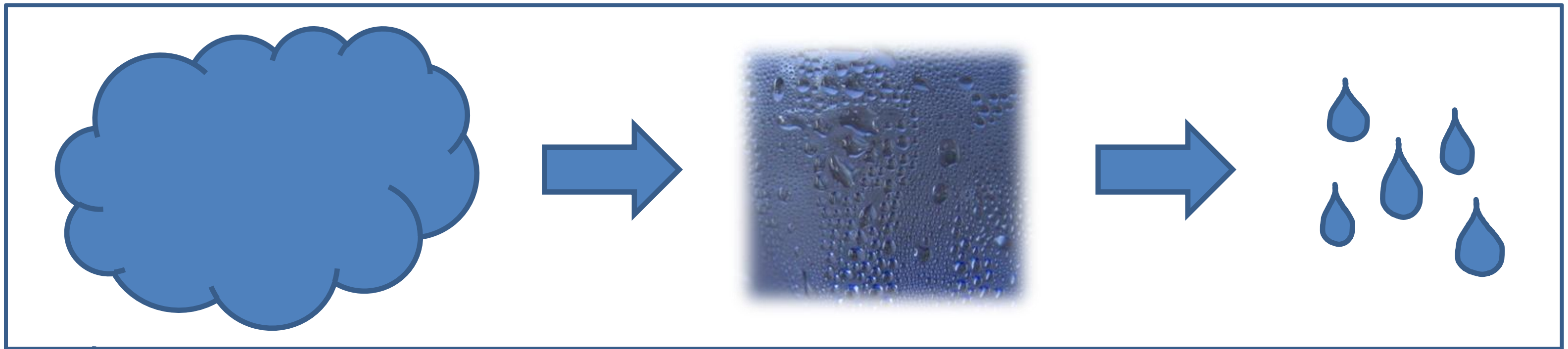
Condensing the cloud: running CIEL on many-core

Malte Schwarzkopf Derek G. Murray Steven Hand

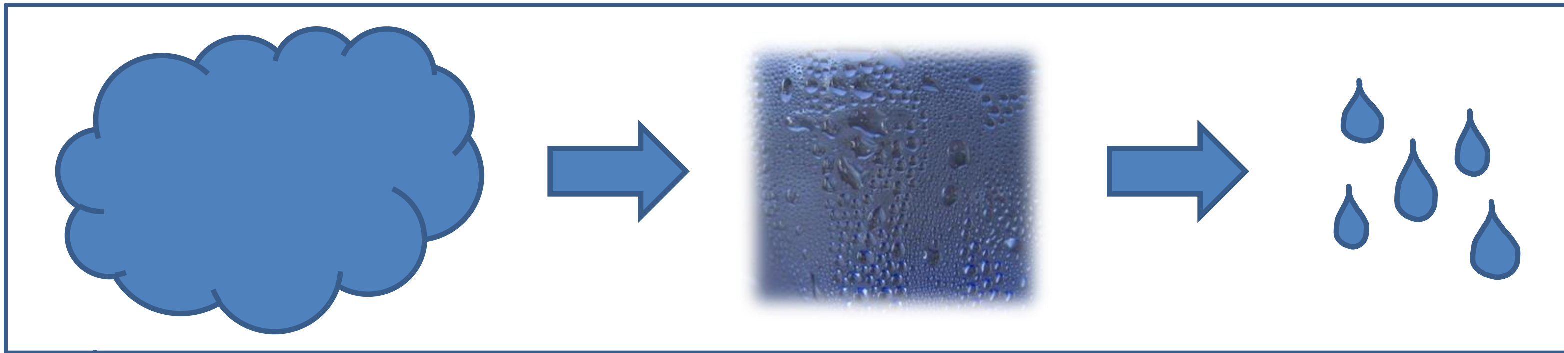
University of Cambridge



Condensing the cloud: running CIEL on many-core

Malte Schwarzkopf Derek G. Murray Steven Hand

University of Cambridge

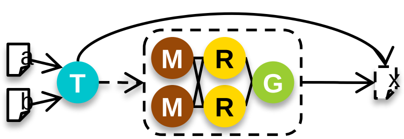


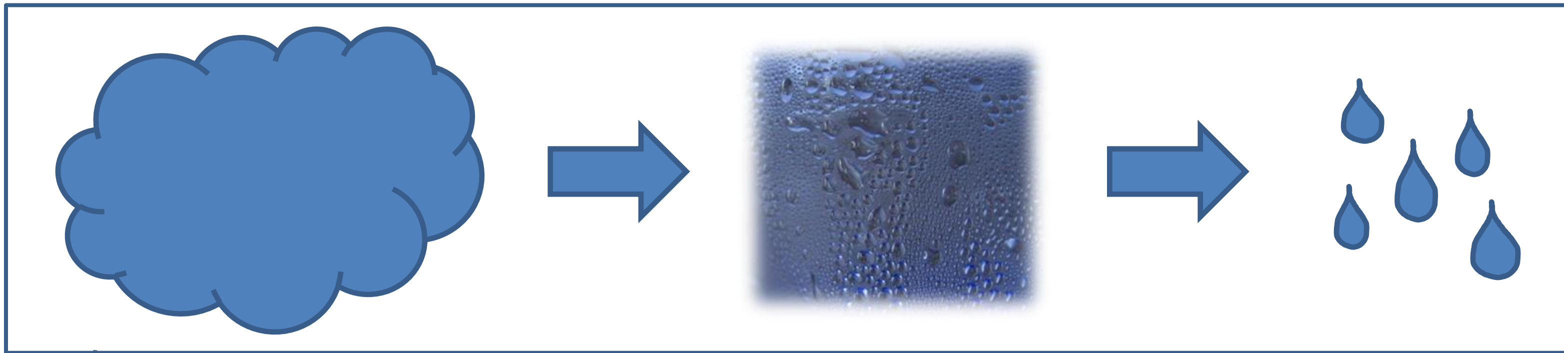
Condensing the cloud: running CIEL on many-core

Malte Schwarzkopf Derek G. Murray Steven Hand

University of Cambridge

CIEL

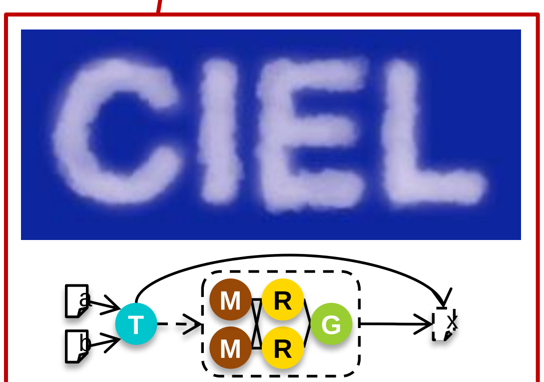


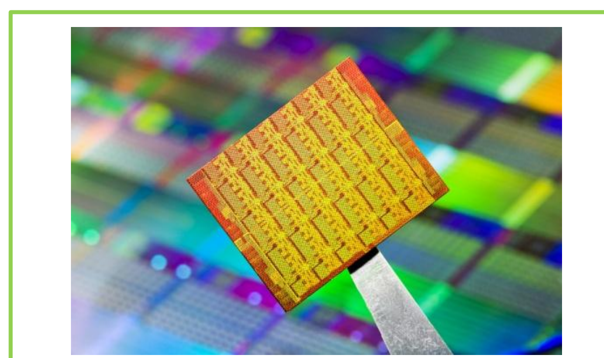


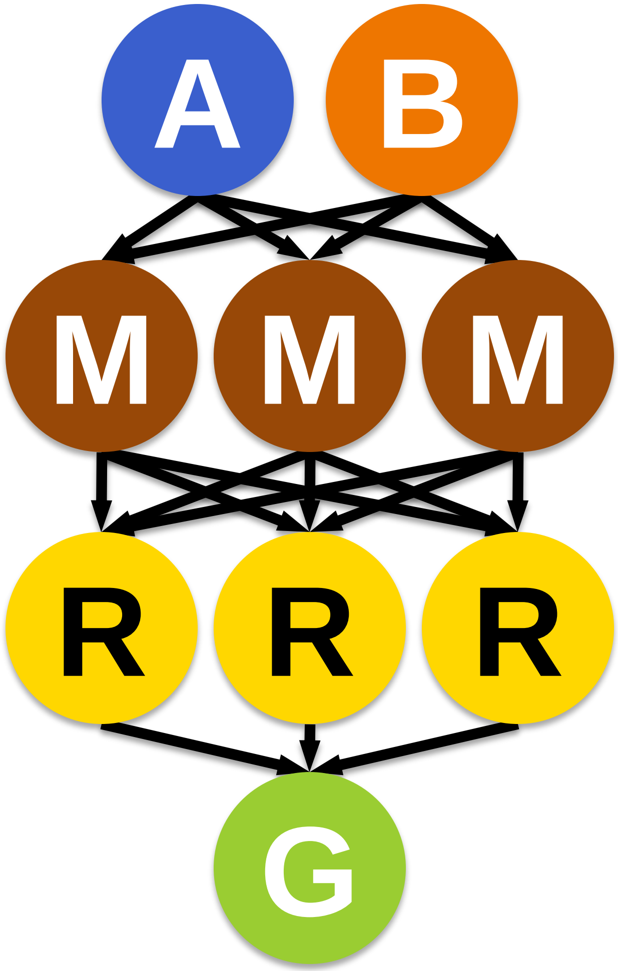
Condensing the cloud: running CIEL on many-core

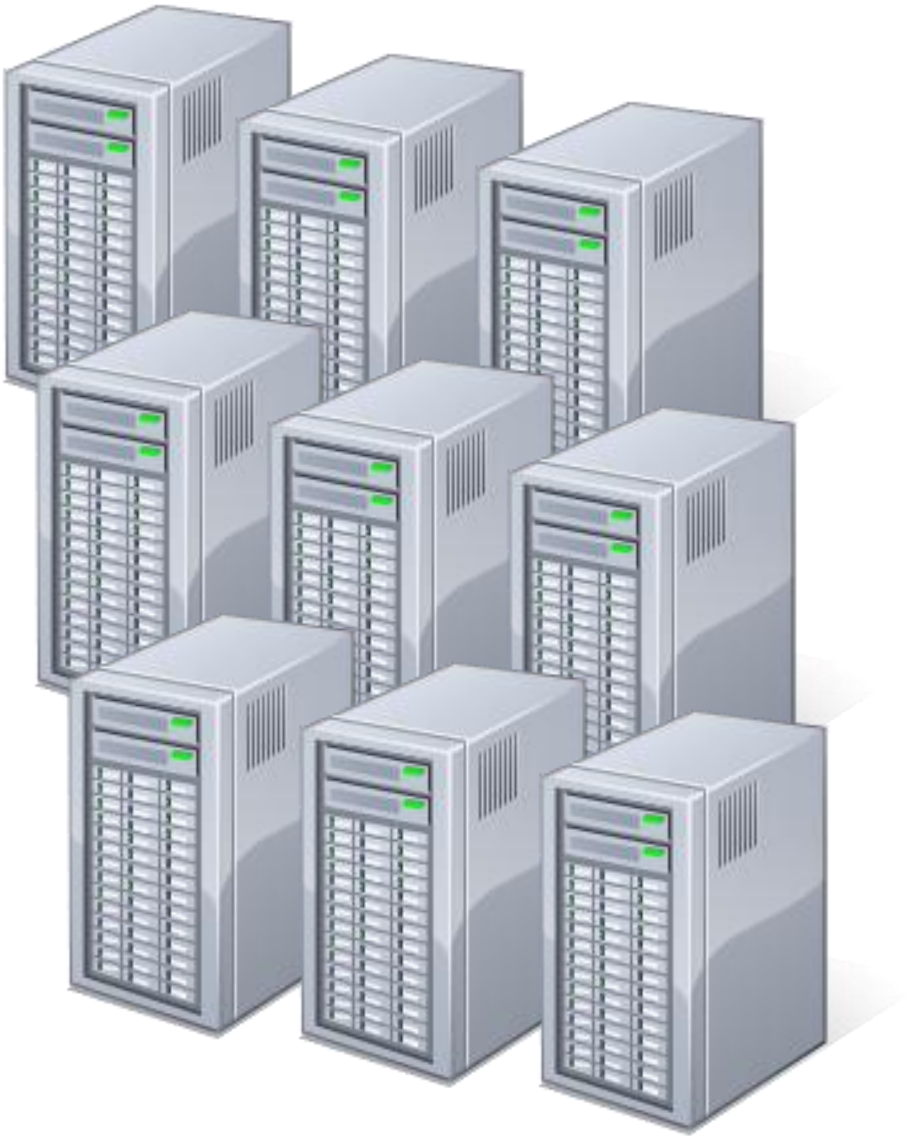
Malte Schwarzkopf Derek G. Murray Steven Hand

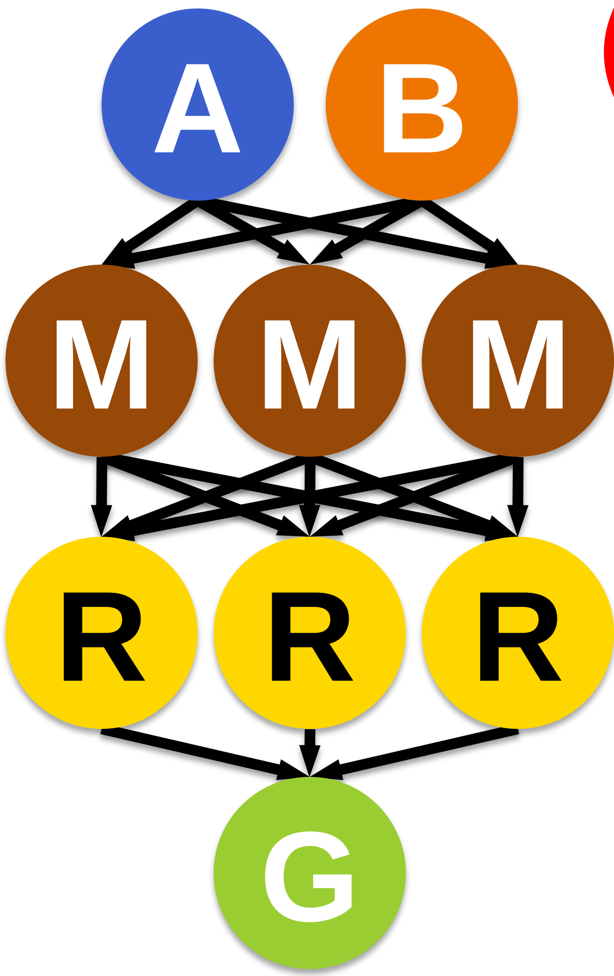
University of Cambridge

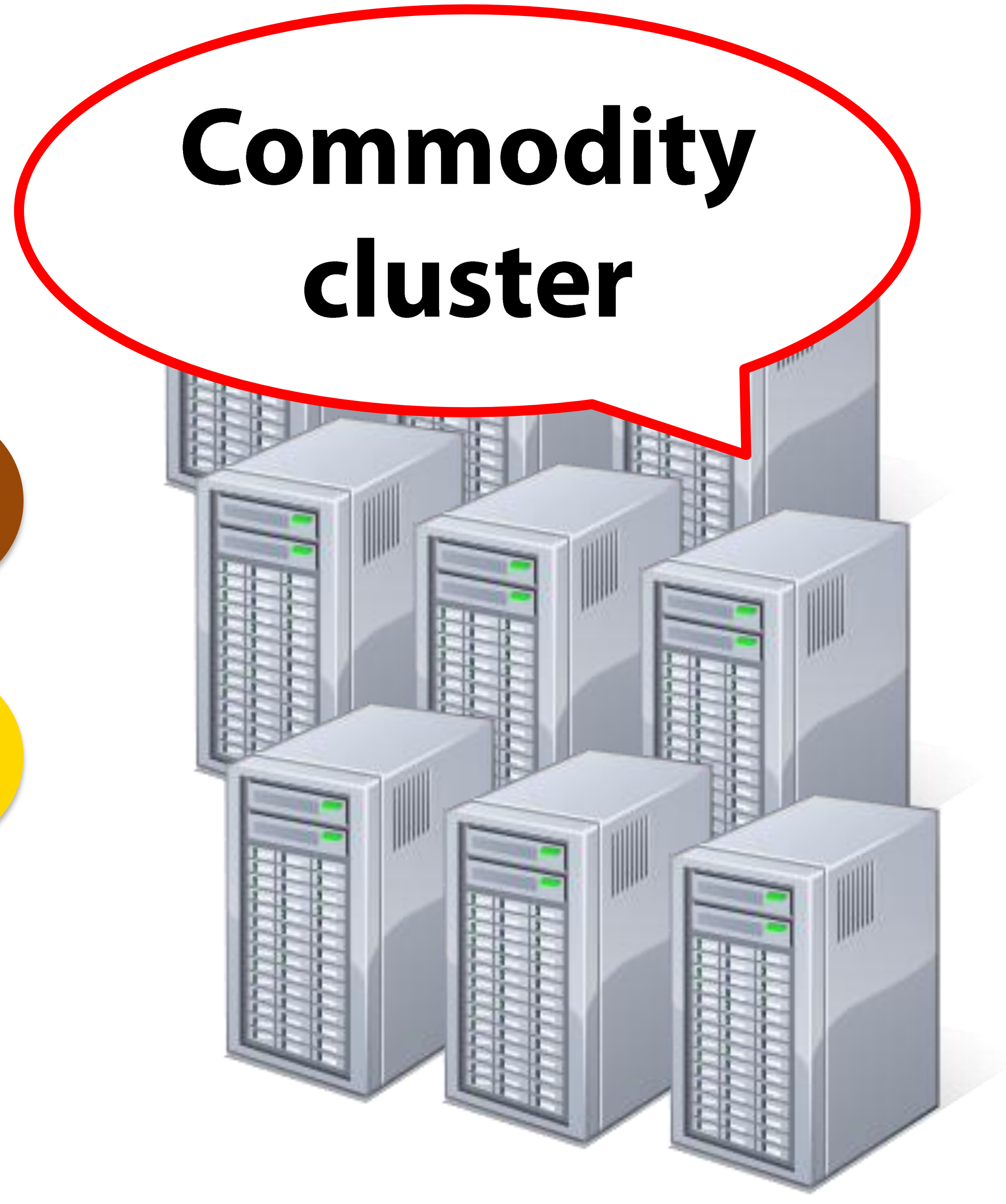


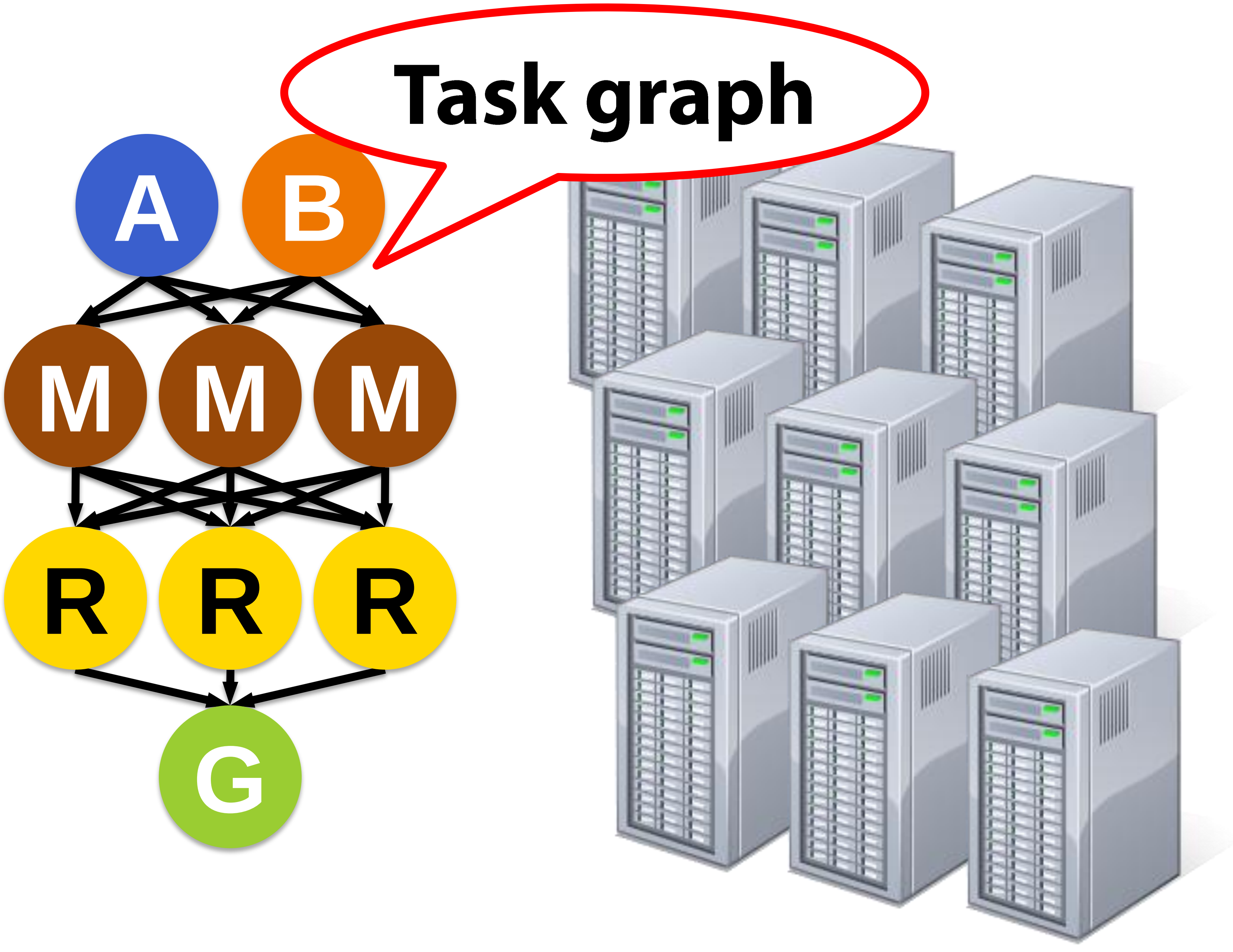



A
B
M
M
M
R
R
R
G




A
B
M
M
M
R
R
R
G


Commodity cluster


Task graph
A
B
M
M
M
R
R
R
G

Elastic resources

Elastic resources

Fault tolerance

Elastic resources

Fault tolerance

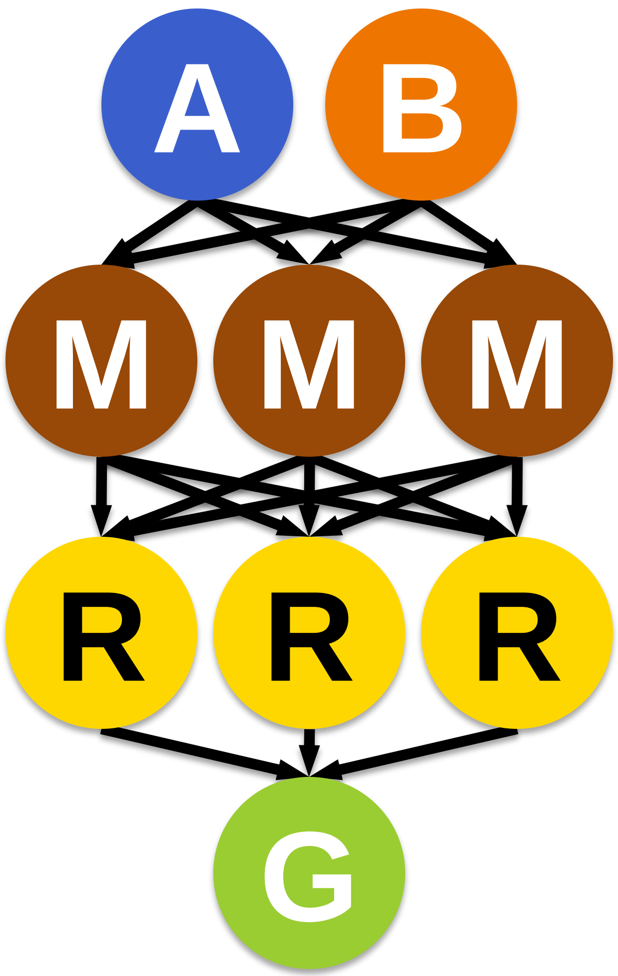
Sequential code

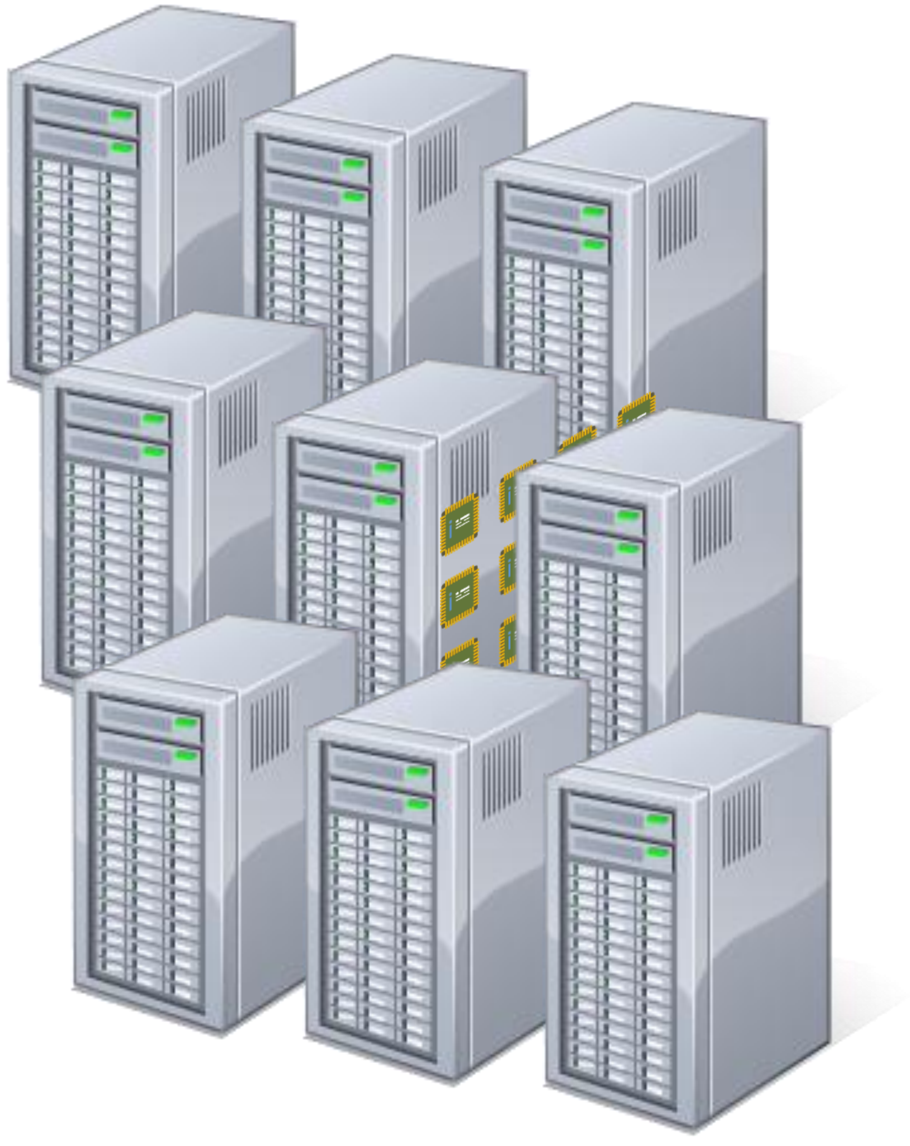
Elastic resources

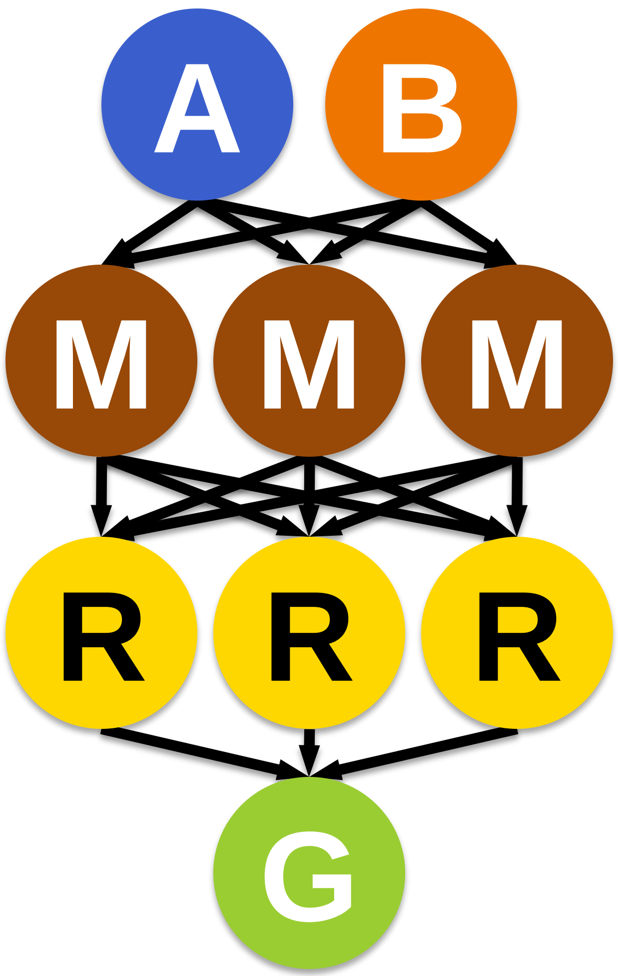
Fault tolerance

Sequential code

Deterministic parallelism


A
B
M
M
M
R
R
R
G



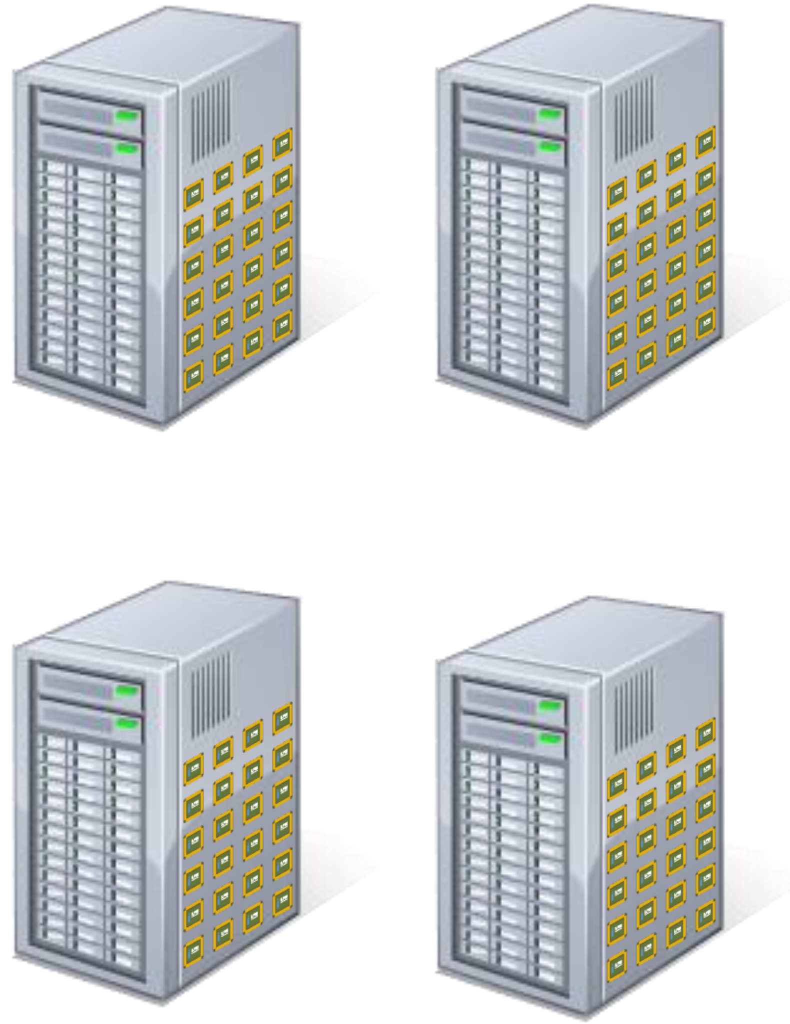

A
B
M
M
M
R
R
R
G

hadoop

CIEL

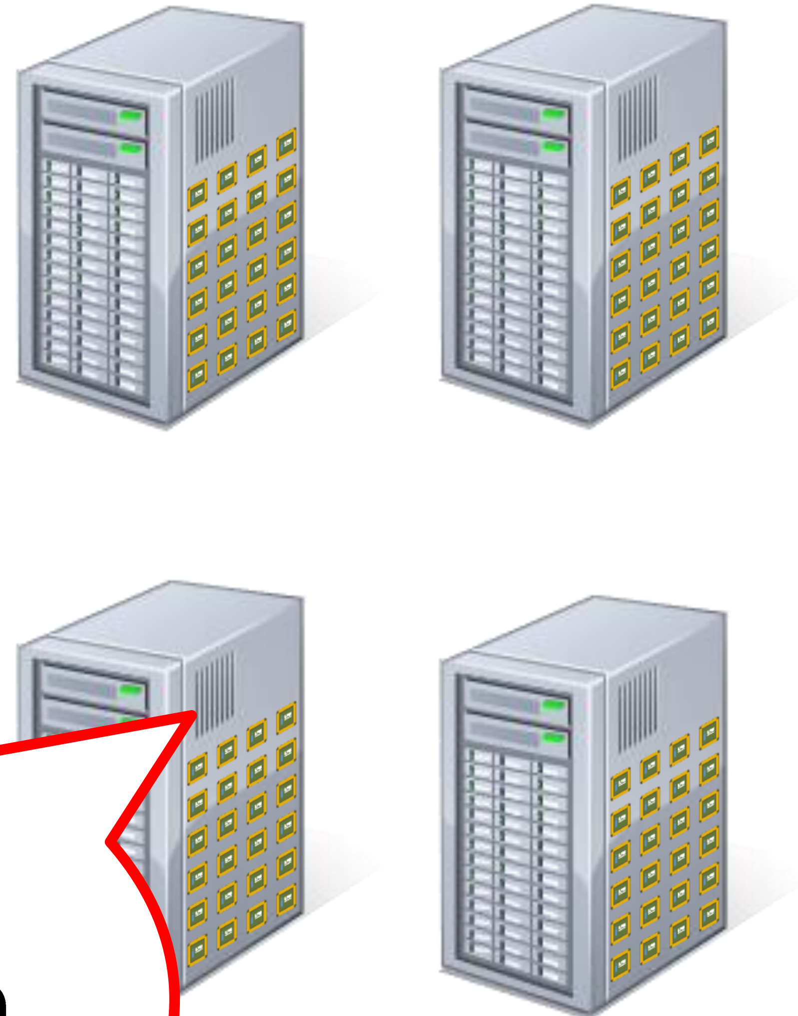
hadoop

CIEL



hadoop

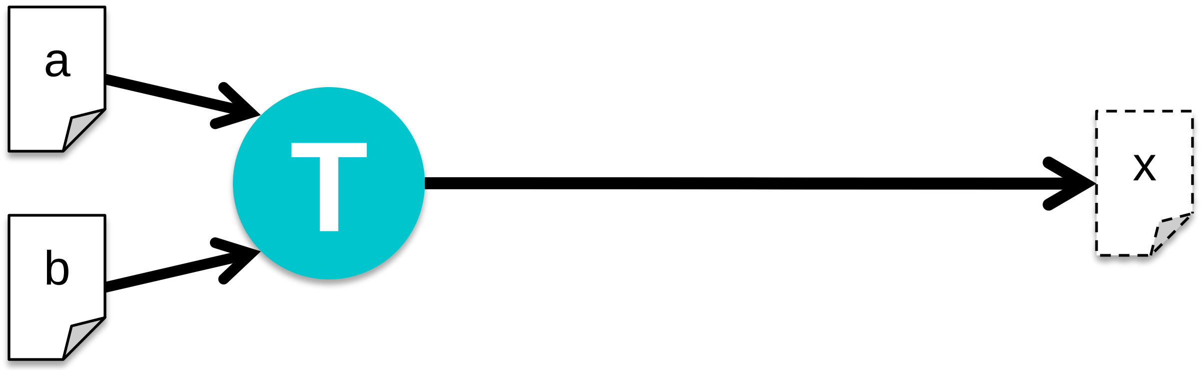
CIEL



Many-core workers in a cluster?

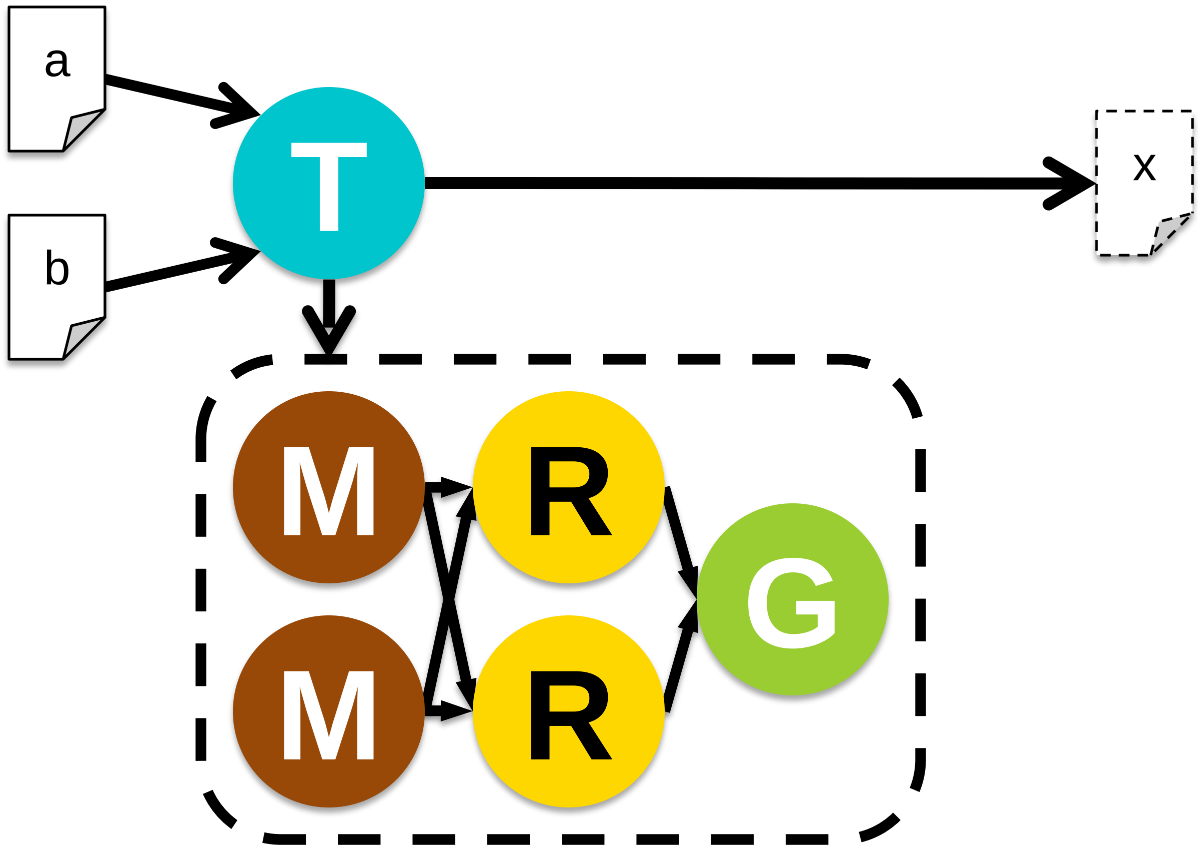
CIEL: dynamic task graphs

- Allow tasks to spawn more tasks



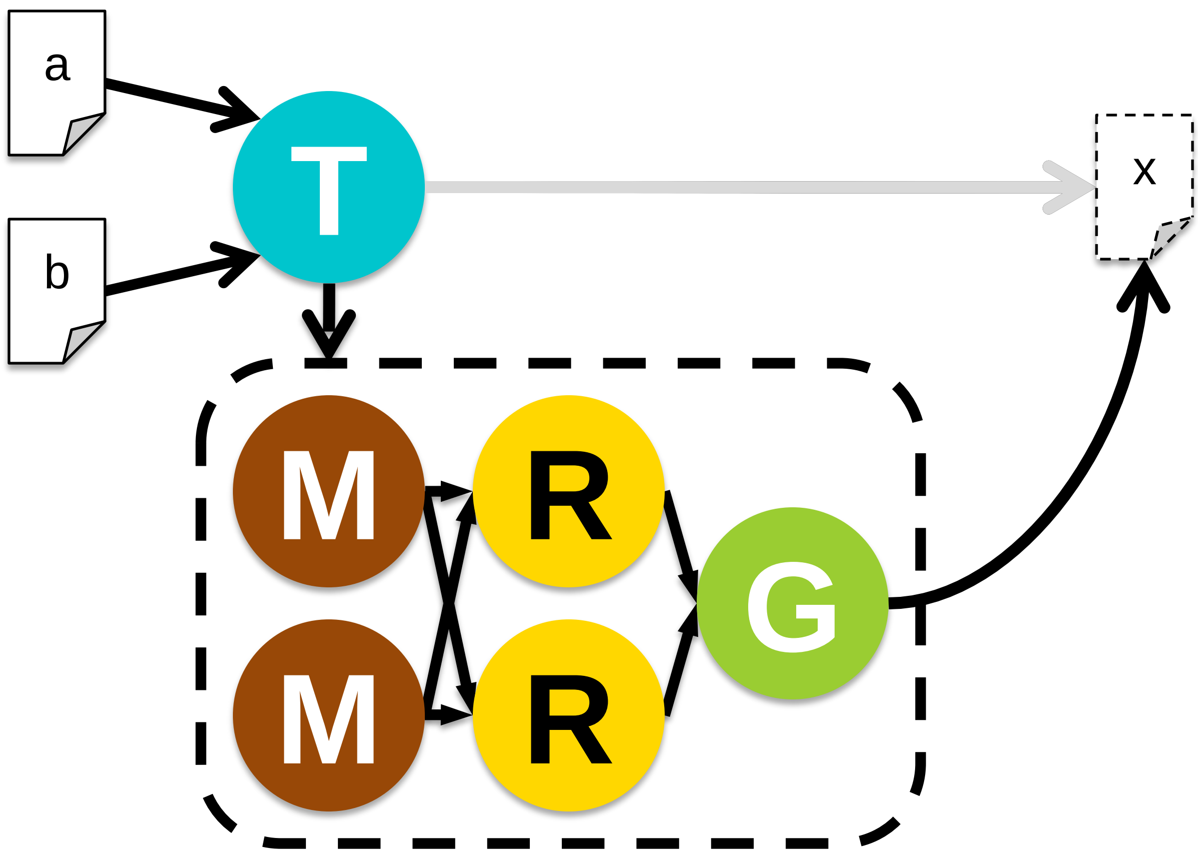
CIEL: dynamic task graphs

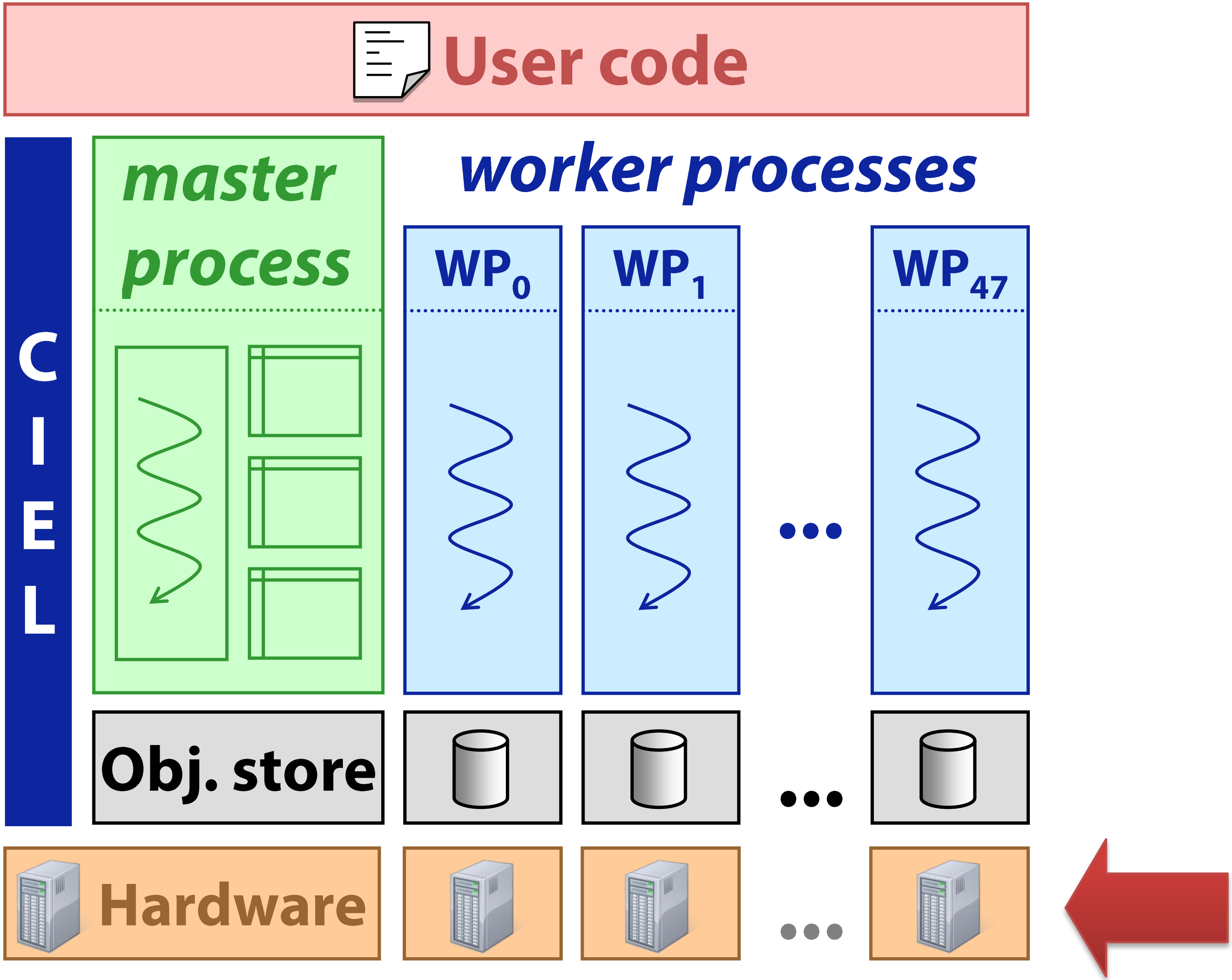
- Allow tasks to spawn more tasks

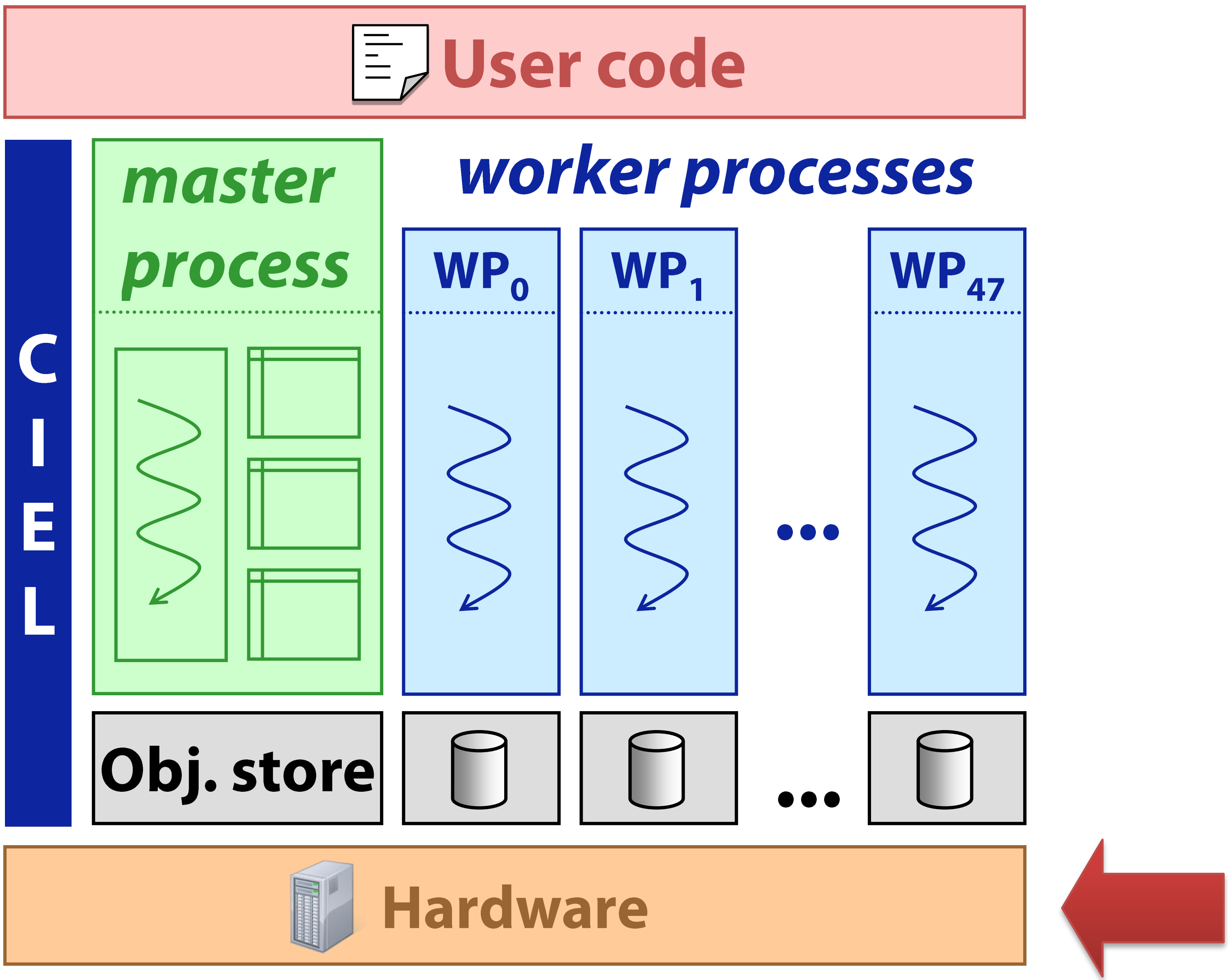


CIEL: dynamic task graphs

- Allow tasks to spawn more tasks




User code
master process
worker processes
WP0
WP1
WP47
C I E L
Obj. store
Hardware



User code
CIEL
master process
worker processes
WP_0
WP_1
WP_{47}
Obj. store
Hardware

Test environments

[1]

4x AMD Opteron 6168 „Magny-Cours“

- dodeca-core
- 1.9 GHz (64-bit), 512 KB L2, 12 MB L3
- 64 GB DDR3 @ 1333 MHz

[2]

Intel Single-Chip Cloud (SCC)

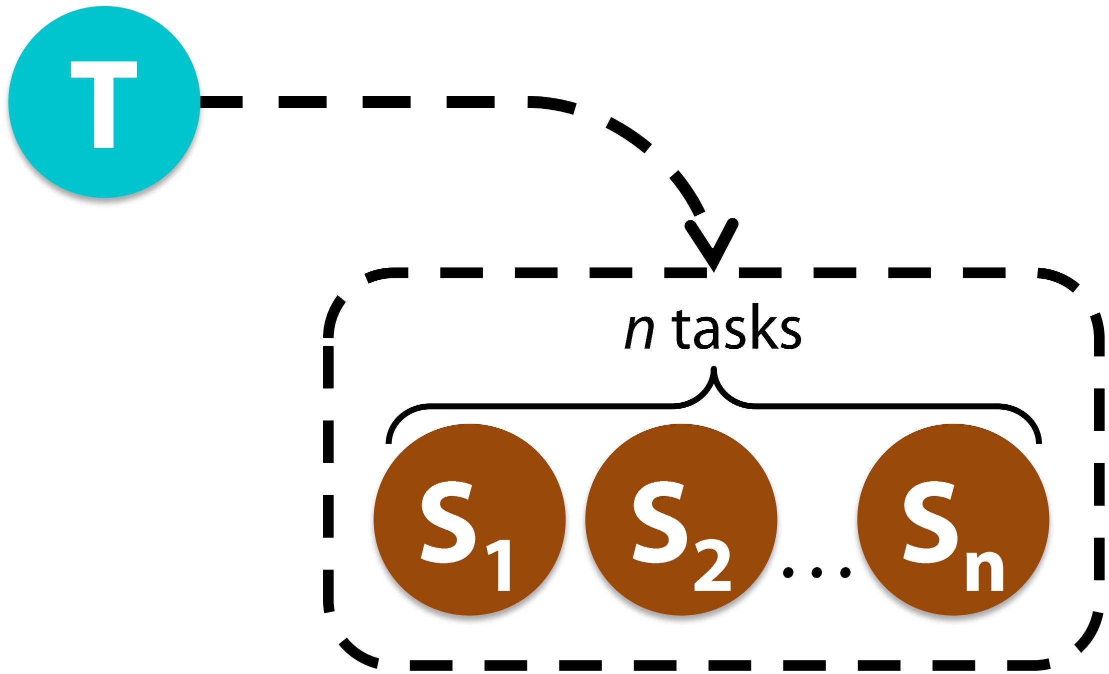
Experimental!

- tetracontakaiocta-core
- 533 MHz (32-bit P54C), 16 KB L1, 256 KB L2
- 64 GB DDR3 @ 800 MHz

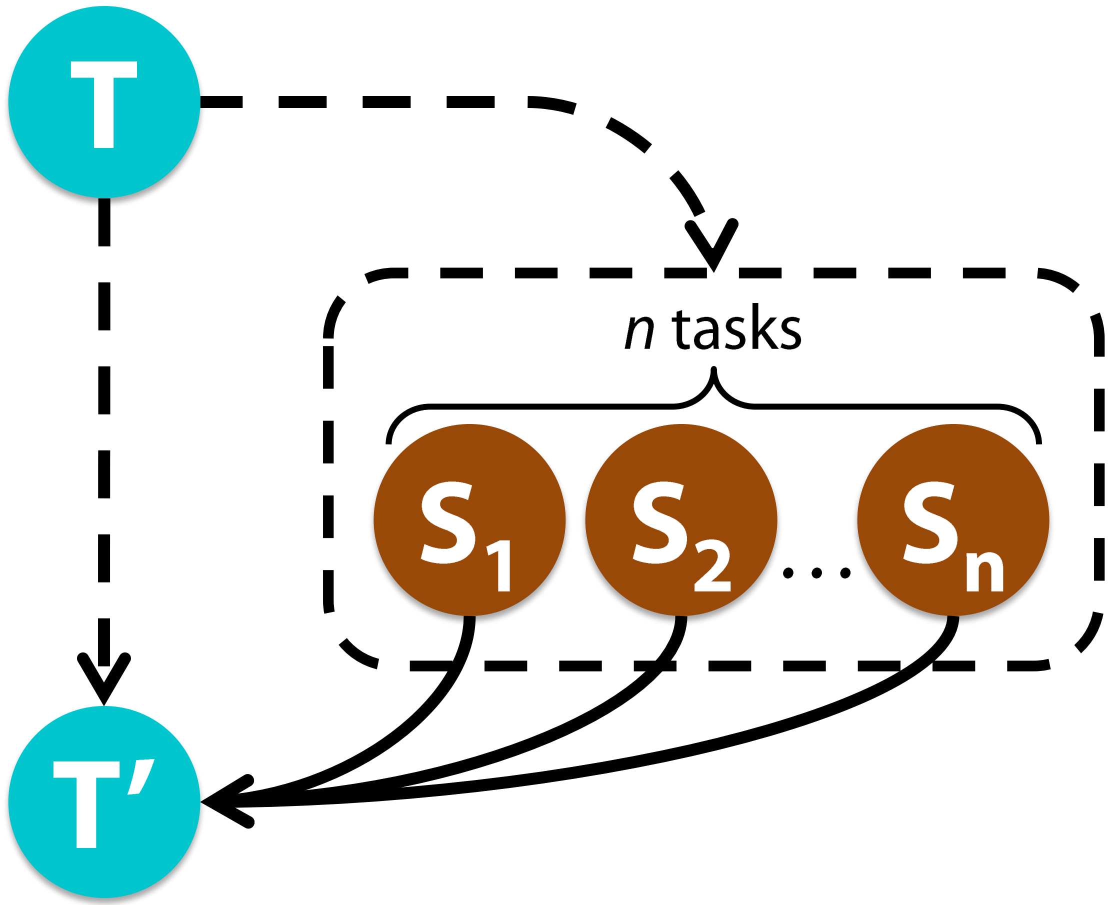
[1] Image from http://en.wikipedia.org/wiki/File:Opteron_logo.png

[2] Image from http://www.intel.com/communities/pix/marc/scc-v-wafer2.jpg

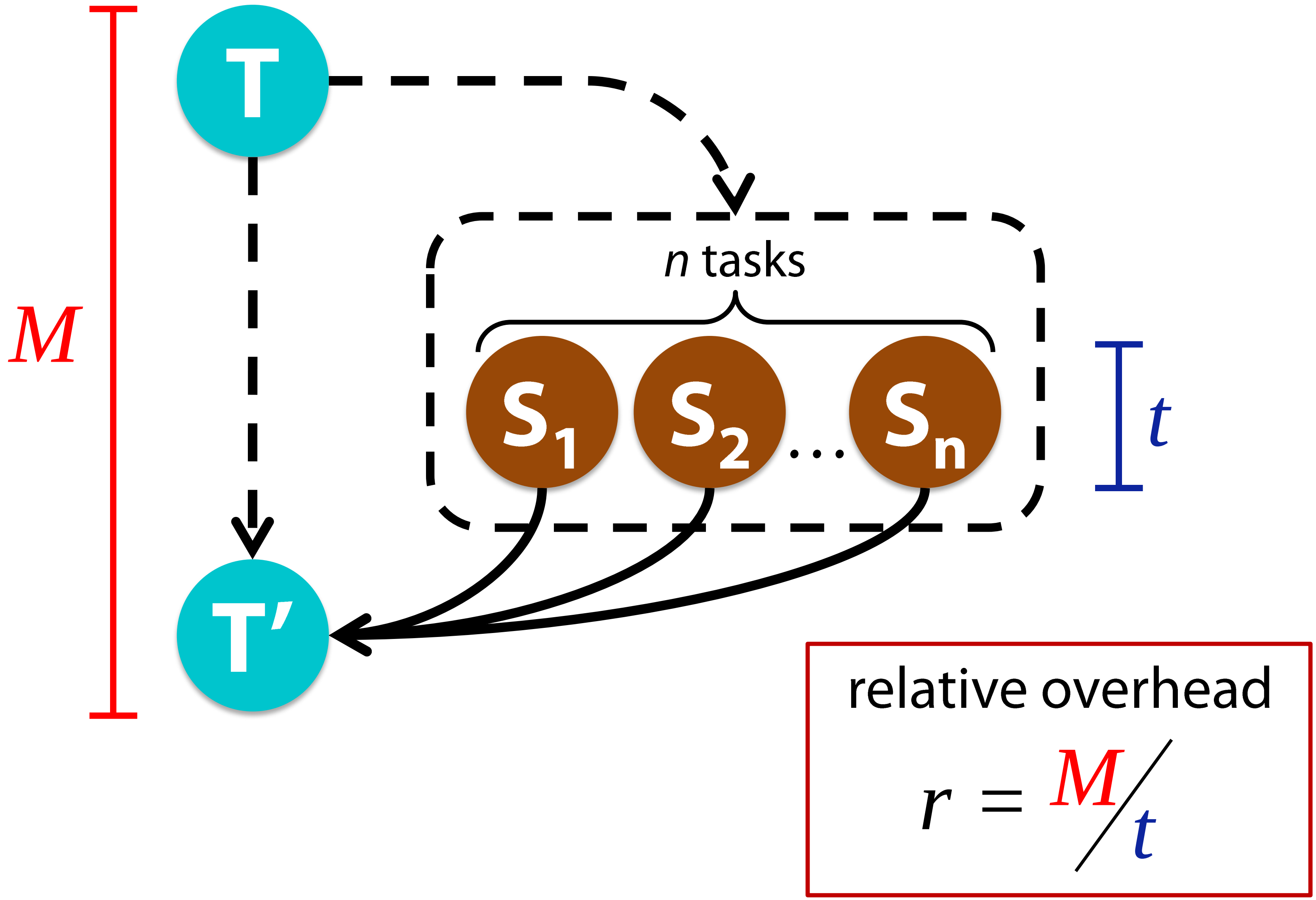
`timespin` µ-benchmark



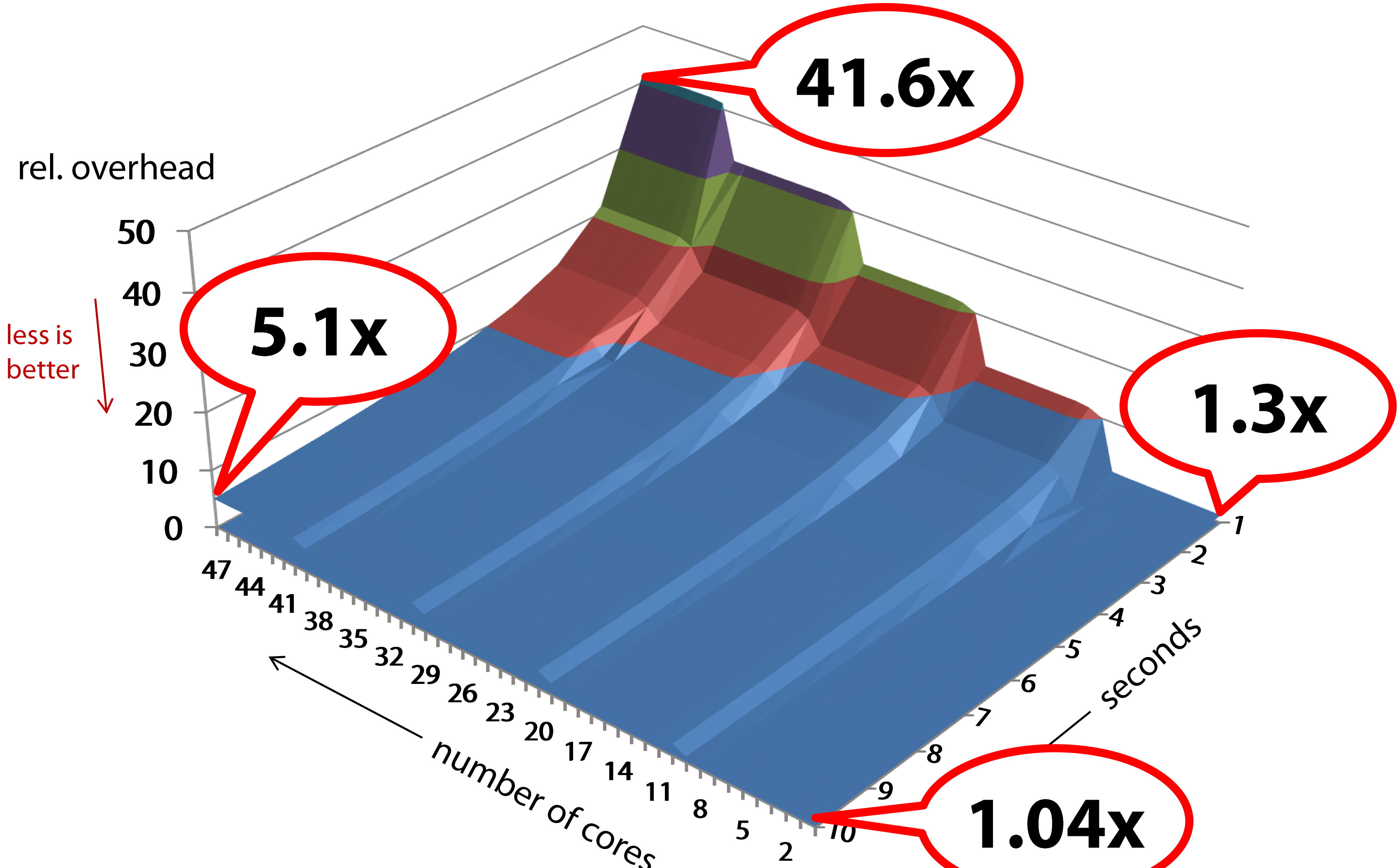
timespin μ-benchmark



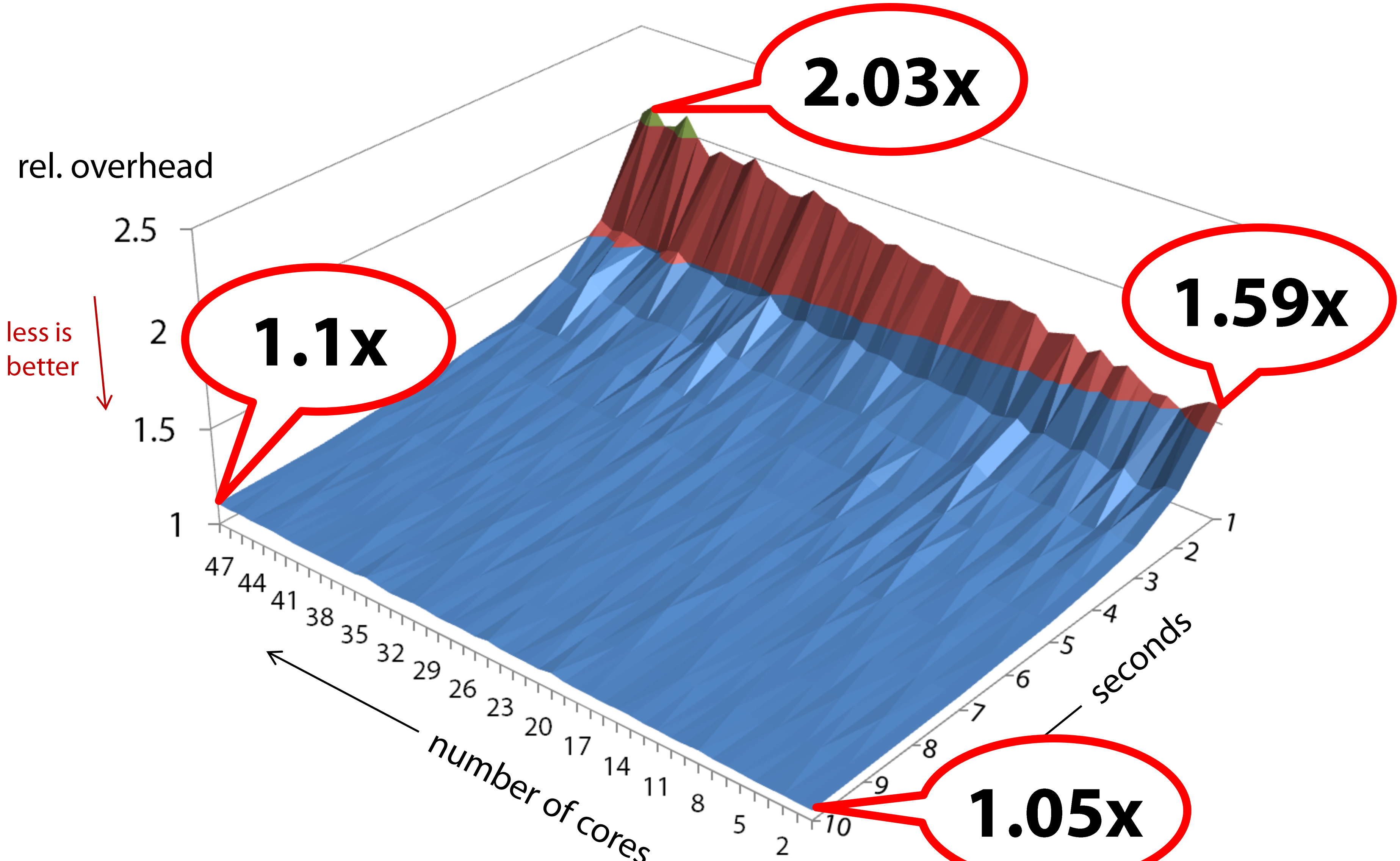
timespin μ-benchmark



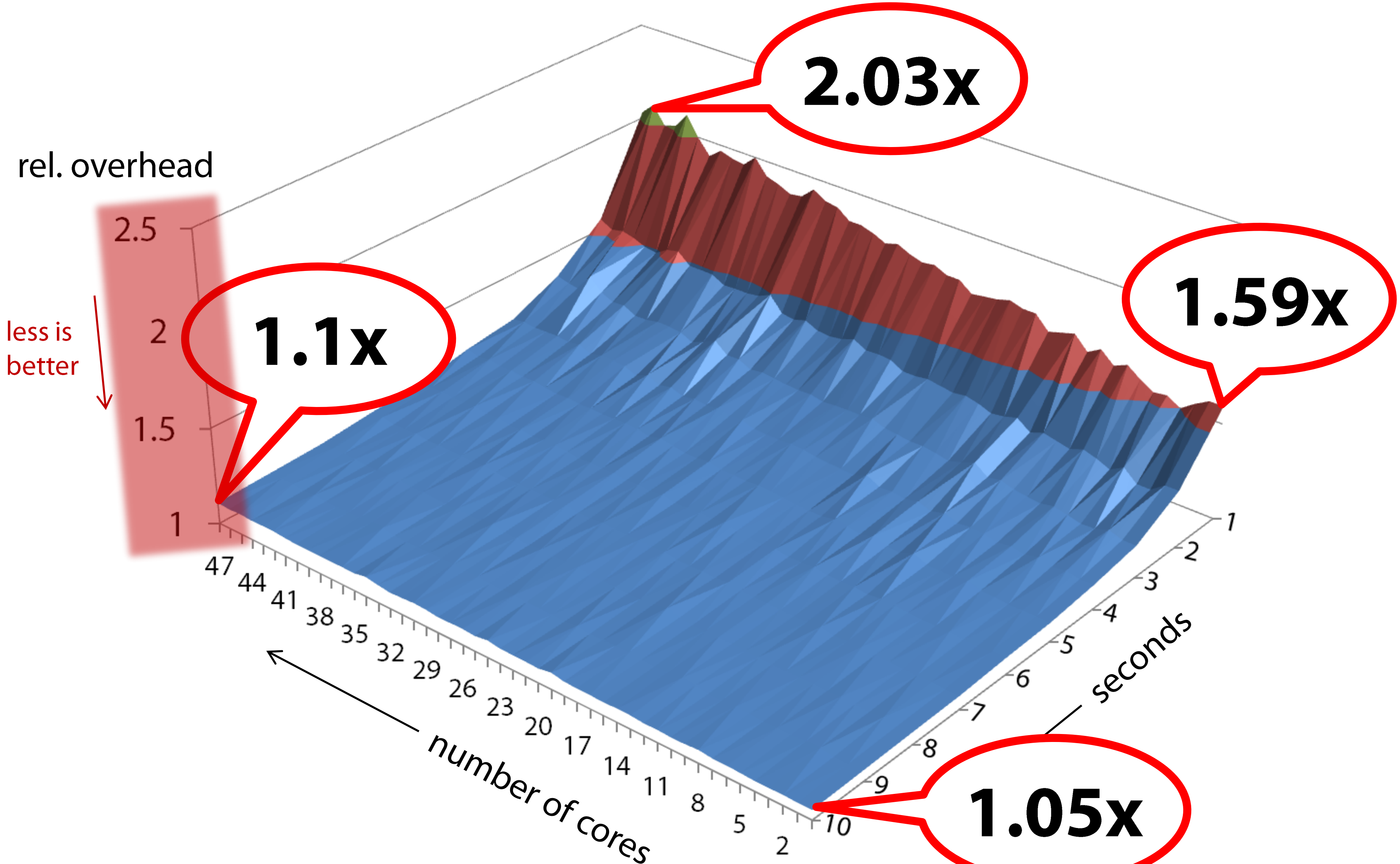
`timespin` on unmodified CIEL



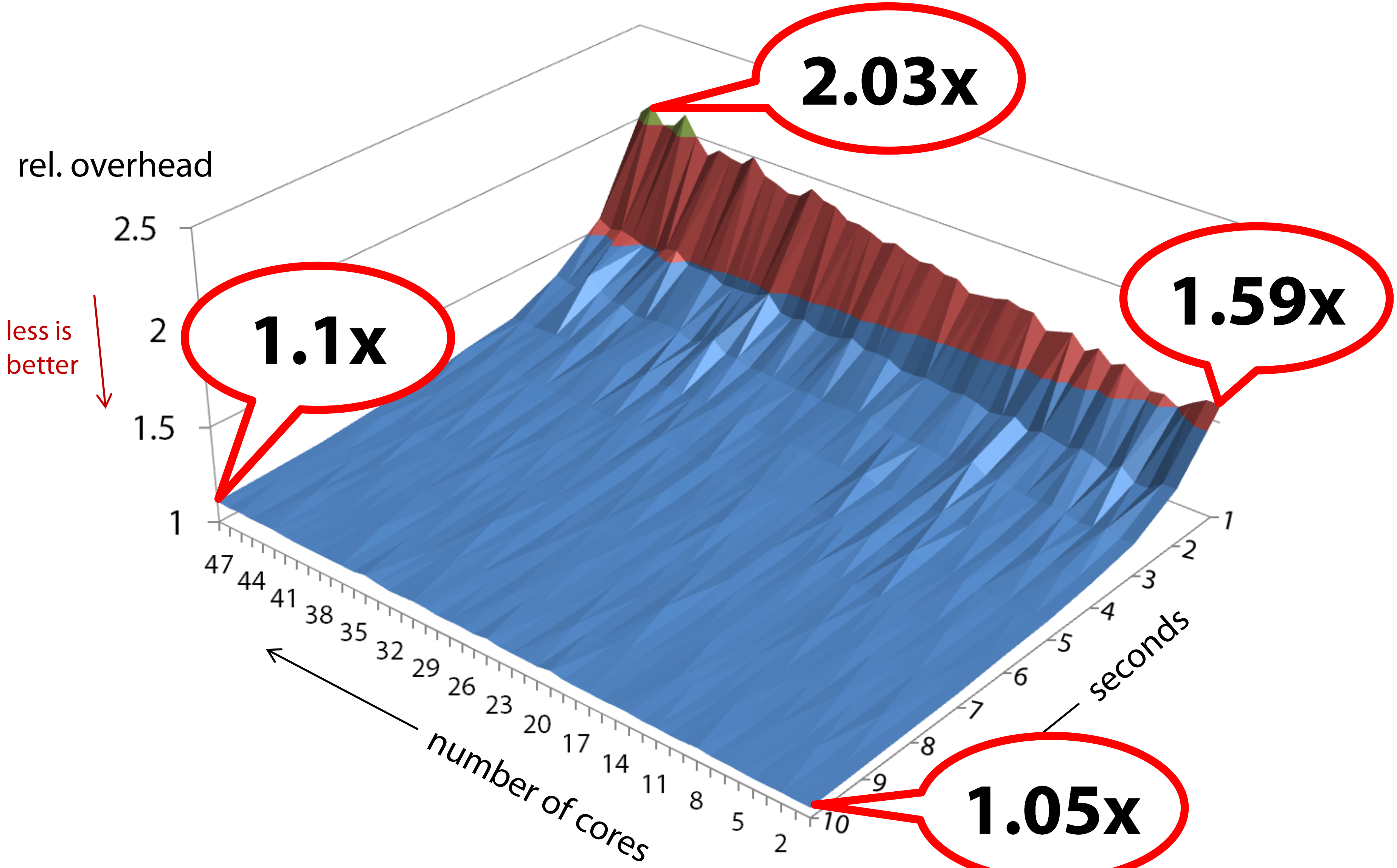
timespin, using lighttpd

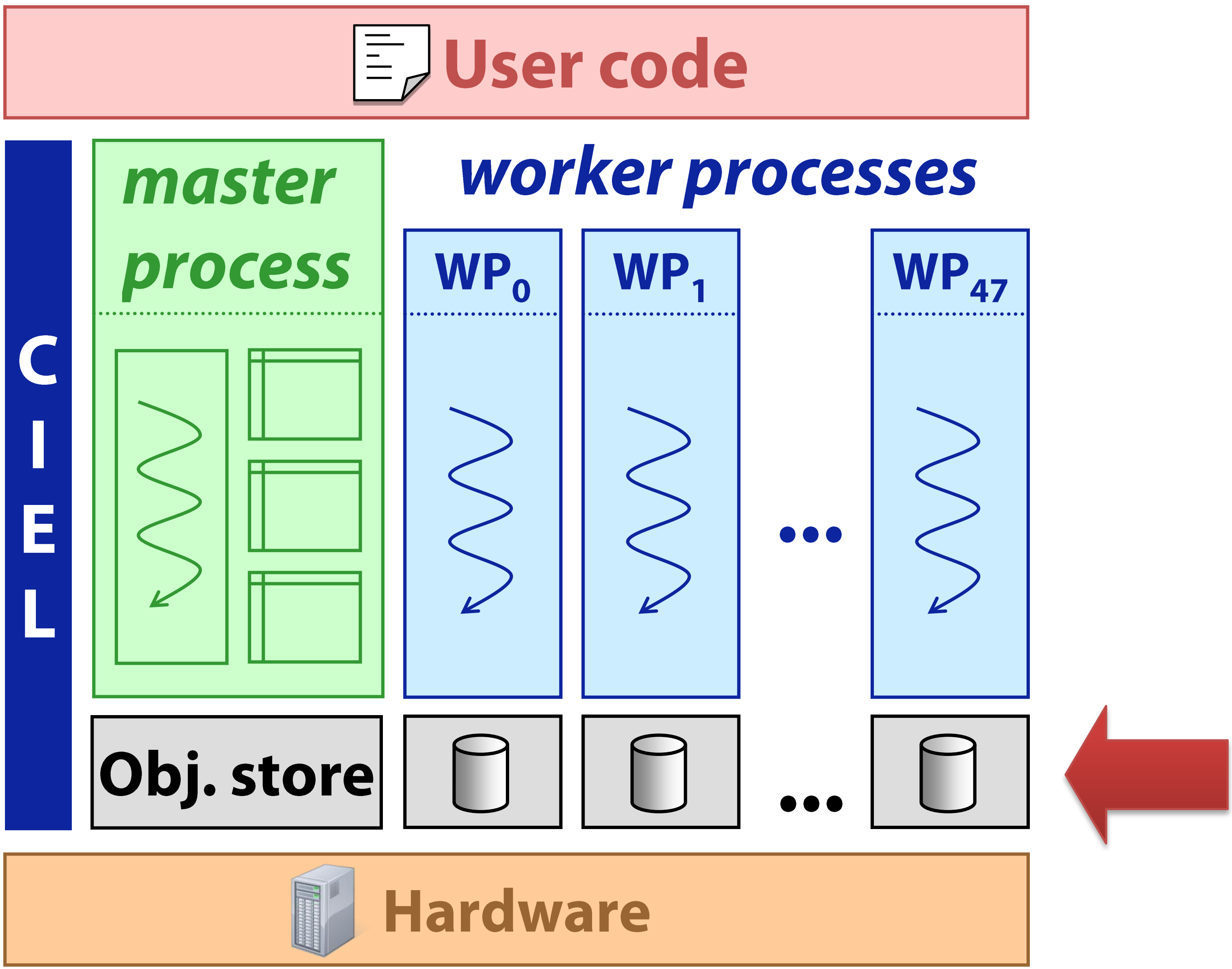


timespin, using lighttpd

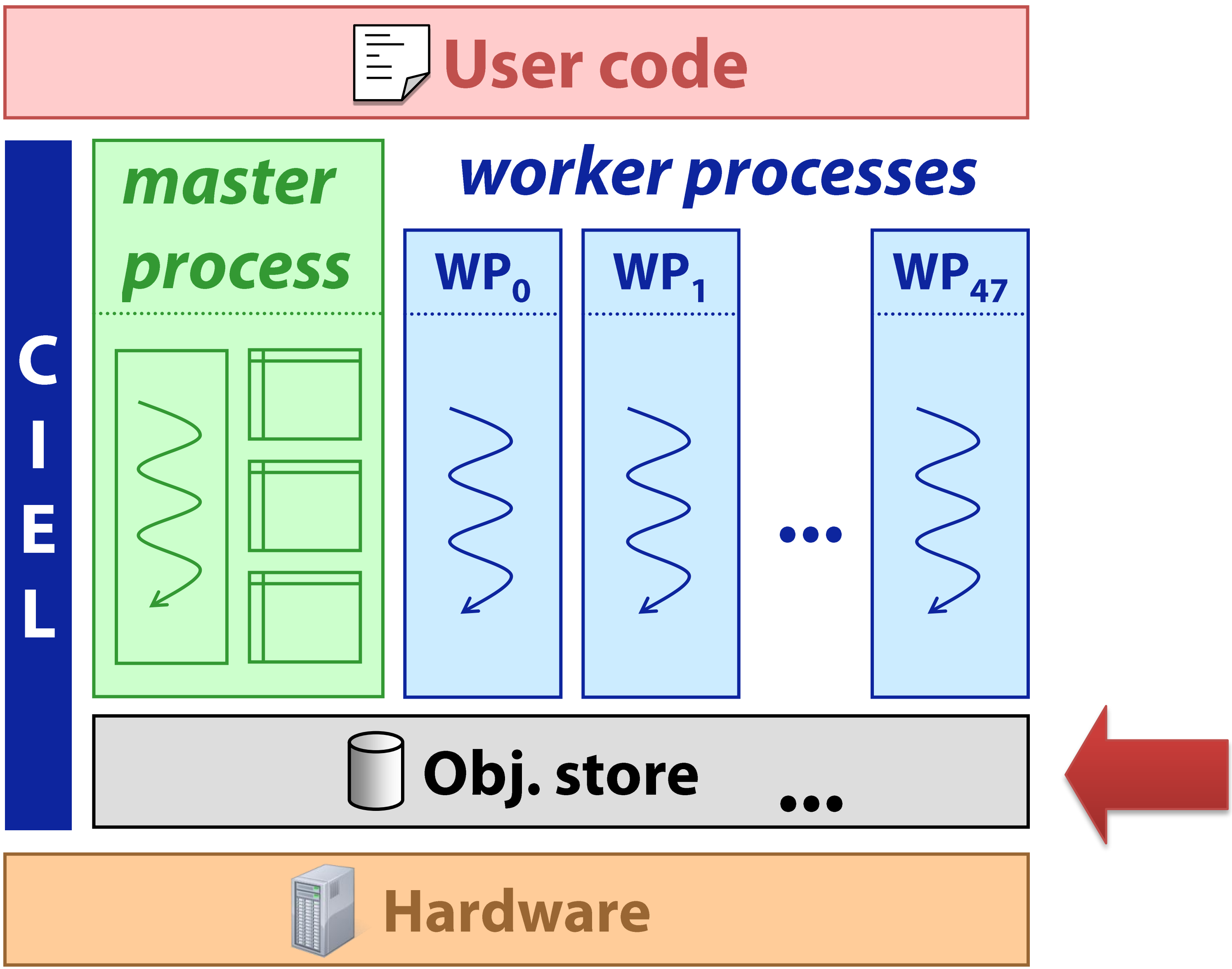


timespin, using lighttpd

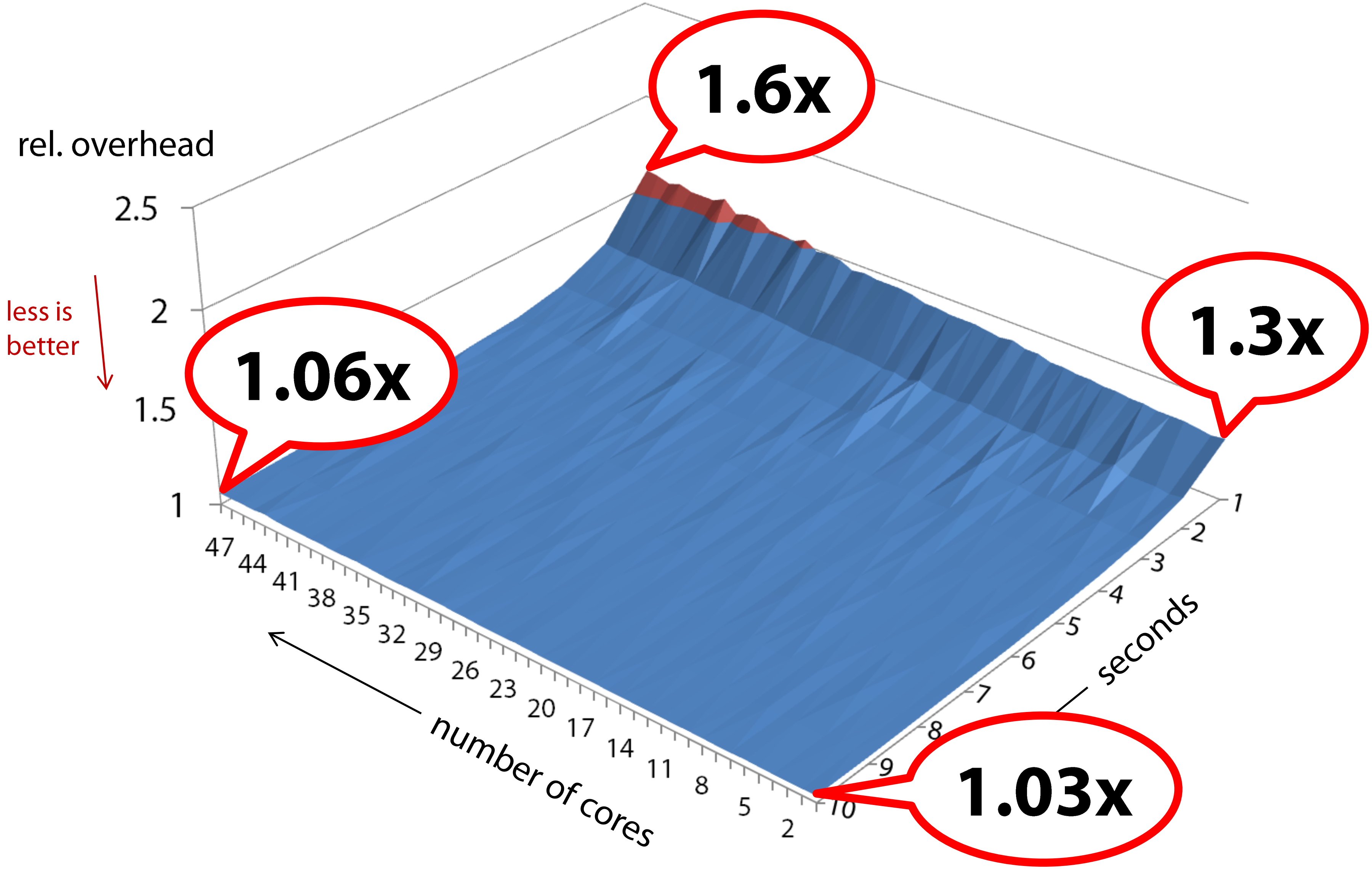


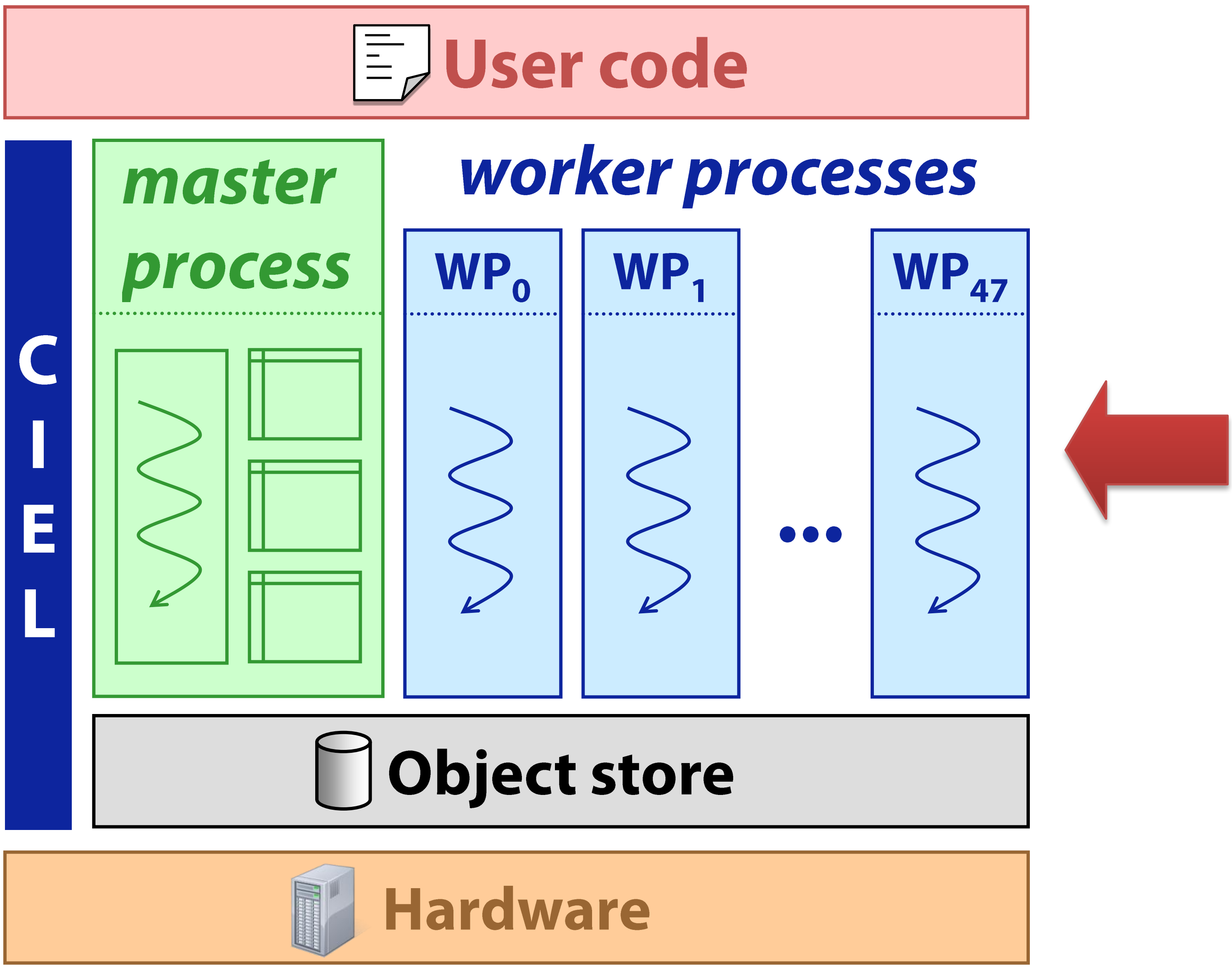


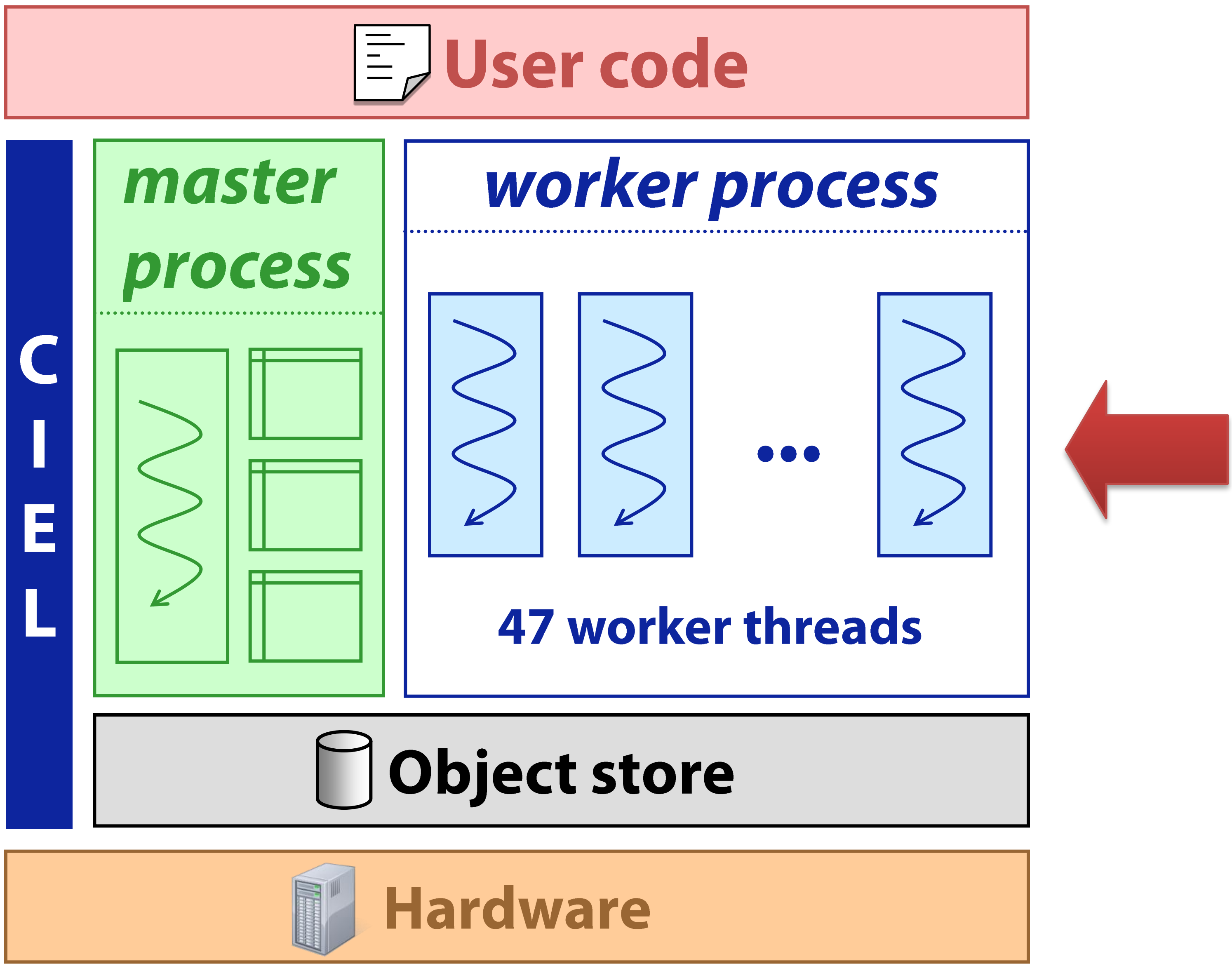
User code
master process
worker processes
CIEL
WP0
WP1
...
WP47
Obj. store
...
Hardware


User code
C I E L
master process
worker processes
WP_0
WP_1
...
WP_{47}
Obj. store
...
Hardware

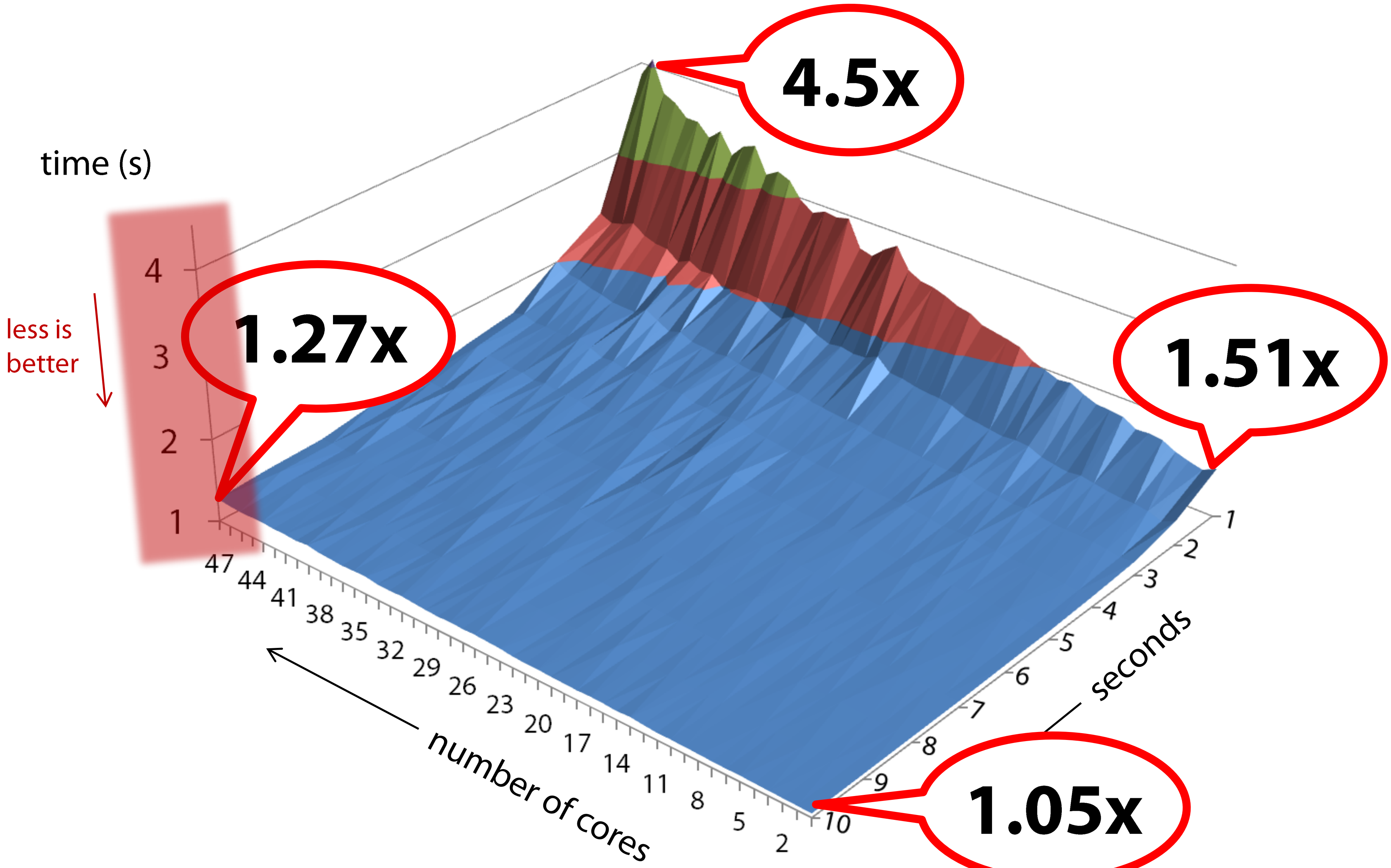
timespin, shared object store

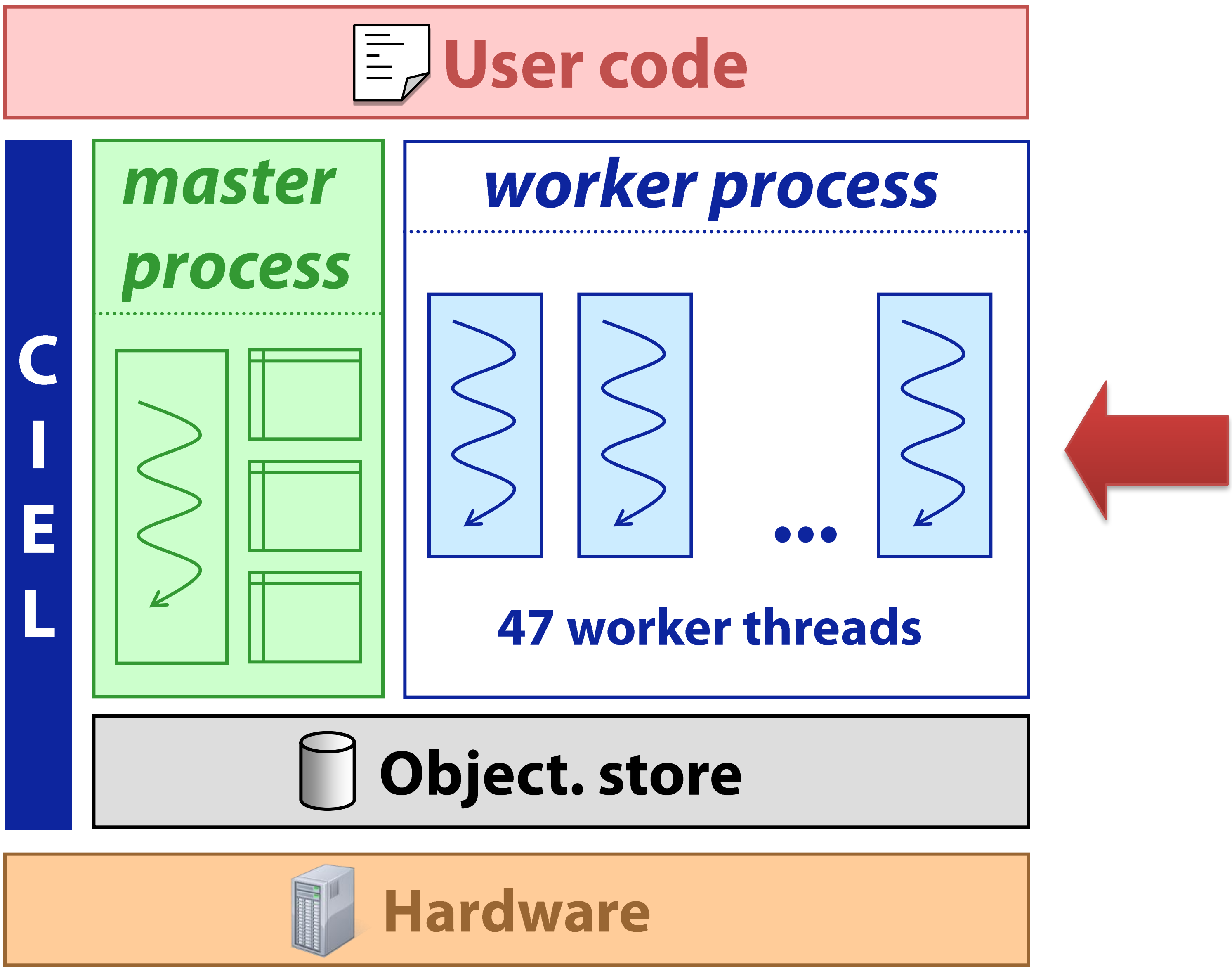


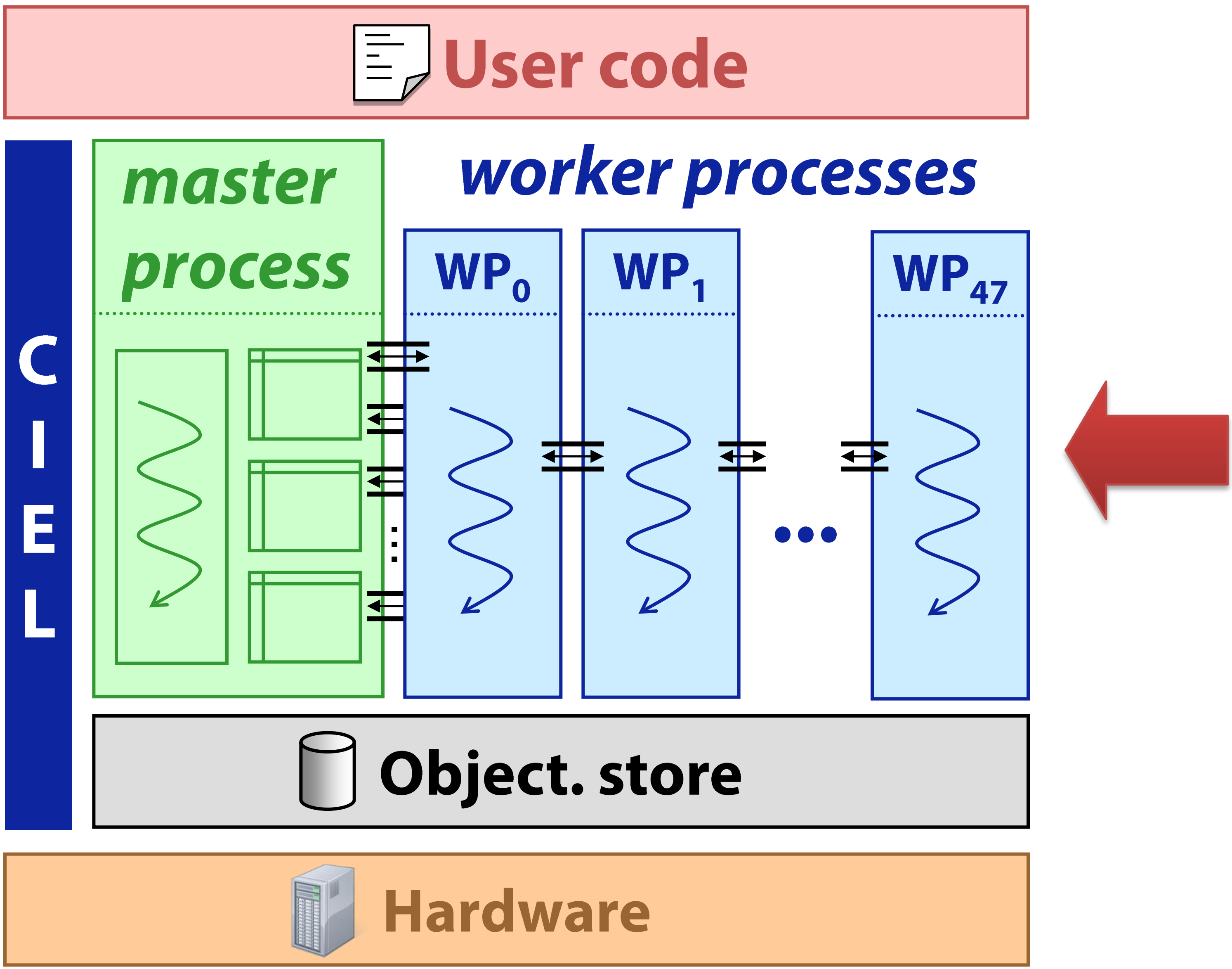

User code
master process
worker processes
WP0
WP1
...
WP47
CIEL
Object store
Hardware


User code
CIEL
master process
worker process
...
47 worker threads
Object store
Hardware

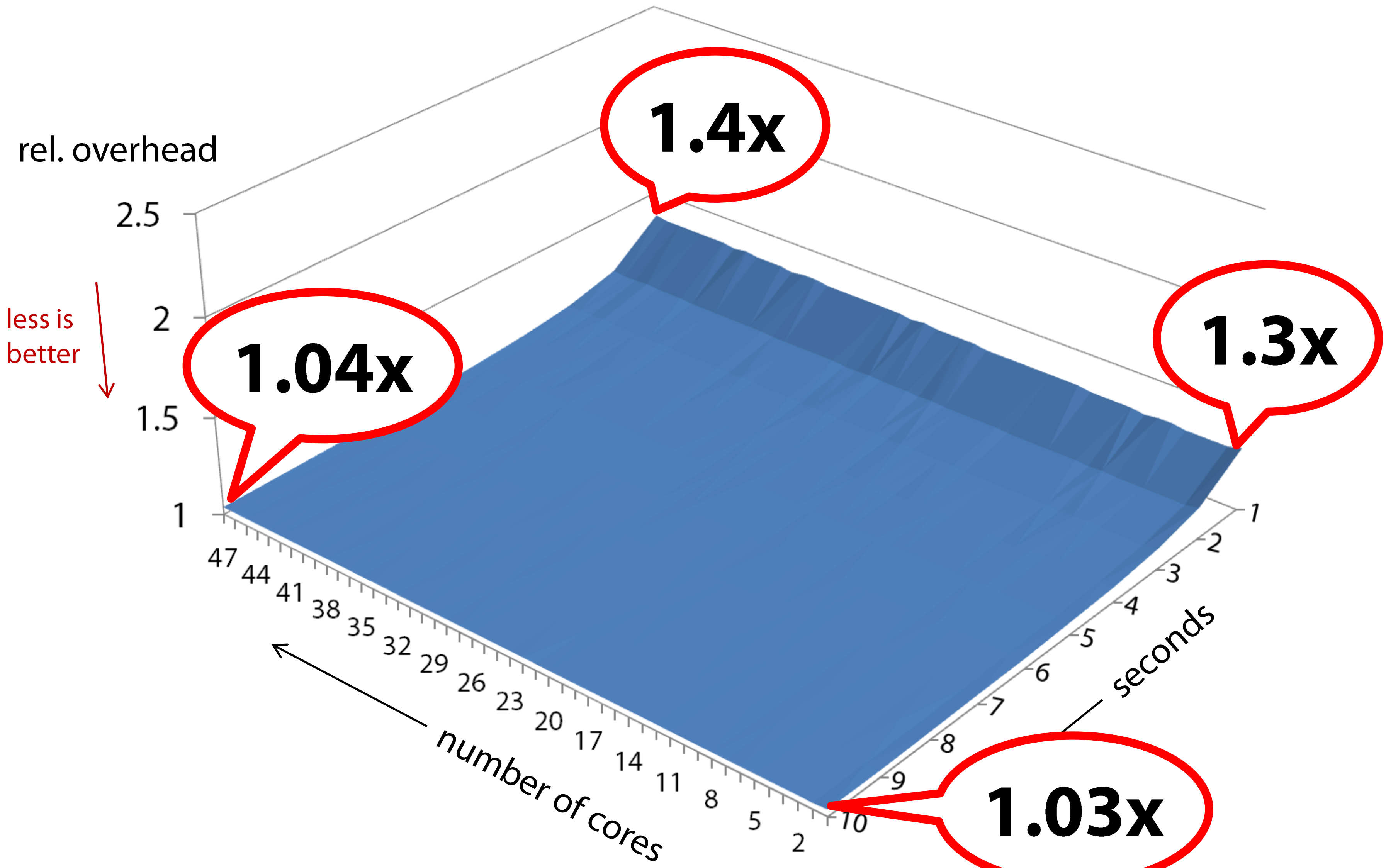
timespin, using multi-worker



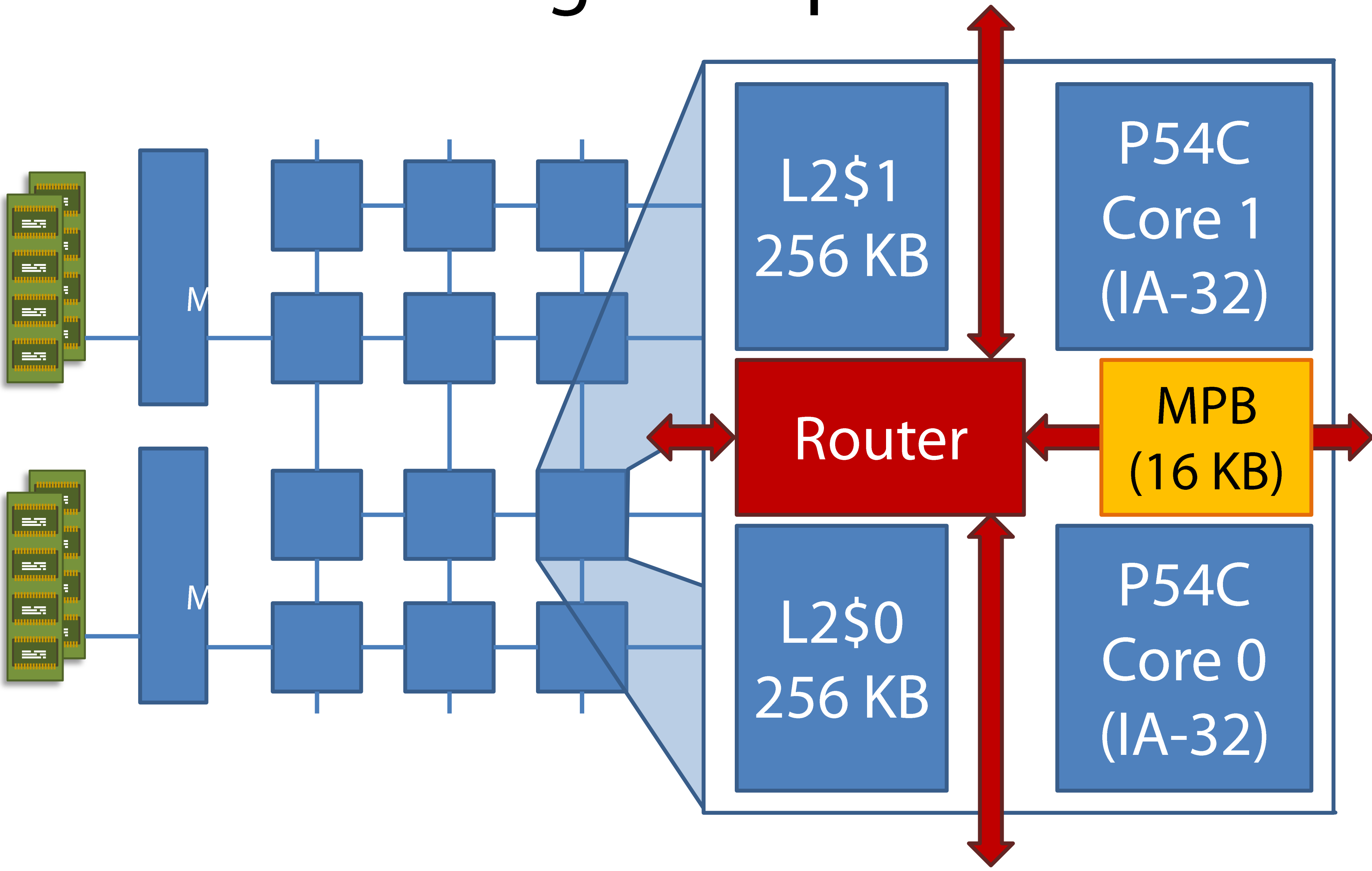

User code
CIEL
master process
worker process
...
47 worker threads
Object. store
Hardware


User code
C I E L
master process
worker processes
WP0
WP1
WP47
...
Object. store
Hardware

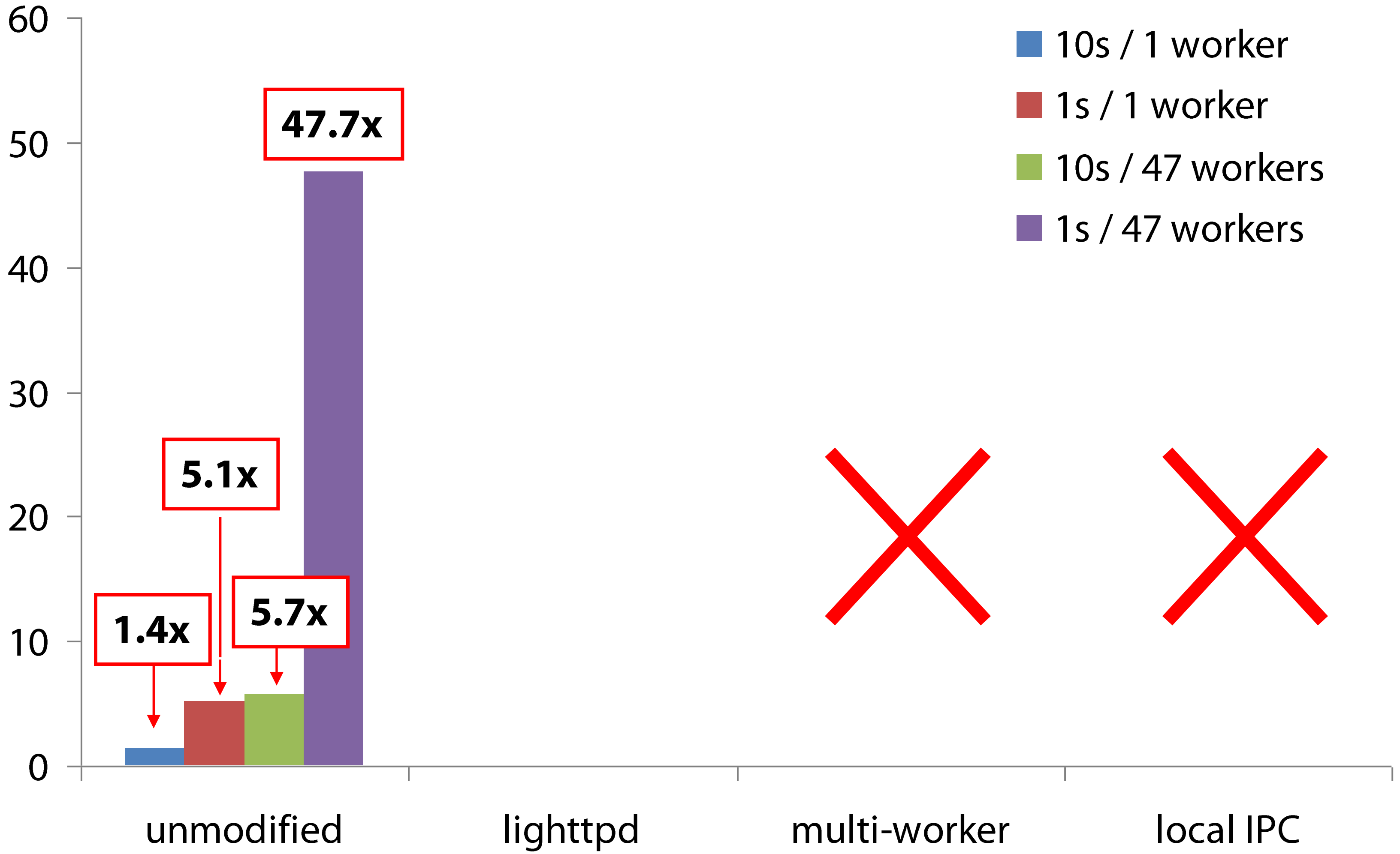
timespin, local IPC mode



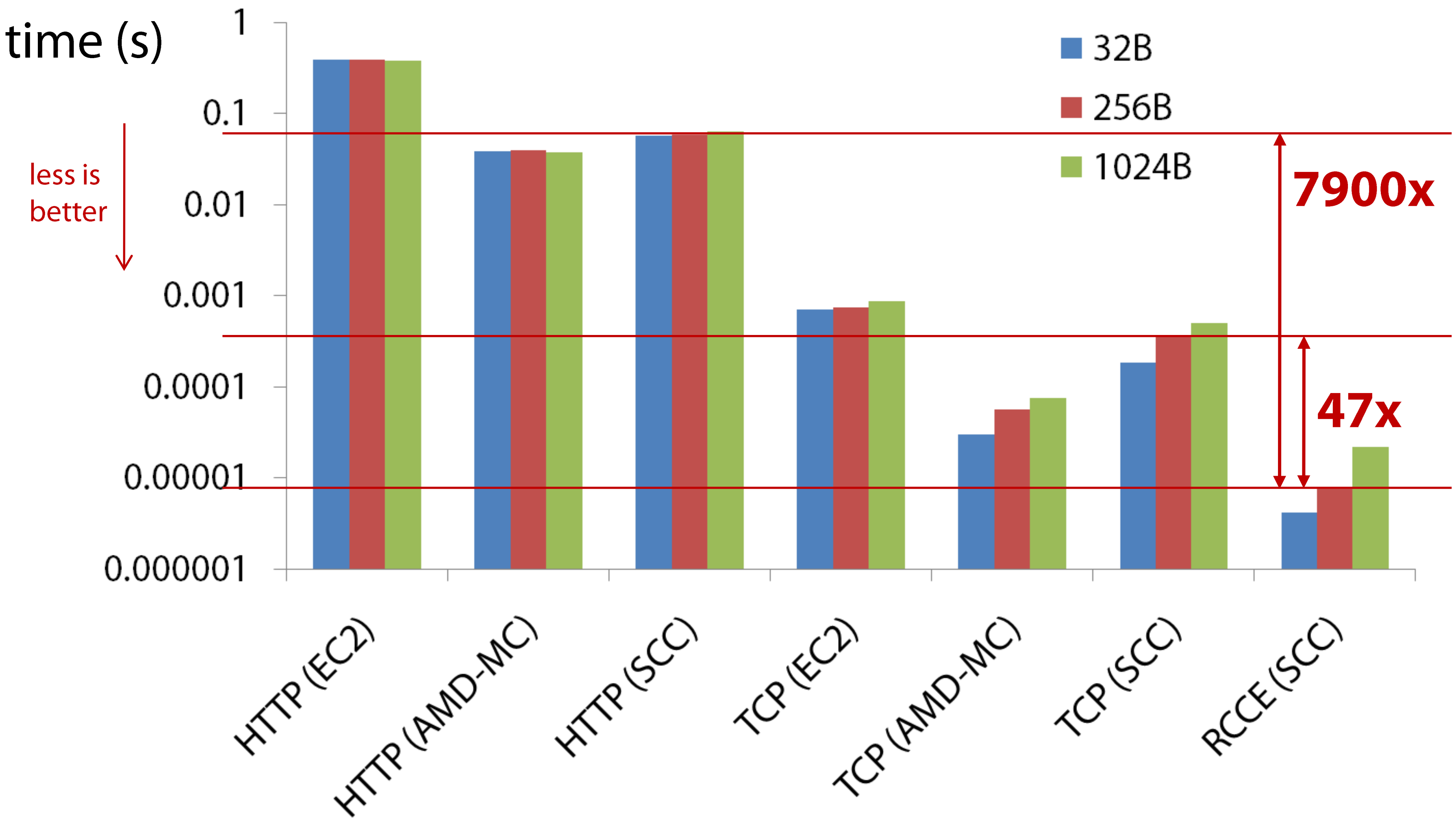
Intel Single-Chip Cloud



`timespin` on the SCC



Message latency



Challenges and Opportunities

Contention vs. sharing

Challenges and Opportunities

Contention vs. sharing

I/O multiplexing

Challenges and Opportunities

Contention vs. sharing

I/O multiplexing

Fault tolerance

Conclusions and summary

- Investigated performance of the CIEL many-core
- Works unmodified, but fine-grained tasks suffer
- Started to address various challenges
- **Next**: multi-scale version for hybrid clusters

http://www.cl.cam.ac.uk/netos/ciel/