Statement of Research

Surendar Chandra

My research is motivated by a vision of ubiquitous access to multimedia content with a light footprint on resource consumption. Media objects are large. Contemporary users access these objects from a wide variety of devices with different resource constraints. My research is attempting to bridge a critical gap between the enormous resource demands of media objects and the need for ubiquitous access. Towards this end, I have addressed challenges in dynamically adapting objects to the prevailing resources using quality aware transcoding. I reduced the energy foot print for streaming multimedia using popular formats in a variety of network conditions. I devised a next generation collaboration file system that is suited for contemporary wireless users. I developed peer-to-peer mechanisms that consider the unique characteristics of the shared objects. I devised mechanisms to address the long term storage needs for a video sensor network. I also developed a fully decentralized security infrastructure that provides data availability, access control, and versioning with tunable levels of certainty. Recently, I am developing a flexible multiuser and multidisplay collaborative screencasting system that is targeted towards corporate workgroups.

My research philosophy centers on choosing practical problems with experimentally demonstrable limitations and then developing techniques to address these challenges. My research approach uses experimental systems techniques; I draw inspiration from methodologies rooted in multimedia, mobile computing, storage, wireless networks and computer security. Frequently, my research has bucked conventional wisdom by demonstrating the limitations of state-of-the-art deployed systems. For example, my group empirically analyzed the wireless user behavior in a number of different locales to show the extent by which prior group collaboration systems will experience poor performance among contemporary wireless users.

1 Dynamic adaptation using Quality Aware Transcoding

Users access the Internet from a wide variety of devices with different resource constraints. Transcoding objects had been a popular mechanism for content adaptation. Earlier systems blindly performed transcoding operations by choosing an arbitrary fidelity level leading to an increase in file size as well as unacceptable fidelity loss. I quantified the fidelity tradeoff characteristics and developed the ability to predict resource requirements for transcoding JPEG images. I developed applications of this quantification knowledge to dynamically allocate available resources on a per-client and per-request basis. I demonstrated how a web service can utilize informed transcoding to gracefully degrade content fidelity to different classes of users and in response to changing client demands. I also showed the value of this technology in a multimedia capture device to manage the local battery and storage resources. My work had generated seven publications with 416 citations in Google scholar.

2 Reducing energy requirements for multimedia streaming

Wireless streaming media are popular and place enormous demands on the battery resources of mobile devices. I analyzed the wireless network interface energy consumption for receiving popular streaming formats through deployed infrastructure networks under varying network conditions. I showed that the IEEE 802.11 power saving mode of operation did not provide energy savings. Analyzing the stream dynamics of these popular streaming formats, I pioneered history-based client and server side strategies to reduce the energy consumed for streaming by transitioning the WNICs to a lower power consuming *sleep* state. I applied these techniques to ad hoc networking scenarios. I also showed the limitations of energy conservation techniques for placeshifting systems. This work lead to nine papers and 222 citations.

3 Group collaboration system among contemporary wireless users

Collaboration mechanisms have had a tremendous impact on human communications. Synchronous collaboration mechanisms such as chat services and asynchronous mechanisms such as email are popular.

Newer systems such as Facebook and Google+ are reaching even more of humanity. The expected performance of these systems depend on the availability behavior of the group members. Contemporary users are wireless and operate from home, office and from a *third place*. Hence, with Xuwen Yu, I performed an extensive analysis of the availability behavior of wireless users from a number of locales. I performed a long term analysis of wireless users at Notre Dame using application level monitoring mechanisms. I analyzed wireless users at a conference and at a research lab. I also analyzed the availability trends over three years for users on a city-wide hotspot federation. I quantified the variability of user availability behavior amongst these locales. I also discovered significant user churn in all these scenarios. These factors have a significant impact on the performance of prior collaboration systems which had originally analyzed their own performance using simulations and limited deployments.

We showed that prior group collaboration systems that either used centralized or distributed approaches will achieve poor performance; maintaining a single shared copy of the conversation is untenable for these weakly connected workgroups. Instead, we developed a moderated collaboration mechanism that maintains multiple copies of the shared object. These replicas are propagated at a rate that is only limited by the wireless user availability patterns. All these copies are presented seamlessly via a file system interface. The file system interface also allows the user to use conventional tools to operate on the shared documents. The various versions of the shared document eventually converge through independent moderation operations. The system automatically logs the provenance of causal reads in order to quantify whether updates from a particular user had been incorporated into the final version. Our distributed approach eases deployment concerns by not relying on the university to provide the storage infrastructure. Except for small groups, our distributed approach achieves performance similar to a server based system. Our prototype *flockfs* exhibits acceptable file system performance and update propagation latency.

4 Content aware peer-to-peer (P2P) systems

P2P systems have been the preferred mechanism to federate storage resources. My group analyzed the contents and queries from a popular P2P system to highlight the fundamental flaws in the assumptions used by prior systems. Since 2003, we collected traffic on the popular Gnutella network. Based on the observed system behavior, we developed an unstructured P2P overlay that used the peer sharing capacity and network characteristics to generate an overlay using local information. Our compact and well connected graph provided better search mechanisms using attenuated Bloom filters and random walkers.

Also, an analysis of the queries and annotations of objects that are stored showed that the query terms and object annotations exhibit a Zipf like distribution. In practice, as compared to the success rates of 62% achievable using assumptions of uniform distribution, Zipf distributed objects were successfully found only 5% of the times. Also, the relative popularity in the object annotations did not correlate well with its popularity in the query workload. Almost half the queries had no matching objects in the system regardless of the overlay or search mechanism used to locate the objects. We developed a P2P middleware that transparently transformed the user queries in order to improve the query success rates. Our experimental analysis showed a success rate of over 60% for rewritten queries in Gnutella networks.

5 Autonomic management of scalable long term storage

We addressed the problem of continual resource management in a system that stored contents across a large number of devices for long durations. We are concerned with resources such as total storage space, object reliability, availability and security. For example, to manage the storage space, systems require the ability to continually increase the total storage space and match the storage requirements of new objects. Traditional storage systems provide persistence for all stored objects to perpetuity, relegating object reclamation to applications. However, not all of the data are equally important; less important objects could be automatically replaced by more important ones. We developed a simple and expressive temporal *impor-*

tance function. We have used this abstraction to manage object persistence. This information allows the storage system to selectively weaken the performance guarantees offered by the storage. Using extensive simulations and observations of a university wide lecture video capture and storage application, we showed that our abstraction allows the users to control the amount of persistence for each individual object. We were investigating mechanisms that can weaken the performance of other resources as well.

6 Security primitives for Ubiquitous storage

Sensor networking and P2P applications operate without any trusted infrastructure and can tolerate weaker security guarantees. With Ashish Gehani, we developed a fully decentralized security infrastructure that provided data availability, access control, and versioning with tunable levels of certainty. The stringency of traditional security guarantees for these operations are minimally relaxed. In exchange the protocols are resilient and can continue to operate correctly despite a fraction of the storage nodes acting maliciously.

7 Screencasting for multiuser and multidisplay collaboration

Screencast is the process of capturing a sequence of images of the display screen. It had been widely used for real time sharing of displays as well as for lecture distribution. Using empirical data from a variety of usage scenarios, I quantified ways in which screencasts behave differently than traditional streaming videos. The updates are created at variable rates and are only limited by the processing capacity of the mobile devices. I devised mechanisms to transform temporal redundancy into spatial redundancy and achieved higher compression ratios for lossless mechanisms. I am building a flexible screencast mechanism for collaboration among a wide variety of devices. I am also adopting this technology for lecture capture. The system will be publicly available.

8 Multimedia technologies to enhance education

There has been considerable evidence on the importance of videos for lecture reviews; both in Computer Science as well as in other fields. Prior efforts had used skilled videographers for the video capture. Several projects had also automated some of this capture work-flow. However, these efforts were expensive; faculty members must still depend on the university to allocate its scarce resources for capturing the lectures.

For seven semesters, I applied my expertise in multimedia and storage technologies to capture and annotate all of my lectures. I distributed videos of my lecture using web based mechanisms, podcasts, iTunes U, YouTube and Google video. I convincingly showed that technology improvements can allow any instructor to capture and produce the videos with minimal effort. I showed evidence that technology improvements are allowing students to consume high definition videos; from about 13.2% to 29.6% of the entire requests. Over the years, my approach consumed over 60 TB of the university's network resources. I investigated the strengths of several viable alternatives. I showed that the real challenge in distributing videos was in choosing the distribution mechanism which balances the desires of the instructor to freely distribute the video and the strain that their choice can place on the campus network. I presented my positive educational experiences as well as the pitfalls of video capture in a colloquium at the Kaneb Center for teaching and learning at Notre Dame. By participating in a university wide iTunes U panel, I shared my experiences with other faculty and the administration as well. The lecture video access traces also serve as the workload for my storage research. I intend to build on these experiences to build a viable large scale capture, archival and distribution mechanism.