Description and Lookup of Media-Stream Adaptation Services

Andreas Schorr, Franz Hauck
Dept. of Distributed Systems
University of Ulm, Germany
{andreas.schorr, franz.hauck}@uni-ulm.de

Andreas Kassler
Dept. of Computer Science
Karlstad University, Sweden
kassler@ieee.org

Abstract: In this paper, we propose a new application of RDF that enables the description of services offered by so-called media-stream adaptation nodes (MSANs). An MSAN can manipulate a stream by changing media format and quality on-the-fly during an ongoing streaming process. An accurate description of the offered services is necessary, because different clients may have very specific requirements that cannot be fulfilled by every MSAN. In this paper, we propose an RDF-based vocabulary that enables an MSAN to provide such an accurate description of its services. We also demonstrate how clients can formulate search queries to find a services provider that fulfills their specific requirements.

1 Introduction

The *Resource Description Framework (RDF)* [W3C04b] is a powerful tool mainly used in the context of the Semantic Web. By enriching traditional web content with machine-readable metadata, the Semantic Web facilitates automated information gathering and allows automated agents to perform complex tasks on behalf of the user thus enabling a much better usage of services offered on the Web. While RDF has traditionally been used for representing metadata related to *web resources*, it can just as well be used for the description of other kinds of services not necessarily related to the Web.

The dynamic adaptation of multimedia content in distributed heterogeneous environments is a key enabler for next generation ubiquitous and pervasive services. Media streaming solutions and systems need to be adaptive to bridge the heterogeneity of networks and devices and to cope with the best effort nature of the current Internet. This will lead to the notion of *media-stream adaptation (MSA) services*. An MSA service can manipulate a stream by changing media format and quality on-the-fly during an ongoing streaming process to provide the best quality for the available ressources. Such a service may be located on a proxy node inside the network if the end-terminals are not able to perform adaptation themselves. In this paper, we propose a new application of RDF that enables the description of services offered by such an adaptation node. Several adaptation-node architectures have been proposed [AMZ95, Y^+96, KS03], but in these proposals, the adaptation nodes usually act as *media gateways*. The gateway and the services that it can provide are assumed to be known a priori by the clients or by a management architecture that controls the gateway on behalf of the clients. Our proposal allows to offer a new kind of MSA service that can be publicly announced and dynamically discovered by clients.

The paper is structured as follows. In Section 2, we shortly discuss existing methods for service description and analyze their applicability for the description of MSA services. Section 3 provides an overview of the operations that may be offered by an MSA service provider. In Section 4, we introduce an RDF vocabulary for the description of media stream adaptation services using the *RDF Vocabulary Description Language* (also known as *RDF Schema*) [W3C04b]. In Section 5, we show how clients can extract information about MSA services from a service registrar using the *SPARQL Query Language for RDF* [W3C06]. We conclude the paper in Section 6.

2 Describing Services

In recent years, several service discovery technologies have been developed, e.g., Salutation [Sal98], Service Location Protocol (SLP) [G^{+}99], Jini [Sun99], Universal Plug and Play (UPnP) [Uni00]. Each technology comprises a language for service description as well as methods used to find services whose description match certain attributes specified by the user. Many existing languages (e.g., SLP Templates [GPK99]) can only define service descriptions that consist of simple key-value pairs. This is sufficient for the description of many simple services such as a printer service with the following possible attributes: *resolution*, *paper size*, or *pages per second*. Here, the attribute values would have simple data types like *integer* or *string*. However, SLP templates fail to describe services such as MSA services, whose description requires more complex data structures and the possibility to express relations between attributes.

RDF allows to make statements about "resources", each statement consisting of subject, predicate, object. A "resource" can be a website, a service, or any other thing that can be uniquely identified. An RDF statement is represented by a graph consisting of node-arc-node (which correspond to subject-predicate-object). Simple graphs, each one representing a single statement, can be concatenated to form arbitrary graphs which represent more complex statements. Services offered by media stream adaptation entities can be described by statements such as "transcodes an MP3 audio stream into G.711 in less than 20 ms". Therefore, RDF seems to be a natural choice for the description of MSA services. Out of the above mentioned technologies, UPnP also uses a more structured, XML-based description model. However, some statements may appear multiple times in one service description. Such a statement would have to be repeated multiple times because of the hierarchical tree structure of the simple XML documents used in UPnP. In RDF, on the other hand, a subgraph can be referenced multiple times without repeating it.

RDF and other description formats define only a language syntax and formal semantics of the basic language constructs, but they do not define *vocabularies* (also called *ontologies*) for the description of resources belonging to a specific category. Typically, different resource categories require different vocabularies. A vocabulary for the description of media-stream adaptation services has not been proposed in the literature before. The idea of publicly announcing MSA service descriptions so that clients can automatically find adaptation services that match specific requirements has already been mentioned within the scope of the project IST-Daidalos [S+05, GL+05], but these earlier proposals describe only the general architecture of a pervasive service discovery service. They do not define concrete vocabularies for specific services like MSA.

We specified our vocabulary by means of RDF Schema [W3C04b]. Alternatively, the Web Ontology Language OWL [W3C04a] could be used for the definition of such an RDF vocabulary. OWL allows to add additional semantics to a vocabulary (ontology) which cannot be expressed with RDF Schema, e.g., disjointness of classes, cardinalities of properties, etc. Nevertheless, as we will demonstrate in Section 5, searching for MSA services that match specific requirements works well with our vocabulary. OWL does not provide any particular benefits in our specific application scenario. On the other hand, the usage of OWL would add additional complexity to the processing of service descriptions.

3 Media-Stream Adaptation Services

In this section, we give an overview on the services a media-stream adaptation node (MSAN) may offer and describe the service parameters that a client has to know in order to decide whether a certain service provider fulfills the client's requirements. While MSANs may adapt streams belonging to non-interactive sessions like video-on-demand (VoD) or live-broadcast as well as interactive sessions like voice-over-IP (VoIP) or video-conferencing, the proposal in this paper refers to adaptation services for *realtime media streaming*. Here, the receiver starts decoding media data while the sender is still transmitting. An MSAN may offer the following services:

- Media Adaptation
  - *Transcoding*: Conversion from one media format into another one, e.g., from MPEG-2 to H.263, or from high bit-rate MPEG-4 to low bit-rate MPEG-4.
  - *Spatial scaling*: Reduction of video frame size.
  - *Temporal scaling*: Reduction of video frame rate or audio sampling rate.
  - *SNR scaling*: Reduction of the quality (the signal-to-noise ratio, SNR) of a media stream. Depending on the media codec, either a certain target bit-rate or a certain quality level (or both) can be achieved.
  - *Channel scaling*: Reduction of the number of audio channels.
  - *Mixing*: Mix several incoming media streams (e.g., audio) into a single stream.
  - *Media translation*: Translate from one media type into another one (e.g., text into speech or vice versa).
- Network Flow Adaptation
  - *Multipoint session*: Create multiple adapted versions of a single media stream and distribute to multiple downstream nodes.

– *Protocol adaptation*: Convert from one protocol stack used by the upstream node into another protocol stack supported by the downstream node.
– *Adaptation of error control*: Use different (or additional) application layer error control schemes in upstream and downstream direction.
– *Conversion between RTP profiles*: Convert from one RTP profile (Realtime Transport Protocol) used by the upstream node into another one supported by the downstream node.
– *Rate control*: Apply specific rate-control schemes for controlling the amount of network traffic in downstream direction.

Some adaptation nodes may offer identical adaptation operations but have different hardware capabilities or use different adaptation techniques for achieving the same result. As a consequence, delay, jitter, quality reduction, and costs caused by the adaptation process can vary on different MSANs. As clients may have very strict requirements on some of these parameters (such as maximum end-to-end delay below 150 ms), not every MSAN will be able to fulfill each client's requirements. In some (but not all) cases, parameters like processing delay are variable or depend on the media content. The delay, for instance, may vary if input and output formats make use of *bi-directional predictive video coding*, where the order in which the video frames occur in a stream will not be identical to the display order of the video frames. If an MSAN adapts such a stream, the transcoder may have to re-order video frames once again (depending on the combination of input and output format) thus generating additional delay. Also, there exist different types of transcoders, some of which will re-order the video frames for a given combination of input and output formats, whereas others will not re-order the frames for the same combination of formats. If re-ordering occurs, the resulting delay can be different for different media streams encoded with the same codec. As a conclusion, the description of each individual adaptation operation must include the parameters *delay*, *jitter*, *quality reduction*, and *costs* and must indicate whether these parameters are content-dependent or not.

We also have to take into account that there exist two completely different approaches for client-MSAN interaction. In a terminal-driven scenario, the client could instruct the MSAN which adaptation operations to perform, e.g., "transcode from MPEG-2 to MPEG-4, reduce the picture size by factor two, and use a target bit-rate of 400 kBit/s". In an MSAN-driven scenario, the client informs the MSAN about the *usage environment description (UED)* [VT05] of the media streams (i.e., user preferences, capabilities and restrictions of the involved terminals and networks). Here, the MSAN decides on its own which adaptation operations to apply. A standardized XML-based format for the representation of UED is defined in MPEG-21, Part 7: Digital Item Adaptation (DIA) [VT05]. Finally, the service description must also contain information about the way clients have to interact with an MSAN. For instance, different MSANs may support different signalling protocols like *Session Initiation Protocol (SIP)* [R$^+$02] or *Media Gateway Control Protocol (MEGACO)* [G$^+$03] for session setup and control. Similarly, they may support different formats for the description of the session content, e.g., *Session Description Protocol (SDP)* [Jac98] or *SDP new generation (SDPng)* [K$^+$05].

4 An RDF Schema for Media-Stream Adaptation Services

In this section, we introduce an RDF vocabulary for the description of media-stream adaptation services (denoted as MSAS vocabulary). Since the vocabulary is quite large, we cannot show the whole RDF schema here. Instead, we show several extracts from an example service description and describe a selection of the classes and properties defined by the MSAS schema. The URI for the vocabulary namespace is `http://mqos.de/ns/msas-schema-v1.rdf`. The complete MSAS schema is accessible from the Web through the same URI. In the following text and figures, we use qualified names with the prefix `msas` assigned to the MSAS vocabulary namespace.

Figure 1 shows an extract of an RDF graph describing a fictitious MSAN. To distinguish the blank nodes in the graph from each other, increasing numbers starting from 1 are assigned to them as blank node identifiers. Blank node `_:1`, which is an instance of the `msas:Contact-List` class, aggregates multiple `msas:contact-info` properties, which describe how to access the services (class names are not explicitly shown in the figures). Since these properties contain structured information, the property values are again modelled as blank nodes (`_:2` and `_:3`), each one being an instance of the class `msas:Contact` and aggregating properties which contain information about a single service access method. In the depicted example, the MSAN services can be accessed by using the signalling protocols SIP or MEGACO. The resource `msas:sip` (an instance of the class `msas:Sig-Proto-Id`) is defined in the MSAS schema and identifies the Session Initiation Protocol; the resource `msas:megaco` identifies the MEGACO protocol. The SIP URI of the MSAN is `sip:a@b.c`, and the MSAN listens for SIP messages at port `5060`. The property `msas:transp-layer` describes which transport-layer protocols can be used to transport session-layer protocols SIP and MEGACO. In the depicted example, SIP can use either UDP or TCP, Megaco is restricted to use TCP.

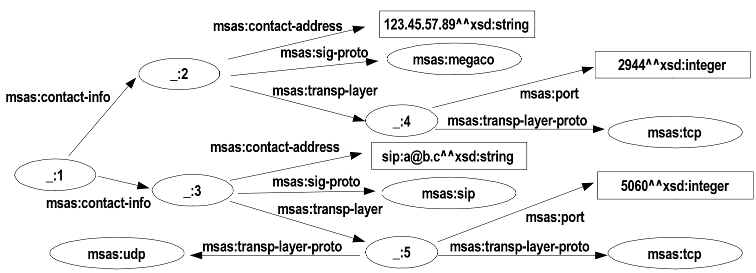

Figure 1: Description of MSAN contact information

Figure 2 shows another extract of the MSAN description. Here, blank node `_:1` is an instance of the class `msas:Media-Adapt-Op` which represents a single media-adaptation operation offered by the MSAN. For simplification, we included in the figure only a subset of the properties that describe the operation. Additional properties not shown in Figure 2 would provide information about jitter, quality reduction, and costs. The properties `msas:in-format` and `msas:out-format` define the input and output media format for the adaptation process, the property `msas:scale-ops` provides a description of possible scaling operations. Transcoding and scaling are performed together as a single media-adaptation operation, and descriptive attributes such as the `msas:delay` property refer to this combined operation as a whole. It is possible that different scaling operations cause different delays for a given combination of input and output media formats. For instance, a special transcoder module may provide SNR scaling with very low delay, whereas spatial scaling would generate a much higher delay. In such a case, two different instances of the `msas:Media-Adapt-Op` class would have to be created, one that includes only SNR scaling, and another one that includes only spatial scaling, and the `msas:delay` property of each `msas:Media-Adapt-Op` instance would indicate the respective delay. URIs for the identification of media formats are defined in the MPEG-7 Media Description Schemes [ISO01] standard, which includes the *Audio Coding Format Classification Scheme (ACFCS)* and *Visual Coding Format Classification Scheme (VCFCS)*. We have assigned the prefix `vcf` to the namespace URI `urn:mpeg:mpeg7:cs:VisualCodingFormatCS:2001:` of the VCFCS. In the depicted example, the input format identifier is `vcf:2.1`, which denotes *MPEG-2 Video Simple Profile*. The output format is `vcf:3.1`, which stands for *MPEG-4 Visual Simple Profile*. The processing delay does not depend on the media content and amounts to 50 ms.

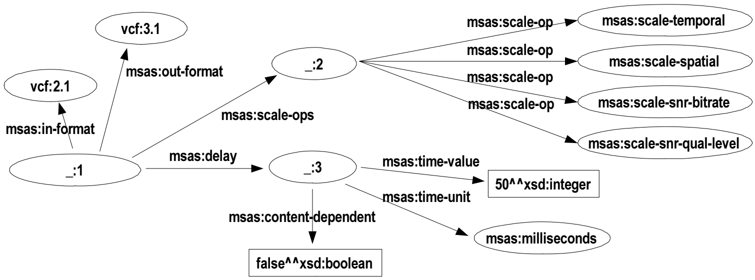

Figure 2: Description of a single media-adaptation operation

For a complete description of a single MSAN, the subgraphs shown above are connected to a single node that represents the MSAN itself. A complete description of an MSAN would contain additional properties that cannot be shown due to space restrictions. Some of them have simpler structures. For instance, whether an MSAN can process MPEG-21 Usage Environment Description (see Section 3) can be expressed by a single boolean property value. A full description of the fictitious MSAN is accessible through the URI `http://www-vs.informatik.uni-ulm.de/proj/qos/examples/msan-ex1.rdf`.

5 Search Queries

We assume that multiple MSANs register their service descriptions at a central service discovery server (SDS), as proposed in [S+05]. The main application of our vocabulary is then to search for an MSAN that can provide a specific adaptation service while fulfilling certain requirements. Either clients in need of an adaptation service or network elements such as SIP proxies can formulate search queries that refer to the SDS's database, which contains all registered service descriptions. We propose to formulate search queries for MSA services by means of the *SPARQL Query Language for RDF* [W3C06].

Within this paper, we can only show one representative example for a complex search query. Here, the client wants to know the contact information of MSANs which are able to transcode a stream encoded with *MPEG-2 Video Simple Profile* into *MPEG-4 Visual Simple Profile* with adaptation delay below 100 ms. Furthermore, the client needs an MSAN that can communicate via the MEGACO protocol. At most three results shall be returned, in ascending order of the delay. The corresponding SPARQL query would be:

```
PREFIX msas: <http://mqos.de/ns/msas-schema-v1.rdf#>
SELECT ?address ?port ?time
WHERE { ?msan   msas:contact-info-set  ?cis .
        ?cis    msas:contact-info      ?ci .
        ?ci     msas:sig-proto         msas:megaco ;
                msas:contact-address   ?address ;
                msas:transp-layer      ?transp .
        ?transp msas:port              ?port .
        ?msan   msas:media-adapt-ops   ?ops .
        ?ops    msas:media-adapt-op    ?op .
        ?op     msas:in-format         <urn:mpeg:mpeg7:
cs:VisualCodingFormatCS:2001:2.1> ;
                msas:out-format        <urn:mpeg:mpeg7:
cs:VisualCodingFormatCS:2001:3.1> ;
                msas:delay             ?delay .
        ?delay  msas:content-dependent false ;
                msas:time-value        ?time .
        FILTER (?time < 100) . }
ORDER BY ?time
LIMIT 3
```

A possible answer is depicted below. Three MSANs have been found that match the search criteria. The fastest one can perform the conversion at a maximum delay of 50 ms.

```
address 134.60.77.210     port 2944   time 50
address 134.60.218.199    port 2944   time 75
address 134.88.99.100     port 12345  time 99
```

6 Conclusion

The availability of media-stream adaptation services in distributed heterogeneous environments is a key enabler for next generation ubiquitous and pervasive systems. In this paper, we introduced a new application of the Resource Description Framework that enables the description of MSA services. By publicly announcing MSA-service descriptions, clients can find a specific service provider that fulfills their individual requirements. We demonstrated how certain properties of a media-stream adaptation node can be described by means of the proposed vocabulary and how clients can formulate search queries for finding an appropriate media-stream adaptation node. We have implemented prototypes of an adaptation node, a service discovery server and corresponding clients [GL+05]. However, the existing prototypes use an older version of the MSAS vocabulary and clients use a proprietary protocol for extracting information from the RDF database of the SDS. We are currently working on an enhanced implementation that uses the mechanisms proposed in this paper.

References

[AMZ95] E. Amir, S. McCanne, and H. Zhang. An application level video gateway. In *Proceedings of ACM Multimedia '95*, November 1995.

[G+99] E. Guttman et al. *RFC2608: Service Location Protocol, Version 2*. IETF, June 1999.

[G+03] C. Groves et al. *RFC3525: Gateway Control Protocol Version 1*. IETF, June 2003.

[GL+05] Teodora Guenkova-Luy et al. Multimedia Service Provisioning in a B3G Service Creation Platform. In *Proceedings of IPSI-Pescara-2005*, Pescara, Italy, July 2005.

[GPK99] E. Guttman, C. Perkins, and J. Kempf. *RFC2609: Service Templates and Service: Schemes*. IETF, June 1999.

[ISO01] ISO/IEC JTC1/SC29/WG11. Information Technology – Multimedia Content Description Interface – Part 5. International Standard 15938-5:2001, ISO/IEC, October 2001.

[Jac98] V. Jacobson. *RFC2327: SDP: Session Description Protocol*. IETF, April 1998.

[K+05] D. Kutscher et al. Session description and capability negotiation, February 2005. Work-in-progress, *draft-ietf-mmusic-sdpng-08*.

[KS03] Andreas Kassler and Andreas Schorr. Generic QoS aware Media Stream Transcoding and Adaptation. In *Proceedings of Packet Video*, Nantes, France, April 2003.

[R+02] J. Rosenberg et al. *RFC3261: SIP: Session Initiation Protocol*. IETF, June 2002.

[S+05] Vincenzo Suraci et al. Design and Implementation of a Service Discovery Architecture in Pervasive Systems. In *IST Mobile Wireless Summit*, Dresden, Germany, June 2005.

[Sal98] Salutation Consortium. *White paper: Salutation Architecture*, 1998.

[Sun99] Sun. *Technical White Paper: Jini Architectural Overview*, 1999.

[Uni00] Universal Plug and Play Forum. *Universal Plug And Play Device Architecture*, 2000.

[VT05] A. Vetro and C. Timmerer. Digital Item Adaptation: Overview of Standardization and Research Activities. *IEEE Transactions on Multimedia*, 7(3), June 2005.

[W3C04a] W3C. *OWL Web Ontology Language Overview, Recommendation*, February 2004.

[W3C04b] W3C. *Resource Description Framework (RDF), Recommendation*, February 2004.

[W3C06] W3C. *SPARQL Query Language for RDF, Candidate Recommendation*, April 2006.

[Y+96] Nicholas J. Yeadon et al. Filters: QoS Support Mechanisms for Multipeer Communications. *IEEE Journal of Selected Areas in Communications*, 14(7):1245–1262, 1996.

Workshop Mobile and Embedded Interactive Systems (MEIS'06)