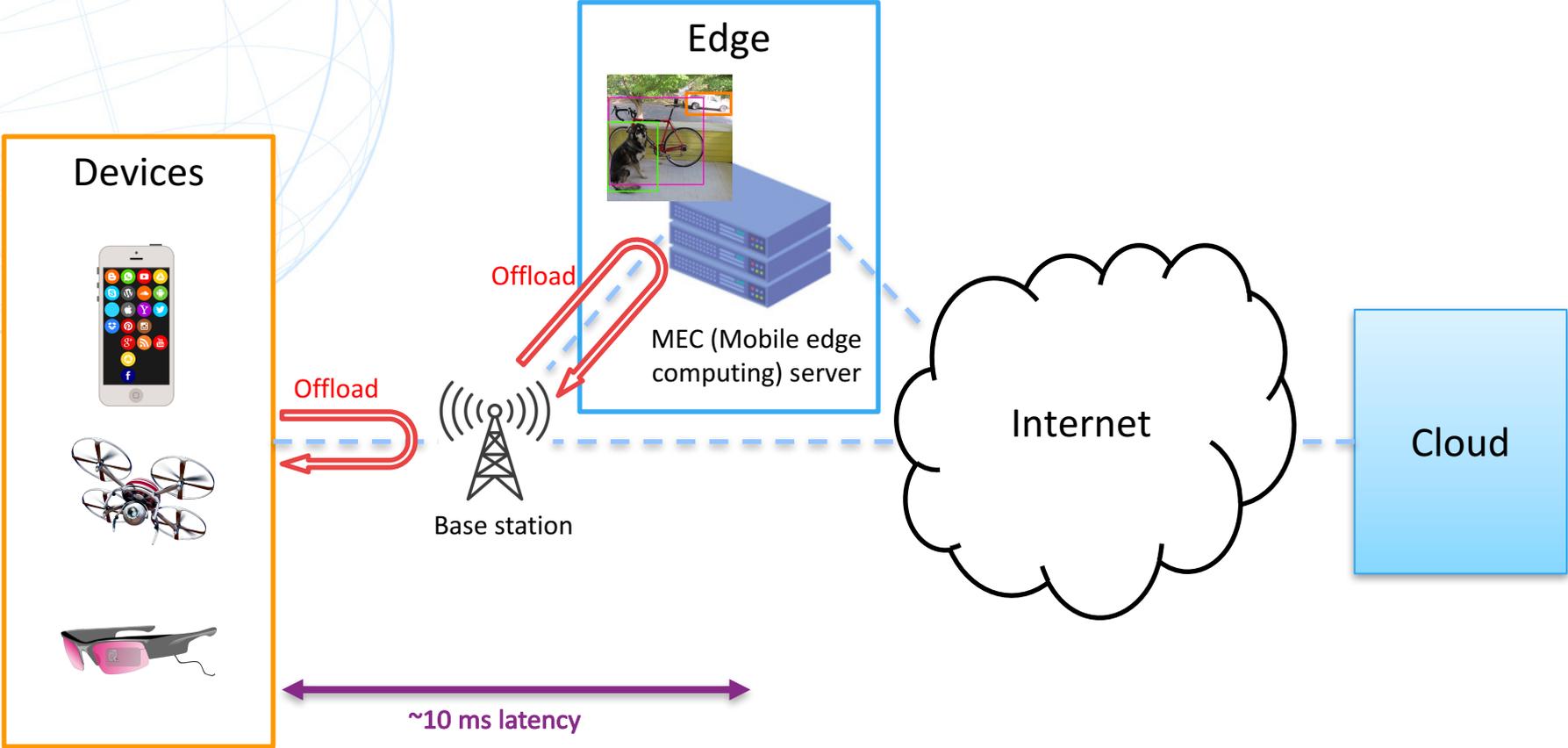


TVM for edge computing platform

NTT Software Innovation Center

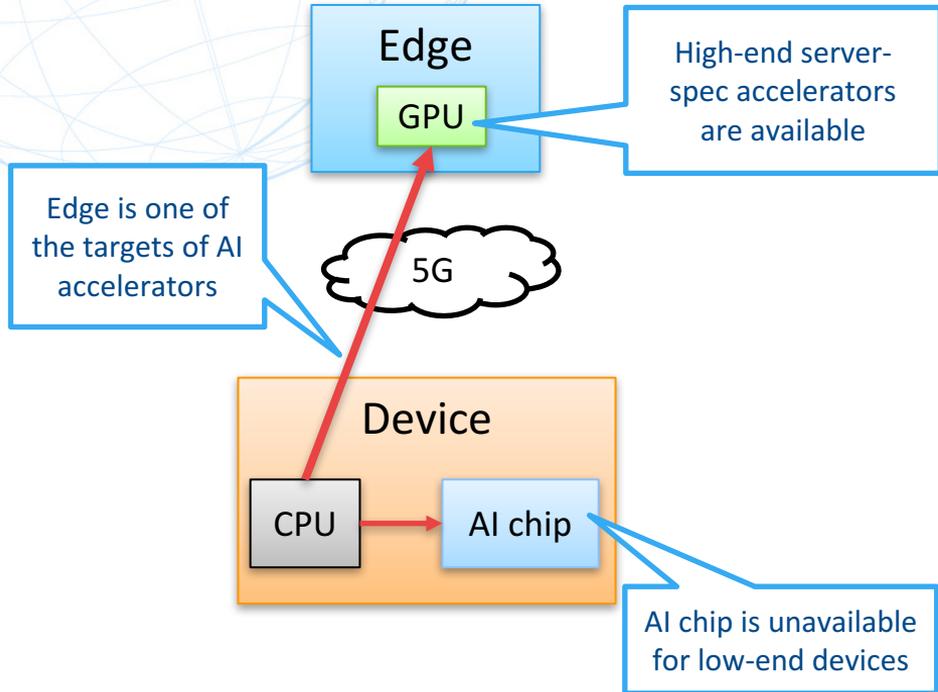
Kazutaka Morita

Inference in 5G era

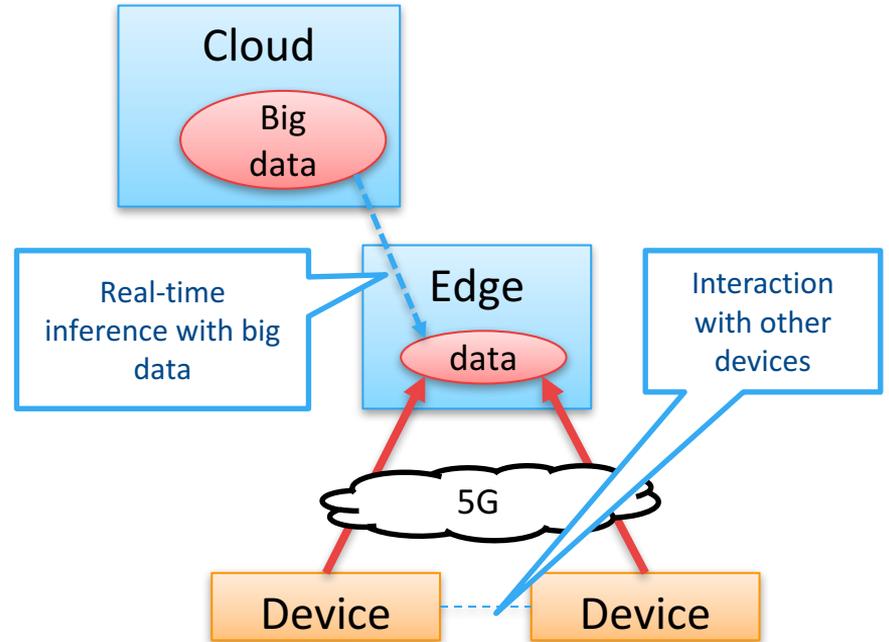


Benefits of offloading inference

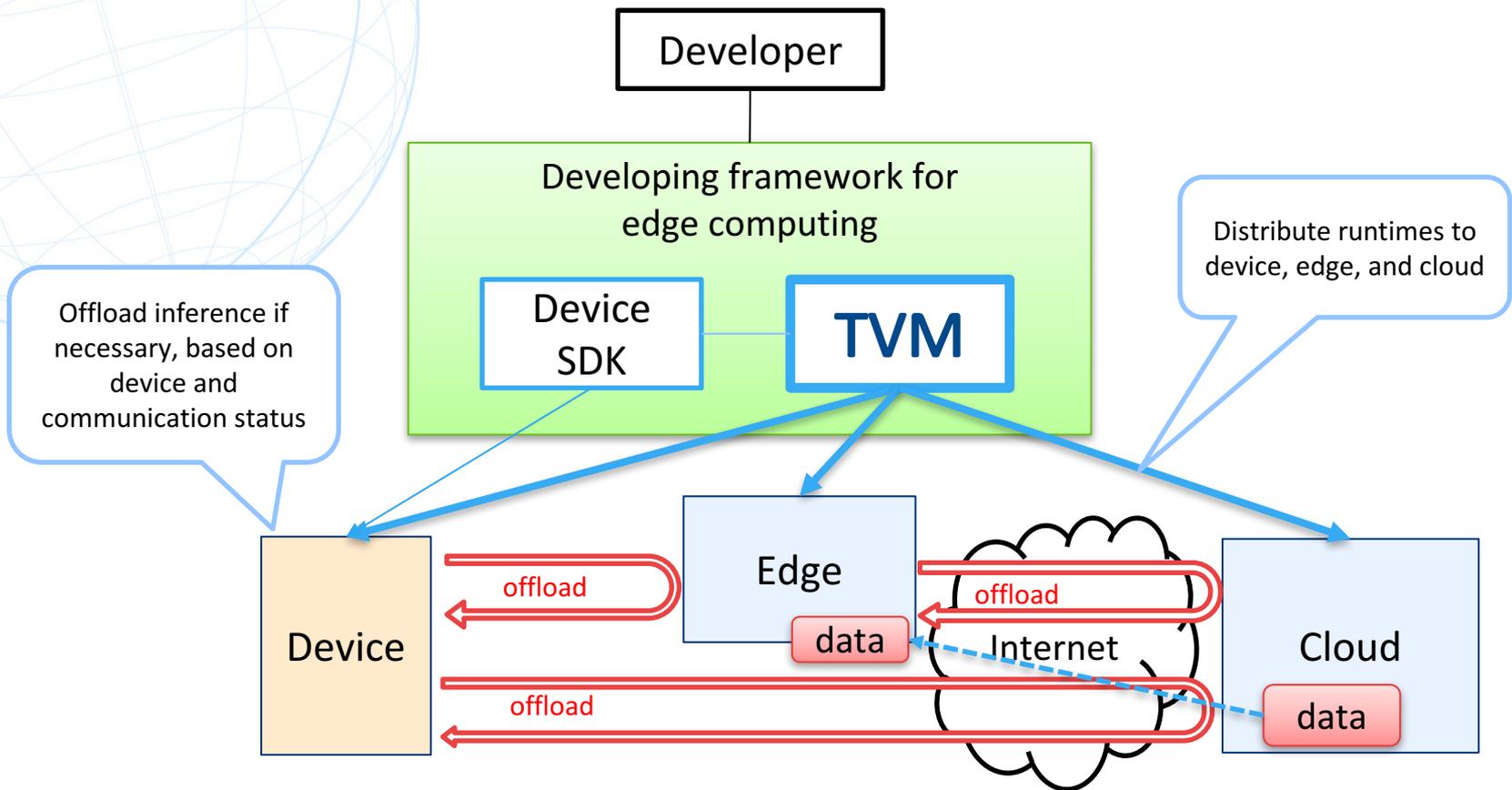
Computing resource



Inference with data

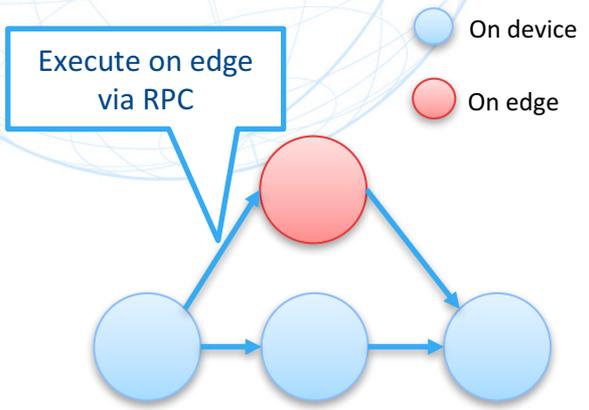


Edge computing platform with TVM



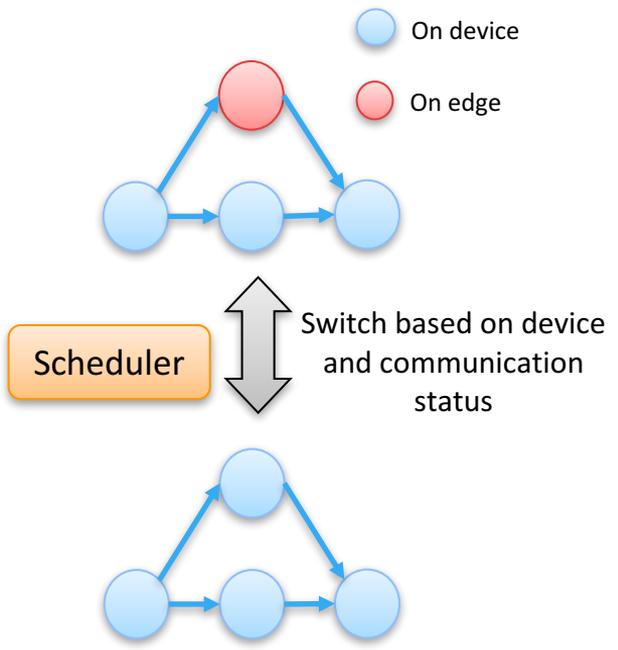
What are required for TVM?

Heterogeneous runtime with offloading support

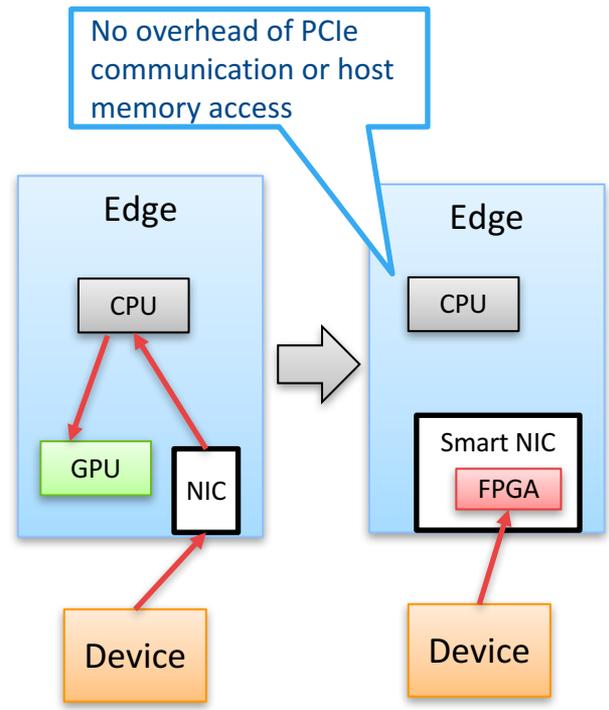


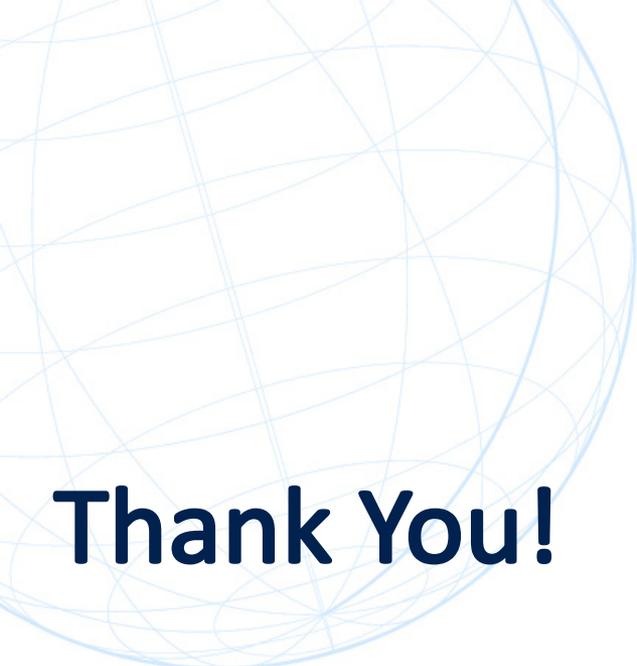
Auto tuning support would be also nice

Dynamic runtime



Smart NIC support





Thank You!

Email: kazutaka.morita.fp@hco.ntt.co.jp