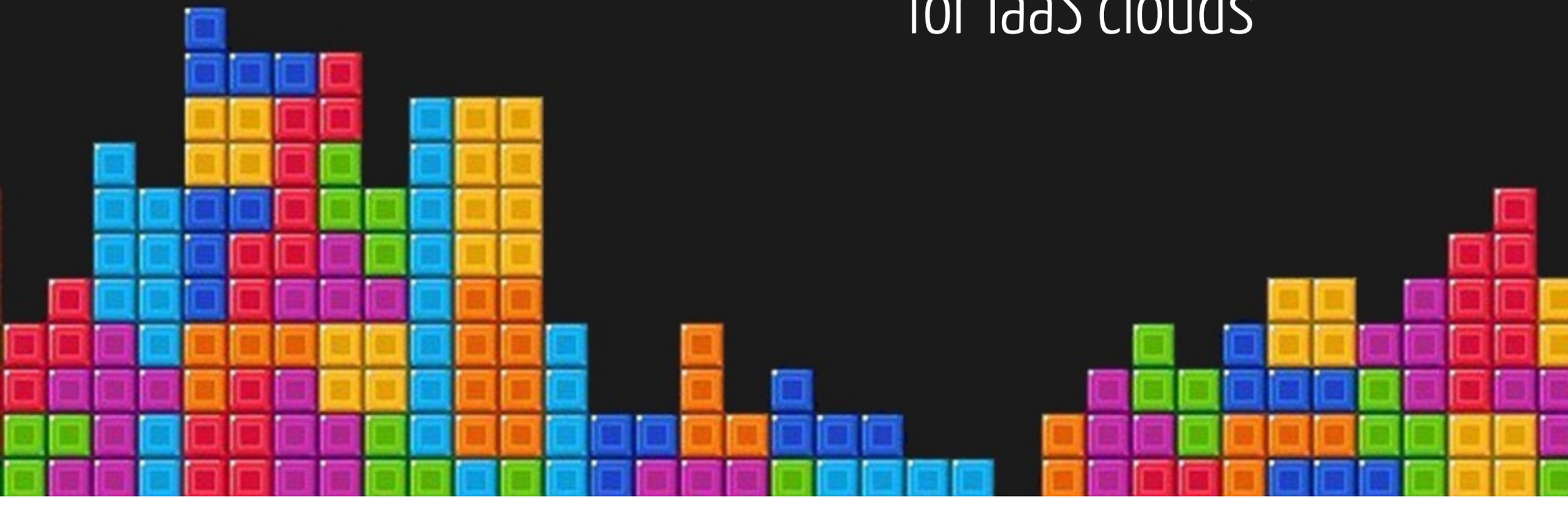
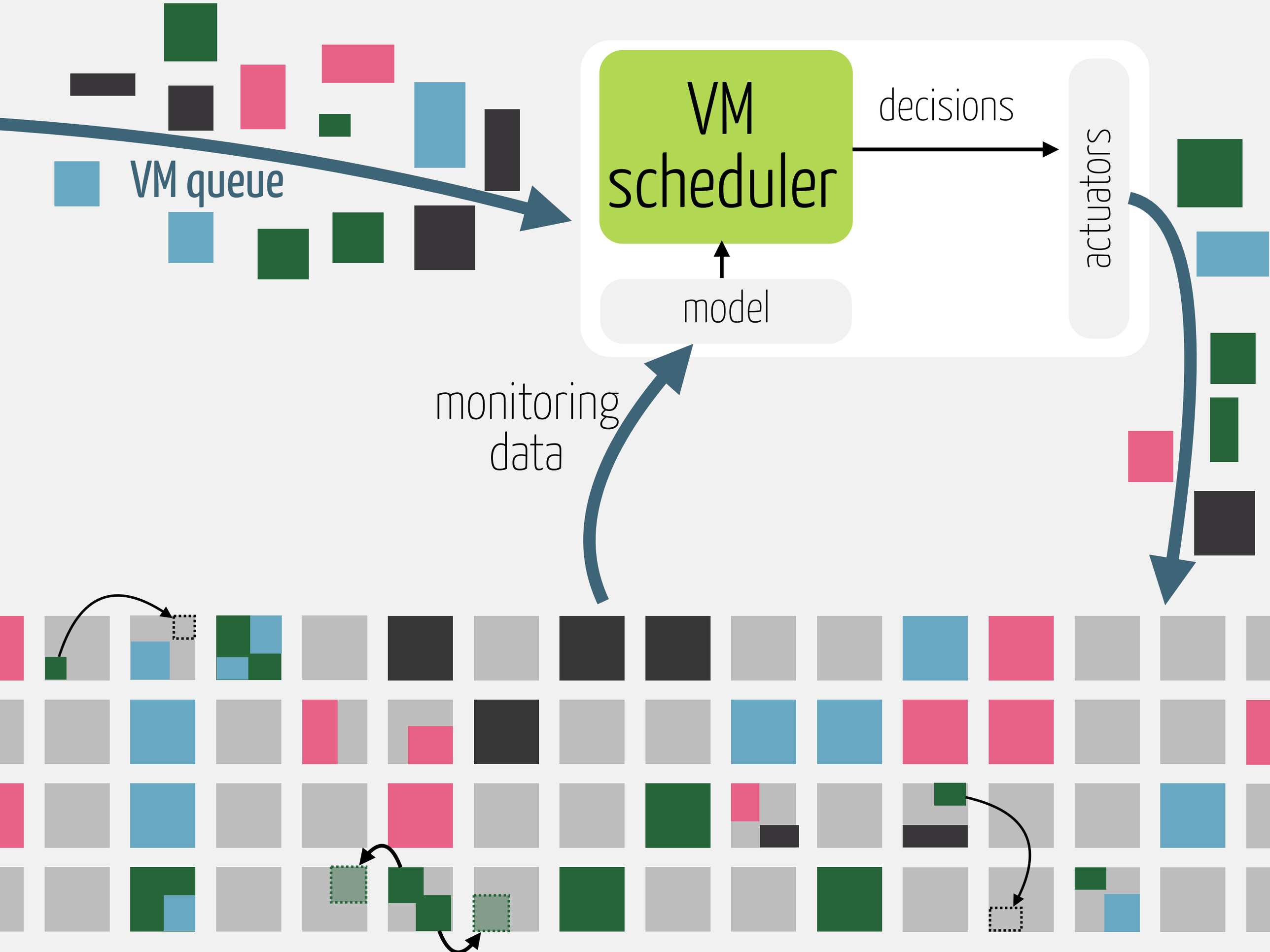
Fabien Hermenier
http://fhermeni.github.io

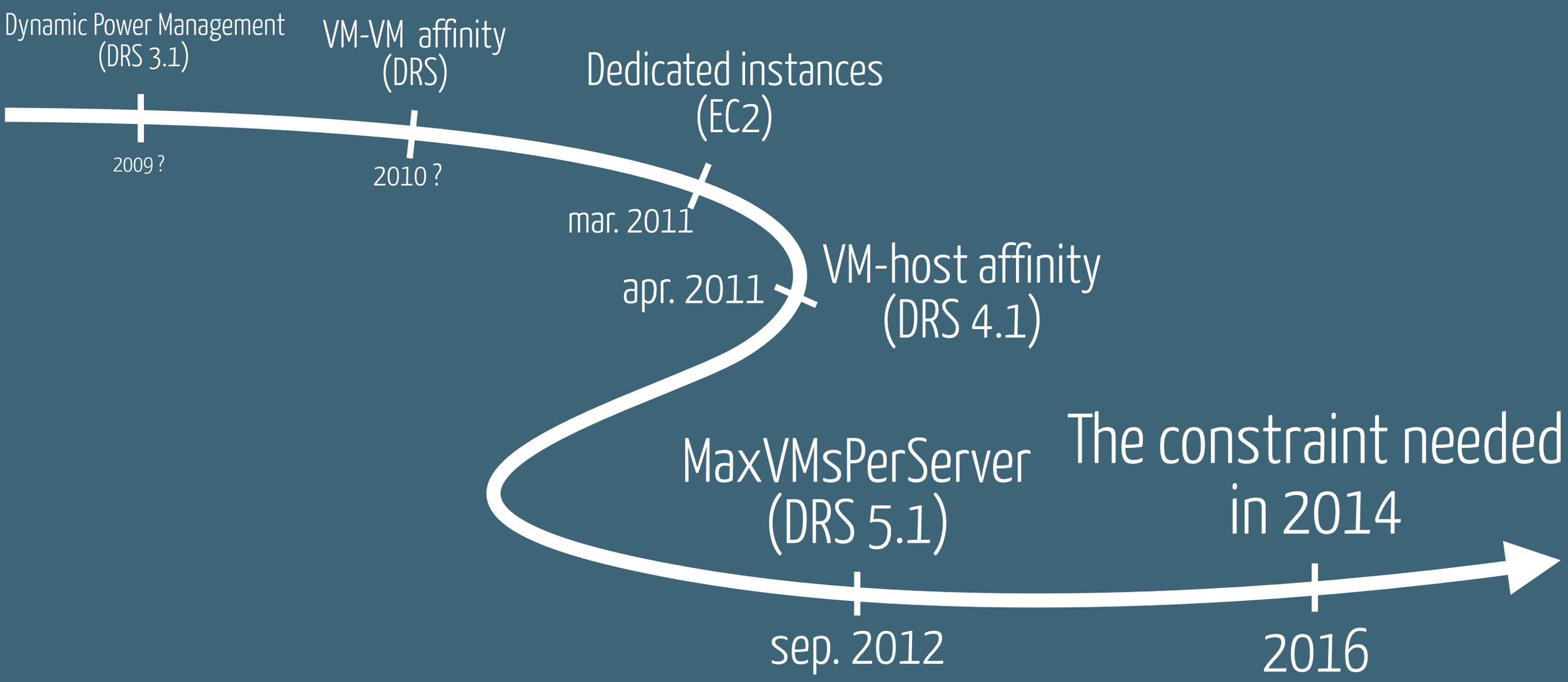
A flexible VM placement algorithm for IaaS clouds

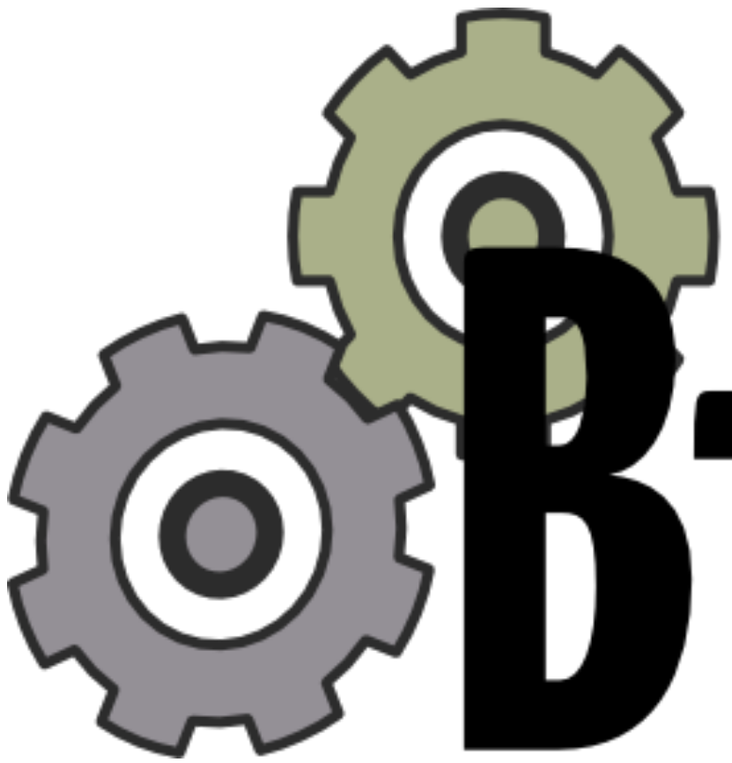


cnrs dépasser les frontières
Université Nice Sophia Antipolis
i3s sophie antipolis


VM queue
VM scheduler
decisions
actuators
model
monitoring data

SLOs and infrastructures evolve

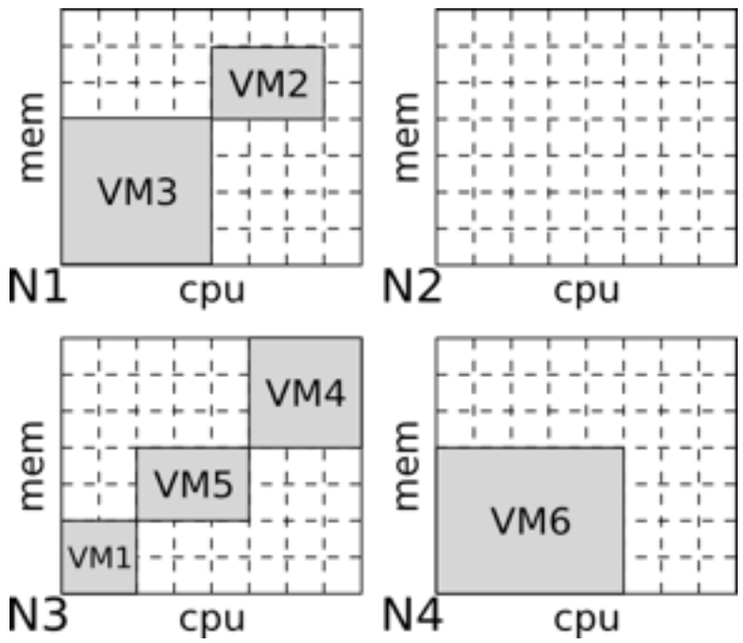




BtrPlace

adapt the VM placement depending on pluggable expectations

network and memory-aware migration scheduler, VM-(VM|PM) affinities, resource matchmaking, node state manipulation, counter based restrictions, energy efficiency, discrete or continuous restrictions

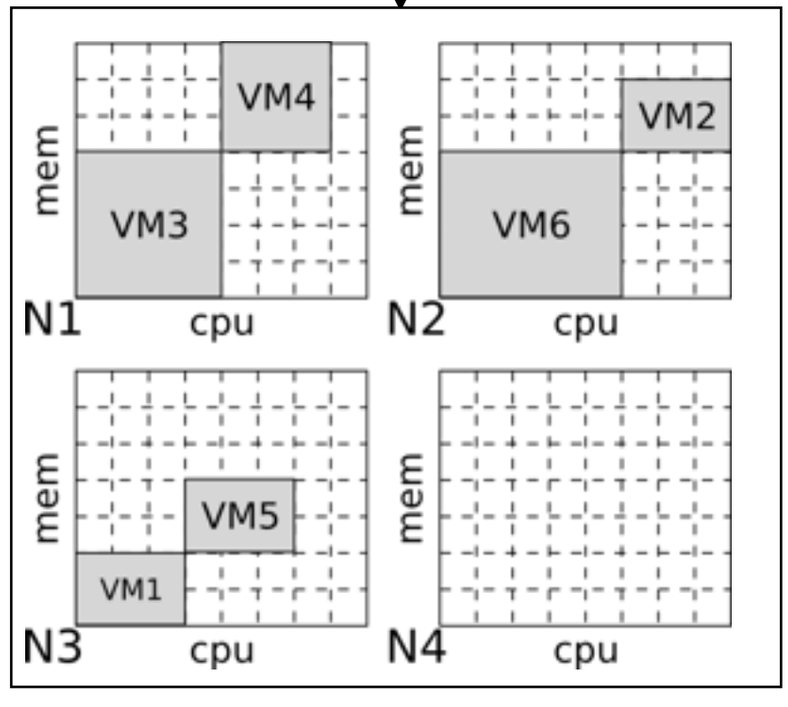



```
spread(VM[2..3]);
preserve(VM1,'cpu', 3);
offline(@N4);
```

BtrPlace

The reconfiguration **plan**

```
0'00 to 0'02: relocate(VM2,N2)
0'00 to 0'04: relocate(VM6,N2)
0'02 to 0'05: relocate(VM4,N1)
0'04 to 0'08: shutdown(N4)
0'05 to 0'06: allocate(VM1,'cpu',3)
```

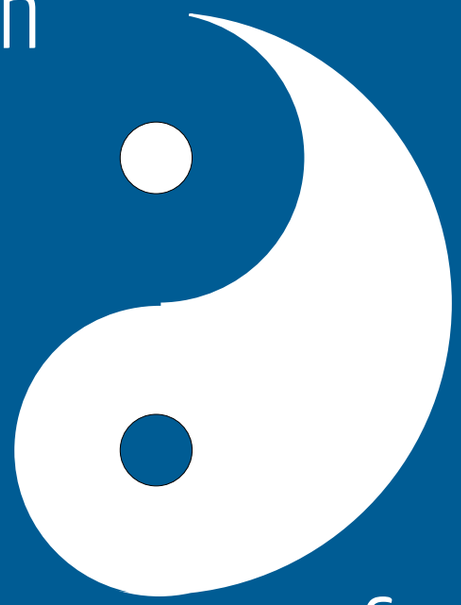


CHOCO

An Open-Source java library
for constraint programming

$$
\begin{aligned}
\mathcal{X} &= \{x_1, x_2, x_3\} \\
\mathcal{D}(x_i) &= [0, 2], \forall x_i \in \mathcal{X} \\
\mathcal{C} &= \left\{ \begin{array}{r} c_1 : x_1 < x_2 \\ c_2 : x_1 + x_2 \geq 2 \\ c_3 : x_1 < x_3 \end{array} \right.
\end{aligned}
$$

deterministic composition
high-level constraints



the right model
for the right problem

BtrPlace CSP

models a reconfiguration plan

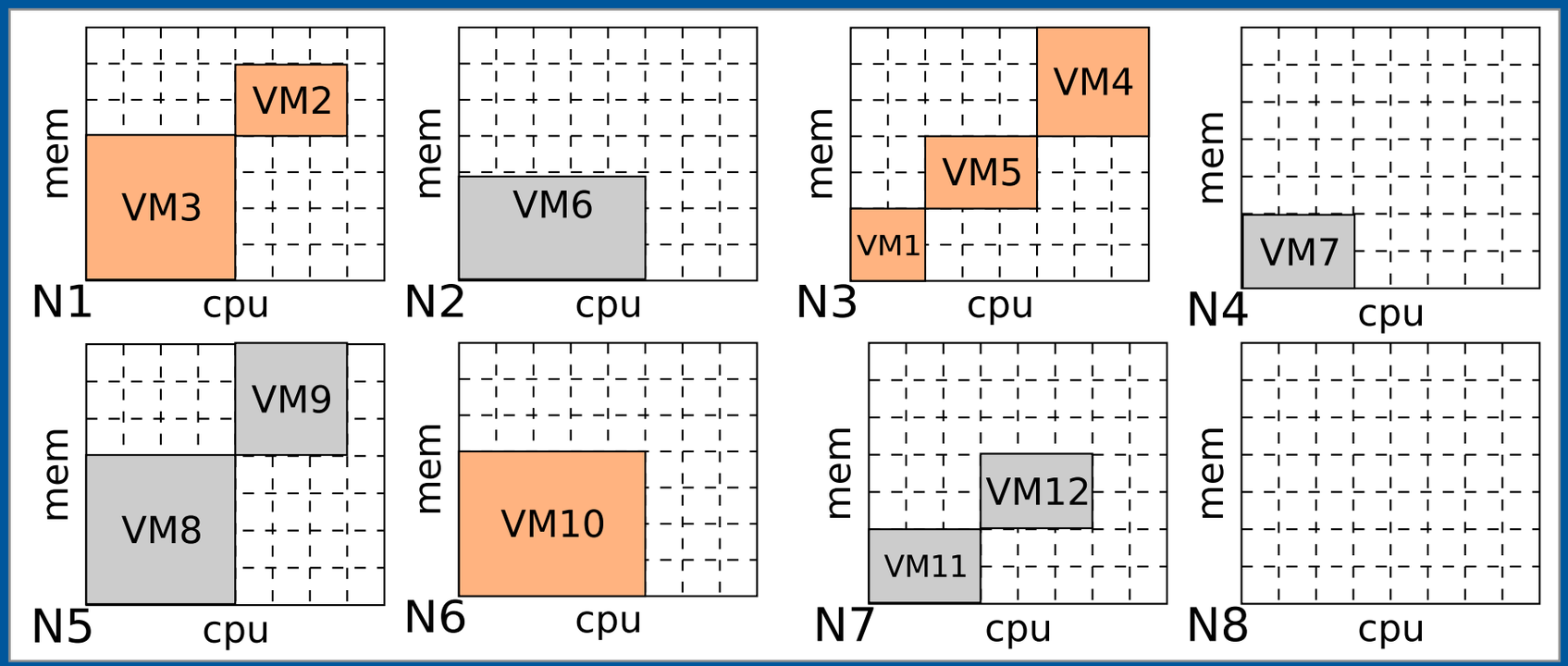
1 transition model per element

pre-defined variables as hooks

$$boot(v \in V) \triangleq \quad D(v) \in \mathbb{N}$$
$$st(v) = [0, H - D(v)]$$
$$ed(v) = st(v) + D(v)$$
$$d(v) = ed(v) - st(v)$$
$$d(v) = D(v)$$
$$ed(v) < H$$
$$d(v) < H$$
$$h(v) \in \{0, .., |N| - 1\}$$

$$relocatable(v \in V) \triangleq \ldots$$
$$shutdown(v \in V) \triangleq \ldots$$
$$suspend(v \in V) \triangleq \ldots$$
$$spread(vs \subseteq vms) \triangleq$$
$$i, j \in vs, i \neq j, host(i) \neq host(j)$$

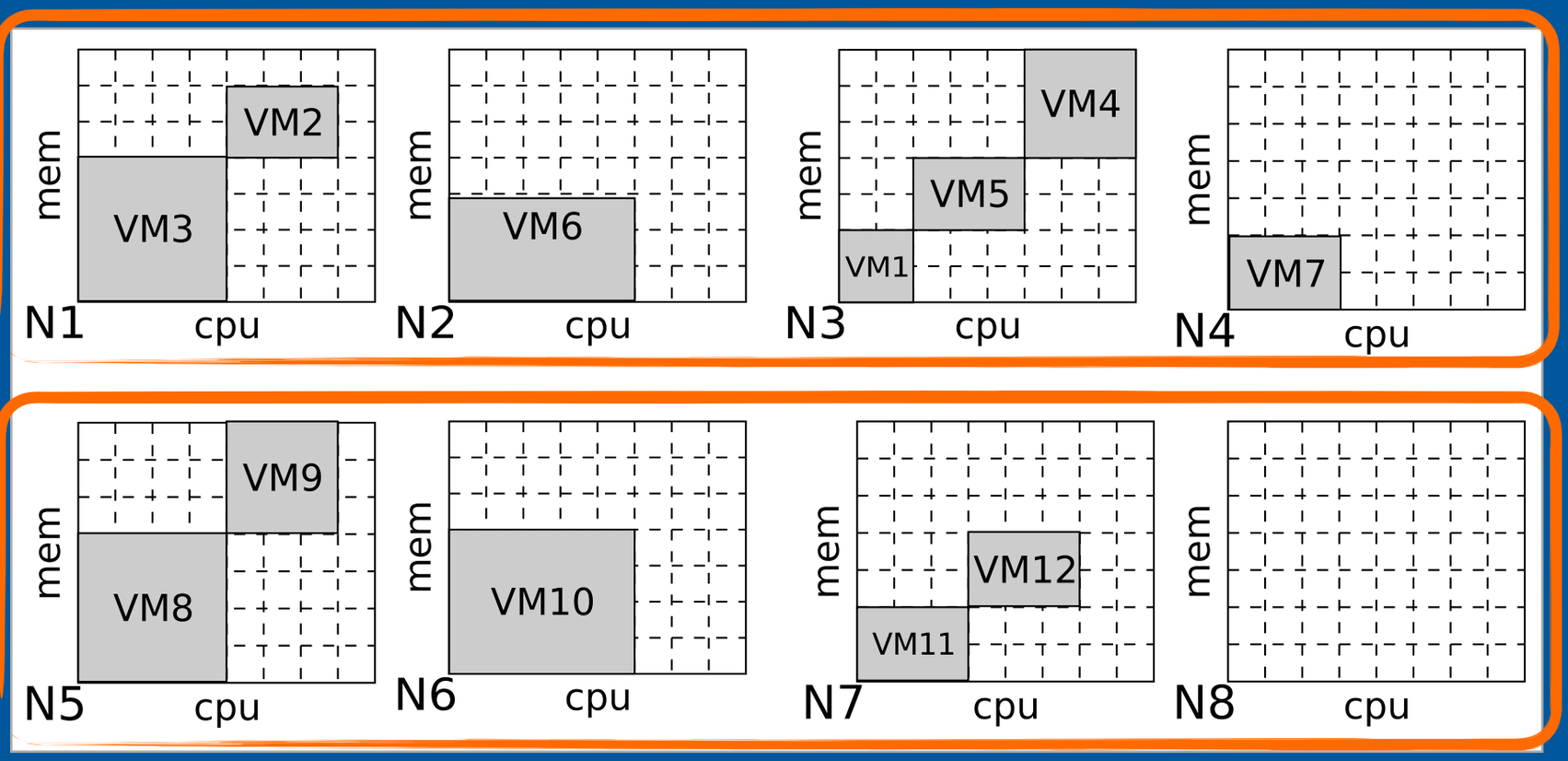
Static analysis for performance




```
spread({VM3,VM2,VM8});
lonely({VM7});
preserve({VM1},'ucpu', 3);
offline(@N6);
ban($ALL_VMS,@N8);
fence(VM[1..7],@N[1..4]);
fence(VM[8..12],@N[5..8]);
```

local search to focus on supposed mis-placed VMs

Static analysis for performance

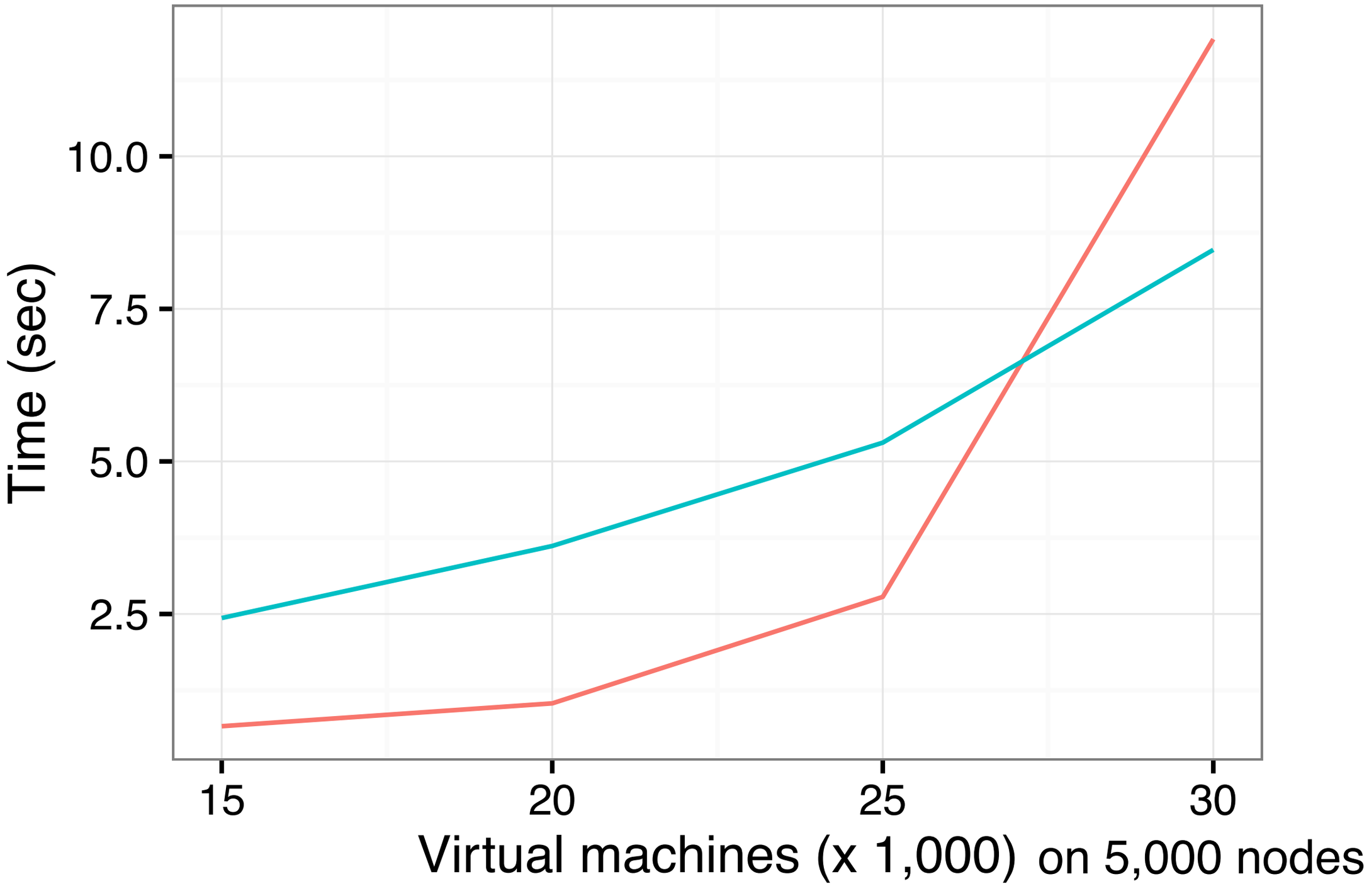



```
spread({VM3,VM2,VM8});
lonely({VM7});
preserve({VM1},'ucpu', 3);
offline(@N6);
ban($ALL_VMS,@N8);
fence(VM[1..7],@N[1..4]);
fence(VM[8..12],@N[5..8]);
```

self-partitioning to solve independent sub-problems in parallel

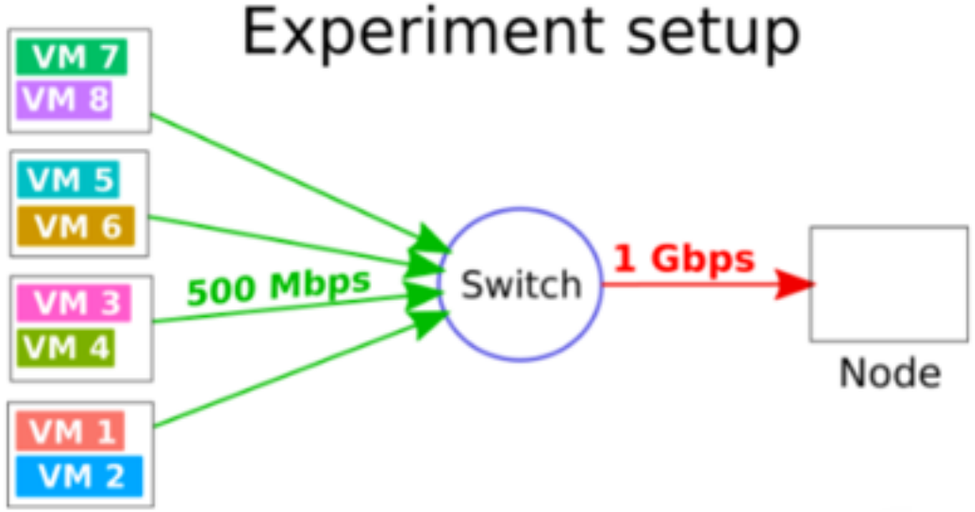
solving latency

load spike
hardware failure

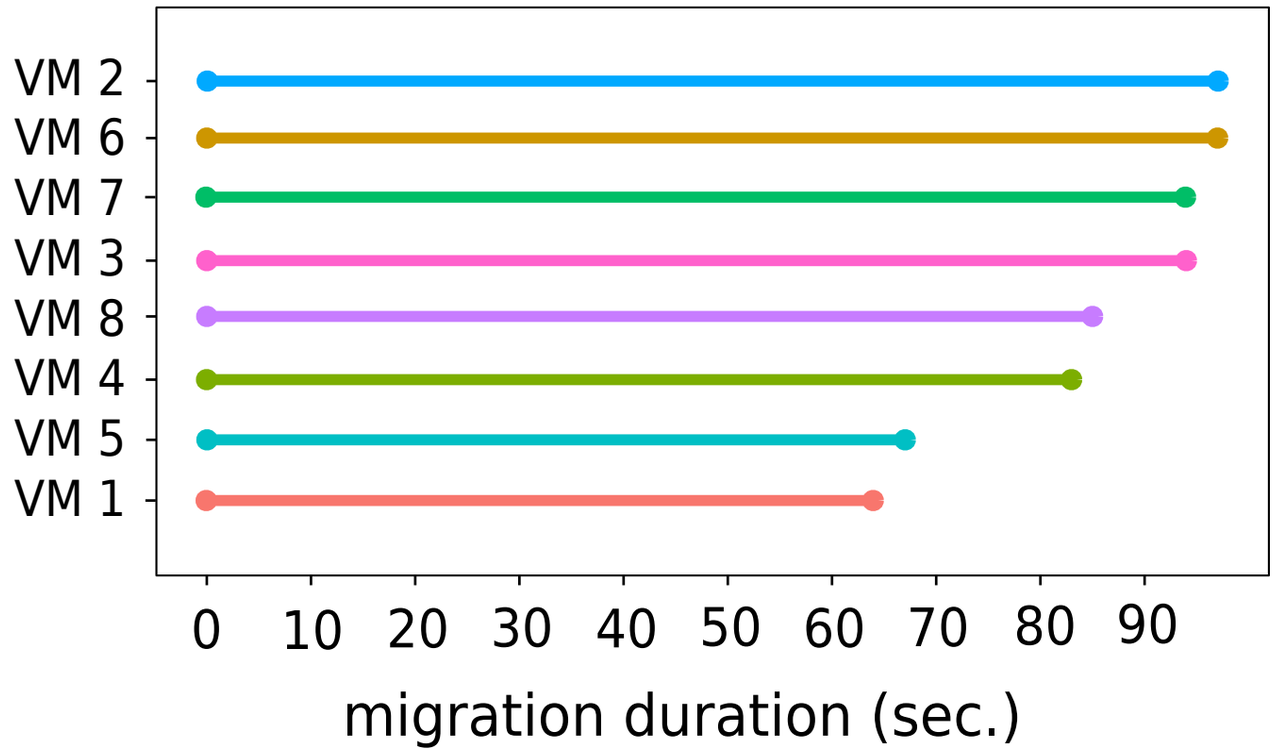


Extensibility in practice

looking for a better migration scheduler



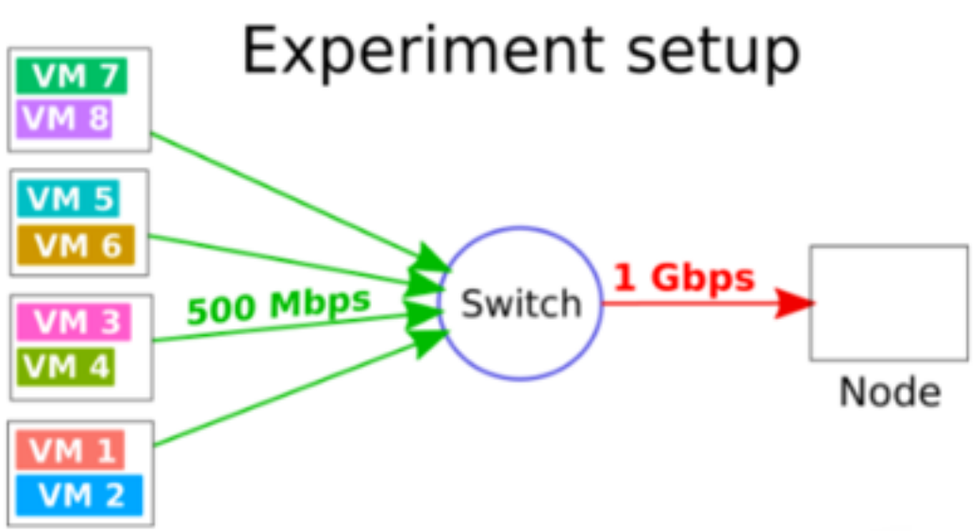
[btrplace vanilla, entropy, cloudsim, …]



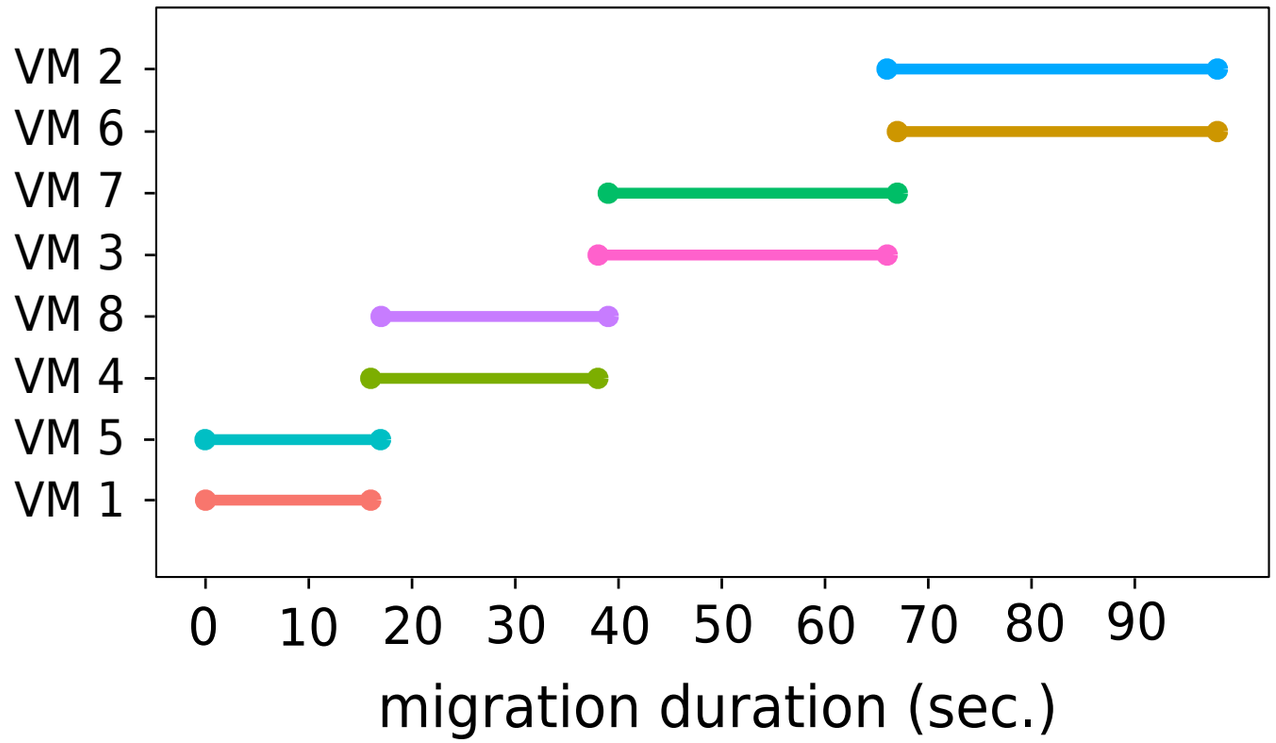
network and workload blind

Extensibility in practice

looking for a better migration scheduler



[btrplace UCC 15]



network and workload aware

Extensibility in practice
solver-side

Network Model — blocking heterogeneous network

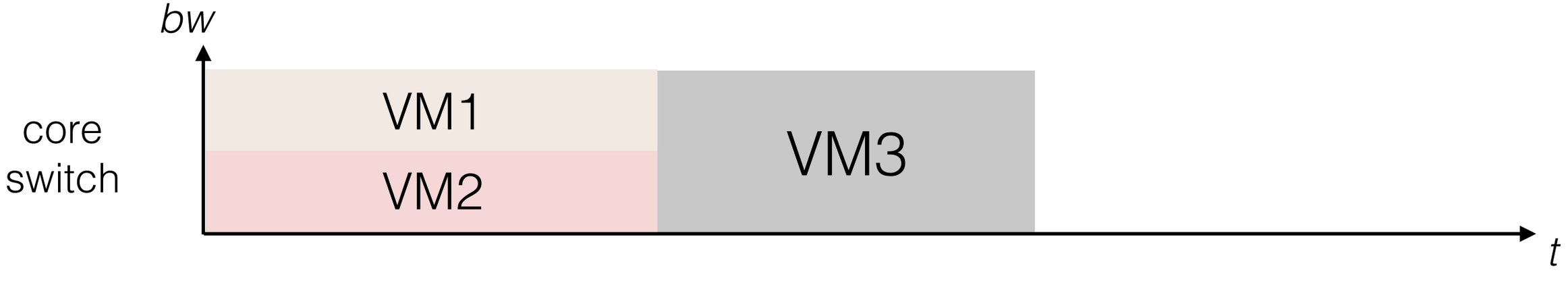
cumulative constraints; +/- 300 sloc.

Migration Model — memory and network aware

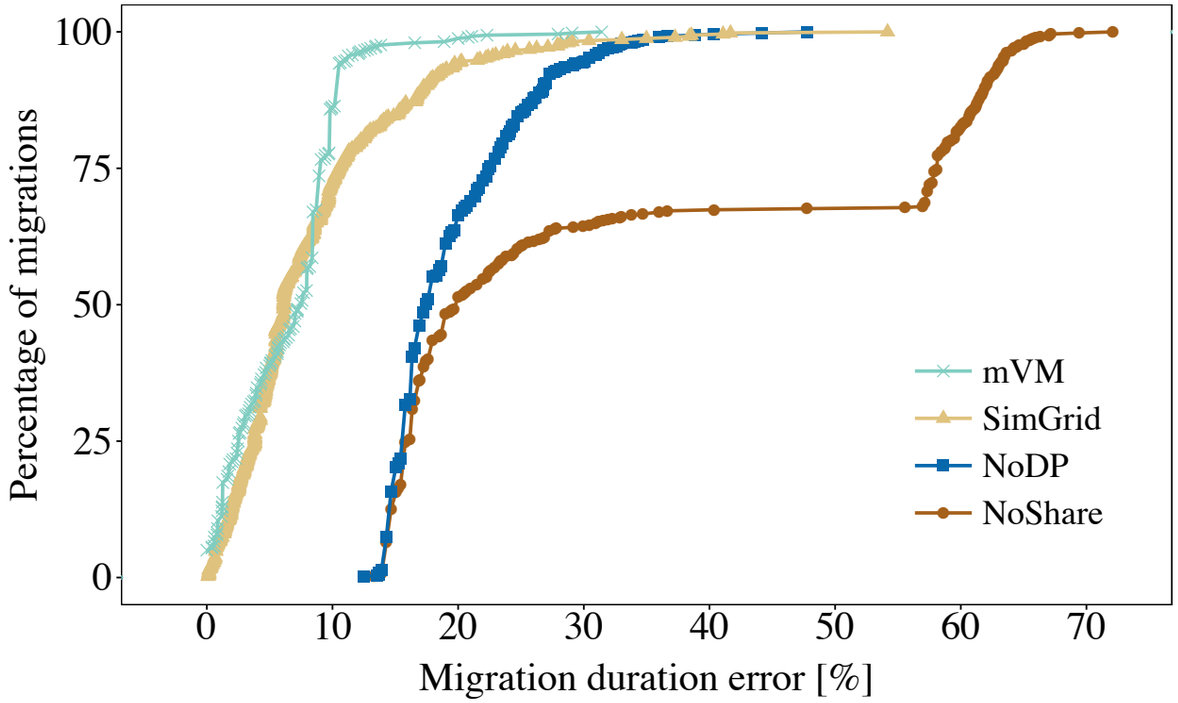
+/- 200 sloc.

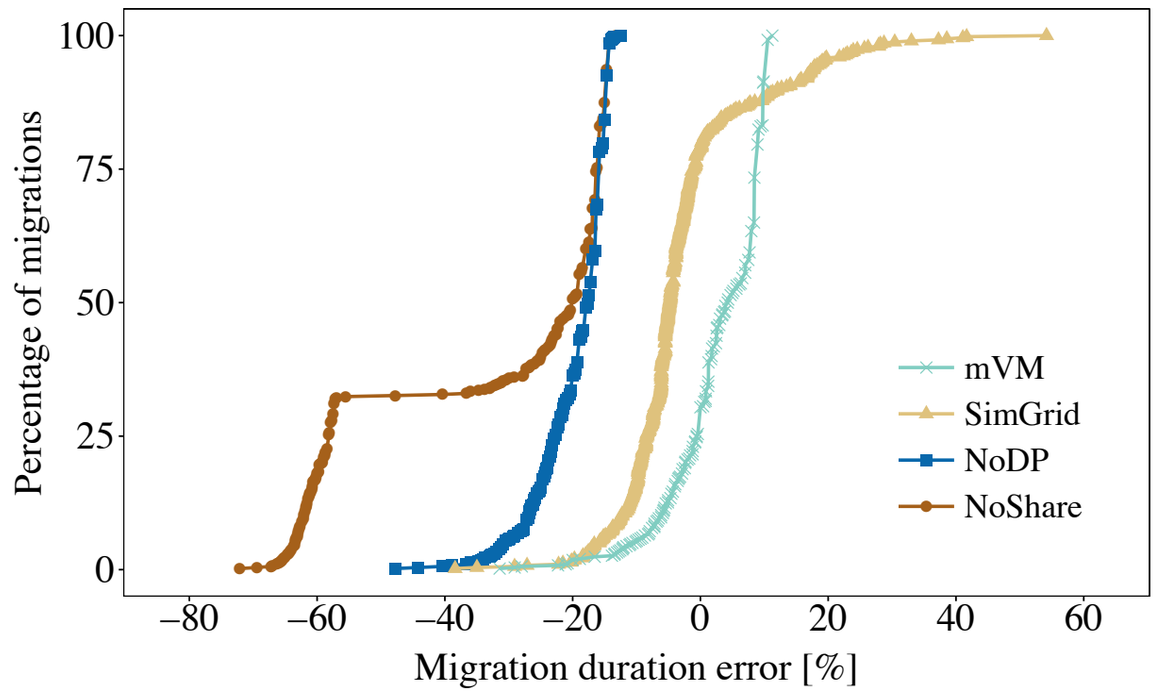
Constraints Model — restrict the migration models

+/- 100 sloc.



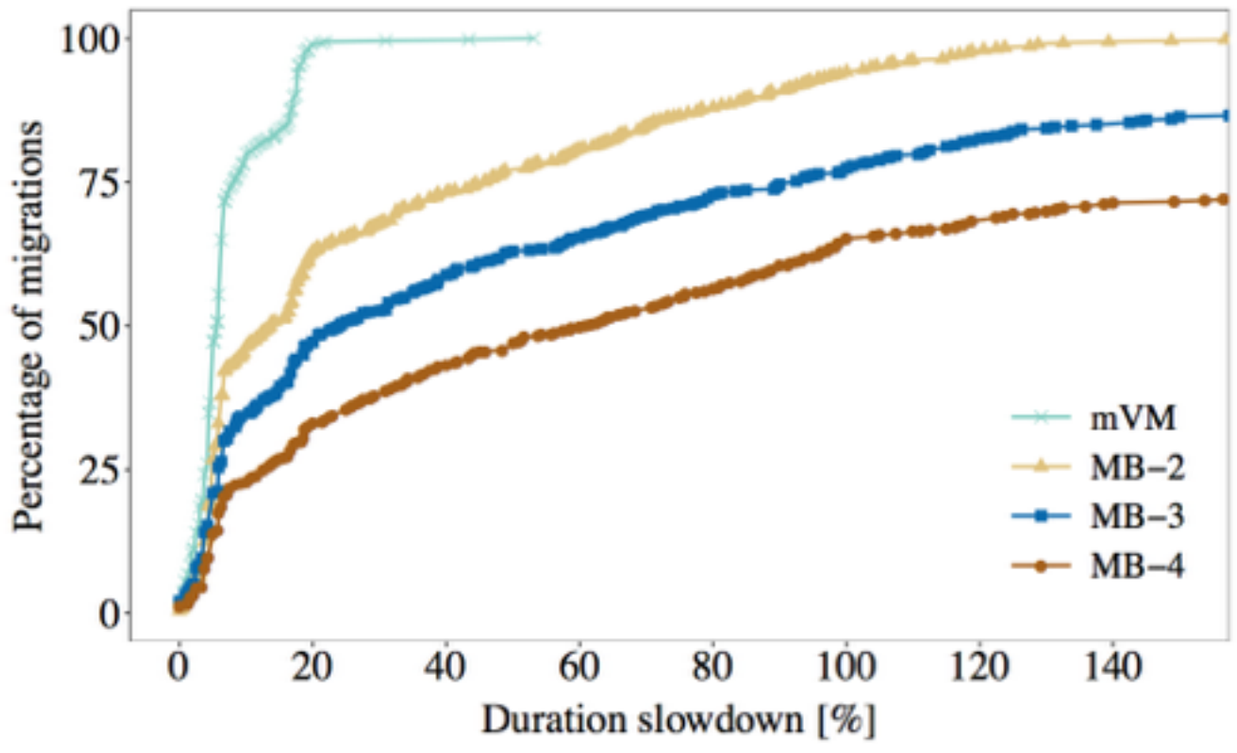
model ± 96% accurate

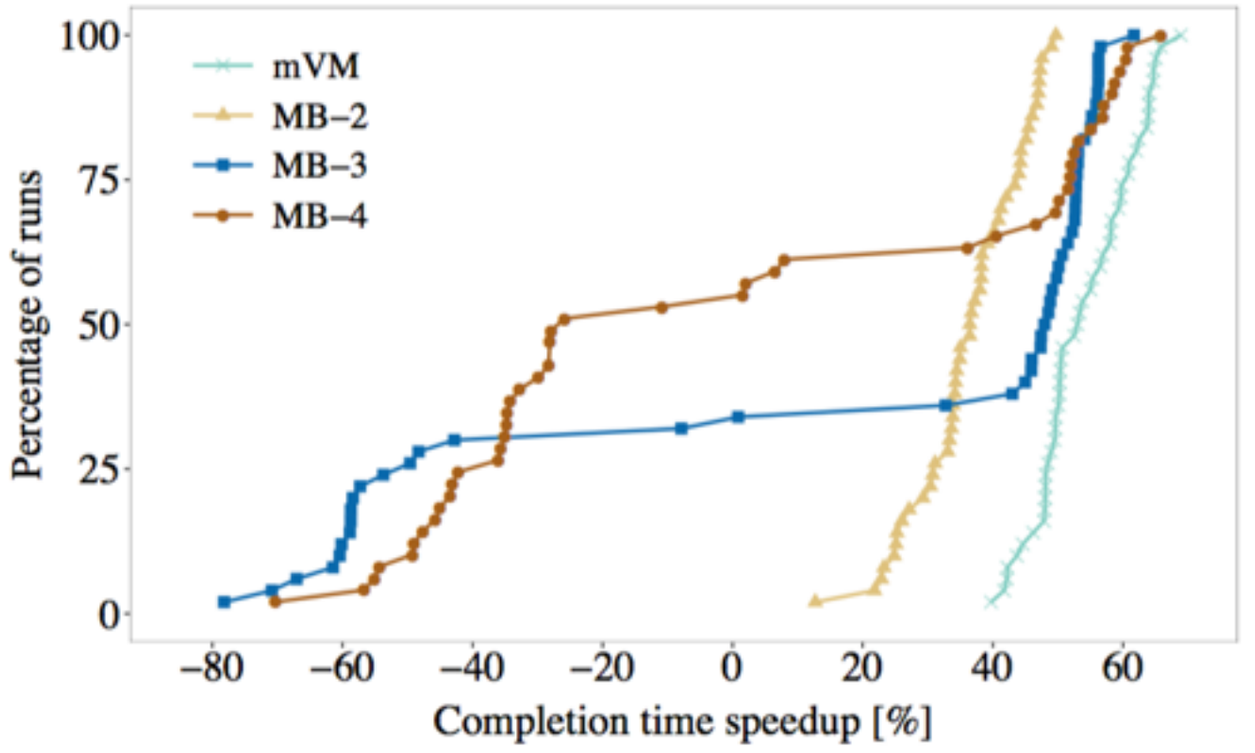




Scheduler	mVM	MB-2	MB-3	MB-4
Mean migration time (sec.)	45.55	57.22	113.2	168.6
Mean slowdown (%)	7.35%	29.69%	141.3%	259.2%

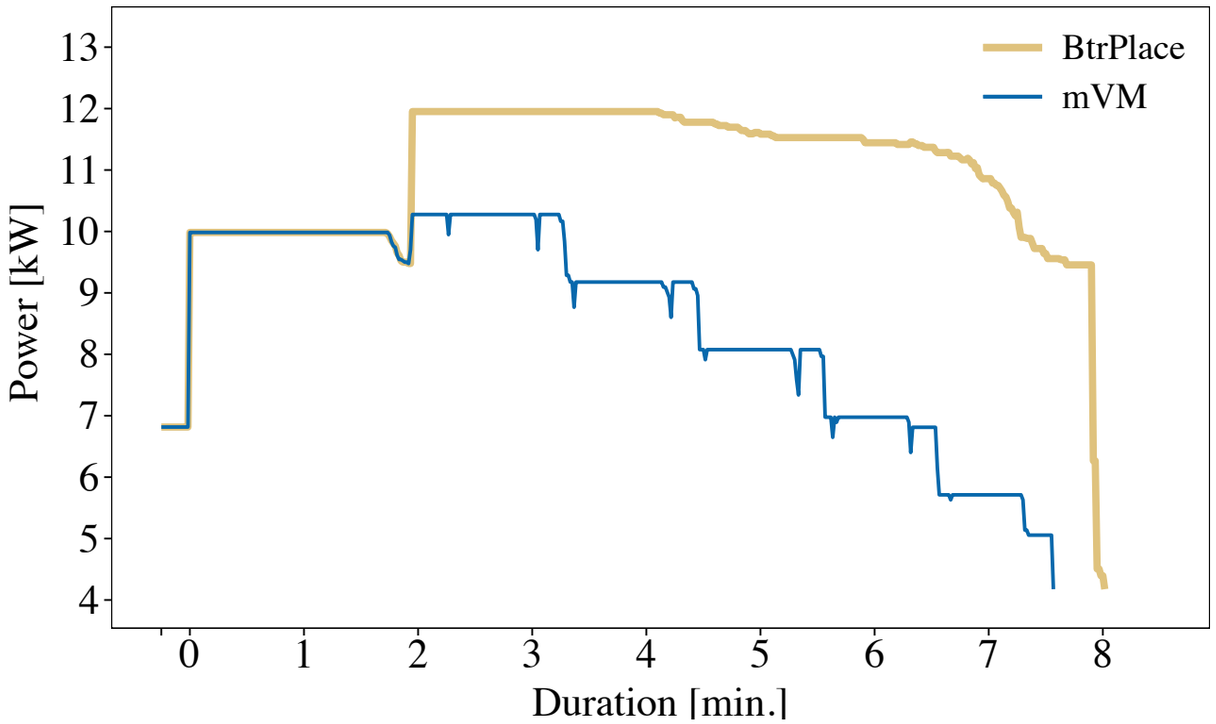
Scheduler	mVM	MB-2	MB-3	MB-4
Mean completion time (sec.)	212.8	295.9	394.6	479.4
Mean speedup (%)	54.18%	36.42%	15.94%	-2.64%

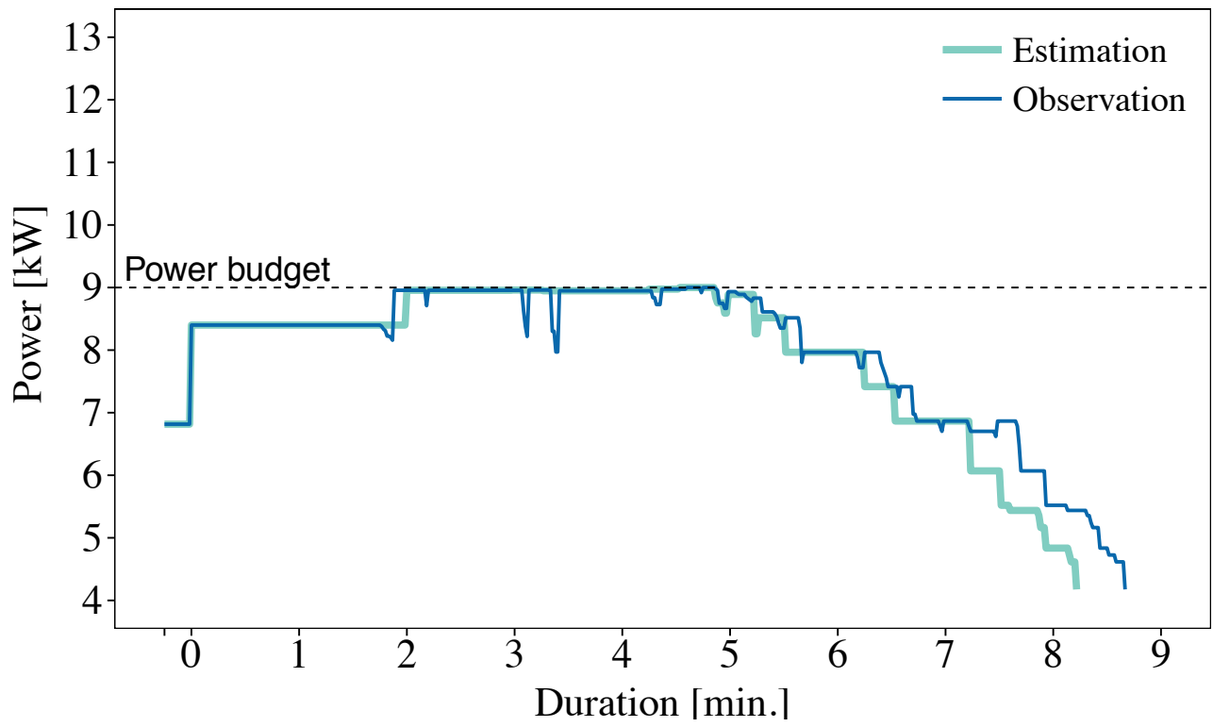




speed up

Energy-aware scheduling





decommissioning 2x24 servers to 24 more powerfull

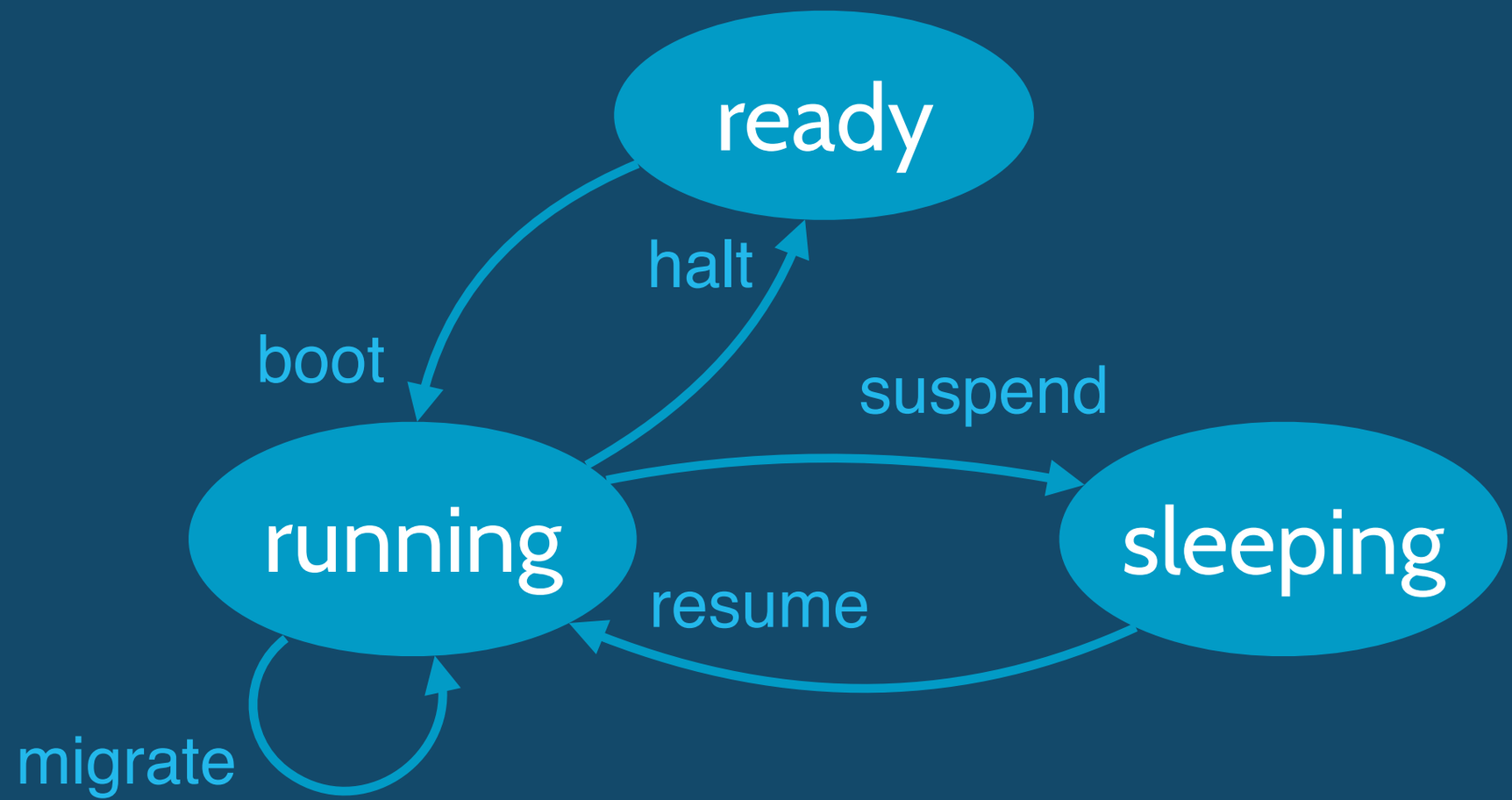
Grid'5000

32.45 years/cpu
of experiments

custom OS images
full cluster reservation
trafic shaping
g5k-subnet
storage reservation
power sensors

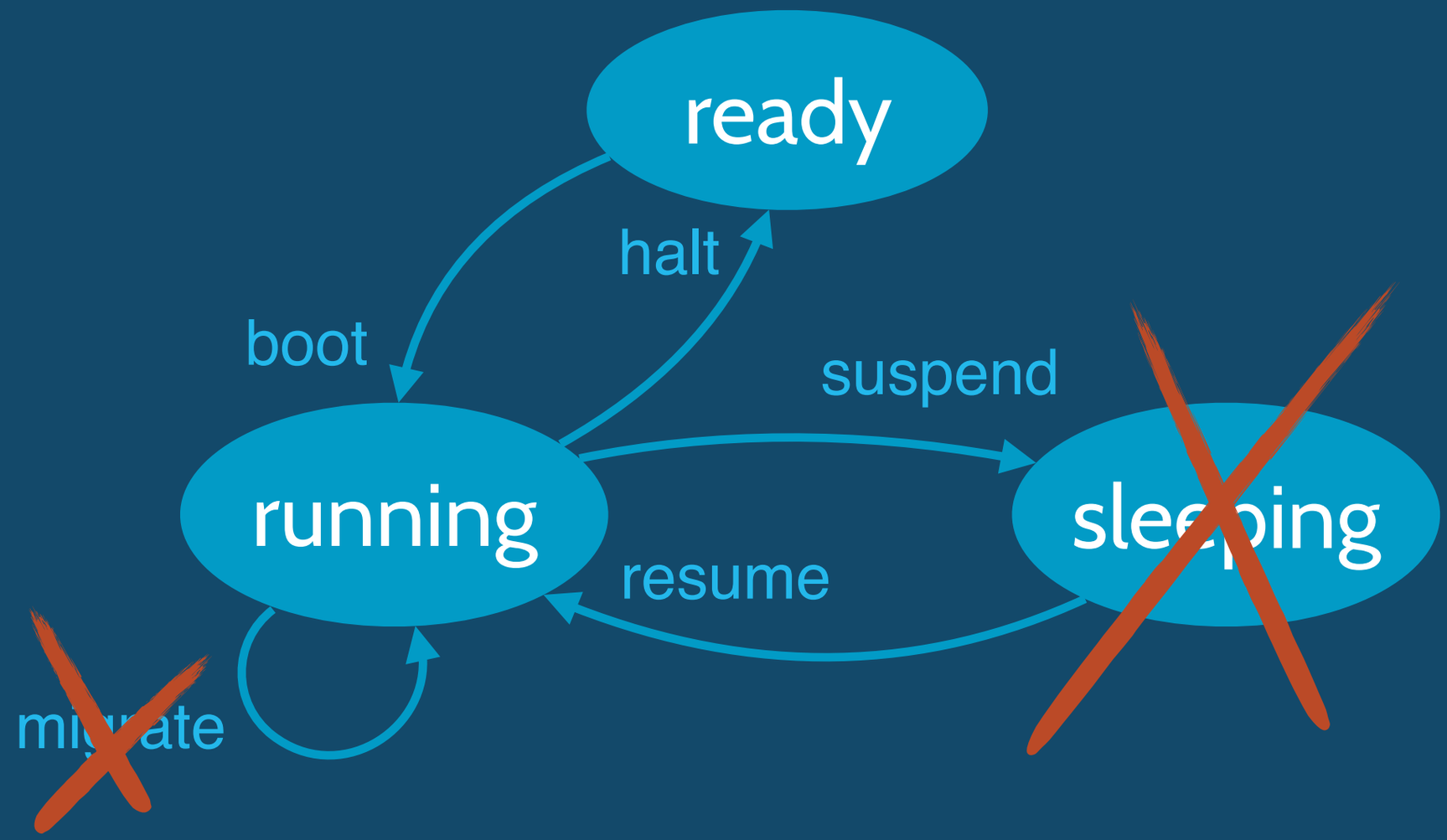
supporting docker

from a VM to a container life-cycle



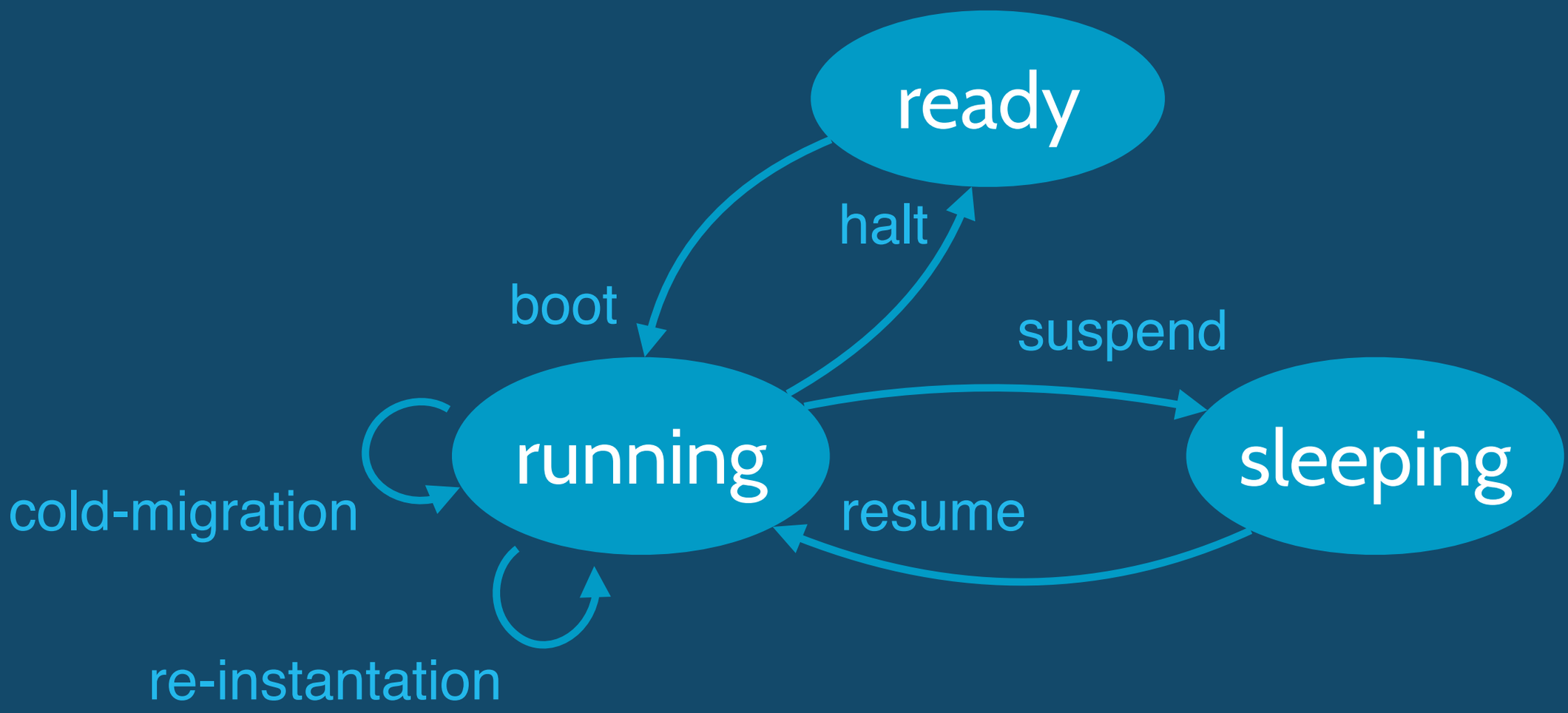
supporting docker

simplify the life-cycle using constraints
or customized life-cycles ...



supporting docker

... or overcome the limitations



adhoc VM schedulers do not fit evolving requirements move to a flexible scheduler to be proactive

http://**BtrPlace**.org

production ready live demo stable user API documented

tutorials issue tracker support chat room