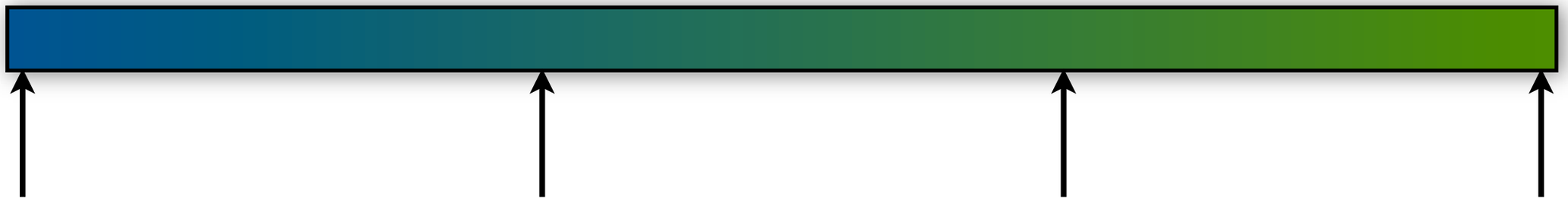
RAMCloud: multiRead

Ankita Kejriwal

The Big Question

What is the ideal distribution of work between the client and the server?



Single Read/Write

Multi Operations
e.g., multiRead

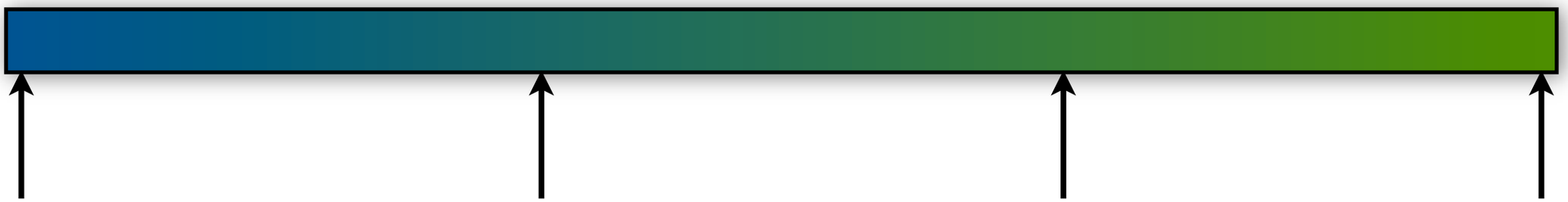
Fixed Advanced Operations
e.g., sql-like ops: avg, sum, min

Upload Code

Design Considerations

- Low level operations (Read-Write):
  - Arbitrary computations are possible
  - Simplifies server implementation
- Higher level operations:
  - Fewer RPCs exchanged => lower latency
  - Less network bandwidth
- Issues with higher level operations:
  - Need data locality
  - Potential security implications (with uploading code)

The Big Question



Single Read/Write

Multi Operations
e.g., multiRead

Fixed Advanced Operations
e.g., sql-like ops: avg, sum, min

Upload Code

Current State

This Project

Motivation

- Batch multiple requests to a single master in one multiRead RPC
- Parallelize requests to multiple masters

multiRead API

Array of
{table Id, object Id}
for each object
to be read

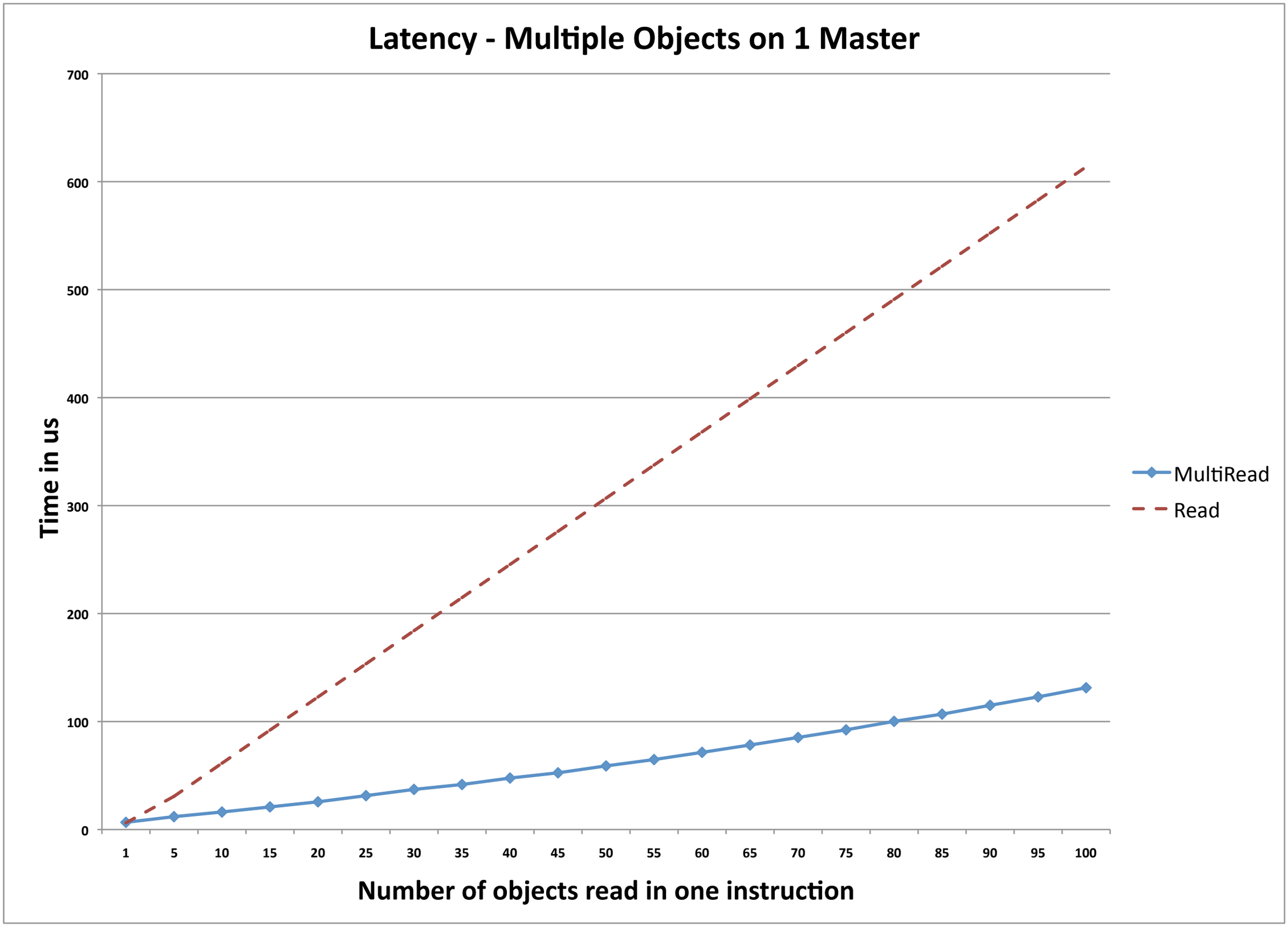
For each object:
{value, version,
status}

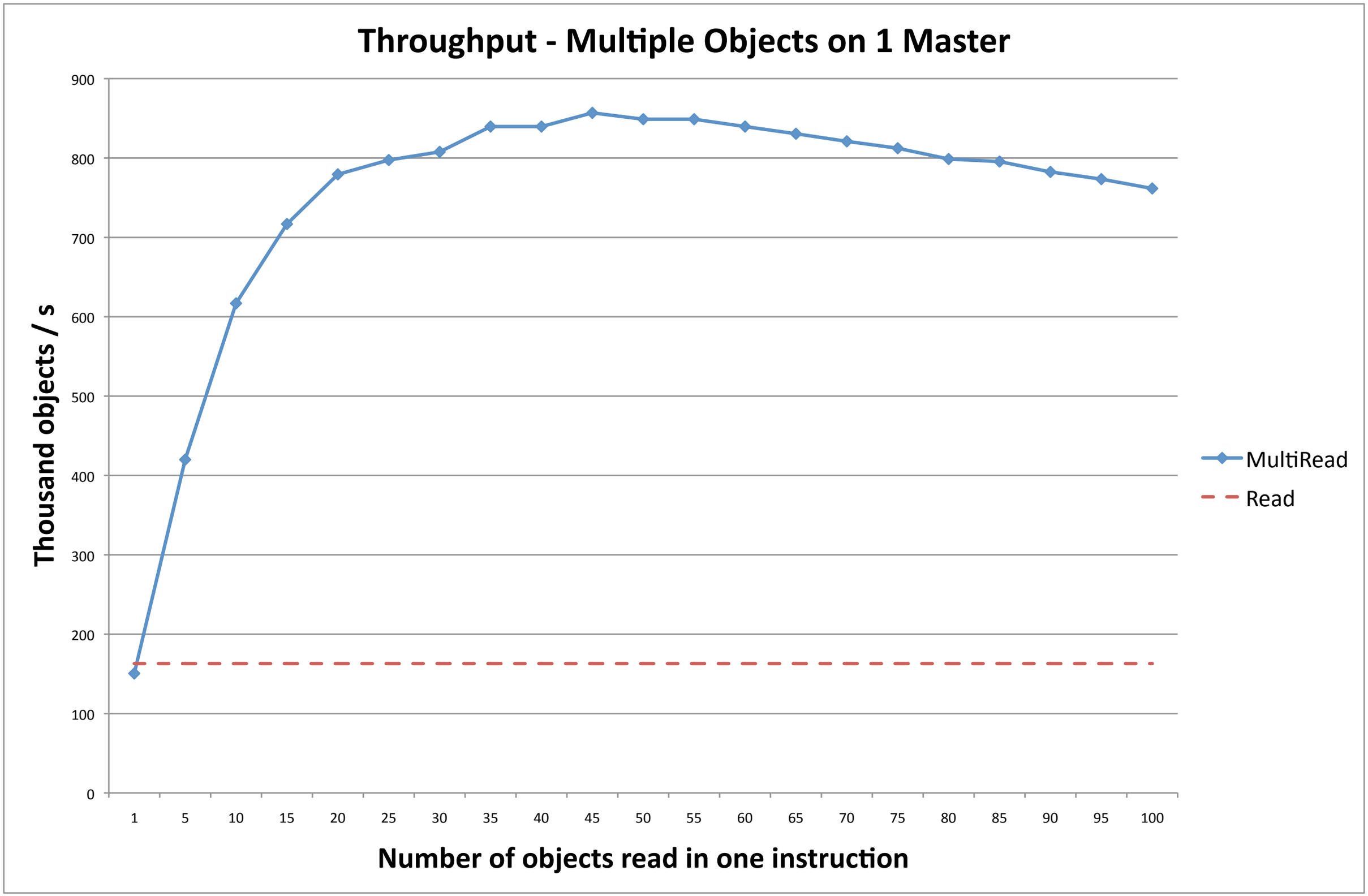
Interesting Issues in Implementation

- Error handling
  - Individual read requests can fail
- RPC Packet Overflow
  - Possible Solution: Client sends another request for remaining objects

Benchmarking

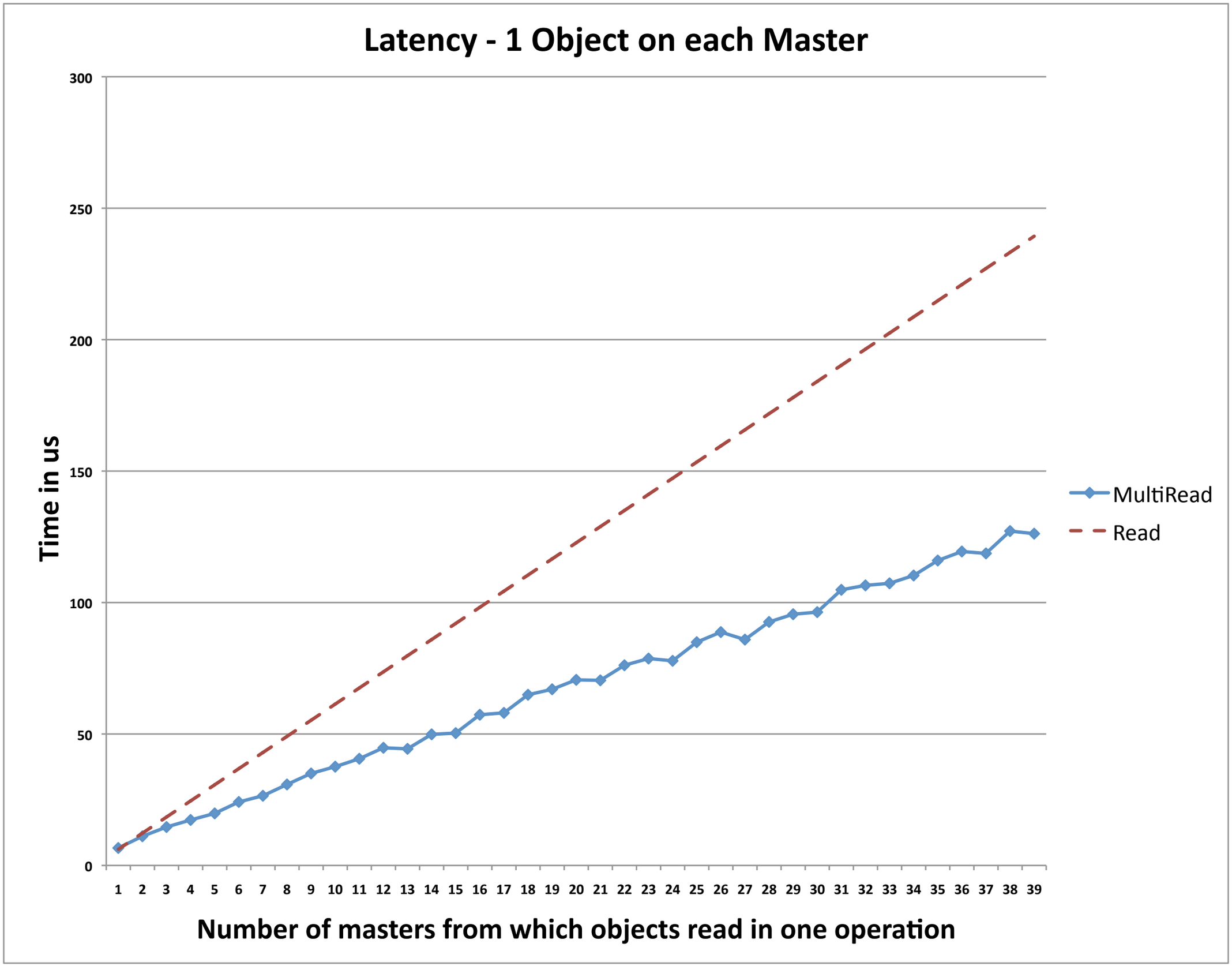
- multiRead and read
- one table, one master
- multiple 100B objects
- random object Ids

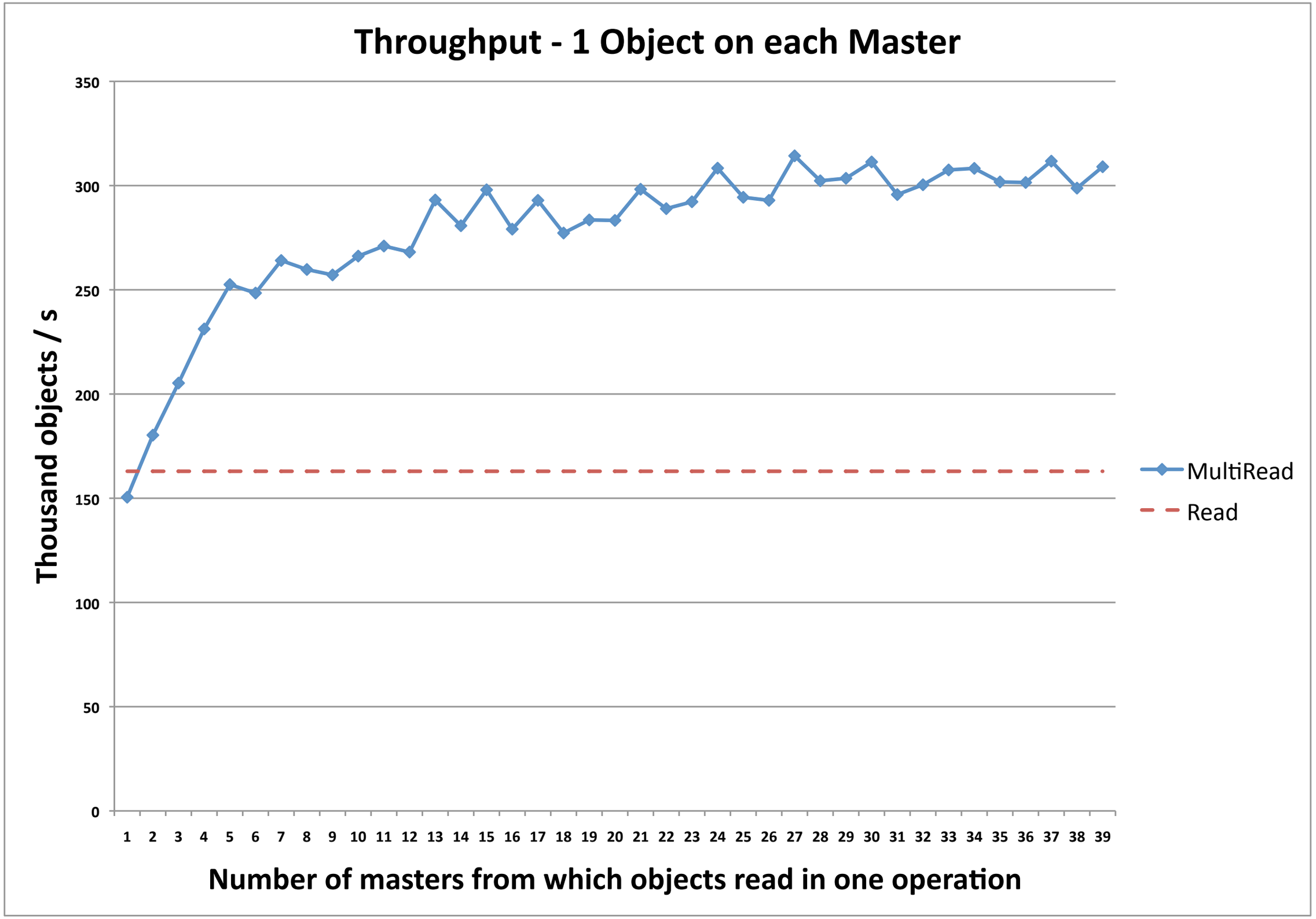

Latency - Multiple Objects on 1 Master
Time in us
700
600
500
400
300
200
100
0
1
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
Number of objects read in one instruction
MultiRead
Read


Throughput - Multiple Objects on 1 Master
Thousand objects / s
900
800
700
600
500
400
300
200
100
0
1
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
Number of objects read in one instruction
MultiRead
Read

Benchmarking

- multiRead and read
- multiple tables, multiple masters
- one 100B object on each master


Latency - 1 Object on each Master
Time in us
300
250
200
150
100
50
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39
Number of masters from which objects read in one operation
MultiRead
Read


Throughput - 1 Object on each Master
Thousand objects / s
350
300
250
200
150
100
50
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39
Number of masters from which objects read in one operation
MultiRead
Read

Summary & Future Work

- Summary:
  - Batched requests to 1 server: up to 5.2x speedup over read
  - Parallel requests to multiple servers: up to 1.9x speedup over read

- Future Work:
  - MultiRead - Understanding what limits performance and whether it can be improved; RPC Overflow; More benchmarks; Optimizations
  - Design, Implementation and Benchmarking of MultiWrite and other advanced operations

Questions?