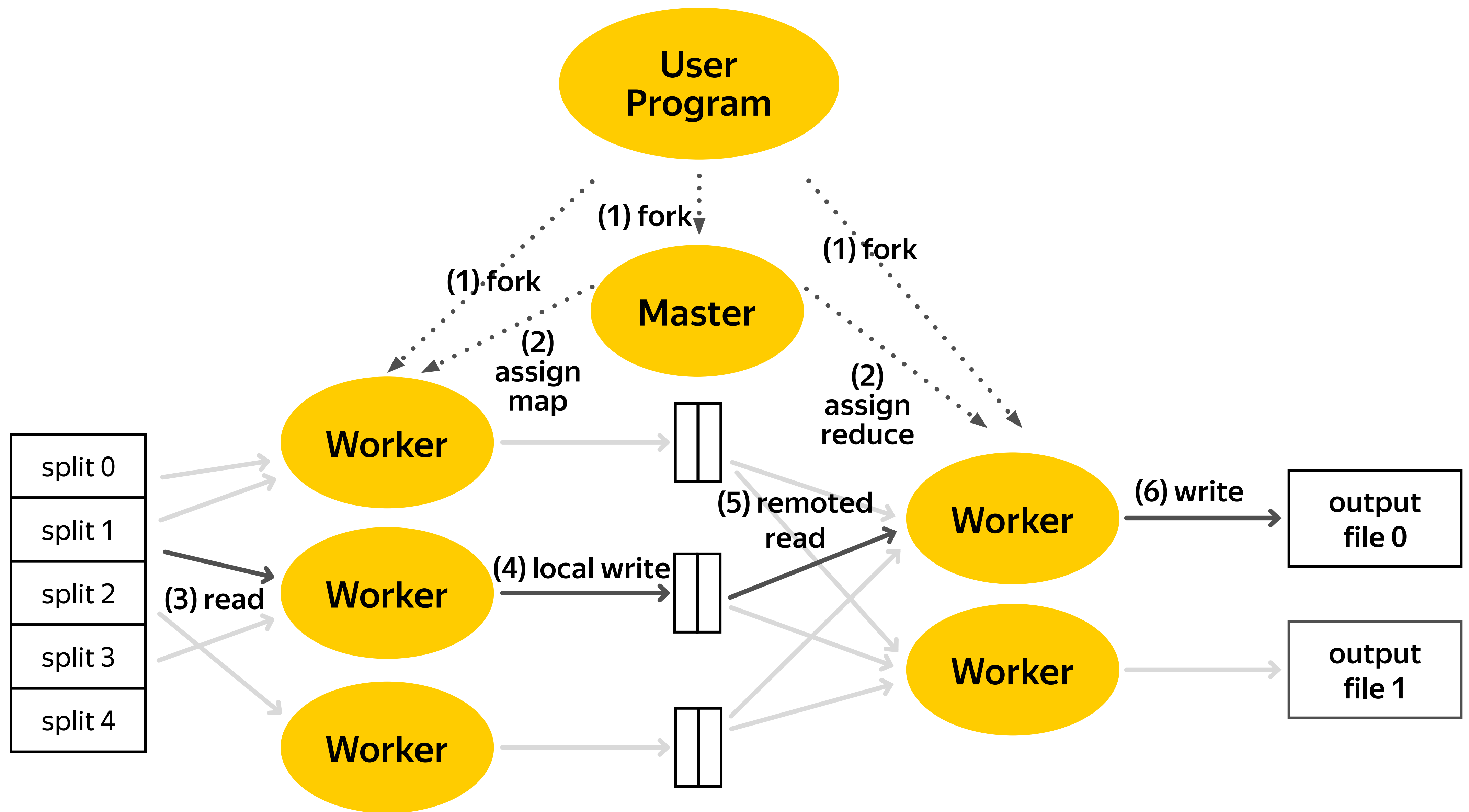


Yandex

MapReduce

Fault Tolerance



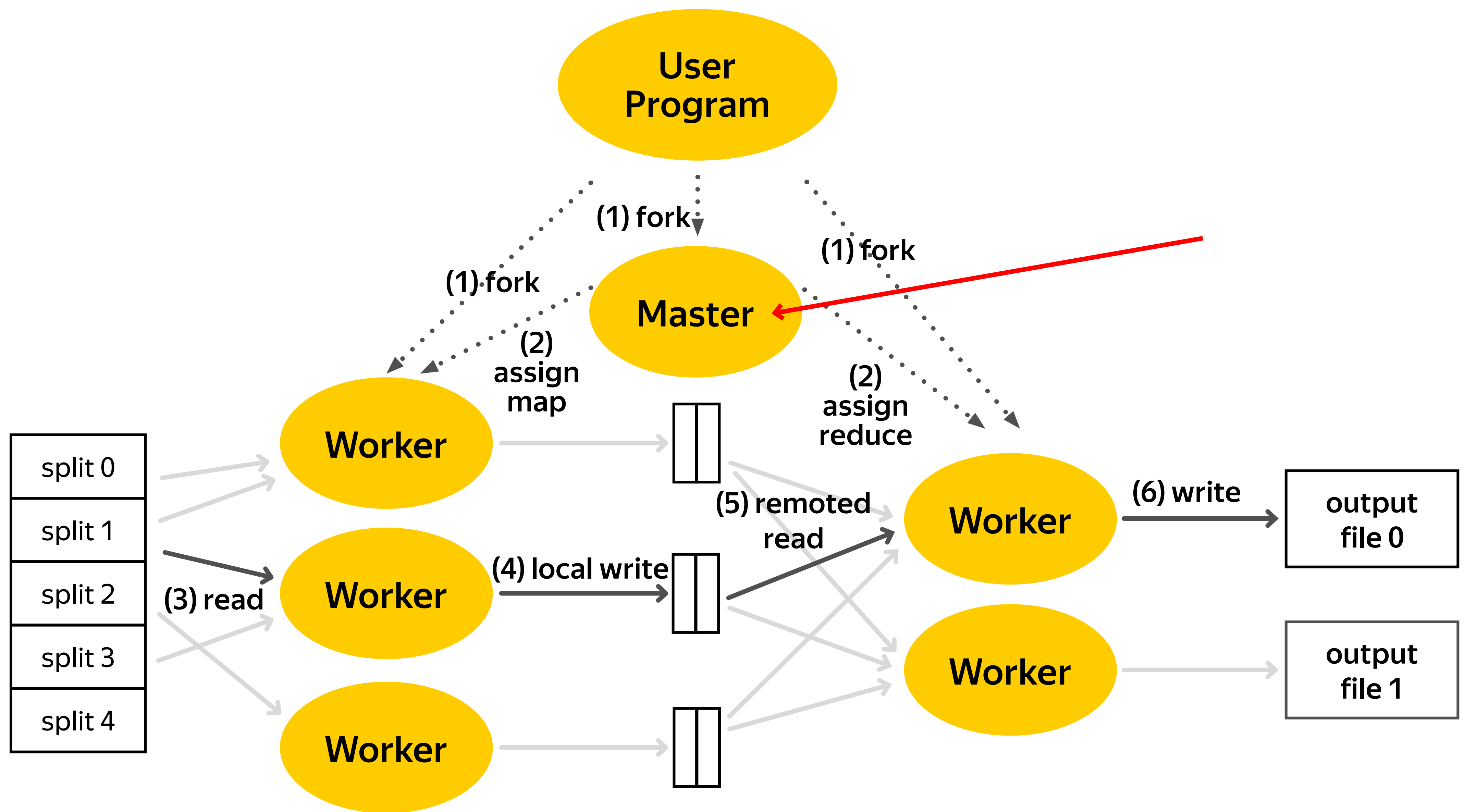
Input
files

Map
phase

Intermediate files
(on local disks)

Reduce
phase

Output
files



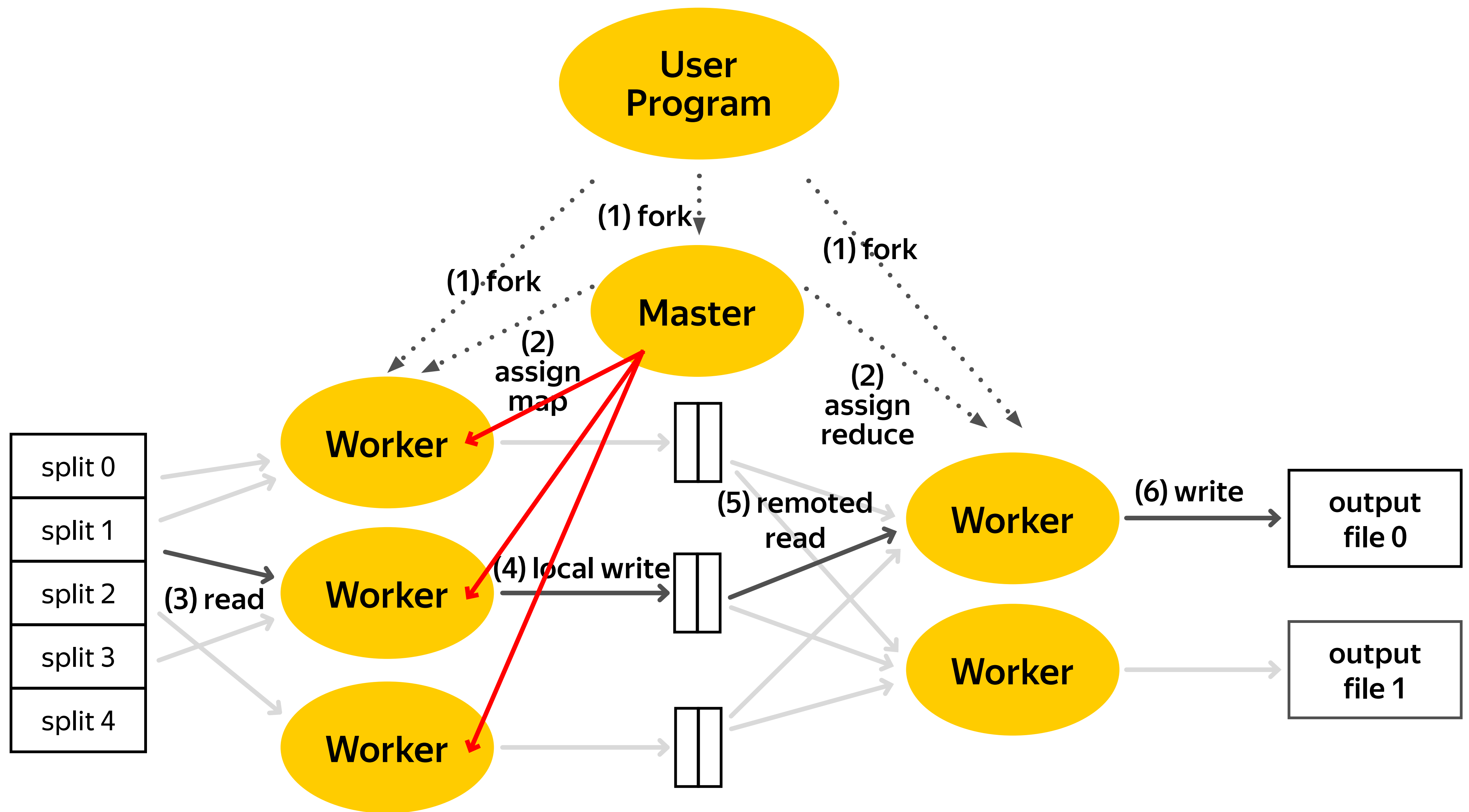
Input
files

Map
phase

Intermediate files
(on local disks)

Reduce
phase

Output
files



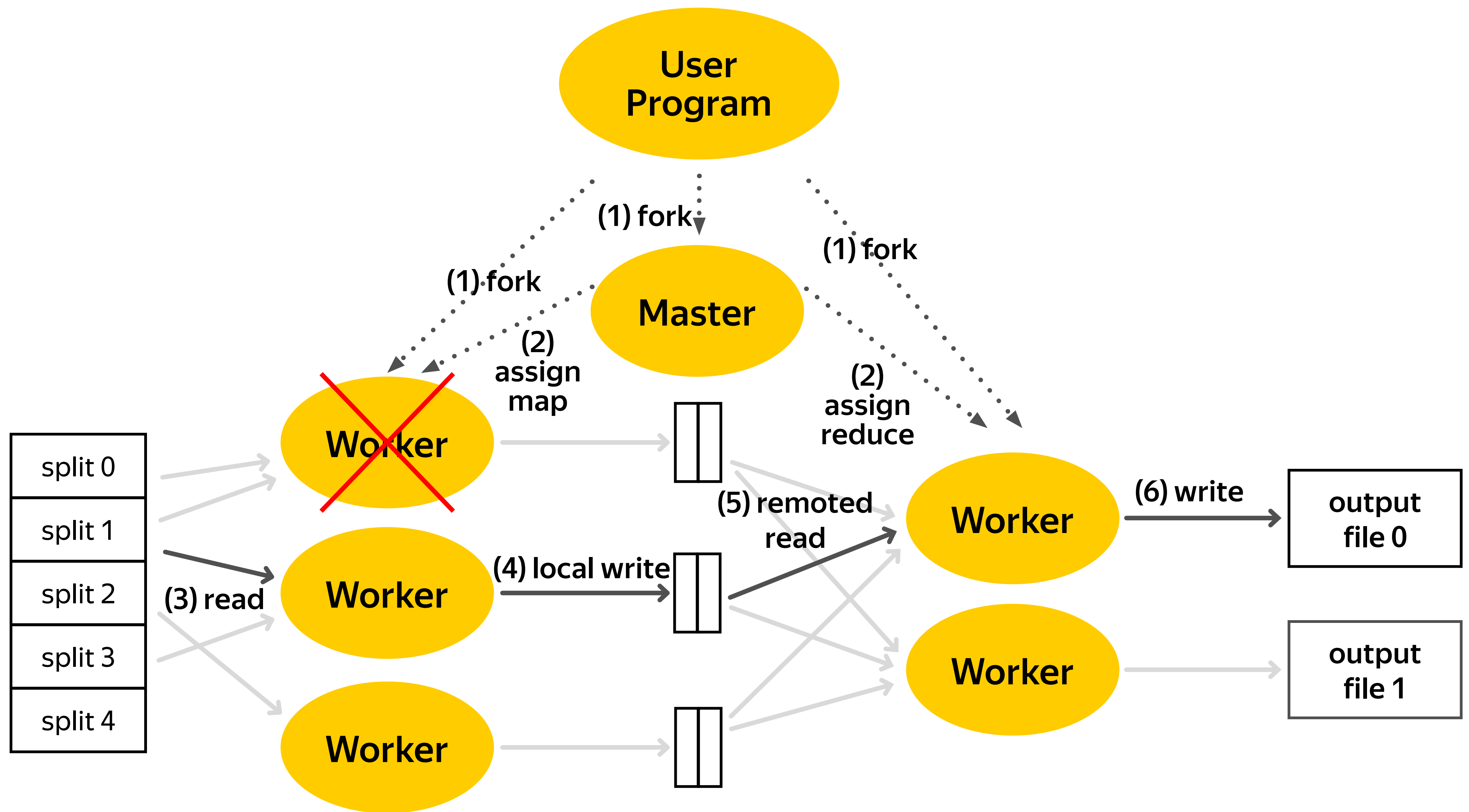
Input
files

Map
phase

Intermediate files
(on local disks)

Reduce
phase

Output
files



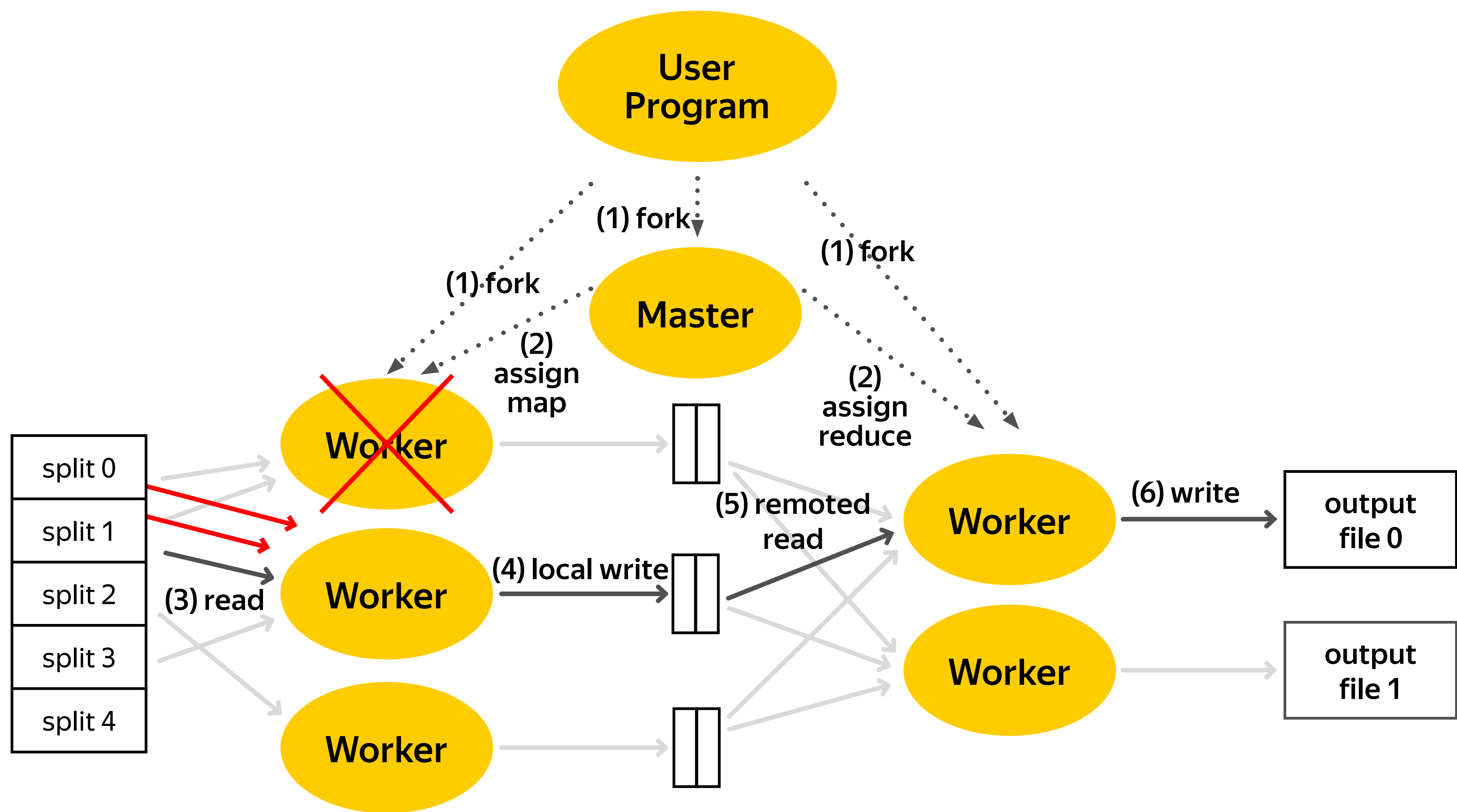
Input
files

Map
phase

Intermediate files
(on local disks)

Reduce
phase

Output
files



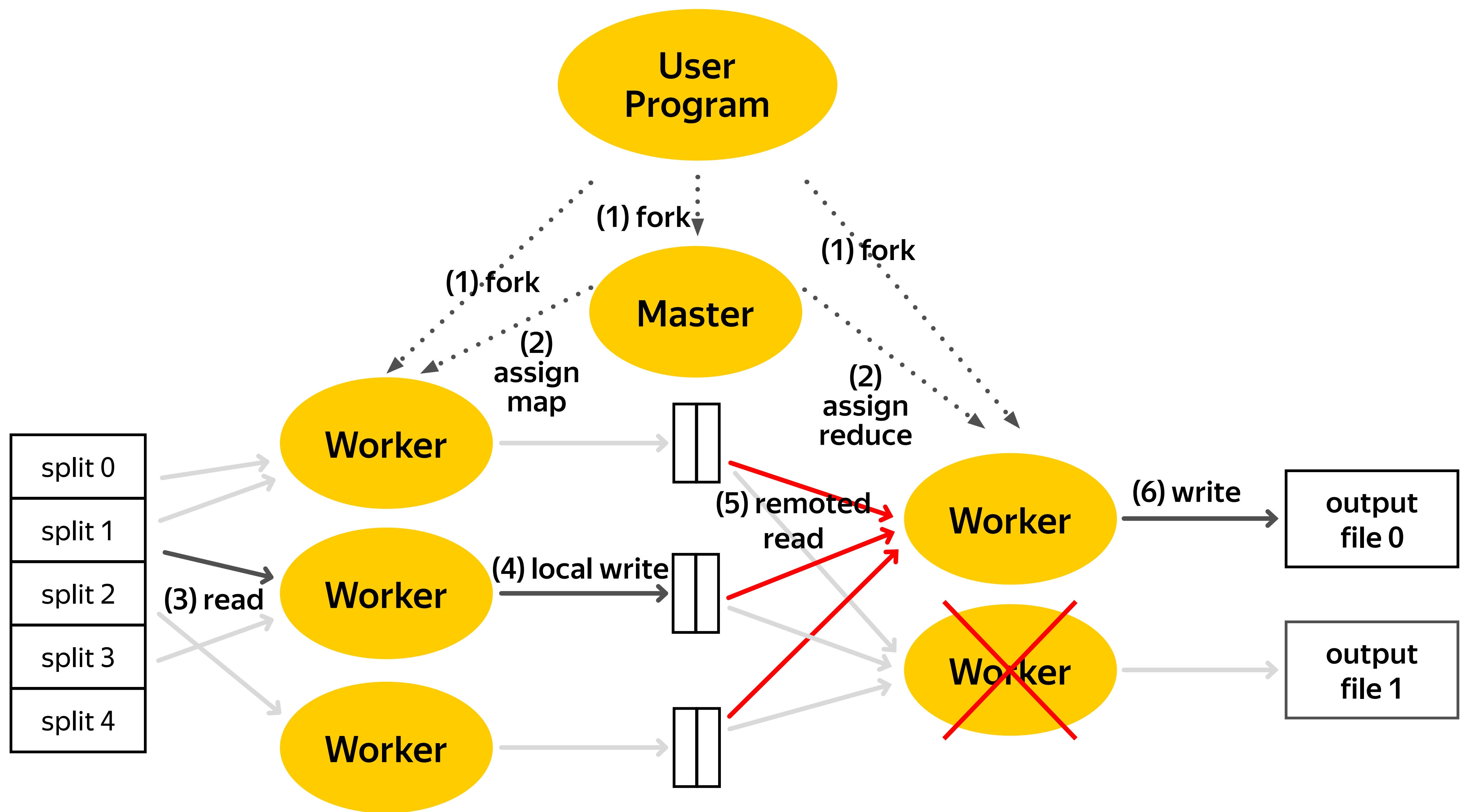
Input
files

Map
phase

Intermediate files
(on local disks)

Reduce
phase

Output
files



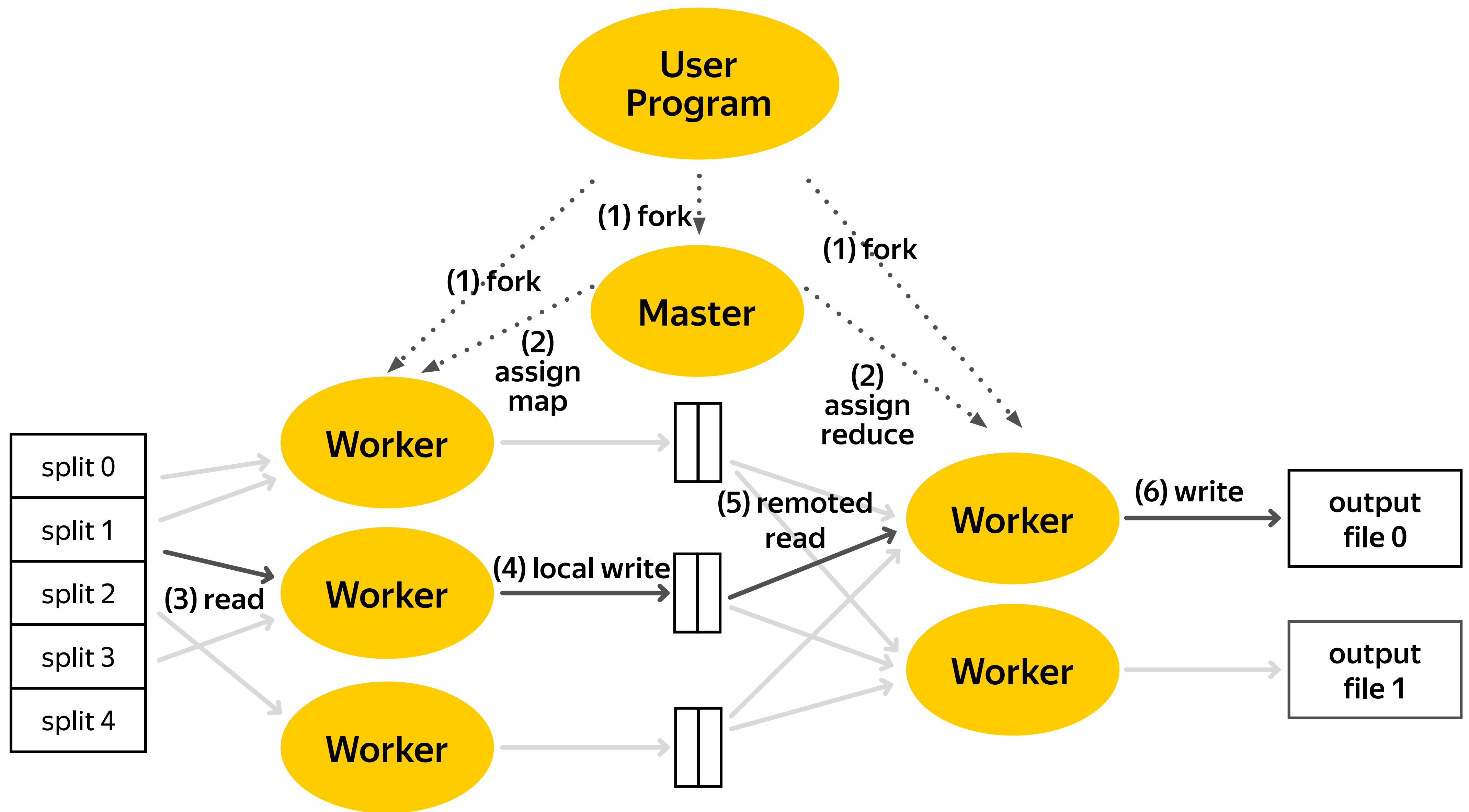
Input
files

Map
phase

Intermediate files
(on local disks)

Reduce
phase

Output
files



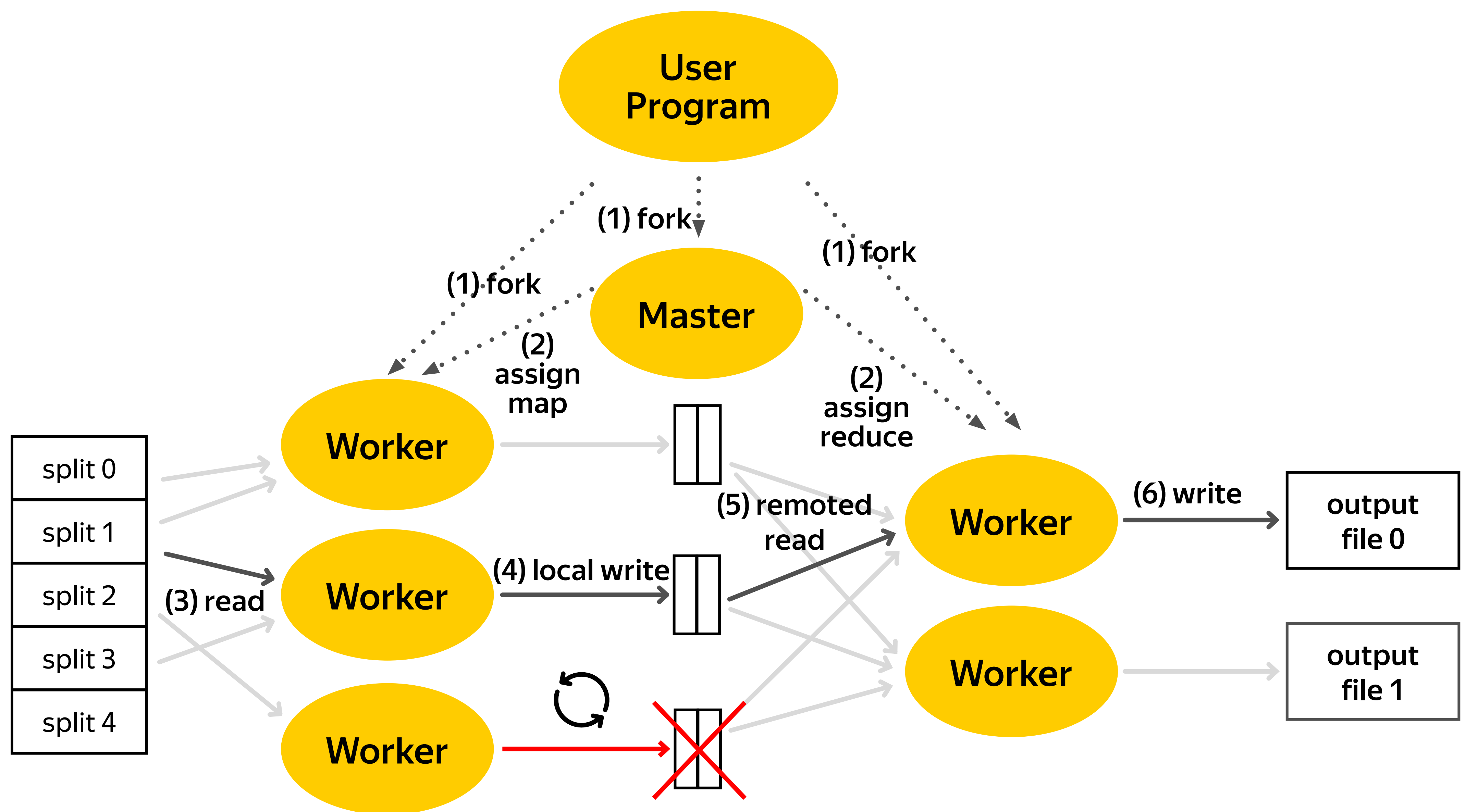
Input
files

Map
phase

Intermediate files
(on local disks)

Reduce
phase

Output
files



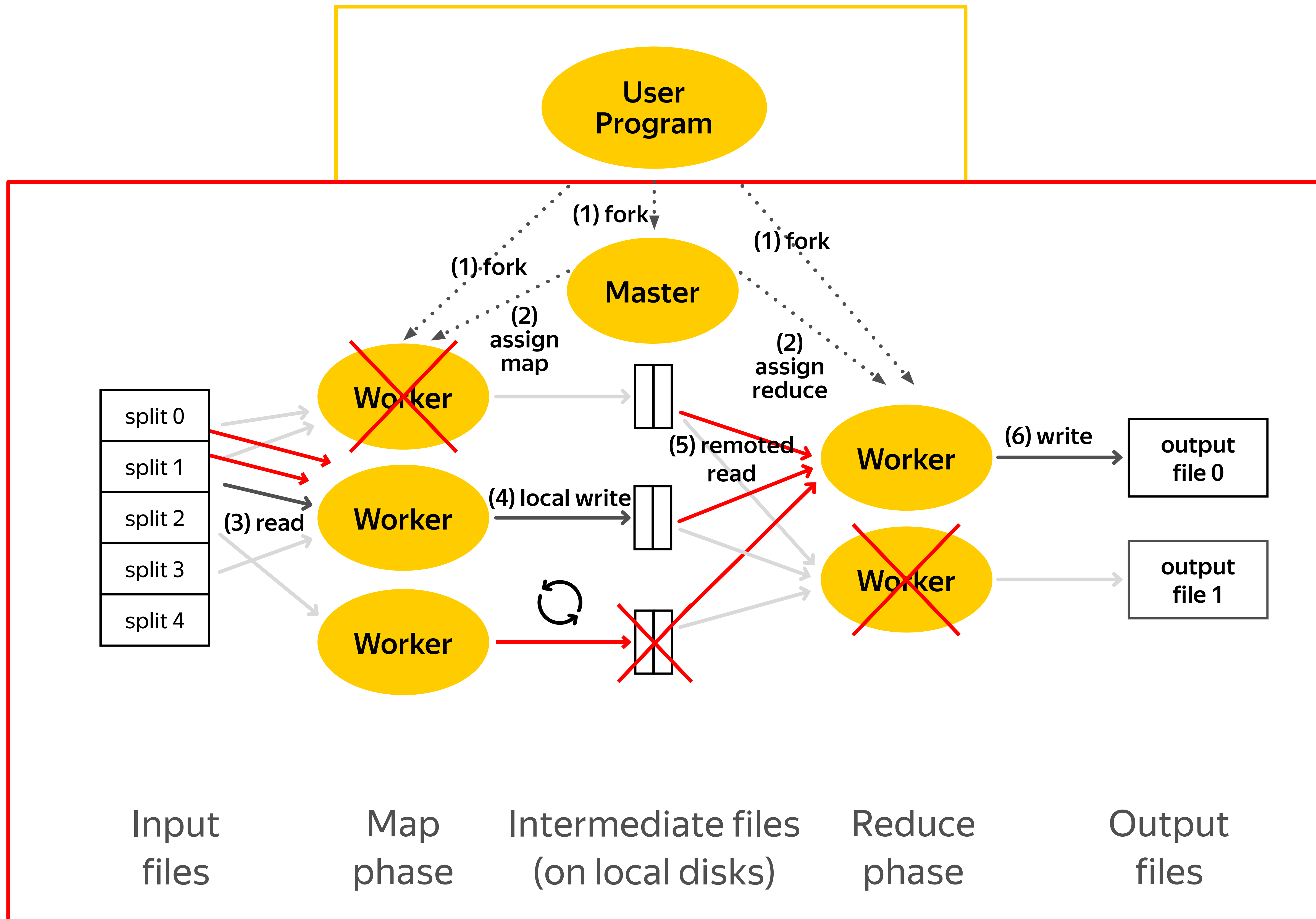
Input
files

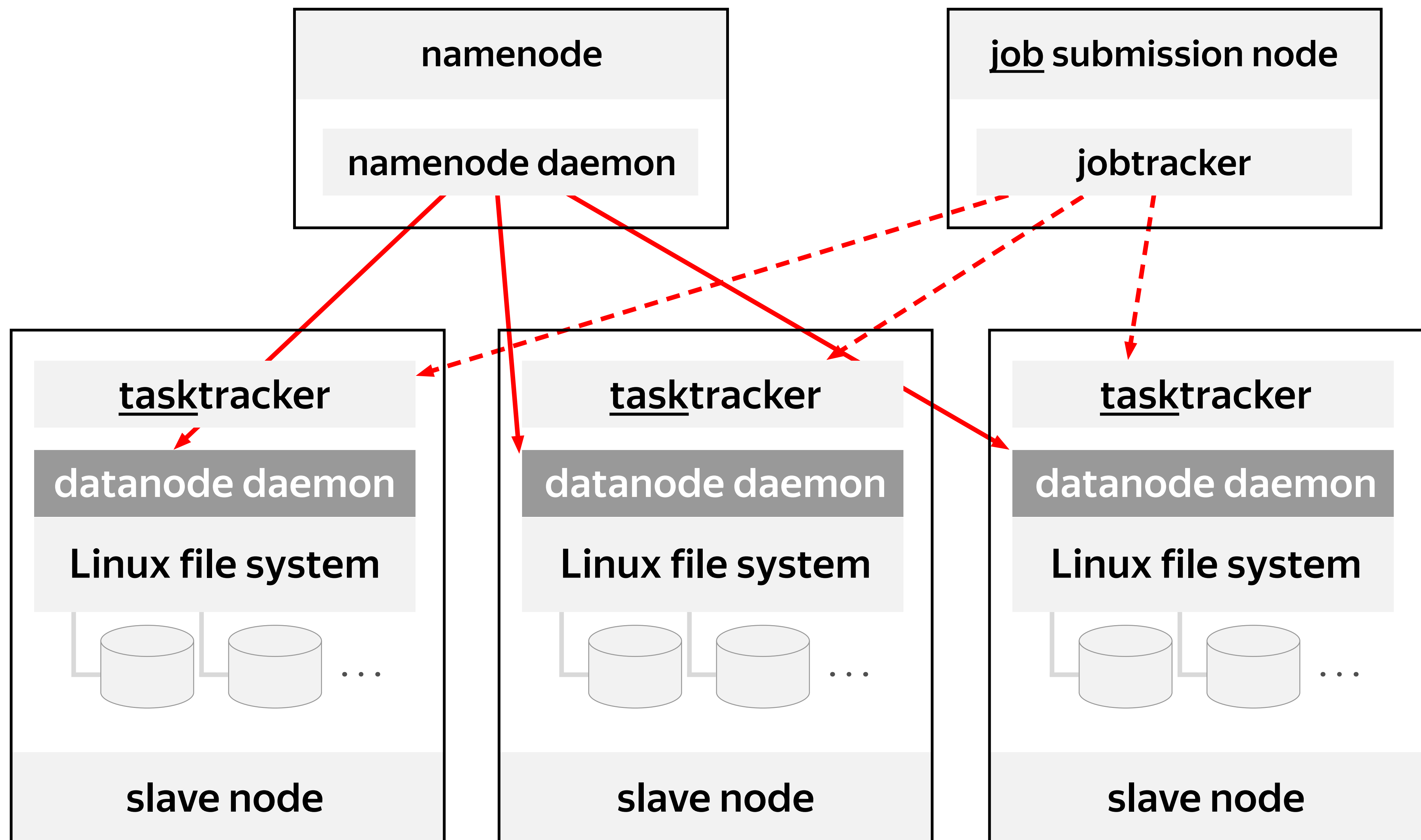
Map
phase

Intermediate files
(on local disks)

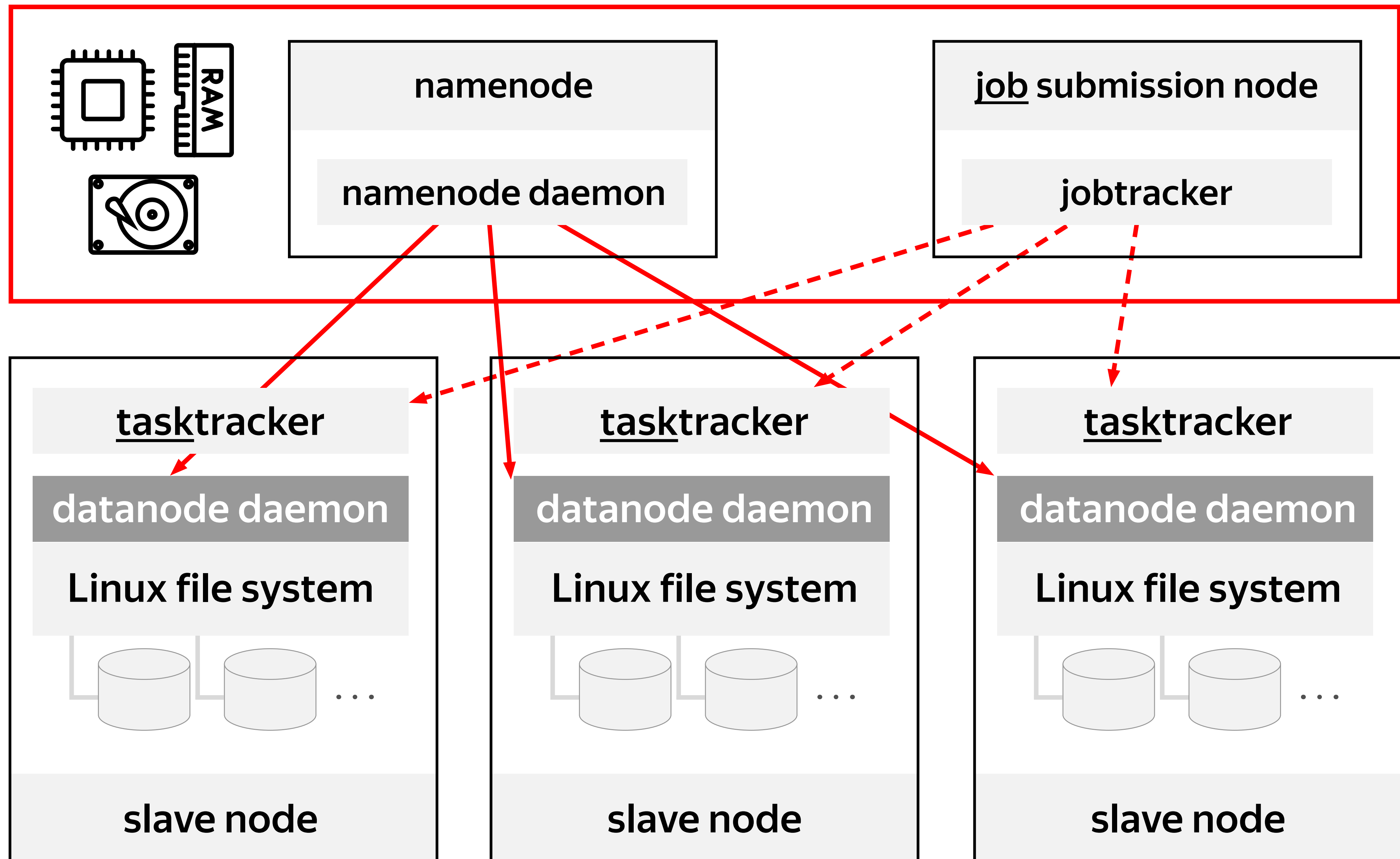
Reduce
phase

Output
files

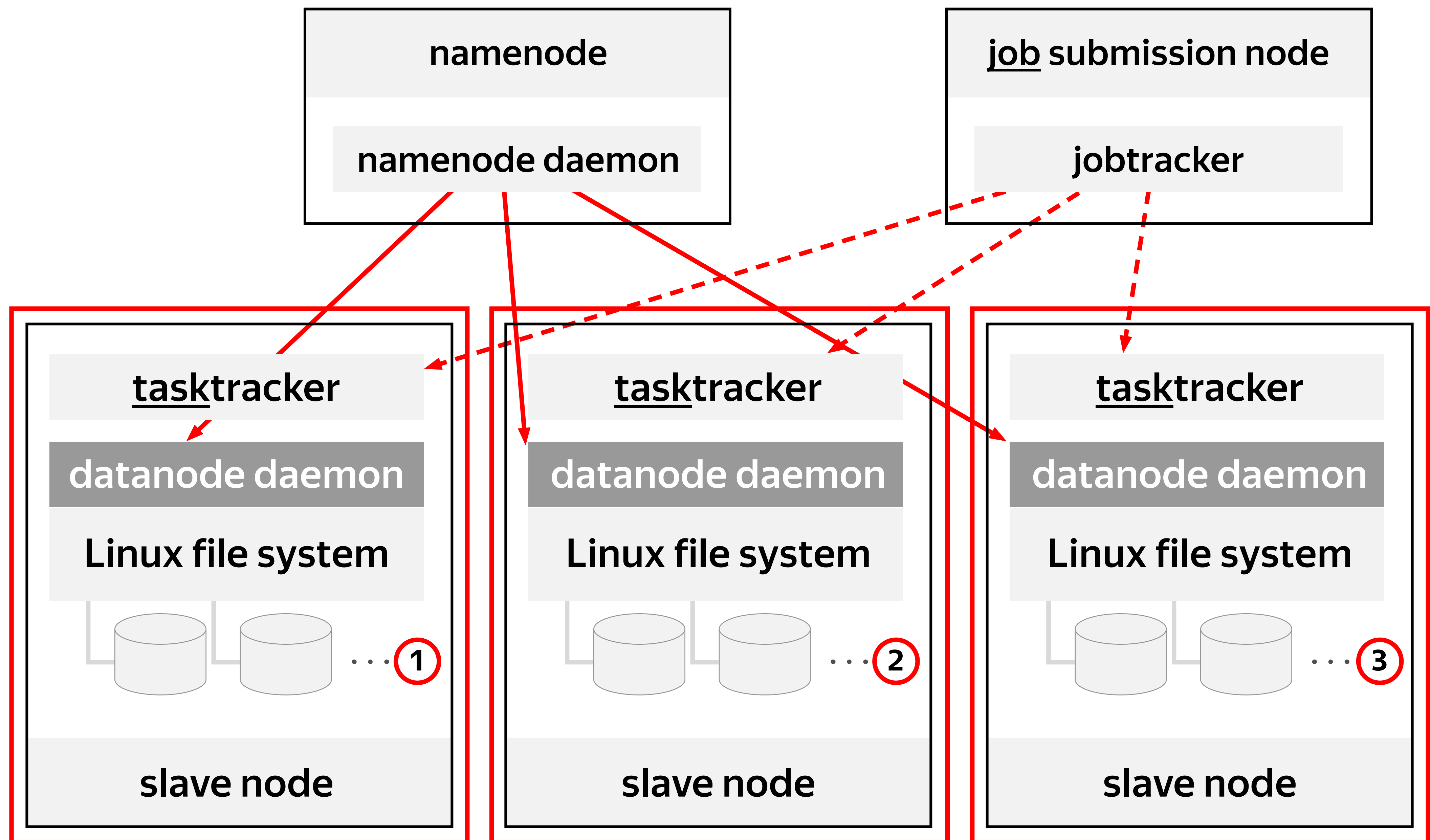




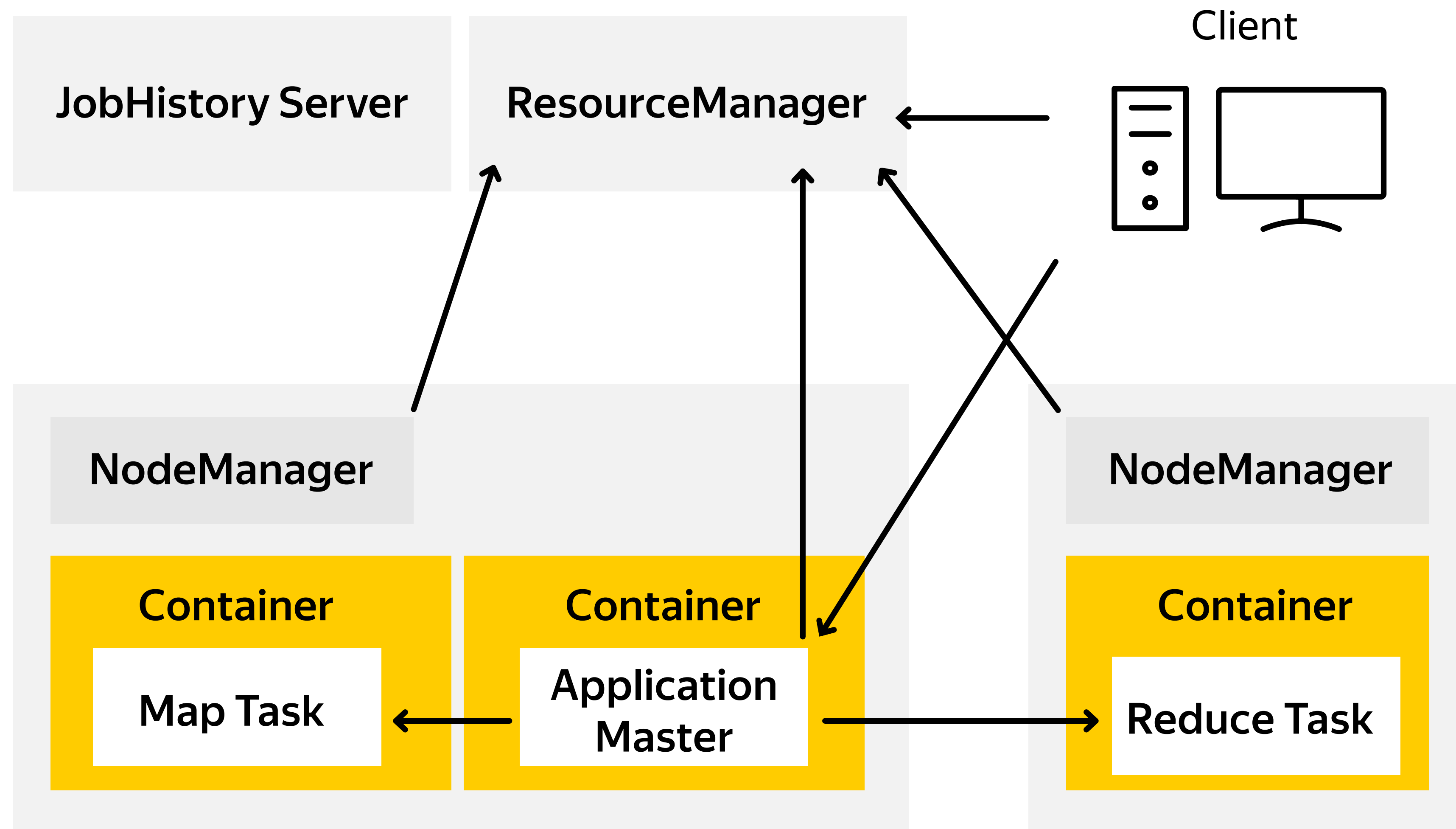
Hadoop MRv1



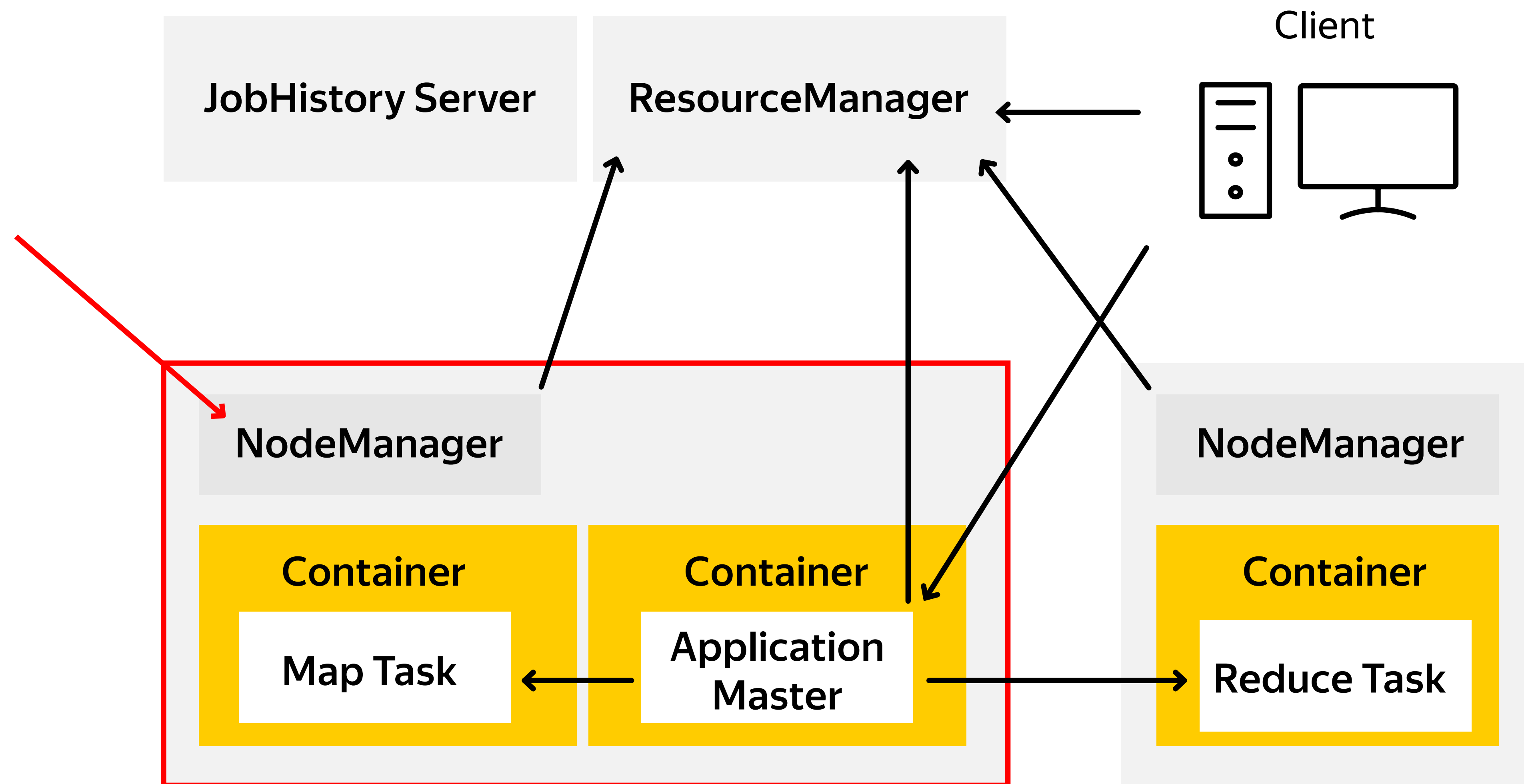
Hadoop MRv1



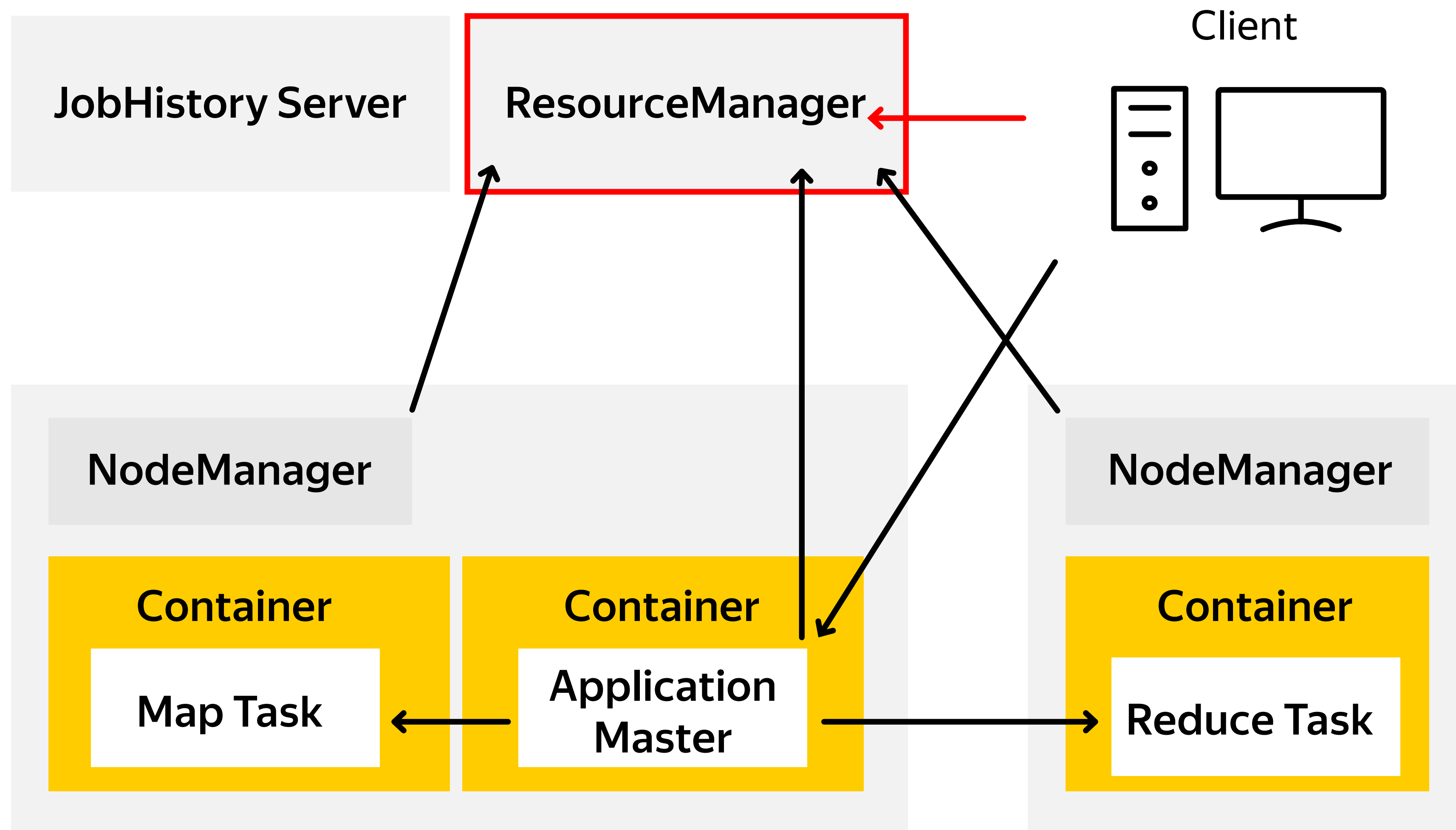
MRv1 vs YARN



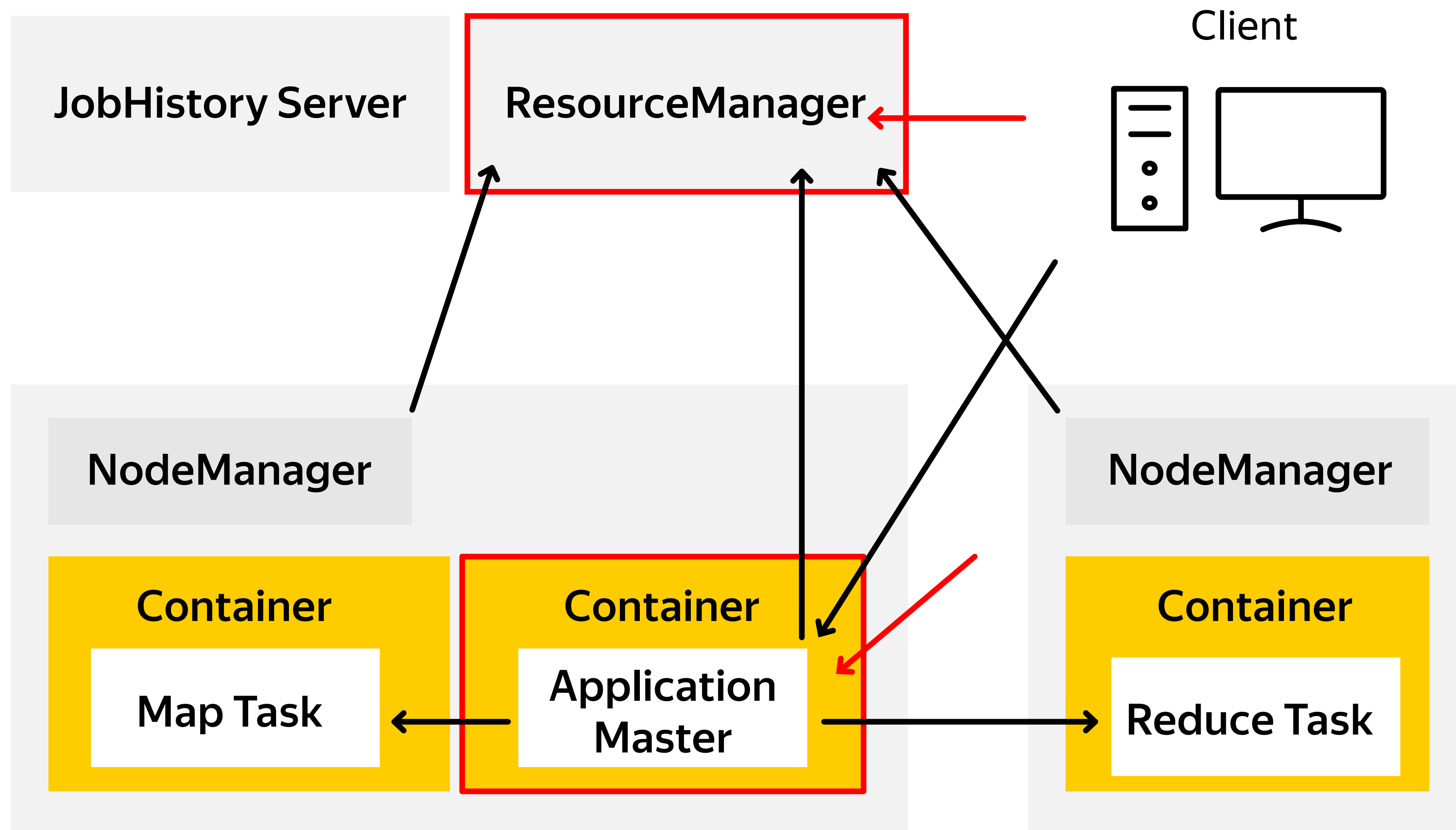
MRv1 vs YARN



MRv1 vs YARN

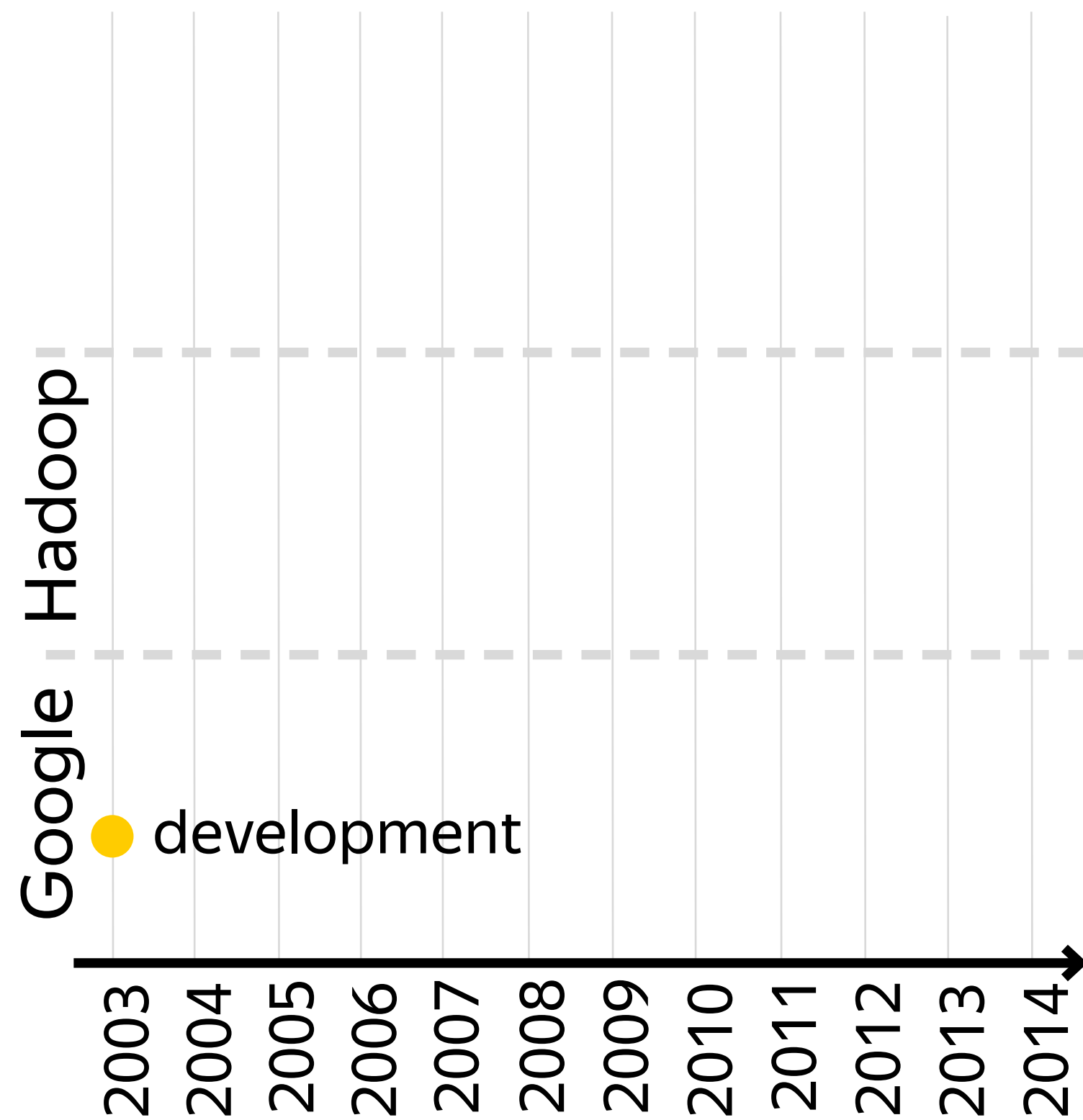


MRv1 vs YARN

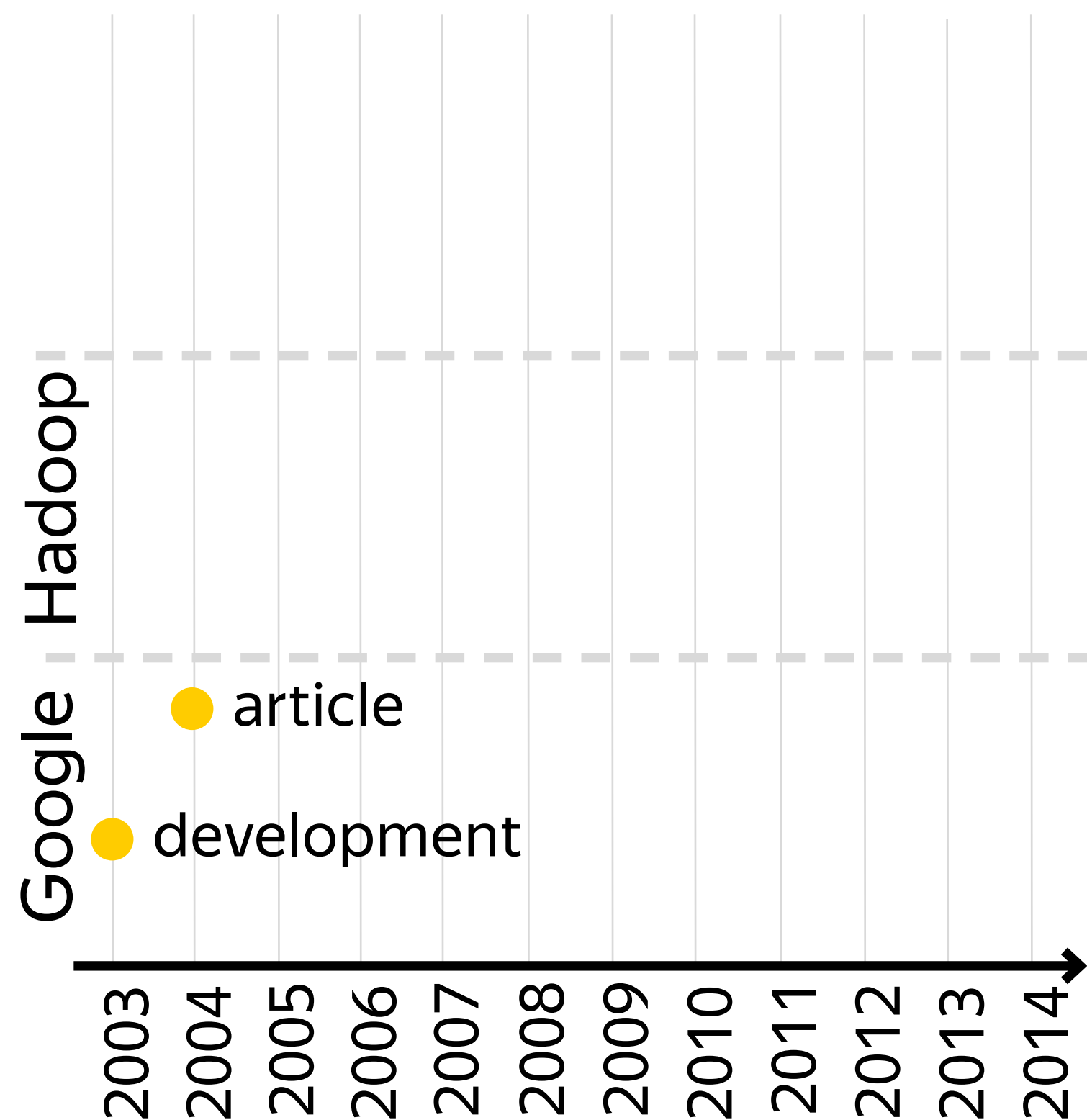


MapReduce Frameworks: History Timeline

- › [2003] Google MapReduce (development)

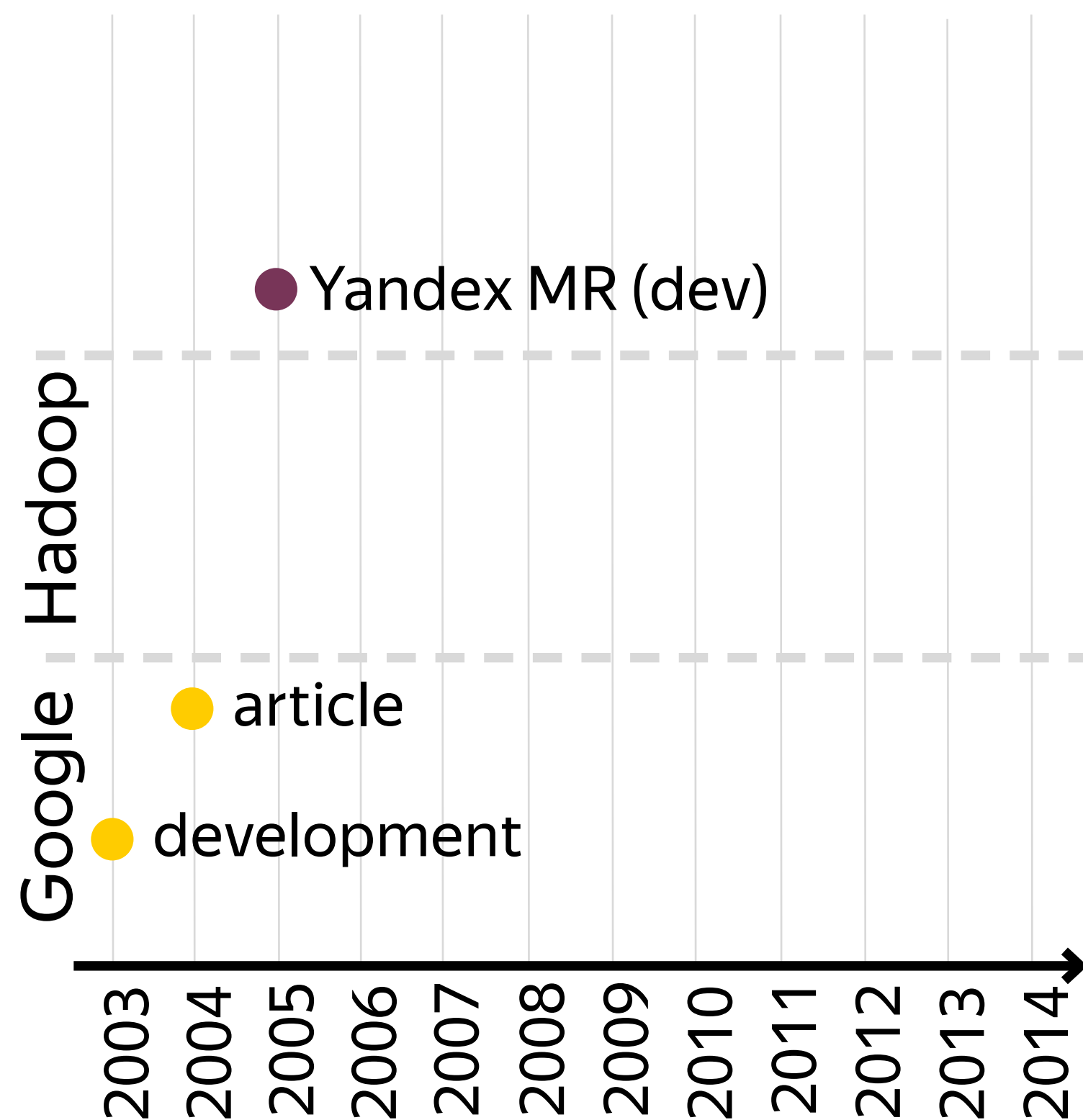


MapReduce Frameworks: History Timeline



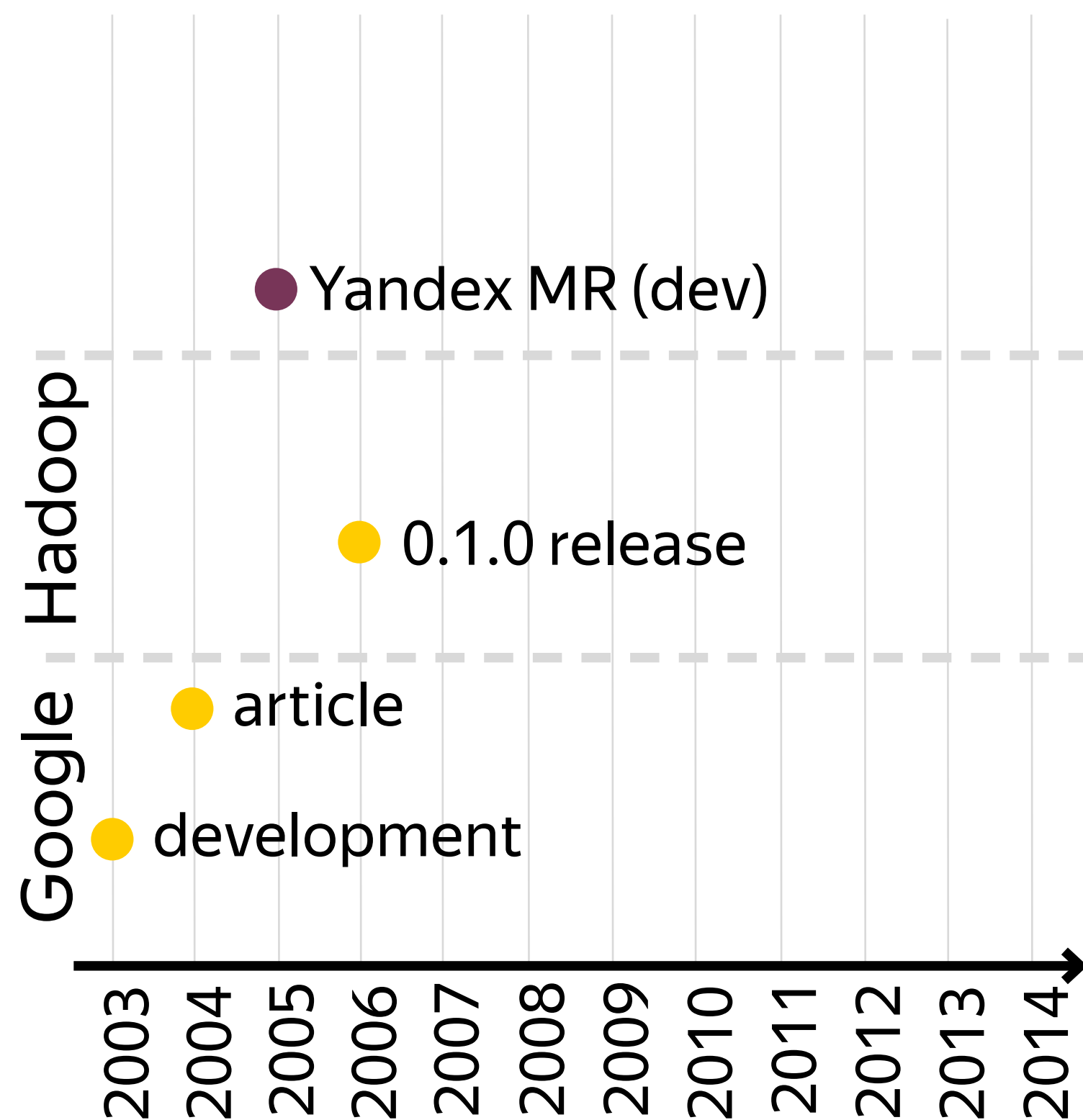
- › [2003] Google MapReduce (development)
- › [2004] Google MapReduce (article)

MapReduce Frameworks: History Timeline



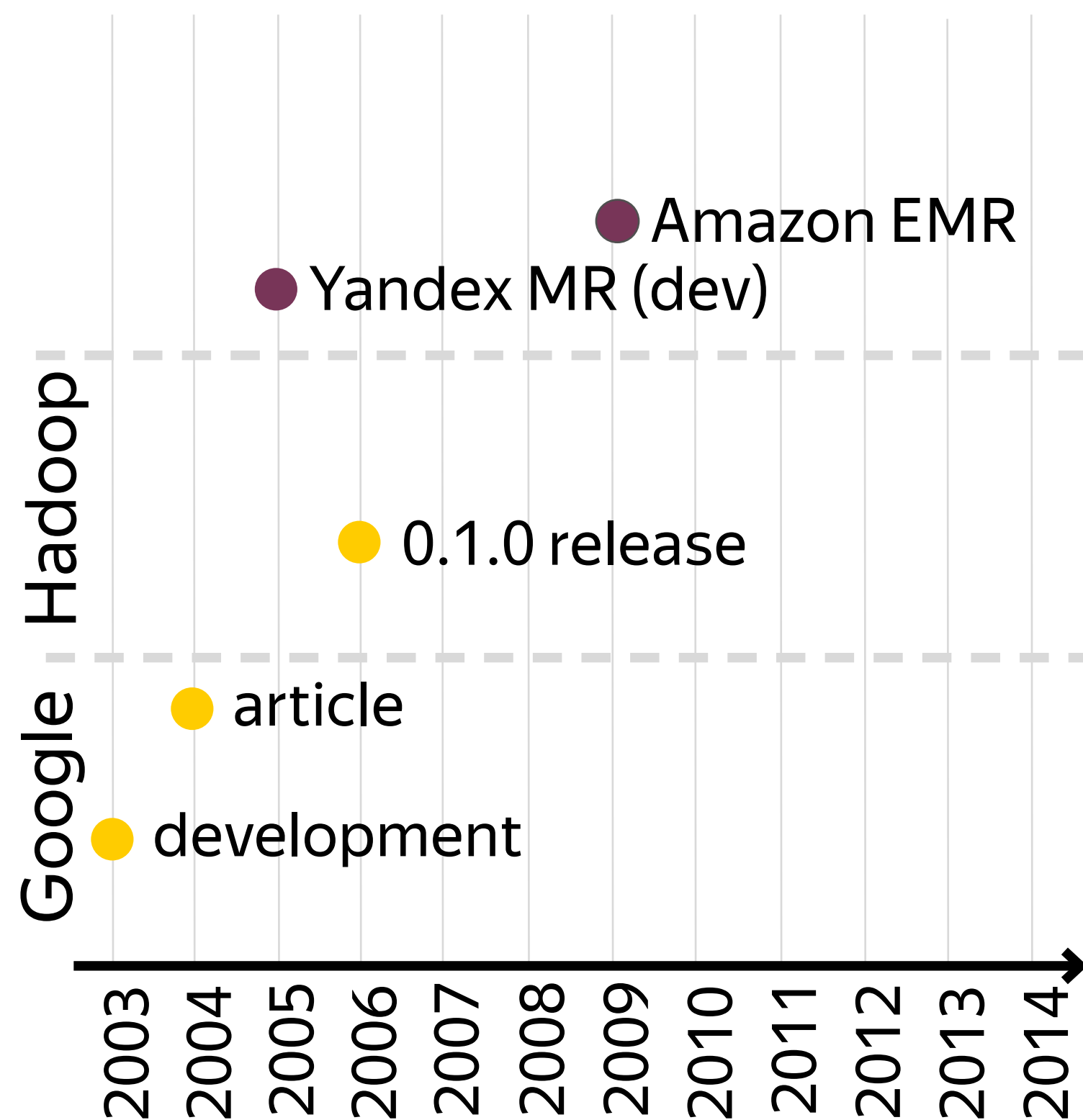
- › [2003] Google MapReduce (development)
- › [2004] Google MapReduce (article)
- › [2005] Yandex MapReduce (development)

MapReduce Frameworks: History Timeline



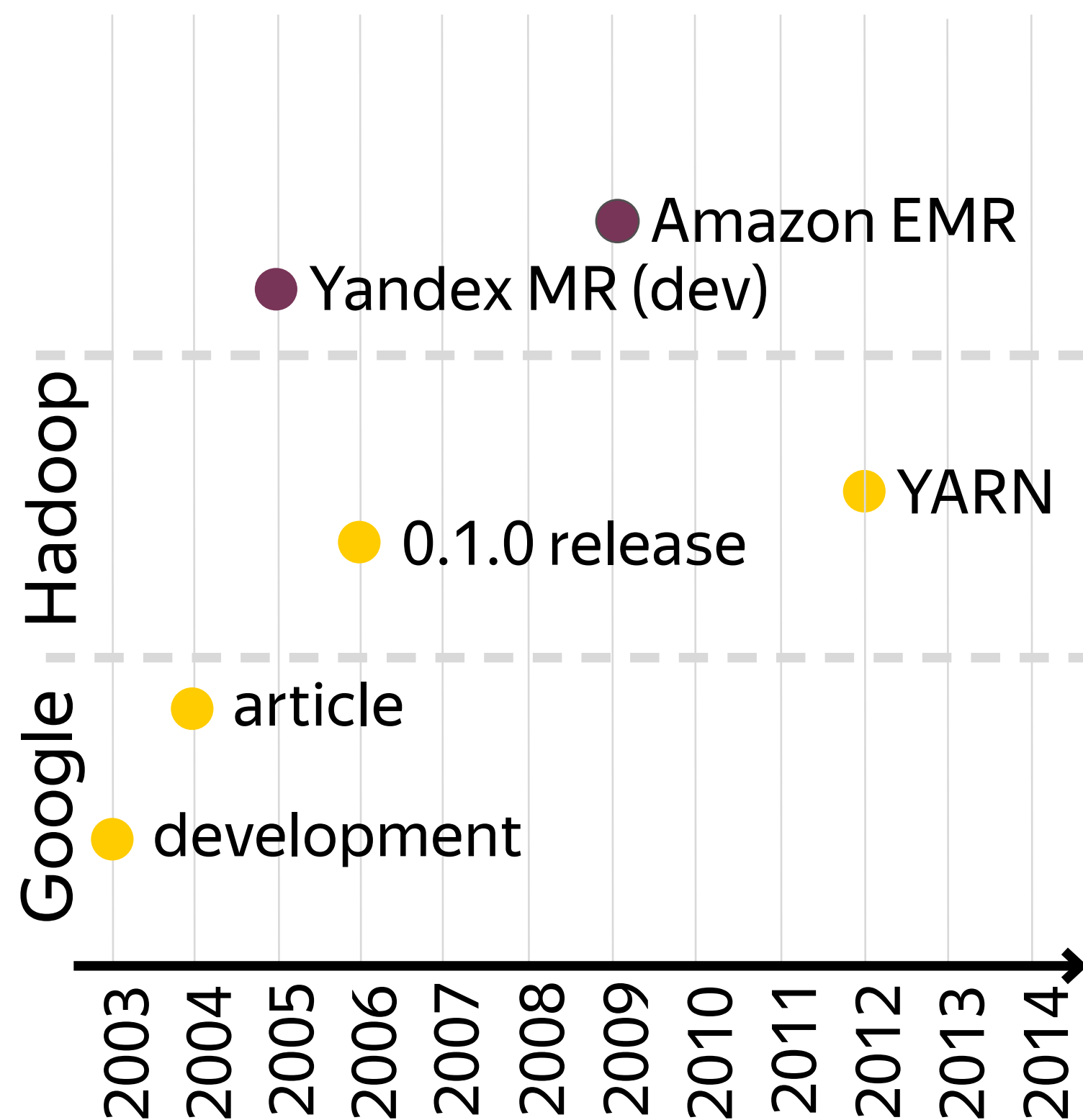
- › [2003] Google MapReduce (development)
- › [2004] Google MapReduce (article)
- › [2005] Yandex MapReduce (development)
- › [2006] Hadoop 0.1.0 release

MapReduce Frameworks: History Timeline



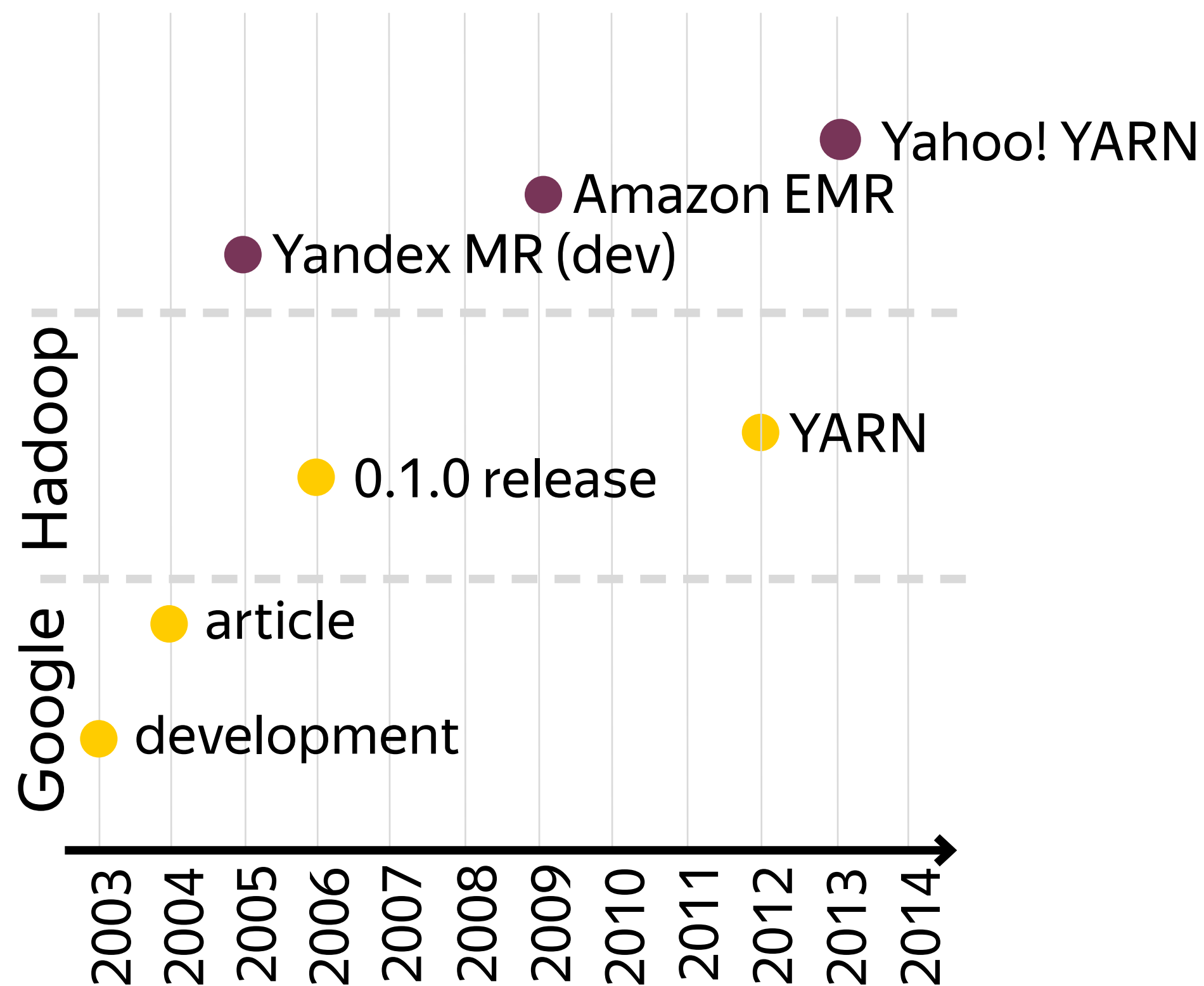
- › [2003] Google MapReduce (development)
- › [2004] Google MapReduce (article)
- › [2005] Yandex MapReduce (development)
- › [2006] Hadoop 0.1.0 release
- › [2009] Amazon EMR (Hadoop inside)

MapReduce Frameworks: History Timeline



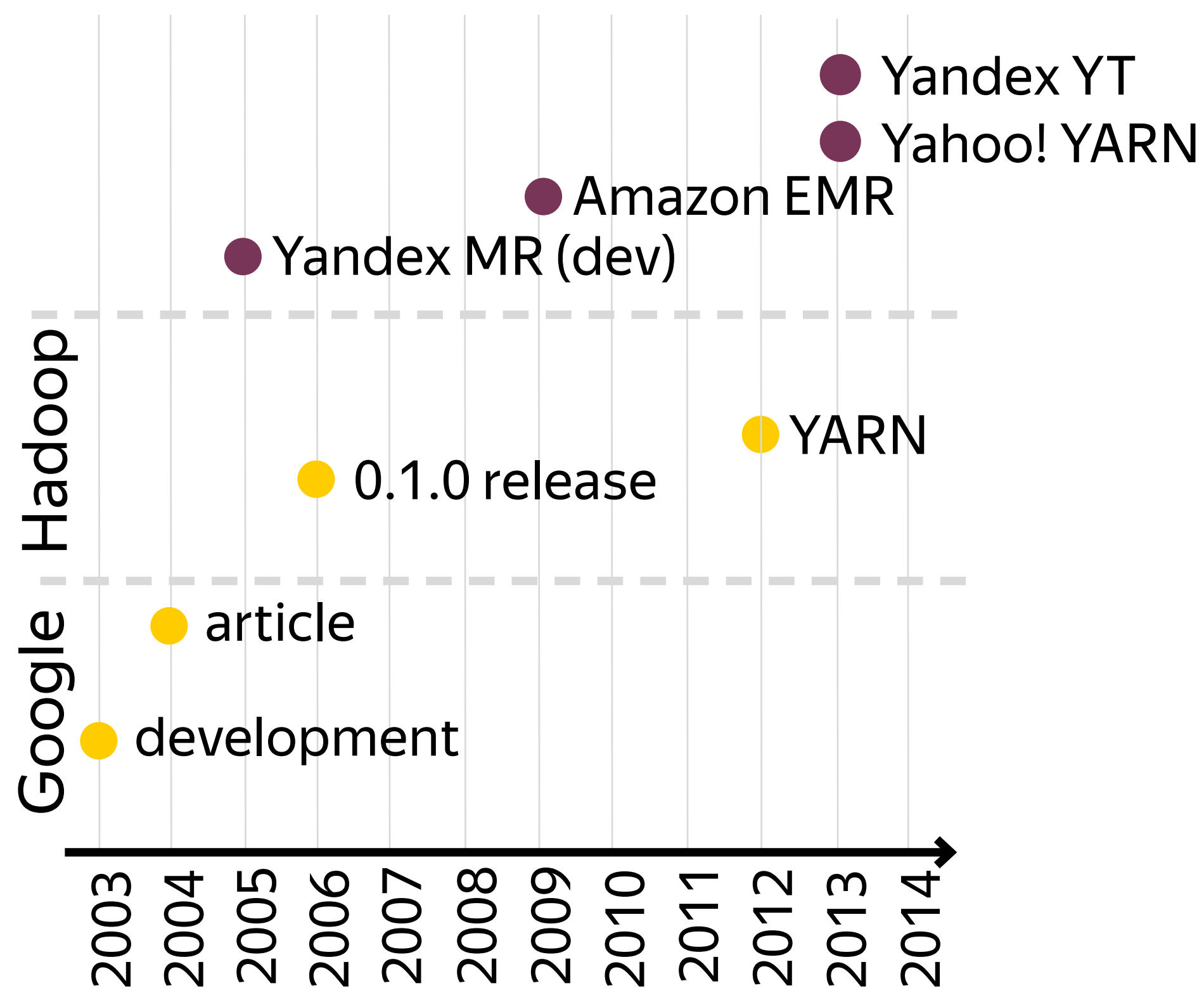
- › [2003] Google MapReduce (development)
- › [2004] Google MapReduce (article)
- › [2005] Yandex MapReduce (development)
- › [2006] Hadoop 0.1.0 release
- › [2009] Amazon EMR (Hadoop inside)
- › [2012] MapReduce —> YARN

MapReduce Frameworks: History Timeline



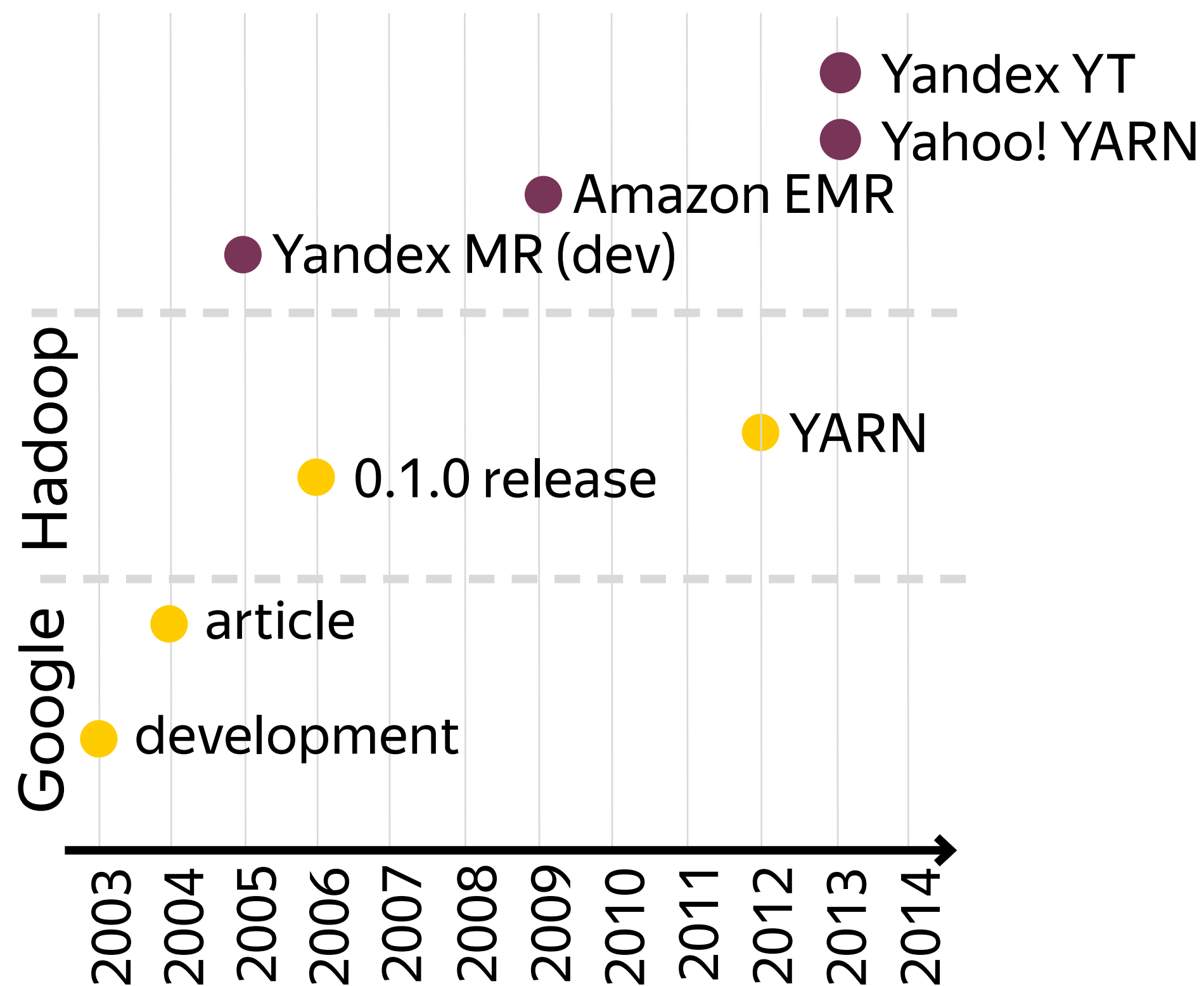
- › [2003] Google MapReduce (development)
- › [2004] Google MapReduce (article)
- › [2005] Yandex MapReduce (development)
- › [2006] Hadoop 0.1.0 release
- › [2009] Amazon EMR (Hadoop inside)
- › [2012] MapReduce → YARN
- › [2013] Yahoo! YARN deployed in production

MapReduce Frameworks: History Timeline



- › [2003] Google MapReduce (development)
- › [2004] Google MapReduce (article)
- › [2005] Yandex MapReduce (development)
- › [2006] Hadoop 0.1.0 release
- › [2009] Amazon EMR (Hadoop inside)
- › [2012] MapReduce → YARN
- › [2013] Yahoo! YARN deployed in production
- › [2013] Yandex YT

MapReduce Frameworks: History Timeline



- › [2003] Google MapReduce (development)
- › [2004] Google MapReduce (article)
- › [2005] Yandex MapReduce (development)
- › [2006] Hadoop 0.1.0 release
- › [2009] Amazon EMR (Hadoop inside)
- › [2012] MapReduce → YARN
- › [2013] Yahoo! YARN deployed in production
- › [2013] Yandex YT
- › ...
- › MapReduce in MongoDB, Riak, ...

MapReduce Framework

- › You can **explain** what will happen if Mapper or Reducer dies;
- › You know what JobTracker and TaskTracer in MRv1, ResourceManager and NodeManager in YARN are.

BigDATAteam