

# Evolving FinOps Maturity Inter's Journey Toward Intelligent Kubernetes Automation



## Inter

Founded in the 1980s, Inter has consolidated itself as one of the world's leading financial technology players. Listed on the Nasdaq and a pioneer as Brazil's first 100% free digital bank, Inter currently operates a Super App serving approximately 40 million customers. With a strategic goal to reach 60 million users by 2027, the institution demands high-performance technological infrastructure capable of sustaining exponential growth with governance and technical efficiency.



## THE EVOLUTION TO DYNAMIC RESOURCE MANAGEMENT

The strategic partnership between Inter and RealCloud has been yielding solid results since its first phase in 2024, focusing on the adoption of Spot Optimization practices. The success of this initiative helped Inter advance its FinOps maturity, reaching a 70% utilization rate of Spot instances within its computing capacity. However, the rapid maturation of cloud-native environments brought new challenges as the infrastructure became more dynamic and complex. Maintaining efficiency began to require constant manual attention from engineering teams, who had to "chase" the ideal resource adjustments uninterruptedly.

*"Achieving 70% savings in compute with spot instance automation was a milestone, but true FinOps maturity comes from eliminating manual fine-tuning. Automating rightsizing transformed our efficiency from a reactive task into a native pillar of our cloud governance."*

**Victoria Ribeiro**  
FinOps Head



The scenario revealed a critical gap in infrastructure management: while resource allocation remained static, consumption and demand in the real world were inherently dynamic. At Inter, the rightsizing process — the fine-tuning in the allocation of CPU and memory for each workload — demanded a speed incompatible with non-automated processes. Engineers dedicated precious time adjusting resources to accompany demand fluctuations, and yet, this frequently resulted in waste of capital with idle resources or risks to performance during access peaks. It became evident that, to sustain the scale of a Super App, the best solution would be the total automation of cloud resources in run-time.

## STRATEGY AND TECHNOLOGICAL SYNERGY

As a strategic partner and global FinOps implementer, RealCloud acted as an advisor to design Inter's next leap in maturity, introducing a Kubernetes Pod Rightsizing solution. The solution was anchored in the ScaleOps platform as the technology capable of managing this complexity autonomously. The implementation focused on the impact that Pod sizing has on Node allocation, creating a cycle of exponential efficiency. By automating management in execution time, the solution began identifying and executing fine-tuning for container requests instantly, eliminating the latency of human decisions.

*"The integration of ScaleOps allowed our Kubernetes clusters to perform request adjustments in real-time, without human latency. Now, each microservice consumes exactly what is necessary, ensuring maximum performance with zero waste."*

**Matheus Meneses**  
Principal Platform Engineer



This orchestration guaranteed that the infrastructure reacted in real-time to the release of resources, executing bin-packing and the down-scaling of instances in a precise manner. The transition to the autonomous model allowed Inter to go beyond simple observability, implementing an intelligence that guarantees that each microservice uses exactly what is necessary for its performance. This approach removed the opportunity cost of the engineering team, allowing the focus to be redirected toward innovation, while the infrastructure operated in a self-sufficient and optimized way

## RESULTS AND BUSINESS IMPACT

The rightsizing automation delivered tangible results that reinforce the sustainability of Inter's business model. In productive environments, the partnership achieved an additional reduction of 33% in computing costs.

...generating a projected annual savings exceeding US\$ 1 million. In addition to the financial impact, the technical efficiency was notable, with the optimization of cluster density reducing the total volume of instances, previously in the range of 3,000 to an average of 2,500, operating with greater resource intelligence and without any prejudice to the end-user experience.

*"Reducing 33% of costs and optimizing 500 instances without affecting user experience proves that automation is the path to profitability. This annual savings exceeding US\$ 1 million reaffirms our financial technological vanguard."*

**Carlos Pedrosa**  
IT Director



The success of this stage solidifies RealCloud as a continuous consulting arm for Inter, with a roadmap that contemplates optimizations in other layers beyond Compute and advanced workload positioning strategies. The partnership proves that cloud maturity is not a static destination, but a process of constant innovation. By automating efficiency, Inter ensures that its technology is a catalyst for profitability, maintaining its lead in the global financial sector.



+1 321 370-6878

### Sobre a RealCloud

A RealCloud é líder em soluções de gestão e otimização de nuvem na América Latina, ajudando empresas a alcançar eficiência operacional com redução de custos e aumento de performance. Com clientes como VTEX, iFood e Hotmart, a RealCloud já auxiliou na economia de mais de US\$ 440 milhões em gastos com nuvem. Saiba mais em: [realcloud.ai](https://realcloud.ai)

