



KubeCon



CloudNativeCon

North America 2024

Accepting Mortality

Strategies for Ultra-Long Running Stateful Workloads in K8s

Sebastian Beyvers
Maria Hansen



What is Ultra-Long Running?

- Duration varies by context (hours to months)
- Rule of thumb: Exceeds runtime of normal workloads by 10x or more
- Factors beyond time: Available resources & the impact of failures
- Our use-case (metagenomic analyses) take up to 4 months



Choosing the Right Tool

- Pick tools based on **state management**, **resource use**, and **K8s fit** – not just performance
- Prefer robust checkpointing with automatic recovery
- Steady resource consumption without major spikes matters more than pure speed
- A well integration into K8s lifecycle, probes, and monitoring should be preferred

Reasons for Pod Failures

Preemption

API-Initiated
Eviction

Resource Limits

Node Pressure

Hardware Failure

Avoidable Failures: The Human Factor

Unavoidable Failures: The Forces of Nature

Configuring K8s

Pod Priority

High priority class for important pods

Pod Disruption Budget

maxUnavailable: 0

Node Affinity

Dedicated high-mem nodes after eviction

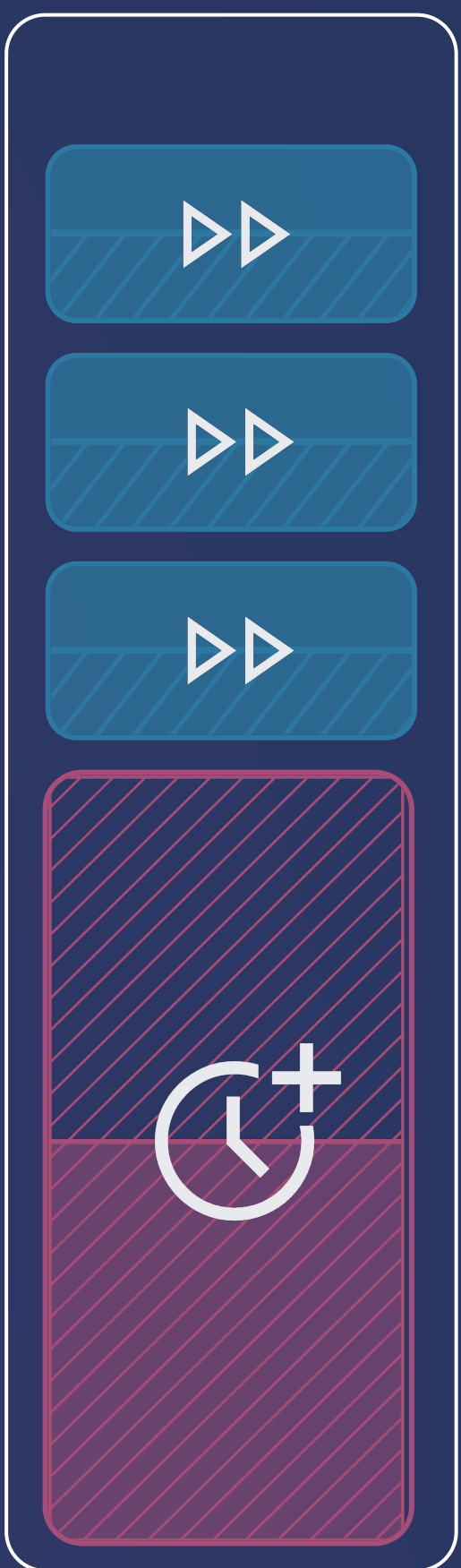
Pod Anti-Affinity

Only one long-running pod per node

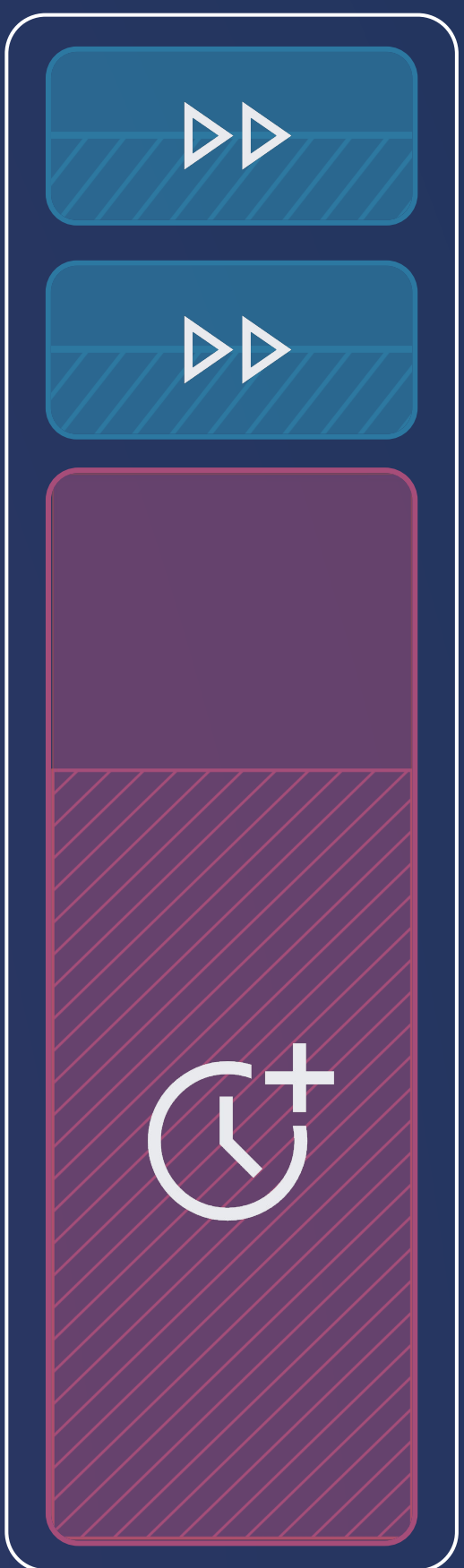
Resources

| | CPU | | RAM | |
|----|--------|----------|--------|----------|
| | Limits | Requests | Limits | Requests |
| ⌚ | - | 1.5x | - | 1.5x |
| ▶▶ | - | 1.0x | 1.2x | 0.8x |

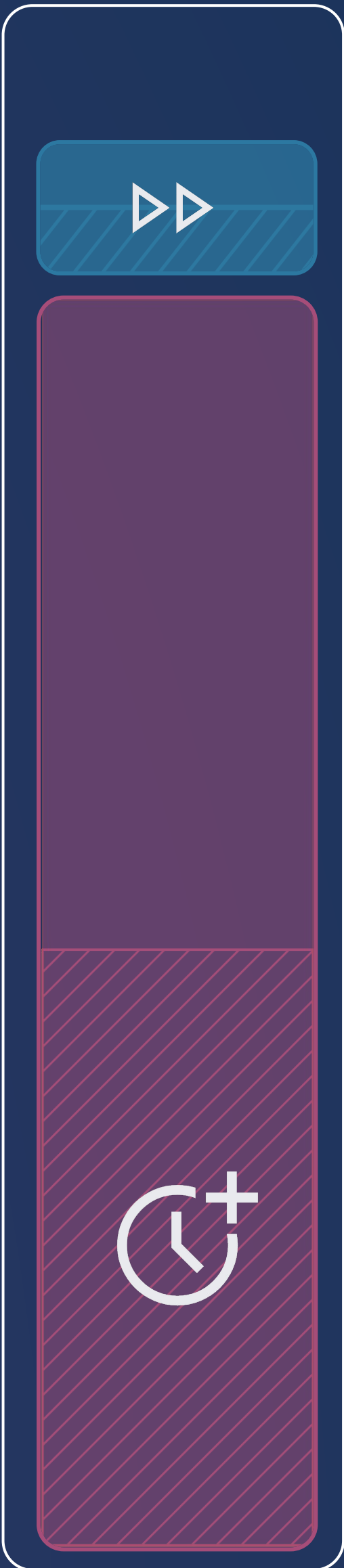
Node placement by memory (requests):
Long-running pods gradually preempt shorter-running pods and automatically migrate to high-memory nodes when regular nodes are exhausted.



Node 1



Node 2



High-mem
Node



Container Necromancy: CRIU to the rescue!

- Create custom container with CRIU support wrapping your workload (crik)
- Checkpoint: Capture process state (memory, FDs) to persistent storage
- Restore: Rebuild FDs and process state via CRIU
- Manual process with limited runtime support in CRI-O / containerd 2.0
- No standard restore API
- Key challenges: PID / TTY restore, running process side-effects
- **Use only as a last resort when native checkpointing is not possible**

▶▶ Regular pods

Request

⌚ Long runners

Memory usage

💡 Our write-up!



Feedback? 😊

