FastRoute

A Scalable Load-Aware Anycast Routing Architecture for Modern CDNs

Authors:
Ashley Flavel, Pradeepkumar Mani, David A. Maltz, and Nick Holt, *Microsoft;*
Jie Liu, *Microsoft Research;* Yingying Chen and Oleg Surmachev, *Microsoft*

Presented by: Wojciech Kordalski

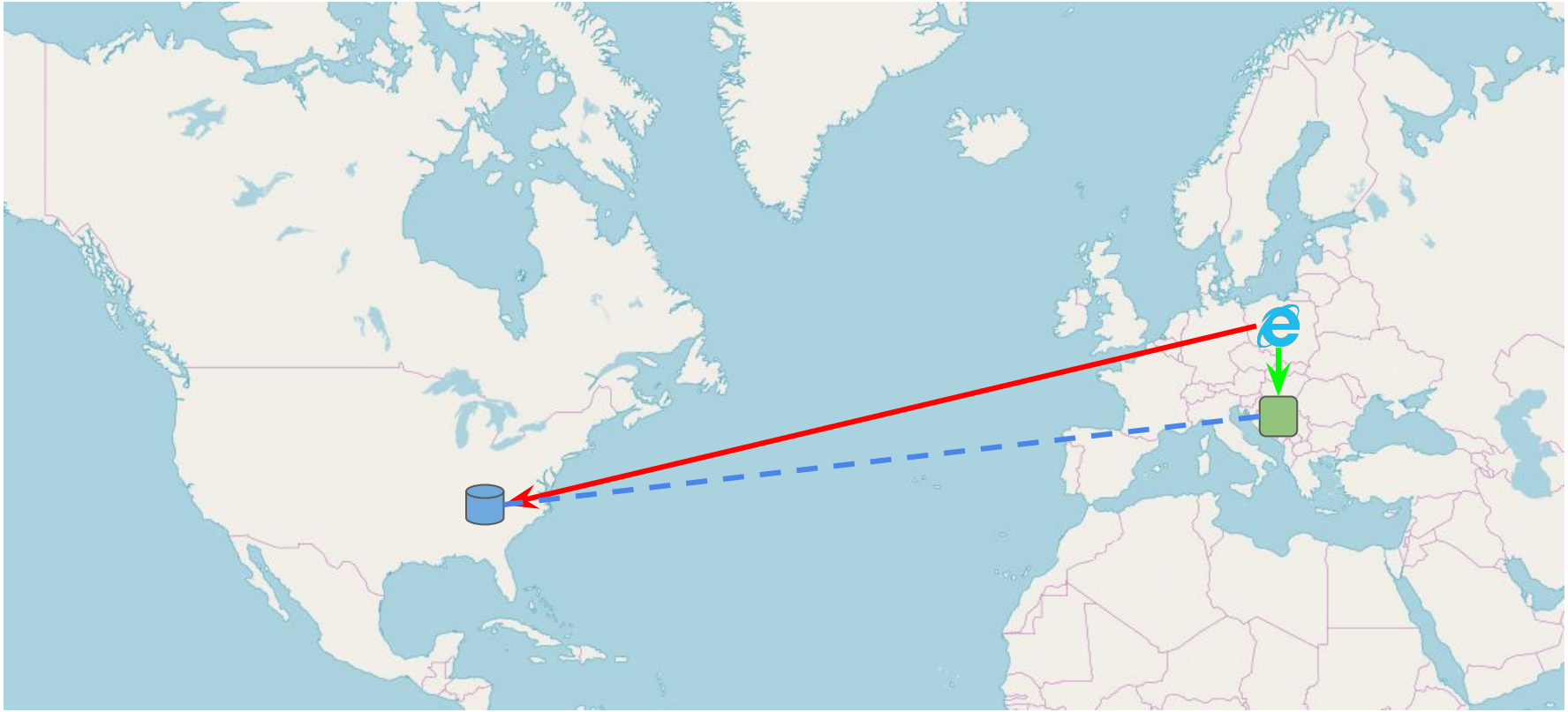
1. CDN
2. Proxy selection techniques
3. FastRoute architecture
4. In production
5. Future work

CDN – Why do we need it?

CDN – Why do we need it?

To improve user-perceived application performance.

CDN – Why do we need them?



1. CDN
2. Proxy selection techniques
3. FastRoute architecture
4. In production
5. Future work

Internet Map

We project the whole address space onto the set of available proxies.
DNS returns different IP addresses of the proxy depending on the IP of the user.

- Flexible control over on which proxy user traffic lands.

But:

- It requires global knowledge to analyse user latency data as well as real-time load and health metrics of nodes to decide where to route traffic.
- Lack of granularity of DNS based responses.
- Support for IPv6 introduces additional complexity.

Anycast routing

Anycast is a routing technique that utilizes the fact that routers running “the de-facto standard inter-domain routing protocol in the Internet” (BGP) select the shortest path to a destination IP prefix.

Consequently, if multiple destinations claim to be a single destination, routers independently examine the characteristics of the multiple available routers and select the shortest one.

Multiple hosts are created in different locations responding on the same IP address. The Internet routing protocols will route packets to the closest host.

Anycast DNS

- DNS server is co-located with each proxy.
- All DNS servers responds on the same IP.
- The BGP protocol chooses the closest DNS server.
- The DNS server returns IP address of “its proxy”.

Anycast DNS

- DNS server is co-located with each proxy.
- All DNS servers responds on the same IP.
- The BGP protocol chooses the closest DNS server.
- The DNS server returns IP address of “its proxy”.

But: “closest to the recursive DNS that handles user request” instead of: “closest to the user”.

Anycast TCP

- All proxies respond on the same IP address.
- The Internet routing protocols determines the closest one.

Anycast TCP

- All proxies respond on the same IP address.
- The Internet routing protocols determines the closest one.

Advantages:

- It is simple.
- Each proxy runs independently to others.
- It is used by many modern CDNs including Edgecast and CloudFlare.

Disadvantages:

- No control over on which proxy the user traffic lands.

1. CDN
2. Proxy selection techniques
3. FastRoute architecture
4. In production
5. Future work

FastRoute

1) deliver a low-latency routing scheme that perform better than existing CDN
2) build an easy-to-operate, scalable and simple system

FastRoute

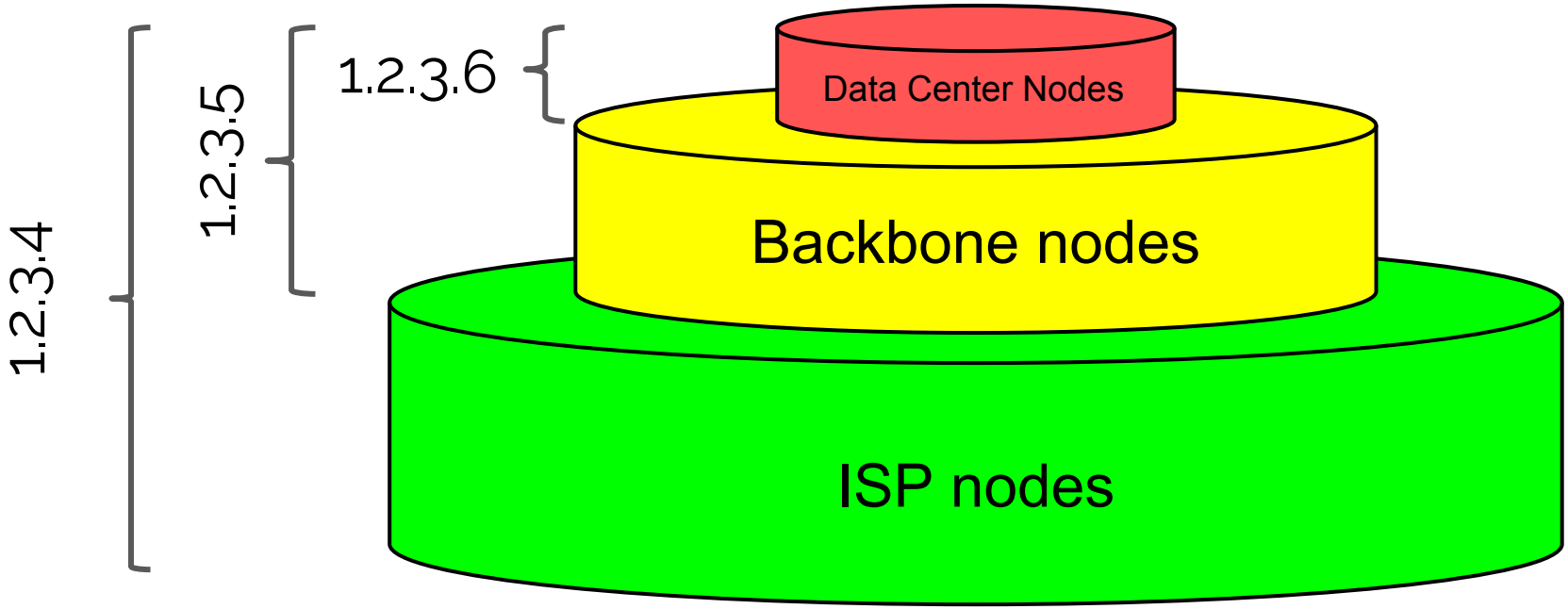
1) deliver a low-latency routing scheme that perform better than existing CDN
2) build an easy-to-operate, scalable and simple system

FastRoute – “Anycast TCP”-like solution,

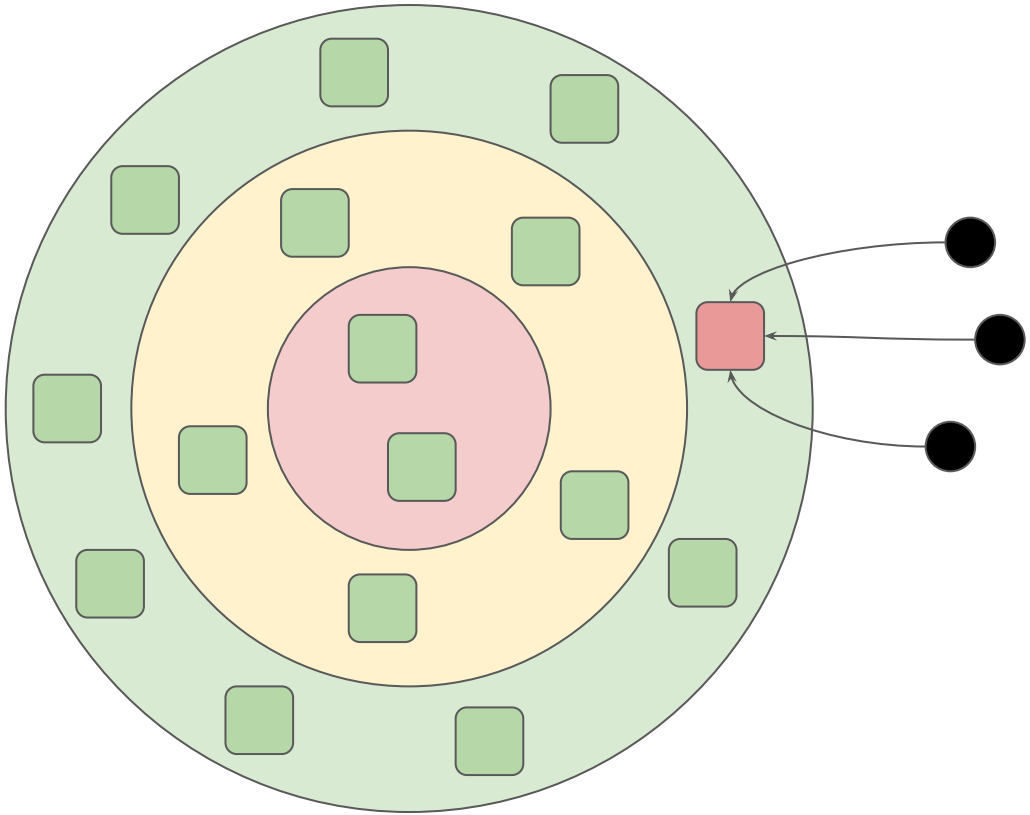
but with enough control over where the user traffic lands

to handle situations when some proxies are overloaded.

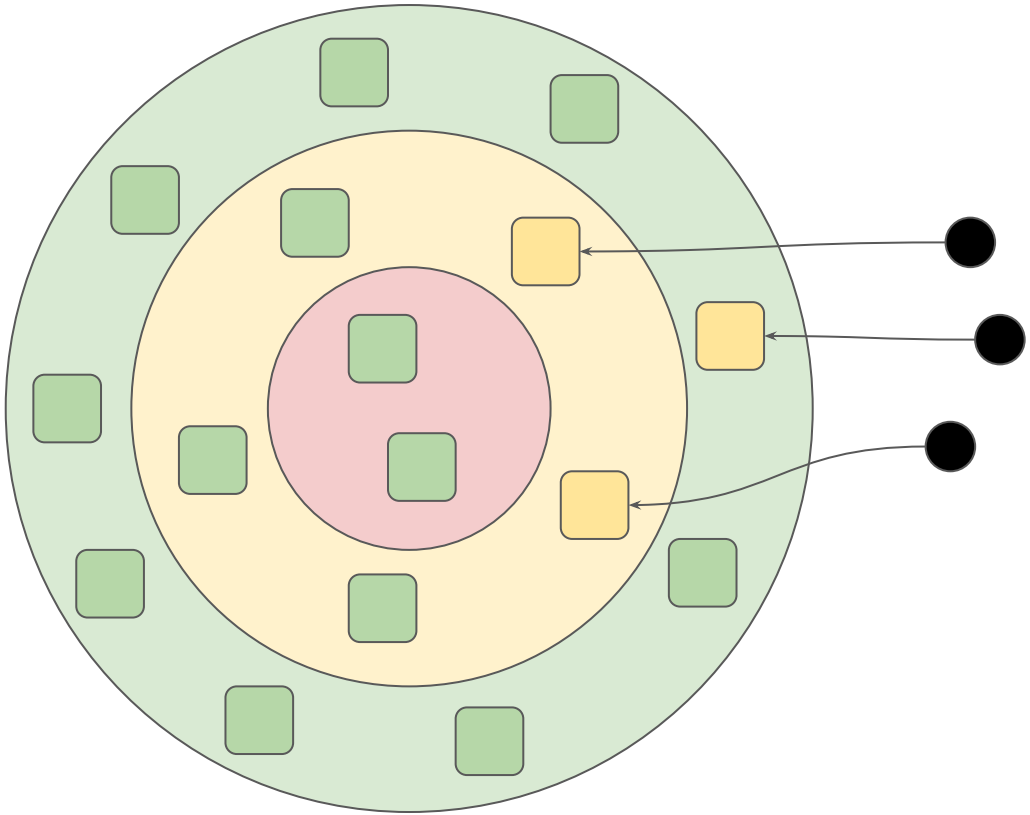
How can we handle where user traffic lands?



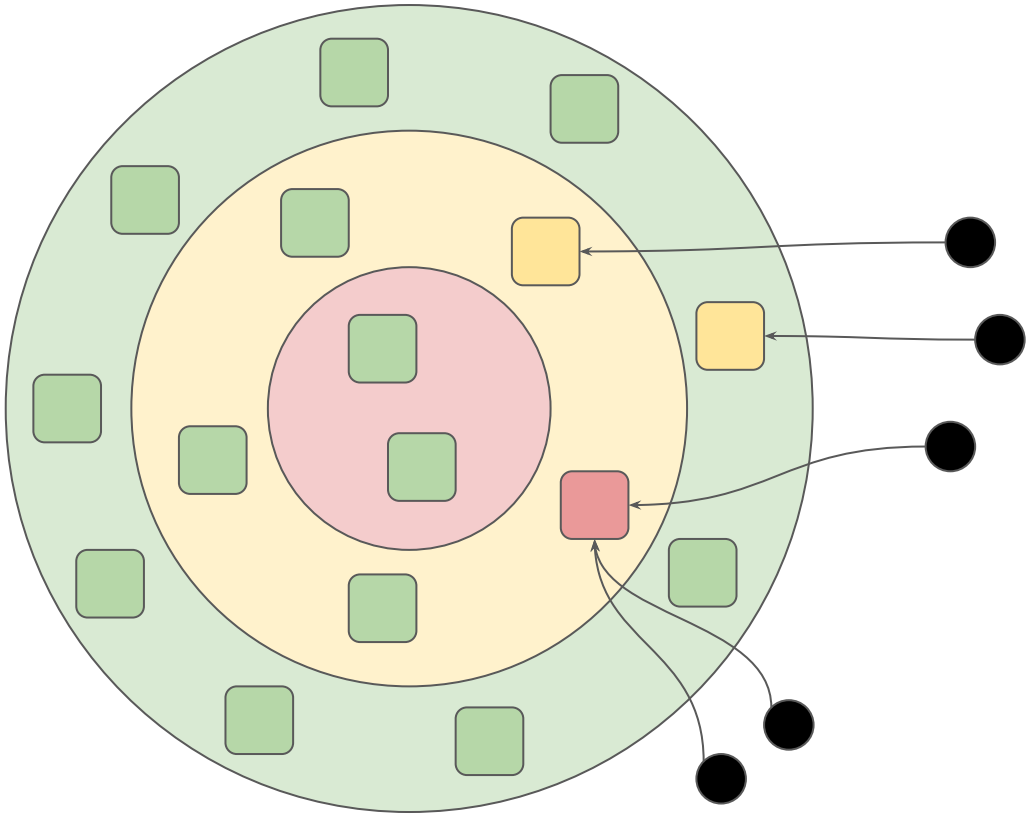
How can we handle where user traffic lands?



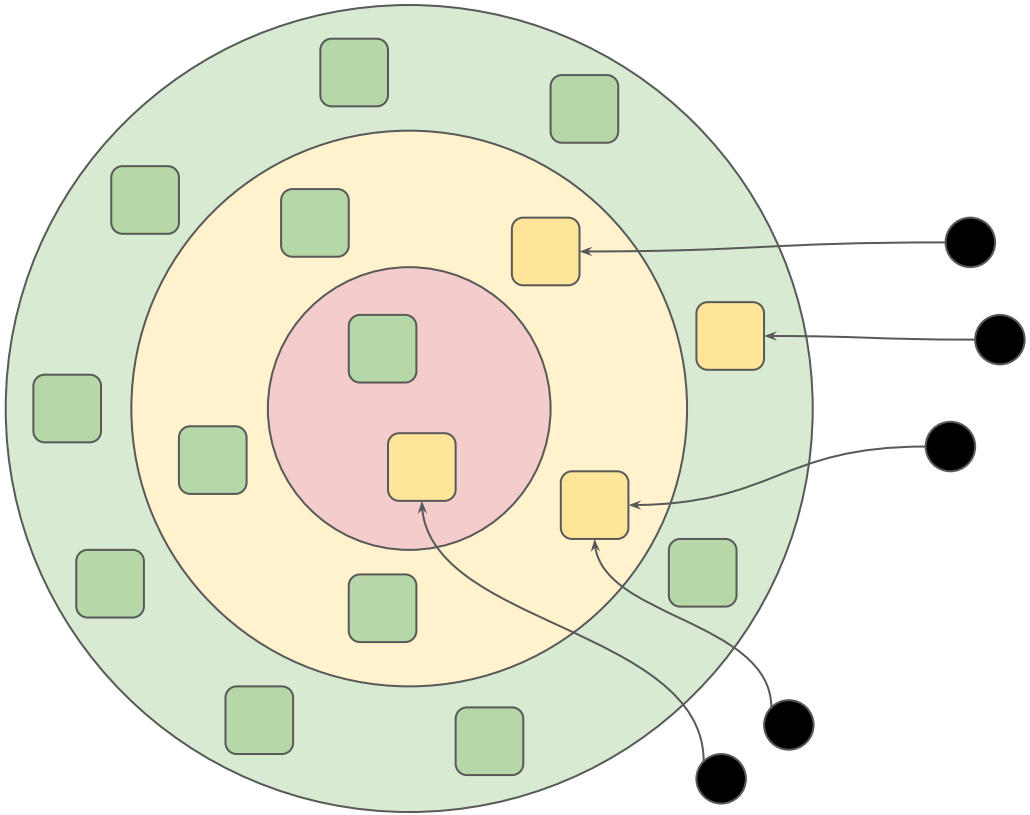
How we can handle where user traffic lands?



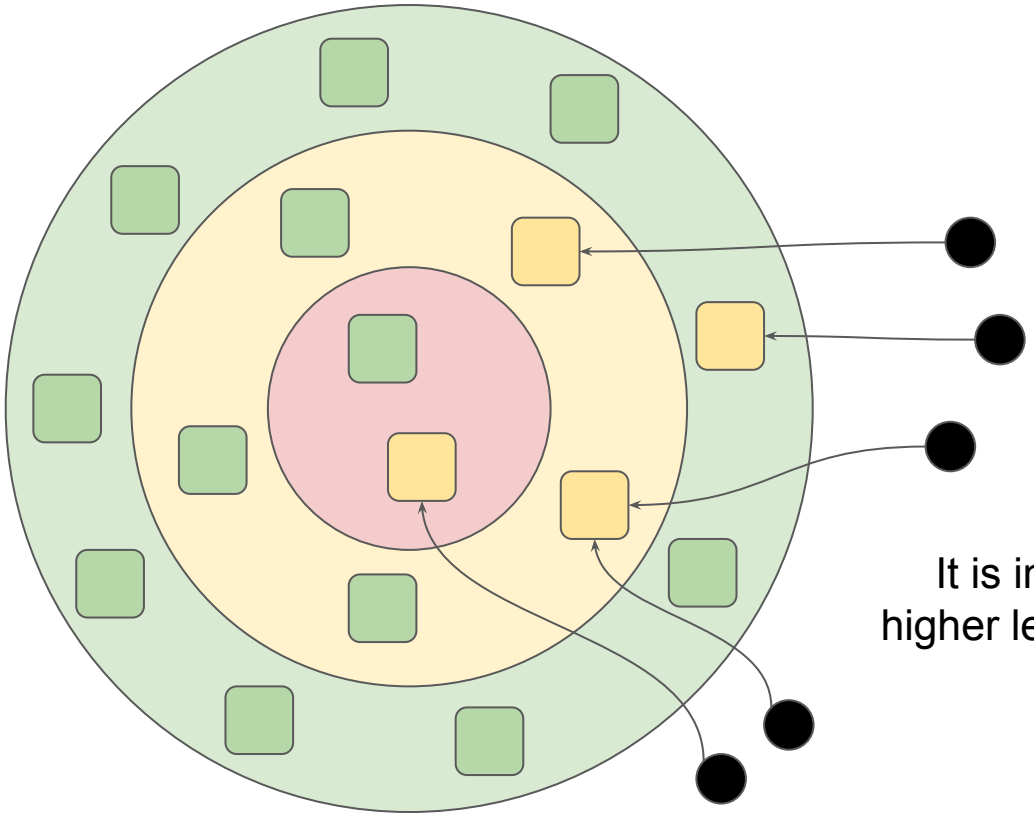
How we can handle where user traffic lands?



How we can handle where user traffic lands?



How we can handle where user traffic lands?



It is important that we redirect traffic to the higher level nodes (not the same level ones). This prevents oscillatory behavior.

FastRoute node

Load balancer

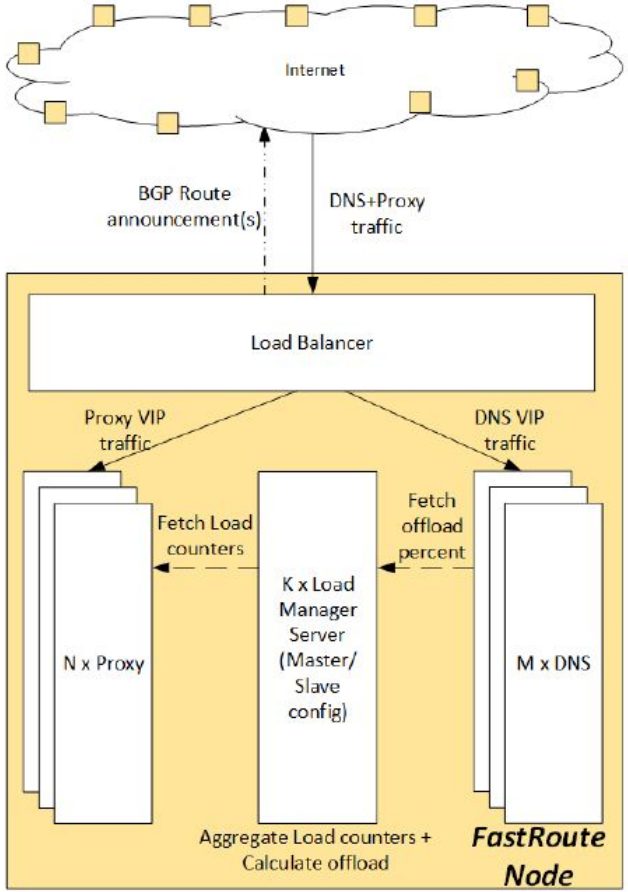
Splits work between N x proxies and M x DNS servers.

Proxy

Handles user traffic.

Load manager

Aggregates information about proxies' load.

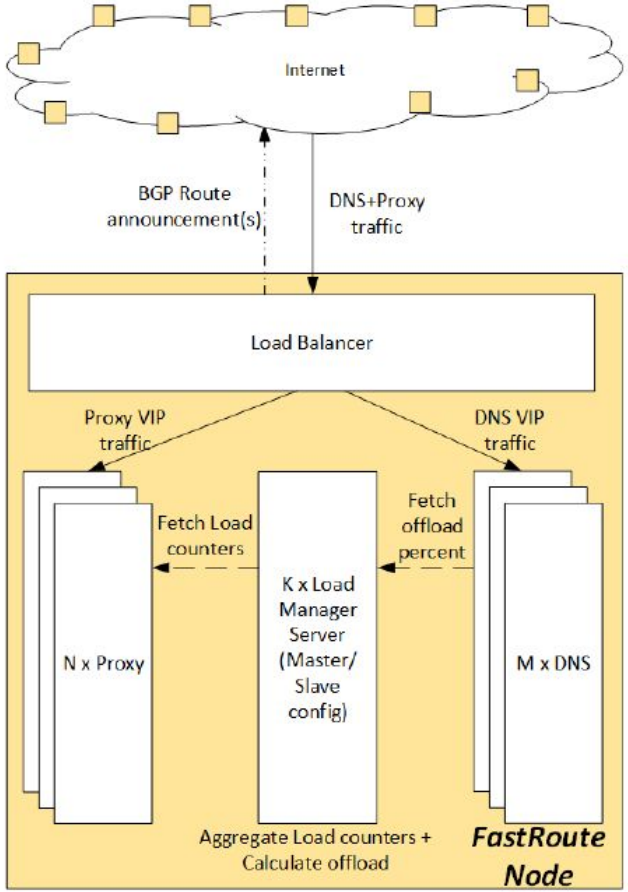


FastRoute node

DNS

Responds for DNS queries with IP address of the layer proxies or redirects to DNS in the higher layer.

Probability of returning redirection (offloading probability) depends on proxies' current load.



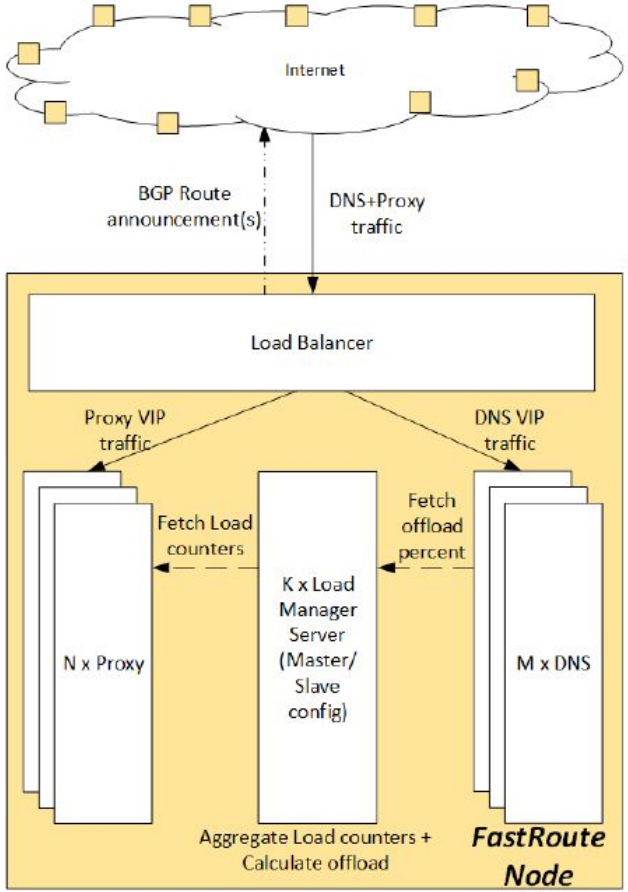
FastRoute node

DNS

Responds for DNS queries with IP address of the layer proxies or redirects to DNS in the higher layer.

Probability of returning redirection depends on proxies' current load.

There's no real-time communication between nodes – each node runs independently.



BGP actions instead of different DNS responses

We could take BGP-level actions to modify routing instead of returning DNS responses.

Why don’t we do it this way?

1) modifying BGP causes TCP sessions breakages (as all connections are affected)
2) it causes less gradual shift of traffic patterns than DNS-based approach

Major assumption

The DNS request for a user lands in the same location as HTTP request (we will call them: *self-correlated* and such traffic is *controllable*)

Major assumption

The DNS request for a user lands in the same location as HTTP request (we will call them: *self-correlated* and such traffic is *controllable*)

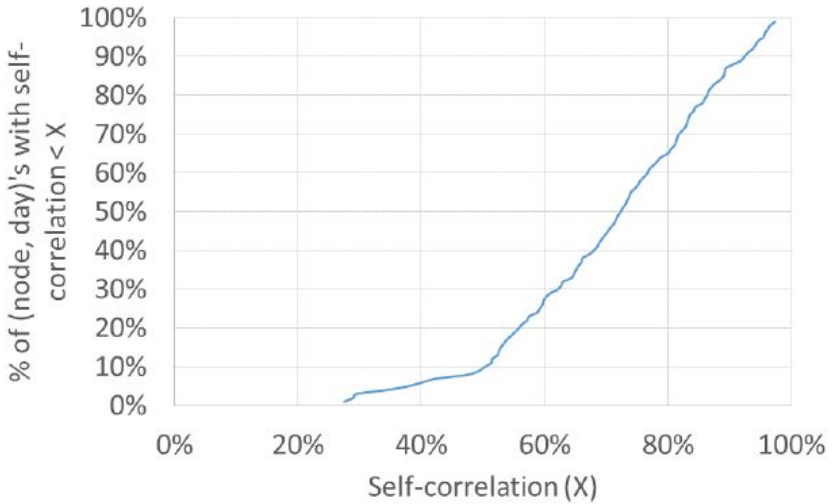
It is not guaranteed for all requests.

Major assumption

The DNS request for a user lands in the same location as HTTP request (we will call them: *self-correlated* and such traffic is *controllable*)

It is not guaranteed for all requests.

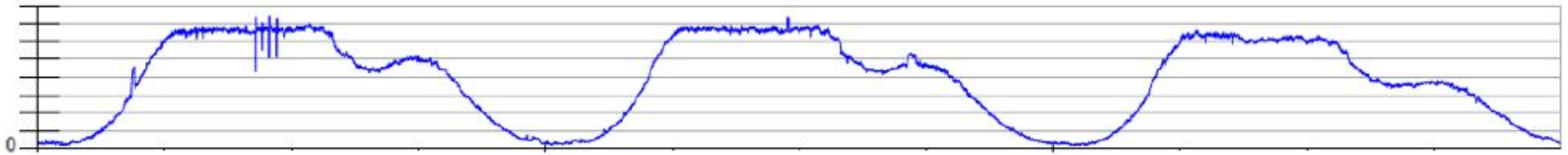
But usually the self-correlation is high enough.



Offload algorithm

We need to divert load in two situations:

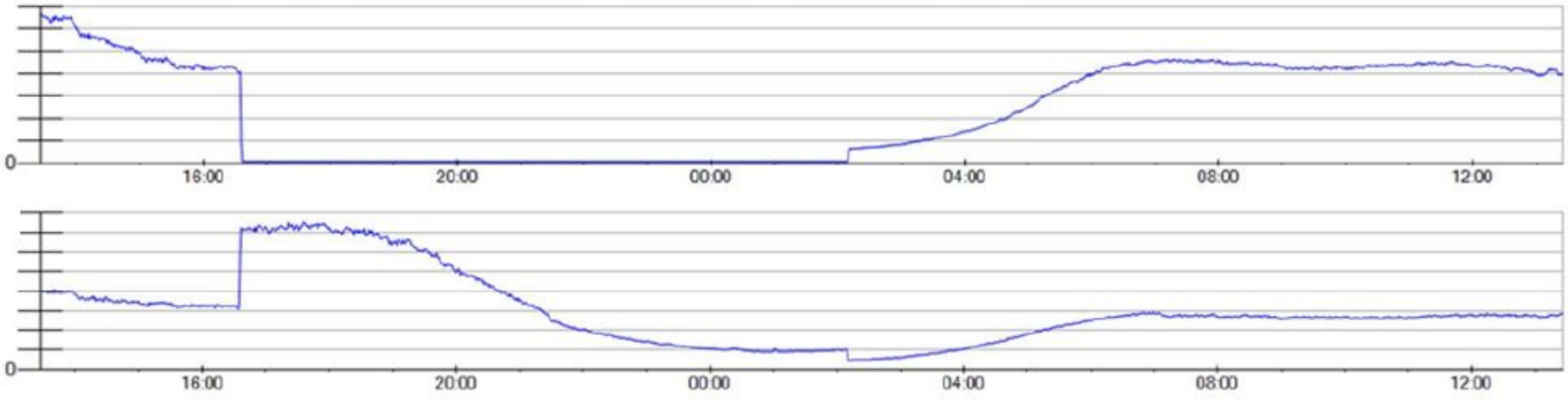
1) The load is increasing/decreasing slowly due to the natural user patterns throughout the day.



Offload algorithm

We need to divert load in two situations:

2) There is step change in load caused by nearby proxy going down. The traffic handled by that node is moved to others.



Offload algorithm

The control over user traffic is limited by:

1) The TTL on a DNS response.
2) Local DNS servers have differing number of users behind them.
3) Not the whole traffic is controllable.

Offload algorithm

The algorithm:

- When the node’s current load is higher than the configured threshold, increase the offload probability super-lineary.
- When the node’s current load is lower than configured threshold, decrease the offload probability lineary.

Less important domains are offloaded before more important ones.

When node is overloaded after offloading everything, manual intervention is needed to force highly cross-correlated nodes to start offloading.

Summary

1) FastRoute use Internet routing protocols to choose nearest proxy.
2) FastRoute nodes are grouped into layers.
3) Nodes are running independently, and decides locally whether to “forward” traffic to higher layer.
4) Each node has usually control over the majority of incoming traffic.

1. CDN
2. Proxy selection techniques
3. FastRoute architecture
4. In production
5. Future work

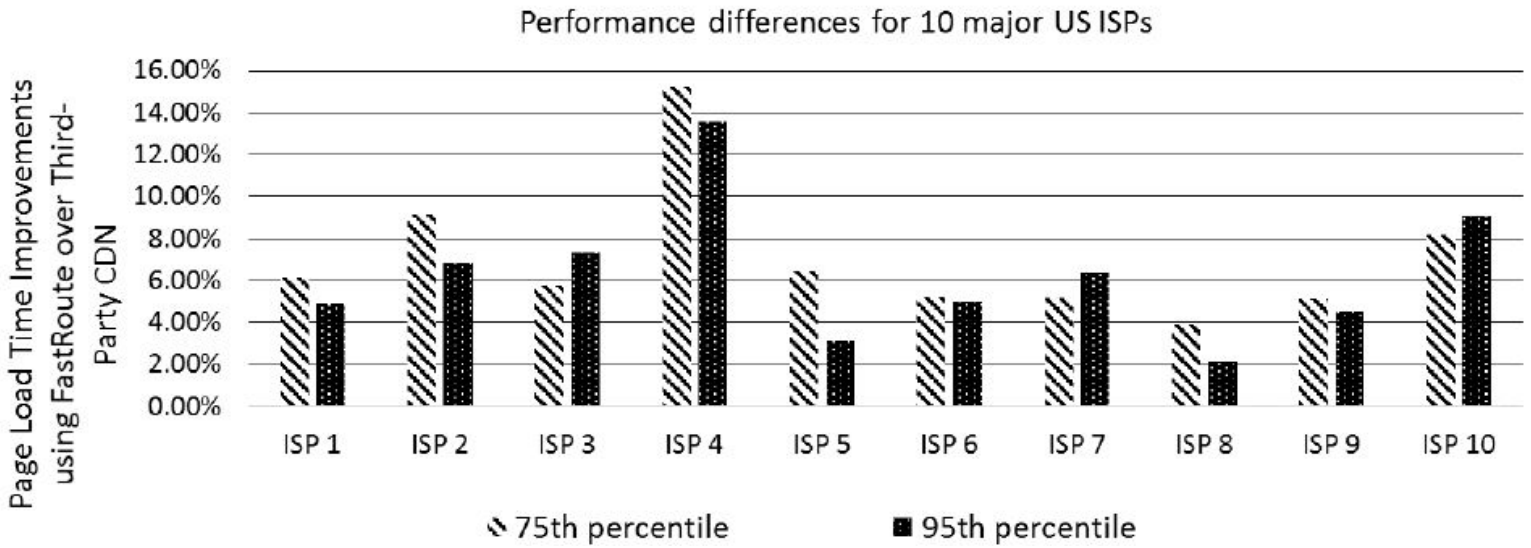
Non-Production test

Microsoft placed small images on their sites. The images were downloaded via FastRoute and third-party CDN by 5% of users.

FastRoute delivered image frequently faster than currently used CDN.

Production test

Moved small random group of US users to FastRoute.
The test was repeated on different user groups, for different time durations.



Load management

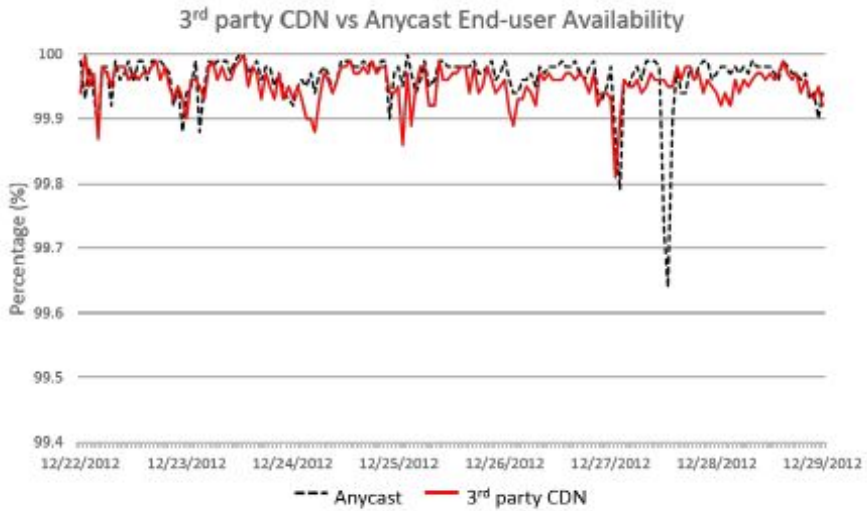
As in 2015, FastRoute was running for 2 years.

There are few overload situations per week. They were caused by proxies going down, spiky user traffic patterns, bugs in the FastRoute code.

None of the situations required manual intervention.

Availability

Availability was measured by placing code downloading small image in Bing toolbar via FastRoute (12 nodes running) and third party CDN.



The availability during the whole week was: 99.96% (3rd party CDN) vs 99.95% (FastRoute)

1. CDN
2. Proxy selection techniques
3. FastRoute architecture
4. In production
5. Future work

Future work

1) Prioritizing and multiplexing user traffic.
2) Analyzing the impact of self-correlation of DNS traffic and proxy traffic when supporting IPv6
3) Understanding of the degree of sub-optimality introduced due to making local decisions, compared to making globally optimal decisions centrally.
4) Studying the distributed load management algorithm from control-theoretic perspective.

Questions