

The information in congestion prices: milliseconds to years

Ramesh Johari
Stanford University
IETF 78, 29 July 2010

Congestion pricing

Congestion pricing is a vehicle to help ensure *scarce network resources are allocated to those who value them most.*

You are probably already familiar with one form of “congestion prices.”

A user during congestion

- Users complain to the network manager
- Complaining more/louder than others earns more network resources



There is *an implicit congestion price*:

How much time are you willing to spend complaining to earn priority during congestion?

Information in prices

We are interested in prices that are more *informative*.

An “efficient” congestion pricing scheme provides useful information at three timescales:

- 1) “short” (milliseconds to seconds)
- 2) “medium” (minutes to hours to days)
- 3) “long” (months to years)

Short timescales

An allocation is “efficient” if and only if
for every user:

The marginal utility for additional capacity
= the sum of prices along the path used

[under some technical conditions]

During congestion events, the price of a resource is
set so that resources are consumed by those
who value them most.

The marginal value of capacity

The view of congestion prices as the *marginal value of capacity* is quite powerful.

In particular, congestion prices are useful *tools* for decisions made on longer timescales.

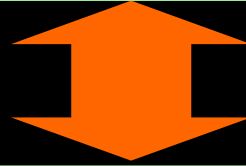
Medium: Traffic engineering

Long: Provisioning

Medium timescales

Traffic engineer

Sets routes, given traffic matrix



Content distribution network

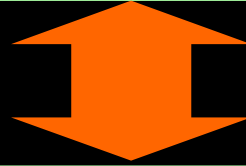
Picks servers, given latencies

If TE is unaware of how CDNs direct users to servers,
their interaction can perform very poorly —
worse than if the TE had chosen to do nothing at all!

Medium timescales

Traffic engineer

Sets routes, given traffic matrix



Content distribution network

Picks servers, given latencies

***But if the correct congestion prices are signaled,
then the interaction of traffic engineers and CDNs
becomes efficient:
always route on *cheapest* paths.***

Long timescales

If prices are the *marginal value of capacity*, they can be used to suggest where *provisioning* is needed in the network.

What happens if congestion prices are not accurate?

Braess' paradox: the naïve instinct to increase capacity in response to local congestion can *actually make all users worse off*

If congestion prices are accurate, then efficient upgrades are simplified: increase capacity where prices are high.

Moral

Every scheme for bandwidth allocation in times of congestion has a built-in value judgment.

Determining the “best” allocation requires identifying which uses are most valuable.

Congestion pricing identifies the *marginal value of capacity*.

This identification also has other benefits, for traffic engineering and provisioning.