Chapter X

XML Compression for Web Services on Resource-Constrained Devices

Christian Werner
University of Lübeck, Germany

Carsten Buschmann
University of Lübeck, Germany

Ylva Brandt
University of Lübeck, Germany

Stefan Fischer
University of Lübeck, Germany

ABSTRACT

Compared to other middleware approaches like CORBA or Java RMI the protocol overhead of SOAP is very high. This fact is not only disadvantageous for several performance-critical applications, but especially in environments with limited network bandwidth or resource-constrained computing devices. Although recent research work concentrated on more compact, binary representations of XML data only very few approaches account for the special characteristics of SOAP communication. In this article we will discuss the most relevant state-of-the-art technologies for compressing XML data. Furthermore, we will present a novel solution for compacting SOAP messages. In order to achieve significantly better compression rates than current approaches, our compressor utilizes structure information from an XML Schema or WSDL document. With this additional knowledge on the “grammar” of the exchanged messages, our compressor generates a single custom pushdown automaton, which can be used as a highly efficient validating parser as well as a highly efficient compressor. The main idea is to tag the transitions of the automaton with short binary identifiers that are then used to encode the path through the automaton during parsing. Our approach leads to extremely compact data representations and is also usable in environments with very limited CPU and memory resources.
INTRODUCTION

The text-oriented data encoding of XML (extensible markup language) is the reason for SOAP messages causing significantly more overhead than the binary message formats of Java RMI (Remote Method Invocation) and CORBA (common object request broker architecture). In an earlier work, we compared different approaches for realizing remote procedure calls (RPCs) and showed that SOAP over HTTP (hypertext transfer protocol) causes significantly more traffic than similar technologies. Using SOAP, the data volume is about three times higher than with Java RMI or CORBA (Werner, Buschmann, & Fischer, 2005).

Fortunately, most of today’s wired networks are fast enough to provide sufficient bandwidth for all applications. However, there are still some application domains with tight bandwidth limits: For example, in cellular phone networks, it is still common to charge the customer according to the transmitted data volumes. Another very common example is a dial-up connection over older technologies like modem or ISDN (integrated services digital network) links. Although their bandwidth is very limited, they are still in use in many enterprise networks. Additional limitations are imposed by foreseen application domains of Web services such as ubiquitous computing. In such energy-constrained environments, the radio interface is usually a main power consumer and therefore tight restrictions apply to the transmitted data volumes on mobile devices.

In order to address the problem of excessive XML message sizes in these domains, a lot of research effort went into the development of binary (and therefore more compact) representations of XML data.

In order to preserve the universal compatibility of binary-encoded XML, standardization is a very important issue: The World Wide Web Consortium (W3C) founded the W3C XML Binary Characterization Working Group in March 2004. Its members analyzed various application scenarios and created a survey of the existing approaches in this field (W3C, 2005b). Furthermore, this working group has specified a set of requirements that are important for binary XML representations. The most requested features are compactness, the possibility of directly reading and writing the binary format, independence of certain transport mechanisms, and processing efficiency.

Another major outcome of this working group was a set of 18 typical use cases for binary XML representations with a detailed analysis of their individual requirements. In all use cases, the property compactness, which is in the focus of this article, was of major importance or was rated as a nice-to-have feature. In 10 of 18 cases, it was even rated as mandatory.

The W3C XML Binary Characterization Working Group has finished its work, and its successor, the Efficient XML Interchange Working Group (W3C, 2005a), took up the work in December 2005. It focuses on interoperability aspects of binary XML and published a first working draft of the efficient XML interchange (EXI) format (W3C, 2007) in December 2007. Although not in the focus of this article, we are currently working on an implementation of this data format. To the knowledge of the authors, there are no other implementations of the EXI format available up to now.

In this article, we elaborate on how to encode SOAP messages efficiently with an approach that has been developed independently of EXI. It exploits the fact that Web service messages are described by an XML grammar that is known to both the sender and the receiver (usually in the form of a WSDL [Web services definition language] file). A large part of the information contained in a message can be inferred from this grammar. This a priori known part can therefore be omitted during transmission; this leads to very promising compression results.