

# The Information Discovery Graph: An Internet Multimedia Session Directory

Nathan Sturtevant, Nelson Tang, Lixia Zhang  
{nathanst,tang,lixia}@cs.ucla.edu  
<http://irl.cs.ucla.edu/IDG>

## Background: Multimedia Session Directories

Multimedia session directories store information about available multimedia sessions. Users can connect to these directories and learn of available sessions in which they are interested. On the MBone, SDR is the name of the session directory that is used. SDR uses the Session Announcement Protocol (SAP) to make multimedia sessions available.

SAP is a very simple protocol. It uses a single global multicast channel to announce multimedia sessions, reserving a constant amount of bandwidth for the announcements. Users merely listen to this well-known channel and keep listening until they find interesting sessions.

## The Problem: Poor Scalability

The design of SAP has a number of problems when scaling to the presumed large number of users and multimedia sources. By keeping the amount of announcement bandwidth constant, as the number of sources increases, they announce themselves with less frequency. This increases the start-up delay for users to find interesting sessions, especially since the announcements are not cached anywhere. Finally, the announcements are made in an arbitrary order; there is no categorization of the sessions by their content.

## Our Solution: The Information Discovery Graph

To address these problems, we propose the Information Discovery Graph (IDG), a scalable multimedia session directory. The IDG was designed to address the challenge of scaling to a large number of users and a large number of multimedia sessions. It is made up of a hierarchy of information **managers**, each maintaining resource information for specific topics or areas. Multimedia **data sources** register themselves with the IDG system and keep the information up-to-date. The managers automatically arrange themselves into a hierarchy according to the semantic content of their data sources. **Users** can then browse or search the hierarchy for content related to their interests.

The hierarchical design of the IDG enables fast searching that scales as the number of data sources grows. By caching parts of the hierarchy, user startup delay is further reduced. Additionally, the IDG is self-organizing, which dynamically keeps the taxonomy topically relevant. Since IDG managers can encapsulate existing SDR traffic, the IDG can be incrementally deployed.

## Current Status and Continuing Research: Refinement and Implementation

We have developed a simulation and an early prototype of the IDG system. Preliminary simulation results have confirmed our theoretical models of reducing the time for a user to locate a session, through temporal caching. We are continuing to work on developing spatial caching, as well as other models, to also reduce the amount of global bandwidth the IDG uses. We are also developing better self-organizing behavior of the managers to ensure a responsive and efficient taxonomy. As we refine our designs, we are implementing a complete prototype, which we hope to deploy and test in more real-world environment.