



Linpro AS

The leading Norwegian provider of
Linux and OSS solutions and services



- Norway's leading provider of solutions and services based on open source software and Linux
- More than 10 years of experience in consulting, development, integration and IT operations
- Over 50 courses, including certification courses from Red Hat, Novell and MySQL

- Senior Software Developer
 - Multiframe, Varnish
 - SIG for C, C++, kernel development and embedded systems
- FreeBSD committer since 1998
 - OpenPAM, OpenSSH, libfetch, pseudofs...

HTTP acceleration in the 21st century

Dag-Erling Smørgrav
Senior Software Developer
Linpro AS

2007 eZ Conference

- **Introduction**
- A little bit of history
- Design and architecture
- Integration issues
- Pretty pictures
- Miscellaneous

- HTTP accelerator, aka. reverse proxy
 - Place in front of your CMS to improve response time and reduce server load
- Designed for modern computers, modern operating systems and modern workloads
- Very flexible configuration system
 - but the default settings are conservative and reasonable
- Open Source (two-clause BSD license)
- Runs on FreeBSD, Linux
- Jointly owned by Linpro and VG
 - an independent foundation is being set up

- Short answer: very
- Long answer: we don't really know, but:
 - < 20 system calls per request on average
 - $< 20 \mu\text{s}$ to handle a request on reasonable hardware
 - Reached 700 Mbps on a single server (dual Opteron blade with GigE)
 - Never been Slashdotted yet...

1000 req/s @ 100 Mbps

Sample	Min	Max	Median	Average	Std. dev.
Full	12.9	3001.0	16.2	71.1	338.5
90% fractile	12.9	26.0	15.9	16.3	1.7

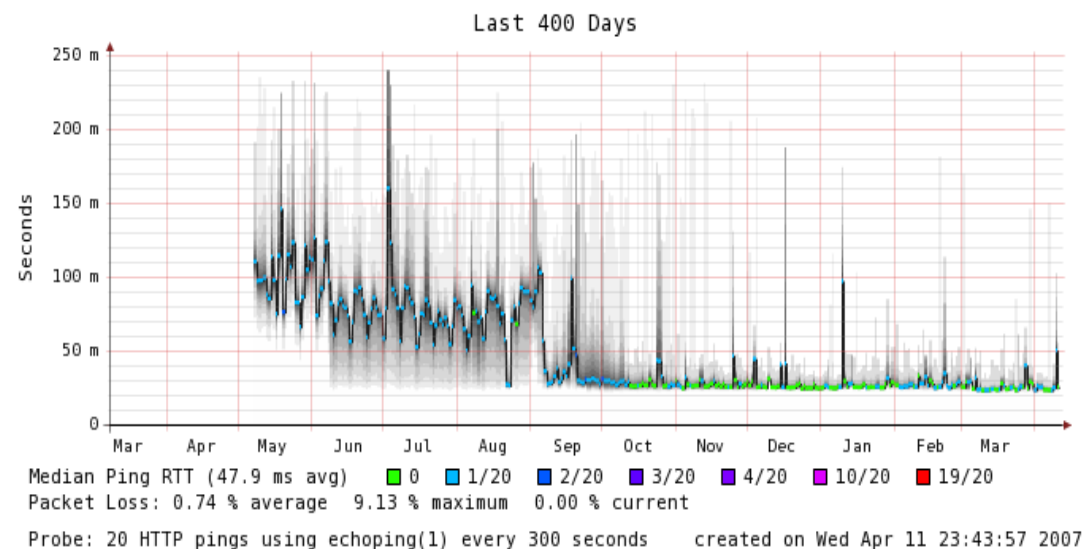
All numbers are in microseconds

- Introduction
- **A little bit of history**
- Design and architecture
- Integration issues
- Pretty pictures
- Miscellaneous

- Verdens Gang (vg.no) wanted to replace Squid
 - poor performance
 - development had stagnated
- Also wanted to “give something back” to the OSS community
- Linpro C / C++ SIG wanted a showcase project
- Poul-Henning Kamp wanted to show what “his” kernel (FreeBSD) can do when used correctly

- Fully financed by VG (approx. \$100k)
- Primary goal: to accelerate vg.no using significantly less hardware than with Squid
- Secondary goal: to lay the foundation for a successful OSS project
 - Varnish should not become stagnant like Squid
- (...or was it the other way around?)

- Response time (from reader's perspective) reduced from 50-150 ms to 20-30 ms
 - pretty much equal to round-trip time – the actual response time is too low to measure across the network...
- Number of proxy servers reduced from 13 to 2
 - one would be enough, but redundancy is good



- Financed by API and Escenic
 - API: 50 online newspapers, hosted by Linpro
 - Escenic: Norwegian Java-based CMS used by API
 - Additional sponsors welcome :-)
- Support for `<esi:include>`
 - The only open source ESI implementation!
- More power, more flexibility
- Protect against and recover from backend failure
- Load balancing
- Administration and monitoring tools

- Introduction
- A little bit of history
- **Design and architecture**
- Integration issues
- Pretty pictures
- Miscellaneous

- Designed primarily for 64-bit multi-core / multi-processor machines
 - Also supports 32-bit single-core systems
- Work *with* the kernel, not *against* it
 - Squid tries to second-guess the VM system
- Use every trick in the book
 - epoll, kqueue, accept filters...
- Minimize the amount of work done in the fast path
 - around 18 system calls per request in average

- Split-process model
 - Parent process monitors child
 - Child process does all the work
- Heavily threaded
 - One thread per client connection
 - Very low contention between threads
- Single memory-mapped cache file
 - Let the virtual memory system handle paging
- Log data stored in shared memory
 - Don't waste time writing logs, let others take care of it

- Cache policy encoded in VCL (Varnish Configuration Language)
 - Simple domain-specific language, vaguely similar to C or Perl
 - Inspect and modify request headers
 - Select backend
 - Inspect and modify response from backend
 - Much more...
- Also a few run-time parameters
 - min / max threads, idle timeouts etc.


```
backend default {  
    set backend.host = "127.0.0.1";  
    set backend.port = "90";  
}  
  
sub vcl_recv {  
    if (req.request == "POST" && req.http.content-length == 0) {  
        lookup;  
    }  
}  
  
sub vcl_fetch {  
    if (obj.ttl < 60s) {  
        set obj.ttl = 60s;  
    }  
}
```

- Introduction
- A little bit of history
- Design and architecture
- **Integration issues**
- Pretty pictures
- Miscellaneous

- Object lifetime
- Content negotiation
- Cookies

- How long the current version is valid
 - directly affects hit rate
- Several headers:
 - Expires
 - Cache-control: s-maxage
 - Cache-control: max-age
- Can be overridden in VCL
 - set `obj.ttl` in `vcl_fetch()`

- Language, Encoding, MIME type...
- Client sends Accept and / or Accept-Encoding, Server sends Vary
- Varnish 1.0.4 does not support Vary, but you can fake it in the development version:

```
sub vcl_hash
{
    if (req.http.Accept-Encoding ~ "gzip") {
        set req.hash += "gzip";
    } else if (req.http.Accept-Encoding ~ "deflate") {
        set req.hash += "deflate";
    }
}
```

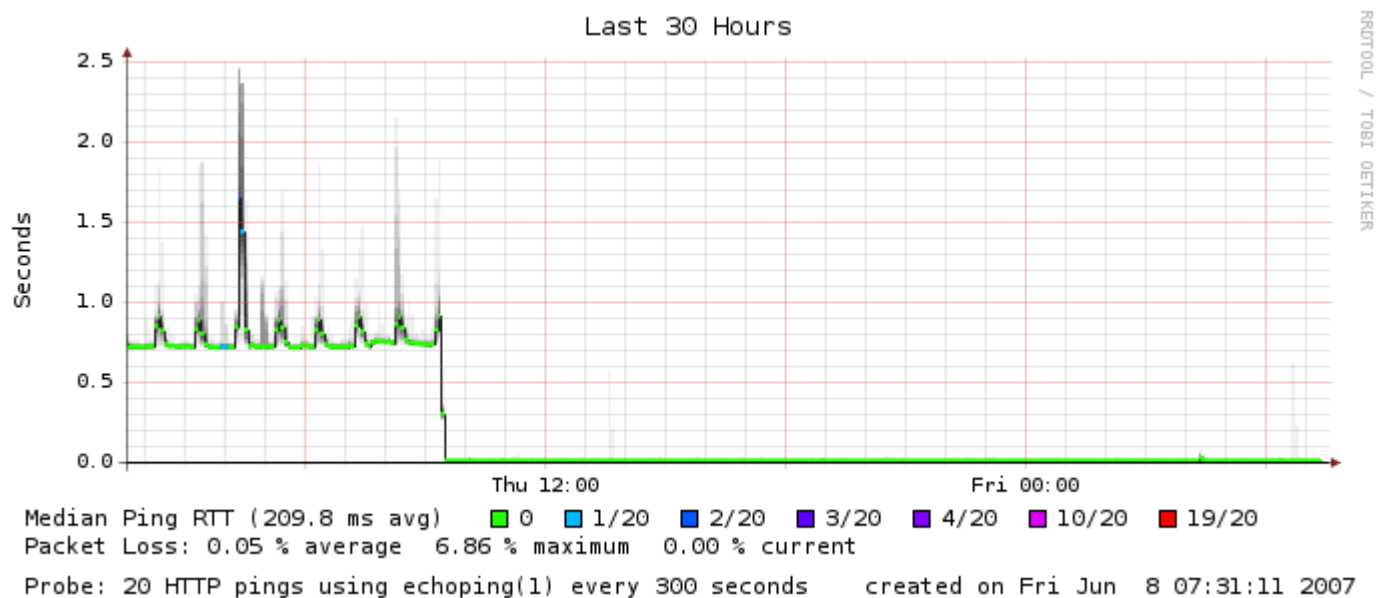
- Store data on the client
 - usually a session ID which refers to data stored on the server
- Often used unnecessarily
- Very rarely used correctly (with Vary)
- Workaround:
 - Evil: unconditionally strip Set-Cookie and Cookie headers
 - Better: strip Set-Cookie except on the login page, pass Cookie through

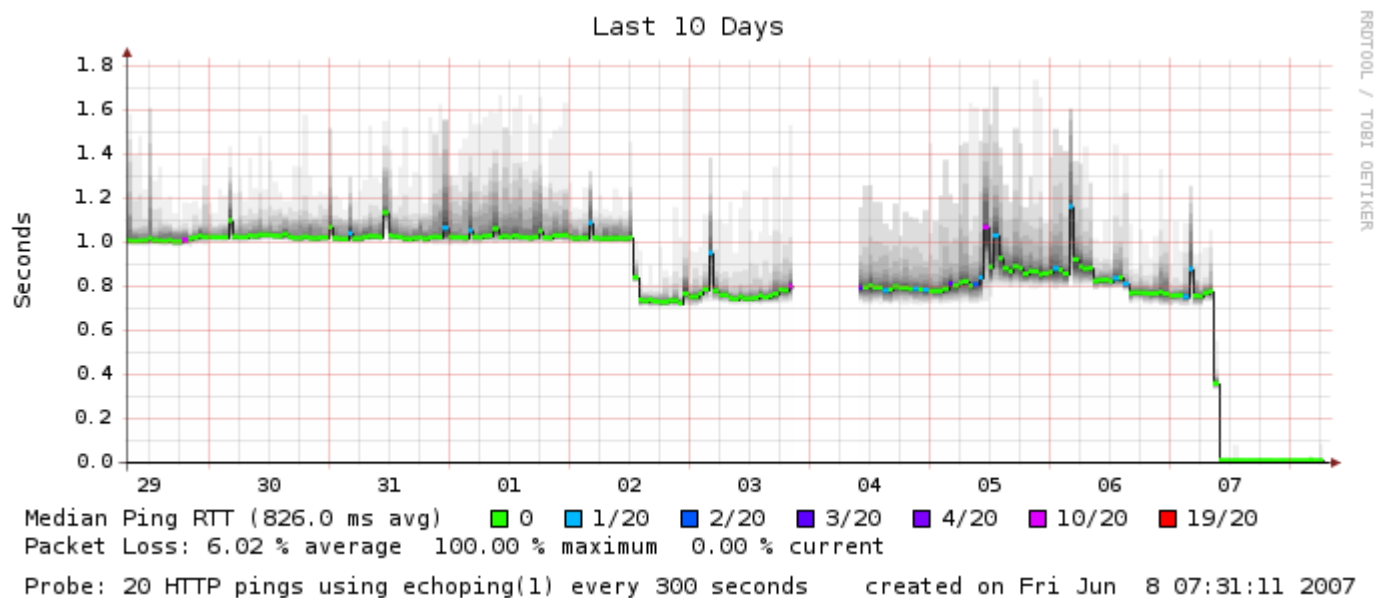
```
GET / HTTP/1.1
Host: ez.no
User-Agent: fetch libfetch/2.0
Connection: close
```

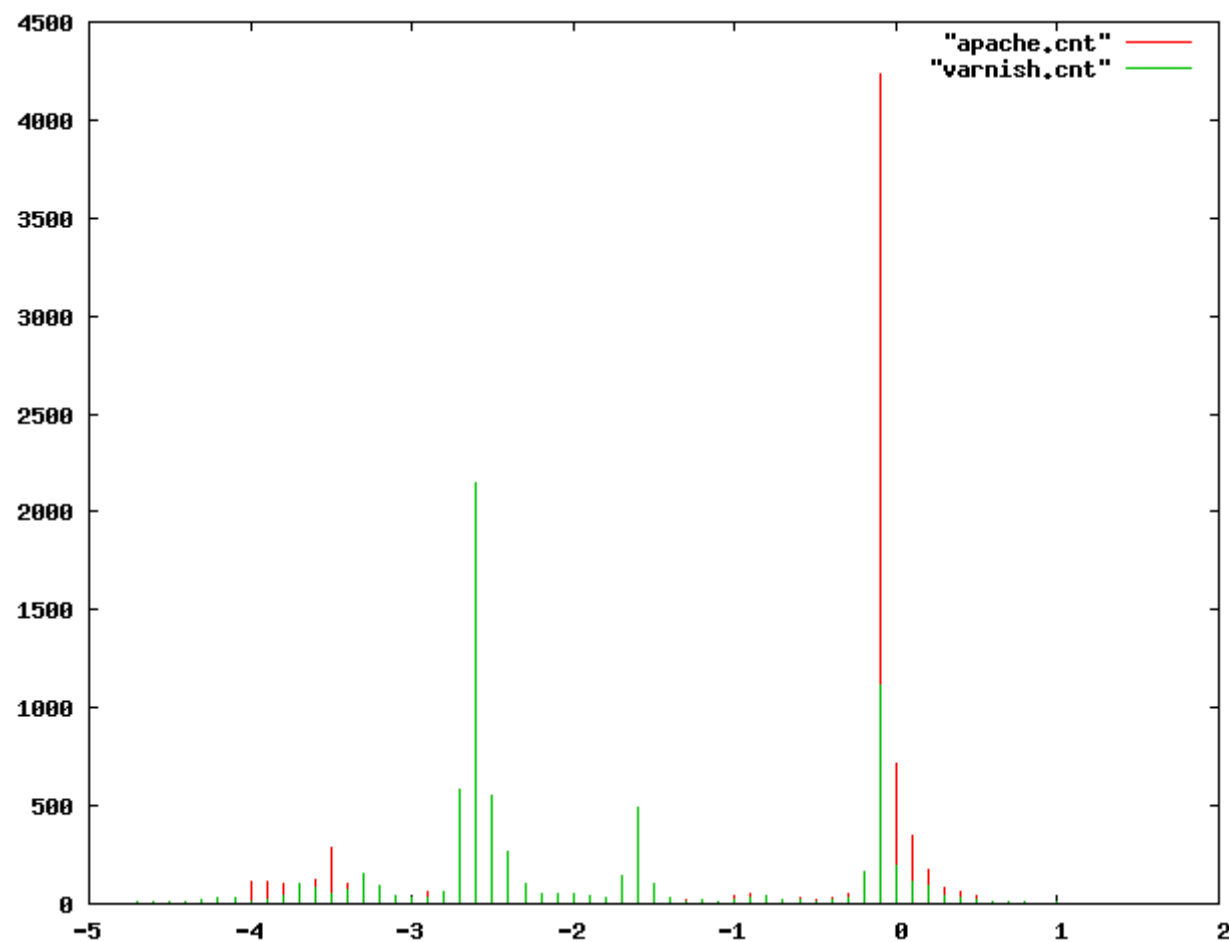
```
HTTP/1.1 200 OK
Date: Thu, 07 Jun 2007 11:09:37 GMT
Server: Apache/1.3.37 (Unix) PHP/4.4.4 mod_ssl/2.8.28 OpenSSL/0.9.8c
X-Powered-By: eZ publish
Set-Cookie: PHPSESSID=bfb2c71f4e8374c6eaeab5623d1bef; path=/
Expires: Mon, 26 Jul 1997 05:00:00 GMT
Cache-Control: no-cache, must-revalidate
Pragma: no-cache
Set-Cookie: PHPSESSID=bfb2c71f4e8374c6eaeab5623d1bef; path=/; domain=ez.no
Set-Cookie: PHPSESSID=bfb2c71f4e8374c6eaeab5623d1bef; path=/
Last-Modified: Thu, 07 Jun 2007 11:09:37 GMT
Content-language: en-GB
X-Server: ez6
X-View-Cache: yes
Connection: close
Transfer-Encoding: chunked
Content-Type: text/html; charset=iso-8859-15
```

- Introduction
- A little bit of history
- Design and architecture
- Integration issues
- **Pretty pictures**
- Miscellaneous









- Introduction
- A little bit of history
- Design and architecture
- Integration issues
- Pretty pictures
- **Miscellaneous**

- and, hopefully, some answers as well

- <http://varnish.projects.linpro.no/>
- <http://www.varnish-cache.org/>
- http://en.wikipedia.org/wiki/Varnish_cache
- <http://people.freebsd.org/~phk/>
- <http://www.linpro.no/>
- <http://www.vg.no/>
- <http://www.api.no/>
- <http://www.escenic.no/>