Content Networking: Research Challenges of Future Content Distribution

Andreas Mauthe
Computing Department, Lancaster University, UK
andreas@comp.lancs.ac.uk

Thomas Plagemann
University of Oslo, Department of Informatics, Norway
plageman@ifi.uio.no

ABSTRACT

Content distribution is one of the large research areas in multimedia communications. In the past research efforts have mainly concentrated on communications aspects. This is now complemented by efforts to better exploit content characteristics to deal with the multitude of requirements placed onto future content networks. This paper gives an overview of the developments in content networking to this date and outlines the research challenges networked audiovisual systems face in future. The particular focus is on content networking and delivery systems. The results presented in this paper are based on the discussions within the European Network of Excellence E-Next working group on Content Networking.

Key words: content distribution networks, content networking, multimedia communication, content management

I. INTRODUCTION

With the advent of the Web and the possibility to transmit multimedia content over the Internet content distribution has become a major issue. *Content Distribution Networks* (CDN) are being developed to efficiently deliver content to end-users. A number of CDNs already exist as part of different infrastructures (e.g. Akamai in the Internet but also traditional TV broadcast networks). Still, with constantly evolving networks, applications and user requirements the developments in this area are very much on-going work.

Over the years CDNs have become more and more sophisticated. First generation CDNs have mainly focused on Web documents using caching and replication techniques (predominantly for the Web) to provide a better service [9, 15, 22]. The second generation of CDNs has been mainly concerned with continuous media issues such as audio and video streaming, video on demand (VoD) schemes, scalable video support, etc. [14]. It is expected that the next generation of CDNs will focus on additional functionalities, e.g. to facilitate the creation, modification, active placement and management of content within the content infrastructure. This ex-

tends the functionality beyond traditional delivery and is therefore called content networking (CN) [14]. Content networking is not only concerned with distribution aspects but also with content management, retrieval, content creation and adaptation support.

The challenges arising from this have been addressed by Working Group 3, Content Networking, of the European Network of Excellence (NoE) E-Next[1]. Within this working group a research roadmap has been identified comprising the future research topics in content networking. The result of this work is reported here.

In this paper the current developments in the area of CDN are outlined in section 2. In section 3 research challenges in the content networking domain are introduced. The research roadmap in section 4 shows the envisaged further developments in this area. Finally, the paper is concluded in Section 5 with a short summary.

II. STATE-OF-THE-ART IN CONTENT NETWORKING

CDNs have been discussed for some time and design aspects have been researched to a certain extent [21]. However, the definition of CDN is still not very clear. In [14] a CDN is defined as a networked infrastructure that supports the distribution of content while optimising the delivery by exploiting the specific properties and characteristics of content. Content in this context has two components, viz. the encoded data (e.g. video, audio, images, documents, Webpages, etc.) and metadata (used for identifying, describing and locating content). Thus, a CDN should not regard content files as transparent data but make use of the available content related information in order to improve the delivery and management within the infrastructure. The developments in the area of CDN have been mainly driven by the communication community. However, another important area in this context is content management and media encoding.

Within the communication domain initially three major aspects have been researched, i.e. the use of multimedia servers to support the timely delivery of continuous media (cf. [10, 13]), QoS aspects, and network caching. The work on QoS issues has resulted in two alternative proposals. The Integrated Services approach (IntServ) together with RSVP uses reservation and admission control on top of IP to guarantee QoS [7, 8]. The idea underlying Differentiated Service (DiffServ) takes a more lightweight approach utilising different service classes and trying to exploit the predictable behaviour of Internet users [3]. Caching has been successfully introduced in the Internet for HTML documents in the late 1990. These are usually conventional text and image based documents. In order to satisfy the increasing use of multimedia content on the Web, new forms of caches are required. Research has developed schemes to also support continuous media. For instance SOCCER [6] uses stream segmentation, dynamic caching combined with self-organising, cooperative caching techniques. Middleman also uses a co-operative technique by employing proxy servers as aggregates that cache video files within a LAN [1].

Probably the currently most popular CDNs are Peer-to-Peer (P2P) systems based on application level overlays [19]. They are commonly used to exchange content by file transfer. P2P systems can be roughly divided into unstructured systems (e.g. Freenet, Gnutella [4]), and structured systems, usually using a Distributed Hash Table (DHT) approach. The former includes centralised schemes (such as the original Napster) and hybrid approaches (e.g. JXTA). Examples for the latter are for instance Chord [20], CAN [16] or Pastry [18]. Issues related to P2P systems are for instance a general lack of streaming support (addressed by [5, 17, 23]), restricted metadata and search capabilities, and missing (pro-)active replication [14]. These are all active research areas.

Another area of interest in the context of content networking is content management systems (CMS). These are distributed systems that manage all kinds of different content types [11]. Within an organisation they support the entire content lifecycle from the original idea to the finished product (including archiving). At present CMS are mainly used within a specific application domain. However, it is envisaged that content management capabilities are becoming an integral part of the content networking infrastructure. This should

include better support for interactive content use, more appropriate content search and location capabilities, and a more comprehensive support of the entire content lifecycle reaching from production to delivery.

Standardisation within the CDN domain has so far been rather fragmented (mostly related to specific topics, e.g. QoS provision). Apart from MPEG-21[2] there are currently no major standardisation activities addressing the problem area comprehensively. Though, MPEG-21 has a much broader focus and is concerned with the entire, fully electronic workflow of digital multimedia content, including creation, delivery and trading. One of the seven key elements, viz. Terminal and Networks, is addressing issues related the functional interoperability between heterogeneous networks and devices, which is related to content networking.

III. FUTURE RESEARCH CHALLENGES

A study of ongoing research shows that it is still very much driven by communication issues. However, the research in the area of content distribution has also led to the identification of a number of challenges that need to be addressed.

3.1 Ongoing research

A major issue is the architecture of content networks (CN). There are various possibilities to design such systems. However, they can be all assessed according to the kind of operations they support and the way content is handled. In order to support the different needs of applications it is necessary to apply the classical principle of systems design to CNs, i.e., the separation of mechanisms and policies.

To achieve a better understanding of how the research community is approaching the problem and to advance the state-of-the-art a questionnaire has been distributed in 2004 to all 42 E-Next member organisations [14]. Within the research projects that have been identified there is a clear consensus about the underlying networking infrastructure that should be chiefly considered in CN research: all assume an IPv4 based network and some do not exclude a later transition to IPv6, all consider best-effort networks and some combine them with DiffServ. Further, most research projects target large scale systems.

The applications that are supposed to use the CNs that are developed in different projects comprise classical video-on-demand as well as news-on-demand with high user interactivity, distributed content management systems for professional use, medical digital video libraries, media-on-demand for education, and 3D virtual worlds for a large number of users that interact through multicast streams. The tools that are within the research projects comprise various programming and simulation tools, of which ns-2 is clearly the favourite of most researchers. However, many researchers have realized that simulation is just an early step in the development and PlanetLab [12] as a test-bed is regarded also as very important.

3.2 Challenges

The ongoing discussion within E-Next shows that there are many challenges that need to be addressed in future CN research [14]. In this section, the once are discussed that are deemed to reflect the major issues, i.e.:

1. How to handle unexpected resource demand and network conditions
2. How to design and develop a research CN
3. The number and kind of CDNs needed to fulfil the requirements of different applications
4. How to take advantage of application and content semantics
5. How to deploy future large scale services
6. How to deal with heterogeneity at the device, application and user level

3.2.1 Measurement effort

CNs rely on CDN services to distribute and manage content within the network. Fundamentally, the emergence of CNs does not necessarily require any new CDN mechanisms apart from the necessary customization required for each application so that CDN elements (e.g. surrogate servers, proxies, etc.) become aware of the content and application semantic. A challenge lies in the ability of the CDN elements to react fast enough to changes in network conditions to avoid any service

disruption. The more interactive the service is, the more stressed the CDN elements are. In the case of CNs the objective is not primarily to offer a better service than the legacy best effort service but to offer a consistent service throughout users sessions' lifetime. With respect to this constraint one of the biggest challenges for CNs is the ability of CNs to continuously monitor network conditions. Network monitoring is also important in the context of wireless and ad-hoc networks because it converges to a whole IP solution. Wireless users might get high download throughput (e.g., in 3G networks) but may suffer highly fluctuating network conditions due to mobility or fading effects.

3.2.2 Towards a research CN

The need for a CN research platform stems mainly from two aspects of CN research. On the one hand there is the need to collect usage statistics to inform system design and engineering decisions. Here the challenge is to "look into the future" by deploying new services to a user population and on a platform reflecting a real world deployment, rather than restrict testing to the labs. To be effective such a CN should not only be very stable but also operate within a realistic scenario.

On the other hand, researchers will also want to be able to quickly deploy and test new services and mechanisms as part of a normal system research cycle. Furthermore, because of the performance issues encountered in CN research, researchers should be given the opportunity to experiment with low-level operating systems mechanisms (e.g. disk scheduling, network protocols, etc.).

3.2.3 How many CDNs are required?

If one considers CDNs as a set of functionalities and active devices within the network and CNs as an upper layer that aims at transforming application and users needs and information into metadata, then two natural questions arise from this global picture:

1. To which extent can the application and user requirements be transformed into metadata that allow CDNs to operate without any further control of the application?

2. How many CDN/CN architectures do we need?

Concerning the latter question, the two answers at the extreme ends of the spectrum are *one CDN architecture per application versus one CDN architecture shared by all applications*. As usual in such cases, it can be expected that the correct answer lies in between the extremes. Accordingly, it is expected that a few CDN architectures will emerge. While the number of distinct CDN architectures is an important research issue, from an operational point of view the number of operated CDNs is more relevant. It is predicted that this will be large. Issues such as scalability will have to be addressed in this context.

3.2.4 Exploiting application and cntent semantics

Current CDN approaches are either implicitly exploiting application and content semantics (e.g., caching strategies for Web content), use a rudimentary sub-set only (e.g., P2P file sharing applications using file names), or largely ignore it. However, the goal of a CDN is to appropriately exploit them to optimize the delivery of content within a CN. The challenge in this context is how to find the right level of abstraction and balance of application knowledge within the communication sub-system. Metadata plays a crucial role in this context. The knowledge represented by the metadata is also highly relevant to feed self-organisation and guide the structuring of CN.

3.2.5 Large scale issue

One of the main issues will be to deliver high quality content in a scalable manner since scalability is necessary to maintain low operational costs. With respect to the scale of the problem, the following questions need to be addressed together:

- How to handle heterogeneity of receivers in terms of network fan-in, fan-out, and resources available at the terminal (e.g., CPU, screen size).
- How to handle scalability for broadcast events to a large audience using scalable congestion control.
- How to support low-latency streaming of live multimedia flows (e.g., minimize the number of hops, select not overloaded peers, etc.).

3.2.6 Heterogeneity

Heterogeneity is an issue at all levels of the CN architecture. At the user level the question is what kind of content does the user want to be presented

with in which way. This does not only reflect on formats but also on the presentation. The different applications using content also provide various functionalities ranging from simple display to sophisticated interactive capabilities. At the device level the range is from high capability high resolution displays to small handheld wireless devices. Two of the major questions in this context are how to accommodate these requirements technically and how the content itself can be transformed to satisfy them. An interdisciplinary approach has to be taken in this context.

IV. ROAD MAP TO FUTURE CONTENT INFRASTRUCTURES

In the previous section the most immanent and urgent research issues in the area of content networking have been presented. Obviously, there are many more research challenges in this area. The aim of this section is to give a more complete and also more long-term view alongside with an overview on the research challenges that have to be addressed within the coming years. Figure 1 presents the challenges in an application context on a time scale from 1995 to 2020.

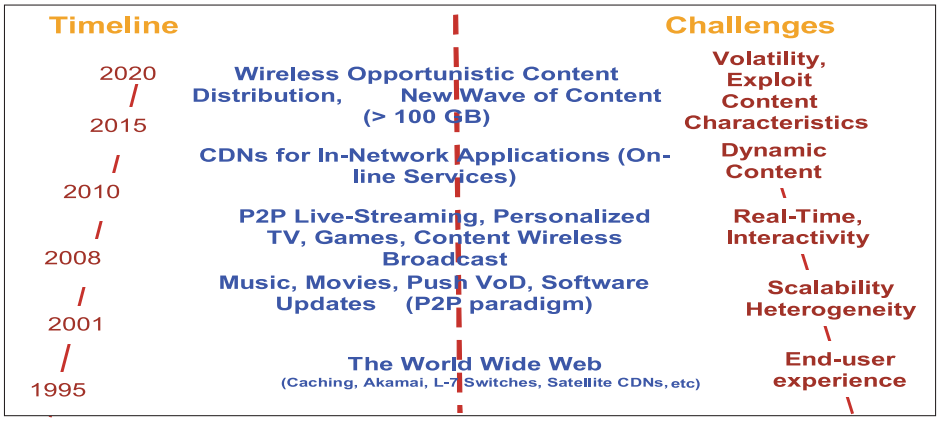

Fig.1 Applications and CDN research challenges

P2P live streaming, personalized TV, Games and wireless broadcasting have been identified as important applications in the near future which require (due to their interactivity) sufficient real time support. The next big challenge after this is the handling of dynamic content e.g., for "in-network applications" such as online-services. In the long-term, volatility also has to be addressed. Further the exploitation of content characteristics within the content network to enable the new wave of content (e.g., wireless opportunistic content distribution) will be a major issue.

This application centric view is related in Figure 2 to the research issues that have been discussed in the previous section (i.e. they are considered from a system view). This figure also takes a long-term perspective and discusses the challenges in the context of various potential system and technical developments over the coming years.

The three major threads coming together in CN research are communication systems, content management, and the exploitation of content characteristics and media semantics for a more effective content handling. Within all these areas specific research topics have been identified that will contribute to the further development of CN and content infrastructures. Today's research in CDN related communications research is, for instance, dealing with streaming, network caching, QoS and P2P issues.

Content management research is currently looking at highly distributed infrastructures (possibly using service based concepts). This is mainly related to the fact that production becomes more fragmented and distributed. Also, this approach reflects the organisational structure better. Last but not least, content will in future not only be produced by one organisation to be made available at a single hand-off point. There will many different contributors to the content creation process. This is already acknowledged by the new user paradigm of prosumer, i.e., a consumer who also produces content and vice versa.

The third research thread is currently mainly looking at Ontologies and content description and issues related to content analysis.

Ultimately we see the different areas converging via two stages. It is envisaged that by 2010 1st Generation Content Networks are in operation that integrate CDN and content management functionality, i.e., the networked infrastructure will not only be concerned with the efficient delivery of content but will also deal efficiently with load balancing, autonomously manage its placement within the system. Furthermore, it will also allow more appropriate search and take into account context by using content adaptation.

At the same time there will be more integrate media creation, production and delivery using distributed content management environments. This infrastructure will on the one hand use sophisticated automatic content analysis and description schemes. On the other hand it will be far more distributed than today's systems. A service oriented approach seems to provide the most suitable abstraction in this context. Here the different functionalities are represented by independent services with open interfaces. The necessary workflow support will also be provided through a special service in this context.

Through further research in areas such as proactive content placement, Data Stream Management Systems (DSMS), interactive CN, context aware content delivery, content adaptation, automatic content classification we envisage the systems to converge to a content infrastructure that comprises support for content creation, handling and delivery. This infrastructure will optimally support the predicted more interactive types of content usages as well as the prosumer paradigm. It will also provide a better platform for professional content creation since it allows instant access to content production and distribution facilities from almost anywhere: Thus, it will enable faster production and distribution.

The in section 3.2 identified challenges are guiding the research carried out along the roadmap. The non-functional characteristics (such

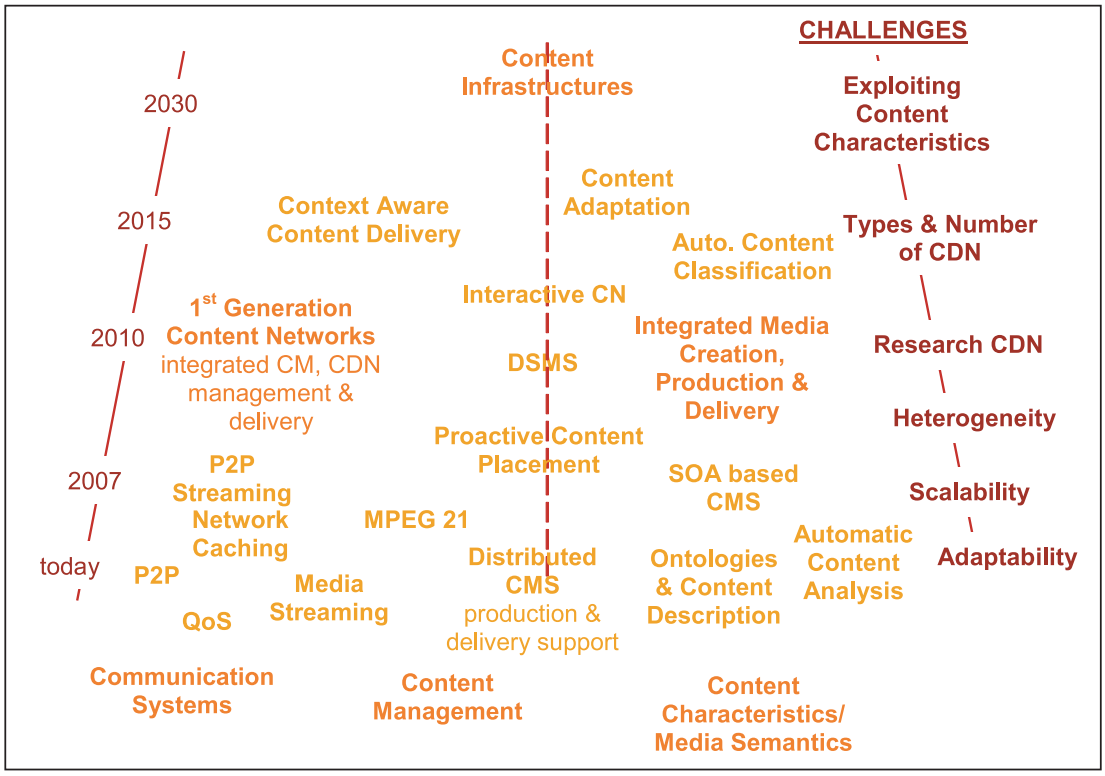

Fig.2 Systems and challenges

as support for heterogeneity, better scalability and adaptability) have to be considered in the different sub-areas.

Many of the individual issues are already being tackled within specific contexts. A "meta-challenge" therefore will be to come up with a framework in which the different approaches can be integrated. This framework needs to be flexible enough to incorporate the results of the different research strands. At the same time it needs to facilitate the development towards one common infrastructure. This is the idea behind the content infrastructures concept. The introduced roadmap is providing a way to achieve this aim.

V. CONCLUSIONS

Content delivery and content networking is one of the large research areas in which multiple domains and aspects are coming together. This paper gives and overview of the developments so far and outlines the research challenges networked audiovisual systems face in future. The particular focus is on content networking and delivery systems. The contribution reflects the discussion of the NoE E-Next working group on Content Networking.

The introduced roadmap indicates the expected developments over the coming years (until 2030). It is envisaged that content infrastructures will by then emerge as systems that allow ubiquitous content access and support for content creation, handling and delivery.

ACKNOWLEDGEMENTS

This paper reports the results of working group 3, Content Networking of the European NoE E-Next (http://www.ist-e-next.net/). We gratefully acknowledge the contribution of all the partners who have actively participated in the discussions.

Further, we would also like to acknowledge the sponsoring of the project through the European Commission (E-Next is an EU IST project within Framework VI, IST-2001-34056, FP6-506869)

REFREENCES

[1]. S. Acharya, B. Smith, "*MiddleMan: a video caching proxy server*", Proceedings of Network and Operating Systems Support for Digital Audio and Video (NOSSDAV'), 2000.

[2]. Coding of Moving Pictures and Audio, N4801, "*MPEG-21 Overview*" version 4, ISO/IEC JTC1/SC29/WG11, Fairfax, USA, 2002.

[3]. M. Fine, K. McCloghrie, "*Differentiated Services Quality of Service Policy Information Base*", The Internet Engineering Task Force, 2003.

[4]. *Gnutella Web Site*, http://www.gnutella.com, 2006.

[5]. M. Hefeeda, A. Habib, B. Botev, D. Xu, B. Bhargava, "*PROMISE: Peer-to-Peer Media Streaming Using CollectCast*" Proceedings of ACM Multimedia, Berkeley, CA, 2003

[6]. M. Hofmann, E. Ng, K. Guo, S. Paul, H. Zhang, "*Caching Techniques for Streaming Multimedia over the Internet*", Bell Labs Technical Memorandum, 1999.

[7]. IETF Integrated Services Working Group, "*Integrated Services Charter*", The Internet Engineering Task Force, 2000.

[8]. IETF RSVP Working Group, "*RSVP Charter*, The Internet Engineering Task Force, 2000.

[9]. I. Lazar, W. Terrill, "*Exploring content delivery networking*", IT Professional 3 (4), 2001.

[10]. P. Lougher, D. Shepherd, "*The design of a storage server for continuous media*, The Computer Journal 36 (1), 1993.

[11]. A. Mauthe, P. Thomas, "*Professional Content Management System: Handling Digital Media Assets*", Wiley Academic Publisher, 2004.

[12]. *PlanetLab*, http://www.planet-lab.org, 2006.

[13]. T. Plagemann, V. Goebel, P. Halvorsen, O. Anshus, "*Operating System Support for Multimedia Systems*", Computer Communications Journal, Elsevier, 23 (3), 2000.

[14]. T. Plagemann, V. Goebel, A. Mauthe, L. Mathy, T. Turletti, G. Urvoy-Keller, "*From content distribution to content networks - issues and challenges*", Computer Communications, Elsevier, 29 (5), 2006.

[15]. M. Rabinovich, O. Spatscheck, "*Web Caching*

and Replication", Addison-Wesley, Reading, MA, 2002.

[16]. S. Ratnasamy, P. Francis, M. Handley, R. Karp, S. Schenker, "*A scalable content-addressable network*", Proceedings of the 2001 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, 2001.

[17]. R. Rejaie, A. Ortega, "*PALS: Peer to Peer Adaptive Layered Streaming*", Proceedings of the International Workshop on Network and Operating Systems Support for Digital Audio and Video Monterey, California, 2003.

[18]. A. Rowstron, P. Druschel, "*Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems*", IFIP/ACM International Conference on Distributed Systems Platforms (Middleware), 2001.

[19]. R. Steinmetz, K. Wehrle, "*Peer-to-Peer Systems and Application*", Springer-Verlag, LNCS3485, 2005.

[20]. I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, H. Balakrishnan, "*Chord: A scalable peer-to-peer lookup service for internet applications*", Proceedings of the 2001 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, 2001.

[21]. D. Verma, "*Content Distribution Networks*", Wiley, New York, 2002.

[22]. A Vakali, G. Pallis, "*Content delivery networks: status and trends*", Internet Computing, IEEE , 7 (6),2003.

[23]. M. Zink, A. Mauthe, "*P2P Streaming using Multiple Description Coded Video*", Proceedings of the 30th EUROMICRO 2004, Rennes, France, 2004.

BIOGRAPHIES

Andreas Mauthe *is Senior Lecturer at Lancaster University and has been working in the area of distributed and multimedia systems, content management and content distribution for more than 15 years. Prior to joining Lancaster University he was heading a research group at the Multimedia Communications Lab (KOM) of the Technische Universit? t Darmstadt working on content management systems content distribution networks, and different aspects of peer-to-peer systems. Andreas has worked in industry for four years where as Chief Development Officer (CDO) he was in charge of the asset & content management system media archive(r). He received a PhD from Lancaster University in 1998.*

Thomas Plagemann *received his Diploma in Computer Science from the University Erlangen-N ü rnberg (Germany) in 1990, and his Doctor of Technical Science from Swiss Federal Institute of Technology (ETH) Zurich (Switzerland) in 1994. In 1995, he was honoured with the Medal of the ETH Zurich for his doctoral thesis, in which he developed the Da CaPo communication subsystem. From 1994 to 1996 Thomas Plagemann was a researcher at UniK - Center of Technology at Kjeller and Telenor R&D (Norway). Since 1996, Thomas Plagemann is Professor at the University of Oslo where he heads the Distributed Multimedia Systems Group. His research interests include multimedia middleware, protocol architectures, QoS, operating system support for distributed multimedia systems, and mobile and delay tolerant networks. He has published many papers in his field and has served as programme committee member of many major international conferences and workshops. He is member of the Editorial Board of ACM Transactions of Multimedia Computing, Communications and Applications and he serves as Editor-in-Chief for the ACM/ Springer Multimedia Systems Journal.*