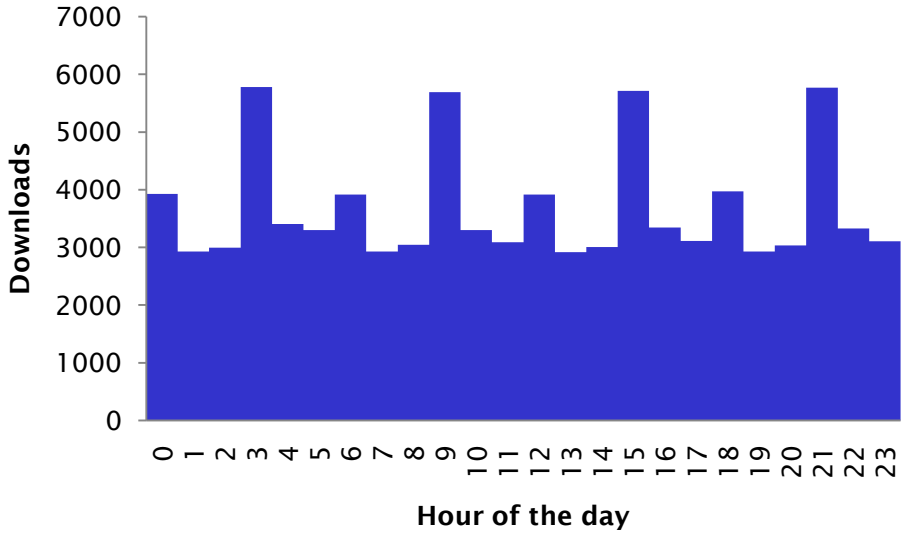
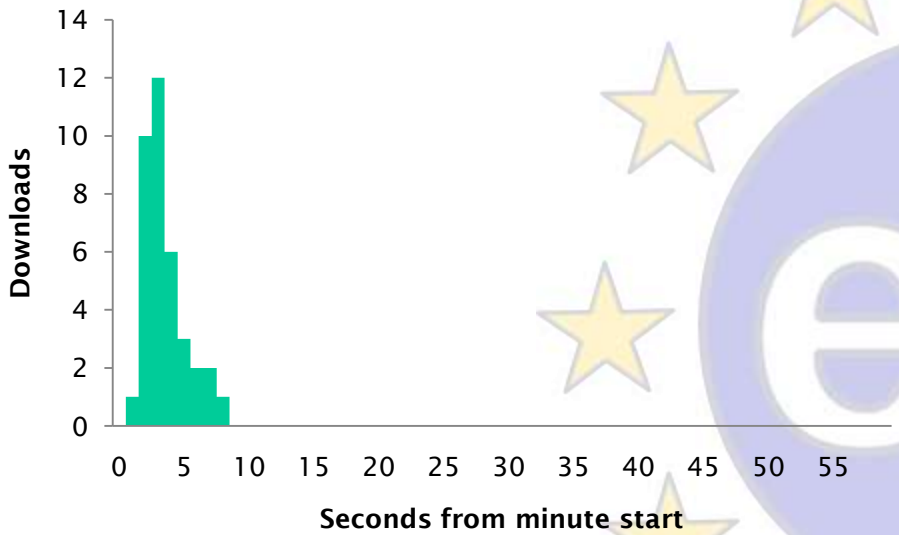
Web Cacheability of CRLs

David Groep, Jan 26th, 2009

Web Cachability, why?

Downloads clustered in first seconds of the minute

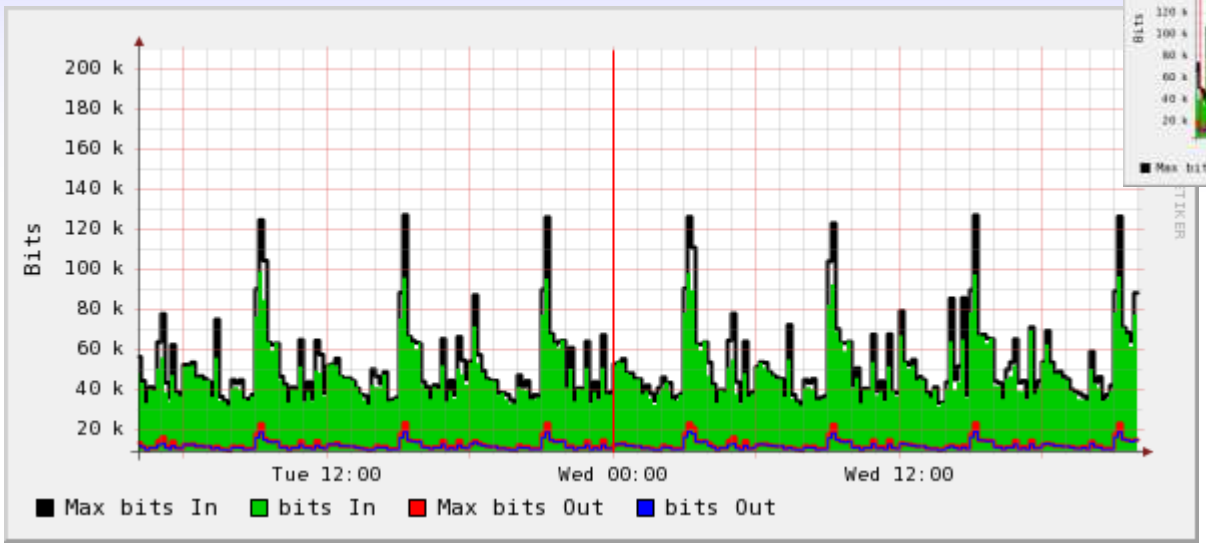


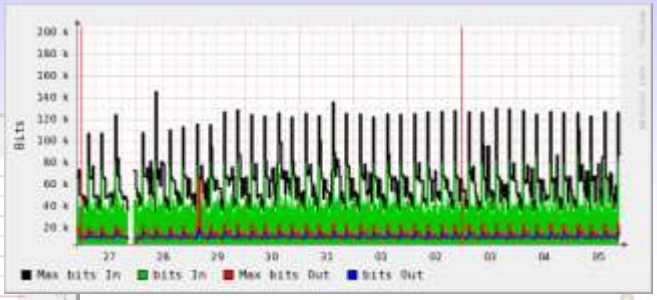


Data: DutchGrid CA

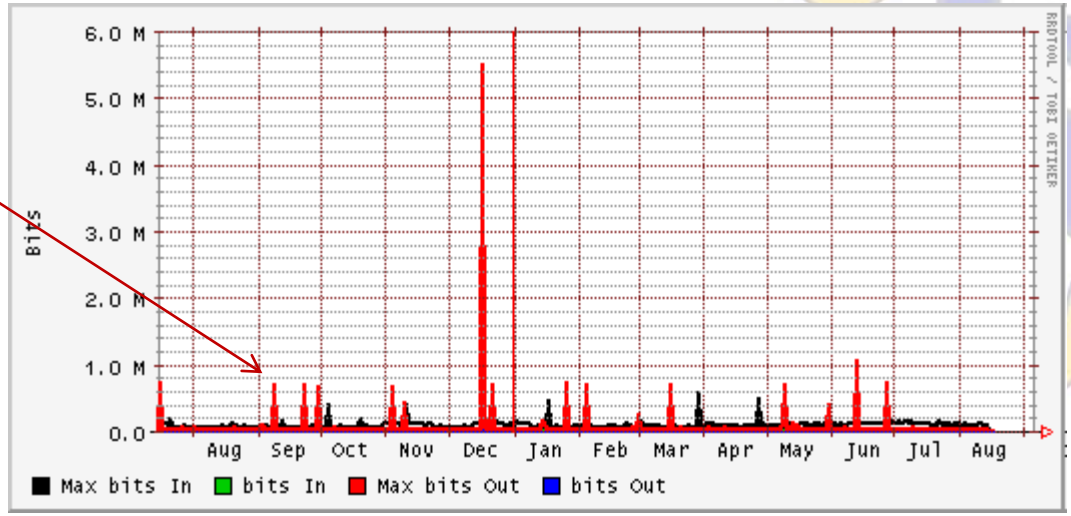
Statistics: 88452 downloads per day per CA
14084 distinct IP addresses
average 4 downloads per day per host

Network traffic



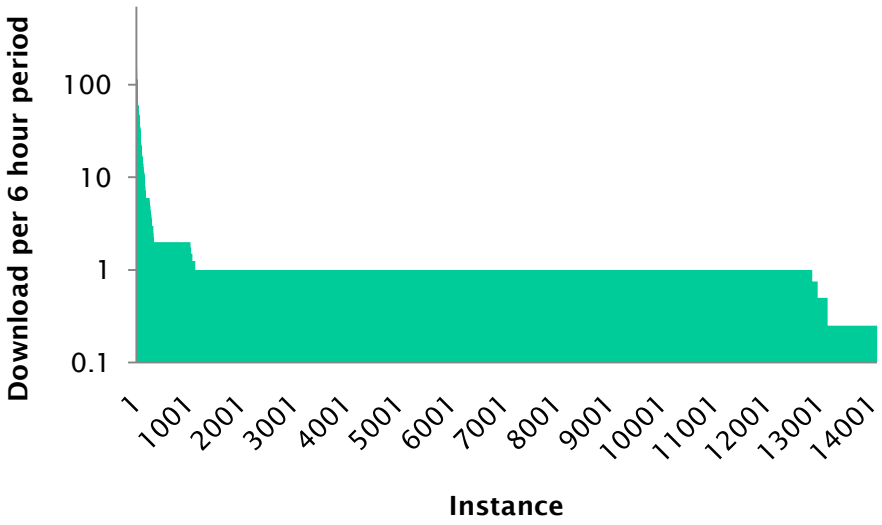


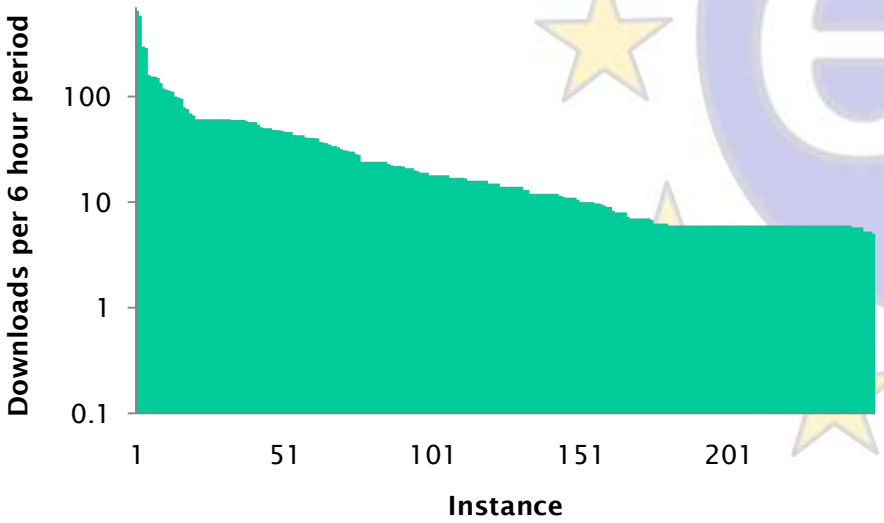
Site cache misconfigurations or new sites



There Are Caches

- Majority of IPs download individually every 6hours
- But there are at least 300 sites that cache!





Web Cacheability

- Good

```
$ HEAD -S http://ca.dutchgrid.nl/medium/cacrl.pem
200 OK
Cache-Control: max-age=3600
Connection: close
Date: Wed, 05 Nov 2008 21:31:48 GMT
Accept-Ranges: bytes
Server: Apache
Content-Length: 4728
Content-Type: text/plain
Expires: Wed, 05 Nov 2008 22:31:48 GMT
Last-Modified: Tue, 04 Nov 2008 10:07:05 GMT
Client-Date: Wed, 05 Nov 2008 21:31:48 GMT
Client-Response-Num: 1
```

Web Cacheability

- Reasonable, but relies on remote site cache setup

```
$ HEAD http://ca.grid-support.ac.uk/cgi-bin/importCRLpem
200 OK
Connection: close
Date: Wed, 05 Nov 2008 21:15:02 GMT
Accept-Ranges: bytes
ETag: "164011-6b3a-1c7652c0"
Server: Apache/2.0.46 (Red Hat)
Content-Length: 27450
Content-Type: text/plain; charset=UTF-8
Last-Modified: Wed, 09 Jan 2008 13:31:31 GMT
Client-Date: Wed, 05 Nov 2008 21:15:02 GMT
Client-Peer: 130.246.143.144:80
Client-Response-Num: 1
```

Web Cacheability

- Update your CRL URL, but answer is reasonable

```
$ HEAD http://www.dutchgrid.nl/ca/medium/cacrl.pem
HTTP/1.1 301 Moved Permanently
Date: Mon, 03 Nov 2008 08:59:13 GMT
Server: Apache
Location: http://ca.dutchgrid.nl/medium/cacrl.pem
Content-Length: 247
Content-Type: text/html; charset=iso-8859-1
```

- Update your CRL URL, answer wastes the cache

```
$ HEAD http://www.lip.pt/ca/lip-crl.pem
HTTP/1.1 302 Found
Date: Mon, 03 Nov 2008 09:04:08 GMT
Server: Apache/2.2.6 (Fedora)
Location: http://ca.lip.pt/lip-crl.pem
Content-Length: 287
Connection: close
Content-Type: text/html; charset=iso-8859-1
```

Web Cacheability

- Uncacheable

```
$ HEAD 'http://crls.services.cnrs.fr/get.fcgi?ca=CNRS-Projets&cmd=getpem.crl'
200 OK
Connection: close
Date: Wed, 05 Nov 2008 21:17:05 GMT
Server: Apache/2.2.8 (Fedora) DAV/2 mod_ssl/2.2.8 OpenSSL/0.9.8b
Content-Length: 593
Content-Type: octet/stream
Client-Date: Wed, 05 Nov 2008 21:35:44 GMT
Client-Peer: 195.220.197.22:80
Client-Response-Num: 1
Content-Disposition: attachement; filename="CNRS-Projets.crl"
```

(no Last-Modified nor Expires header)

And Last-Modified header should be there to allow HEAD requests

Web Cacheability

- Rather 'interesting'

```
$ HEAD http://www.cs.tcd.ie/Grid-Ireland/gi-ca/1e43b9cc.r0
200 OK
Cache-Control: max-age=259200
Connection: close
Date: Mon, 03 Nov 2008 09:05:40 GMT
Accept-Ranges: bytes
Server: Apache
Content-Length: 4505
Content-Type: text/plain
Expires: Thu, 06 Nov 2008 09:05:40 GMT
Last-Modified: Tue, 28 Oct 2008 15:42:44 GMT
Client-Date: Mon, 03 Nov 2008 09:05:40 GMT
Client-Peer: 134.226.32.57:80
Client-Response-Num: 1
```

Web Caching configuration (apache)

- Apache 2.x configuration - within your (virtual) host section

```
<FilesMatch "cacrl.(pem|der|cer)$">
   ExpiresActive On
   ExpiresDefault "access plus 1 hours"
   Options -Includes
</FilesMatch>

<FilesMatch "cacert.(pem|der|cer)$">
  ExpiresActive On
  ExpiresDefault "access plus 1 days"
  Options -Includes
</FilesMatch>
```

Changes to Fetch-CRL?

- Fabio Hernandez' fetch-crl utility is getting 'old'
  - Some fixes but a shell script is too limited to be really 'smart'
  - But then not all possibilities are yet used!
    (like the failover capability for downloads, should we do that?)
- New features in 2.7
  https://dist.eugridpma.info/distribution/tests/fetch-crl-2.7.0/
  - retain original download with 'Last-modified' time as obtained from the web server, so as to do HEAD instead of GET requests. … if you define a cache directory for the tool to use
  - Make cRLAgingThreshold of 24 hours the default
- Wish list for a successor tool
  - Make the tool 'stateful' so it can give adapt to network trouble and do parallel downloads
  - Support for multiple root certs with the same DN (.0,.1 system)
  - Read directly from the .info file (multiple URLs per CRL)


eu
gridpma