ENTERPRISE MULTIMEDIA STREAMING: ISSUES, BACKGROUND AND NEW DEVELOPMENTS
D. S. Turaga¹, M. van der Schaar² and K. Ratakonda¹
¹IBM T. J. Watson Research Center, Hawthorne, NY, 10523
²University of California, Davis, CA 95616
{turaga, ratakond}@us.ibm.com, mvanderschaar@ece.ucdavis.edu

1. INTRODUCTION: ISSUES IN ENTERPRISE MULTIMEDIA STREAMING
An enterprise network is a private network linking sites within an organization (typically a company). It may be widely dispersed, multifaceted, and includes all the networking services and equipment of the organization. While the streaming of multimedia over enterprise networks is gaining momentum, there is a disconnect between the needs of such enterprise streaming systems and current research in streaming media. Current research has focused mainly on the Internet, and while enterprise multimedia streaming does faces many challenges similar to Internet multimedia streaming, there are additional requirements and advantages specific to the enterprise streaming model. Firstly, since the enterprise owns and can control the network, the infrastructure, e.g. content distribution networks (CDN) can be established according to the enterprise needs. Hence, issues of design and management of the infrastructure and enforcement of network policy are specific to the enterprise network, and can differ significantly from the Internet case. Secondly, since data in such enterprise networks is commercially valuable, care needs to be taken to ensure that the multimedia streams do not disrupt the data flow. This requires the design of appropriate flow control mechanisms and fairness criteria that ensure multimedia streams are “friendly” to other streams. Other requirements on the multimedia streaming include the need for adaptation to account for varying bandwidths and user requirements, and the need for secure streaming that is resistant to attack. Importantly, these requirements are driven by the multimedia applications deployed over the enterprise network, many of which require interactivity, and have very stringent constraints on the delay and achieved multimedia quality. We partition all these different requirements and issues into three broad categories that we label: Infrastructure, Streaming and Applications. We summarize these issues and their details in Table 1.

<table>
<thead>
<tr>
<th>Broad Area</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure</td>
<td>Design: including server placement and content distribution</td>
</tr>
<tr>
<td>Management: policy based management, enforcement of policy</td>
<td></td>
</tr>
<tr>
<td>Delivery: including content discovery and request-routing</td>
<td></td>
</tr>
<tr>
<td>Streaming</td>
<td>Multimedia adaptation to meet network variations, user preferences etc.</td>
</tr>
<tr>
<td>QoS reservation, Flow control, fairness, friendliness to other streams</td>
<td></td>
</tr>
<tr>
<td>Secure streaming</td>
<td></td>
</tr>
<tr>
<td>Applications</td>
<td>E-commerce, In-service training, Video-conferencing, Immersive environments, Defense applications, and broadcast</td>
</tr>
</tbody>
</table>

While some of these issues have been studied in the past, the interplay between them, in the context of enterprise streaming, has not been sufficiently investigated in current research. In this paper we provide an overview of the relevant research on these streaming issues and attempt to highlight their limitations in the context of enterprise multimedia streaming. Due to space limitations, and the vast amount of literature available on the above issues, instead of providing a comprehensive review, we describe some of the key research areas and provide references to relevant papers in each area. This paper is organized as follows. We present background research on the Infrastructure in Section 2, describe research on Streaming and adaptation issues in Section 3 and finally describe some Applications in Section 4. We conclude with a list of papers, submitted to the special session on enterprise multimedia streaming at ICME 2005.

2. BACKGROUND: INFRASTRUCTURE
In this section we provide an overview of research on establishing infrastructure for enterprise networks, the design, management, and policing of the network. We include research on CDNs, overlay networks, issues of server placement, content distribution, request routing, policies for management and enforcement mechanisms, and server scheduling. In our description we use the terms server, surrogate and proxy interchangeably.
2.1. Content Distribution Networks

CDNs facilitate and optimize the management and delivery of static or streaming content over IP networks that include enterprise corporate intranets, extranets, and also the Internet. By ensuring that content is distributed and delivered closer to end users, CDNs overcome network bottlenecks and performance issues while providing greater accounting and management control over how the networks are used. For enterprises seeking to optimize their own networks to handle streaming media, and for service providers wanting to offer content delivery as a managed service, CDNs offer key advantages [1]:

- Built on scalable components, CDNs can grow to accommodate bandwidth-intensive applications, such as eLearning and corporate Webcasts.
- By pre-positioning content close to the users likely to request it, CDNs significantly reduce bandwidth usage.
- By providing fast and reliable access to content, CDNs improve employee productivity and acceptance of new applications.
- Optional content filtering enables the enterprise to block objectionable content.
- Through caching, load-balancing, routing optimization etc., CDNs increase content availability and protect against server overload.
- By empowering the network to manage, share, and distribute content more effectively, CDNs are designed to reduce server hosting costs at enterprise data centers.

The CDN infrastructure may be wired or wireless, and in this overview we do not distinguish between the two because they are similar in terms of our defined broad categories, although they differ in the details due to the underlying characteristics of wireless channels and associated issues such as mobility etc. An overview of wireless CDN research is provided in [2]. The CDN infrastructure deployment can also be achieved using some of the techniques developed for overlay networks. Overlay networks use application layer networking in conjunction with the deployment of powerful nodes/proxies in the network to achieve some of these goals of enterprise streaming. Most of the current efforts on overlay networks are aimed at improving the QoS provided by the network infrastructure, enabling multicast routing and peer-to-peer file sharing. Existing studies such as RON [3], Yoid [4], and OPUS [5], primarily focus on performing application-layer unicast and multicast routing, based on network measurements to improve path performance and failure resiliency. Recent research in distributed content delivery has also included end-system approaches [6], [7]. However, further research needs to be performed in extending these ideas to provide the wide-range of services required by enterprise multimedia streaming applications.

2.2. Server placement and content distribution

This problem involves determining how many servers to use, where to position them, and how to distribute content across them [8]. It has been shown that this is very similar to the Facility Location Problem [9], which has been described as an NP-hard problem. The work in [10] tries to solve the problem by minimizing the number of autonomous systems traversed when clients request objects from the server. The authors compare four replicating strategies, and conclude that a single CDN server with knowledge of the entire system is capable of making the best decision as to the placement of servers. The work in [11] solves this problem for a tree topology network, which is a rather unrealistic topology for many current networks. The work in [12] describes the placement and replication strategy used in a decentralized peer-to-peer (p2p) style system. In this work, requests for content travel between nodes and content returns to the requesting node via the search path, with each node along that path caching a copy of the content. If another request arrives at one of these nodes within a short period of time, then the node should be able to serve that new request. Over time, as fewer requests reach the original source node, and many reach the new source node, the content on the original node is removed to make way for newer more popular data. In these systems content moves dynamically from one location on the network to another location which is closer to the requesting clients. Other work on peer-to-peer media communications is presented in [13].

There are different approaches to content distribution including replication, and the use of multiple descriptions of the content. The work in [14] presents ongoing research on replica placement in adaptive-CDNs (A-CDNs). Algorithms to optimize the placement of replicas in the surrogates of an A-CDN are designed based on a static model for cost-quality-optimized adaptive content networking. Other work on placing mirrors or replicas across servers is included in [15] and [16]. In multiple description (MD) CDNs [17], each unit of content (e.g. a movie) is divided into multiple complementary descriptions that are spread across a number of servers. As an example, for MD with two descriptions, each server may host 0, 1, or both descriptions. An important special case is when each server hosts only one description, which leads to the notion of coloring where each server needs to be assigned a particular color corresponding to a unique description. The goal of this coloring problem is to color the servers so that a complete set of descriptions (e.g. both descriptions for MD with two descriptions) are close to every client.

2.3. Content discovery and request-routing

Content discovery is the process of identifying servers/proxies that possess the desired content, or relevant portions of it. Proposed discovery mechanisms in p2p networks are based on flooding and forward routing algorithms. In flooding mechanisms, as used in Gnutella

---

**Facility Location Problem**

A problem in which the goal is to determine the optimal location for facilities to serve a set of clients, given that each facility has a fixed capacity and the cost of serving a client increases with the distance between the client and the facility.

**Content discovery and request-routing**

The process of identifying servers/proxies that possess the desired content, or relevant portions of it.
[18], peers flood a network with queries to discover the desired content. While such a model supports robustness and extensive reach, it does not scale well as the size of the network (number of peers) grows. Forward routing mechanisms are used in systems such as Chord [13], where keys are assigned to peers using consistent hashing, enabling other peers to locate a key efficiently. Recent work on efficient and distributed content discovery has also been presented in [19]. The research in this field has mainly focused on downloading of data (media or otherwise) and needs to be extended to allow for streaming applications. Request-routing involves receiving client requests, and determining the server or set of servers which can be used to service this request. The IETF Content Distribution Internetworking Group (CDN-WG) has begun the process of specifying the requirements and standardizing the protocols for CDNs [20]. The work in [21] presents an overview of the different request routing techniques that have been developed by researchers. The request-routing system uses dynamic information about network conditions and load on the surrogate servers to balance the load while directing client requests. This system may also interact with the accounting system and the distribution system to dynamically redistribute the content. Different request-routing solutions include DNS-based, transport-layer based, content-layer based, application-layer based and NAT-based routing. However, a majority of these schemes have not been designed for streaming media, and extending these schemes for enterprise streaming of multimedia is an interesting area of research.

2.4. Policies for CDN management
Network policies are a set of rules to administer, manage, and control access to network resources [22]. The policy framework defined by the IETF/DMTF consists of a policy management tool, policy repository, policy enforcement point (PEP), and policy decision point (PDP). Work on developing network policies for CDN management is presented in [23]. However, a majority of these policies are not specific to the requirements of multimedia streaming, and further research is necessary to design optimal multimedia streaming CDN management policies.

2.5. Server scheduling
Work in this area focuses on designing optimized scheduling and batching algorithms at the servers for video streaming in CDNs. In [24] the authors study batching and patching with pre-caching at the proxy for multiple heterogeneous videos, however they restrict all assets in the system to be managed using a single scheme, irrespective of the difference in their access patterns. Other related work also exists in [25] and [26]. In [27], the authors propose a unified mathematical framework under which various server scheduling and proxy cache management algorithms for video streaming can be analyzed. They incorporate known server scheduling algorithms (batching, patching, and batch patching), and proxy management algorithms (full caching, partial caching, and no caching, with an option to cache patch bytes or not) in this framework and analyze the minimum backbone bandwidth consumption under the optimal joint scheduling/caching strategy.

3. BACKGROUND: STREAMING
In this section, we highlight the issues of multimedia streaming, flow control and adaptation. We present an overview of research on multimedia adaptation, and research on flow control for fairness (in order to prevent multimedia flows from disrupting enterprise data flows), and finally some research on secure multimedia streaming.

3.1. Multimedia Adaptation
Multimedia adaptation mechanisms are used to tailor the bitstreams to specific bandwidth variations and user preferences based on content characteristics [28], [29]. For live and interactive applications, a real-time encoder adapts the rate or resiliency of the bitstream based on end-to-end network feedback [30]-[32]. For off-line multimedia streaming, adaptation is performed by switching between multiple pre-encoded non-scalable streams or by transcoding through requantization, DCT coefficient dropping, frame dropping, and spatial resolution reduction [33]-[37]. Scalable video compression [38]-[41] has also been proposed to address the dynamic bandwidth variations within IP networks.

Numerous solutions have been proposed for efficient multimedia streaming over the Internet [42]-[47]. Many of these multimedia adaptation schemes are applicable to the enterprise streaming scenario, however they need to be modified to account for the network management, distributed content, and QoS support specific to the enterprise network.

3.2. Issues of Flow Control and Fairness
As mentioned earlier we would like the multimedia streaming to not interrupt data flows in enterprise networks. Typical multimedia streaming solutions use protocols such as UDP [48] for their transport. However such protocols are inherently unfair to data traffic that mainly uses TCP for transport. Hence, a large amount of work has been performed on designing TCP-friendly streaming solutions [49] and a survey of such approaches can be found in [50]. However, a majority of these schemes make assumptions on the TCP performance model, and are not designed to be friendly to TCP streams at all time scales. Additionally enterprises often require all streams to go across firewalls, and streaming multimedia over TCP is necessary in many of these cases. Finally, the reliable transmission provided by TCP can assist some of the error control and protection schemes designed for multimedia streaming. Hence, we think that TCP streaming is going to be of interest to many enterprise networks. The work in [51] develops models to capture TCP streaming performance and shows that TCP generally provides good streaming performance when the achievable TCP throughput is roughly twice the media
bitrate, with only a few seconds of startup delay. A number of existing research efforts that use TCP for streaming [52]-[55] combine client-side buffering and rate adaptation to deal with the variability in the available TCP throughput. Research on multimedia adaptation usually assumes undifferentiated, best-effort service from the network, and relies on signaling mechanisms such as packet loss rates for feedback. Enterprise networks can be designed with enhancements for QoS support. Hence, usage and QoS dependent pricing can be used as a natural incentive to drive adaptive behavior by applications to enable a fair use of the network resources. In [56], the authors present a framework for the adaptation of multimedia application sending rate and/or choice of network services in response to dynamic network pricing and changes in application requirements. The adaptation is based on the (monetary) value of a service as perceived by the user, relative to the price charged by the network. A survey of prior work on dynamic pricing for QoS enabled networks is included in [57].

3.3. Secure streaming

Securing a digital video stream [58] involves the following: conditional access, authentication, copy control, and content tracking. Conditional access enforces that the stream is viewed under the rules specified by the content owner (usually after the viewer has paid a fee). Authentication is the verification of the integrity of the video stream, the identity of the source, and the identity of the receiver. Copy control is a mechanism for controlling the number of copies of a video stream a user is allowed to create, and content tracking is the embedding of a signature or serial number into the video stream for subsequent identification. The implementation of these security measures often involves cryptographic algorithms, digital signatures (or certificates), and watermarking. Hence there are several open research issues regarding the security of digital multimedia content especially if it is streamed over networks to different clients, and some of these issues are studied in [59]-[60].

4. BACKGROUND: APPLICATIONS

In recent times digital media has been acknowledged as a standard data type and it is expected that it will become the medium for business-business, business-employee and business-customer communications [61]. Enterprises are, thus, beginning to explore the use of digital media in improving the effectiveness and efficiency of their business processes. There are several applications that are currently deployed, and are increasingly likely to be deployed in the future, over enterprise networks in which multimedia streaming is an essential component. This is because, besides driving the development of new applications and technologies, multimedia streaming has several tangible advantages for effective organizational management. These include:

- **Improved communications**: when deployed for corporate communications and inter-departmental collaboration. Streaming can help companies get their messages out to larger audiences, or narrowly defined groups, more quickly and cost effectively.
- **Centralized Information Archive**: when applied to training, education, product launches or investor announcements, multimedia streaming facilitates the creation of a central archive of information that can be accessed at a later date.
- **Reduction in Travel Costs**: Real-time Access to Information: streaming makes information available in real-time. This is critical for information rich market segments such as financial institutions.

We hence believe that multimedia streaming will be used by corporations especially for corporate communications, in-service training (e-learning), product launches, performance reporting and delivery of real-time financial information for E-commerce.

Besides the applications for organizational management, multimedia streaming over enterprise networks is also being deployed in the development of Business TV, and immersive environments for gaming, etc. The defense network infrastructure can also significantly benefit from streaming multimedia especially for surveillance, monitoring etc. Future research into enterprise multimedia streaming will both drive and be driven by the successful deployment of some of these applications, and the identification of other novel applications. Finally, we want to mention that real solutions for multimedia streaming and distribution need to be combined with provisioning and management of mixed-media information and high-level semantic analysis to satisfy business scenario requirements.

5. CONCLUSIONS

In this paper we describe the enterprise multimedia streaming problem and highlight its requirements and specific issues. We group these issues into three broad areas that we label: Infrastructure, Streaming and Applications. We present an overview of current research in several diverse areas that is relevant to solutions for enterprise streaming, and highlight directions for further development. Through this paper we hope to generate interest in designing novel solutions and applications for enterprise streaming. Finally, we list papers that will be presented in the special session on Enterprise Multimedia Streaming at ICME 2005, below:

- Introductory paper: “Enterprise Multimedia Streaming: Issues, Background and New Developments”.
- Pascal Frossard and J. Chakareski, “Rate-Distortion Optimized Bandwidth Adaptation for Distributed Media Delivery,” EPFL.
REFERENCES


