

Hadoop's Adolescence: A Comparative Workloads Analysis from Three Research Clusters

Kai Ren, Garth Gibson (CMU), YongChul Kwon, Magdalena Balazinska, Bill Howe (UW Seattle)

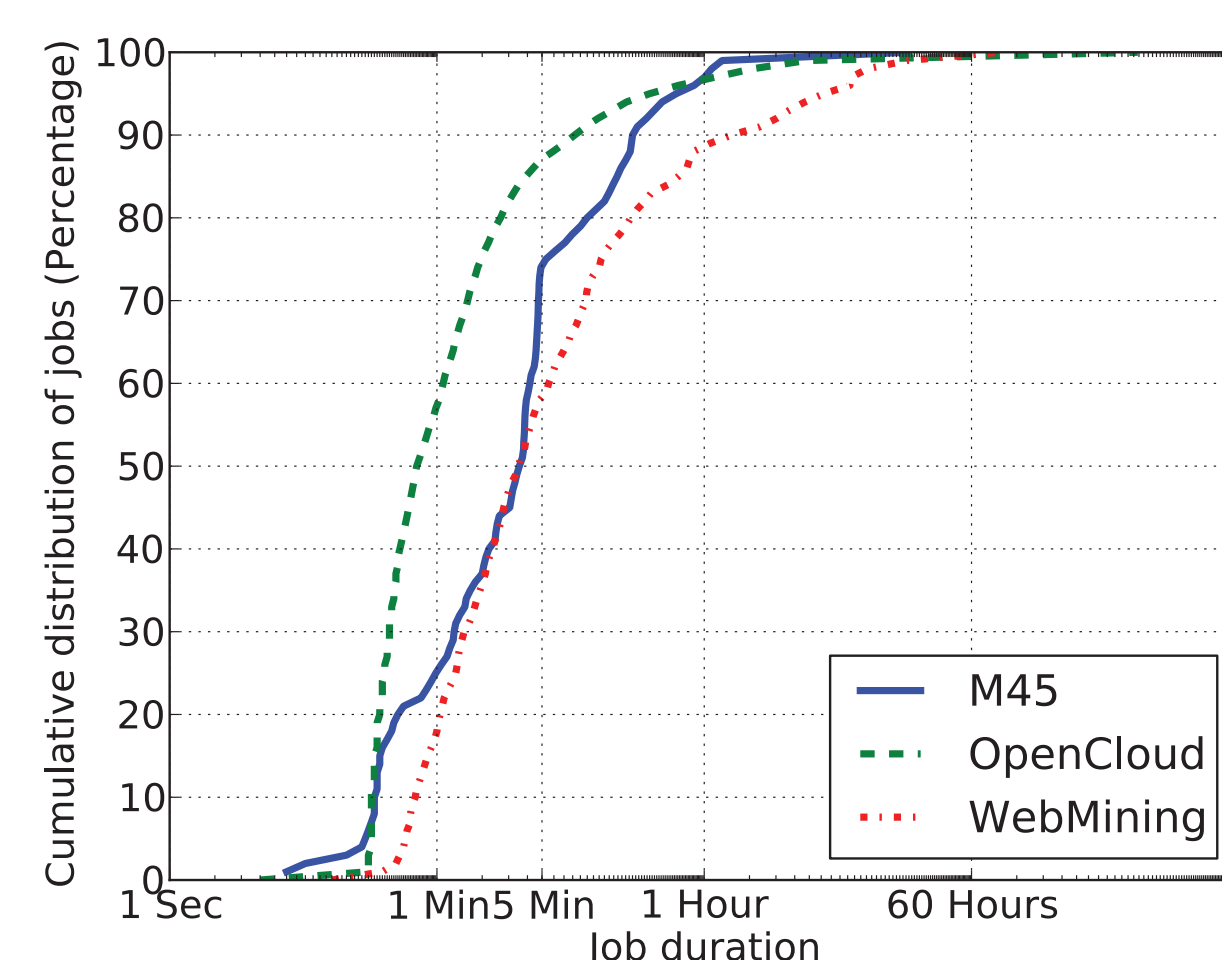
Data Collection

- **OpenCloud:** A research cluster at CMU, consisting of 64 nodes, and each node has 2x 2.8GHz quad-core CPU, 16GB RAM, 10 Gbps Ethernet NIC and four Seagate 7200 RPM SATA disks.
- **M45:** A production Hadoop cluster provided by Yahoo! for academic research. It consists of 400 nodes, each with 2x 1.86GHz quad-core CPU, 6GB RAM, and four 7200 RPM SATA disks.
- **WebMining:** 9 nodes cluster from an anonymized research group. 4x quad-core CPU Xeon E5630, 32GB RAM, and four 1.8TB SATA disks.

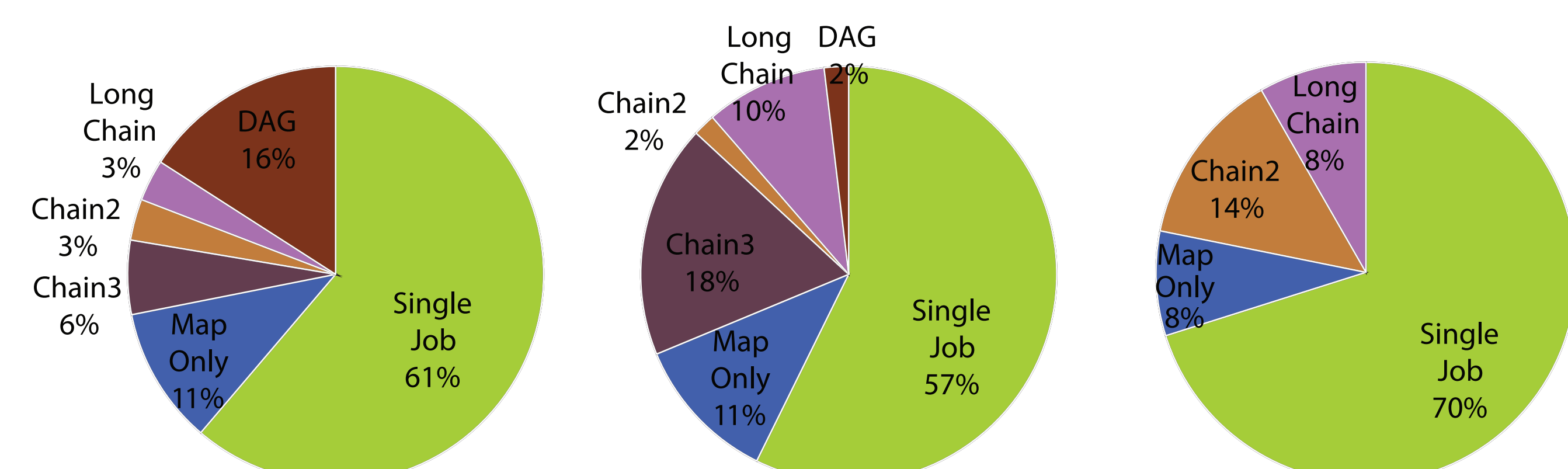
Cluster	Duration	# Jobs	# Users
OpenCloud	1.5 years	58321	78
M45	5 months	43609	20
Web Mining	5 months	1229	5

Application Workload Profiling

- **Job Size:** Small jobs are prevalent in all three workloads. Typical small jobs include debugging, ad-hoc query and etc.



- **Job Structure:** we analyze the dependency between jobs by looking at whether a job uses another job's output as input



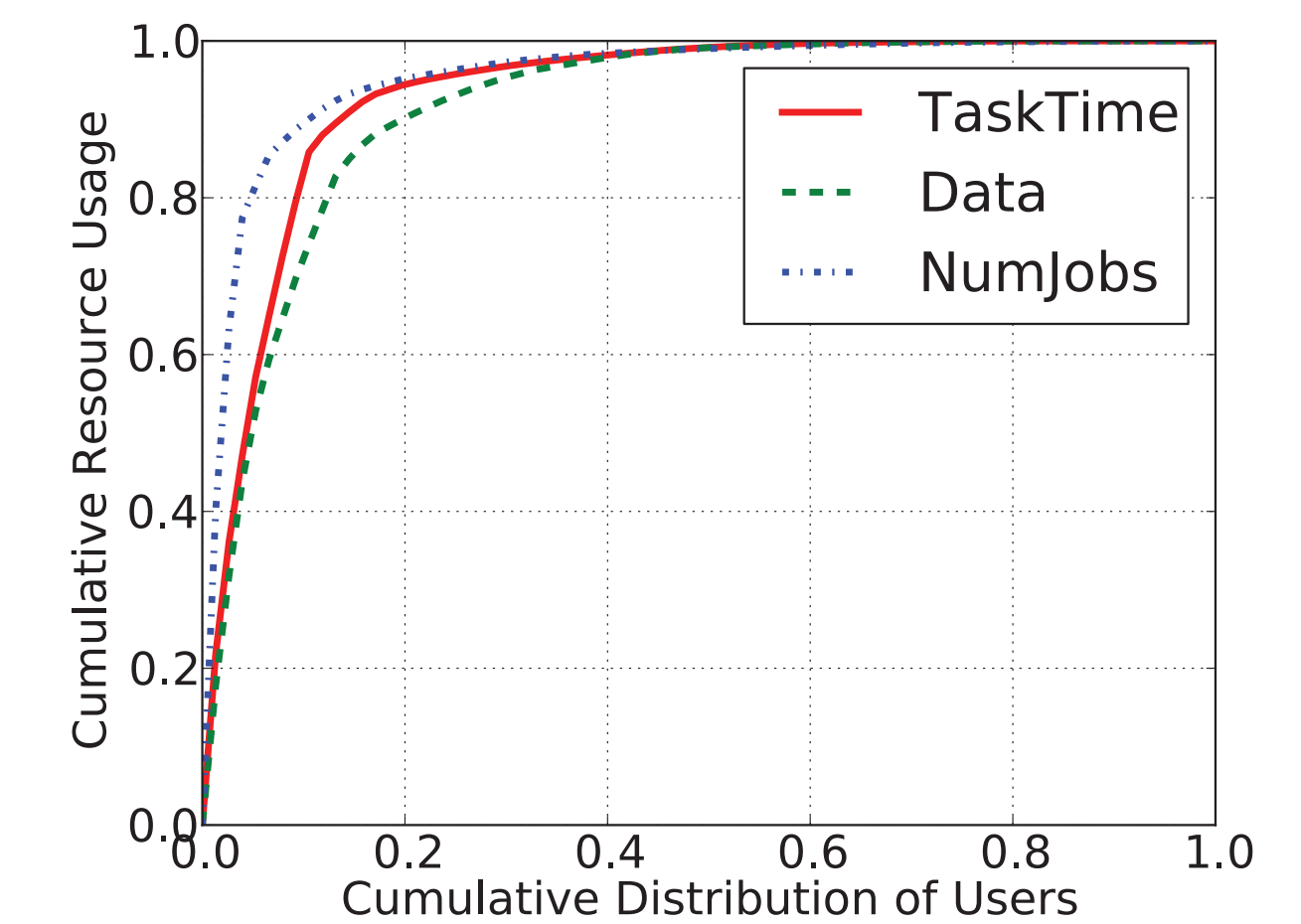
- **Job Types:** Native Java interface in MapReduce are most popular. Pig and Streaming job which can ease programming are also popular.



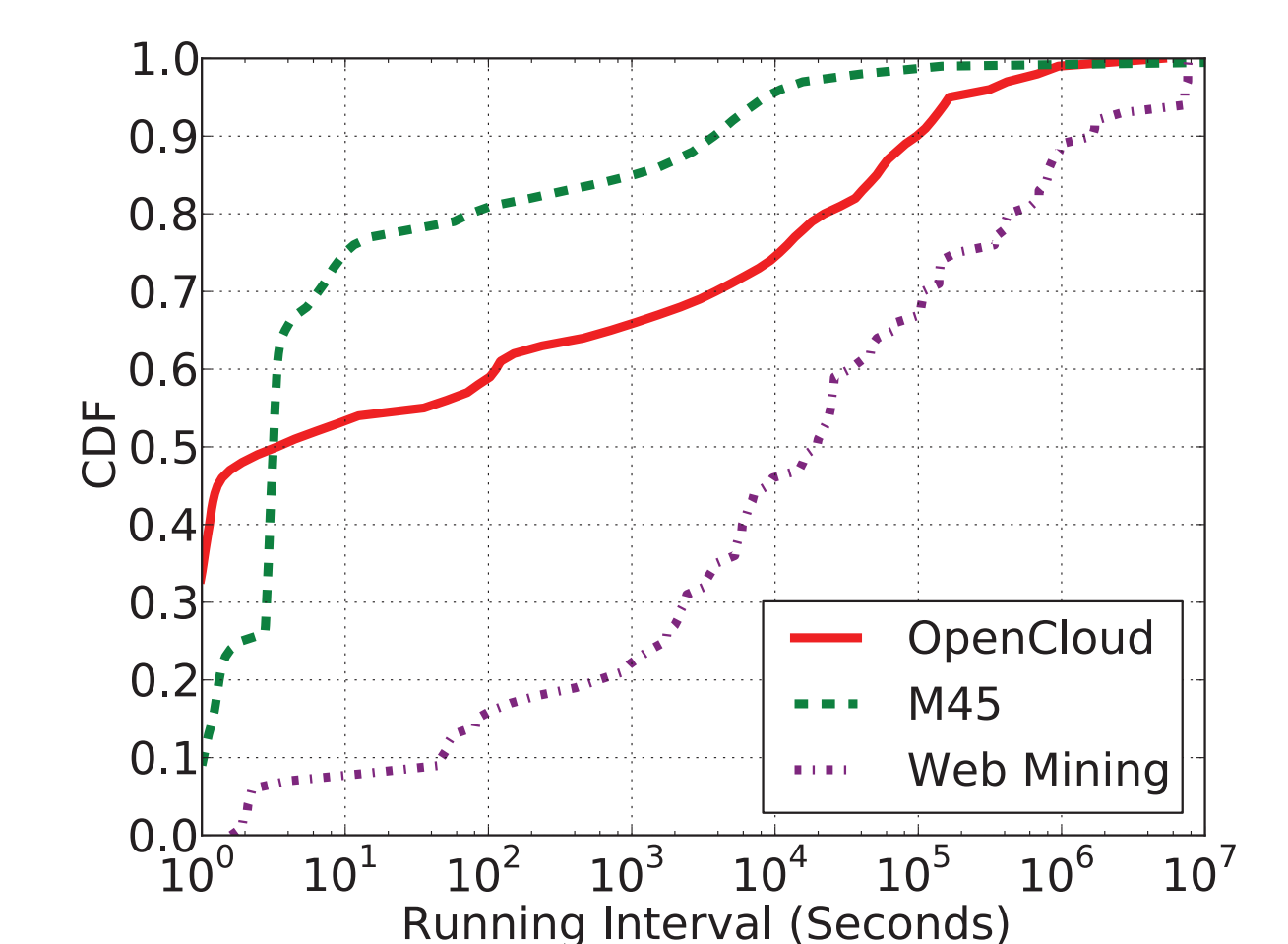
- **Job Configuration Tuning:** Users tend to tune parameters directly related to failures, rather than performance consideration. For example, tuning JVM option to avoid "Out of memory" errors.

User Behavior Analysis

- **User Usage:** Resource usage of each user follows power-law distribution. A few of users dominate the utilization of the entire cluster.

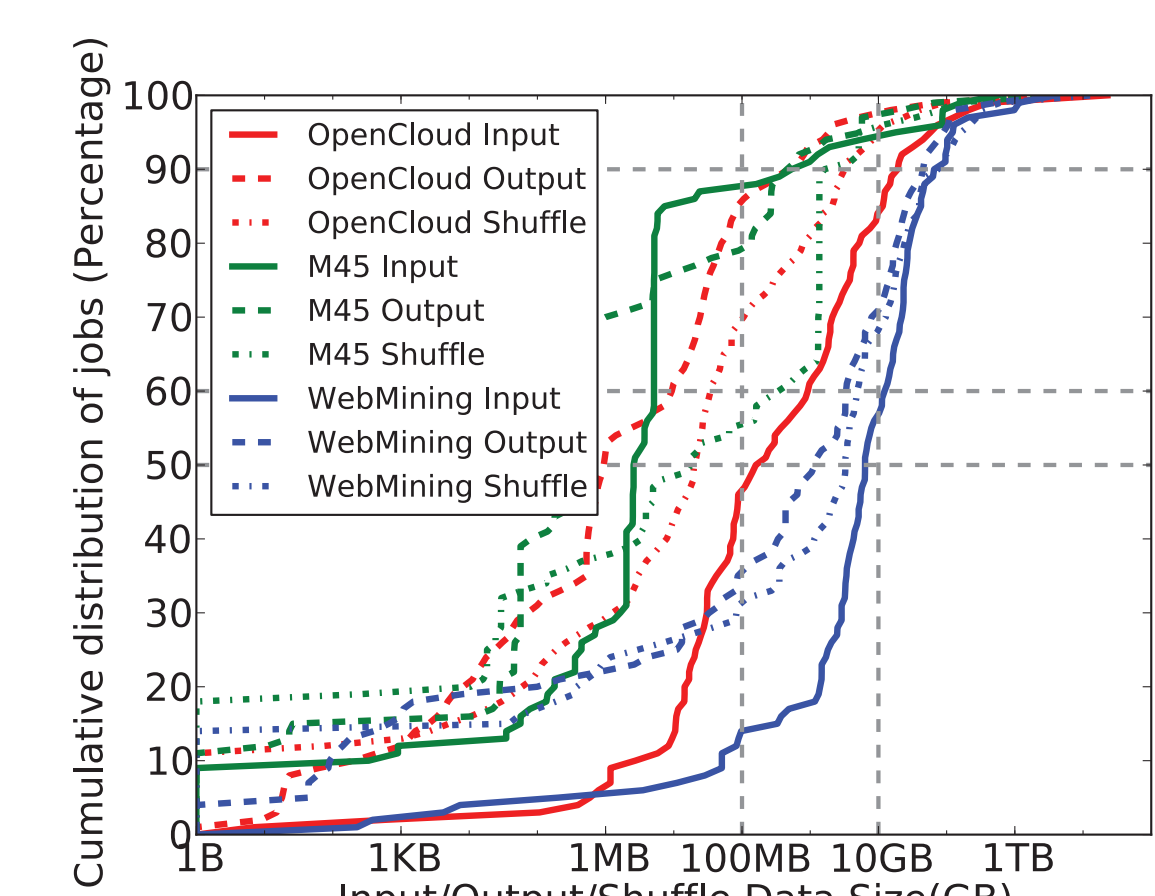
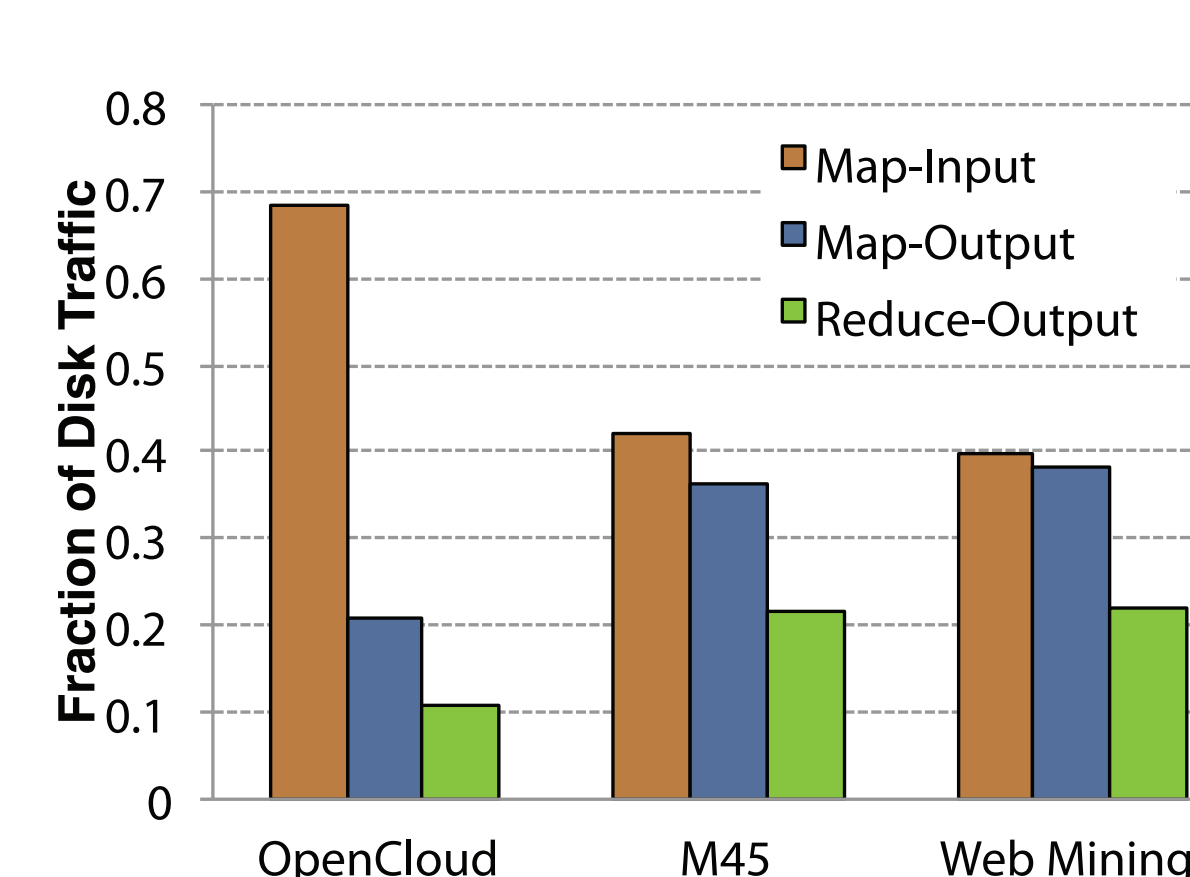


- **Job Submission:** The distribution of submission intervals between jobs within a batch shows at least 20% jobs are interactive, which are not supported well by Hadoop.

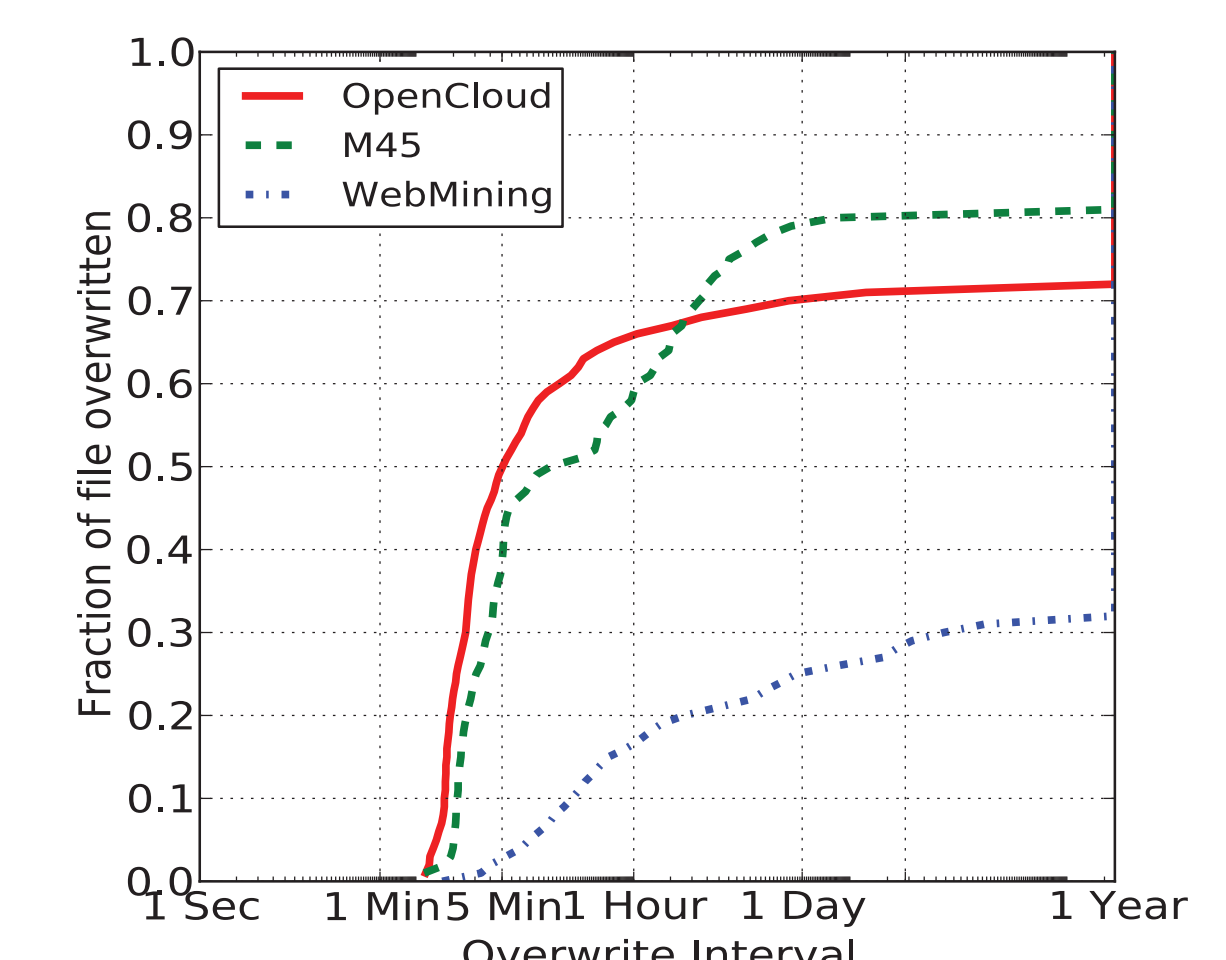
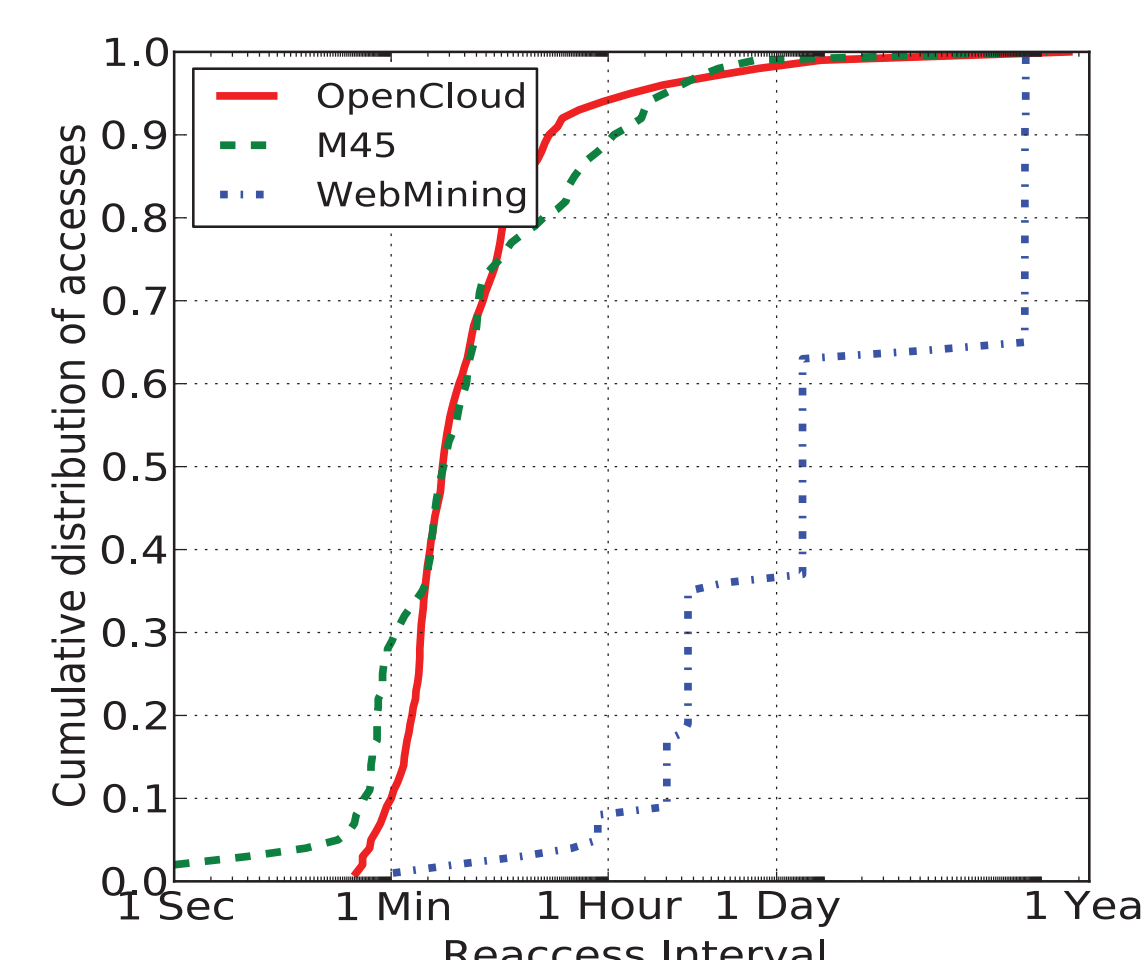


Storage Workload Analysis

- **Data Characteristics:** Map-Input dominates the I/O traffic, thus optimizing map task locality is crucial. Many jobs have input data less than 10 GB which confirms the prevalence of small-scale jobs.



- **Access Pattern:** 90% data re-access happen within 1 hour. Since lots of small I/O, simulation shows a global cache with 1GB per node can reduce 60~70% disk I/O. About 50% jobs have their data overwritten within 5 minutes.



Conclusion

- The use of Hadoop for academic research is still in its adolescence.
- Many applications are independent, single-step, small-scale jobs that may be amenable to simpler non-Hadoop solution.
- Underuse of Hadoop features, extensions, and optimization tools.
- Techniques like caching may help with I/O performance.

