Facebook's photo storage

Tomasz Wysocki

Based on paper by Doug Beaver, Sanjeev Kumar, Harry C. Li, Jason Sobel, Peter Vajgel "Finding a needle in Haystack: Facebook's photo storage".

Problem is to

- store lot of images (260 bilion)
- be ready for new ones (1 bilion / week)
- serve them (up to 1 milion / second)

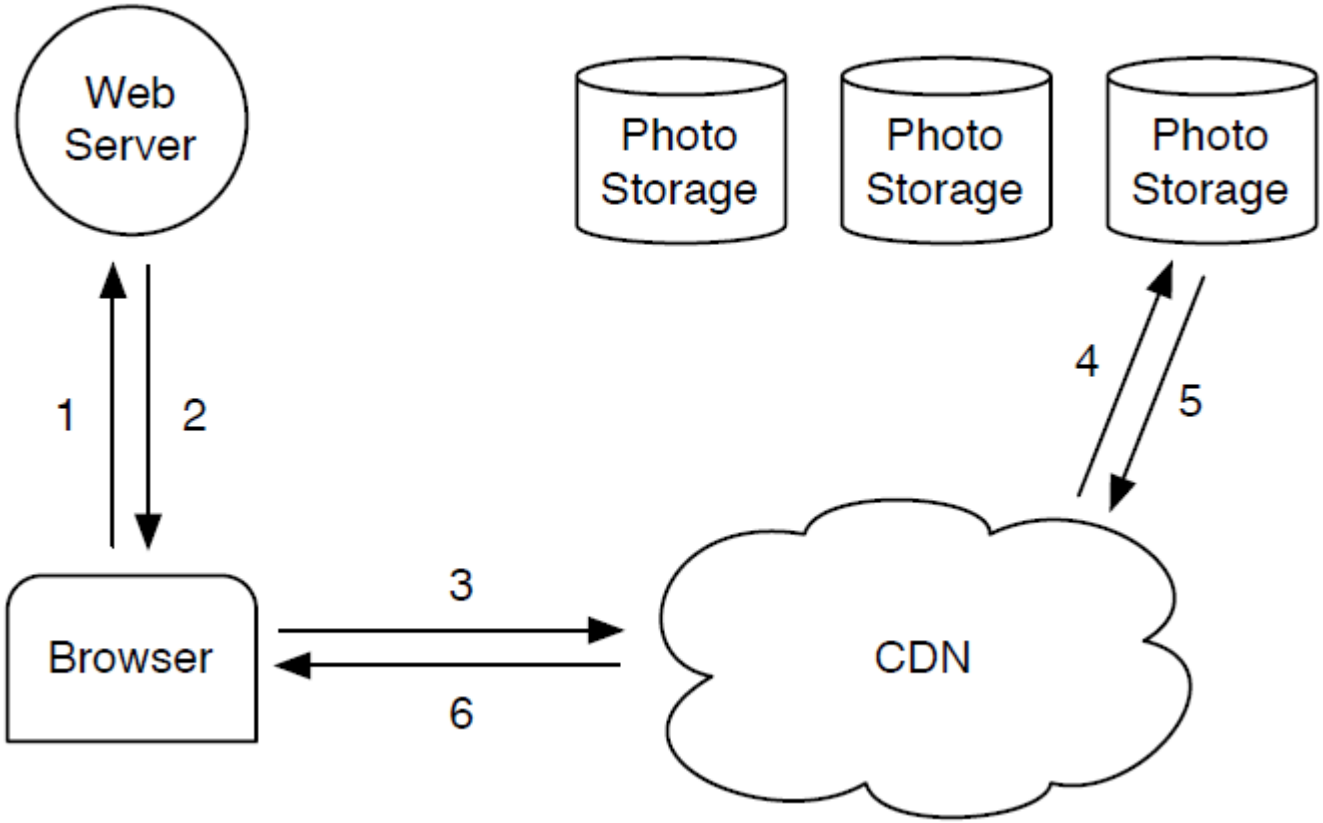
Solution must be

- high throughput
- have low latency
- fault-tolerant
- cost-effective

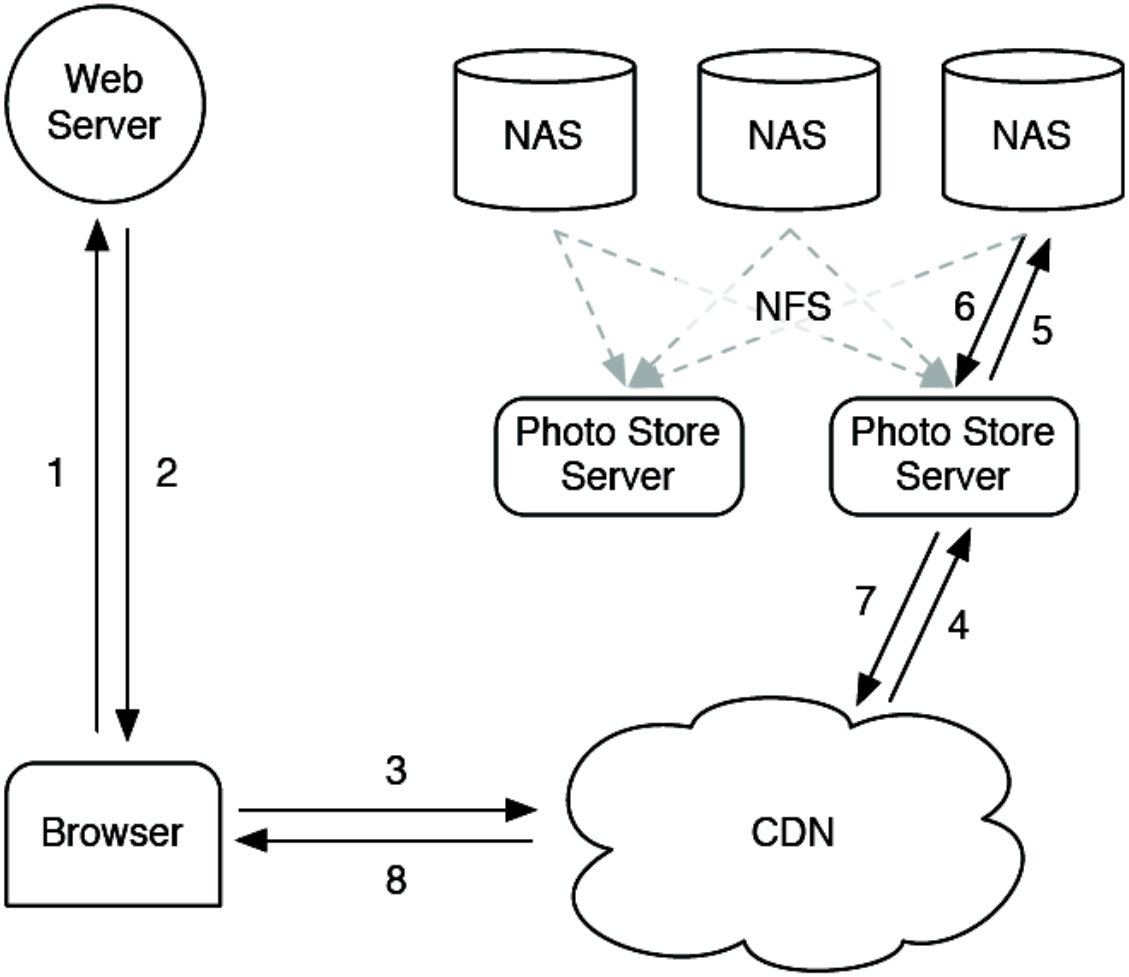
Constraints

	Normal hard drive	**Facebook photos**
Create	often	very often / never
Delete	often	rarely
Read	very often	very often
Modify	often	never*

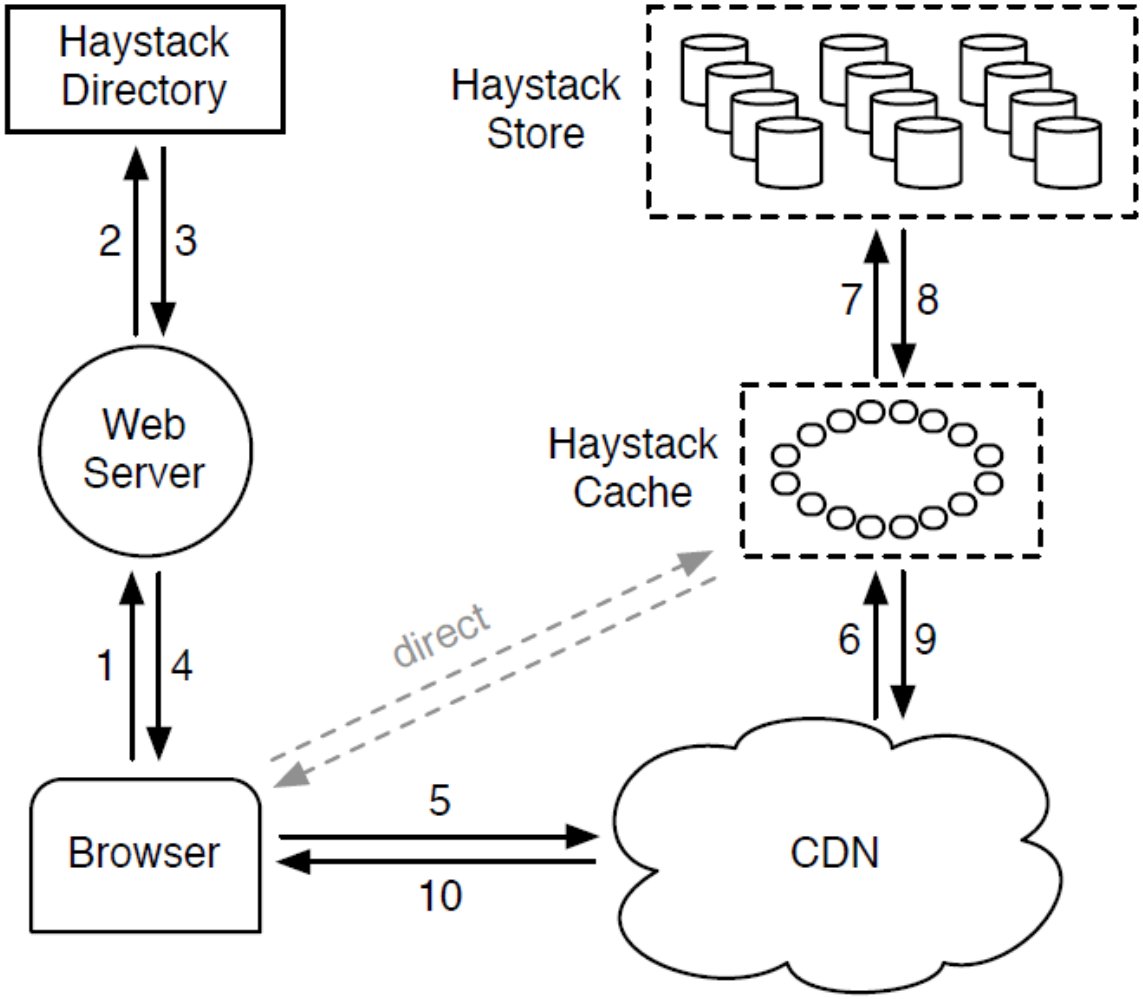
Usual design



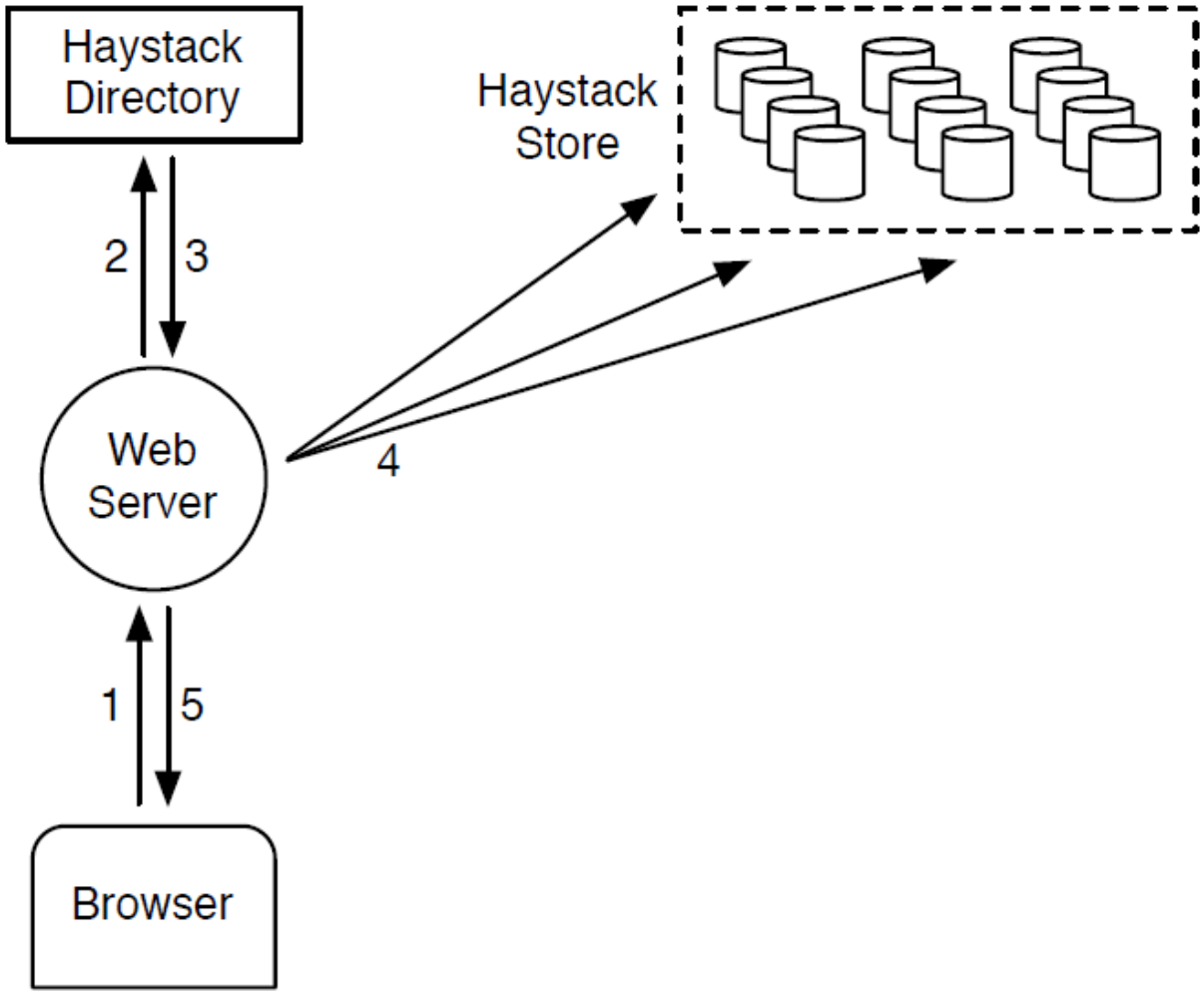
Facebook NFS approach



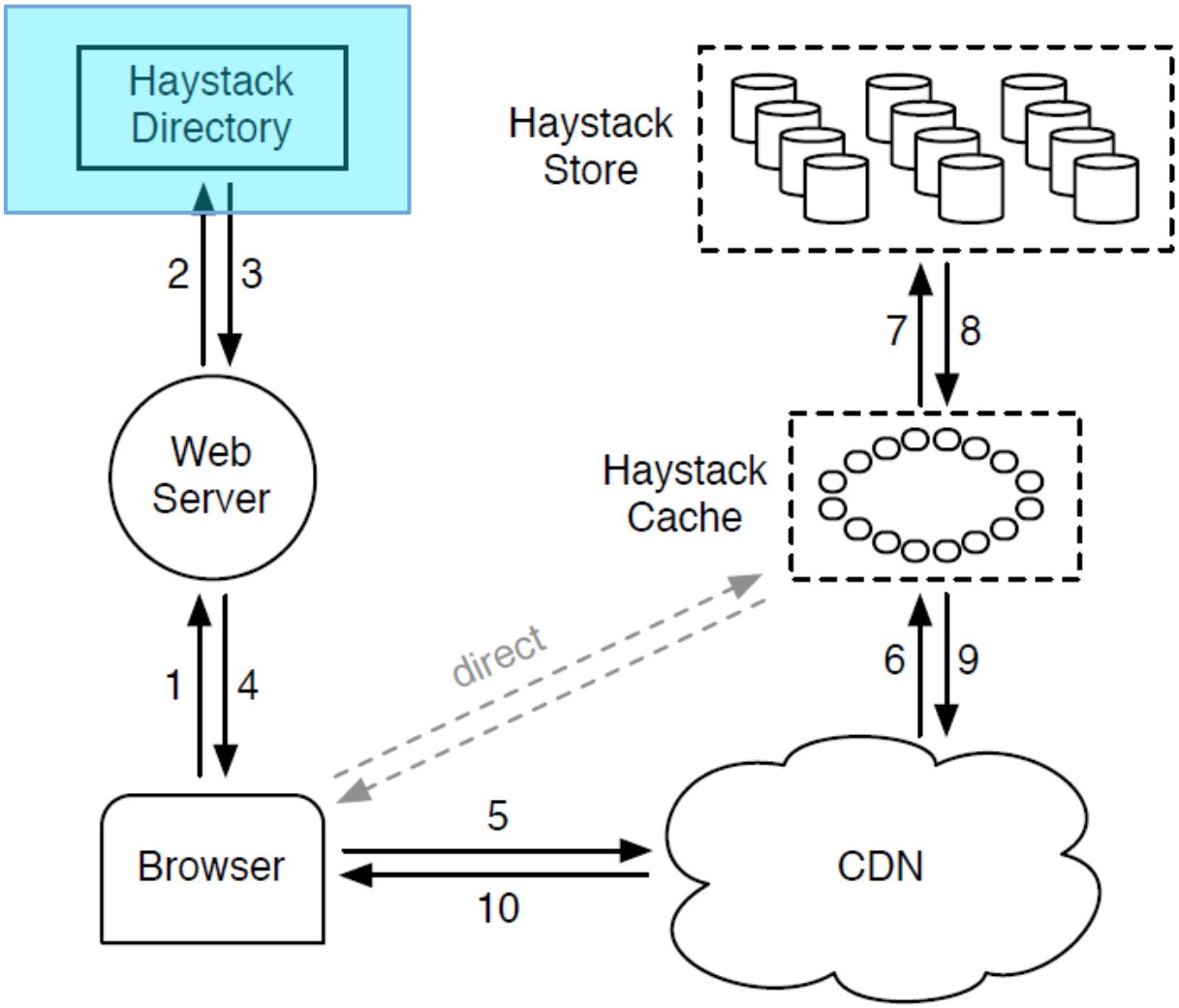
Haystack


Haystack Directory
Haystack Store
2
3
7
8
Web Server
Haystack Cache
direct
1
4
6
9
Browser
5
10
CDN

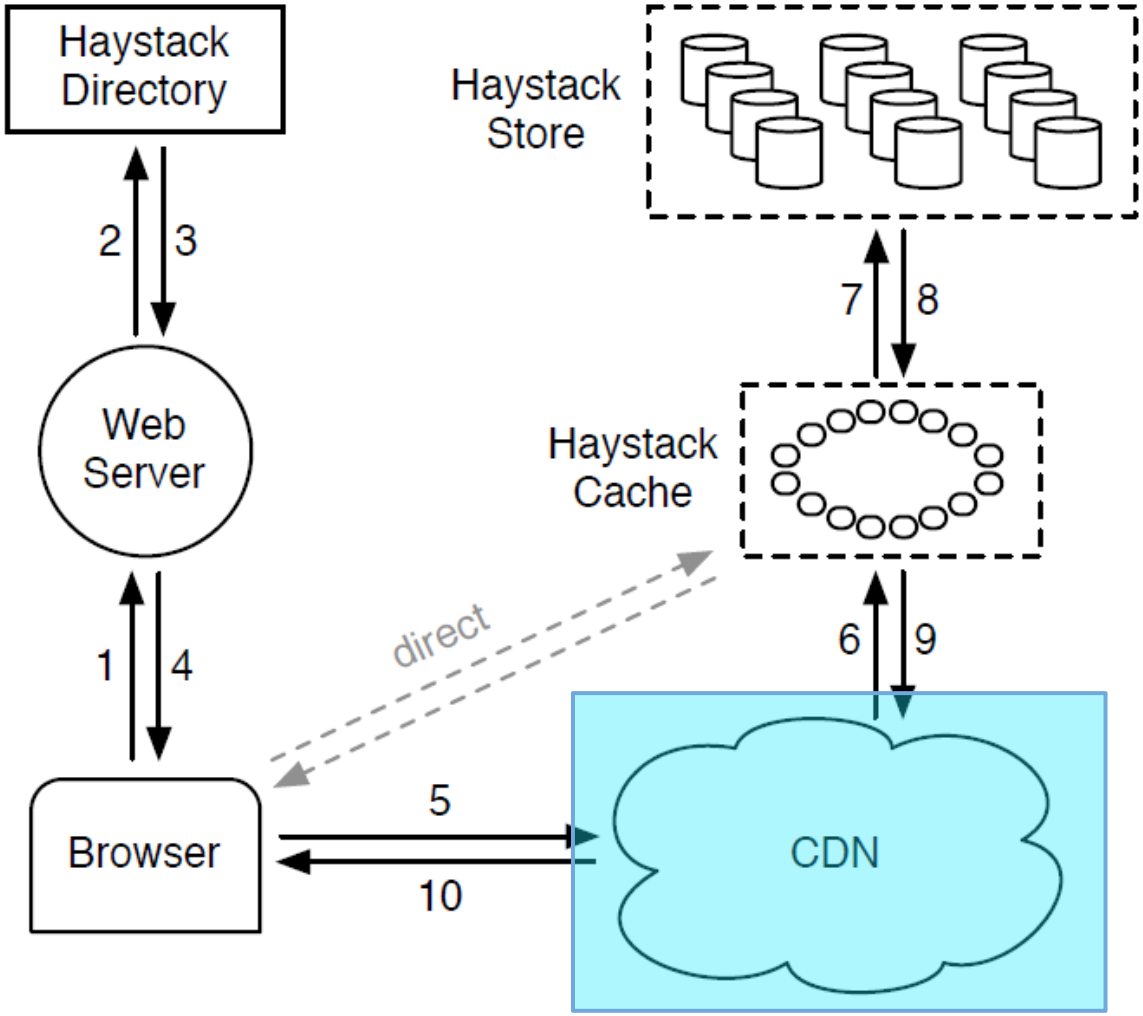
Haystack - uploading



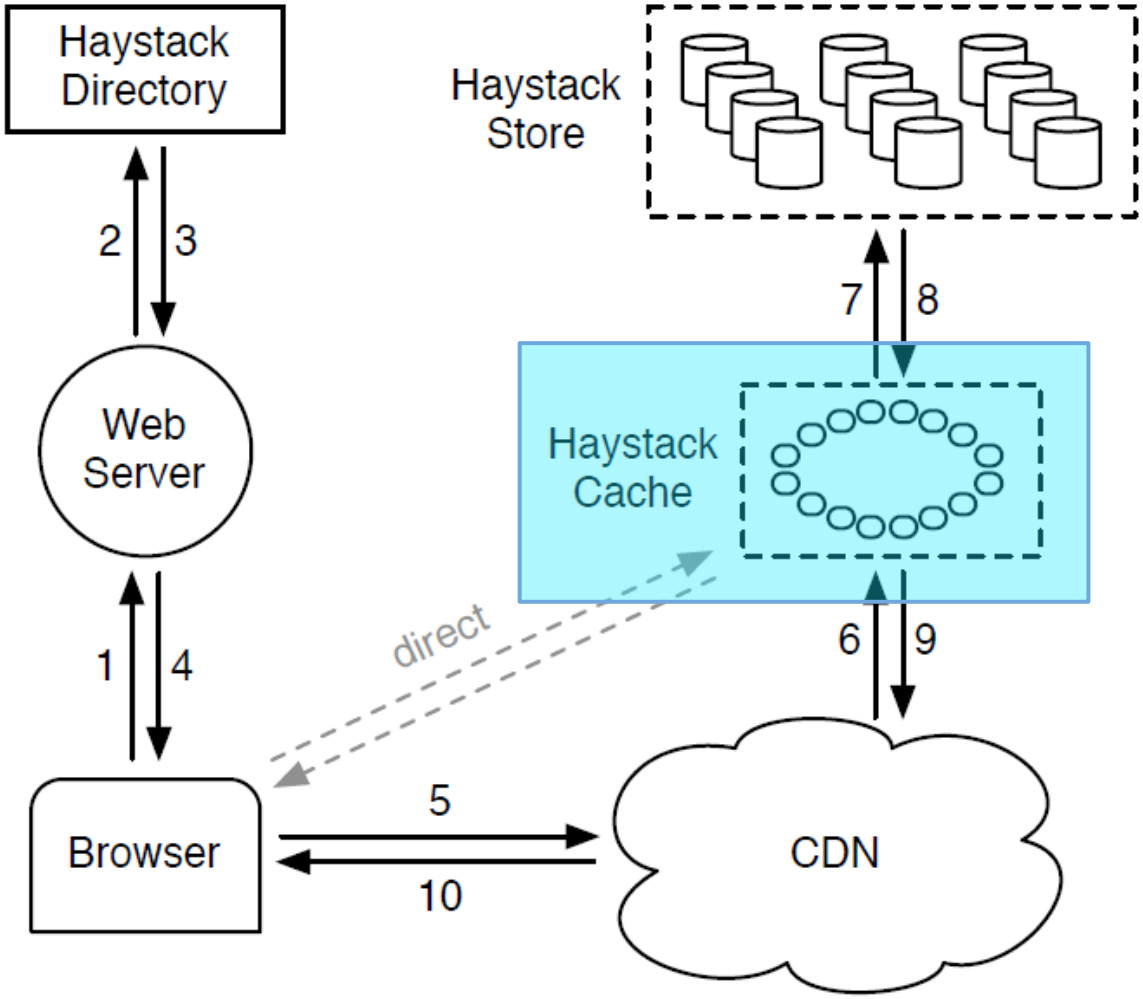
Haystack Directory



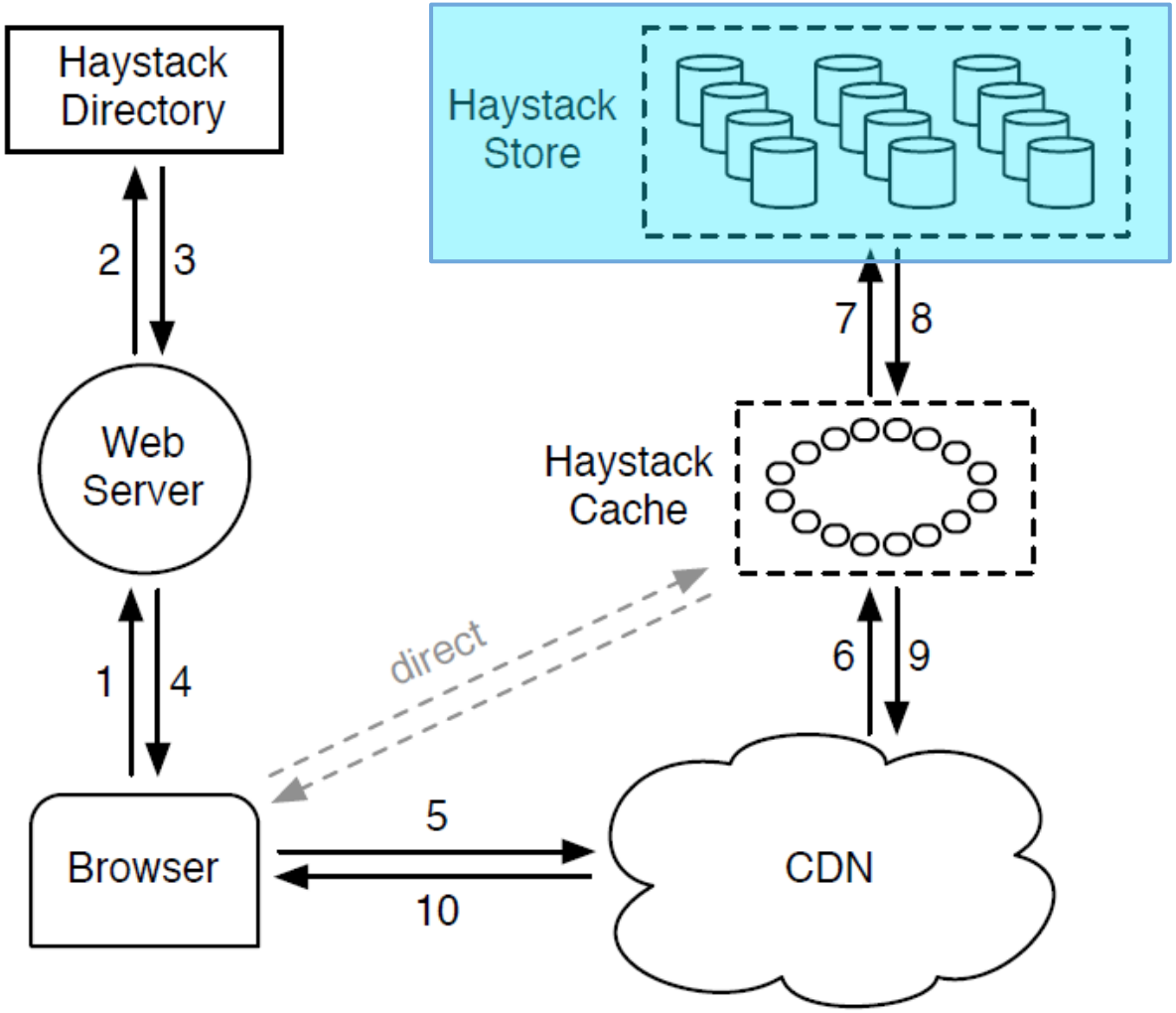
CDN


Haystack Directory
Haystack Store
2
3
7
8
Web Server
Haystack Cache
direct
1
4
6
9
5
Browser
CDN
10

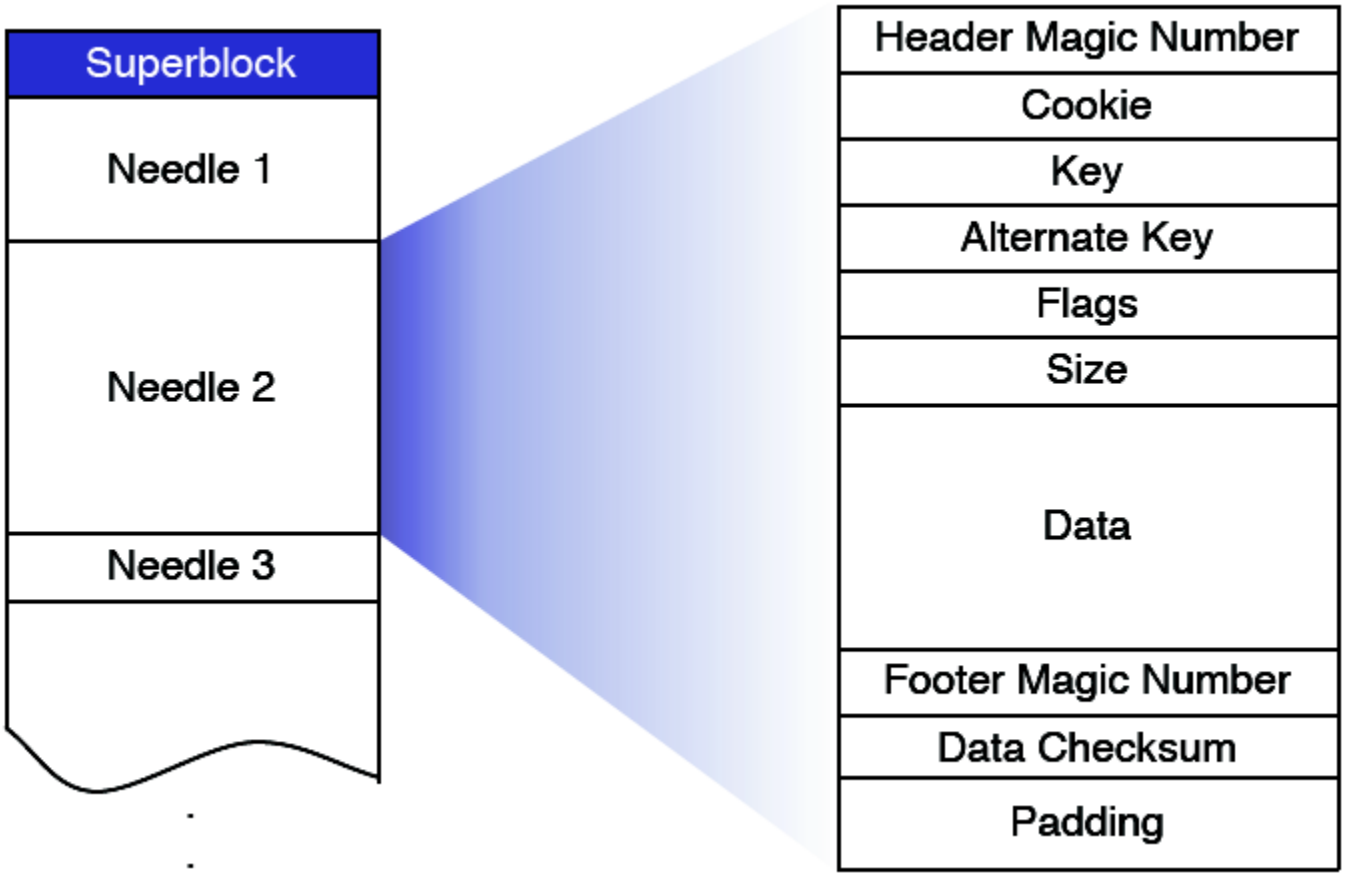
Haystack Cache



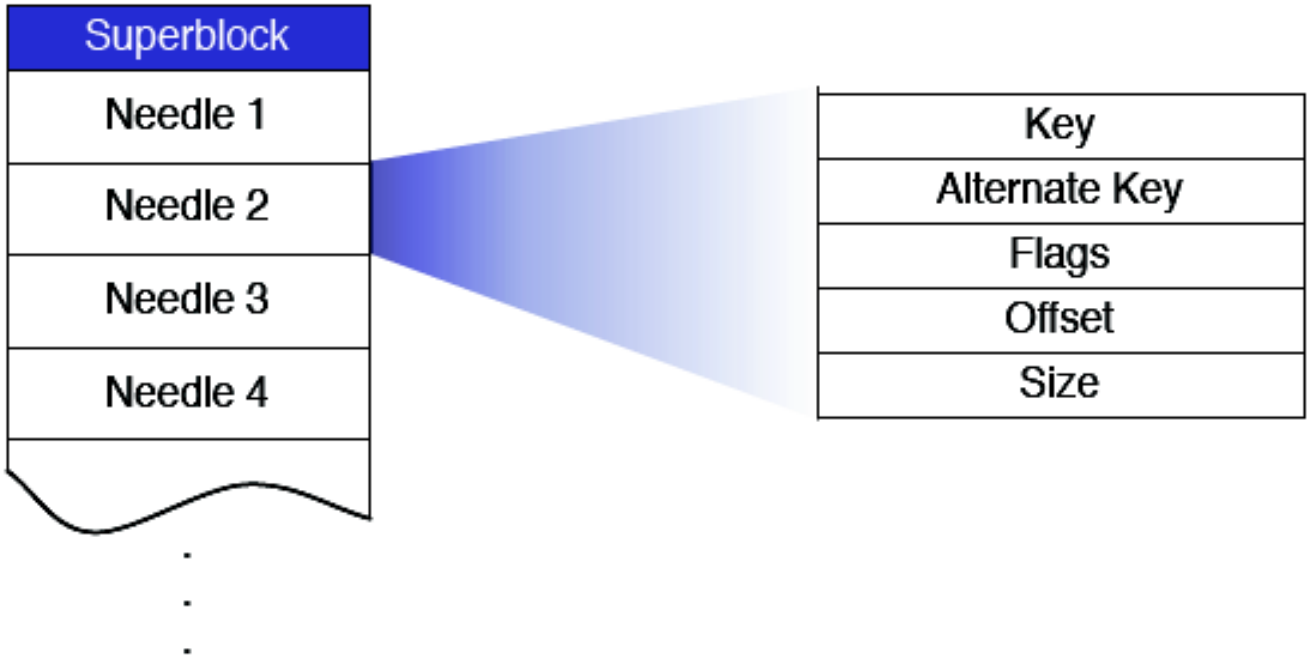
Haystack Store



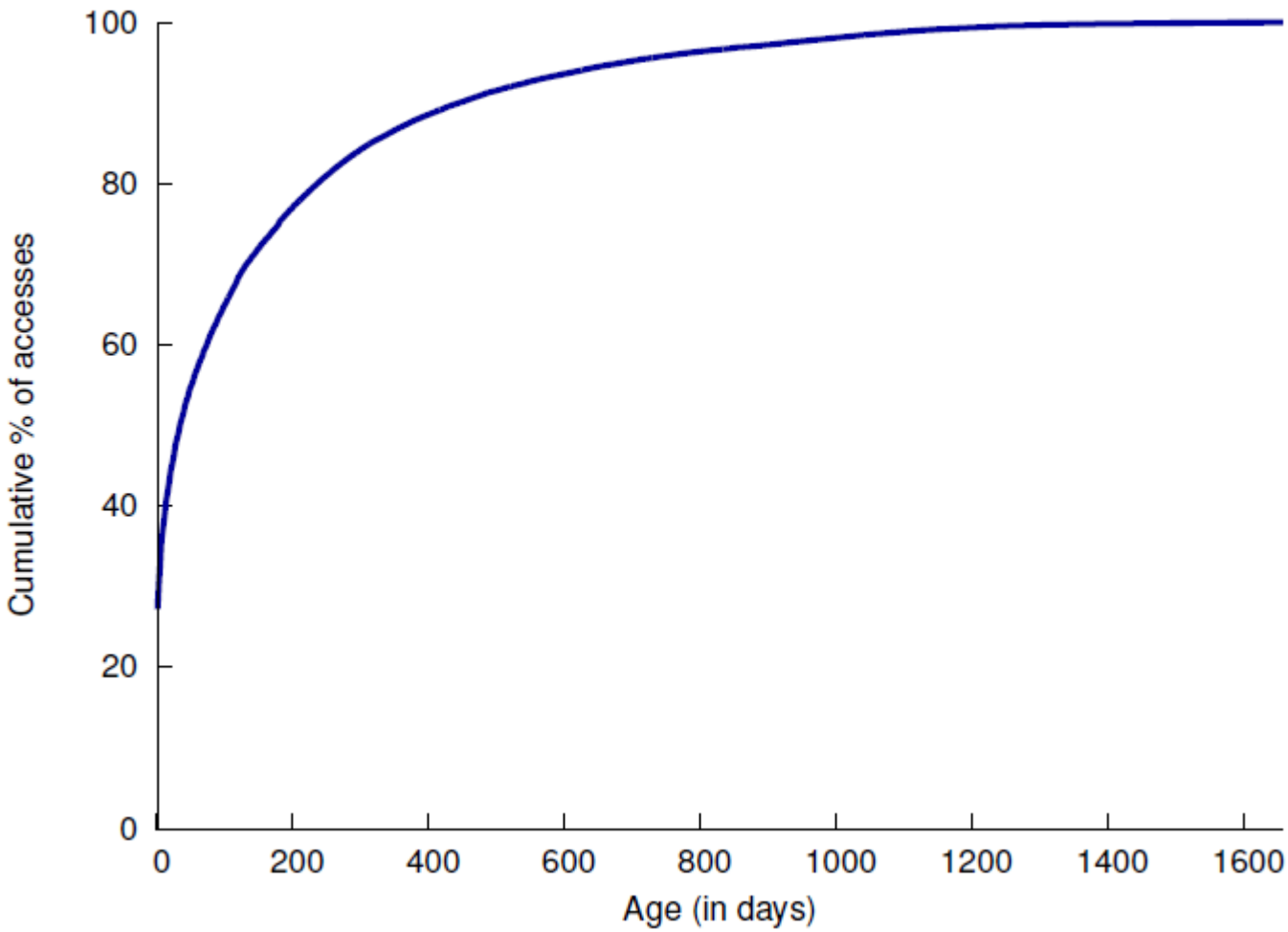
Haystack Store file



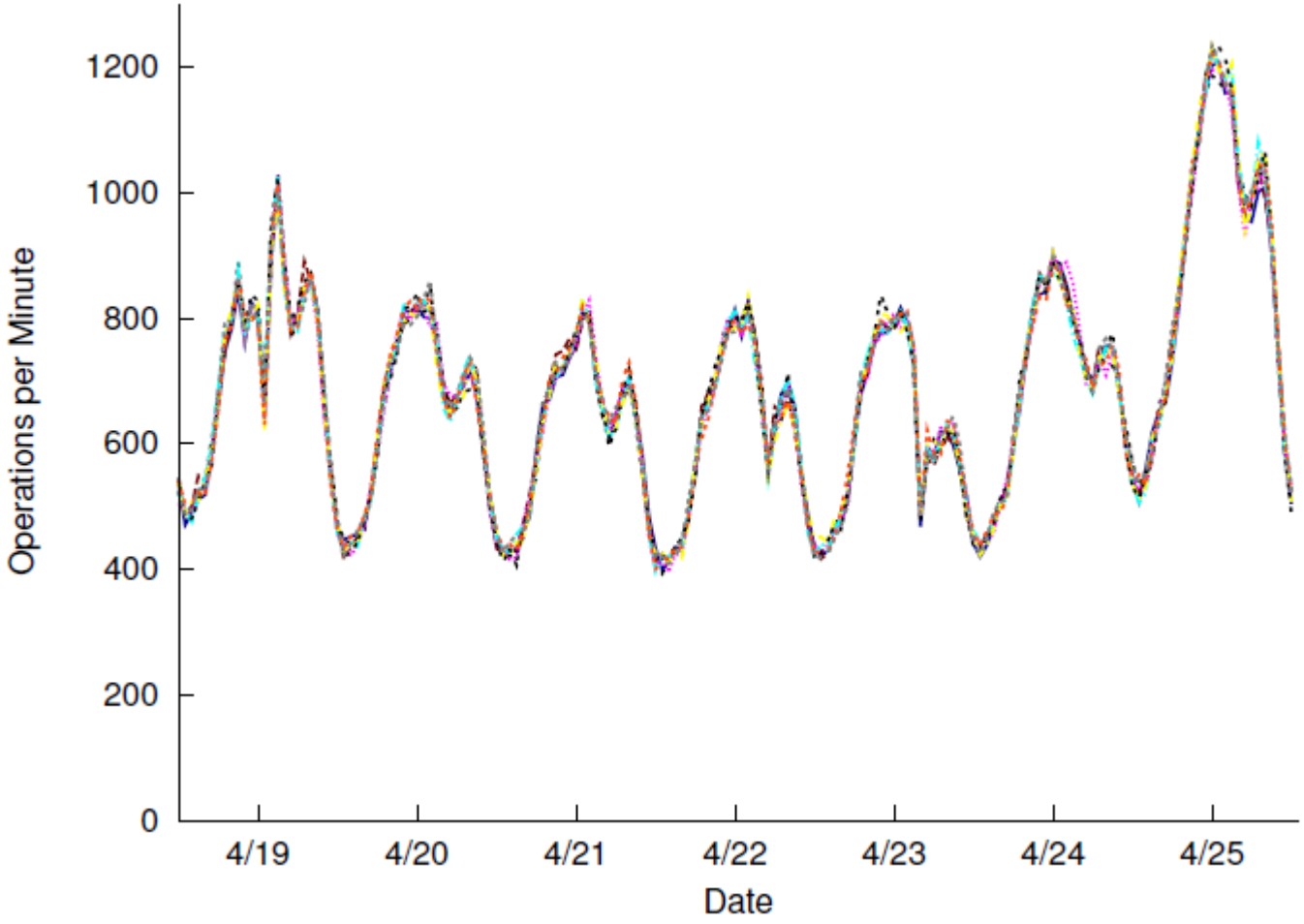
Haystack Index



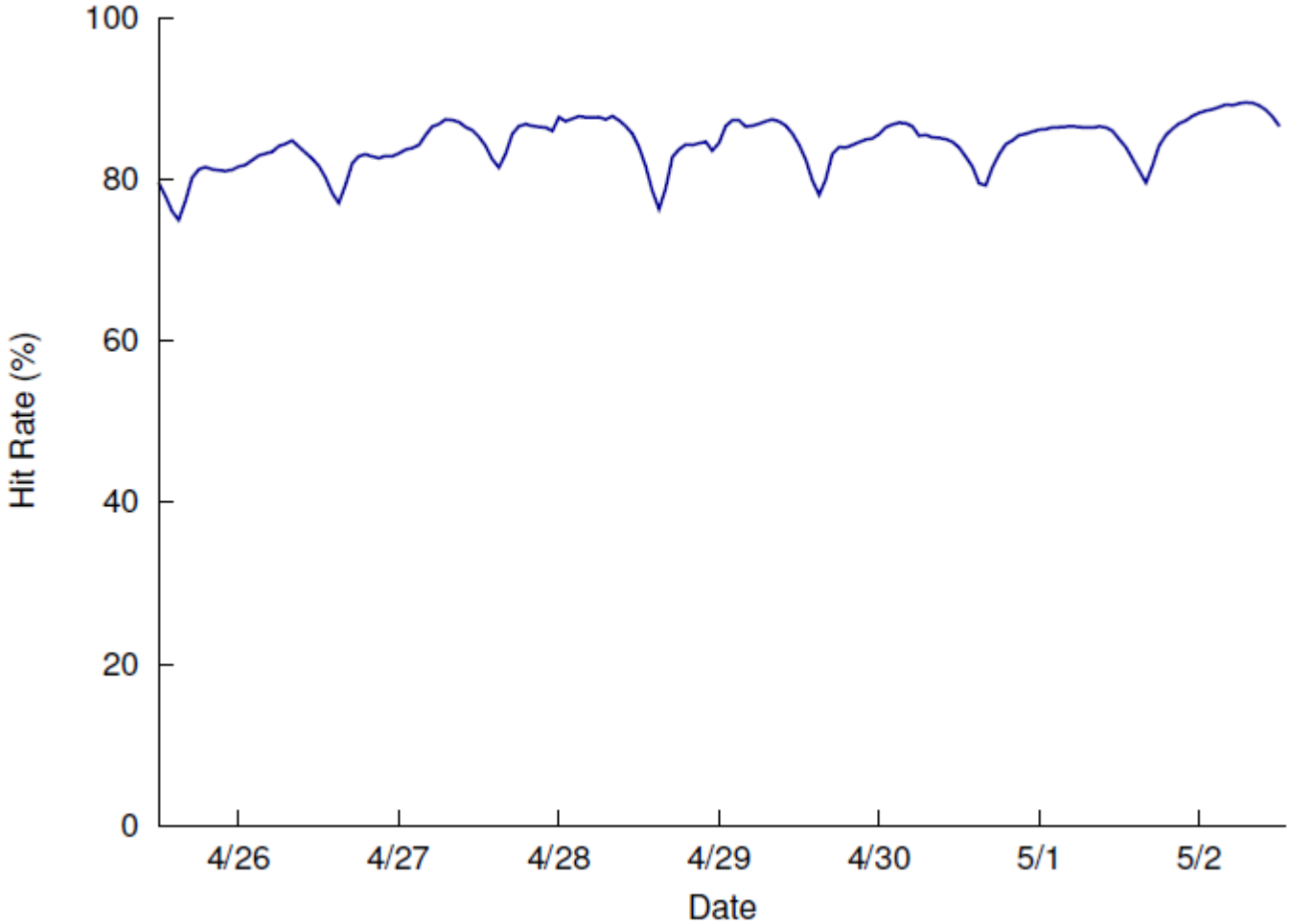
When file is accessed



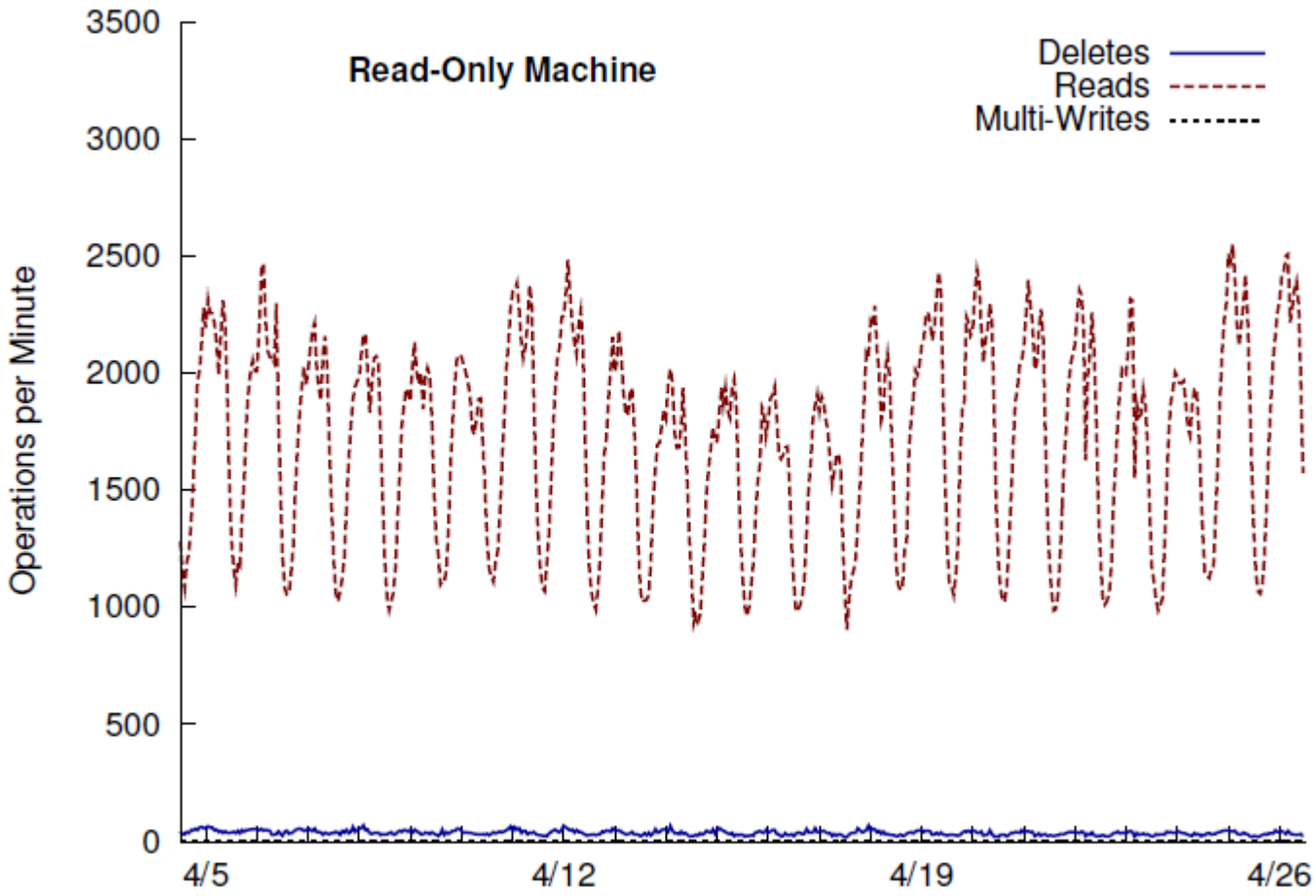
How load is distributed (write)



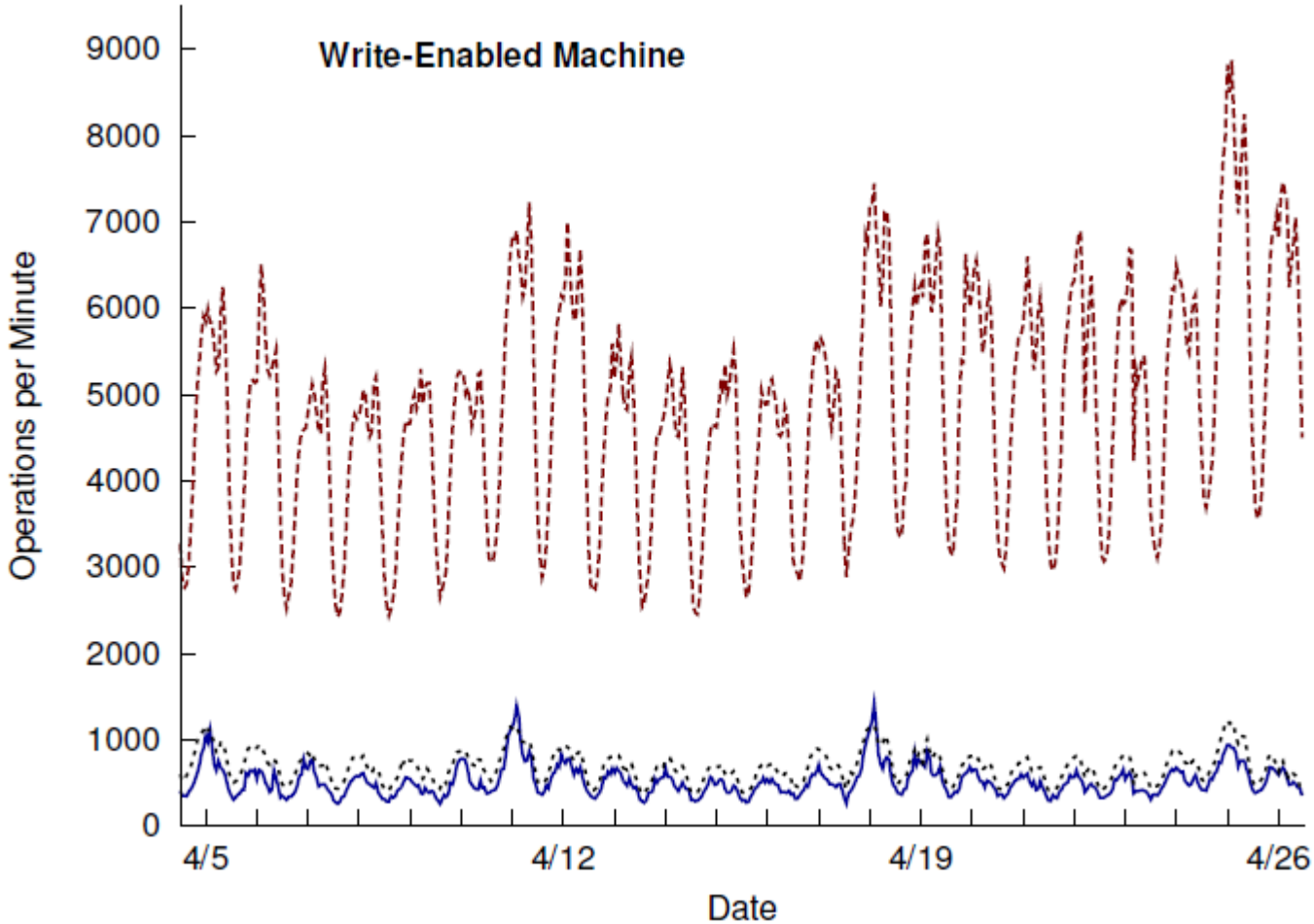
Cache hit rate



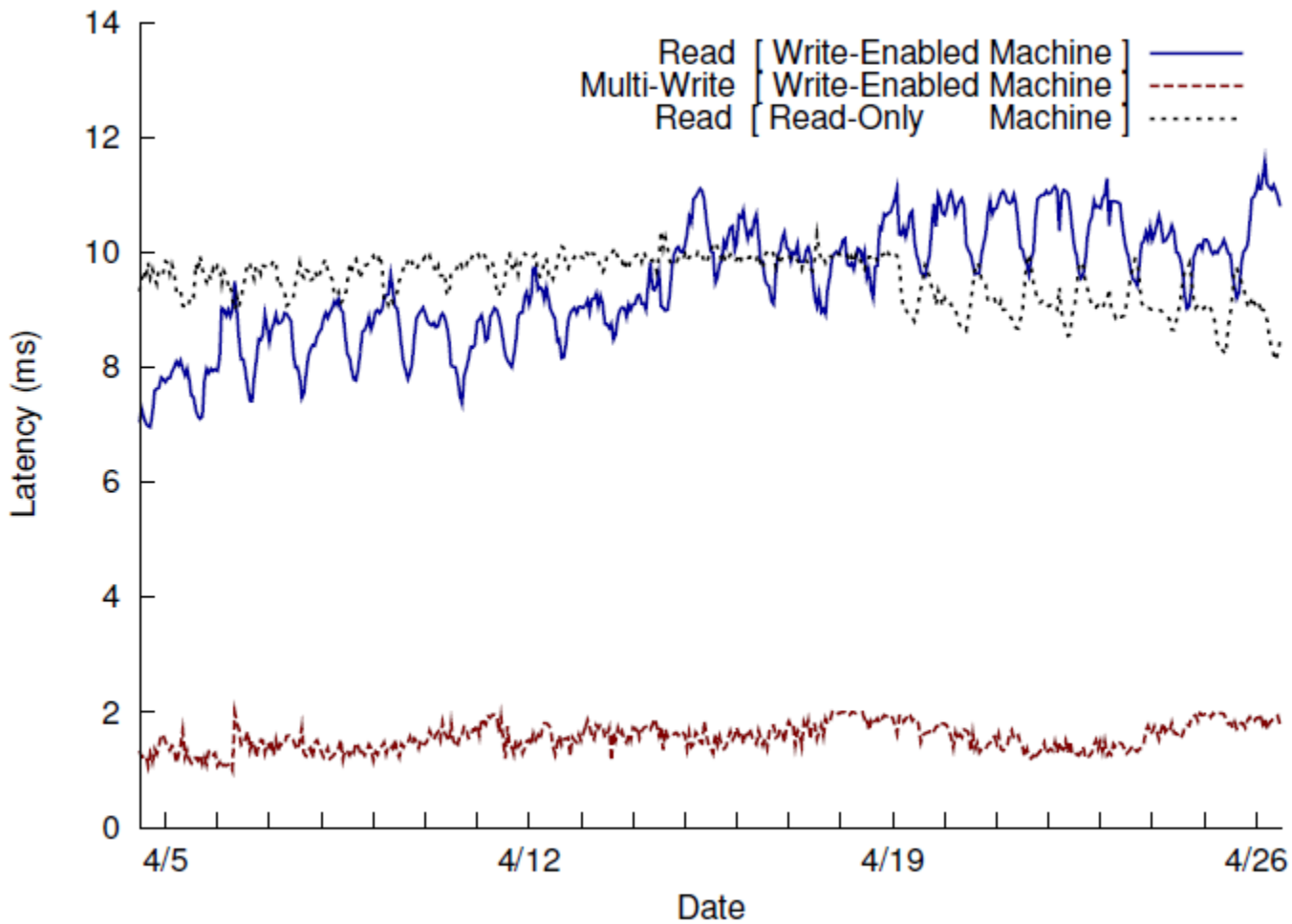
Read-Only Machine



Write-Enabled Machine



Latency


Read [Write-Enabled Machine]
Multi-Write [Write-Enabled Machine]
Read [Read-Only Machine]
Latency (ms)
0
2
4
6
8
10
12
14
4/5
4/12
4/19
4/26
Date

One more thing