

GOALS

Leverage CPU load forecasts for Kubernetes predictive autoscaling

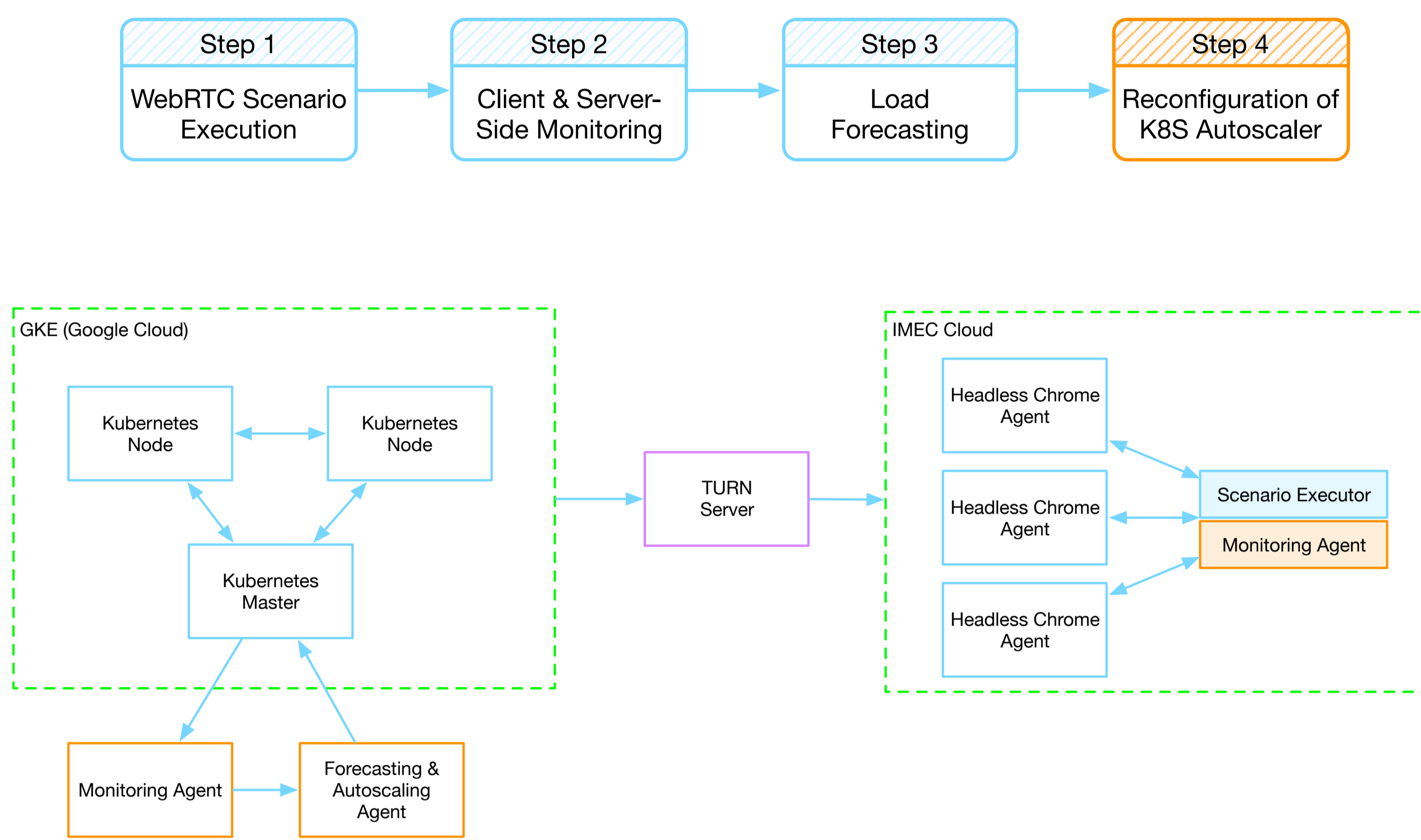
Comparison of Kubernetes stock and predictive autoscaling policies with various WebRTC load patterns

CHALLENGES

Adapt VM-based predictive autoscaling to Kubernetes

Experiment with different forecasting methods and compare their performance against the performance of Kubernetes stock autoscaler

DEMO SETUP



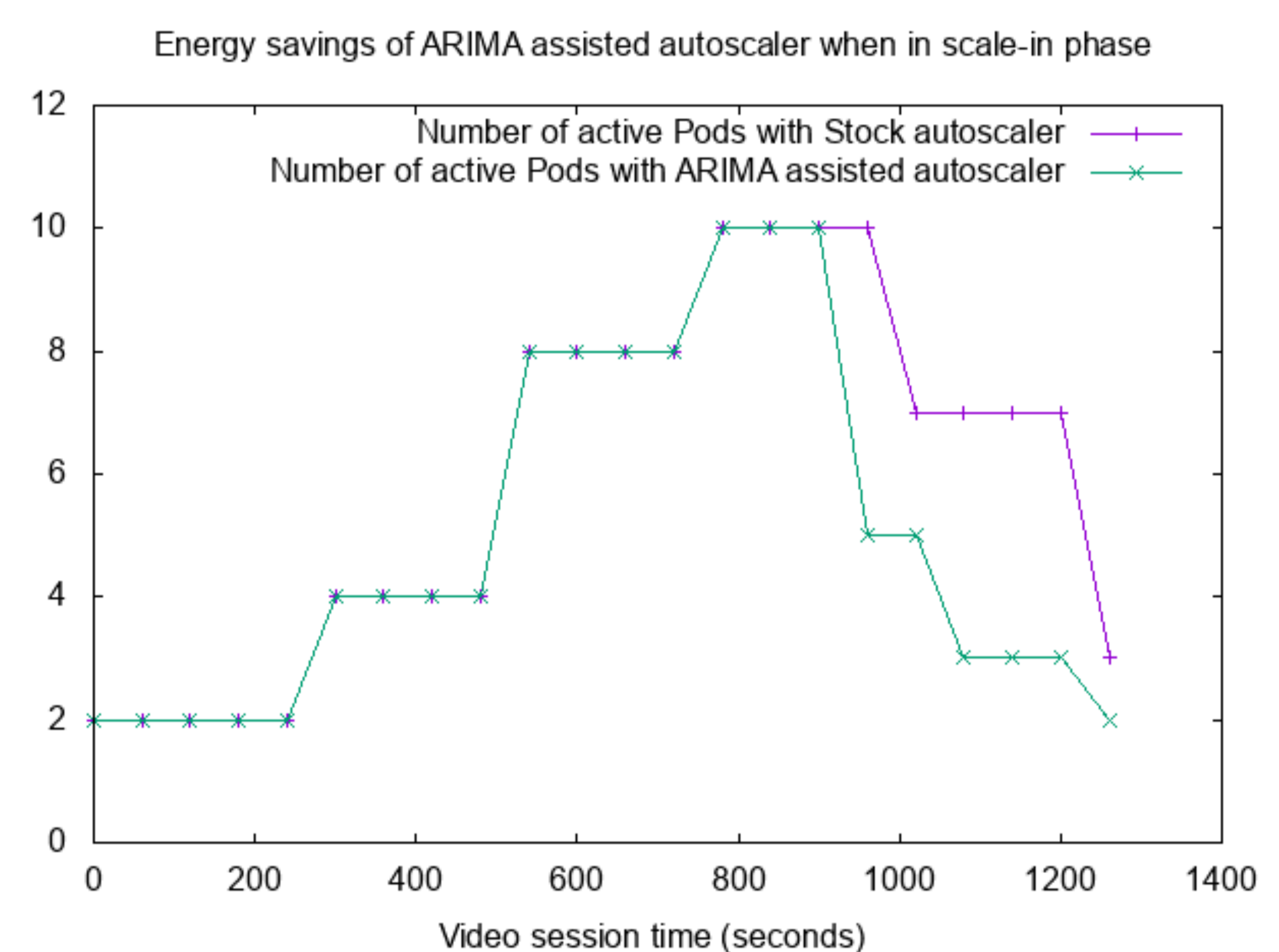
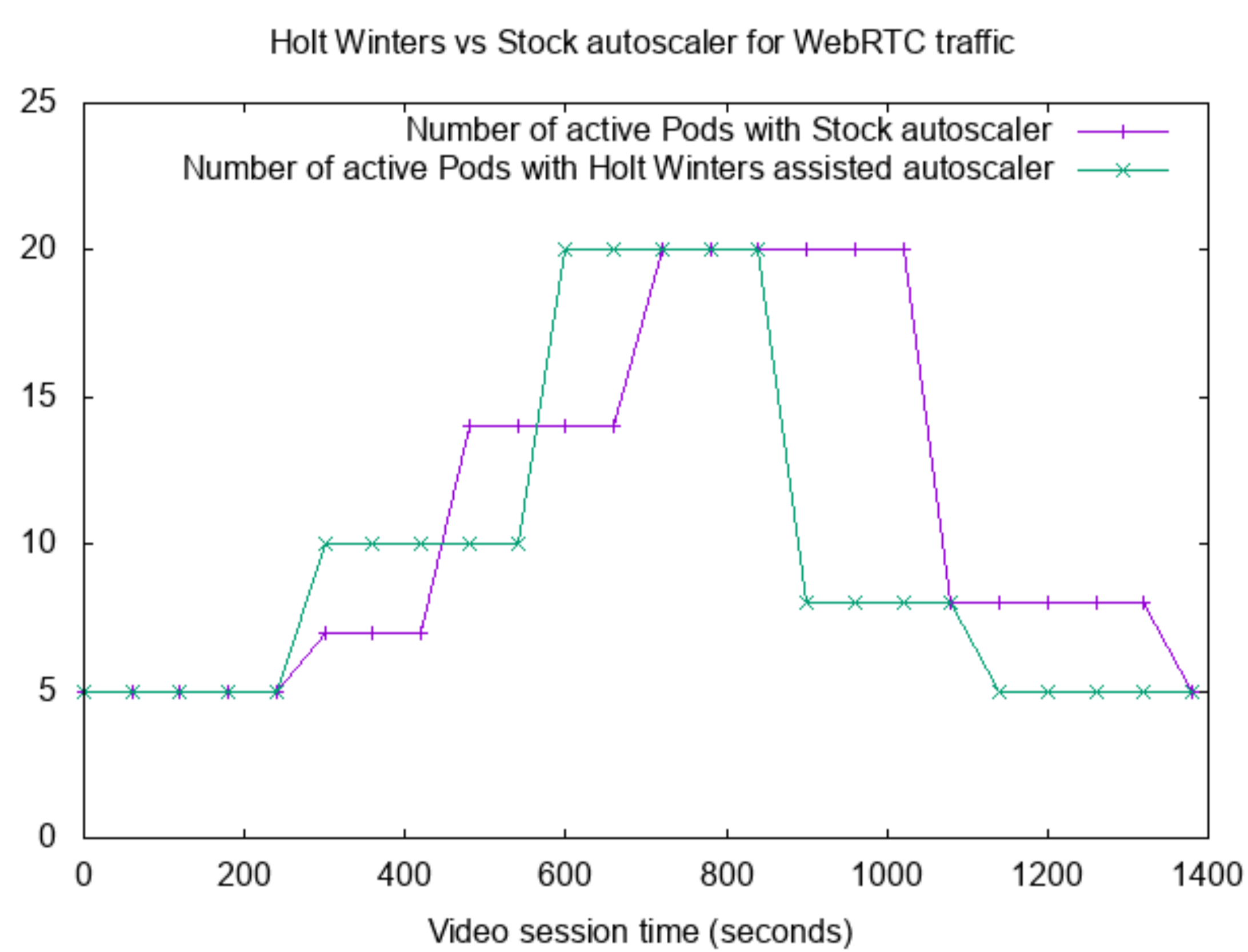
RESULTS

20-30% energy saving through proactive shutdown of idle Pods during the *scale-in* phase

5-15% QoS improvement in terms of 'good sessions' during the *scale-out* phase

Holt Winters has performed better than i) ARIMA and ii) Recurrent Neural Networks in the majority of experiments

MORE RESULTS



CONCLUSIONS

Predictive models are effective in improving the performance of Kubernetes stock autoscaler

This project provided Modio with significant knowledge which will be capitalized as a new feature of our commercial product, Qiqbus, https://modio.io/qiqbus_modio

POST MORTEM

Thanks to the software tools and hardware provided by Fed4FIRE+ we were able to implement and validate an autoscaling agent applicable to Kubernetes for intelligent autoscaling in NFV 5G environments