The new ISO 20022 financial messaging standard and its encodings

ISO 20022

ISO 20022 is an international standard, published by the International Organization for Standardization (ISO), that aims to enhance communication interoperability between financial institutions, their market infrastructures, and their end-user communities.

ISO 20022 defines a unified standardization approach, which consists of a common development methodology, a common process, and a common repository that are open for use by all financial standards initiatives.

This effort is motivated by the belief that we would all benefit from huge cost savings if we could use a single message standard for all our financial communications. All financial standards initiatives will be able to take advantage of this new approach to the standardization of financial messages. Standards developers will be able to avoid the duplication of efforts, and the developers of financial applications will no longer need to support many different overlapping standards.

At the core of the ISO 20022 philosophy is the idea of *syntax-independent business modeling*. It requires that all the relevant aspects of the business (actors, processes, information exchanges, etc.) as well as all data objects and messages be described using abstract *models*.

One of the advantages of the *syntax-independent business modeling* approach is that as new encoding technologies emerge, it will be possible to use them to optimize message exchanges without having to modify the abstract models that describe the messages.

Another advantage of this approach is that it provides a common way to describe messages belonging to disparate financial message standards. The new edition of ISO 20022 adds extra metamodel features that better support standards such as FIX and FpML. Once a set of messages that exist in two or more financial standards have been modeled and registered as ISO 20022 messages, it will be easier for software vendors to create adapters that translate messages in real time between any two of those standards, or between any one of them and ISO 20022. This will facilitate the coexistence of multiple financial standards in the short run, and the convergence towards ISO 20022 in the long run.

Each financial institution has its own set of internal data objects supporting various business concepts. One of the goals of ISO 20022 is to identify and standardize the data objects that are shared between institutions and to store them in a data dictionary that is part of a common Repository managed by a Registration Authority. Using those standard data objects as building blocks, a community of users or an organization can develop ISO 20022 syntax-independent message models and submit them for registration into the ISO 20022 Repository. A registration process is specified on the ISO 20022 website and has been agreed to by ISO and the ISO 20022 Registration Authority, currently the Society for Worldwide Interbank Financial Telecommunication (SWIFT).

The first edition of ISO 20022 (2004) included rules for automatically generating an XML schema from a message model. In the new edition of the standard, analogous rules for automatically generating an ASN.1 schema (see below) were added. Therefore both XML Schema and ASN.1 (see below) are now directly supported by ISO 20022. When an organization submits a set of messages for registration into the ISO 20022 Repository, a complex registration procedure is initiated. As part of that procedure, message models are automatically transformed into message definitions in XML Schema as well as in ASN.1, and both the message models and the generated message definitions are added to the Repository. The content of the ISO 20022 Repository is published at http://www.iso20022.org.

ASN.1

ASN.1 (Abstract Syntax Notation One) is a data specification and encoding technology jointly standardized by ISO, IEC (International Electrotechnical Commission), and ITU (International Telecommunication Union), and widely used across several industries (cellular telephony, signaling, network management, Directory, Public Key Infrastructure, videoconferencing, aeronautics, Intelligent Transportation, and so on).

ASN.1 is:

- a formal notation for specifying the logical structure of data that is to be exchanged between two endpoints; and
- several standard sets of encoding rules for encoding data whose logical structure is specified in ASN.1 notation.

A message type specified in ASN.1 notation is independent of hardware platform, operating system, programming language, and local representation. The ASN.1 notation is used to create ASN.1 schemas, which are text files containing the definition of one or more message types of arbitrary complexity. An instance of a message type is encoded using one of the standard sets of ASN.1 encoding rules (DER, PER, OER, XER, etc.) for the purpose of transmission.

In ASN.1, the description of the logical structure of a message is separate from the details of the encoding. As a result, protocol designers can focus on specifying the essential properties of the data to be exchanged without being distracted by the details of the encoding. Message schemas written in ASN.1 are concise, since they describe only the logical structure of the data. The reader of an ASN.1 schema can quickly grasp the structure of the data and the aspects that are most relevant to the logic of the protocol. Since the encodings of ASN.1 are standardized, most users of ASN.1 do not need to be concerned with the exact bits on the wire produced by the encoding rules in use. Most of the times a user will simply trust the encoder/decoder provided by the chosen ASN.1 development tool to produce and decode the bits on the wire correctly.

Given an ASN.1 schema specifying the messages of a protocol, ASN.1 development tools can automatically generate source code, encoder/decoders, and other artifacts that will facilitate implementation and testing. Many general-purpose ASN.1 tools, both commercial and free, are available for several operating systems and programming languages. The ITU website contains a fairly

comprehensive section about ASN.1 at http://www.itu.int/ITU-T/asn1/introduction/index.htm, including a list of ASN.1 tools.

The various standard sets of encoding rules of ASN.1 have different characteristics. The encoded messages they produce can be more or less compact, more or less self-descriptive, and more or less time-efficient. Usually, a protocol designer will indicate which encoding rules must be used for a particular ASN.1 schema. In the case of an ASN.1 schema generated from an ISO 20022 message model, the ISO 20022 standard recommends the use of the Packed Encoding Rules (PER), aligned variant, but other encoding rules may be used.

PER is recommended by ISO 20022 because it is very compact and fast. When used to encode an ISO 20022 message, PER achieves optimal utilization of the available bandwidth and of the computing resources at both ends of the communication. The Octet Encoding Rules (OER) are usually less compact but faster than PER. The Basic Encoding Rules (BER) and the Distinguished Encoding Rules (DER) are also less compact than PER but their encoded messages can often be understood, to some extent, even by someone who has an outdated or incomplete version of the underlying ASN.1 schema.

BER and DER are used in a variety of applications such as public key infrastructure and smart card standards. PER is mostly used in telecommunications applications. OER is mostly used in automotive applications.

The Extended XML Encoding Rules (E-XER) produce XML encodings that closely match those described by an XML schema. When one generates both an XML schema and an ASN.1 schema from an ISO 20022 message model by following the mapping procedures specified in ISO 20022, any valid XML schema instance will also be a valid E-XER encoding of the ASN.1 schema, and vice versa. ISO 20022 applications based on ASN.1 will be able to take advantage of this fact by using PER, OER, or DER to efficiently exchange binary messages with other ISO 20022 applications that understand ASN.1, and by using E-XER to encode and decode XML messages when communicating with ISO 20022 applications that do not support ASN.1.

The use of the binary encoding rules of ASN.1 (PER, OER, and DER) will support applications in which bandwidth utilization and encoding/decoding speed are critical, such as new card standards, the use of mobile phones as payment devices, embedded devices, and low-latency and high-frequency securities, derivatives, and commodities trading.

Performance Tests

OSS Nokalva conducted a series of performance tests to measure the relative performance of the ASN.1 encodings and of the XML encoding of ISO 20022 messages, using our ASN.1/C Tools for the ASN.1 encodings and the gSOAP toolkit for the XML encoding. We used a set of ISO 20022 message instances from the ISO 20022 Repository. The message instances belong to the following business areas:

- acmt – Account Management – 20 message instances
- caaa – Acceptor to Acquirer Card Transactions – 15 message instances
- camt – Cash Management – 106 message instances

- pacs – Payment Clearing and Settlement – 18 message instances
- pain – Payment Initiation – 7 message instances
- catm – Terminal Management – 3 message instances

For all the above business areas except the last, we used all the message instances that were present in the ISO 20022 Repository on March 24, 2012. For the Terminal Management business area, we used the message instances that were present in the Repository on December 11, 2012.

We first generated an ASN.1 schema from each of the registered message models by following the standard mapping procedure specified in Part 8 of the new edition of ISO 20022, with minor modifications due to differences in the UML metamodel between the old and the new edition of ISO 20022.

We then measured the time taken to encode and decode each message instance, using each of the following encoding rules:

- ASN.1 OER (Octet Encoding Rules)
- ASN.1 DER (Distinguished Encoding Rules)
- ASN.1 APER (Packed Encoding Rules, aligned variant)
- ASN.1 UPER (Packed Encoding Rules, unaligned variant)
- ASN.1 E-XER (Extended XML Encoding Rules)
- XML as per the XML Schema registered in the ISO 20022 Repository
- XML as above, compressed using ZLIB (compression level 1)
- XML as above, compressed using ZLIB (compression level 6)
- XML as above, compressed using ZLIB (compression level 9)

For all the ASN.1 encodings except OER (i.e., DER, APER, UPER, and E-XER), we used the Time-Optimized Encoder/Decoder (TOED) that is part of OSS Nokalva's ASN.1/C Tools, version 9.0.3.1. For OER, we used a pre-release version of the ASN.1/C Tools. For the XML encodings, we used the open-source tool gSOAP, version 2.8. We chose gSOAP over other popular software tools for XML because it supports C/C++ data bindings—an encoding/decoding interface style similar to that used by our ASN.1/C tools.

We ran all the tests on an Intel Core i5-2400 3.1 GHz desktop PC with 4 GB of RAM and Windows 7 SP1.

The results of the tests are shown in the following figures.

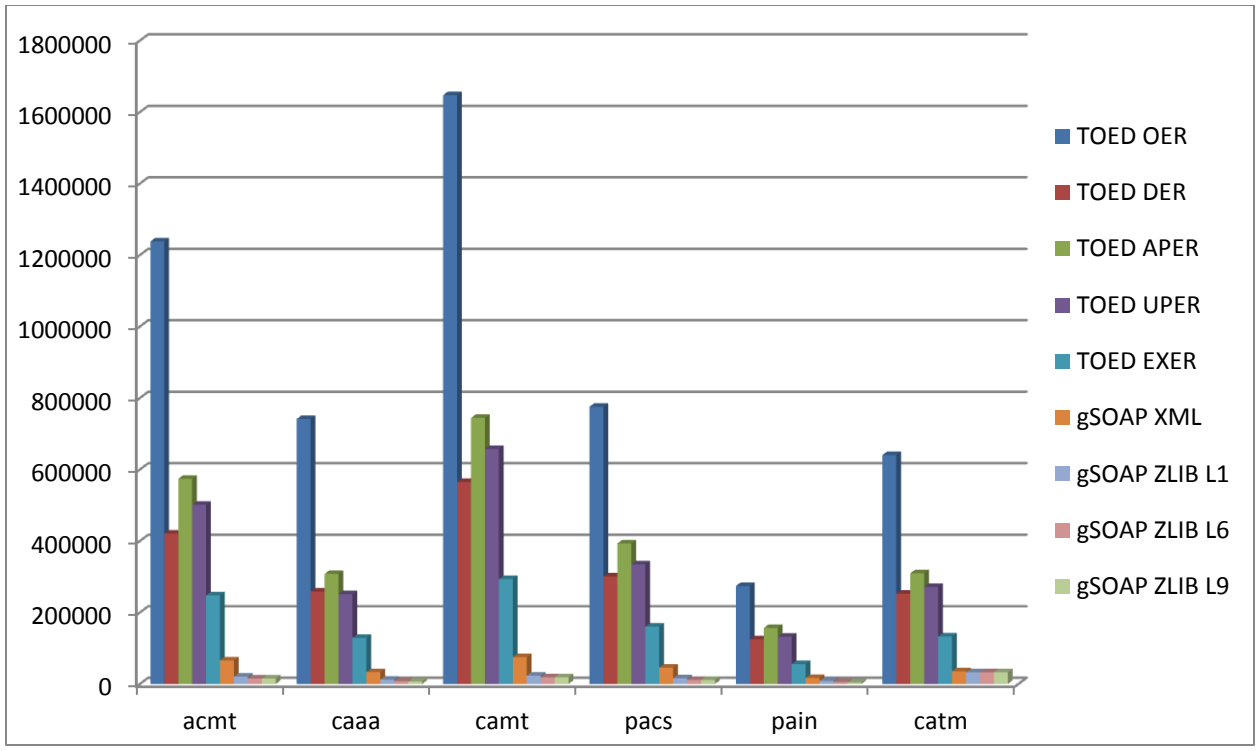

Figure 1 – Number of encoding operations per second

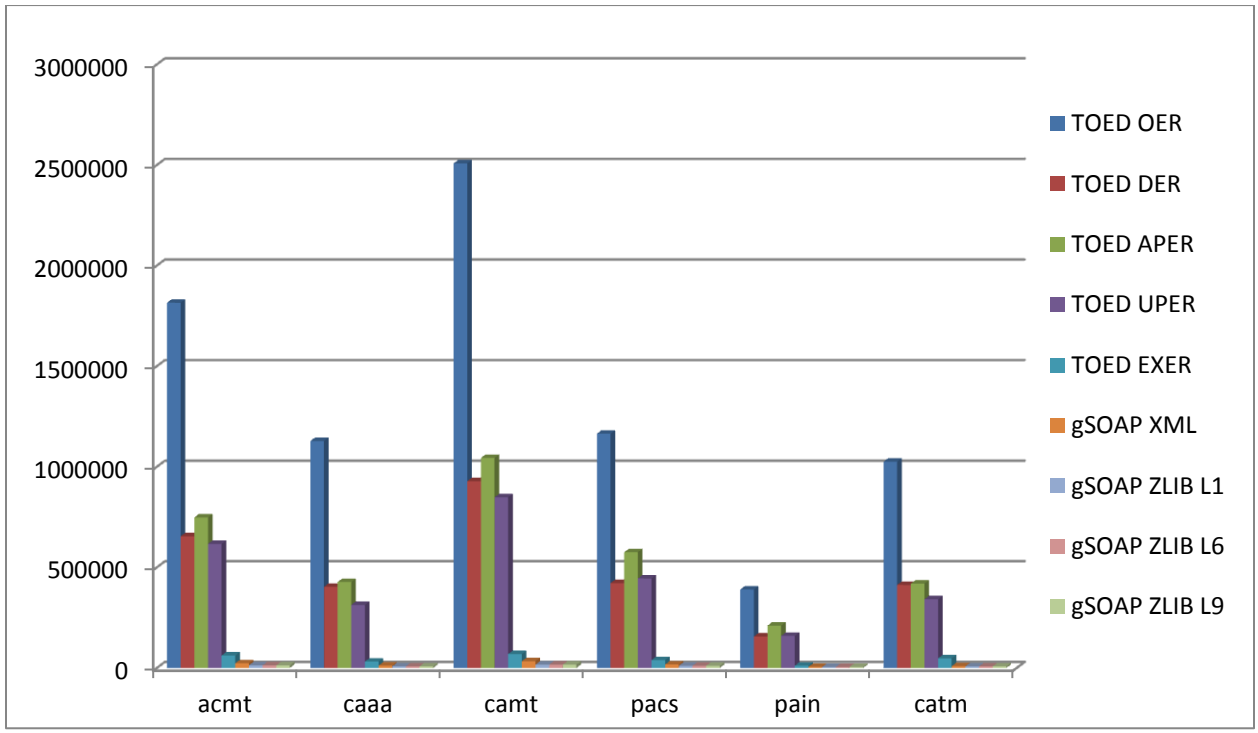

Figure 2 – Number of decoding operations per second

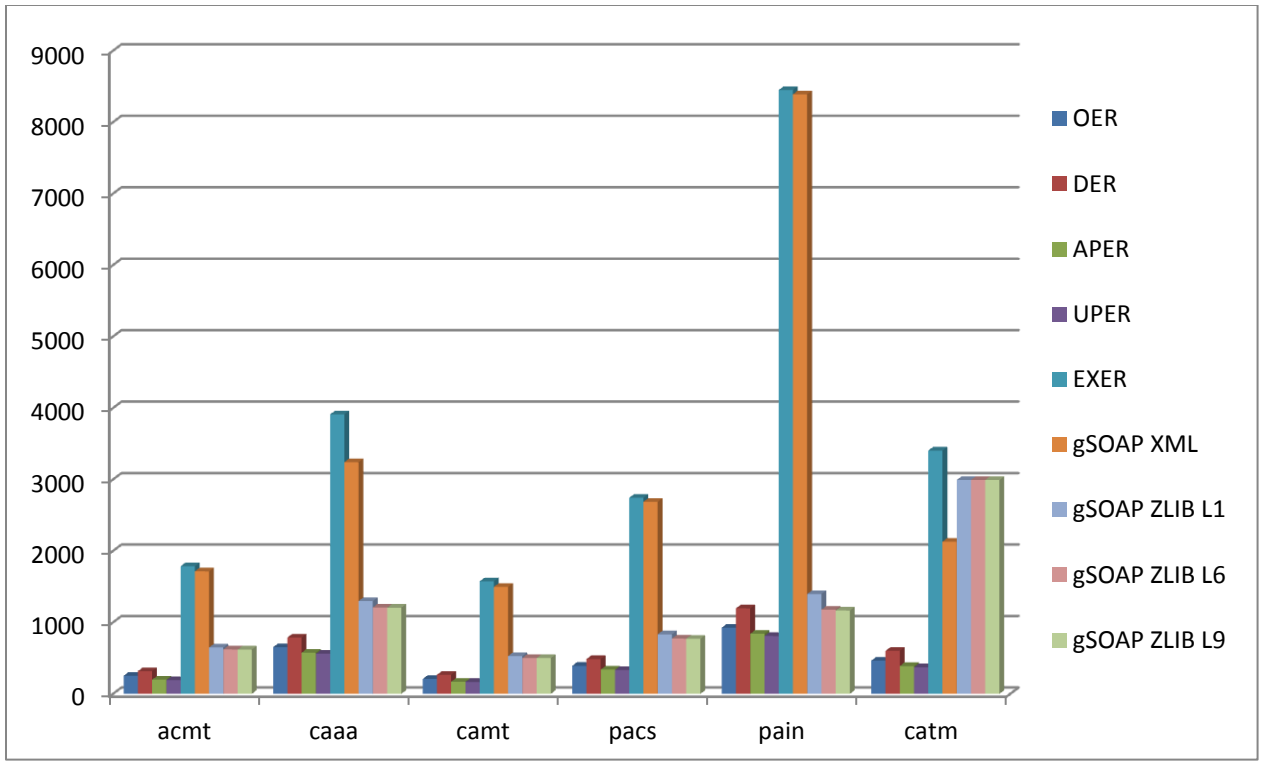

Figure 3 – Size of the encoded message (bytes)

Figure 1 shows the results of the tests for *encoding* operations. On the vertical axis is the number of encoding operations per second. The height of each column is the harmonic mean of the measured number of *encoding* operations per second calculated over all the message instances belonging to a given business area, for a given encoding.

Figure 2 shows the results of the tests for *decoding* operations. On the vertical axis is the number of decoding operations per second. The height of each column is the harmonic mean of the measured number of *decoding* operations per second calculated over all the message instances belonging to a given business area, for a given encoding.

Figure 3 shows the size of the encoded messages. On the horizontal axis is the size of an encoded message in bytes. The length of each bar is the arithmetic mean of the size of the encoded message calculated over all the message instances belonging to a given business area, for a given encoding.

Our test results show that the users of ISO 20022 who choose to encode their messages in an ASN.1 binary encoding can experience a significant increase in performance compared to the use of XML.

The increase in *encoding* speed over XML was a factor of 6-8 for DER, a factor of 7-9 for PER Unaligned, a factor of 9-10 for PER Aligned, and a factor of 16-22 for OER. The increase in *decoding* speed over XML was a factor of 24-45 for DER, a factor of 25-37 for PER Unaligned, a factor of 31-45 for PER Aligned, and a factor of 66-111 for OER. The encoded messages are also significantly smaller (4 to 10 times smaller) in an ASN.1 binary encoding than in XML.

Our results also show that the use of ZLIB compression associated with the XML encoding in the gSOAP tool have a significant impact on encoding and decoding speed but the resulting message sizes are still larger than those that can be attained using an ASN.1 binary encoding.

The E-XER results are reported for completeness. The differences in encoding/decoding speed measured between the E-XER encoding and the gSOAP tool reflect differences in the design of the two tools and not the characteristics of the encoding, which is an XML encoding in both cases.