

// SEPT 2023

PENGUIN SOLUTIONS

The Easy Button For Training Compute

PENGUIN[™]
SOLUTIONS

Tesla **without** Penguin AI Factory Partnership



Training compute should soon not be so much of a limiting factor.

Very difficult bringing the 10k H100 cluster online btw. Similar experience to bringing our now 16k A100 cluster online.

Uptime & performance are low at first, then improve with lots of work by Tesla & Nvidia.



Elon Musk @elonmusk
30 August 2023

Meta **with** Penguin AI Factory Partnership



Working in partnership **with our implementation partner, Penguin Computing, we improved our overall cluster management.** By the time we completed the second phase of building RSC, availability stayed above 95 percent on a consistent basis. This was no small feat given that we added a 10K GPU cluster while concurrently running multiple research projects.

We now have a template for building large GPU clusters that is repeatable and reliable.



∞ Meta
18 May 2023

How the Easy Button was built

- 7.5 years ago Penguin wins a RFP for a company by the name of Facebook
- After winning the contract, each subsequent year requests were made for additional functionality
- Phase 1 request was for improved telemetry
- Phase 2 request was for RCA using the additional data collected
- Phase 3 request was for Human out of the Loop remediation
- Phase 4 request was for managing multiple GPU-clusters at different physical locations using the same control plane

Scyld Clusterware: AI Factory Enablement at Scale

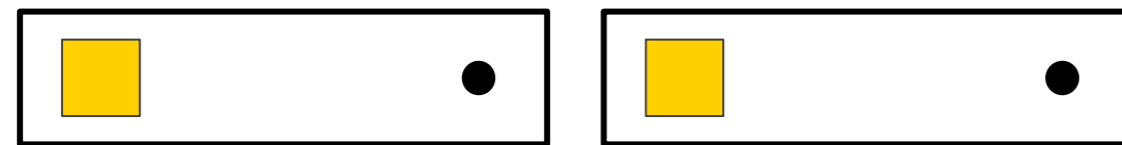
USERS,
SCIENTISTS &
ADMINISTRATORS



Scyld Clusterware SW

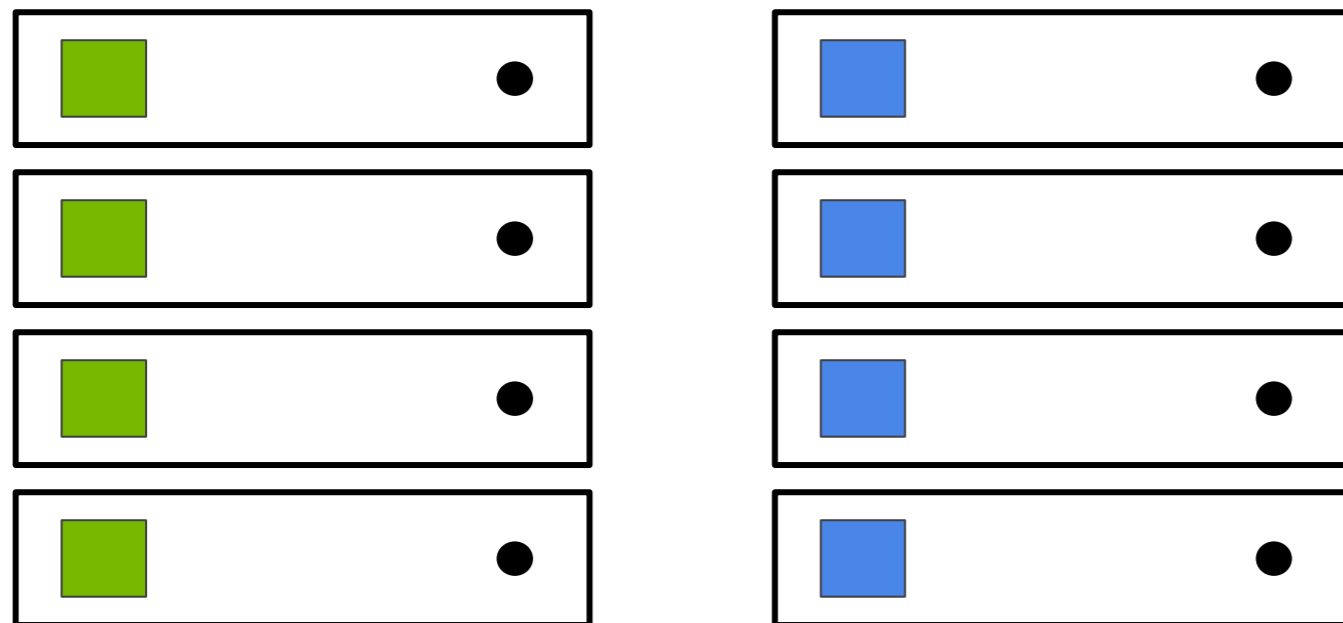
PENGUIN
SOLUTIONS

On-Prem AI Factory



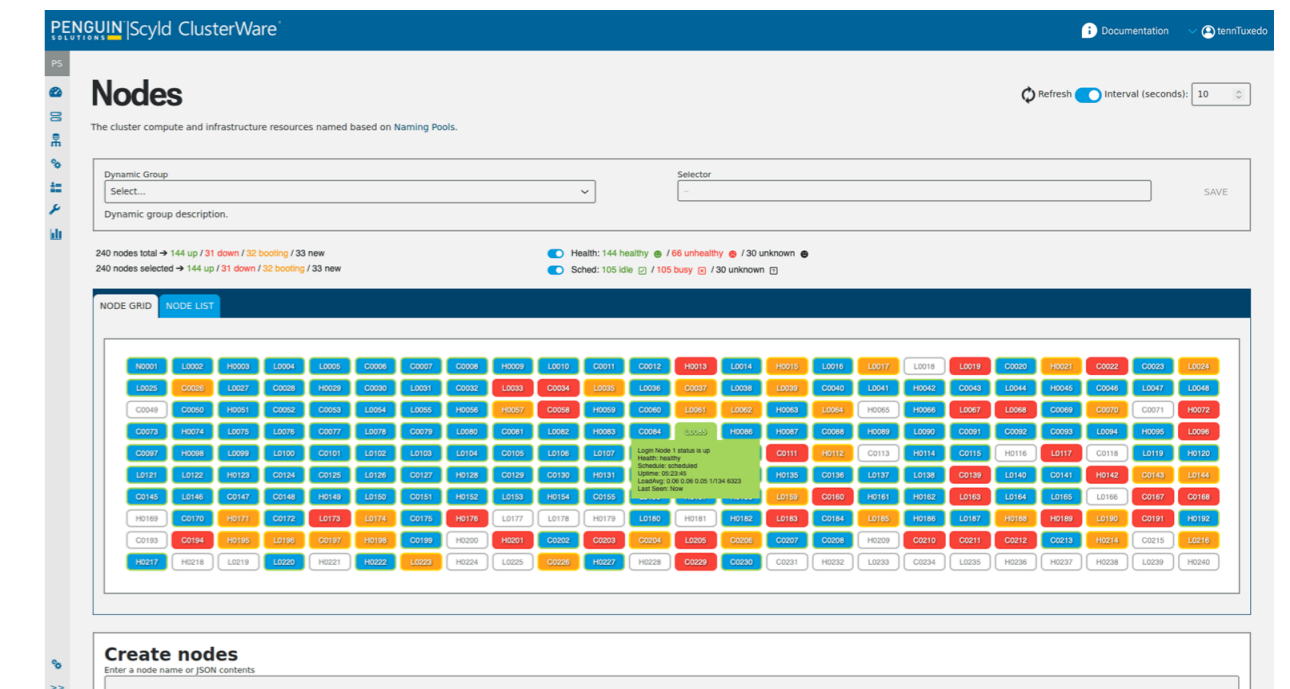
Compute

Storage



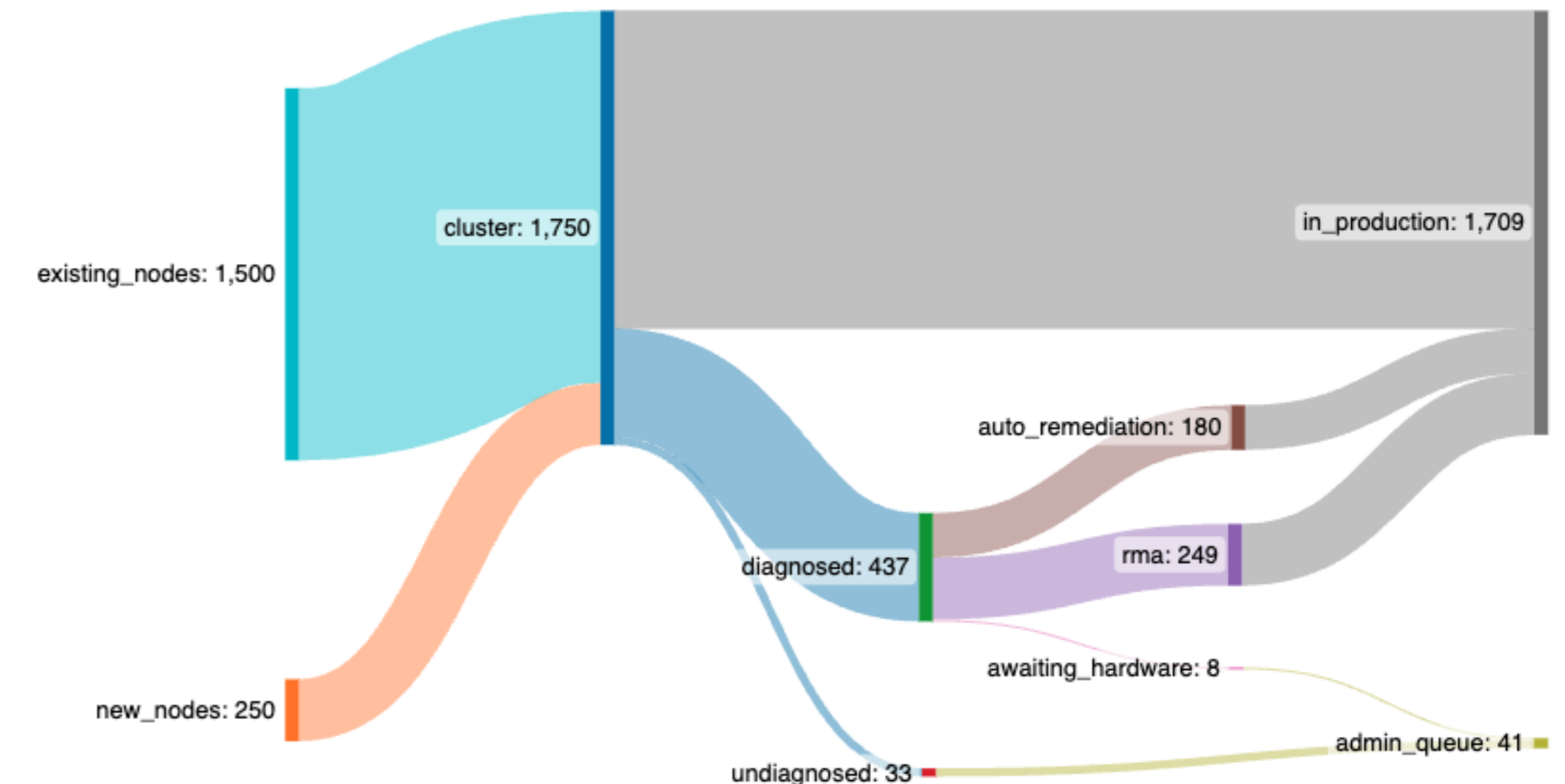
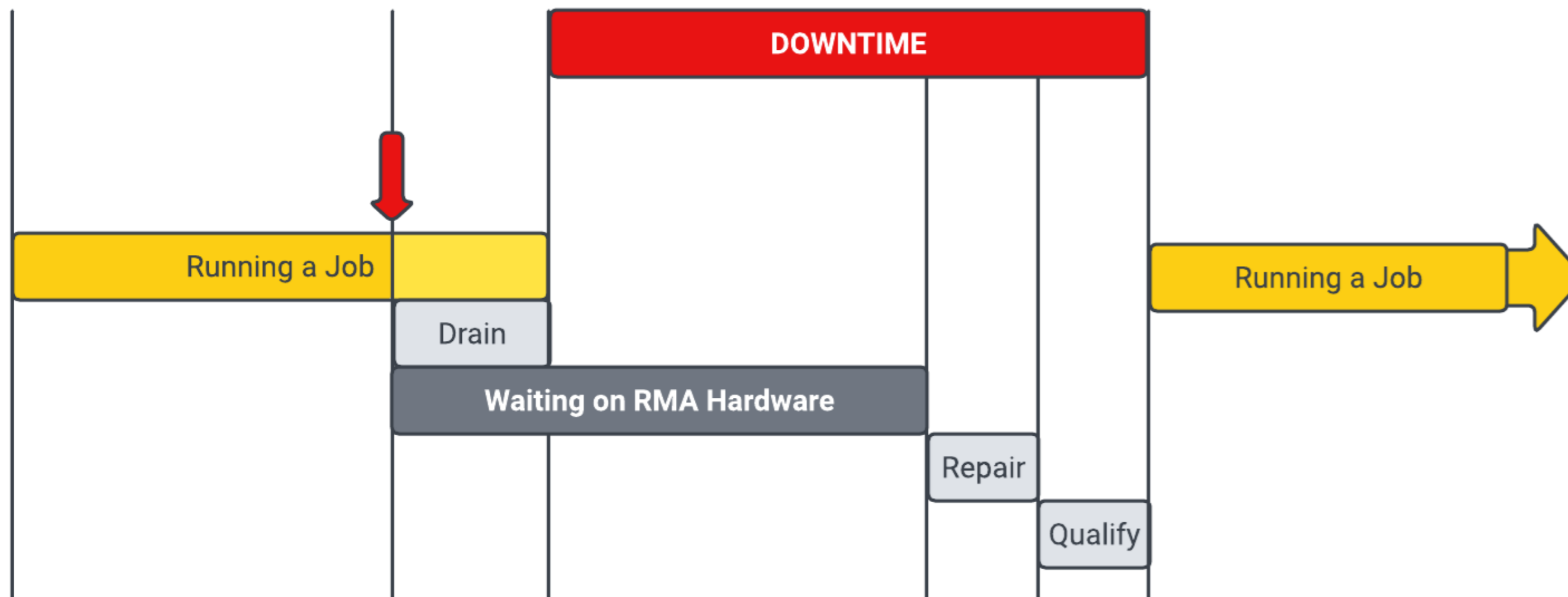
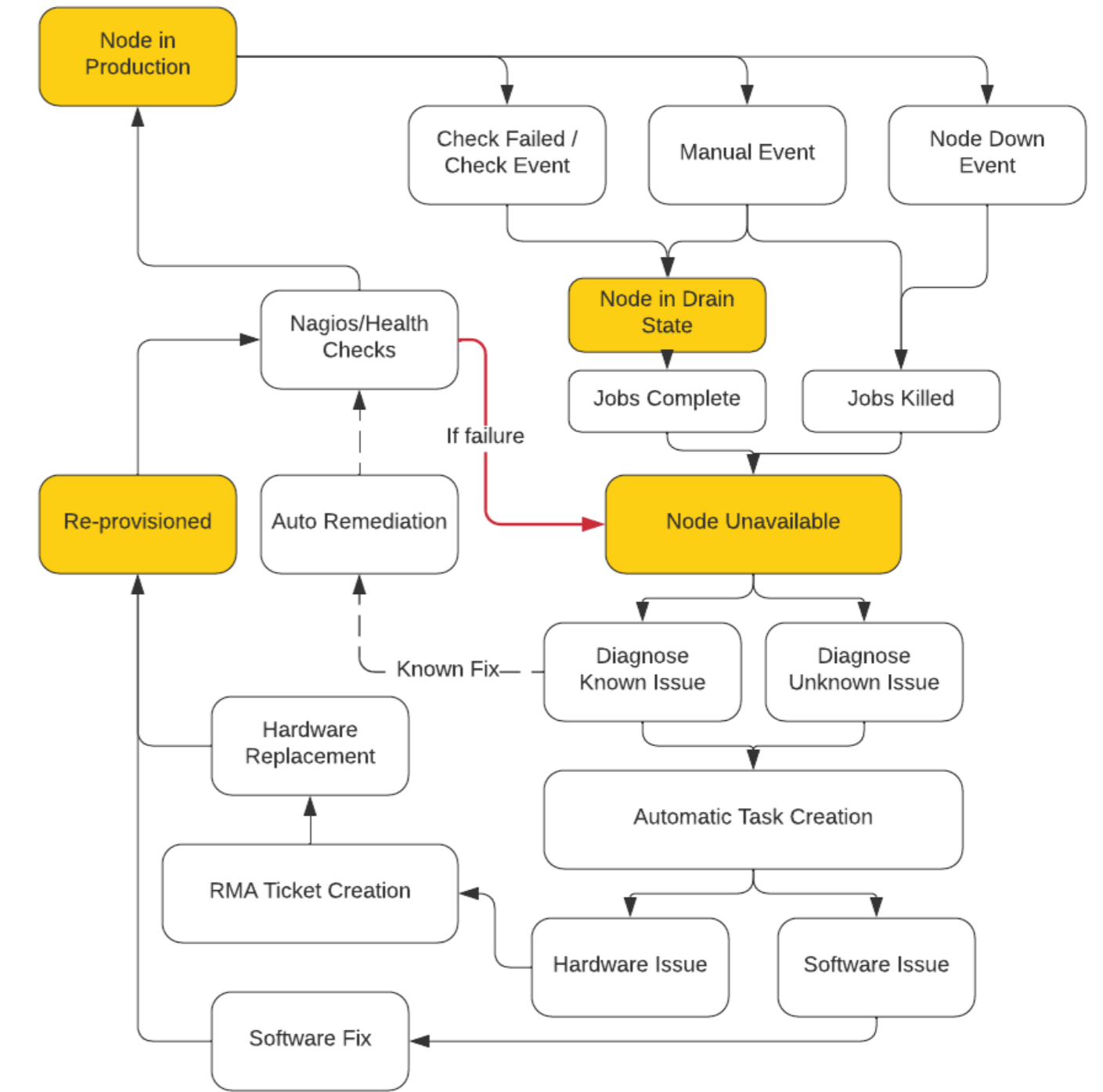
Cluster Management Software that allows Administrators to efficiently manage hardware and software resources to get the most of your AI jobs.

- Tame Complexity
- Monitoring & Alerting
- Rapid Provisioning & Extensibility
- Supported Schedulers (Slurm, Kubernetes, etc.)
- Highly Secure



Automating GPU Infrastructure Management

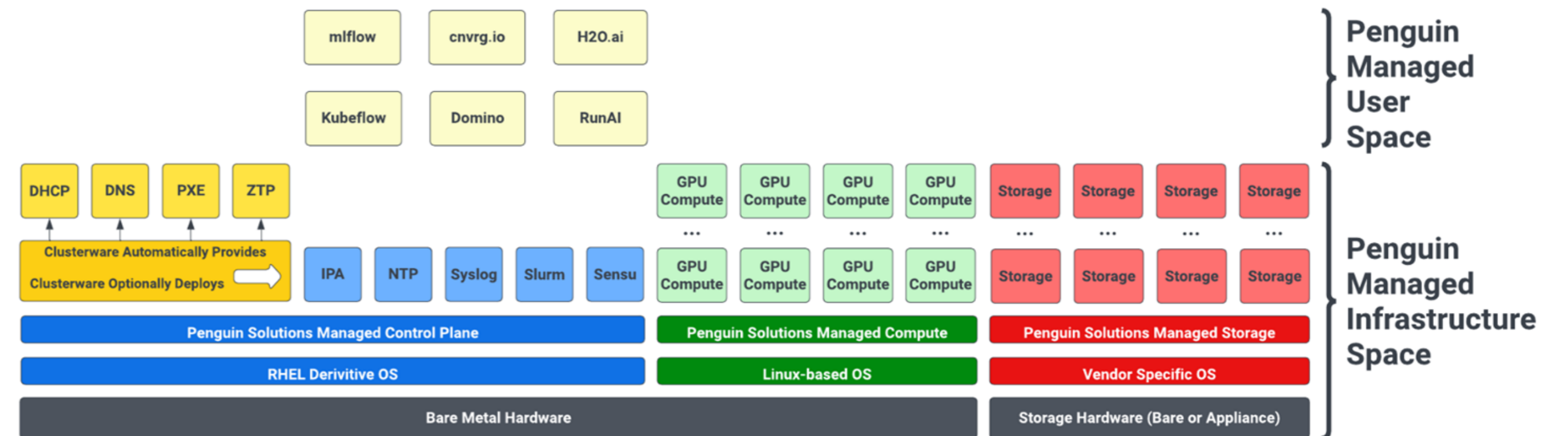
- Human Out-of-the-Loop: Automation wherever feasible
- Automated triage and remediation of known node failures
 - Streamline repair and re-qualification of hardware prior to returning to production.
 - Reduce downtime with early detection and RMA of failed hardware
- Rapid deployment of updates and configuration changes
- Automated detection of wiring and InfiniBand issues
- GPU and network performance monitoring
- Flexible templates to meet customer needs



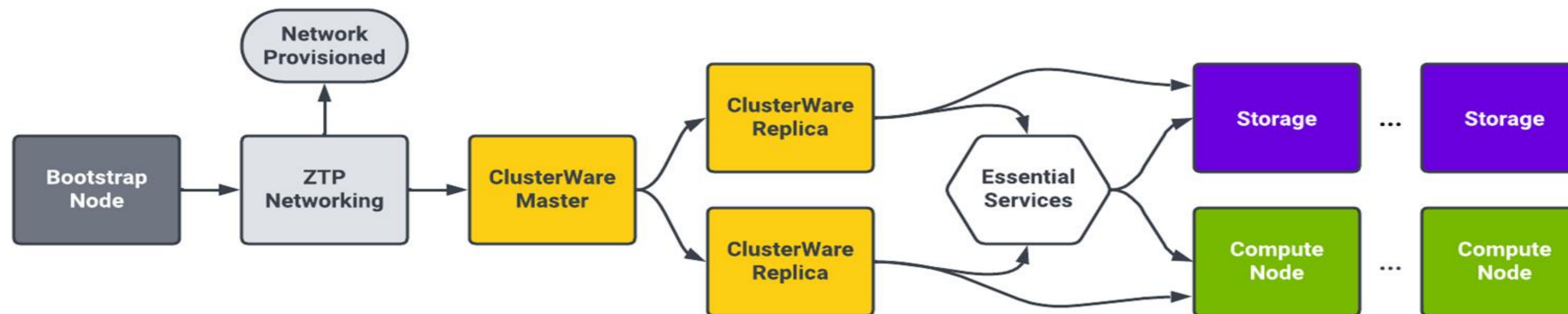
Automated GPU Infrastructure Deployments

“The state of *everything* should be defined in code.”

Automating and managing infrastructure through code *instead of manual processes*.



- Administrators spend time fixing problems instead of laboring over repetitive tasks.
- Automatically bringing nodes back into production after qualifying tests post-repair.



Start with nothing...

...end with a running cluster.

// SEPT 2023

THANK YOU

WE APPRECIATE YOUR TIME

PENGUIN[™]
SOLUTIONS