

# Progetto per Assegno di Ricerca CIRI ICT

## Digital Twin per applicazioni a microservizi in ambienti Kubernetes

Kubernetes si è imposta come una delle piattaforme di riferimento per l'orchestrazione di container, aiutando a gestire e scalare applicazioni distribuite in modo efficiente e automatizzato. L'uso di Kubernetes consente alle aziende di ridurre i costi e aumentare l'efficienza delle proprie applicazioni distribuite, fornendo un'infrastruttura automatizzata e scalabile. In particolar modo, questa piattaforma ben si sposa ad essere utilizzata per supportare applicazioni in ambienti edge. L'*edge computing* rappresenta una delle aree di interesse più promettenti sia per il mondo accademico che per l'industria, aprendo tutta una nuova serie di sfide e opportunità di business. Anche gli enti di standardizzazione, come ETSI, 3GPP e GSMA, stanno seguendo questa tendenza, cercando di definire l'architettura dei sistemi *edge* e le opportune soluzioni per farli funzionare. Ad esempio, il 3GPP nella release 17 ha avviato uno studio per esplorare come le applicazioni ospitate in un ambiente di *edge computing* possano essere utilizzate dai dispositivi di utente in una rete 5G. ETSI ha avviato l'*Industry Standardization Group* per il *Multi-access Edge Computing* (MEC) in cui viene definita un'architettura per un sistema MEC - avente similitudini con il lavoro fatto in NFV-MANO – e un insieme di interfacce standardizzate per regolare le interazioni tra i suoi componenti.

L'applicazione di questo paradigma in ambienti 5G e, in futuro, 6G necessita di stringenti controlli della qualità di servizio offerta dalle applicazioni containerizzate. Per questo motivo modelli che sfruttano gemelli digitali (Digital Twin) delle applicazioni possono essere utilizzati per realizzare ambienti di simulazione il più aderenti possibile con l'ambiente reale. Questi ambienti possono quindi essere utilizzati per predire il comportamento del sistema reale in diverse condizioni di carico oppure per effettuare il training offline di modelli di intelligenza artificiale, ad esempio agendo come ambiente di training per algoritmi di Reinforcement Learning.

Questo progetto di ricerca intende indagare algoritmi di intelligenza artificiale per la descrizione statistica del tempo di risposta di microservizi in ambienti Kubernetes. In primo luogo, si ipotizza di studiare algoritmi ad apprendimento offline, quindi basati su serie storiche preacquisite dal sistema. In seguito, l'algoritmo verrà esteso per supportare un apprendimento online, tramite l'acquisizione in tempo reale di metriche provenienti dal sistema grazie all'utilizzo del software Istio per la realizzazione di Service Mesh in ambienti Kubernetes.

Piano di attività:

- studio della letteratura sugli algoritmi di intelligenza artificiale per la descrizione statistica del tempo di risposta di microservizi;
- studio dello stato dell'arte sull'adozione di tecniche di intelligenza artificiale per l'orchestrazione di microservizi, con particolare attenzione ai deployment Kubernetes;
- studio ed applicazione sperimentale (emulativa/simulativa) di algoritmi ad apprendimento offline, basati su dati preacquisiti;
- studio ed applicazione di tecniche di apprendimento online, basati su dati acquisiti in tempo reale, mediante impiego ed estensione di software preesistente per la realizzazione di diversi scenari di deployment in ambito Kubernetes.