Chapter 5.
Traffic Management Decisions in the United States

5.1 Introduction

Since it became popular among the American public, fixed-line broadband Internet service has been primarily offered without application-specific network management. While some cable operators adopted specific techniques to manage peer-to-peer traffic, those practices largely ceased in the wake of the FCC's *Comcast Order* (FCC 2008). The large telephone companies did not make similar forays early on, and as the net neutrality debate intensified over time, they had even less incentive to pursue network management solutions that involved managing particular Internet applications.

This result was not solely the product of engineering designs or regulatory imperatives, but of the intermingled forces of both. The cable operators that deployed peer-to-peer management techniques were responding to a very specific engineering challenge that years of emergency bandwidth capacity upgrades showed no signs of fixing. But the context in which they made their solution choices was highly shaped by their unique regulatory history – or lack thereof. Compared to the telcos, the cable broadband operators had existed in a highly favorable and recently established regulatory environment prior to their adoption of peer-to-peer management solutions. For some operators, this environment imbued organizational paradigms for traffic management decision-making with a flexible, freewheeling character. It was under these circumstances that the solution that was ultimately denounced by the FCC was chosen and deployed. Only in response to the *Comcast Order* did these operators begin to institutionalize broader internal review and oversight of their traffic management decisions.

The telcos' networks were based on different access technologies with different characteristics. Their need to intervene to manage specific Internet applications was less urgent and they therefore largely refrained from doing so. But the telcos also had a radically different orientation toward their regulatory environment. They had fought for decades to

have common carriage obligations lifted from their data and broadband services offerings, finally achieving that goal in 2005. They had good reason to tread cautiously so as not to upend these hard-fought gains. They had also accrued thorough regulatory review structures – policies, personnel placements, and means for internal cross-departmental dialogue – that they readily applied to internal traffic management debates. Throughout the entire history of their broadband offerings, the telcos' traffic management decisions were grounded in the combination of the unique engineering constraints of DSL and fiber mediums and deeply embedded organizational processes to address regulatory risk.

This chapter analyzes in detail how the commingled forces and technology and regulation shaped traffic management decisions and decision-making in the United States. The combined decisions of the cable and telephone operators have resulted in fixed broadband networks on which over-the-top Internet applications have not generally been managed, in stark contrast to the United Kingdom.

5.2 Regulatory History and the *Policy Statement* in Context

As discussed in Chapter 1, by the mid-2000s the telephone companies had been engaged in nearly 40 years' worth of uphill battles to relieve themselves of regulatory obligations on data services, with each successive round of the *Computer Inquiries* resulting in the removal of additional regulations. Cable operators, by contrast, had never been regulated as common carriers during the short existence of cable broadband service. The two industries' contrasting experiences provided the backdrop for a number of events in the summer of 2005 that profoundly altered the landscape of broadband Internet regulation in the United States.

First, the Supreme Court handed down its decision in the *Brand X* case (National Cable & Telecommunications Assn. v. Brand X Internet Services 545 U.S. 967 (2005)), classifying broadband Internet service offered over cable networks as an information service not subject to common carrier obligations. Shortly thereafter, in its *Wireline Broadband Order* (FCC 2005b), the FCC ascribed the same classification to wireline broadband Internet service

(offered via DSL), creating parity with cable and ending the telephone industry's decades-long struggle to free its provision of data and broadband Internet services from regulation under Title II. To counterbalance this deregulatory step, the FCC simultaneously issued its *Broadband Policy Statement* (FCC 2005c), laying out a framework of principles declaring broadband consumers' rights to access the legal content and services of their choice, to connect the legal devices of their choice to the network, and to enjoy competition among broadband service providers. In making all four principles subject to "reasonable network management," the FCC formally recognized broadband network management for the first time.

The immediate operational effects of the *Policy Statement* were few. For the major fixed-line network operators, the *Policy Statement* largely reflected the status quo of industry practices at the time – it was "how the market had evolved in the US anyway," as one telco employee described it. Broadband offerings at the time were not characterized by significant limits on the content or services that customers could access or the devices they could attach to the network. Providers' previous restrictions on certain applications, in particular virtual private networking, had largely been lifted in the years before the *Policy Statement* was adopted (Coalition of Broadband Users and Innovators 2003; National Cable and Telecommunications Association 2003). As a result, there were few immediate consequences of the *Policy Statement* – none of the large operators needed to make drastic changes to the way they ran their networks in order to feel confident that they were in compliance with the *Policy Statement* in the months after it was adopted.

While the *Policy Statement* may have been simple to operationalize at the start, there was significant uncertainty about how it would be interpreted and enforced going forward. The ways in which the cable and telephone industries approached this uncertainty were shaped in significant part by their respective legacies of regulatory oversight at the Commission. The *Policy Statement* did not arise out of a vacuum, but rather was a by-product of a series of struggles in the courts and at the FCC. The same can be said for the respective mindsets that

the cable and telephone companies took to interpreting the *Policy Statement* and making traffic management decisions as time wore on.

The greatest uncertainty was about whether the *Policy Statement* was enforceable. When the FCC published the *Statement*, Chairman Martin's accompanying press release explained that "policy statements do not establish rules nor are they enforceable documents" (Martin 2005a). Commissioner Copps, the FCC's most forceful advocate for nondiscrimination rules, was more equivocal, but still left some uncertainty as to his opinion of the *Statement*'s enforceability:

> While I would have preferred a rule that we could use to bring an enforcement action, this is a critical step. And with violations of our own policy, I will take the next step and push for Commission action. A line has been drawn in the sand. (Copps 2005)

The question of enforceability eventually would prove to be the most pivotal of the US net neutrality debate.

Perhaps only second to enforceability in its lack of clarity was the meaning ascribed to "reasonable network management." The *Policy Statement* provided no further explanation as to what "reasonable" might mean in the context of network management, nor which practices or kinds of practices constituted network management functions in the eyes of the Commission. The network management provision was, after all, in a footnote. The focus of the entire effort in developing the principles lay elsewhere, as explained by a former FCC staffer who had been at the Commission during the drafting of the Four Freedoms:

> We thought at the time, in 2004, that the real action in that four freedom document was in the first freedom, right? This is the consumer entitlement to use applications and content of their choice. You know if you'd asked us then where we were going to have to adjudicate or where we were going to have to enforce, it was all going to be around that. It's only when I guess Kevin Martin looked at the *Comcast* case that all of those anxieties around the way consumers would access lawful content were poured into the network management bucket. . . . But we didn't think at that time that that's where the enforcement focus would be.

Network management was not the focal point of the Commission's action, and therefore was not well specified.

Thus, although most network operators could view the *Policy Statement* as largely mirroring existing industry practices, they were still faced with significant uncertainty as to how the Commission might react and whether enforcement action was a possibility should they decide to change their practices, particularly their traffic management practices. As the next section shows, the approaches that the cable and telephone industries took to this uncertainty were highly influenced by their respective regulatory histories. The cable broadband industry's short and successful record of dealings at the Commission allowed for the creation of corporate environments where risk-taking and deployment of new traffic management approaches seemed justified under broad interpretations of "reasonable network management." The telcos' hard-fought deregulatory status created the opposite effect: caution so as to prove that fears about lack of FCC oversight were unfounded, and a sense of the need to examine thoroughly the potential regulatory consequences of any new practices on the network.

5.3 Industry Approaches to Traffic Management

By the mid-2000s, US broadband subscribership was growing steadily. Nationwide, year-over-year growth in subscribership was 30-50% (see, for example, data collected by the FCC (2004)). Traffic volumes, both on average and during the peak usage period (between when work and school let out and when people went to sleep), were likewise growing steadily, with reports of 25-35% year-over-year growth (Erman et al. 2009; J. K. Smith 2008).

On many networks, a disproportionate amount of traffic growth was attributable to peer-to-peer file-sharing traffic. Popular peer-to-peer applications are characterized by their ability to efficiently transfer large files unattended: users can configure their peer-to-peer clients to download specific files in advance and then leave those clients running in the background without the need for user interaction. This means that even if the fraction of a network's broadband subscribers using peer-to-peer applications is small, the proportion of peer-to-peer traffic can be large. Furthermore, peer-to-peer networks only function if peers both upload

and download, because peers are the sources of all files to be shared. As a result, peer-to-peer traffic tends to be roughly symmetric, with comparable amounts of traffic being sent in the downstream and upstream directions.

Few public studies of the impact of peer-to-peer traffic on broadband networks in the US in the mid-2000s exist, but evidence from other countries suggests that peer-to-peer traffic was accounting for approximately 60% of residential downstream broadband traffic in that time frame (Cho et al. 2006; Plissoneau, Costeaux, and Brown 2005). Because residential broadband is generally offered as an asymmetric service with downstream bandwidth many times that of upstream bandwidth, peer-to-peer traffic was having an even more disproportionate impact in the upstream direction (Martin and Westall 2007). Several cable operators interviewed for this thesis reported peer-to-peer upstream traffic levels in the range of 80% of total traffic on cable networks in the mid-2000s.

The cable and telephone industries reacted differently to the challenges that the growth of peer-to-peer traffic presented. Numerous stakeholders have pointed to differences in access technology to explain the differential responses – that the cable companies, with more shared infrastructure closer to end customers and less upstream bandwidth capacity than either DSL or fiber networks, saw more intense effects from peer-to-peer traffic (Glover et al. 2008; Sandvine 2010b; Teplitz et al. 2010). Interviewees often explained cable's more drastic efforts to control peer-to-peer traffic in light of these differences.

However, technology differences were far from the only factors that distinguished the cable and telco approaches to managing peer-to-peer and network traffic in general. The cable companies' more cavalier approach to regulatory ambiguity created greater flexibility as to how traffic management decisions were made and opened the door for adoption of peer-to-peer-specific management solutions. While the technical characteristics of the telcos' infrastructure put less pressure on them to deal with growing bandwidth demands in novel ways, they also took far more care in considering the regulatory and public perception

consequences of application-specific management, ultimately shying away from it for the most part. These distinctions in approach are explored in the next two sections.

5.3.1 Cable Industry

Managing Peer-to-Peer Traffic

There is no doubt that peer-to-peer traffic was putting serious strains on cable broadband networks in the 2004-2005 time frame, particularly in the upstream direction. Cable interviewees described how freshly upgraded network links would reach 80-90% utilization and require new upgrades in a matter of months. In some instances bandwidth requirements doubled every year for several years.

Cable broadband relies on significant shared infrastructure in the access network. Dozens of subscribers in the same neighborhood may share the same local fiber optic node, and hundreds of subscribers may share the same Internet Protocol port further up in the network. Popular peer-to-peer file-sharing applications place a particular strain on this kind of architecture because they were designed to maximize the amount of data being exchanged at any one time, in part by opening many simultaneous connections to other peers. The result in the mid-2000s was that with even a small number of avid peer-to-peer users sharing a particular network link, bandwidth was quickly becoming saturated as it was rolled out. One cable provider described the urgent tenor of the situation as follows:

> The example that comes to my mind is the case in which there was a university town . . . and the kids had come back to school somewhere around late August/early September time frame. And [the network engineers] saw this huge influx of traffic, even much more so than they would have expected after the kids have left at the end of the spring semester. I think there were emergency crews out there putting out new fiber nodes, splitting networks, and doing it all on an emergency basis because the network had quickly gotten bogged down with a traffic load that it couldn't handle. And it was a traffic load that it could have handled with the same population of people, say, three months before. And I think that that was the first time that, at least in my experience, the reality of sudden congestion or sudden rise of traffic started to become very real.

The upstream contention problem was acute for cable. Performance of other applications was suffering as a result. Some operators noted that upstream contention was taking a

disproportionate toll on real-time applications such as VoIP that require minimal latency in both directions in order to provide reasonable call quality. Martin and Westall (2007) found that just 15 active BitTorrent users on a cable link shared among 400 total users could cause VoIP call quality to fall below a usable performance threshold. Operators were viewing the growth of upstream traffic as unsustainable from a capacity planning perspective. They were in search of a fix.

One solution that gained significant traction involved deploying deep packet inspection (DPI) technology to limit the number of upstream peer-to-peer connections at particular points on the network. This was the approach that Comcast pursued that eventually became subject to regulatory scrutiny. Sandvine, a network equipment vendor, supplied the DPI equipment that Comcast deployed in an "out-of-line" fashion: Internet traffic was copied and sent to the Sandvine devices for inspection, rather than having the devices sit directly between end users and the rest of the Internet (Zachem 2008). The equipment was configured to identify protocols associated with heaviest traffic usage, all of which were peer-to-peer protocols, including BitTorrent. When the number of upstream-only connections associated with any of the protocols reached a pre-defined threshold in a particular location on the network (indicating potential congestion on the upstream), the device would send a TCP "reset" packet to both sides of the peer-to-peer file exchange, causing the exchange to cease. This solution approach was viewed by many observers as discriminatory because it involved selecting traffic associated with particular Internet applications for different treatment from all other traffic (Ammori et al. 2007a; Copps 2008; Goldberg, Kumar, and Monahan 2007).

Comcast first trialed the Sandvine equipment in May 2005 and moved to network-wide commercial deployment that lasted from 2006 to 2008 (Zachem 2008). According to numerous interviewees, a number of other cable operators pursued this or very similar application-specific approaches within a similar time frame. The limited existing network-based research into detectable interference with BitTorrent traffic supports this conclusion as well (Dischinger et al. 2008).

In some cases, the configurations of the devices on the network and the identified protocols differed. For example, the cable provider RCN used Sandvine equipment to manage peer-to-peer traffic in two different ways that did not involve TCP reset packets (Kiddoo 2010). First, RCN redirected requests for peer-to-peer downloads from other networks onto its own network to avoid having to pay the transit costs associated with transferring the data from another network to one of its own customers. Second, the company limited the number of simultaneous upstream peer-to-peer connections that could occur within each geographic market in which the ISP operated. Because the Sandvine devices were installed "in-line" – directly between subscribers and the rest of the Internet – these connection limits could be enforced without the need to send TCP reset packets. Additional upstream requests above the pre-determined threshold could simply be blocked by the Sandvine device. Interviewees likewise confirmed that deployments varied between operators and that some providers chose to use the in-line approach to peer-to-peer management rather than having individual connections terminated via a TCP reset. The out-of-line approach was much easier for subscribers and other external parties to detect because the TCP reset packets had an easily identifiable signature.

According to interviewees and available public information, application-agnostic approaches to managing network traffic, such as usage-based billing or tactics based on per-user traffic volume, do not appear to have been widely deployed. One solution, as described by a cable interviewee, took a less application-focused approach, but could still be considered "application-specific" under the definition of that term provided in Chapter 1. The solution involved identifying traffic based on characteristics other than the application with which it was associated. In this approach, routers on the cable network would identify the network to which each packet was destined. If the destination network was another residential network (as opposed to a network used primarily to host web sites, for example), that was the first clue that the traffic was likely peer-to-peer. If the packet's size and transport protocol were also determined to be characteristic of peer-to-peer file transfers, the packet was then assumed to

be part of a peer-to-peer exchange. Traffic flows containing such packets were then rate-limited in the upstream, helping to control upstream utilization.

This approach did not involve identification and differential treatment of specific applications, but it did discriminate on the basis of the destination of the traffic. Furthermore, some of the effects were likely similar to approaches that did single out specific peer-to-peer applications. Thus this example fits with the broader trend of large cable operators pursuing application-specific traffic management solutions. And as described by the cable interviewee, it was deployed in tandem on the network with another solution that did target specific peer-to-peer applications using Sandvine equipment.

Motivations Behind Application-Specific Management Decisions

By the time applications-specific solutions were being deployed, there was a growing debate in the press and in Congress about net neutrality and the potential for discriminatory conduct in the wake of the FCC's deregulatory activities. Why would cable operators take up application-specific traffic management solutions under these circumstances?

One feature that all of the approaches described above had in common and that made the TCP reset approach particularly compelling was that they were viewed as highly targeted, pre-packaged solutions that could alleviate, on relatively short order, what was becoming an emergency bandwidth problem for cable operators. The theory behind the TCP reset approach was to narrowly focus on the network's pain point – high upstream utilization caused in large part by peer-to-peer traffic – while having a minimal impact on the peers whose file transfers were interrupted. Most peer-to-peer clients are designed to automatically re-start a download when it gets interrupted (assuming another peer elsewhere has made the same file or file portion available). Assuming that the downloading peer would re-start by connecting to a peer on a different ISP or in a less congested part of the cable operator's network, the TCP reset approach would have the effect of moving upstream peer-to-peer traffic out of the trouble spot while introducing only a small delay in the time it took for the downloader to

finish the file transfer. Uploaders were assumed to be mostly indifferent as to whether their uploads were interrupted or how quickly they finished.

Thus the TCP reset solution appeared likely to solve cable operators' specific problem immediately and persistently without significantly degrading the performance of peer-to-peer transfers. The other approaches described above – in-line peer-to-peer management and management based on traffic characteristics other than application – were similarly designed to narrowly focus on peer-to-peer traffic, although their performance effects may have been more variable since they involved dropping packets or preventing connections for some period of time rather than terminating entire peer-to-peer connections. To operators, all of these approaches seemed more focused on solving the existing problem than other conceivable nondiscriminatory approaches. Usage-based billing, for example, would have required accounting infrastructure and customer education without guaranteeing that it would actually cause heavy users to change their behavior or how long such a change might take. Capacity upgrades were ongoing, but they were perceived to be insufficient to keep up with traffic demand given their cost. By comparison, the approaches that the cable operators pursued appeared to offer more certainty, more immediate results, and more lasting efficacy.

Furthermore, solutions based on DPI were extremely attractive to the operators' marketing teams, which saw great appeal in using DPI not only for traffic management, but to understand how their customers were making use of their Internet service and how that usage might be parlayed into new product opportunities. At several ISPs, marketing teams purchased the equipment with little oversight from engineering staff. As one interviewee explained:

> [T]he people on the engineering side, if you ask them about what the experience of . . . installing [the DPI] was, there wasn't any of the typical consultation with engineering. It was, "We've decided to buy this system, you'll be installing it. Only you guys will know about it. Here's the budget, please get it done as soon as possible." And [engineering] looking at it and being like, "What the fuck is this? How do we configure this? What is this?" They had no vote in it at all, no influence.

Thus in some cases the fact that the DPI platforms were capable of providing marketing units with intelligence about network usage was an even more compelling reason to make a DPI purchase than any reason related to traffic management. As Marsden (2010) has argued, DPI's multiple uses make it an attractive investment. Nondiscriminatory traffic management solutions by their very nature could not have provided the same kinds of application-level insights as DPI-based platforms.

Finally, cable operators took a somewhat cavalier approach to their regulatory circumstances. Cable interviewees emphasized the streamlined structures of their organizations, describing operations as "bootstrapped" and "lean" with a "fly-by-the-seat-of-your-pants" culture. The cable operators had (compared to the telcos) relatively few personnel resources dedicated to developing interpretations of ambiguous regulatory policy and applying those interpretations to internal traffic management decision-making. In some cases regulatory staff exercised an extremely light touch, as exemplified by the following exchange with a cable public policy executive:

> Interviewer: [W]hat was the timing of your role – at the end of something you would come in and say that we should think about the public policy implications of this? Or was it kind of earlier? . . . How did it integrate into a decision to roll something out on the network? Or did it integrate at all?
>
> Respondent: Generally no. Other than ensuring that the business units and the engineering teams were acquainted first with the Four Freedoms as articulated by Michael Powell. And we always felt comfortable in asserting that we were operating our network consistent with those freedoms. And then in the *Policy Statement* articulated by Kevin Martin and his Commission, you know that consumers could go anywhere on the Internet that they wanted, use any content, connect any device that didn't interfere with the network . . . . But other than ensuring that those principles were known and understood and part of our daily practice, we didn't have much occasion to discuss specific network management decisions.

Even in cases where regulatory and public policy staff had the opportunity to develop internally their conceptions of "reasonable network management," those conceptions were sometimes quite broad. For example, some cable interviewees expressed that after the *Policy Statement* took effect, they assumed that any practice that served the purpose of providing a good user experience, or that was not anti-competitive, should be considered "reasonable."

The mere existence of the carve out for network management, and in a statement of questionable enforceability no less, gave some operators the confidence to pursue application-specific approaches. One cable interviewee explained this in discussing his company's dealings with Chairman Martin's staff regarding the reasonable network management footnote:

> [T]hat footnote was very important to us and we worked with them on that. And also I don't think it was in the document itself, but maybe it was in his press release or his statement where he specifically said, "this isn't enforceable." So those two things gave us a fair amount of comfort. Everything that was in there we felt that we were already doing at the time. So we were thinking that this set a framework, put things to rest.

During the *Open Internet* proceeding, a number of cable operators provided clear articulations of just how broadly they conceived of the term "reasonable." Cox exhorted the Commission to "establish a presumption that properly disclosed network management practices are reasonable, rebuttable only by evidence that the management tools are an artifice for anti-competitive conduct" (Wilson et al. 2010, 32). Time Warner Cable argued that the definition of "reasonable network management" should consist of any practice employed to reduce congestion, prevent the transfer of unlawful or unwanted traffic, or "any other network management practice that is intended to improve service quality or performance rather than to achieve any anti-competitive objective" (Teplitz et al. 2010, 71). Where cable operators were internally assuming such standards under the *Policy Statement*, the application-specific approaches seemed justifiable.

In short, the cable industry was facing a significant upstream engineering problem that even an accelerated schedule of capacity upgrades was not able to fully handle. A solution presented itself that was not only narrowly targeted, but had additional benefits in the eyes of marketing teams, some of whom made the actual equipment purchases. All of this came to pass in the absence of rigorous oversight from policy and regulatory staff.

Where policy and regulatory oversight did exist, interpretations of existing regulatory ambiguity and the likelihood of regulatory intervention gave operators tremendous leeway in

choosing traffic management solutions. Such an approach to regulatory risk made sense in the context of the cable industry's brief and successful regulatory history. Cable broadband operators had never been subject to nondiscrimination obligations under Title II. Just weeks after Comcast began its Sandvine trial in 2005, the Supreme Court affirmed this regulatory classification. The cable industry could logically conclude that discriminatory traffic management approaches, including the TCP reset solution, were well within the bounds of what the FCC might consider "reasonable." That conclusion ultimately proved to be untrue and has shaped broadband Internet regulation ever since.

5.3.2 Telephone Industry

Absence of Application-Specific Management

By and large, the telcos refrained from pursuing the kinds of application-specific traffic management approaches that cable companies had deployed. This difference in approach was partly motivated by differences in access technologies – the urgent need to find a palliative for peer-to-peer upstream congestion simply did not materialize on DSL and fiber networks. But the telcos' decisions were likewise shaped by their specific regulatory history and obligations, just as the lack of such factors shaped the decisions of the cable companies.

The telcos were seeing the same kinds of broadband traffic growth in the early and mid-2000s as the cable companies saw, including surges in peer-to-peer traffic, but their networks bore the brunt of the growth differently. Residential DSL networks offer each subscriber a dedicated link in the access network, pushing the point at which multiple subscribers' traffic gets aggregated further up in the network where greater bandwidth is available to manage the aggregate load. As a result, DSL subscribers were much less likely to see the performance of their access network connections suffer when a small number of subscribers in the neighborhood chose to make aggressive use of peer-to-peer applications. Furthermore, DSL offerings tended to be less asymmetric than cable offerings, with the ratio of advertised upstream to downstream bandwidth around 1:4, whereas for cable it tended to be closer to

1:10 (Greenstein and McDevitt 2011). Thus the specific impact of the growth of peer-to-peer on upstream bandwidth was more muted for the telcos.

In addition to offering DSL broadband, beginning in 2003 the nation's two largest telephone companies began deploying new broadband products that supported video, voice, and Internet services over fiber networks: AT&T's U-Verse and Verizon's FiOS. These services provided tens to hundreds of times the access network capacity offered by DSL or cable at the time, giving the operators a much larger bandwidth buffer with which to handle traffic upsurges and relieving them of the urge to pursue fine-grained management of individual Internet applications.

In the absence of the kind of emergency bandwidth crunch that the cable companies were experiencing, the telcos were left without a meaningful impetus to begin managing peer-to-peer or any other specific Internet applications on their DSL networks. Overall, there is little evidence that the large telcos were discriminating against Internet applications in any significant way from the mid-2000s through the end of the decade. All telco interviewees emphasized this point.

It may be tempting to explain away these claims by way of interviewees' wariness to admit to engaging in practices that could reflect poorly on their companies, but their claims provide a sharp contrast to those from cable employees, many of whom openly discussed discriminatory practices even in cases where those practices had not attracted public attention. Telco interviewees made convincing arguments, often citing specific corporate executives, policies, or decision instances that governed traffic management choices. One telco executive put it this way: "What I can say, and as a company [we] have stated this, is that we do not impede any public Internet traffic. Our opinion is a bit is a bit."

In the case of fiber deployments, network operators did take steps to manage the interaction between their own IP-based video offerings and Internet traffic, but did not apply those management techniques to Internet applications themselves. The video services offered as

part of the fiber deployments were designed to compete with cable television service, and as such the telcos sought to match the user experience of a customer clicking through cable channels. But unlike cable, the fiber networks were engineered to have capacity shared between Internet traffic and linear video and/or video-on-demand traffic. To ensure that video traffic would not squeeze out Internet bandwidth in a household with multiple people watching different HD or SD programs in different rooms, the network operators adopted management strategies that involved prioritizing their own video streams over Internet traffic and always reserving a portion of capacity dedicated to Internet traffic (AT&T 2012; Verizon 2012). These management strategies had the potential to impact a subscriber's Internet service when multiple video streams were in use, but they involved no specific management of Internet traffic. This was a deliberate choice that the network operators made.

Motivations Behind Traffic Management Choices

These choices were shaped as much by regulatory circumstances as by technological design. Compared to cable in the mid-2000s, the telcos arguably had a higher risk of having an enforcement action brought against them on the basis of a claim of discriminatory conduct. Until August 2005, the telcos were operating in many cases under the nondiscrimination obligations of Title II and in all cases under the remaining obligations from the *Computer Inquiries*. It was far from clear what these obligations meant for traffic management – the *Madison River* enforcement action (FCC 2005a) was the only remotely relevant case concerning Internet applications, and it dealt with explicit blocking of all traffic associated with a particular application, providing little guidance about whether more nuanced approaches to traffic management might be considered discriminatory.

With the *Wireline Broadband Order* (FCC 2005b), the remaining obligations were lifted and replaced with the *Policy Statement* of questionable enforceability. But just months later, Verizon and AT&T – comprising 70% of the nation's DSL market (Noam 2009) – each agreed to abide by the *Policy Statement* for two years as conditions of their respective mergers with MCI and SBC (FCC 2005d; FCC 2005e). While they may have had no more

clarity than the cable operators about what "reasonable" network management was, they had far more certainty that if they were seen to be violating the principles of the *Policy Statement*, the FCC could take action against them.

In AT&T's case, the FCC extended that commitment even longer and imposed further net neutrality conditions when the company acquired BellSouth in late 2006 (FCC 2006). Although the new conditions, which lasted until the end of 2008, did not speak to traffic management specifically, they required that the merged entity "maintain a neutral network and neutral routing in its wireline broadband Internet service," (FCC 2006, 154) further increasing the risk that the FCC might take action should it learn of "non-neutral" practices.

As a result, the risk of regulatory enforcement could never be confidently disregarded. This risk did not completely quash the impetus to pursue traffic management approaches that involved differential treatment of different services – otherwise the video prioritization and bandwidth reservation techniques employed on fiber networks never would have materialized. But it did result in the institutionalization of corporate practices designed to mitigate regulatory risk. These organizational processes were integral to the telcos' decisions not to pursue application-specific management, and they reflect a stark contrast in approach to regulatory risk as compared to the cable operators.

Within the telephone companies, regulatory review was thoroughly embedded in the process of traffic management decision-making. As one telco engineer explained, "our policy folks always have their antennas up" on how developments in the policy space might affect activities on the engineering side of the company. One policy executive likewise explained that as a policy team, "ideally you come in as early as possible or practicable" to review new product developments. Another policy-focused interviewee described how "every time they rolled out a new feature . . . it was pretty much par for the course that [the engineers] would bring me in and we would talk to them and advise them on whether we thought there were

issues." Internal regulatory oversight is one feature that all the large telcos appear to have had in common.

This oversight was not only manifest in the usual exchanges between regulatory or policy teams and other corporate divisions, but also in bureaucratic forms that were developed even before net neutrality became an issue of public debate. One company convened an internal policy council that brought together regulatory, policy, marketing, legal, and technical staff to discuss traffic management decisions and ensure company-wide understanding of new traffic management capabilities. Another operator had policy personnel embedded from the very beginning within the team developing new broadband products so that a policy perspective could be provided as engineering, architectural, and business decisions were made along the way. Regulatory and policy oversight was enshrined in corporate organizational structures.

Might decades of regulation provide some explanation for why the telcos invested this effort in regulatory review? Year after year of operating under a considerable regulatory regime engendered a sensitivity to regulatory risk within these companies that shaped them profoundly, right down to the ways in which policy personnel were organized internally. In some cases, the risk of regulation appears to have penetrated even further, down to the very core of corporate philosophy. A number of telco employees spoke about how corporate philosophy, as articulated by senior executives, circumscribed the set of traffic management choices available to engineering teams. The precise interplay of the forces of regulatory risk and corporate philosophy is difficult to discern – which came first, the chicken or the egg? But as one former senior telco executive explained, there is no doubt about their combined contribution to the decision not to manage Internet applications:

> It's been [this company's] philosophy for as long as I've been promoting and pushing this philosophy that our job is to build capacity to satisfy the customer. And when you take that and you couple it with the regulatory – I'm going to call it overhang but it's probably not the right word – of net neutrality concerns, threats, we have stayed as far away from attempting to modulate network traffic as possible. If we're going to for some reason have congestion in the network, everyone's going to suffer from that congestion until we can physically get it fixed, either through true hardware or optimization techniques to get everybody back together. So . . . we have done no traffic shaping, traffic blocking, if you will.

Such philosophies were naturally not uniform across the industry or its executives. Former AT&T CEO Ed Whitacre famously ignited controversy when suggesting that rather than getting a "free lunch," large applications providers should be required to pay for preferential traffic treatment (O'Connell 2005), and his rhetoric was echoed by others in the industry, including Verizon senior vice president John Thorne (Mohammed 2006) and Qwest CEO Richard Notebaert (Reardon 2006). But there was clearly a gap between this rhetoric and the reality of how the telcos actually chose to manage fixed network traffic. When it came to traffic management techniques that had no revenue potential attached, the notion of staying away from modulating particular applications clearly pervaded the corporate mindset in some cases.

In sum, the telcos' initial and enduring approach to traffic management – "a bit is a bit" – can be explained by the combination of technology, regulatory risk, and, in some cases, corporate philosophy. In some instances they had application-specific tools available that they used to manage the performance of their video services, but they chose not to extend those tools to over-the-top services. While the technical imperative to implement application-specific traffic management responses was lacking, that did not prevent the telcos from injecting regulatory oversight into traffic management discussions early and often, in stark contrast to large cable operators.

5.4 The *Comcast Order* and its Aftermath

Cable operators – or some cable operators, at least – were about to drastically revise their approaches, however. Comcast did not disclose its use of the Sandvine TCP reset solution when it was deployed on the network. In May 2007, Robb Topolski, a Comcast customer, wrote about his peer-to-peer connections getting reset in the DSL Reports online forum (Topolski 2007). Over the following months, a number of parties conducted investigations into interference with peer-to-peer traffic and published their findings online, culminating

with an Associated Press (AP) article published in October 2007 that documented nationwide tests the AP ran to demonstrate Comcast's use of the TCP reset technique. The following month, public interest groups and others petitioned the FCC to take action in response. After months of dramatic public debate (discussed in Chapter 8), Comcast began developing FairShare, a "protocol-agnostic" approach to traffic management. In August 2008 the Commission ordered Comcast to cease its use of the TCP reset solution by the end of the calendar year. By January 2009, FairShare had replaced the TCP reset solution throughout the Comcast network.

Rather than managing peer-to-peer uploads, as the TCP reset solution had done, FairShare was designed to reduce the impact of heavy users during times of high network utilization. Comcast deployed software to monitor network utilization in the aggregate and on a per-subscriber basis (Zachem 2008). When a network segment was nearing a congested state, the software would determine if one or more subscribers had been using disproportionate bandwidth, and those subscribers' traffic would be assigned to a lower priority level than the rest of the subscribers sharing the network segment. Thus this approach mitigated the impact of heavy users on the rest of the user base without regard to which applications were in use or which applications might be the source of disproportionate traffic. And rather than terminating individual connections, it allowed all heavy users' traffic to be exchanged, but ensured that other users' traffic would get priority. When the heavy subscribers' usage returned to normal levels, their traffic would return to normal priority.

For other cable companies that had been using similar approaches to Comcast, the FCC proceeding and the public attention surrounding it by and large spelled the end of application-specific traffic management. RCN agreed to cease using its in-line peer-to-peer management techniques in 2009 as part of a class action settlement that arose as a result of the public attention that Comcast's traffic management had accrued (Kiddoo 2010). One cable executive explained the effects of the public debate as follows:

> [W]e were looking at pushing [TCP resets] further as the FCC and Comcast had their brouhaha. And it was determined that Comcast had to stop. And in that case, sort of as the technology policy intersects with Washington policy, the uncertainty that Washington caused put a hold on us deploying further, and eventually caused us to stop. So we stopped packet resets . . . at year end whenever that year was when the FCC had ordered Comcast to stop.

The Washington policy debate may have put an end to management focused on peer-to-peer applications, but that did not mean that the kinds of traffic demands that had spurred its adoption in the first place had gone away. Cable operators still needed to manage that demand, and by and large they turned back to capacity upgrades, opting to "let standard capacity planning go back in place," as one engineer put it. But operators did not necessarily find themselves back in emergency node-splitting mode.

Even as cable operators had been pursuing the TCP reset approach, a new technical standard for cable modem design, known as Data Over Cable Interface Specifications (DOCSIS) 3.0, was being finalized. By bonding together multiple cable channels for use in delivering Internet traffic, DOCSIS 3.0 could supply cable broadband networks with several times more capacity both upstream and downstream than what was available previously. It required that equipment both in subscribers' homes and in the cable network be upgraded, but once those upgrades occurred, increased bandwidth was immediately available.

At the same time, patterns of network usage were beginning to shift. Web-based streaming video sites like YouTube and Hulu were maturing and attracting increasing numbers of visitors. Cisco estimated that 20% of US Internet traffic in 2007 was YouTube traffic (Cisco 2008). The relative portion of network traffic attributable to peer-to-peer applications was shrinking as a result, at least in the downstream if not in the upstream (Erman et al. 2009). Heavy peer-to-peer use could still wreak havoc on an under-provisioned network link, but it was no longer the main driver of traffic growth.

Thanks to both of these developments – and with the TCP reset solution off the table – cable operators could return to more pure capacity-focused approaches to managing network load

without heading straight back into the urgent upgrade cycles of years past. As one engineer enthusiastically explained, "DOCSIS 3.0 is the knight in shining armor that arrived onto the scene in the nick of time."

This is not to say that other traffic management approaches were not discussed and tested. The most prominent example was that of Cox, which in 2009 ran a trial of a system that distinguished between what the company considered to be "time-sensitive" applications and "non-time-sensitive" applications and prioritized the former over the latter in the upstream direction (Wilson et al. 2010). During times of congestion, non-time-sensitive applications (peer-to-peer applications, newsgroups, and other similar bulk file transfer applications) would have their speeds reduced to ensure that the performance of time-sensitive applications (web, VoIP, streaming, and all other traffic not classified as non-time-sensitive) was unaffected. The trial ran for a limited number of months on a portion of the Cox network and was never deployed network-wide.

In sum, with the watchful eye – or raised eyebrow? – of the FCC looming over them, cable operators by and large abandoned application-specific management techniques, taking advantage of DOCSIS 3.0 to help them handle the evolving traffic profiles on their networks.

5.4.1 Broader Effects of the *Comcast* Proceeding

The effects of the *Comcast* proceeding reached far beyond the particular choice of traffic management strategy for some cable operators. Across the broadband industry, it swiftly raised awareness of the potential for regulatory backlash associated with application-specific traffic management. The telcos were by and large not engaging in practices that might earn them the FCC's ire, and the *Comcast Order* reinforced their decisions to steer clear of any practice that might raise questions before the Commission. This was true even after the expiration in 2007 and 2008 of the merger conditions that had made FCC enforcement a more plausible threat for Verizon and AT&T. Comcast's ordeal at the FCC was enough to give the telcos pause even in absence of those conditions.

As traffic demands evolved over the years after the *Comcast Order*, the telcos leveraged their existing organizational structures to ensure that proposals for new traffic management practices received sufficient regulatory review. For some cable operators, the *Comcast* proceeding crystallized a need to develop similar processes. Numerous cable interviewees noted that their legal and policy staff became more involved in traffic management decision-making during and after the proceeding. The ordeal "had the effect of bolstering government and legal affairs teams . . . at least their ability to impact network management decisions," as one cable engineer described it. Another noted the "huge amount of time" he began to spend visiting with regulatory and policy staff in Washington.

The organizational adjustments were not limited to increased interaction between policy teams and engineers. One interviewee cited a trickle-down effect, where increased caution on the part of policy staff caused the finance department to view new research or development work with skepticism, withholding budget money for engineers to experiment with new approaches to traffic management in the lab. Engineering thus became constrained with both regulatory oversight and tighter budgets. But the opposite phenomenon also materialized: where marketing teams had been driving the charge towards DPI-based platforms that could support both traffic management and network intelligence, there was a distinct shift in leadership back towards the engineering side of the business. One cable engineer recounted such a change vividly:

> [E]ngineering was sort of back in the driver seat on all of the major decisions related to what we do in the network. And if someone in engineering wanted to raise their hand to . . . our leadership and say, "Hey, this isn't the right thing to do," or "This is going to really blow up in our faces," or "This isn't solving, really, the underlying problem," we sort of have that veto authority then. Whereas before, [the marketing teams] were like, "That's great. Just install the shit. You know, you're the network plumber. Shut the fuck up." That was kind of the view. . . . We, from that point forward, had an equal seat at the table.

Clearly, there was a belief within some organizations that allowing the marketing units too much control over these kinds of decisions had contributed to a choice of traffic management technology that was much derided by the public and regulators.

The *Comcast* proceeding and its aftermath profoundly shaped broadband and its regulation in the US, and many of its other effects are dealt with in later chapters. As for the substance and process of traffic management decision-making, it put an end to what application-specific management did exist in the US. It put broadband providers of all stripes on high alert that the FCC was taking a hard line against discriminatory treatment of Internet traffic. And it prompted some organizational transformation within certain cable operators, introducing greater regulatory oversight into traffic management decision processes and rearranging the power relationships between different corporate units. As such it left an indelible mark on the way that US broadband providers, and cable operators in particular, select the methods they use to manage Internet traffic.

5.5 Conclusion

The period following the *Comcast Order* was tumultuous from a legal and regulatory perspective (see Chapter 8). Comcast sued to have the order overturned. Barack Obama was sworn in as President just months later, having campaigned for office on a platform that included explicit support for net neutrality. The FCC, under its new chairman Julius Genachowski, launched a proceeding to develop open Internet rules. While it was ongoing, the D.C. Circuit upheld Comcast and ruled that the *Policy Statement* was unenforceable, causing the FCC to explore reclassifying broadband Internet service under Title II. Leaders in Congress and at the FCC held meeting after meeting in search of a compromise set of rules that a wide range of stakeholders would support. The Congressional talks broke down, paving the way for the FCC to adopt the *Open Internet* rules without reclassifying broadband. Legal challenges immediately ensued.

Amidst all of this uncertainty and upheaval, traffic management practices remained largely unchanged. Network operators were leery of ruffling feathers in Washington, and regulatory and policy staff had increased insight into potentially controversial engineering changes. US broadband users rode out the decade largely free of application-specific management.

The experience of US broadband users in the first decade of the 21st century, characterized by a relative lack of application-specific traffic management, arose out of the interplay between technological and regulatory forces. The engineering specifications of different access network technologies contributed to the difference in the degree of urgency with which network operators considered managing specific Internet applications. But this urgency, or lack thereof, was couched within existing organizational conceptions of regulatory risk and organizational institutions for addressing that risk. The decisions that operators made were no more determined by the architecture of their networks than by the amount of time and resources they had spent to obtain a particular regulatory classification. The combination of these forces secured a relatively "neutral" traffic management experience for US broadband users. As the next chapter will demonstrate, the same cannot be said for the United Kingdom.