

Hull: Less is more

Ссылка: <https://people.csail.mit.edu/alizadeh/papers/hull-nsdi12.pdf>

В данной статье поднимают вопрос создания системы в которой появляется tradeoff небольшого кол-ва Bandwidth на чрезвычайно низкий Latency.

Авторы утверждают, что большая часть систем ориентируются на уменьшение Bandwidth. Однако для некоторых задач может потребоваться low-latency. Преимущественно их работа важна для Data center, так как в них многие детали для уменьшения Bandwidth уже учтены и это позволяет нам сконцентрироваться на улучшение Latency.

Важными сферами применения является высокочастотный трейдинг, научные распределенные вычисления и облачные вычисления.

Одним из главных нововведений является разработка новой структуры данных – Phantom Queue. Данная структура, хоть и не является очередью по своей сути, позволяет нам контролировать поток сообщений внутри сети. DSTCP – улучшенная версия TCP используемая всюду в качестве основного протокола.

Исходя из данных которые авторы приводят в конце работы мы видим небольшое ухудшение Bandwidth, однако серьезный рост Latency с лихвой покрывает потери.

Личное мнение: статья сложная и требует долгого и вдумчивого чтения, однако тема весьма интересная, что несколько облегчает восприятие.