



Performance at Mass.gov



Rob Bayliss

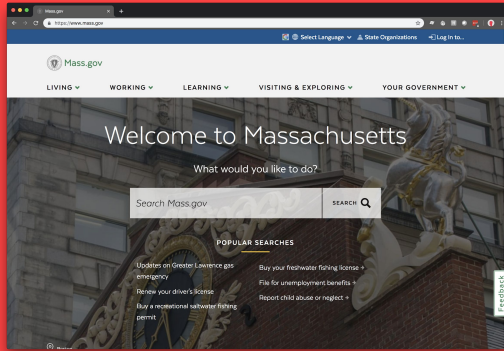
CTO/Last Call Media



Moshe Weitzman

Drupal Architect

A word about Mass.gov



Serves the Commonwealth of Massachusetts

Stakeholders are:

- Constituents (as visitors)
- State Organizations (as visitors and publishers)

Receives ~15M pageviews/month

Changes are released twice weekly by a team of developers.

Backend Performance:

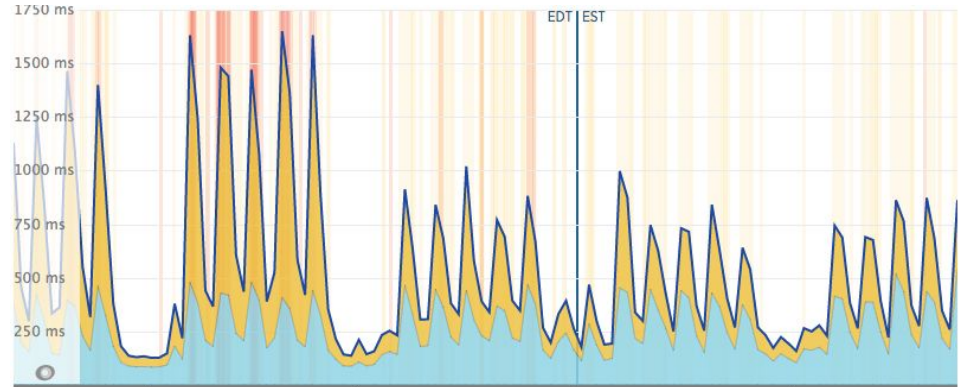
State employees need to be able to publish content in **near real time**.

Constituents need **fast access** to information with **no disruptions**.

The DevOps team needs to balance **freshness** with **reliability**.

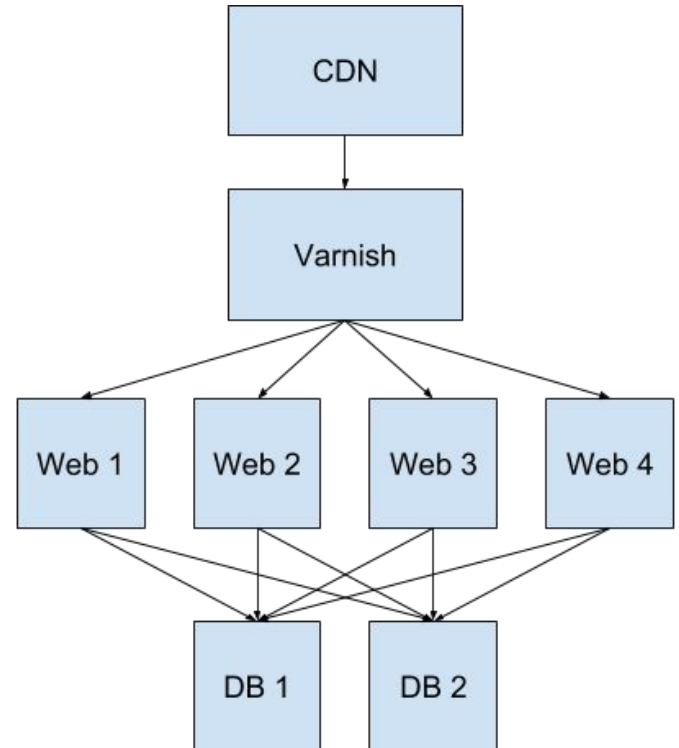
The Traffic Profile

- Visitors come in all day (heaviest 8AM - 8PM)
- Editors work between business hours (8AM - 5PM)
- Releases happen twice weekly (after 8PM)



The Operating Environment

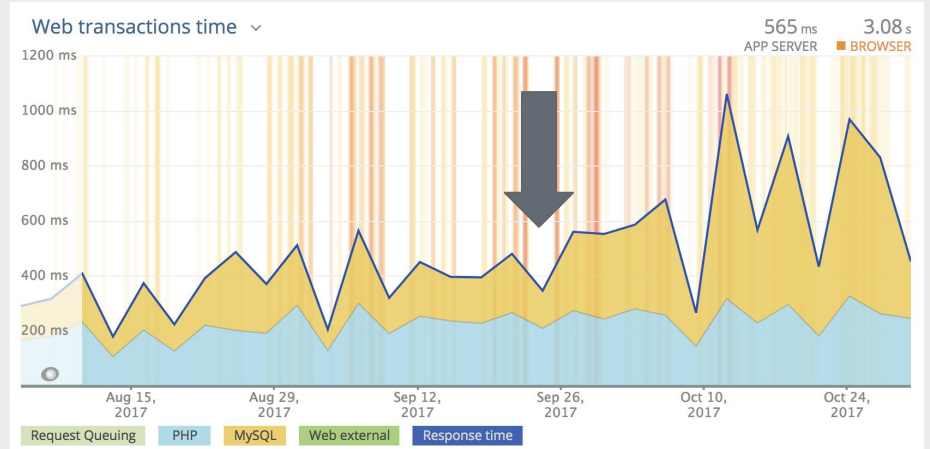
- CDN caches everything for anonymous traffic
- Varnish caches everything for anonymous traffic
- Webs are only hit when there is a miss on CDN and Varnish



September, 2017

Site “Launch”

DNS is switched to change the URL of the Drupal site from *pilot.mass.gov* to *mass.gov*



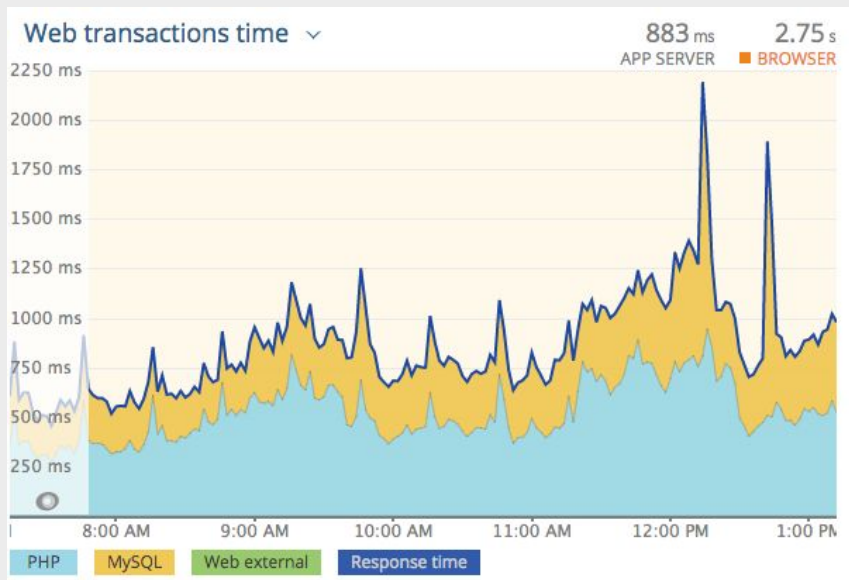
November 30, 2017

Purge module is rolled out

Varnish cleared on content
update

Editors see changes in < 1 hour

Freshness improves, but
performance take a big hit.



December, 2017

Many missing cache tags are fixed.

The screenshot shows a web browser displaying the 'Early Childhood Education & Care' page. The 'Learning' link in the 'MORE ABOUT:' section is highlighted with a red box. Below the page, the browser's developer tools are open, showing the Network tab with a list of resources and the Headers tab for the selected resource. The Headers tab shows various cache-related headers, with 'X-Drupal-Cache-Tags' and 'X-Drupal-Dynamic-Cache: MISS' highlighted in red boxes.

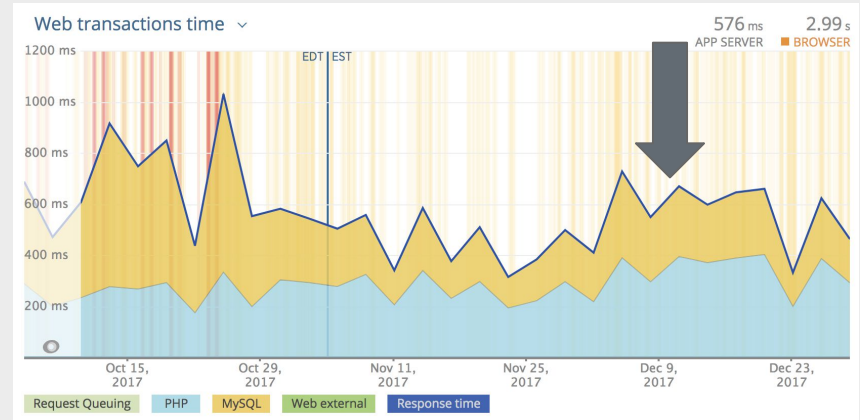
Network Tab Headers:

- X-Content-Type-Options: nosniff
- X-Drupal-Cache-Contexts: language_interface route theme timezone url.path url.query_args_wrapper_format url.site user.permissions user.roles.anonymous
- X-Drupal-Cache-Tags: block_view config:block.block.ajaxpattern config:block.block.ajaxpattern_2 config:block.block.contentflags onfig:block.block.currentrevisionstate config:block.block.decisiontreeadmin config:block.block.footer1 config:block.block.footer2 config:block.block.footer3 config:block.block.mainnavigation config:block.block.messages tabs config:block.block.pagefeedback ra config:block_list config:flag.flag.watch_content config:system.menu.footer-1 config:system.menu.footer-2 config:system.menu.footer-3 config:system.menu.main config:system.menu.utility config:system.site config:user.role.anonymous file:28831 file:28836 ttp_response node:14586 node:14636 node:16956 node:16466 node:16586 node:16651 node:16691 node:16786 node:16741 node:16751 node:16761 node:16956 node:17061 node:17211 node:17216 node:17256 node:17276 node:17286 node:17296 node:17316 node:17566 node:17571 node:17576 node:17681 node:17631 node:17631 node:17656 node:19711 node:21166 node:21186 node:22431 node:22446 node:22536 node:251 node:38511 node:3736 node:3566 node:3821 node:4281 node:4286 node:53196 node:57366 node:64861 node:8261 node_view paragraph 51486 rendered taxonomy_term:47171 user:16
- X-Drupal-Dynamic-Cache: MISS
- X-Frame-Options: SAMEORIGIN
- X-Generator: Drupal 8 (https://www.drupal.org)
- X-Powered-By: PHP/7.1.17

December, 2017

Drupal cache lifetime
increased to 3 hours

Attempt to take advantage of the purge work by
holding pages that *haven't changed* for up to 3 hours.



January, 2018

Backend disruption due to
accidental self-DDOS

Work on the emergency alerts system disrupts the site after it is deployed with a cache-busting query parameters in an AJAX request.

Takeaways:

- Every PR needs to be considered for performance impact.
- Having a system in place to run some level of load test or benchmarking on changes before production is key if uptime matters.
- Even front end work can affect backend performance.

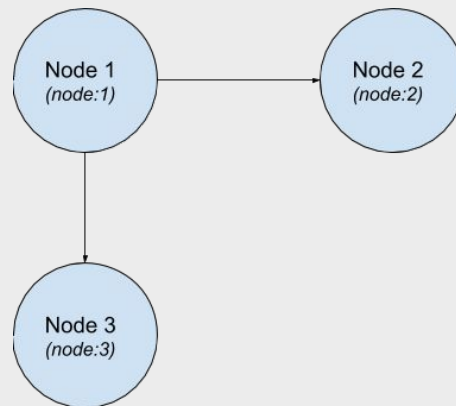
January, 2018

Strip out the *node_list* cache tag from most pages, replacing with “relationship” tag clearing.

The absence of the *node_list* tag is backed up by Behat tests, and relationship clearing is also tested.

```
* - Visit SD and observe that SP _is_ reflected in the "Related to" section.
*
* To work around this, we clear tags for all referenced entities when a
* referencing entity is saved. See https://github.com/massgov/mass/pull/1747
* for a full discussion of the issue.
*/
function mass_fields_entity_insert(EntityInterface $entity) {
    mass_fields_entity_clear_referenced($entity);
}

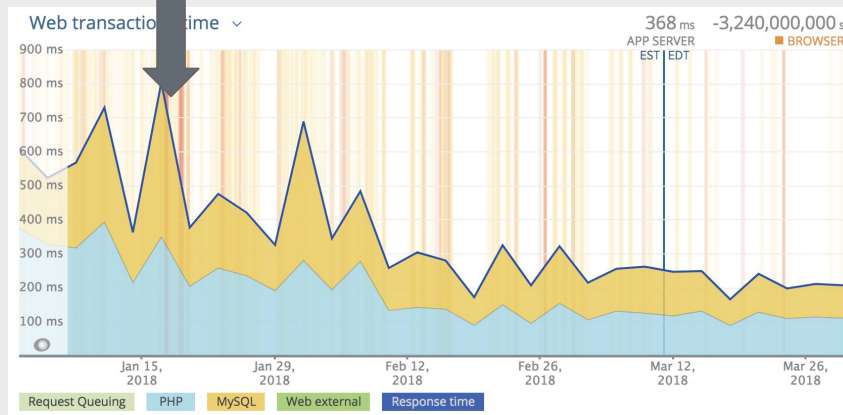
/**
 * Implements hook_entity_update().
 */
function mass_fields_entity_update(EntityInterface $entity) {
    mass_fields_entity_clear_referenced($entity);
}
```



January, 2018

Strip out the *node_list* cache tag from most pages, replacing with “relationship” tag clearing.

The absence of the *node_list* tag is backed up by Behat tests, and relationship clearing is also tested.



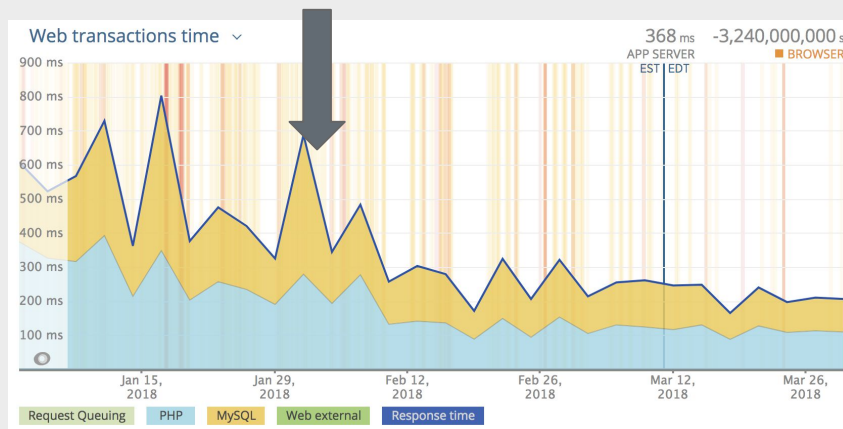
January, 2018

Metatags module is patched to prevent 2x token generation.

This performance enhancement was submitted and accepted on D.O.

Takeaways:

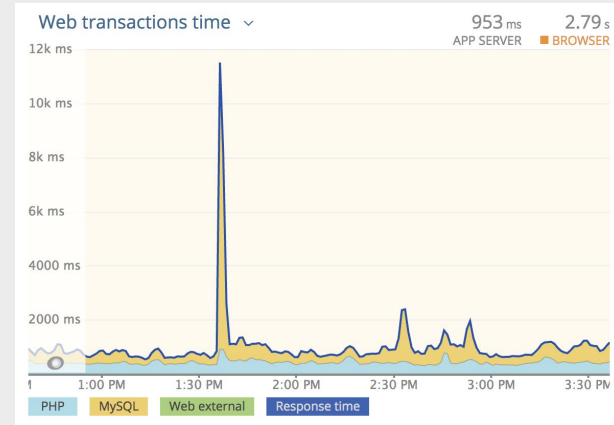
- Contrib modules aren't immune to performance problems.



February, 2018

Backend disruption due to heavy traffic on admin views

Symptoms: Users unable to log in, slowness on uncached pages of public site.



February, 2018

Backend disruption due to heavy traffic on admin views

Symptoms: Users unable to log in, slowness on some pages of public site.

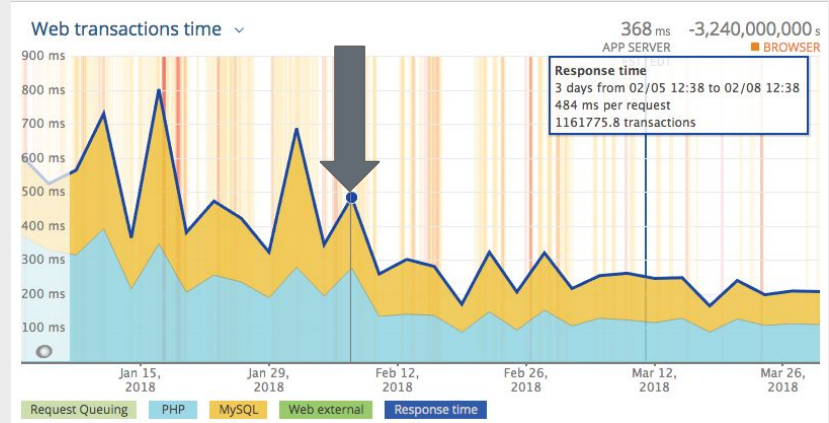
Takeaways:

1. Your bottlenecks will change as traffic patterns change.
2. Be proactive about fixing small issues before they become big issues - we spotted this issue the month prior, but didn't prioritize.
3. Full pagers are evil because they require count queries that are always as bad or worse than the views query.

February, 2018

Upgrade to PHP 7

“Free” performance win celebrated by everyone.



February, 2018

New Relic Monitoring
Customizations.

You can't improve what you
can't measure

Most time consuming

/dynamic.cache.view	38.4%
/entity.node.canonical:service_details	7.39%
/media_entity_download.download	6.15%
/redirect.canonical	3.95%
/Drupal...ontroller\NodeViewController->view	3.77%
/entity.node.canonical:regulation	3.04%
/entity.node.canonical:service_page	2.71%
/redirect.redirect	2.45%
/entity.media.canonical:document	1.97%
/entity.node.canonical:curated_list	1.92%
/entity.node.canonical:decision	1.85%
/entity.node.canonical:how_to_page	1.85%

February, 2018

New Relic Monitoring
Customizations.

You can't improve what you
can't measure

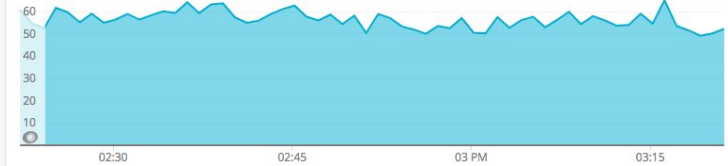
% Dynamic Cache

Since 60 minutes ago



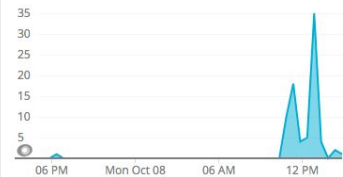
% Dynamic Cache

Since 60 minutes ago



Node Edits

Since 1 day ago



What kinds of content is being edited?

Since 1 day ago

NAME	TRANSACTIONS
WebTransaction/Custom/entity.node.edit_form:guide_page	62
WebTransaction/Custom/entity.node.edit_form:service_page	24
WebTransaction/Custom/entity.node.edit_form:curated_list	16
WebTransaction/Custom/entity.node.edit_form:service_details	5
WebTransaction/Custom/entity.node.edit_form:info_details	4
WebTransaction/Custom/entity.node.edit_form:how_to_page	4

February, 2018

Audit and remediate cache
hogging items

A	B	C	D	E	F
Bin	Writes	Reads	Write %	Size (mb)	Cardinality
bootstrap	313,776.00	4,950,000.00	6.34%	1.33	22.00
config	103,000.00	8,980,000.00	1.15%	8.53	6,063.00
container	26.00	2,120,000.00	0.00%	0.03	1.00
data	1,882,850.00	13,300,000.00	14.16%	890	487,652.00
default	24,500.00	5,490,000.00	0.45%	14.64	2,122.00
discovery	111,000.00	13,900,000.00	0.80%	7.13	471.00
dynamic_page_cache	661,000.00	3,320,000.00	19.91%	4116	35,718.00
entity	726,000.00	8,520,000.00	8.52%	2481	420,816.00
menu	132,000.00	1,340,000.00	9.85%	35.44	27,206.00
render	692,000.00	4,040,000.00	17.13%	1536	87,685.00
Total	4,646,152.00	65,960,000.00	7.04%	9090.1	1,067,756.00

February, 2018

Audit and remediate cache
hogging items

Anatomy of a cache item:

```
response:[languages:language_interface]=en:[request_format]=html:[route]=mass_map.map_page5a04d4631e658a97aee299dca5490f20606afa3b1d74727d2789c61256bf616f:[theme]=mass_theme:[url.path]=/visit-massachusetts-state-Qe00jmfLF-ho00WMXuD9XMrqNVUDXc0sEZrAm7KYfas
```

This cache item uses the following “contexts”:

- *languages:language_interface*
- *request_format*
- *route*
- *theme*
- *url.path*

February, 2018

Audit and remediate cache
hogging items

A simple example:

```
response:[route]=entity.node.canonical15a04:[url.path]  
=/visit-massachusetts-state-parks
```

This represents the same content as:

```
response:[route]=entity.node.canonical15a04:[url.path]  
=/node/123
```

You will create a new cache entry every time a context you care about changes. Be a nihilist... care about less stuff!

February, 2018

Audit and remediate cache
hogging items

Problematic contexts:

1. *url.path* - prefer *route* wherever you can
2. *url.query_args* - prefer *url.query_args.X* instead
3. *headers.** - there aren't many valid use cases for this
4. *cookies.** - Don't use cookies in your backend code.

February, 2018

Audit and remediate cache
hogging items

Rule of thumb: cache global elements with the broadest/least contexts possible, using a #lazy_builder if you need a single, simple element to change based on some conditions.

Did you find what you were looking for on this webpage? *

Yes No

If you need to report child abuse, any other kind of abuse, or need urgent assistance, please [click here](#).

SEND FEEDBACK

February, 2018

Audit and remediate cache
hogging items

(a word about the dynamic
page cache)

Rule of thumb: If you have a reverse proxy cache in front of your site, use *dynamic_page_cache* instead of *page_cache*. This will save a lot of cache space due to url variations that don't matter to your site (eg: */?utm_source=Twitter*).

	No Cache	Dynamic Page Cache	Page Cache
Homepage	438	156.4	50.4
1	405	178	60
2	437	138	52
3	574	160	57
4	426	182	39
5	348	124	44
topics/parks-recreation	517	166	46.6
1	538	135	46
2	499	138	56
3	589	145	47
4	509	225	45
5	450	187	39
Cache Size (bytes)	0.00	256,000.00	272,000.00
Avg Response Time (ms)	477.5	161.2	48.5
% Original time		33.76%	10.16%

February, 2018

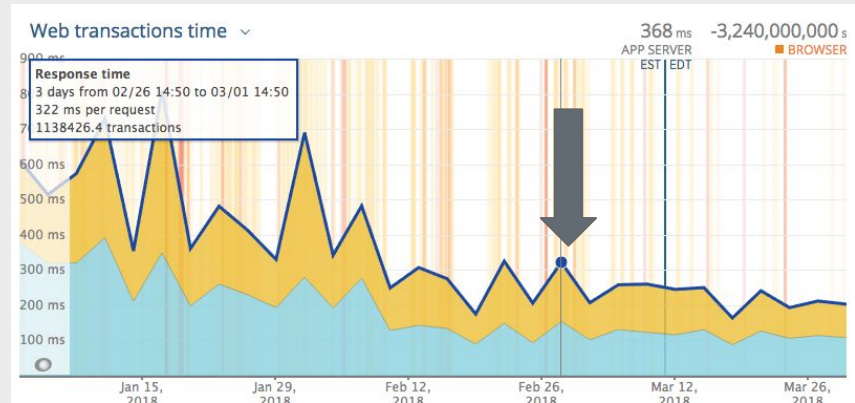
Begin Crawling the *CD*
environment nightly.

We intended this as a nightly error check, but it provides a pretty good daily performance audit for pre-production code, too!



February, 2018

CDN Cache Lifetime goes from 60 minutes to 30 minutes



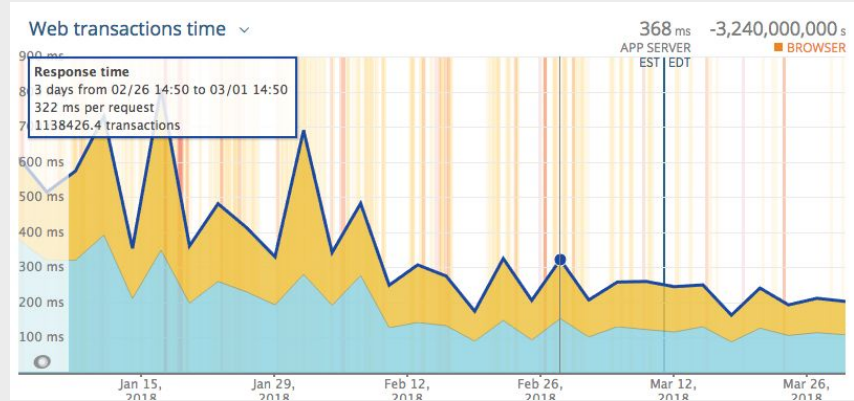
Overall Effect

Time to publish is cut in half,
with average response time
cut by over 50%

Nov 1: 526ms

Mar 29: 227ms

This has resulted in a 0 disruptions due to performance issues since we completed our work.



Frontend Performance:

Constituents need **fast access** to information across many devices and varying network connections.

Our research:

0-1s load time: 1.6% bounce rate

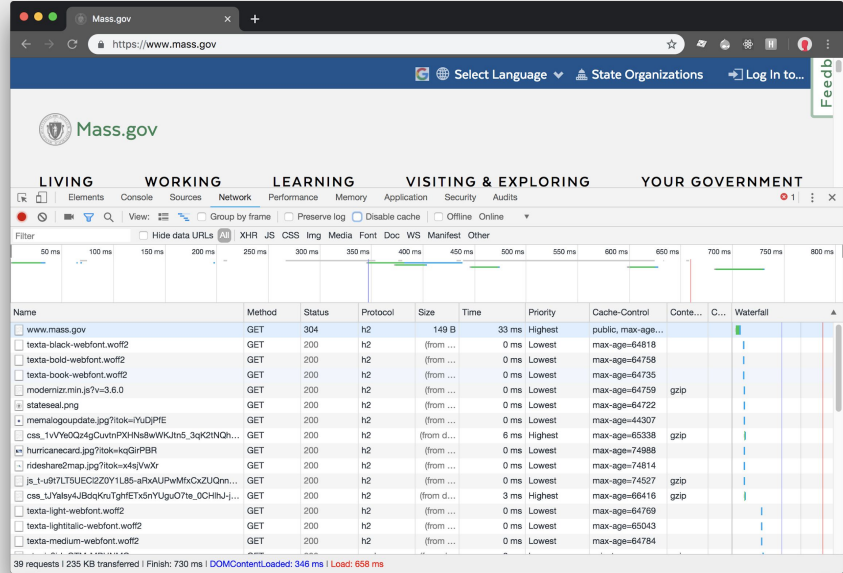
1-3s load time: 10.4% bounce rate

3-7s load time: 22.7% bounce rate

7-13s load time: 24.7% bounce rate

April, 2018

Enable HTTP/2



April, 2018

Switch from directly inlining SVG to embedding with a *use* statement and inlining each used SVG 1x per page.

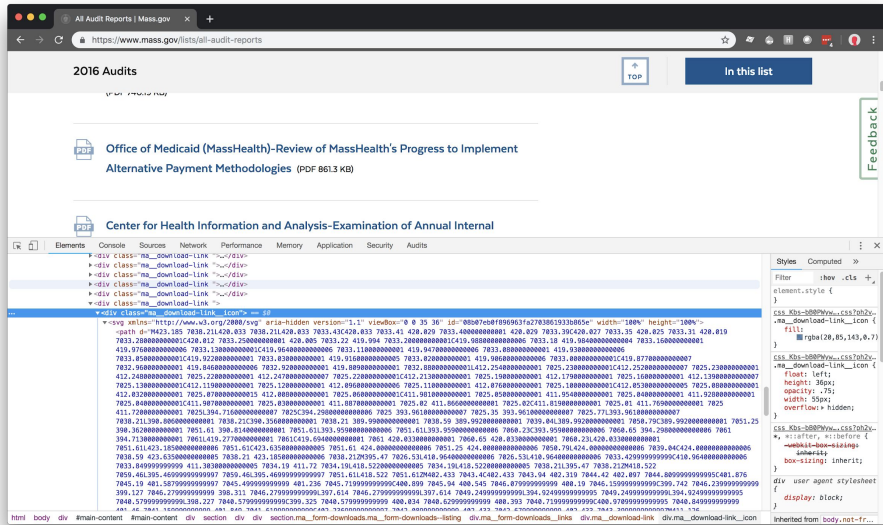
We discovered SVG was taking up a large % of the initial HTML response due to duplicated icons on the same page.

We came up with a solution to embed each icon that was used on a page 1 time in the header, and reference it multiple times with an SVG *use* statement.

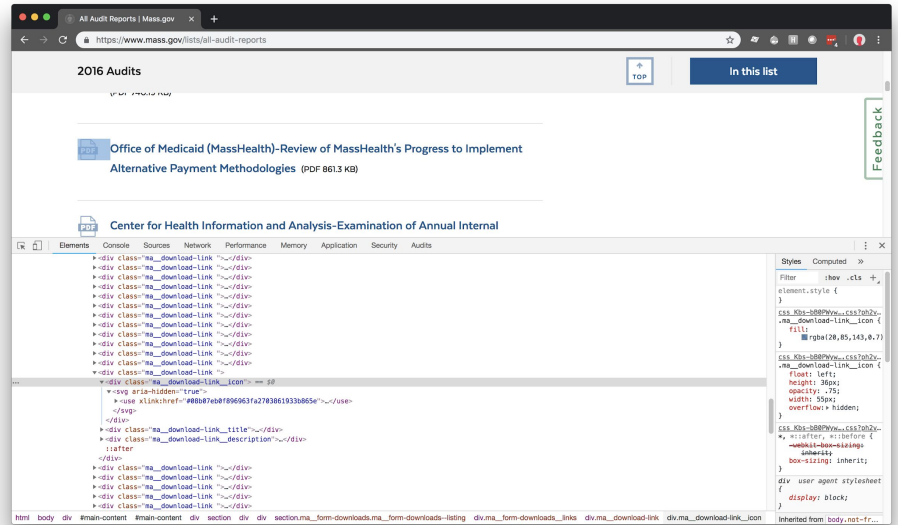
```

2 styleguide/source/_patterns/02-molecules/emergency-header.twig
View
... @@ -1,7 +1,7 @@
1 {% set alertPrefix = emergencyHeader.prefix ?: 'Emergency Alerts' %}
2 <h2 class="ma_emergency-header">
3 <span class="ma_emergency-header_label">
4 -   {% include "@atoms/05-icons/svg-alert.twig" %}
5 +   {{ icon('alert') }}
6 </span>
7 </span>
8 <span class="ma_emergency-header_title">{{ emergencyHeader.title }}</span>
9 </h2>

```



Size: > 1MB, Load: > 30s



HTML 78Kb, Load: 2.6s

April, 2018

Switch from directly inlining SVG to embedding with a *use* statement and inlining each used SVG 1x per page.

Backend cache sizes also dropped tremendously, since we were no longer storing all of that inline SVG in *cache_render*, and *cache_dynamic_page_cache*.



April, 2018

Audit, deduplicate, and clean up javascript

No longer loaded:

- 1 extra copy of Modernizr
- 1 extra copy of Handlebars
- Several deprecated custom JS files
- Google Maps (only loaded where needed)
- Google CSE (lazy loaded where needed)



April, 2018

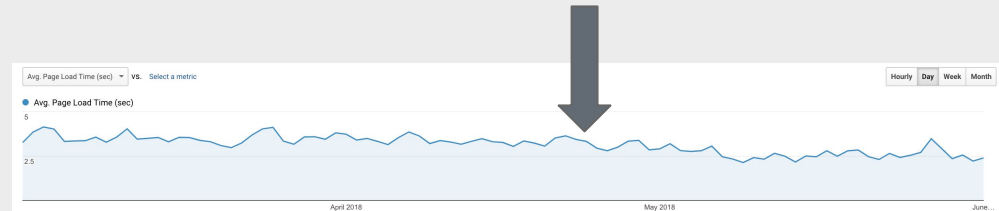
Reformulate JSONAPI AJAX request made on every page to be more specific about what's it's asking for

Before (781k before gzip):

```
https://www.mass.gov/jsonapi/node/alert?include=field_target_pages_para_ref,field_alert&filter[status][value]=1
```

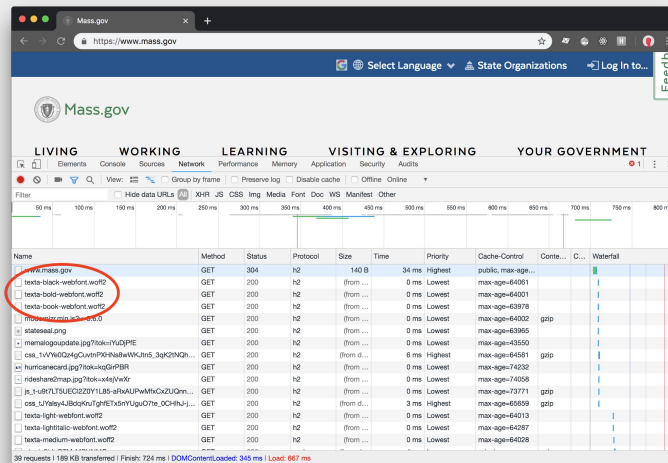
After (16k before gzip):

```
https://www.mass.gov/jsonapi/node/alert?page[limit]=250&sort=-changed&include=field_target_pages_para_ref,field_alert&filter[status][value]=1&fields[node--alert]=title,changed,entity_url,field_alert_severity,field_alert,field_target_pages_para_ref,field_alert_display&fields[paragraph--emergency_alert]=id,changed,field_emergency_alert_timestamp,field_emergency_alert_message,field_emergency_alert_link,field_emergency_alert_content&fields[paragraph--target_pages]=field_target_content_ref
```



April, 2018

Add WOFF2 webfont variants and use preloading to shorten the critical asset chain.



The screenshot shows the network tab of a browser's developer tools for the URL https://www.mass.gov. The table below represents the data shown in the network tab, with the 'texta-book-webfont.woff2' resource circled in red.

Name	Method	Status	Protocol	Size	Time	Priority	Cache-Control	Conte...	G...	Waterfal
https://www.mass.gov	GET	304	h2	140 B	34 ms	Highest	public, max-age=...			
texta-book-webfont.woff2	GET	200	h2	(from ...)	0 ms	Lowest	max-age=44021			
texta-book-webfont.woff2	GET	200	h2	(from ...)	0 ms	Lowest	max-age=44001			
texta-book-webfont.woff2	GET	200	h2	(from ...)	0 ms	Lowest	max-age=43978			
https://www.mass.gov/...	GET	200	h2	(from ...)	0 ms	Lowest	max-age=44002	gfp		
statesand.org	GET	200	h2	(from ...)	0 ms	Lowest	max-age=43965			
mamaloguodata.jpg?0x=NYUJPIE	GET	200	h2	(from ...)	0 ms	Lowest	max-age=43550			
css_1Vf6eQz4yCvvnP9XNabWkH5_3qC2N2h...	GET	200	h2	(from d...)	6 ms	Highest	max-age=44581	gfp		
humancarecard.jpg?0x=KqPBR	GET	200	h2	(from ...)	0 ms	Lowest	max-age=74232			
idashwebmedia.jpg?0x=KwKwW	GET	200	h2	(from ...)	0 ms	Lowest	max-age=44508			
_js_1-6R7LTSUCZZDY1LBS-ahAUjPwAtthCzUjGm...	GET	200	h2	(from ...)	0 ms	Lowest	max-age=73771	gfp		
css_17faley4Jl8oqKufgE7EdmYUqio7He_DChJ...	GET	200	h2	(from d...)	3 ms	Highest	max-age=44599	gfp		
texta-light-webfont.woff2	GET	200	h2	(from ...)	0 ms	Lowest	max-age=44013			
texta-light-italic-webfont.woff2	GET	200	h2	(from ...)	0 ms	Lowest	max-age=44287			
texta-medium-webfont.woff2	GET	200	h2	(from ...)	0 ms	Lowest	max-age=44028			

39 requests | 189 KB transferred | Finish: 124 ms | DOMContentLoaded: 345 ms | Load: 687 ms

April, 2018

Add WOFF2 webfont variants and use preloading to shorten the critical asset chain.

- WOFF2 is about 25% smaller.
- Preloading allows the browser to initiate fetching the fonts before it even requests the CSS file that declares their usage. Fonts are needed for the first render.
- All browsers that support WOFF2 also support preloading.



May, 2018

Use imagick for “on the fly”
image optimization

This simple toggle cut some of our larger hero images
by over 100K!

Select an image processing toolkit

Imagick image manipulation toolkit

GD2 image manipulation toolkit

▼ IMAGICK IMAGE MANIPULATION TOOLKIT SETTINGS

JPEG SPECIFIC SETTINGS

Quality

%

Higher values mean better image quality but bigger files. Quality level below 80% is not advisable when using ImageMagick.

Use Google Pagespeed Insights image optimization.
See the [guidelines](#) for further information.

Tip: Generated images can be converted to the JPEG format using the Convert effect.

Imagick resize filter

Define the resize filter for image manipulations. If you're not sure what you should enter here, leave the default settings.

Strip images of all metadata.
Eg. profiles, comments, ...



May, 2018

Use a static feedback form rather than a javascript embed

Did you find what you were looking for on this webpage? *

Yes No

If you need to report child abuse, any other kind of abuse, or need urgent assistance, please click here.

[SEND FEEDBACK](#)



May, 2018

Optimize the state seal and use it for both the header and footer.

- The “logo” is really the seal + some text, whereas it used to be an image.
- The footer now uses the same seal.
- Cut ~100k from every uncached pageview.



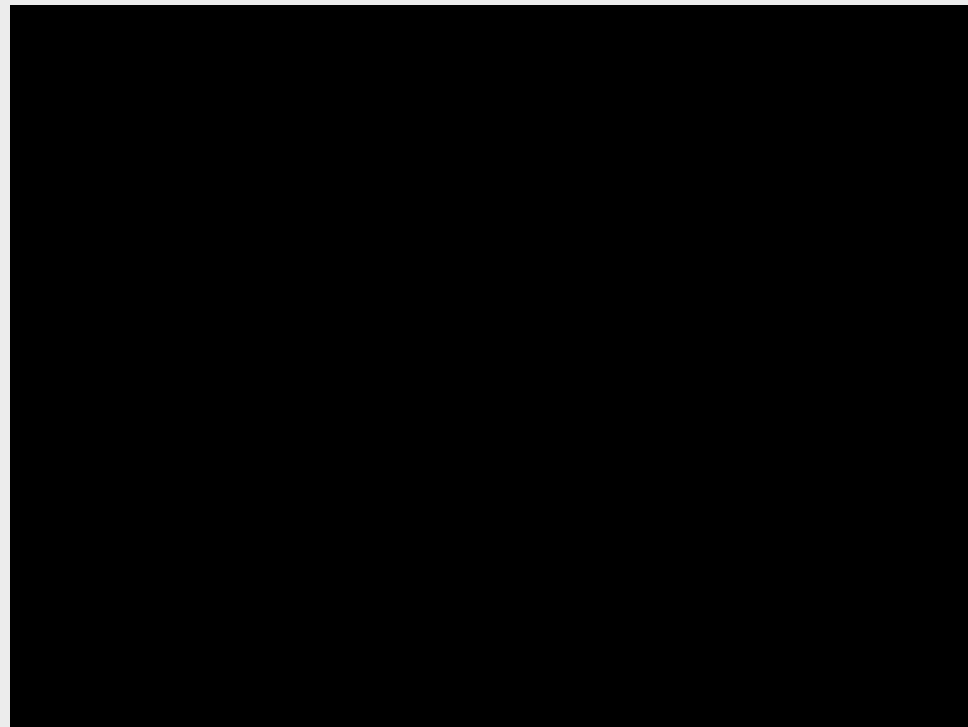
May, 2018

Speedcurve diff of individual
"Curated List" page

SIZES		SIZES
TOTAL SIZE 1124KB	45% SMALLER	TOTAL SIZE 613KB
HTML SIZE 20KB	5% SMALLER	HTML SIZE 19KB
CSS SIZE 92KB	39% SMALLER	CSS SIZE 56KB
JS SIZE 648KB	47% SMALLER	JS SIZE 341KB
IMAGES SIZE 138KB	75% SMALLER	IMAGES SIZE 34KB
 FONTS SIZE 199KB	29% SMALLER	 FONTS SIZE 142KB
FLASH SIZE 0KB	0% NO CHANGE	FLASH SIZE 0KB
VIDEO SIZE 0KB	0% NO CHANGE	VIDEO SIZE 0KB
OTHER SIZE 26KB	19% SMALLER	OTHER SIZE 21KB

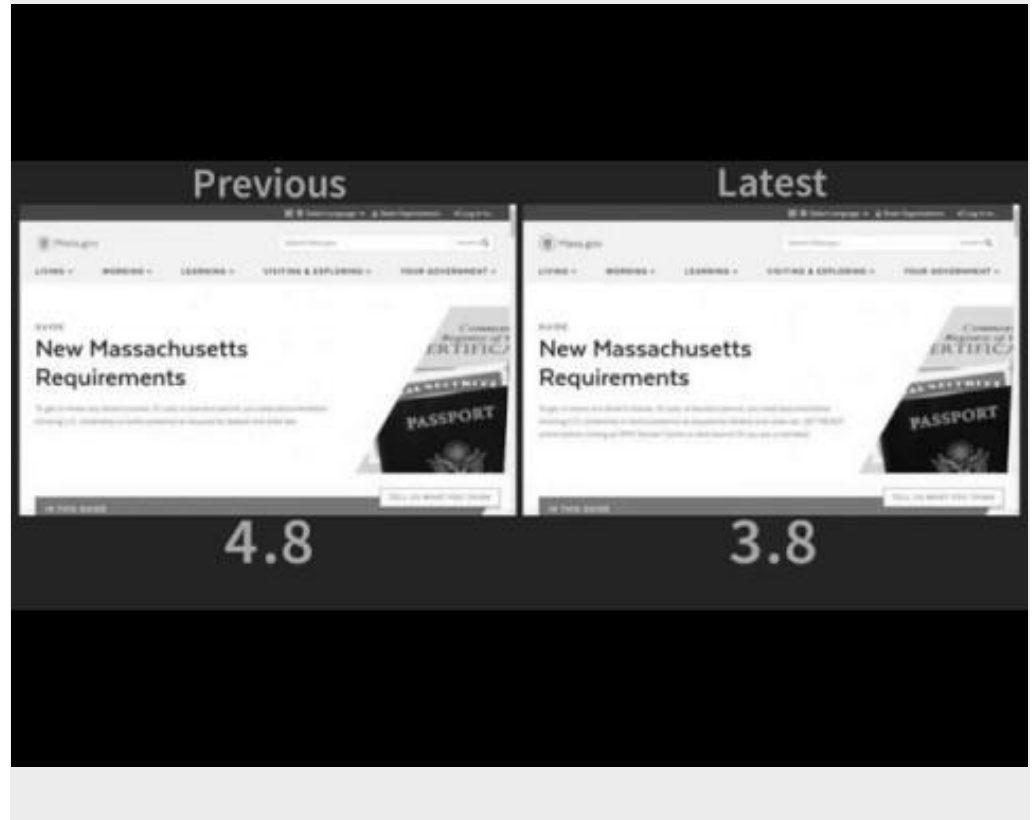
May, 2018

SpeedCurve visualization of
homepage load



May, 2018

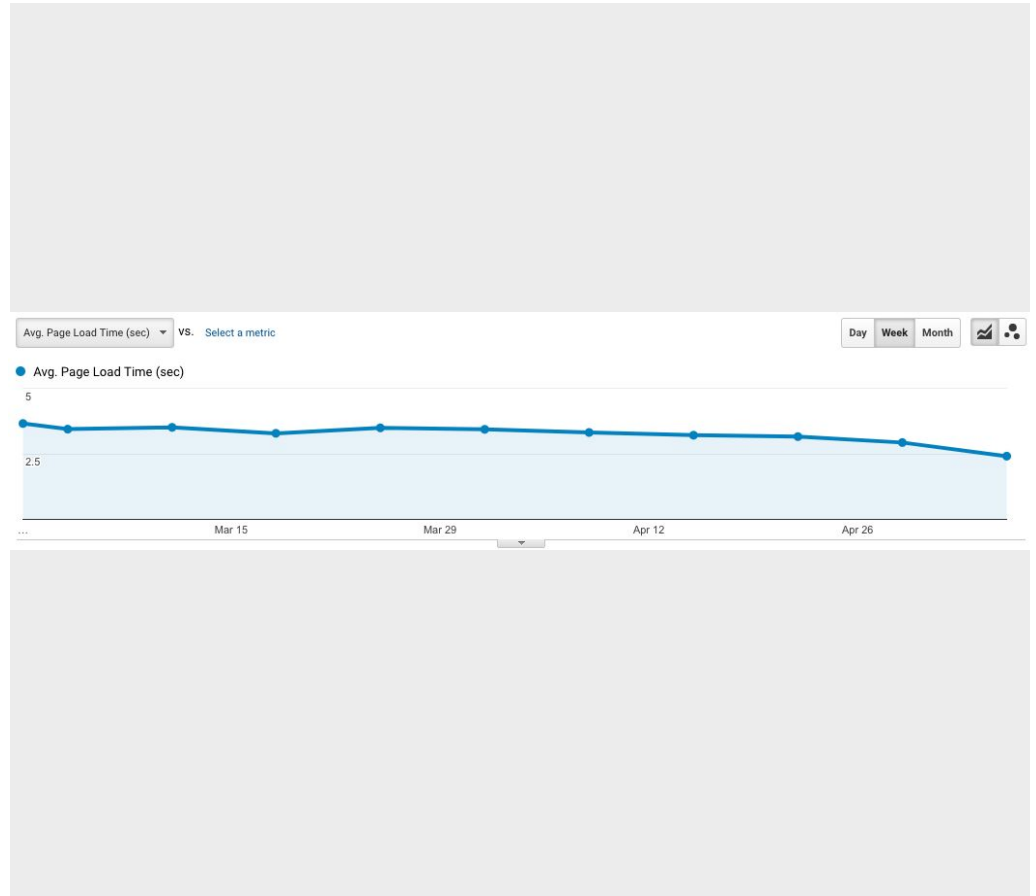
SpeedCurve visualization of
“REALID” guide



Overall Effect

35% improvement (3.6s vs 2.4s)

(as per Google Analytics)



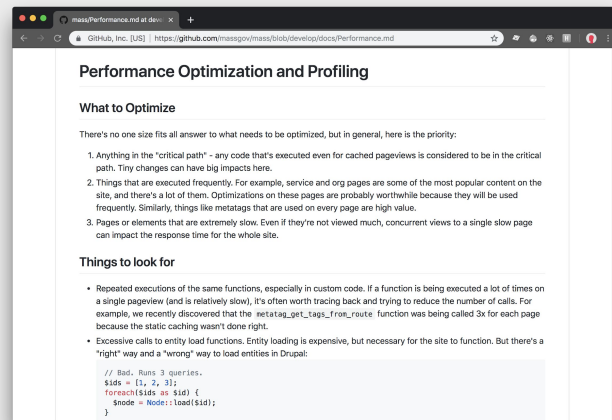
Questions?

Psst... Like what you see here? Massachusetts Digital Services is hiring!
<https://www.mass.gov/digital-services/jobs>
Visit @MassGovDigital on Twitter or Medium

December, 2017

Performance Documentation
is added to the repository

Allows us to **share** a common set of performance standards with the rest of the team. Provides a **living standard** for how we think about performance on the site.



February, 2018

Other, minor work

Batch load entities wherever you can:

```
// Bad. Runs 3 queries.
$sids = [1, 2, 3];
foreach($sids as $sid) {
    $node = Node::load($sid);
}

// Good. Runs 1 query.
$sids = [1, 2, 3];
foreach (Node::loadMultiple($sids) as $node) {

}
```

```
// Bad. Runs as many queries as there are items.
foreach($node->field_my_reference as $item) {
    $itemEntity = $item->entity;
}
// Good. Runs 1 query.
foreach($node->field_my_reference->referencedEntities() as $itemEntity) {

}
```

February, 2018

Other, minor work

Fix Expensive Queries, including:

- Anything more complicated than a simple equality check or LIKE (avoid substring matching).
- COUNT queries
- Anything that's showing up in your slow query logs