

adaptTo()

APACHE SLING & FRIENDS TECH MEETUP
10-12 SEPTEMBER 2018

The underappreciated power of content invalidation

Jakub Wądołowski, Cognifide (@jwadolowski)



There are only two hard things in Computer Science:
cache invalidation and naming things.

Phil Karlton

HTTP caching

What can be cached?

- HTTP response (body & headers)
- What makes response cacheable?
- Static assets only?
- Body encoding matters
 - gzip, deflate, br, ...

Caching layers

- AEM (custom in-memory cache)
- Dispatcher
- CDN
- Web browser

Invalidation strategy

Invalidation models

- TTL-driven
- Event-driven
- Mix of both

Dispatcher invalidation

- 3 primary techniques
 - by folder level (statfile)
 - TTLs
 - resource only
- Extras
 - refetching
 - custom script

```
/cache
{
    /docroot "/opt/cache/publish"

    /rules
    {
        /0000 { /type "deny" /glob "*" }
        /0060 { /type "allow" /glob "/content/myproject/*.html" }
        /0090 { /type "allow" /glob "/content/dam/*.png" }
        /0092 { /type "allow" /glob "/content/dam/*.jpg" }
    }

    /invalidate
    {
        /0000 { /type "deny" /glob "*" }
    }

    /invalidateHandler "/etc/httpd/conf/dispatcher/invalidateHandler.sh"
}
```

A thick, hand-drawn yellow arrow originates from the right side of the slide and points diagonally down and to the left, specifically towards the `/invalidate` section of the configuration code.

Invalidation request

GET /dispatcher/invalidate.cache HTTP/1.1

→ CQ-Action: Activate

→ CQ-Handle: /content/adaptto/2018/en/home

→ CQ-Action-Scope: ResourceOnly

Referer: about:blank

Content-Length: 0

Content-Type: application/octet-stream

User-Agent: Jakarta Commons-HttpClient/3.1

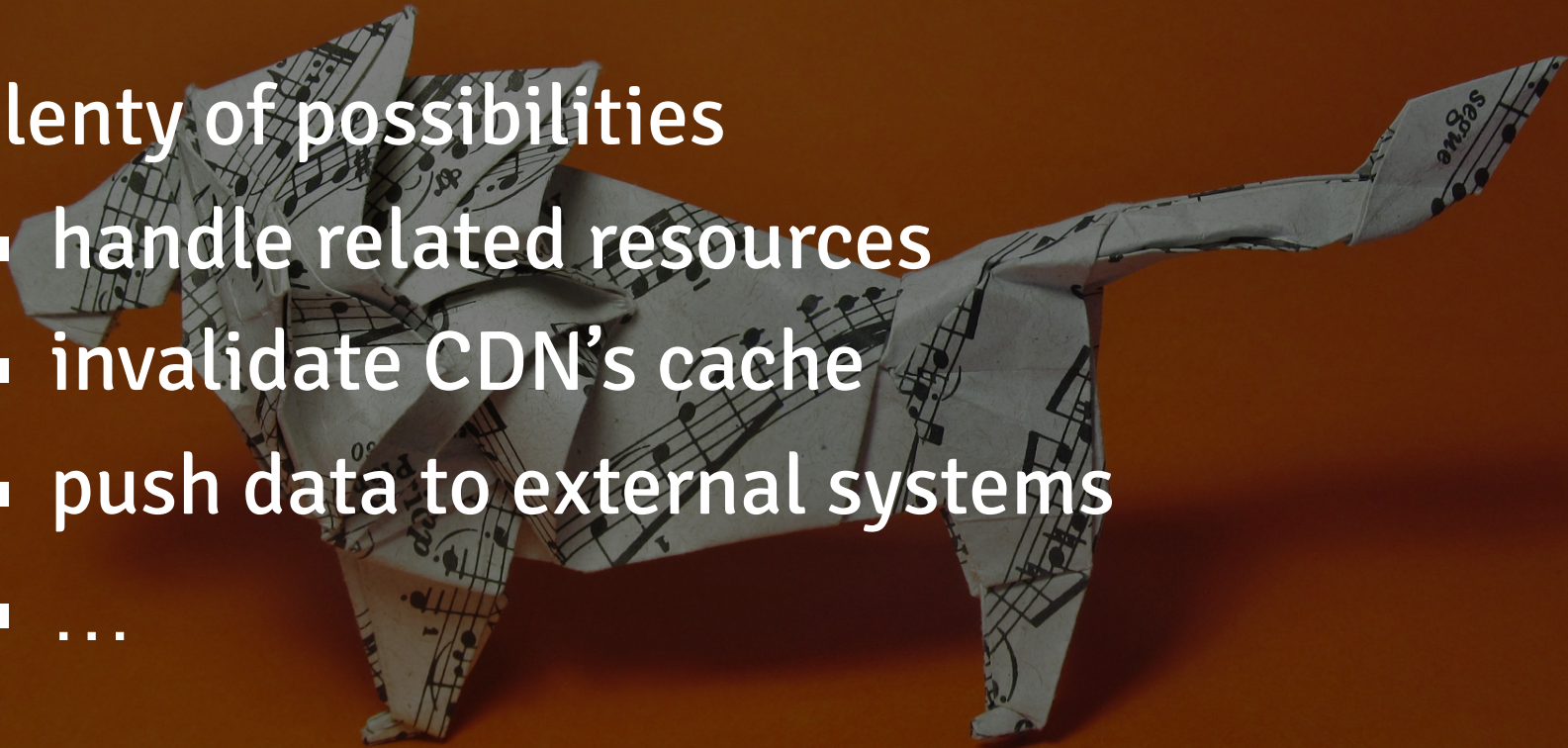
Host: adapt.to



```
/etc/httpd/conf/dispatcher/invalidateHandler.sh \  
$1 \ # CQ-Handle  
$2 \ # CQ-Action  
$3  # CQ-Action-Scope
```

Nothing fancy, huh?

- Plenty of possibilities
 - handle related resources
 - invalidate CDN's cache
 - push data to external systems
 - ...



CDN invalidation

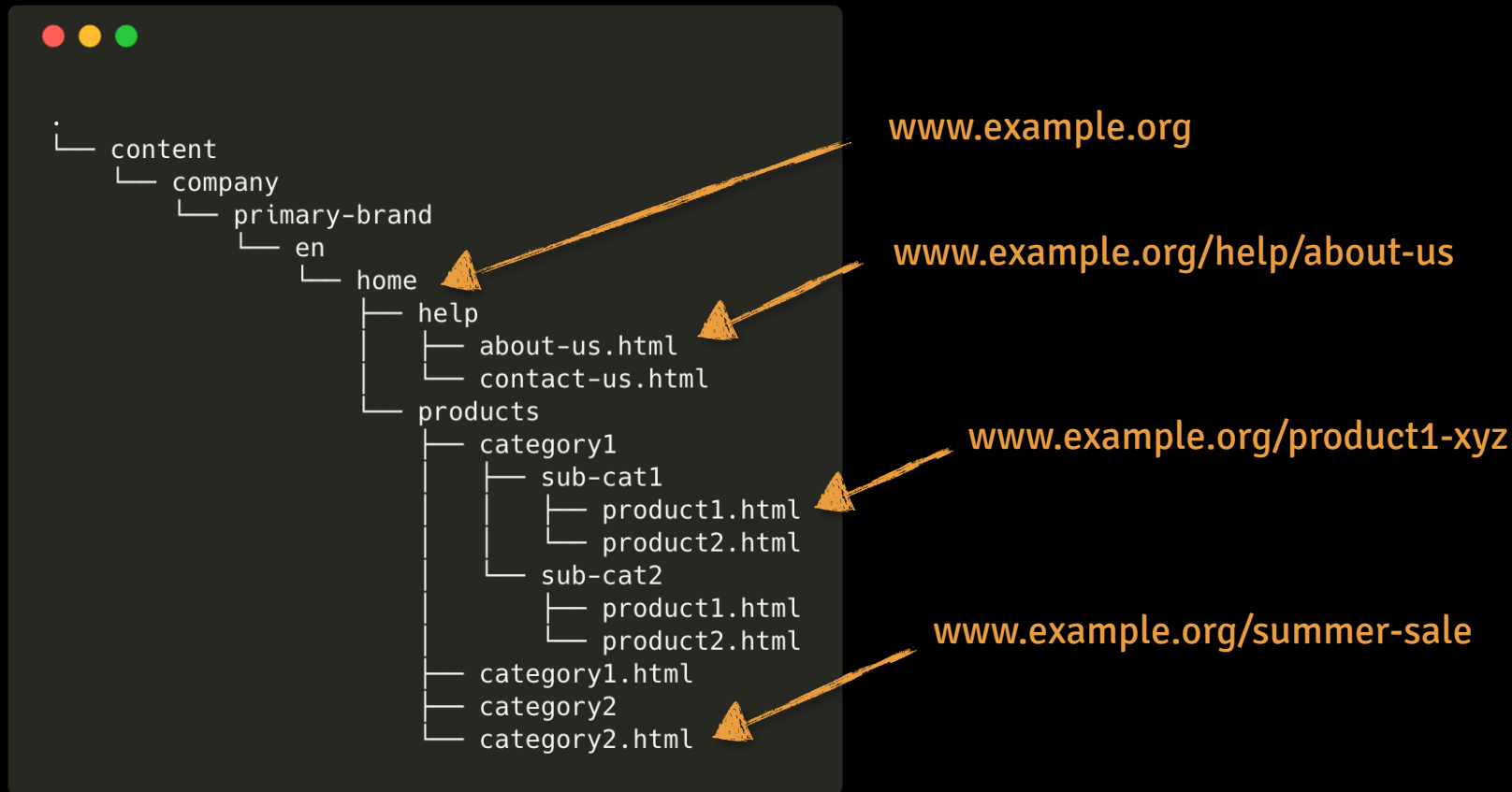
The plan

- High hit ratio & cache coverage
- Long TTLs (30+ days)
- AEM content changes reflected quickly
- Precise invalidation

Content view

- Dispatcher mirrors AEM content structure
- SEO-optimized customer-facing URLs

Dispatcher cache



CDN cache



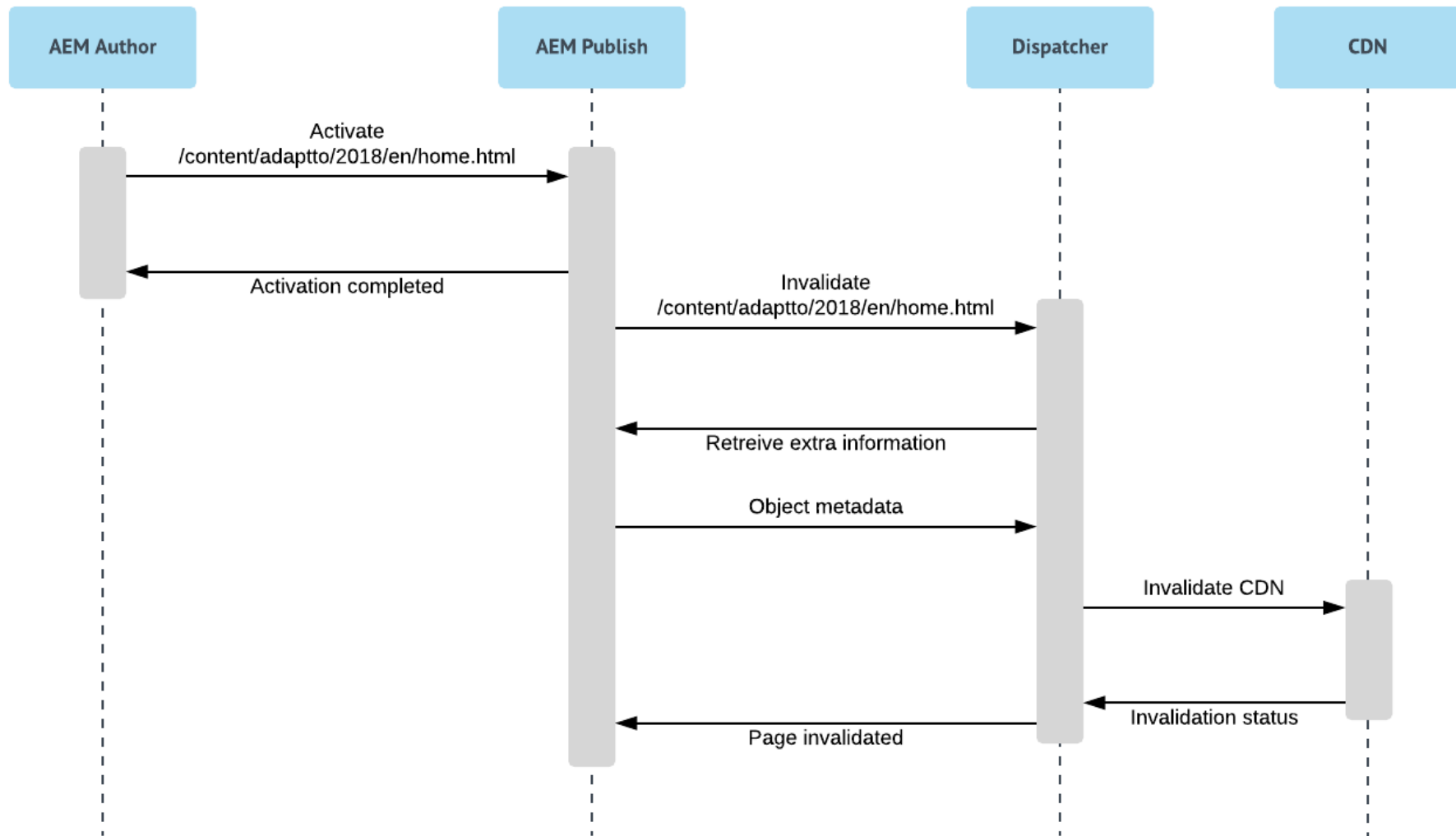
```
{  
  "homepage_key": "homepage.html_response",  
  "contact-us_key": "contact-us.html_response",  
  "schedule_key": "schedule.html_response",  
  "...": "...",  
  "venue_key": "venue.html_body"  
}
```


CDN cache keys



```
key = hash_function(protocol, host, path)
```

```
key = hash_function(  
    "https",  
    "adapt.to",  
    "/2018/en.html"  
)
```



Invalidation APIs

- Multiple endpoints
 - feature/authentication differences
 - performance may vary
- Invalidation scope
 - everything
 - URL(s)
 - tag

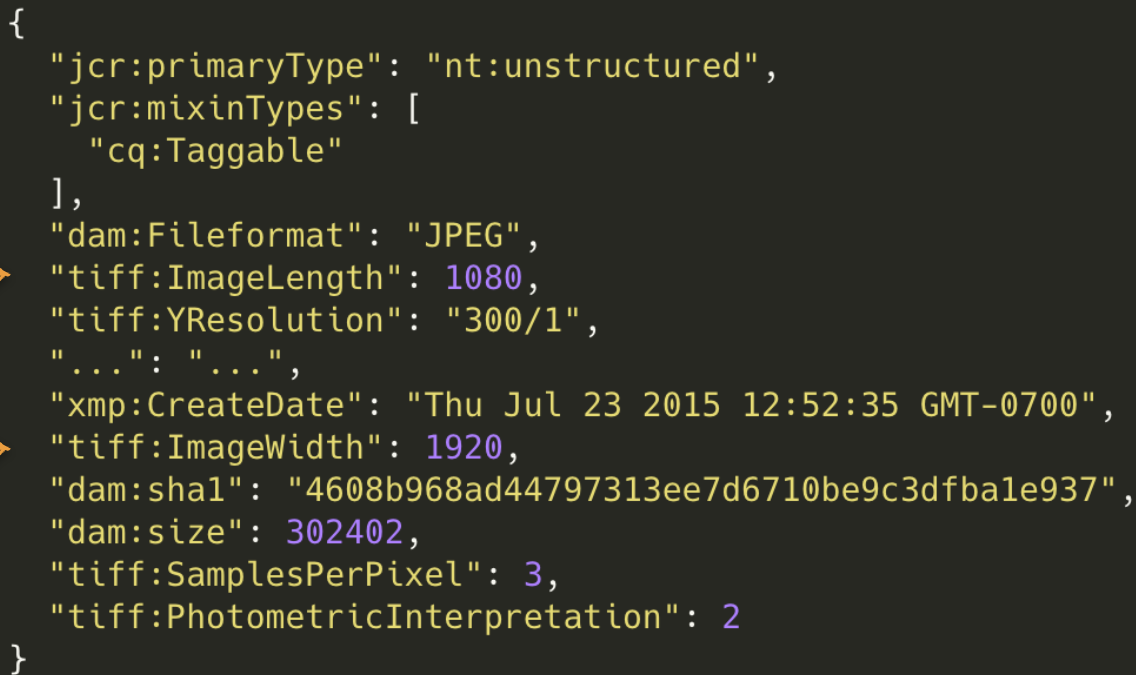
DAM assets

- Sounds trivial?
- Renditions
- Alt texts
- Non-standard properties

/content/dam/adaptto/image.jpg/jcr:content/renditions.1.json

```
{
  "jcr:primaryType": "nt:folder",
  "jcr:createdBy": "admin",
  "jcr:created": "Fri Apr 27 2018 10:01:28 GMT+0100",
  "cq5dam.web.1280.1280.jpeg": {
    "jcr:primaryType": "nt:file",
    "jcr:createdBy": "admin",
    "jcr:created": "Fri Apr 27 2018 10:01:28 GMT+0100"
  },
  "original": {
    "jcr:primaryType": "nt:file",
    "jcr:createdBy": "admin",
    "jcr:created": "Fri Apr 27 2018 10:01:28 GMT+0100"
  },
  "cq5dam.thumbnail.319.319.png": {
    "jcr:primaryType": "nt:file",
    "jcr:createdBy": "admin",
    "jcr:created": "Fri Apr 27 2018 10:01:28 GMT+0100"
  }
}
```

/content/dam/adaptto/image.jpg/jcr:content/metadata.json



```
{
  "jcr:primaryType": "nt:unstructured",
  "jcr:mixinTypes": [
    "cq:Taggable"
  ],
  "dam:Fileformat": "JPEG",
  "tiff:ImageLength": 1080,
  "tiff:YResolution": "300/1",
  "...": "...",
  "xmp:CreateDate": "Thu Jul 23 2015 12:52:35 GMT-0700",
  "tiff:ImageWidth": 1920,
  "dam:sha1": "4608b968ad44797313ee7d6710be9c3dfba1e937",
  "dam:size": 302402,
  "tiff:SamplesPerPixel": 3,
  "tiff:PhotometricInterpretation": 2
}
```

DAM invalidation script

```
declare INVALIDATED_PATH="${1}"
declare DAM_ASSET="/content/dam/."+

# -----

# Get all the renditions from AEM
function renditions() {}

# Generate API payload:
function api_payload() {}

# Send a request to Fast Purge API v3
function ccuv3_invalidate_by_url() {}

# -----

if [[ "${INVALIDATED_PATH}" =~ ${DAM_ASSET} ]]; then
    ccuv3_invalidate_by_url $(api_payload $(renditions ${INVALIDATED_PATH}))
fi
```

Akamai API request

```
POST /ccu/v3/invalidate/url/production
```

```
Content-Type: application/json
```

```
Host: api.akamai.com
```

```
User-Agent: HTTPie/0.9.9
```

```
Authorization: EG1-HMAC-SHA256 <signed_token>
```

```
{  
  "objects": [  
    "https://example.org/content/dam/images/image1.jpg.rendition.480.360.jpg",  
    "https://example.org/content/dam/images/image1.jpg.rendition.900.675.jpg",  
    "https://example.org/content/dam/images/image1.jpg.rendition.1152.864.jpg",  
    "https://example.org/content/dam/images/image1.jpg.rendition.2880.1233.jpg"  
  ]  
}
```


Alt text (1/3)



```

```

Alt text (2/3)

- Get all the pages that use given asset
- Query Builder API



```
path=/content/adaptTo  
property=fileReference  
property.value=/content/dam/adaptto/images/1.png
```

Alt text (3/3)

```
{
  "hits": [
    {
      "path": "/content/adaptto/2018/home/jcr:content/content/box_123/image",
      "excerpt": "",
      "name": "image",
      "title": "image"
    },
    {
      "path": "/content/adaptto/2018/home/contact-us/jcr:content/content/box_789/image",
      "excerpt": "",
      "name": "image",
      "title": "image"
    }
  ]
}
```

Vanity URLs

- Each page can have a number of easy-to-remember URLs
- Those pages are cacheable too

Vanity URLs (1/3)

Vanity URL

Vanity URL

/homepage



/start-here



/yet/another/vanity



Add

Vanity URLs (2/3)

Agent Settings

Settings Transport Proxy Extended Triggers Batch

Name
informative name of this agent

Description

Enabled ☒

Serialization Type
▼

Retry Delay
time in milliseconds


Agent User Id
Leave empty to use system user. This user's keystore will be used for SSL connections.

Log Level
▼

Use for reverse replication ☐

Alias update ☐

OK Cancel

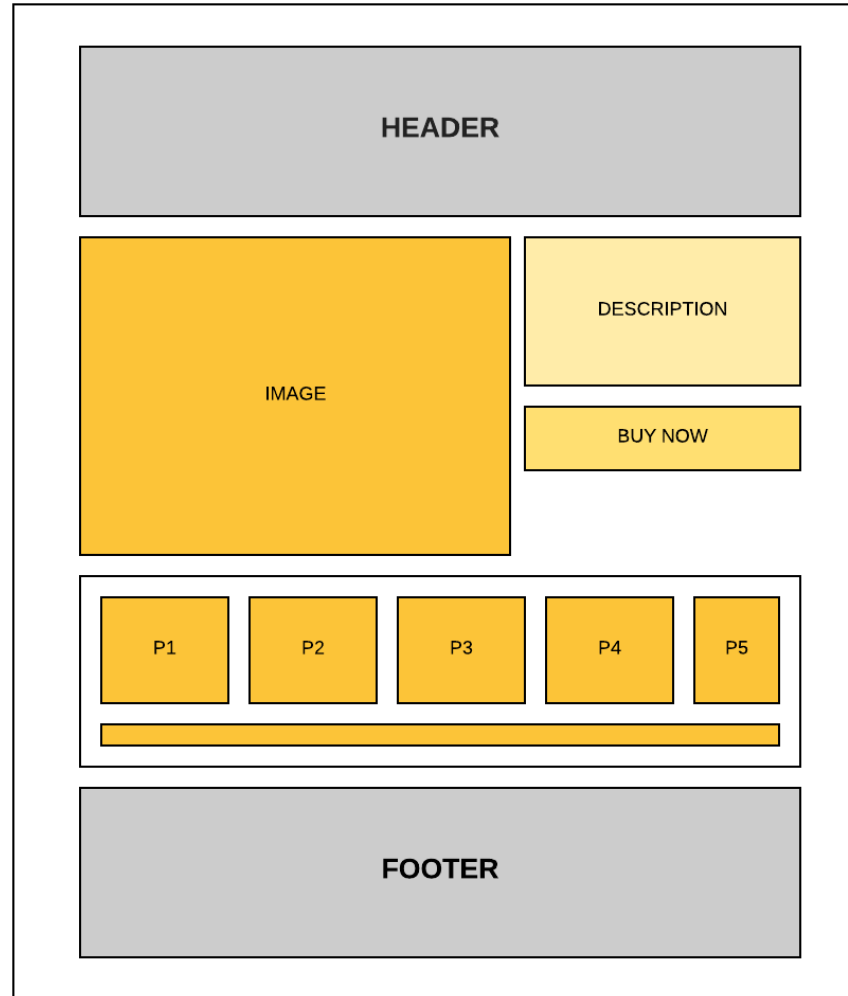
A large red arrow points from the right side of the dialog box towards the "Alias update" checkbox, which is currently unchecked.

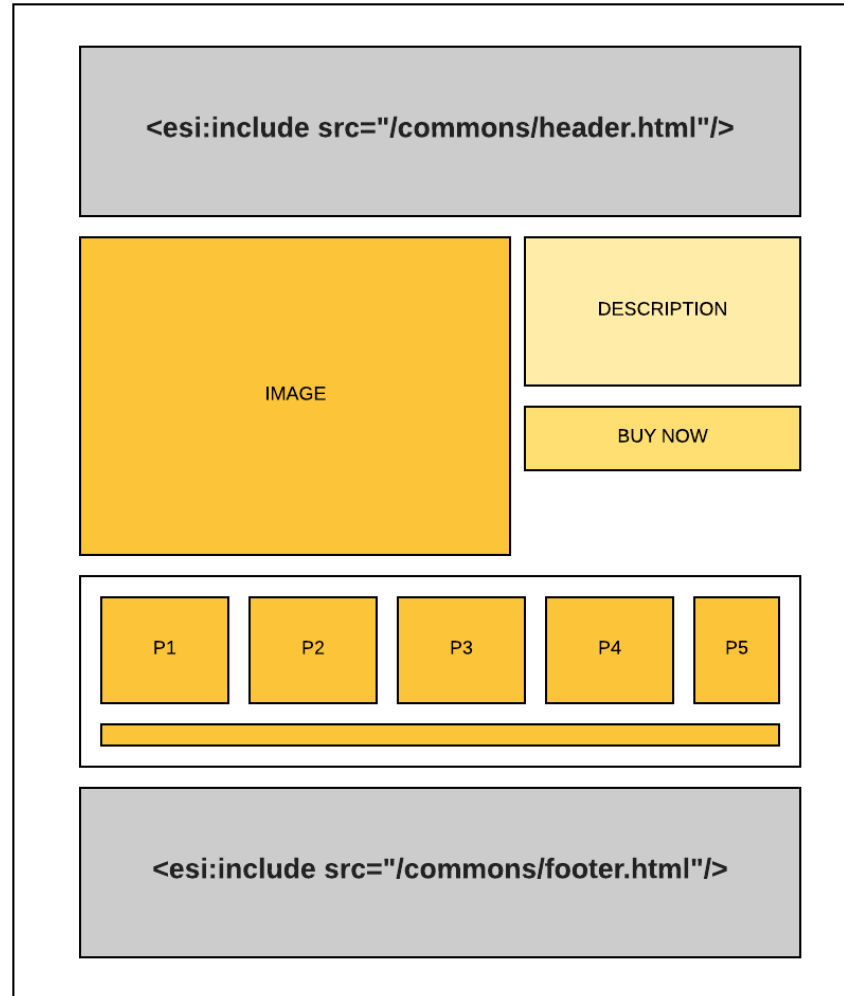
Vanity URLs (3/3)

```
[Sun Sep 01 13:16:53 2018] Activation detected: action=Activate [/start-here]
[Sun Sep 01 13:16:53 2018] Activation detected: action=Activate [/yet/another/vanity]
[Sun Sep 01 13:16:53 2018] Activation detected: action=Activate [/homepage]
...
[Sun Sep 01 13:16:53 2018] Activation detected: action=Activate [/content/adaptto/de/de-de/home]
[Sun Sep 01 13:16:53 2018] Touched /mnt/var/www/html/publish/.stat
[Sun Sep 01 13:16:53 2018] Evicted /mnt/var/www/html/publish/content/adaptto/de/de-de/home.html
[Sun Sep 01 13:16:53 2018] Evicted /mnt/var/www/html/publish/content/adaptto/de/de-de/home.html.h
```

Reusable components

- Common elements (i.e. header)
- SDI (Sling Dynamic Include) to the rescue
 - SSI (Server Side Includes)
 - ESI (Edge Side Includes)

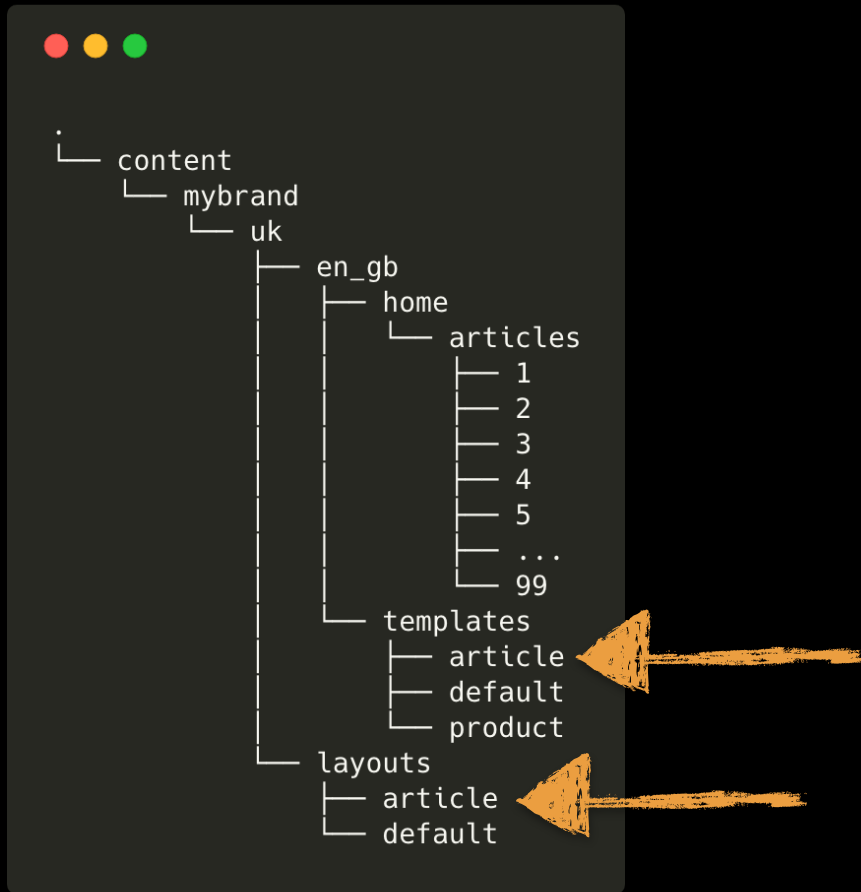


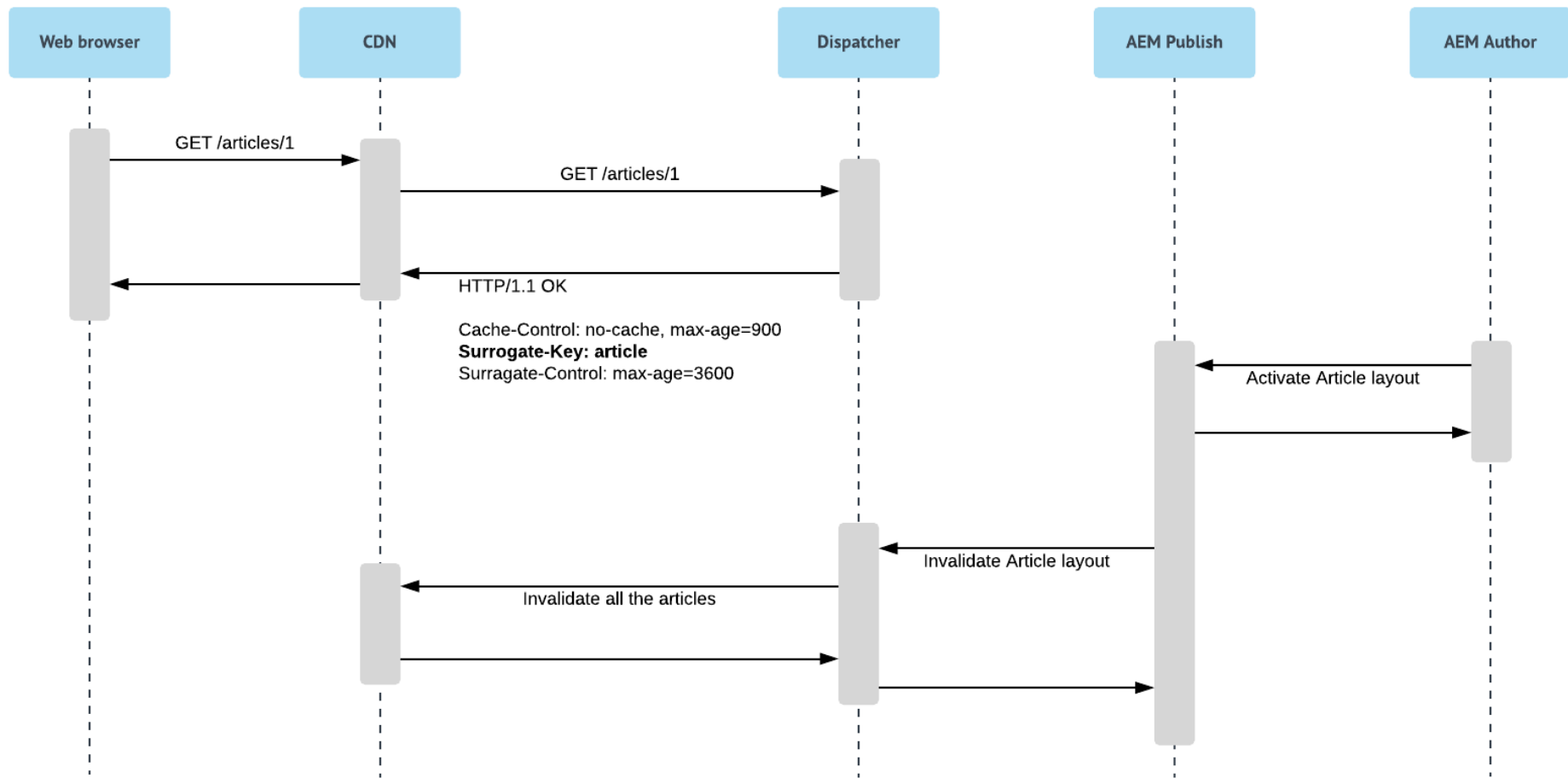


Tag-based invalidation (1/4)

- Not everything can be wrapped into SSI/ESI
 - Layout/templates
 - HTML markup configs
- Is QueryBuilder the only option?

Tag-based invalidation (2/4)





Tag-based invalidation (3/4)



```
POST /service/<SERVICE_ID>/purge/article
```

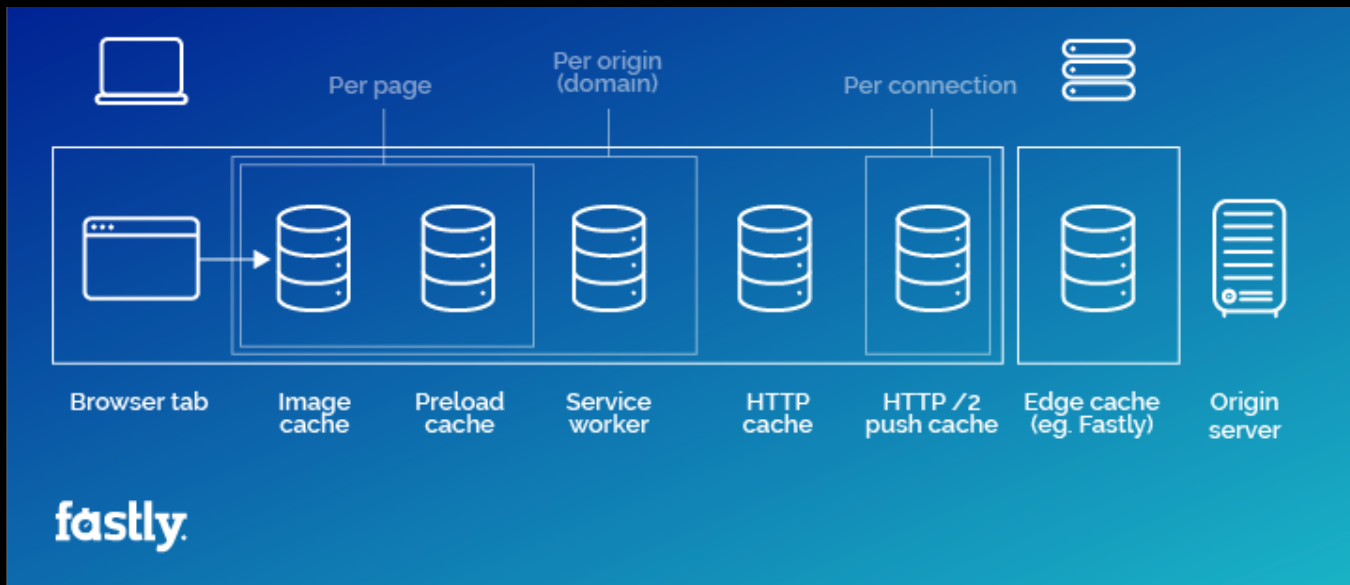
```
Fastly-Key: <auth_token>
```

```
Accept: application/json
```

```
Fastly-Soft-Purge: 1
```

Browser caching

Browser cache setup

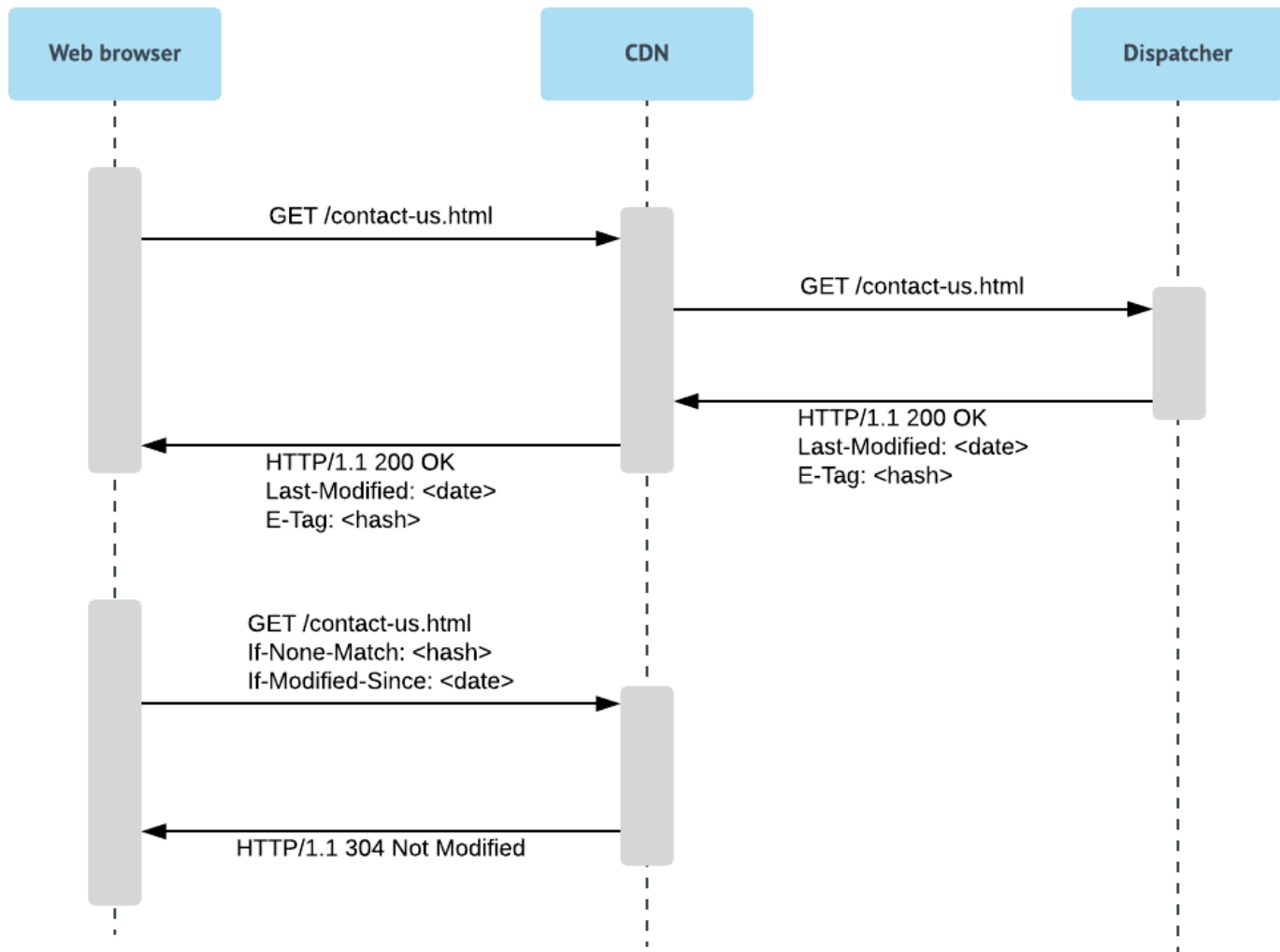


Versioned clientlibs

- Generally a good idea
- Does it fit well into CDN setup?

Cache-Control

- public/private
- max-age=<seconds>
- must-revalidate
- no-cache



We barely scratched the surface...

- API throttling
- Error handling
- Payload limits
- Non-obvious content references

Lessons learned

- purge all is dangerous
- make invalidation as precise as possible
- plan SSI/ESI since the very beginning
- QueryBuilder calls may require extra indexes



Thank you!