

Developing Video Streaming Service Across Campuses: The VIVA Experience

Introduction/Abstract

The Virtual Library of Virginia (VIVA) is a publicly supported consortium of libraries at 71 public and private, non-profit higher education institutions in Virginia USA. Members include research universities, liberal arts colleges and 2-year colleges. [VIVA](#) has been in existence for over a dozen years. VIVA is guided by three fundamental principles: equitable access to information, cooperation across institutions and cost effective use of public resources. In 2006 VIVA licensed the rights in perpetuity for close to 400,000 students and faculty at Virginia colleges and universities to use more than 500 hours of video content from the US Public Broadcasting Service (PBS). The license with PBS requires that the video be streamed and that users authenticate with a user id and password. PBS does not host the video; the video must be hosted by or within the consortium.

This paper will provide an overview of the issues associated with providing access to large online video collections using VIVA as a case study. The VIVA project began with the idea of developing a support network and best practice standards that would enable member schools to host online video licensed by VIVA on their own campuses. A task force is working on system specifications, technical standards and process models for encoding and serving streaming multimedia content. One of VIVA's members, the University of Virginia has agreed to host a streaming service and act as a Shibboleth Service Provider for other VIVA members. Shibboleth uses open source software to authenticate and authorize access to services for institutions that have joined a federation and developed policy agreements regarding access to services. It offers the promise of enabling sharing among schools that have diverse approaches to identity management and varying levels and types of technical infrastructure and expertise. Now VIVA is also pursuing a Shibboleth streaming video federation whose success will depend upon collaboration between the library and IT communities at large and small institutions. The paper will conclude with a discussion of the challenges associated with developing a Shibboleth implementation plan and a digital video support strategy for the wide variety of institutions that belong to VIVA.

Educational Video Online

Online distribution of commercial entertainment media has been the subject of recent attention in the press. There has also been attention to projects which use non-commercial streaming video in education (for example, [Research Channel](#)) and on conversion of analog video to digital formats for archival and educational use ([Ubois 2005](#)). However there is very little in the literature or the press about online distribution of commercial educational video content in higher education. Large scale licensing and online distribution of commercial educational video by libraries or library consortia is a relatively new phenomenon. This is in part due to the rapidly evolving technology

needed to support streaming media and in part due to the unsettled nature of markets and licensing models. The first and largest online video project involving a library consortium in the US was the [Digital Media Center](#) at OhioLINK a consortium of academic libraries in the state of Ohio. The OhioLINK project (Bauer 2004) takes advantage of a common authentication scheme and central distribution service to provide the higher education community in Ohio with access to over 1100 full length videos, most distributed by Films for the Humanities and Sciences (now part of FMG, the Films Media Group). Another large scale educational video streaming project, involving a single campus has been done by VIVA member, James Madison University. ([Clark, Maxfield & Saunders 2005](#)) and involves over a thousand titles from FMG, WGBH, Annenberg and smaller distributors.

At a strategic planning retreat in 2005, the VIVA Steering Committee considered issues associated with licensing of online educational video. For most of its 12 year history VIVA has invested the bulk of a multi-million dollar annual budget on e-collections hosted by vendors. The 2004-06 biennium budget from all sources is \$13,372,674 USD. Of that budget 89.9% is allocated for collections. In its earliest days technology investments represented a significantly larger portion of the overall budget. In its first biennium some 80% of VIVA budget was invested in technology to level the playing field for institutions that did not have infrastructure to support e-resources offered by VIVA. (Alberico 2002) A decade ago many schools had to clear technological hurdles before they could offer online database access to students and faculty; a similar situation now exists for multimedia content. One of the issues facing VIVA as it enters the shifting landscape of digital media is to understand market trends. Scenarios were developed for possible outcomes within VIVA.

Scenario 1: VIVA licenses, vendor hosts.

After a few years of startup experience with VIVA-hosted content, most video titles licensed by VIVA on behalf of non-profit higher education in Virginia will be hosted by vendors. VIVA continues to support titles for which it has purchased rights by serving content from computers at member schools. This scenario reflects the current situation for e-journals and other mostly text content licensed or purchased by VIVA since it was founded.

Scenario 2: VIVA purchases and centrally hosts its own content.

VIVA develops a centralized multimedia distribution service and metadata; metadata are available to members who wish to integrate with web sites and integrated library systems. Whenever possible, VIVA opts to purchase rather than license multimedia content. VIVA establishes download rights as an essential term in its multimedia licenses.

Scenario 3: VIVA outsources hosting to its members.

VIVA members provide multimedia distribution and file management services for the portion of multimedia titles not hosted by vendors. VIVA develops distributed approaches to providing access to multimedia. VIVA members insure that no member is without access. A federation allows schools to divide responsibility by encoding format, vendor or content type. Members share responsibility for cataloging and user support.

Scenario 4: Each school hosts on its own; VIVA provides support.

VIVA develops multimedia specifications, provides training and a support network to all members. VIVA promulgates information about best practices including detailed specifications on hardware, software and authentication mechanisms needed to support streaming. Each school is responsible for hosting, cataloging and managing its own VIVA audio, video and image collections.

After discussing scenarios and considering available video the VIVA steering committee established a strategic goal of providing access to media collections and decided to pursue licensing a large video collection from the US Public Broadcasting Service (PBS). The [PBS collection](#) was appealing because it included quality video covering a broad range of subject areas of interest in the consortium. Furthermore PBS was willing to supply digital video source files that would enable VIVA to develop a testbed for resolving collection management, technical, access, user support and content distribution issues.

The PBS collection consists of 476 titles comprising approximately 560 hours of video in MPEG-4 format compressed at a bit rate of ~2 Mbps. The collection was delivered on DVD. Many titles in the collection are multi-episode series. License terms prohibit VIVA from offering the PBS videos for downloading but allow any VIVA institution or group of institutions to stream video content to the VIVA user community. The license requires individuals to provide id and password before initiating a stream. The license prohibits mixing, mashing or developing derivative works from PBS media files. The storage estimate for the files comprising the collection is ~2Tb. This assumes a MPEG-4 master file for each video episode, and two files encoded for streaming, one at 300Kbps and one at 800Kbps.

Online Video Licensing Issues

Business models and delivery systems for online digital video in the higher education marketplace are about ten years behind models and delivery systems for e-journals and online reference databases over the past decade. The number of educational video publishers and distributors who are willing to enter that market is still relatively small. Licensing of media content occurs within a different context from that of reference databases and electronic journals that have been the mainstay of VIVA for most of its

history. The current licensing environment for online media is dominated by licensing approaches originating from the entertainment industry rather than in scholarly publishing. Difficulties that rights issues, access technologies and market forces present to learning institutions have been well documented ([McGeveran & Fisher 2006](#)) In the higher education market a limited number of vendors offer online video, even fewer host streaming services and fewer still offer a user download option in standard licensing practice. A notable exception among major distributors is the Films Media Group whose [FMG On Demand](#) service offers options for vendor hosted streaming, for local streaming, for downloading and for outright purchase of digital video content for online delivery.

It remains to be determined whether the online educational video market will move to pay per view approach, a vendor-hosted subscription model or a rights purchase arrangement with the licensee assuming responsibility for hosting the streams. An issue in licensing documentary video is subsidiary rights. Documentaries often include music, film segments and other source content where rights are owned by third parties or where ownership of rights is undetermined. Rights issues have been a major impediment to both licensing and online distribution. Consortia and individual institutions who want to provide their communities with access to online media must be prepared to address complex licensing challenges associated with online distribution channels, rights management and pricing models. In the end the decision about whether to license an online video collection will be determined by a complex cost benefit equation involving compromise between vendors and educational institutions.

In the case of the VIVA PBS license, the terms that prohibit offering video episodes to users for download were counter balanced by the ability to make the collection available to all VIVA members in perpetuity at reasonable cost. PBS covers subjects unlikely to become dated allowing time for less technologically advanced schools to catch up with regard to network capacity and support infrastructure.

Multimedia Task Force

Once the decision to license a major video collection for online distribution was made, a task force whose members represented the types of institutions in VIVA was charged with addressing the challenges of providing equitable, cost effective access to rich media content. The following charge was developed for the task force:

- Serve as a forum for exchanging information about technical issues related to multimedia collections
- Develop expertise among VIVA members about the technologies, non-textual information content, evolving markets and distribution channels for multimedia collections.
- Develop recommendations on technical specifications and technical support for multimedia content licensed by VIVA

- Develop recommendations on cataloging, organizing and preserving digital, multimedia collections licensed by VIVA
- Develop recommendations on distribution models and access control for VIVA multimedia collections
- Identify human resources, technology investments and organizational structures within VIVA institutions that will be needed to support streaming media
- Develop recommendations for user support including minimum technical requirements, training and documentation
- Consider service integration issues; develop recommendations on how to integrate multimedia content from VIVA with existing library services and with campus learning environments

One of the first tasks undertaken was to conduct a survey to determine interest and experience level of members in working with online video. The survey determined that more than a third of respondents were hosting video on their campuses and multiple approaches were being used. It also indicated a high level of interest in a centrally hosted service. Finally it became clear that the cable plants on almost all campuses were adequate to support local streaming; at the same time many campuses had inadequate pipes to the edge to allow them to take advantage of a centrally hosted service. The task force also developed an online forum which enabled communication among members and provided space for sharing software, technical information and encoding samples.

Technical Challenges

Decisions on selection of a video player, video streaming server software and encoding formats are closely related. There is no single solution that will meet the needs and preferences of all users. Compromise is a given.

Media Player and Server Selection

One challenge for any streaming media project is to identify minimum desktop specifications, as well as browser, and media player recommendations. The challenge is confounded by quick turnaround of new player releases. The player is where learners interact with media content; the player selected will determine the encoding scheme needed to deliver video streams that can be read by the player. Four free player technologies were considered for the VIVA PBS project: *Windows Media Player*, *QuickTime*, *RealPlayer* and *Flash*. It should be noted that Flash video runs within the browser; a Flash player was developed to demonstrate that technology for the project.

Questions to be answered by any institution considering large scale streaming include: How much flexibility to provide in terms of specifying browsers and media players students and faculty need to use? Is it preferable to deliver media content through a player or through a browser? Should the number of players or browsers supported be

limited? Are there specific players or browsers to avoid? What would it take to support multiple browsers and players?

It was also important to identify a streaming server solution that could be supported on VIVA member campuses. The recommendation would need to include hardware, operating system, streaming software and storage specifications. A goal for server recommendations was to get as close to a turnkey option as possible. Another goal was configuration recommendations to optimize video streaming. Servers associated with each of the four players were also considered.

Encoding Issues

Deciding on a media player and an encoding format are closely intertwined. In keeping with VIVA's emphasis on equitable, cost effective access open formats and open source, non-proprietary software were preferred. Early on a decision was made that VIVA could only support one encoding format. Individual schools could opt for a locally preferred format if they wanted to. Files would be encoded for streaming with multiple bit streams in order to support users on both high and low bandwidth networks. Test encodings were prepared for *Windows Media*, *Real Media*, *QuickTime* and *Flash*. Files representing each scheme at a variety of bit rates, frames per second (fps) and resolutions were made accessible from a single web site. The same source was used to enable comparison. Data on file size, encoding time, and each encoding parameter were recorded and associated with individual files.

Factors considered when evaluating encoding formats included quality of video streams under different network conditions, player characteristics and expertise needed to manage the type of server needed for each format. Other factors included expense and time required for encoding tools and the server software needed for each format. Further complicating these factors are the multiple "flavors" available within a particular encoding format and the often complex relationships between encoded media files, media server software and media players. Some of the players could play files in formats other than their native format and there were nuances associated with files associated with particular players and servers. For example, with the QuickTime player our preference was to use the more open .mp4 file type rather than the relatively more proprietary .mov file type but the performance we saw from the .mov files was superior and the tools needed to encode in that format were less expensive and more powerful. Similarly, the files encoded for Real Media were encapsulated allowing multiple bit streams to be wrapped in a single larger file. Another challenge was developing an approach to distributing encoded files between schools. After considering options including FTP and peer-to-peer transfers a decision was made to copy files to high capacity external hard drives and ship them from one institution to another.

The decision making process for choosing a player, server and encoding scheme was simplified when the network administrator at the University of Virginia offered to support

a 3-year pilot by hosting a streaming service with authentication and authorization handled by open source Shibboleth software. The only video server that would work in conjunction with Shibboleth was Darwin Streaming Server. Darwin Streaming Server is an open source Shibboleth-enabled application that works with Apple QuickTime and MPEG-4 formats on Windows or UNIX platforms. The pilot agreement settled the decision to go with the QuickTime player and some flavor of MPEG-4. Schools that want to take advantage of a central service could use a Shibboleth approach and those wanting to host video locally could use either Darwin or QuickTime server.

Now VIVA is pursuing a dual strategy of providing members with recommendations on hosting streaming video content and developing a centrally hosted solution. Each path on this parallel effort is based on a common set of encoding standards and those encoding standards assume the MPEG-4 scheme. Each strategy is also based on an understanding of the circumstances of VIVA members as reflected in surveys. While Shibboleth simplified decision-making on choice of server, player and encoding format it introduced new concerns with regard to efficient use of network bandwidth.

Shibboleth

Shibboleth ([Morgan et al 2004](#)) requires cooperating pieces of software at the Service Provider (SP) or *Target* end of a connection and at the Identity Provider (IdP) or *Origin* end. Additional software manages trust relationships between Service Providers and Identity Providers in a federation and routes service requests based on a requestor's response to the question "Where are you from?" (WAYF)

Once Shibboleth identifies the home institution of a person requesting a service the Shibboleth Identity Provider (IdP) software at their institution interacts with the local identity management and directory systems on their campus; requestors login using the approach used on their own campuses and Shibboleth Identity Providers are then able to pass assertions about them to Shibboleth Service Providers. Service Providers provide access to services based on user attributes and policies governing which users or types of users have access to which resources and services. Background on Shibboleth may be found on the Internet2 Shibboleth web site at: <http://shibboleth.internet2.edu/>

Federations codify policies and trust relationships that govern which services are available to whom at which member institutions. Federation management includes managing Shibboleth WAYF (Where Are You From) service used to route service requests to appropriate Shibboleth Identity Providers. In the Shibboleth world the metadata representing policies and trust relationships between institutions is managed at the national level. One agency responsible for federation management in the US is [InCommon](#). Roles performed by InCommon include serving as a repository for metadata about trust relationships, functioning as a certificate authority and managing WAYF services for US higher education and research institutions that use Shibboleth to control access to online services and resources. For the VIVA project, InCommon is

seen as the best way to take advantage of federation capabilities while maintaining maximum flexibility to support future sharing agreements between campuses. Schools that want to join InCommon apply for membership and sign a participation agreement, paying a one-time \$700 initiation fee and \$1,000 per calendar year.

VIVA institutions wanting to participate in a federation need to install Shibboleth Origin software configured to interact with their local attribute authority and identity management solution. The survey done by the VIVA task force identified LDAP as the most widely used identity management solution in the state. To support a Shibboleth authenticated streaming video service VIVA and the University of Virginia (UVA) will collaborate on recommending standards, preparing documentation, building a support network and offering Install Fests to help members get up and running with the Origin software. The VIVA project will employ software that was developed at New York University for the [Database of Recorded American Music \(DRAM\)](#) to enable Shibboleth access to audio from Darwin Streaming Server.

Federation Requirements

Without a critical mass of institutions a Shibboleth pilot won't provide equitable access or justify the work involved. 87% of survey respondents indicated willingness to explore Shibboleth and 76% were willing to implement service within a 12-18 month timeframe. Only 36% of respondents were hosting a production video service. The majority of VIVA schools are not sure yet what to do with streaming video. Schools not involved in hosting collections are not sure what that would require. To help VIVA schools make the decision about whether to participate in a federation, VIVA is identifying resources and technical capability necessary for success. The assessment will be done for Shibboleth and local hosting. It is not an either/or situation; 36% of respondents are already involved in local hosting and the richness of the PBS content makes the Shibboleth option appealing to 87%.

Bandwidth Issues

One issue to explore is the implications of Internet pipe size on a VIVA federation. Size of pipes to some campuses might be an issue, especially for video streamed from external domains. The biggest schools are on Internet2 and National Lambda Rail (NLR) with pipes in the 1-10Gbps range. Pipe size available at other VIVA schools varies considerably. The 6 VIVA schools on Internet2 and National Lambda Rail (NLR) network are unlikely to experience bandwidth problems in a Shibboleth environment. However the host at UVA has limited ability to pipe information to schools using the commodity Internet. Those schools have access to less bandwidth but the amount of bandwidth on incoming pipes might not be the only problem; contention for access on the portion of the host's outgoing network available for commodity Internet traffic is limited and needs to be monitored. In the unlikely event that streaming service to

participants on the commodity Internet exceeded capacity at UVA, traffic would need management, for example, by limiting concurrent users.

Almost all survey respondents indicated a 10-100Mbps capacity on campus which should be adequate for campus use of the video, assuming that the buildings with that capacity are those where videos would be viewed. It should be noted that if even if that problem is resolved on the server end, schools with very low bandwidth connectivity might still experience problems receiving streams in the range of 300-800Kbps along with all of their other inbound and outbound traffic. The single biggest unanswered question related to a hosted service is what the usage patterns will look like and whether demand is likely to exceed the ability of networks at member schools to deliver video to their users. While that does not seem likely it cannot be known without testing.

First Steps

This project requires collaboration between institutions to develop policies, technical cooperation involving IT departments on different campuses and willingness on the part of libraries to provide support and integration with campus information services. It is a level of intra and inter institutional cooperation that is consistent with VIVA's mission but it will be a long range challenge.

A small group of institutions will soon begin testing the service. A decision was made to involve schools that are less likely to require technical support in the early phases of the project while also including a representative slice of VIVA's membership. IT organizations at participating schools must be willing to install Shibboleth Origin software on their campuses and join the InCommon federation. A Shibboleth Origin requires a dedicated UNIX box running Apache and Tomcat and a commitment of someone's time to install and maintain the software. VIVA and the University of Virginia will co-sponsor an Install Fest to help with that task. Schools must also have the ability and the willingness to enter into an authentication agreement with federation partners and to tie their Shibboleth origin to their campus authentication service. Seven schools including a mixture of research universities, comprehensive universities and one private liberal arts college have committed to phase one of the project.

Encoding standards suitable for high and low bandwidth distribution at bit rates of 300 and 800 kbps and viewable with the free QuickTime player have been finalized:

<p>800k Stream 512x384 Frame Size Up to 30 Frames per Second Video - H.264 @ 700kbit Audio - AAC @ 96kbit Stereo</p>	<p>300k Stream 400x300 Frame Size 30 Frames per Second Video - H.264 @ 256kbit Audio - AAC @ 40kbit Mono</p>
---	---

The software that enables the Shibboleth Service Provider software to interact with the Darwin Streaming Server has been implemented and testing has begun.

Encoding is a slow process; it can take several hours to encode a single video. Encoding with available hardware and software could take a significant amount of time. To reduce turnaround time for encoding the entire collection VIVA approved funding to outsource encoding to a service bureau and issued an RFP. A contract for the MPEG-4 transcoding work has been awarded. In addition, source video files were distributed on high capacity hard drives to two VIVA schools for encoding in Real and Windows formats that can be shared with other schools which prefer those formats for local hosting.

Cataloging Issues

It is anticipated that VIVA will pursue a parallel strategy for supporting local and centralized hosting when it comes to cataloging, file management and supporting resource discovery. Recommendations about file structure, data structure and descriptive meta-data for media content are being developed. Strategies for developing, importing, exporting and sharing metadata and cataloging information in order to avoid duplication of effort in VIVA libraries are being pursued.

There are established methods for cataloging audiovisual content (Weitz 2001). VIVA will supplement accepted cataloging principles with approaches that lend themselves to resource discovery for streaming video content. As is the case for many digital collections it is likely that schools will want to integrate catalog data in their OPAC with discovery tools, descriptive meta-data and access options tailored to characteristics of a streaming media on the web (Calhoun 2006). Other issues center on handling individual episodes within multi-episode series and links to web sites maintained by PBS. Early work involves developing model catalog records for the same PBS content used for encoding tests. The record structure supports series and episode level description for multi-episode series as well as links to associated web sites. File naming conventions have also been developed. VIVA is also exploring approaches to embedding descriptive meta-data in the files as the files are encoded. A goal is to embed only the simplest and most basic meta-data. Schemas being considered include unqualified Dublin Core, MPEG-7 and i-Tunes. The embedding work will be done by the vendor who will be doing the encoding work.

Quality Control and User Support

Video quality is a subjective experience involving variables from encoding scheme to available bandwidth to the user's computer. Quality perception can be affected by playback capabilities, for example, features of player VCR controls. Bandwidth constraints dictate tradeoffs. Blocking and jitter are common problems; adjusting parameters to fix problems in one area can exacerbate problems in the other. Content characteristics and anticipated use patterns influence encoding decisions. VIVA's goal

is to find a happy medium resulting in viewing experience suitable for PBS content and acceptable to a broad cross section of users.

Quality assurance for streamed video is a priority and requires testing of encoding quality and from the point of view of the user viewing a video on a network. Unfortunately quality testing for streaming video cannot be easily automated though there are protocols available. For the purpose of quality testing from the user's perspective VIVA is considering use of a construct known in the industry as a Hypothetical Reference Circuit (HRC). ([Winkler & Campos 2003](#)) An HRC includes all of the parameters required for a successful viewing experience from the encoded file, to the server, to the network to the viewer's player software and workstation. VIVA will define HRCs that include two standard encoding schemes. To test for quality under the Shibboleth solution, files on Darwin streaming server will reflect different source videos under conditions typical of schools across the state. One goal is to stress test the circuit in a controlled situation to pinpoint problems. This will help determine, for example, whether a 300k stream will work for VIVA schools with relatively low bandwidth pipes. Once fault points have been identified through testing, the consortium can adopt design and support strategies that address known issues

Conclusion

No blueprint exists for consortium licensing and hosting of digital video. Any project involves a complex mix of licensing, technical, policy and organizational issues. Challenges are great; the promise of improving learning by making rich educational media available to large audiences is even greater. A VIVA federation offers future potential for shared management of and access to multimedia content. The PBS collection is a potential high payoff, low risk starting point for a federation because access to rich educational content can be extended to a broad audience who need only authenticate themselves at their own institutions. Once a federation is established the basic elements will be in place for other shared services as well as for access control to services and resources supplied to VIVA by Shibboleth-compliant vendors. Shibboleth offers intriguing possibilities for integrating online learning environments across institutions, allowing students and faculty from different places to participate in learning communities in the same way they now share access to quality online information provided by VIVA. To the extent that it lays in place the infrastructure for inter-institutional sharing of online services, this means that it is goal worth pursuing.

References

- Alberico, R., Winter 2002, 'Academic Library Consortia in Transition', in Strategic Alliances and Partnerships: New Opportunities for Higher Education, New Directions in Higher Education Series. Jossey-Bass, vol. 2002 no.120, pp. 63-72.
- Bauer, C.& Carlin, J, 2003, 'The Case for Collaboration: The OhioLINK Digital Media Center', Journal of Library Administration, vol.29, no. 2-3, pp.81-82.
- Calhoun, K., March 17, 2006, 'The Changing Nature of the Catalog and its Integration with Other Discovery Tools', Library of Congress, Washington. Available at <http://www.loc.gov/catdir/calhoun-report-final.pdf>
- Clark, J., Maxfield, S. & Saunders, B., October 2005, 'Online Video Collections: Transition from Media Checkout to Media Streaming', EDUCAUSE Annual Conference, Orlando. Available at <http://www.educause.edu/ir/library/pdf/EDU05226.pdf>
- McGeveran, W. & Fisher, WW., August 2006, 'The Digital Learning Challenge: Obstacles to Educational Uses of Copyrighted Material in the Digital Age', Berkman Center Research Publication No. 2006-09. Available at <http://ssrn.com/abstract=923465>
- Morgan, R. L., Cantor, S., Carmody, S., Hoehn, W. & Klingenstein, K. 2004, 'Federated Security: The Shibboleth Approach'. EDUCAUSE QUARTERLY, vol. 27; pt 4, pp. 12-17. Available at <http://www.educause.edu/LibraryDetailPage/666?ID=EQM0442>
- Ubois, J. Marc, 2005, 'New approaches to television archiving' First Monday, vol.10, no. 3, Available at http://firstmonday.org/issues/issue10_3/ubois/index.html
- Weitz, J., 2001, 'Videorecording cataloging: problems and pointers' Cataloging and Classification Quarterly, vol. 31 , no.2, pp.53-83.
- Winkler, S. & Campos, R., 2003, 'Video quality evaluation for Internet streaming applications', in Proc. SPIE/IS&T Human Vision and Electronic Imaging, 5007, pp. 104--115, (Santa Clara, CA), Jan. 21--24, 2003.