Proceedings, International Symposium on Multimedia Communications and Video Coding (MCVC-95), Brooklyn, NY, October 1995 (invited paper).

ARCHITECTING VIDEO-ON-DEMAND SYSTEMS: DAVIC 1.0 AND BEYOND

Alexandros Eleftheriadis

Department of Electrical Engineering
Columbia University
New York, NY 10027, USA
eleft@ctr.columbia.edu

INTRODUCTION

The culmination and convergence of several different technologies, including video compression and networking, are today making possible the design (and soon the deployment) of commercial interactive video services. Several commercial trials have been announced, are in progress, or are already completed. The multi-disciplinary nature of video-on-demand services, which includes content providers, service providers, network providers, cable TV distribution companies, as well as the computing and consumer electronics industries, has made it extremely difficult to concentrate within a single organization the expertise needed to design and implement such services. Attesting to that is the significant activity that has been created within the industry in terms of alliances and partnerships.

As with any prospective service of such anticipated scale, it is imperative to ensure that open and interoperable systems are used. This not only alleviates the problem of having to design vertical solutions, but also helps to stir healthy competition (technical and economic) within well-defined sub-system domains (e.g., among content providers or network providers). Most importantly, it allows for—and actually encourages—a fast-paced evolutionary path by easing the process with which innovations and new research results can be incorporated: having well-defined interfaces enables the substitution of existing subsystems with more sophisticated or better performing ones, without having to redesign the whole system.

We discuss the design of Video-on-Demand (VoD) systems, focusing on the efforts of the Digital Audio-Visual Council (DAVIC[1]). DAVIC, established in June 1994, is an association of more than 200 industrial and academic organizations from around the world, pursuing the definition of interfaces and protocols for interoperable VoD systems. DAVIC's scope is extremely broad, addressing the complete spectrum of vertical (bottom-up) and horizontal (end-to-end) specifications. It primarily performs a "systems integration" function, interconnecting several different technologies to create a complete and coherent system. As such, it attempts to distill all current experience about the design of VoD systems (actively monitoring and/or collaborating with related activities, such as ISO, ITU, the ATM Forum, IMA, and the IETF), as well as create solutions for problems that still remain unresolved. As a result, several recent developments in standardization activities (both official and industry-based) have been incorporated. This includes the ISO/IEC MPEG-2 and MHEG-5 specifications, ATM Forum and ITU-T specifications on ATM, the recently designed IP Version 6 (most often referred to as IPng, for "next generation"), CORBA 2.0, etc.

DAVIC expects to release its first set of specifications (DAVIC 1.0) in December 1995. A common and very important theme in DAVIC is to define a *single* solution for each functionality desired, thus eliminating the need to support multiple (and typically conflicting) specifications.

VoD is just one from a number of possible applications that can be supported from a video-capable communication system. In fact, by proper provisioning, a large number of different applications can be accommodated without overburdening the implementation. Furthermore, an open approach allows the introduction of new services that may not have been anticipated by the system's designers.[2, 3]

The structure of the paper is as follows. We first give an overview of a System Reference Model, that defines the basic components of a VoD system. We then describe in detail each of these components, namely the server, the delivery system (core and access networks), and the subscriber terminal, as well the necessary protocols operating between them. We conclude by discussing media encodings and future directions of the DAVIC effort. The description is necessarily brief, and it should be noted that it is based on the current draft status of the DAVIC specifications; the reader is encouraged to consult the actual text of the specifications for more detailed and up-to-date information.

SYSTEM REFERENCE MODEL

The notion of a VoD system is quite broad, and has been used to denote quite different application domains. For our purposes, VoD indicates an interactive digital audio-visual service with a prevalent video component. Two fundamental design problems are the delivery of high-quality digital video to the end-user, and the handling of interactivity. What makes them particularly challenging is the desire for scalability, i.e., to be able to efficiently support a large number of users, as well as the interoperability constraint. The latter mandates the definition of several different and independent subsystems, the specification of their interfaces, and the verification of their interoperation. The interoperability constraint can be applied at all levels, from inter-regional systems (across regulatory domains), to individual protocol components.

The DAVIC system architecture is based on four basic entities: the end-user, the delivery system provider (or network provider), the service provider, and the content provider. The end-user interacts with the system using a subscriber terminal, or Set-Top Unit (STU). STU connectivity is provided by the delivery system provider using a possibly tiered approach: a core network that forms the backbone of the delivery system, and an access network that connects the STU to the core network (typically covering the last few miles). This tiered architecture is necessitated by the design of the physical networks that are already available to end-users (cable TV coax and telephone twisted pair), although an end-to-end ATM solution has distinct advantages.[4] Systems that can be directly connected to the core network (e.g., computers using high-speed connections) can have a much simpler architecture, incorporating the access network functionality (service selection etc.) at the end-system. Finally, the service providers offer access to stored or live content, while the content providers generate the programming material provided by the service.

Within each system, and across all of them, we can distinguish a number of different information flows and distinct interfaces. Figure 1 indicates the overall architecture, in the form of DAVIC's System Reference Model (SRM). The figure depicts a number of different information flows (S1–S4) and associated interfaces or reference points (denoted by A). The latter designate the entire protocol stack at the particular interconnection point they refer to. In the following sections each of the components is described in more detail. Our interest is the interfaces and protocols between components and not the component implementation details.

In order to properly define the operational aspects of the system, it is important to identify a common set of functionalities that are needed to implement a desired set of applications. The initial set of applications targeted for support in DAVIC 1.0 includes movies-on-demand, teleshopping, broadcast, near video-on-demand, delayed broadcast, games, telework, and karaoke-on-demand; the common functionalities needed to support these functions have been taken into account in the overall design of the system.

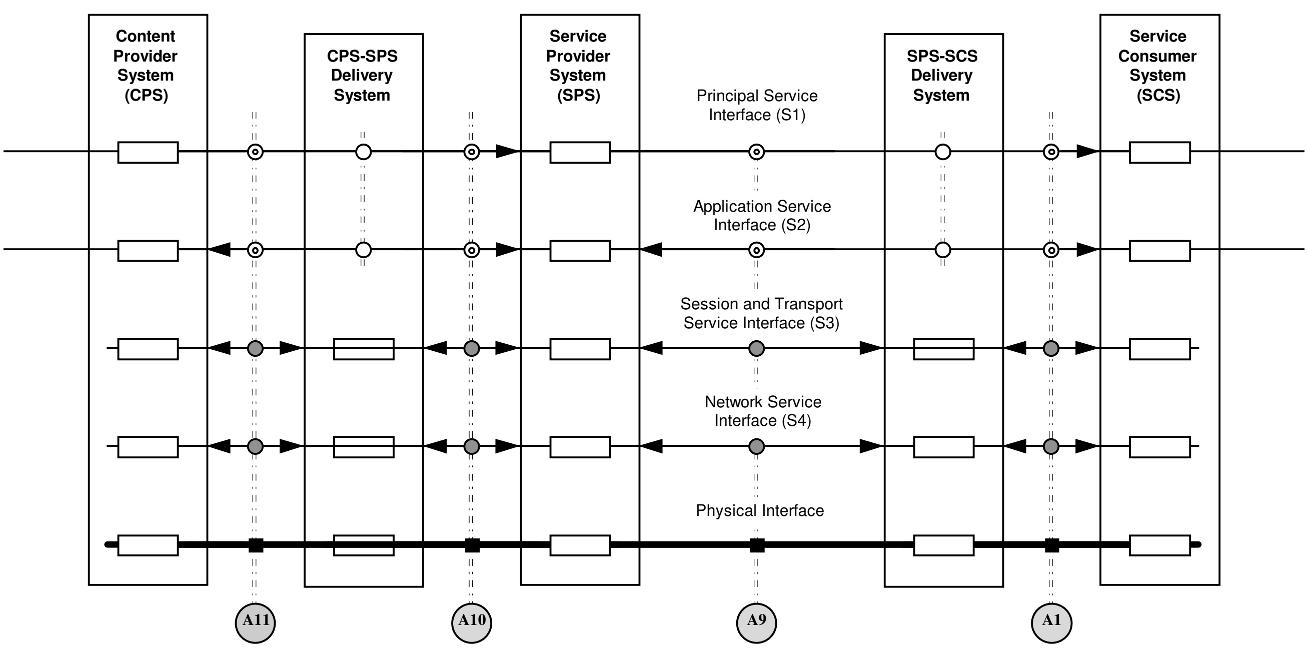

Figure 1. DAVIC System Reference Model. Also shown are interfaces (A1, A9, etc.), and information flows (S1–S4) across system components.

SERVICE PROVIDER SYSTEM

The server has four basic responsibilities: service gateway functionality, stream services, application services, and content services (Figure 2). Gateway functionality is a "brokerage" service that provides the means for registering and decommissioning services, allows the client (STU) to discover the existence of a service (e.g., an application for accessing movies-on-demand), and establishes and manages sessions. The stream services provides the repository and source for streams, i.e., content bitstreams. This service is initiated from the service gateway, when a content selection has been made by the user (e.g., to playback a movie). The application service is the key vehicle with which application functionality is provided, and an important source of added-value by individual service providers. Finally, the content services are responsible for content management, including loading and unloading between the content provider and service provider, as well as between service providers. These core services can be used to define additional ones, such as client profiling, file , and download services.

The application service interface (S2 flow) is based on an object-oriented architecture, centered around OMG UNO[5] (Universal Networked Objects) 2.0 and the default Internet Inter-ORB Protocol (IIOP). The session service interface (S3) is based on DSM-CC U-N (ISO/IEC 13818-6, Digital Storage Media Command and Control, User to Network part). Hence OSI Layers 4 and 3 for these interfaces utilize TCP and UDP over IP. The principal service interface is based on the MPEG-2 TS (Transport Stream) delivery. More details about protocol specifics are given later on.

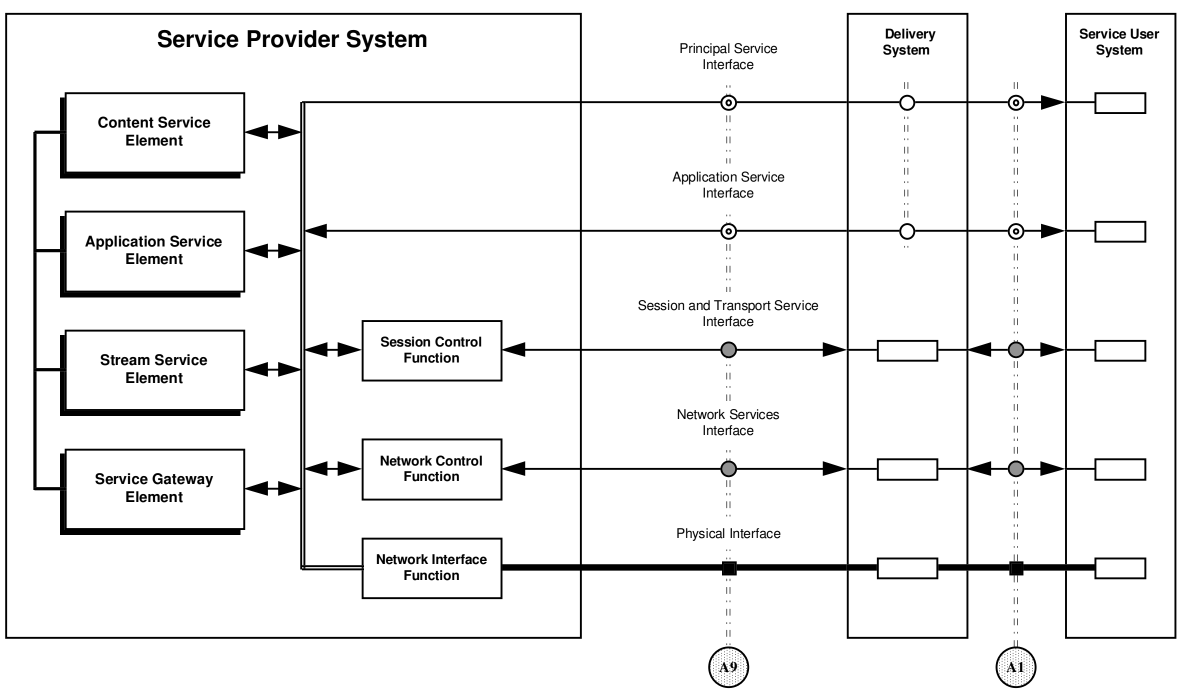

Figure 2. Service provider system architecture. Internal interfaces (those that are not visible through the A9 point) are left unspecified, as they are implementation-dependent and do not affect interoperability.

All of the server internals are purposely left unspecified, since they do not affect interoperability (the only needed specification is for the interfaces exported across the A9 interface as shown in Figure 2). In addition, due to the challenges involved in server design, they are highly platform-specific.

DELIVERY SYSTEM

Delivery systems can include both physical media-based (e.g., CD-ROM) as well as networked systems. In DAVIC 1.0 only networked systems are considered, for both wired (e.g., CATV) and wireless (terrestrial and satellite broadcast) networks. The delivery system architecture follows a potentially tiered approach that includes a core network that forms the backbone of the system and an access network that bridges client STUs with the core network. The overall architecture is shown in Figure 3.

Core network switching and multiplexing is based on ATM. The Access Node (Figure 3) has the responsibility of adapting between the ATM-based core network and the specifics of the access network used. DAVIC provides support for a variety of access network types, including Asymmetric and Very high speed Digital Subscriber Lines (ADSL, VDSL), Fiber-to-the-Curb (FTTC) and Fiber-to-the-Home (FTTH), as well as ATM-based and non ATM-based Hybrid Fiber Coax (HFC).

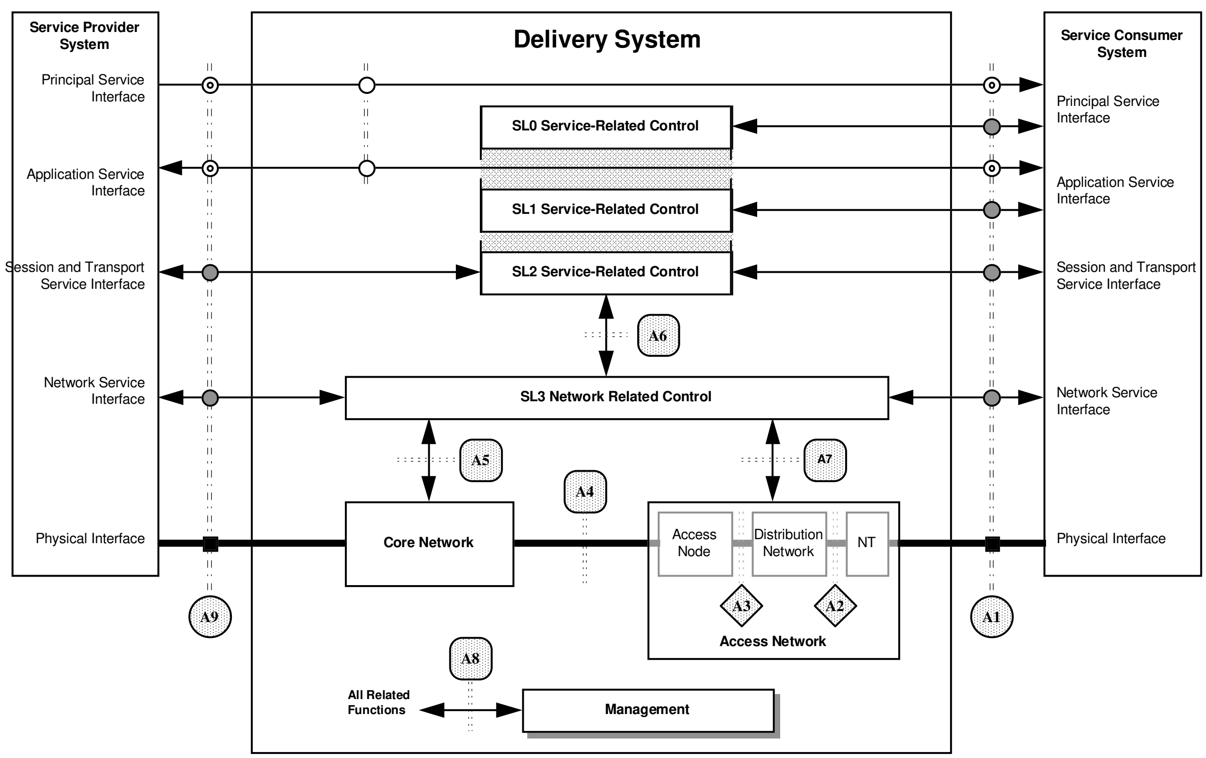

Figure 3. Delivery system architecture. Both core and access networks are included. The access network involves an access node for interconnection with the core network, the distribution system, and network termination (NT). NT may be either active, or passive if the actual termination is performed at the STU. Note that the A2 and A3 interfaces will not specified by DAVIC.

Satellite systems can function both as core networks and access networks; in the former case they have to comply with the A4 interface; while in the latter with the A1 interface. Obviously, in order to provide user control, satellite delivery systems have to be augmented with appropriate reverse channels; the mechanism for this is not yet determined by DAVIC.

Regarding to the in-house network, i.e., the part between the NT and the STU(s), two options exist. In an active NT configuration, a bus topology can connect several STUs to the same connection; in a passive configuration (where the STU provides termination functionality) the in-house network becomes more complex and is not specified.

Network-related control (S4 information flow, at the Network Service Interface) is on ITU-T Q.2931, as these are public networks. The access network beyond the A4 interface can be based either on ATM or on MPEG-2 TS delivery. A mapping function is then necessary in order to assign virtual paths/channels to MPEG-2 TSs. With respect to network management functions, two candidates under consideration are CMIP defined by ITU-T M.3010 and SNMP defined by RFC1157. The former is likely to be used for the core and access networks, while the latter is preferred by end service providers.

SERVICE USER SYSTEM

The STU is arguably one of the most interesting components of a VoD system, as it represents a significant percentage of the opportunities for added value for service and content providers. It is also the only component (hardware-wise) that belongs to the domain

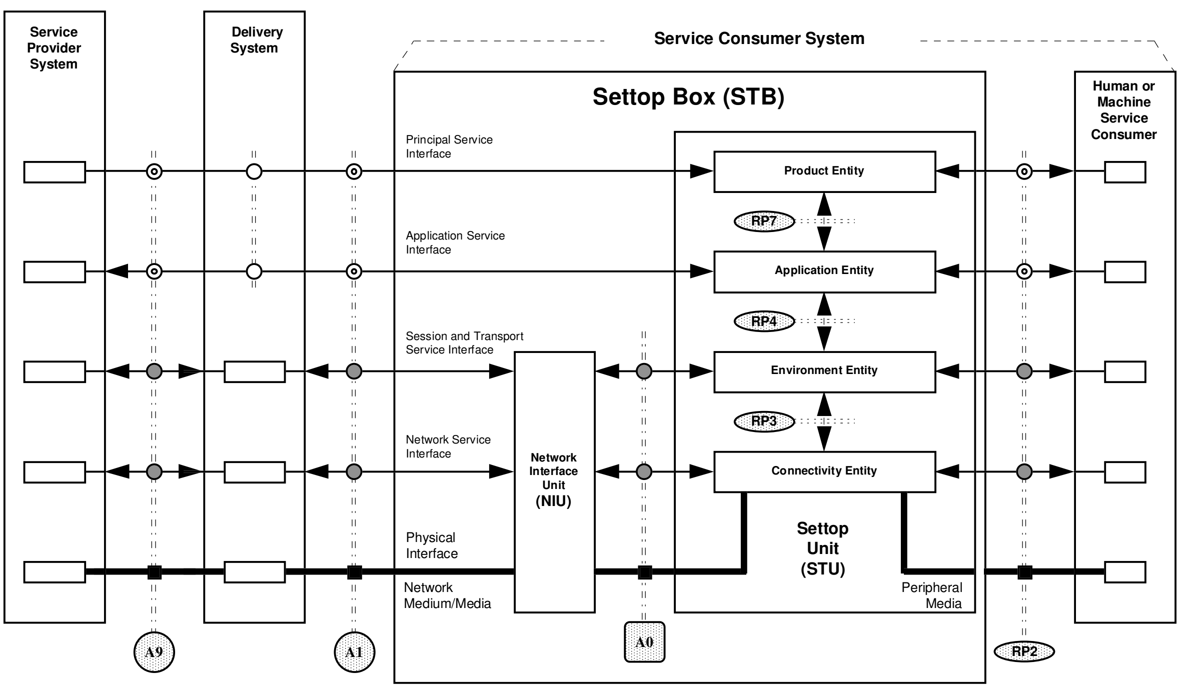

Figure 4. Set-top Unit (STU) Architecture. Note that the actual Set-Top device also includes the Network Interface Unit (NIU), to form the Set-Top Box (STB). To facilitate application interoperability, the architecture defines internal STU interfaces, at Reference Points (RPs) 3 and 4.

of consumer electronics: users will be able to select among several different models and obtain the one of their choice.

The STU accepts a signal compliant to the A1 interface (high-speed downstream channel, Figure 4) through the physical interface. The Network Interface Unit (NIU) then processes the signal based on the type of A1 interface used, for example QAM demodulation for HFC or satellite networks, and then passed through the A0 interface (if one exists). The high-speed data is processed by an MPEG-2 demux, with any relevant control messages (possibly out-of-band) passed on to the appropriate entities, as shown in Figure 4. Conditional access functionality is also provided at the STU/NIU, but its specification is not included in DAVIC 1.0.

In order to facilitate interoperability at the application level, a standard application interface has been adopted. Between the RP3 and RP4 reference points we have the STU Operating System (OS), on which a Run-Time Engine (RTE) resides (these form the Environment Entity). The RTE exports a well-defined interface across RP4, to be used by application developers. The application itself is based on MHEG-5[6] (residing on top of RP4), which was designed to support the distribution of interactive applications across platforms of different types and brands. MHEG-5 is actually a subset of MHEG-1, in which platforms with minimal resources are taken as primary implementation targets. This portable application format is stored at the server, for downloading to the STU. In order to allow for application-specific functionality, extensions to the OS (and thus the RTE) as well as the application code are allowed. This provides platform-independent content representation, without sacrificing flexibility or the capability for creative differentiation. Note that MHEG-5 does not provide a scripting language, but it allows using one to perform more complicated tasks from the generic ones the standard supports. MHEG-5 object communication between the server and the STU is performed using the DSM-CC U-U (User-to-User) specification. Although the mapping of object and content references is defined in DAVIC 1.0, that of API actions is not (although work is under way to ensure that sufficient functionality is provided to support high-level API run-time facilities).

PROTOCOL CONFIGURATIONS

The preceding discussion described the individual architecture of each of the components of a VoD system, as conceived in the DAVIC model. The glue that interconnects these components are the individual protocols that operate between them. Due to the large number of interfaces and variations thereof (due to the multiplicity of access networks), we cannot exhaustively list them all here. We thus provide only indicative examples only for the server.

The basic stack structure is centered around ATM/AAL5, MPEG-2 TS, and TCP, UDP, and IP. The ATM Forum's 5/8 mapping is currently used for mapping MPEG-2 TS packets to AAL5 PDUs, but the subject is far from being a closed issue.[7] ATM is used within the core network and, depending on the access network, may extend end-to-end. MPEG-2 TS is used for the downstream channel (from the server to the STU), carrying audio-visual information as well as other data in the private data part of the bitstream. IP is used for application command and data information transfer (e.g., DSM-CC U-U), as well as network management (SNMP). Figure 5 indicates, for example, the protocol stack to be used for the S1 information flow (downstream traffic) across the A9 interface (server to core network); note that various physical layers are acceptable. Figure 6 shows the corresponding stack for the S2 information flow, consisting of application command and data information. Similarly, Figures 7 and 8 depict the protocol stack for the S3 (session control) and S4 (network control) information flows.

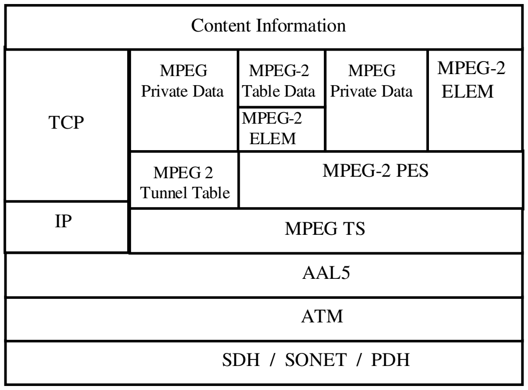

Figure 5. Protocol stack for S1 information flow (high-speed downstream) at the A9 interface.

DSM-CC U-U
OMG IDL
OMG UNO
TCP
IP
AAL5
ATM
SDH/SONET/PDH

Figure 6. Protocol stack for S2 information flow (application command and data transfer) at the A9 interface.

DSM-CC U-N
TCP/UDP
IP
AAL5
ATM
SDH/SONET/PDH

Figure 7. Protocol stack for S3 information flow (session control) at the A9 interface.

Q.2931
SAAL
ATM
SDH/SONET/PDH

Figure 8. Protocol stack for S4 information flow (network control) at the A9 interface.

Protocol stacks for other corresponding interfaces are similar, with modifications primarily to the lower layers. For example, for an ADSL-based access network the only difference in Figure 5 would be the substitution of the physical layer.

CONTENT REPRESENTATION

Content representation for multimedia information is based, as mentioned before, on the MHEG-5 specification. For monomedia encoding, DAVIC uses the following formats: MPEG-2 for video (MP@ML) and still picture encoding, MPEG-1 Layers I and II for single and dual channel audio (32, 44.1, and 48 kHz), DAVIC-specific formats for linear (PCM) audio and graphics bitmaps, and Unicode (ISO/IEC 10646) for character encoding (with the Latin-1 subset being mandatory). With respect to graphics, the specified format is error-resilient and extensible, and supports: CLUT1, 2, 4, 8 and RGB16 formats; square, 4:3, and 16:9 pixel aspect ratios; transparency; 50% translucency; and full or half-screen resolutions.

BEYOND DAVIC 1.0

The preceding discussion only described some of the key features of the architecture of a DAVIC system. Several components were omitted, including security, local session and initialization protocols between the STU and the access network, usage information protocols (for billing etc.), as well as profile and level delineation. DAVIC is already looking forward, and several issues have been identified for coverage in specifications after version 1.0. These include definition of the A10 interface (between content provider and server), server MIB and management, satellite return channel specification, A0 and A1 connectors and STU data ports, scrambling, key management, authentication, and copyright enforcement, 3-D graphics, etc. More sophisticated functionalities are included in DAVIC's long-term workplan, including distributed servers, symmetric channels, multi-point connections, mobile access, home networks, and service scalability.

Undoubtedly, the all-encompassing scope and aggressive agenda of DAVIC are unique within the standardization community. Furthermore, the selection of standards-based solutions is a sound approach in building a consensus-based system. The strong and continuously increasing support it enjoys from the entire spectrum of the computing and communications industries are very encouraging signs for its success. A critical step in this respect will be the rate of adoption of the DAVIC specifications in actual products in the months to come. Another key issue will be whether or not it will be able to attract and absorb recent developments that have contributed to the rapid growth of the Internet, and provide a unified delivery platform. The different design philosophies and cultures may prove critical in this respect.

REFERENCES

1. DAVIC 1.0 Specifications, Revision 4.0 (draft), Hollywood, CA, September 1995.
2. Y. H. Chang, D. Coggins, D. Pitt, D. Skellern, M. Thapar, C. Venkatraman, An open-systems approach to video on demand, *IEEE Comm. Mag.*, Vol. 32, Nr. 5, pp. 68–80, May 1994.
3. S.-F. Chang, A. Eleftheriadis, and D. Anastassiou, Development of Columbia's video-on-demand testbed, *Image Communication*, Special Issue on Video-on-Demand and Interactive Television, (to appear, 1995).
4. T. Kwok, A vision for residential broadband services: ATM-to-the-Home, *IEEE Network. Mag.*, Vol. 9, Nr. 5, pp. 14–28, October 1995.
5. Object Management Group, Universal Network Objects Version 1.0, Common Object Request Broker Architecture (CORBA) Version 2.0.
6. ISO/IEC CD 13522-5 (Committee Draft), Multimedia and Hypermedia Experts Group (MHEG), July 1995.
7. S. Dixit and P. Skelly, MPEG-2 over ATM for video dialtone networks: issues and strategies, *IEEE Network. Mag.*, Vol. 9, Nr. 5, pp. 30–40, October 1995.