

Real-time queues are those in which customers have deadlines. Each arriving customer has a lead time, the time until the customer's deadline elapses, and the lead time runs down at unit rate per unit time. Each customer brings a certain amount of work for the server of the queue, and one can model the lead time/work profile of the queue by a measure on the real line that assigns mass equal to work at the position of the lead time. As the queue evolves, this so-called lead-time profile is a measure-valued process on \mathbb{R} . As the queue enters heavy traffic, the limiting lead-time profile measure-valued process can be determined as a function of the limiting workload process for the queue.

In this talk, we discuss real-time queues in which customers renege (leave the queue) when their deadlines elapse. We show in this case that the limiting workload process is a doubly reflected Brownian motion and we characterize the limiting lead-time profile process as a function of the limiting workload process. The analysis relies on a newly discovered formula for the operator that maps paths taking values in \mathbb{R} into doubly-reflected paths taking values in $[0, a]$.

This is joint work with Lukasz Kruk, John Lehoczky, and Kavita Ramanan.