Team-based Message Logging: Preliminary Results

Esteban Meneses, Celso L. Mendes and Laxmikant V. Kalé

Parallel Programming Laboratory
University of Illinois at Urbana-Champaign

Contents

- Message Logging
- Team-based Approach
- Experimental Results
- Conclusions

1253 separate node crashes on Jaguar during 537 days (Aug-22-2008 to Feb-10-2010)

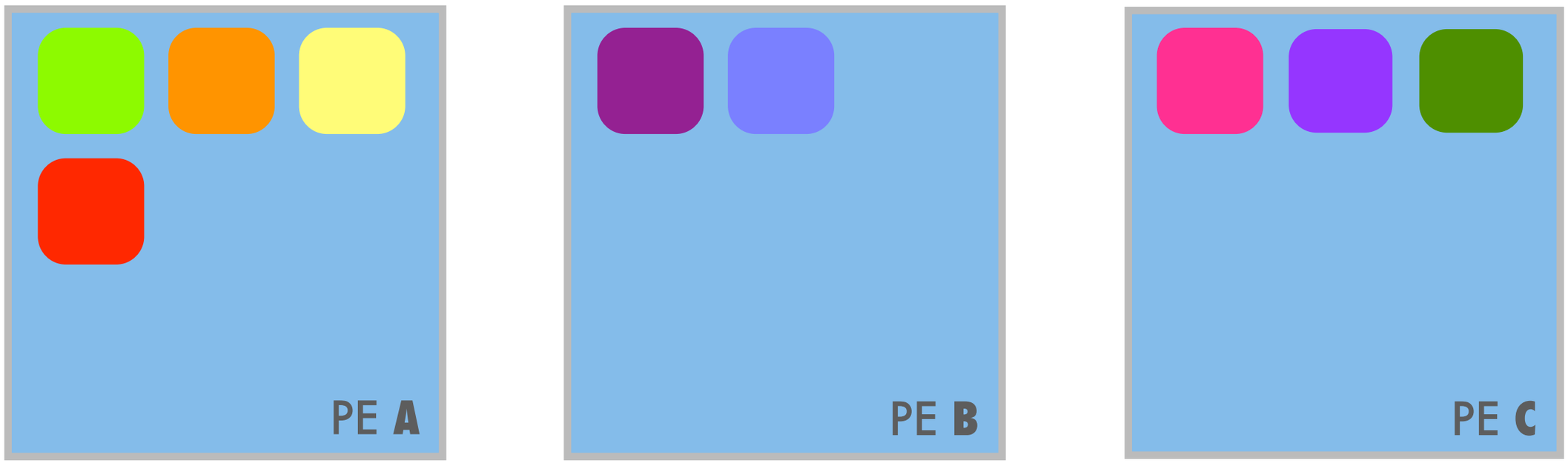
2.33 failures per day

Sequoia will have 1.6 million cores and an exascale machine around 100 million cores...

We will see failures all the time

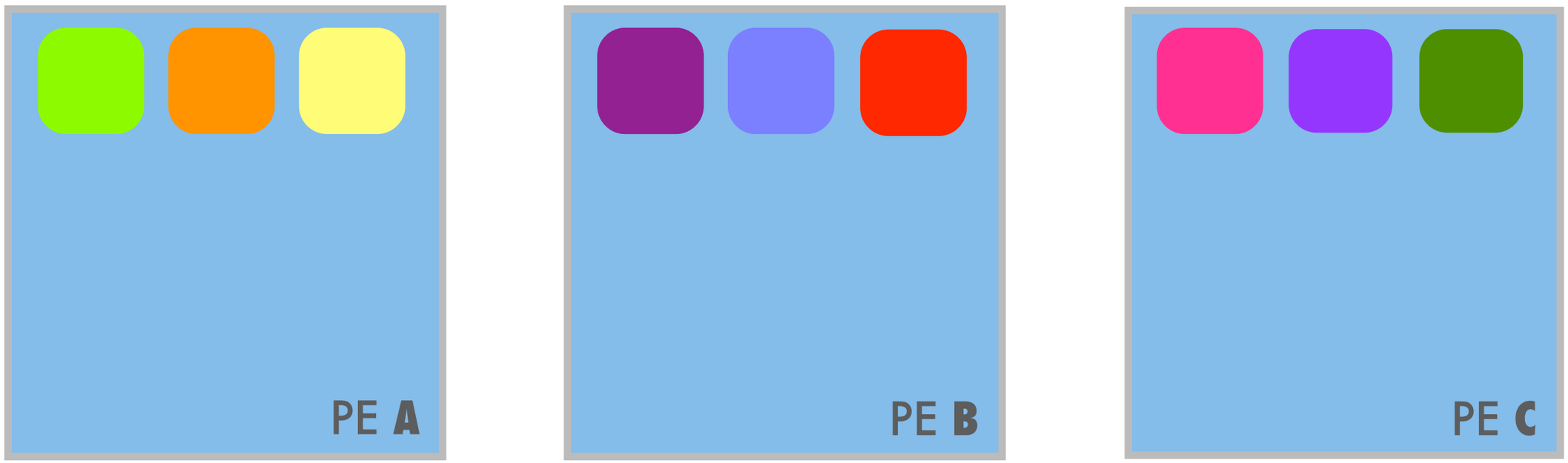
Charm++

- Object-based over-decomposition.
- An intelligent RTS assigns objects to processors.



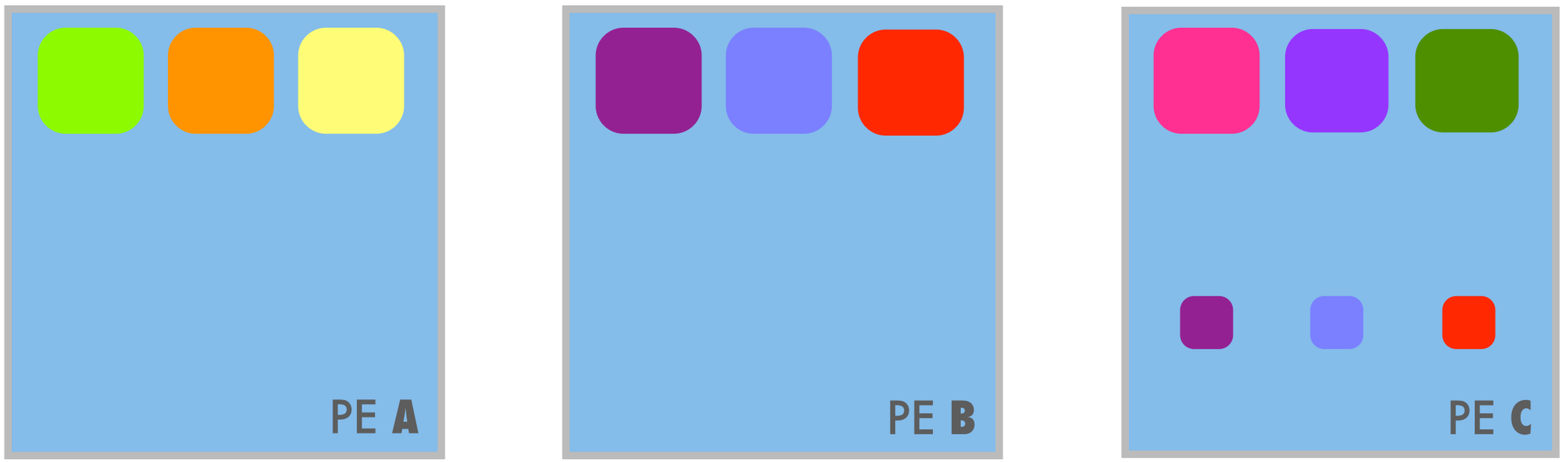
Charm++

- Object-based over-decomposition.
- An intelligent RTS assigns objects to processors.



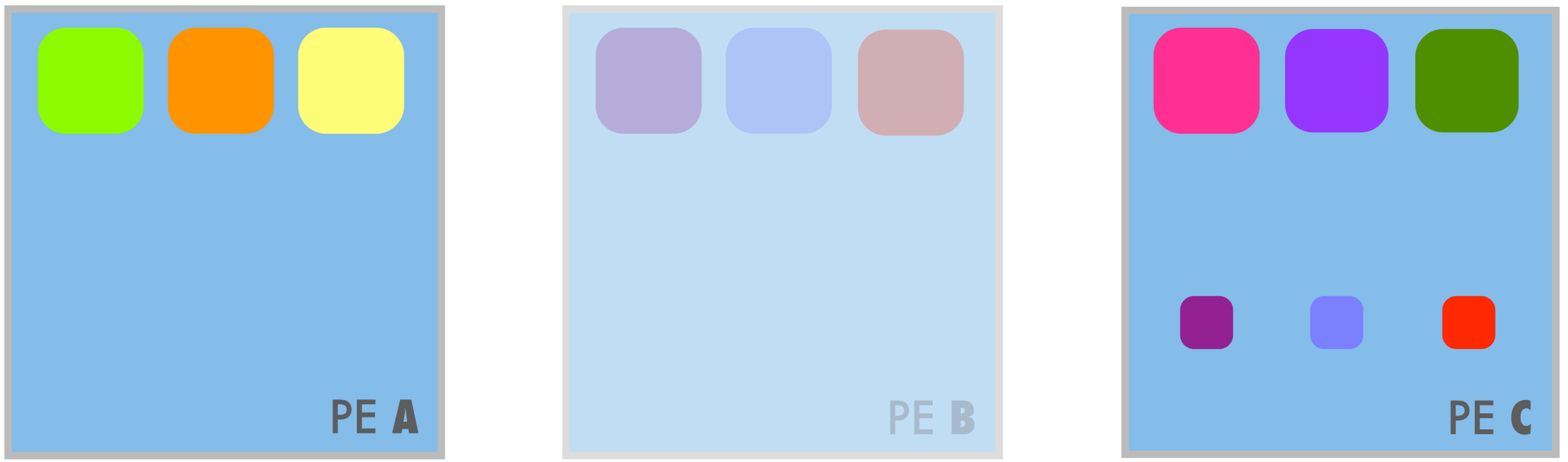
Charm++

- Object-based over-decomposition.
- An intelligent RTS assigns objects to processors.

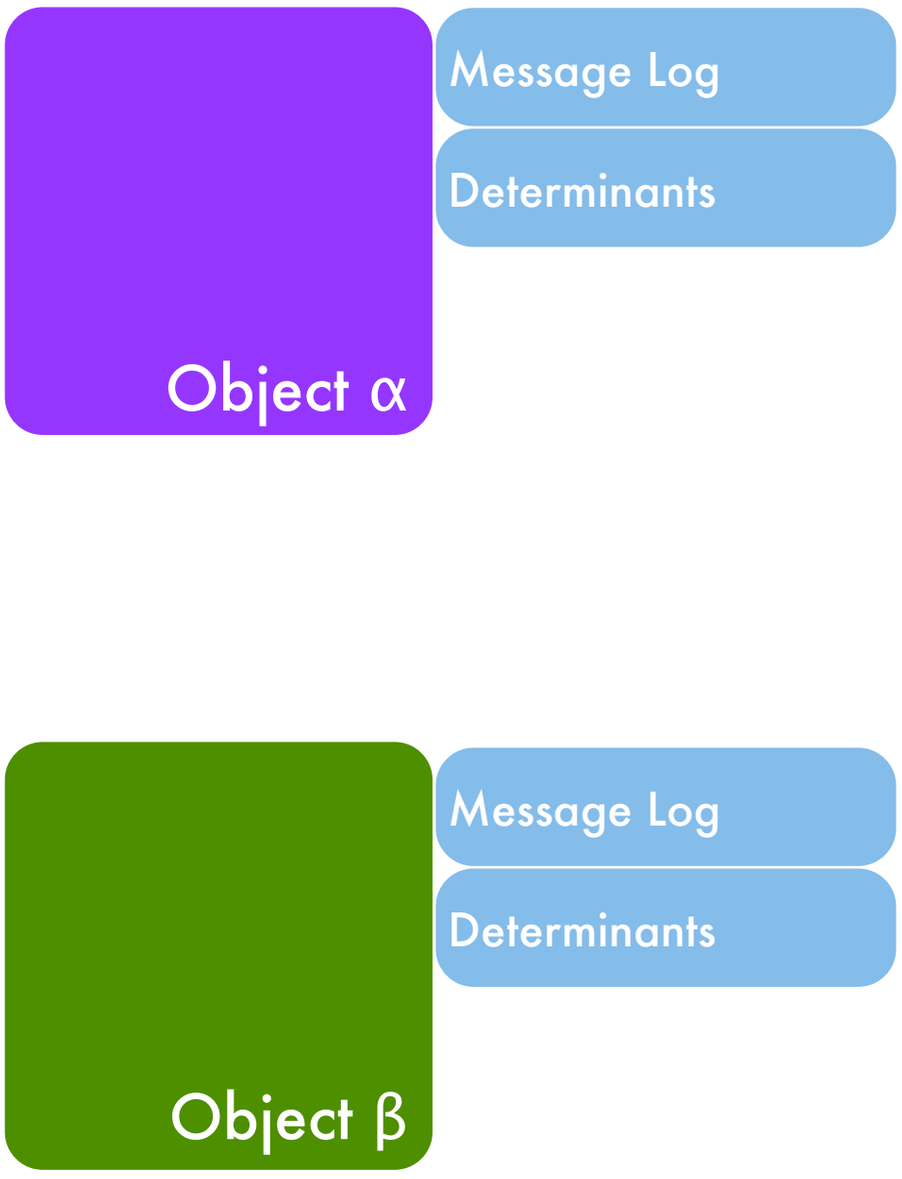


Charm++

- Object-based over-decomposition.
- An intelligent RTS assigns objects to processors.

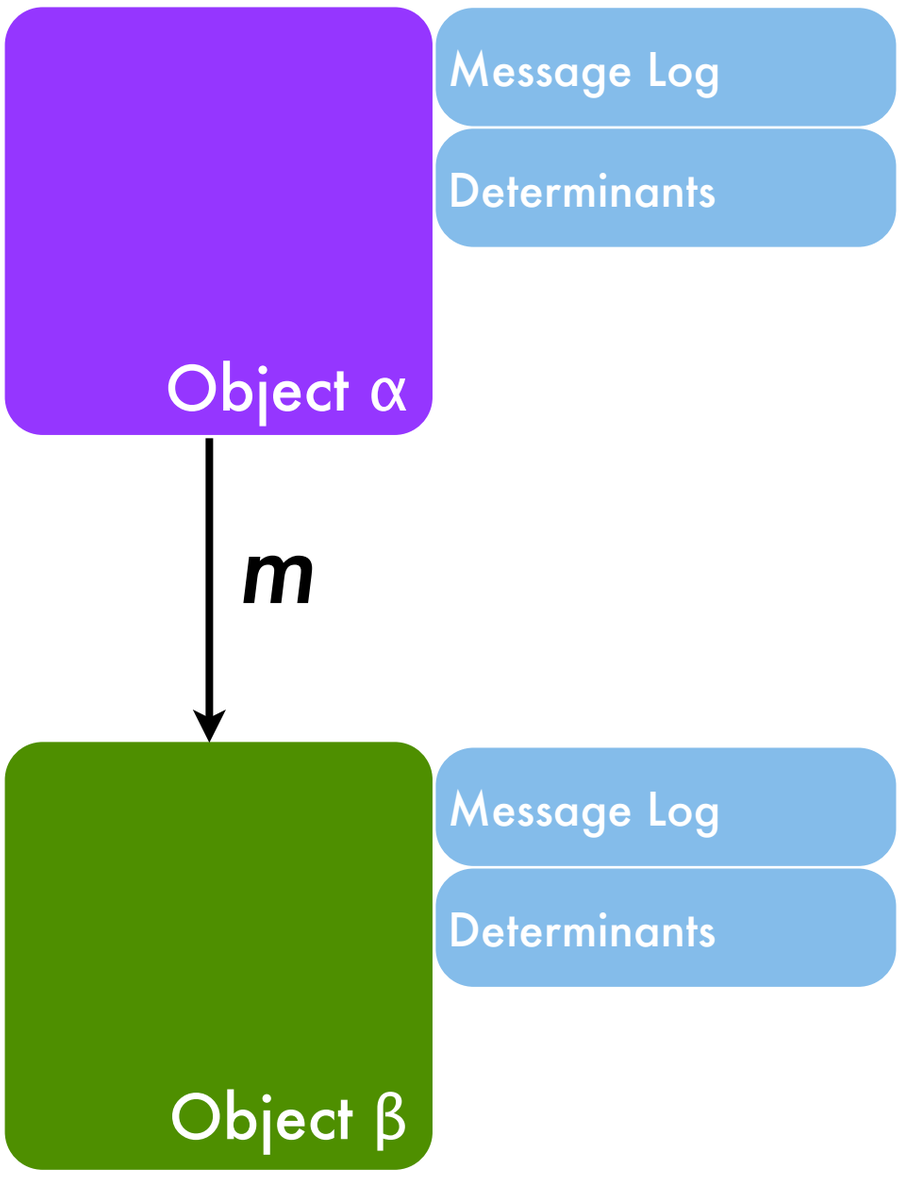


Message Logging



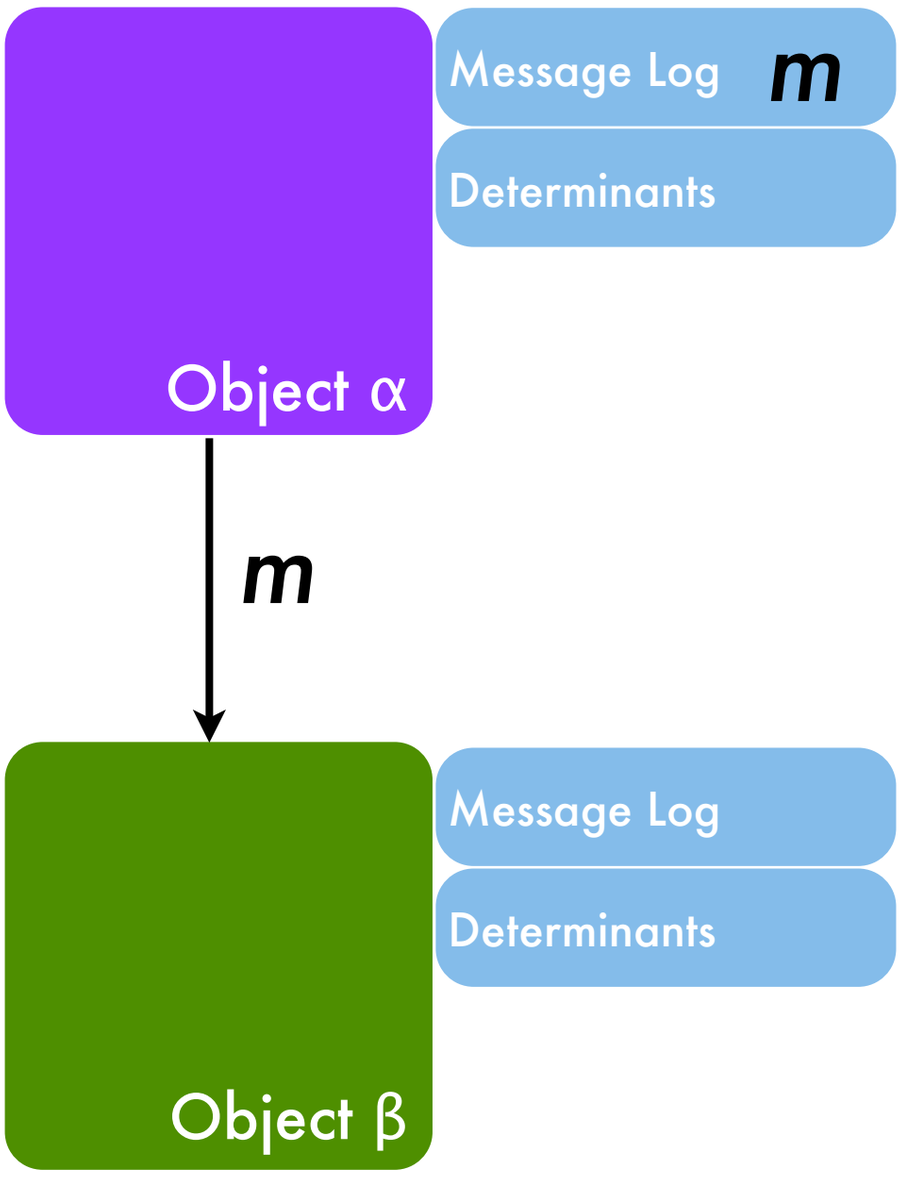
- Sender-based.
- Piecewise Deterministic.
- Protocols: Pessimistic and Causal.

Message Logging



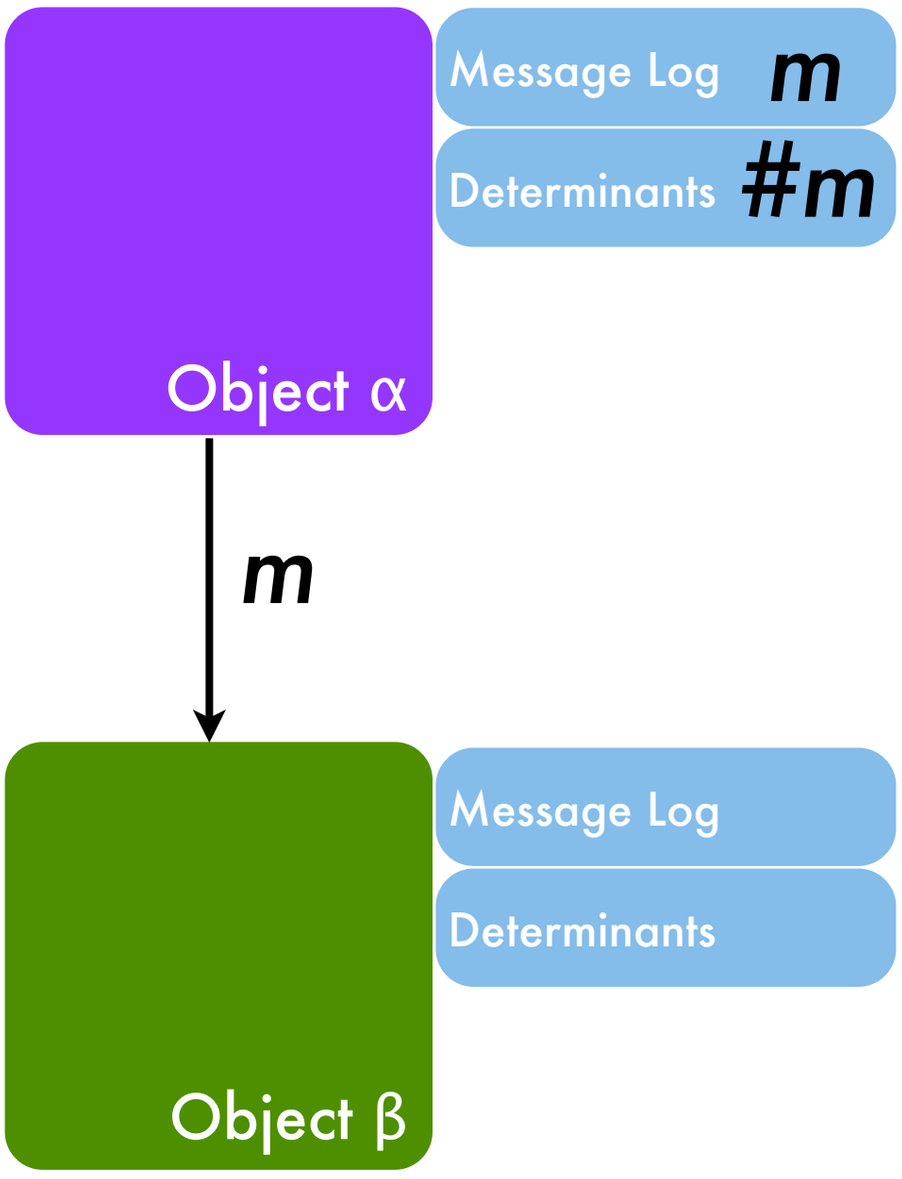
- Sender-based.
- Piecewise Deterministic.
- Protocols: Pessimistic and Causal.

Message Logging



- Sender-based.
- Piecewise Deterministic.
- Protocols: Pessimistic and Causal.

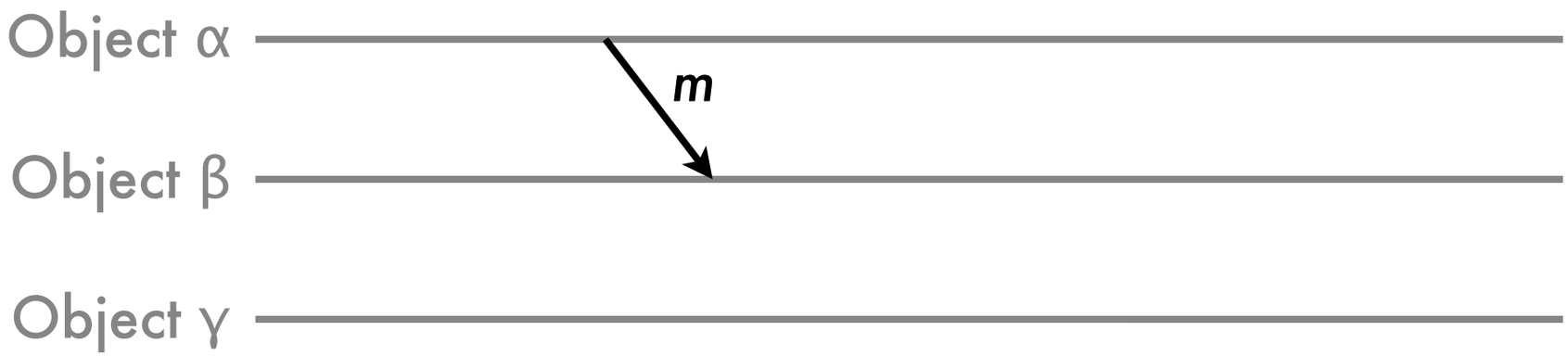
Message Logging



- Sender-based.
- Piecewise Deterministic.
- Protocols: Pessimistic and Causal.

Message Logging Protocols

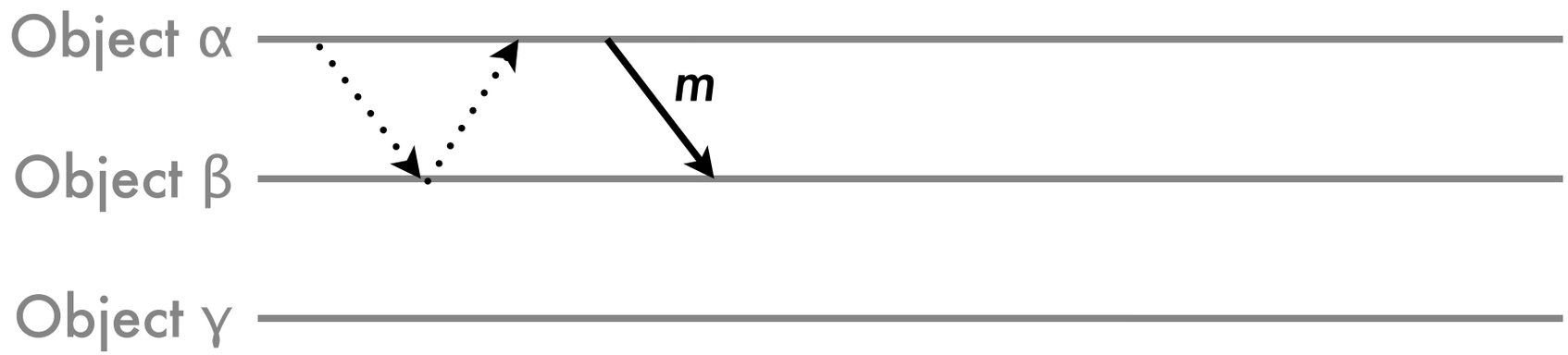
Pessimistic



Causal

Message Logging Protocols

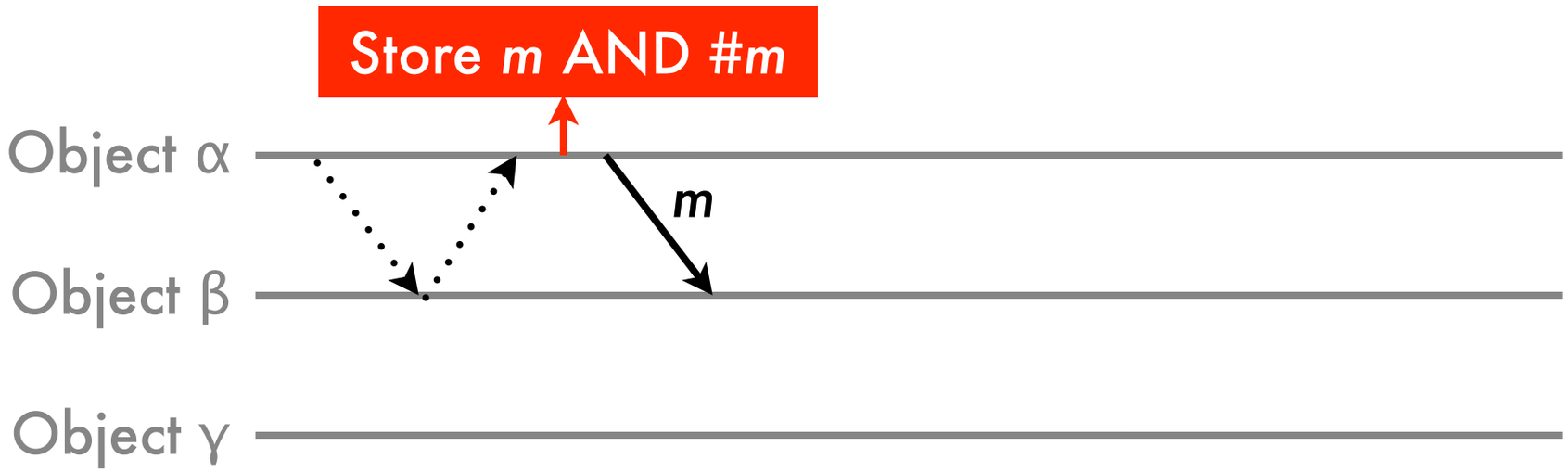
Pessimistic



Causal

Message Logging Protocols

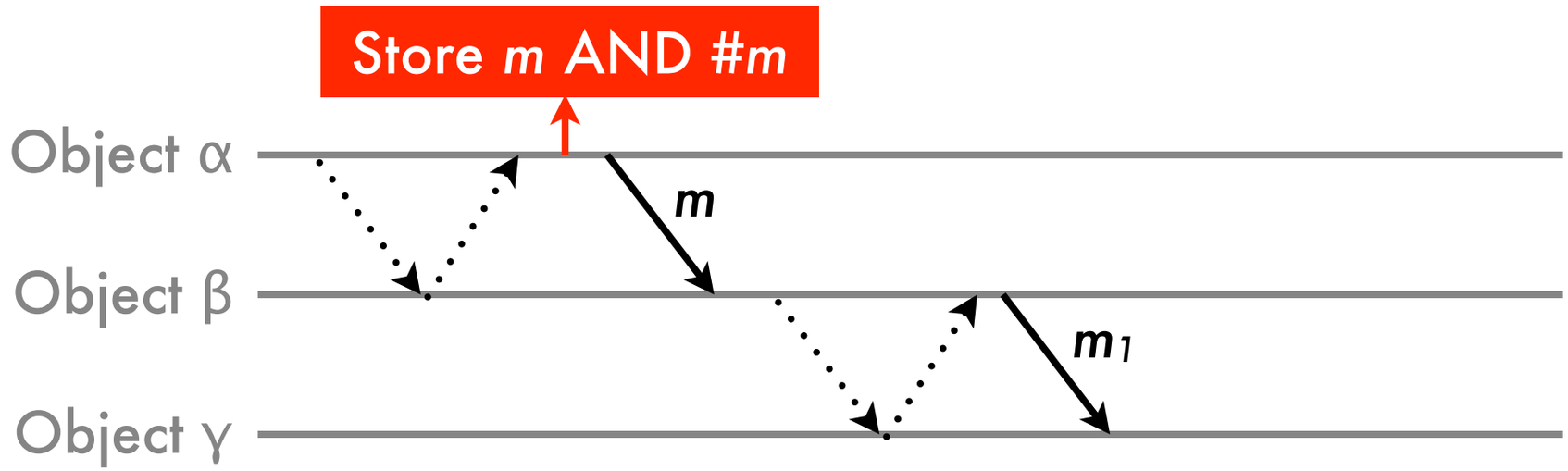
Pessimistic



Causal

Message Logging Protocols

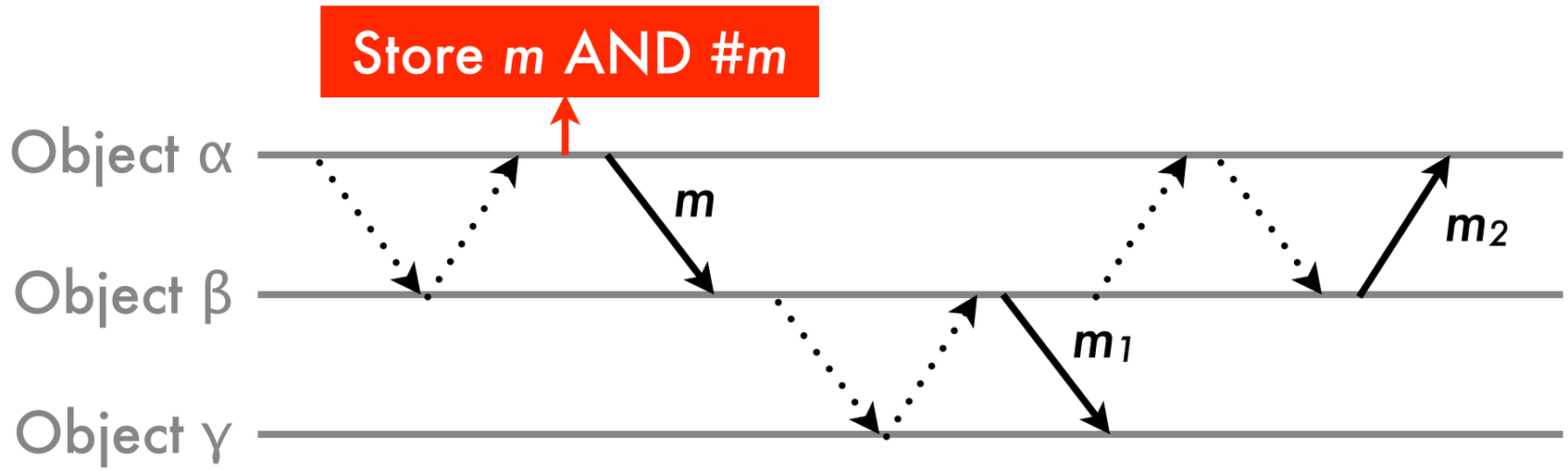
Pessimistic



Causal

Message Logging Protocols

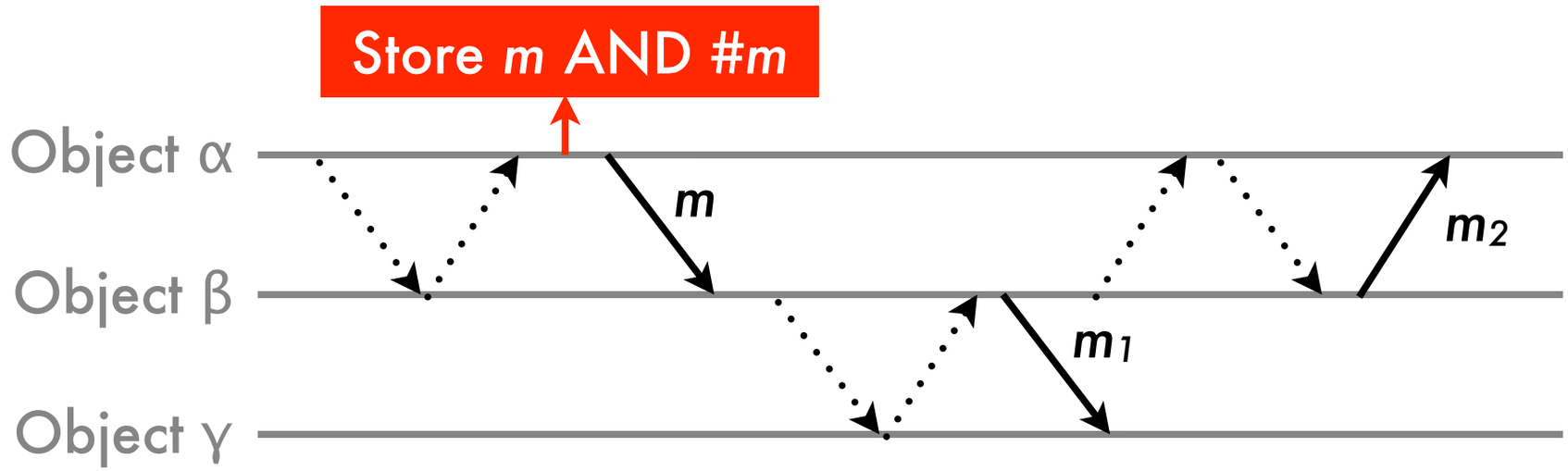
Pessimistic



Causal

Message Logging Protocols

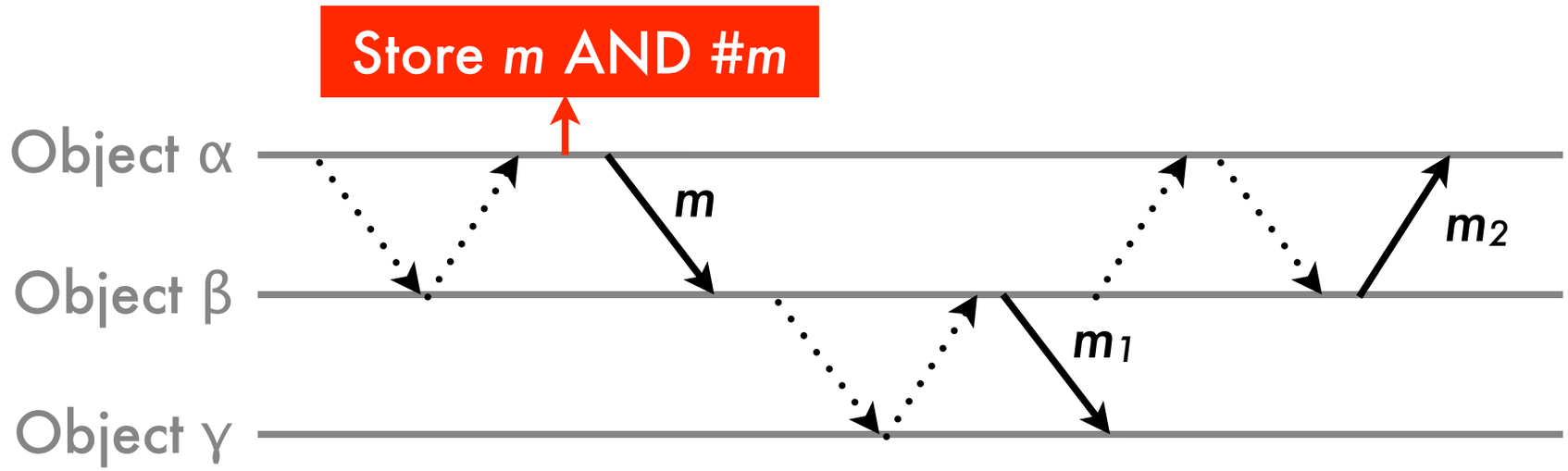
Pessimistic



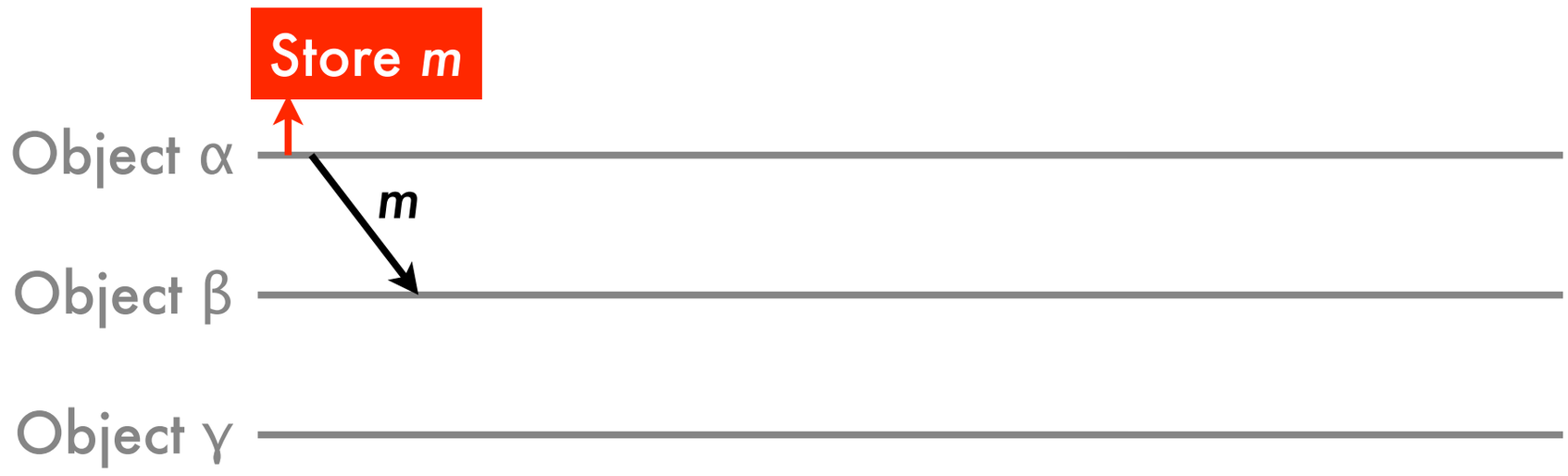
Causal

Message Logging Protocols

Pessimistic

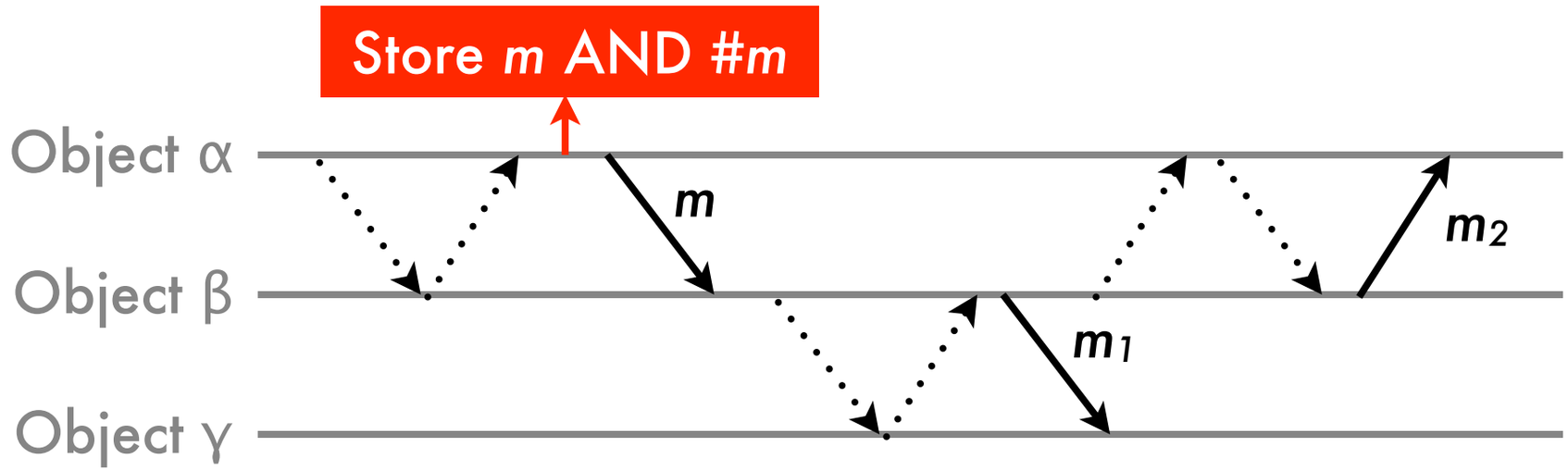


Causal

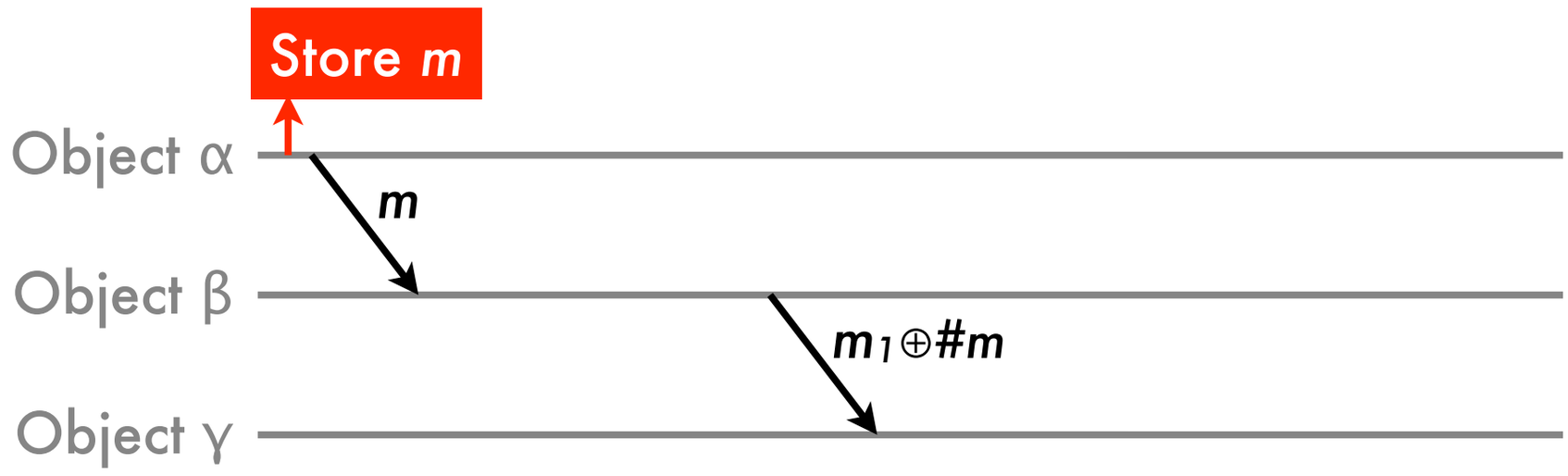


Message Logging Protocols

Pessimistic

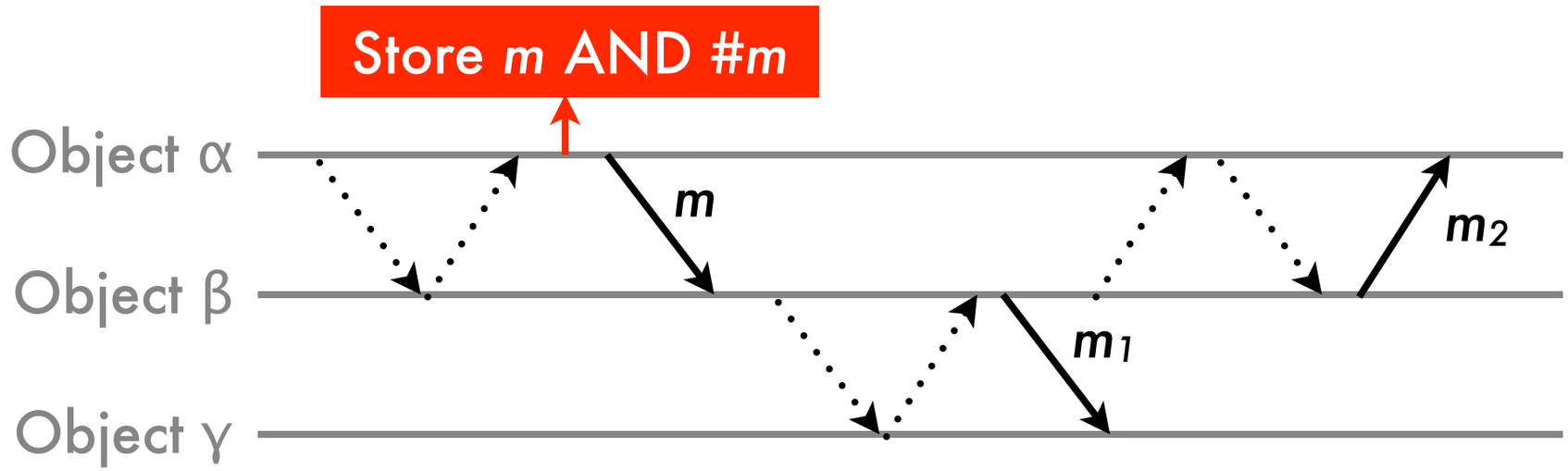


Causal

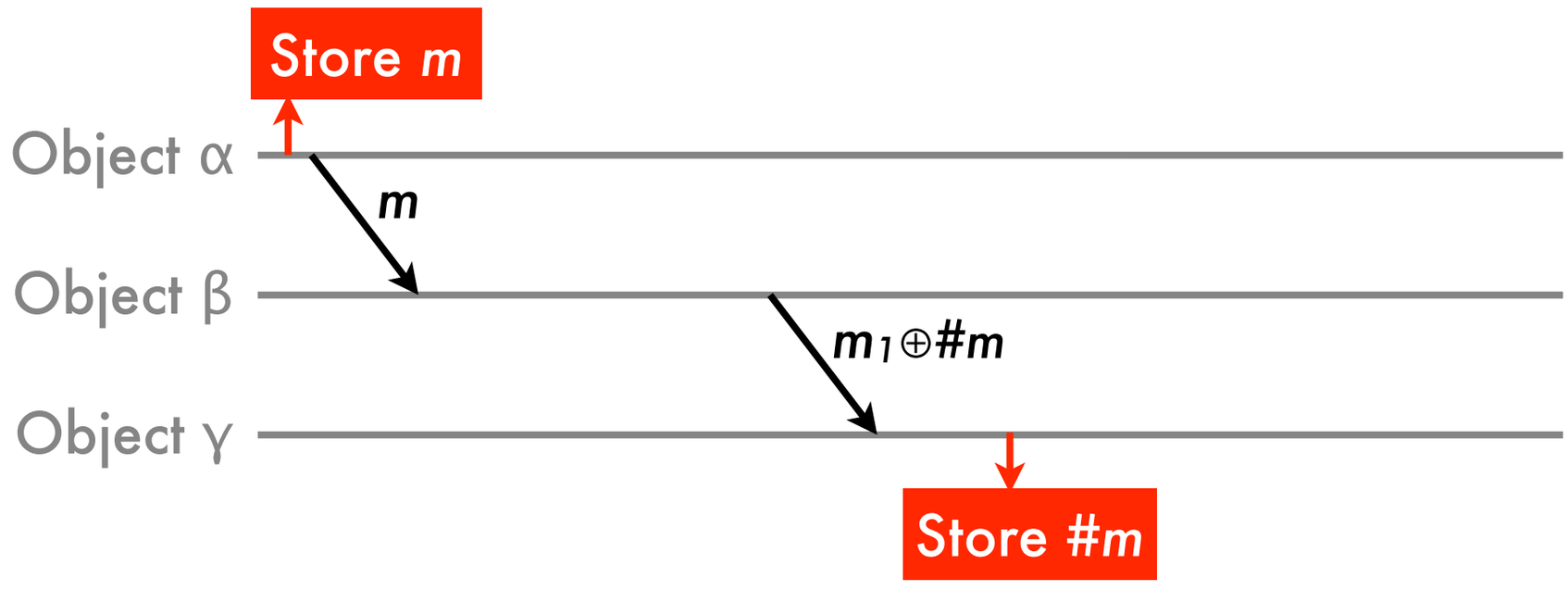


Message Logging Protocols

Pessimistic

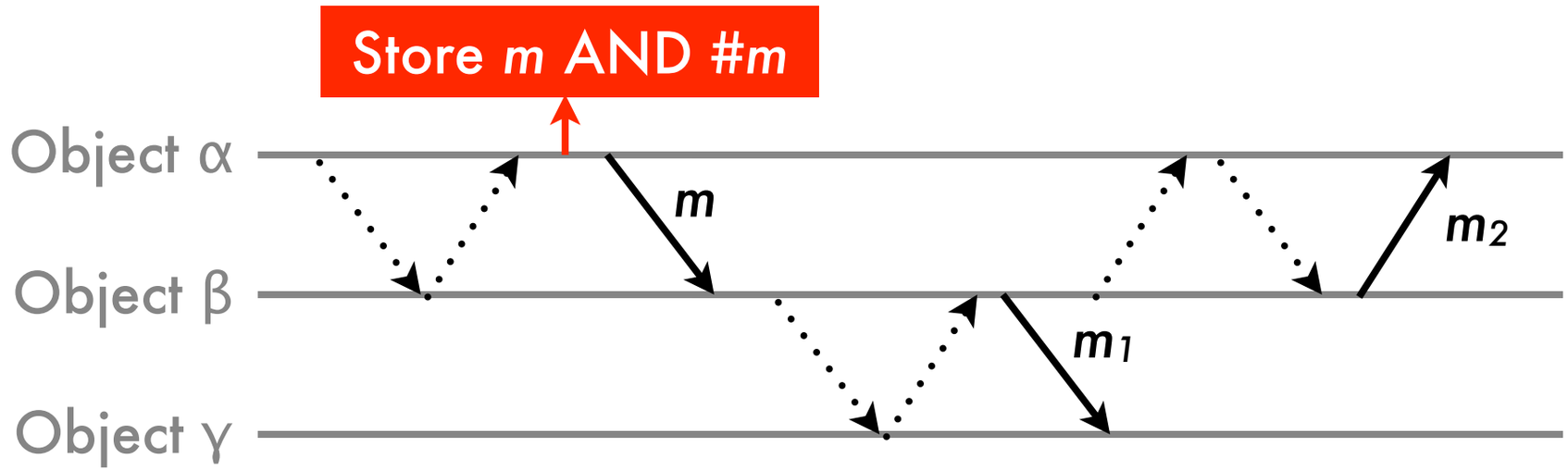


Causal

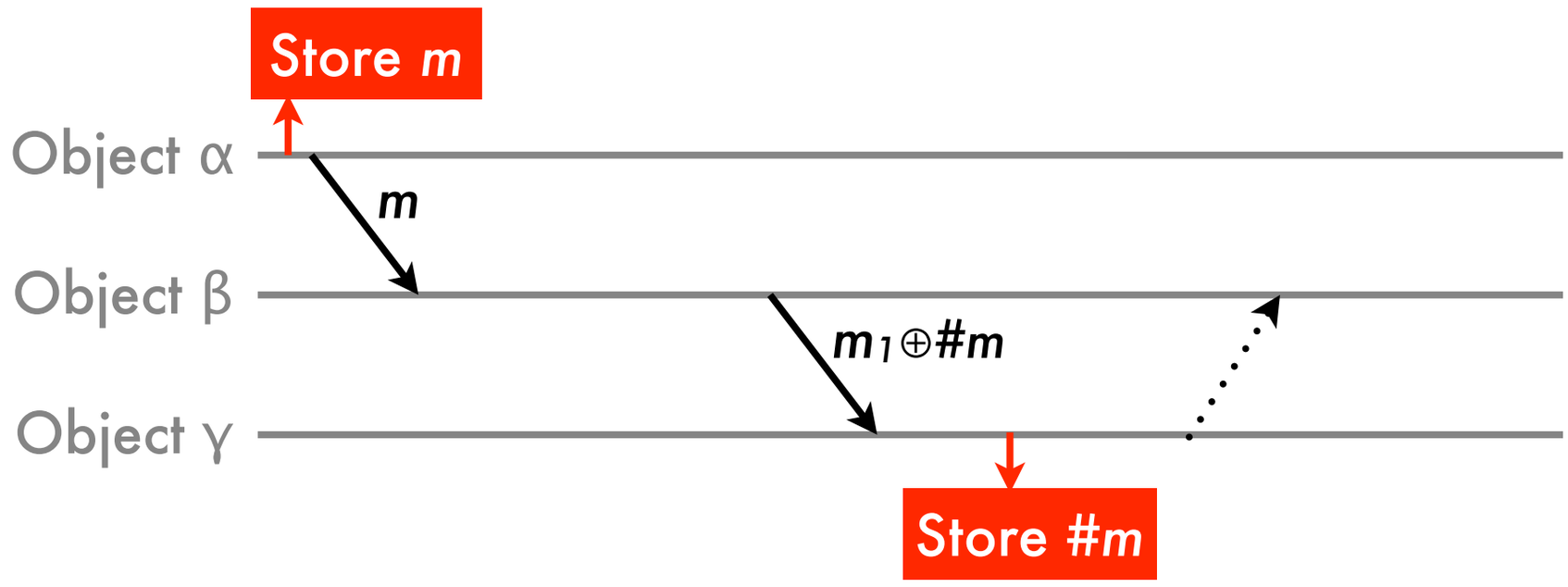


Message Logging Protocols

Pessimistic

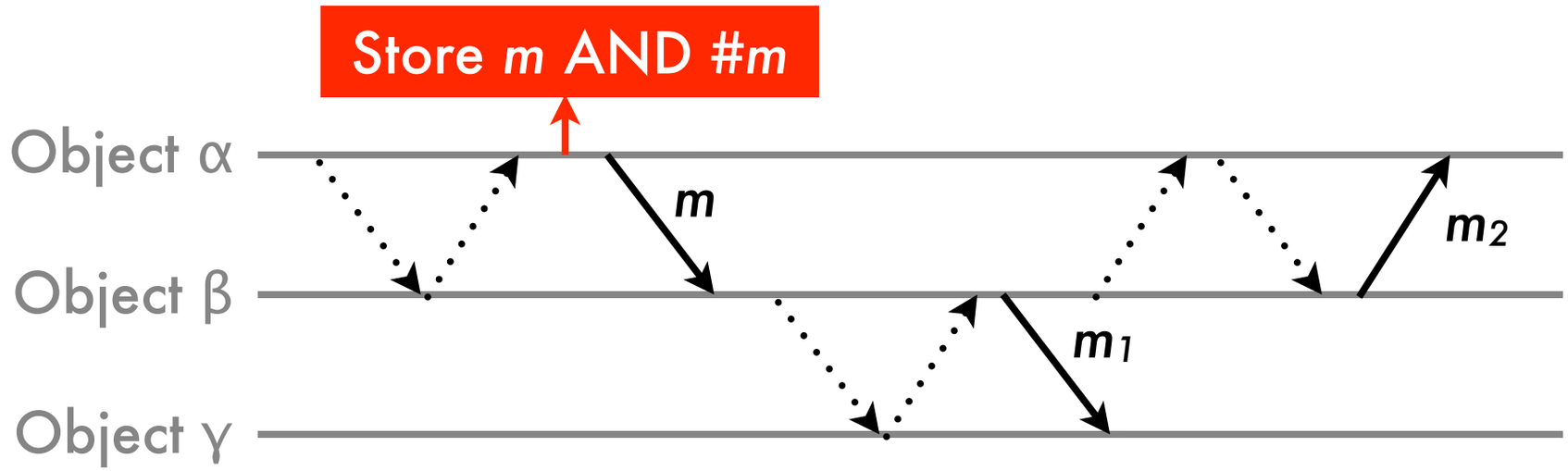


Causal

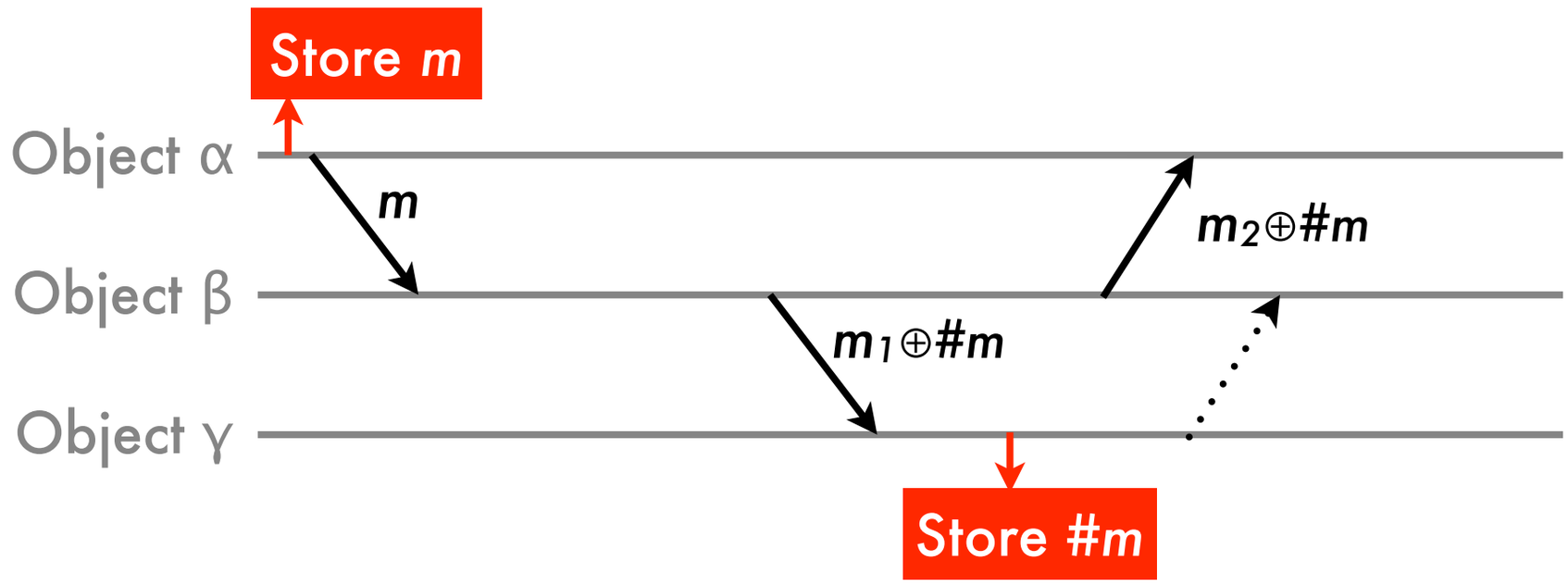


Message Logging Protocols

Pessimistic



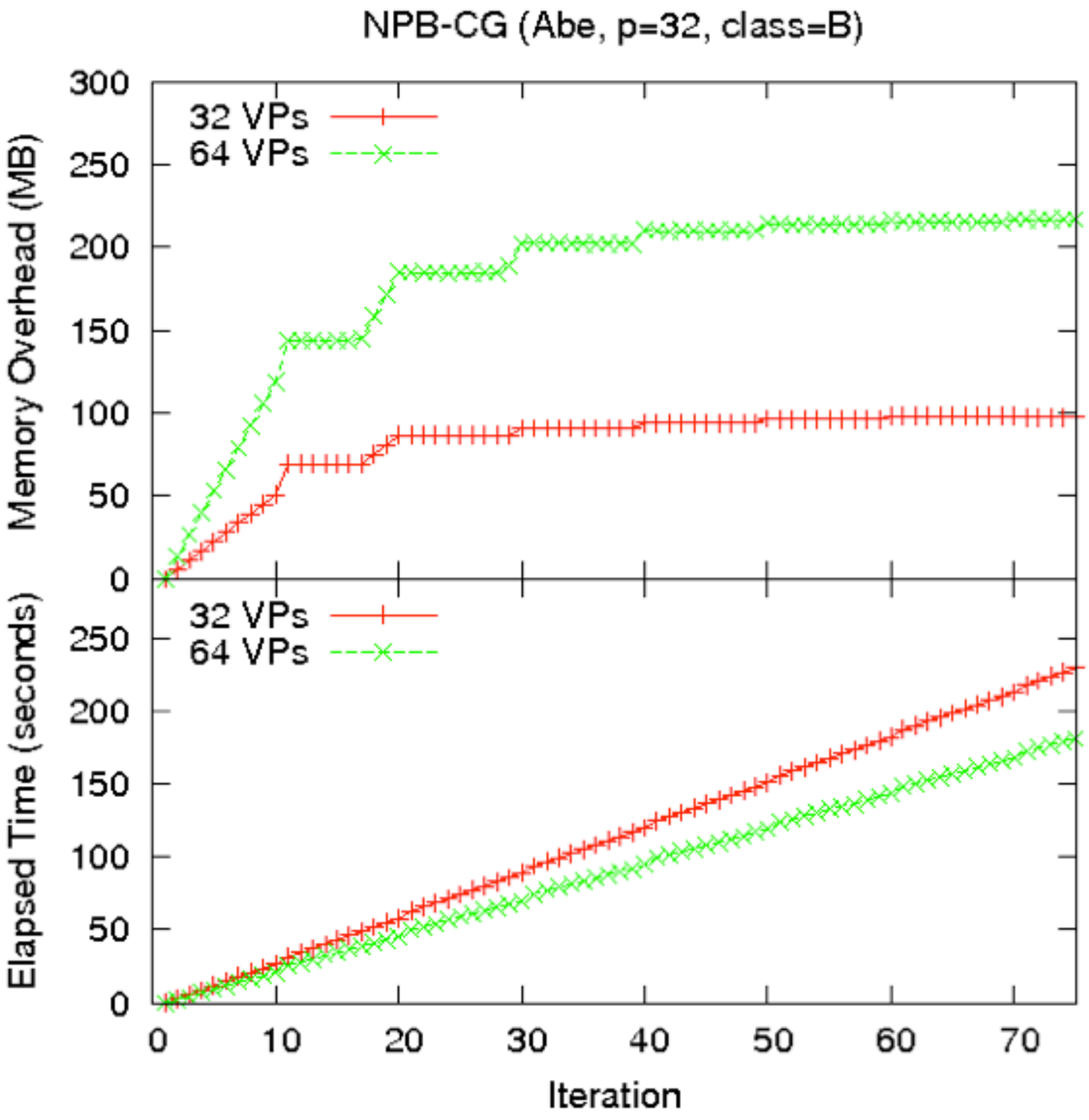
Causal



Virtualization

Resilience 2010, Melbourne, Australia

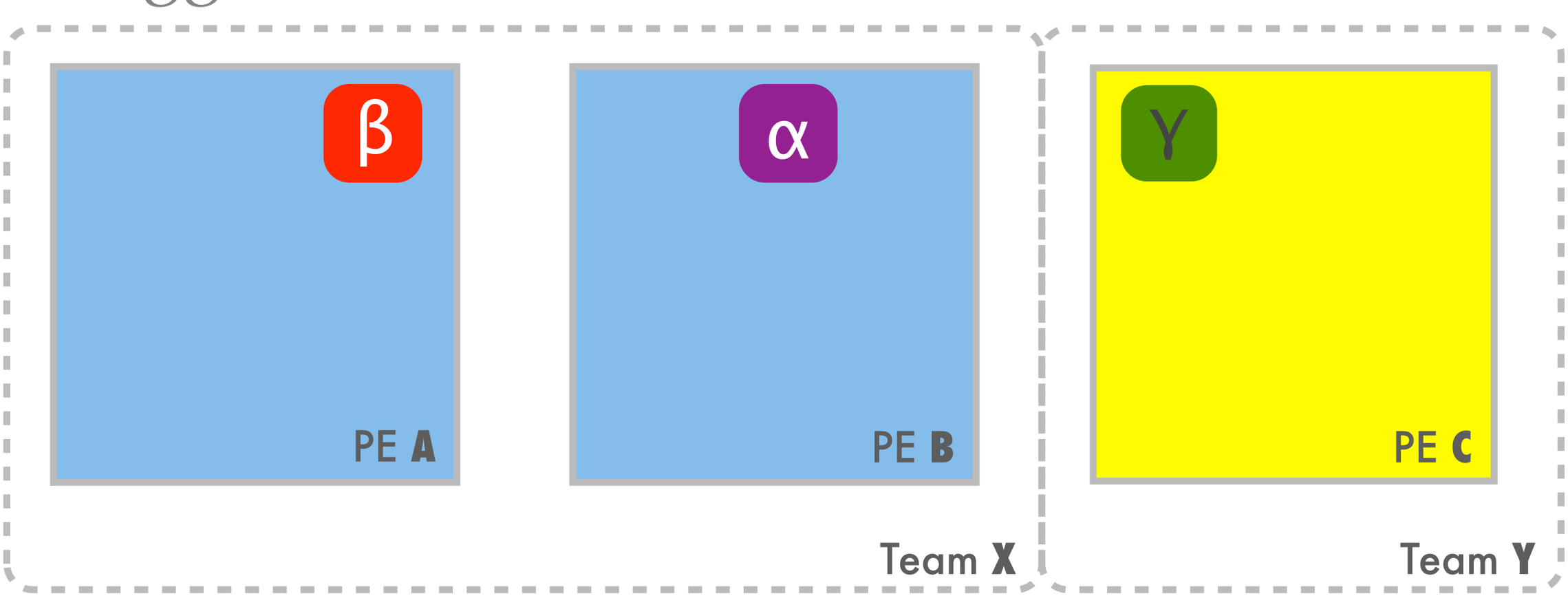
Virtualization



- Higher virtualization ratio:
- Hides latency overhead.
- Increases number of objects and messages.

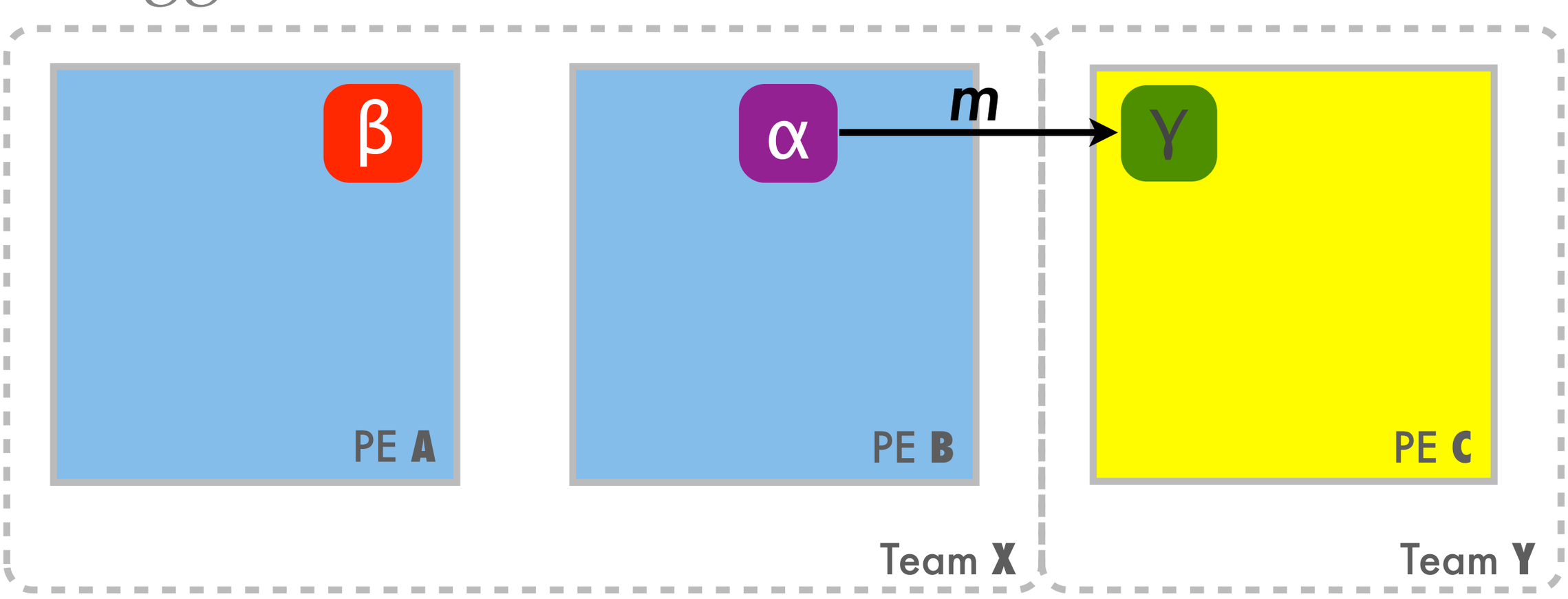
Team-based Approach

- Goal: reduce memory overhead of message log.
- Only messages crossing team boundaries are logged.



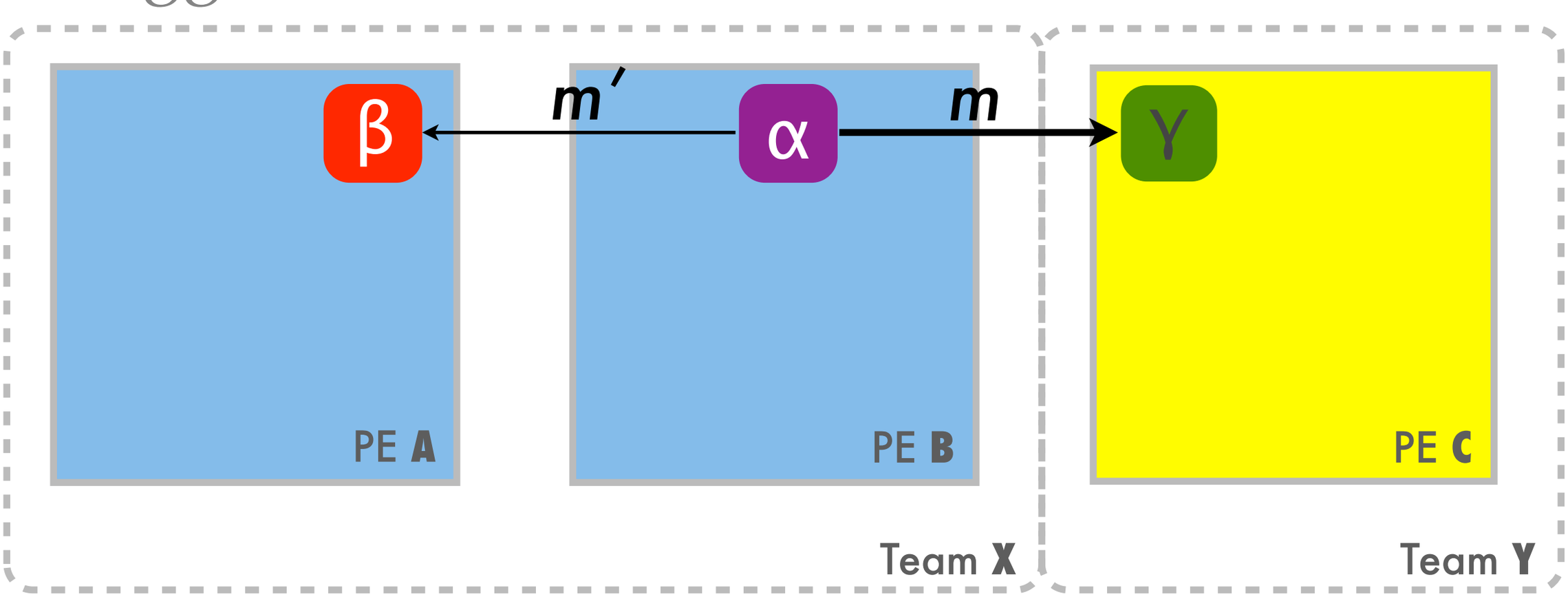
Team-based Approach

- Goal: reduce memory overhead of message log.
- Only messages crossing team boundaries are logged.



Team-based Approach

- Goal: reduce memory overhead of message log.
- Only messages crossing team boundaries are logged.



Processor Teams

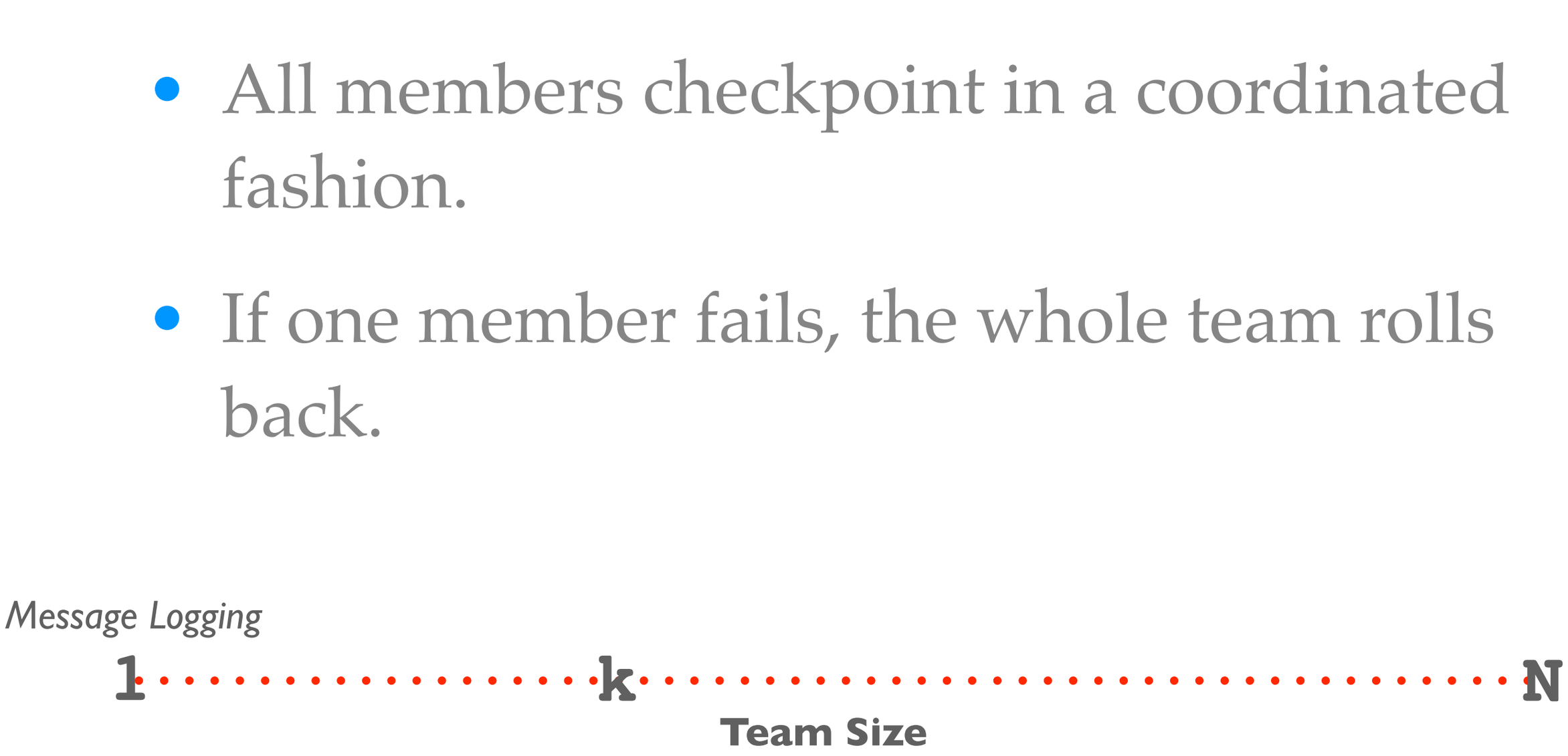
- Each team acts as a recovery unit:
  - All members checkpoint in a coordinated fashion.
  - If one member fails, the whole team rolls back.

Processor Teams

- Each team acts as a recovery unit:
  - All members checkpoint in a coordinated fashion.
  - If one member fails, the whole team rolls back.

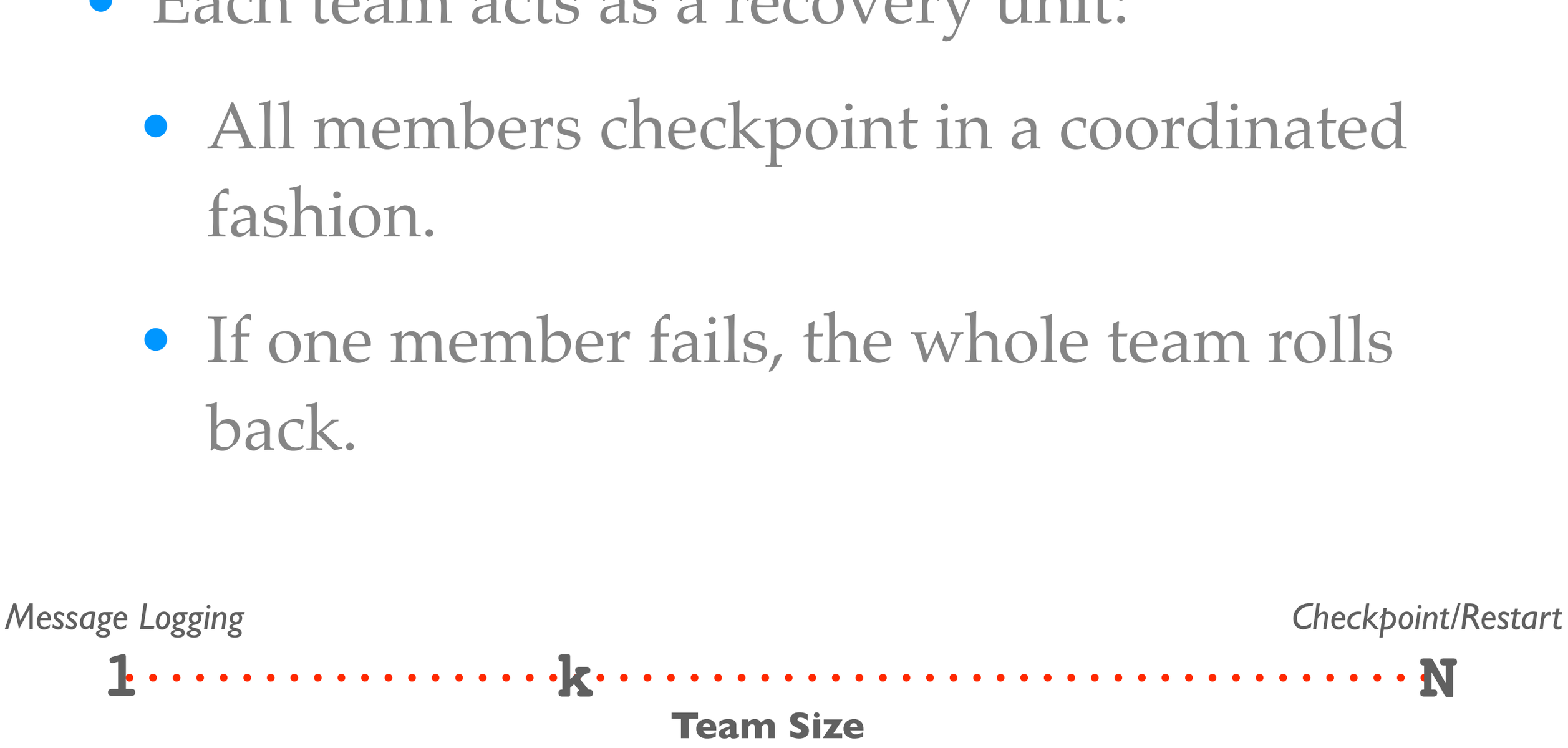
Processor Teams

- Each team acts as a recovery unit:
  - All members checkpoint in a coordinated fashion.
  - If one member fails, the whole team rolls back.



Processor Teams

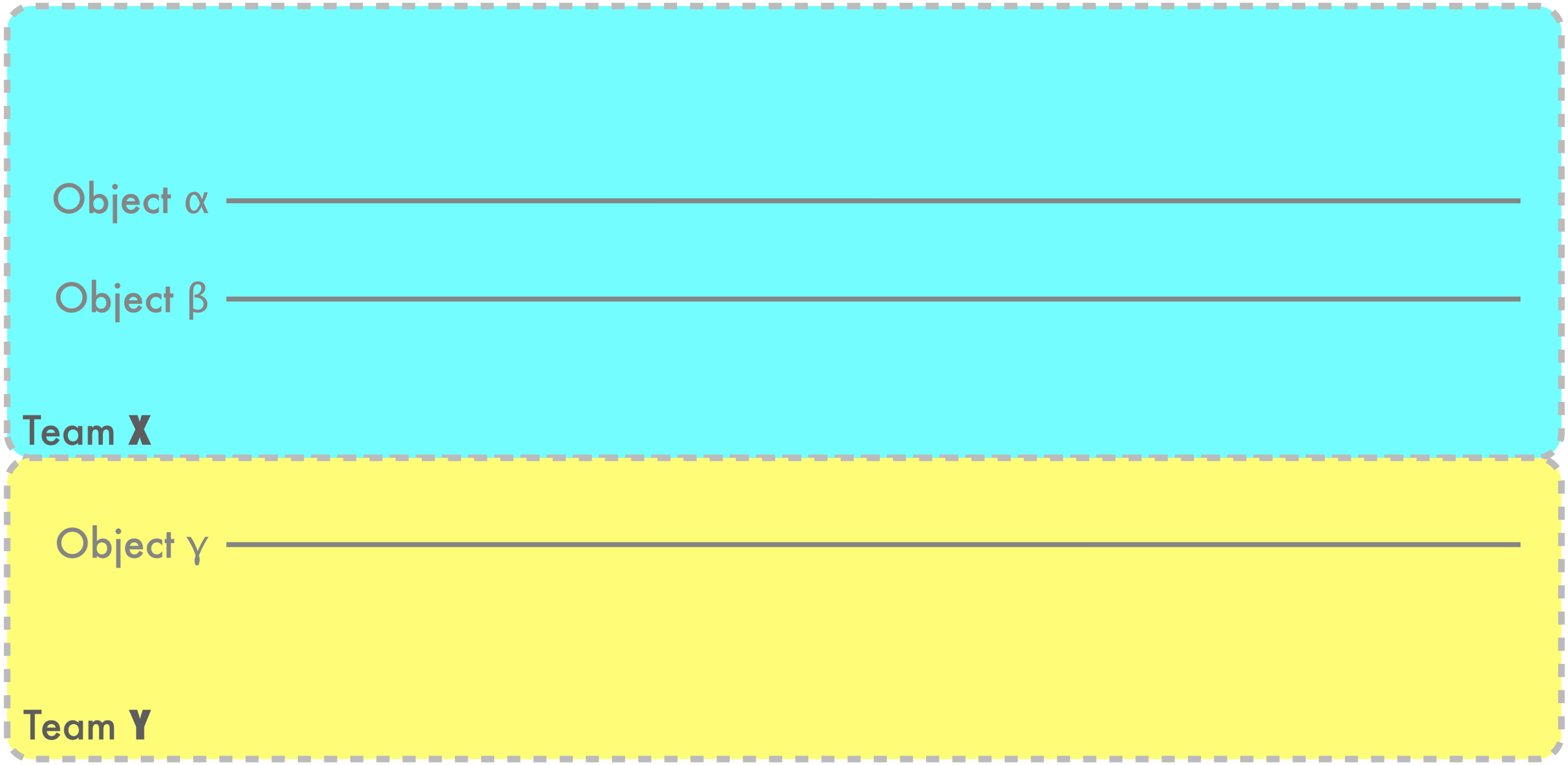
- Each team acts as a recovery unit:
  - All members checkpoint in a coordinated fashion.
  - If one member fails, the whole team rolls back.



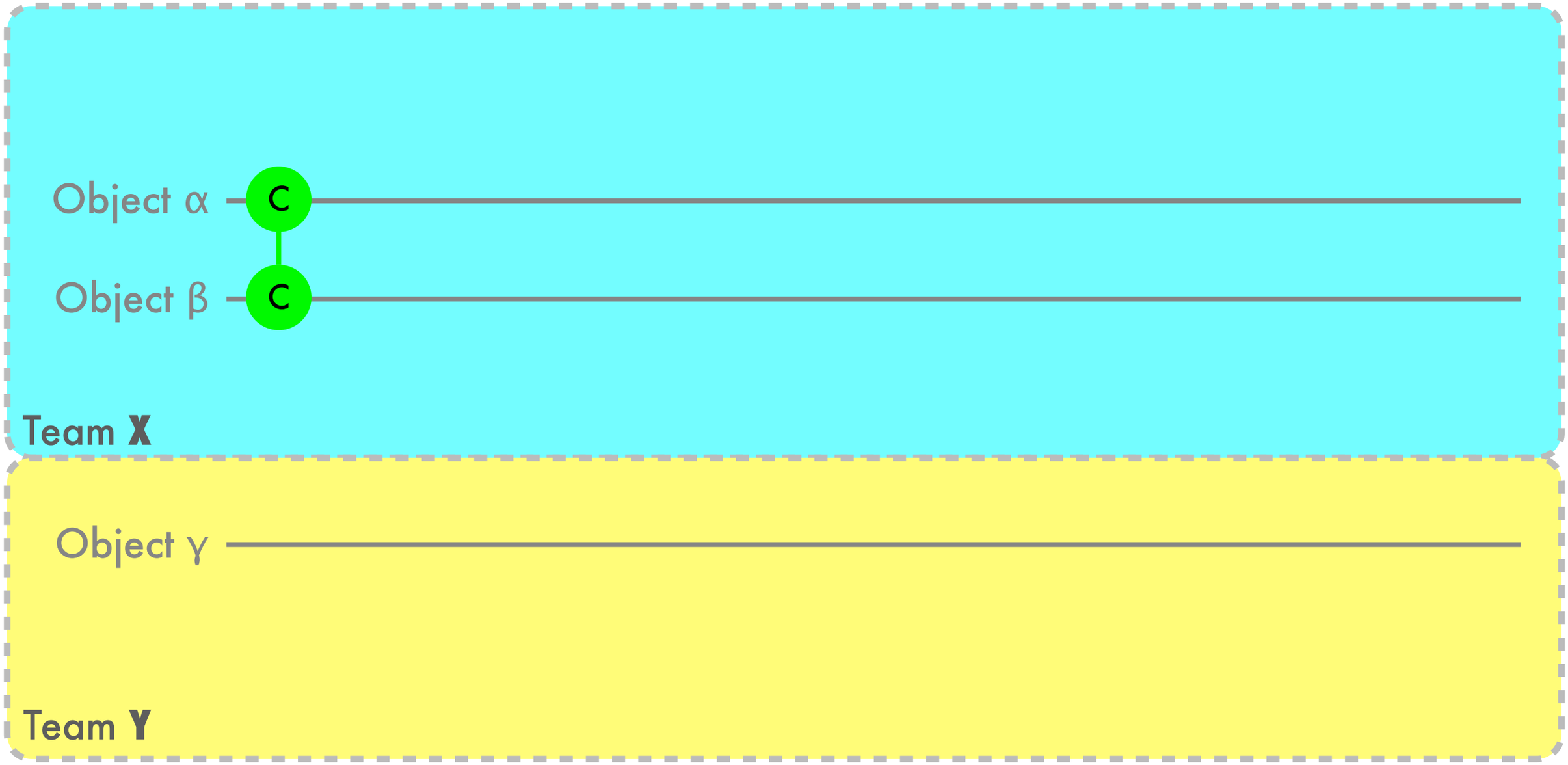
Processor Teams

- Each team acts as a recovery unit:
  - All members checkpoint in a coordinated fashion.
  - If one member fails, the whole team rolls back.

Recovery

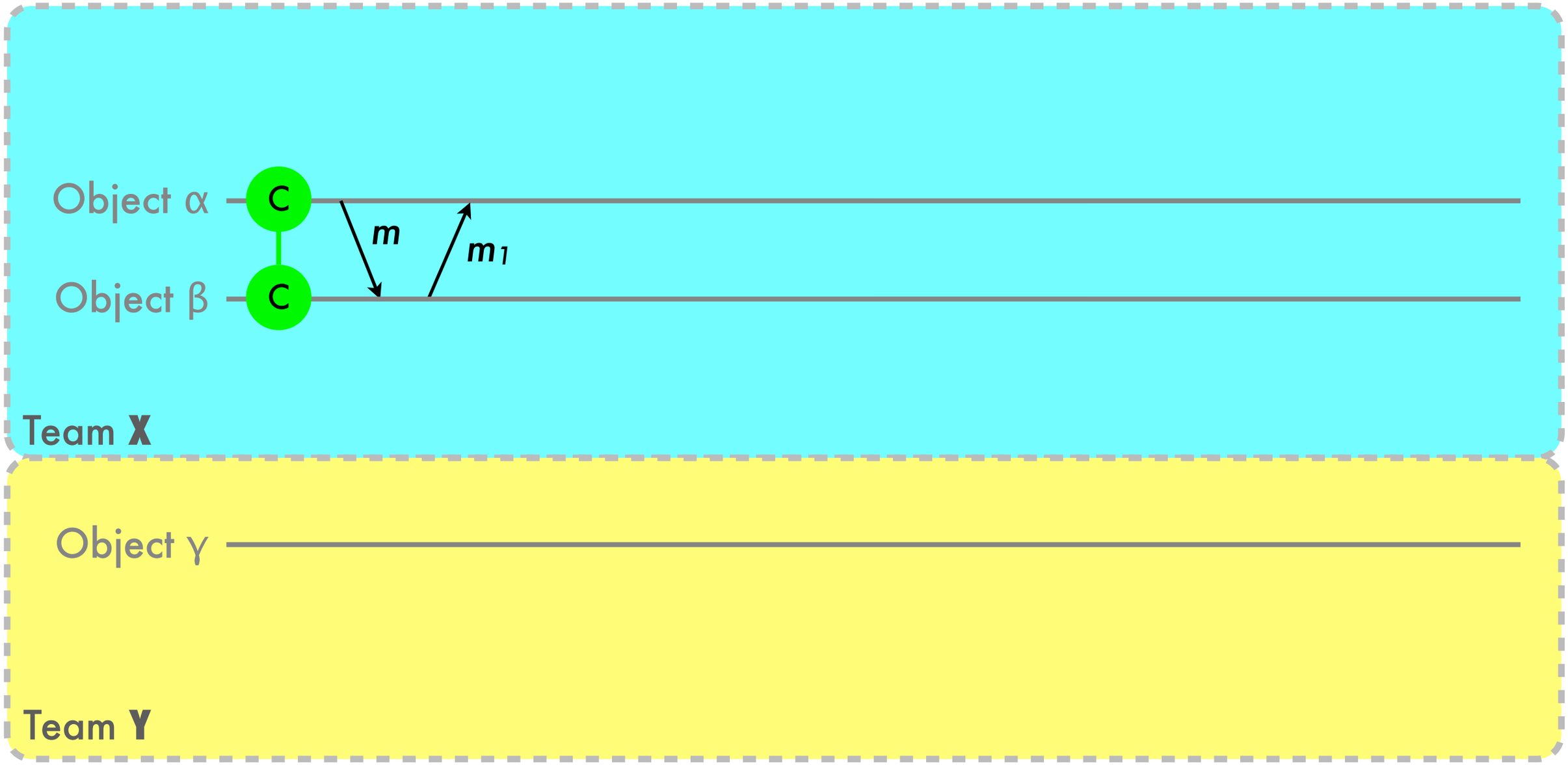


Recovery

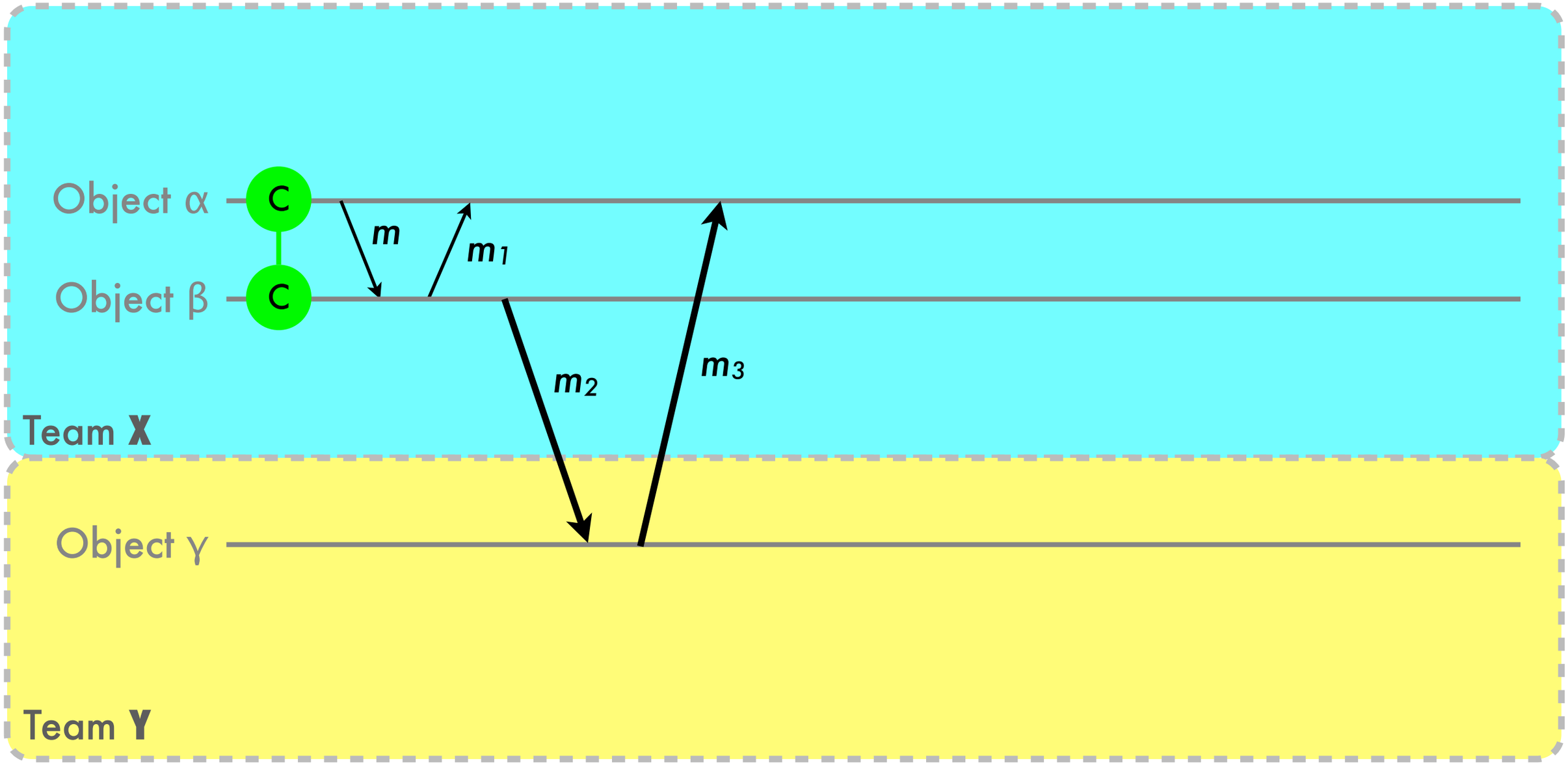


Resilience 2010, Melbourne, Australia

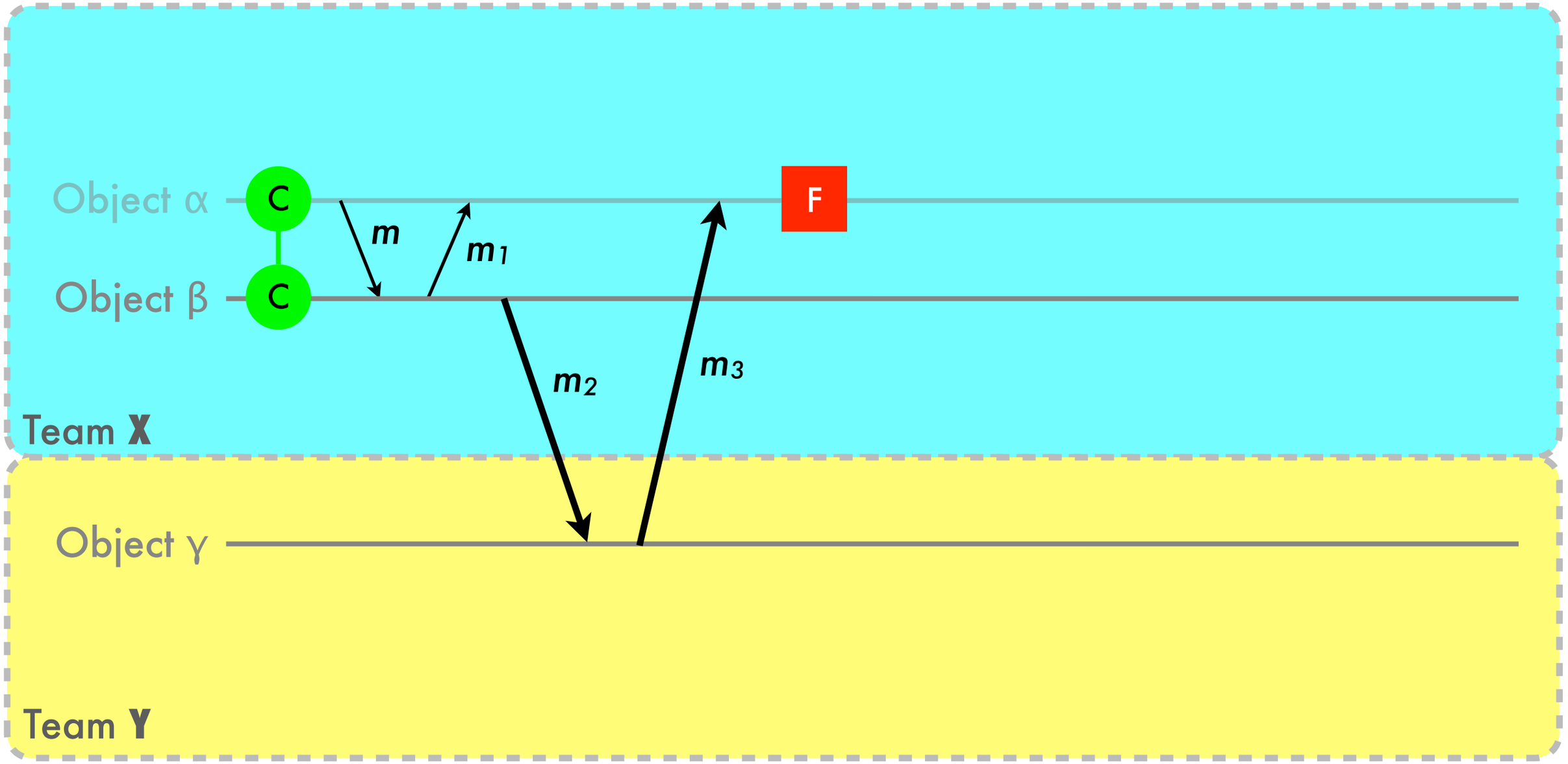
Recovery



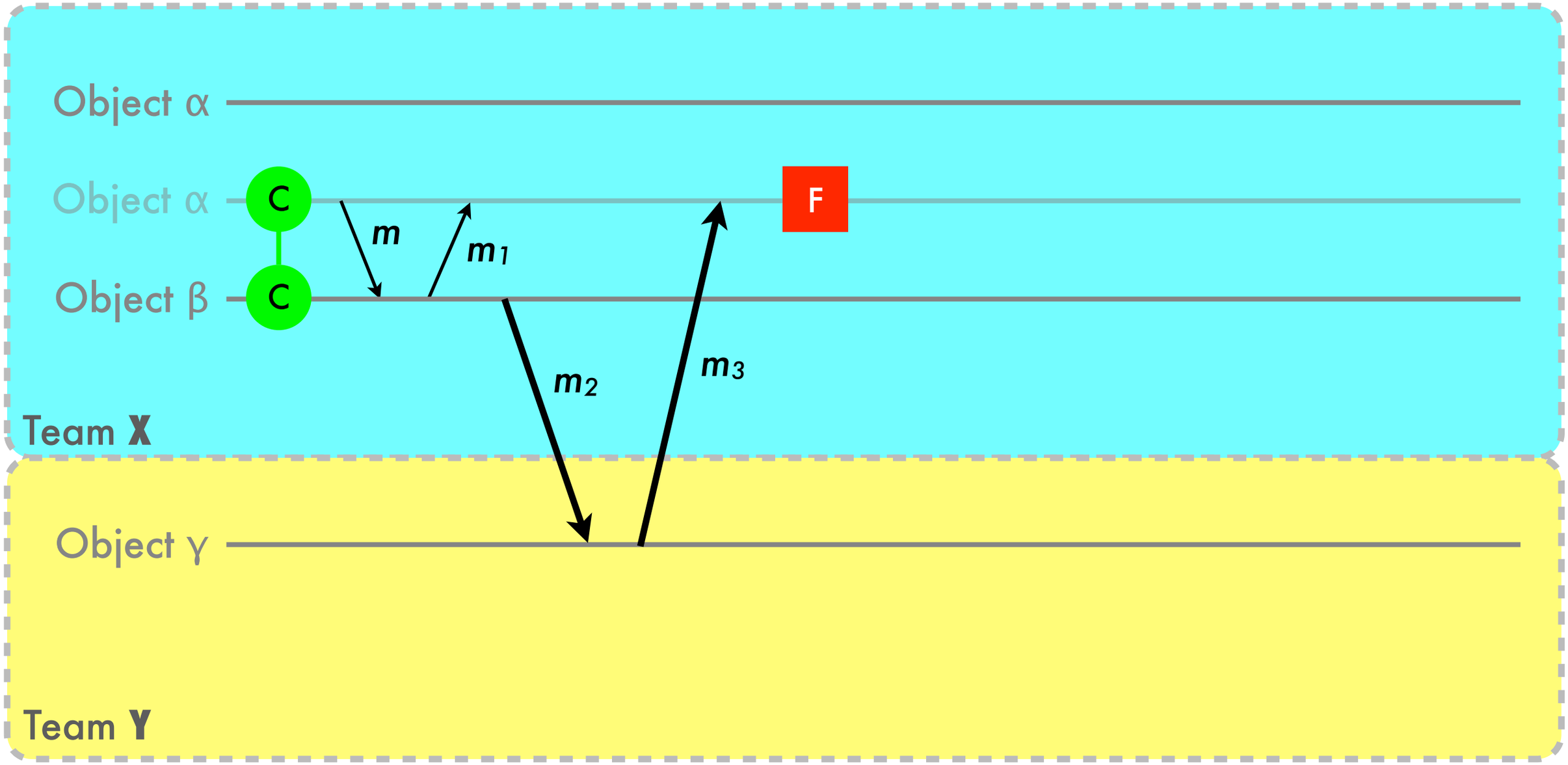
Recovery



Recovery

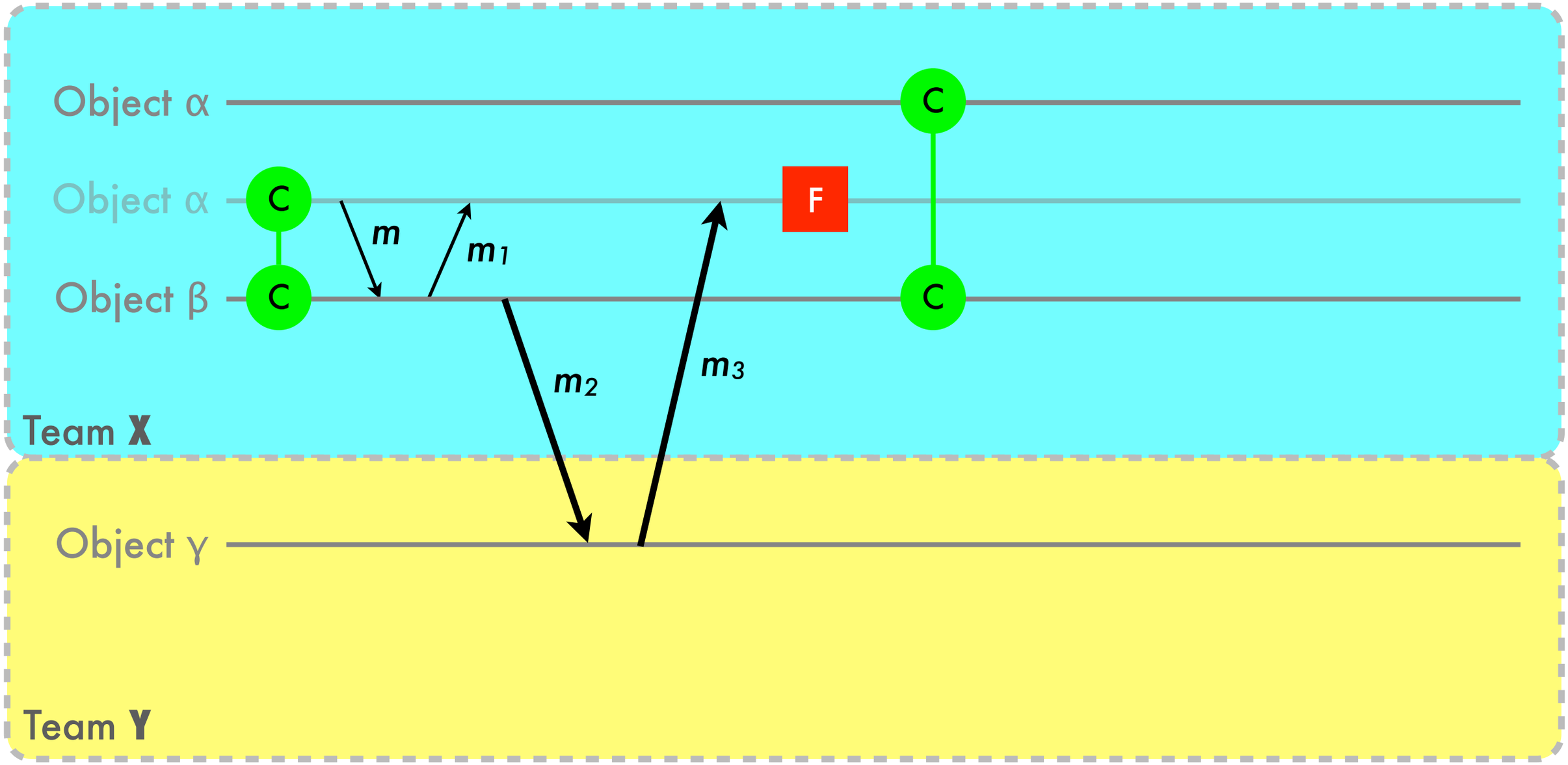


Recovery

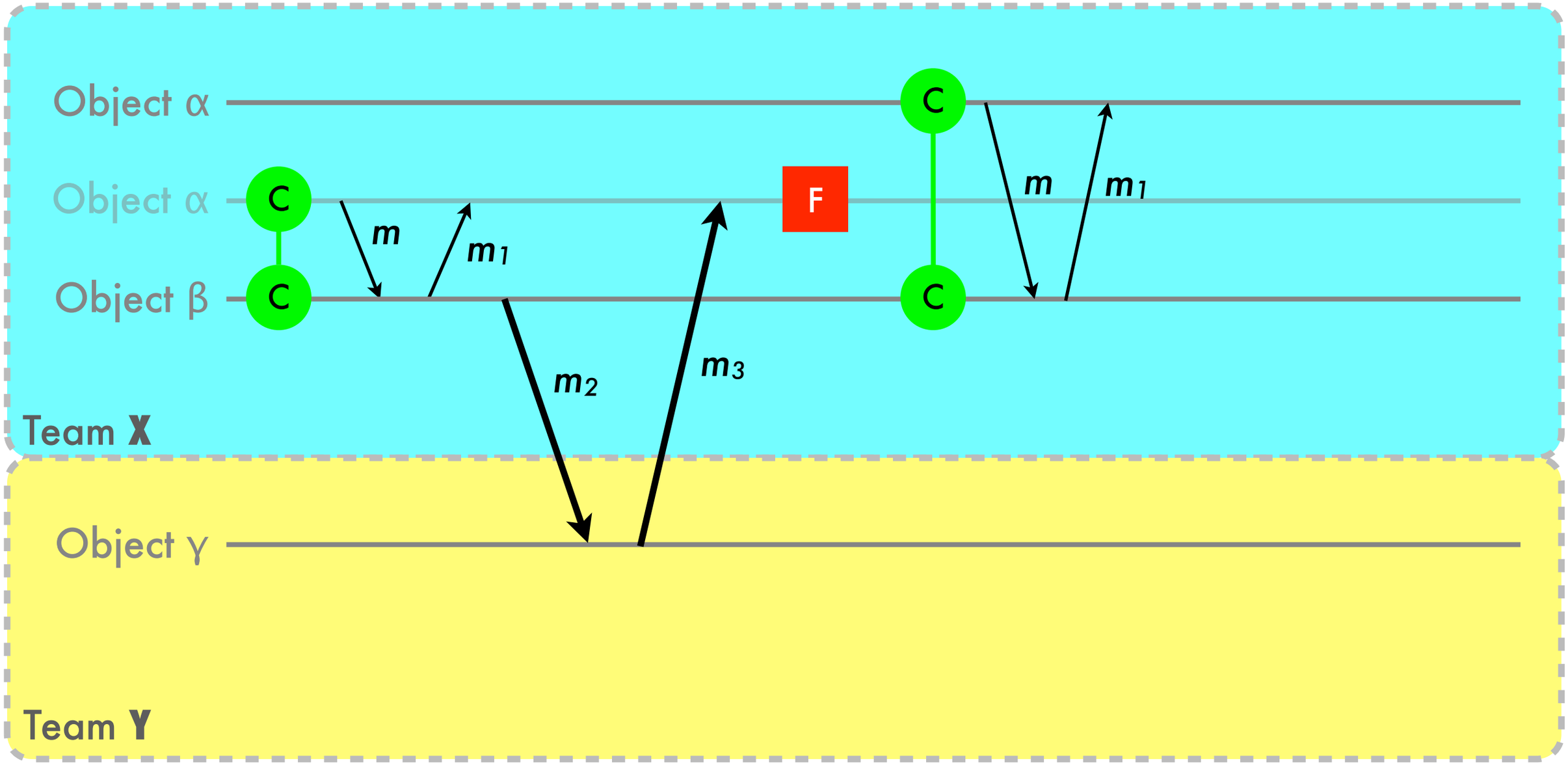


Resilience 2010, Melbourne, Australia

Recovery

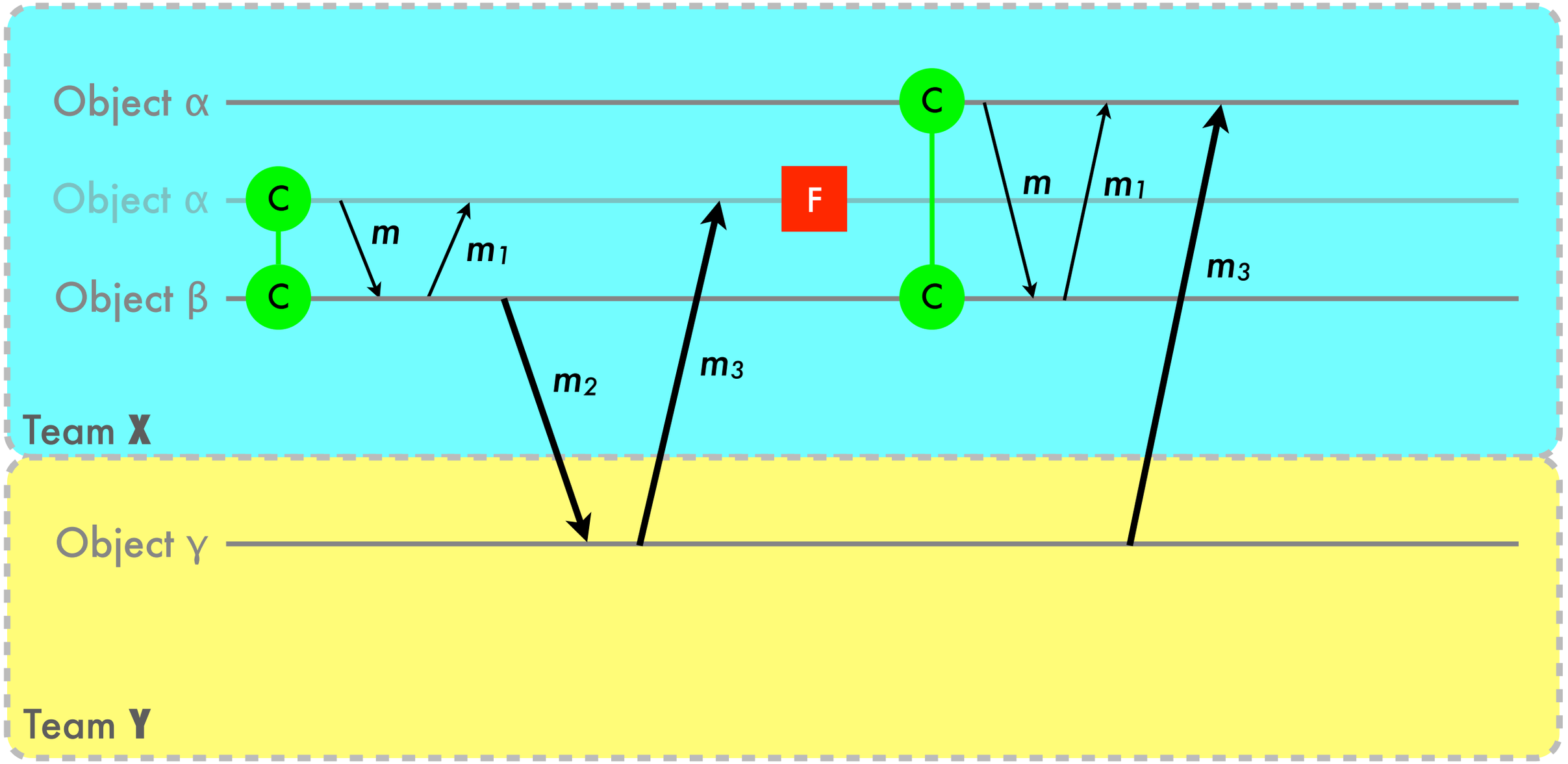

Object α
Object α
Object β
C
C
C
C
F
m
m1
m2
m3
Team X
Object γ
Team Y

Recovery



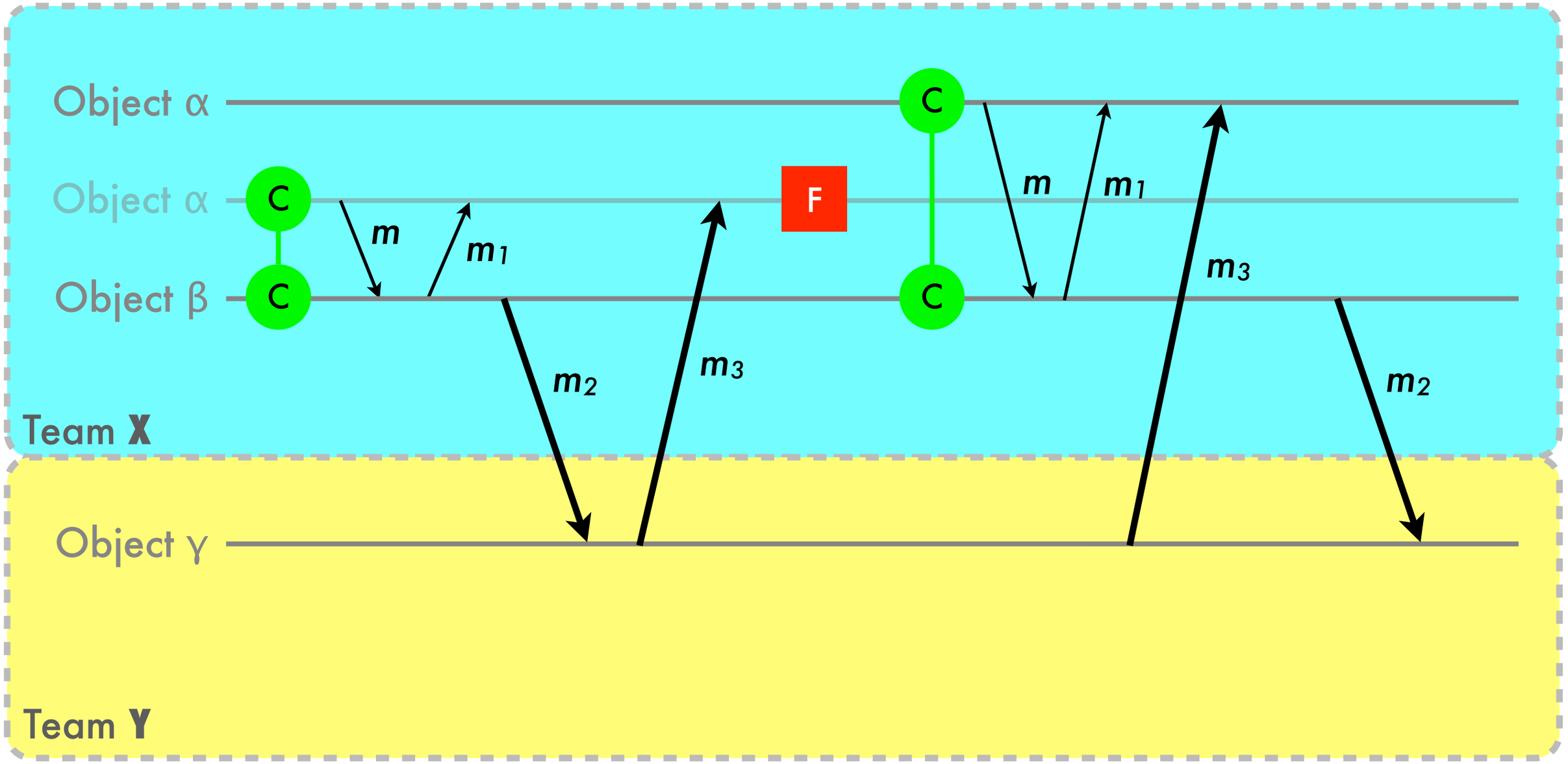
Resilience 2010, Melbourne, Australia

Recovery



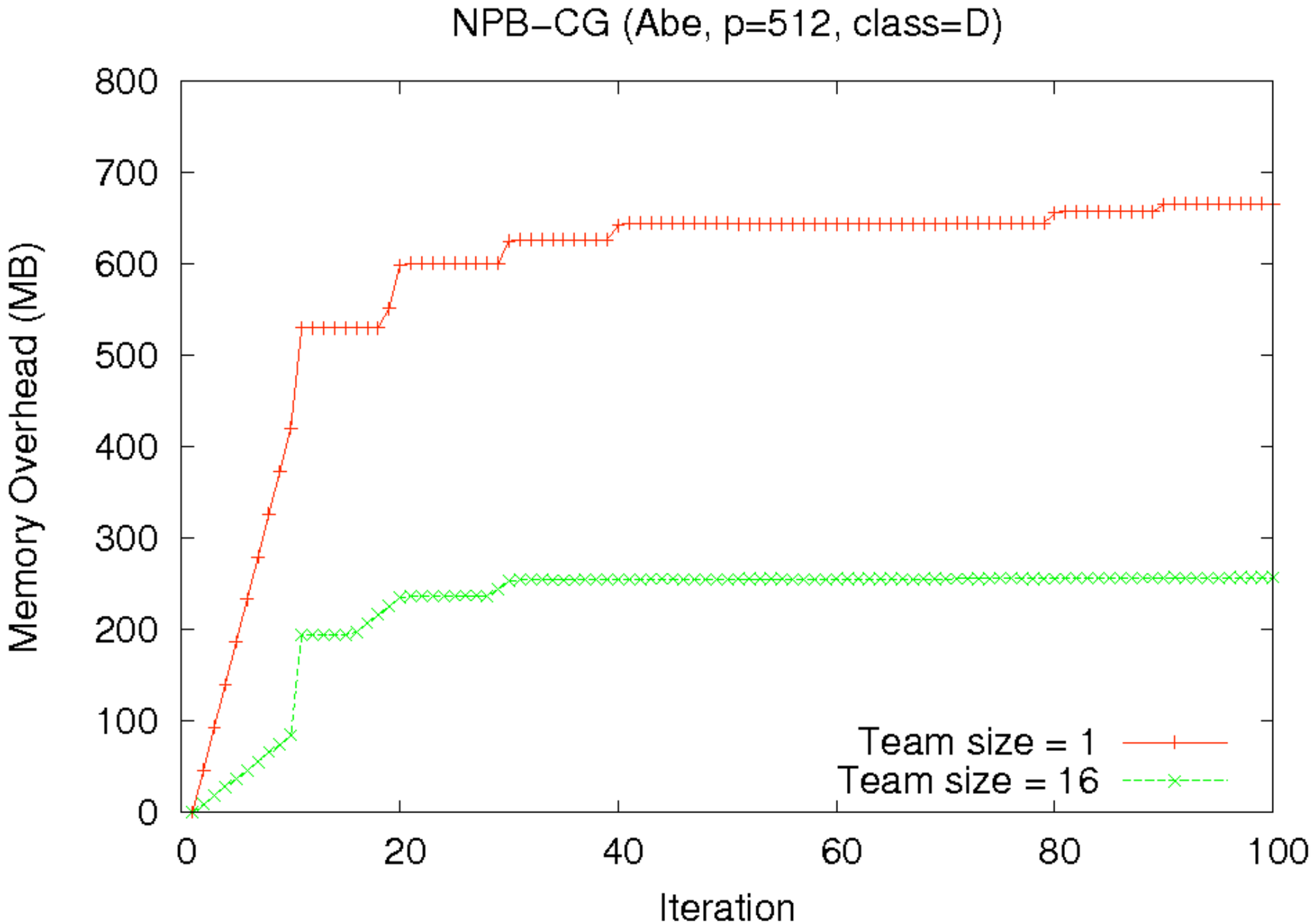
Resilience 2010, Melbourne, Australia

Recovery

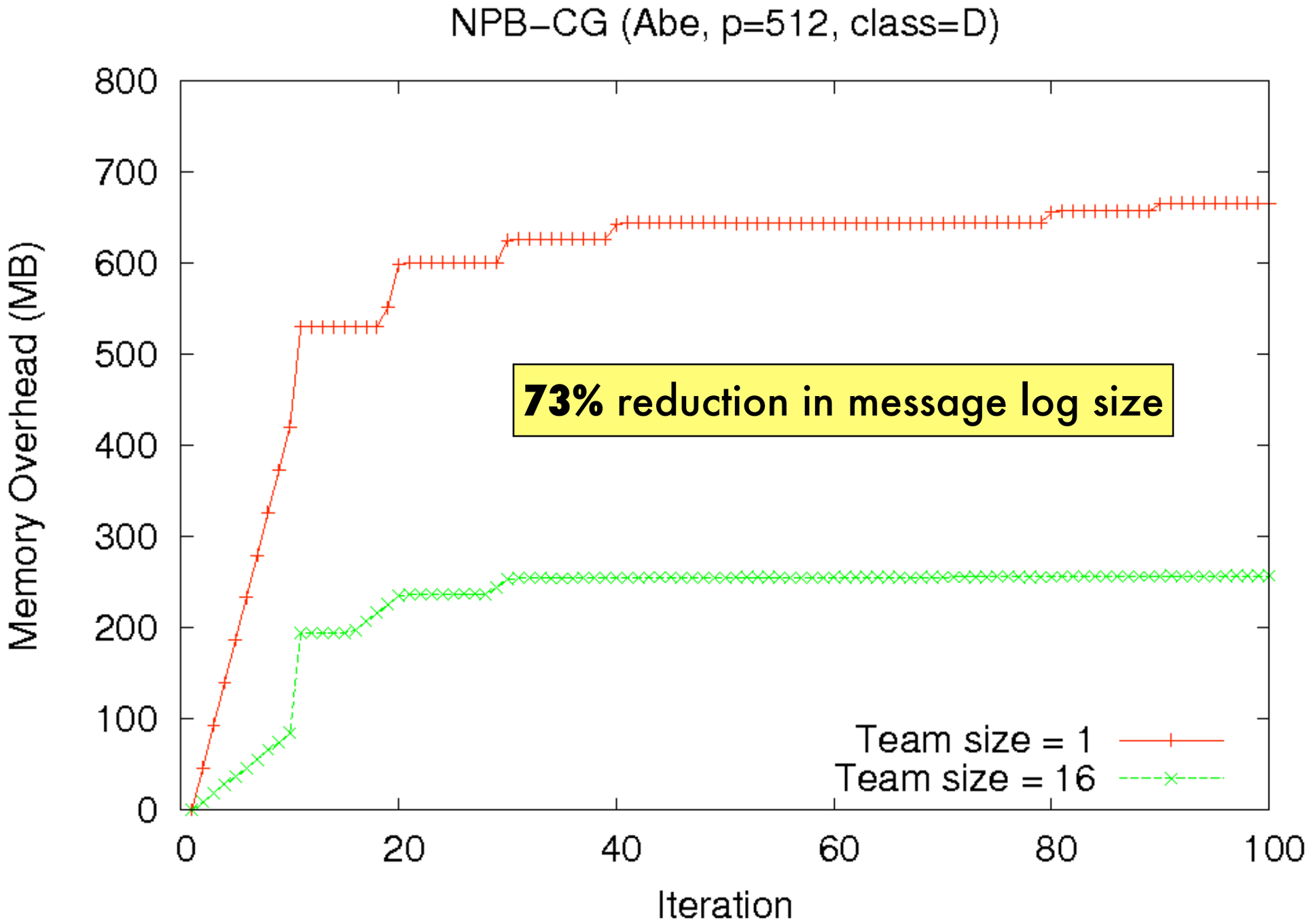


Resilience 2010, Melbourne, Australia

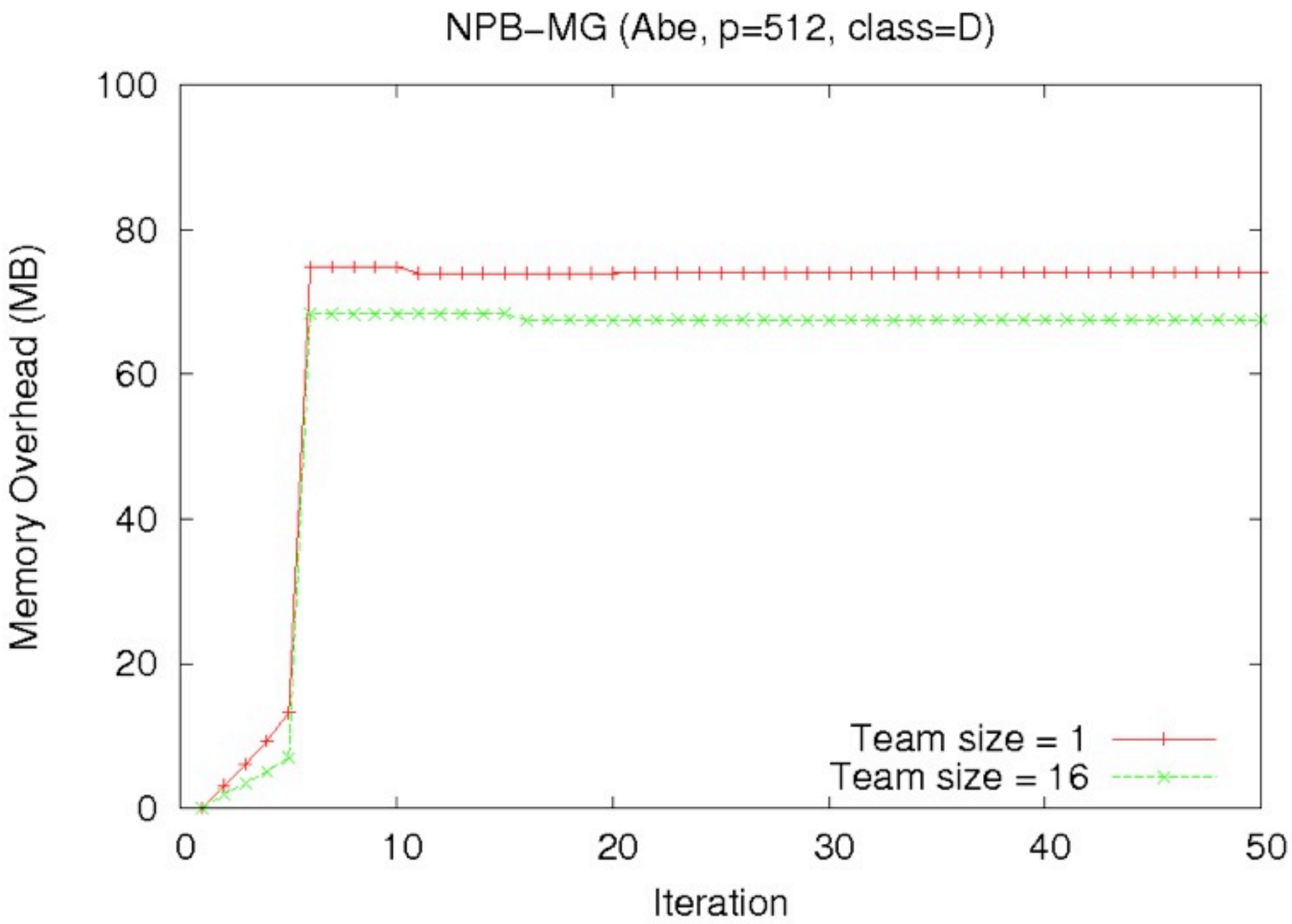
NPB-CG


NPB–CG (Abe, p=512, class=D)
Memory Overhead (MB)
800
700
600
500
400
300
200
100
0
0
20
40
60
80
100
Iteration
Team size = 1
Team size = 16

NPB-CG

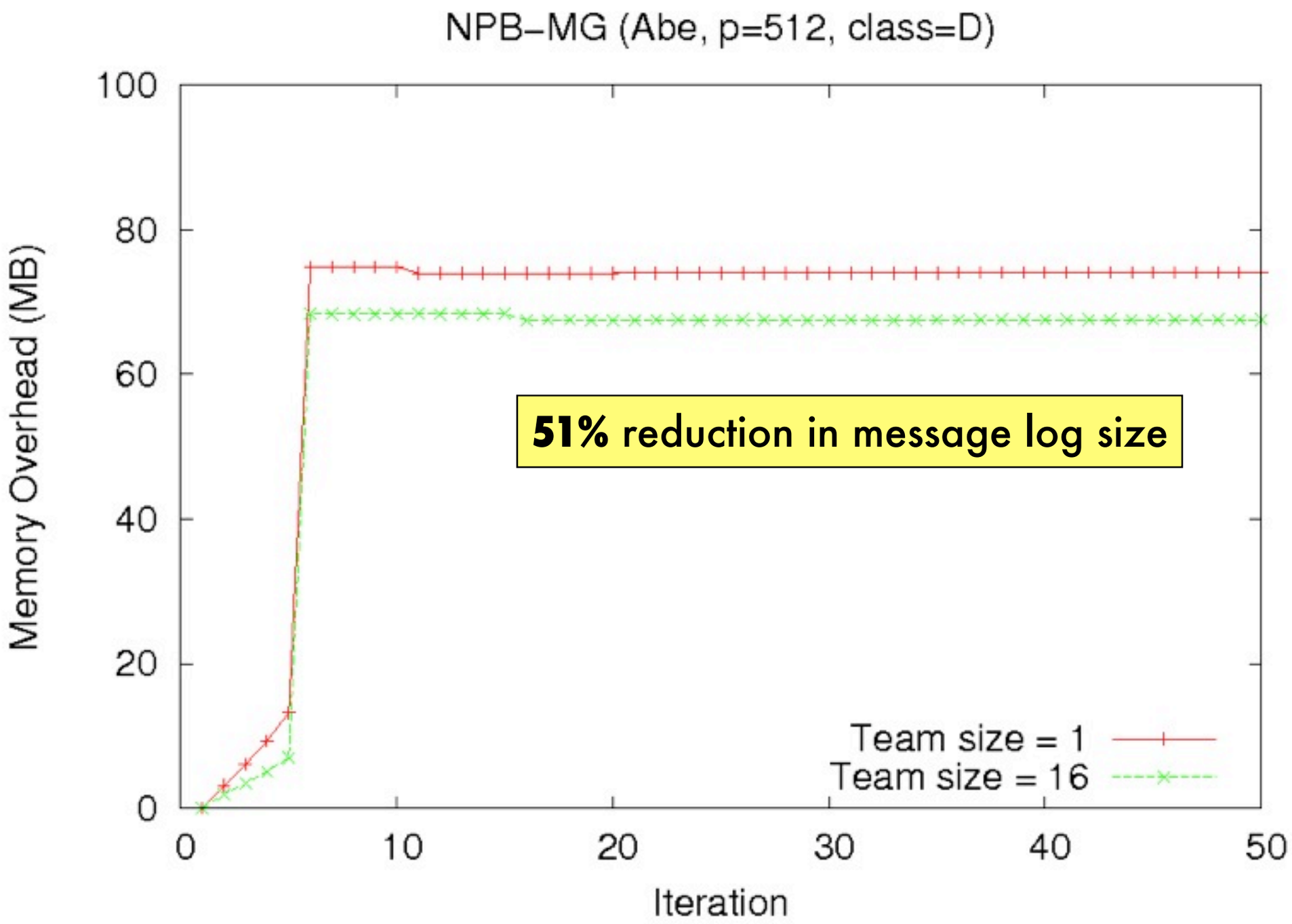

NPB–CG (Abe, p=512, class=D)
Memory Overhead (MB)
800
700
600
500
400
300
200
100
0
0
20
40
60
80
100
Iteration
73% reduction in message log size
Team size = 1
Team size = 16

NPB-MG



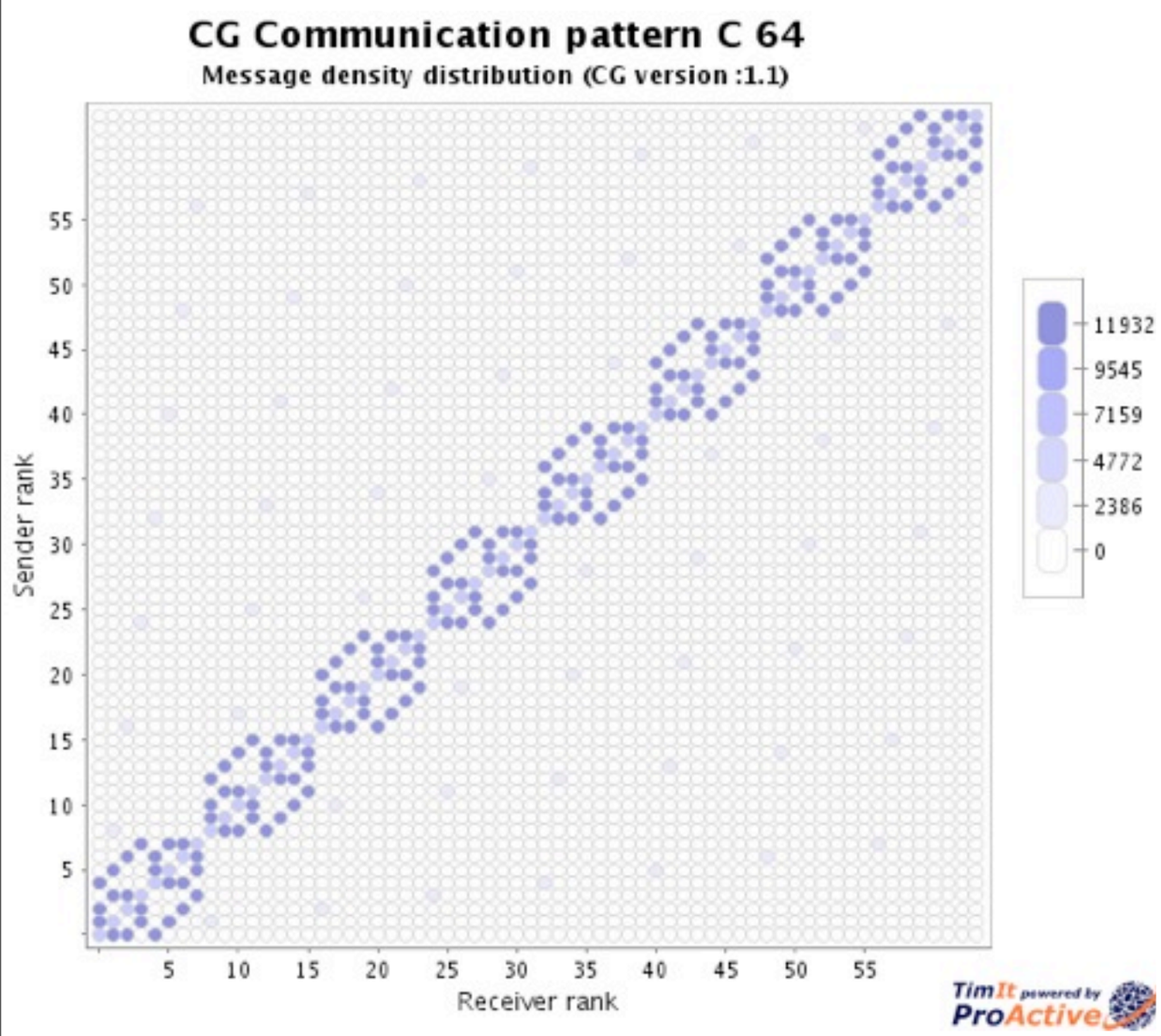
Resilience 2010, Melbourne, Australia

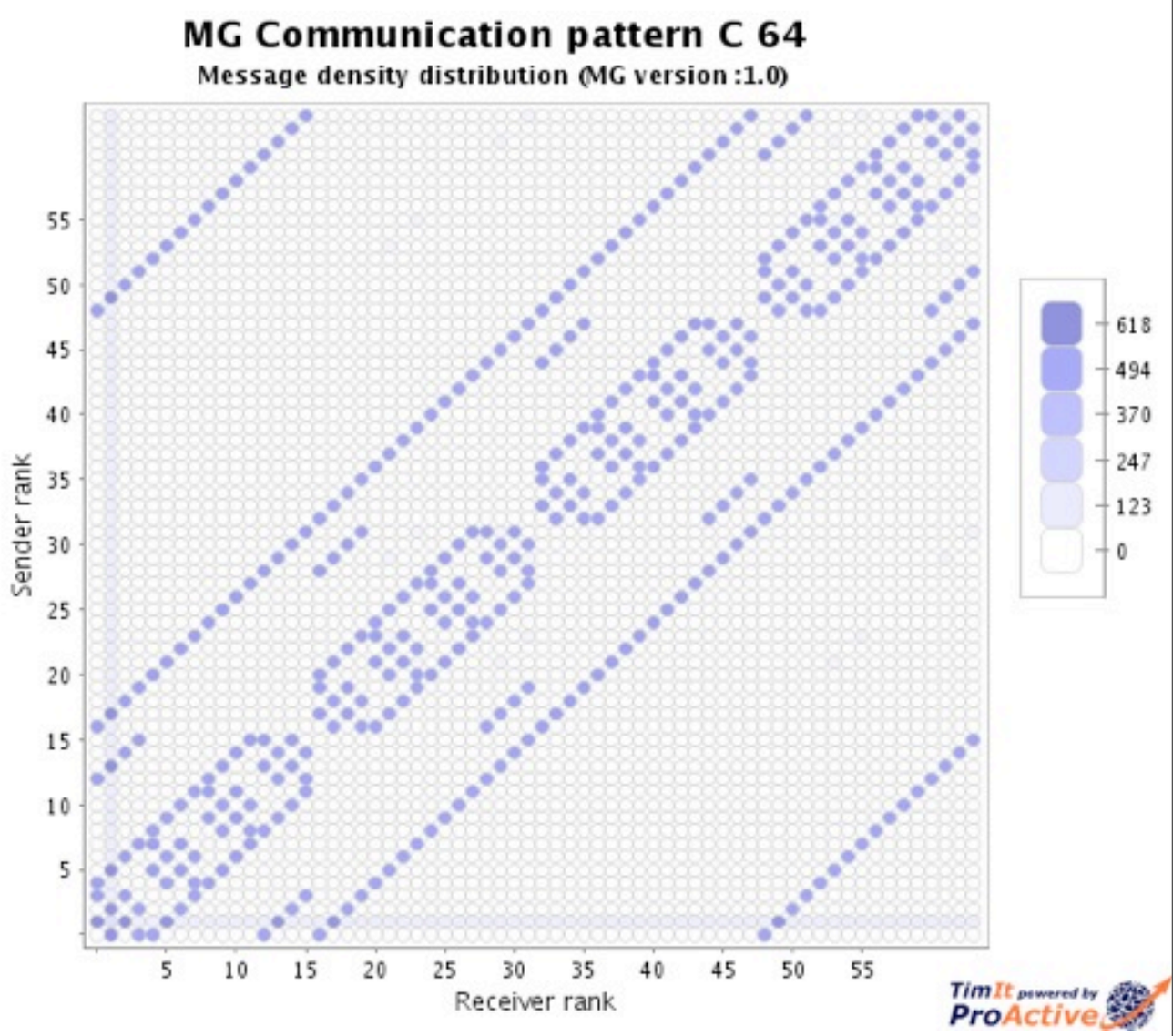
NPB-MG



Resilience 2010, Melbourne, Australia

Communication Pattern

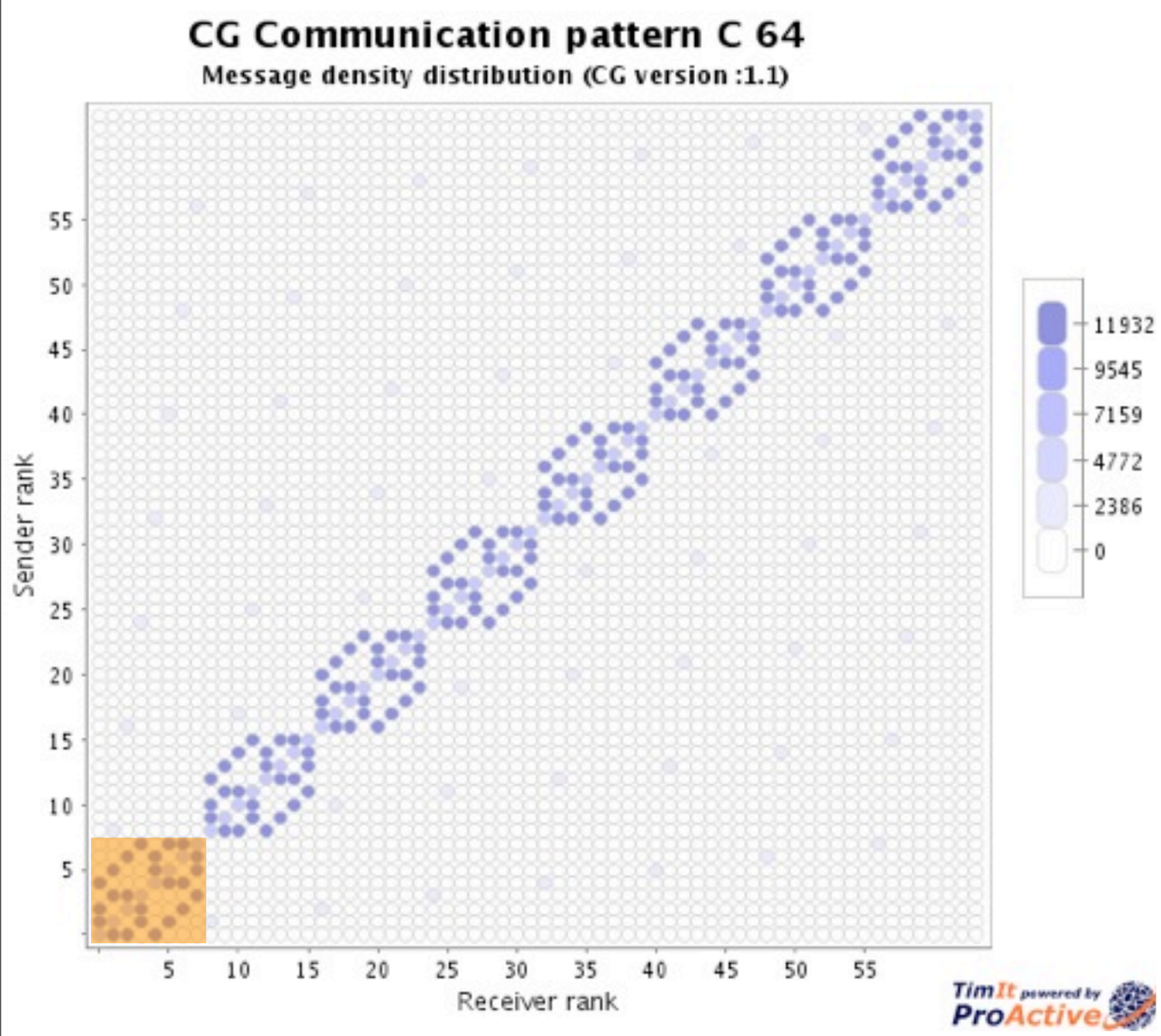


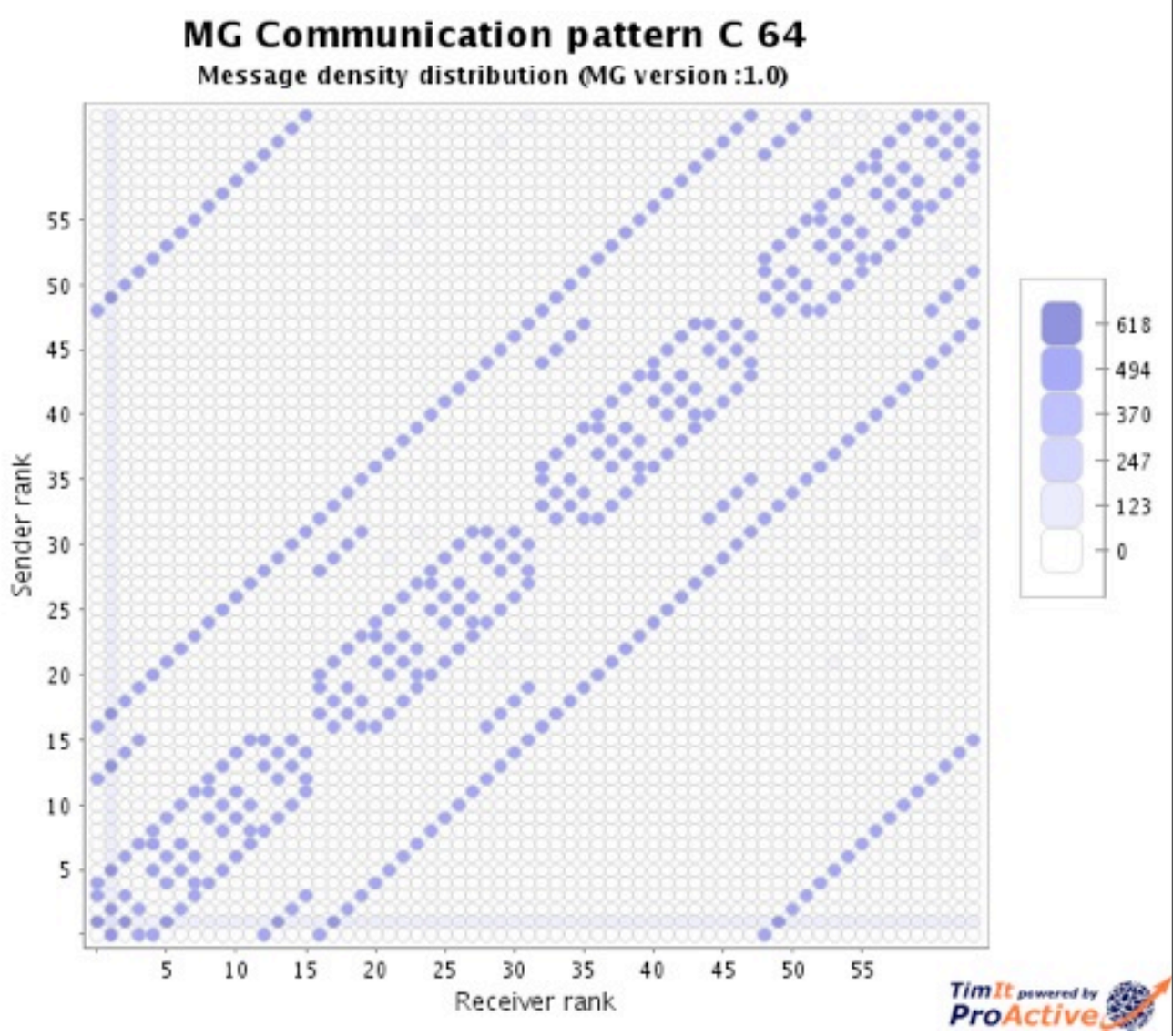


Reproduced with permission from http://proactive.inria.fr

Resilience 2010, Melbourne, Australia

Communication Pattern

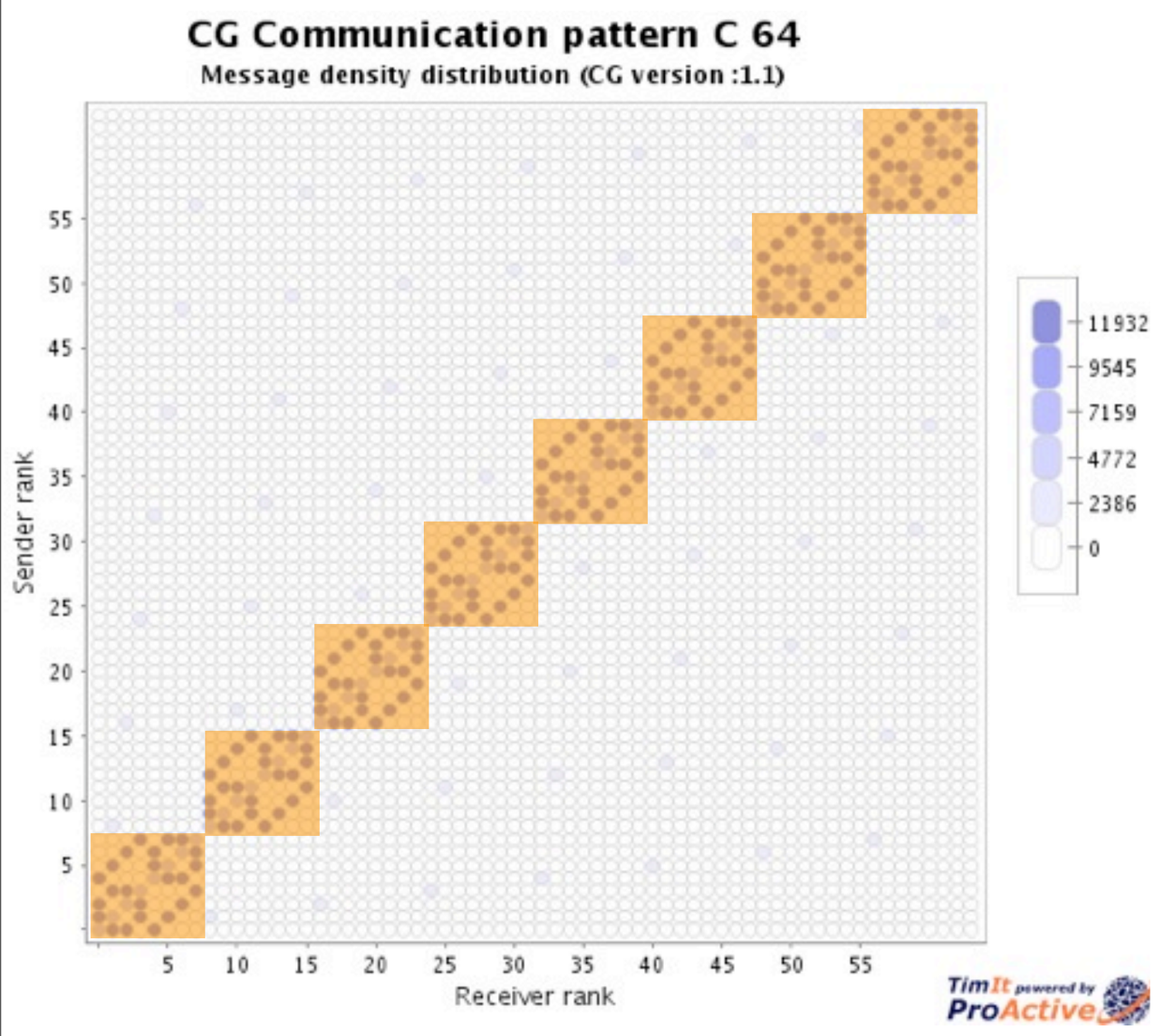


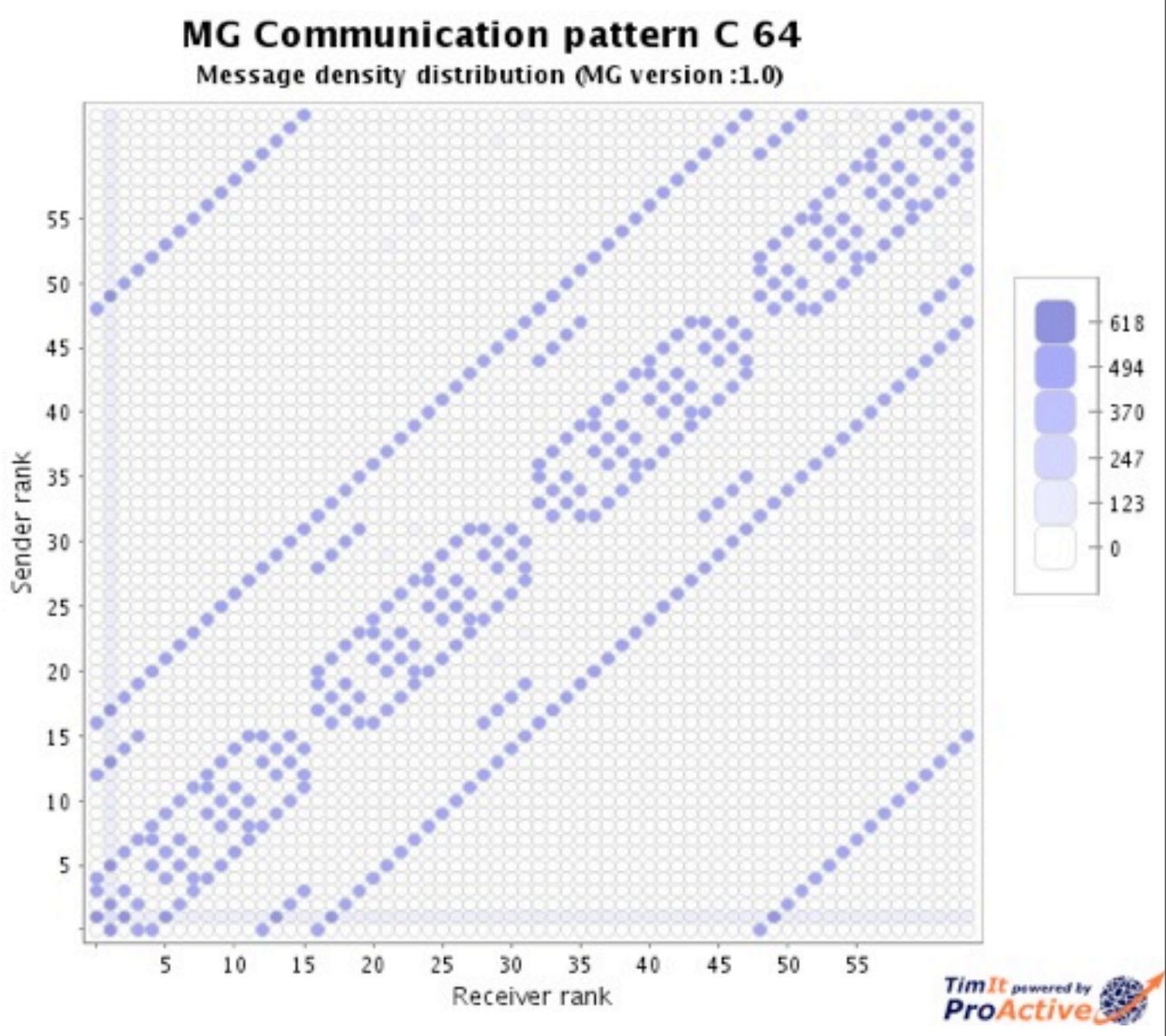


Reproduced with permission from http://proactive.inria.fr

Resilience 2010, Melbourne, Australia

Communication Pattern

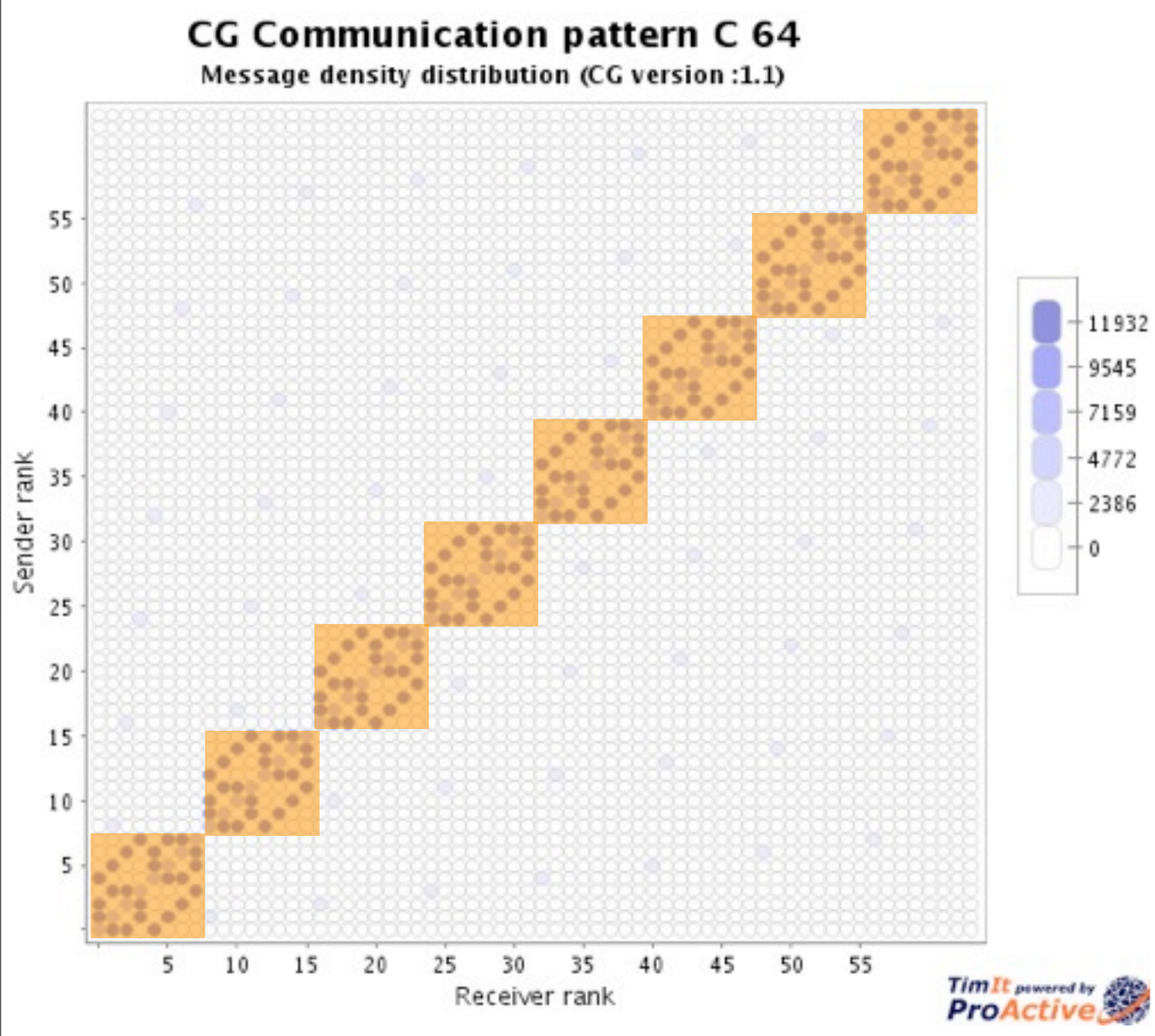


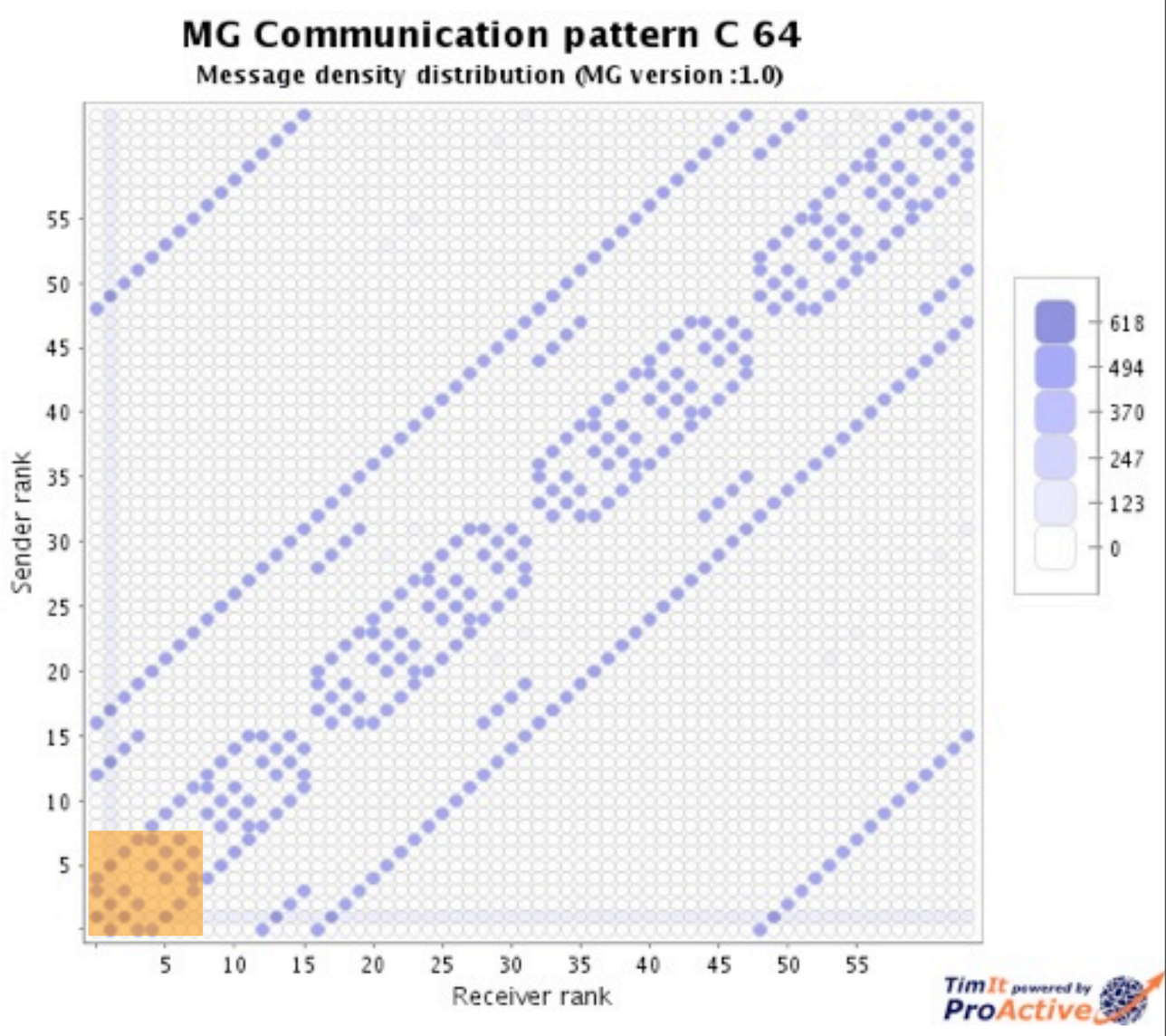


Reproduced with permission from http://proactive.inria.fr

Resilience 2010, Melbourne, Australia

Communication Pattern

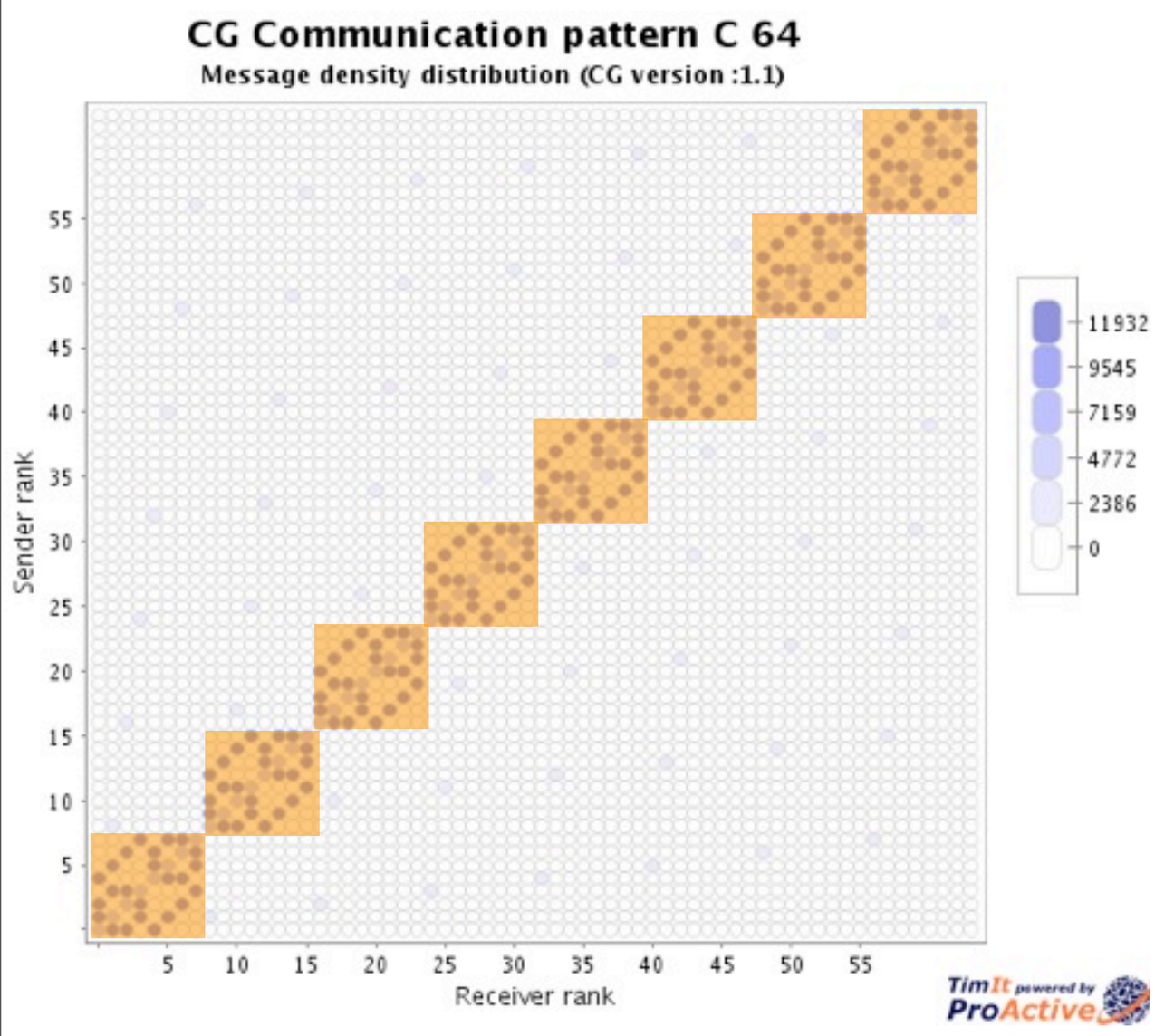


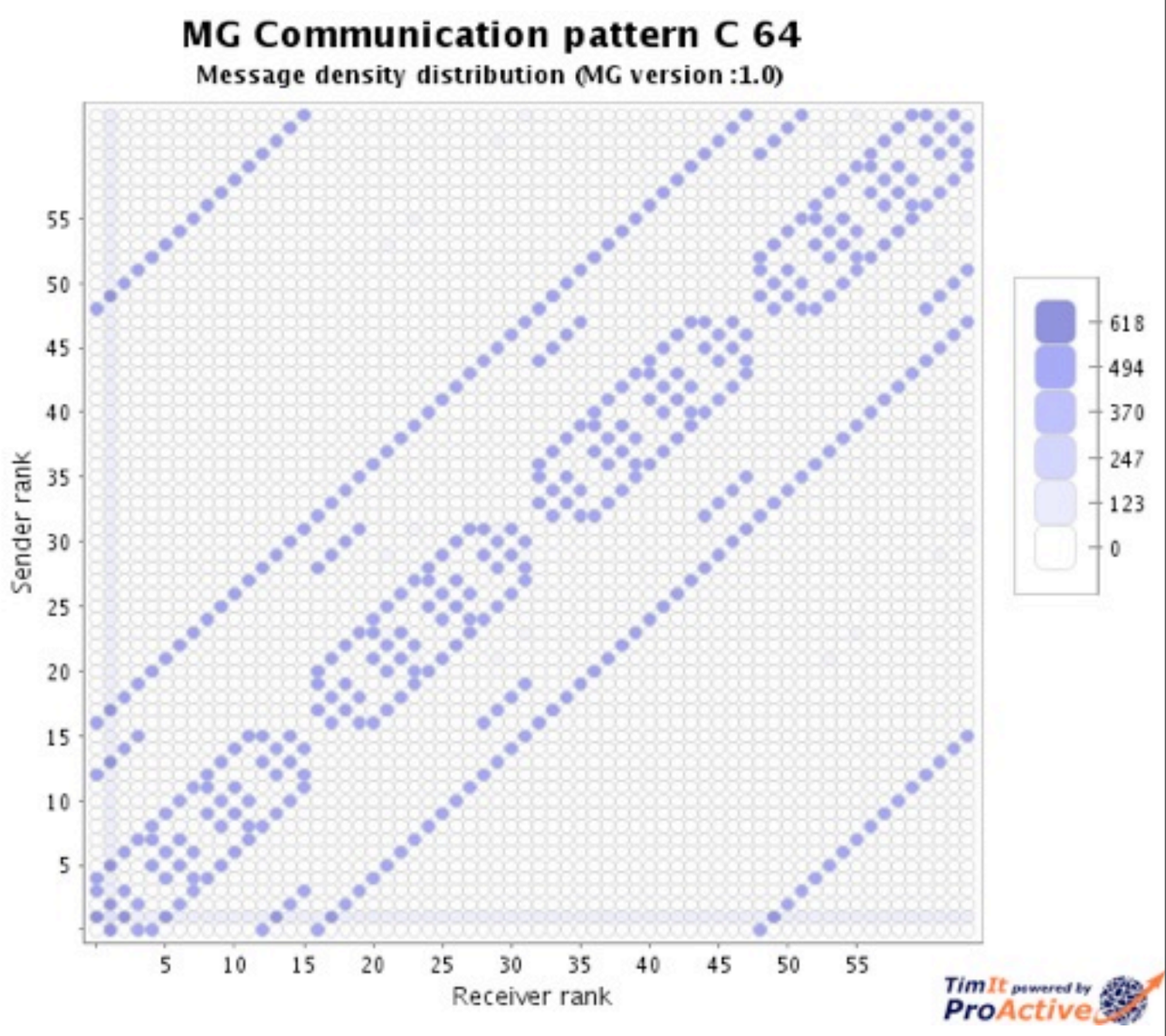


Reproduced with permission from http://proactive.inria.fr

Resilience 2010, Melbourne, Australia

Communication Pattern

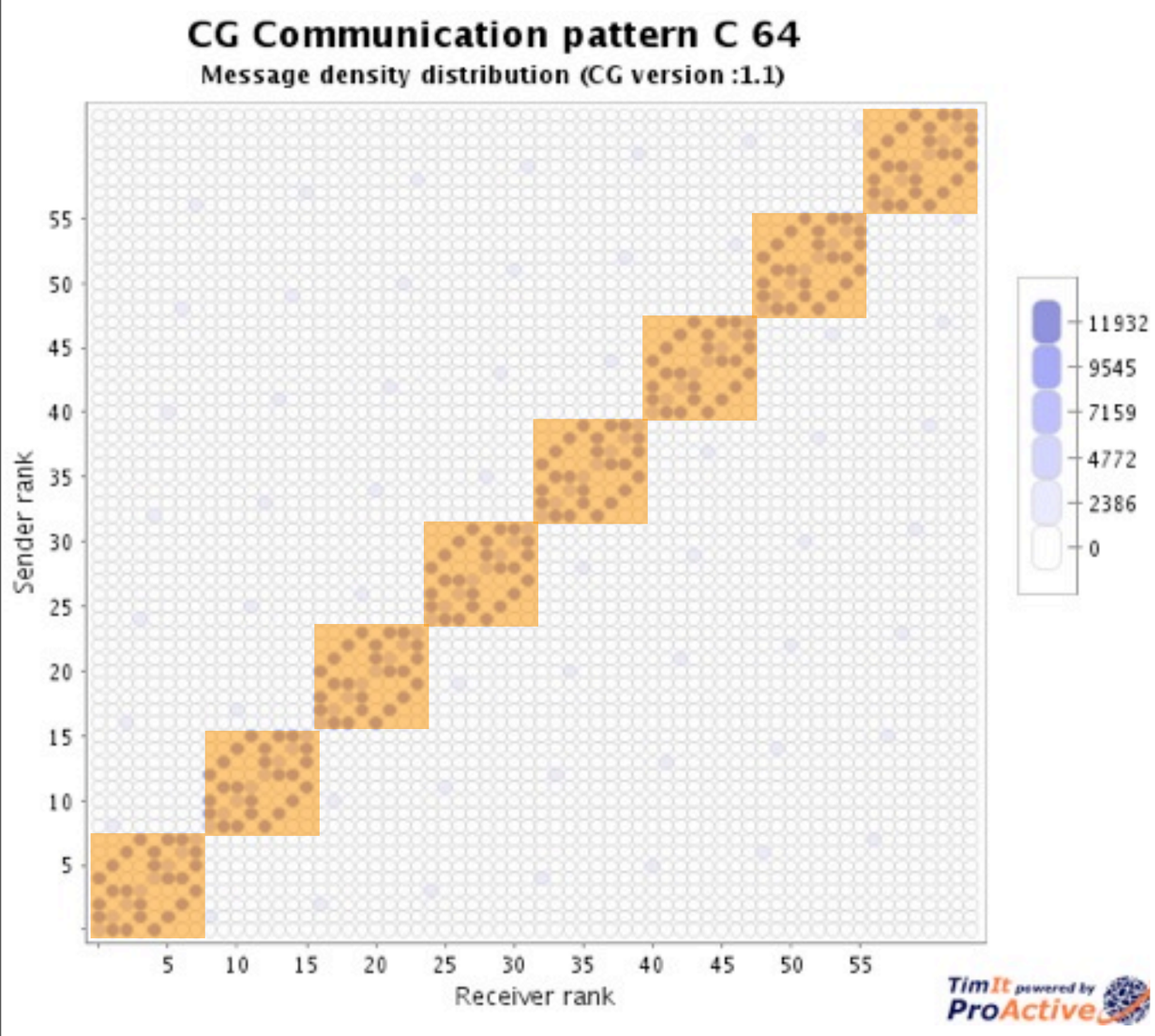


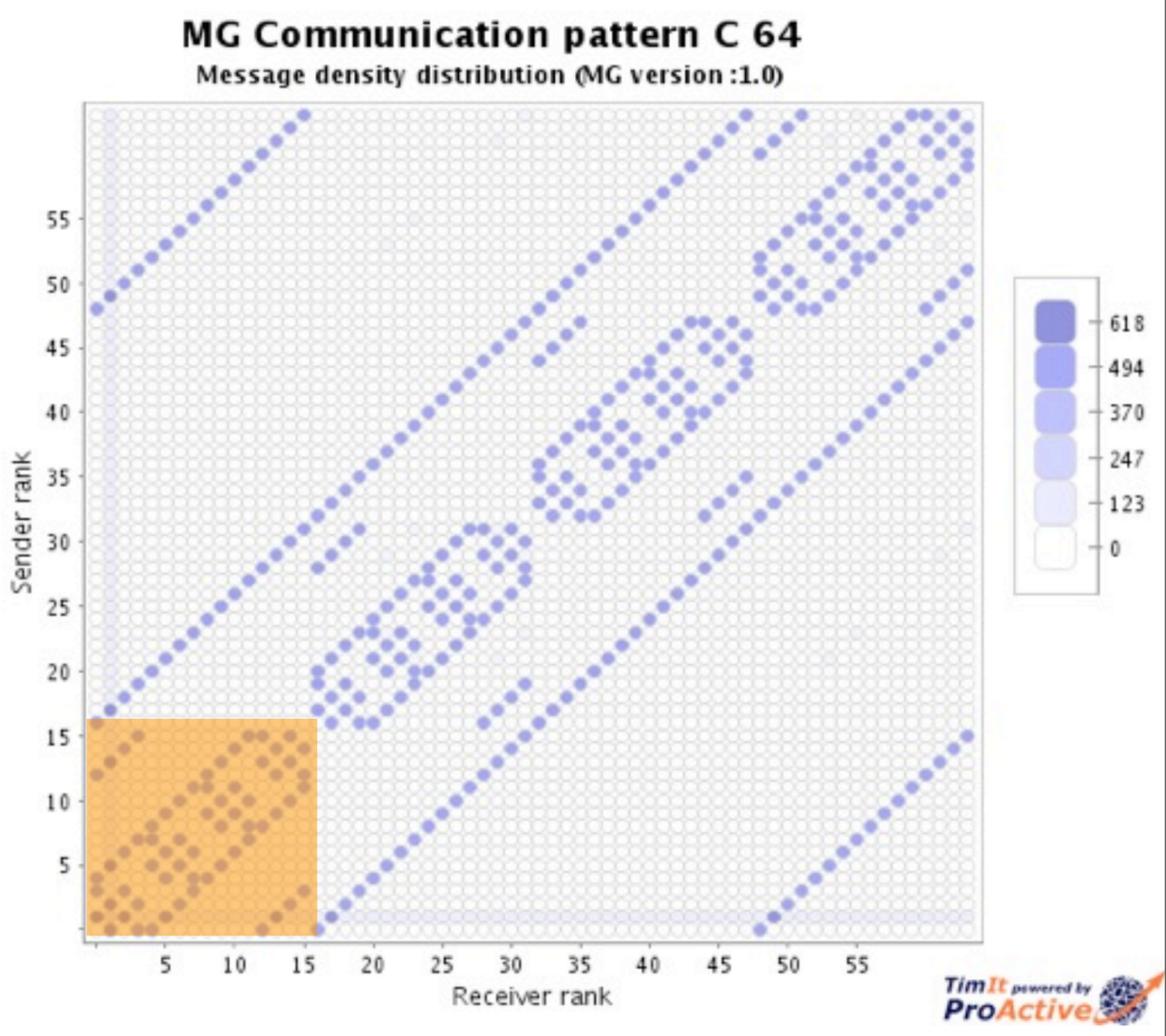


Reproduced with permission from http://proactive.inria.fr

Resilience 2010, Melbourne, Australia

Communication Pattern

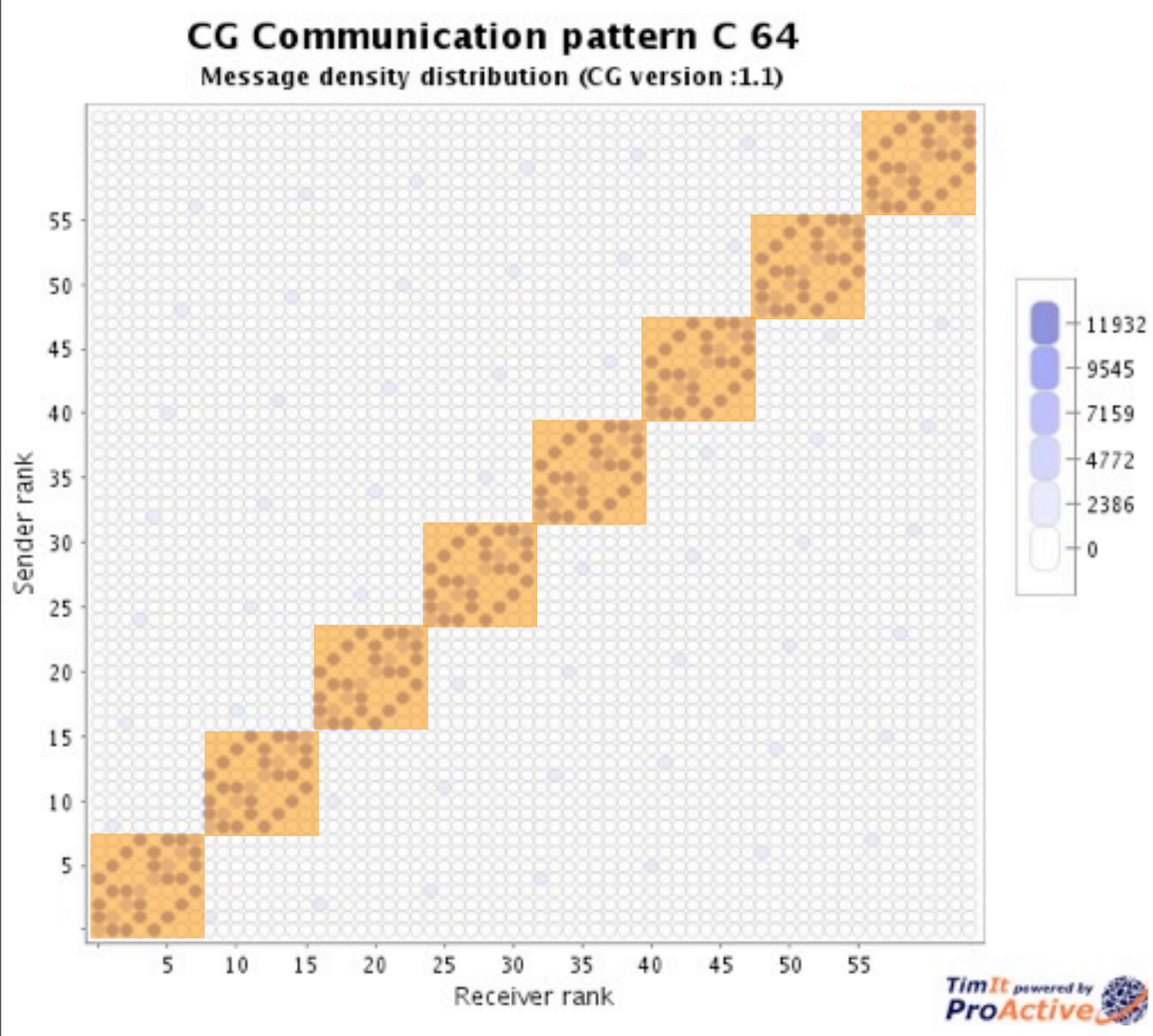


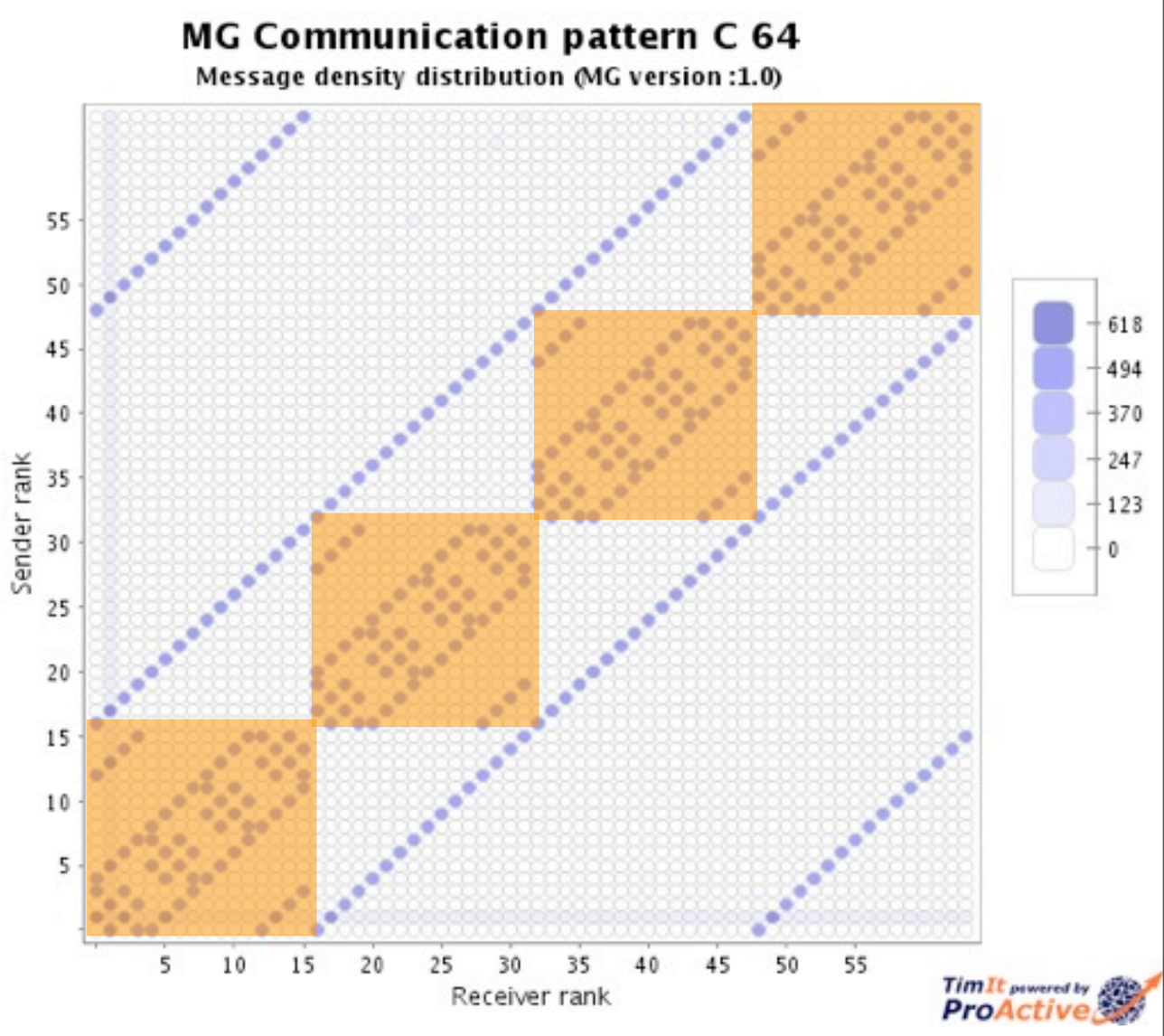


Reproduced with permission from http://proactive.inria.fr

Resilience 2010, Melbourne, Australia

Communication Pattern

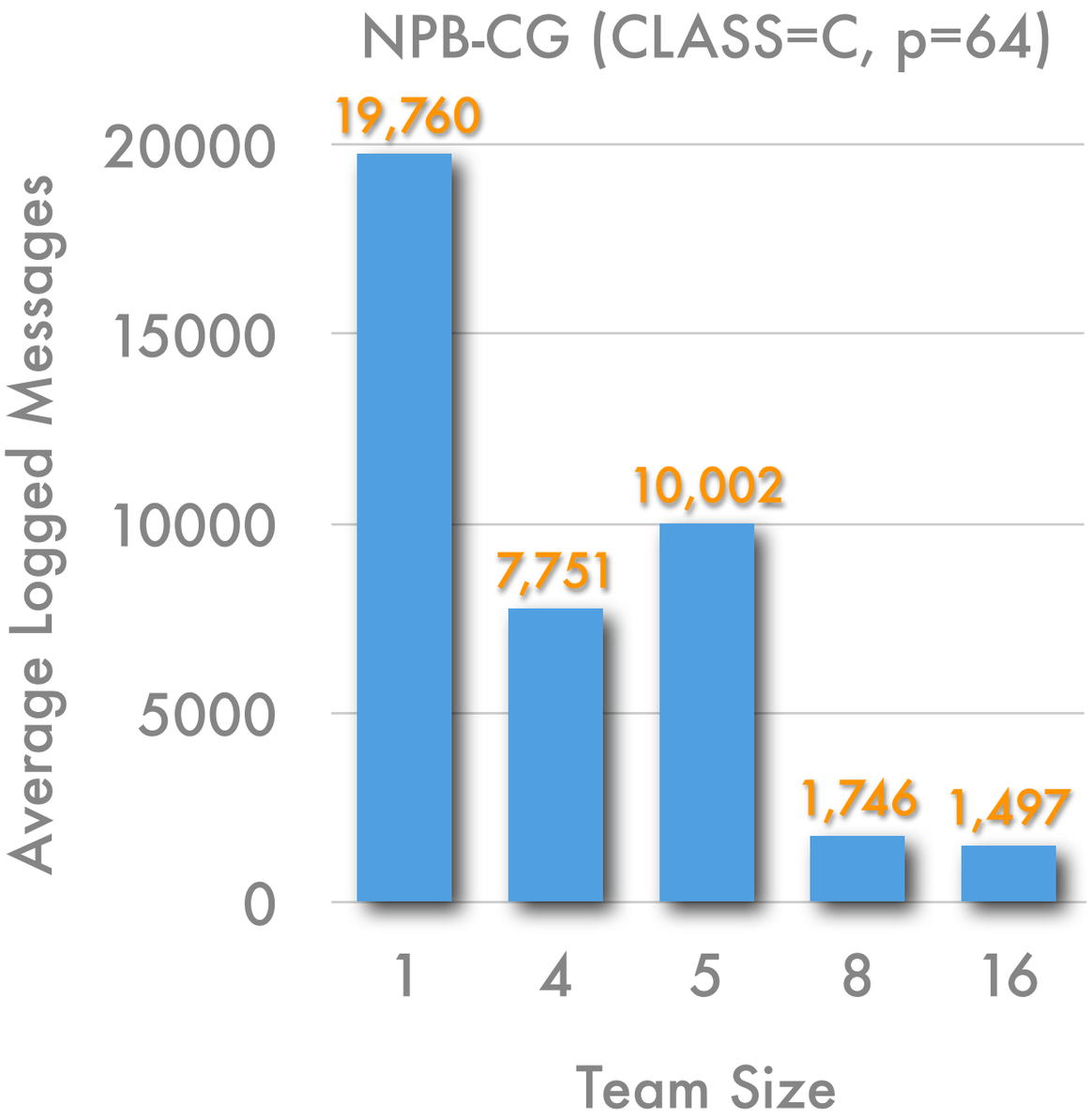


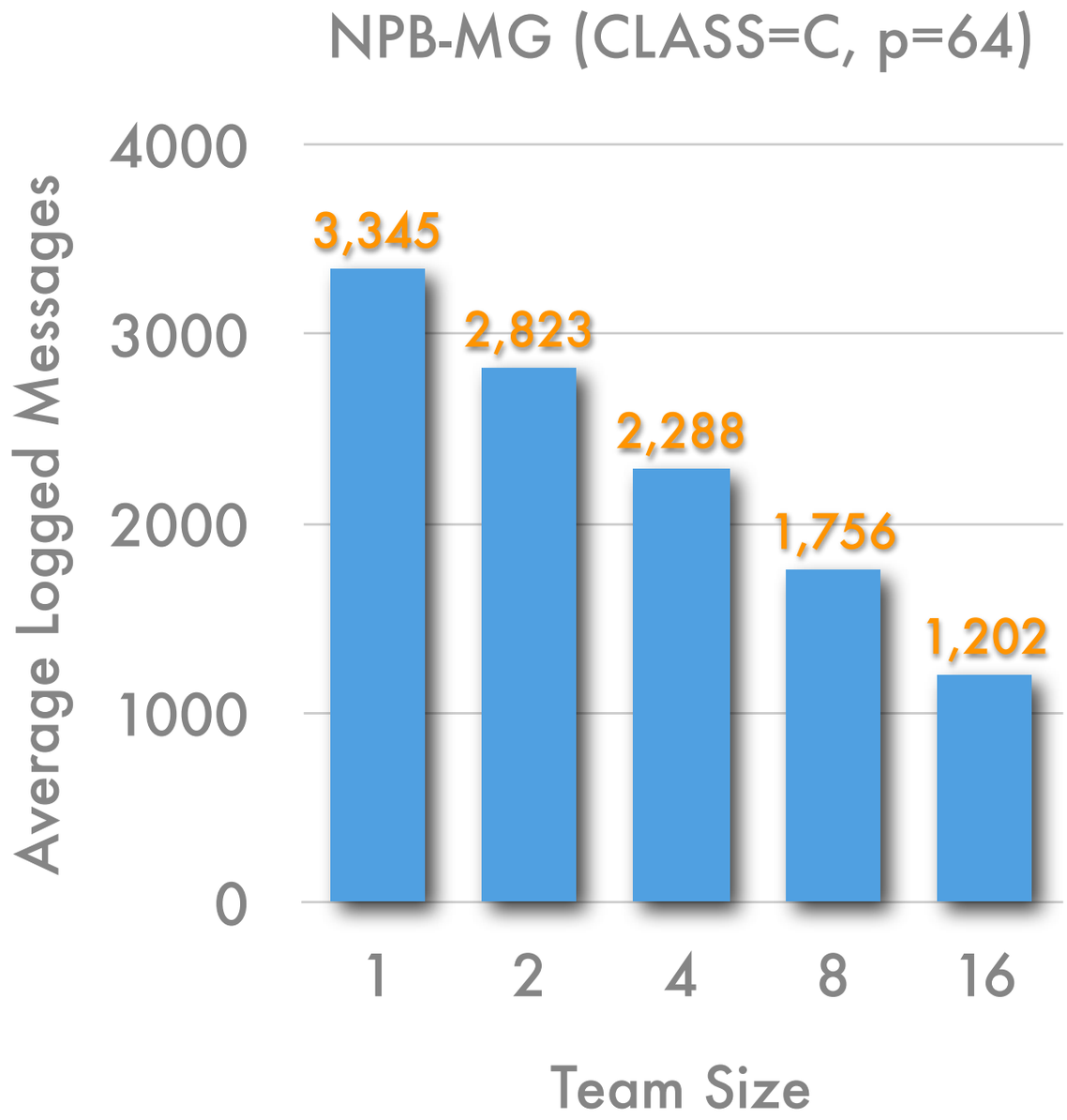


Reproduced with permission from http://proactive.inria.fr

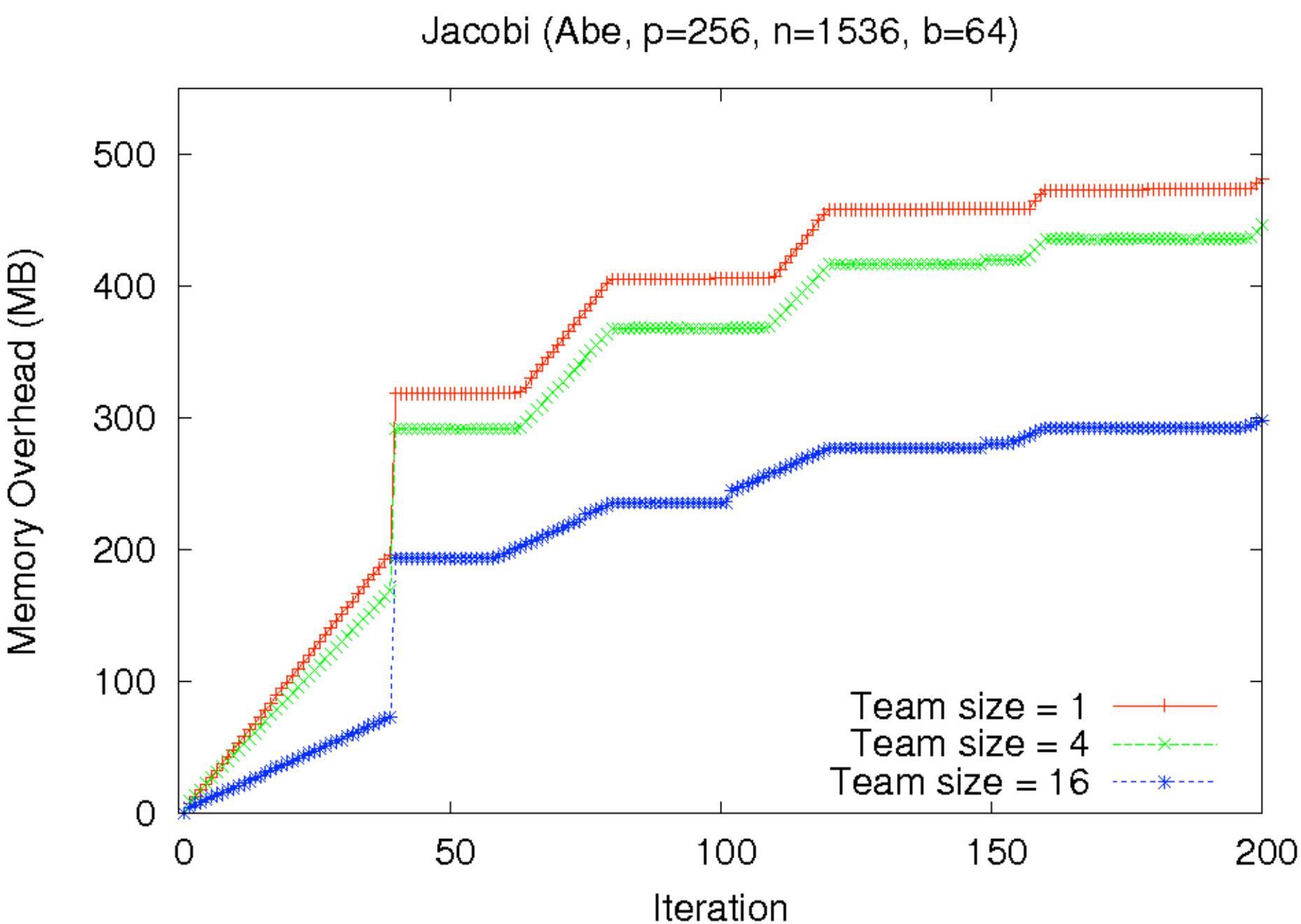
Resilience 2010, Melbourne, Australia

Message Log Reduction



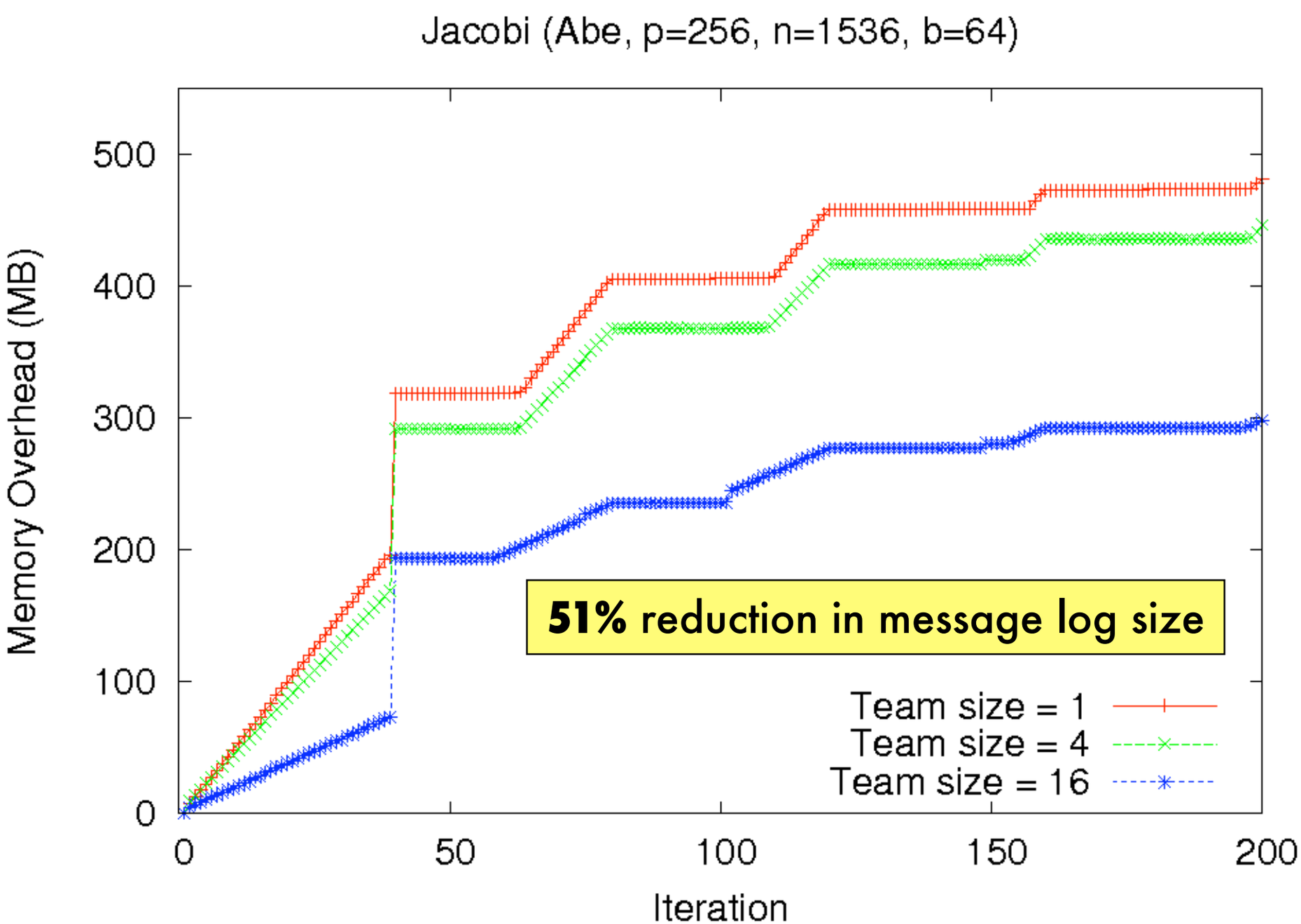


Stencil 3D



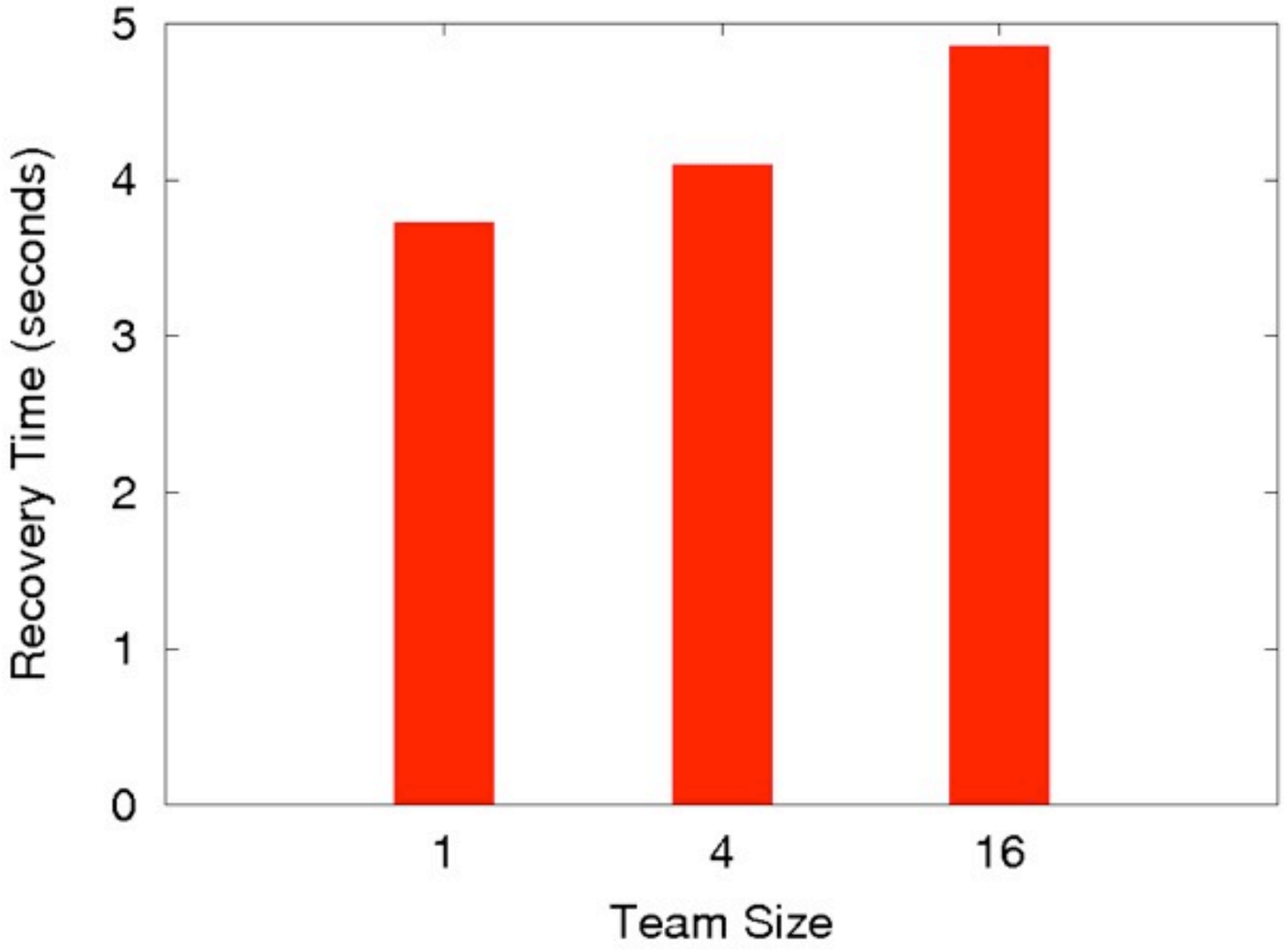
Resilience 2010, Melbourne, Australia

Stencil 3D



Resilience 2010, Melbourne, Australia

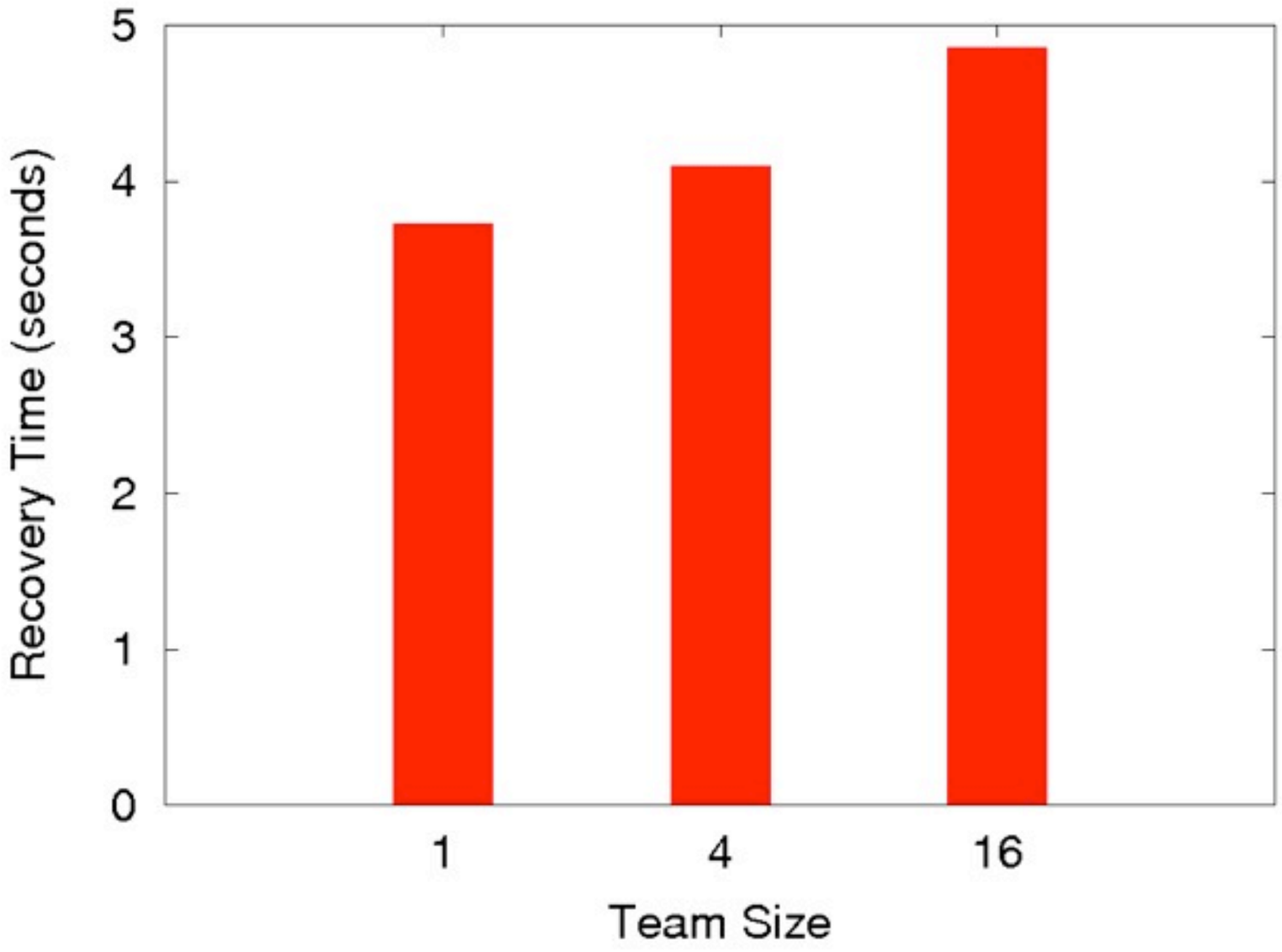
Recovery Time



Resilience 2010, Melbourne, Australia

Recovery Time

30% increase in recovery time



Resilience 2010, Melbourne, Australia

Research Questions

- Highly connected objects should belong to the same team.
  - Exploit communication graph, dynamic groups, team-aware load balancer.
- Processor teams vs object teams?
- Overlapping teams?

Conclusions

- Team-based approach can substantially reduce the memory overhead.
- **Contribution**: team size acts as a middle point between two traditional techniques.
- It can be used in conjunction with major message logging protocols.

Future Work

- Enrich **Team**-based Approach.
  - Smarter team formation.
  - Coupling with load balancer.
  - Dealing with correlated failures.
- **SMP**-aware fault tolerance.
- Larger Charm++ **applications**.

Acknowledgments

- **US Department of Energy**, FastOS Program (Colony-1 and Colony-2 projects).
- **NSF/NCSA**: Deployment efforts specific for Blue Waters.
- **Machine allocation**: TeraGrid MRAC: NCSA, TACC, ORNL.
- **Greg Bronevetsky** from LLNL.
- **Franck Cappello** from INRIA/Illinois.

Q&A