

# HTTP Content-Coding in VO Services

Mark Taylor (Bristol)

IVOA Interop  
Sesto

16 June 2015

`$Id: ccoding.tex,v 1.10 2015/06/16 09:23:59 mbt Exp $`

## Outline

- What is HTTP Content-Coding?
- Is it useful?
- Should services be using it?
- Should clients be using it?

# Content-Coding

## RFC 2616 (HTTP/1.1):

Sec 3.5: **Content Coding** token values

- ▶ Defines known coding tokens (gzip, compress, deflate, identity)

Sec 14.3: **Accept-Encoding** request header

- ▶ Presented by client to indicate willingness to receive encoded response

Sec 14.11: **Content-Encoding** response header

- ▶ Included by server to indicate that response is in fact encoded

# Example

Use `curl` to submit a content-coded HTTP request:

```
% curl -s -v --compress http://cu9tap.obspm.fr/tap/tables > /dev/null
* About to connect() to cu9tap.obspm.fr port 80 (#0)
* Trying 145.238.176.25... connected
* Connected to cu9tap.obspm.fr (145.238.176.25) port 80 (#0)

> GET /tap/tables HTTP/1.1
> User-Agent: curl/7.19.7 (x86_64-redhat-linux-gnu) libcurl/7.19.7 ...
> Host: cu9tap.obspm.fr
> Accept: */*
> Accept-Encoding: deflate, gzip
>

< HTTP/1.1 200 OK
< Date: Mon, 18 May 2015 13:09:15 GMT
< Server: TwistedWeb/12.0.0
< Content-Type: text/xml
< Vary: Accept-Encoding
< Content-Encoding: gzip
< Content-Length: 6760
<

{ [data not shown]
* Connection #0 to host cu9tap.obspm.fr left intact
* Closing connection #0
}
```

# Why?

If client and server agree to use compressed response (e.g. gzip):

- Network traffic is reduced
- Computational load at service and client is increased
  - ▷ Usually a good deal for the client (especially on slow connections)
  - ▷ Service may be able to cache compressed static documents
- XML gzips well
  - ▷ VOSI TableSet:  $\sim \times 10$  or better
  - ▷ TABLEDATA-serialized VOTable:  $\sim \times 8$
  - ▷ BINARY2-serialized VOTable:  $\sim \times 2$

Implementation:

- Client-side: conceptually fairly easy, not provided out of the box in Java, but libraries exist
- Server-side: ?

# Status

## Which TAP services honour Accept-Encoding: gzip requests?

- for /tables endpoint:
  - ▶ <http://cu9tap.obspm.fr/tap>,  
<http://tools.asdc.asi.it/TAP>
- for TAP query responses:
  - ▶ <http://cu9tap.obspm.fr/tap>
- I didn't find any others (though I didn't look all that hard)
- ... it looks like most services do not

# Questions

- Why are not many services using this?
  - Server CPU more valuable than bandwidth?
  - Hard to implement?
  - Never thought about it?
  - Security??
- Is it worth client code presenting Accept-Encoding?
- Is it harmful for client code to present Accept-Encoding?
  - *is there any reason I shouldn't do it?*
- Should more services be encouraged to use Content-Encoding?