Active Elements for High-Definition Video Distribution

Petr Holub*, Eva Hladká†, Jiří Denemark†, and Tomáš Rebok†
*Institute of Computer Science and †Faculty of Informatics
Masaryk University in Brno
Botanická 68a, 602 00 Brno, Czech Rep.
Email: hopet@ics.muni.cz, eva@fi.muni.cz, jirka@ics.muni.cz, xrebok@fi.muni.cz

***Abstract*— Active Elements (AEs), an extension of user-empowered programmable (active) routers, provide support for multimedia distribution in collaborative environments. They can be organized in distributed systems to mitigate the scalability problem for very bandwidth and/or processing power demanding distribution patterns. We present a prototype implementation of Active Elements based on JXTA peer-to-peer framework. As shown in this paper, the AEs are flexible enough to be used for distributing high-definition uncompressed video, while also providing additional processing possibilities when distributing lower bandwidth streams like compressed HDV video streams.**

I. Introduction

The general problem of synchronous (on-line or interactive) processing lies in providing environment with as low latency of processing and data distribution as possible. We are building user-empowered network support for synchronous multi-point data distribution, that is both scalable to support a large number of clients and also robust with respect to outage of network links and other elements inside the network. This research extends our development of active router and reflectors, providing a general Active Element concept based on similar ideas and principles.

Distribution of multimedia data over IP network leads to a multicast schema. However, as the native multicast solution is not always reliable or even available, other distribution schemes were developed following approach of multicast virtualization. They are usually based on a central distribution unit—a reflector, like the H.323 Multi-point Control Unit (MCU)—that may be organized into a cascade (network) when a higher number of clients needs to be supported. The cascade of H.323 MCUs is usually statically configured and does not offer a user-empowered approach. Another well known example of multicast-like schema is the distribution used in the Virtual Room Videoconferencing System (VRVS) [1]. This is provided as a service and user data traffic is managed by VRVS administrators. The successor of VRVS called Enabling Virtual Organizations (EVO) [2]—is based on self-organization of system of reflectors, again not empowering the end-user with tools to change the distribution topology. There are also other simpler UDP packet reflectors available like rcbridge [3], [4], reflector [5], and Alkit Reflex [6].

The paper is organized as follows: Section II briefs Active Elements (AEs) approach, Section III summarizes our collaborative tool based on high-definition (HD) video, Section IV describes results of using AEs for uncompressed HD video distribution while Section V deals with distribution and processing of compressed HDV streams, and Section VI contains concluding remarks and future work outline.

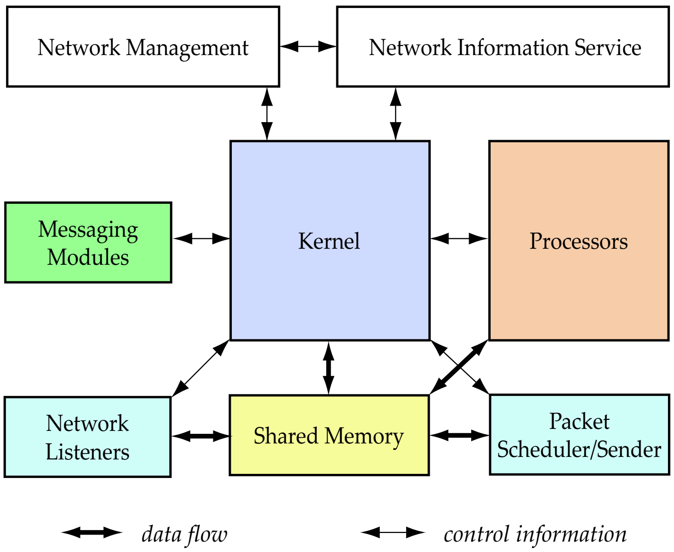
II. Active Elements

Real-time virtual collaboration needs a synchronous multimedia distribution network that operates at high capacity and low latency. Such a network can be composed of interconnected service elements—so called *Active Elements* (AEs) [7]. They are a generalization of the user-empowered programmable reflector that is a programmable network element replicating and optionally processing incoming data usually in the form of UDP datagrams, using unicast communication only. If the data is sent to all the listening clients, the number of data copies is equal to the number of the clients, and the limiting outbound traffic grows with $n(n-1)$, where n is the number of sending clients. The reflector runs entirely in user-space and thus it works without need for administrative privileges on the host computer, which can be understood as implementation of user-empowered principle.

The AEs add networking capability, i. e. inter-element communication, and also capability to distribute its modules over a tightly coupled cluster. Only the networking capability is important for scalable environments discussed in this paper. The network management is implemented via two modules dynamically linked to the AE in run-time: Network Management (NM) and Network Information Service (NIS) as shown in Figure 1. The NM takes care of building and managing the network of AEs, joining new content groups and leaving old ones, and reorganizing the network in case of link failure. NIS gathers and publishes information about the specific AE (e. g. available network and processing capacity), about the network of AEs, about properties important for synchronous multimedia distribution (e. g. pairwise one-way delay if available or RTT otherwise, estimated link capacity), and also information on content and available formats distributed by the network.

For the out-of-band control messages, the AE network uses self-organizing principles already successfully implemented in common peer to peer (P2P) network frameworks, namely for AE discovery, available services and content discovery, topology maintenance, and also for control channel management. The P2P approach satisfies requirements on both robustness

Fig. 1. Architecture of Active Element with Network Management and Network Information Service modules.



and user-empowered approach and its lower efficiency has no significant negative impact as it routes administrative data only. In prototype implementation, this has been implemented using JXTA P2P substrate [8].

A. Re-balancing and Fail-Over Operations

The topology and use pattern of any network changes rather frequently, and these changes must be reflected in the overlay network, too. We consider two basic scenarios: (1) re-balancing is scheduled due to either use pattern change or introduction of new links and/or nodes, i. e. there is no link or AE failure, and (2) a reaction to a sudden failure. In the first scenario, the infrastructure re-balances to a new topology and then switches to sending data over it. On the contrary, a sudden failure in the second scenario is likely to result in packet loss (for unreliable transmission like UDP) or delay (for reliable protocols like TCP), unless the network distribution model has some permanent redundancy built in. The probability of failure of a particular link or AE is rather small, despite high frequency of failures in global view of large networks. Thus the two fold redundancy might be sufficient for majority of applications, and the redundancy may be increased for the most demanding applications.

B. Data Distribution in Network of AEs

Separation of control plane from data distribution plane allows for modular implementation of distribution models with different properties. We have studied a number of different data distribution models for the AE network [7], which feature different performance to robustness properties:

- *2D full mesh*—the simplest model which features very high robustness so that AE outage only influences the clients that are directly connected; it also minimizes number of hops inside the overlay network,
- *3D layered mesh*—this model improves performance over the 2D model while retaining the recovery behavior and minimization of number of hops inside the overlay network,
- *3D layered mesh with intermediate AEs*—additional improvement over the 3D layered model, which can be seen as a transition to spanning trees,
- *redundant (minimum) spanning trees*—the model which allows maximum flexibility, efficient recovery from the network outages, allows optimizing data distribution with respect to saturation of lines; extension to multiple pre-computed redundant spanning trees brings about capability of very fast recovery.

III. High-Definition Interactive Collaborative Environments

A. Uncompressed High-Definition Video Transport

Enabled by current high-speed networks, high-definition (HD) video transmissions have become an essential tool for many applications. Providing truly interactive collaborative environment is still very challenging, because it requires severe limitations on processing to achieve acceptable level of interactivity (ideally less than 100 ms) and thus use of uncompressed video is the most convenient. That however imposes high demands on underlying networking infrastructure especially for multi-point data distribution, as each video stream has 1.5 Gbps. During 2005, we have developed and successfully demonstrated a prototype of low latency multi-site collaborative environment based on uncompressed HD-SDI video according to the SMPTE 292M standard [9].

The whole system comprises two basic parts: client applications and network distribution and processing. We have developed the client tools based on DVS Centaurus[1], Chelsio 10 GE cards and UltraGrid software by Colin Perkins and Ladan Gharai [10], extending it with full 1080i HD support (1920×1080 resolution; previously only lower 720p resolution was supported) and full software display [11] including field de-interlace algorithm and color space down-sampling.

The network distribution requires some service for multi-point distribution in order to supply data from each participant to all the other participants. To remove dependency on native multicast, we are relying on virtual multicast implemented by the AEs. We have demonstrated it's usability over 10 GE networks where each AE was able to duplicate each 1.5 Gbps stream on common dual-AMD64 PC from one partner to two other partners.

B. HDV Compressed High-Definition Video Transport

HDV [12] is a proprietary MPEG-2 based compression scheme developed by SONY and JVC that is transmitted in the MPEG-TS envelope over the IEEE-1394 interface in a similar way to the DV format for standard definition resolution. The resolution of the video is 1440×1080, with 50 or 60

[1] http://www.dvs.de/english/products/oem/centaurus.html

interlaced fields per second, 8-bit color space, 4:2:0 color space sampling, and 60:1 inter-frame compression resulting in approximately 25 Mbps video stream. We have implemented a tool [13] for FreeBSD 5/6 to read HDV data sent over the IEEE-1394 interface. The data are then rendered using the VideoLAN Client (VLC) tool [14] and sent over the network either using VLC or some other tools like `netcat` [15]. Because of the MPEG-2 format the HDV transmission suffers from higher latency; end-to-end latency measured using laboratory setup described above is as high as 1.9 s, leaving this transport mostly for unidirectional applications like HD streaming only.

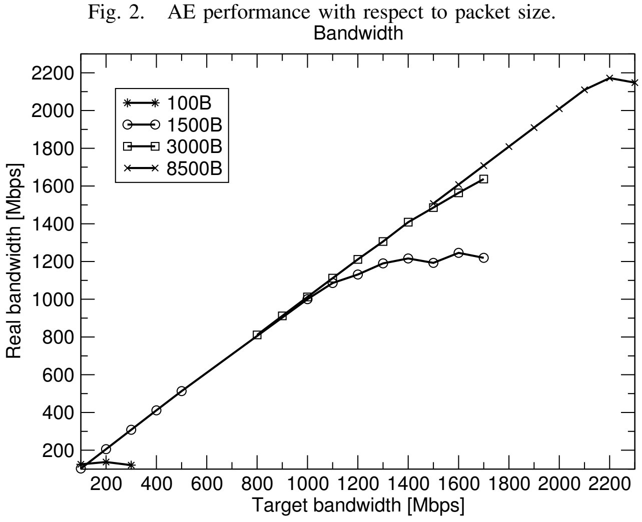
IV. AEs for Uncompressed HD Distribution

A. Performance of AEs

We have made several tests to measure the performance of the AE with respect to size of UDP packets used for video transmission. As a testbed we have used two back-to-back connected dual AMD64 computers with 10 GE Chelsio T110 LR NIC cards in 133 MHz PCI-X slots. The following three software components were running on two computers:

- computer A: UDP packet generator providing a UDP stream to computer B at a given bit rate consisting of packets of a desired size,
- computer B: AE receiving the UDP stream from the computer A and replicating it into the two equivalent output streams which are sent back to the computer A,
- computer A: two packet analyzers running in parallel capable of detecting lost and out-of-order packets and average bit rate received from AE on the computer B; each analyzer examines one stream.

Fig. 2. AE performance with respect to packet size.



Performance measurement results of the AE are shown in the Figure 2. Exact numbers describing some key points from the graph are summarized in Table I: (a) gives maximum bandwidth of a single stream with packet loss <0.01% with respect to packet size in use, while (b) gives packet loss with respect to bandwidth of a single stream given 8,500 B packet size, which we have chosen as a standard for our uncompressed HD application. The performance evaluation confirms the necessity to use Jumbo Ethernet frames for the AE to be able to replicate uncompressed HD-SDI video streams at 1.5 Gbps. Setting appropriate MTU on all the hosts and all over the path is the only part that requires administrative privileges, thus violating the user-empowered paradigm.

TABLE I
PERFORMANCE OF THE ACTIVE ELEMENT.

(a)

packet size [B]	*max. bandwidth* [Mbps]
100	100
500	300
1500	400
3000	800
6000	1700
9000	2000

(b)

bandwidth [Mbps]	*packet loss* [%]	*CPU load* [%]
1800	0.0	52
1900	0.0	55
2000	0.0099	60
2100	0.037	76
2200	1.74	80
2300	7.07	84

B. Real-World Performance

The HD transport together with AE for data distribution has been demonstrated during iGrid 2005 workshop [11] (CZ101 and US127 demos) and at SC 2005 conference. Because no more than 3 sites were participating, there was no real need to build the network of AEs. However, in order to demonstrate feasibility of this approach, we have also built a 3D layered mesh of AEs with one intermediate AE where AE cascading was used for multiplicating one stream to one site in multiple copies, so that one site saw one stream multiple times. Because a single AE running on a high-end machine is unable to create more than two copies of the data at 1.5 Gbps because of traffic limitations on given hardware, a AE network is required for supporting more than 3 sites, which is not very efficient.

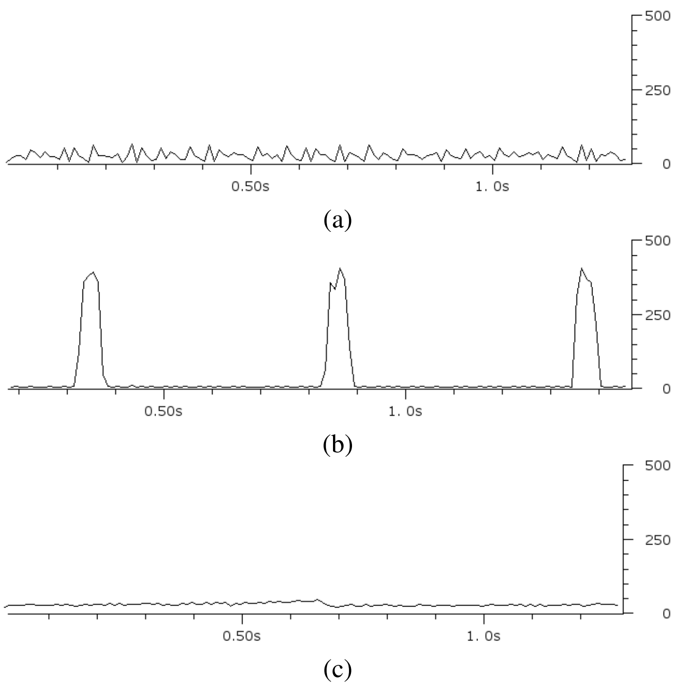
We also had to solve a problem with bursty HD traffic, which lead to loss of many packets due to the implicit thread switching. The original AE implementation used two different threads—one for the network listener and the second for sender. With implicit thread switching, the listener thread could have been stopped and the sender thread activated regardless of whether more data were coming. The current implementation gives explicit precedence to the listener which yields its precedence to the sender only when no more data are to be read. The only negative outcome we have observed with this approach is higher CPU load on the AE.

V. AEs for HDV Distribution

The high processing capacity of the AE can be also used for other purposes, e. g. for traffic shaping in case of very bursty data distribution with otherwise only moderate bandwidth requirements. We demonstrate this potential on the distribution of bursty HD traffic in the HDV format.

The HDV stream is highly compressed, requiring only 25 Mbps bandwidth. However, the HDV data generated by commonly used tools are very bursty, with peeks almost 20 times above the theoretical bandwidth requirements. Burstiness is introduced by the sending application—either VLC (Figure 3a) or even worse using `netcat` (Figure 3b).

Fig. 3. Bursty traffic produced by HDV transport with VLC (a) and `netcat` (b). Traffic after smoothing by AE is shown in (c). Horizontal axes show time and vertical axes show number of packets received at each time instant.



We have enabled a traffic smoothing module in the AE for smoothing based on sliding average of the incoming bandwidth. The resulting data flow is very smooth with no bursts at all as shown in Figure 3c. The penalty for this is small increase of transmission latency (depending on burst characteristics and smoothing interval, ranging from tens to hundreds milliseconds) because of packet buffering. The observable advantage of having smooth data flows is the receiving VLC being capable of rendering the stream with close to zero image defects, which is impossible for bursty streams above 10 Mbps.

VI. Conclusion

In this paper the Active Element and networks of AEs have been introduced to provide a flexible multi-point data distribution environment. We have discussed the advantages of this solution and we have described the processing power, robustness and fail-over capabilities.

Further, the AEs were used to support a simple collaborative environment based on the HD video transmission. The flexibility of our approach has been demonstrated on using the same elements (only with different configuration) for the distribution of uncompressed HD video streams with 1.5 Gbps stressing the data duplication capacity of the AEs, and for the distribution of bursty HDV traffic, where the smoothing extreme traffic bursts again shown processing capacity of the AEs—in this case as traffic shapers. Though not very efficient, we also demonstrated that AE network can support higher number of clients even with the uncompressed HD streams.

The future work will be focused on supporting real multipoint distribution of high speed multimedia streams. This is required when a larger number of clients must be connected, either because of higher number of participating sites or because each site sends multiple HD video streams (e. g. visualization and videoconferencing, stereoscopic streams etc.). We will also continue developing the concept of distributed AE to provide scalable multi-point replication of high demanding data.

Acknowledgment

This research is supported by a research intent "Optical Network of National Research and Its New Applications" (MŠM 6383917201). We would also like to thank to Michal Procházka, Miloš Liška and Lukáš Hejtmánek for their help with implementation of HD transport tools.

References

[1] "Virtual Room Videoconferencing System (VRVS)," http://www.vrvs.org/.
[2] P. Galvez, "From VRVS to EVO (Enabling Virtual Organizations)," in *TERENA Networking Conference 2006*, Catania, Italy, May 2006, *Accepted.*
[3] M. Buchhorn, "Designing a multi-channel-video campus delivery and archive service," in *The 7th Annual SURA/ViDe Conference*, Atlanta, GA, USA, Mar. 2005.
[4] "rcbridge," http://if.anu.edu.au/SW/rcbridge.html.
[5] O. Hodson, "UDP packet reflector/forwarder," http://www.cs.ucl.ac.uk/staff/s.bhatti/teaching/z02/reflector.html.
[6] "Alkit reflex," http://w2.alkit.se/reflex/.
[7] P. Holub, E. Hladká, and L. Matyska, "Scalability and robustness of virtual multicast for synchronous multimedia distribution," in *Networking - ICN 2005: 4th International Conference on Networking, Reunion Island, France, April 17-21, 2005, Proceedings, Part II*, ser. Lecture Notes in Computer Science, vol. 3421/2005. La Réunion, France: Springer-Verlag Heidelberg, Apr. 2005, pp. 876–883.
[8] P. Holub and E. Hladká, "Ubiquitous user-empowered networks of active elements," in *TERENA Networking Conference 2005*, Poznań, Poland, June 2005, p. 3.
[9] Society of Motion Picture and Television Engineers, "Bit-serial digital interface for high-definition television systems," SMPTE 292M-1998.
[10] C. Perkins, L. Gharai, T. Lehman, and A. Mankin, "Experiments with delivery of HDTV over IP networks," in *12th International Packet Video Workshop*, Pittsburgh, PA, USA, Apr. 2002.
[11] P. Holub, L. Matyska, M. Liška, L. Hejtmánek, J. Denemark, T. Rebok, A. Hutanu, R. Paruchuri, J. Radil, and E. Hladká, "High-definition multimedia for multiparty low-latency interactive communication," *Future Generation Computer Systems*, 2006, *Accepted.*
[12] "HDV information," http://www.hdv-info.org/.
[13] P. Holub, "HDV capture for FreeBSD 5/6 operating system," http://sitola.fi.muni.cz/~hopet/HDV/ and http://docs.freebsd.org/cgi/getmsg.cgi?fetch=37672+0+archive/2005/freebsd-firewire/20050206.freebsd-firewire.
[14] "VideoLAN Client (VLC)," http://www.videolan.org/.
[15] "netcat," http://netcat.sourceforge.net/.