DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF SOUTHERN DENMARK, ODENSE

COMPUTER SCIENCE COLLOQUIUM

Characterizing YouTube Workload at a Campus Network

Anirban Mahanti
Department of Computer Science and Engineering
Indian Institute of Technology (IIT), Delhi, India

Thursday, 23 October, 2008 at 10:15
Auditorium U49E

Abstract:

This talk discusses the traffic characteristics of the popular Web 2.0 video sharing service, YouTube. Over a three month period, we observed over 23 million transactions between users on a large university network and YouTube, including more than 600,000 video downloads. At the same time, we collected statistics on the globally popular videos on YouTube.

In this talk, we examine usage patterns, file properties, file popularity and referencing characteristics, and transfer behavior of YouTube. We also study several YouTube user session characteristics such as session durations and inter-transaction times. Where appropriate, we will discuss the similarities/differences between YouTube workload and traditional Web and media streaming workloads. Implications of our findings on network management, content distribution strategies, and design of next-generation synthetic Web workloads are considered. For example, we show that as with traditional Web, caching could improve the end user experience, reduce network bandwidth consumption, and reduce the load on YouTube's infrastructure. Unlike traditional Web caching, however, Web 2.0 provides additional meta-data that should be exploited to improve the effectiveness of strategies like caching. We conclude the talk with a brief discussion of current and future work.

While most of this talk will focus on YouTube workload characterization, a brief overview of my research interests and recent, significant, research contributions will also be presented.