CHAPTER 2

BACKGROUND

This thesis builds on prior research in three areas: traffic engineering, content delivery, and the interaction of the two. We first review the major classes of traffic engineering strategies (Section 2.1). Next, we discuss common techniques used by content delivery systems, including strategies for content placement and for request redirection (Section 2.2). Finally, we survey research on the interaction between network and content delivery, and find that the interaction between traffic engineering and either overlay routing or request redirection have been the focus of past research (Section 2.3).

2.1 Traffic engineering

The goal of traffic engineering (TE) is to avoid congestion hotspots in the network by optimizing routes based on network topology and expected traffic demand. In the context of large Internet service provider (ISP) networks, traffic engineering decides both intra-domain (within the ISP) and inter-domain routing (across ISPs). We focus here on intra-domain routing and refer the reader to [48, 104] for a survey of inter-domain traffic engineering.

We classify traffic engineering schemes based on the frequency at which they update routing. By this attribute, TE schemes can be grouped into three categories: (1) *Oblivious TE* uses static routes that are seldom updated [20, 22]. (2) *Offline TE* updates routes periodically, e.g., every few hours or every few days, based on recent history of traffic demand [50, 49]. (3) *Online TE* updates routes at timescales of hundreds of milliseconds, reacting instantaneously to traffic demand changes [70].

TE schemes are evaluated based on link utilization based metrics, e.g., a widely used metric is maximum link utilization [117]. A TE scheme is usually compared against the optimal solution that minimizes the given metric by solving a multi-commodity flow optimization [70, 50]. By this measure, oblivious routing schemes perform poorly and can be shown to be arbitrarily worse compared to the optimal strategy [50]. For many ISPs networks, simple oblivious routing schemes are sub-optimal by a small constant factor [117, 122, 70]. Offline TE schemes, while sub-optimal, perform superior to oblivious TE schemes. e.g., Fortz and Thorup show that offline TE delivers up to 2$\times$ better on AT&T network backbone. Online schemes have been shown to achieve near-optimal performance, but they are rarely used in production networks.

In practice, offline TE based on Open Shortest Path First (OSPF) and Multiprotocol Label Switching (MPLS) are commonly used [117, 122, 50, 45]. Routes computed by OSPF traffic engineering must follow shortest-weight paths, therefore OSPF TE provides limited functionality to split traffic among multiple paths. MPLS TE overcomes this limitation by enabling traffic between two nodes to be split in arbitrary ratios among multiple paths. Therefore, MPLS TE gives better results than OSPF TE [117, 122].

Prior work on traffic engineering is based primarily on evaluation of link utilization based metrics, and has largely ignored the impact of traffic engineering on user-perceived performance. Further, the comparison of TE schemes has not taken into account the interaction with content delivery. This thesis contributes in answering these questions. In Chapter 3, we provide a comparison of traffic engineering schemes focusing on user-perceived metrics such as file download times, and VoIP call quality. In Chapter 3 and Chapter 4, we evaluate TE schemes while accounting for the interaction between TE and content delivery and show that this interaction helps simpler TE schemes, such as oblivious TE or OSPF TE, perform closer to the optimal TE strategy in terms of user-perceived metrics as well as TE metrics.

2.2 Content delivery

Content delivery systems seek to provide a high-quality experience to users accessing content in all regions at all times. A canonical example of a content delivery system is a content delivery network (CDN). State-of-the-art CDNs operate geo-distributed datacenters, and use a combination of edge caching, intelligent server selection, and path and protocol optimizations for delivery of several types of content, e.g., video, bulk downloads, and interactive websites [41, 86]. Given their geo-distributed deployment, the decisions of content placement, i.e., locations at which a content is placed, and request redirection, i.e., which location is best positioned to serve a user's request, are central to the functioning of a CDN.

2.2.1 Content placement

Placement strategies depend on whether content is static or dynamic. Dynamic content has limited cacheability therefore placement strategies for static content are not always applicable for dynamic content.

2.2.1.1 Static content placement

Static content, such as videos, audio, images and software updates, contribute to a vast majority of traffic in the Internet [85, 33]. The placement of static content is commonly handled by a caching strategy. A simple and widely used caching strategy is least recently used (LRU) cache replacement [119]. Caching strategies are effective because they exploit geographic and temporal locality of requests, resulting in high cache hit rates in many cases [38, 109, 57]. An alternative to caching is a planned placement approach, which prepositions content at a set of locations based on prior knowledge of demand. Planned placement is effective in scenarios where workloads are predictable over long intervals, e.g., hours, or days [19].

2.2.1.2 Dynamic content placement

Several applications today generate dynamic content such as stock prices, weather information, price catalogs. Such dynamic content is typically stored at a small number of fixed locations across the globe, mostly for fault tolerance objectives [7]. Due to a limited number of fixed replica locations, content accesses from regions away from the replica locations incurs high latency [7, 14]. Extensively replicating dynamic content is costly due to bandwidth and server resources used in propagating updates to all locations.

Placement strategies for dynamic content is an active research area. A naive placement strategy of replicating all data at all locations would incur high update costs. The alternatives provided by current systems either require manual configuration to decide placement or result in sub-optimal latency. For example, Spanner [36] provides configuration options to manually select which locations should a given subset of data be replicated. DHT-based systems automatically decide placement but result in high latency because replica are chosen randomly [99, 98]. Volley [14] uses a placement heuristic to select a single best location for each data item, but it would result in sub-optimal latencies when a data item is popular across many geographic regions.

This thesis addresses the problem of dynamic data placement across geo-distributed data centers while enhancing prior work in this area (Chapter 6). Our system, Auspice, automatically makes data placement decisions (unlike Spanner), places data replicas based on demand-locality (unlike DHT-based replication), and creates multiple replicas of each object (unlike Volley) and limits update propagation costs.

2.2.2 Request redirection

Request redirection strategies complement placement strategies by selecting the server location that is best suited to process a user's request. These strategies have been extensively studied and form the heart of CDN technology today. To quote from a report by Akamai, *"the system directs client requests to the nearest available server likely to have the requested content."* where the "nearest" server is one whose round trip latency as well as packet losses are small, and an "available" server is one that is lightly loaded considering all resources, i.e., network, CPU and disk [41].

Request redirection is implemented using three processes: (1) *Monitoring:* Probe messages sent intermittently help monitor network characteristics and server load and identify congested regions of network and overloaded server locations [56, 118]. (2) *Estimating distances:* The measured statistics are combined to compute a distance function that reflects the proximity of a server location to users in a geographic region [118]. (3) *Informing the user:* The user is informed of selected server/s either via DNS resolution or via HTTP redirection as described in [41] and [23].

2.3 Interaction between network and content delivery

Studying the interaction between network and content delivery has been a topic of much interest in both systems and theory communities. Several related questions have been put forth. Do these interactions negatively affect objectives of networks and content delivery systems? What is the sub-optimality caused due to these interactions in the worst case, and for typical topologies and traffic demands? How to leverage these interactions to improve traffic engineering and content delivery objectives?

Yet, we don't fully understand these interactions because prior research has studied the interaction of network routing with only a subset of content delivery decisions. Much prior research has focused on two aspects: the interaction of overlay routing and network routing [106, 95], and the interaction of request redirection and network routing [69, 42, 54, 121]. While placement decisions are critical to user-perceived performance, there has been little research on how content placement interacts with network routing.

2.3.1 Interaction between traffic engineering and overlay routing

Several results show the negative interaction between selfish overlay routing and network routing [106, 95], however it appears that selfish overlay routing is not used by most of the Internet traffic. For example, traffic from CDN edge server to the client always follows network routing. Further, overlay routing yields "marginal" benefits ($<$ 30%) over network routing for 79%-96% of paths depending on which geographic region is being considered [97], which suggests that traffic between CDN servers forming an overlay network follows network routing in most cases. For this reason, this thesis does not model the interaction between overlay and network routing.

2.3.2 Interaction between traffic engineering and request redirection

Recent research has investigated the interaction between request redirection and traffic engineering, without considering the role of placement strategies. This interaction is commonly studied in the context of Internet service providers (ISPs) and content providers (CPs) with geo-distributed datacenters. Both analytical results [69, 42] and system implementations [54, 121] have shown that there is value for joint optimization of request redirection and traffic engineering, and cooperative strategies can help traffic engineering metrics and also reduce user perceived latencies. Commonly, these efforts assume that all content is available at all locations, ignoring the fact that content availability at a location depend on placement strategies. Therefore, in this thesis, we account for the effect of content placement along with request redirection, in studying the interaction between network and content delivery.

2.4 Summary

In this chapter, we reviewed prior research on traffic engineering and summarized different classes of traffic engineering schemes. We also reviewed common techniques used for content delivery, namely strategies for placement of static content and of dynamic content, and request redirection techniques used by CDNs. Finally, we reviewed research on interaction between network and content delivery, and found that much prior research in this area has focused on two aspects, the interaction of overlay routing and network routing, and the interaction of request redirection and network routing.