AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

NET210

Building low-latency websites with Amazon CloudFront

Kamil Bogacz (he/him)
Edge Specialist Solutions Architect
Amazon Web Services

Philippe Cheng (he/him)
Software Development Manager
Amazon Web Services

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Performance importance and challanges

Introduction to Amazon CloudFront

Caching best practices

Application-level improvements

Web delivery protocols

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Evolution of the web

* 3 MB
(2017)

700 KB
(2010)

Page bloat



Device diversity



Client-side rendering



Web builders managed services

*Source: HTTP archives

2006 Amazon.com

"Every 100 ms in delay costs 1% of sales"

Deloitte: Milliseconds make millions **2020**

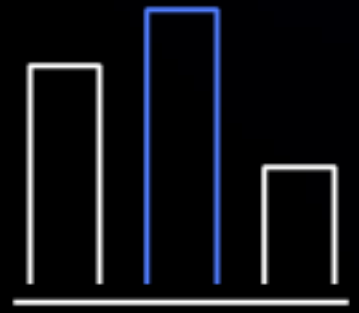
"With a 0.1 s improvement in site speed, we observed that retail consumers spent almost 10% more"

Source: https://www2.deloitte.com/ie/en/pages/consulting/articles/milliseconds-make-millions.html

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

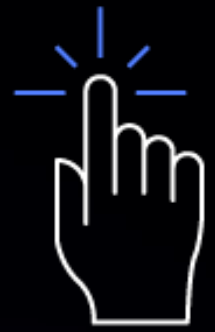
Impact on SEO



Performance impacts page score



Manage redirects



Optimized URL paths



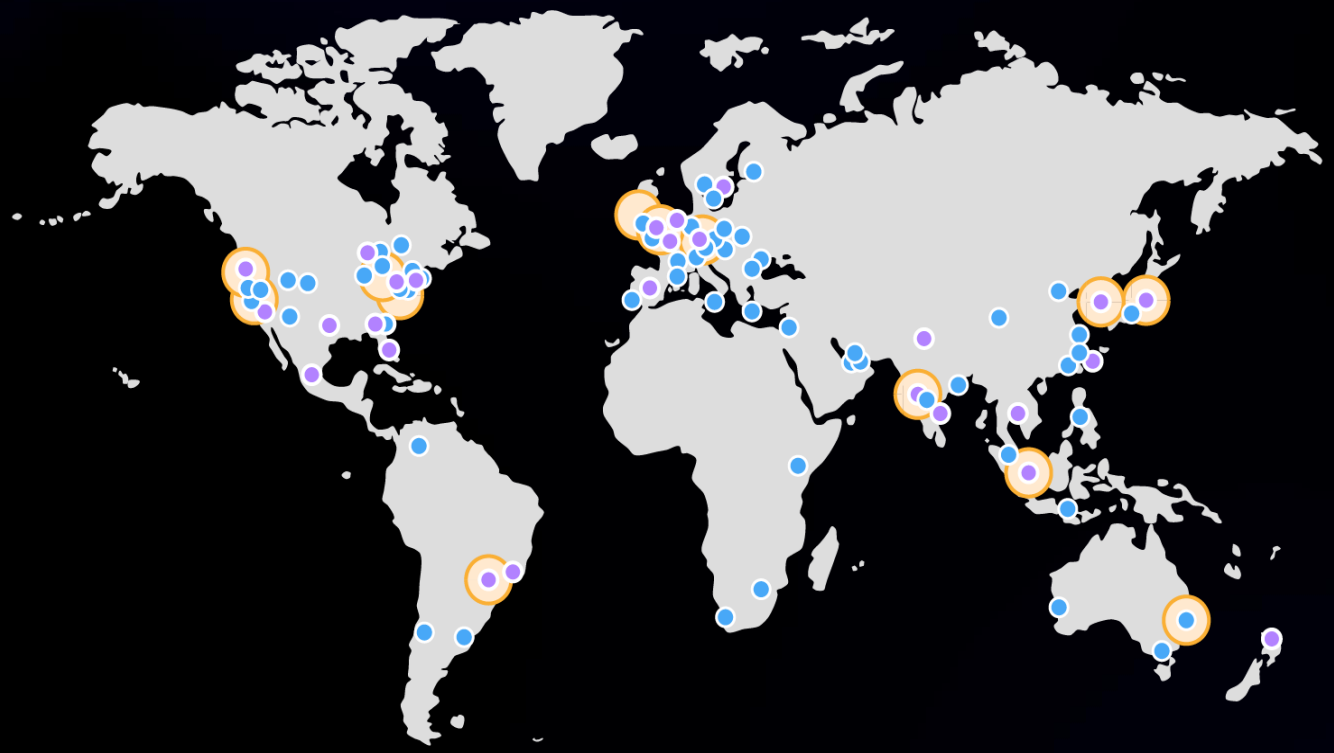
Crawler-optimized pages

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon CloudFront

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

CloudFront: Performance at the global scale



300+ points of presence

47 countries

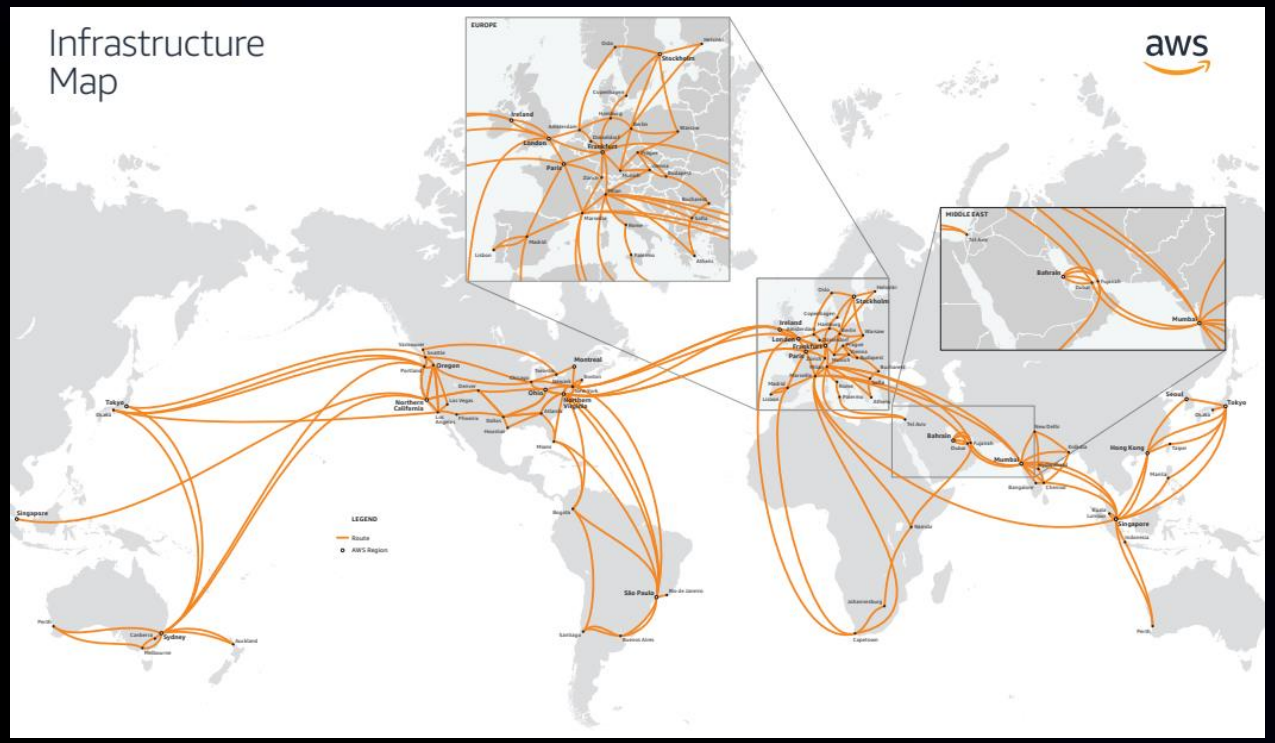
90 cities

80% YoY infrastructure growth

Embedded locations deep in ISP networks

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

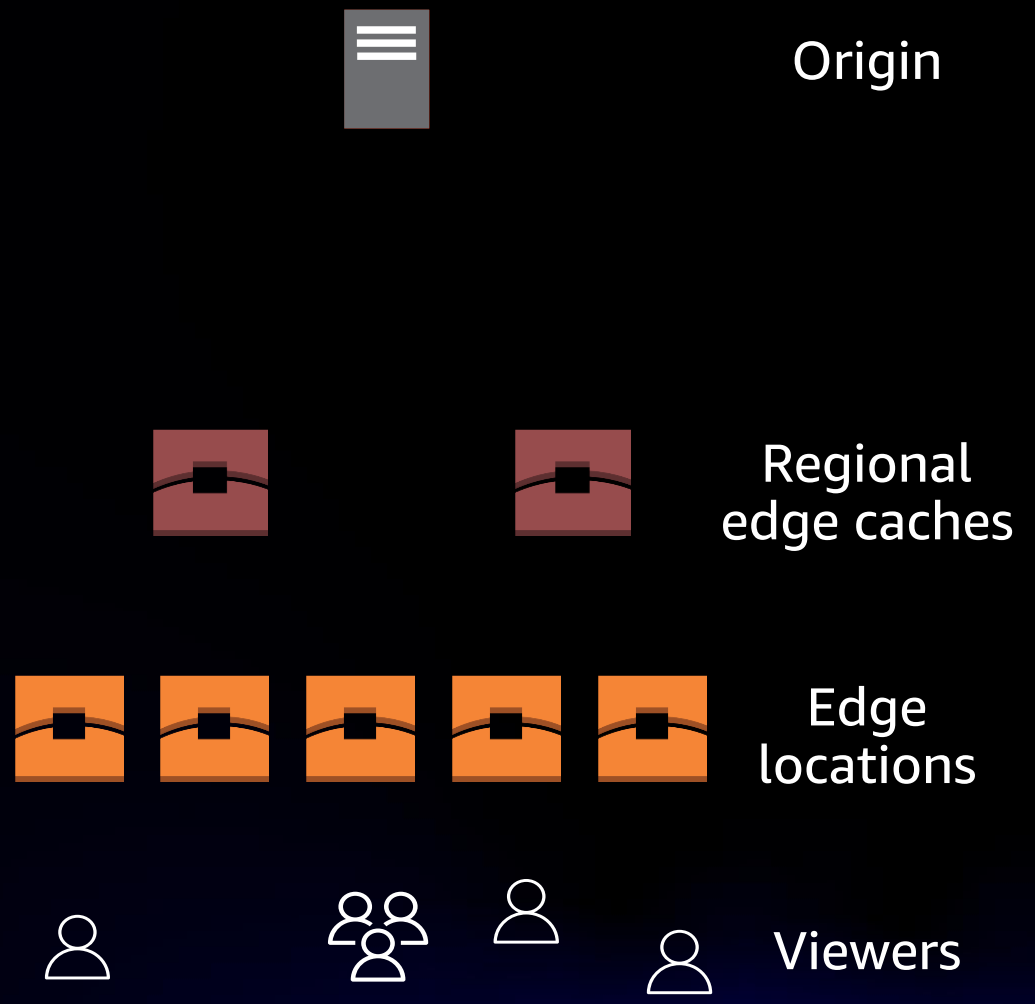
AWS network advantage



- **Private backbone network**
  - High throughput
  - Fully redundant
- **High-density POPs**
  - Well-connected to eyeball ISPs
  - Cache content close to the viewers
- **Regional edge caches**
  - Most large cache width
  - Automatic scaling

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

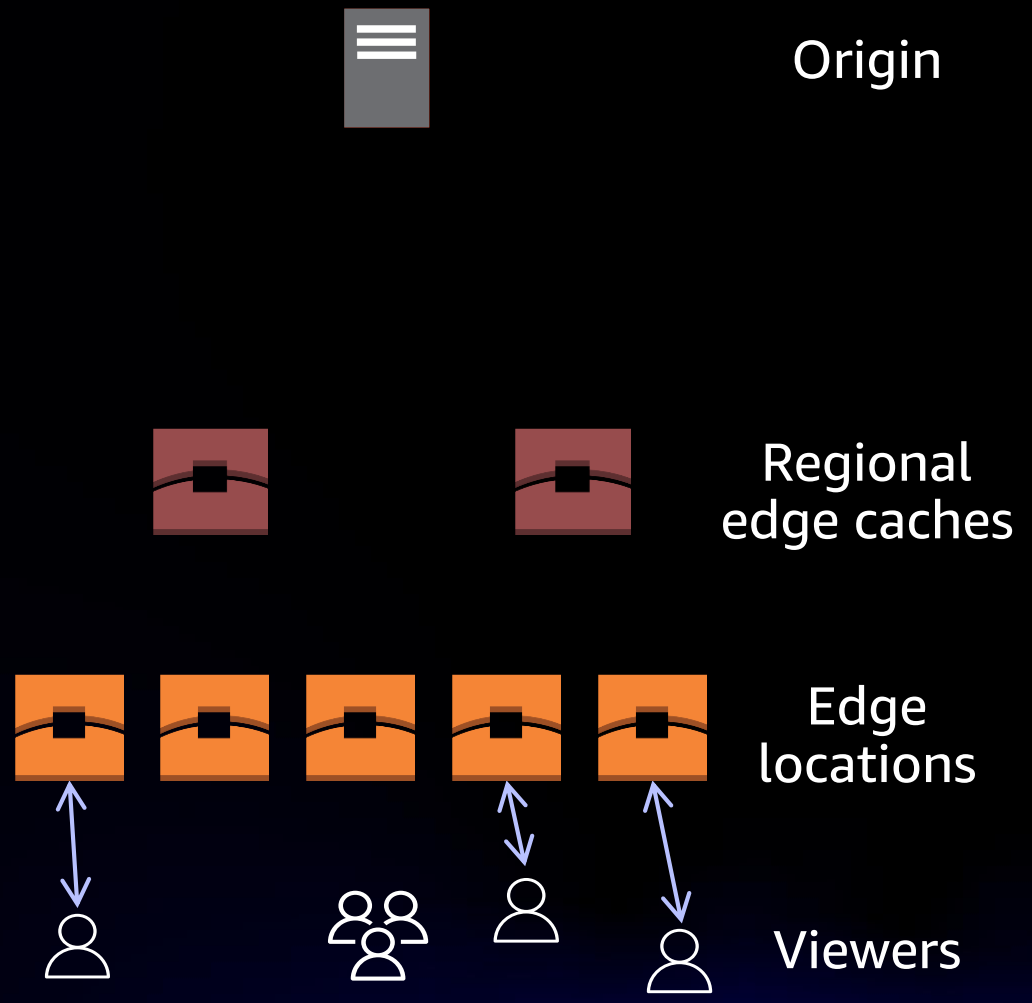
CloudFront request flow



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

CloudFront request flow

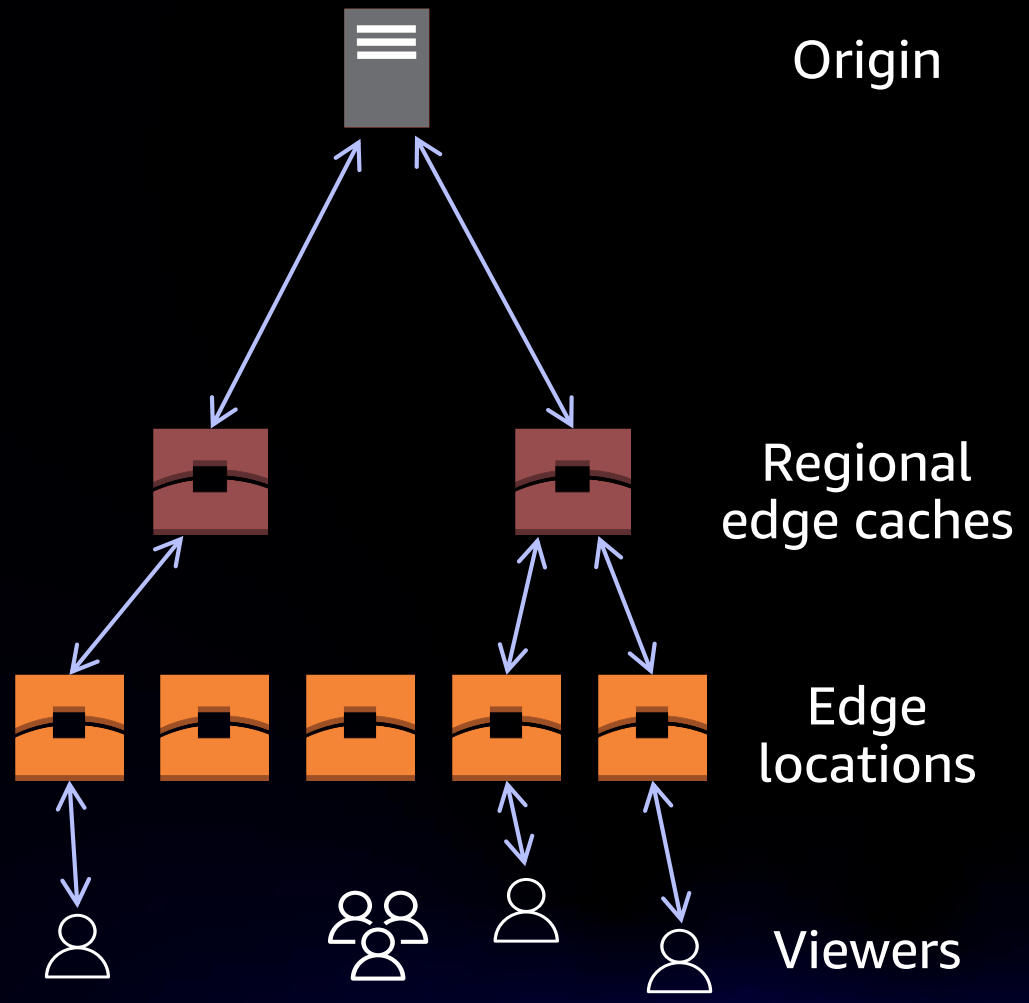
- Viewer requests routed to best PoP



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

CloudFront request flow

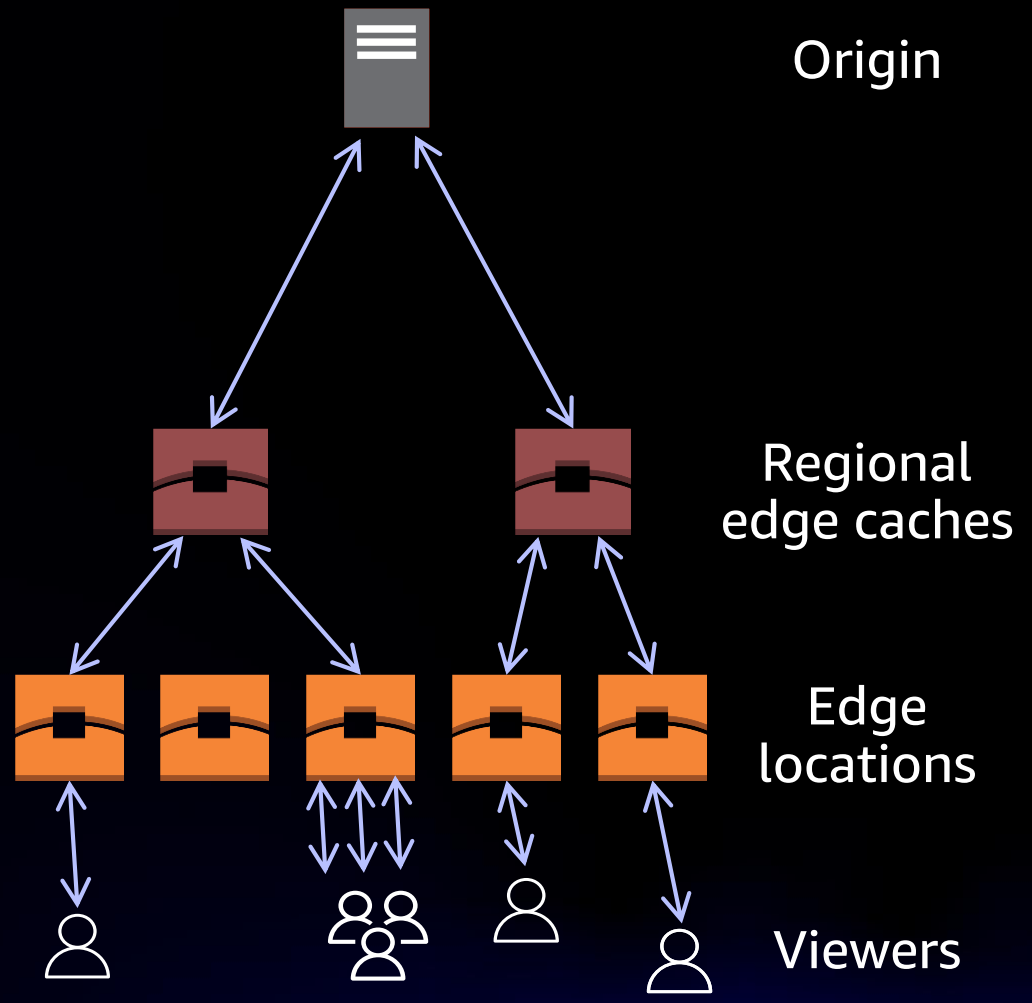
- Viewer requests routed to best PoP



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

CloudFront request flow

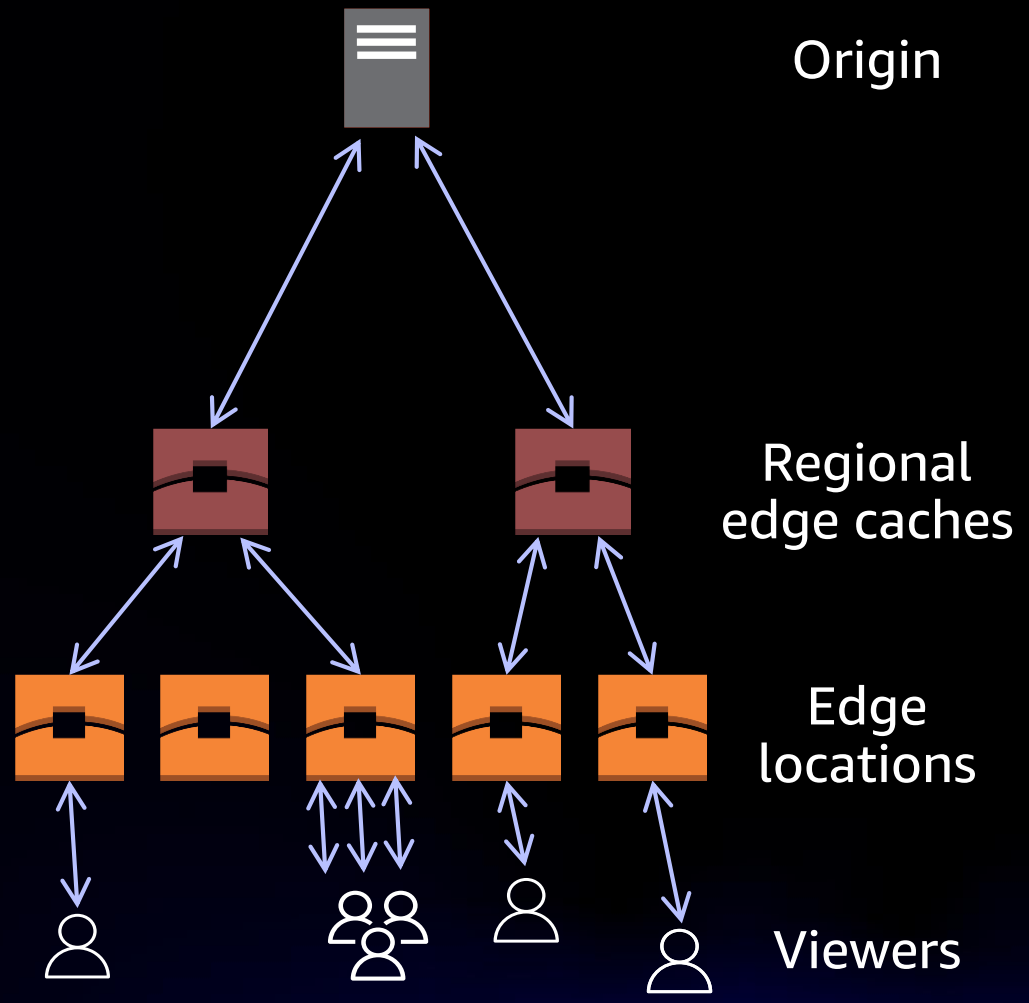
- Viewer requests routed to best PoP
- Request collapsing minimizes requests upstream



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

CloudFront request flow

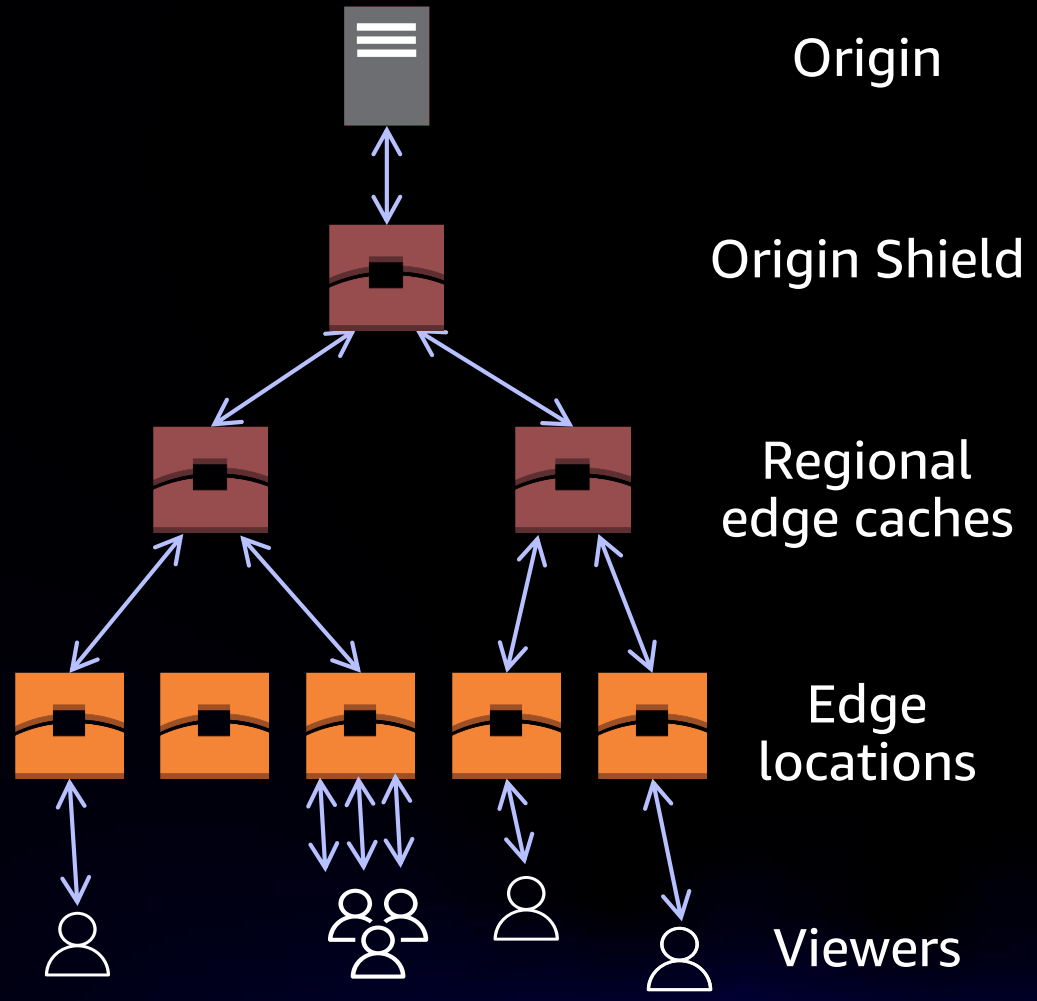
- Viewer requests routed to best PoP
- Request collapsing minimizes requests upstream
- Content cached in edge locations



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

CloudFront request flow

- Viewer requests routed to best PoP
- Request collapsing minimizes requests upstream
- Content cached in edge locations
- Origin Shield further reduces load on origin



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Layers of performance

- Caching
- Application
- Protocols



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Caching

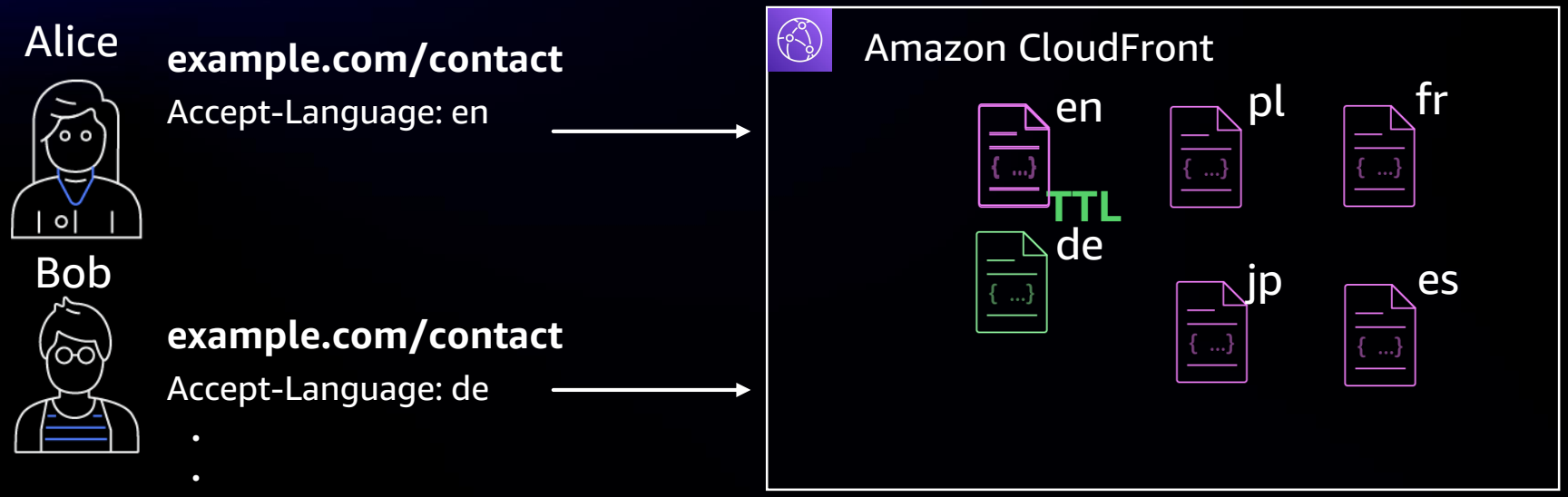
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cache hit ratio

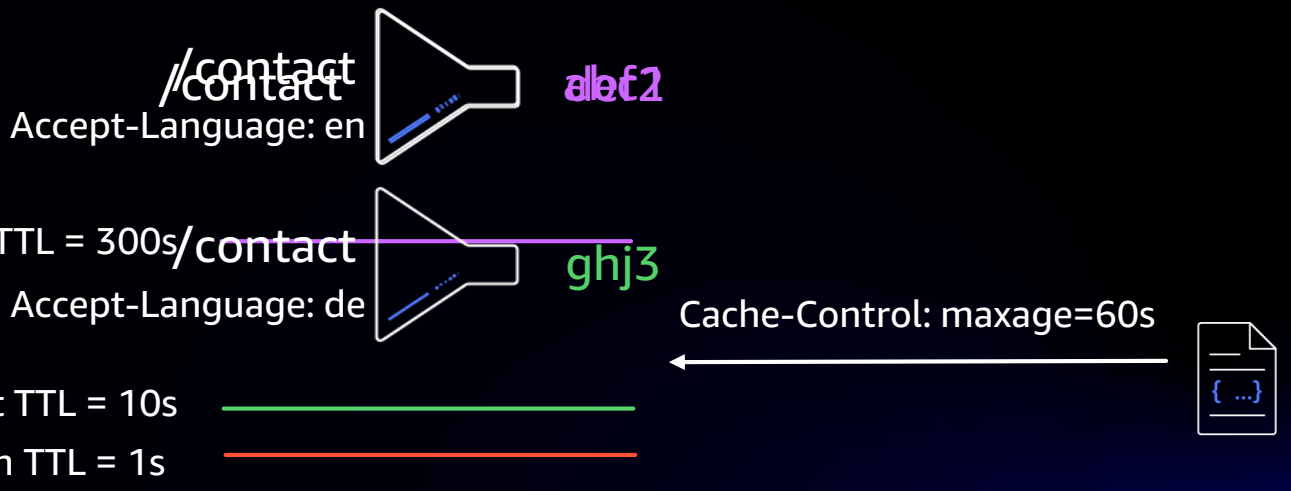
$$\text{Cache hit ratio} = \frac{\text{requests served from the cache}}{\text{total requests}}$$

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Principles of caching

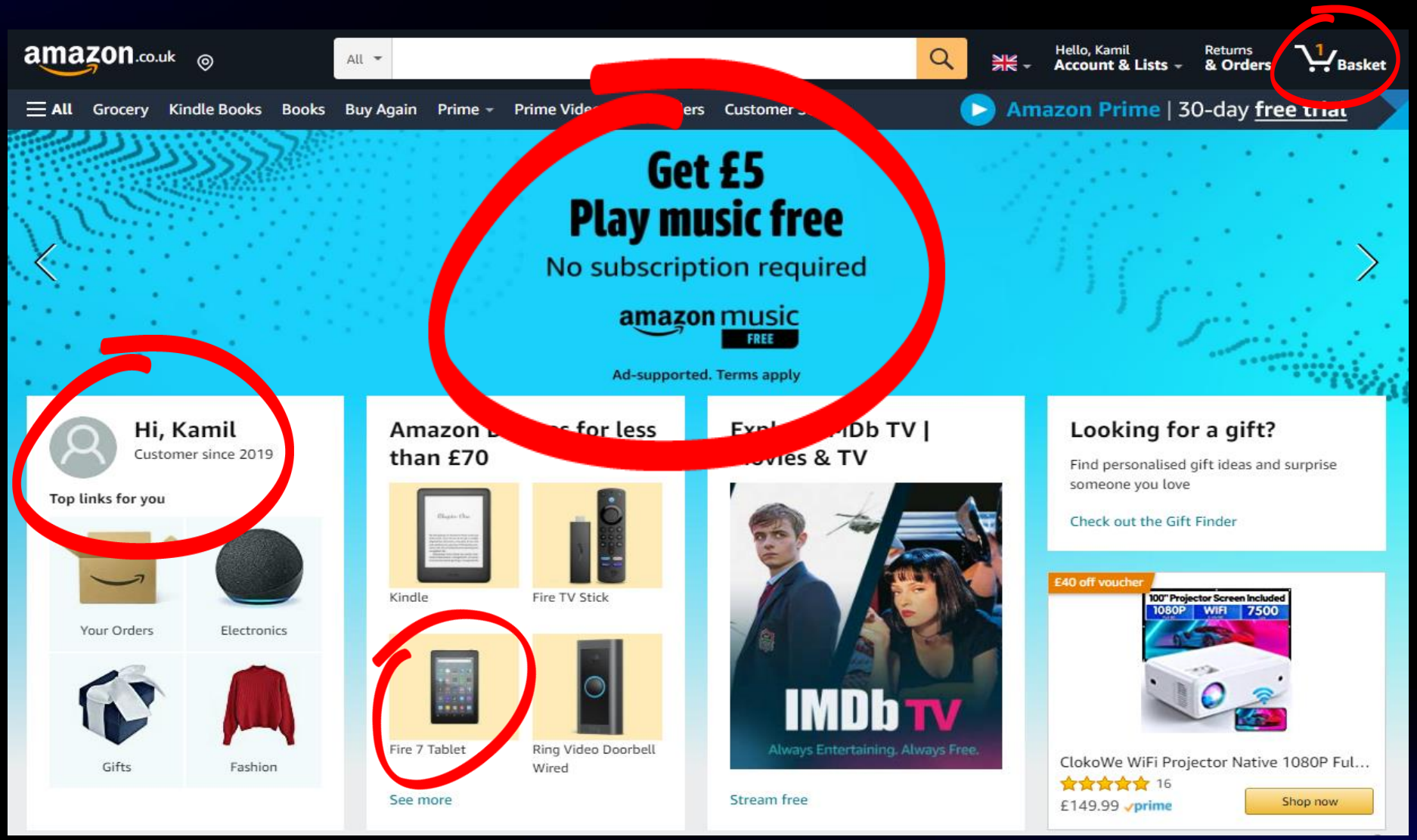






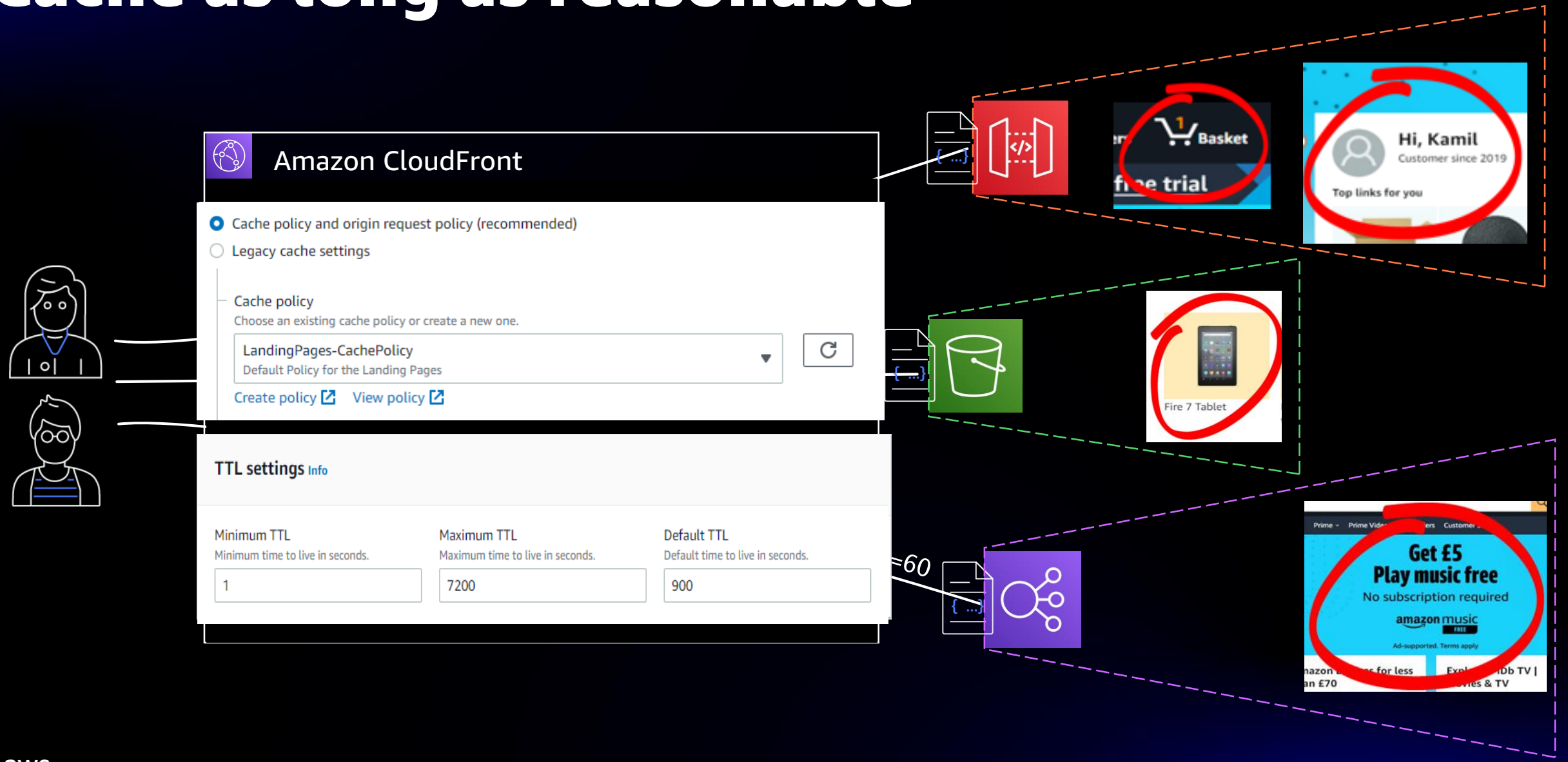
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cache as much as possible



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

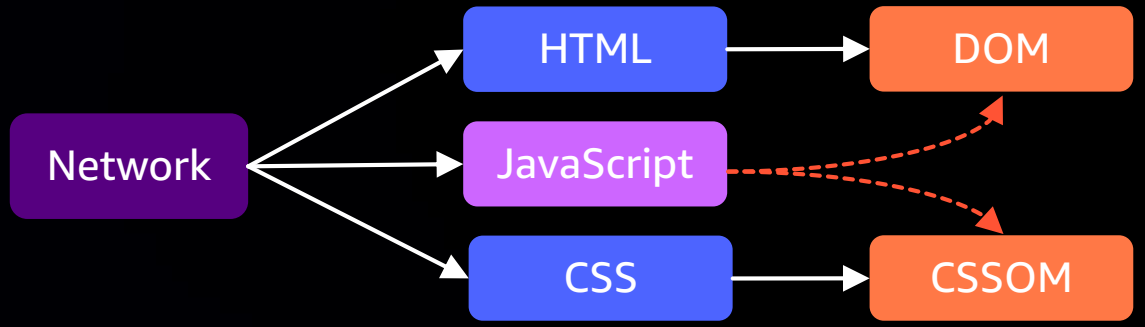
Cache as long as reasonable



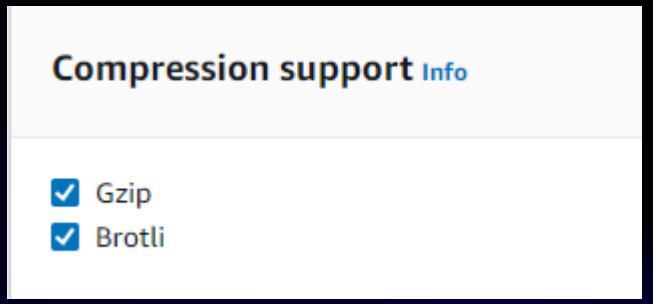
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Compression advantage

- Look after critical rendering path



- Provide your text files in multiple compression formats
  - Brotli can bring up to 25% of improvement over Gzip
- Adjust cache key policies accordingly



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

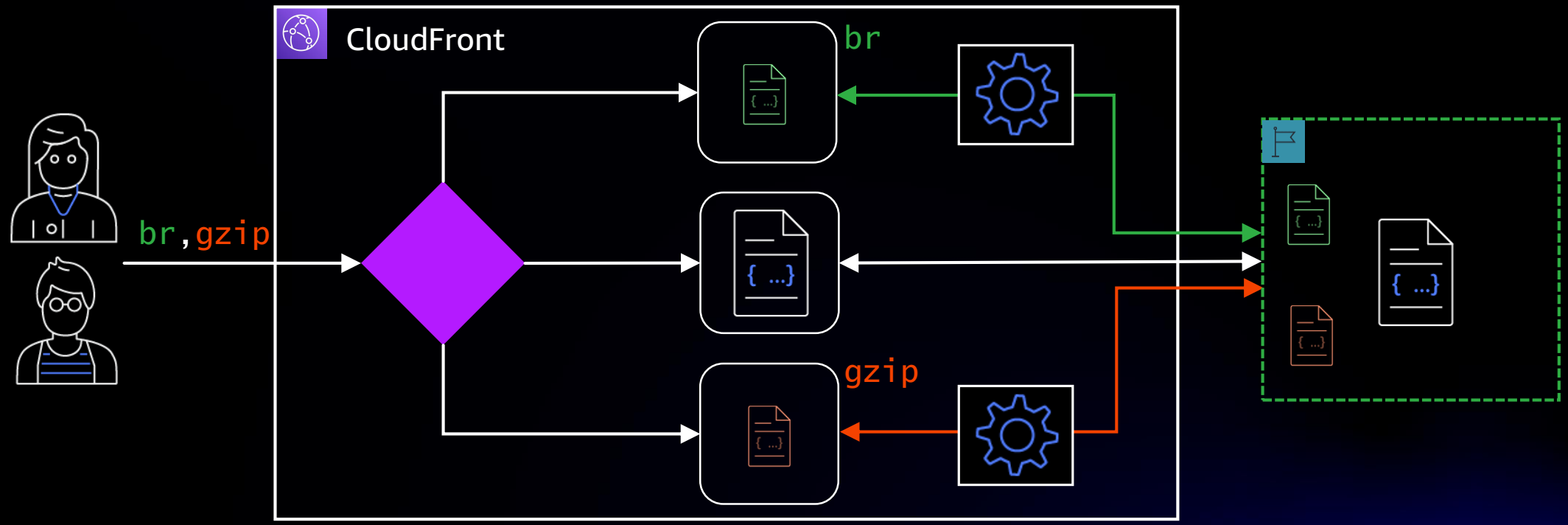
Compression in CloudFront



```
Accept-Encoding: br,gzip,compress
Accept-Encoding: br,deflate,gzip
Accept-Encoding: gzip;q=0.8,*;q=0.1
Accept-Encoding: deflate,gzip
```



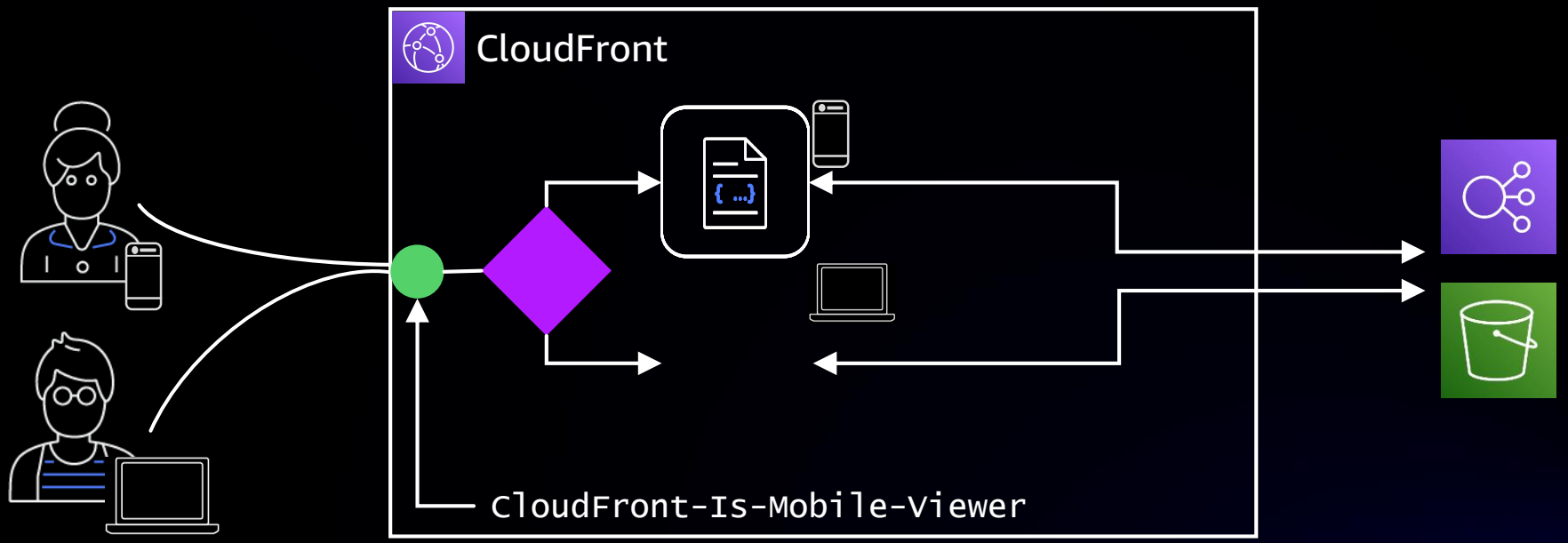
```
Accept-Encoding: br,gzip
Accept-Encoding: gzip
Accept-Encoding: br
Accept-Encoding: identity
```



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Adapt content to the viewer

- Consider principles of *responsive design*
  - Serve device-optimized version of the page and assets
- **Do not** vary the cache on the User-Agent



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

CloudFront-generated headers

Device characterization

- CloudFront-Is-Android-Viewer
- CloudFront-Is-Desktop-Viewer
- CloudFront-Is-IOS-Viewer
- CloudFront-Is-Mobile-Viewer
- CloudFront-Is-SmartTV-Viewer
- CloudFront-Is-Tablet-Viewer

Viewer's location

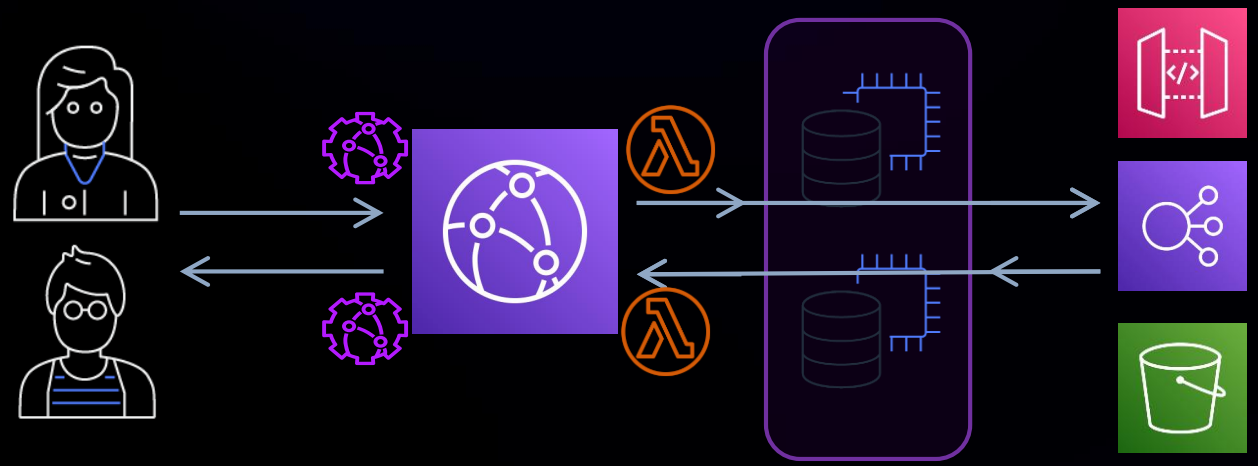
- CloudFront-Viewer-City
- CloudFront-Viewer-Country
- CloudFront-Viewer-Country-Name
- CloudFront-Viewer-Country-Regiono
- CloudFront-Viewer-Country-Region-Name
- CloudFront-Viewer-Latitude
- CloudFront-Viewer-Longitude
- CloudFront-Viewer-Metro-Code
- CloudFront-Viewer-Postal-Code
- CloudFront-Viewer-Time-Zone

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Application

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Edge serverless applications



When to use it?

- Undifferentiated functionality
- Improve cacheability and origin offload
- Route to the most optimal endpoint

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

When serverless meets CDN

WHY WOULD YOU NEED COMPUTE AT THE EDGE?



Performance gain



Agility



Origin independence



Cost optimization

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Edge serverless compute options

Lambda@Edge

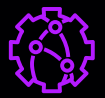


- Long-running workloads
- More complex use cases
- Scaling considerations

Use cases

- Response body manipulation
- Server-side rendering
- Integrations with SaaS vendors
- Advanced multi-origin routing

CloudFront Functions



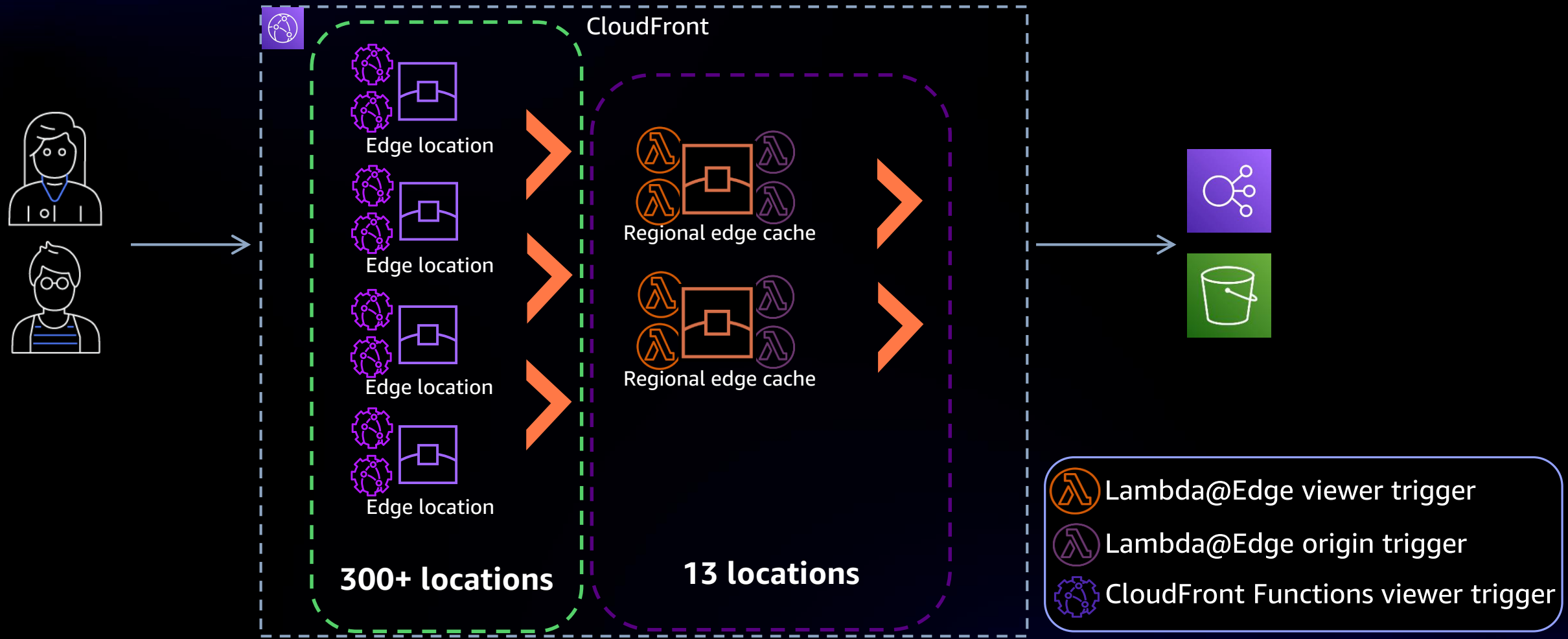
- Lightweight functions
- Stateless and self-contained logic
- Highest scale applications

Use cases

- Header and URL manipulations
- Redirects at the Edge
- Request authorization
- Viewer prioritization

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

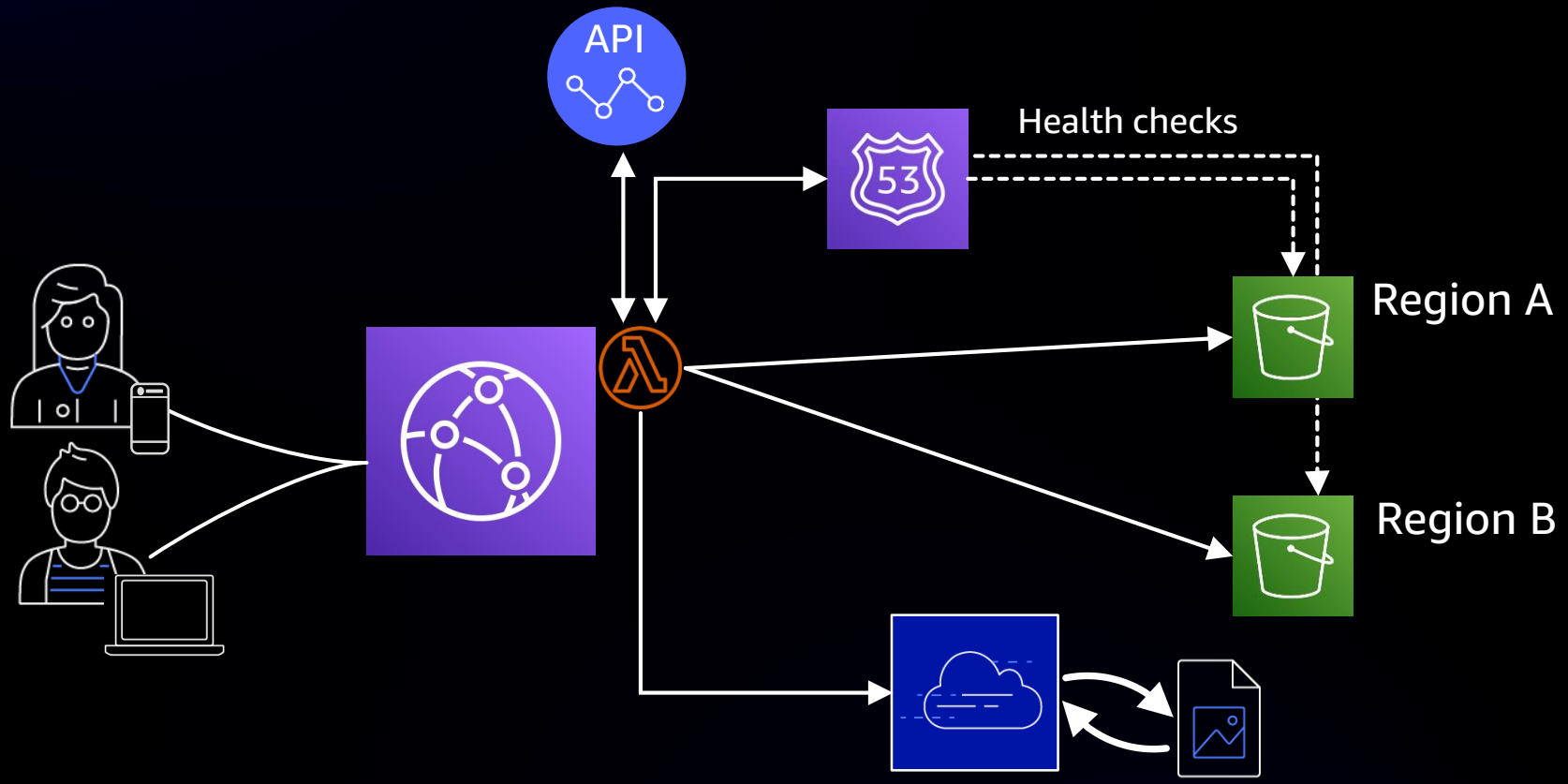
CloudFront serverless triggers



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Edge serverless example

contentful

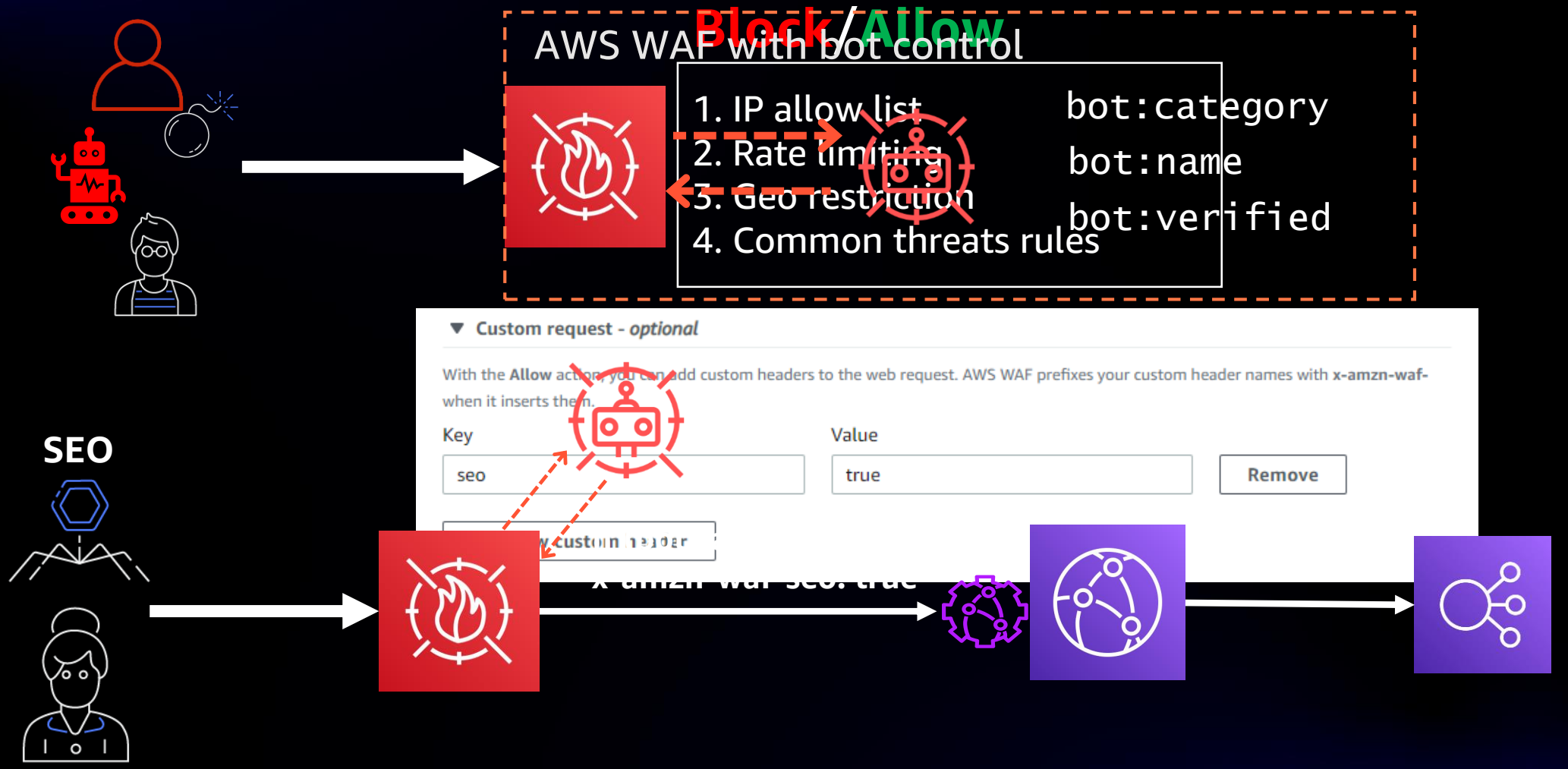


"To achieve faster delivery and response times, CloudFront and Lambda@Edge play a vital role in delivering assets from Contentful content platform worldwide with low latency, while running custom code simultaneously."

aws

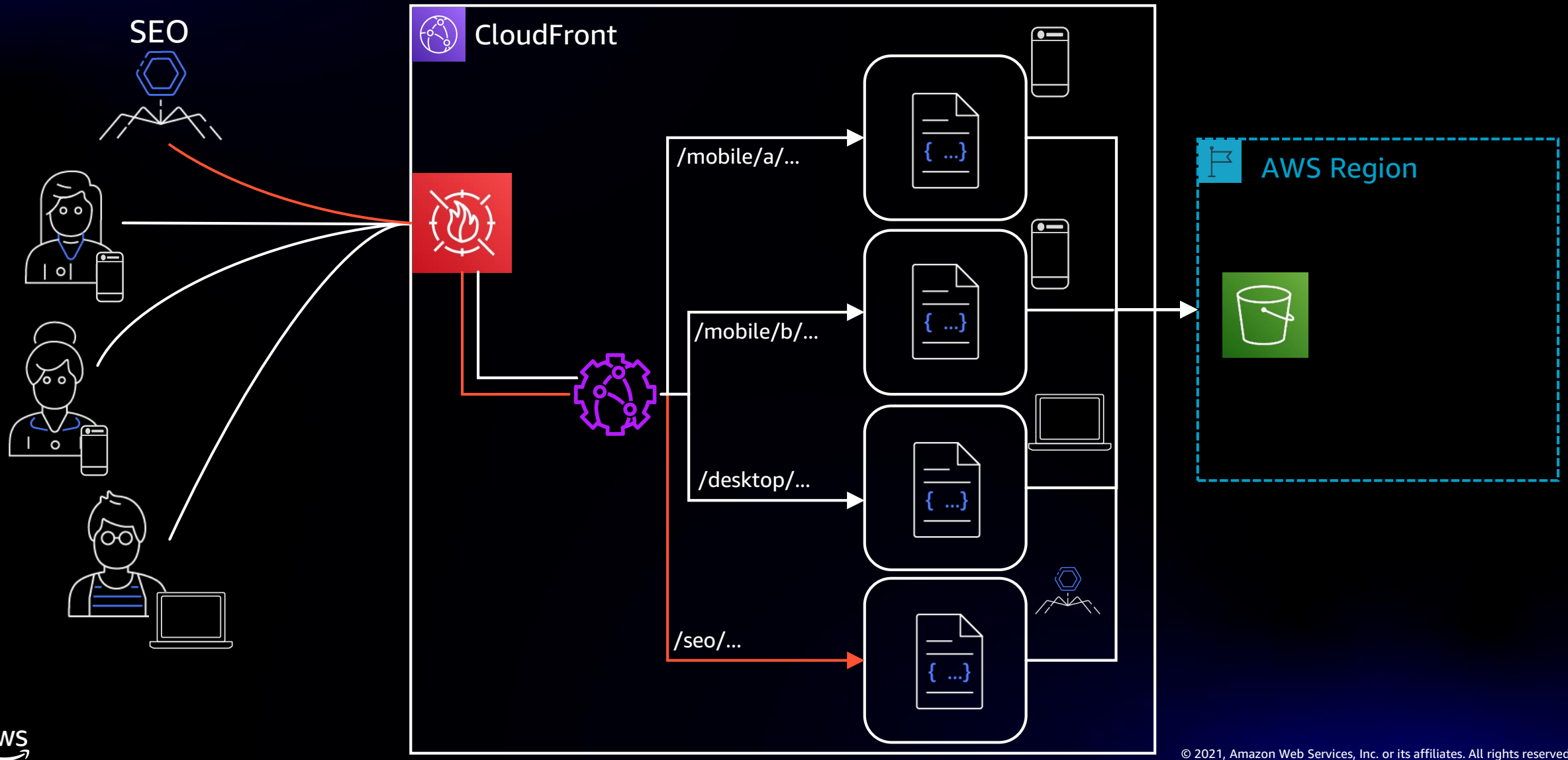
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS WAF – more than web security



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Optimize for different audiences



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

CloudFront > Distributions > E2NWJXDC52BUHZ

E2NWJXDC52BUHZ

General | Origins | Behaviors | Error pages | Geographic restrictions | Invalidations | Tags

Behaviors

Save | Move up | Move down | Edit | Delete | Create behavior

Filter behaviors by property or value

	Preced...	Path pattern	Origin or origin group	Viewer protocol policy	Cache policy name	Origin request policy name
○	0	Default (*)	bogaczkb-abc-demo.s3.eu-wes...	HTTP and HTTPS	Managed-CachingOptimized	ABC-Demo

© 2008 - 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

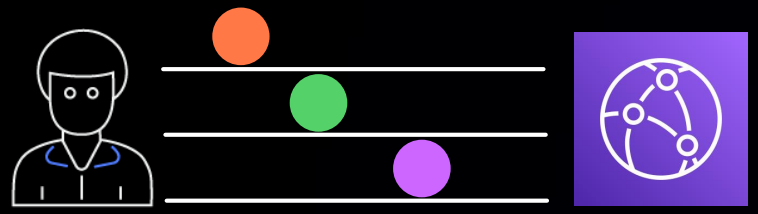
Protocols

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

HTTP evolution

HTTP/1.1

- Parallel TCP connections
- One request at a time
- Prone to head-of-line blocking
- Ideal for large downloads



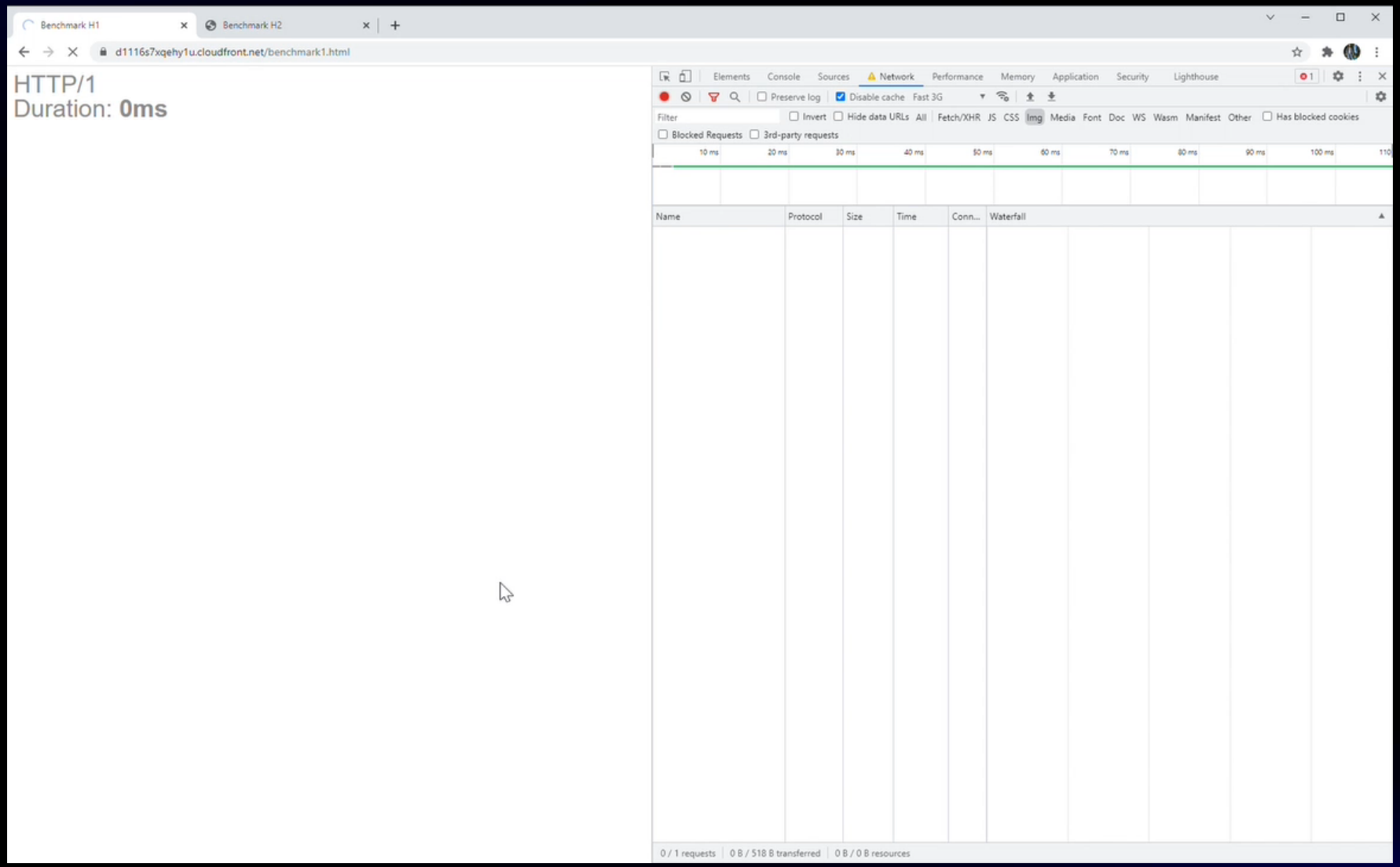
HTTP/2

- Single TCP connections
- Request multiplexing
- Header compression
- Streams prioritization
- Ideal for web content



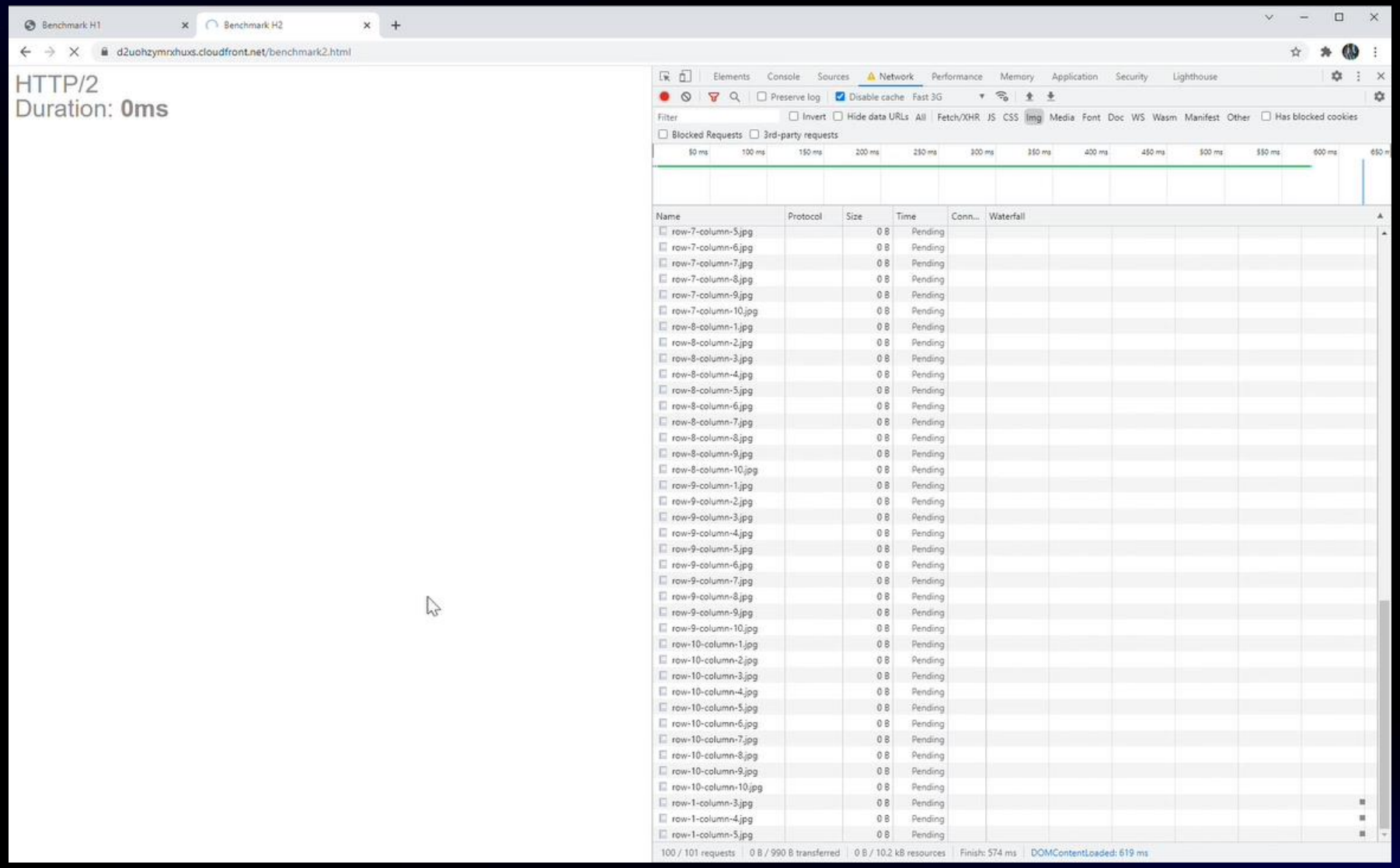
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

HTTP 1.1



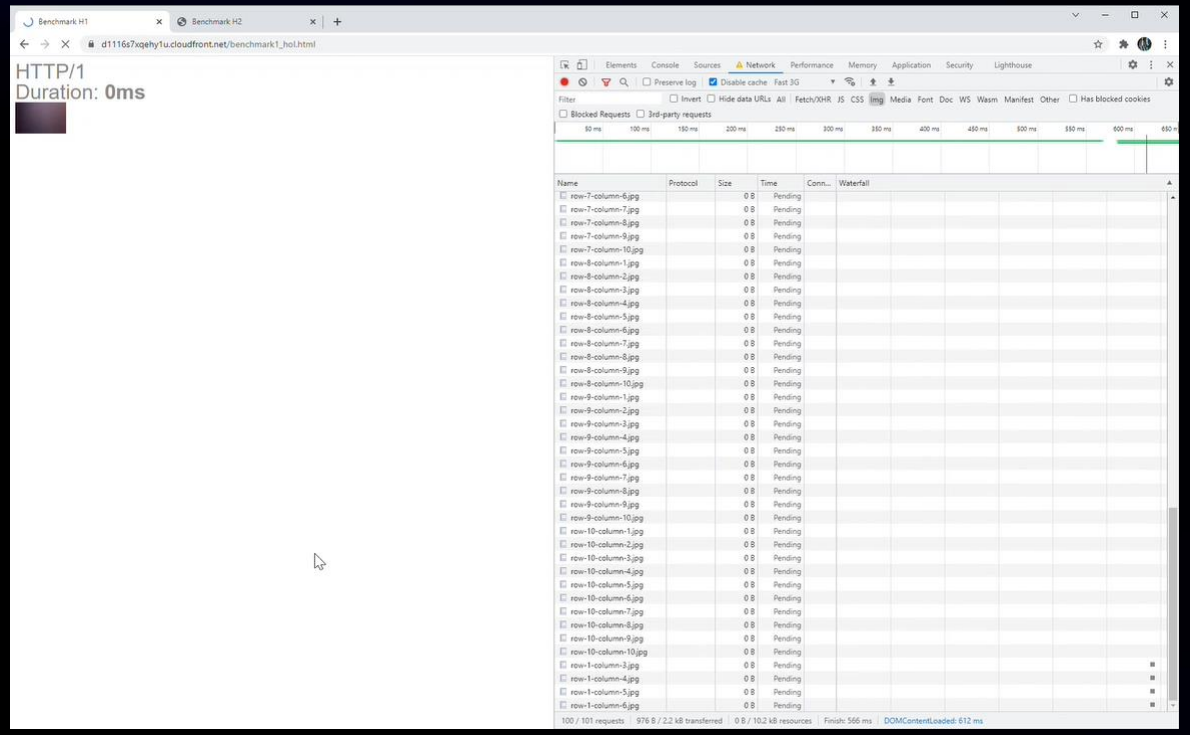
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

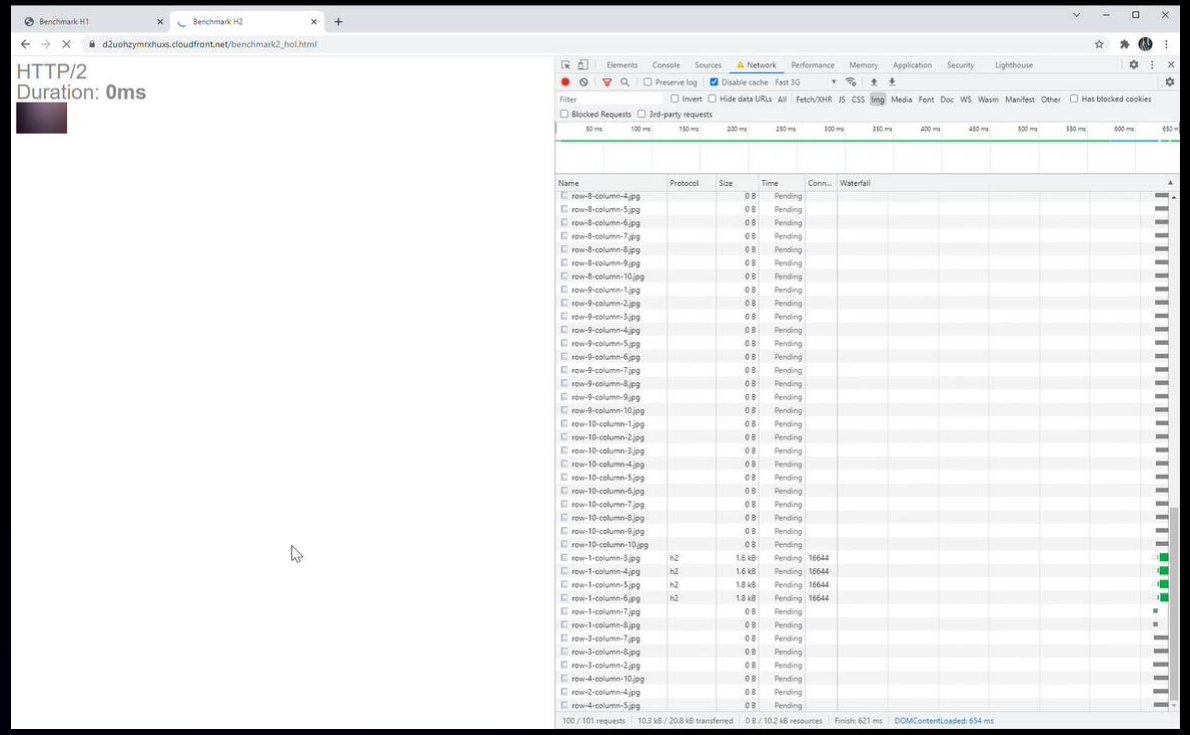
HTTP/2



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

HTTP head-of-line blocking

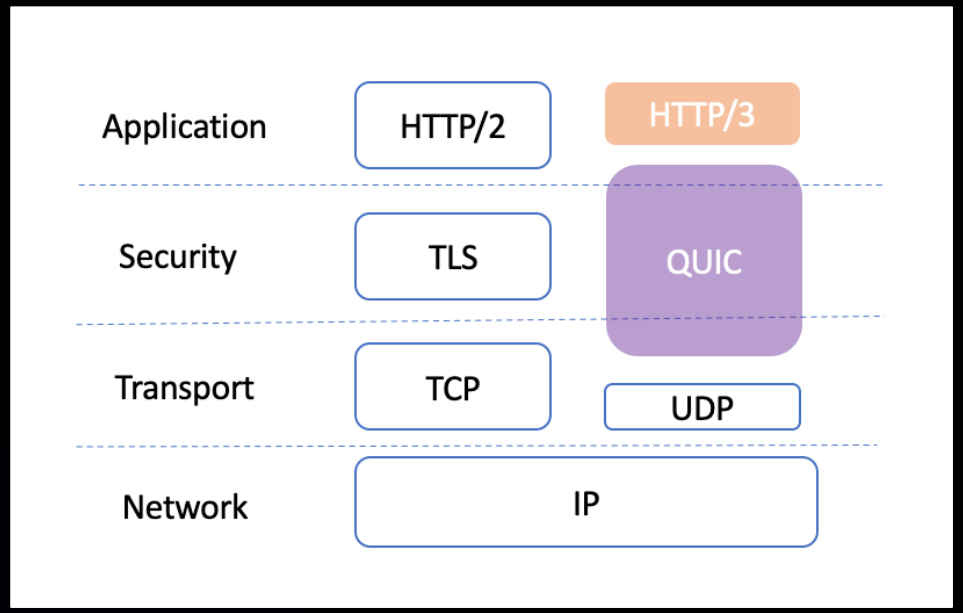




© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Next iteration: HTTP/3

- Built on top of QUIC, a transport protocol to make web connections faster
- UDP – simple, fast, connection-less
- Head-of-line blocking avoidance – independent streams
- Uses TLS 1.3 – 1-RTT handshake and more secure, 0-RTT for resumed session
- Connection migration – improved performance for mobile and emerging markets



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Recap

- CloudFront addresses challenges of modern websites
- Stay close to the viewers with global network of high-density POPs
- Define tailored caching strategy for different content and devices
- Fine-grain traffic characterization with native headers and AWS WAF
- Complement origin functionalities with edge serverless
  - Improve caching efficiency with traffic segmentation
  - Complement origin functionalities

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Kamil Bogacz
bogaczkb@amazon.com

Philippe Cheng
chengphi@amazon.com

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.