

# Position Paper: The Convergence of Video on IP and HTTP

Bruce Davie      Richard Maunder  
Cisco Systems, Inc.  
{bsd, rmaunder}@cisco.com

January 6, 2011

“Convergence” has been both an industry buzzword and something of a holy grail for many years. We define convergence to be the ability to carry widely diverse services, which at one time required specialized, service-specific infrastructure, on a single network. Voice services are now well along such a converged path thanks to the success of voice over IP, and video is heading the same way.

Convergence of many services onto a single network provides two broad classes of benefits. The first class relates to cost reduction. Network operators often focus on the immediate cost savings of consolidating a set of diverse offerings onto a common infrastructure, with attendant reduction in capital and operational costs. However, we believe a more significant long term benefit comes from the potential for greater innovation and cross-fertilization between services. For example, voice over IP applications benefit from the ability to treat voice as another form of data that can be managed via an email program or Web interface. This cross-fertilization enables new services that were not previously possible, or even envisaged, in siloed systems.

A closer examination of video suggests that the convergence of video onto IP networks, while advanced in some respects, is far from complete. While it is certainly true that a great deal of video runs over IP, today there are different architectural approaches for enterprise and service provider networks; different technologies used for “linear” (live) and “nonlinear” (e.g., video on demand) content; a diverse, constantly shifting, set of video codecs, container and DRM options; and differences in the delivery of content in “over-the-top” environments (e.g., Hulu, YouTube, iPlayer) versus managed environments, such as the IPTV services offered by broadband service providers. We believe that the historically high bandwidth and processing requirements for digital video have caused every video environment to be treated as a “special case”, and that this poses a potential impediment to realizing the full benefits of convergence.

Recent technological developments could lead to a new level of unification in the video space. One such development is the trend towards the use of HTTP and “adaptive” streaming techniques. Initially targetted towards over-the-top, nonlinear content, adaptive streaming is quickly showing promise for live content and managed networks as well. However, this trend raises some important questions and challenges. Of particular importance to the W3C community is that adaptive streaming over HTTP to

date has been developed in a proprietary manner, with comparable but incompatible approaches from a variety of vendors. While there are benefits to allowing for innovation in the development of adaptive streaming techniques, the lack of a standard adaptive streaming approach may also hinder adoption (some devices will be unable to play some content) or add cost to systems (end devices need to understand multiple protocols, content delivery systems need to store multiple container formats). Just as with the debate over default codecs for HTML5 video, there may be room for debate as to how much of adaptive streaming needs to be standardized or whether some sort of default should be recommended. We are interested in discussing the options in this space. It will also be necessary to determine whether the W3C is the appropriate place for any standardization, as a number of other standards bodies have shown signs of interest in the problem area.