

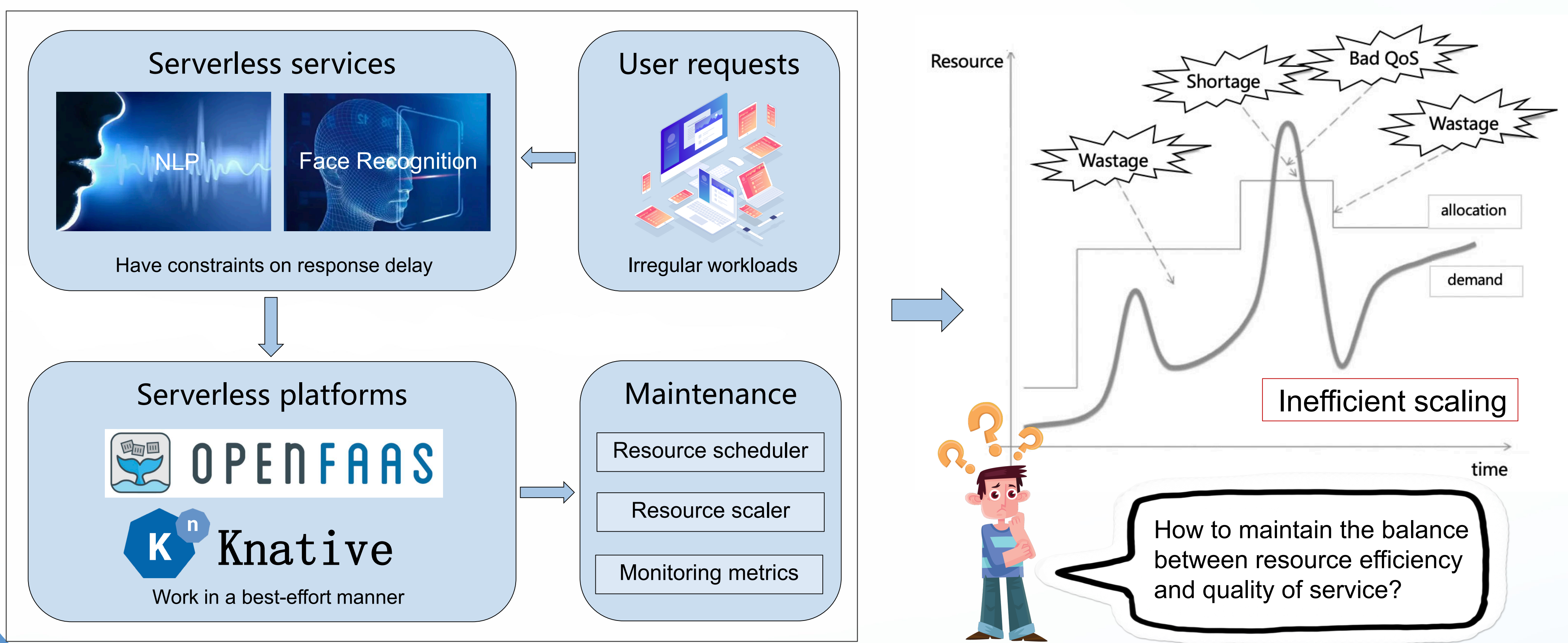
Adaptive Auto-Scaling of Delay-Sensitive Serverless Services with Reinforcement Learning

Zhiyu Zhang, Hao Zeng, Tao Wang, Wenbo Zhang

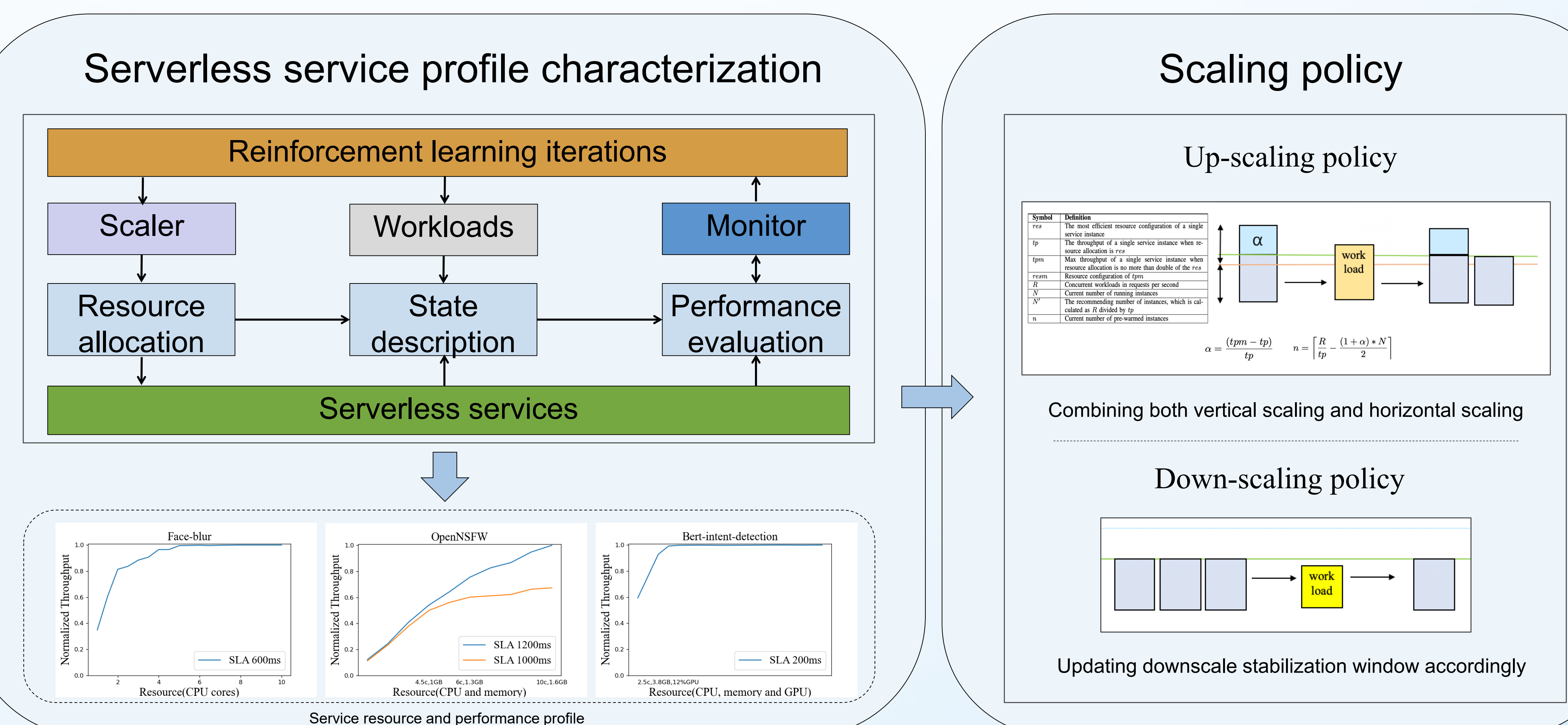
In 2022 46th Annual IEEE International Computer Software and Applications Conference, COMPSAC 2022.

Contact: Hao Zeng, zenghao20@otcaix.iscas.ac.cn

Auto-scaling for newly deployed serverless services is challenging



Characterizing serverless service profile to guide the scaling policy



Effectively ensure quality of service while achieving a reduction of cost

