Why we made XZES40 Transformer.

Useful, Challenging, and Open Source

This project was chosen because it provided students the opportunity to create a real-world impact, solve an interesting problem, and contribute to the Free, Libre, and Open Source Community (FLOSS).

The software may be used by individual, organization, and enterprise users to XML perform document transformations. This is a necessary and time consuming task which can now be done faster and *in the cloud*. It proved technically challenging in that it required an understanding of C programming, in-memory caching, parallel computation, application networking, and extensive library usage. The final product was shared with the FLOSS community, giving both our client, the Apache Software Foundation, and any the rest of the world a tool which is useful, usable, and free (*in more than one way*).

Applications and uses

XML is a machine readable data markup language used to store any well-formed topic. XML Stylesheets are used to express transformations of a given XML document. These are used together to store large amounts of data, and "ask" the data a question like "How many employees in this XML file took less than three days of vacation, and make more than $50,000 per year?"

XML doument transformation is used for largely utilitarian purposes in a wide veriety of fields. Police departments, corporate offices, academic institutions, and countless other organizations need and use XML documents and transformations.

The XZES40 application provides an easy to use and access platform for transforming such documents over the web.

CS CAPSTONE PROJECT: XZES40 TRANSFORMER

High Performance XML/XSLT Transformation Server

Optimized for Speed and Convenience

The XZES40 Transformer takes two files as input, an XML document and XSLT stylesheet, and a variable number of custom build parameters or dependency files, and generates a new XML document. While this type of application is not new the concept has been improved with key optimizations and made accessable over the internet via a Web interface.

The strengths of this system are are that it does not need to be installed locally on a user's system, but is still able to perform it's tasks in a timely manner.

The application uses the Apache Xalan and Apache Xerces C++ libraries to perform XML document transformations and a Python CGI script running through an Apache HTTPD server to respond to incomming requests. The transformer daemon is optimized to perform fast transformations by caching previously encountered XML documents and performing document transformations in parallel when possible. The Python CGI script communicates with the Daemon over a local network socket; multiple requests are sent to the daemon and each one is serviced in a POSIX thread. Once the transformation is complete it is sent back to the user.

The web service runs on a remote server atop existing techologies so it can be run, managed, and modified by anybody with Linux and Apache HTTPD experience.

Conclusions/Outcomes

The developed application can generate a new XML file given an XML input, Stylesheet input, and custom build parameters via a Web interface, all without error.

The minimium requirements were completed, like transforming XML documents, as well as many of the optimizations and benefits we set out to complete. The most directl benefit of these being a web interface to access the transformer over the internet. In addition to this we successfully completed a simple in-memory caching system and parallel transformation of the documents. In the end we gave users an application fast enough to be competitive with enterprise software, while still being convenient to usable via the browser.

Results

The following is a table which represents the benchmark results of transforming 19 files 3 times each on the Altova enterprise XML transformer, Xalan Open Soruce transformer, and our XZES40 web Transformer. The files tested came from the Xalan project source code repository.

filename	XZES40	Xalan	Altova
builddocs.xml	0.109s	0.204s	0.187s
buildlibs.xml	0.109s	0.214s	0.193s
charter.xml	0.110s	0.221s	0.190s
ommandline.xml	0.110s	0.201s	0.178s
download.xml	0.113s	0.205s	0.169s
extensions.xml	0.108s	0.196s	0.177s
faq.xml	0.108s	0.188s	0.165s
getstarted.xml	0.108s	0.193s	0.162s
index.xml	0.107s	0.189s	0.151s
Install_save.xml	0.108s	0.212s	0.162s
install.xml	0.110s	0.208s	0.147s
overview.xml	0.110s	0.179s	0.155s
programming.xml	0.108s	0.169s	0.143s
samples.xml	0.106s	0.177s	0.135s
secureweb.xml	0.105s	0.199s	0.139s
test-faqs.xml	0.109s	0.185s	0.136s
usagepatterns.xml	0.121s	0.205s	0.147s
average	0.110s	0.196s	0.161s

Average XML Transformation Speeds
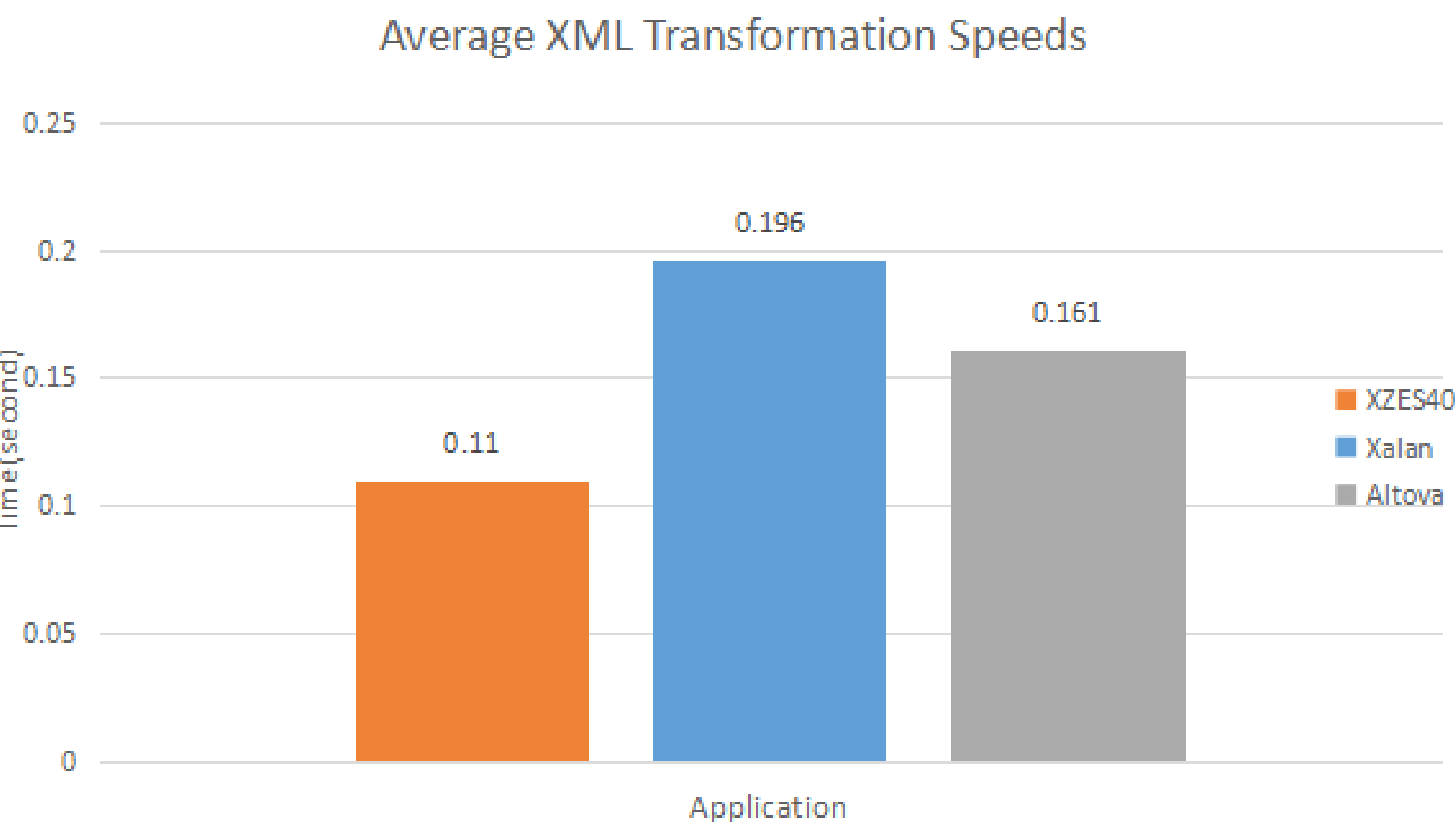
The team behind XZES40 Transformer.

XZES40 Transformer

Choose an XML file, XML stylesheet, and [optional] custom parameters. Click "Transform Files" to download your transformed documents.

Upload your XML file: Choose File No file chosen
Upload your XSLT file: Choose File No file chosen

Add XML Parameter   Transform Files

Made in association with the Oregon State University Capstone program and the Apache Software Foundation.

For more information please visit github.com/xzes/xzes40-transformer

Sponsor

Steven Hathaway
    The Apache Software Foundation
    Email: shathaway@apache.org

Team Members

Elijah C. Voigt
    Email: voigte@oregonstate.edu

Shuai Peng
    Email: pengs@oregonstate.edu

Zixun Lu
    Email: luzi@oregonstate.edu

About the Team

Elijah, Shuai, and Zixun are are fourth year Computer Science students at Oregon State University. Elijah is an Applied CS student focusing on mathematics and computer security, hoping for a career in systems engineering and games development. Shuai is a Systems CS student and is passionate about understanding and designing complex systems, looking for a career in software engineering. Zixun is a double major in CS and Business, intending to be an entrepreneur.

The team also thanks and appreciates Steven Hathaway who gave much of his time and patience to the project.