A Model-Based Namespace Metadata Benchmark for HDFS

Cristina L. Abad[1], Yi Lu[2], Roy H. Campbell[2], Nathan Roberts[3]

[1] Escuela Superior Politécnica del Litoral

[2] University of Illinois at Urbana-Champaign

[3] Yahoo, Inc.



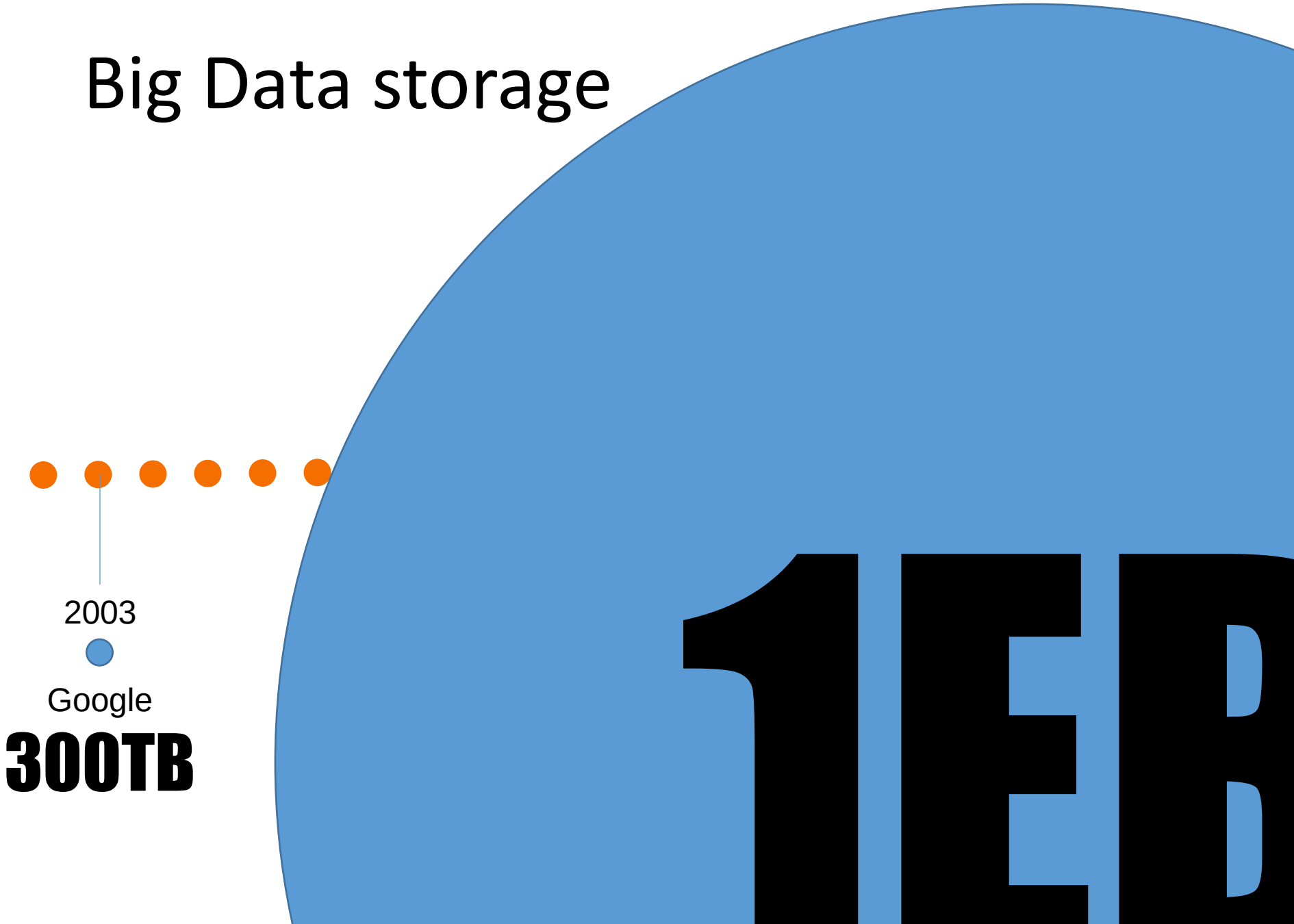


ILLINOIS

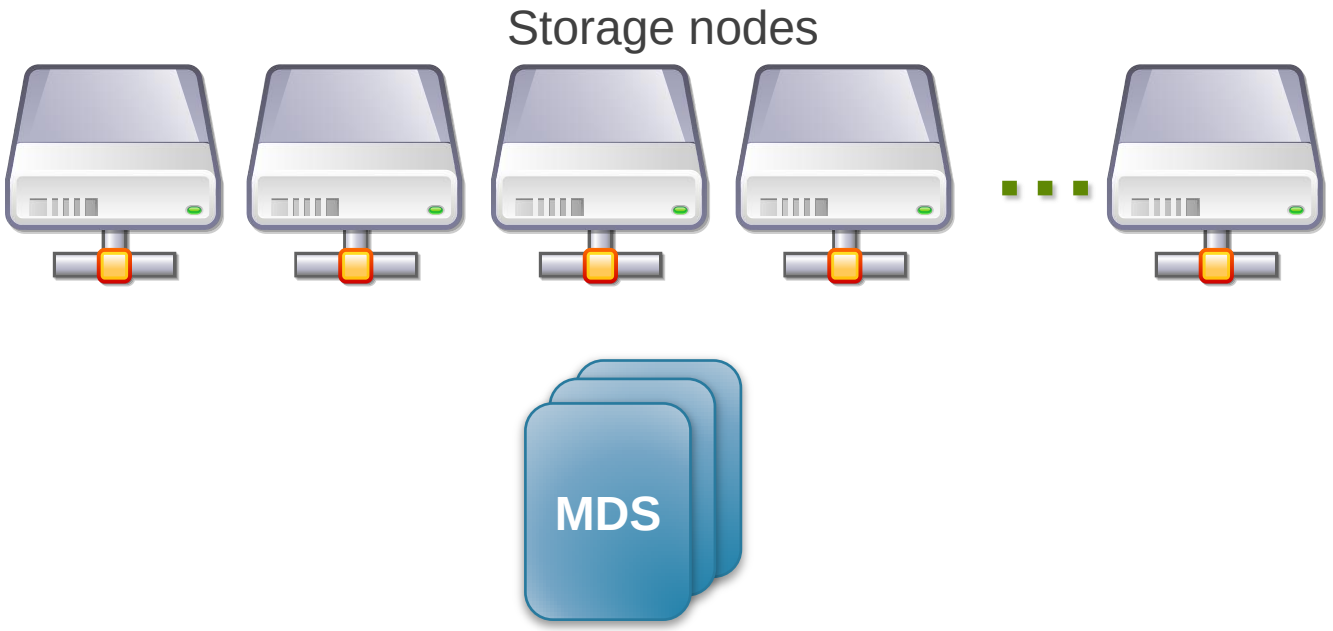
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

YAHOO!

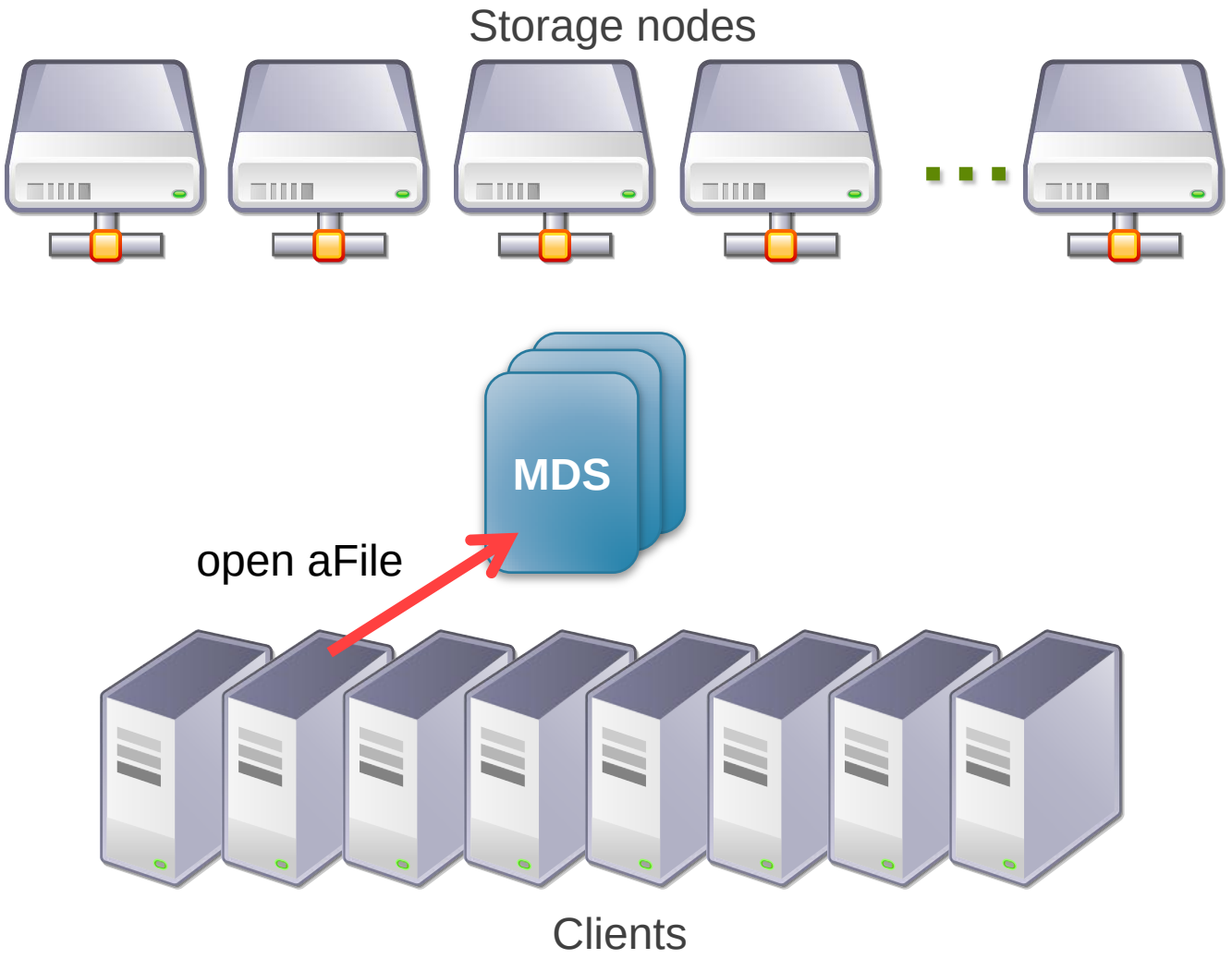
Big Data storage



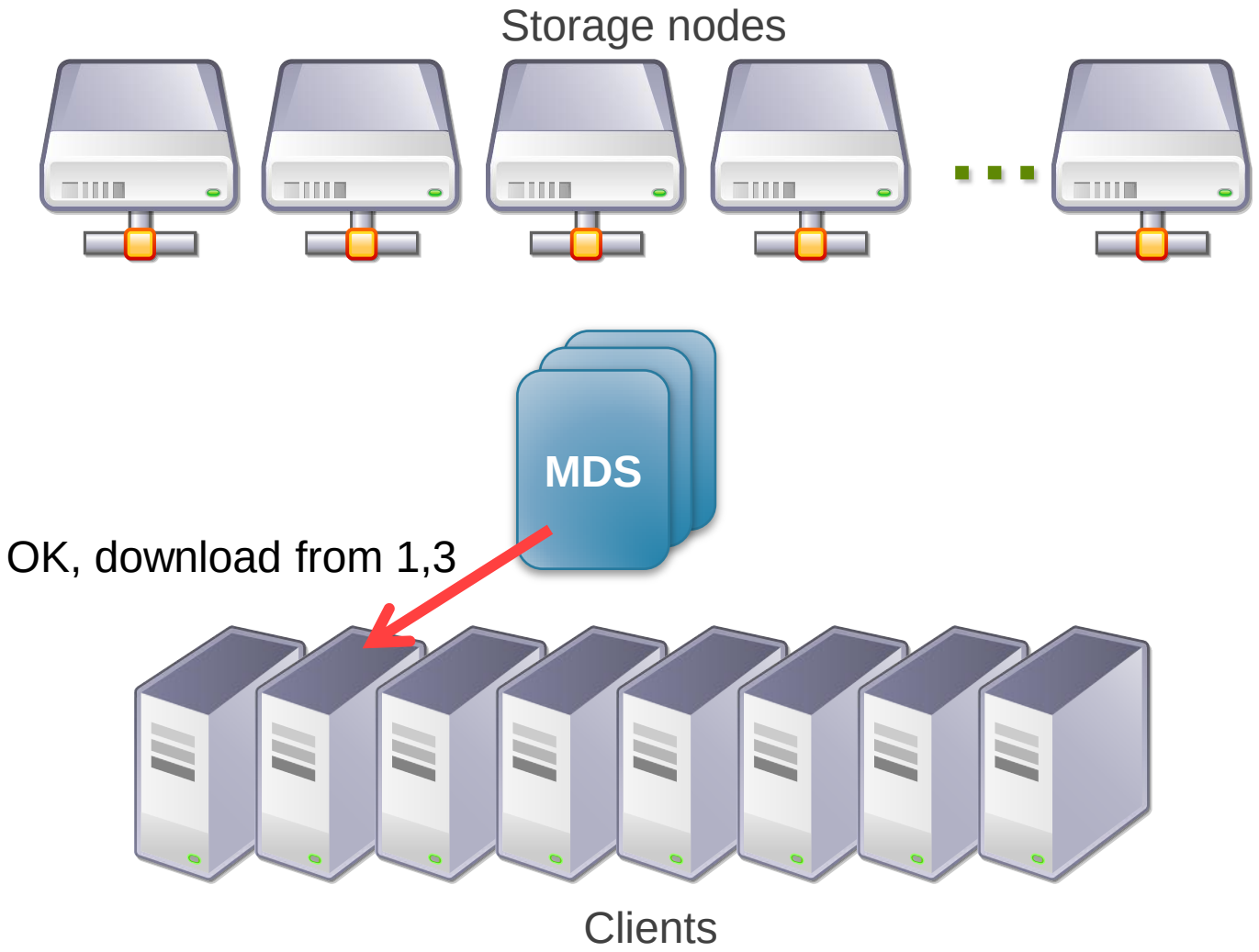
Metadata management



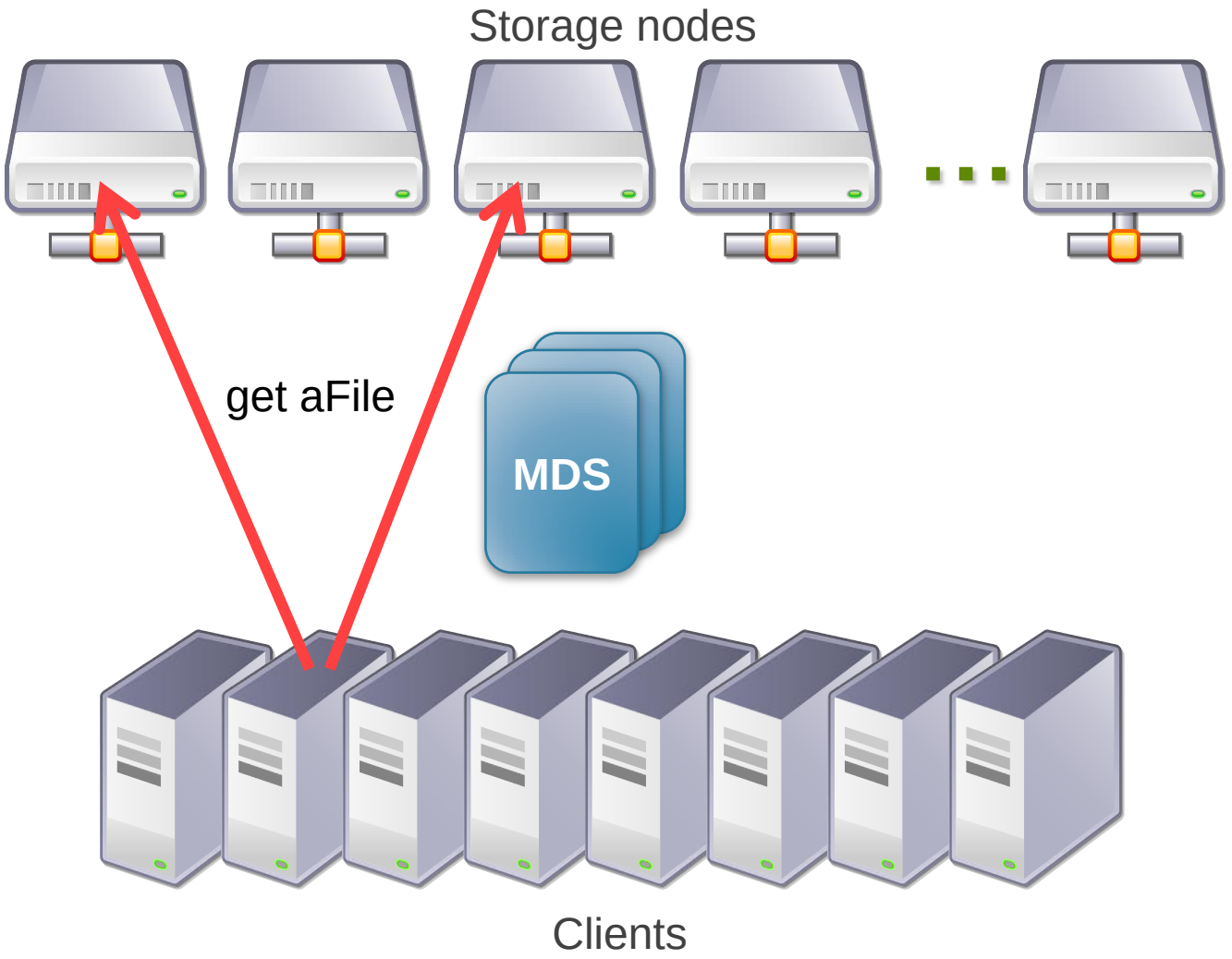
Metadata management



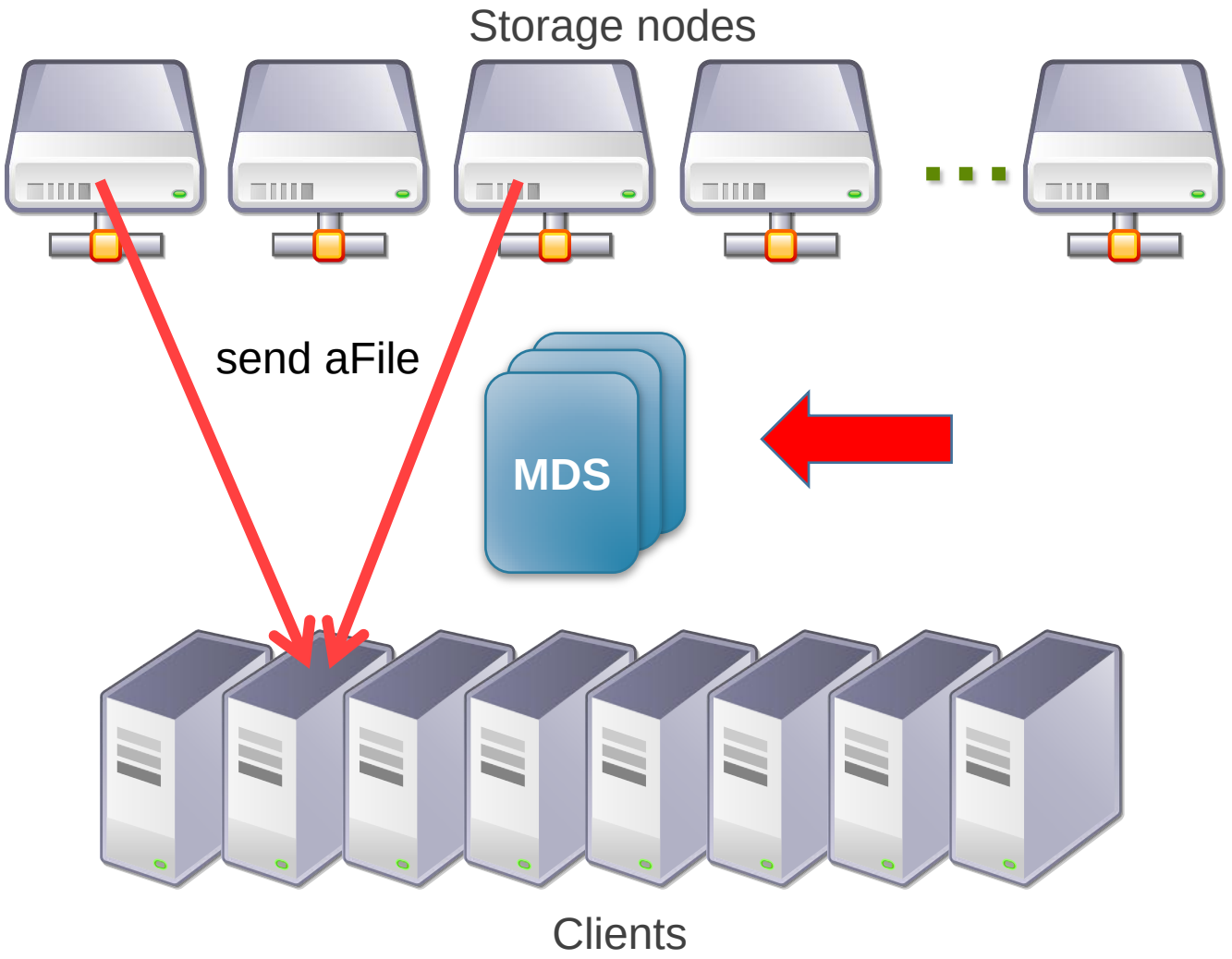
Metadata management



Metadata management



Metadata management



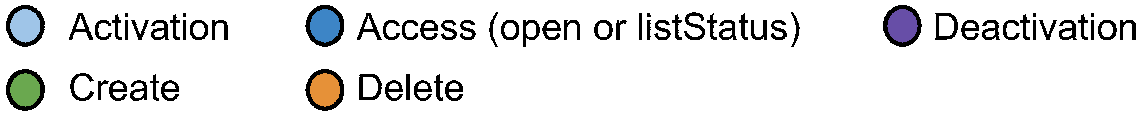
Benchmarking Medatada Servers

- Desirable properties:
  - Metadata-intensive workloads
  - No IO load (isolate MDS evaluation)
  - Realistic workloads
  - Type-aware workload generation
  - Hierarchical namespaces

MimesisBench

- Realistic namespaces
  - Namespace Generation Module [IEEE/ACM UCC 2012]
- Preserve access patterns
  - Short-term temporal correlations
  - Temporal correlations due to long-term file popularity
  - Built-upon our model for object request streams [IFIP Performance 2013]
- Support for most common metadata operations
  - Create, delete, list, open
- Pre-existing namespace
- Realistic accesses on a hierarchical namespace

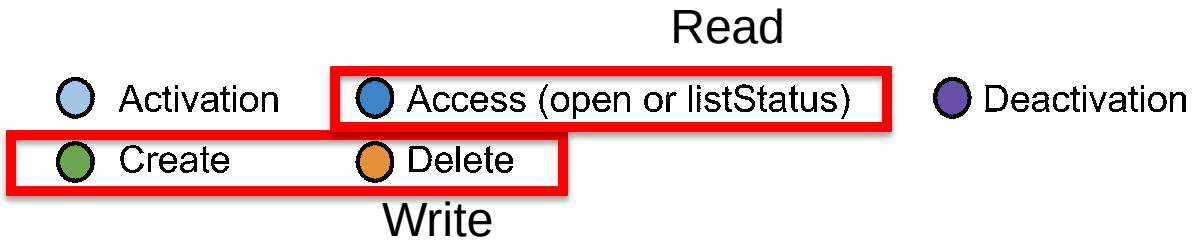
MimesisBench: Access model



- Original renewal process model:



MimesisBench: Access model



- Original renewal process model:



- Extended to include creates, deletes and listStatus:



Design

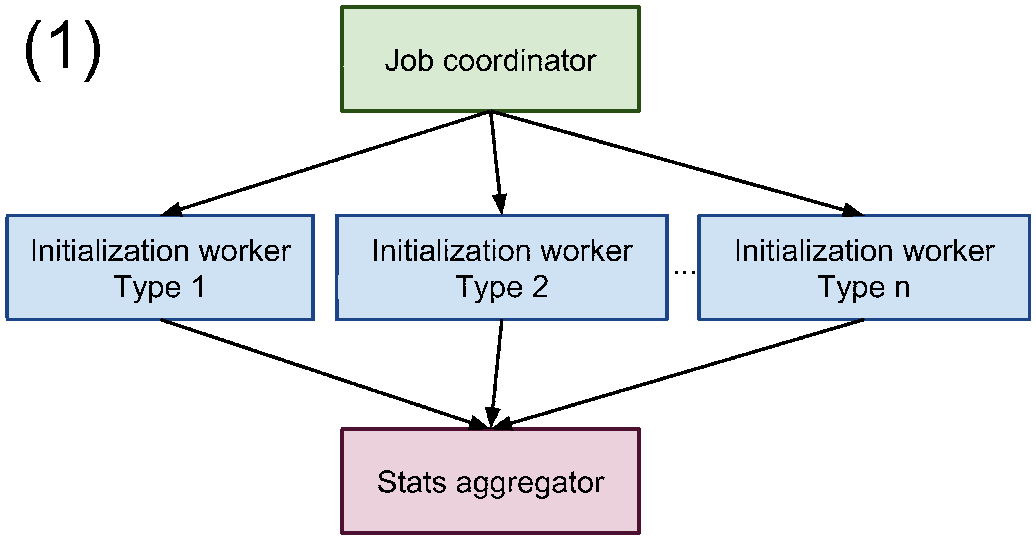
Phase (1):

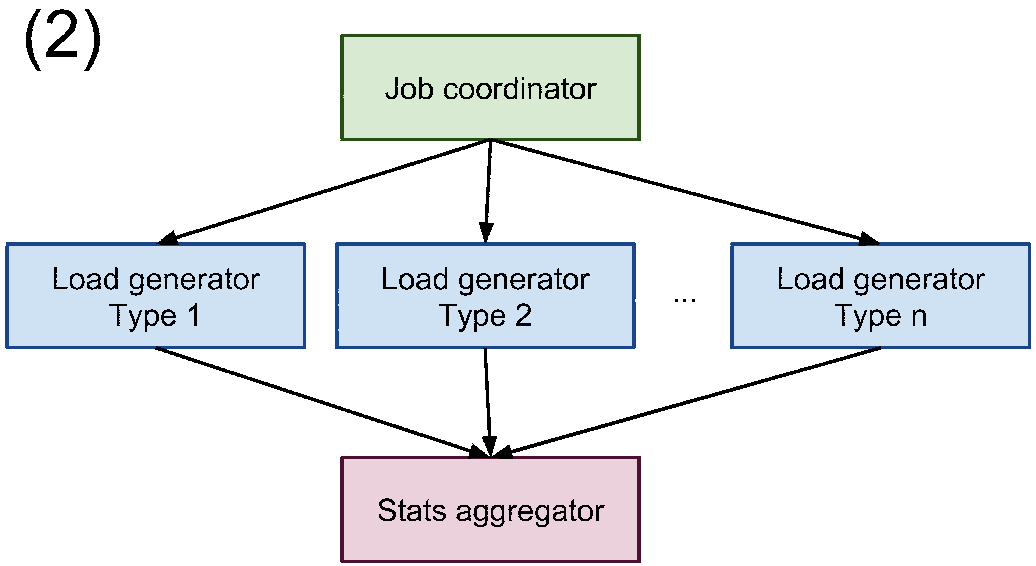
Create pre-existing namespace

Phase (2):

Issue model-based workload

- Multi-threaded producer-consumer approach





Scaling workloads

- Multi-dimensional workload scaling across dimensions:
  - Number of files
  - Number of files of a particular type
    - By modifying the weights of the file types
  - Time (interarrivals)
    - Can be accelerated or decelerated
  - Active span (of the files)
    - Increases/decreases the access counts of the files

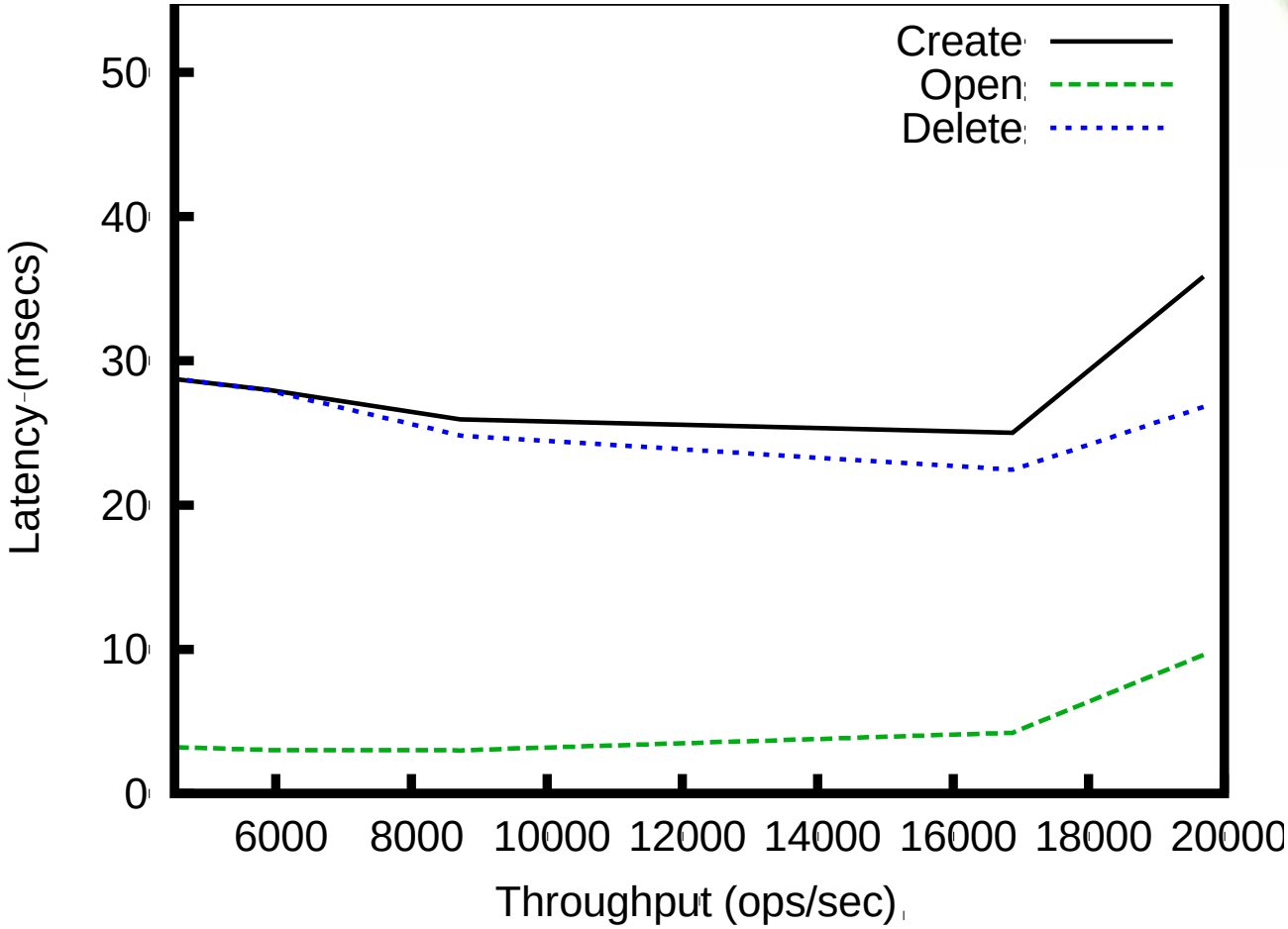
Using MimesisBench to evaluate HDFS performance



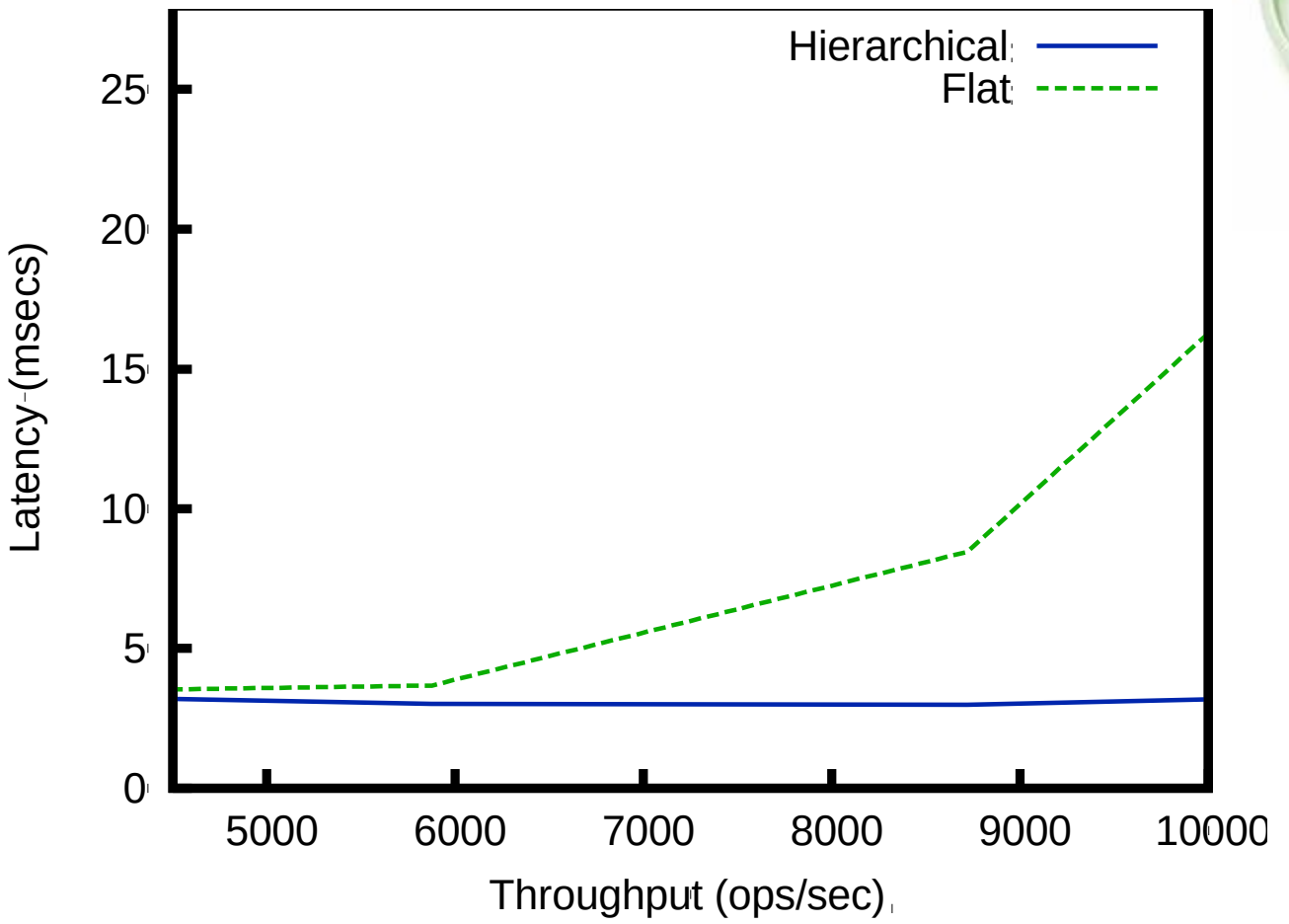
- Latency of opens, creates and deletes
- Throughput
- Flat vs. hierarchical namespaces
- Isolating workloads
- Evaluating an improvement (NN locking)

- k = 30
- Popularity CDF difference < 4%

Operation latency



Flat vs. hierarchical namespaces



Evaluating an improvement: HDFS-5239



- Allow FSNamesystem lock fairness to be configurable
- Throughput improvement:

	Fair lock	Unfair lock	Improvement
Hierarchical	19 696 ops/sec	21 899 ops/sec	11%
Flat	10 255 ops/sec	11 299 ops/sec	10%
Read-mostly	53 324 ops/sec	103 331 ops/sec	94%
Write-heavy	8 694 ops/sec	8 701 ops/sec	1%

MimesisBench summary

- Model-based approach for realistic workload generation
- Can be used with any storage system that is able to interface with Hadoop (HDFS, Lustre, etc.)
- Multi-dimensional workload scaling
- Autonomic type-aware model allows sensitivity and ‘what-if’ analysis
- Stable results (very small coefficient of variation)