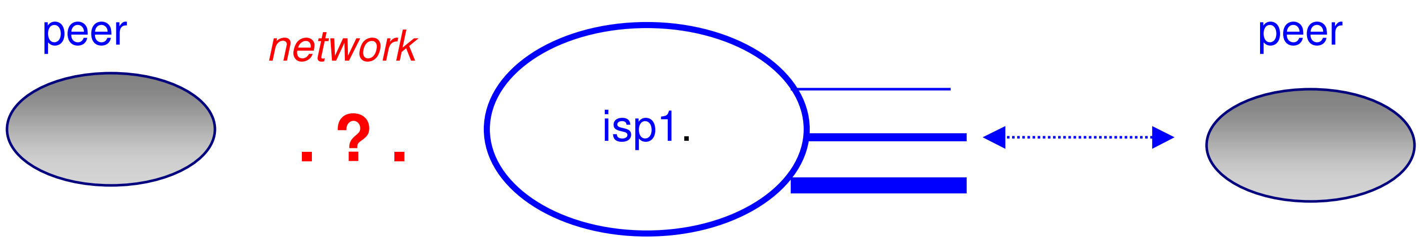
Performance limitations (2)

- bottleneck is suspected to be the access ISP – **provider selection**
- more difficult to verify – long term decision



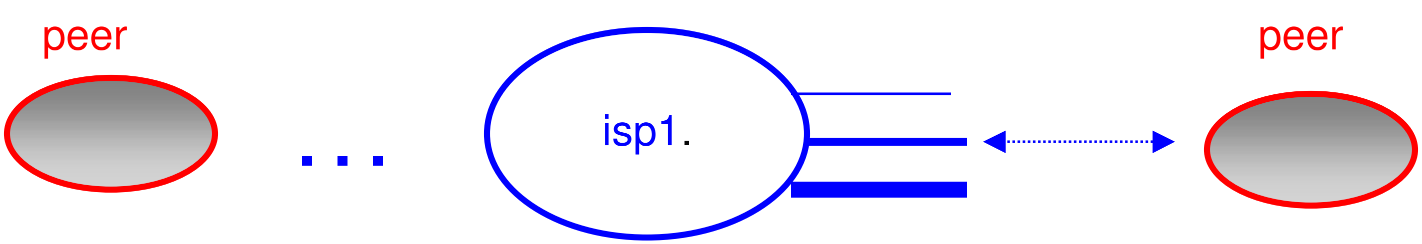
Performance limitations (3 *)

- bottleneck is somewhere further inside the the network – request re-routing from isp1 (assuming this is an option) - no solution



Performance limitations (4 *)

- bottleneck is the endpoint (server) (could afford to transmit faster but by following the standard, may impose unnecessarily delays).



current practices - distributed servers

- performance enhancing techniques (increase server's responsiveness)
- load-balancing – traffic localization.
- load–balancing techniques : addressing server overload, routing based, DNS-based
- DNS based:

  collect BGP tables,

  use the AS path length per prefix as distance metric

  use distance to determine nearest server

  look up client's address in the table,

  return matching server address to the query point
- policy distribution ? different redirections to different clients.

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

Requirement for new types of Service Level Agreements, SLAs (customer-provider contracts) by the ISPs.

Potential “deregulation” of congestion control rules creates opportunities for:

- performance enhancements (‘local’ scope)
- commercial exploitation (when used as a charging mechanism)

TAPAS

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Transport Options (2) - purpose

- performance enhancement (in networks where more aggressive transmission behaviour most likely will not cause congestion problems)
- **Example 1:** In FreeBSD TCP used a very large initial cwnd for all connections on the same LAN

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

- **Example 2**: logical extension….“relax” congestion control rules for all TCP connections within a corporate intranet.

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 (DHCP), to customers/subscribers have different tariffs for *static* and *dynamic* allocations.
- Analogy :

  Offer 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)

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

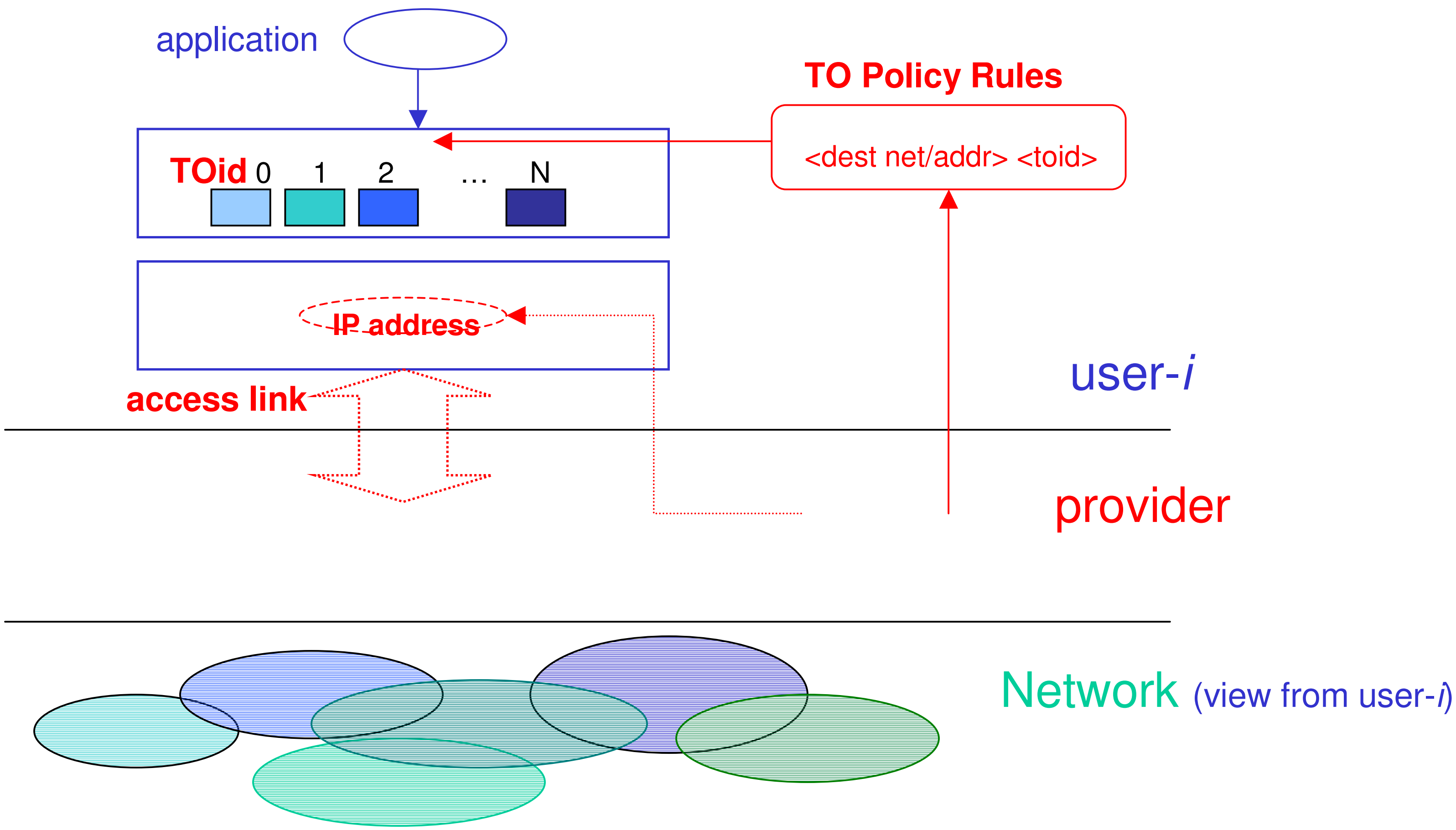
- preconfigured **policy rules** at the connection end-points determine which TOid will be used, settle for the minimum of the two
- Policy rules :: table entries in a configuration file, with two parts

  <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



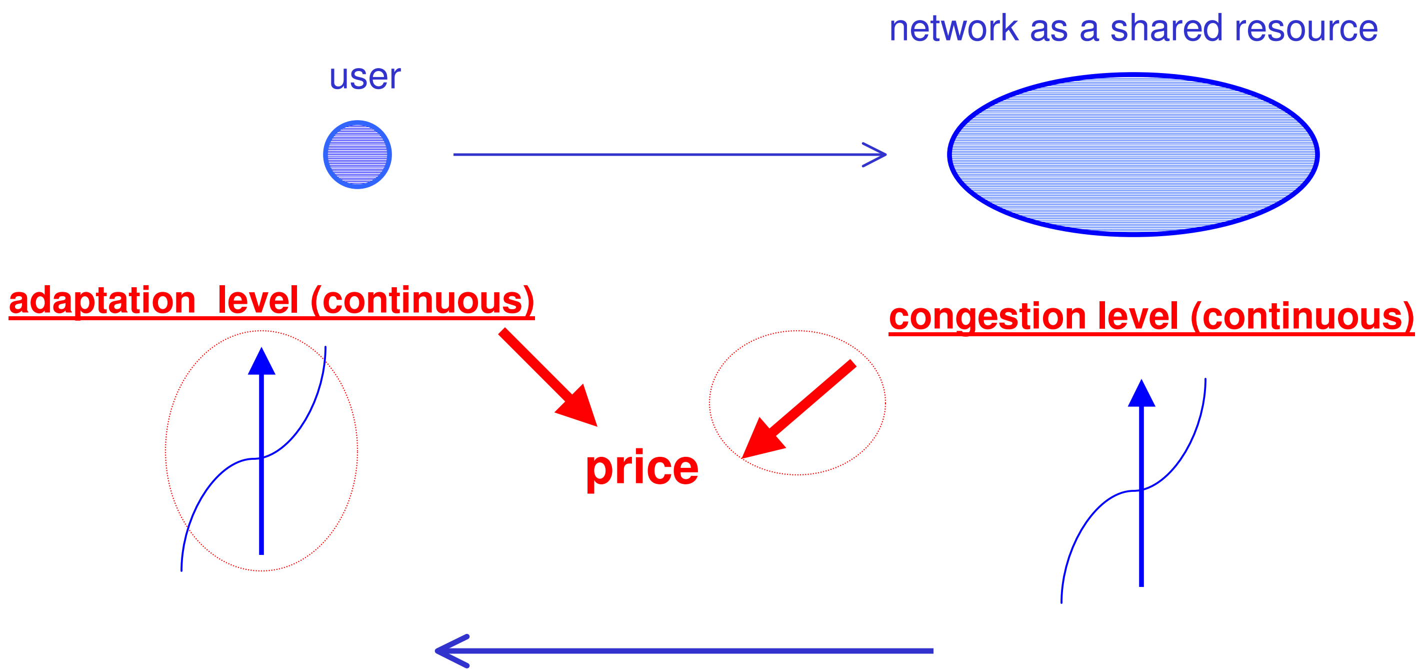
TAPAS

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Transport Options (4) – evaluation

- Comparing the Transport Options model to Congestion Pricing.
- Benefits to providers, users
- Problems

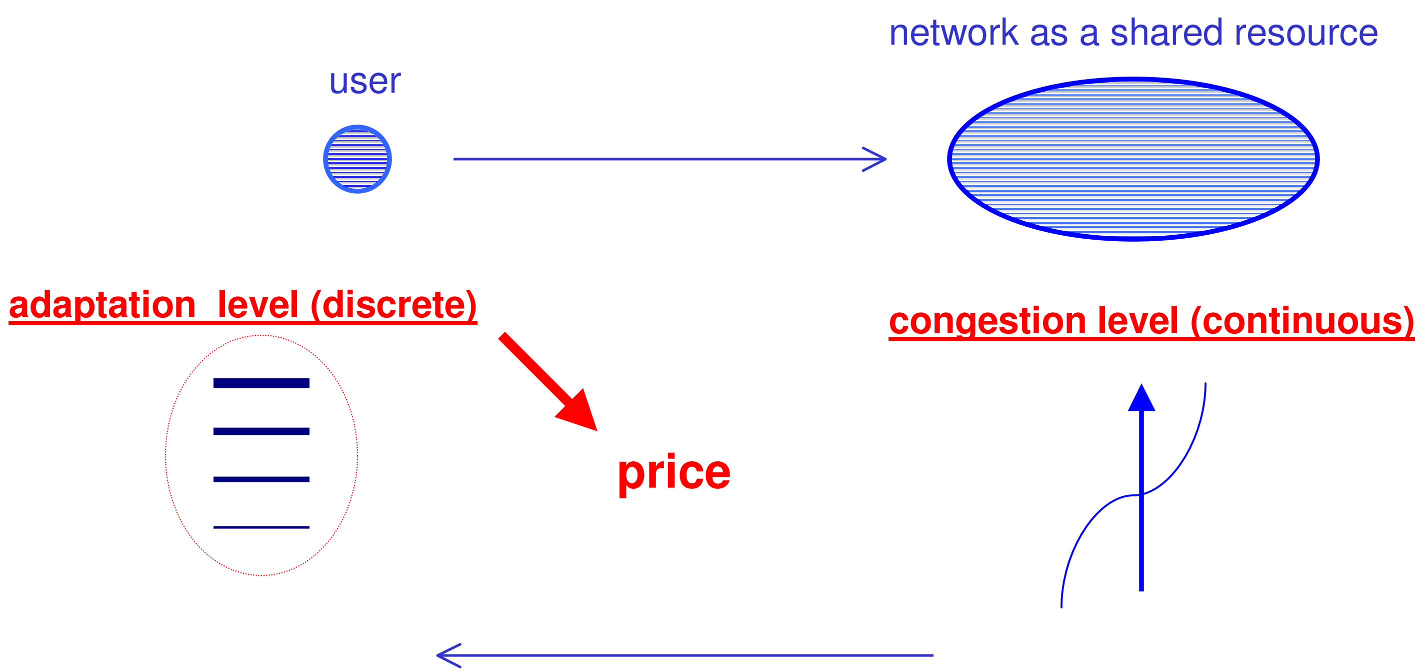
Congestion Pricing vs.Transport Options



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

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

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 …

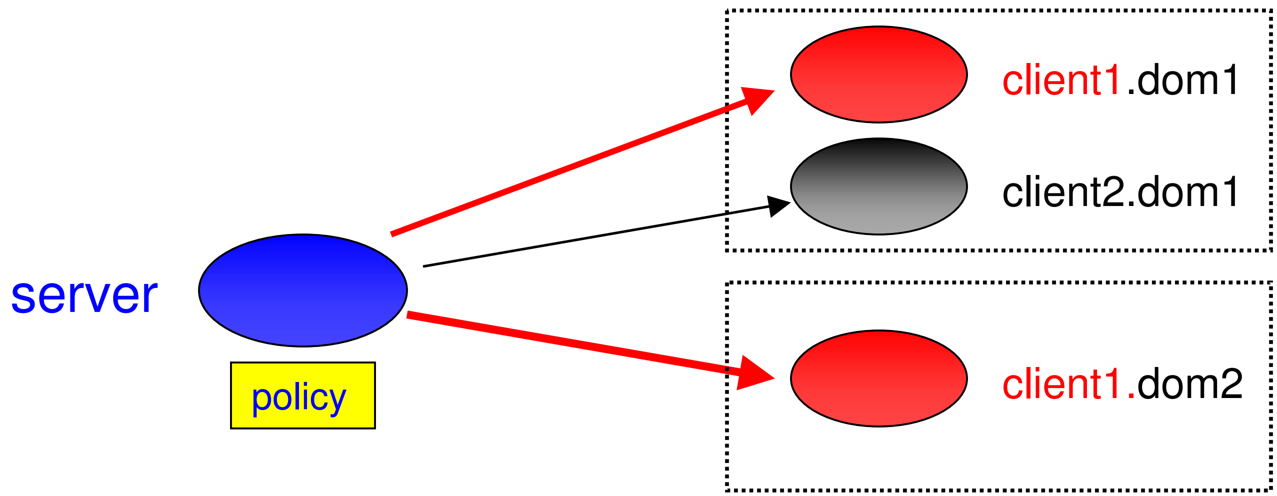
  no quantitative guarantees (is this a problem?)

Transport Options

- performance enhancement - introduce differential charging
- approach orthogonal to techniques such mirroring, load-balancing
- applies after the remote endpoint has been determined
- providers distinguish between two types of customer: server (snd-type) and client (rcv-type) - different issues involved

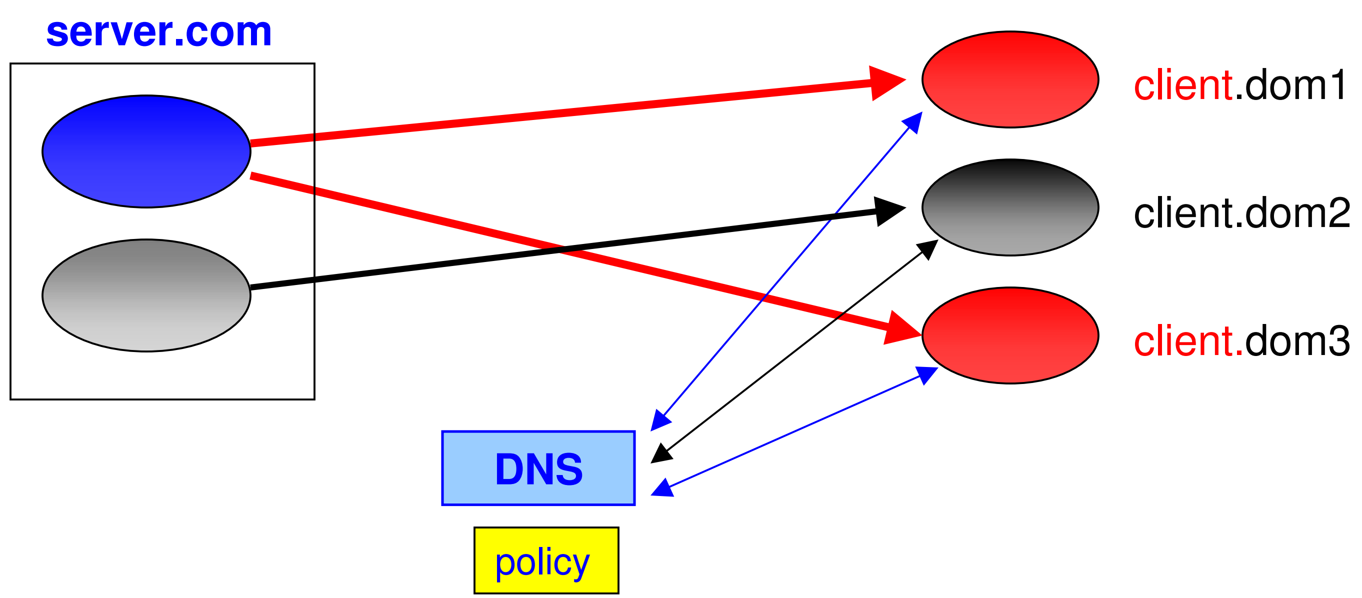
Transport Options – server side

- *server* determines how to transmit to each client according to policy
- two levels of service { “***standard***”, “***premium***”}
- implementation e.g. **(cwndInit = 1, a=1)** or **(cwndInit = 7, a=2)**



Transport Options – server side

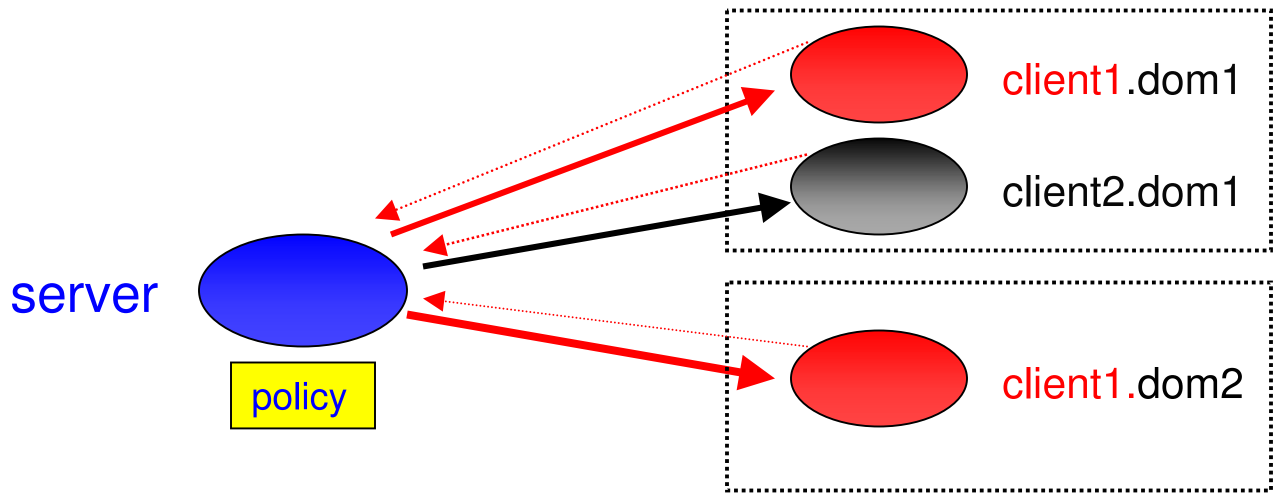
- DNS re-directs client (or domain) to appropriate server according to policy
- a server is configured to always serve either ***standard*** or ***premium***
- assume server overload is not a problem.
- redirected servers are often “near” the client



UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Transport Options – client side

- *client* signals to the server the type of service
- server decides whether the client is eligible for the requested type of service.
- provider may



Provider interconnection issues (1)

- Q1 : should the access ISP for the server care about how server clusters operate? or just traffic level will suffice?
- Q2 : should ISPs (at next level care about their neighbour's practices or traffic level at their interconnection links.
- Q3 : average traffic levels (measured over long time intervals) may appear the same, but the increased burstiness of the traffic (smoothed in the averages) may increase packet loss (?).

Provider interconnection issues (2)

- Assumption 1 : congestion (persistent packet loss) occurs only at inter-ISP links and never inside an ISPs domain.
- Assumption 2 : the provider "incorporates" all traffic it accepts from its neighbours (treats it as its own). Bilateral ISP relationships.
- Assumption 3: Content provider (ASP, "server") has sort of "limited liability" (only accountable for the level of traffic to its access ISP - not responsible for the effects of its traffic further down the path).
- *If* we accept above Assumptions we formulate sort of an "irrelevance hypothesis" : Transport Options should make no difference to current inter-ISP operation practices.
- require quantitative evidence of the effects of increased burstiness (in hope that we fail to reject the hypothesis)
- encouraging evidence exists already (studies of increased initial window RFC)

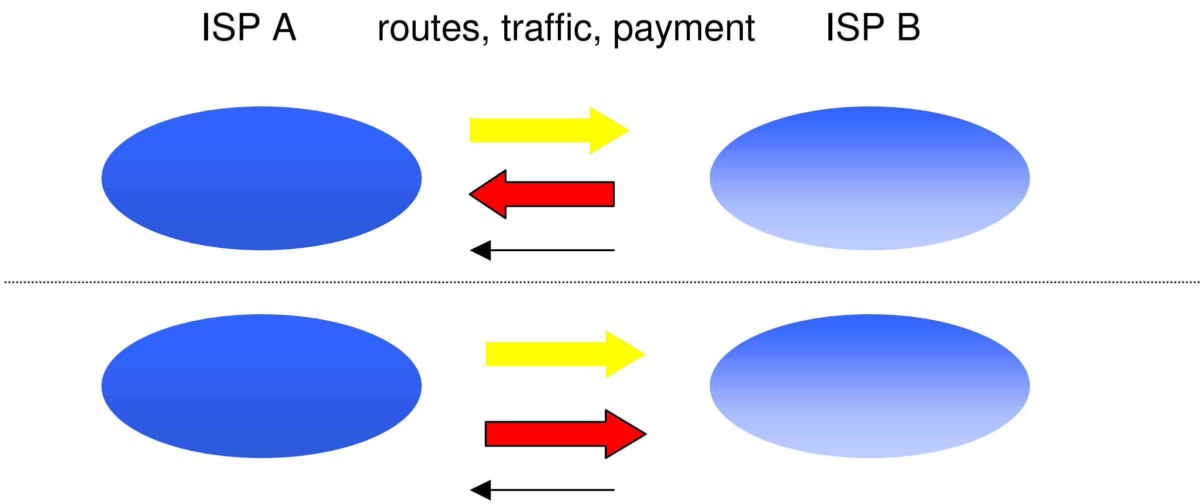
TAPAS

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Provider interconnection issues (3)

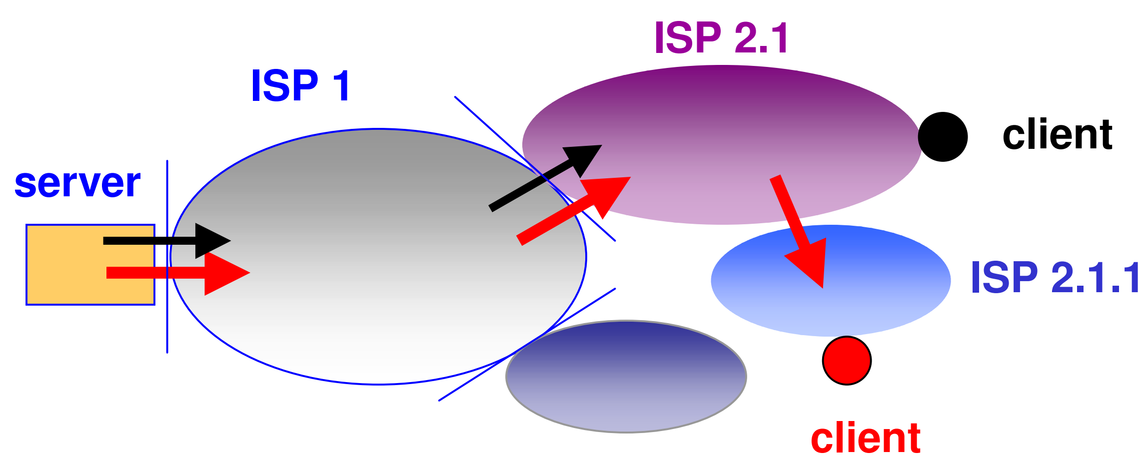
- simulate the effects of burstiness
- prepare a workload which maps to long-term, average utilisation say ~ 30%
- fixed number of flows, packets/flow
- run the workload with standard and premium control behaviours
- observe transfer times, packet loss rates

Provider interconnection issues (4)



Provider interconnection issues (4)

- ISP2.1 accepts traffic for its own **client** and the **client** of ISP2.1.1
- it is possible (depending on **server** policy) that the ISP2.1.1 **client** receives ‘’better’’ service than the ISP2.1 **client**
- if ISP2.1 considers that the levels of transit traffic it accepts from ISP1 penalises its own customer may have to re-negotiate with ISP1 (not 2.1.1)



Summary

- Transport Options - improve performance - allow differential charging – not harmful to the network – very simple to deploy.
- slides / paper online at

  http://www.cl.cam.ac.uk/~pg281/tapas/

  http://www.cl.cam.ac.uk/~pg281/research/

UNIVERSITY OF CAMBRIDGE
Computer Laboratory