HPA Optimization & Predictive Elasticity

Subject: Eliminating Vision AI Blindspots in Industrial Safety

1. Executive Summary

In industrial environments like SAIL, Vision AI serves as a crucial safety mechanism. Standard infrastructure scaling struggles to respond to real-time chaos. This whitepaper proposes transitioning to Predictive Horizontal Pod Autoscaling (HPA), using perimeter telemetry to eliminate safety coverage gaps.

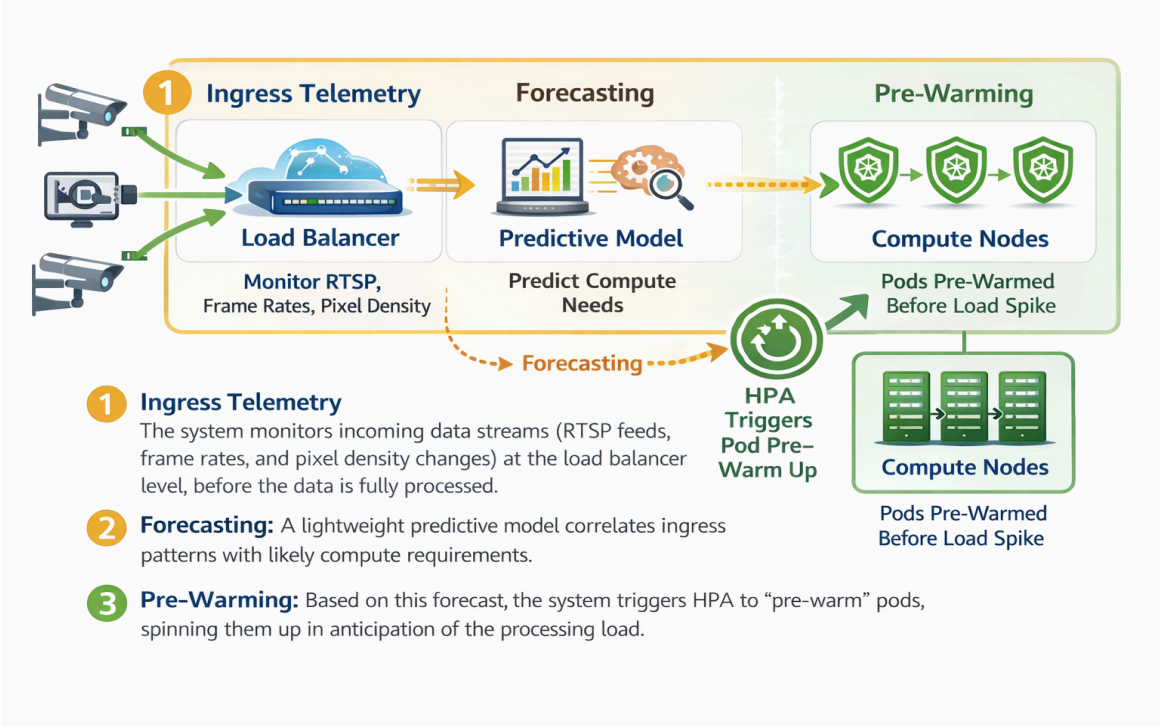
2. The Infrastructure Problem: The "Blackout Blindspot"

Deploying Vision AI in industrial environments faces infrastructure challenges due to extreme volatility. Legacy systems with reactive autoscaling fail due to:

1. **Lag Time**: Standard HPA reacts after CPU/Memory spikes, missing critical events.
2. **Resource Exhaustion**: Sudden spikes cause OOM crashes, leading to system failures when monitoring is most needed. (This results in a "Blackout Blindspot"—a period of unmonitored blindness caused by scaling delays.)

3. Senrysa Solution: Perimeter-Driven Predictive HPA

3.1 Architectural Workflow



4. Performance Impact: Zero-Latency Scaling

By decoupling scaling triggers from current resource usage, we achieve a proactive infrastructure state.

Metric	Legacy Reactive Scaling	Predictive HPA	Impact
Readiness Speed	High Latency (Reactive)	75% Faster (Proactive)	Pods ready before load hits.
Stability	Frequent *OOM Crashes*	Zero OOM Crashes	Resources pre-provisioned.
Coverage	Intermittent "Blindspots"	100% Uptime	Continuous safety monitoring.