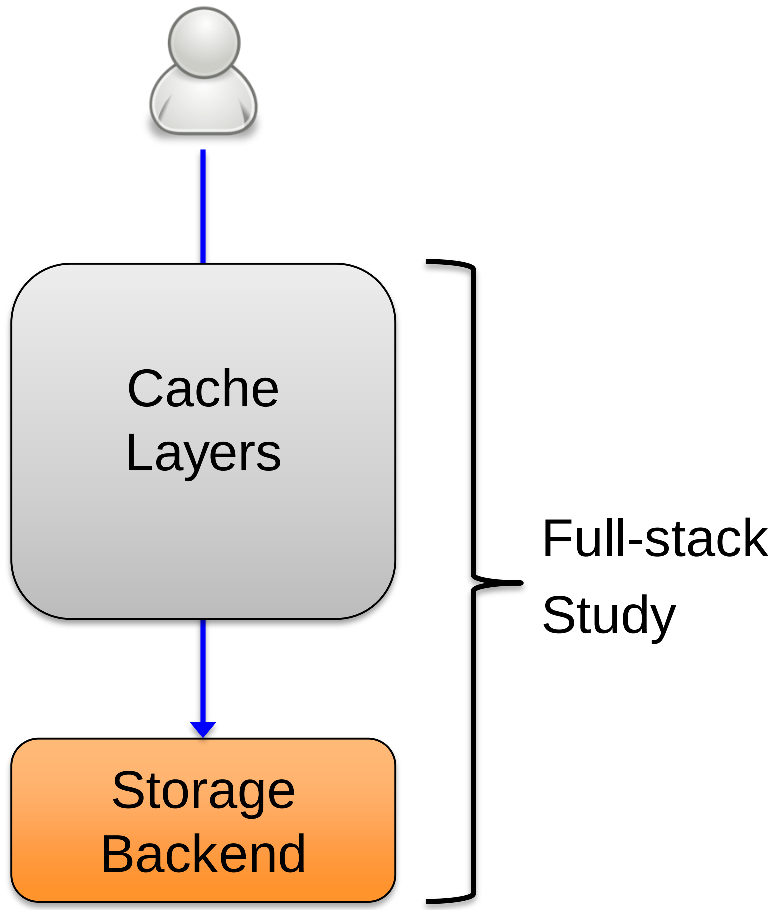
An Analysis of Facebook Photo Caching

Qi Huang, Ken Birman, Robbert van Renesse (Cornell),
Wyatt Lloyd (Princeton, Facebook),
Sanjeev Kumar, Harry C. Li (Facebook)

250 Billion* Photos on Facebook



* Internet.org, Sept., 2013

Preview of Results

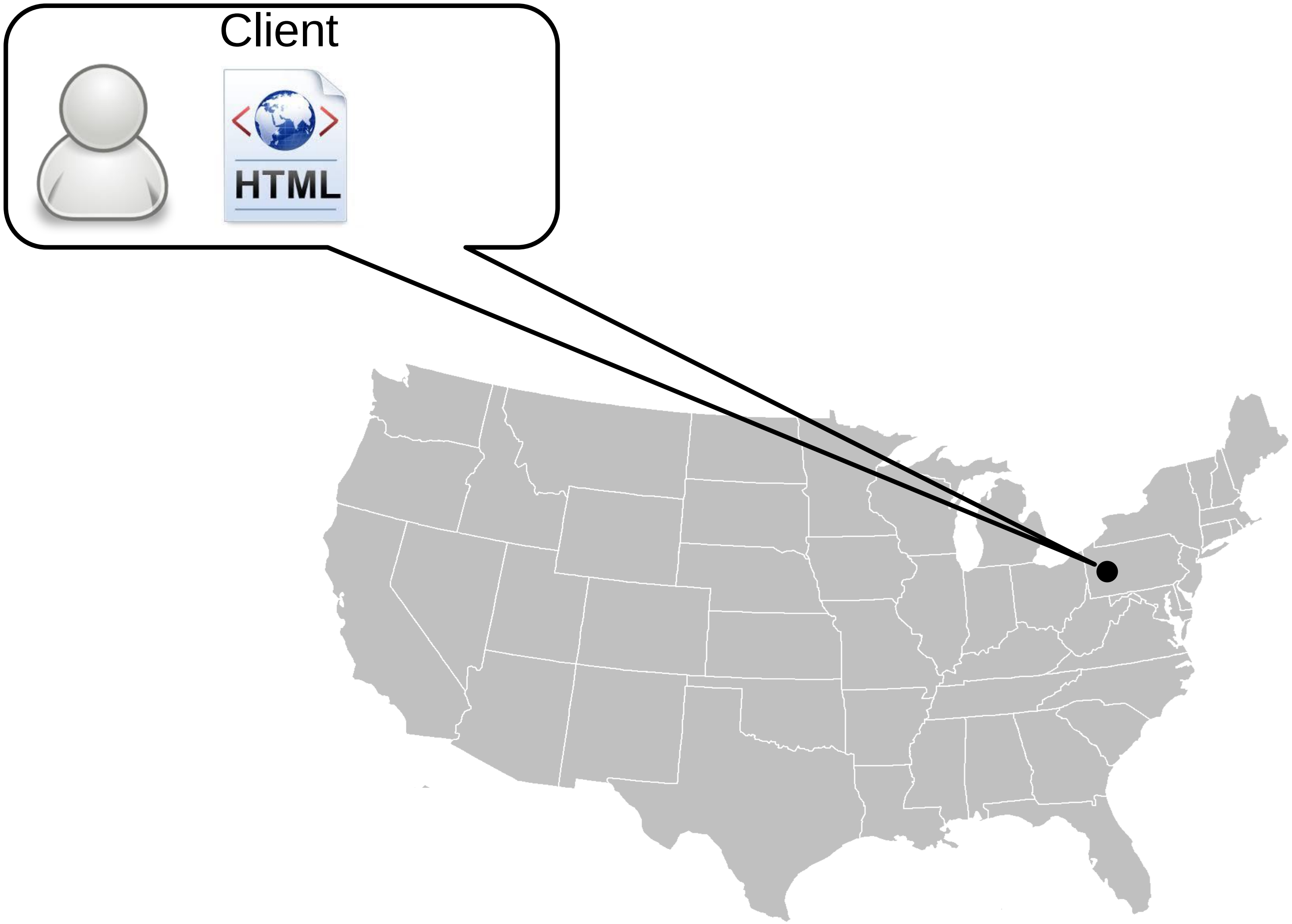
Current Stack Performance

- Browser cache is important (reduces 65+% request)
- Photo popularity distribution shifts across layers

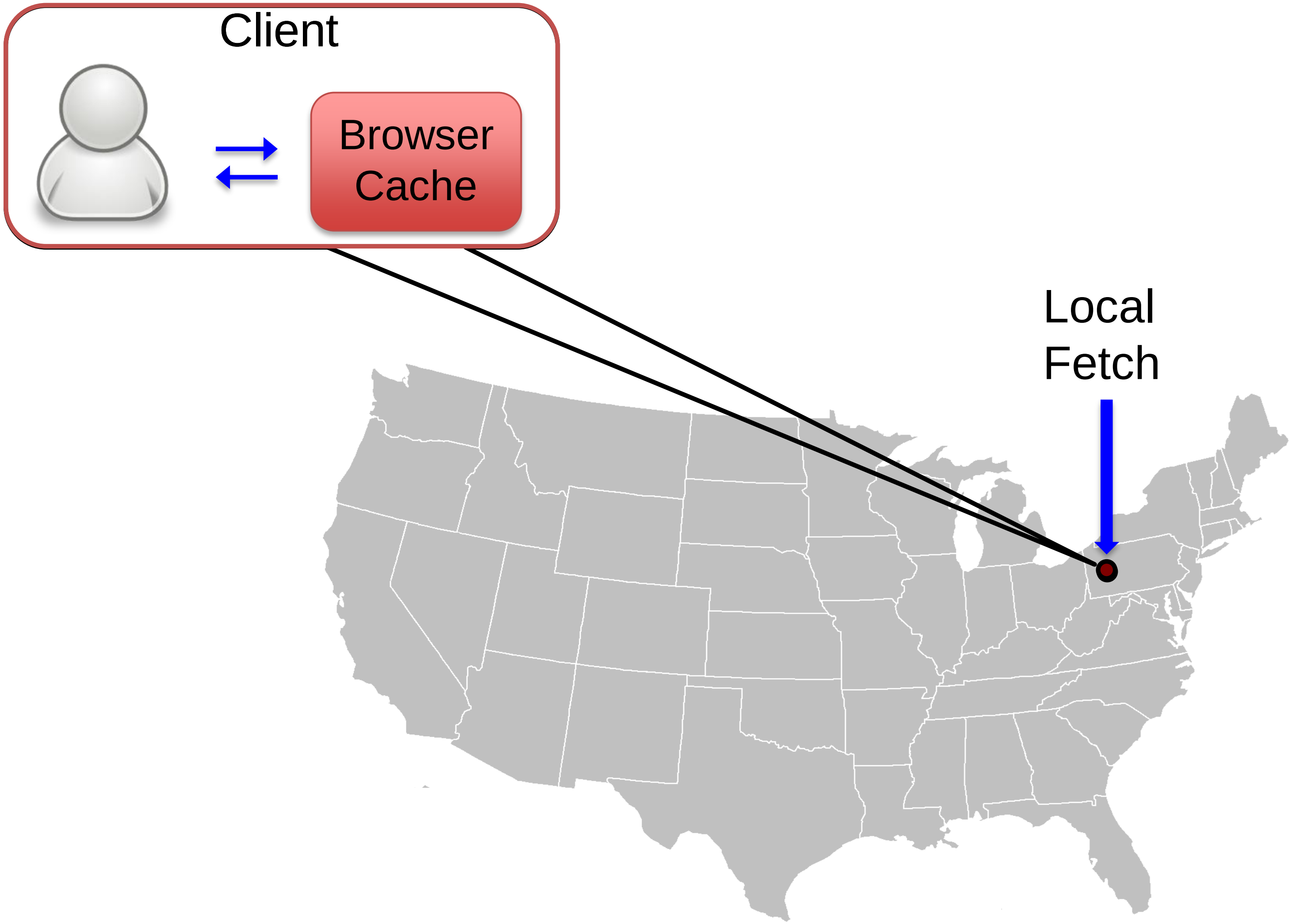
Opportunities for Improvement

- Smarter algorithms can do much better (S4LRU)
- Collaborative geo-distributed cache worth trying

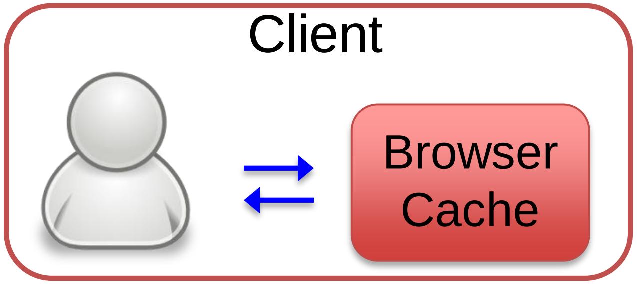
Facebook Photo-Serving Stack



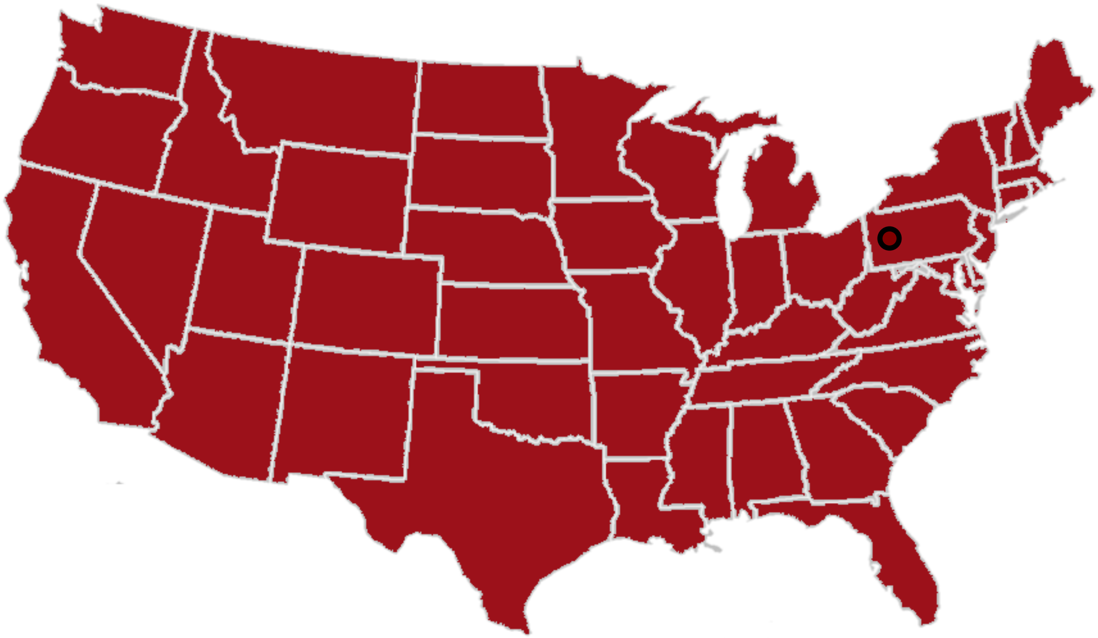
Client-based **Browser Cache**



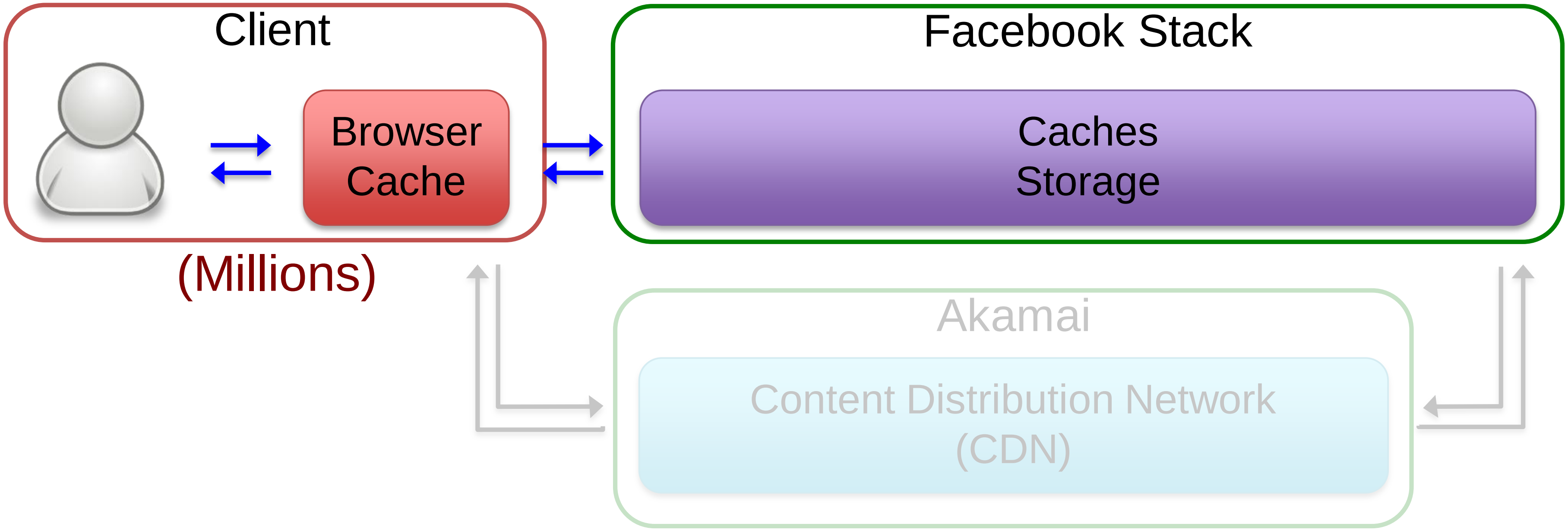
Client-based Browser Cache



(Millions)

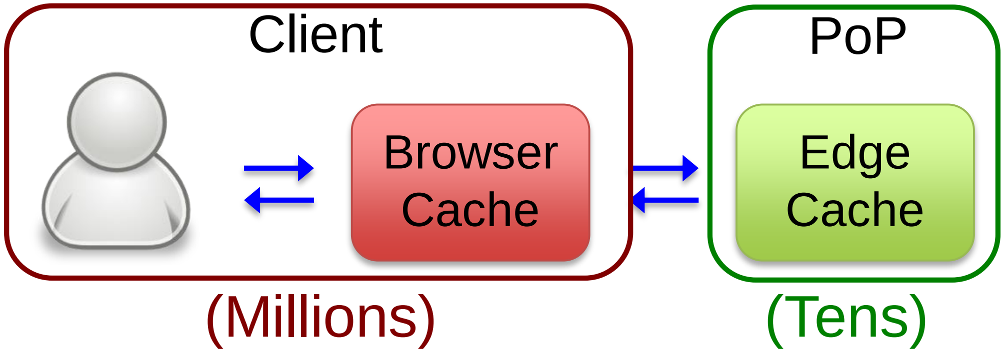


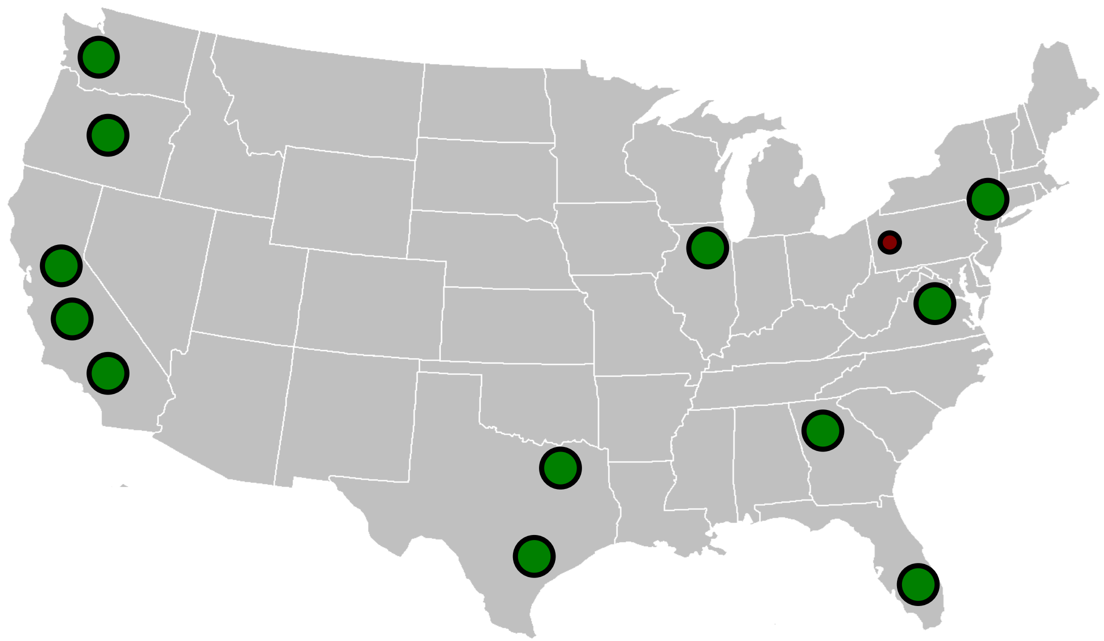
Stack Choice



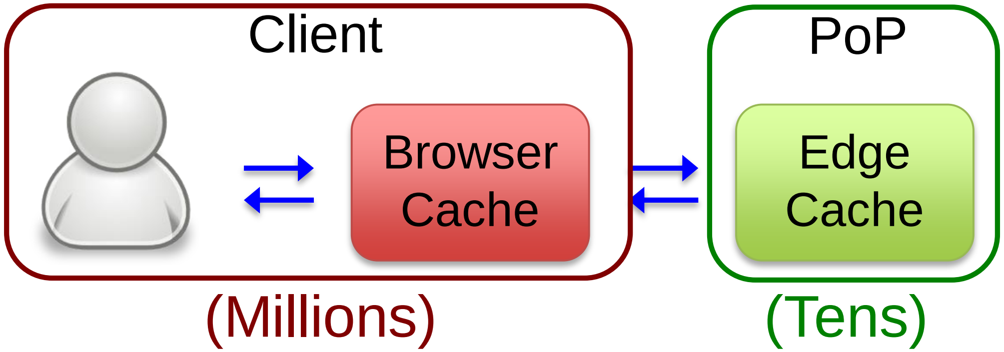
- **Focus: Facebook stack**

Geo-distributed Edge Cache (FIFO)





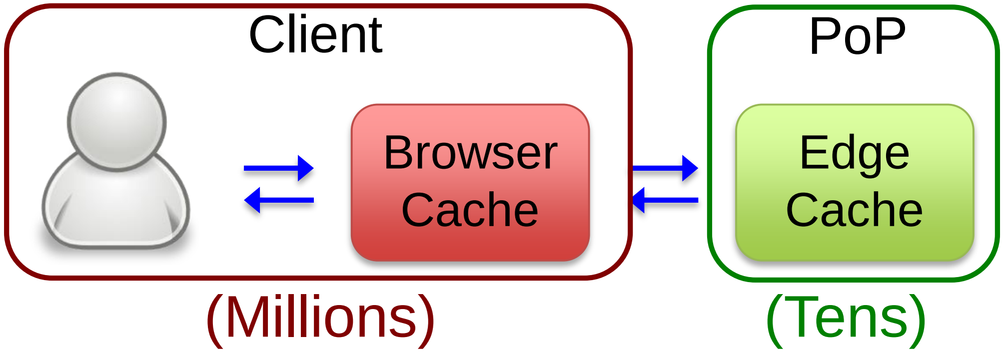
Geo-distributed Edge Cache (FIFO)

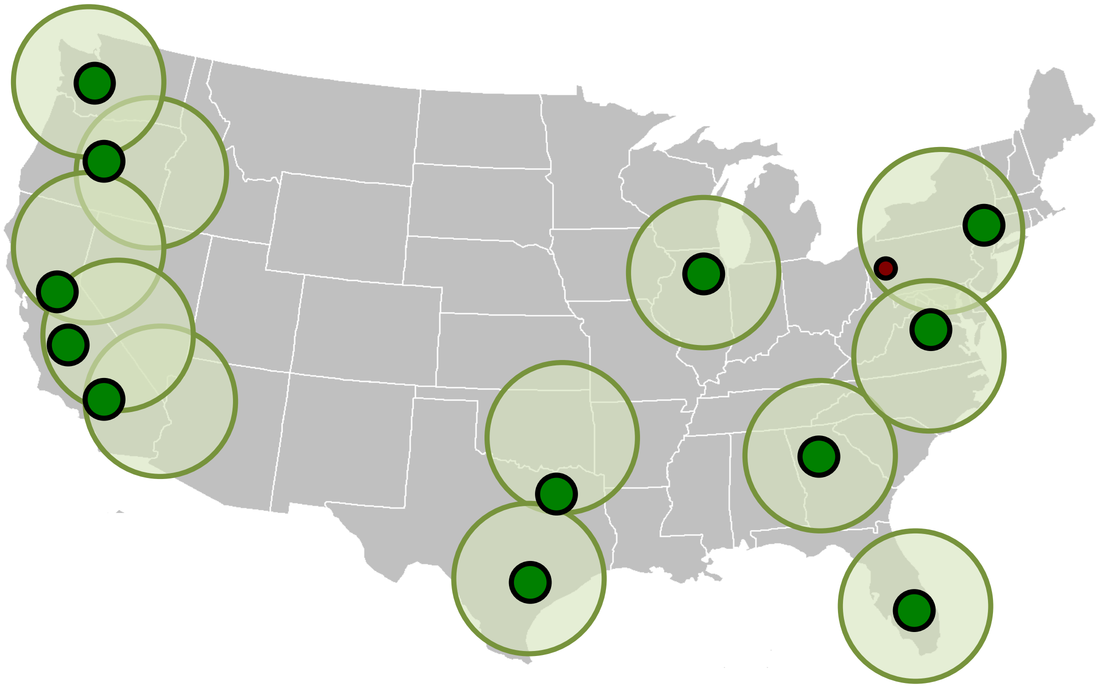


Purpose

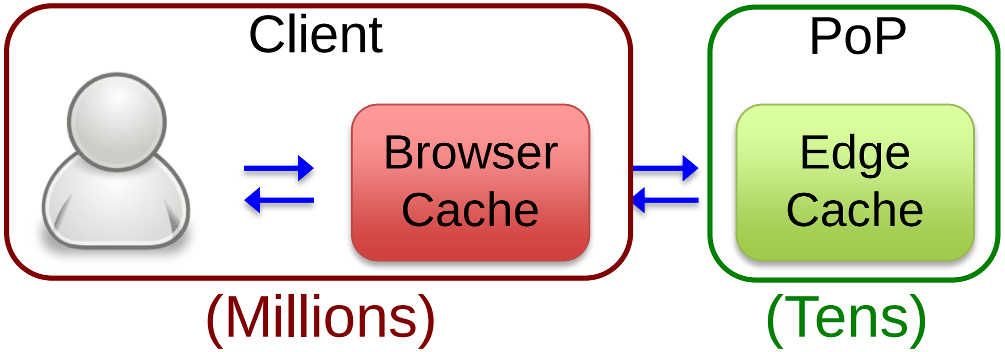
1. **Reduce cross-country latency**
2. **Reduce Data Center bandwidth**

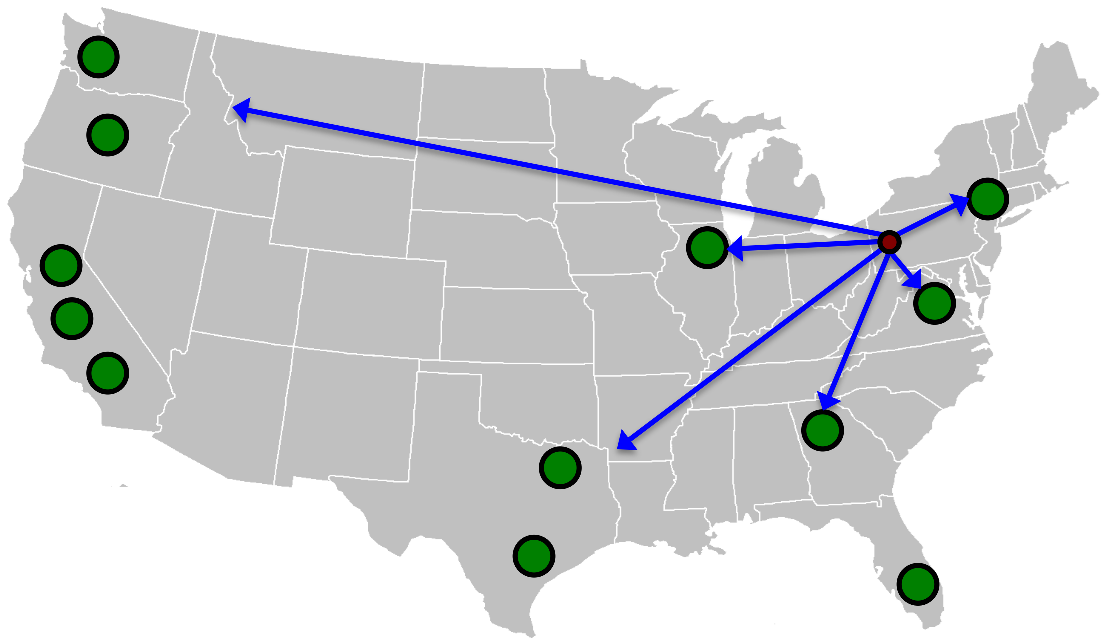
Geo-distributed Edge Cache (FIFO)



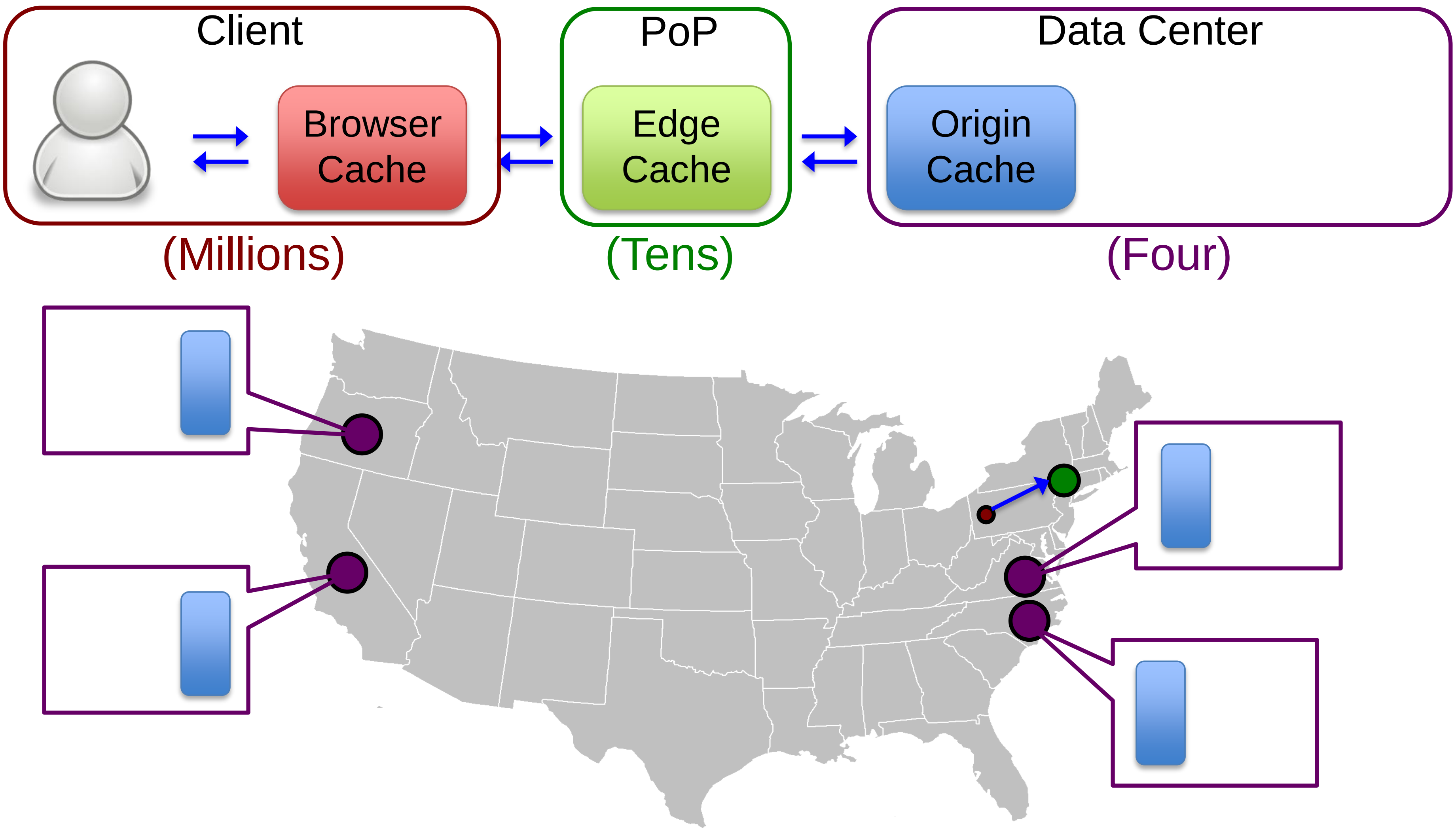


Geo-distributed Edge Cache (FIFO)

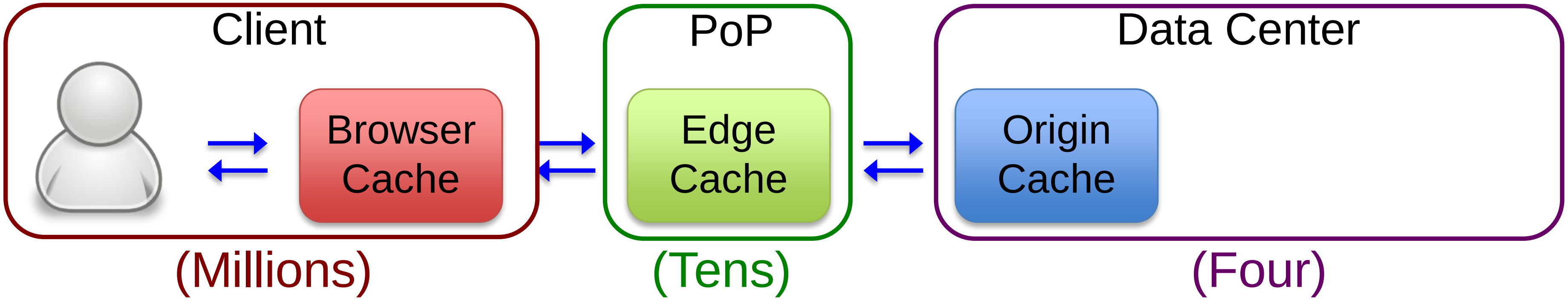




Single Global Origin Cache (FIFO)



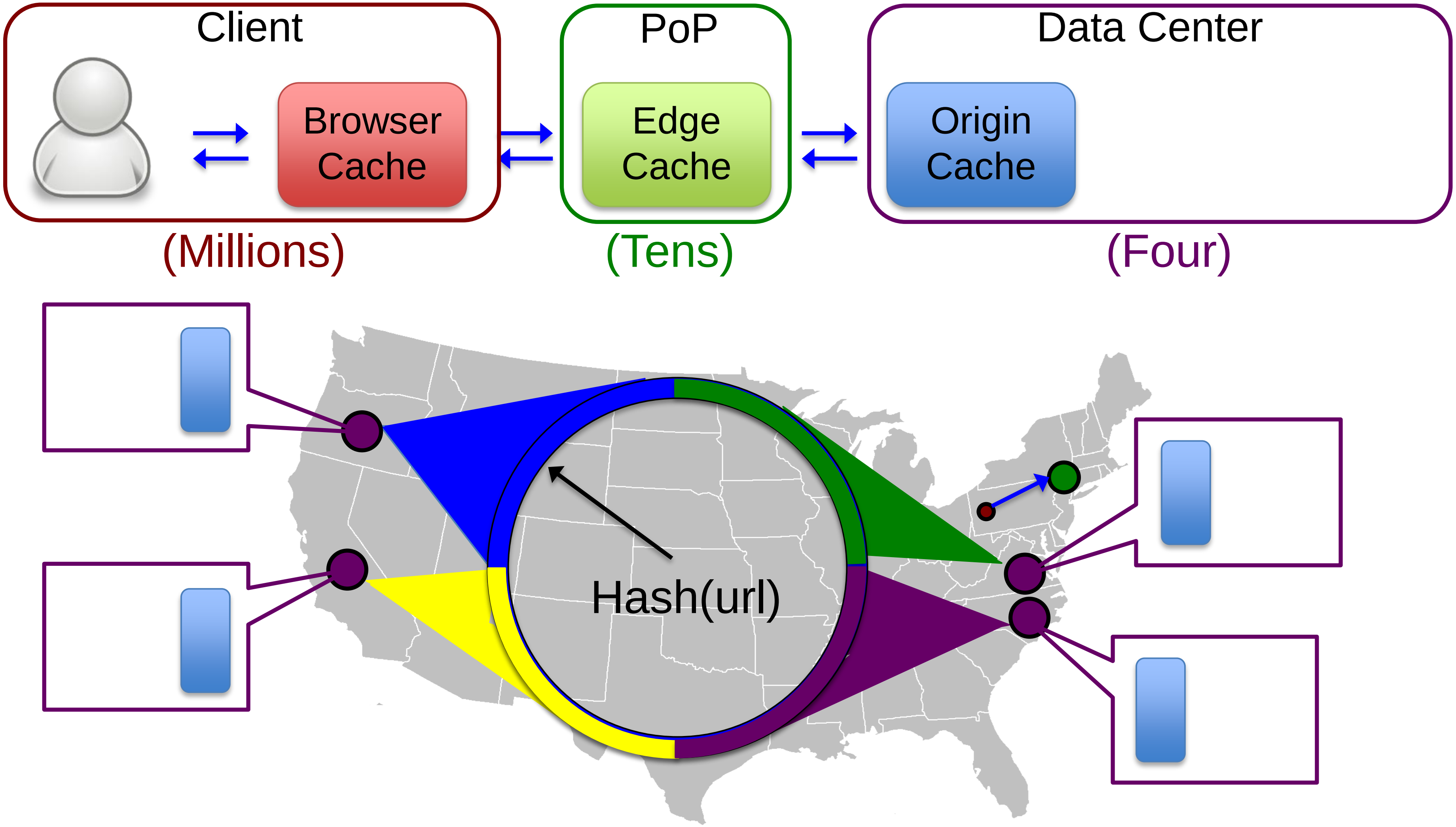
Single Global Origin Cache (FIFO)



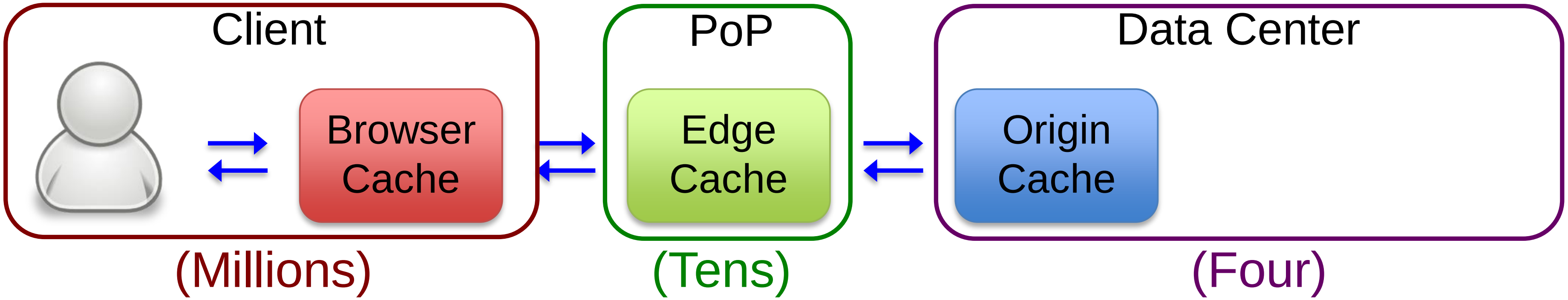
Purpose

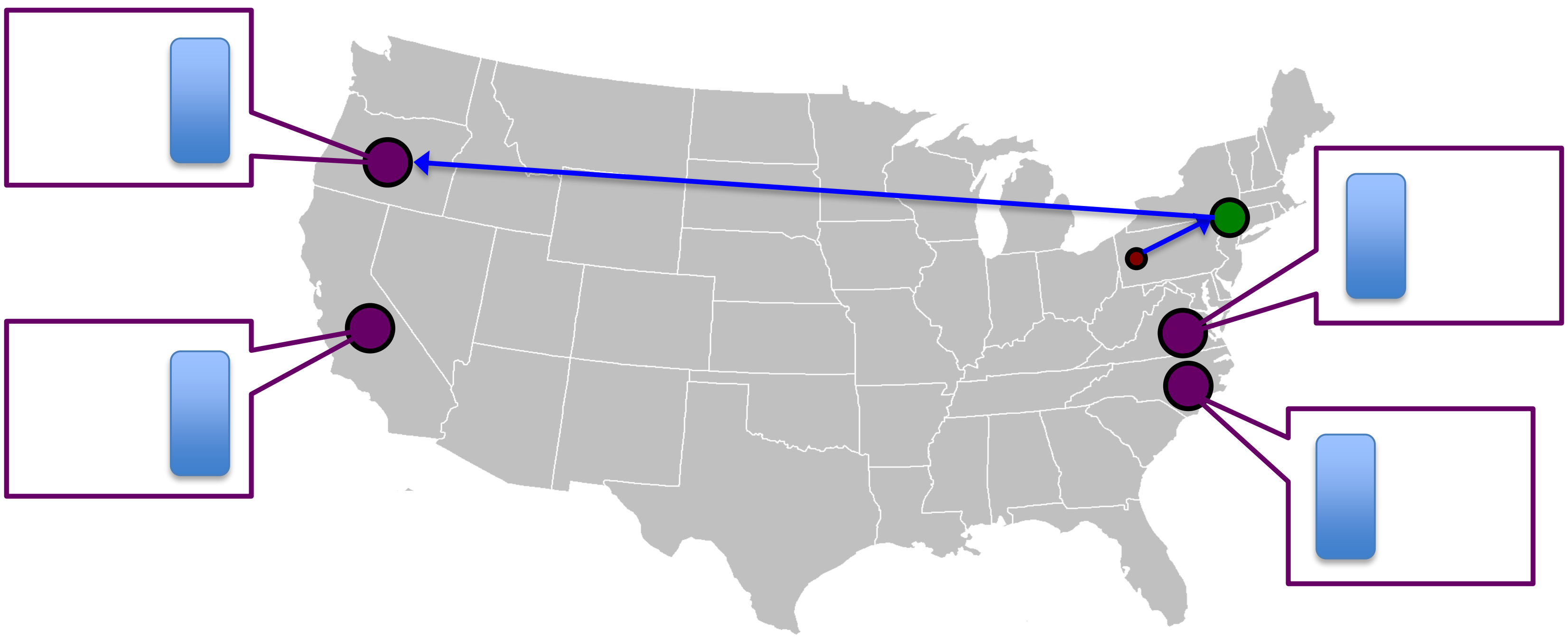
1. **Minimize I/O-bound operations**

Single Global Origin Cache (FIFO)

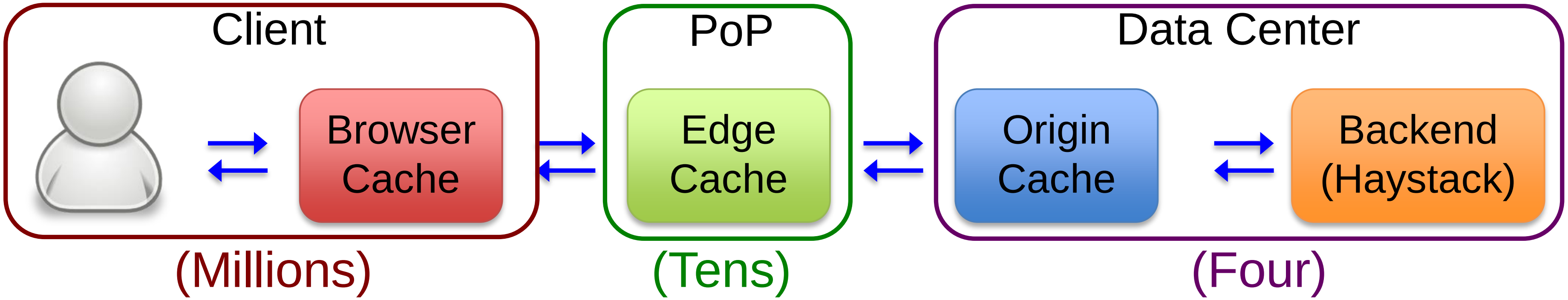


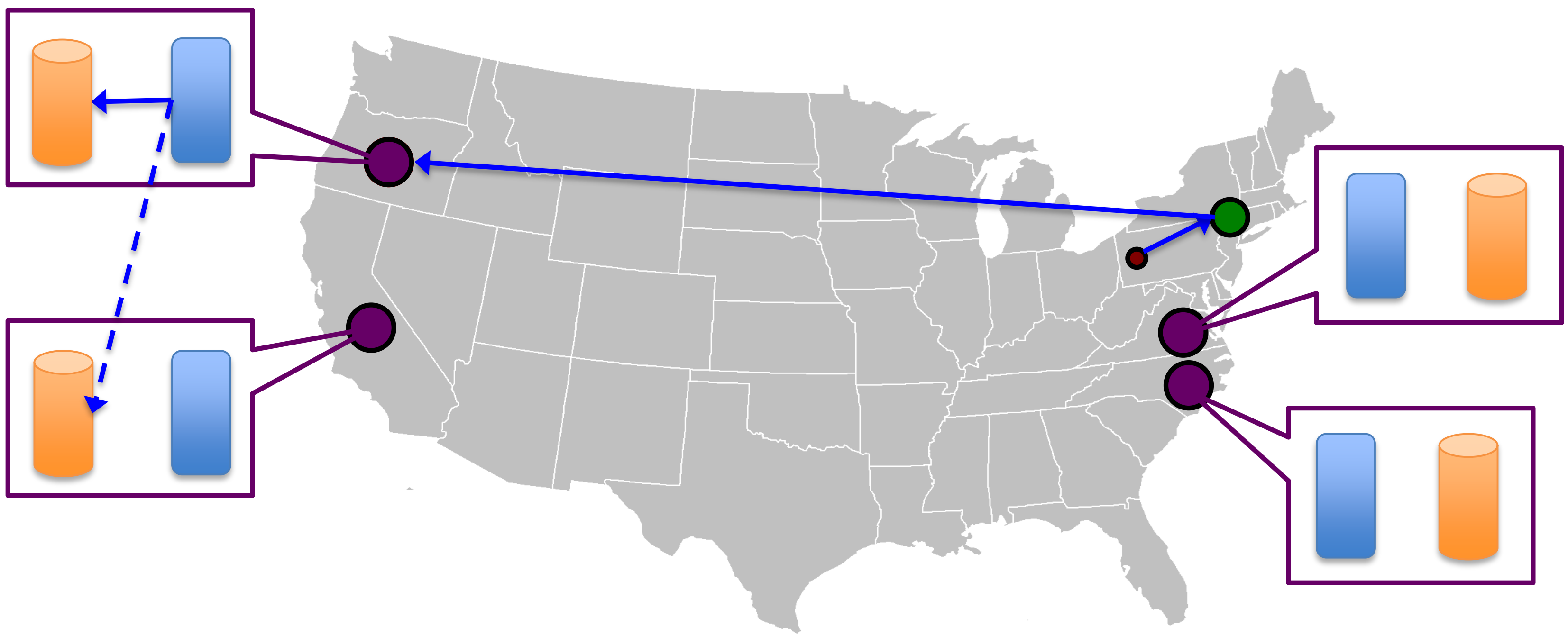
Single Global Origin Cache (FIFO)





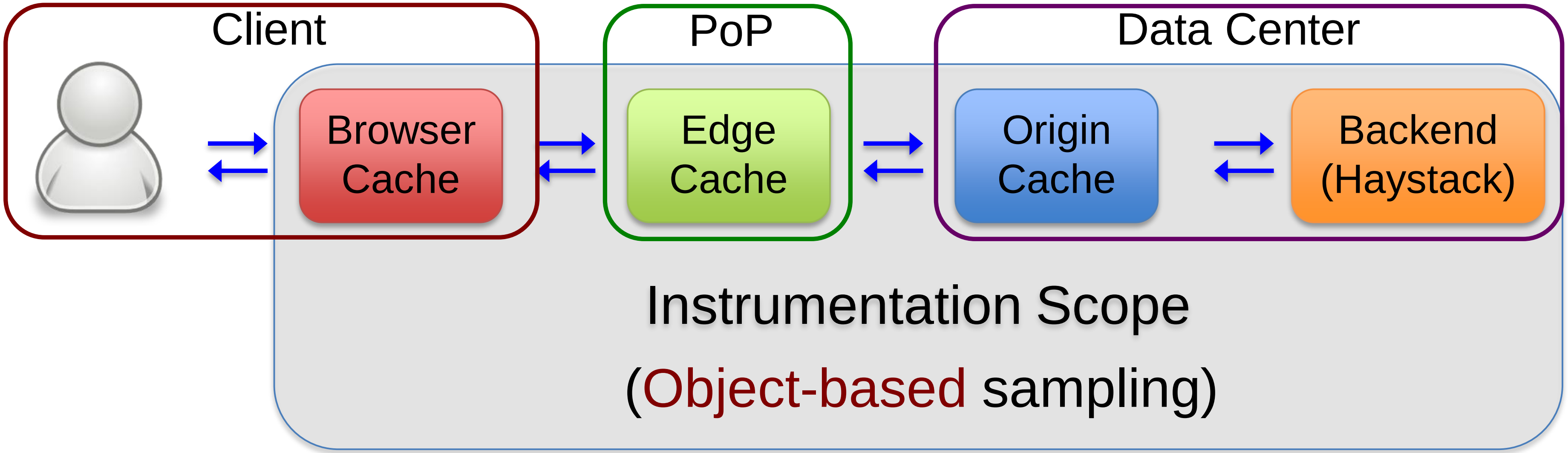
Haystack Backend





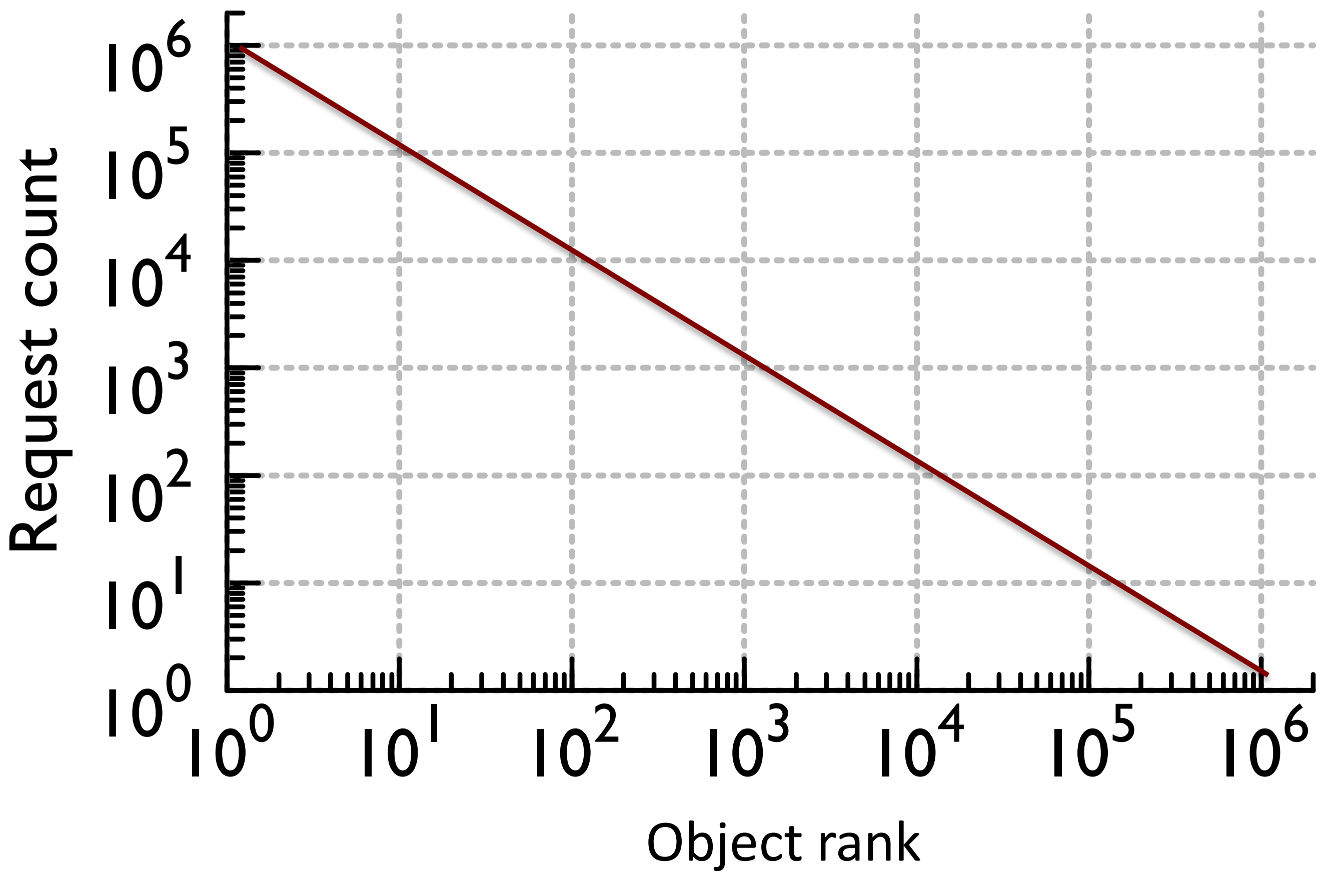
How did we collect the trace?

Trace Collection

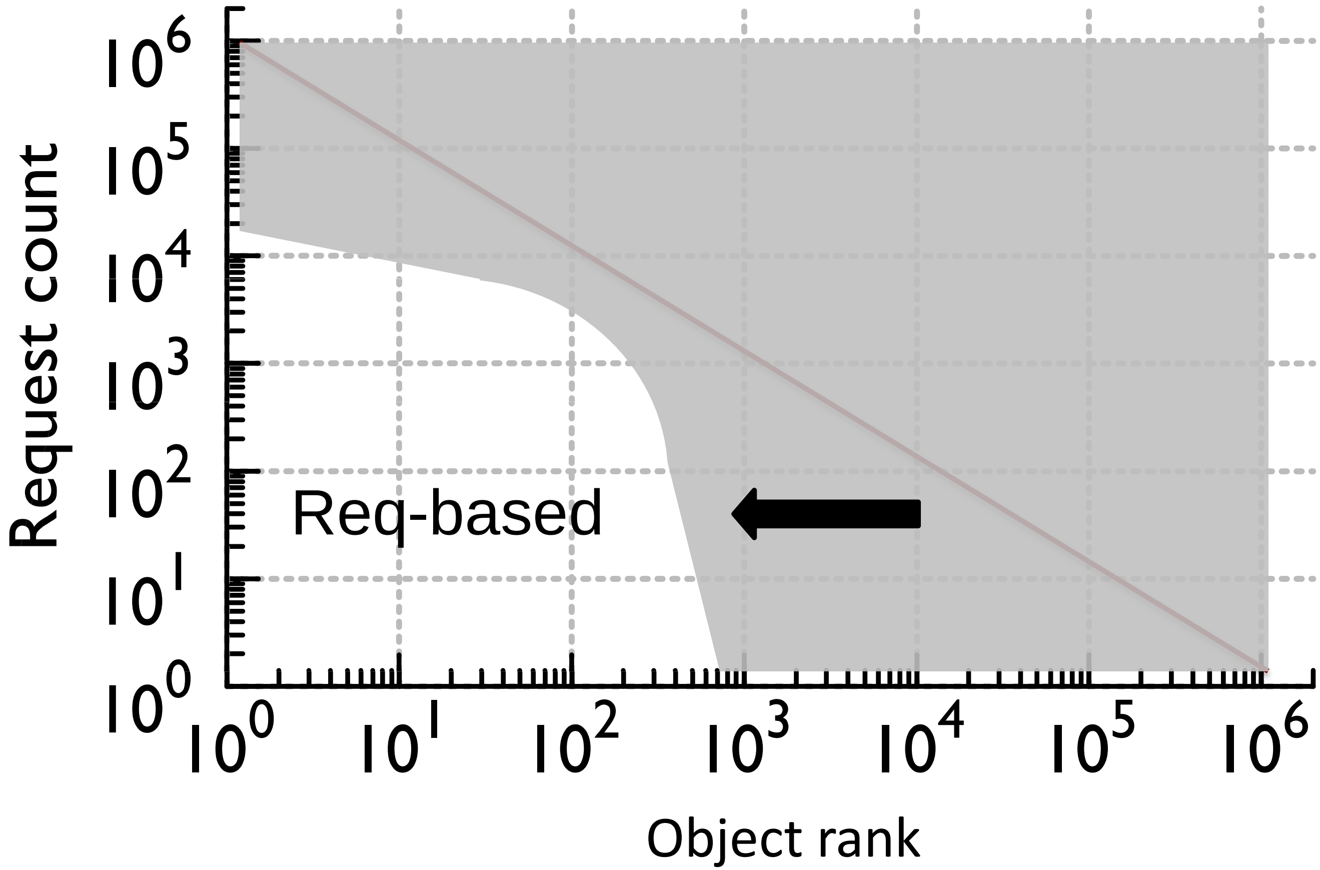


- Request-based: collect X% of requests
- Object-based: collect reqs for X% objects

Sampling on Power-law

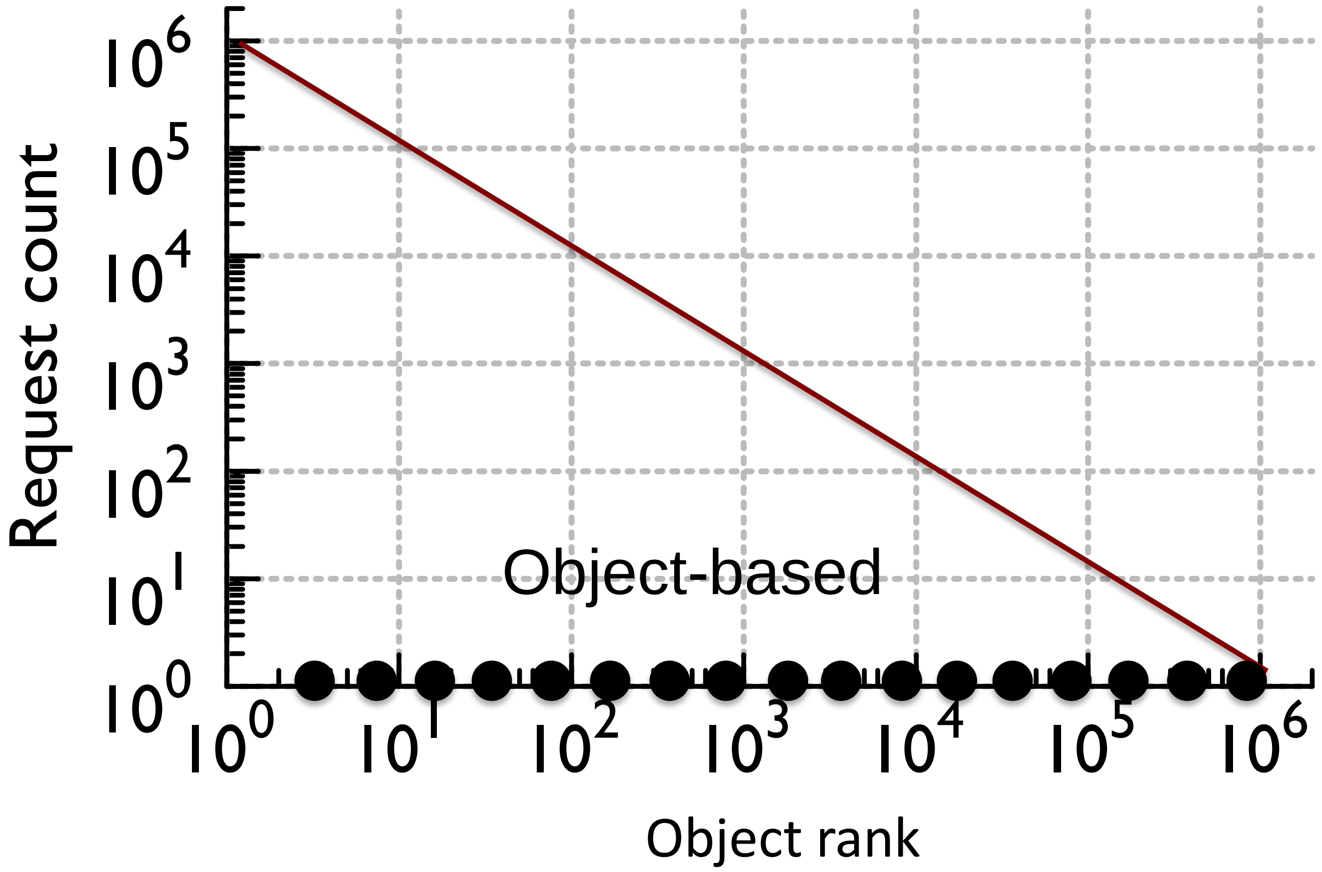


Sampling on Power-law



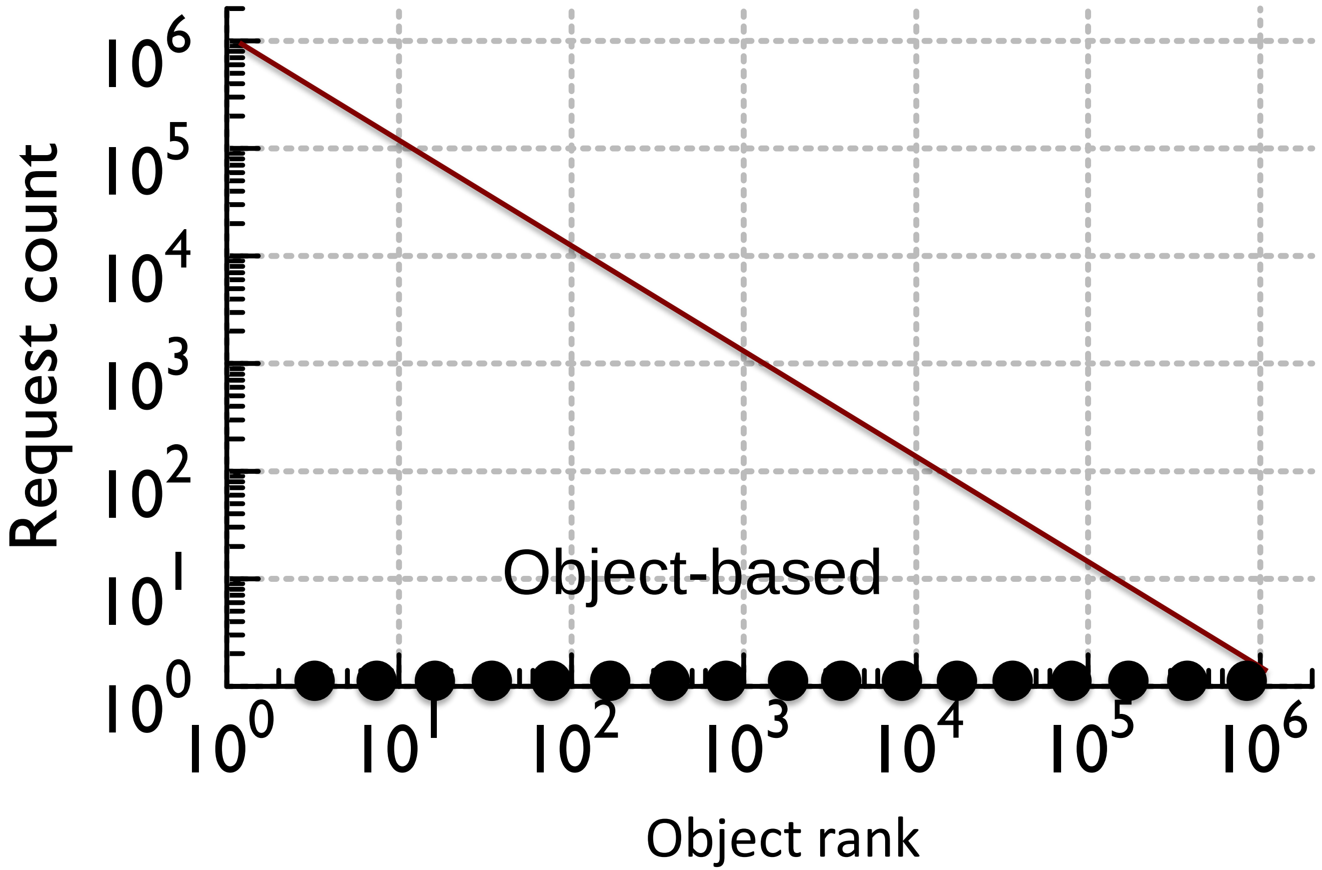
- Req-based: bias on popular content, inflate cache perf

Sampling on Power-law



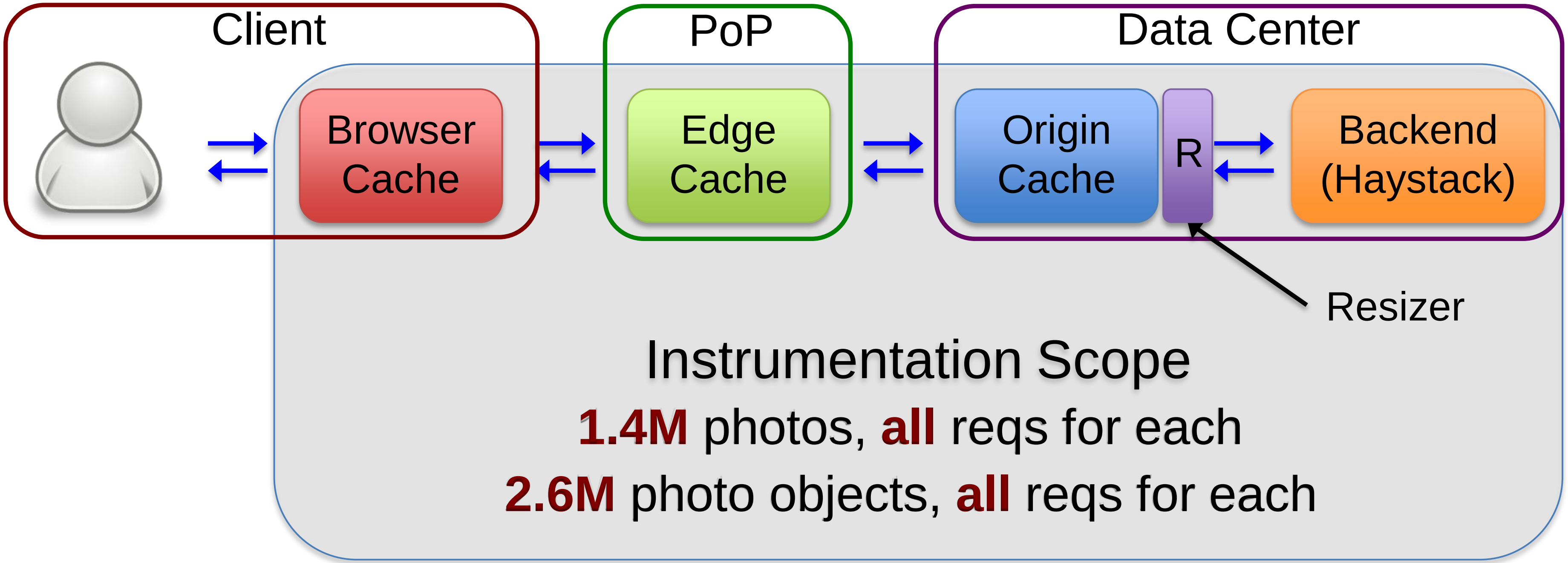
- Object-based: fair coverage of unpopular content

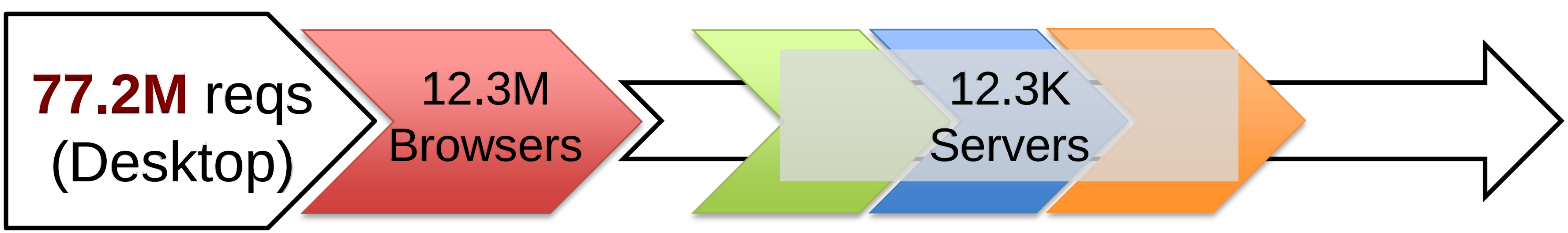
Sampling on Power-law



- Object-based: fair coverage of unpopular content

Trace Collection





Analysis

- Traffic sheltering effects of caches
- Photo popularity distribution
- Size, algorithm, collaborative Edge
- In paper
  - Stack performance as a function of photo age
  - Stack performance as a function of social connectivity
  - Geographical traffic flow

Traffic Sheltering

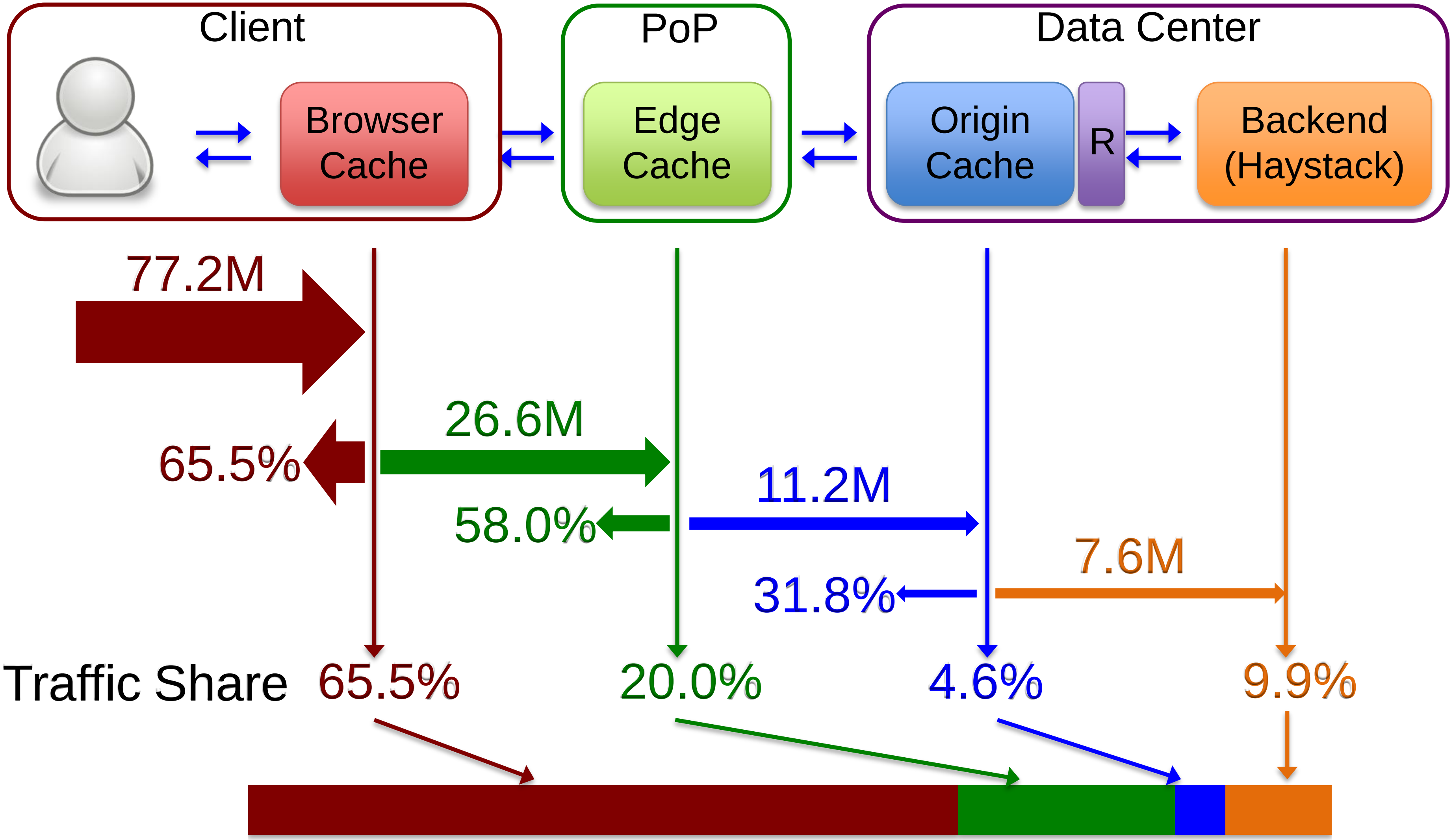
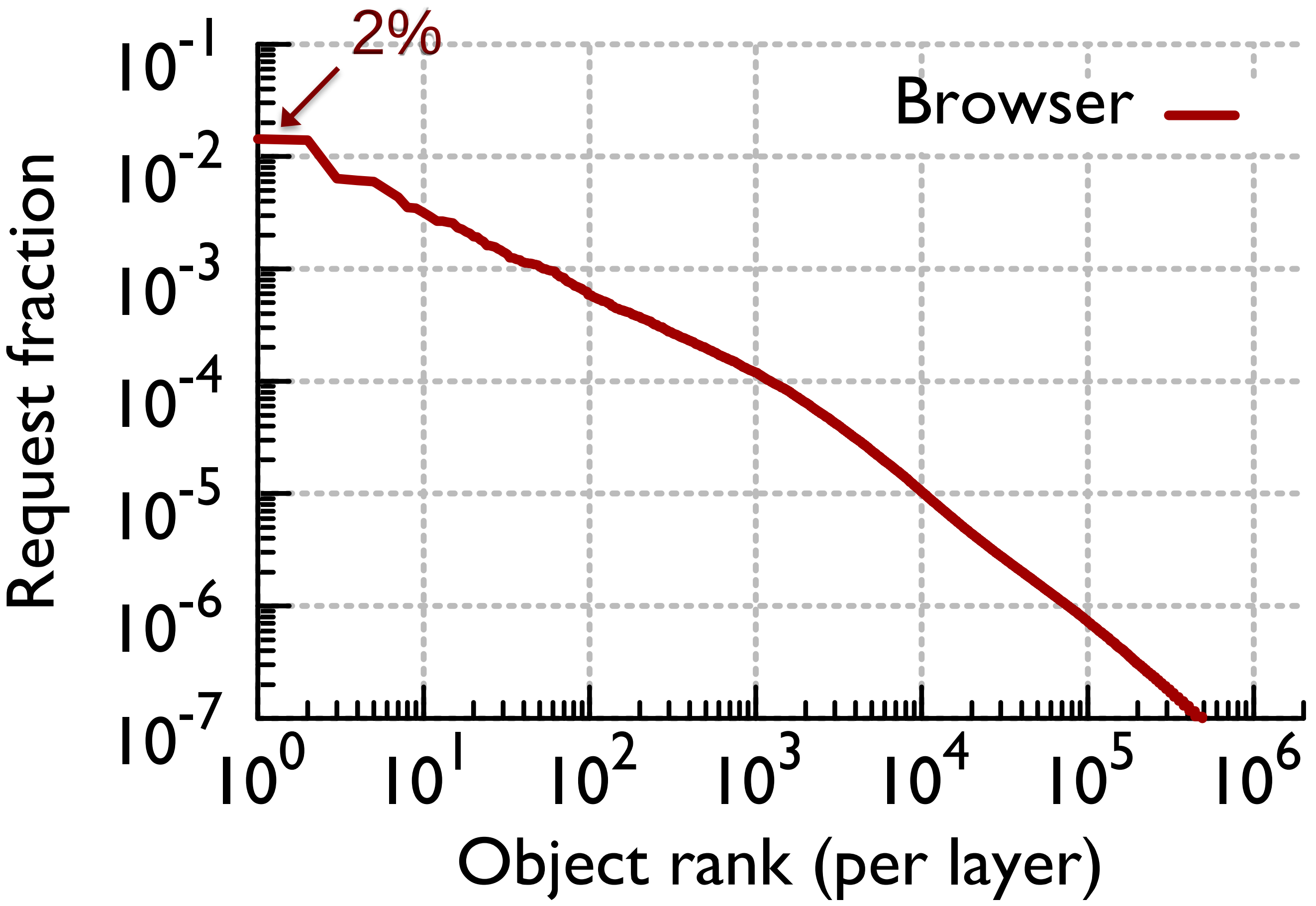


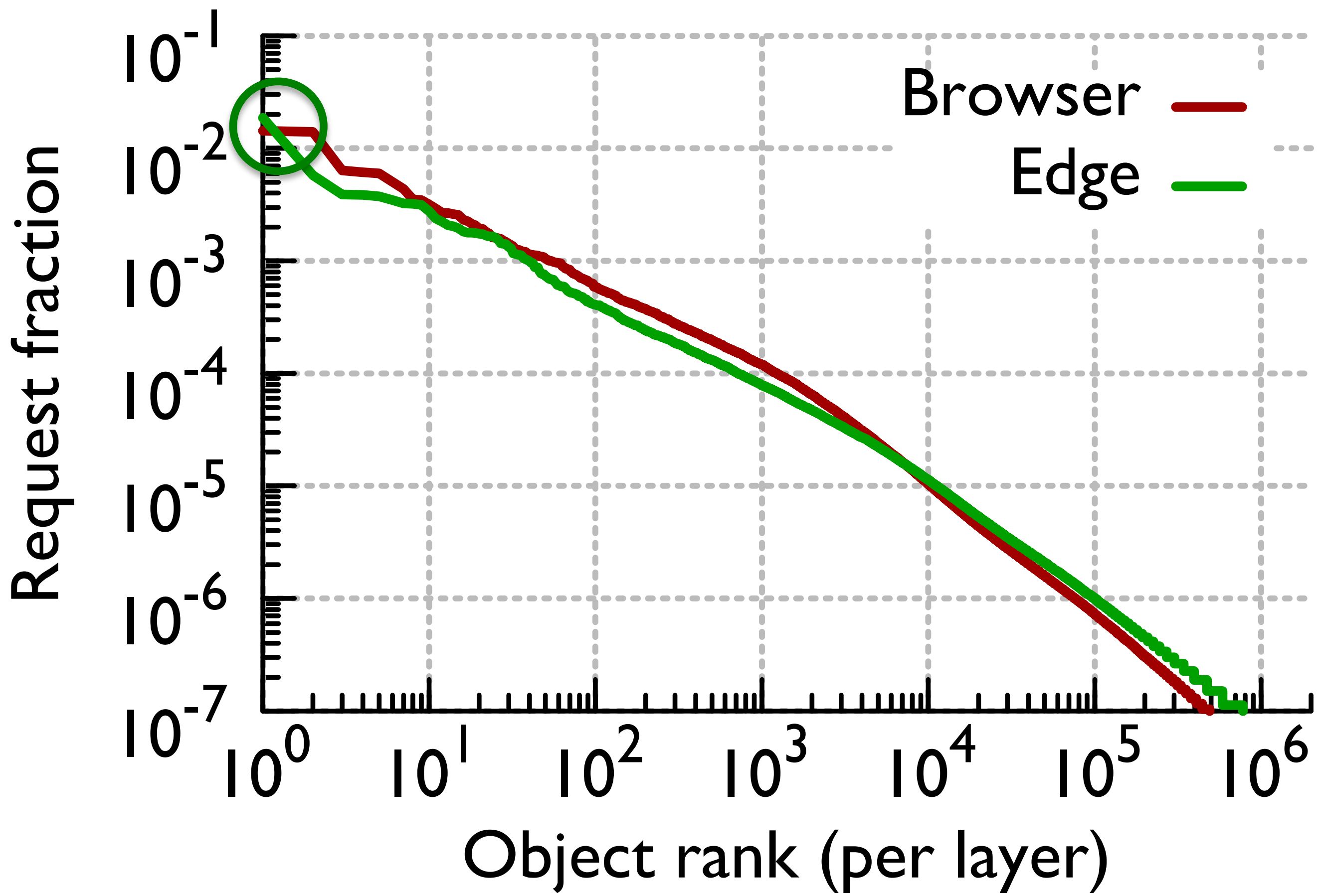
Photo popularity and its cache impact

Popularity Distribution



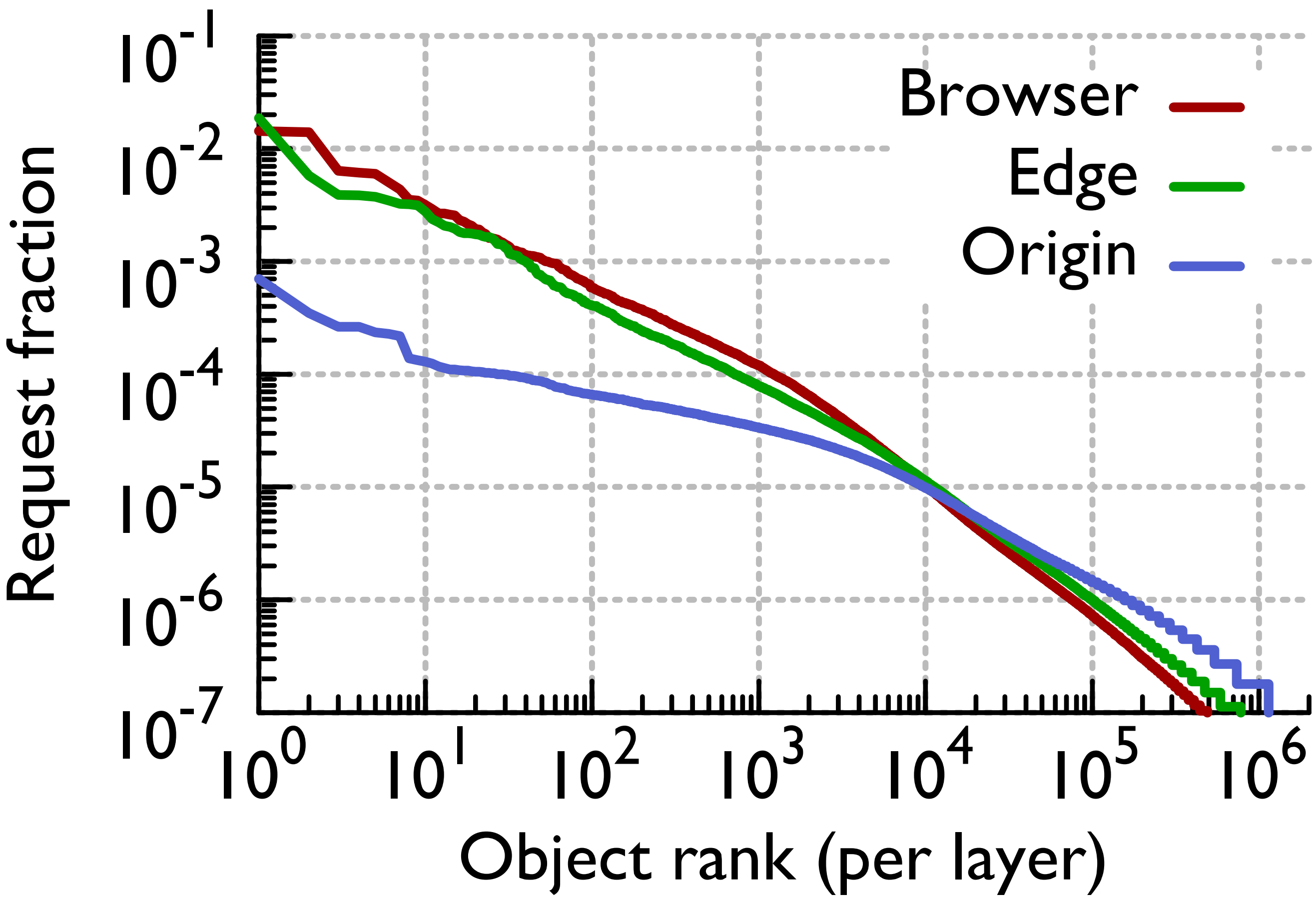
- Browser resembles a power-law distribution

Popularity Distribution



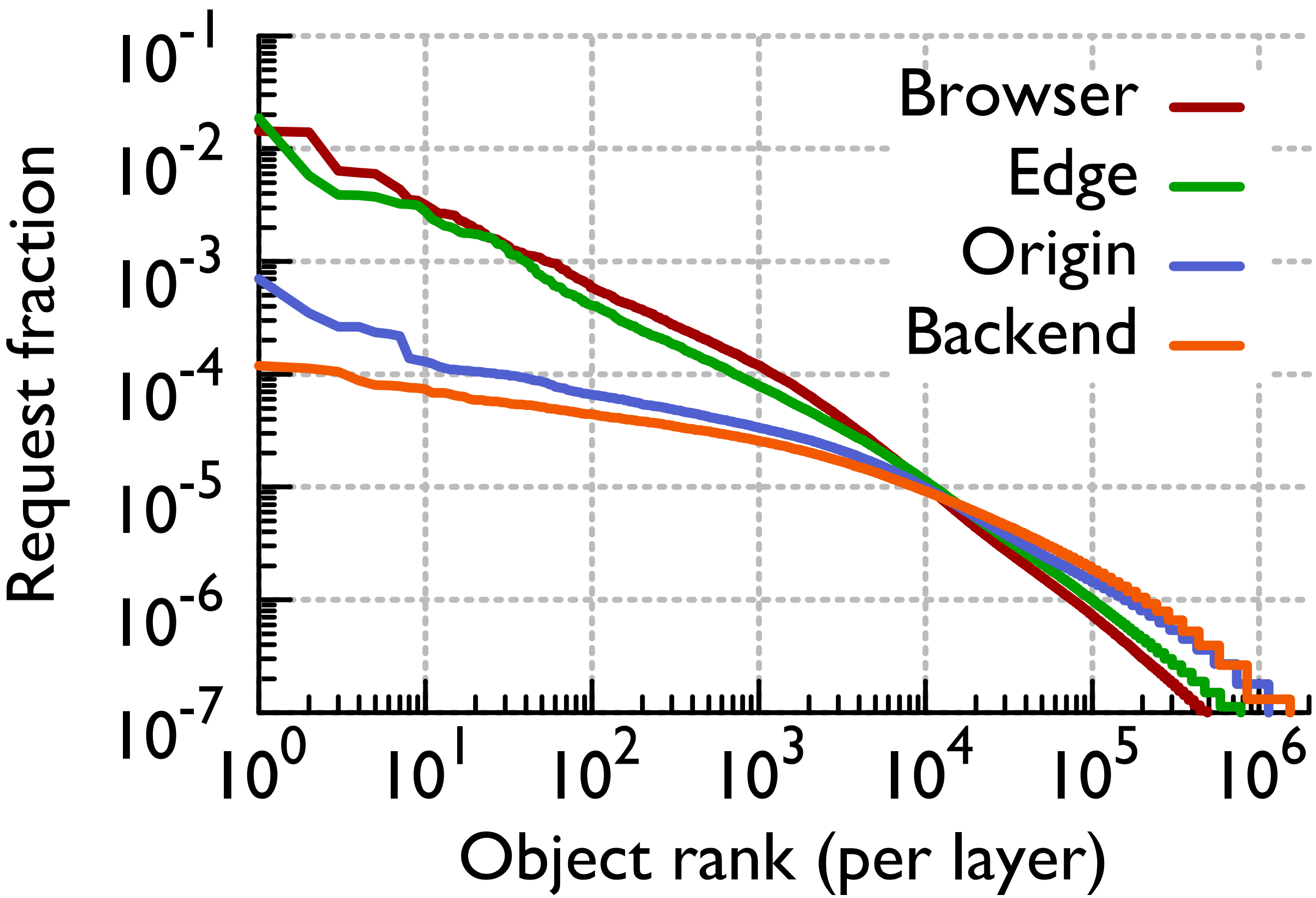
- “Viral” photos becomes the head for Edge

Popularity Distribution



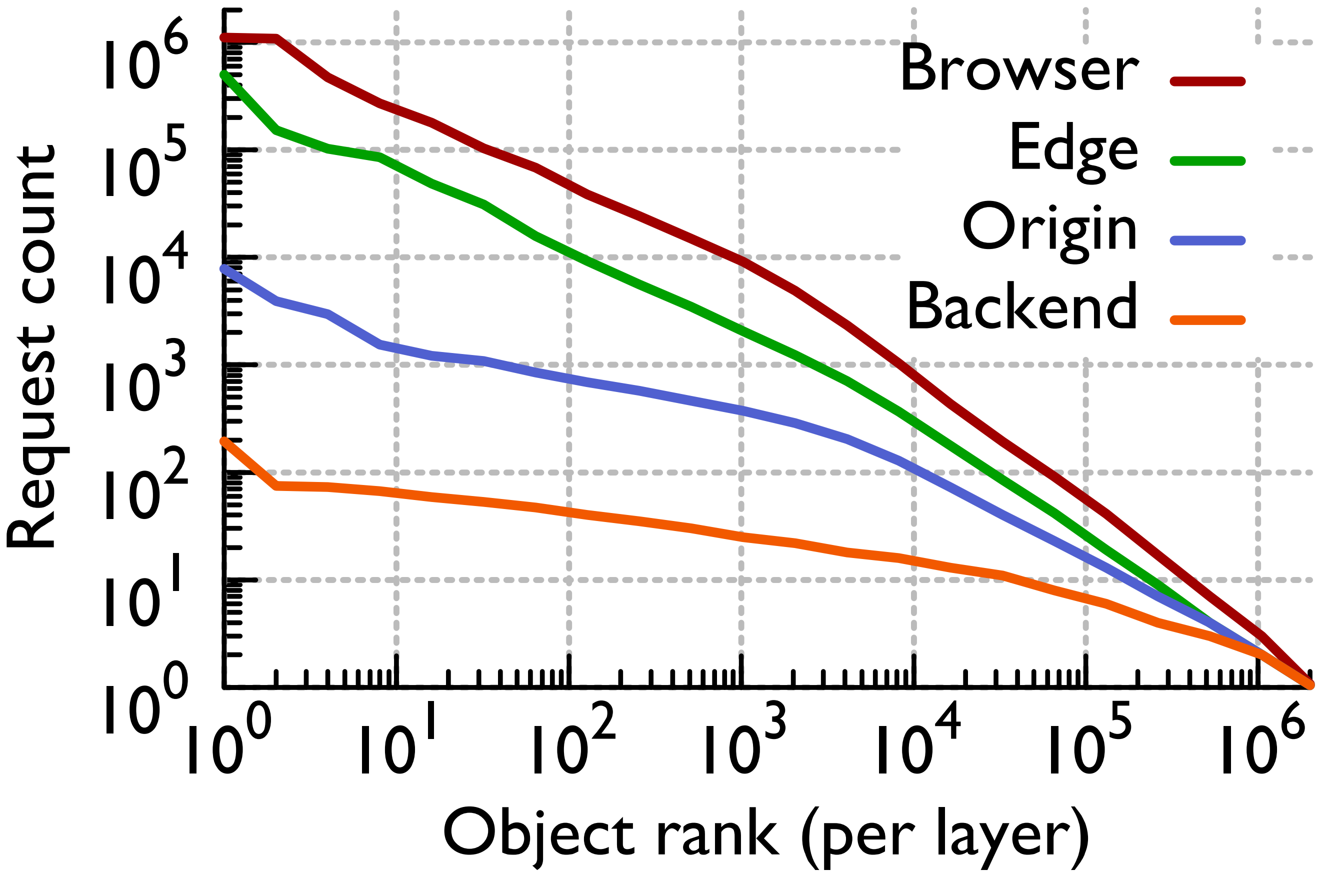
- Skewness is reduced after layers of cache

Popularity Distribution



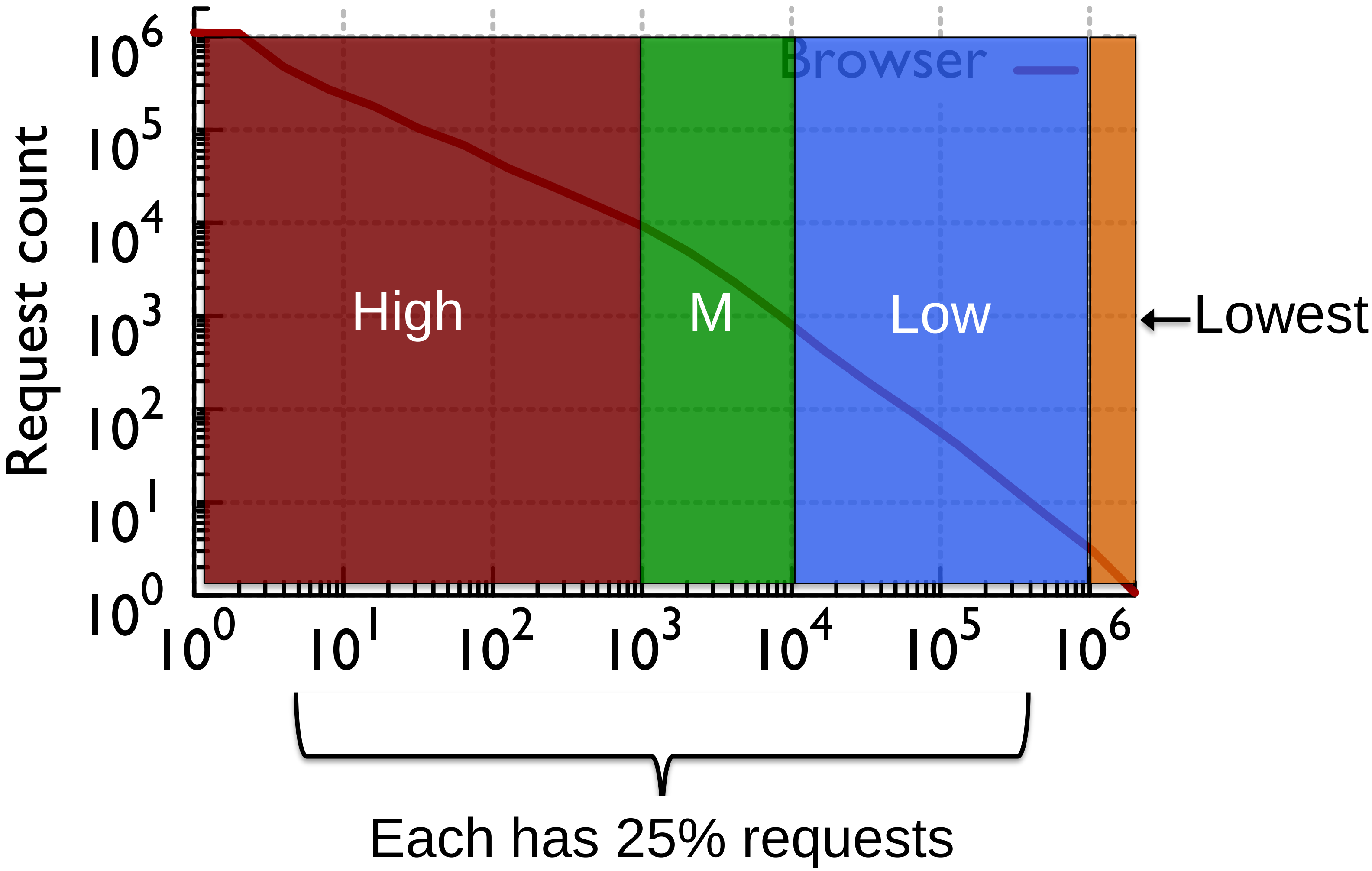
- Backend resembles a stretched exponential dist.

Popularity with Absolute Traffic

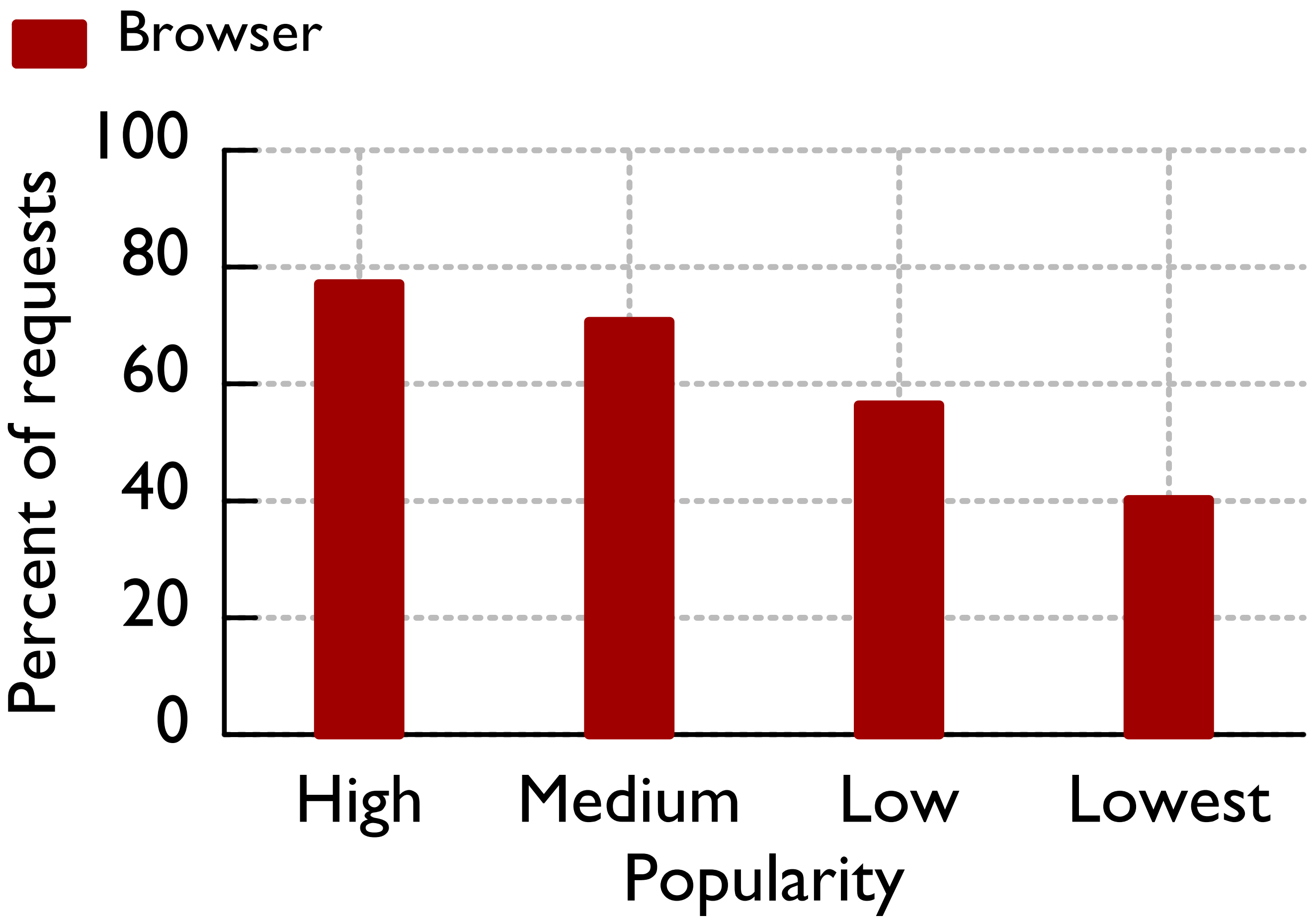


- Storage/cache designers: pick a layer

Popularity Impact on Caches

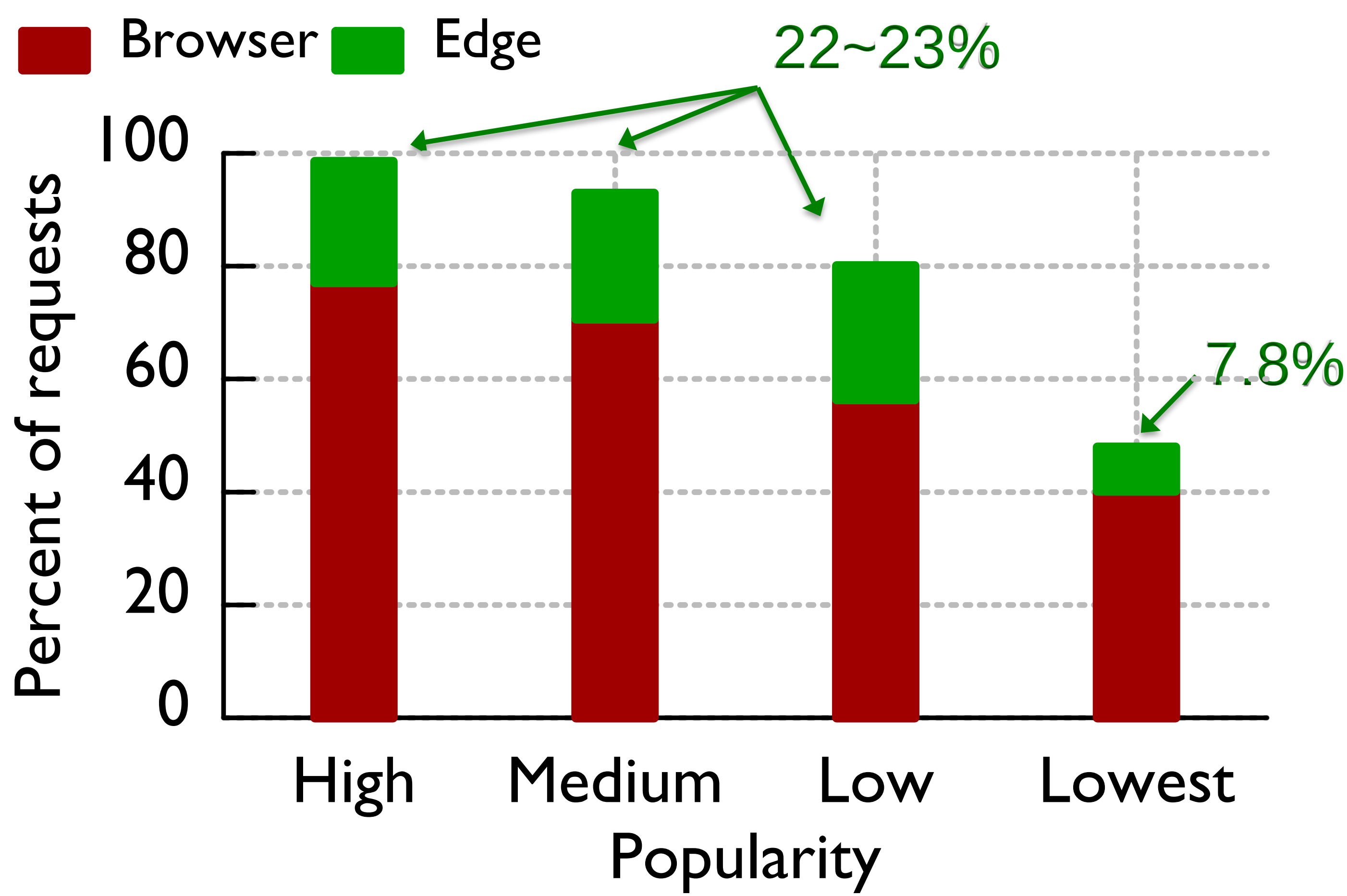


Popularity Impact on Caches



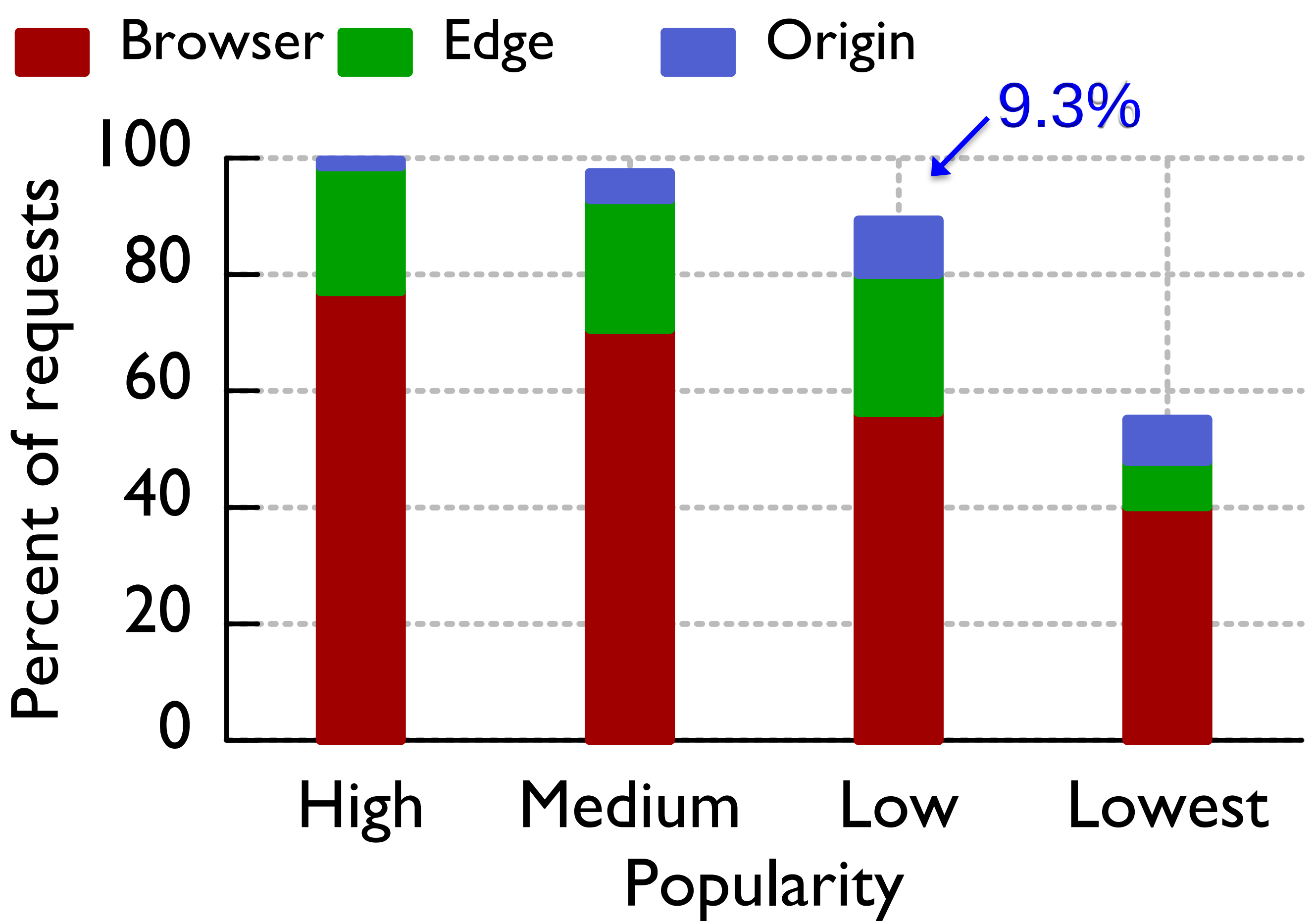
- Browser traffic share decreases gradually

Popularity Impact on Caches



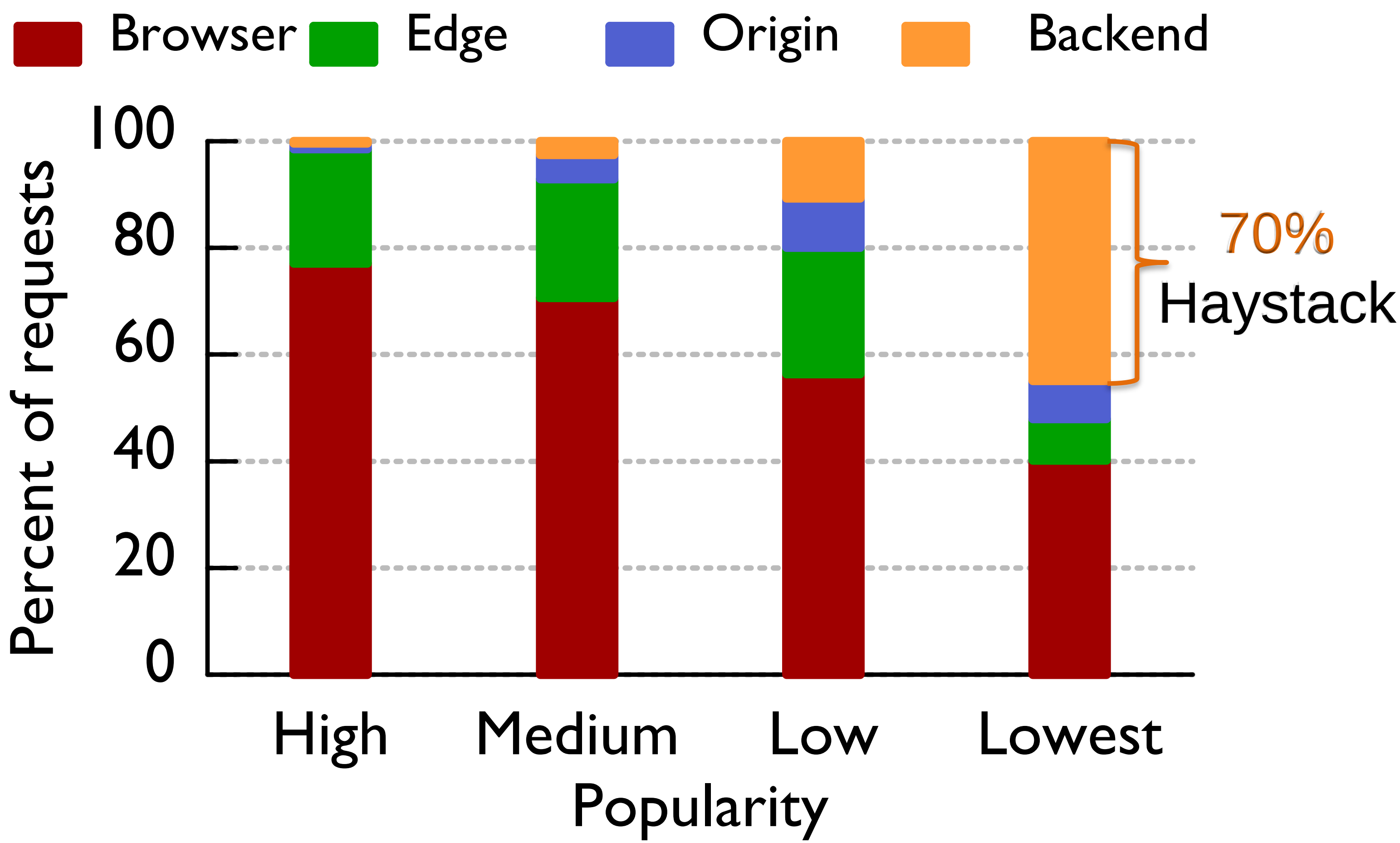
- Edge serves consistent share except for the tail

Popularity Impact on Caches



- Origin contributes most for “low” group

Popularity Impact on Caches



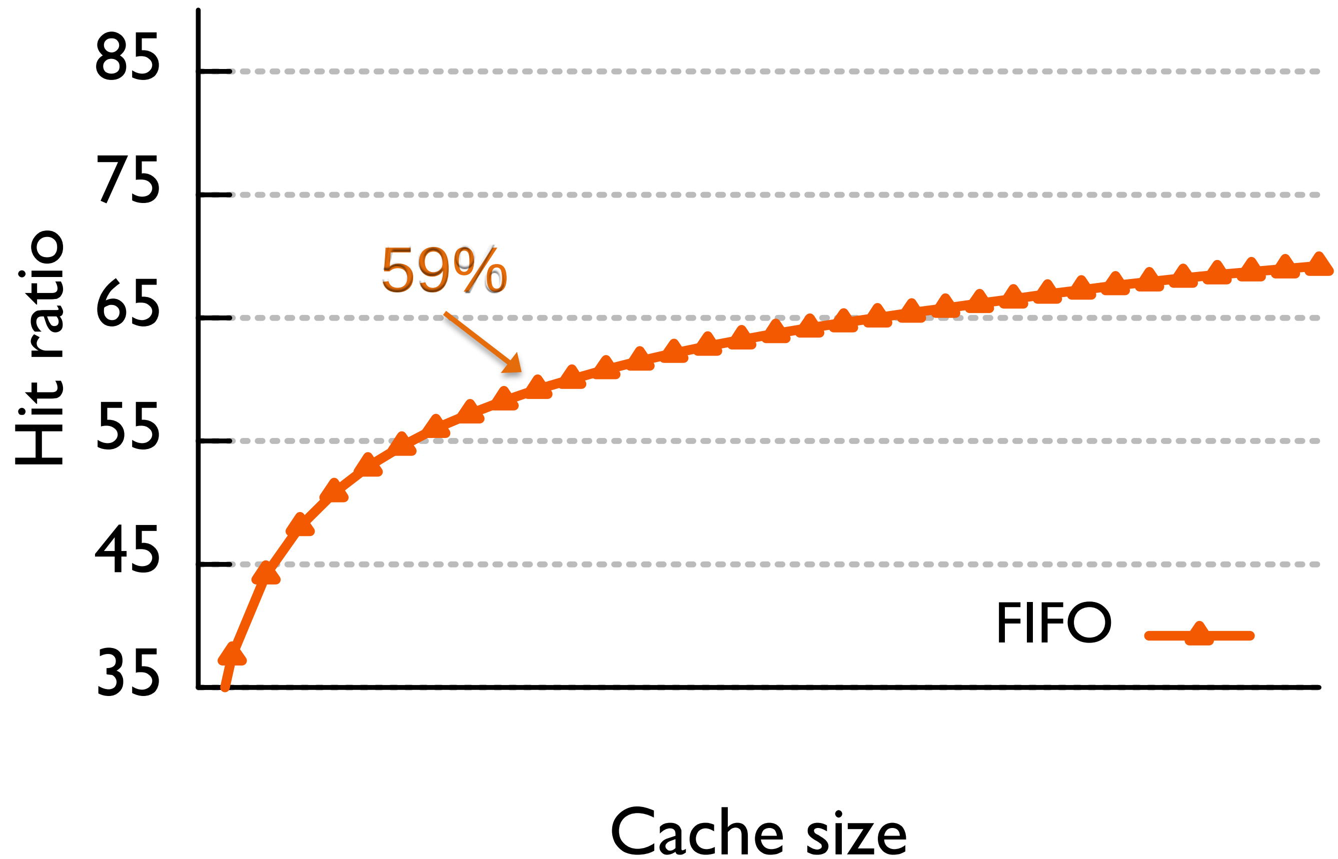
- Backend serves the tail

Can we make the cache better?

Simulation

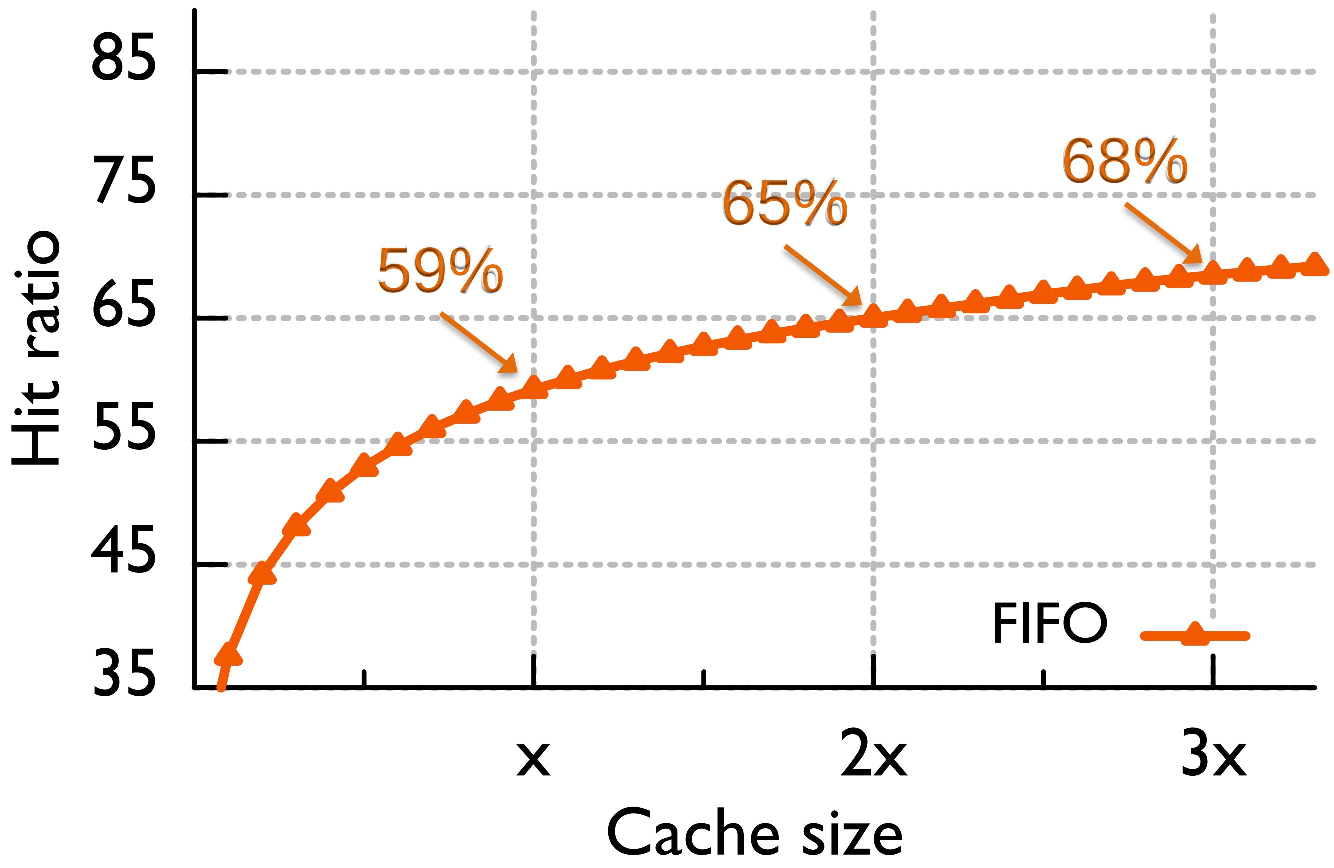
- Replay the trace (25% warm up)
- Estimate the base cache size
- Evaluate two hit-ratios (object-wise, byte-wise)

Edge Cache with Different Sizes



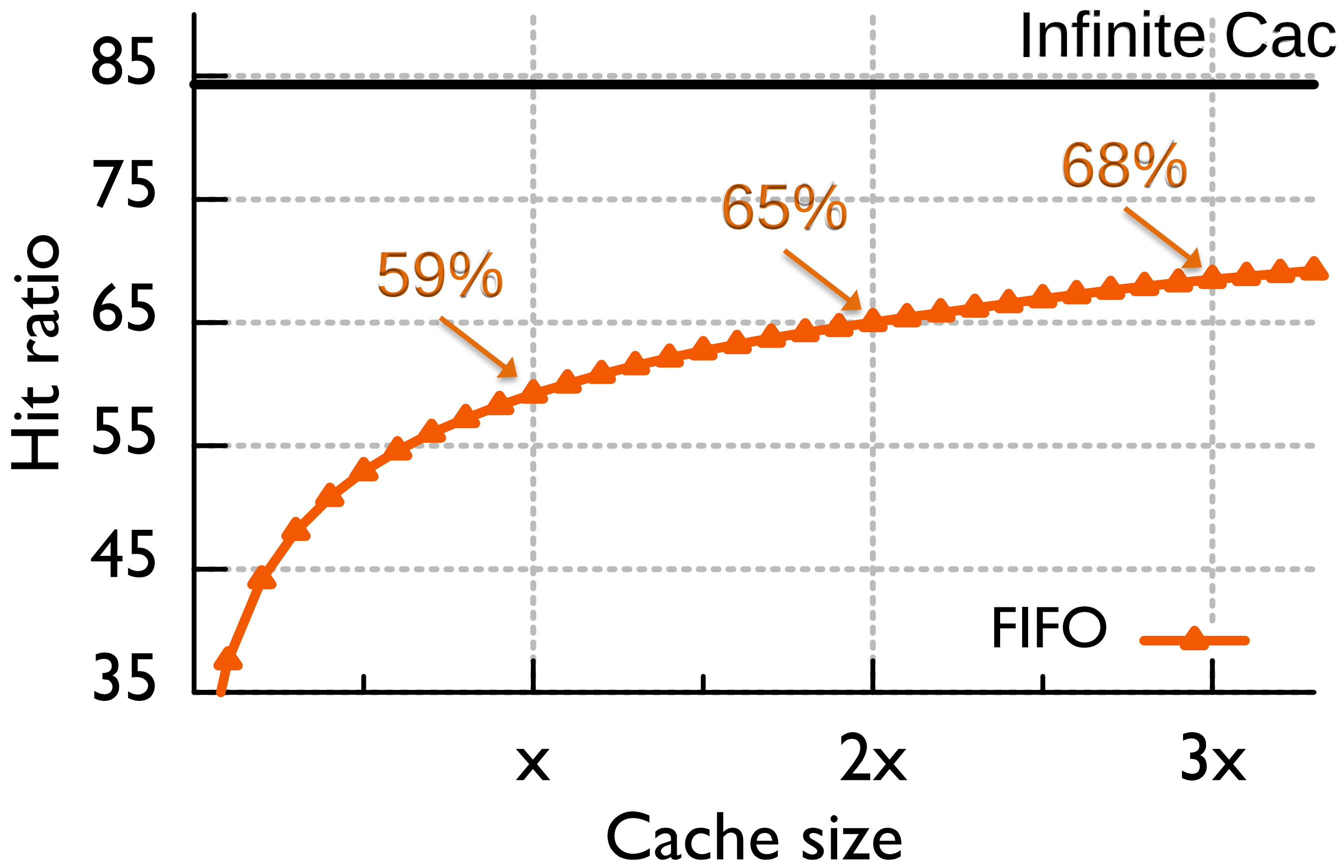
- Picked San Jose edge (high traffic, median hit ratio)

Edge Cache with Different Sizes



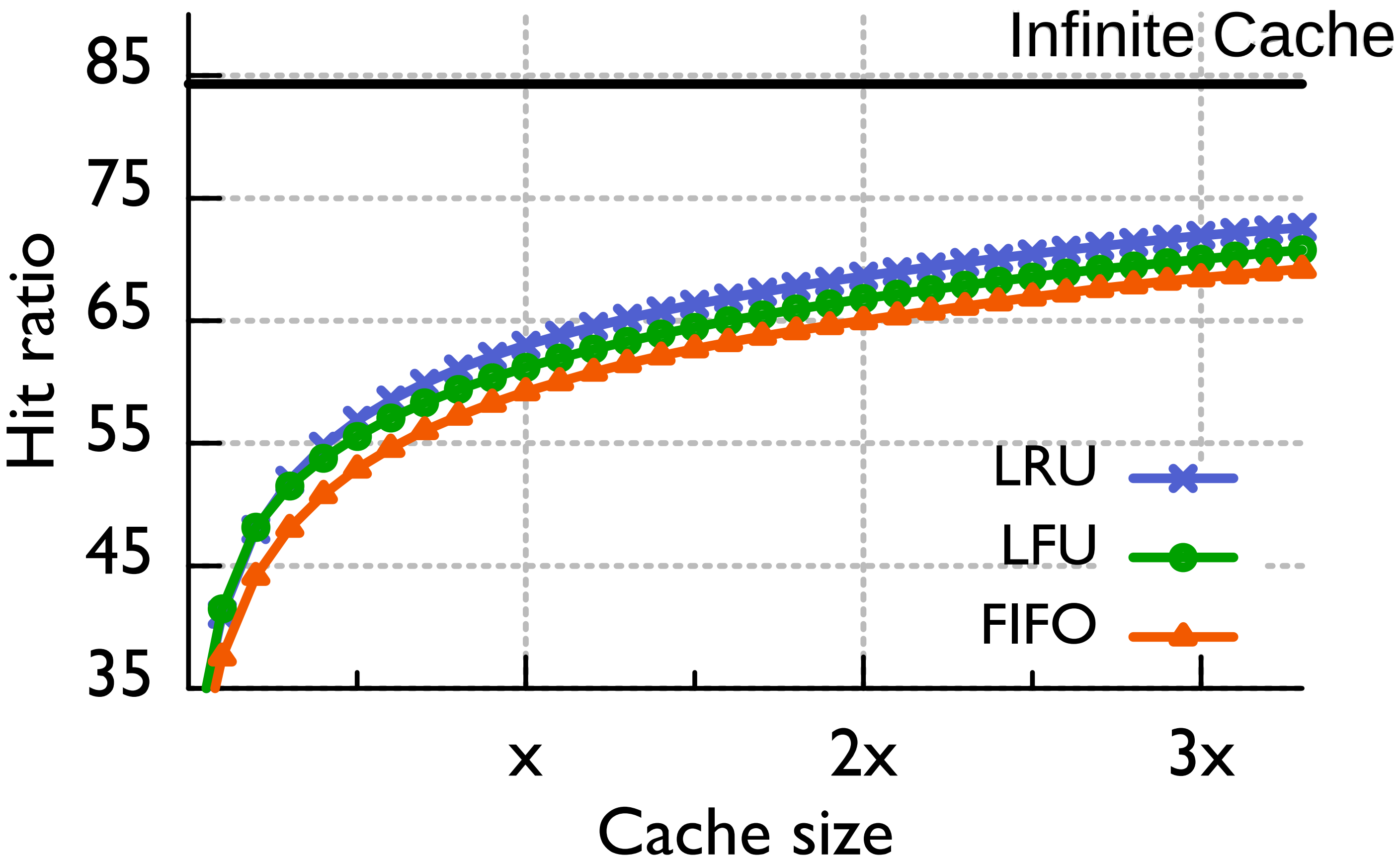
- "x" estimates current deployment size (59% hit ratio)

Edge Cache with Different Sizes



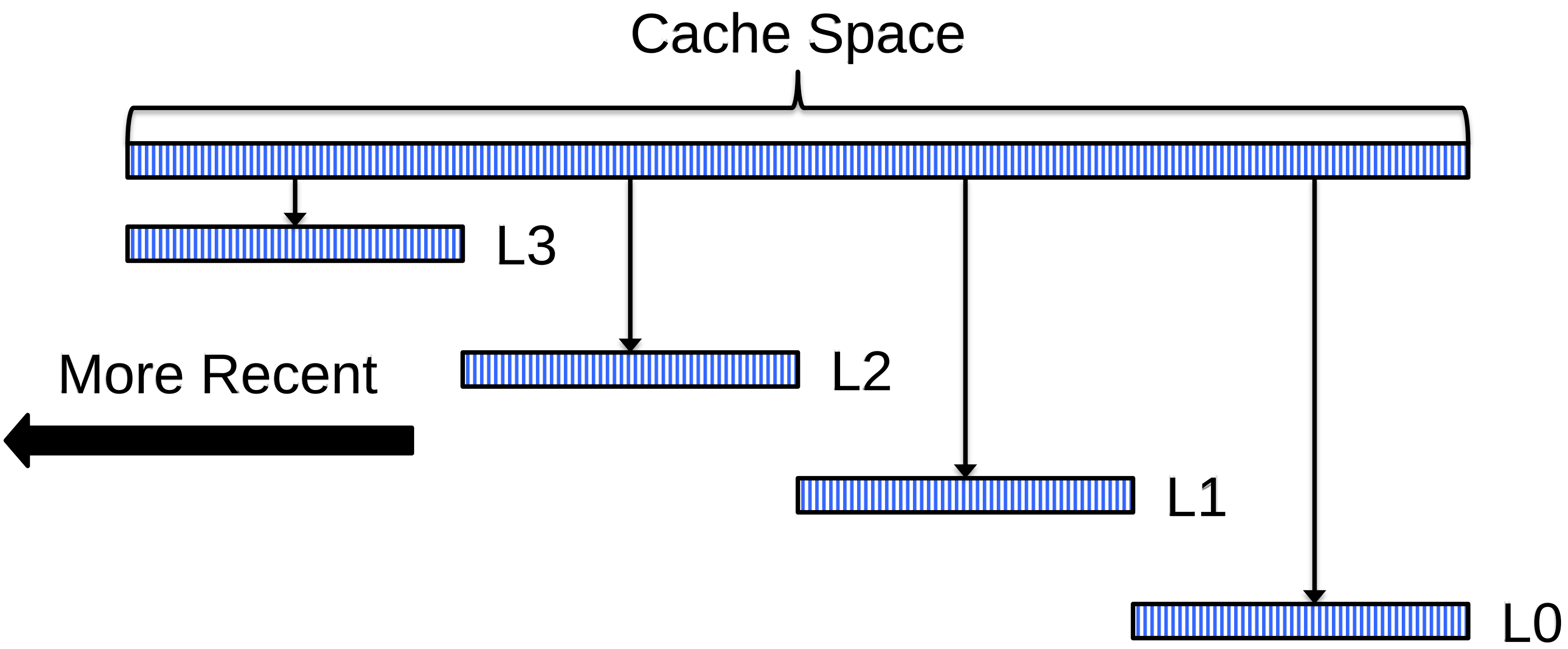
- "Infinite" size ratio needs 45x of current capacity

Edge Cache with Different Algos

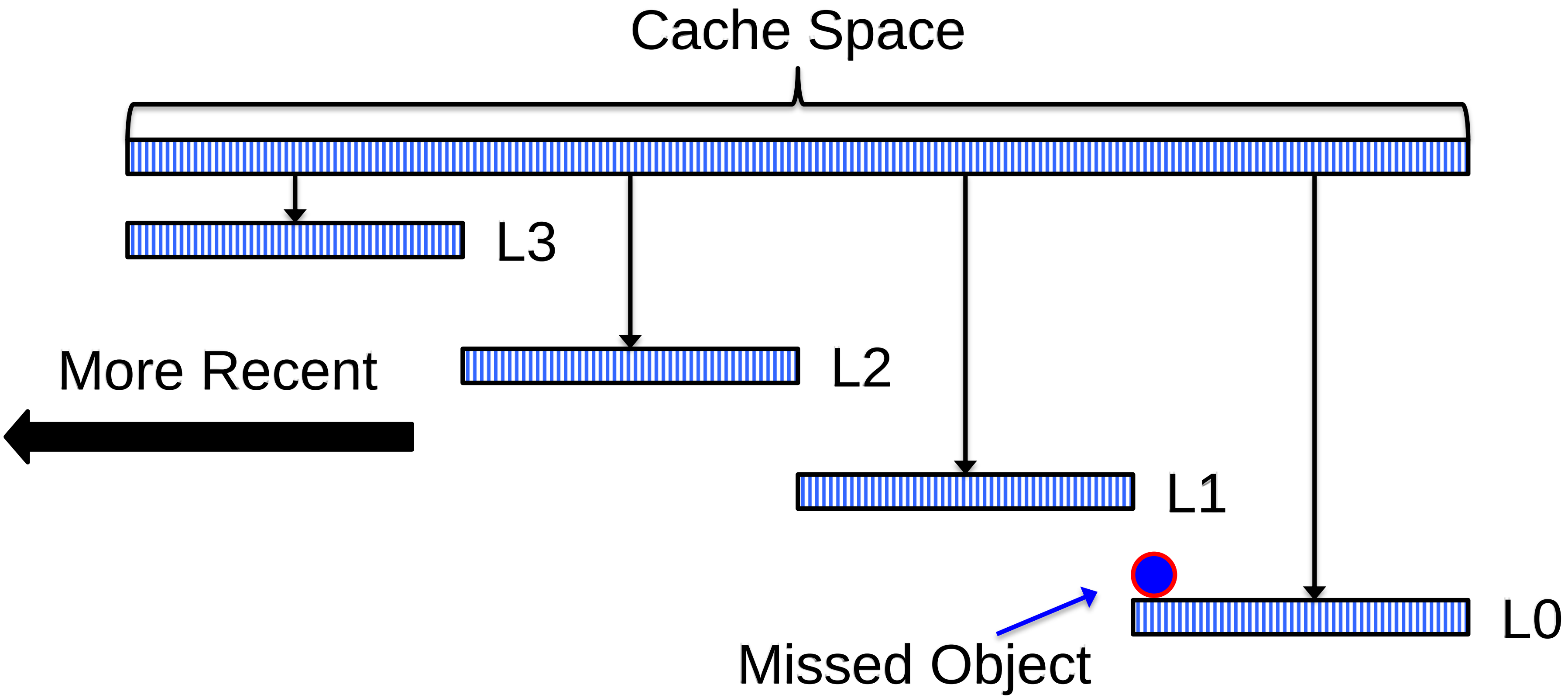


- Both LRU and LFU outperforms FIFO slightly

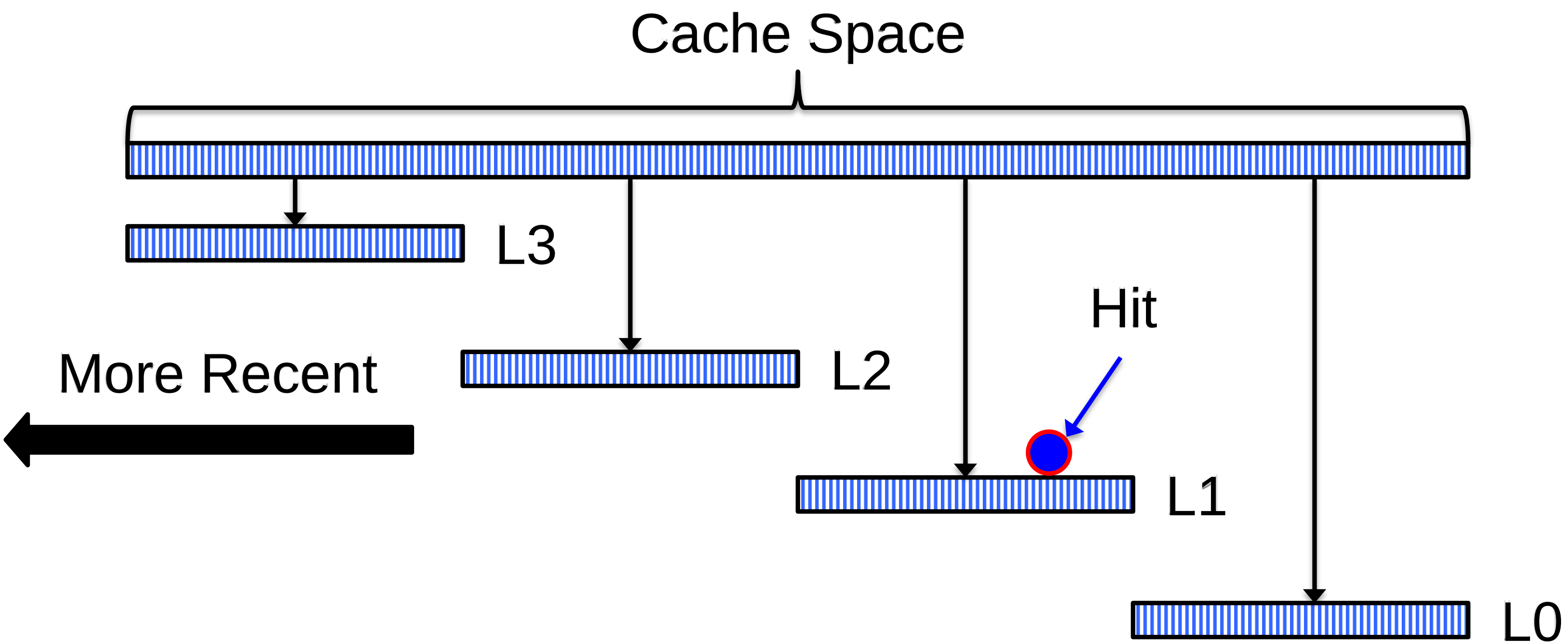
S4LRU



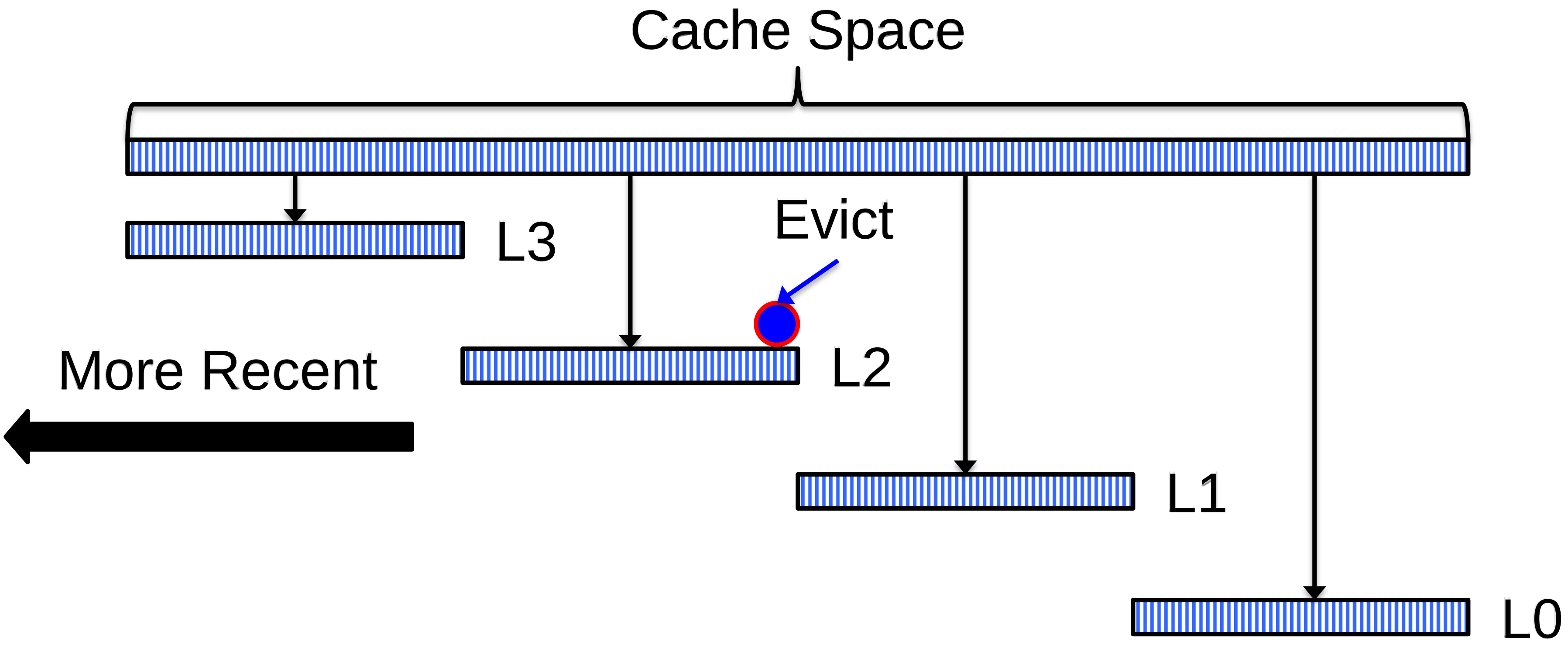
S4LRU



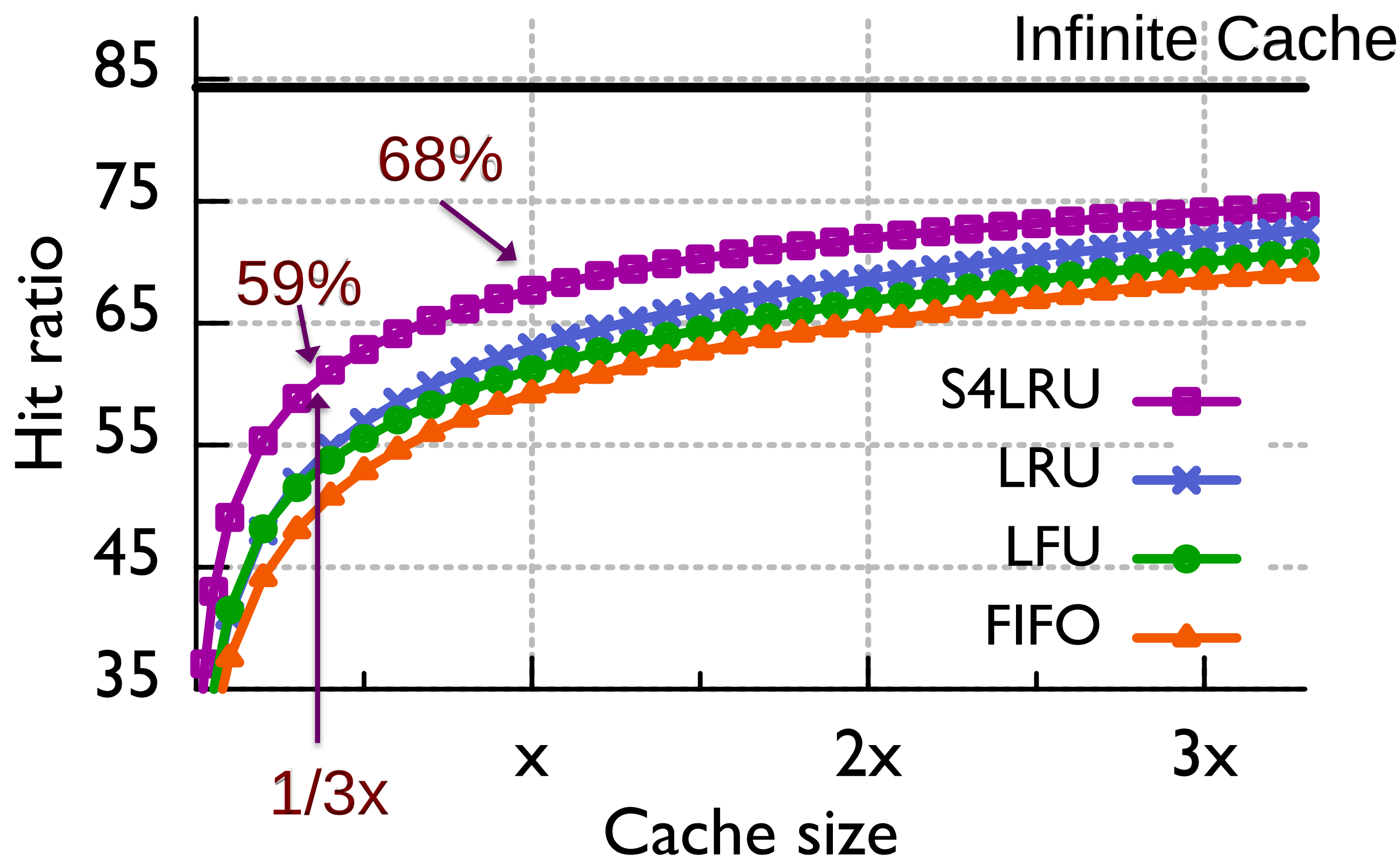
S4LRU



S4LRU

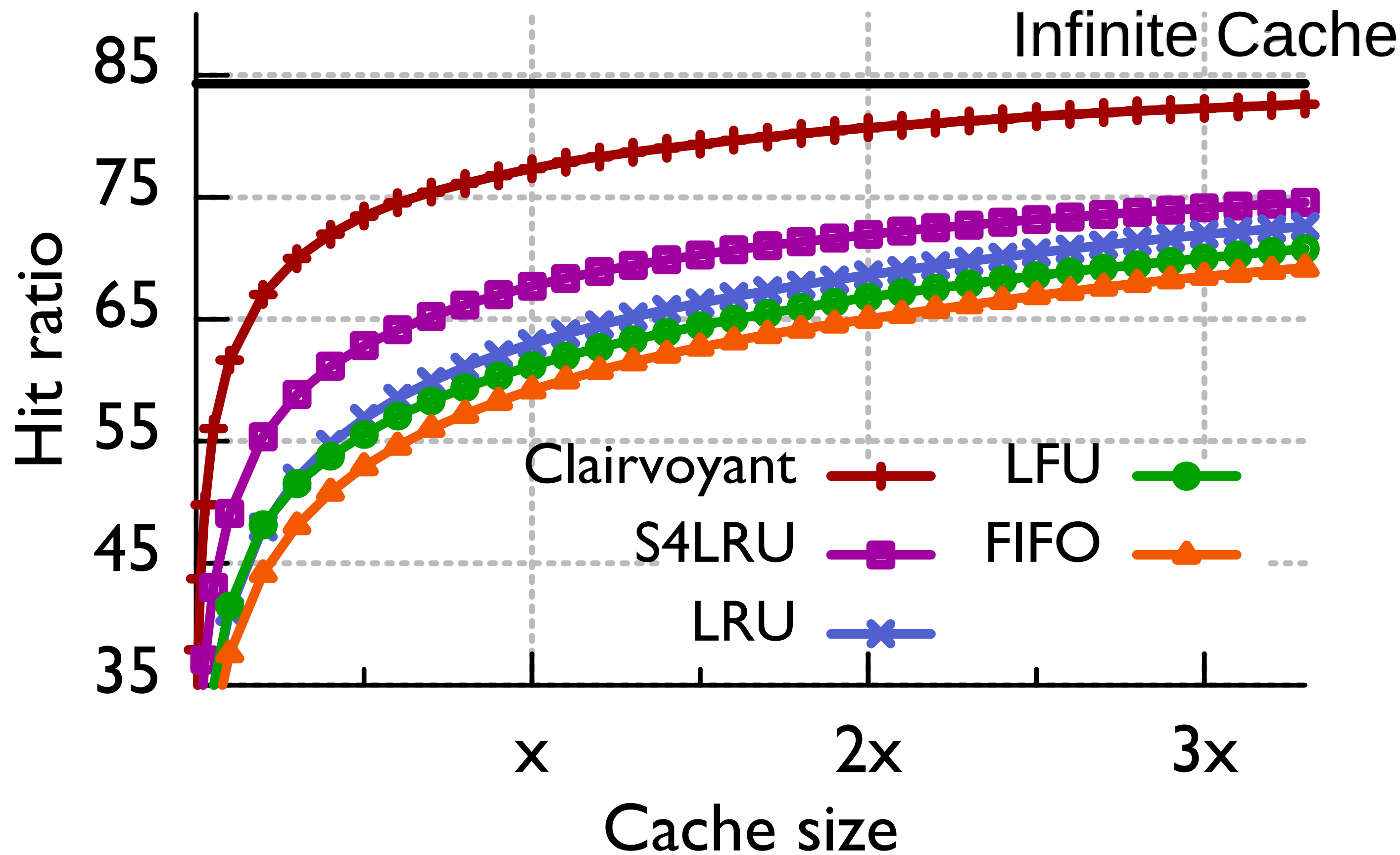


Edge Cache with Different Algos



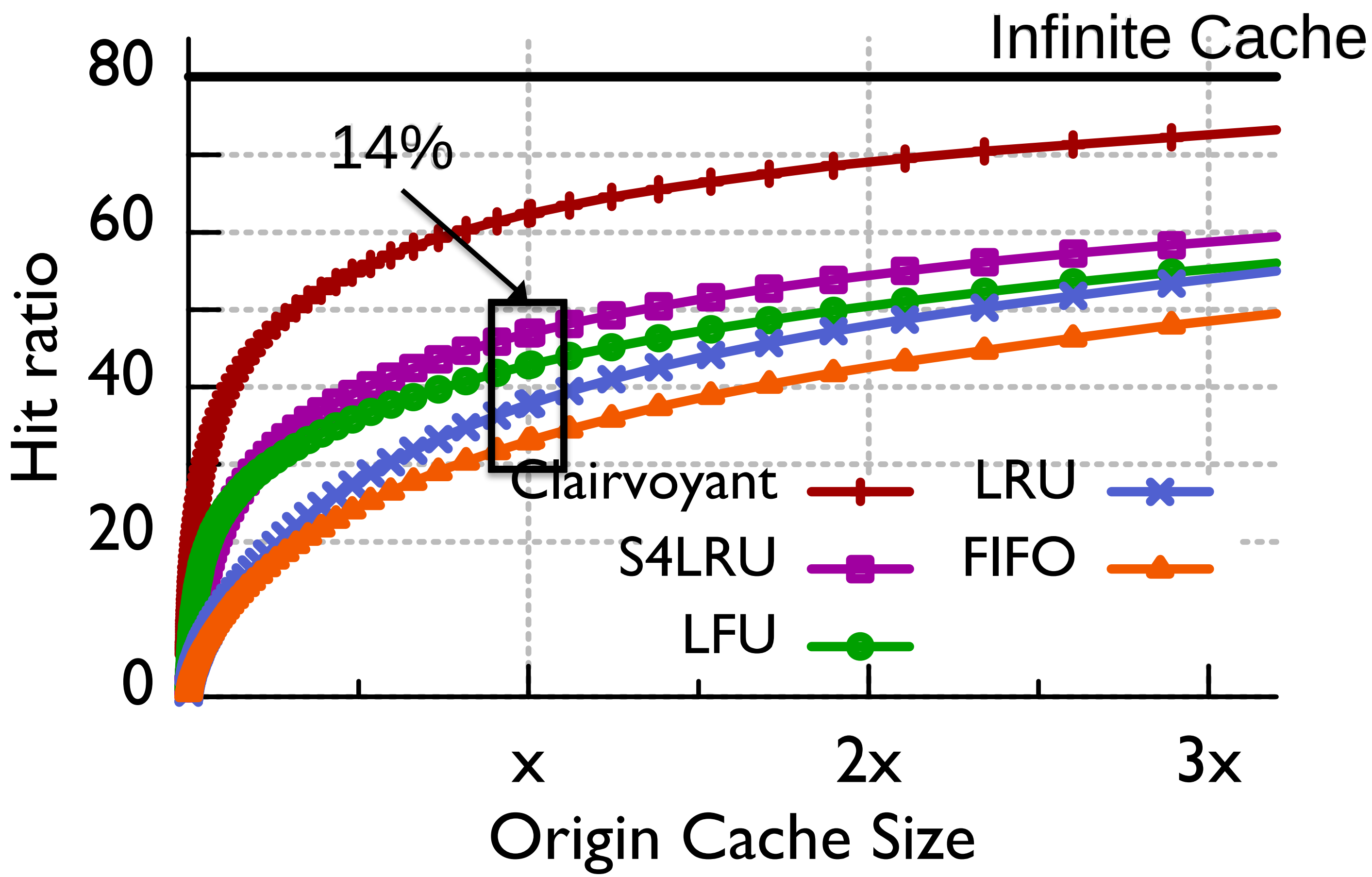
- S4LRU improves the most

Edge Cache with Different Algos



- Clairvoyant (Bélády) shows much improvement space

Origin Cache



- S4LRU improves Origin more than Edge

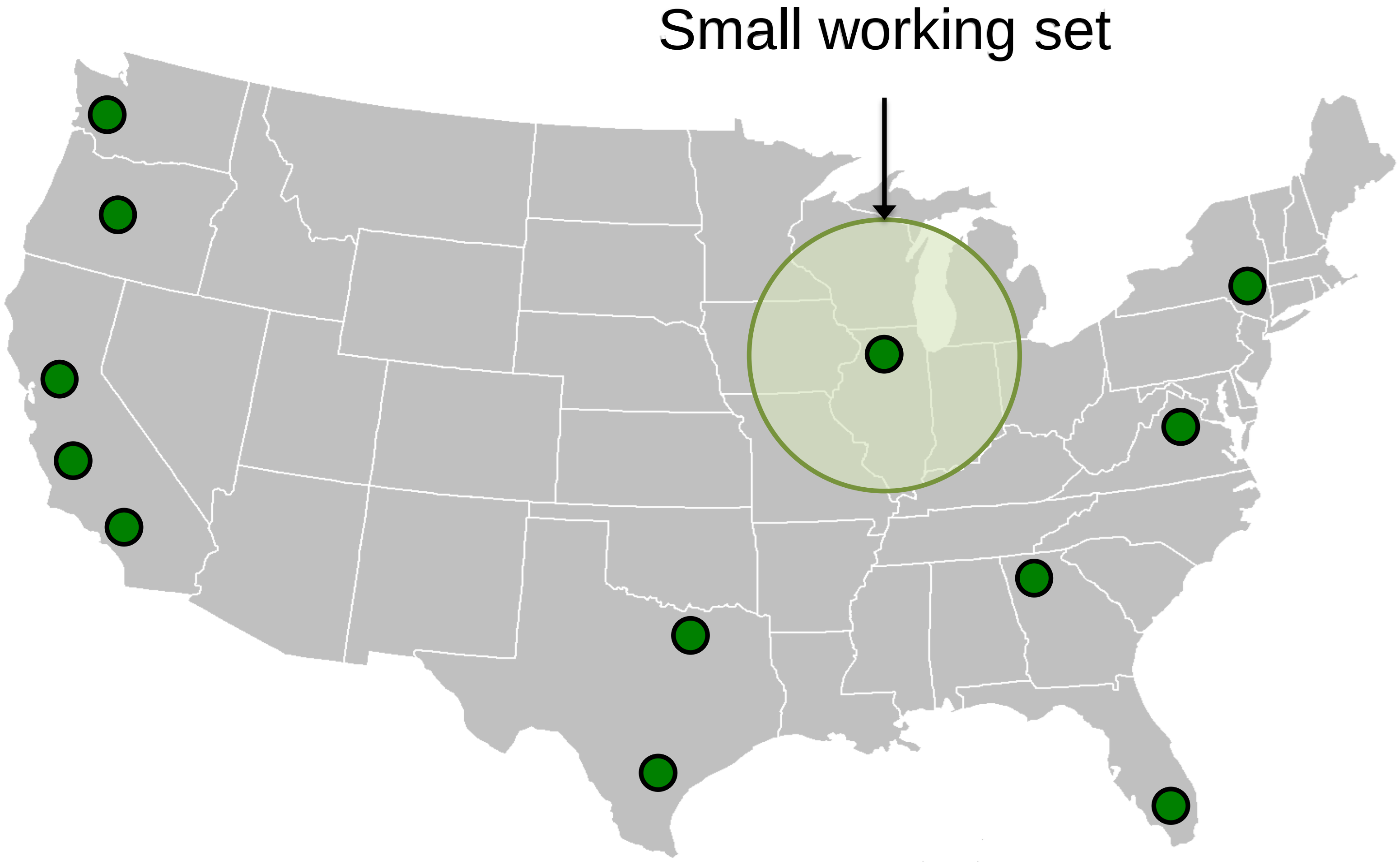
Which Photo to Cache

- Recency & frequency leads S4LRU

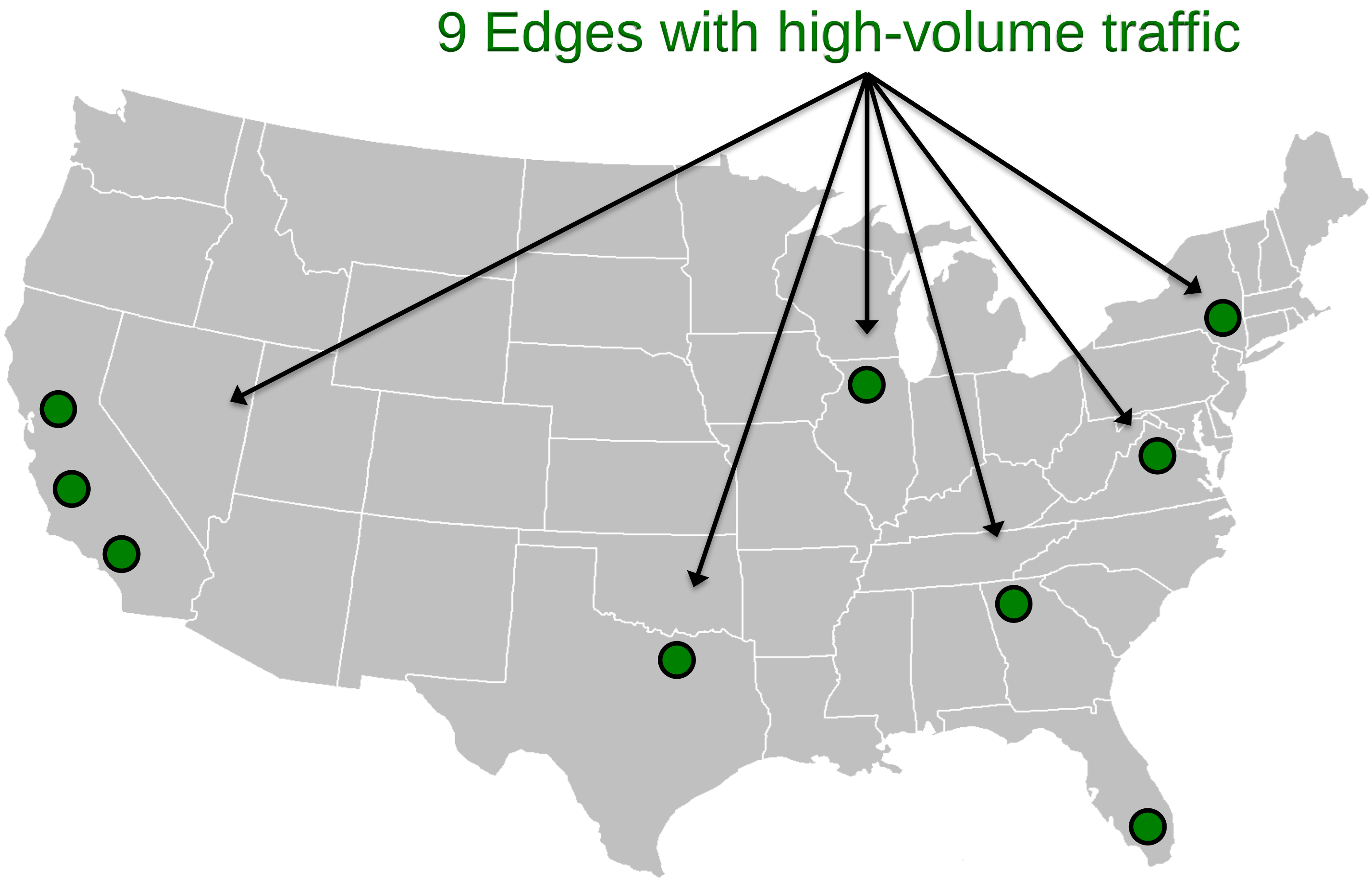
- Does age, social factors also play a role?

Collaborative cache on the Edge

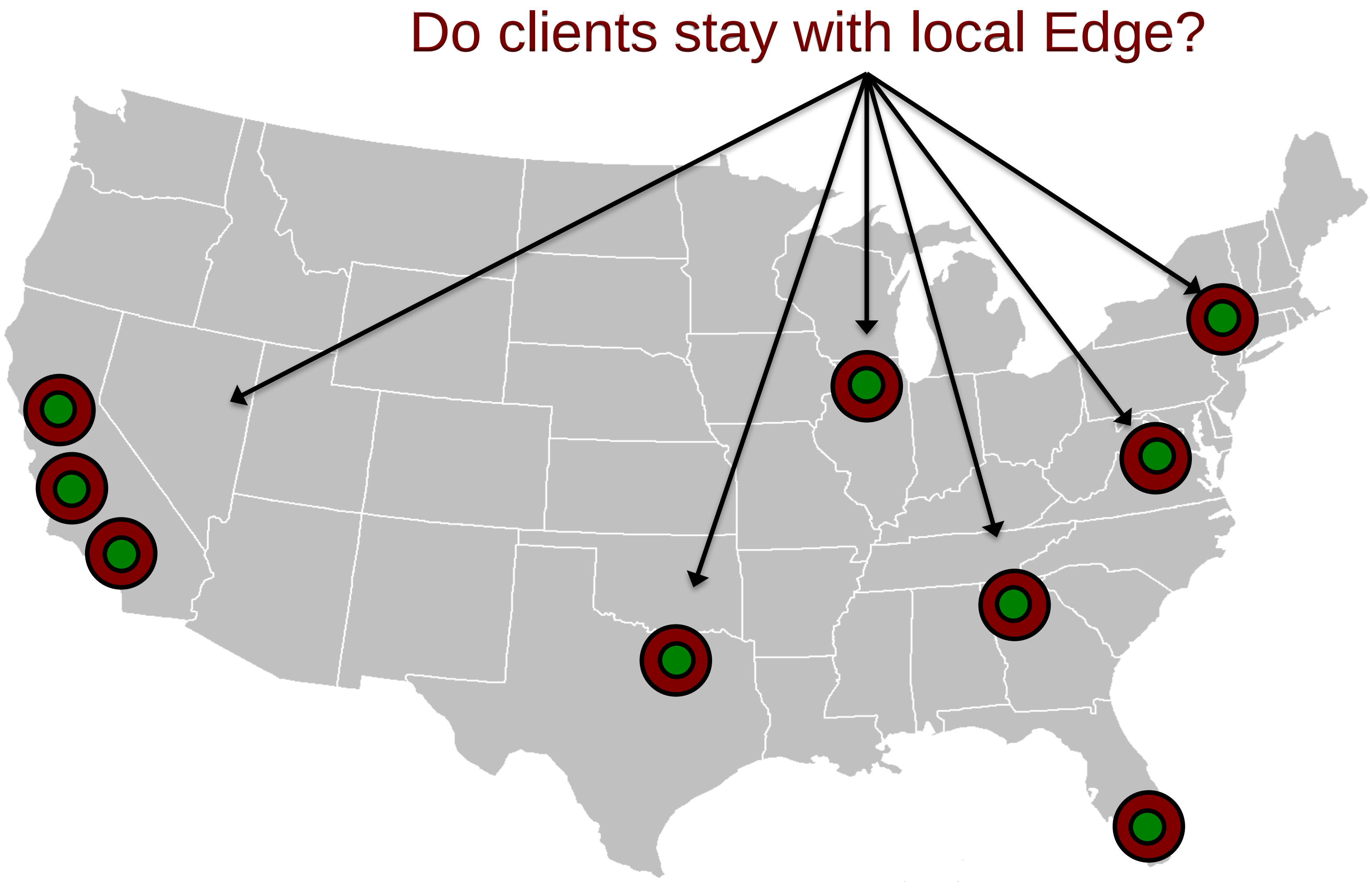
Geographic Coverage of Edge



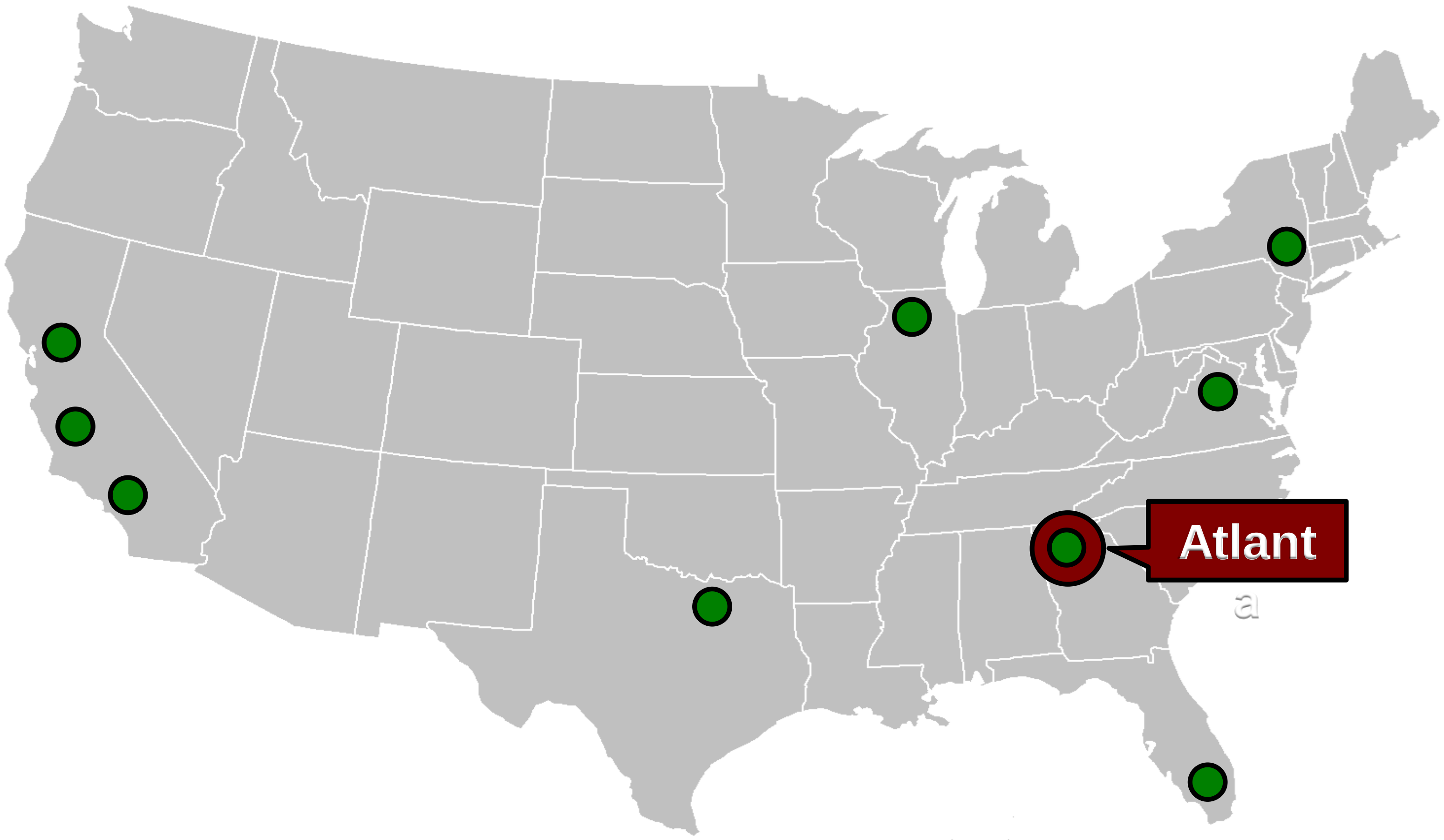
Geographic Coverage of **Edge**



Geographic Coverage of Edge

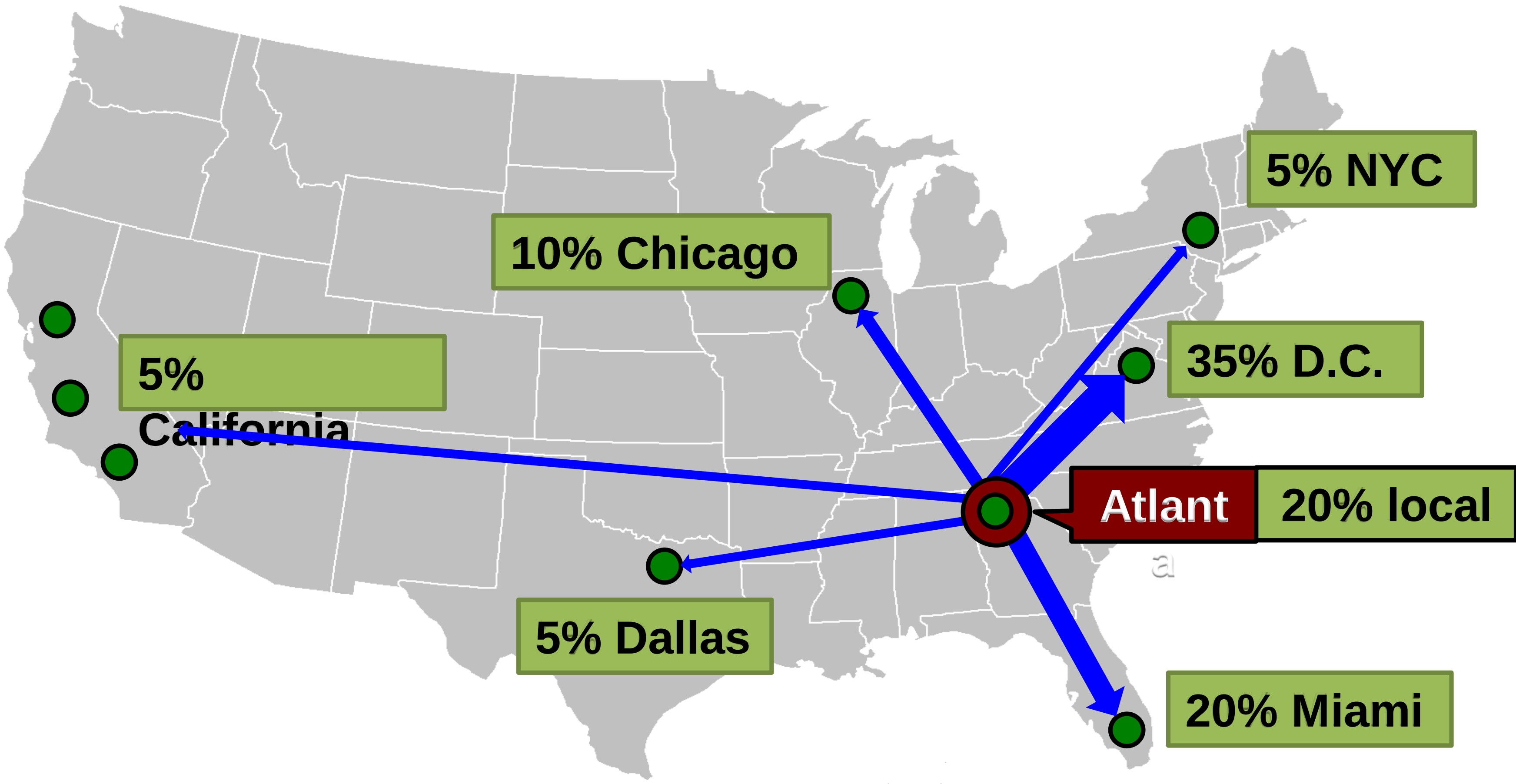


Geographic Coverage of Edge



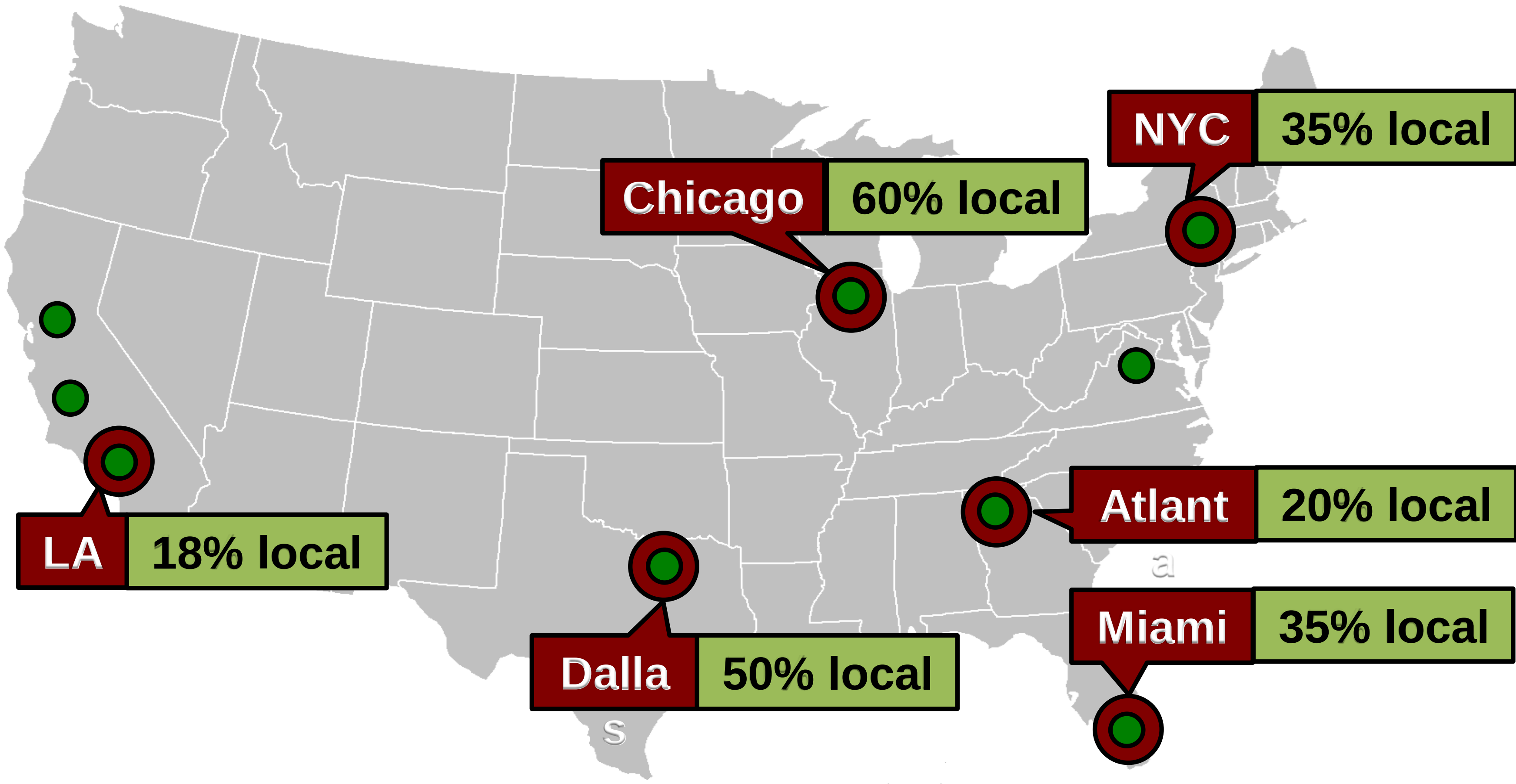
Geographic Coverage of Edge

- Atlanta has 80% requests served by remote Edges

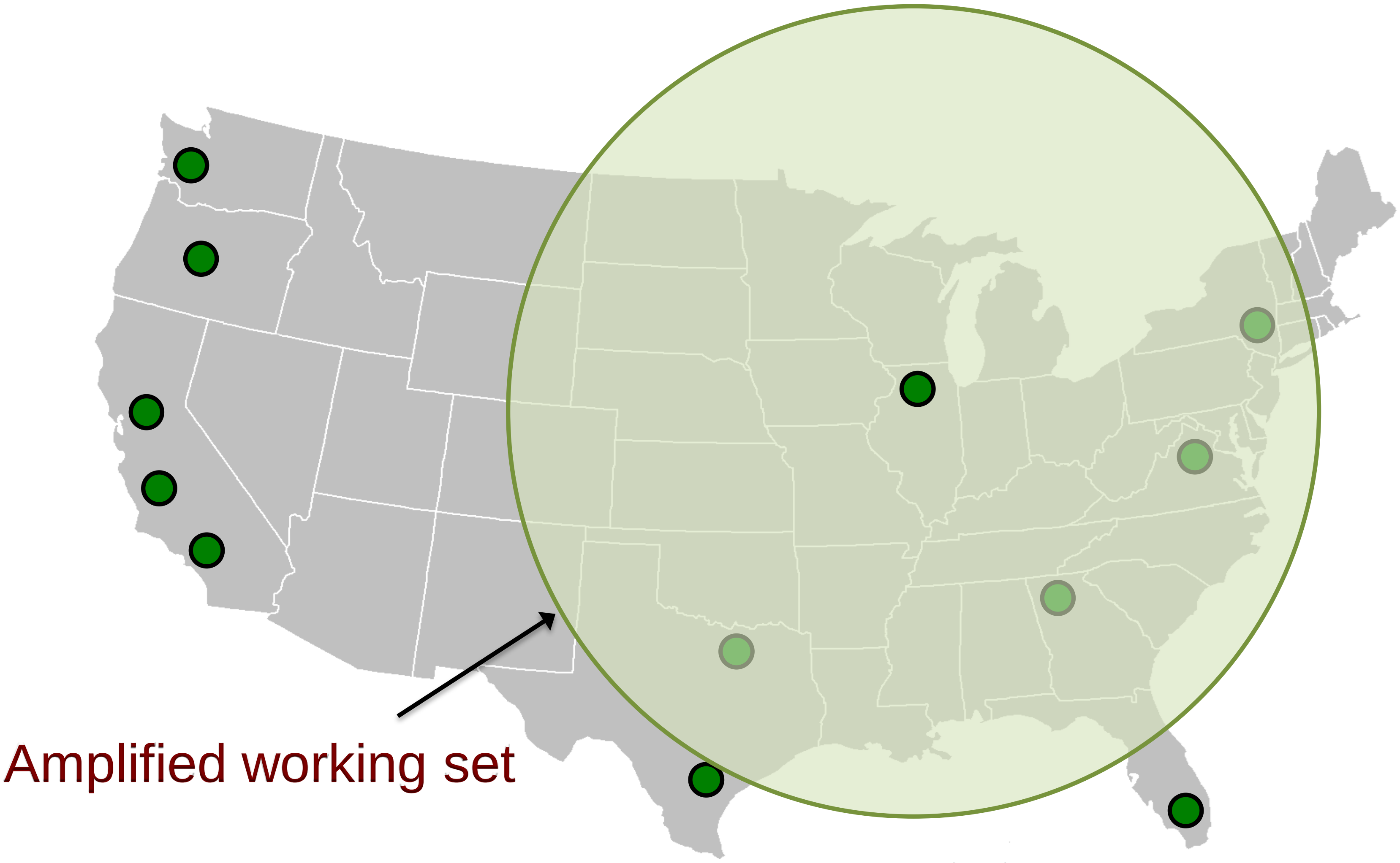


Geographic Coverage of Edge

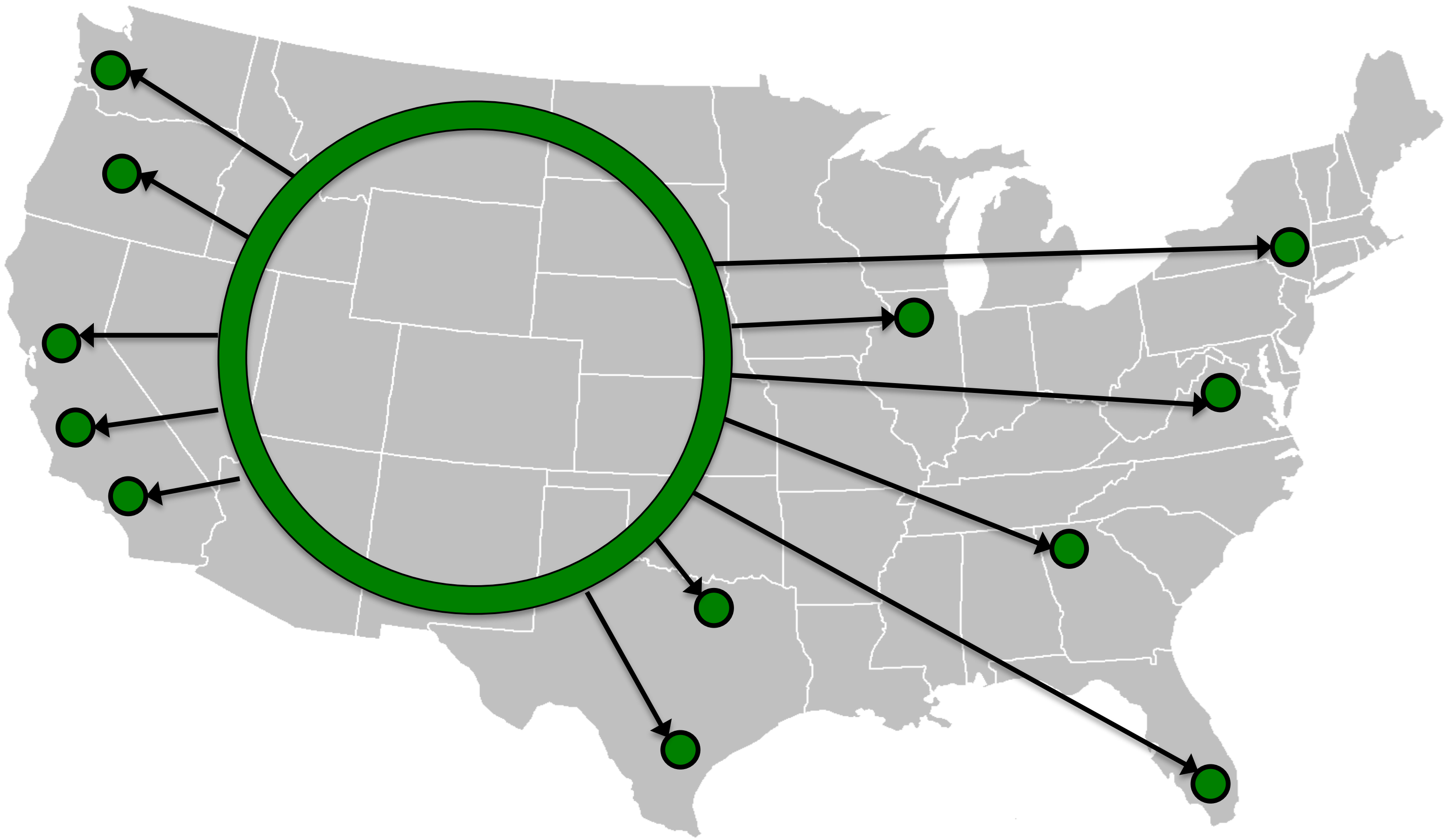
- Substantial remote traffic is normal



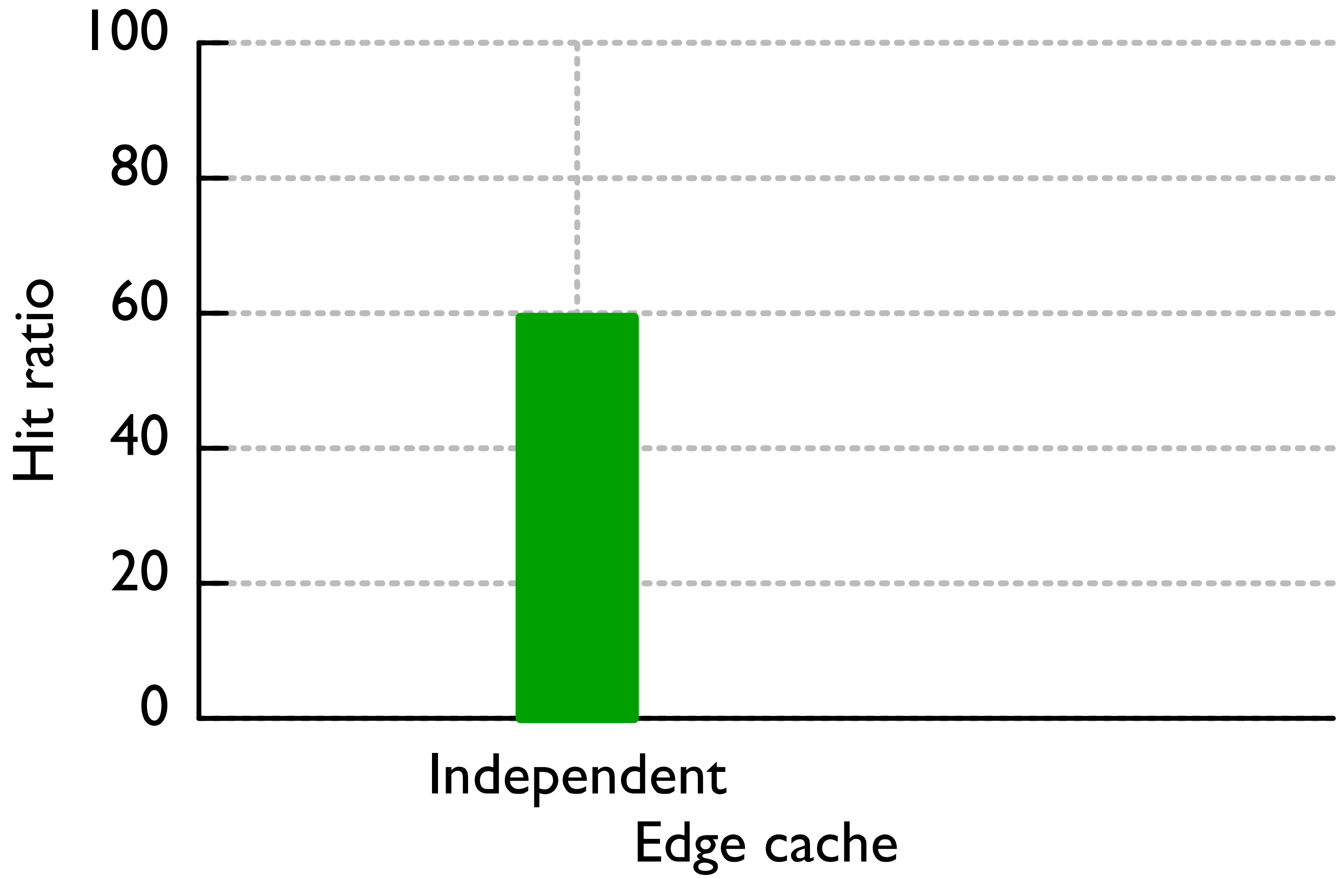
Geographic Coverage of Edge



Collaborative Edge

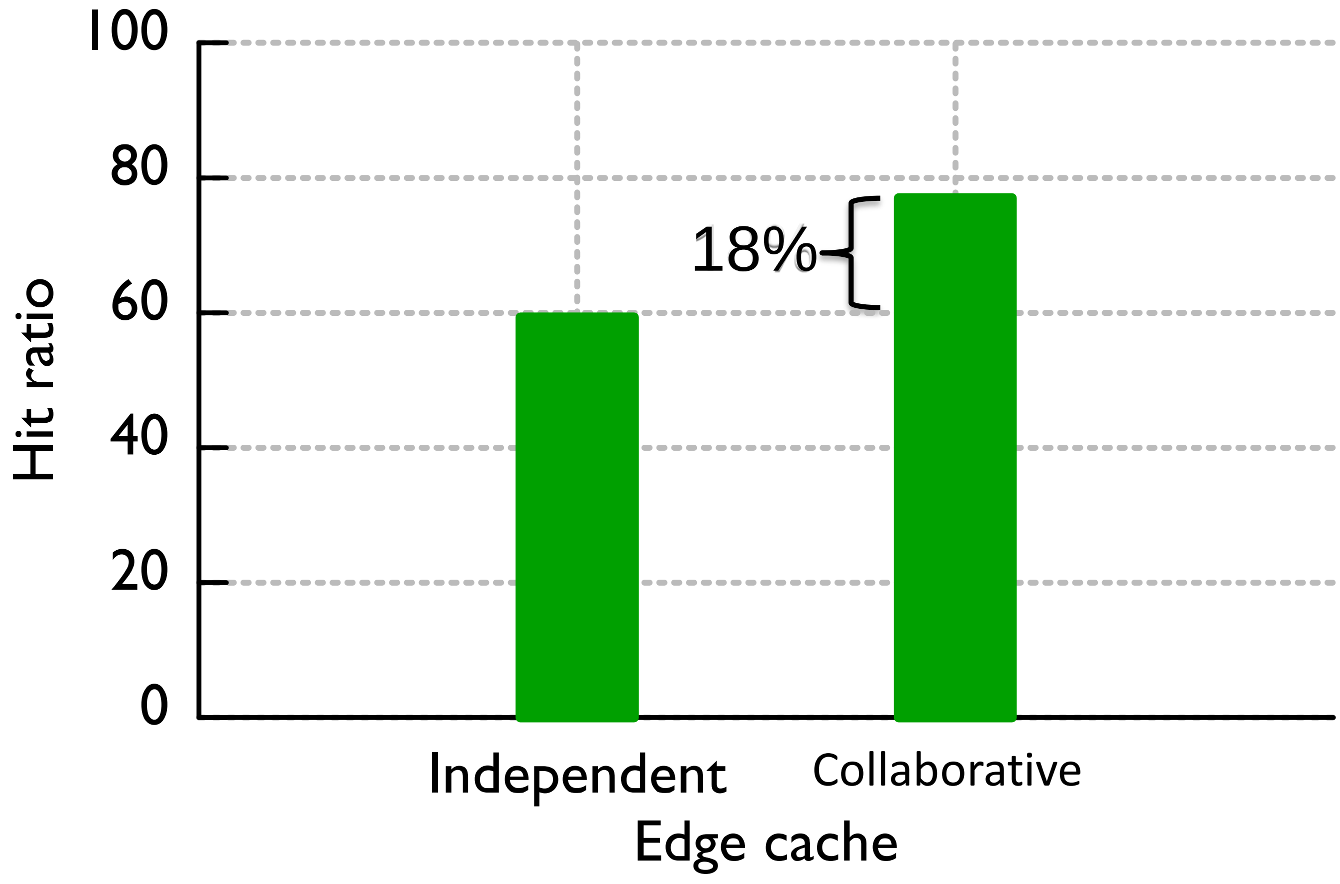


Collaborative Edge



- “Independent” aggregates all high-volume Edges

Collaborative Edge



- "Collaborative" Edge increases hit ratio by 18%

Related Work

Storage Analysis

BSD file system (SOSP ’85), Sprite (SOSP ’91), NT (SOSP ’99), NetApp (SOSP ’11), iBench (SOSP ’11)

Content Distribution Analysis

Cooperative caching (SOSP ’99), CDN vs. P2P (OSDI ’02), P2P (SOSP ’03), CoralCDN (NSDI ’10), Flash crowds (IMC ’11)

Web Analysis

Zipfian (INFOCOM ’00), Flash crowds (WWW ’02), Modern web traffic (IMC ’11)

Conclusion

- Quantify caching performance
- Quantify popularity changes across layers of caches
- Recency, frequency, age, social factors impact cache
- Outline potential gain of collaborative caching