Integrating distributed multimedia systems and interactive television networks

Alex Allister Shvartsman

Massachusetts Institute of Technology, Laboratory for Computer Science,
545 Technology Square, NE43-340, Cambridge, MA 02139, USA

Abstract

Recent advances in networks, storage and video delivery systems are about to make commercial deployment of interactive multimedia services over digital television networks a reality. The emerging components individually have the potential to satisfy the technical requirements in the near future. However, no single vendor is offering a complete end-to-end commercially-deployable and scalable interactive multimedia applications systems over digital/analog television systems. Integrating a large set of maturing sub-assemblies and interactive multimedia applications is a major task in deploying such systems. Here we deal with integration issues, requirements and trade-offs in building delivery platforms and applications for interactive television services. Such integration efforts must overcome lack of standards, and deal with unpredictable development cycles and quality problems of leading-edge technology. There are also the conflicting goals of optimizing systems for video delivery while enabling highly interactive distributed applications. It is becoming possible to deliver continuous video streams from specific sources, but it is difficult and expensive to provide the ability to rapidly switch among multiple sources of video and data. Finally, there is the ever-present challenge of integrating and deploying expensive systems whose scalability and extensibility is limited, while ensuring some resiliency in the face of inevitable changes. This proceedings version of the paper is an extended abstract.

Keywords: Interactive Television, Distributed Systems, Multimedia, Networks, Systems Integration, Object-Oriented Systems.

1. INTRODUCTION

Computer and electronics manufacturers and software systems providers are rapidly developing commercial components as telecommunications companies are investing billions of dollars in the deployment of networks and systems in preparation for the rollout of interactive television services [28]. Many of these activities are further being stimulated with the encouragement of the FCC (Federal Communications Commission) and by the continued attention given to the future National Information Infrastructure.

While the individual components are becoming available to provide coverage of the delivery systems ranging from video and application servers to delivery networks to set-top terminals, no vendor is offering a complete end-to-end massively deployable interactive multimedia digital/analog television system. A major part of deploying such a commercial system consists of integrating a large set of maturing components and sub-assemblies. These challenges are magnified by the lack of complete portfolios of established standards, by the long gestation of technology that is being developed under intense competitive pressures and by the sometimes conflicting goals of optimizing systems and networks for delivering video while enabling highly interactive distributed multimedia applications. The standards will make the job of integrating such systems easier, but it will take some time for effective set of standards to emerge – successful standards usually follow successful implementations and not vise versa. Delivering better-than-VCR quality video, e.g., 5 Mbps (megabit per second) MPEG-2 [1], requires that maximum performance is extracted from the magnetic disk storage used for these purposes. Commercial components used to implement high bandwidth delivery networks still require relatively long setup times. Thus, while it is now possible to deliver continuous high quality video streams to end users, it is difficult and expensive to provide the capability to rapidly switch between multiple sources of video and associated data and many technical problems remain to be addressed [9]. While the vendors of equipment have concentrated on solving the video delivery problem, the telecommunications companies also need the ability to delivery highly interactive multimedia applications without which little can be offered to the consumers. Integration of interactive television systems requires imposing creative synergy on the basically static video delivery capabilities with dynamic interactive applications.

 0-8194-1976-6/96/$6.00

In defining the software service applications it is important to consider the short term integration and deployment, as well as the longer term strategy. For early technology adopters, the issues of delivery platforms and network integration sooner or later dominate the scene. The challenge is how to enable the delivery of adequate video and interactive applications while scaling back on the valuable in the long term application architecture and how to protect the overall investment against premature obsolescence. The alternative is to opt for later adoptation of mature, standard or commodity systems, but risk a possibly diminished market share.

The material presented in this paper is based on actual definition, development and integration efforts, however the presentation is vendor-independent and service-provider-independent. While this report is necessarily preliminary, we believe that the issues we present will have to be dealt with by any near term ITV integration and roll-out effort. The technology is evolving rapidly and the state-of-affairs will change, yet based on our experience in the past twelve months these changes are slower than predicted.

This paper is structured as follows. In Section 2 we survey the main ITV services currently under development. In Section 3 we present a typical architecture for ITV services. In Section 4 we discuss development and integration issues. We conclude with a general discussion in Section 5.

2. DISTRIBUTED MULTIMEDIA AND ITV SERVICES

Interactive television (ITV) services differ from the traditional passive television in the way the multimedia content (video, audio and data) is delivered and in the degree of interactivity. The content is delivered based on interactive requests and not according to a pre-established schedule. The viewers are able to alter the sequencing and in some cases the delivered content in the sense that the delivery is customized in response to viewer actions.

2.1 Interactive multimedia services

The following services are targeted by several interactive television initiatives: Video on demand, News on demand, shopping, distance learning, and entertainment and games [6]. We now cover these services in more detail.

Video and movies on demand: Video-on-Demand (VOD) and Movies-on-Demand (MOD) services are considered by many to be the flagship of any interactive television service. The basic service includes the ability of a consumer to select a video program or a movie from the list of available titles and view the program on demand, that is, without prior reservations or scheduling. The consumer is able to use "VCR-like controls", which means the viewer is able to (at least) pause the program, rewind it, and fast forward through it. The provider of such service needs to maintain an extensive on-line library of video titles to appeal to the largest audience possible. The underlying video delivery network must be able to provide dedicated channels of sufficient bandwidth to the customer premises. At the same time it is expected that such service will provide video quality that noticeably exceeds the quality provided by a home VCR*.

Near-Video-on-Demand (NVOD) is a closely related service that offers somewhat lesser flexibility to the consumers at a lower cost to the provider and the consumer. The idea here is to supplement the dedicated on-demand channels to consumers with the frequent periodic broadcasts of the most popular titles. However, in some cases NVOD is harder to implement than VOD because digital broadcast setup is more complicated than a single dedicated session.

News and reference services: New-on-Demand (NOD) and reference services use technologies similar to VOD, but instead of a passive on-line library of titles, NOD provides a sophisticated news retrieval and reference services that combine live and archived video, access to textual data and still photography from a variety of sources [15]. The information is delivered to the consumer based on the interactive requests or based on the scoping and filtering criteria defined by the consumer. Peak video delivery requirements here are the same as for VOD, but a NOD service imposes significant processing requirements to carry out to search and cross reference the available data and information in response to the consumer requests. The searching and filtering activities may need to be persistent in the sense that the processing will be performed not only during the actual use of the NOD service, but between NOD sessions.

---

* These requirements are difficult to satisfy. Digitized video requires from 150 to 250 Mbps communications bandwidth. Using video compression reduces this requirement to about 1.5 Mbps and provides home VCR-class video quality. To provide higher quality, the video needs to be compressed to about 3 Mbps. To transmit such video to customer premises, the communications requirements is about 50 times that of the conventional telephone service bandwidth of 64 Kbps. A VOD library will need to contain hundreds of video titles to be viable. At 3 Mbps and assuming 100 minutes per title, the on-line storage capacity is about 225 Gbytes per each 100 titles. A VOD service may have peak usage at about 10% of the subscribers. To deliver the kind of video we describe, the delivery network must be able to provide up to 300 Mbps combined bandwidth to every 1,000 subscribers. The challenge here is both to be able to extract the video data from the on-line libraries and to deliver the resulting video streams to the consumers [3,5,8,22].

Interactive shopping and electronic commerce: Home shopping will provide a customizable shopping environment that will far exceed the functionality of the electronic catalogue shopping and passive TV shopping available today. Consumers will be able to rapidly and effectively focus on the products and services that are of direct interest to them. The shopping service will provide a multimedia interface that will allow users, in addition to receiving text-oriented information, to observe and examine the products/services visually and to see the products of interest in action. This will be accomplished in a consumer-directed setting with the users specifying what amount of information and what types of product presentation are preferred by them. Parts of the interactive shopping services will also be similar to the NOD service. The shopping service will have a persistent component that will continue searching for products requested by the users and filtering out unwanted information during and between the interactive shopping sessions.

Entertainment and games: Interactive entertainment may become a frequently used service. Interactive entertainment may take the form of NOD-like exploration of the topics of interest to the consumers. World-Wide Web provides one browsing-oriented interaction paradigm. ITV may substantially improve on this by providing higher quality multimedia component by taking advantage of the higher available communications bandwidth. Games will initially consist of relatively simple applications that will be down-loaded to the set-top device, thus not incurring significant costs associated with the use of the server and network facilities. In the longer term, the games will become more sophisticated. This means that the server facilities will need to be involved to provide greater processing and storage needs of the games. At the same time games will evolve from purely human/machine paradigm to multiple human/machine paradigm. This will require both the consumer-to-server and consumer peer-to-peer communications. Such communications will not be available in the first wave of ITV deployment, but may become commercially viable in the future.

Distance learning: Educational interactive programming and distance learning and are areas where ITV has significant potential [26]. Current indications are that these services have the potential of being very popular, however they might not have sufficient commercial value to their providers. It is expected these that the deployment of such services will be motivated by regulatory or promotional factors.

2.2 Some commercial and consumer issues

We now briefly review the main issues driving the vendors of ITV equipment and systems, future providers of ITV services, and the issues of acceptance of ITV services by their future consumers.

The future ITV service providers include the current providers of cable, telecommunications, and telephony services. These enterprises are working under significant marketing pressures to offer digital/interactive television services. There are also the time considerations associated with the FCC approvals of ITV trials and deployment. The task of the provider of the ITV network *and* service is also made more difficult by the FCC requirement that the video delivery networks allow equal access to other ITV service providers [7]. In some cases this also requires that the development of a regulated delivery network is performed separately from the unregulated service development.

The vendors of ITV components include network technology providers, digital video services vendors, interactive server and software vendors and set-top device vendors. All vendors are working under the pressures of market and under the pressures of their consumer base, i.e., the future ITV service providers. The technology is emerging relatively rapidly, but the pressures apparently exceed the vendors' ability to deliver production-grade, commercially viable products. The vendors are forced to commit to aggressive schedules or fear losing a potential market share to real and imaginary competition. The vendors, in their attempt to market the emerging technology and to please their customer base are overextending themselves to the point where their ability to forecast the availability of products is negatively affected.

The consumers, in the end, will determine the commercial success of the services offered by the ITV providers. One of the highly speculative questions is which services will be popular and which will not generate much interest. While VOD service has high priority in the current efforts, it is not clear whether the viewers will find the service affordable and use it often enough to make it profitable. There are also indications that games and interactive entertainment will be most popular. On the other hand, the expanding availability and the improving quality of other electronic services and interactive applications using more traditional networking will continue to undermine future ITV services.

3. AN ARCHITECTURE FOR INTERACTIVE TELEVISION PLATFORMS

We now present an architecture for ITV systems and applications and set the stage for presentation of integration and technology issues in the next section. The architecture is presented at the level of abstraction suitable for the context of this paper, but more importantly at the level of abstraction that takes an integration-oriented view (for example, although it is important to understand a video server architecture, we do not present it here, since it is a fixed point in the integration process). We examine the architecture along two dimensions. We first take a high level network view of an ITV architecture. Next we take a distributed software systems and frameworks view of the architecture.

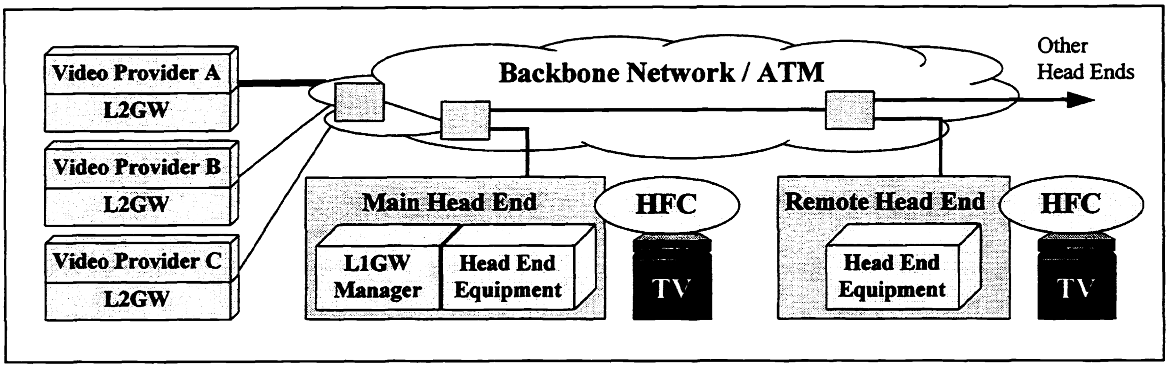

Figure 1. High level view of the ITV networks and system components.

3.1 High level networks and video delivery platforms view

The emerging ITV delivery systems have a number of similarities that are due to the state of the technology or that are caused by the regulatory issues. In Figure 1 we illustrate one of the approaches. The regulated part of the ITV deliver network, called the Level 1 network, provides the basic ITV service delivery capabilities to the customer premises. The Level 1 network is also called the video dial tone (VDT) network. The Level 1 network includes the backbone digital network (in our case based on ATM, asynchronous transfer mode [10]), the necessary head end equipment and the Hybrid Fiber/Coax local distribution network (HFC [19]), and the Level 1 gateway (L1GW) functionality that controls the connection establishment between the set-top and video service providers. Level 2 consists of the non-regulated components. Level 2 includes the video service providers, shared or dedicated Level 2 gateways (L2GW) that control the establishment of the application connections, implement service application loading to the set-top boxes and control access to the ITV servers.

The ATM backbone and HFC local delivery approach is one of several possible architectures. As of this writing, the approach satisfies the established technical requirements and cost-effectiveness constraints. For other possibilities see the surveys in [12,14]. The head end facilities serve as the gateways between the ATM and HFC networks. The head end is also used to detect and forward upstream communications from the customer premises. Analog TV programming can also be inserted at the head end.

As we have already suggested, this platforms view is hardly surprising given the available technology and regulatory requirements. In the next section we consider software systems and frameworks view that isolates the application systems from the physical platforms view and that offers significant flexibility to the ITV service developers and providers.

3.2 Systems and frameworks view

We now present a frameworks-oriented architecture (Figure 2). The main advantage of such an approach is that it provides a homogeneous environment to software applications across the server platforms and the set-top. Once the general purpose distribution services are provided by the frameworks, the software development can concentrate on ITV applications and supporting systems. Software will also remain the most flexible component in the distributed system and it is important to maximize this flexibility from the very start. This is important during both the development and deployment stages. We discuss the advantages of the frameworks approach in more detail in Section 4. We now briefly describe the main components in this view starting with the set-top application and the top layer.

The set-top application provides the user interface and serves as the client for the server-based ITV applications. The client is down-loaded upon the establishment of a Level 2 connection with the server. Because the set-top resources are limited, parts or all of the client may be reloaded as defined by the application or upon switching from one ITV service to another. The server applications include of the multimedia application proper and supporting systems, e.g., billing.

The client and server application frameworks define and implement the environment within which the applications operate. The client framework includes a set-top operating system and the media player subsystem. The frameworks allow for applications to be effectively distributed on multiple servers for performance, load balancing, and faulttolerance. The frameworks also enable integration with external systems. The actual distribution facilities are provided to the frameworks by an ORB (Object Request Broker) or an ORB-like mechanism where a standard ORB is not available, e.g., on the set-top. This enables the transparent communication between the set-top client and the ITV services provided by the applications running on the server complex. There are many benefits in adopting an object-oriented paradigm provided by an ORB, but other approach, e.g., RPC (remote procedure call), can also be used. The video delivery subsystem deals with the delivery and consumption of video signals. Because of the performance issues and the continuous nature of video data, the delivery of video is normally done out-of-band to the ORB.

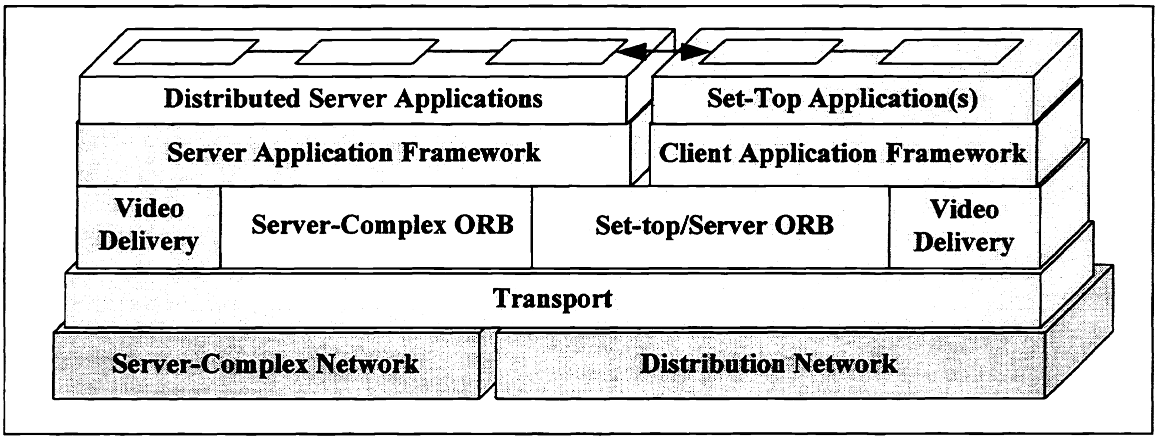

Figure 2. Frameworks architecture.

The transport layer provides end-to-end connections for server-to-server and client-to-server communications. It hides the non-homogeneity of the underlying networks. The distribution networks includes the ATM backbone and the HFC local distribution. The server-complex network is used by the servers for any out-of-band communication needed to provide the ITV service. It is not explicitly shown in Figure 1, but it uses internal service provider networks and it may be included in parts of the backbone network.

The networks-oriented view of the ITV systems (such as the view in Figure 1) is complemented by the application and frameworks view (such as in Figure 2) to enable the successful development and integration of ITV services. The networks and individual components form the necessary foundation, but it will be the well-integrated viewer-oriented applications that determine the success of an ITV service.

4. INTEGRATION AND TECHNOLOGY ISSUES

While the development of the enabling technology and support systems is normally done by the vendors with expertise in the relevant technologies and systems, the formulation and development of the ITV applications is addressed at the level of the future service provider. The service provider needs to direct the service definition activities to consider both the short and longer term strategy. The short term strategy deals with the deployment of an initial service as soon as possible using the emerging technology (as discussed in Section 2.2). The longer term strategy deals with the definition of extensible and scalable architectures that enable the delivery of multiple ITV services. It is tempting to concentrate on the longer term strategy assuming mature state-of-the-art systems, but given the unpredictable state of technology and uncertain acceptance of ITV by the future consumers, it is important to have a simplified, quicker-to-market approach. For the shorter term strategy, the challenges are (a) enabling ITV services while scaling back on the long term strategic architecture, and (b) building and deploying the expensive large scale systems today while paying due attention to the resiliency of the systems in the face of the inevitable change. Regardless of the strategy, the integration of immature and not fully standardized systems, applications and networks becomes one of the high priority issues. We next present a more detailed view of the networks and systems architecture introduced in Section 3.1 then we discuss important integration issues that should be expected in any ITV deployment and integration effort.

4.1 Networks view

We now take a functional view of the delivery network architecture (Figure 3) to set the context for the discussion (references are provided for the readers requiring greater detail). The major components are the video provider complex, the ATM backbone, the Level 1 and 2 gateways (L1GW and L2GW), the head end, the HFC, the set-top (STB), and finally an off-the-shelf TV.

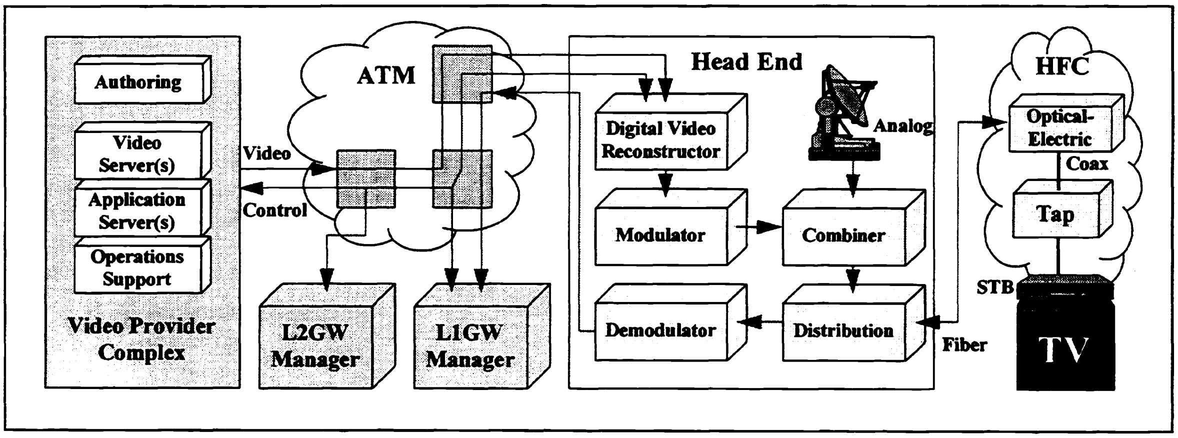

Figure 3. An expanded view of ITV systems and delivery networks.

The video provider complex is the main Level 2 entity and it includes one or more ITV services and an authoring environment. The services are implemented as distributed applications residing on the video and special-purpose servers. See [16,21,25] for information on video servers, their features and possible architectures. As we briefly described in Section 3.1, L1GW and L2GW provide connection establishment and session management for the regulated Level 1 and unregulated Level 2.

The head end contains digital video reconstruction equipment, modulation/demodulation equipment, signal combination components and the optical distribution output components (see [11,19] for detailed description). The reconstruction equipment is used to reconstruct the video stream MPEG-2 transport packets arriving in ATM cells. The packets are then modulated the packets onto an RF carrier (so are the downstream signaling messages to the set-top). Analog broadcast, digital video and signaling data are then combined for optical distribution to the HFC network(s). The upstream signaling from set-tops is also demodulated at the head end. HFC local distribution network consists, at a high level, of optical-electrical converters, amplifiers and taps to the customer premises. The set-top device residing on customer premises can be provided either by the ITV service provider or the consumer.

4.2 Service applications view

We now examine the top layers of Figure 1 in more detail. Development of software and integration of distributed multimedia resources in an ITV system of significant size must recognize the *object-based* nature of the physical and logical entities comprising the enterprise. The systems must deal with numerous classes of objects, large numbers of object instances, a very large number of attributes and the correspondingly large number of events and changes in the values of objects' attributes. It is important for the systems not only to operate within the prescribed parameters at any given instant of time, but also *evolve* in time by incorporating new object classes and mutating existing classes. The system must also be resilient in the face of both the planned and unplanned changes in the enterprise networks, computing resources and other network-based resources. Finally the system must be able to deal with the evolving *distribution* requirements, with *scalability* of the dynamic evolving enterprise, and provide for application distribution *transparency*.

The application-oriented view in Figure 3 is predicated on the use of a framework for the client/server applications. The properties expected in a state-of-the-art framework include [23]: (a) a heterogeneous set of hardware, software and data components, (b) size and geography that can vary over a large range, (c) interconnected uniform set of services, e.g., naming, remote invocation, registration, time, security, (d) global availability of the services. Among the necessary features we also count an integrated information repository and facilities for class definitions, module integration, repository-based application integration, and object-oriented dispatching of operations. The framework interfaces must be structured to enable the evolution of the framework itself without the need for the existing applications to track these changes. The frameworks-based approach is particularly effective for integration of custom software applications and external systems. The framework can provide an object-oriented view of components that by design streamlines the integration of compliant components [24]. The framework empowers the application developers by offering a standard set of tools for developing distributed systems such as DCE [18] and CORBA [17]. Such frameworks allow the designers to concentrate on the application and not on the problems of integration, provision of object features, implementation of distribution and concurrency. The framework addresses all of this once for all applications.

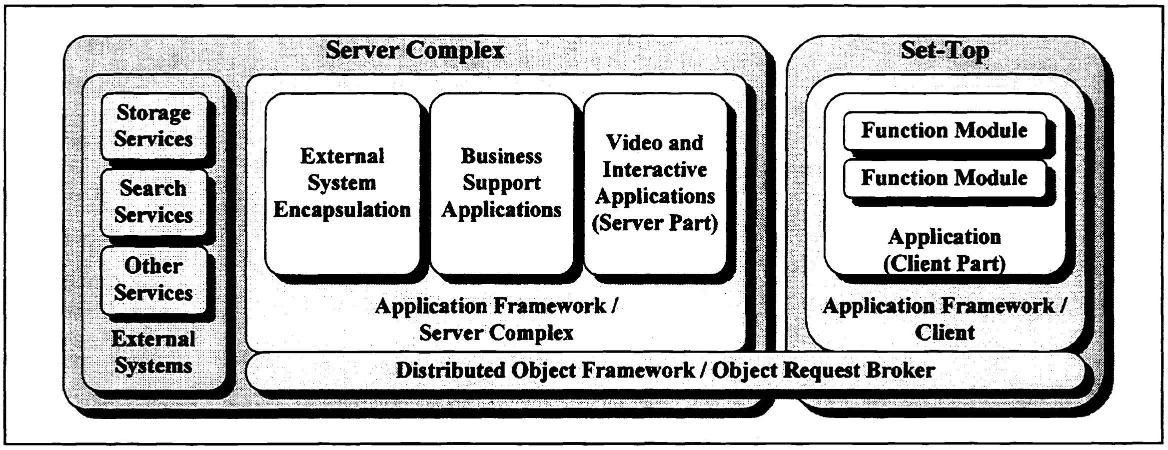

Figure 4. Frameworks-based view of applications and external systems integration

The custom applications are developed within the framework using a consistent and uniform view of the resources, services and applications comprising the client set-top and the server-complex. External systems are used to provide functions that are not cost-effectively developed as custom applications. These external systems might be directly integratable within the software framework. Where this is not possible, external systems are encapsulated by thin or value-adding wrapper within the framework. Such external systems may include databases, file and object storage systems, search services, and operations support systems. The application framework also extends into the set-top environment. The set-top applications are able to access the server-part of the applications as if these were objects local to the set-top.

4.3 Function-adding integration

Since the vendors are developing new software products now, the future deployers of these systems are able to influence the design and the functionality of these systems. There are both advantages and disadvantages in doing so. Normally a vendor would have a vision that may or may not be consistent with the plans of the deployers. If the vendor's vision is more coherent and far-sighted, it is better to align the deployment plans with that vision. On the other hand, the future providers of ITV services may have a better strategy. In that case the vendor may benefit from aligning their product with the customer's strategy. However, because of the advanced nature of this development it is hard to comparatively evaluate different strategies that are being pursued by several vendors and future ITV service providers.

One alternative is to defer to the vendor's vision and extend or refine the functionality of the (future) off-the-shelf software product by additional in-house development. This is one area where the choice of the overall architecture, software platforms and integration frameworks is extremely important. The goals of software integration are not limited to pure integration of systems. System isolation and encapsulation is equally important. The integration effort must attempt to both enable the use and deployment of the emerging systems and to ensure the overall system resiliency in the face of change. The integration of software components should not be limited to thin encapsulation of the components for the purposes of inter-operation. In some cases, it is important to pursue *value-adding integration* that not only ties the components into a well integrated system, but extends components and changes the component view of the system to align this view with the defined architectural approach and overall strategy.

4.4 Incremental development and integration

Given that the application requirements and the associated technological feasibility are still evolving and that the existing technology is still not sufficiently mature, the development and integration processes are most effectively carried out using spiral and incremental approaches. Spiral development helps assure that the development can commence early and that the systems under development can successfully evolve to completion while incorporating new technology and responding to changing requirements. Using incremental development and integration we can start with a relatively small configuration that makes many simplifying assumption, e.g., relative to the views in Figure 3 and Figure 4, and then incrementally incorporate additional necessary sub-systems en-route to the full deployment environment.

4.5 Using simulated and development environments

If the development of the ITV-enabling software and hardware systems needs to progress concurrently with the service application development and the integration of such systems, then the application and integration will necessarily use pre-production version of the enabling components or, simulated components and simulated development environments. Simulated components are used when even the pre-production versions are scarce or not available. The use of such components helps establish quantitative performance and sizing metrics. They are also used as surrogates in early integration with other components. Development environments are used to simulate the execution environment when the development platforms are distinct from the deployment platforms. Such simulated systems may be provided by the vendors or development for internal use by the ITV service developers. The use of simulated environments allows for the development and integration activities to proceed but has associated risks and difficulties. Since the emerging components are normally not well-defined, the simulations must make assumptions that may or may not be consistent with the eventual production-grade components. Even when the simulated tools are provided by the vendors, additional effort will be required to maintaining consistency between versions.

4.6 Integrating early product versions

During the pre-release use of the emerging products, the vendors are understandably reluctant to commit to formal specifications of their products. This means that during the integration of such products they are being viewed as essentially black boxes. When problems are encountered, it is very difficult to focus on the possible components that are causing the problems and diagnose the reasons. When a black-box component is used, any major externally observable problem leads to significant delays in its rectification because the resolution requires that the already scarce and thinly spread vendor resources need to be diverted from mainline development activities to diagnose problems. Another problem is that the vendors may not be able to provide support and maintenance for early version once another version is available. However, the newer version may be incompatible with the environment in which the older version is deployed and this requires that the older version remains in use until the new one is re-integrated.

4.7 Video, application and special purpose servers

The development of commercial video servers has been proceeding at a slower pace than anticipated. The servers still are far from the projected video output stream capacity and the ability to store video content (see Section 4.14) and are only now beginning to approach production-grade quality suitable for cost-effective deployment. Integrating video servers within an ITV service has to date been very challenging. The situation is better with the servers that are used to implement services that are closer to traditional off-the-shelf interactive or transaction-oriented computer systems.

4.8 Set-top devices

The digital set-top is a component without which ITV deployment is impossible. The set-top is a special purpose computer that includes a CPU, RAM memory for applications and supporting frameworks, and ROM containing parts of an operating system and communications facilities. The specialized modules include RF tuning, video/audio processing, graphics, modulation and communication modules. For a set-top model in the development stages, only a limited number of prototypes may be available for integration and testing. The prototypes are of course much more expensive than the eventual production units whose target cost is in the few hundreds of dollars. In addition to normally expected changes, the compatibility of prototypes with the production units may be affected by changes due to cost considerations and evolution of the design for high volume manufacturability. When prototypes are available, their use may be confined to the labs, since their deployment on customer premises may require regulatory and safety certification.

4.9 Test equipment issues

The availability of integrated test equipment is somewhat less than satisfactory. Although test equipment is available to assist with protocol analysis and network integration troubleshooting, for the newer areas, such as MPEG, the functionality of the available equipment falls short of expectations. It is often the case that several pieces of test equipment are required to provide sufficient coverage. Using such a portfolio us in many cases difficult because, as the result, the tools used by the integration and test engineers are not well integrated.

4.10 Near term backbone network evolution

The broadband networks used as the back ITV delivery networks will only be fully proved with massive deployment of

ITV services. The backbone networks based on ATM that we consider must resolve a number of interoperability issues as full protocol suites are being developed. In the area of ATM switching, the network may include switches of two types, PVC (permanent victual circuit) and SVC (switched virtual circuit). PVC establishes pre-provisioned connections while SVC allows for the connections to be established in real-time through signaling. Since PVC, once provisioned, does not require time-consuming setup it is fast but cannot be adapted in real-time to dynamically changing communication requirements. Due to costs and availability, PVC is normally used in the initial network deployment (and when the communications patterns are well understood and are fixed). With the future availability of cost-effective SVC solutions, the backbone network will need to migrate to SVC in the areas where dynamic switching is an advantage. This involves re-integrating the network, upgrading some ITV components and deploying additional signaling protocols.

4.11 Co-existence with other services

Other telecommunications services such as telephony and analog cable service will need to co-exist with the digital ITV services. It may be easier to provide a new approach using exclusively digital services, however compatibility and co-existence with the existing services must also be considered. There are regulatory pressures to offer multiple services over the same network to justify costs being passed on to consumers. Some of the networks that are being deployed now may be used by other services even before the digital service is deployed. The challenge is to ensure that resources can be used cooperatively and without interference. Significant testing and integration is required to guarantee this.

4.12 Content acquisition, preparation and maintenance

The ability to provide extensive video libraries is gated by expedience with which video content can be acquired and prepared. Content providers are still working on the business and copyright issues. The content preparation process is bound to be time consuming and have lengthy lead times. The automated approaches to loading new content, providing high concurrency access to popular content and archiving less popular content need to defined and implemented. Other issues include insufficient standardization (see Section 5.4) and limited server storage (see Section 4.14).

4.13 Business support systems

Another area that needs to be addressed by any commercial ITV system is its billing model. The model has to be implemented to provide flexible Level 1 and Level 2 billing structures. When a billing system is developed, it needs to be integrated with any other system that is capable of generating billable events or performing billable activities. When a vendor-provided system is used, it needs to be normalized to be consistent with the enterprise billing model. Here we need to consider time-oriented patterns based on the duration of usage (similar to long distance telephone billing), per-use patterns (similar to some pay-per-view billing where a fixed charge is applied after a preview but regardless of whether the viewing was complete), or hybrid patterns. One also needs to consider value-pricing of services. For Level 1 billing, the tariffs may also need to be approved from the regulatory standpoint. This may greatly complicate the integration effort and may also require component suppliers to change their respective billing models.

4.14 Scalability of integrated systems and post-integration scalability

An additional challenge in attempting to integrate pre-production grade components is that, in addition to the quality and functionality deficiencies of the early versions, there are performance and scalability issues that might not be resolvable using the available versions of components. Will application servers be able to support viewer interactivity? Will video servers be able to deliver sufficient numbers of video streams? Will the Level 1 and Level 2 gateways provide adequate response to connection management activities? Will the network switches be able to provide adequate levels of service under the design loads? These questions will not be fully answered until the (near-)production components are available together with verifiable performance parameters. Until then, it is difficult to design the server complex and the delivery network with a high degree of certainty that the designed configuration will support the required services.

In many areas touched on above, the forecasting of the eventual component performance levels is a difficult task. The vendors of the equipment have been working hard to build the components with the required performance characteristics, but there is still a long way to go. In many cases, the performance estimates have been downgraded by significant factors*. Developing and analyzing performance models is one way to proceed, but the work will not

* To illustrate the situation, consider a video server that may have been designed to deliver 300 MPEG streams at 3 Mbps. If we took a snapshot today, the vendor may have downgraded the number of streams to 100, while the pre-production versions can deliver about 10 streams. This means that using the 10 streams today we need to be able to predict the system performance up to a year from now using the full range of streams.

necessarily yield meaningful results because so many parameter are in flux. Another approach is to invest in a simulation setup. This can consist of both hardware and software systems. Doing so is not easy since additional equipment and resources have to be dedicated and this will affect the main development and integration activities.

5. Discussion and Issues

We have presented a broad view of the integration issues in distributed multimedia and ITV systems. The challenges facing the ITV industry and its suppliers are still formidable. Below we address additional relevant topics.

5.1 To buy or to build software-intensive components

An enterprise intending to implement an ITV service will need to make several buy-vs-build decisions. Most of such decisions will be in the software systems area. Given that most ITV components are not available off-the-shelf, it is expected that in some cases the decision to "buy" may delay the implementation of a robust service because of the diminished control over the development. The decision to build implies more control, but also greater responsibility and expense. To build, one must also possess the necessary development resources or be able to expediently acquire such resources. For an enterprise that is used to ordering and installing off-the-self systems, the proposition that the enterprise needs to also become a developer and integrator of non-trivial software systems is not one always easy to deal with.

5.2 On Technology Assessment and Integration Planning

An integration effort can be greatly aided by continued and focused technology assessment. This activity starts during the general technology evaluation leading to the selection of a group of vendors and continues during integration while covering the selected vendors as well as other potential vendors. The forecasting and the actual progress that vendors are making needs to be monitored to enable the meaningful planning of the integration activities. The vendor-provided delivery dates, even for the pre-production equipment and systems, are usually unstable and tend to be pushed out. This does not only postpone the planned integration and test activities, but it results in frequent changes to the order in which certain components at certain functionality levels are introduced in the integration effort. In some cases this causes the entire effort to be replanned. To optimize the planning, it is prudent to maintain a trade-off between engineering activities that are based on a reasonable planning horizon and reactive activities that are the result of the changes in vendor delivery schedules. It is important to assign dedicated resources to each vendor supplying major components. These resources should consist of technical and management liaison personnel. For a major vendor, it is recommended that a dedicated technical liaison engineer is assigned. The liaison personnel will need to maintain regular teleconferencing contact with the vendor(s) and periodically visit the vendor facilities to understand the state of the product development and conduct quality and engineering progress assessments. This ensures that the integration effort is well informed on what partial and pre-production deliveries can be made and when. The benefit is, of course, that (re)planning is done based on factual information and that the integration completion forecasts avoids last minute surprises.

Finally, given the preliminary state of technology, it may be necessary in some cases to introduce additional vendors or even replace vendors as the result of changes in the course of integration. The on-going technology assessment effort ensures that when such changes are unavoidable, the decisions are made expediently and from the position of knowledge.

5.3 Involving the Vendors

The vendors of individual components are primarily and naturally interested in developing and delivering their respective components. When it comes to integration, the cooperation among the disparate vendors in many cases takes the back seat to vendors' individual concerns. It is important that the vendor cooperation is sustained from the very early stages down to the final configuration of the end-to-end system. Vendor commitment to integration also, sooner or later, ensures that by solving integration problems vendors directly help themselves and reduce the longer term maturation costs of their individual products. In the new technology area such as ITV, the service providers, integrators and vendors will benefit from close cooperation and synergy. ITV services cannot be provided without the available technology and any technology cannot be successful unless it is massively deployed.

5.4 Standardization and commoditization issues

Video providers will eventually arrive at a set of standards (*de-jure* or *de-facto*) for interoperable services, delivery platforms and content formats. This means that most of the builders and deployers of early ITV systems whose approaches were not chosen as the basis for standardization will need to migrate to the standards to be commercially viable. In some hardware areas the risk is low, e.g., ATM/SONET, in others it is possibly high, e.g., set-top or video server interfaces. There are may be limited mitigation actions in the networks/hardware area, but using frameworks-based approach (as we discussed in Section 4.2) will help make systems resilient to change.

- Some of the unresolved standardization issues include signaling among ITV components, content encoding and software interfaces. In the area of signaling standards are emerging, but proprietary signaling is also used among the set-tops, gateways, servers and head-ends, and where the specifications are open they are incomplete. Content encoding is only partially standardized. While MPEG encoding is being broadly accepted for video, it appears that the standard is still open to implementation interpretations. Standards will also needed for correlation of forward/reverse encoded content with the normal play content, general content packaging, authoring tools, and set-top, and server software interfaces and environments. In these areas standards may be supplemented by commoditization of components.

5.5 Set-top issues

ITV is, by its very nature, a client-centric application. The ITV systems architecture and the performance requirements will be dictated by the usage patterns that are established by the end users and by the massively deployable customer premises equipment. The deployment and growth of the ITV services will cause the eventual merging of the television set and the personal computer. Currently there are two opinions on this subject. One posits that the personal computer will evolve to become a TV set. The other assumes that the TV set will evolve to become a personal computer. Regardless of who will be right, ITV revolves around the piece of equipment residing on customer premises. The following issues will determine the competitiveness of set-top devices: (1) they must be sufficiently inexpensive for the ITV services to be commercially viable, (2) they must be sufficiently functional to house the client part of ITV applications and to be able to receive and process data information at the rates expected in ITV and, (3) they must be compatible with the conventional analog cable services and with the television sets that are in the tens of millions of homes.

5.6 Longer term service issues

Reaching the initial design, development and deployment targets is a very challenging task, yet we need to also look beyond the initial targets regardless of how modest or ambitious these targets are. It is important to start architecting the future evolution of the system to prevent early obsolescence and to enable feasible evolution of the deployed system if and when the demands on the service increase. These demands may be generated from two directions: (1) increase in the consumer base, and (2) increase in the service provider base. When the number of consumers in an area served by a particular serving facility exceeds the capacity of the facility, will there be a viable expansion path?

When providing equal access to the Level 1 network, will the second Level 2 provider noticeably degrade the quality of service originally available from the first provider? Note that in the world of telephony we do not associate the availability of multiple long distance providers with the poorer quality service across all such providers. The same might not be true for ITV. The existence of multiple long-distance providers does not increase the frequency of use, but an attractive new ITV service will increase the frequency of use. Also, while a telephone set is a very simple device that is not concerned with what service we are using, a set-top device is much more complicated and it may be required that different client applications are loaded into the set-top device.

5.7 Networks evolution

Emerging architectures such as the one we presented here, given today's technology, assume a tree-like main distribution network. The main video application serving facility is connected to the root of the network. The regional head ends are positioned at the leaves of the network. This makes sense, provided that most video and data is broadcast or multicast from the root and that the VOD or dedicated ITV application usage is relatively low. The scalability of such a target environment is reasonable, provided additional serving facilities can be deployed, provided sufficiently "thick" tree trunks are (or can be) deployed. Of course the (trivial) scalability achieved for new service areas by simply "cloning" delivery systems is also possible. The HFC networks, in addition assume asymmetric communications pattern where the downstream bandwidth is high and the upstream bandwidth is very low.

The longer evolution will include deploying more sophisticated ITV services that will include the need for more intensive server-to-server communications and much higher upstream bandwidth. As the services diversify, it becomes impractical to clone each kind of server. In some cases an integrated distributed multimedia application will require that several servers cooperate in delivering the service. In addition there may be small service providers that will require access to wide areas covered by regulated networks. Such changes invalidate the assumptions made in the tree-like network architecture. The network will need to become more general with several geographically distributed serving facilities that need high connectivity. Finally, applications will sooner or later require much greater upstream bandwidth from the customer premises for applications such as video conferencing and video mail. With such applications, the distinction between the topologies of the video delivery networks and general purpose digital communications networks is going to be diminished (we discussed related integration issues in Section 4.14).

6. ACKNOWLEDGMENTS

I thank Mark Huttemann for valuable discussions. Part of this work was performed at Logica, Inc. Part of this work was performed at MIT supported by ARPA Grant N00014-92-J-4033, NSF Grant 922124-CCR, and ONR-AFOSR Grant F49620-94-1-0199. Author's telephone (617) 253-6182; e-mail: alex@theory.lcs.mit.edu.

7. REFERENCES

[1] C.J. Birkmaier, "An Open Architecture for Digital Communication Systems", *IEEE Multimedia*, pp. 77-83, Fall, 1994.

[2] C. Carroll, "Development of Integrated Cable/Telephony in the UK", *IEEE Communic. Mag.*, vol. 33, no. 8, pp. 48-60, 2995.

[3] M-S. Chen, D.D. Kaudlur and P.S. Yu, "Support for Fully Interactive Playout in a Disk-Array-Based Video Server", in *Proc. ACM Multimedia'94*, pp. 391-398, 1994.

[4] L. Chiariglione, "MPEG: A Technological Basis for Multimedia Applications", *IEEE Multimedia*, pp. 85-89, Spring 1995.

[5] Y.N. Doganata and A.N. Tantawi, "Making a Cost-Effective Video Server", *IEEE Multimedia*, pp. 22-31, Winter, 1995.

[6] P.C. Ellis-Barr, "Systems Integration Issues for Interactive Television Services", *Telecomm. J. Australia*, vol. 6, pp. 31-37, 1994.

[7] FCC, CC Docket No. 87-266, In the Matter of . . . Regulatory Processes for Video Dialtone Services, October 20, 1994.

[8] G.R. Ganger, B.L. Worthington, R.Y. Hou and Y.N. Patt, "Disk Arrays: High Performance, High-Reliability Storage Subsystems", *IEEE Computer*, vol. 27, no. 3, pp. 30-37, 1994.

[9] I. Gopal, "Multimedia Networking: Applications and Challenges", *in 13th ACM Symp. on Distr. Computing*, pp. 9-12, 1994.

[10] A. Iwata, N. Mori, C. Ikeda, H. Suzuki, M. Ott, "ATM Connection and Traffic Management Schemes for Multimedia Internetworking", *Comm. of the ACM*, vol. 38, no. 2, pp. 72-90, 1995.

[11] C. Kaplan, "Interactive Video Services", *Telecommunications, IEE Conference*, Publication No. 404, pp. 323-326, 1995.

[12] R. Karpinski, "Video Dial Tone: Putting the Pieces Together", *Telephony*, pp. 6-12, July, 1994.

[13] B.G. Kim and P. Wang, "ATM Network: Goals and Challenges", *Comm. of the ACM*, vol. 38, no. 2, pp. 39-44, 1995.

[14] T.D.C. Little and D. Venkatesh, "Prospects for Interactive Video-on-Demand", *IEEE Multimedia*, pp. 14-24, Fall, 1994.

[15] G. Miller, G. Baber, M. Gilliand, "News On-Demand for Multimedia Networks", *Proc. ACM Multimedia'93*, pp. 383-392, 1993.

[16] K. Natarajan, "Video Servers Take Root," *IEEE Spectrum*, no. 4, pp. 66-69, 1995.

[17] *OMG, Common Object Request Broker Architecture*, Object Management Group, Framigham, MA, 1992.

[18] *OSF, Introduction to DCE*, Open Software Foundation, Cambridge, MA, 1992.

[19] A. Paff, "Hybrid Fiber/Coax in the Public Telecommunications Infrastructure", *IEEE Comm.*, vol. 33, no. 4, pp. 40-51, 1995.

[20] W. Pugh and G. Boyer, "Broadband Access: Competing Alternatives", *IEEE Communic. Mag.*, vol. 33, no. 8, pp. 34-46, 2995.

[21] R, Rooholamini and V. Cherkassky, "ATM-Based Multimedia Servers", *IEEE Multimedia*, vol. 2 no. 1, pp. 39-52, 1995.

[22] F. Schaffer and J-P. Nessbaumer, "On Bandwidth and Storage Tradeoffs in Multimedia Distribution Networks", in*Proc. IEEE Infocomm'95*, pp. 1020-1026, 1995.

[23] M.D. Schroeder, "A State-of-the-art Distributed System", in *Distributed Systems*, S. Mullender, Ed., ACM Press, 1993.

[24] A.A. Shvartsman and C. Strutt, "A Framework for Distributed Object Management and Generic Applications", submitted to *Theory and Practice of Object Systems (TAPOS)*, 1995 (also see Brandeis University Tech. Report CS-94-176).

[25] S. Taylor, D. Chin and S. Knight, "The Magic Video-on-Demand Server and Real-Time Simulation System", *IEEE Parallel and Distributed Technology*, vol. 3, no. 2, pp. 40-51, 1995.

[26] F.A. Tobagi, "Distance Learning with Digital Video", *IEEE Multimedia*, pp. 90-94, Spring, 1995.

[27] F.A. Tobagi, J. Pang, R. Baird and M. Gang, "Streaming RAID – A Disk Array Management System for Video Files", in *Proc. ACM Multimedia'93*, pp. 393-399, 1993.

[28] M.P. Vecchi, "Broadband Networks and Services: Architecture and Control", *IEEE Comm. Mag.*, vol. 33, no. 8, pp. 24-32, 1995.