THE INTERACTIVE EVOLUTION

Anthony J. Aukstikalnis

GENERAL INSTRUMENT
JERROLD DIVISION

ABSTRACT

For many years CATV systems have been viewed as ideal for the implementation of two-way interactive systems. In the late seventies and early eighties, several revolutionary systems were developed which failed primarily because they were technology driven, rather than market driven. They attempted to address a wide range of interactive applications and became too costly and too complex to operate. Presently, two-way interactive systems concentrate on specific applications, such as Pay-Per-View and Home Shopping, and are developing in a more evolutionary manner. This paper summarizes the current state of the art and explores possible approaches for evolving current embryonic interactive systems into ones with more wide spread applications. Candidate control system architectures are analyzed.

Interactive System Model

Figure 1 shows a basic interactive system model. There are three major system components to the model.

- The consumer who is requesting, receiving and paying for interactive services.
- The service component which delivers the service and receives compensation for it. In a cable based system the service component will consist of the organization providing the service and the cable system operator who provides the communications medium.
- The Customer Service and Billing component which provides the customer service, billing and payment processing functions. This component authorizes the consumer to receive services, collects data regarding the services used, bills the consumer and processes payments.

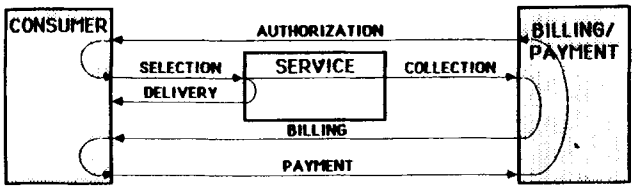

FIGURE 1: INTERACTIVE SYSTEM MODEL. LINES ARE COMMUNICATION LINKS. BLOCKS ARE SYSTEM COMPONENTS.

The communication segment can be separated into a control subsystem and a service delivery subsystem. The control subsystem consists of communications links which integrate the consumer, service and billing/payment components in the interactive system. It comprises the following.

- The authorization link which passes data allowing the consumer to use all or part of the interactive service offerings.
- The selection link, used by the consumer's equipment to request an interactive service.
- The delivery link which provides the requested service to the consumer. In today's interactive systems, the delivery link is usually a 6 MHz video channel.
- The collection link which passes data pertaining to the service a specific consumer has utilized.
- The billing link which periodically, usually on a monthly basis, bills the consumer for services utilized. The U.S. Postal System is currently used for both consumer billing and payment.
- The payment link which the consumer uses to pay his bill. Bill payment implicitly authorizes the consumer to continue to receive the interactive services he desires.

Historical Perspective of U.S. Cable Based Interactive Systems

This paper deals only with U.S. cable based interactive systems. There have been several telephone based interactive system trials, the most notable being the Knight-Ridder/AT&T Viewtron system and the Times Mirror Gateway system.

The earliest cable based interactive system to be widely installed was the Warner/Pioneer Qube system; a system which features pay-per-view programming and opinion polling. Qube uses interactivity to enhance the basic video services. Cox also introduced their Indax system, a more general purpose video and data oriented system, which went to field trial in San Diego, California and Omaha, Nebraska. Manitoba Telephone demonstrated their Omnitel system in Winnipeg which combined video, data and telephone traffic. General Instrument continued development of the Omnitel system in the United States. Other systems developed by General Instrument, Packet Technologies and World Video Library reached the prototype stage.

All of these systems, in one way or another, were technological successes and business failures. They were technology driven, rather than market driven. They all were not cost effective, primarily because of the high costs of running and maintaining the system and its data bases. Today, only Qube remains installed in some systems.

During the past few years, there has been a trend away from general purpose interactive systems to single purpose interactive systems. These systems concentrate on specific applications, such as pay-per-view and home shopping, rather than attempting to address the entire spectrum of two-way interactive applications. Even interactive game shows, such as Money Mania on the Nashville Network, are beginning to appear.

The present systems utilize a full video channel (live video) as the forward path and cable or (more likely) telephone as the return path. The degree of interactivity required is low and is accomplished by low bandwidth data communications or voice conversations. Major characteristics are:

- Video programming with a high entertainment content as the offered service.
- Quick response time, with emphasis on impulsive reactions.
- Cost effective implementation.
- Specific applications as opposed to general applications.

Interactive System Requirements

The major interactive system requirements are response time, conflict resolution, traffic handling capacity, cost effectiveness, simplicity and ability to handle multiple applications.

- Response time is the delay from requesting the service to delivering the service. From the consumer's viewpoint the response time should be as close to zero as possible. One to two seconds is acceptable for most services.
- Conflict resolution is the ability of the system to automatically deliver the desired service without the need for consumer retries due to a "busy signal".
- Traffic handling capacity is the ability of the system to accommodate the consumer population on a peak and average load basis. Bandwidth must be minimized in order to not severely decrease channel capacity for video services.
- Cost effectiveness refers to the ability of the system to be profitable for all participants - the system operator, the service provider and the equipment supplier. Cost effectiveness is a function of the value of the services offered to the consumer, the cost of offering these services, and the capital cost of the equipment involved.
- Simplicity refers to several facets of system implementation and operation. The interactive system must be reliable. It must be simple for the consumer to use, preferably from his easy chair while watching television. Most important, it must be simple for the system operator to operate. Manual operations of a highly complex nature must be eliminated. The more transparent an interactive system is to system operators, the better its chances of succeeding are.

- As we have seen, today's interactive systems focus on one application which makes their operation profitable. The ability to handle multiple applications will be important in the future if the current trends in interactive applications continue however, handling multiple applications must not overly burden system cost factors.

Candidate System Architectures

As shown in Figure 1, an interactive system can be partitioned into a communications segment, a service segment and system components. This partitioning is shown diagramatically in Figure 2. We have already covered the various interactive system components (consumer, service and billing/payment).

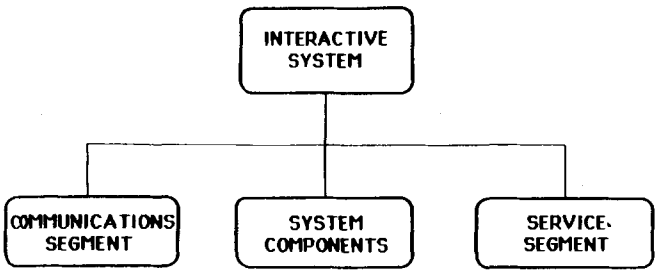

FIGURE 2: INTERACTIVE SYSTEM SEGMENTS.

The communications segment provides the links for system control and service delivery. There are various architectural alternatives for implementing the communications segment, the most common of which are shown in Figure 3.

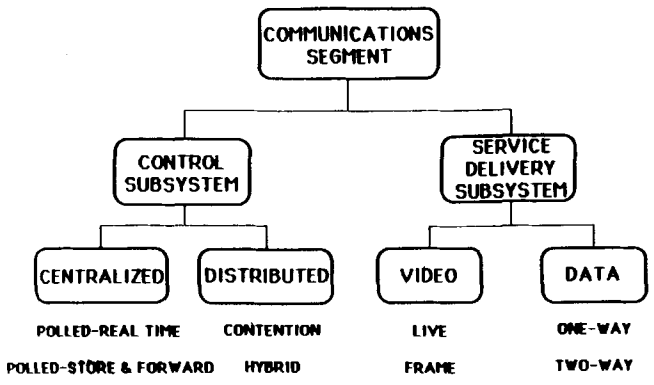

FIGURE 3: COMMUNICATIONS SEGMENT ARCHITECTURE TREE.

The purpose of the control system is to provide the link and protocols necessary to allow an orderly transfer of information between the consumer, service provider and billing/payment components. The control system can use either centralized or distributed control. Centralized control denotes a central facility which interrogates consumer equipment in order to maintain an orderly communications flow. Distributed control refers to a control system in which the consumer equipment initiates communications with central and other facilities. Distributed control requires more complex protocols and algorithms.

Referring to Figure 1, for all control subsystem architectures, the authorization, billing and payment links are very similar. For instance, authorization normally takes place over the downstream data link used for enabling addressable converters. The billing link can either be the U.S. Postal System or the downstream data link. Payment is via the U.S. Postal System.

Each control subsystem architecture will now be summarized, highlighting differences in the selection, delivery and collection links.

Polled-Real Time Control

Polled-real time control utilizes a central facility which addresses consumer equipment in predetermined sequences. Each consumer's equipment is interrogated, or polled, once every main cycle. When the consumer unit is polled, a two-way communication link is established and information is transferred from the polled unit to the central facility. The central facility performs authorization checks and delivers the service to the customer.

Response time is determined by when the request for service is made with respect to the polling cycle and the time it takes for the central facility to process the request. Delivery of service is not instantaneous, and in a large system the response time could take tens of seconds. The inherently slow response time dictates that cable, rather than telephone, be used as the return path.

Delivered service data collection is accomplished during the time that the central facility and the consumer unit are in a two-way communication mode. There is no delay in collecting service use data.

Polled Store and Forward Control

Polled store and forward control also utilizes a central facility which addresses consumers in predetermined sequences. The major difference is that the consumer equipment stores data indicating that a service has been selected and delivered. The consumer equipment is self authorizing. The polling cycle is used only to collect service use data and, therefore, does not have to be continuous, nor is it as time critical as in the real time system.

The store and forward system architecture depends on preauthorization of the consumer for the desired service set. No real time checking of consumer parameters (credit, authorization, etc.) or service availability is performed prior to delivery. Response time is almost instantaneous and there are no real limitations to system size, providing the consumer equipment has adequate memory for services requested and delivered.

The return path can either be via cable or telephone since, while telephone return is slower, there is no system requirement for fast polling cycles.

Contention Control Systems

Contention control was used in most of the distributed control architectures which were developed. Cox's Indax and General Instrument's MetroNet are examples of systems which utilized packet broadcasting with collision detection protocols, like Carrier Sense Multiple Access with Collision Detection (CSMA CD). ALOHA type systems which used a central facility to resolve common communication channel conflicts were also tested. Multiple narrow band channels are required to accommodate reasonable system sizes. The General Instrument MetroNet system, for instance, used multiple 300 KHz, 128 kbs channels which could accommodate 200 active users.

Service provider facility checks are made before a selected service is delivered. Response time is dependent on the number of active users and their average message length. In a well constructed system, delivery of the service is almost instantaneous; the critical parameter being the time a service provider requires to process the request. Collection of service use data is accomplished when the service facility and consumer equipment are in a two-way communication mode.

Contention systems require more complex protocols and algorithms for the consumer equipment. In larger cable systems, frequency agile transmitters and receivers will also be required. Consequently, consumer equipment tends to be relatively expensive.

Hybrid Distributed Systems

The hybrid distributed system uses the Public Switched Telephone Network (PSTN) for the return path (selection) communications and the cable system for service delivery. Customer selection is by means of a telephone keypad or a special keypad which interfaces with the PSTN.

Two-way communications are established over the telephone during the service selection process. The service facility performs authorization checks and delivers the service to the consumer. Data collection for delivered services is accomplished during the time when the service facility and consumer equipment are in a two-way communication mode. Response is dependent on dialing time and the service facility processing. This type of system is usually associated with video on demand, either pay-per-view or video frame dissemination.

Service Delivery Subsystems

Communications links are also required for delivery of requested services. The requested services can be video services, data services or a combination of both. Video services can either be live video (e.g., pay-per-view systems) or video frames (home shopping catalog services).

Data services can either be one-way (teletext) or two-way. One-way data services used on a pay-per-use basis could utilize a store and forward system to collect consumer use data for billing and payment processing purposes.

Two-way data systems allow data exchange between the consumer and the service facility, usually in the form of a request by the consumer and delivery by the service facility.

Today's services are usually live video with at least one one-way data application (X-PRESS). A natural evolution is the delivery of video frames to frame store equipment for catalog home shopping applications and other applications where reasonable resolution is required.

	Response Time	Conflict Resolution	Traffic Handling Capacity	Cost Effectiveness	Simplicity	Multiple Applications Capability	Communication Paths
Polled-Real Time	Relatively slow-depends on polling cycle time	Minimized-talk to one user at a time	Moderate-depends on response time requirements	Inexpensive in-home and system equipment	Simple equipment and operation	Limited	Cable downstream Cable return
Polled-Store and Forward	Fast, self authorizing	None required Self authorizing	High-non real time use data collection	Slightly more expensive in-home equipment	Simple equipment and operation	Limited-services must be on cable at all times	Cable downstream Cable or telephone return
Contention	Relatively fast-depends on number of active users per channel	Collison detection and retry algorithms	Moderate-requires multiple channels and user channel assignment	Relatively expensive in-home and system equip.	Relatively complex equipment and operation	Very flexible -data and video on demand	Cable downstream Cable return
Hybrid	Moderate-limited by telephone set up time	Retry if return path is busy	Relatively low -limited by PSTN	Inexpensive if PSTN costs ignored	Moderately complex if PSTN is considered	Moderate-data and video on demand	Cable downstream Telephone return

TABLE I: COMPARISON OF CANDIDATE CONTROL SEGMENT ARCHITECTURES

Comparison of Control Segment Architectures

Table 1 summarizes the previous discussion of candidate interactive system control architectures. Each of the candidates has strengths and weaknesses. Polled systems are relatively simple to operate and require minimum capital investment. They tend to be better suited for a limited applications set where the system is tailored to the predominent application. Most of the current cable based interactive systems are polled systems. Contention systems are relatively complex to operate, primarily because of their distributed nature, and are relatively expensive. They, however, are very flexible and can be used for a large applications set without severely compromising the efficiency of handling any specific application. Hybrid systems are attractive since they use the well established PSTN for return path communications and leverage off of an established equipment base. Traffic handling capacity is limited unless more expensive equipment complements are used. The system operator also does not have total control over the system.

CONCLUSION

Early attempts to develop general purpose, multiple application interactive systems failed, not technically but on a business basis. Today's interactive systems are tailored to specific applications and, while they are not as technically sophisticated as the early attempts, they are succeeding on a business basis. The beachhead established by pay-per-view and home shopping using polled control systems can be grown to a broader applications base using more sophisticated control techniques. Catalog shopping using full resolution video frame stores, games, stock quotes and broader pay-per-view applications appear to be the next generation of applications. As the applications set grows, distributed control systems will become more attractive - despite their added complexity. The dreams of the late seventies and early eighties are now becoming realities. The question is not whether CATV systems will support interactive services in the future. The question is whether we can evolve the current, narrowly focussed interactive services to a broader range of offerings in a cost effective and profitable manner.