A protocol stack for futuristic multimedia

Leif Arne Rønningen
Norwegian University of Science and Technology, Dept. of Telematics.
O. S. Bragstads plass 2B
N-7491 Trondheim, Norway
tel. 47 9003 0473
leifarne@item.ntnu.no

***Abstract* - A new protocol called the AppTraNet protocol, combines the application layer, transport layer and network layer protocols into one protocol, with only one common header that can be processed efficiently in parallel by hardware. The protocol handles the set-up and management of collaborative distributed scenes, and content transfer. Predicted future networked multimedia collaborations require data rates of up to tenths of Gbps per scene object, and face-to-face time delays lower than 10-30 milliseconds. To guarantee the time delay, the quality of scene objects is allowed to vary with space and time. The adaptation scheme uses quality shaping, based on separate and independent coding of scene sub-objects sent in independent packets, drop of selected sub-objects, and rate-adaptive feedback control. Set-up and other control packets cannot be dropped, but these packets have no real-time requirements. A priority queuing hardware architecture is used in network nodes to handle both types of packet streams. The behavior and performance of the architecture have been verified by simulation, using Simula/DEMOS simulation tools.**

I. INTRODUCITON TO DMP

This paper is about the three-layer Distributed Multimedia Plays architecture, DMP, with focus on the new AppTraNet protocol defined within the architecture. The protocol combines the application layer, transport layer and network layer protocols into one protocol, with only one common header that can be processed efficiently in parallel by hardware/ASICs.

Examples of future services using the new architecture, are jazz sessions, song lessons, musical theatre education, rehearsals and performing arts, coming generations 'Multimedia Home Spaces', games and other entertainment, general education, and business meetings.

To get a complete understanding and motivation for DMP, important parts of the DMP architecture such as the concepts of near-natural virtual quality of scenes, SceneProfiles, traffic generated by scenes, and Quality Shaping of scenes using various traffic and resolution control schemes, should be studied. However, it is impossible to cover everything here, so section I.A is added just to point to aspects 'missing' in this paper.

A. Complementary work

By definition, the near-natural virtual scene has a quality that approaches the natural scene, that is, users should not perceive any difference when experiencing a real scene and the corresponding virtual scene. This is expected to be obtained in say 10 years from now. Recent tests [1] indicate that stereoscopic video at HDTV quality (2k x 1k pixels, 60 Hz progressive scan) has substantially lower perceived quality than the corresponding real scenes. A user is defined as a group of humans or other objects in a real scene, or a network server. The scenario, as exemplified in Fig. 1, is a futuristic, virtual scene that shall support near-natural virtual quality. This prerequisites auto-stereoscopic multi-view, surround multi-channel sound, guaranteed maximum user-to-user time-delay less than 10-30 ms, hierarchic object oriented scenes described by SceneProfiles, scene object quality that varies with time and space, graceful degradation of quality, and a defined security level [2].

Services are to be understood in a broad sense: the total service (with a well defined adaptive quality) received by users from one or more service providers. SceneProfiles give standardized descriptions of how to shoot and present standardized stereoscopic multiview adaptive scenes. Users negotiate SceneProfiles as the first step of establishing a service. For an example, see section 2.G, and for a detailed description see [2].

Fig. 1. Scenario: a 'Near-natural virtual collaboration'. The walls, the ceiling and the floor are displays.

Traffic generated from near-natural virtual scenes is extremely high, up to 10^3 – 10^4 higher than from today's videoconferencing systems. This traffic also is extremely variable during the collaboration. Examples are shown in later sections, and further elaboration can be found in references [2] and [3].

The concept of Quality Shaping was introduced to give graceful degradation of quality when traffic overloads the network or system components fail. The concept builds on controlled dropping of sub-objects from (selected packets), and scaling of scene resolution/composition and coding parameters. The scheme guarantees a maximum user-to-user delay without any reservation of resources. However, to guarantee a minimum quality level, admission control is needed. Controlled dropping of sub-objects as part of Quality Shaping will be treated in this paper. More advanced parts of the scheme are treated in [2].

The division of scenes into sub-scenes, objects and finally sub-objects, see Fig. 2, is of fundamental importance for DMP. This is the basis for making multimedia content packets independent. The paper [3] verifies the principle.

The DMP architecture synchronizes sources. Since maximum delays can be guaranteed, maximum jitter at the destination can also be guaranteed.

For control and management packets only static routing can be permitted, in order not to destroy sequences. To guarantee the maximum delay of content packets, routes of lower delay than the maximum permitted can be selected [2].

Basic design goals of DMP are to simplify and to extend the quality of existing collaborative systems. Video conferencing systems using standards such as the H.323 need a large number of different protocols to work properly. The aim here is to reduce the number of protocols to two, and correspondingly reduce the number of architectural layers to three. To handle the high data rates, data processing have to be performed by hardware (ASICs). Software solutions shall be used for functions without severe real-time requirements.

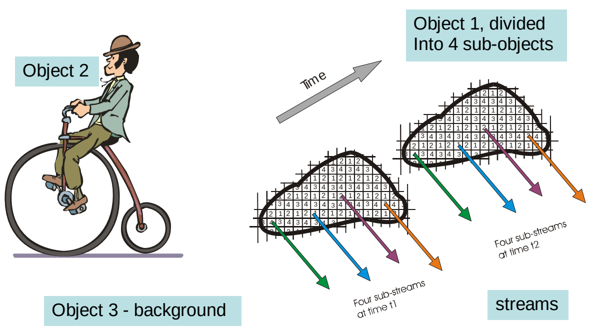

Fig. 2. Object-oriented scene with sub-objects

B. *Related work*

Quality Shaping is based on concepts such as Traffic Shaping [4], [5] and ATM ABR [6], where packet drop and feedback control were applied.

The TCP protocol (over IP) does not support the real-time transfer of audiovisual content. To support voice and video in packet networks the protocol stack RTP/UDP/IP is widely used, but does not support traffic control, and does not guarantee the Quality of Service. On the contrary, UDP traffic streams fill up a channel according to the 'best effort' principle, and suppress TCP traffic [7]. Protocols such as XCP [8], RCP [9], RED [10] (TCP extensions), DCCP (with TFRC) [11], and P-AQM [13] try to control the traffic using feedback control. In paper [13] it is shown that P-AQM outperforms the aforementioned schemes, and provides high link utilization and fairness between TCP and UDP traffic. All these proposals add complexity to the TCP/UDP protocols, and do not support adaptive control of user scenes and traffic (as with DMP).

DMP applies the protocols IPv6 [14] and IPsec [15], slightly adapted. This makes it possible to transmit AppTraNet packets through the existing Internet.

The ability of TCP to support reliable end-to-end transfer is not required by the new architecture, this is guaranteed by the AppTraNet protocol, see below. UDP does not provide functions useful to DMP, and is not applied. Since the AppTraNet protocol also supports necessary functionality for the application layer, other application protocols such as H.323 for videoconferencing (including RTP, UDP and other), or SIP with SDP and MSRP are not used in the DMP architecture.

II. THE THREE-LAYER DMP ARCHITECTURE

In this section, we give descriptions of behavior that involve user equipment and network nodes simultaneously. There are two different node types in the network, AccessNodes and (Core) NetworkNodes. In addition, a number of specialized servers are needed, for example to support collaboration establishment and management.

Packet delays through nodes and processing delays in user equipment can be guaranteed lower than a specified value. Except for output link queues in nodes, there is no buffering or waiting. Information is included in packets so that an audio or visual packet is independent of all other packets. The content of a packet is used to present parts of an object immediately, at the right place and with the right quality (variable, however), not waiting for other parts of the object or other objects. Pre-stored (negotiated) configuring data such as SceneProfiles, are used (without delay) in the rendering process. The pre-stored configuring data can be the result from negotiations when a complex multi-party scene has been set up, or some configuring action taken during the collaboration. In short, objects are automatically synchronized from the source, and are presented within a guaranteed time, with a controlled variable quality due to controlled drop of packets in the network.

A. *Hierarchy and distribution of functional sub-systems*

Fig. 3 and 4 show the overall protocol hierarchy of DMP. Three layers are defined, the Linksical layer, the

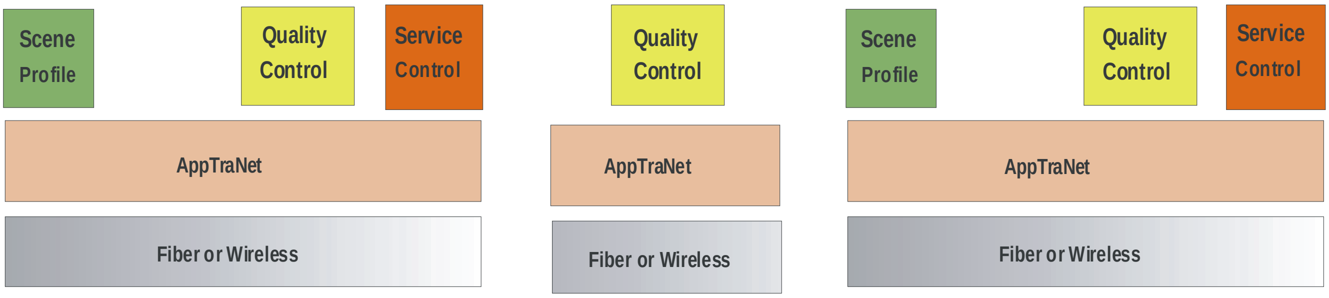

Fig. 3. Three-layer architecture, Source AccessNode, Network Node, and Destination AccessNode

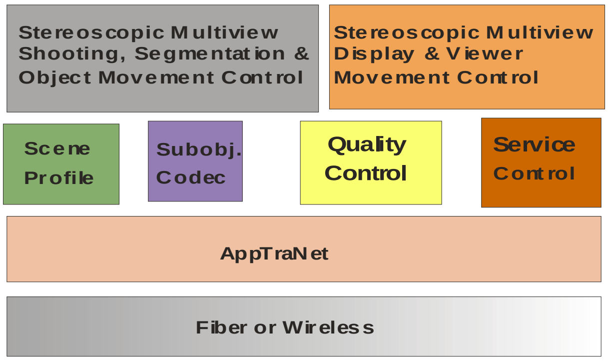

Fig 4. Three-layer architecture, user

AppTraNet layer, and the Application layer.

The AppTraNet protocol has only one combined header, the only above the Linksical layer. The protocol runs on top of the Linksical layer, a combined link/physical layer. This layer assumes optical fibers between network nodes. The protocol data frame includes the AppTraNetPacket as payload, and a preamble for recognizing frame start and bit clock synchronization [2]. Conversion between optical signals and electrical signals are made at each end of optical links.

B. *The AppTraNet layer*

The AppTraNet layer is a combined application, transport and network protocol layer. The DMP architecture uses the IPv6 protocol and IPsec security, and defines a new part supporting specific DMP requirements. To enable efficient design of high-performance hardware, ASICs, there is only one AppTraNet packet header, where parameters can be processed in parallel by physically parallel hardware, ASICs.

C. *The AppTraNet packet*

Some parameters of the combined AppTraNet header are introduced to increase the performance, to support efficient hardware design, and to reduce the complexity of the network. The parameters allocate bits and bytes in the order as shown below. The packet length is 1.5k Bytes.

The IPv6 header is used according to the standard, except for the IPv6 addresses that are used to uniquely identify users [2]. The 48 lsbs of the address uniquely identify AccessNode, and if 2^{10} users can be connected to an AccessNode, there is 38 bits left for the identity of each user individual.

Integrity, authentication and encrypted payload are provided by IPsec AH and ESP in Transport Mode. See RFC 4302 [14] and RFC 4393 for a guide into IPsec. The ISAKMP (Internet Security Association and Key Management Protocol) [15] and IKE (Internet Key Exchange) [16] are used for key exchange (other may be considered).

A drawback could be that the AccessNode and its servers have to be trusted, but in DMP, the service provider generally is trusted (responsible for the services, the quality, the servers, and the network).

The new part (in addition to the IP and IPsec parts) of the protocol header has the following parameters:

PT, Payload type, 8bits.

PacketRate, 16 bits, the current packet rate from the sub-object, indicated by the user [2]

Sequence number, 8 bits, used for controlled dropping of progressive JPEG2000 compression layers, and of sub-objects compressed by the NOC scheme [2]

Timestamp, 32 bits, used for delay measurements from a network node to an access node, and from users to AccessNodes

ServiceID 48 bits, used to identify on-going services [2]

PixelAdrB, 16 bits, addresses the start pixel of the sub-object represented in this packet (rectangle)

PixelAdrE, 16 bits, addresses the end pixel of the sub-object represented in this packet (rectangle)

SPQSP, 32 bits, reference to SceneProfile and QalityShapingProfile, used for collaborations [2]

Reserve count, 8 bits, counts the number of Reserve bytes

Reserve bytes, 0-64 bytes, for future use

D. Packet types, priority and drop

The PT Payload type parameter reserves values for DMP use, defining the payload. Two main groups are defined, multimedia content packets and control packets. The following types are so far defined: visual, audio, graphics, quality shaping, object move, viewer move, scene profiles, addressing, acknowledgement, httpRequest, and httpResponse. Some control packets are used in typical request-response sequences, see II.G. It is up to the application how to handle the situation if packets that need acknowledgement are not acknowledged. Since such requests do not have real-time requirements (normally), and shall have an extremely low probability for being lost, the output link queuing system shown in Fig. 5 is introduced as part of the AppTraNet protocol in all nodes after switching (routing). Control packets enter queue Q2, which holds packets on a very large store. If selected by Sel, with probability p1, the control packet enters Q1 for link output. The module H is an abstraction of the output Linksical layer, sending the packets out at the maximum packet rate given by the link capacity. Control packet transport delays are highly variable, depending on traffic patterns, see section 2.5. The maximum length of Q2 should be so large that reaching the maximum should have an extremely low probability, before the admission control decreases the input traffic to the network [2].

Multimedia content packets, with a maximum guaranteed end-to-end delay, but that can be dropped selectively, are sent to buffer B2. The selector Sel fetches packets from B2 with probability 1-p1, and forwards to Q1. Q1 has a limited length and determines maximum jitter of a transfer through the node. B2 is a very short buffer used for dropping according to the sequence number of the packet. If Q1 is full, packets start queuing in B2. Assuming Near-Natural Object Coding, NOC [2], and nine sub-objects, B2 drops arriving packets as follows:

If B2.length 0-12 then join B2 else

If B2.length 13-18, drop from sub-object 8 else

If B2.length 19-24, drop from sub-object 8 and 7

If B2.length 24-29, drop from sub-object 8, 7 and 6

and so on (sub-object packets 9 are never dropped)

The Sequence number is for NOC coding directly given by the sub-object number, while for JPEG2000 coding both the sub-object number and the progressive layer number have to go into the calculation [2]. Optimizing the maximum length of B2 and Q1, the drop decision lengths of B2, and the value of p1, is part of the QualityShaping scheme design [2].

Since a path through the network includes nodes as shown in Fig. 5 in series, the maximum end-to-end time delay can be guaranteed. If the processing times in user equipment and switches (nodes) are constant, the end-to-end packet delay is a sum of propagation times in the path, the constant processing times in path nodes, and the waiting times in Q1 in path nodes (the waiting time in B2 can be neglected compared to the waiting time in Q1).

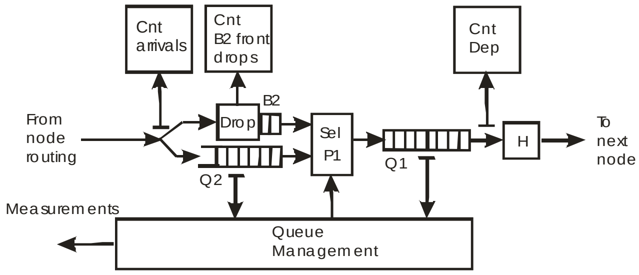

Fig. 5. Dropping and prioritizing of packets in network nodes, AppTraNet layer. Hardware architecture

E. Application layer

Fig. 4 shows six main functional blocks on the DMP application layer for the user scene. The sub-systems SceneProfile, QualityControl and ServiceControl have common functions with AccessNodes, while the three other sub-systems are implemented only in the user equipment. The network nodes handle only one application layer block, the QualityControl.

F. ServiceControl

The ServiceControl, allocated to the AccessNodes, are the ‘beacons’ of the DMP system. The ServiceControl has the top responsibility for setting up, manage, and release collaborations between users, on request from a user. The ServiceControl manages and routes the packet exchange between users and various servers allocated to AccessNodes, and between access- and destination network nodes. It cooperates via user ServiceControl with the Stereoscopic Multiview Display & Viewer Movement Control and the Stereoscopic Multiview Shooting, Segmentation & Object Movement Control in the user terminals. For this purpose, the ServiceControl has a number of supporting servers to its disposal:

- *QualityControl* – handles scene quality shaping by dropping sub-objects.
- *Security Server* – implements IPsec AH and ESP, receives authentication requests from ServiceControl, returns acceptance or not. Encrypts and decrypts payload.
- *Proxy and Address Server* – receives translation requests from the ServiceControl, returns acceptance or not. Stores info about on-going services, user IDs and address relations in the domain. This Server also handles adaptation of protocols and formats for other ITC systems, such as WWW and email.

- SceneProfile and QualityShapingProfile Server – receives SceneProfile check request from ServiceControl, returns acceptance or not. Stores standardized SceneProfiles. Stores standardized QualityShapingProfiles. Supports the ServiceControl to adapt scene quality according to traffic load, using the optimal QualityShapingProfile.

The user has a corresponding ServiceControl, cooperating closely with the ServiceControl of the AccessNode. The following modules are used:

- QualityControl – shapes scene quality by adjusting QualityShaping parameters [2].
- Security Module, implements IPsec AH and ESP, should be a detachable hardware module, e.g. with a smart-card.
- SceneProfile and QualityShapeProfile Descriptions.
- Sub-object encoder and decoder. The NOC lossless compression scheme and JPEG2000.
- A sub-system for Stereoscopic Multiview Shooting, Segmentation and Object Movement Control.
- A sub-system for Stereoscopic Multiview Display and Viewer Movement Control.

A special type of user is the servers storing objects and scenes that are shot or synthesized in advance. The fist four bullet points above are part of such a server, and the behavior is as for a normal user.

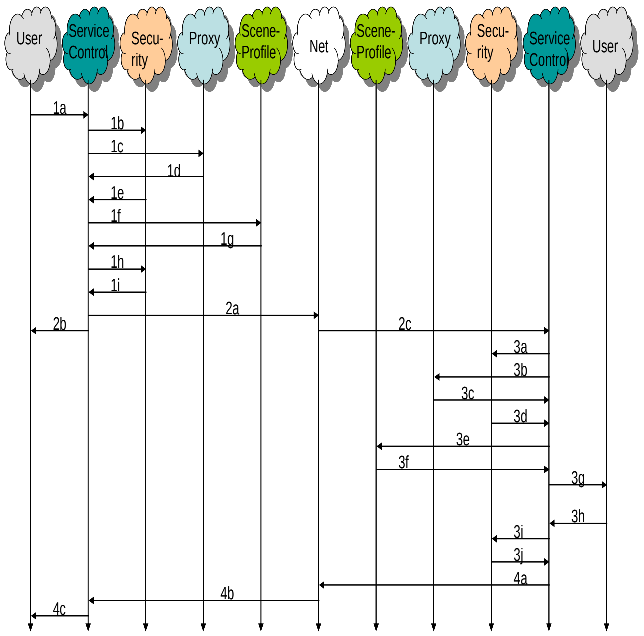

Fig. 6. Set-up of multi-party DMP

G. *Example, establishment of a multiparty DMP*

Fig. 6 shows a successful setup of a one-way collaboration from user A to B. When there are N users in a collaboration, all users have to set up a one-way collaboration to all the others, giving N(N-1) set-ups. The ServiceControls in the AccessNodes handle the setup and the management of the sessions until they are released. All servers and users are uniquely identified by their IP addresses. All possible services are described by standardized SceneProfiles. Users must choose SceneProfiles during set-up of a service, but this can be changed during the session. The SceneProfile defines how to shoot videos and record sound, and how to present on the receiving user display.

Packets *Description*

1a - AH and ESP secured service request (ServiceID, UserIDs, A, B, C, D, E and F, {set of prioritized SceneProfiles})

1b - 1a sent to Security for security check

1c - 1a sent to Proxy for address validation and translation

1d - Positive response to 1c (ACK)

1e - Positive response to 1b (ACK)

1f - 1a sent to SceneProfile for acceptance of scene

1g - positive response to 1f (ACK)

1h - Service request sent to Security for AH and ESP addition

1i - Positive response to 1h (ACK)

2a - AH and ESP secured service request

2b - Positive response to 1a (ACK)

2c - as 2a

3a - 2c sent to Security for security check

3b - 2c sent to Proxy for address validation

3c - Positive response to 3b (ACK)

3d - Positive response to 3a (ACK)

3e - 2c sent to SceneProfile for acceptance of scene

3f - Positive response to 3e (ACK)

3g - 2c sent to user for acceptance

3h - Positive response to 3g (ACK)

3i - Service acceptance sent to Security for AH and ESP addition

3j - Positive response to 3i (ACK)

4a - AH and ESP secured service acceptance

4b - as 4a

4c - as 4b

User A can now start shooting objects in the agreed scene, sort out sub-objects according to the agreed SceneProfile and send audio and vision AppTraNet packets to user B. User A Control activates the Stereoscopic Multiview Shooting, Segmentation and Object Movement Control, which tracks the objects. At user B, multimedia content packets immediately present their content on the display. User B ServiceControl activates the Stereoscopic Multiview Display and Viewer Movement Control, to track the eyes of the viewer to find the object on the display he focuses at any time.

The QualityControl is also activated to adapt the quality of the scene to the traffic load in the network.

H. Example, multimedia content packet transfer

A multimedia content packet always represents only one sub-object. When generated by the sender, the packet is coded, and finds the way through the network, through decoding, and to the display RAM at the receiving end, using only the AppTraNet packet header parameters {IP address, PT, ServiceID, PixelAdrB, PixelAdrE, SPQSP}. The packet is, regarding coding, decoding and display, totally independent of all other packets. The content (pixels) is presented to the user immediately. If a sub-object is not updated after a certain time (typically 10 ms), the missing pixels are constructed by interpolation by the Buffer Manager. In network nodes, as shown in Fig. 5, content packets enter the buffer B2, and are forwarded to Q1 or are dropped in B2. From Q1 the packets are output on the link to the next node.

Fig. 7 shows how multimedia content and control packets are processed at the receiving end. The IP address points to users, and further to individuals which may set up services. The PT parameter determines whether it is a multimedia content packet or a control packet. Multimedia packets first pass the IPsec security module. Then the Drop module drops sub-objects if necessary, according to the Sequence number. The Decoder decodes the packets (NOC or JPEG 2000) and forwards to the router, that in turn sends the packet to the right place in the Scene object buffer, given by the SceneProfile (SPQSP parameter), and the PixelAdrB and PixelAdrE parameters.

Control packets join queue Q2, pass security control, and attends the ServiceControl. (The QualityControl module is described in [2]).

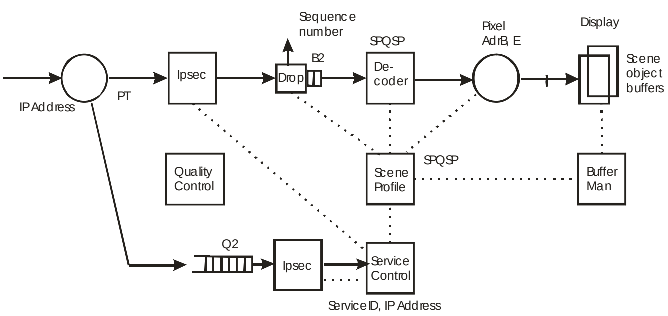

Fig. 7. Incoming packets to users. Hardware/ software architecture (simplified)

III. TRAFFIC MODELING AND PERFORMANCE EVALUATION OF THE AppTraNet PROTOCOL IN NETWORK NODES

The aim of this section is to model and study the performance and behaviour of the AppTranet protocol as part of the network node shown in Fig. 5, by means of simulation.

In a real DMP network, the collaboration is basically independent of the other collaborations, but even if multimedia content packets are independent of all other packets regarding coding, decoding, maximum transfer delay and display, they are highly auto-correlated within the collaboration, due to synchronization of packet transmission from scenes and guaranteed maximum transfer time. By nature, the collaboration can be characterized by transient traffic processes (never stationary). When many independent transient processes merge, the stream is normally expected to get ‘smoother’ slopes and lower auto-correlation. This is given by the Palm Khintchine limit theorem, extended by Cinlar [18]. However, in DMP, it is assumed that all traffic processes are transient. These very complex problems are treated further in [2].

A. Behavior of Network Nodes, modeling and simulation

Four different scene traffic sources have been modeled, a Virtual Dinner, a Song Lesson, an Existing Movie, and a Futuristic Movie, see [2]. The packet rates from each model follow transient slopes. Fig. 9 shows the pattern when five streams from each model start at random times and merge. The packet rate and rate variability are extremely high, from nearly 0 to about 50 Mega packets per second (packet length is 1.5 k Bytes).

Fig. 10 shows the pattern when fifty streams from each model start at random times, drawn from a uniform distribution [0 – 300] seconds, and merge. The packet rate varies between about 60 and 260 Mega packets per second.

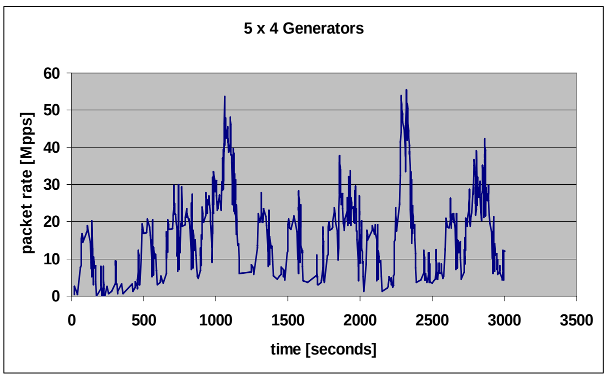

Fig. 9. Merged traffic stream, 5 x 4 streams

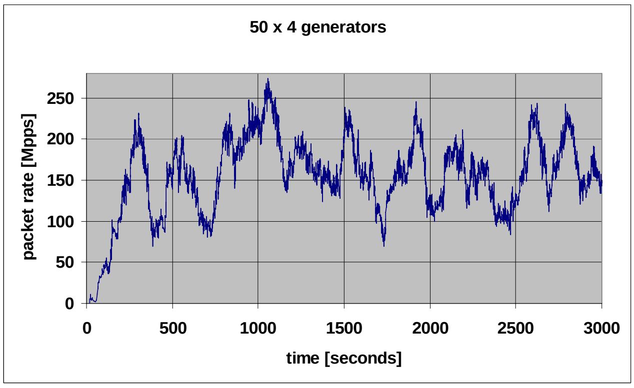

Fig. 10. Merged traffic stream, 50 x 4 streams

For simulation, the DEMOS (Descrete Event Modelling on Simula) simulation language [19] is applied in a process oriented way. A simulation environment called cim [20] has been used for building the simulation, and the simulator has been run on a high-end standard PC.

In the simulation model, it is assumed that the merged stream shown in Fig. 10, is input to and overloads an output link in node C (Trondheim) shown in Fig. 11. The streams merging at the input of an output link queue in node C, are assumed to be a representative selection of input streams to AccessNodes. The characteristics of the traffic as shown in Fig. 10 can therefore be applied as input to a node C link queue.

The scene traffic sources are modeled using DEMOS entity. Other parts of the simulator model the behaviour of Fig. 5. Packets are modeled as entity, and generated according to the packet rate vs time description given by the traffic source models. The packet entities join queues, modeled by DEMOS waitq. The ‘servers’ of the simulator, Drop, Sel and H, are also modeled as entities, and fetch packets from queues. Sel uses DEMOS condq (pointer cq below) and a uniform distribution xdraw to serve B1 and Q2 properly.

An outline of the DEMOS program with entity Sel in detail (assumed to be self-explanatory) is as follows:

```
demos begin
   ,,,declarations,,,,
   entity class ReadPacketRateFromFile,,,,,,;
   entity class GenerateNOCpacket,,,,;
   ,,,other classes,,,

   entity class Sel;
   begin
      ref (entity) ep;
      real p;
loop:
      cq.waituntil(or2(
      and2(Q2.length > 0,Q1.length < Q1n),
      and2(B2.length > 0,Q1.length < Q1n)));
      begin
        p:= xdraw.sample;
        ep :- none;
        if p > p1 then
        begin
           if B2.length > 0 then ep :- B2.coopt;
        end else
        begin
          If Q2.length > 0 then ep :- Q2.coopt;
        end;
        if ep =/= none then
        begin
          ep.schedule(now);
        end;
      end;
      repeat;
   end***Sel***;

   ,,,object generation,,,
   ,,,statistics,,,
   ,,,run time, intervals,,,
end*****demos***;
```

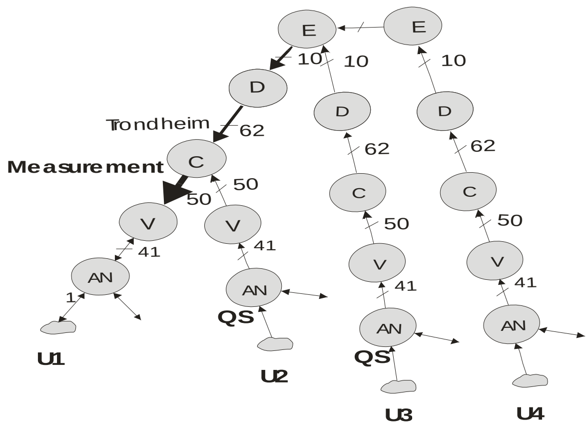

Fig. 11. DMP network, using a 5-level European hierarchy of nodes [2]

The capacity of the output link, modeled by H, is set to 200 Mpps (Mega packets per second). Assuming that the maximum length of Q1 to start with is zero, dropping of packets occurs when the input packet rate exceeds 200 Mpps.

How do viewers perceive the time-varying quality of scene objects? This depends highly on the content characteristics of the video, object movements, zooming, panning, the maximum temporal and spatial quality (resolution) of important objects, the degree of reduced quality due to sub-object drop, the occurrence of reduced quality periods, the duration of the periods, and the interpolation method (missing pixels found by interpolation). Assuming standard test videos with wanted characteristics and of defined maximum quality, and an interpolation method, the perceived quality test can be designed using results from the simulator presented in this paper. The quality variations of the objects are measured indirectly by measuring the drop period occurrences vs drop packet rate vs drop period duration.

Fig. 12 shows the 3D surface diagram for the drop period occurrences vs drop packet rate and drop period length.

The 50 x 4 generators produce overload intervals larger than 800 ms in only 1.2% of all drop period occurrences. 86% of all drop period occurrences are below 50 Mpps overload and at the same time intervals shorter than 600 ms. A dropping scheme as in Fig. 5 with 9 sub-objects, would reduce the packet rate down to around 200 Mpps by dropping one sub-objects, two sub-objects, and rarely three sub-objects.

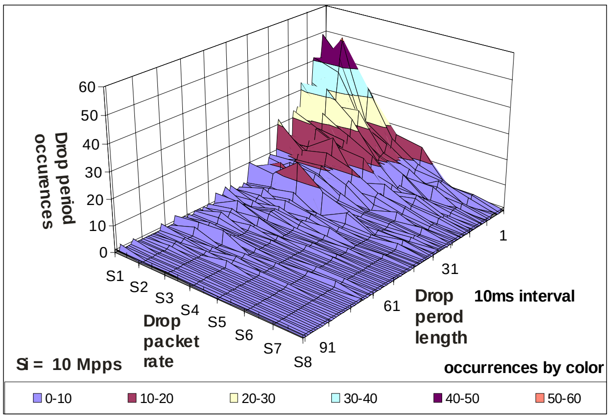

Fig. 12. Drop period occurrences vs drop packet rate vs drop period length.

How this time-varying quality of the scenes, that is, varying temporal and spatial resolution of objects and jitter are perceived by the users, remain to be tested. A perceived quality test can be carried out using a stereoscopic projector wall with sufficient spatial and temporal resolution, and video servers with the necessary play out capacity. The video (and sound) can be edited and constructed off-line, introducing interpolation of dropped sub-objects, as given by the simulations.

Assuming that the jitter of the end-to-end delay should not exceed 10 milliseconds (ms), the sum of Q1 maximum queuing delays over all nodes (maximum ten) in a path must not exceed 10 ms. Queuing delays of 10 ms will add to the propagation delays, which in Europe is smaller than about 30 ms, worst case. In this case the Q1 delays have small effects on the dropping of packets, and using zero queue lengths should be considered. However, in the case when the propagation delay is around 1 ms in a path, the sum of the queuing delays could be increased to say 20-30 ms, which could reduce the dropping noticeably.

The simulations show dropping in only one node of a large network. The merging of traffic streams into the considered output link in node C in Fig. 11 could, in the case of saturated links into C, be modeled just by adding a constant packet rate to merged rate in Fig. 10. If all input links are saturated, the merged packet rate will be constant. More likely is something in between this extreme situation and the one shown in Fig. 10. That is, a merged stream with lower variability of the packet rate than shown in Fig. 10. If the average packet rates are held constant, the drop occurrence periods are expected to be longer, and the drop packet rates lower in the new merged stream. This means a more stable but reduced perceived quality due to sub-object drop. Tests have to be conducted to verify this.

Fig. 13 shows the step response from a simulation when the input packet rate steps from zero to about 40% (mean) above the output link capacity (200Mpps). If the maximum length of Q1 is 56 Mega packets, then the queue Q1 is full after 0.7 ms (seven time intervals) (200 *1.4 – 200)*0.7 = 56), and the buffer B2 starts dropping. By dropping packets from sub-objects 8, 7 and 6, up to 40% reduction of the packet rate is obtained 1.1 ms (eleven time intervals) after the step in input.

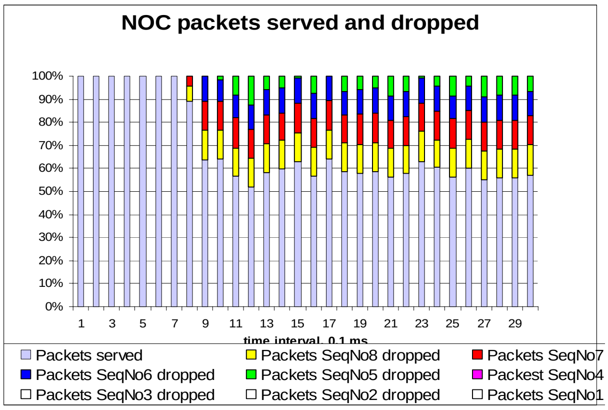

Fig. 13. Simulation – NOC packets served and dropped

IV. Conclusion

A new protocol handling collaboration setup, quality shaping control, and audiovisual transfer, and combining the application-, transport- and network layers into one layer, is proposed. This protocol is designed to support the collaboration data rates expected ten years from now. The protocol is less complex than existing protocol stacks (RTP/UDP/IP, SIP/TCP/IP, or extensions of TCP (such as XCP, RCP, DCCP/TFRC, and RED)). The protocol guarantees a maximum user-to-user delay of audiovisual packets. Due to independent packets from sub-objects, controlled dropping, feedback control and admission control, a minimum quality of a scene object can also be guaranteed. The complexity of overload problems is dramatically simplified compared to existing systems. Graceful degradation of the scene quality is obtained even when the network is severely overloaded.

The next step of future work is to test how the time-variable scene quality, as given by Fig. 12, is perceived by the viewers, and to state the technical requirements in order to obtain near-natural scene quality.

References

[1] E Heiberg, "Stereoscopic video and requirements to spatial and temporal resolution," Master's thesis, NTNU 2007.

[2] L. A. Rønningen, "The DMP system and physical architecture," Technical web report, Dept. of Telematics, NTNU, 2007. http://www.item.ntnu.no/~leifarne/

[3] L. A. Rønningen, A. Lie, "Distributed Multimedia Plays with QoS guaranties over IP," Wedelmusic, Leeds 2003.

[4] L. A. Rønningen, "Analysis of a Traffic Shaping Scheme," Proceedings of the 10th International Teletraffic Congress, ITC'10, Montreal 1983.

[5] J. S. Turner, "New Directions in Communications (or Which Way to the Information Age)," IEEE Commun. Magazine, vol.24, pp. 8-15, Oct. 1986

[6] T. V. Lakshman, P. P. Mishra, K. K. Ramakrishnan, "Tranporting Compressed Video Over ATM Networks with Explisit Rate Feedback Control," INFOCOM '97: Proceedings of the INFOCOM '97. Sixteenth Annual Joint Conference of the IEEE Computer and Communication Societies. Driving the Information Revolution, (Washington, DC, USA), p.38, IEEE Computer Society, 1997.

[7] A. Lie, O. M. Aamo, L. A. Rønningen, "A Performance Comparison Study of DCCP and a Method with non-binary Congestion Metrics for Streaming Media Rate Control, " in Proc. Of 19th International Teletraffic Congress, ITC'19, Beijing, China, Aug-Sep 2005.

[8] D. Katabi, M. Handly, C. Rohrs, "Congestion Control for High-Bandwidths-Delay product Networks,", Proc of ACM Sigcomm, 2002.

[9] N. Dukkipati, M. Kobayashi, R Zhang-Shen, N. McKeown, "Processor Sharing Flows in the Internet," in Proc. Of Thirteenth International Workshop on Quality of Service (IWQoS), Passau, Germany, June 2005.

[10] S. Floyd, V. Jacobson, "Random Early Detection gateways for congestion avoidance," IEEE ACM Transactions on Networking,1(4), August 1997

[11] DCCP Datagram Congestion Control Protocol, IETF 2007, http://www.ietf.org/html.charters/dccp-charter.html

[12] A. Lie, "P-AQM: Low-Delay Max-Min Fairness Streaming of Scalable Real-time CBR and VBR Media," "unpublished", 2007.

[13] RFC 2640, IPv6, www.ietf.org/rfc/rfc2460.txt

[14] RFC 4302 IPsec, IP security. www.ietf.org/rfc/rfc4302.txt

[15] RFC 2408, ISAKMP, Internet Security Association and Key Management Protocol. www.ietf.org/rfc/rfc2408.txt

[16] RFC 4306, IKE, Internet Key Exchange. http://tools.ietf.org/html/rfc4306

[17] E. Cinlar, "Superposition of point processes. In Stochastic Point Processes: Statistical Analysis, Theory, and Applications," ed. P. A. W. Lewis, 549-606, Wiley Interscience, New York, 1972.

[18] DEMOS simulation language http://www.item.ntnu.no/fag/ttm4110/host06/index.html

[19] Cim simulation environment http://www.item.ntnu.no/fag/ttm4110/host06/index.html