1 R2.O2 ON SIMPLIFICATIONS OF THE PROBLEM SETTING

Fig. 1 and Fig. 2 illustrate the overall effectiveness and efficiency for zipf-distributed and uniformly distributed remote data separately. They show similar performance benefits.

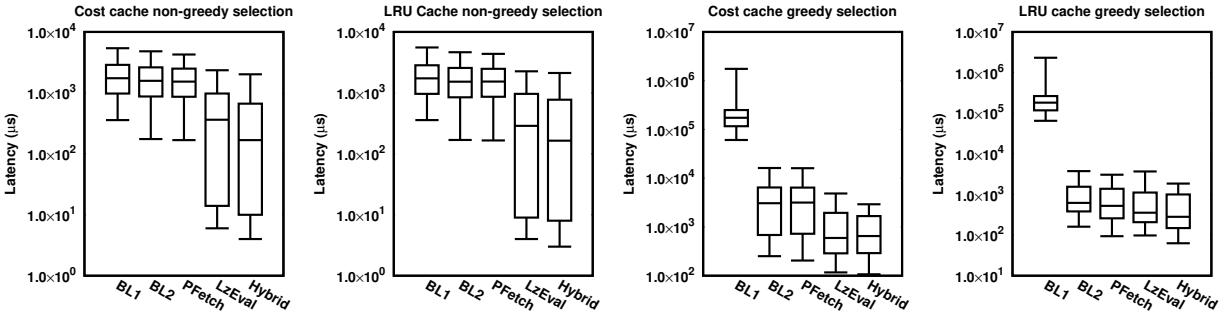

Figure 1: Overall effectiveness and efficiency for zipf-distributed remote data (log scale).

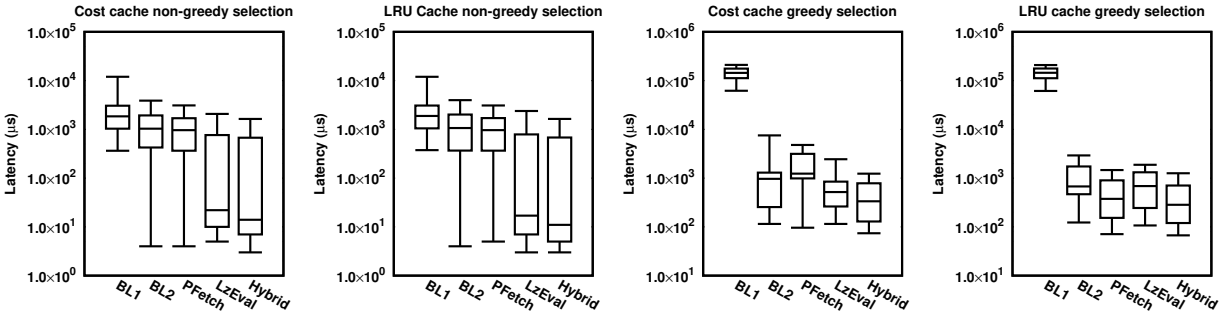

Figure 2: Overall effectiveness and efficiency for uniformly distributed remote data (log scale).

2 R2.O3 ON INCLUDING A BASELINE FOR POST-PROCESSING

Fig. 3 shows lazy evaluation postpone steps' impact on performance. The optimal performance is achieved when postponing four steps. Postponing one or two steps hides little latency. Whereas postponing seven steps generates too much extra partial matches and thwarts all the benefit of lazy evaluation.

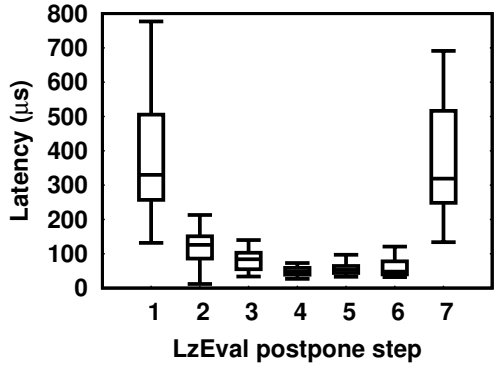

Figure 3: Lazy evaluation postpone steps' impact on performance.