InvisibleHand
Networks, Inc.

technical paper

Efficiency of Dynamic Bandwidth Markets

How IP service providers can increase revenues by 45% and improve gross margins on bandwidth 5 fold.

Efficiency of Dynamic Bandwidth Markets

Nemo Semret[1]

This paper demonstrates and quantifies the efficiency of dynamic market-based pricing of IP bandwidth relative to the 90th percentile usage-based pricing model common in the ISP industry today. Based on a simple model of dynamic pricing, and using real-world traffic and price data, we illustrate how dynamic market-pricing can yield 25%-65% increases in revenue, and 2 to 7-fold increases in gross margins for sellers of IP bandwidth, compared to existing usage-pricing models.

[1] nemo@invisiblehand.net

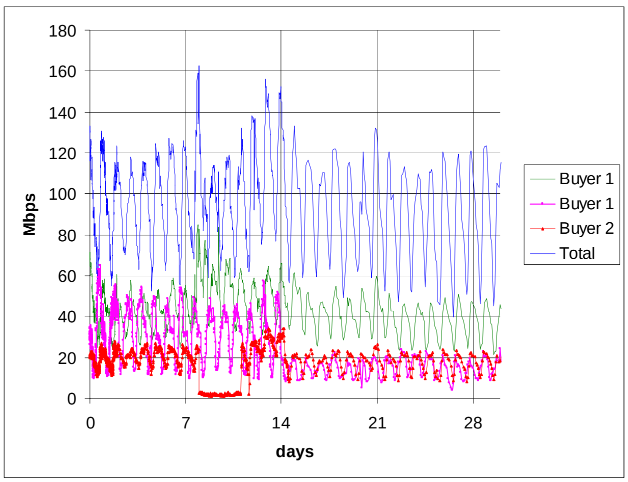
Introduction

We assume a "sender pay" bandwidth market (the analysis applies equally in a "receiver pay" market – the direction does not matter so long as it is consistent throughout), and take the view of a provider (seller) with buyers sending traffic in at a given access point. To best isolate the effect of a dynamic market, we assume the provider's cost of bandwidth is given by the traditional model, and compare different pricing models for the revenue side. This reflects a) the fact that for all providers, the dominant variable cost is transit fees paid to other ISPs for the traffic that has to be sent on to other providers, and b) that what we are interested in is the gain from adopting dynamic pricing. Thus the cost is calculated as the 90th percentile of the aggregate traffic, times a wholesale bandwidth price of \$200/Mbps/month. We will show that the 90th percentile usage pricing that is prevalent in the industry for IP transit bandwidth is inefficient by showing that it is possible to consistently buy in that model and resell at a profit using a more efficient pricing model, in other words there is an arbitrage opportunity.

Base case

Consider a provider with, initially, 3 large buyers. The figure shows the traffic flows of the three buyers – a 30-day trace taken from an actual Internet exchange point, where the buyers are content providers and ISPs.[2] The average level of the total traffic is 95Mbps.

The revenue with 90th percentile billing is given by the 90th percentile of each buyer's flow times the same price of \$200/Mbps/month.



[2] Note that throughout this paper, we use "flow" and "buyer" in the aggregate sense. The discussions and analysis here would not normally be applicable to individual connections (e.g. TCP sessions). A single "flow" here represent hundreds or thousands of individual downloads, streams, etc., and one buyer represents e.g. the total traffic from a website, or from a single service provider.

	Peak	90th %	Mean	Price	Value
Buyer 1	85	57	43	200	$ 11,389
Buyer 2	66	50	34	200	$ 9,973
Buyer 3	34	25	17	200	$ 4,939
					$ 26,301
Aggregate cost	163	122.1	95	200	$ 24,415
Profit					$ 1,886
Margin					7%

With static, usage-based pricing, a profit margin is realized simply through **statistical multiplexing gain**. Indeed, because statistically, it is unlikely that different buyers will peak at the same times, the aggregate traffic is "smoother" (less bursty) than the individual streams: here, the peak-to-mean ratio is around 1.9 for the buyers individually, but 1.7 for the aggregate. On the 90th percentile, the gain is 7%.

Price Elasticity

The key characteristic of demand here is price elasticity, which measures the degree to which demand will respond to changes in price. For example, a price elasticity of 3 means that a decrease in price of 1% will create an increase in demand of 3%. More generally, if there is a total demand for a quantity Q of some good, at a price P, the elasticity K relates changes in demand (ΔQ) with changes in price (ΔP) as follows:

$$\Delta Q/Q = -K\ \Delta P/P$$

Thus, theoretically, a provider can increase traffic by ΔQ simply by lowering price from P to $P(1-\Delta Q/(QK))$. Of course, this assumes there is sufficient liquidity, i.e. there are plenty of buyers and no barriers other than price preventing them from buying more. In practice, in the absence of a real-time market, this occurs over a period of time, as new prices are advertised, new buyers get provisioned, etc.

According to a comprehensive study,

"… for companies provisioning Internet services, the demand for bandwidth appears to be much more elastic than it is for telephone services. That is, for every 50% fall in the price of bandwidth, some ISPs have purchased 100% or more additional capacity."[3]

This corresponds to a price elasticity of K = 2. Another study, done on a wide range of data services from the early 1970s to the late 1990s, also finds a price elasticity of 2.[4]

One can also validate the elasticity value by looking at long term forecasts of demand and price, and backing into an implied elasticity: most estimates of current IP traffic growth are at around 100% per year[5] to 140% per year[6], while prices are estimated to decline at rate ranging from 35% per year[7] to 50% per year[8]. This implies that the price elasticity is between 2 and 4.

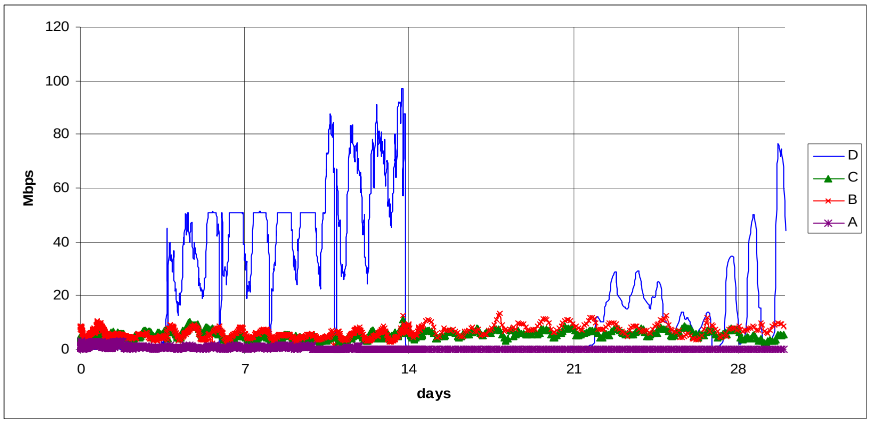
[3] *International Bandwidth 2000,* TeleGeography, Inc., Apr. 2000.
[4] "Beyond Moore's Law: Internet Growth Trends", L. G. Roberts, *IEEE Computer,* Sep. 2000.
[5] "Growth of the Internet", K. G. Coffman & A. M. Odlyzko, http://www.research.att.com/~amo, 2001.
[6] "Bandwidth Explosion", Lehman Brothers Global Equity Research, 2001.
[7] "Bandwidth Explosion", Lehman Brothers Global Equity Research, 2001.
[8] "Shrinking Streams grow bigger", *Wired News,* Nov. 2000.

Growing Traffic with Static Pricing

A provider can increase traffic by simply lowering prices. As noted above, we assume the provider can only change it's revenue side, but the unit cost remains fixed at $200/Mbps/month. By attracting a larger numbers of buyers, the seller hopes to further smooth the aggregate traffic, while still charging buyers based on their individual 90th percentile peaks, and that the resulting multiplexing gain will outweigh the required drop in prices.

Assuming a price elasticity of 2, if the seller described above wishes to grow average traffic by 35%, then prices must be reduced by 17.5%.

We take content providers as the additional buyers since they are generally more price sensitive than other classes of bandwidth buyers[9], and there fore more representative of additional traffic likely to be attracted to by a drop in prices. The next figure shows a combination of 4 buyer flows -- 30-day traces of actual Internet traffic from 4 content providers hosted in a data center.



This additional traffic has an overall mean of 34Mbps. Assuming a price elasticity of demand of K=2, to attract this amount of additional demand, the price must be reduced from 200 to $164/Mbps/month. The revenue with 90th percentile billing is given by the 90th percentile of each buyer's flow times $164/Mbps/month.

	Peak	90th %ile	Mean	Price	Value
Buyer 1	85	57	43	164	$ 9,358
Buyer 2	66	50	34	164	$ 8,194
Buyer 3	34	25	17	164	$ 4,058
Buyer A	3.0	1.1	0.3	164	$ 181
Buyer B	13	10	7	164	$ 1,578
Buyer C	11	7	6	164	$ 1,220
Buyer D	97	51	21	164	$ 8,388
					$ 32,978
Aggregate Cost	250	174.7169	129	200	$ 34,943
Profit					$ (1,965)
Margin					-6%

[9] See for example "Streaming bleeds cash", *Industry Standard,* Sept 25, 2000

In this case, the traffic growth is actually detrimental to the seller, since it goes from a small positive profit margin of 7% to a negative margin. The reason is that, even though statistically, burstiness decreases when more independent flows (buyers) are mixed, in practice, the seller has no deterministic control over when bursts occur. In this case, the new flows (particularly D) are much more bursty (the added traffic has a peak to mean ratio of 3.3), so the new aggregate ends up more bursty than it was before (peak-to-mean of 1.9 instead of 1.7).

In a "luckier" scenario, excluding the worst offender (buyer D), we get:

	Peak	90th %ile	Mean	Price	Value
Buyer 1	85	57	43	187	$ 10,638
Buyer 2	66	50	34	187	$ 9,315
Buyer 3	34	25	17	187	$ 4,613
Buyer A	3.0	1.1	0.3	187	$ 206
Buyer B	13	10	7	187	$ 1,794
Buyer C	11	7	6	187	$ 1,387
					$ 27,953
Aggregate	175	138.8827	107	200	$ 27,777
Profit					$ 176
Margin					1%

The new aggregate is indeed less bursty (peak-to-mean ratio of 1.6) than either previous case, but the multiplexing gain is not enough to offset the price reduction on the revenue-side, so the profit margin of 1% is still smaller than it was with less traffic.

Of course, the change in profit margin will depend on the traffic pattern in each situation. On average, aggregation does smooth traffic and reduce unit costs. The question is: does it reduce costs sufficiently to offset the price decline required to grow traffic? For ISPs today, given that

- there is a positive correlation between price elasticity and burstiness (specifically, additional traffic attracted by lower prices is likely to be more costly), and
- with usage-based pricing, the seller cannot control exactly when and how much the traffic will burst,

the answer is uncertain at best.

Growing Traffic with Dynamic Pricing

Of course, the main problem in the previous case is that, with traditional pricing models, i.e. static usage-based pricing, the new price has to apply to all demand, at all times.

What if prices would automatically drop only when there is lower demand, in order to grow traffic, and conversely, prices would rise when demand is high? In effect, the provider would shift from a model of fixed (unit) prices and unpredictable traffic levels (which affect not only the cost of the seller, but the quality experienced by the buyers), to prices varying in a way which actively helps smooth the traffic and improve the average utilization of the network.

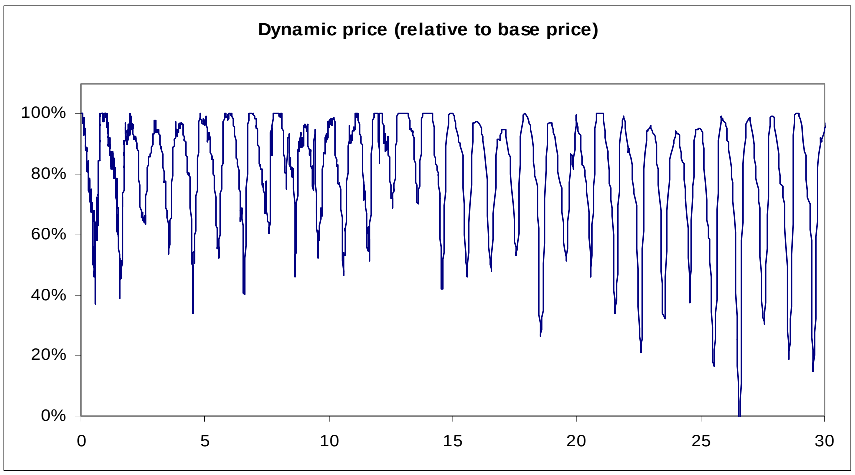
The promise of dynamic pricing is that the seller could obtain the benefits of aggregation (multiplexing gain), while still having the highest sustainable prices.

Consider the following simplified model of dynamic market pricing: at all times, the buyers pay a current price for their current bandwidth allocation, and prices dynamically adjust at all times to the point where the demand is maximized, subject to not increasing the cost. More precisely, let Q(t) be the existing demand (traffic level) at time t, when the price is fixed at P, and let Q_c be the committed amount for the aggregate, which here is the 90th percentile. Then, the additional quantity of demand to be generated is:

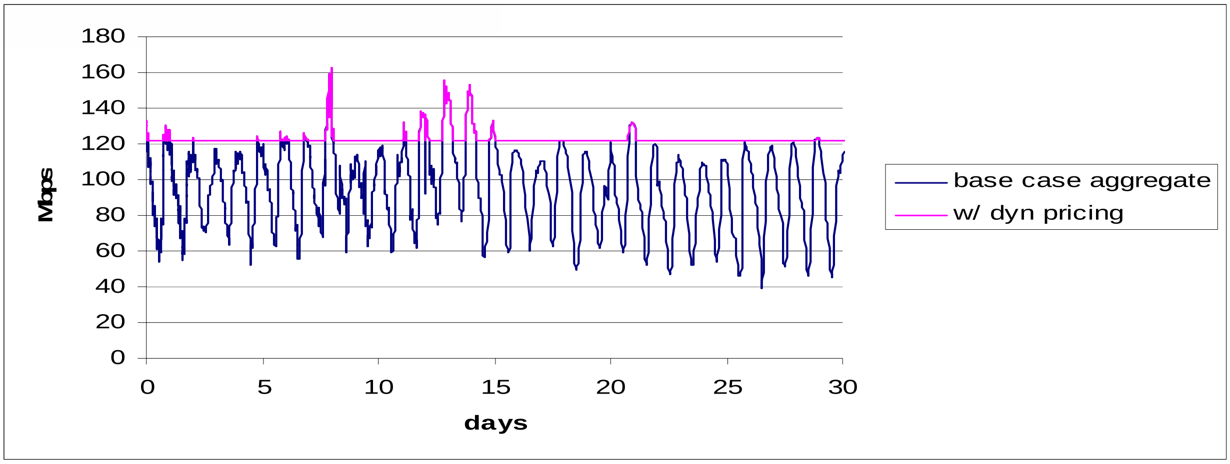
$$\Delta Q(t) = Max(Q_c - Q(t), 0).$$

So the dynamic price at time t is:

$$P(t) = P[1 - \Delta Q(t)/(Q(t)K)].$$



With the traffic pattern of the base case, the 90th percentile is at 122Mbps. The application of dynamic pricing means any time the traffic level is below that, the price drops enough to increase demand to the point where the traffic is raised to 122Mbps. Thus demand is maximized without changing the cost. The next figures show the dynamic price relative to the base price P and the resulting traffic levels.



In the traffic graph, the reference price point P (which is the 100% level in the relative price figure) represents the willingness to pay of the base case buyers, which is already revealed in the base case. Of course, since the pricing model is now dynamic rather than a 90th percentile commitment, the reference price P must be set such that the base case buyers, with dynamic pricing P(t), end up paying the same total as they are in the

base case for the same bandwidth, which in this case means P=$326/Mbps/month. The end result for the seller is then:

Reference price	326
Revenue from base buyers	26301
Revenue from additional traffic	6126
Revenue gain	**23%**
Total revenue	32427
Cost	24415
Margin	**25%**

Thus, with simplified dynamic market pricing model, existing revenue is preserved, costs remain the same, but additional traffic is attracted and revenues increase by 23%, and profit margin increases by a factor of 3.5 times, from 7% to 25%.

Again, the actual values are dependent on the specific traffic patterns. For completeness, we consider now the effect in the secondary case, where the existing traffic is from buyers A, B, C, D rather than buyers 1, 2, 3. With 90th percentile based revenues, the result is:

	Peak	90^{th} %ile	Mean	Price	Value
Buyer A	3.0	1.1	0.3	200	$ 220
Buyer B	13.3	9.6	6.7	200	$ 1,920
Buyer C	11.0	7.4	5.5	200	$ 1,485
Buyer D	97.2	51.0	21.3	200	$ 10,209
					$ 13,835
Aggregate	111.8	65.2	33.8	200	$ 13,041
Profit					$ 794
Margin					6%

The profit margin in the traditional model (i.e. the simple multiplexing gain from aggregation only) is very close to the one in the base case. However, here, since the existing demand is very variable, from the seller's perspective, there is a greater inefficiency, and therefore, more to be gained with dynamic pricing. The end result is then:

Reference price	502
Revenue from base buyers	13835
Revenue from additional traffic	8830
Revenue gain	**64%**
Total revenue	22664
Cost	13041
Margin	**42%**

With this highly variable demand, the increase in profit margin between fixed 90^{th} percentile based pricing and dynamic market pricing is 7 fold.

Note that we assumed a price elasticity of K=2, which is on the low end of the real-world estimates for IP traffic. In that regard, the gains from dynamic pricing are under-estimated, since higher elasticity would mean greater increases in demand for smaller decreases in price.

Dynamic Pricing versus Versioning

As evidenced by the above simplified model, the main cause of inefficiency with fixed-price usage-based pricing is that any new price has to apply to all demand, at all times. A natural question to ask is therefore: are there pricing models other than dynamic market pricing which are also more efficient than the traditional models?

In theory, barriers not directly related to demand can be created to segregate buyers by willingness to pay, and charge different prices for a good that costs the same to produce and deliver for all. This is known in economics as "versioning", and consists essentially of attaching some kind of inconvenience or other penalty to the lower prices: for example, the airline industry's infamous "Saturday night stay-over" restriction is versioning used to separate business travelers with a high willingness to pay from tourists with a low willingness to pay, even though they may be getting the same thing at the same time. Similarly, paperback editions of books are released later than hardcover editions, which allows the publisher to separate out the higher willingness to pay of the most eager readers by selling them hardcover editions at a high price (much greater than can be accounted for by the actual difference in the cost of producing and distributing hardcover versus paperback books *per se*, as evidenced by the fact that as soon as a paperback edition comes out, the hardcover editions price drops much closer to that of the paperback). In reality, many such restrictions, built-in delays, etc., are fig leaves for what a buyer advocate would call price discrimination.

With IP bandwidth, one way to do versioning is to artificially degrade the quality of some traffic:

"*Service providers may have several motivations for imposing caps including […] attempting to increase future profits by generating demand for "premium" classes of service -- customers may be willing to pay more for a higher cap or no-cap level of performance.*"[10]

Another way, more common today, is to rely on human expertise in negotiations to identify willingness to pay, but that remains an expensive approach, whose benefits only outweigh the cost in the largest deals. But even such differentiated pricing would remain static, and thus not able to respond to unpredictable demand.

With real-time dynamic pricing, price can adjust to traffic demand, and with intelligent buying, demand can respond to price. Thus, in a market with sufficient buy-side liquidity, if the seller offers a fixed capacity and lets price drop when there is excess capacity, this will attract additional traffic.

Conclusion

This highly simplified analysis illustrates why, with static pricing, despite tremendous growth in Internet traffic, ISPs are caught between

1. Pursuing growth at the risk of entering a death-spiral of declining prices;
2. Pursuing profitability by focusing only on high-value customers willing to pay premium prices on an underutilized network;
3. Using their core business (data transport) as a loss-leader and seeking profits in "higher value services" such as security, managed hosting etc., often competing at a disadvantage against more focused specialized providers of those services.

Dynamic market-based pricing offers a way out of this quandary; one that provides **25%-65% increases in revenue, and 2 to 7-folds increase in gross margins** (bandwidth only).

The dynamic pricing model we use assumes there is an efficient market mechanism, and perfect liquidity (i.e. if the price is right, a buyer can always be found). With less than perfect liquidity, there may not be the instantaneous response in demand increase when prices decrease. This needs to be factored into any prediction of the efficiency gain from dynamic pricing in a practical scenario.

An overall environment with sufficient liquidity is key to successfully realizing the benefits of dynamic market pricing. Various factors will contribute to liquidity:

- Buyer sophistication: the extent to which existing buyers can shape their own traffic in response to price incentives, through load-balancing, intelligent routing, changing quality, etc.

---

[10] "DSL vs. Cable Modem" http://compnetworking.about.com/library/weekly/aa111200b.htm

- Location: the traffic exchange point: for example, a data center or peering point (with one or multiple network service providers) will have a pool of buyers who can be attracted into a flexible price market;
- The appropriate market mechanisms and software platform for real-time pricing and allocation of bandwidth.

These last two points are discussed in detail in a companion paper focusing on software and network service architectures enabling real-time market pricing of bandwidth[11].

[11] "Merkato Overview", InvisibleHand Networks Technical paper.