A PERFORMANCE EVALUATION OF SOAP VARIANTS FOR GIS WEB SERVICES

Seung-Jun Cha*, Jun Li*, Yun-Young Hwang*, Kyu-Chul Lee*, Yoon-Seop Chang**, Kyung-Ok Kim**
*Department of Computer Engineering, Chungnam National University
**Electronics and Telecommunications Research Institute (ETRI)
{junii, leejun_1224, yyhwang, kclee}@cnu.ac.kr, {ychang76, kokim}@etri.re.kr

ABSTRACT By introducing Web Services, distributed GIS services from different vendors can be dynamically integrated into a GIS application using the interoperable standard SOAP protocol. However, it is debatable whether SOAP can really meet the performance requirements of GIS. This paper presents an experimental evaluation of the performance of different SOAP variants: standard SOAP, SwA/MIME, and SOAP/MTOM. The objective of this paper is to demonstrate that SOAP performance in communicating large volumes of GIS data could be effectively improved by recent standards. Moreover, SOAP/MTOM is identified to be the fastest and the most efficient messaging protocol.

KEY WORDS: GIS, Web Services, SOAP, MIME, MTOM

1. INTRODUCTION

Over the past decade GIS(Geographic Information Systems) technology has evolved from the traditional model of stand-alone systems to distributed models. Distributed GIS services will be implemented more extensively by using Web Services. By introducing Web Services, distributed GIS services from different vendors can dynamically integrated into a GIS application using the interoperable standard SOAP(Simple Object Access Protocol) protocol(Alameh, 2003). However, it is debatable whether SOAP can really meet the performance requirements of GIS.

This paper presents an experimental evaluation of the performance of different SOAP variants: i.e. standard SOAP, SwA(SOAP with Attachments) using MIME(Multipurpose Internet Mail Extension), and SOAP using MTOM(Message Transmission Optimization Mechanism). These standard protocols are evaluated in communicating multiple raster(JPEG) map data and vector(GML) data. The objective of this paper is to demonstrate that SOAP performance in communicating large volumes of GIS data could be effectively improved by recent standards. Moreover, SOAP/MTOM is identified to be the fastest and the most efficient messaging protocol.

The rest of this paper is organised as follows: a brief discussion of OGC Web Services standards and SOAP's variants is given in section 2. Section 3 shows comparisons with related work, and section 4 contains how we conduct the experimental evaluation. The results and analyses of our evaluation are described in section 5. Conclusions are followed in section 6.

2. BACKGROUND

2.1 OGC Web Services Standards

The Open Geospatial Consortium(OGC) is a non-profit, international standard organization that leads the development of standards for geographic data related operations and services(Sayar, et al., 2006).

The OGC Web Services(OWS) initiative has undergone multiple phases – including the mapping service as Web Map Server(WMS), data manipulation operations services as Web Feature Server(WFS), services for allowing access to geospatial coverages as Web Coverage Server(WCS), and OGC Web Service Architecture. WFS provides feature data in vector format encoded in Geographic Markup Language(GML) and WCS provides coverage data in raster format.

2.2 SOAP and It's Variants

SOAP is a platform-independent, extensible and XML-based protocol for distributed computing. A SOAP message consists of three different elements: the SOAP Envelop containing SOAP Body and the optional SOAP Header(Mitra, 2003).

A typical SOAP message is structured as follows:

```
<?xml version='1,0' ?>
<soap:Envelope
  xmlns:soap="http://www.w3.org/2003/05/soap-
  envelope">
<soap:Body>
   <m:data xmlns:m=;http://example.org/people">
     <photo>/aWKKapGGyQ=</photo>
              <wav>Faa7vROi2VQ=</wav>
   </m:data>
  </soap:Body>
</soap:Envelope>
```

SwA/MIME applies MIME attachments to SOAP, using multipart/mime content type and putting the SOAP Envelope in the root MIME part and other related attachments in ensuing MIME parts inside the MIME package. It relies on HREF attribute and Content-ID MIME header to relate attachments to SOAP message parts(Barton, et al., 2000).

WSA/DIME(WS-Attachment using DIME) uses DIME to send binary data and is more efficient than SwA. DIME is a simpler protocol than MIME but there are no plans for standardizing WSA/DIME, which has restricted its development. So we exclude it from the test.

SOAP/MTOM, a part of SOAP 1.2 specification, applies XOP to SwA. Staying in the XML Infoset, SOAP/MTOM-based attachments are equivalent to embedded SOAP elements semantically to the endpoint SOAP Nodes. With many Web Services standards defined on XML Infoset model, its significance lies in the fact that the presence of attachments in SOAP messages is no longer an exceptional or special case(Gudgin, et al., 2005). A SOAP/MTOM message, which is equivalent to the previous SOAP message, can be represented as follows:

```
Content-Type: application/xop+xml; charset=UTF-8;
  type="application/soap+xml; action=\"ProcessData\""
Content-Transfer-Encoding: 8bit
Content-ID: mymessage.xml@example.org
Content-Description: A SOAP Envelope with my picture in it
<?xml version='1,0' ?>
<soap:Envelope
xmlns:soap='http://www.w3.org/2002/12/soap-envelope' >
  <soap:Body>
   <m:data xmlns:m='http://example.org/stuff'>
    <m:photo xmlmime:contentType='image/jpg'>
     <xop:Include
href='cid:http://example.org/me.jpg'/></m:photo>
    <m:wav mlmime:contentType='sound/wav'>
     <xop:Include
xmlns:xop='http://www.w3.org/2004/08/xop/include'
  href='cid:http://example.org/my.wav'/></m:wav>
   </m:data>
  </soap:Body>
</soap:Envelope>
--MIME_boundary
Content-Type: image/jpg
Content-Transfer-Encoding: binary
Content-ID: http://example.org/me.jpg
fd a5 8a 29 aa 46 1b 24
...
```

3. RELATED WORK

Previous researchers have examined the performance of using SOAP in a number of different scenarios. (Ying, et al., 2004) reported that performance could be simply and effectively improved by reducing the payload of a SOAP message using SwA using MIME, WSA using DIME, and XSOAP. (Ng, et al., 2005) argued on the overall performance of SOAP and the effectiveness of compression and binary encoding as ways of improving performance or reducing bandwidth requirements. They used and tested SOAP, gZip, remoting binary, and MTOM for compression and binary encoding.

Our study differs from pervious studies in that we compare performance among standard SOAP, SwA/MIME, and SOAP/MTOM. And there are no studies of comparing SOAP protocols with GIS Web Services.

4. EXPERIMENTAL DESIGN

In our test scenario (Figure 1), the client reads external GIS data and attaches it to the SOAP message, and then the message is sent to the server. The server reads that message and composes simple response message, then sends it back to the client again.

In the real situation of GIS web services, the server should send GIS data, and the client displays the data on the screen. However, the web server is not implemented to send large data(e.g. the Axis web server cannot send over 16Mbytes data when we have tested), the test is performed in the opposite direction. As you know the result of our evaluation later, there is no case that the web server should send over 16Mbytes GIS data in a normal Internet environment.

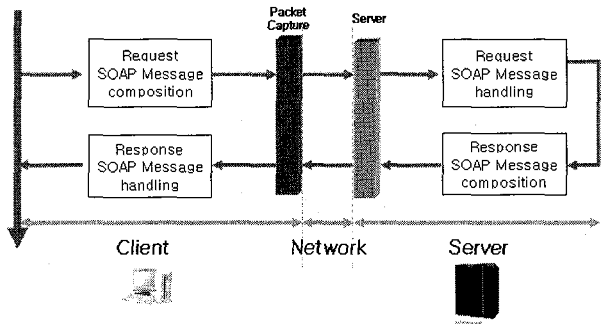

Figure 1. Test scenario

4.1 Evaluation Metrics and Methods

The following performance evaluation metrics and methods are used during our test. A packet capture program in the client-side is used in evaluating message size and serialization/deserialization time.

- **Message size:** It means the size of the SOAP message which includes attachment data at the client. Packet capture program is used for evaluating it.

- **Roundtrip time:** It means the time taken from the point after the SOAP message composition ends at the client to the point the client receives the response message from the server. It is evaluated by using timestamp in the source code.

- **Serialization and Deserialization time:** Serialization is the conversion of an object instance to a data stream of byte values in order to prepare it for transmission. Deserialization is opposed to serialization, and it is for computer's handling. The test evaluates the client-side's serialization/deserialization and the server-side's serialization/deserialization. But the server-side's serialization time and deserialization time cannot be clearly evaluated separately, because its web server cannot be modified to insert test code and the packet capture program cannot be used at the server.

4.2 Evaluation Environments

The configuration of the server systems is:
- Pentium 4 - 2.6GHz processor
- 1024Mbytes of memory (512Mbytes * 2)

• RedHat Linux 9

The configuration of the client systems is:
• Mobile Pentium 1.7GHz
• 512Mbytes of memory
• Microsoft Windows XP Professional

The SOAP implementations used in our test are Apache Axis because Axis 1.4 supports SwA/MIME, Axis 2.0 supports standard SOAP and SOAP/MTOM, and both along with the Xerces(XML parser). Java 1.5.06 is the programming language. Apache Httpd is used for web application, and Tomcat 5.0 is used as the container for Apache Httpd. The Ethereal, a shareware TCP traffic monitor, is used to evaluate metrics. The Server and the client are connected with a hub(NetGear FS2005, 100MBytes LAN), so that it is not affected by other network traffics. Two different operating systems are chosen for testing the platform independence and interoperability of SOAP messages.

4.3 Test Data Design

When we design the test data, we refer to the Google Map. The Google Map provides 12 tiles of image in a window. And the average size of each tile is 20Kbytes. Assume that current Internet speed which is commonly used is 500Kbps. But the test environment is consists of the server and the client, and just one hub. The Ethernet speed is measured almost 90Mbps. The Ethernet speed is 180 times faster than the Internet speed, so the data size should be set 180 times bigger than the Goole Map; 3Mbytes for one tiles and 36Mbytes of one window(12 tiles). Tests are performed by sending from 1Mbytes to around 64Mbytes of data.

Vector test data is composed of GML-typed data, because GIS services serve it for the response. Assume that it needs 1Kbyte polygon data for expressing a building, test are performed by communicating different numbers of buildings. The size varied from 1, 10, 100, 1,000, 5,000 to 10,000. 10,000 numbers of buildings are enough for GIS services to provide in one window.

5. EVALUATION RESULTS

5.1 Tests for Raster Data

The message size result of sending raster data of different size is presented in Figure 2. SwA/MIME and SOAP/MTOM show similar in size. When using standard SOAP, attachment data must be encoded by either Base 64 Encoding or Hexadecimal text Encoding. Base64 is a method of encoding arbitrary binary data as ASCII text. Since Base64 encoding divides three bytes of the original data into four bytes of ASCII text, the encoded data typically is about 33% bigger.

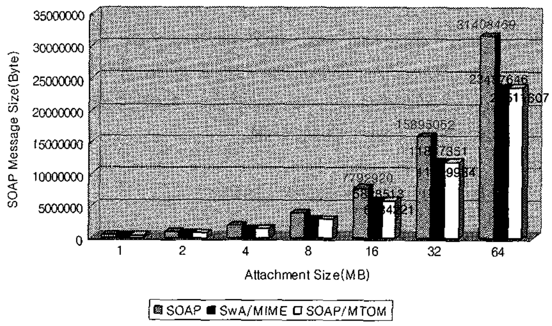

Figure 2. SOAP message size of raster data

When we test the roundtrip time and serialization / deserialization time, SOAP/MTOM achieves the best performance (Figure 3 and 4).

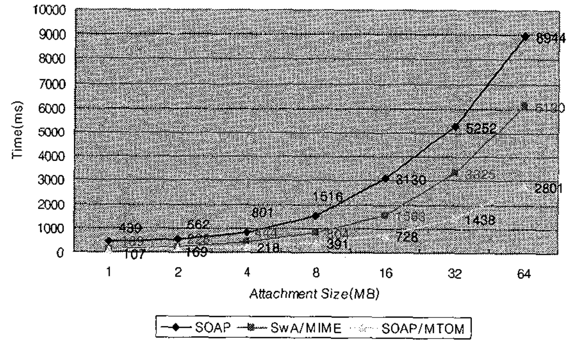

Figure 3. Roundtrip time of raster data

Using standard SOAP the external data is embedded into the SOAP envelope as an element. It takes a long time for serialization and deserialization (Figure 4). SwA/MIME and SOAP/MTOM use a URI as an element value to reference external data. That means the data isn't included in the SOAP envelop. Furthermore, SOAP/MTOM uses XOP optimization mechanism with XML infoset. It gives the recipient the option of using either the original file that may be identified by a URI, or to use a cached copy that accompanies the actual SOAP message. It enhances greatly the speed and of processing as the external data is already present when the recipient is starting processing the message.

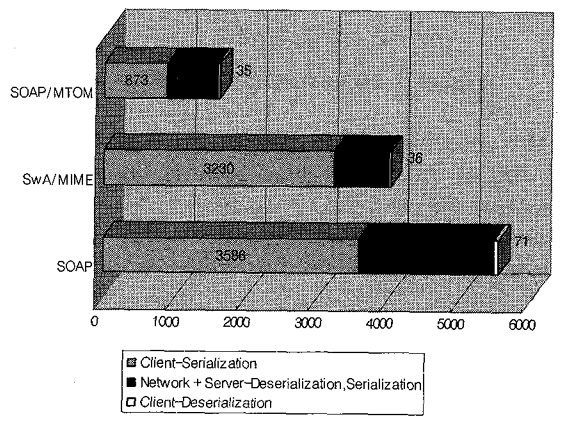

Figure 4. Serialization/deserialization time of raster data

32Mbytes of raster data is used when we perform serialization / deserialization tests in Figure 4.

As a result, SOAP/MTOM takes almost 1500ms when evaluated 32Mbytes of raster data. But standard SOAP takes over three times than that.

5.2 Evaluation for Vector Data

Because vector data is ASCII-type, not binary, no encoding is necessary for attachments. SwA/MIME and SOAP/MTOM's message size is bigger than standard SOAP rather. This is because attachment data are in the boundary, and extra tags are produced additionally. In using SOAP/MTOM, it transfer SOAP message by using XML Infoset, its additional tags are more than SwA/MIME (Figure 5).

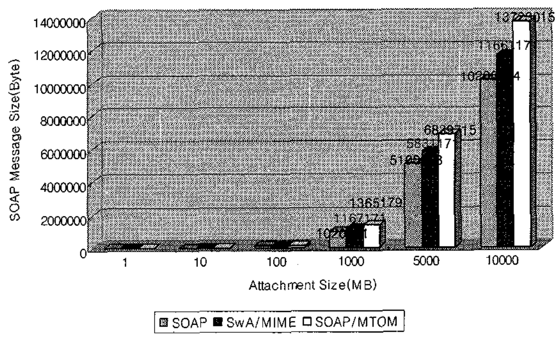

Figure 5. Message size for vector data

The result of the vector data's roundtrip time (Figure 6) is similar to that of the raster data's.

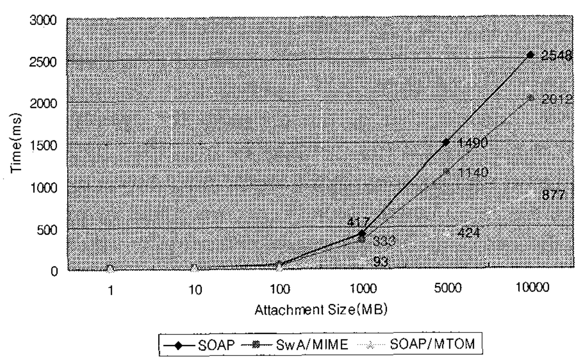

Figure 6. Roundtrip time of vector data.

All attachments are tagged in standard SOAP. SwA/MIME takes less time because deserialization is faster than standard SOAP. As SOAP/MTOM uses XOP optimization and XML Infoset, it has the smallest roundtrip time.

As a result, when sending 10,000 numbers of GML Vector data SOAP/MTOM takes three times less than standard SOAP. It means that SOAP/MTOM also archives good performance even when sending a large amount of vector data.

6. CONCLUSIONS

We have performed evaluation experiments of standard SOAP, SwA/MIME, and SOAP/MTOM using workload that reflect variety range of typical data types of GIS Web services.

Our tests indicate that SOAP performance dealing with raster data is improved in SOAP/MTOM. It is because SOAP/MTOM uses XOP optimizing mechanism so that serialization time and deserialization time is reduced remarkably. The larger data we test, the more gaps we get.

It also works when dealing with vector data. A standard SOAP message is made by embedding vector data as elements or attributes. Serialization/deserialization time of its SOAP message increases according to its number of attachments. Although total message size using SwA/MIME and SOAP/MTOM is bigger than standard SOAP because of tagging, roundtrip time is less than that. It is because SwA/MIME and SOAP/MTOM consist of small SOAP message and attachments exist in boundaries. Therefore, SOAP/MTOM's performance is almost equals in small Vector data, but in large amount of vector data it is better than using standard SOAP and SwA/MIME.

Our future research is to find how to use SOAP/MTOM with AJAX(Asynchronous JavaScript and XML). Current GIS services serve a map image by using AJAX, which is an asynchronous transfer method. We expect that combining web services with AJAX can improve the performance of GIS Web services remarkably.

REFERENCES

Alameh, N., 2003, Chaining Geographic Information Web Services, *IEEE Internet Computing*, pp. 22-29

Barton, et al., 2000, *SOAP Messages with Attachments*, December, 2000 (on-line) http://www.w3c.org/TR/soap-attachments/

Gudgin, M., et al., 2005, *SOAP Message Transmission Optimzation Mechanism*, (on-line) http://www.w3.org/TR/soap12-mtom/

Mitra, N., 2003, *SOAP Version 1.2 Part 0:Primer*, (on-line) http://www.w3.org/TR/2003/soap12-part0/

Ng, A., et al., 2005, A Study of the Impact of Compression and Binary Encoding on SOAP performance, Department of Computing, Macquarie University, Australia

Sayar, A., et al., 2006, Integrating AJAX Approach into GIS Visualization Web Services, Community Grids Lab, Indiana University, Bloomington, Indiana

Ying, Y., et al., 2004, A performance Evaluation of Using SOAP with Attachments for e-Science, School of Computer Science, Cardiff University