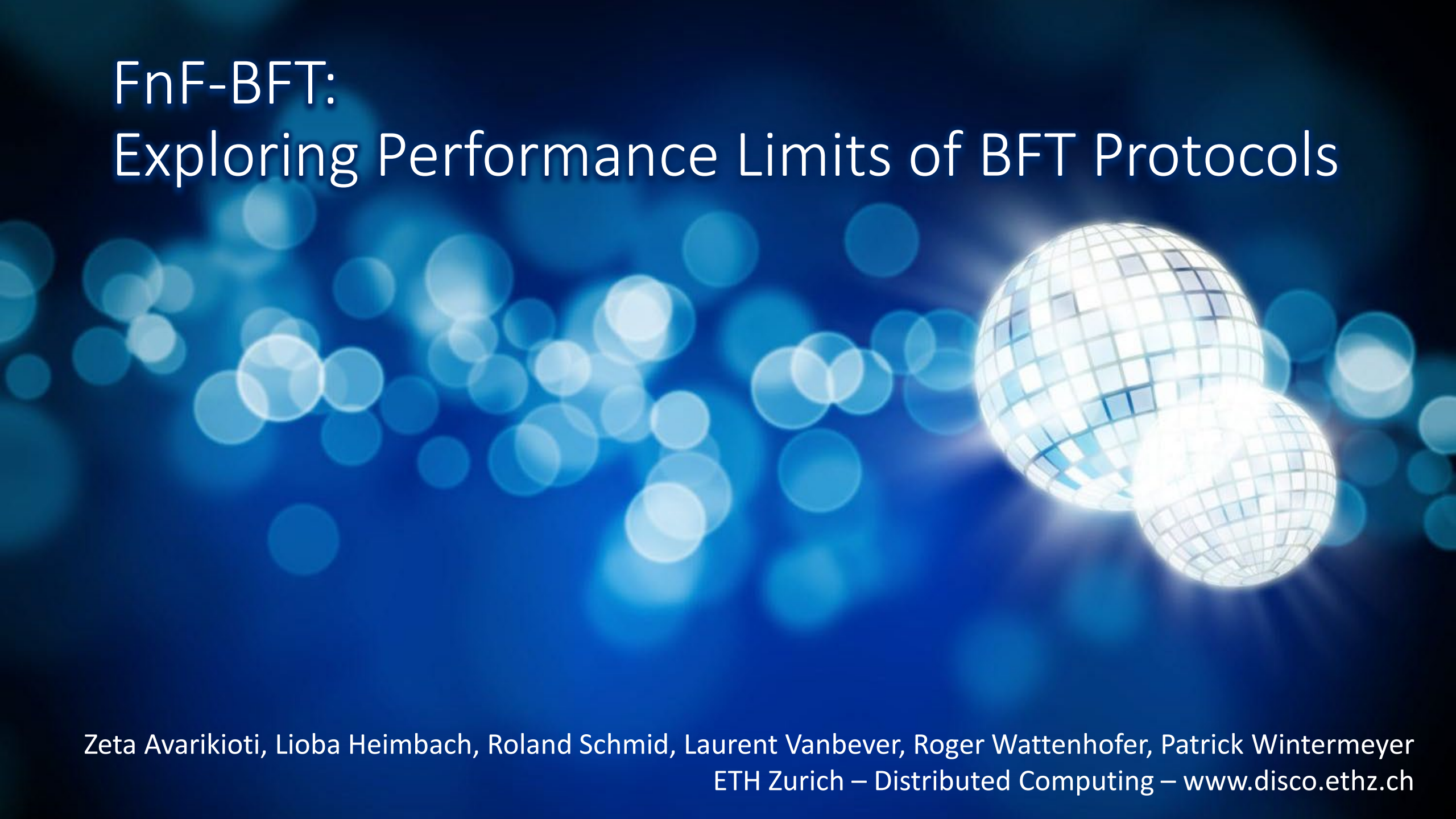
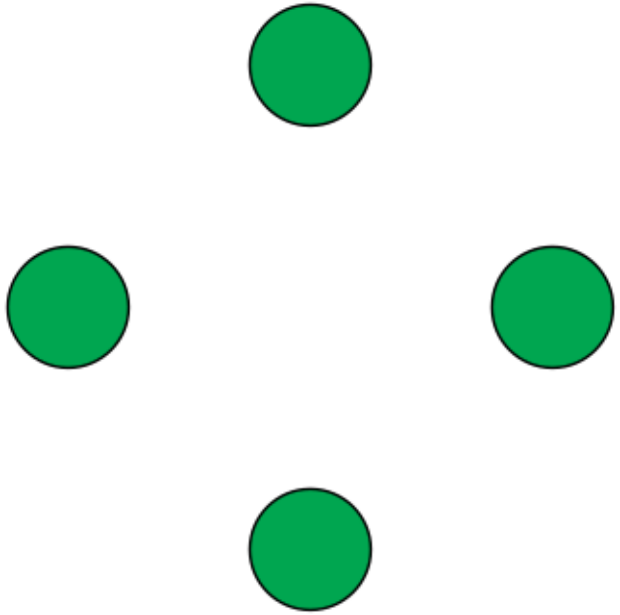
FnF-BFT:
Exploring Performance Limits of BFT Protocols

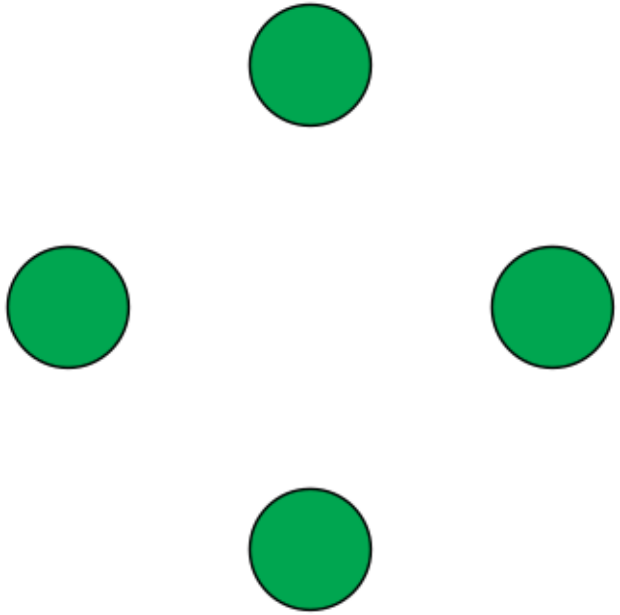


Zeta Avarikioti, Lioba Heimbach, Roland Schmid, Laurent Vanbever, Roger Wattenhofer, Patrick Wintermeyer
ETH Zurich – Distributed Computing – www.disco.ethz.ch

Byzantine fault tolerance

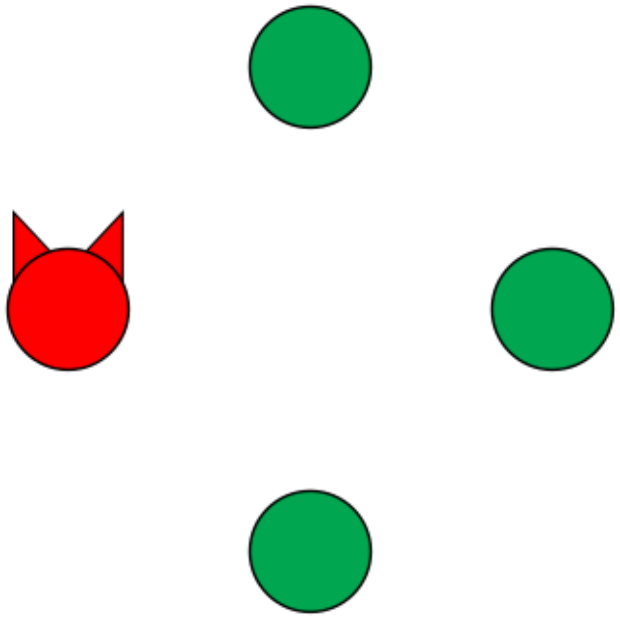


Byzantine fault tolerance



$n = 3f + 1$ replicas

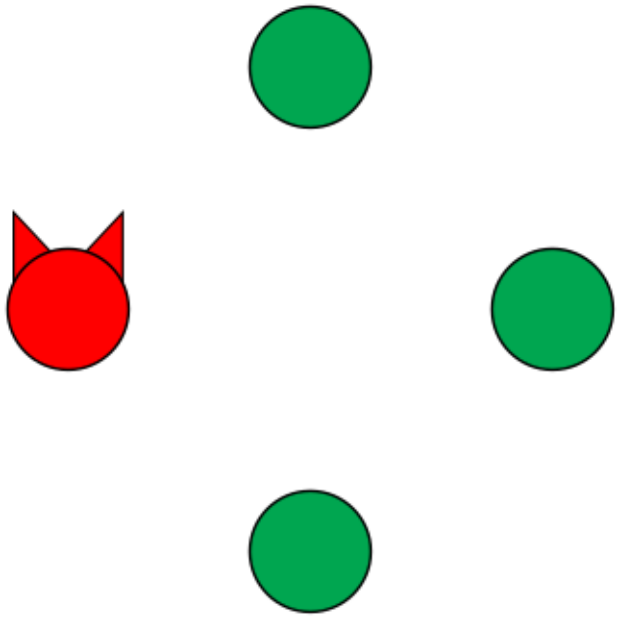
Byzantine fault tolerance



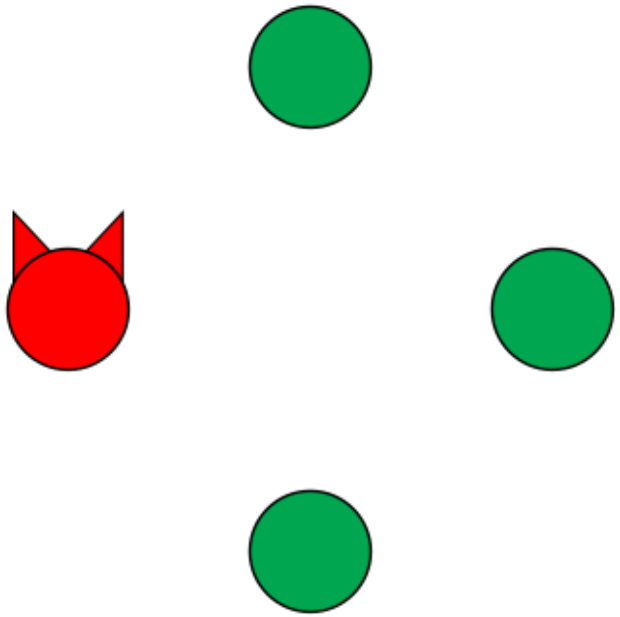
$n = 3f + 1$ replicas

f replicas may byzantine

Properties

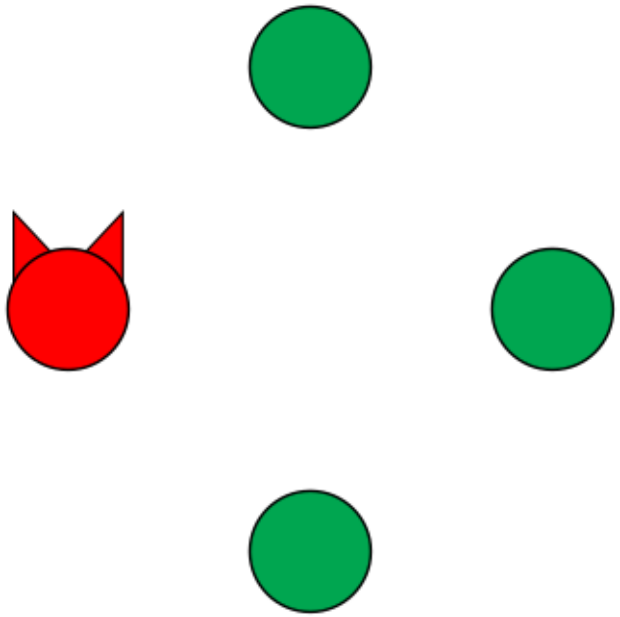


Properties



safety

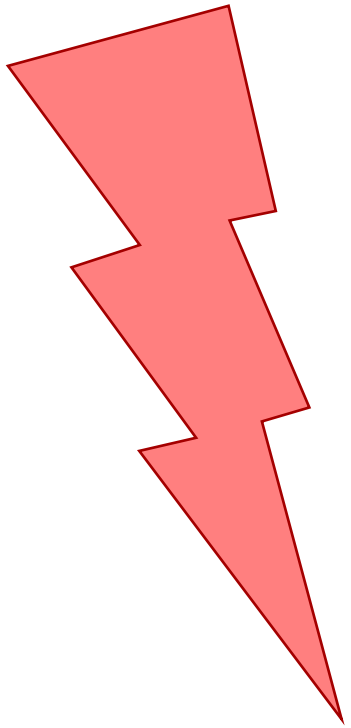
Properties



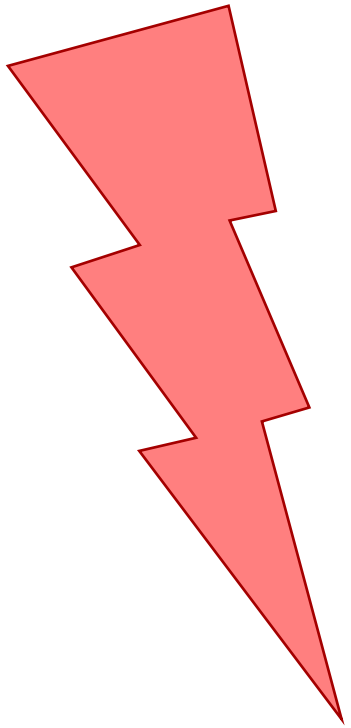
safety

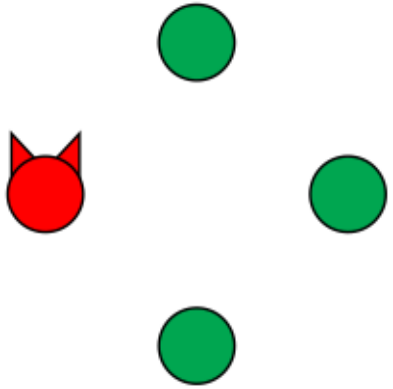
liveness

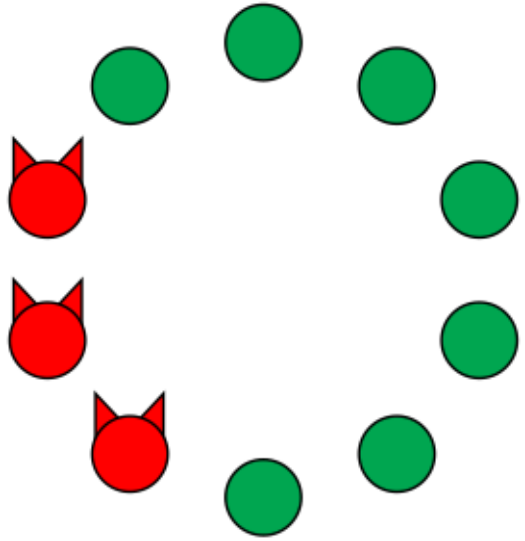
BFT protocols



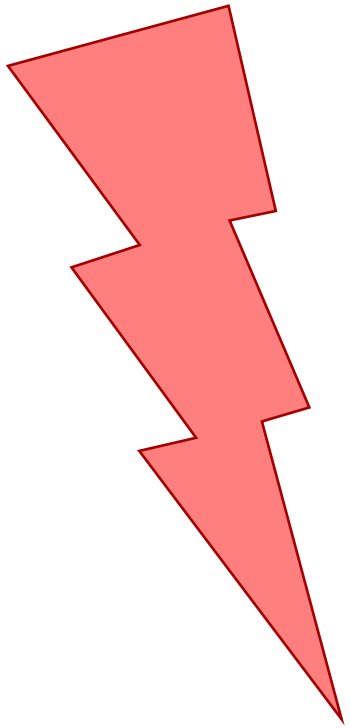
BFT protocols

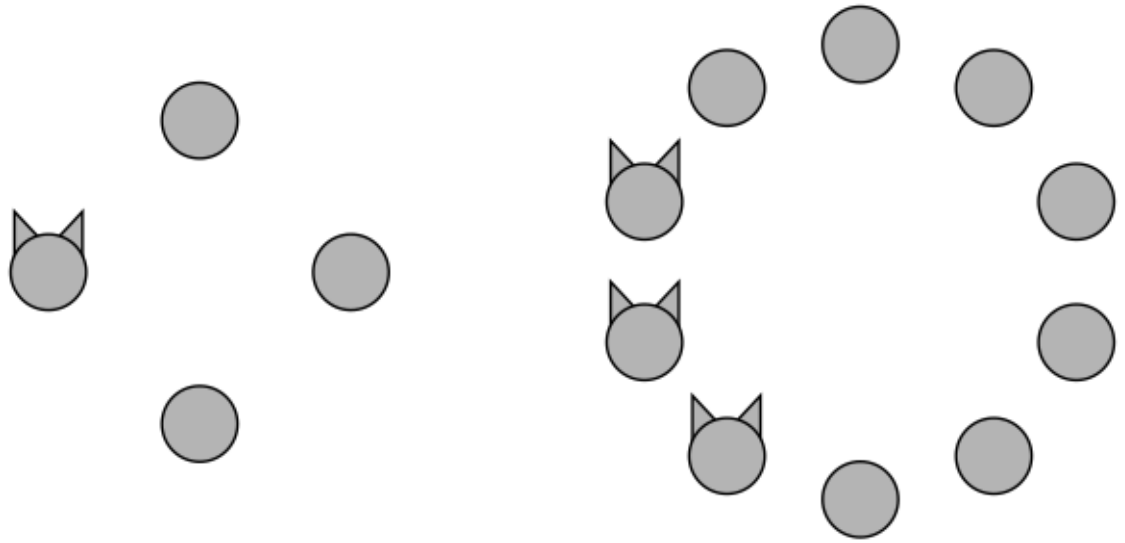


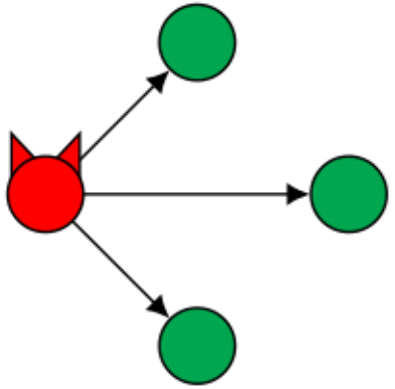


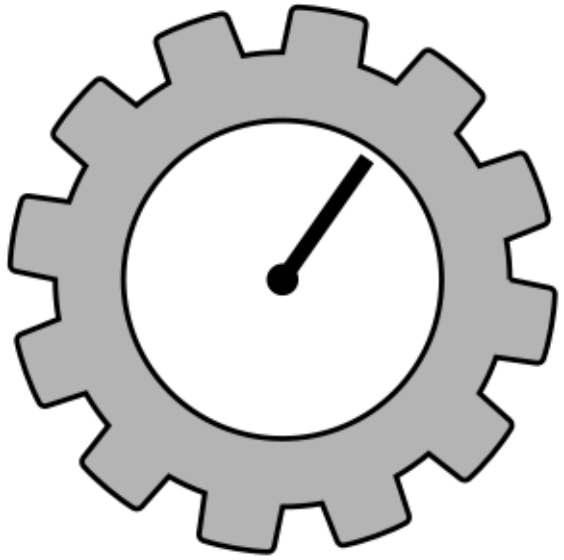


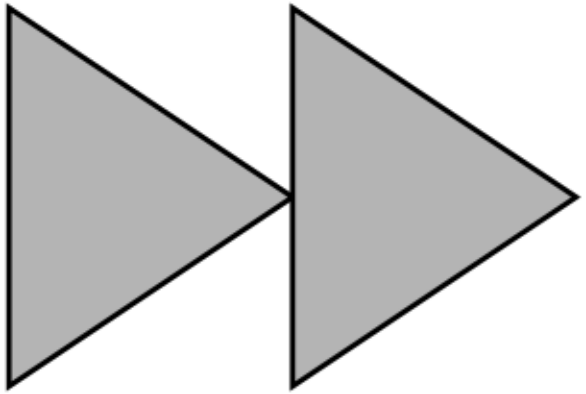
BFT protocols

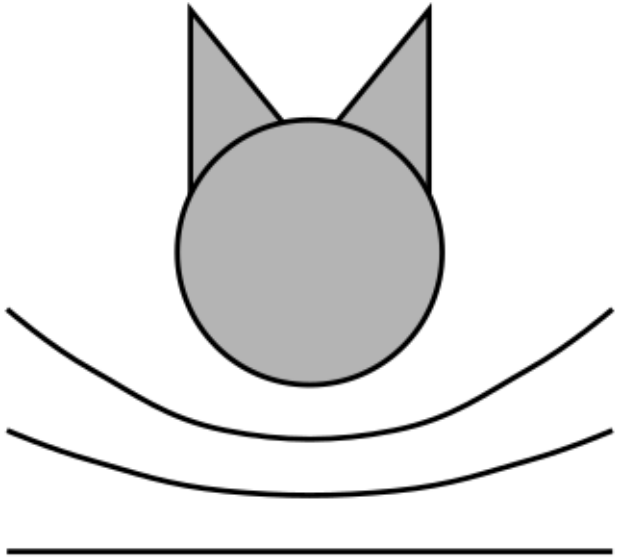


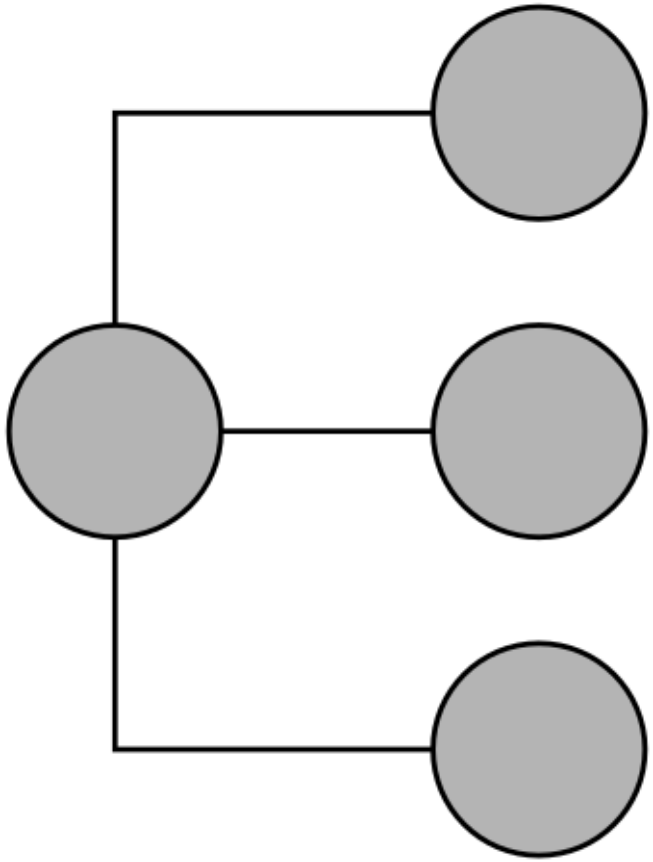


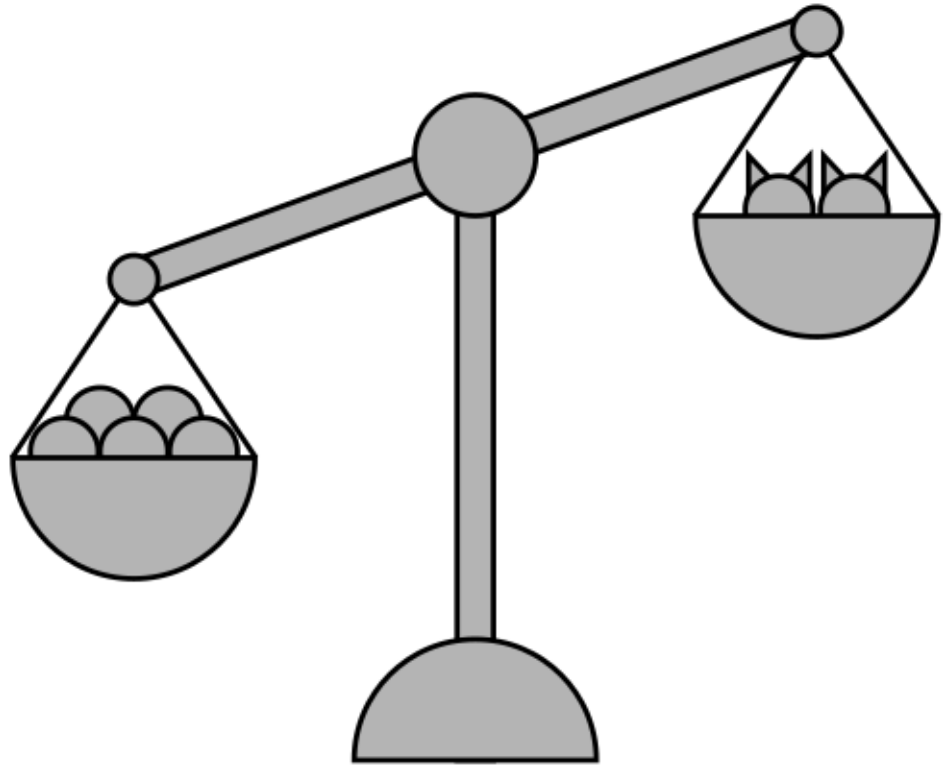




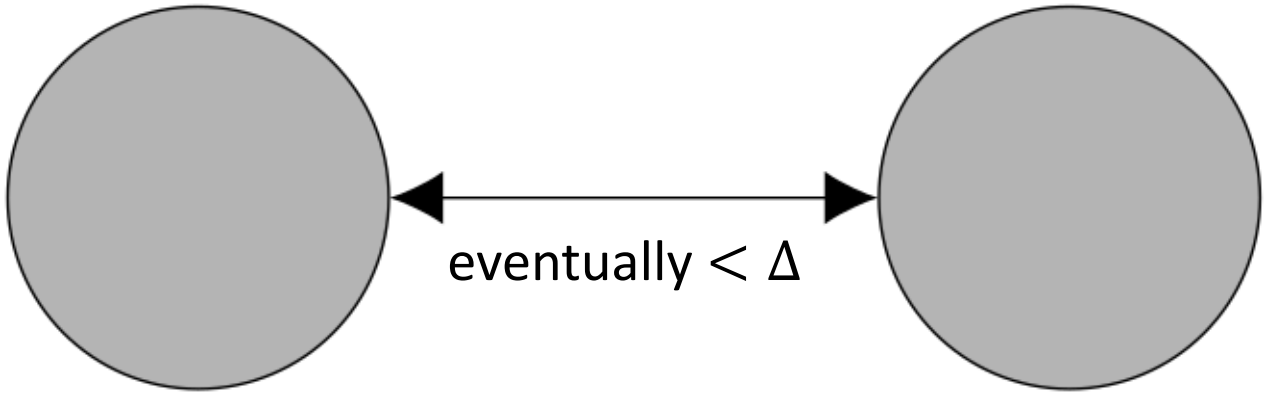




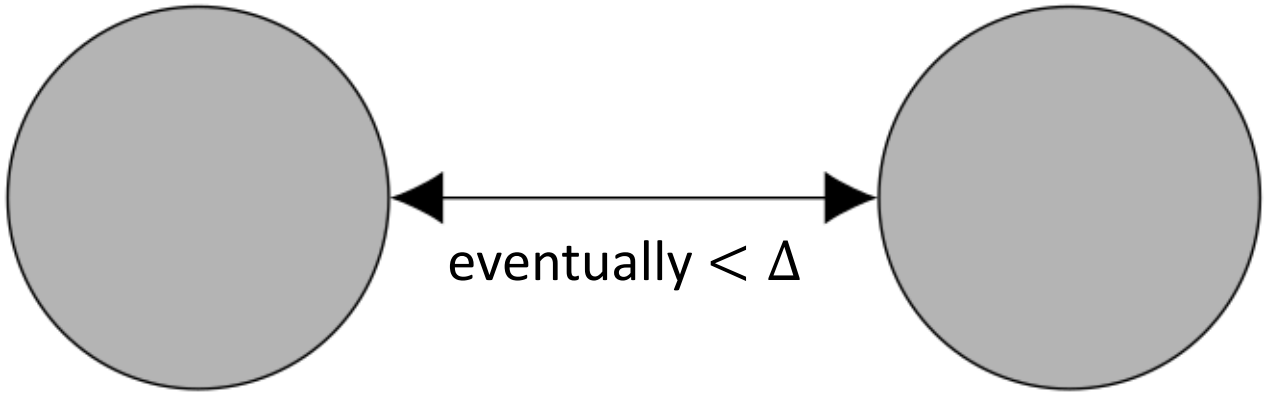




Communication model

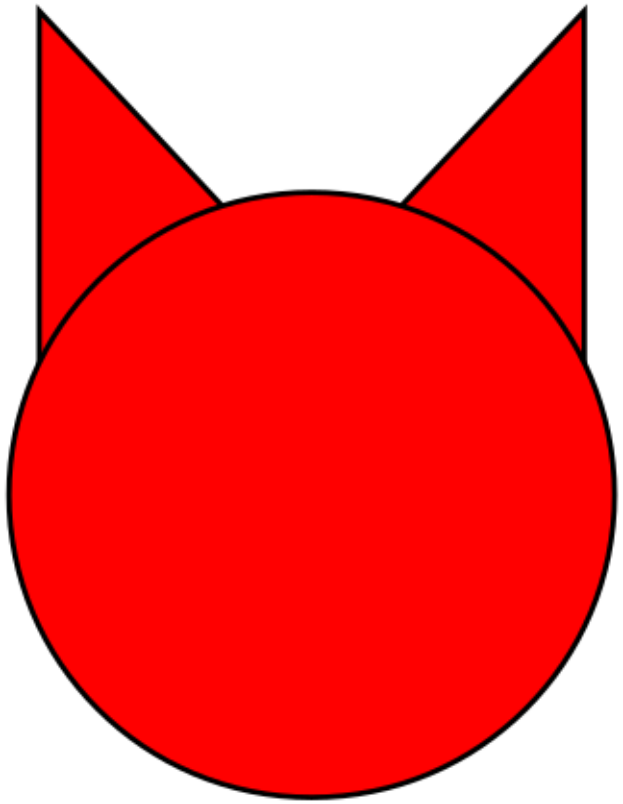


Communication model



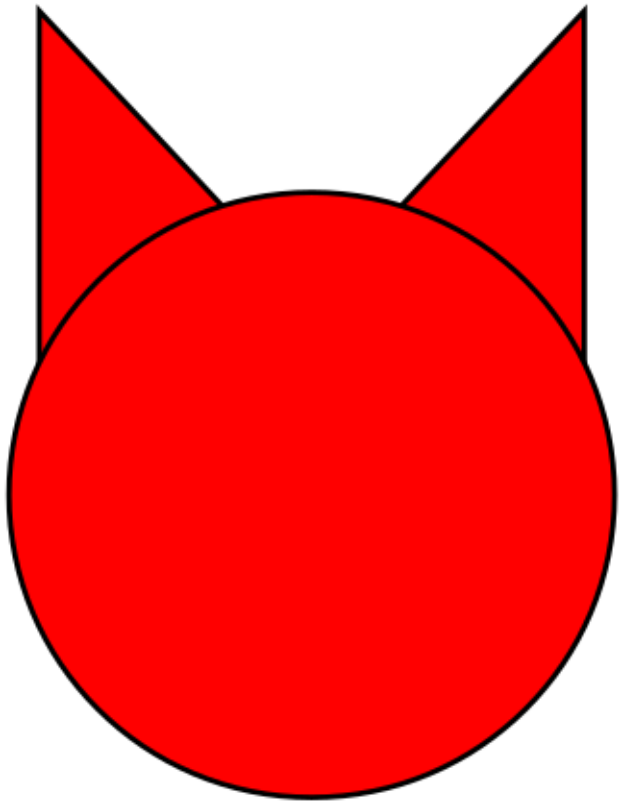
$n = 3f + 1$ replicas, f replicas may byzantine

Model



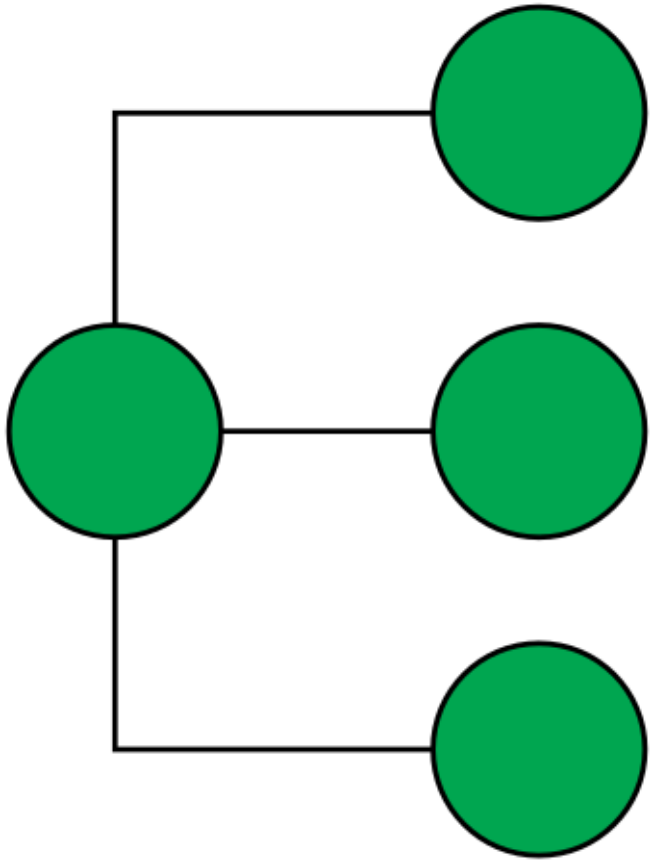
f byzantine replicas

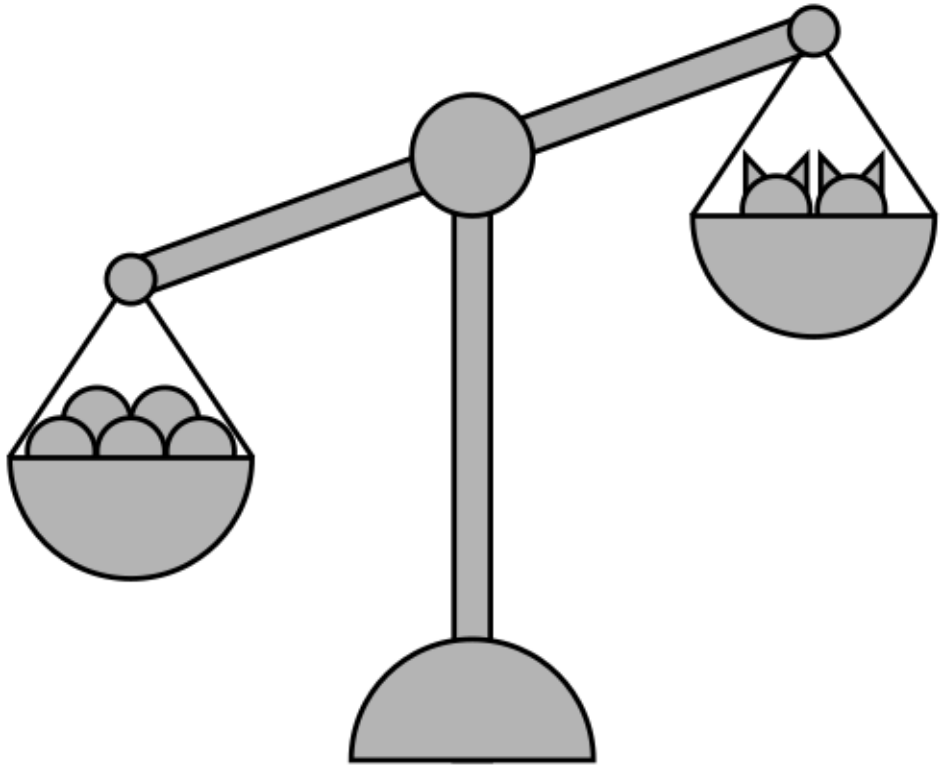
Model



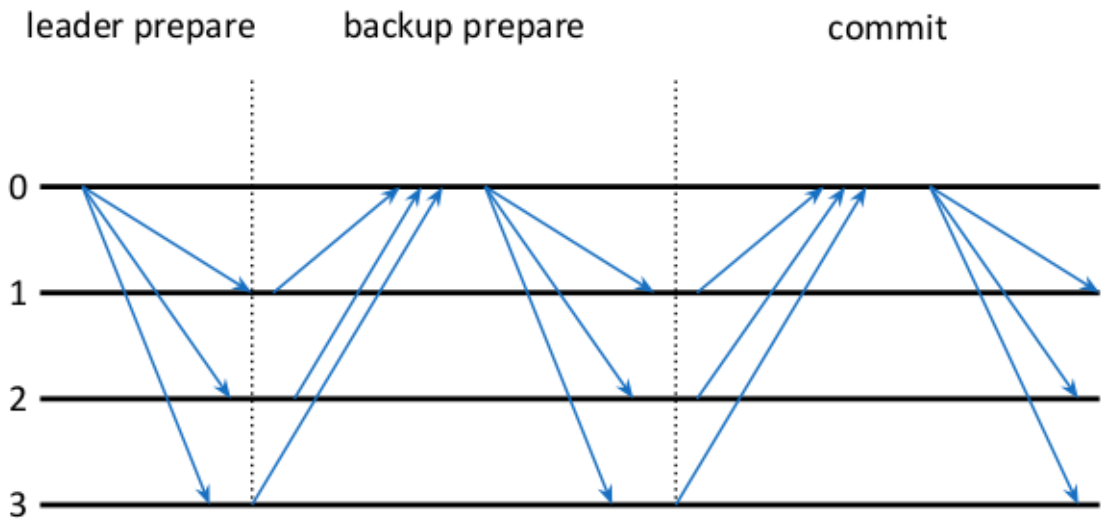
f byzantine replicas

any number byzantine clients

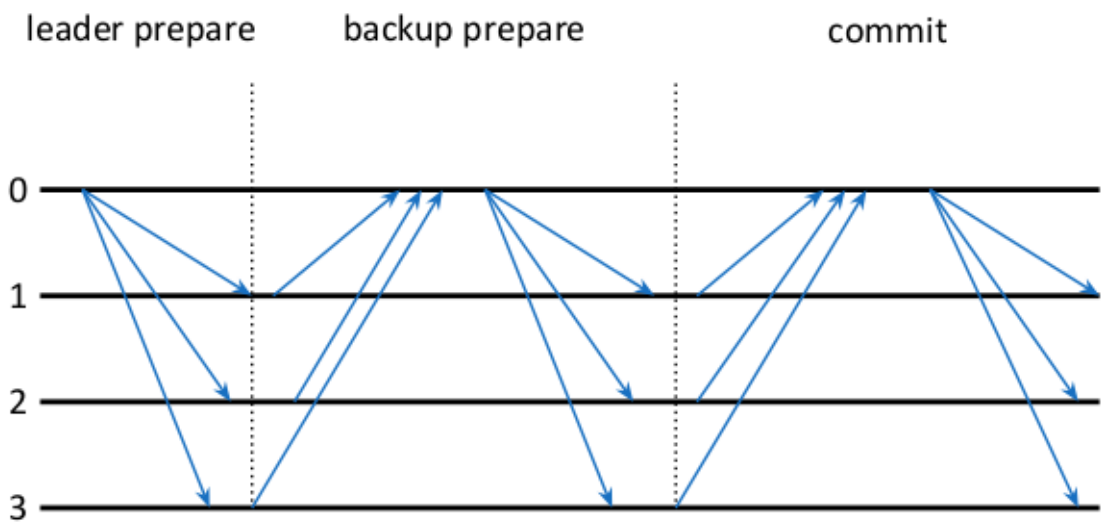


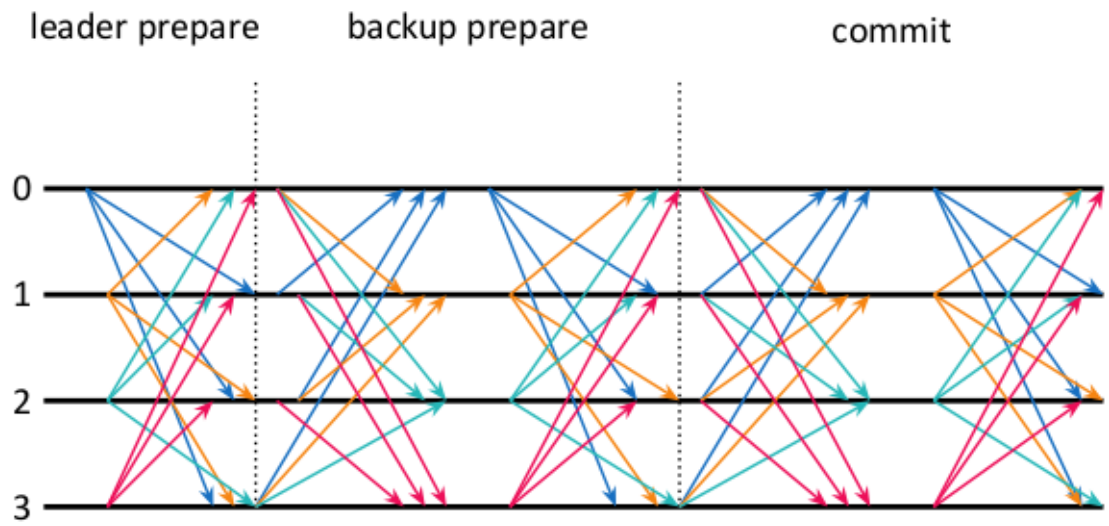


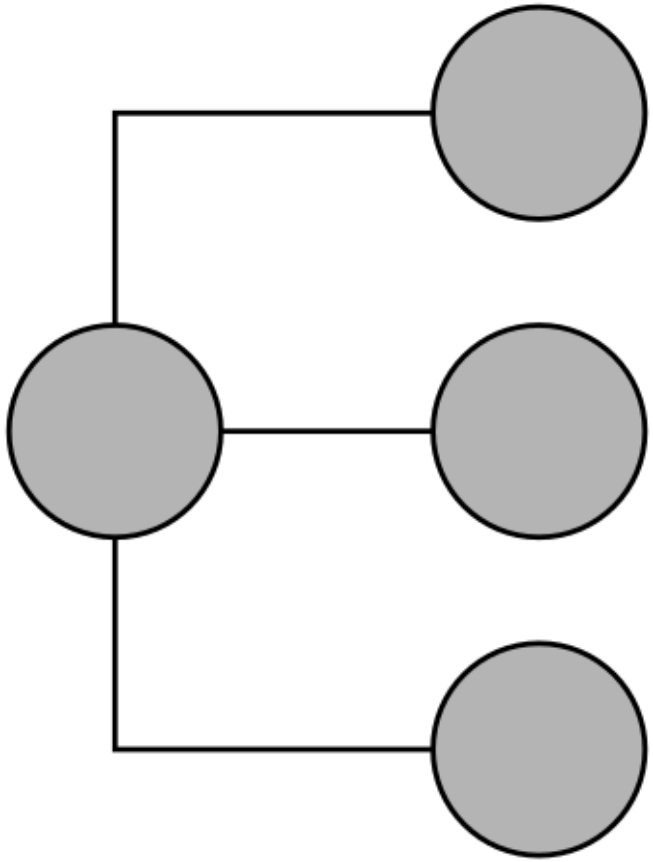
Parallelize

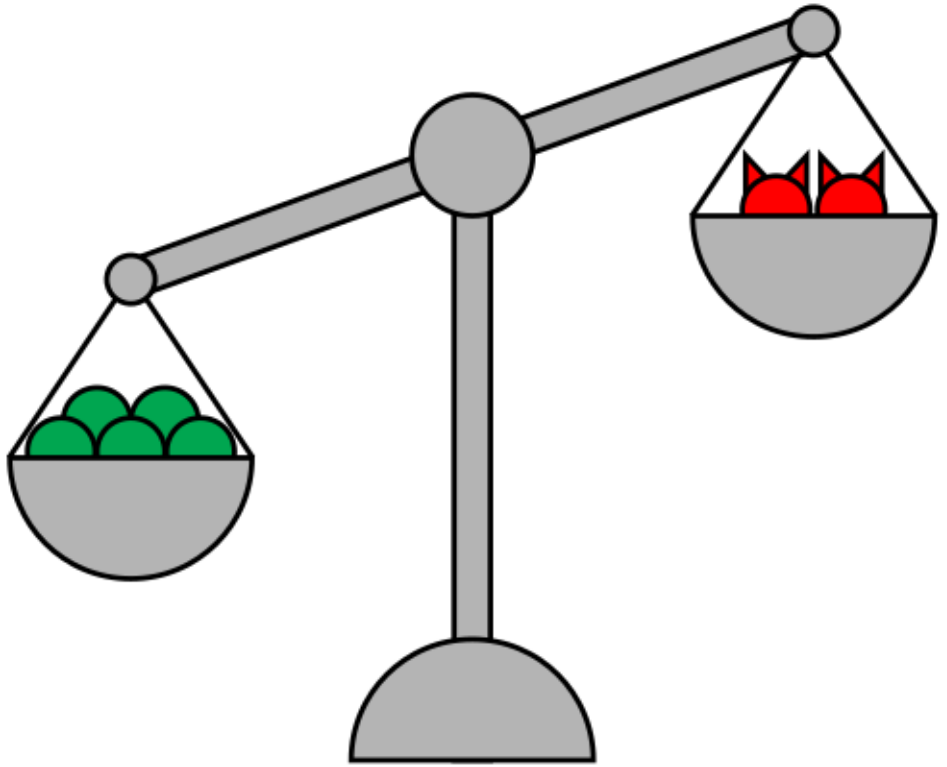


Parallelize

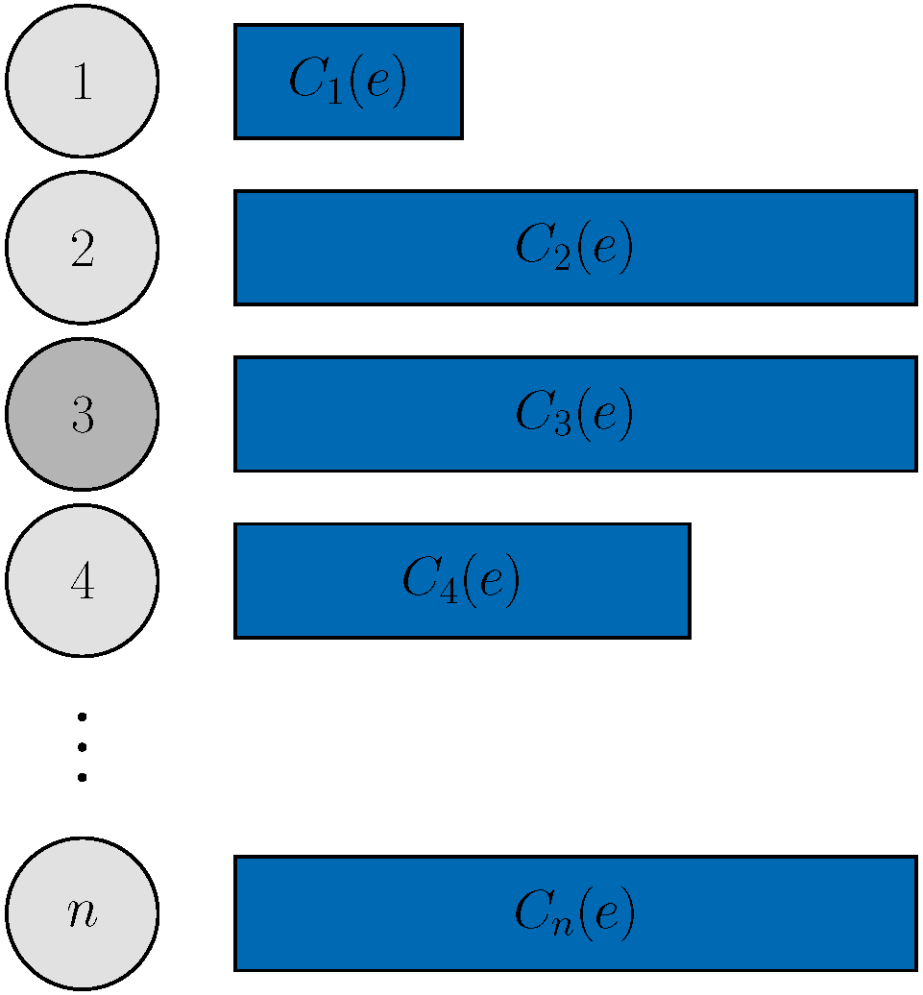




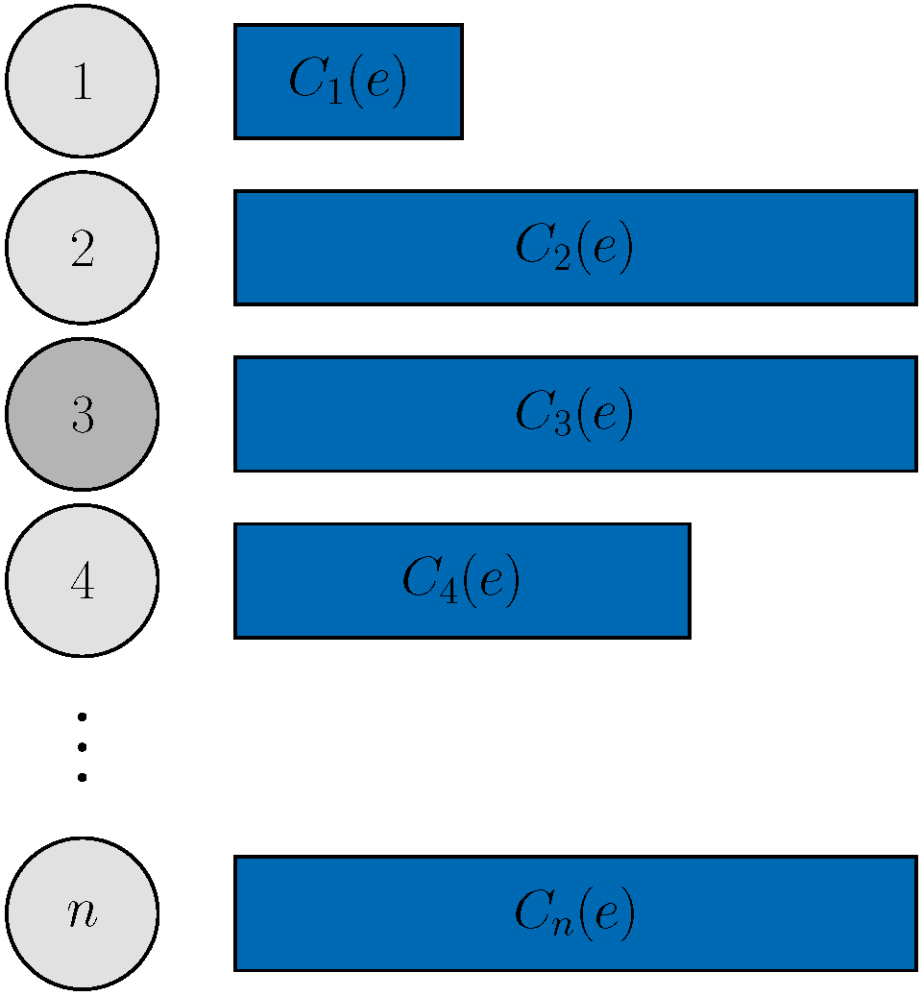




Parallel leaders

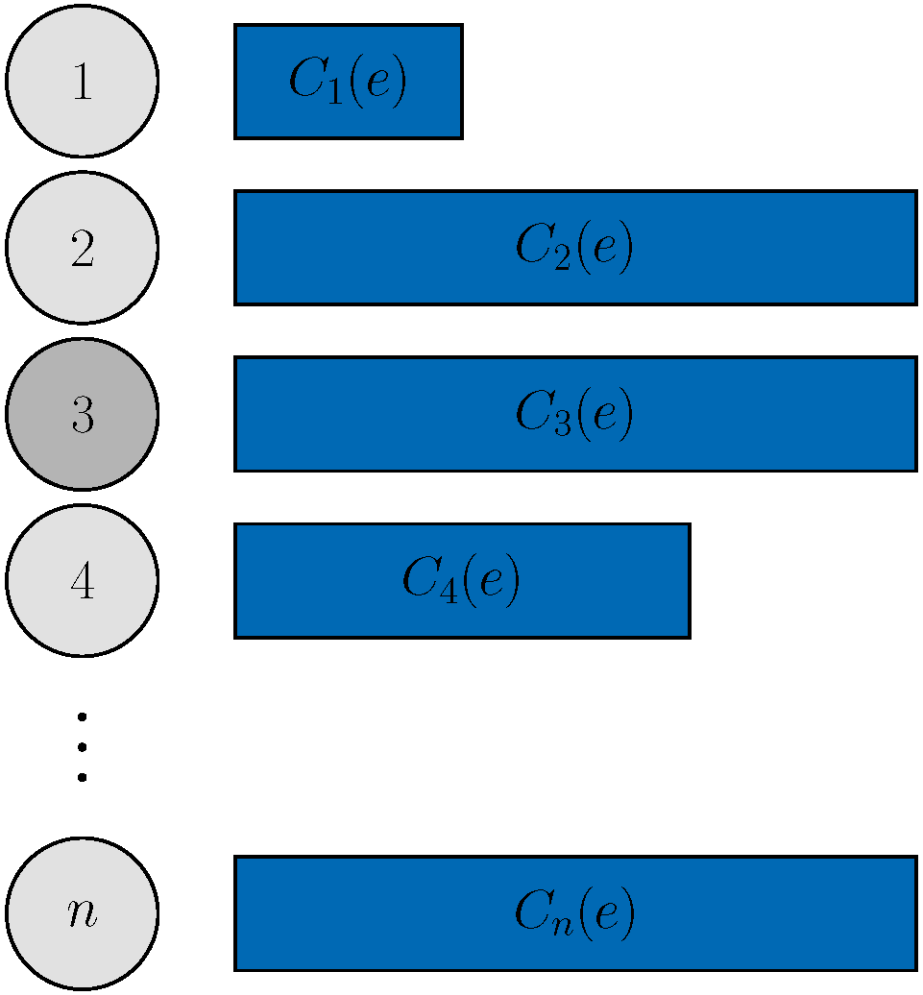


Configuration



$C_v(e)$ number of requests assigned to leader v

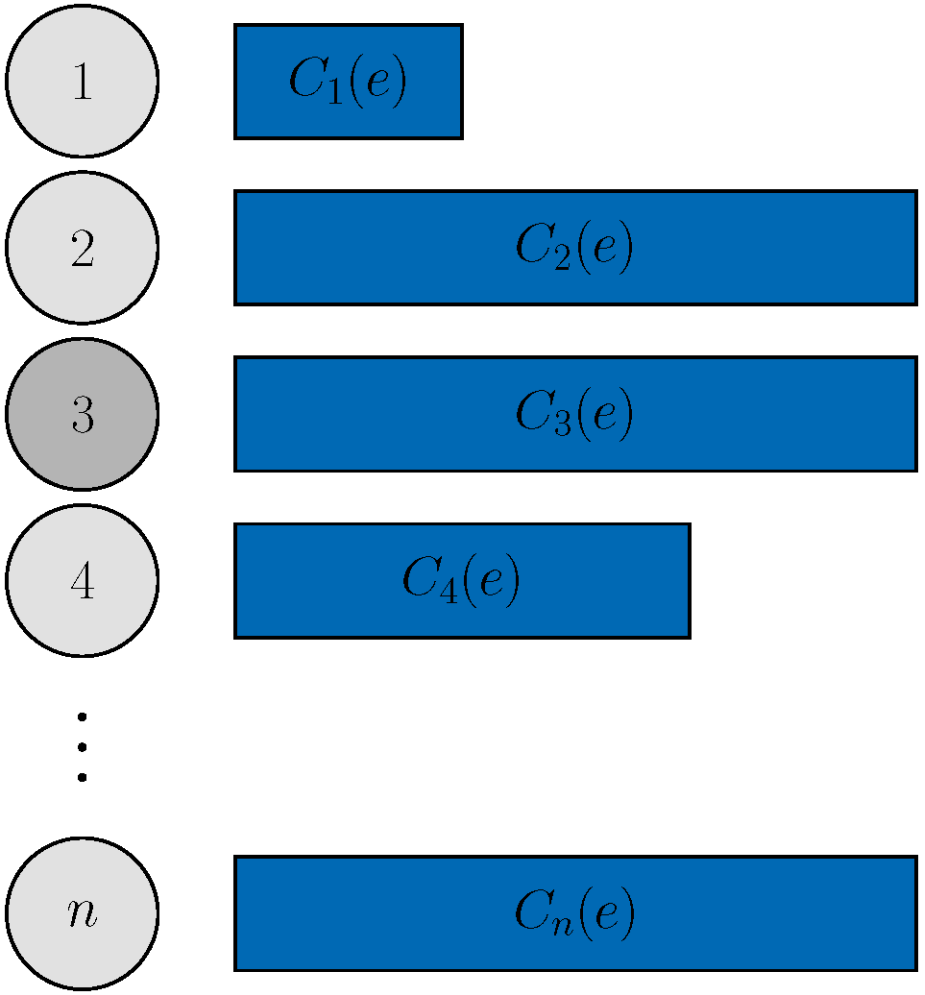
Configuration



$C_v(e)$ number of requests assigned to leader v

$C_v(e)$ initially $C_{\min} \in \Omega(n^2)$ for all $v \in [n]$

Optimization

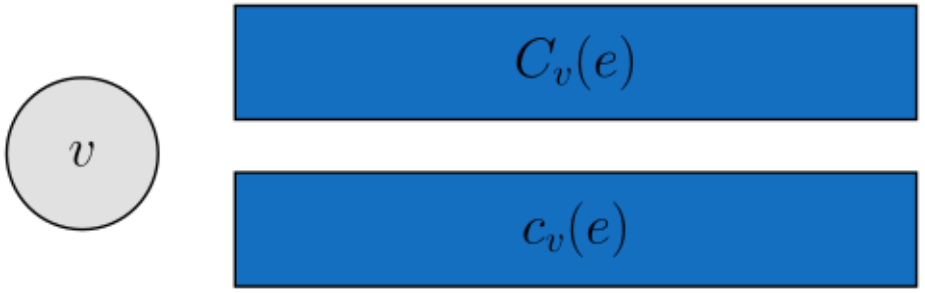


$C_v(e)$ number of requests assigned to leader v

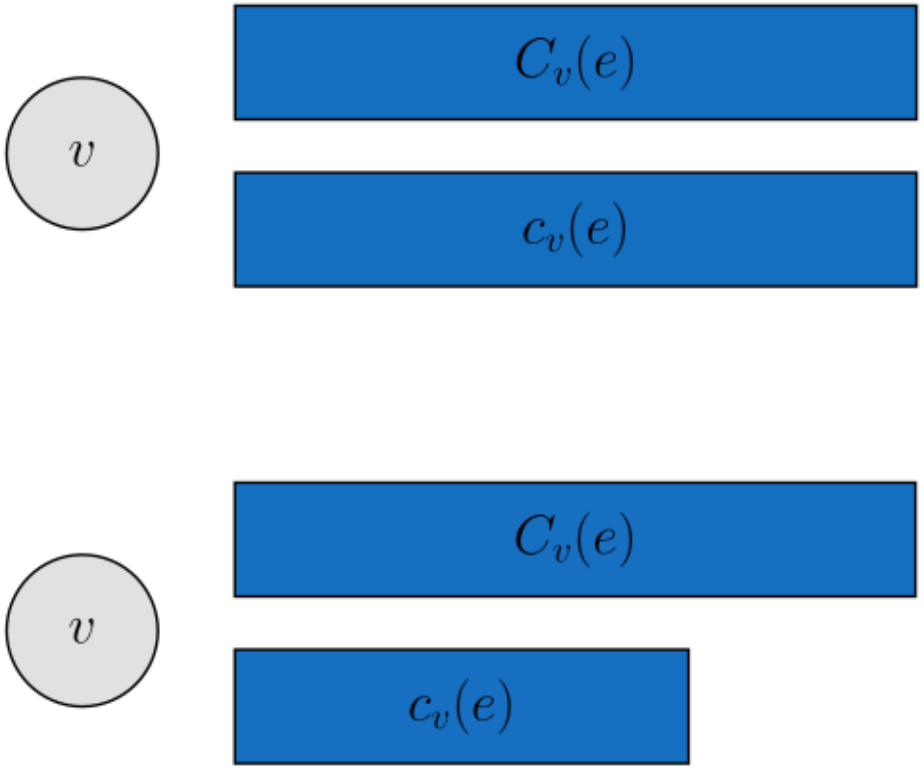
$C_v(e)$ initially $C_{\min} \in \Omega(n^2)$ for all $v \in [n]$

$C_v(e)$ updated with leader v's performance

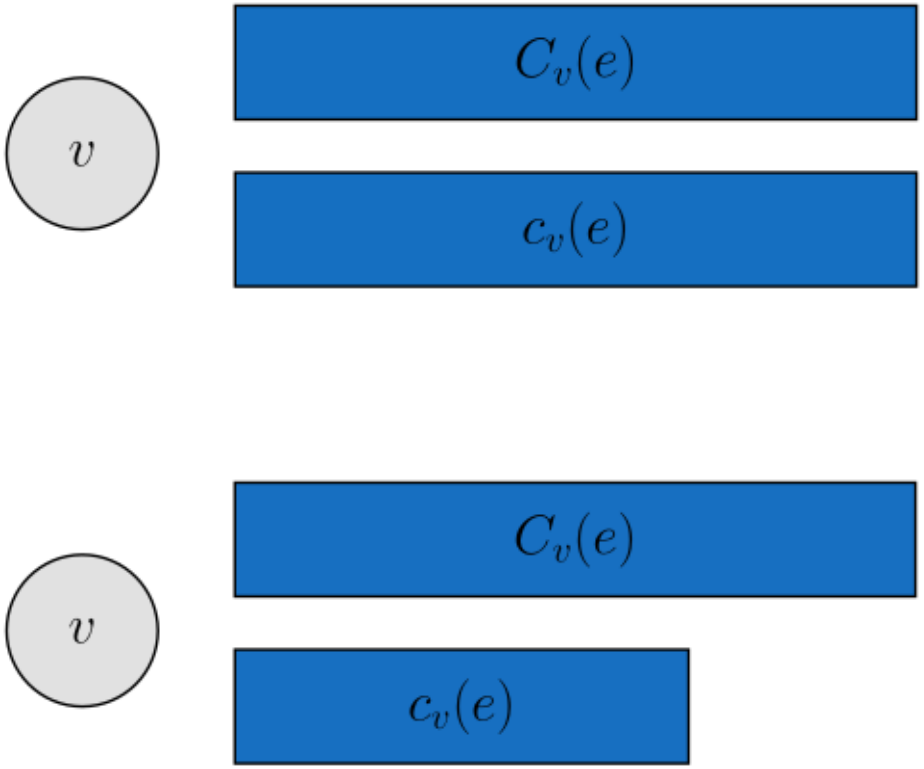
Optimization



Optimization

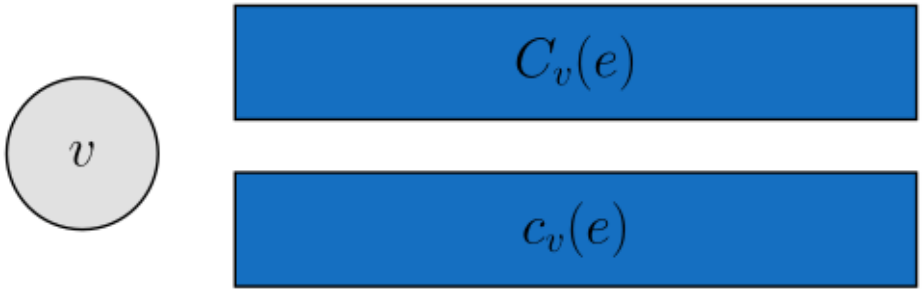


Optimization

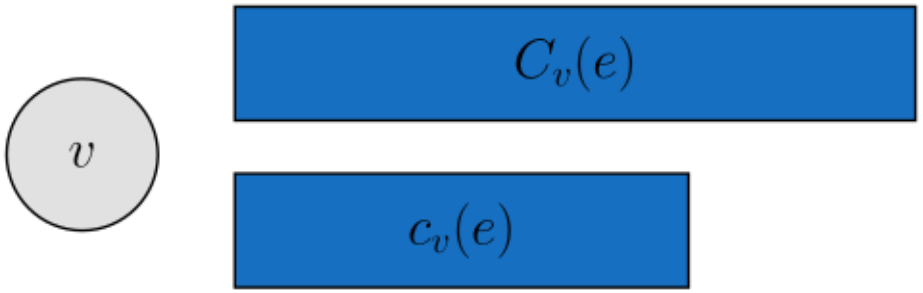


$$C_v(e+1) = 2 \cdot c_{v(e)}$$

Optimization

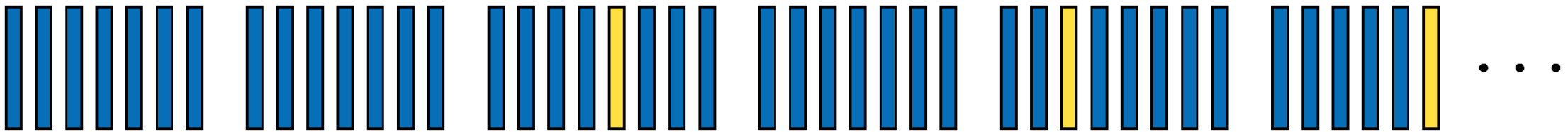


$$C_v(e+1) = 2 \cdot c_{v(e)}$$

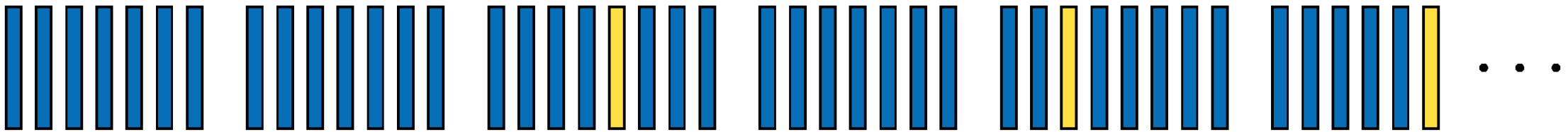


$$C_v(e+1) = \max(C_{\min}, \max_{i \in \{0,\dots,f\}} c_v(e-i)$$

Primary rotation

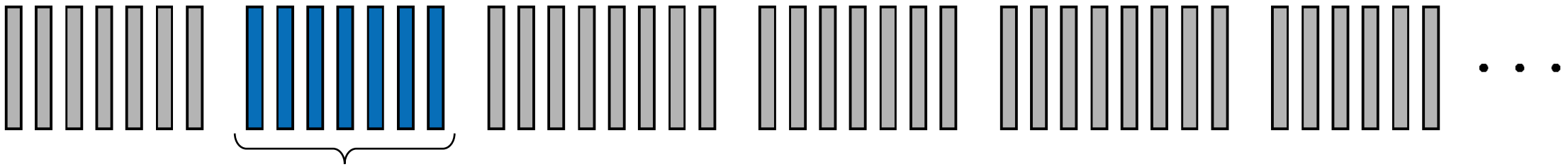


Primary rotation

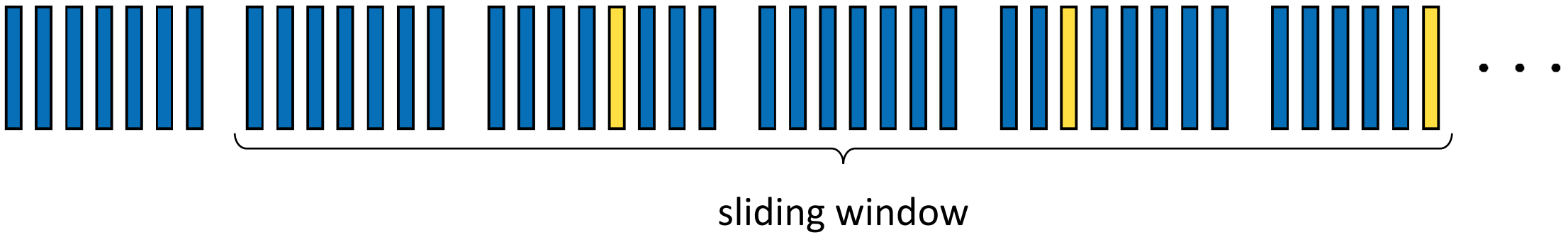


explore and exploit

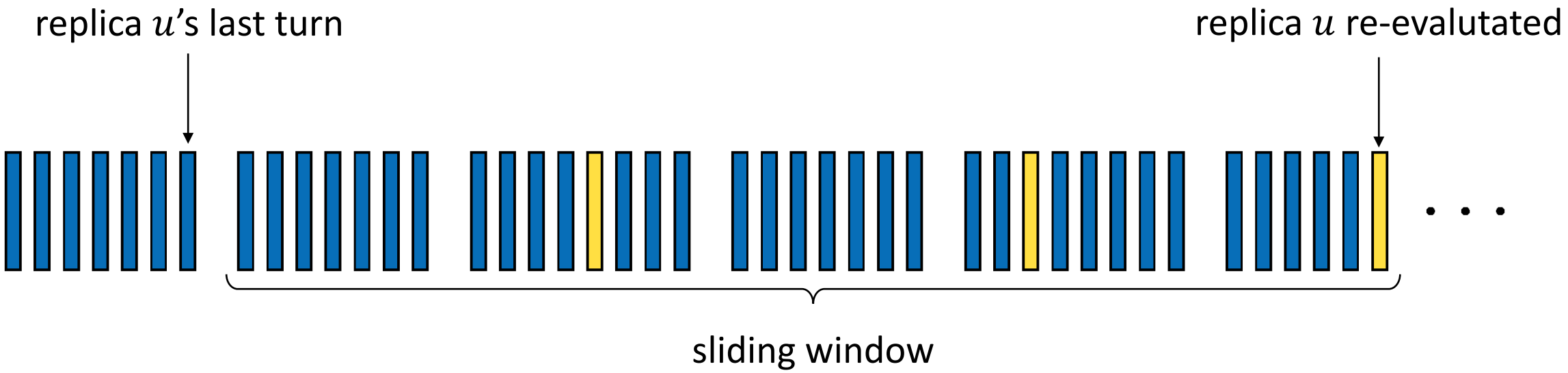
Exploit best primaries

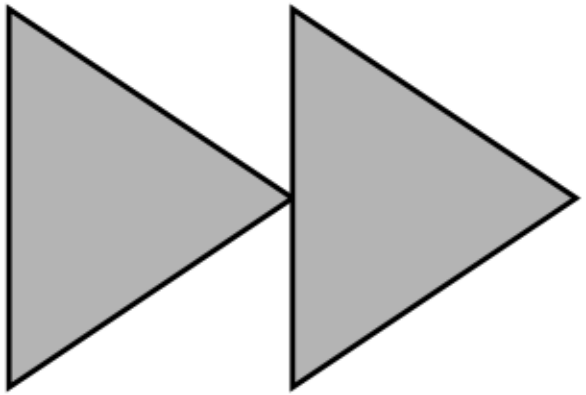


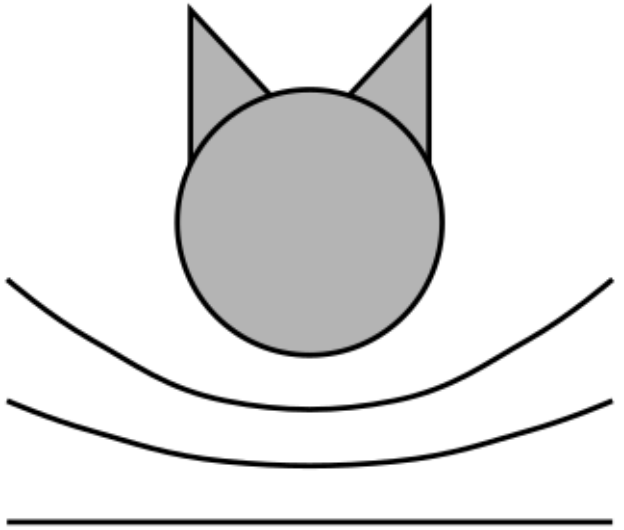
Re-evaluate primaries



Re-evaluate primaries





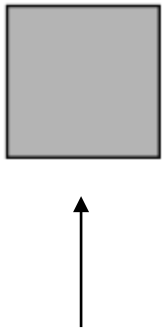


Efficiency

complexity measure: authenticator complexity

Efficiency

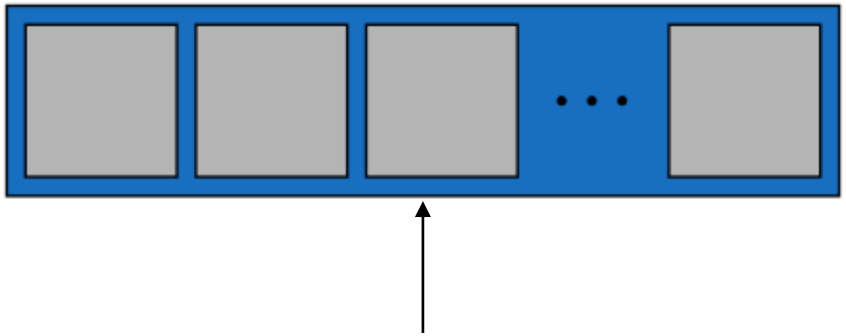
complexity measure: authenticator complexity



request creation cost: $O(n)$

Efficiency

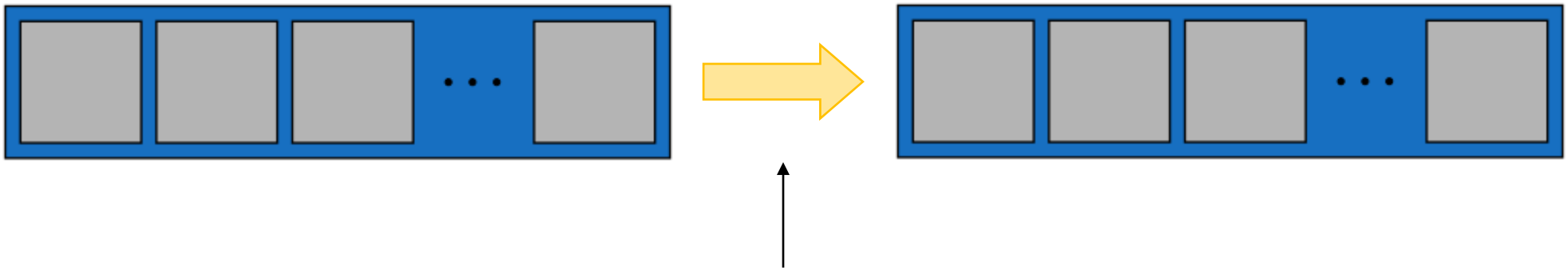
complexity measure: authenticator complexity



average number of requests: $\Omega(n^3)$

Efficiency

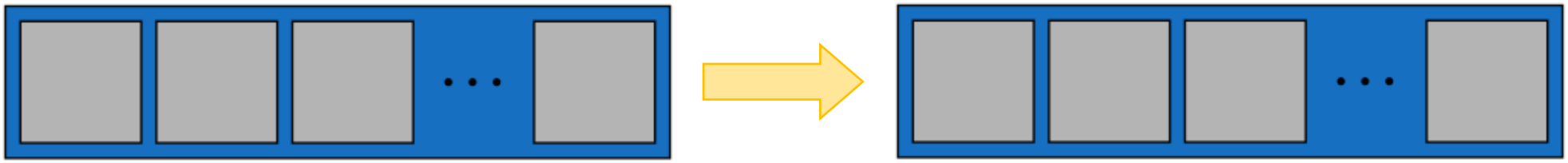
complexity measure: authenticator complexity



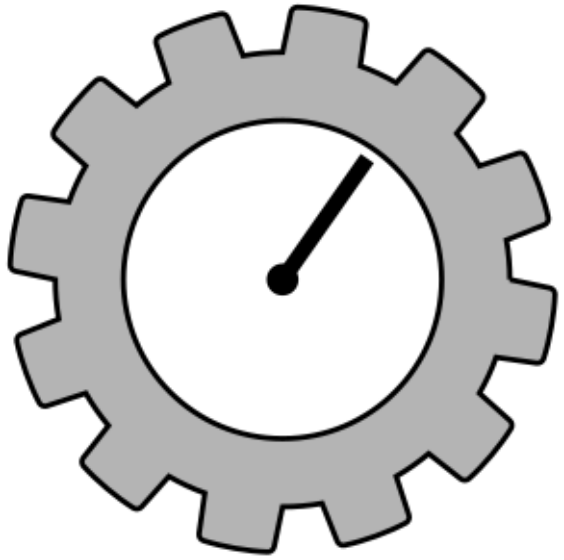
epoch-change cost: $O(n^3)$

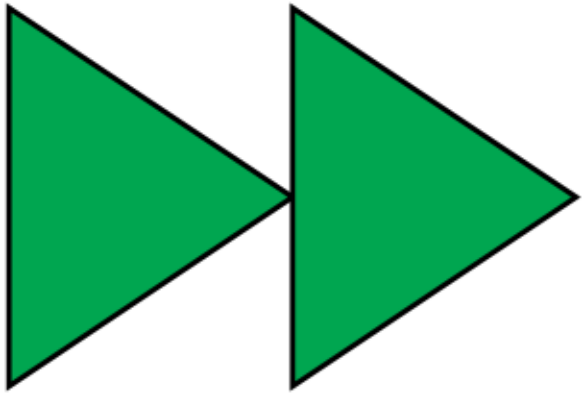
Efficiency

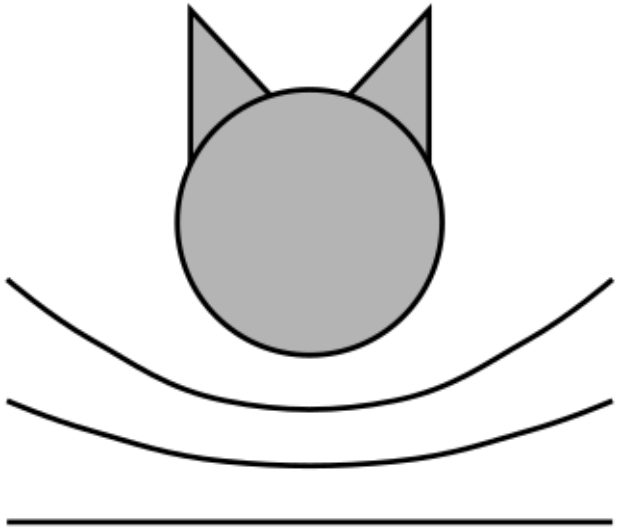
complexity measure: authenticator complexity



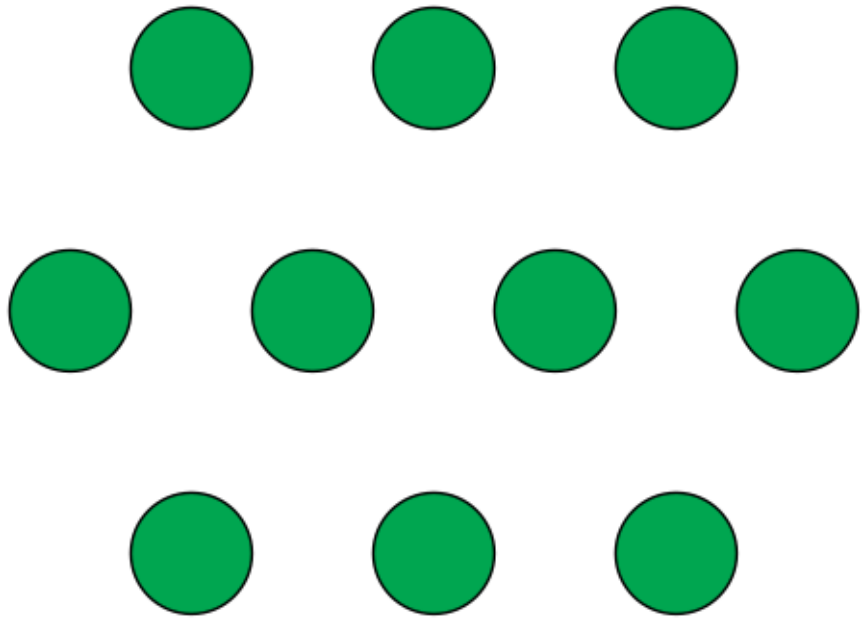
amortized request cost: $O(n)$



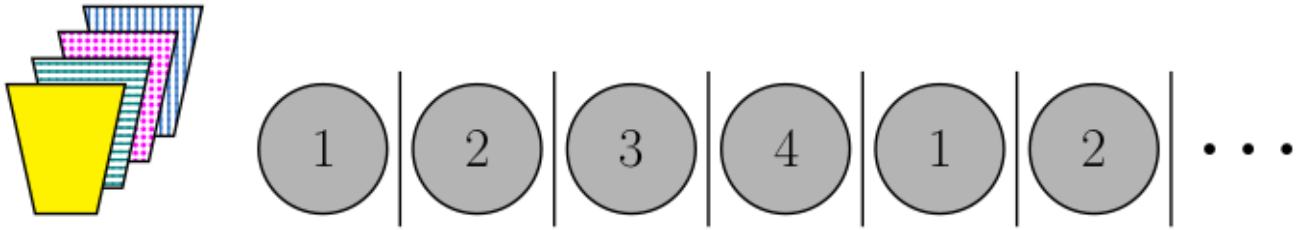




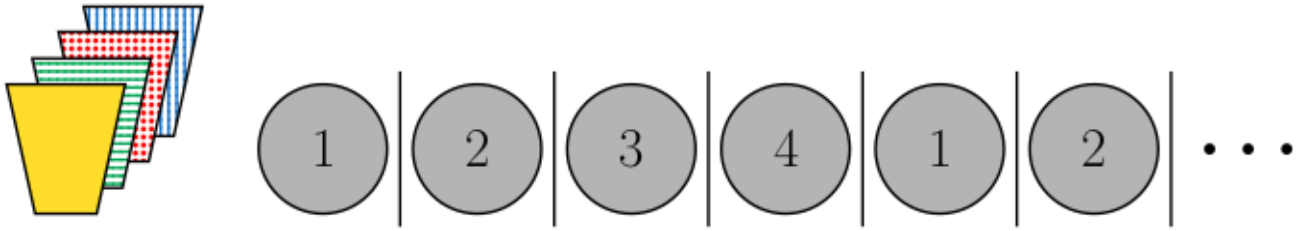
Assumptions



Speedup

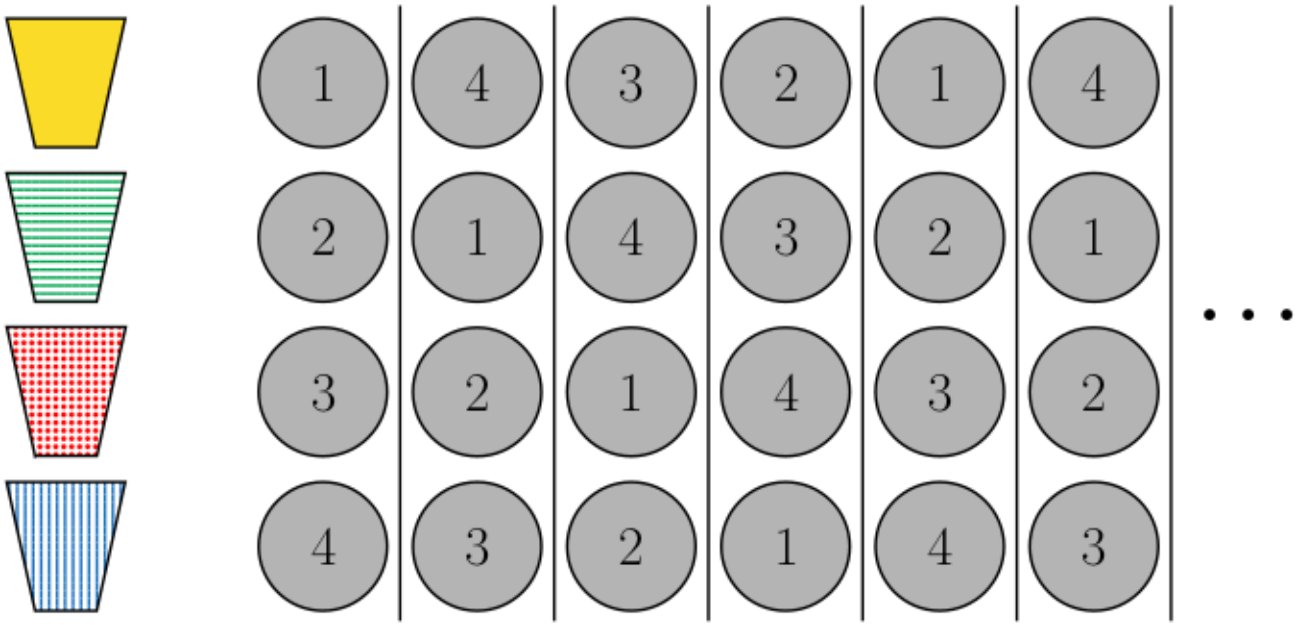


Speedup

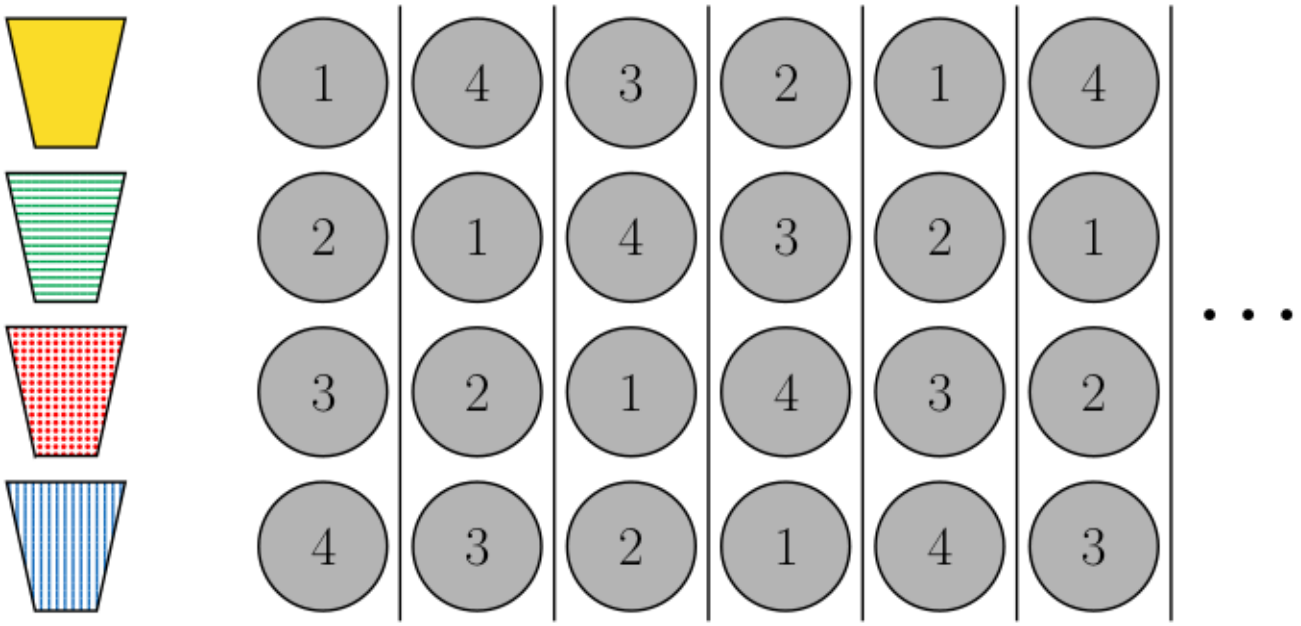


time units per request:
$\Theta(n)$

Speedup

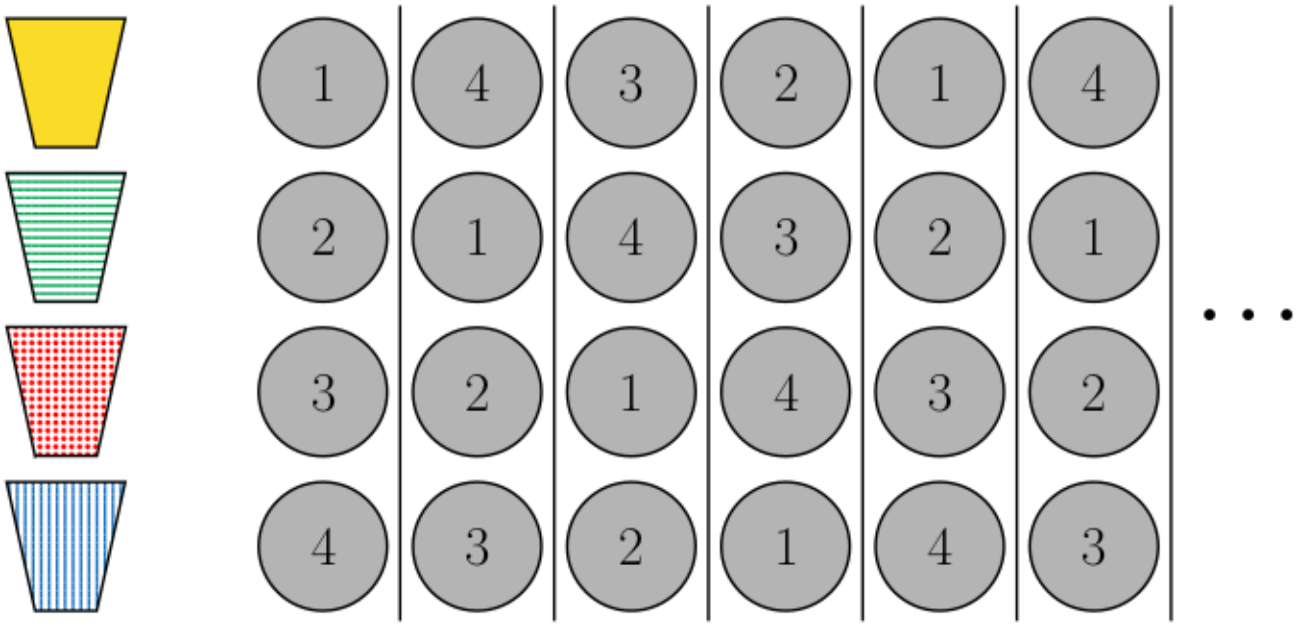


Speedup



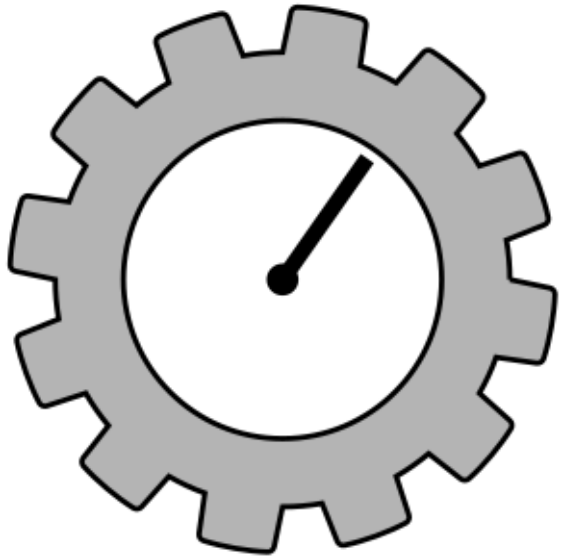
time units per request:
$\Theta(1)$

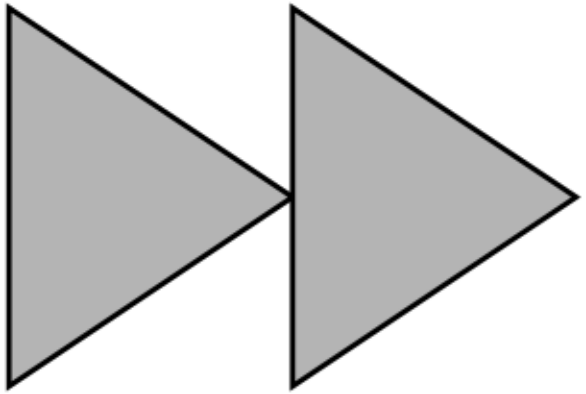
Speedup

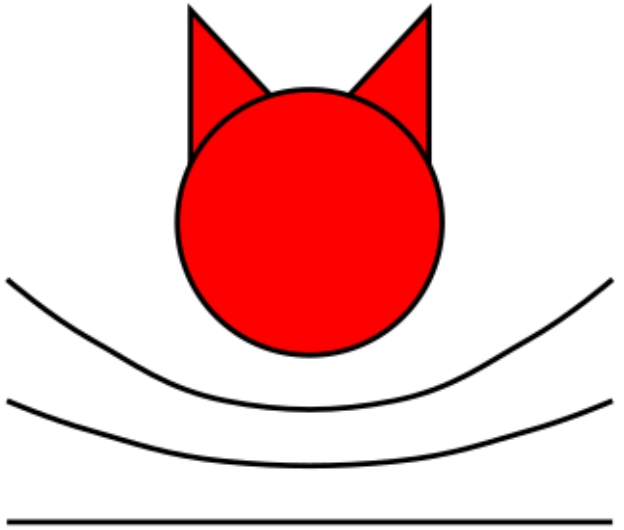


time units per request:
$\Theta(1)$

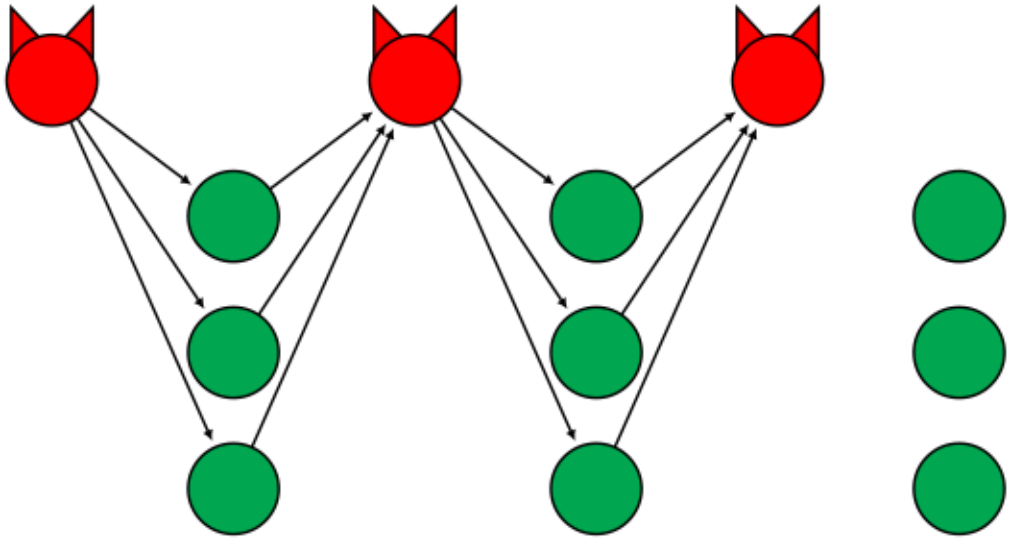
speedup:
$\Theta(n)$



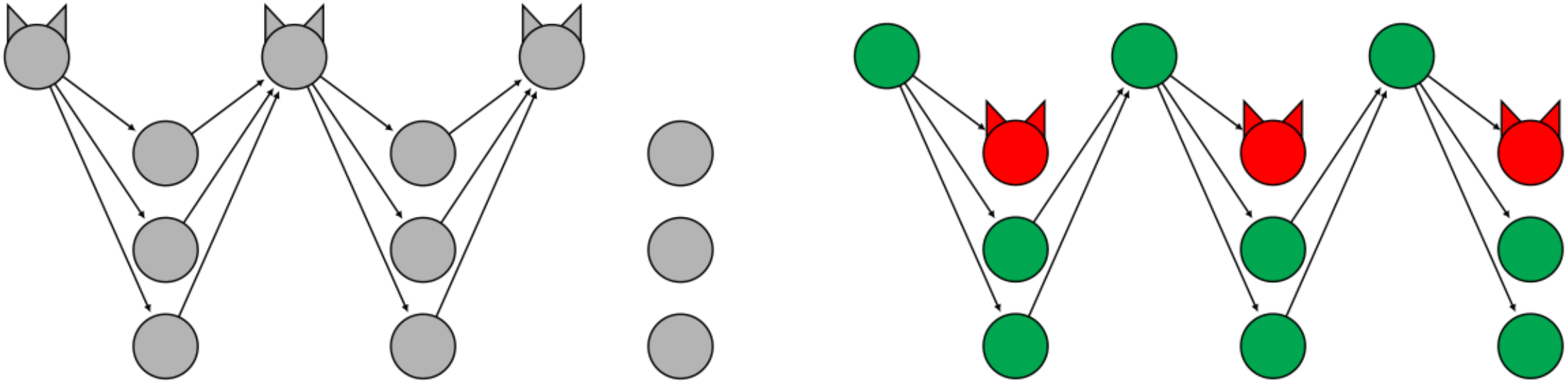




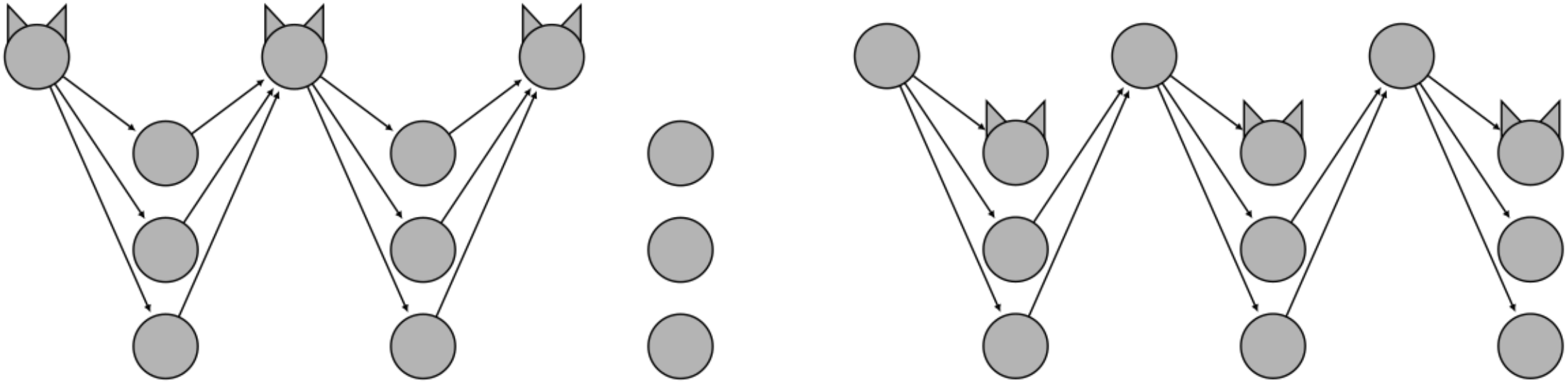
Correct primary epochs



Correct primary epochs

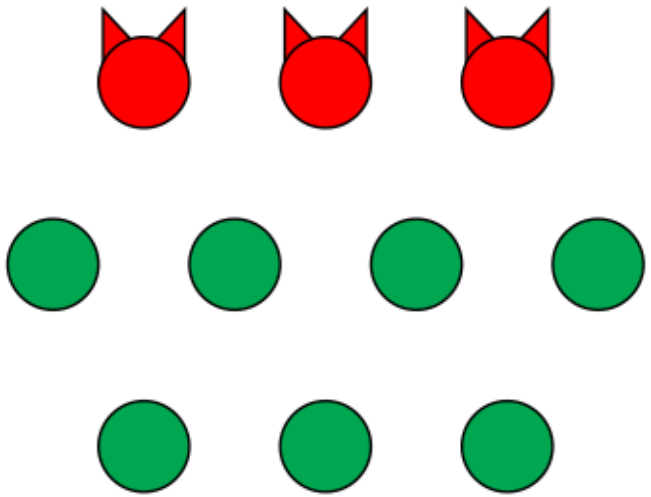


Correct primary epochs

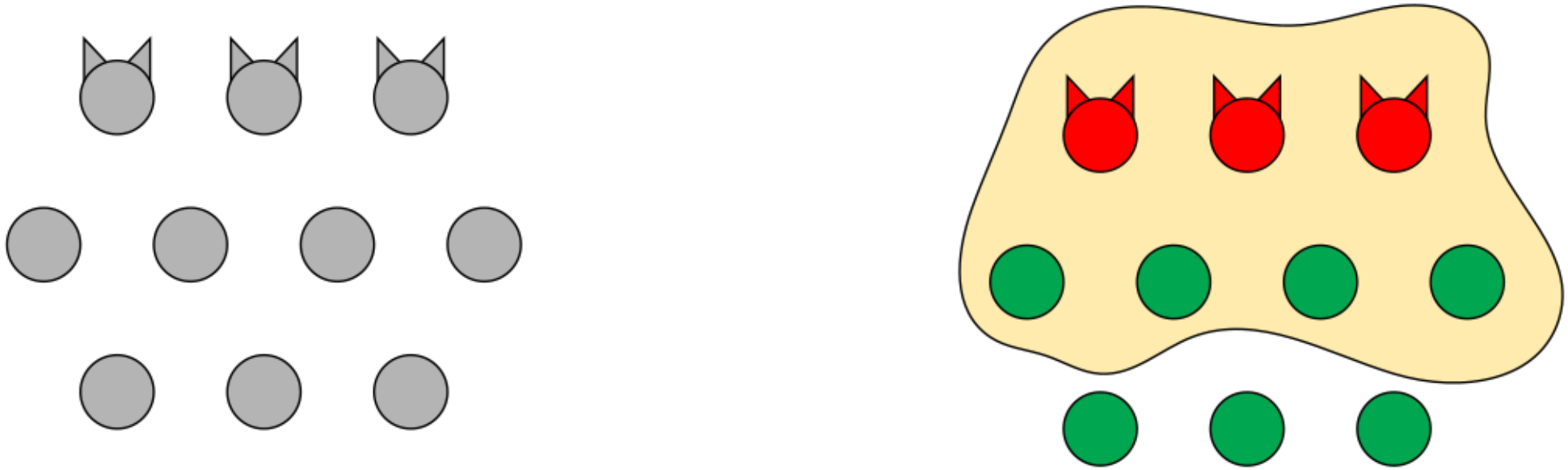


effective utilization: $\frac{8}{9}$

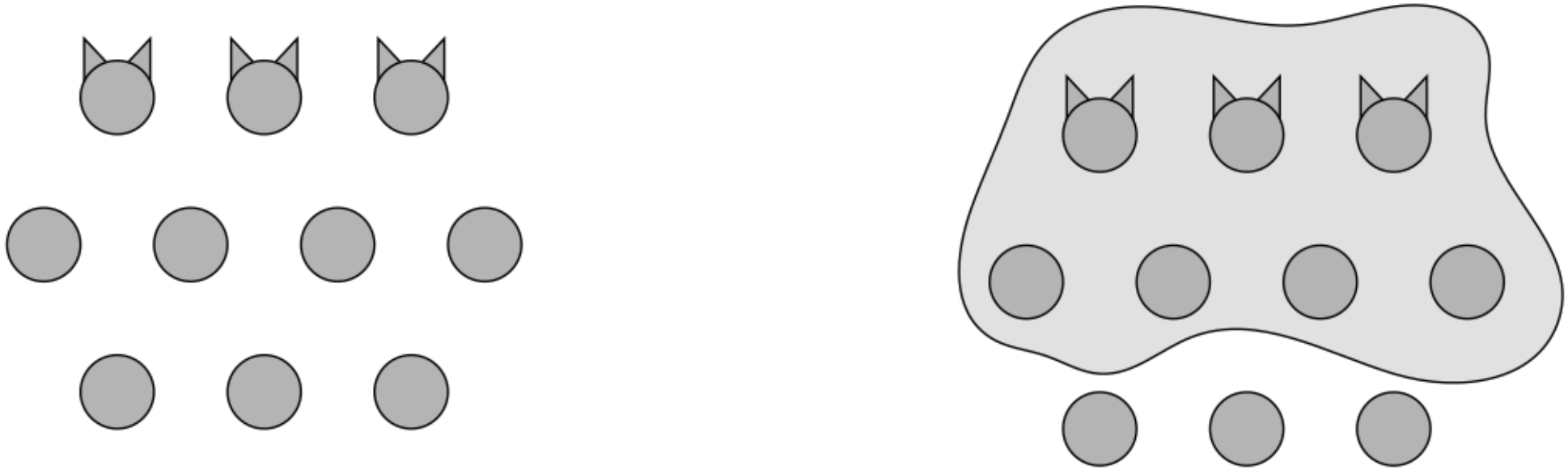
Byzantine primary epochs



Byzantine primary epochs

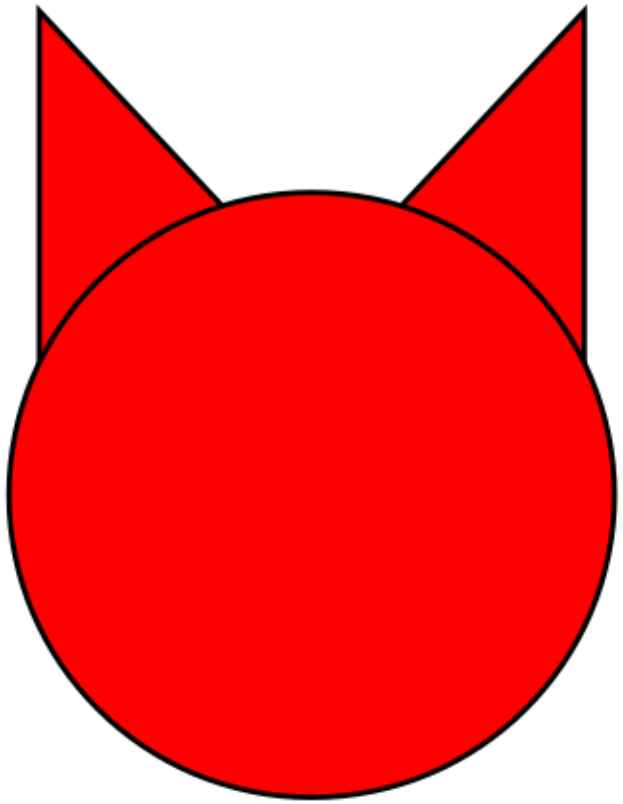


Byzantine primary epochs



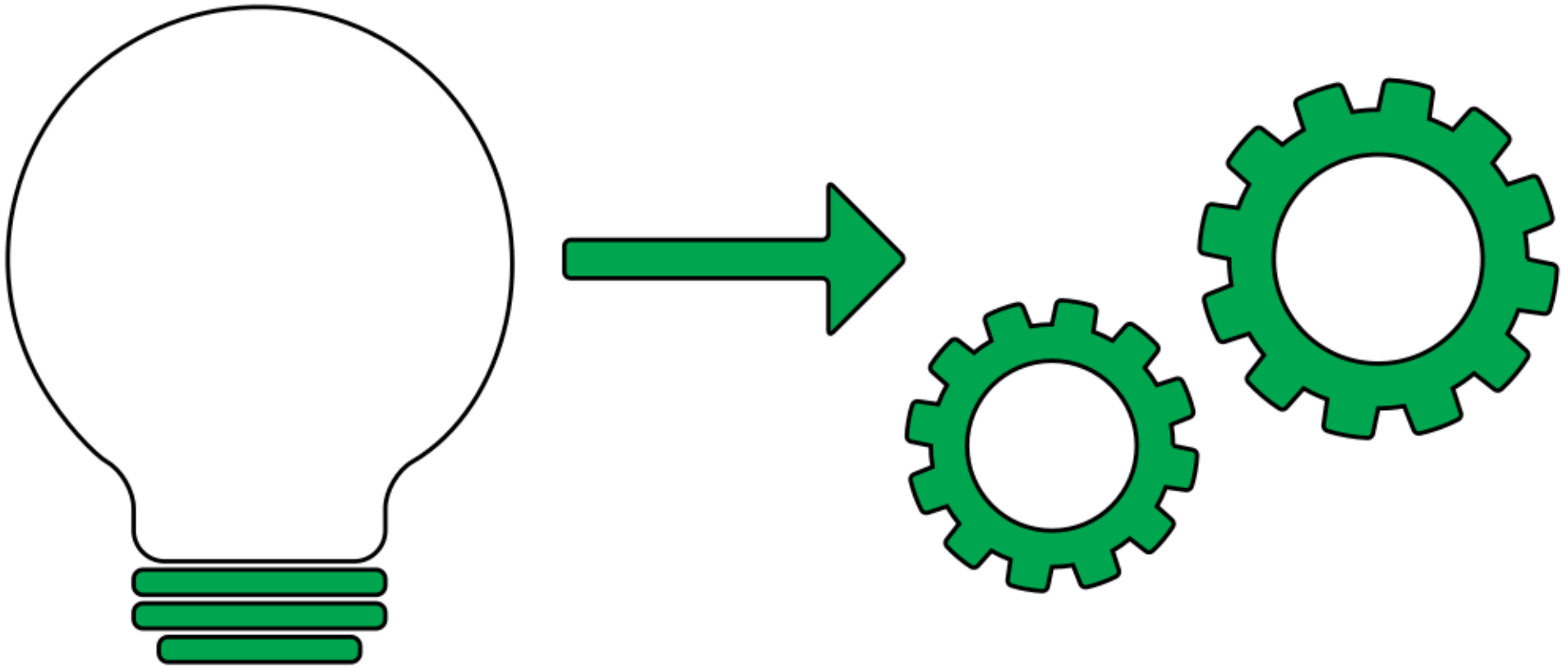
ratio of byzantine primaries: $\frac{f}{g}$

Effective utilization

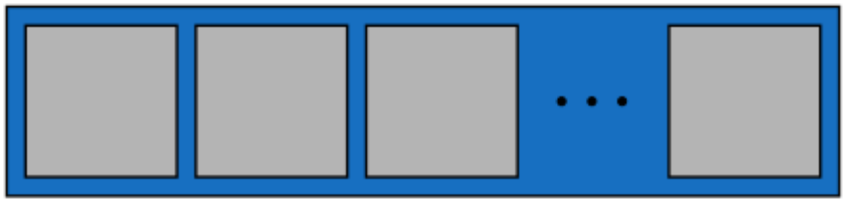


effective utilization:

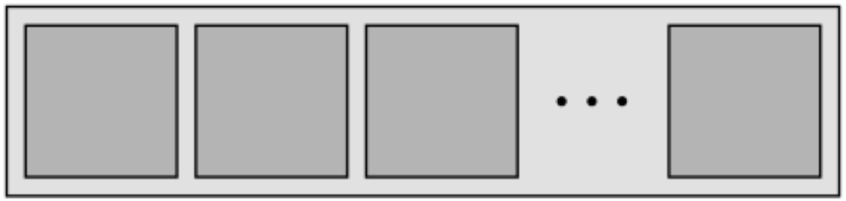
$$\frac{8}{9} \cdot \frac{g-f}{g} \geq \frac{16}{27}$$

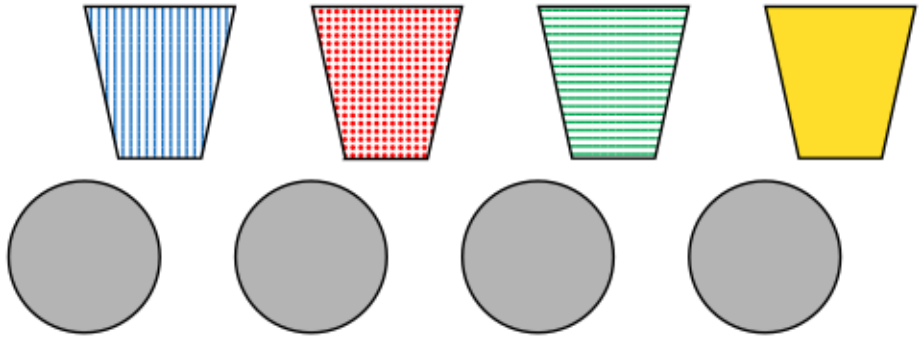


Setup

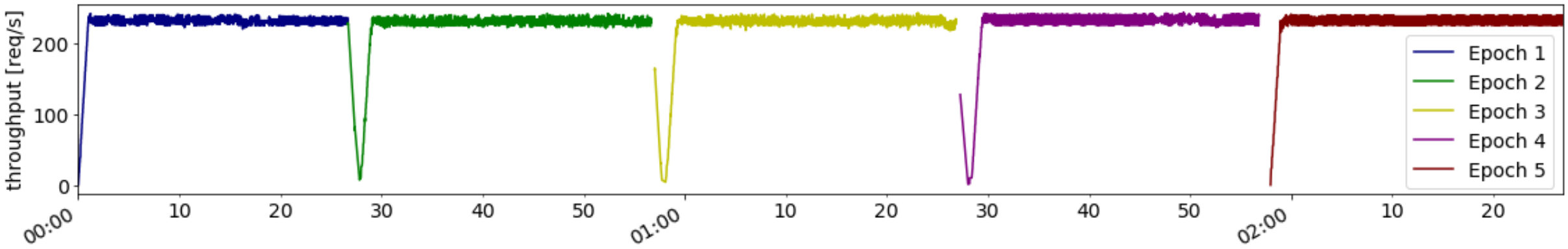


Setup

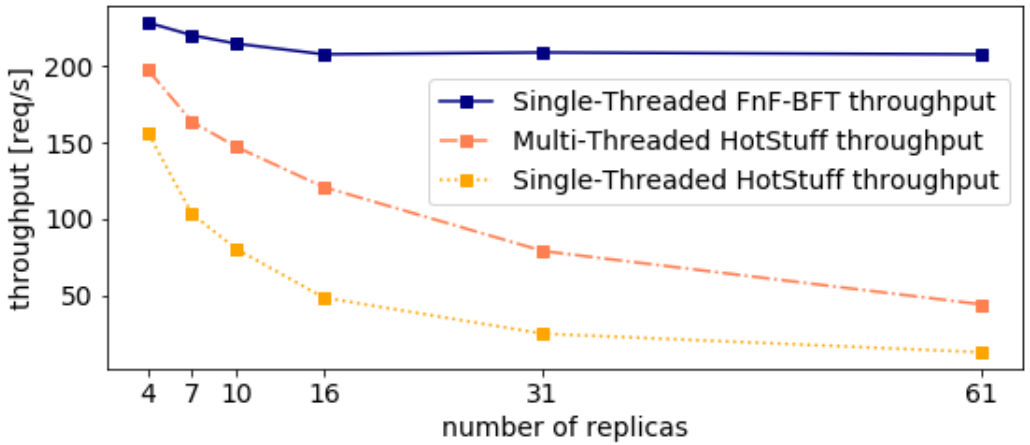




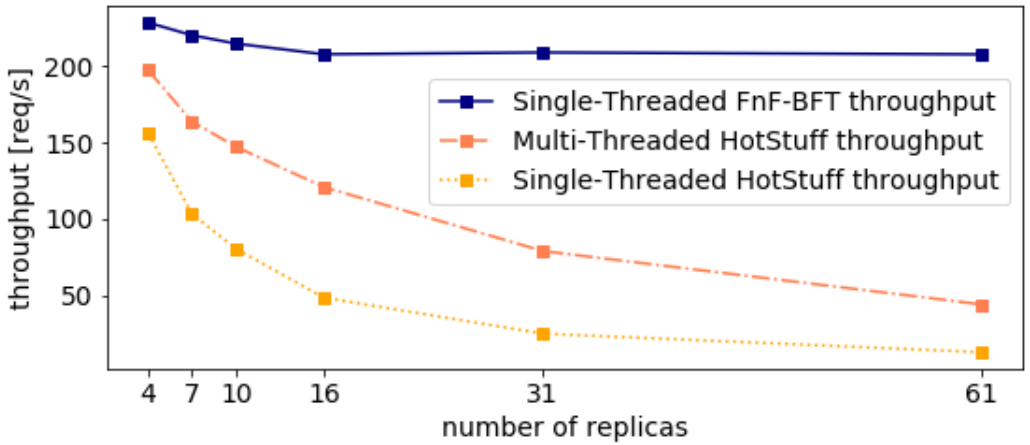
Performance

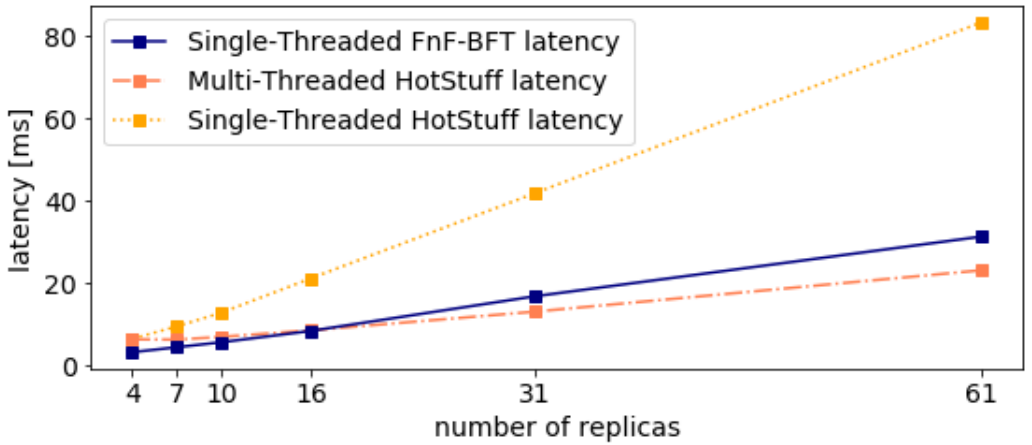


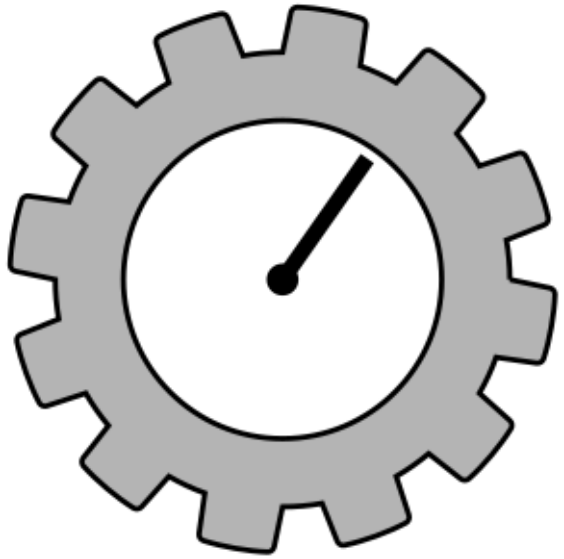
Performance

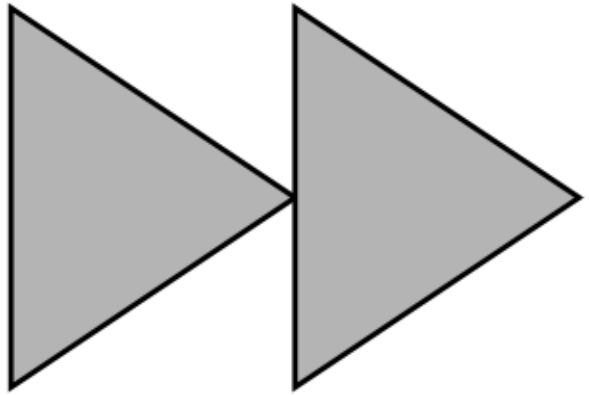


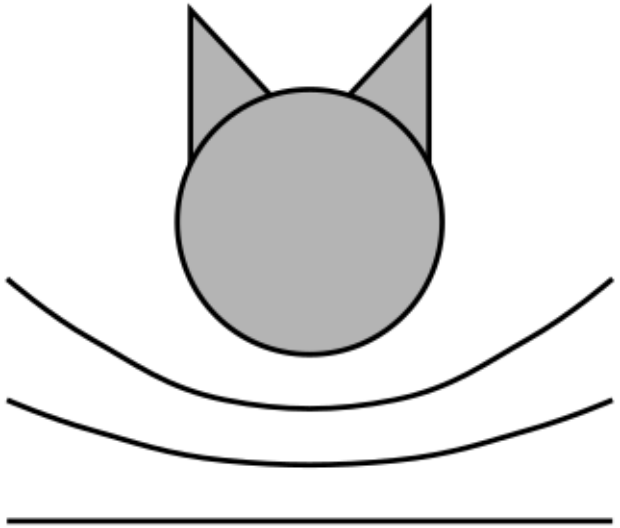
Performance

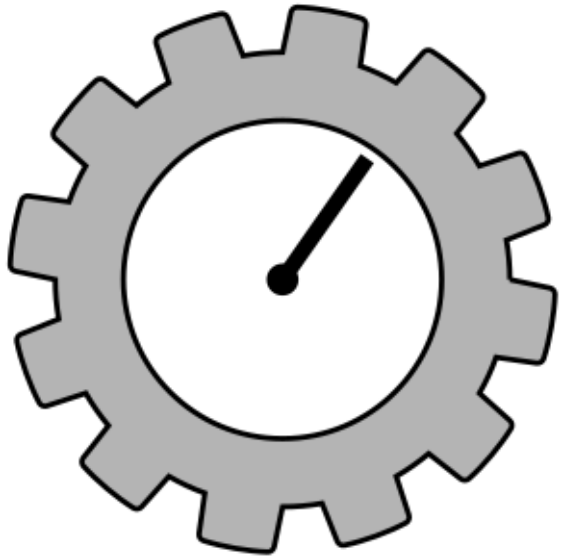


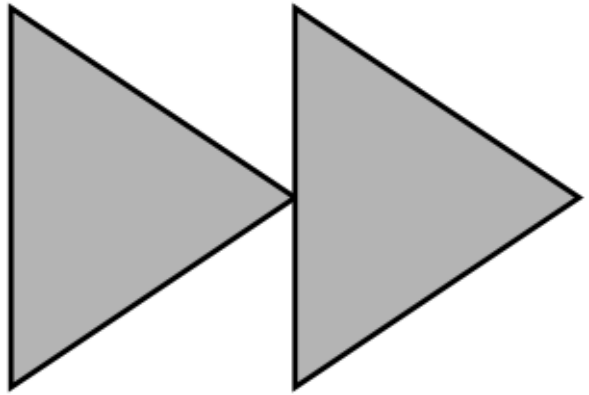


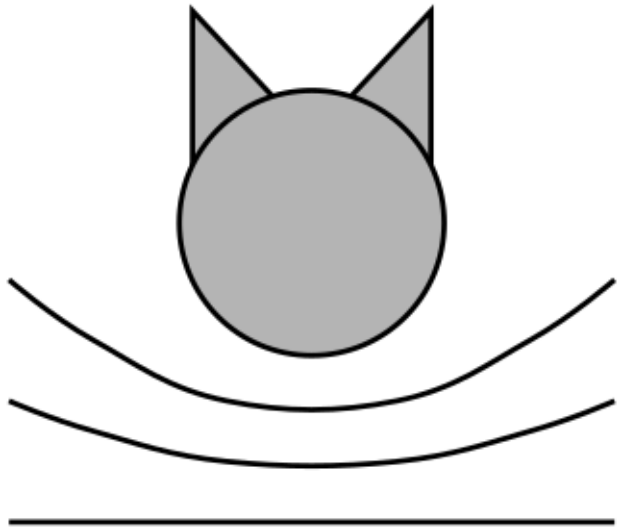




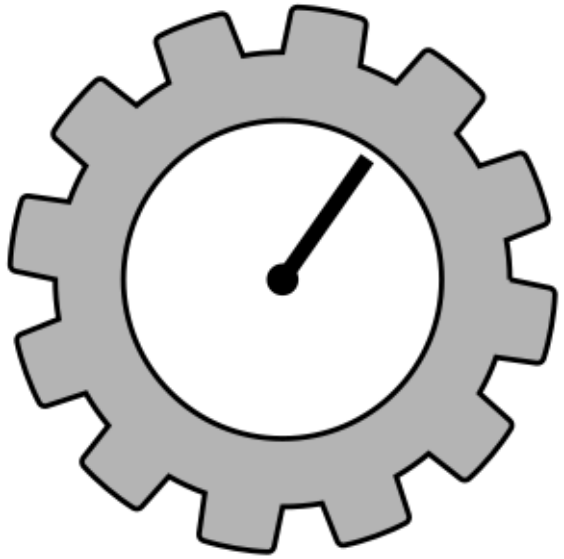




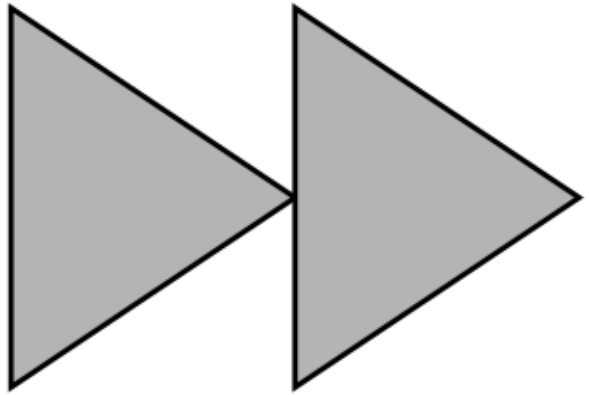




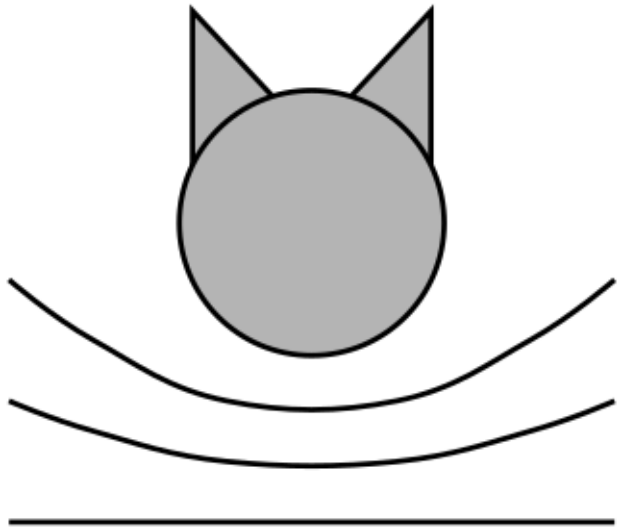
amortized request cost:
$O(n)$

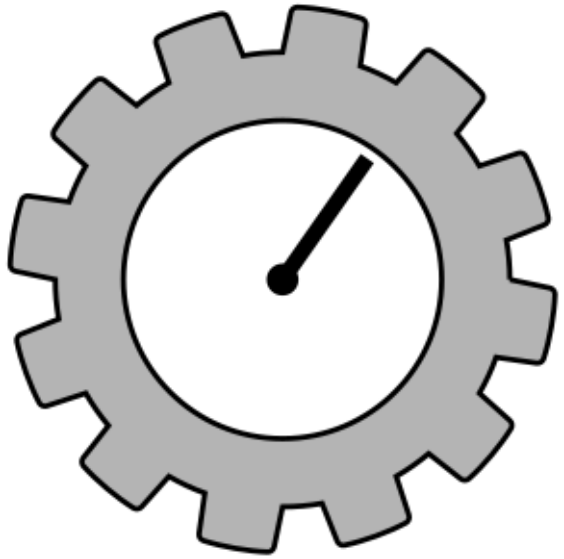


amortized request cost:
$O(n)$

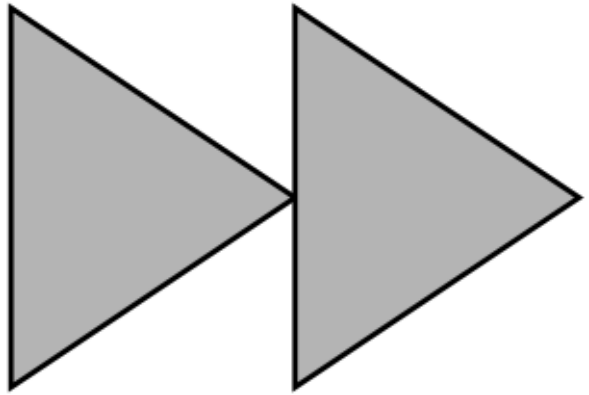


speedup:
$\Theta(n)$

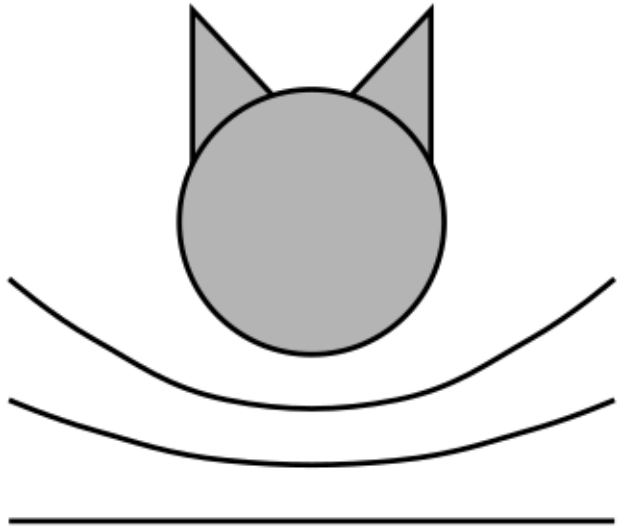




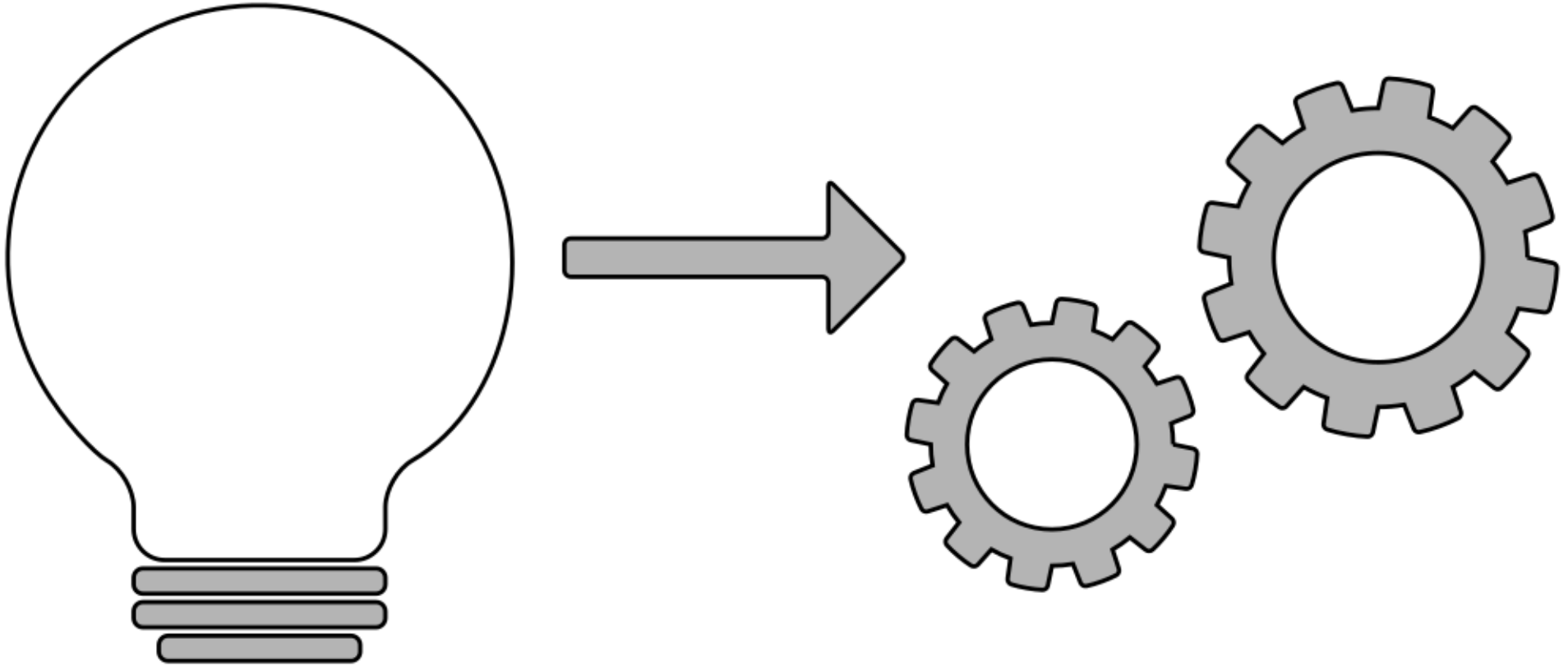
amortized request cost:
$O(n)$

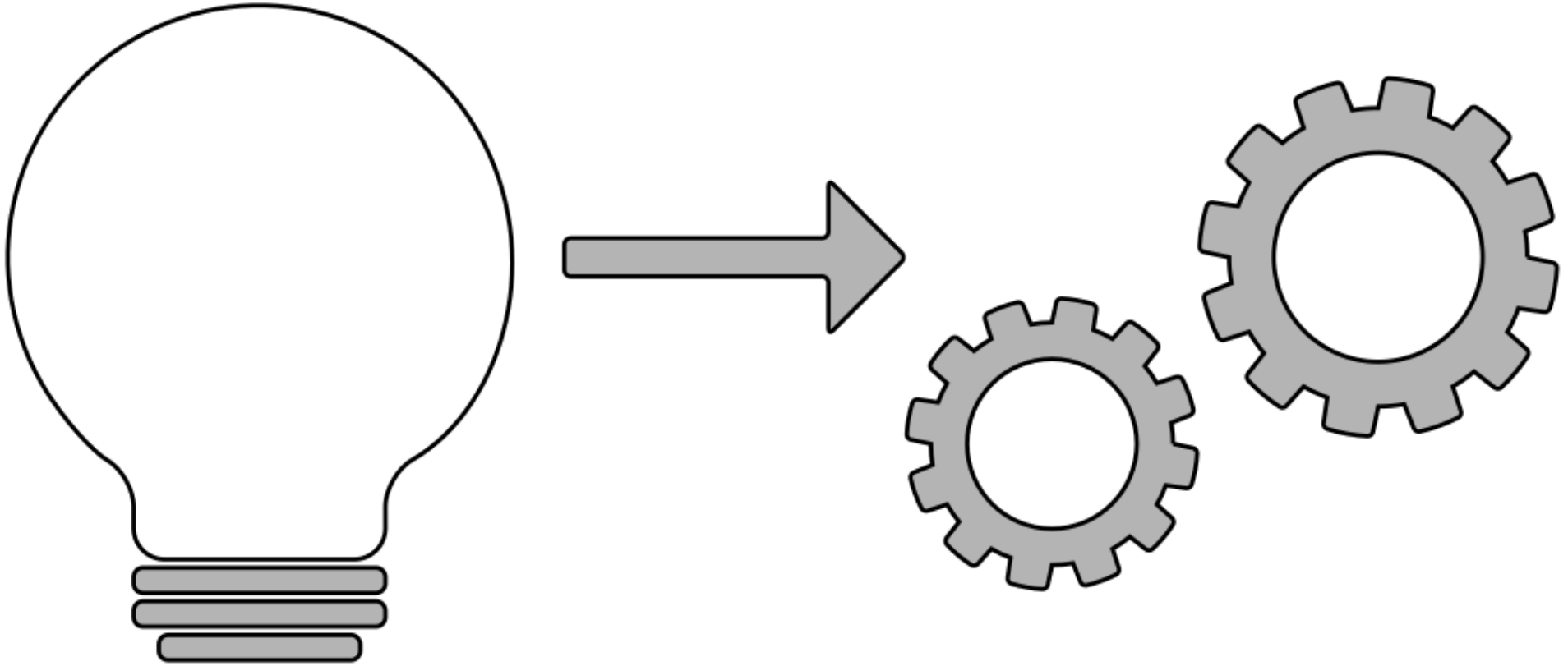


speedup:
$\Theta(n)$



effective utilization:
$\geq \frac{16}{27}$





significantly improved
scaling capabilities

Thank You!

Questions & Comments?

Zeta Avarikioti, Lioba Heimbach, Roland Schmid, Laurent Vanbever, Roger Wattenhofer, Patrick Wintermeyer
ETH Zurich – Distributed Computing – www.disco.ethz.ch

Configuration parameters

Configuration Parameter	Setting
Requests per block	1
Threads per replica	1
Threads per client	4
Epoch timeout	30s
No progress timeout	2s
Blocks per checkpoint (K)	50
Watermark window size ($2 * K$)	100
Initial epoch watermark bounds	10000

Table 1. FnF-BFT configuration parameters used across all experiments.