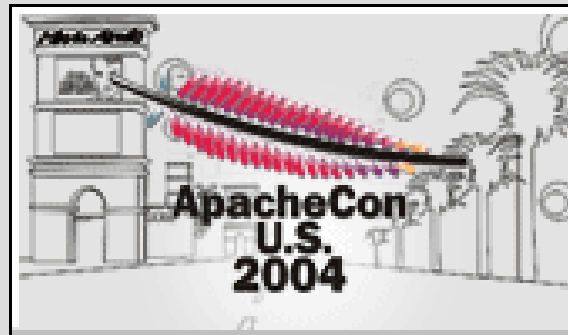


Using XSL and mod_transform in Apache Applications

Paul Querna
chip@OutOfOrder.cc



What is XSL?

- Extensible Stylesheet Language (XSL)
- A family of Standards for XML by the W3C:
 - XSL Transformations (XSLT)
 - XML Path Language (XPath)
 - XSL Formatting Objects (XSL-FO)

XSLT Example

```
<?xml version="1.0"?>
<xsl:stylesheet
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  version="1.0">

<xsl:template match="/">
  <html>
    <head><title>A Message</title></head>
    <body>
      <h1>
        <xsl:value-of select="message" />
      </h1>
    </body>
  </html>
</xsl:template>
</xsl:stylesheet>
```

Data Source...

```
<?xml version="1.0"?>  
<message>Hello World</message>
```

Outputs...

```
<html>
<head>
<meta http-equiv="Content-Type"
      content="text/html; charset=UTF-8">
<title>A Message</title>
</head>
<body>
<h1>Hello World</h1>
</body>
</html>
```

Why is XSLT good?

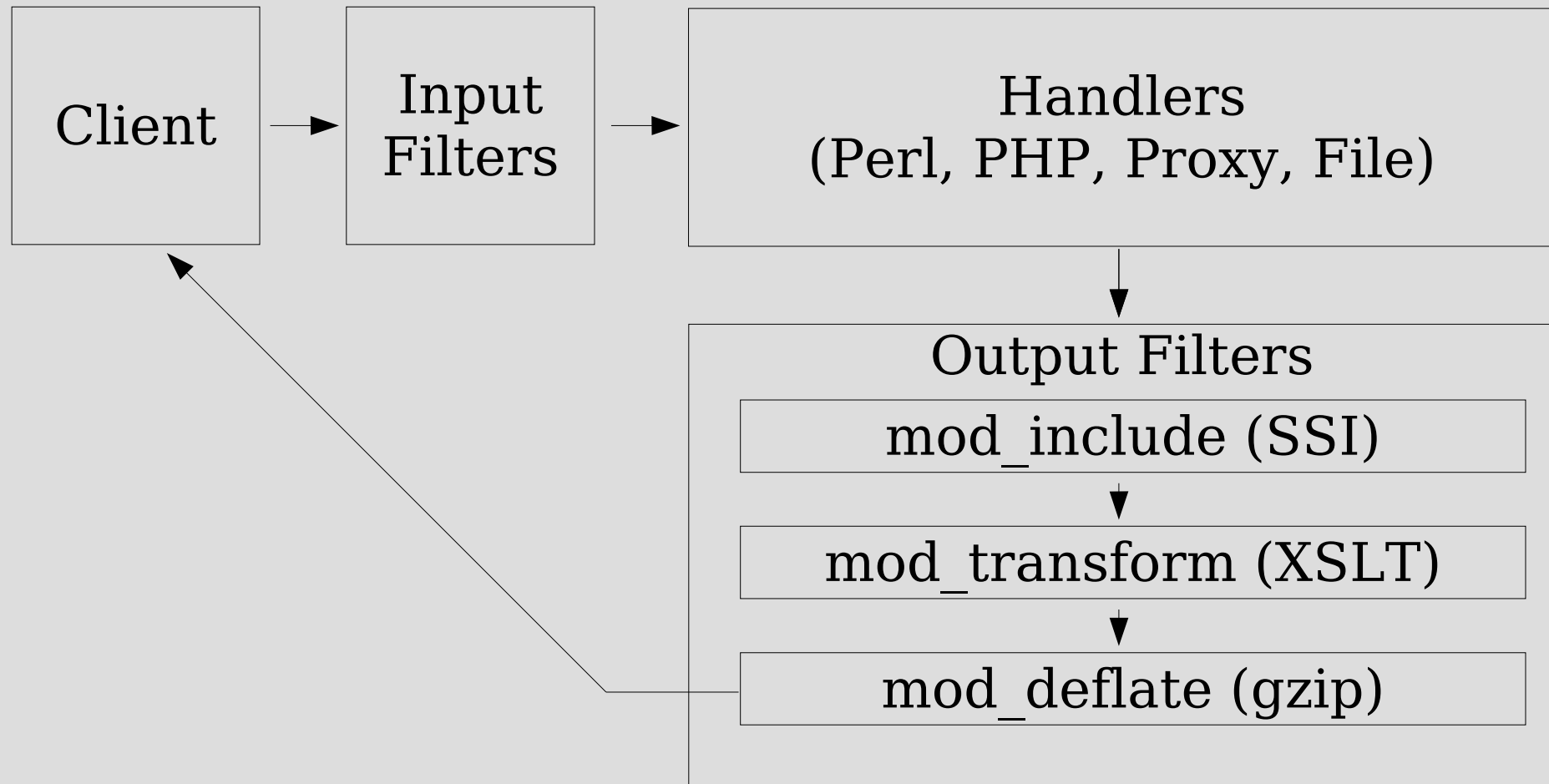
- Mixing Data and Presentation is **bad!**
 - Keeps Data in a clean XML schema
 - Keeps the Presentation of this Data separate
- XSLT is XML
- Easy to Extend
- Put HTML or other Markups directly in the XSLT.
 - Easy for Web Developers to create a template

Why is XSLT bad?

- XSLT is XML
- Complicated XSLT can be slow
- Yet another language to learn

Where does Apache fit in this?

- Apache 2.0 has Filters!



mod_transform

- Uses libXML2 and libXSLT from Gnome
 - C API
 - Doesn't depend on other Gnome Libs.
 - Provides:
 - EXSLT
 - XInclude
 - XPath
 - Xpointer
 - ... and more

Static XML Files

- AddOutputFilter XSLT .xml
- TransformSet /xsl/foo.xsl
 - Only if your XML does not specify a XSL File
- TransformOptions +ApacheFS
 - Uses Sub-Requests to find files
 - Makes mod_transform work like Apache AxKit

Dynamic Sources

- XML Content Types:
 - AddOutputFilterByType XSLT application/xml
- Controlled Content Types:
 - AddOutputFilterByType XSLT applicain/needs-xslt
- Works for Proxied Content, PHP, mod_perl, mod_python, CGI, SSI, etc..
 - mod_filter
 - In 2.1 CVS!

PHP

- Content Type
- With a Patch
 - PHP can manually add output filters
 - `apache_add_output_filter("XSLT");`
 - List Active Filters
 - `apache_get_output_filters();`

PHP Example

```
<?php
    apache_add_output_filter("XSLT");

    echo "<?xml version='1.0'?>\n";
    echo "<?xml-stylesheet type='text/xsl'
href='./demo.xsl'?>\n";
    echo "<demo>\n";

    $filters = apache_get_output_filters();

    foreach($filters as $filter) {
        echo "<filter>$t_val</filter>\n";
    }

    echo '</demo>';
?>
```

Apache Module

- Direct:
 - `ap_add_output_filter("XSLT", NULL, r, r->connection);`
- By Content Type:
 - `ap_set_content_type(r,"application/xml");`
- Passing a libXML2 Document Tree:
 - `mod_transform_XSLTDoc(r, doc) ;`

mod_svn_view

- Uses mod_transform
- Provides a web based view using the Native Subversion C libraries

SVN View

[Repo List] - [gentoo]/ apache-overlay

Property Value

[no properties set]

	Filename	Revision	Username	Message
[dir]	profiles	321	Hollow	* changed apache-mod...
[dir]	scripts	336	beu	* new tool: - simple...
[dir]	docs	269	stuart	Added web-handbook a...
[dir]	dev-util	366	trapni	renamed subversion 1...
[dir]	dev-php	309	beu	update manifest.
[dir]	dist	369	vericgar	Fixing default confi...
[dir]	www-apache	367	trapni	version bump
[dir]	dev-libs	356	chip	APR-1.0 ebuilds now ...
[dir]	net-www	368	Hollow	version bump
[dir]	eclass	358	urilith	Cleaning up the ecla...
[dir]	irc	347	Hollow	* assigned more spec...
[dir]	dev-perl	258	trapni	did an s,extramodule...
[file]	ChangeLog	159	mtindal	Final commit for tre...
[file]	skel.apache1.ebuild	337	Hollow	* many many mod_* eb...

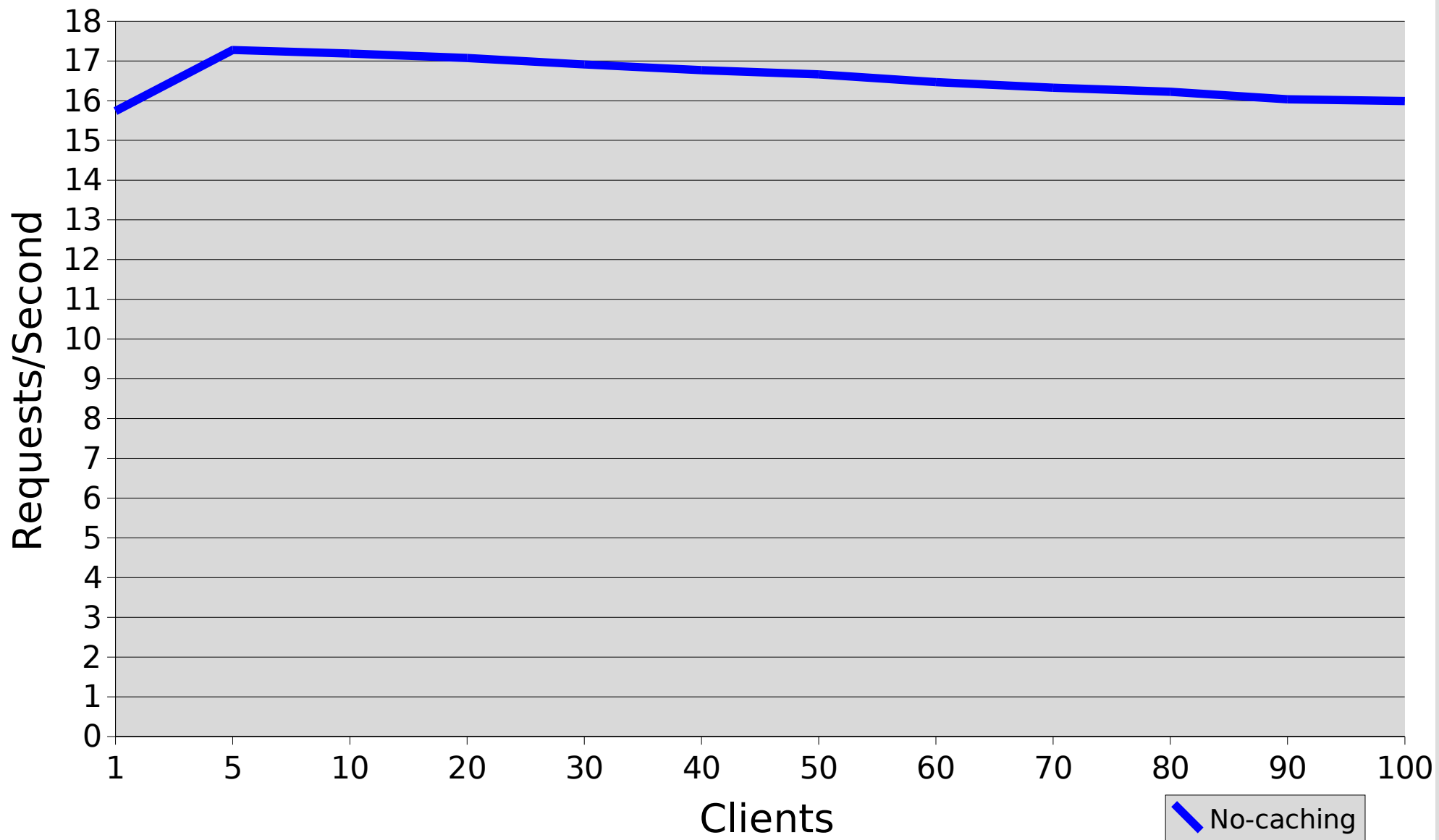
Performance Issues

- Parsing the XSLT with libXSLT is slow
 - mod_transform can cache the XSL Files
- Parsing the XML to a libxml2 is slow
 - Applications can pass a docPtr in memory
- Doing the Transform itself is slow
 - mod_cache can capture the output

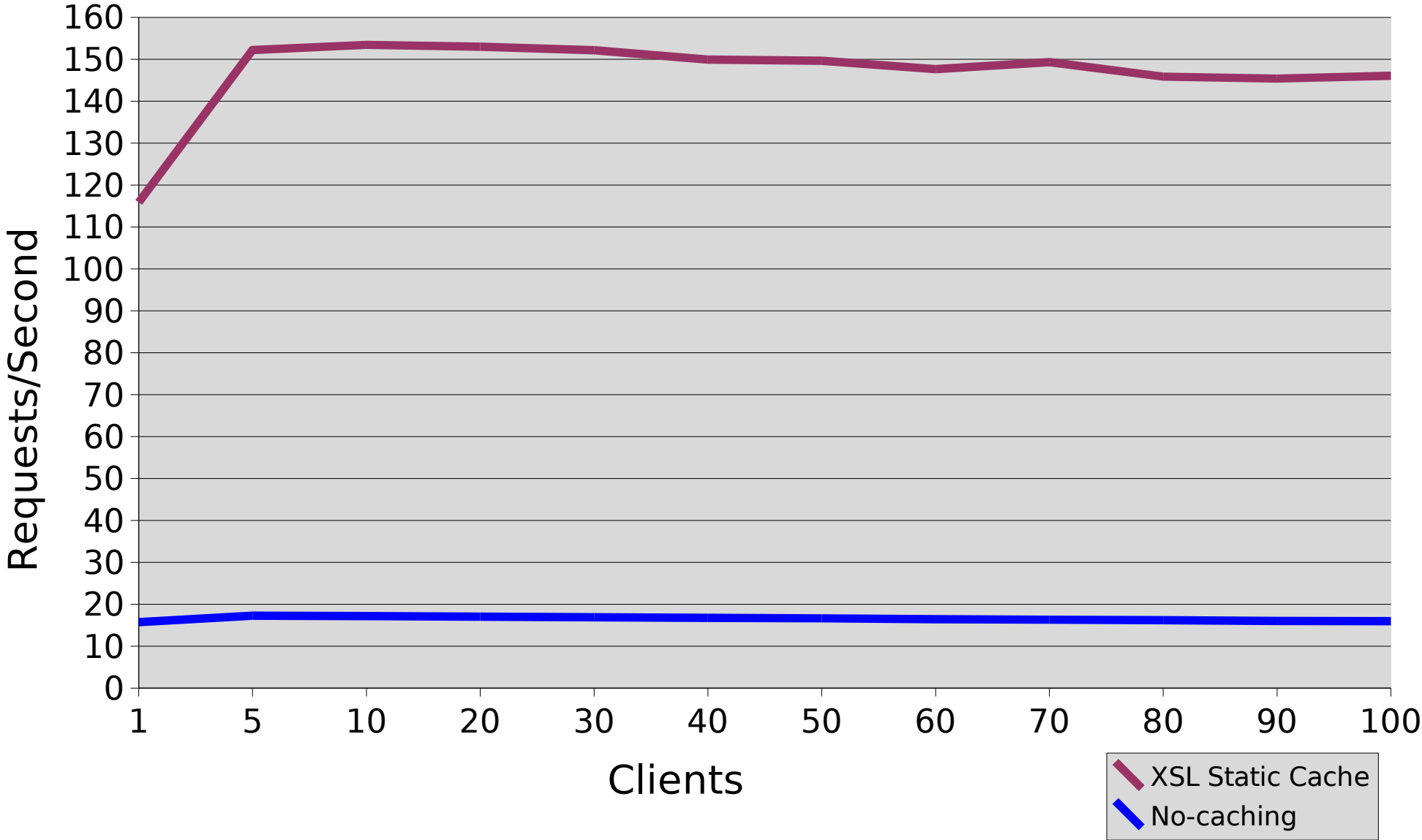
Benchmarks!

- Your Millage May Vary
 - Test to you environment/hardware/software

No Caching



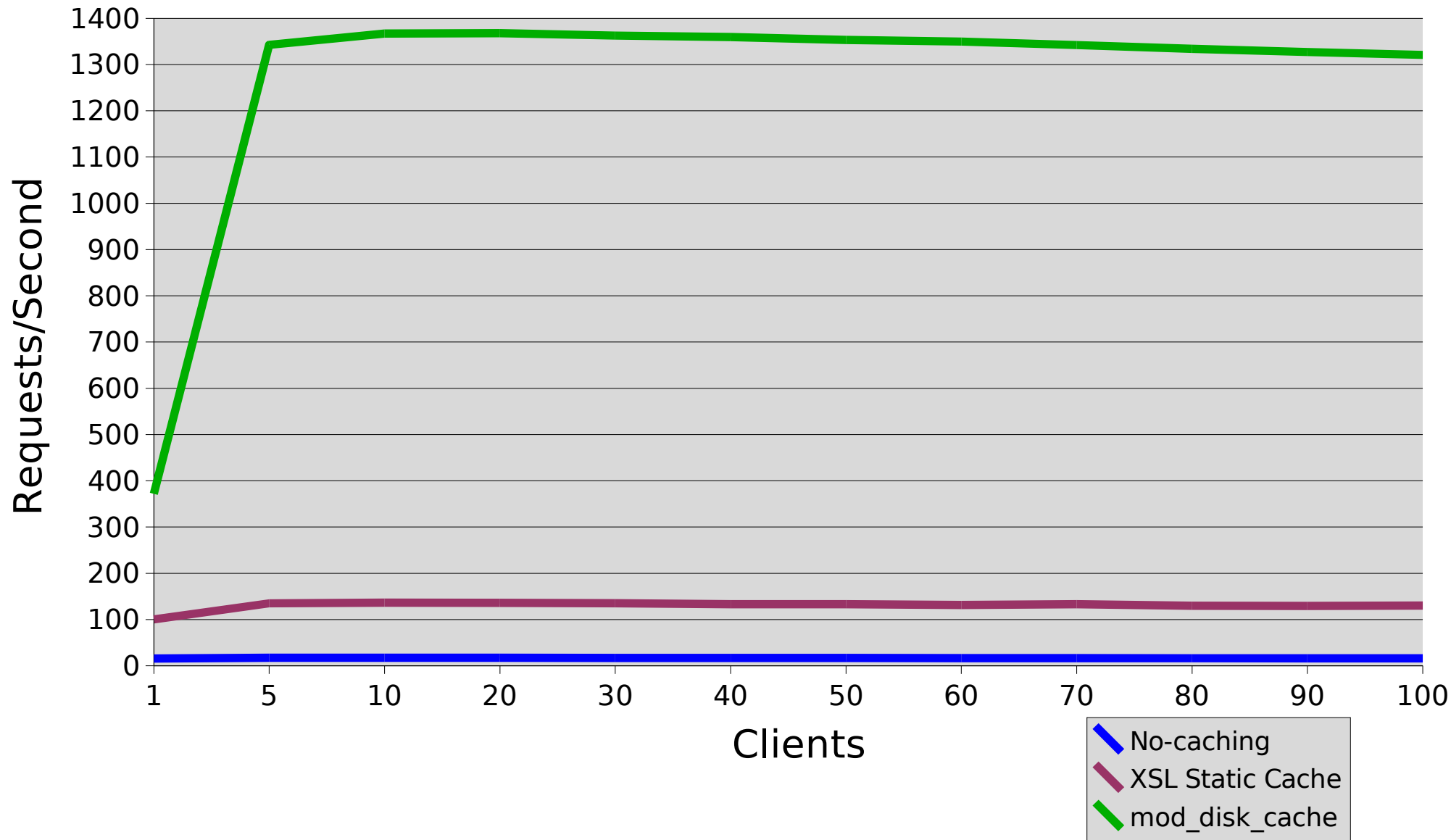
XSL Static Cache



XSL Static Cache Issues

- Does not see changed XSL Files
 - Not a big issue on production servers
 - Only sees changes on an Apache Restart
- 'Lowest Hanging Fruit' to improve performance

mod_cache



mod_cache Issues

- mod_cache:
 - Doesn't have any dependency checks
 - Changed content is not seen!
 - Changed XML Files are not noticed!
 - Changed XSL Files are not noticed!
 - Can be fixed
 - Etag Generation
 - In the ~/TODO
 - Slower, but should still be faster than running the XSL

Alternatives

- mod_php / mod_perl / mod_python
 - Can use libXSLT
- mod_ext_filter
 - Use an external Process
 - Performance?
- Apache AxKit
 - Apache 1.3 Only
 - Requires mod_perl
 - Cannot transform content from other Handlers

Alternatives

- mod_xmlns / mod_publisher
 - Still in development
 - Framework for a custom template system
- Modxslt2
 - Provides Extra Features
 - Access to Query Arguments and POST
 - ~10,000 LOC

Future Development

- Dependency Trees – mod_depends
 - Etags
 - LastModified
- Access to Query Arguments and POST
- Dynamic XSL Caching

Resources

- Slides, mod_transform, PHP Patch:
 - <http://www.outoforder.cc>
- LibXML2:
 - <http://xmlsoft.org/>
- LibXSLT:
 - <http://xmlsoft.org/XSLT/>
- Apache AxKit
 - <http://axkit.org/>

Resources

- mod_publisher / mod_xmlns
 - <http://apache.webthing.com>
 - http://apache.webthing.com/mod_publisher/
- Modxslt2
 - <http://www.mod-xslt2.com/>

Thanks!

- mod_transform co-authors:
 - Nick Kew
 - Edward Rudd

Questions?