Real Time Adaptive Profiling in Storm Topologies
Storm Architecture
Storm Topology
RTAPTS Architecture
Optimization by Queuing Theory

- We can get the emit rate per component ($\mu$ in Queuing theory)
- Thus, we can get the input rate per component ($\lambda$ in Queuing theory)
- We assume a Poisson distribution for the arrival rates, and model the system as a M/M/1 Queue.
- In order to have a stable queue we want $\lambda/\mu < 1$.
- We choose $\lambda/\mu = 0.75$ as the optimal value
- Scale up/down to get the optimal value
Emit rate with no optimization

linear topology spout -> bolt1 -> bolt2 -> bolt3 -> bolt4, with bottleneck in bolt3
Emit rate with RTAPTS optimization

linear topology spout -> bolt1 -> bolt2 -> bolt3 -> bolt4, with bottleneck in bolt3
Cpu Gain vs number of Threads
Scalability Test

Topology with no Bottleneck: total emit rate vs number of workers

- S emit
- B1 emit
- B2 emit
- B3 emit
Bottleneck Behavior

Topology with bottleneck: total emit rate vs number of workers

- S emit
- B1 emit
- B2 emit
- B3 emit

Number of workers