Karlsruhe Institute of Technology

A Bulk-Parallel Priority Queue in External Memory with STXXL

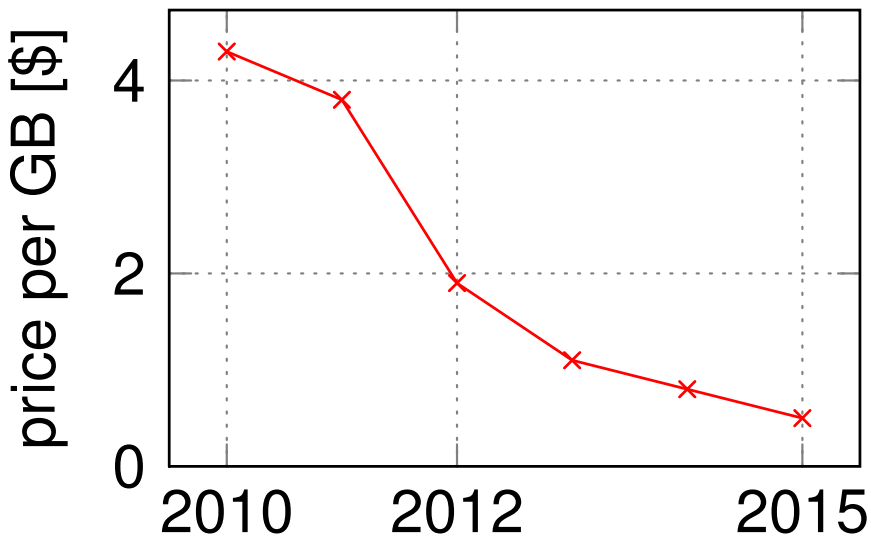
Timo Bingmann, Thomas Keh, Peter Sanders | June 29th, 2015 @ SEA 2015

INSTITUTE OF THEORETICAL INFORMATICS – ALGORITHMICS

Motivation

HDD/SSD transfer rate	
2002	2015
HDD: 60 MB/s	HDD: 180 MB/s SSD: 500+ MB/s

Server SSD price trend



Sorting speed	
Intel Xeon E5-2650 v2 *Sorting 1 GB of random integers (64-bit)*	
1 core	85 MB/s
16 cores	1290 MB/s

Source: Gartner, Market Trends: Evolving HDD and SSD Storage Landscapes

History and Classification

External Memory

Brodal, Katajainen
Worst-case efficient external-memory priority queues (1998)

Sanders
Fast Priority Queues for Cached Memory (2000)

Concurrent

Sundell, Tsigas
Fast and lock-free concurrent priority queues for multi-thread systems (2003)

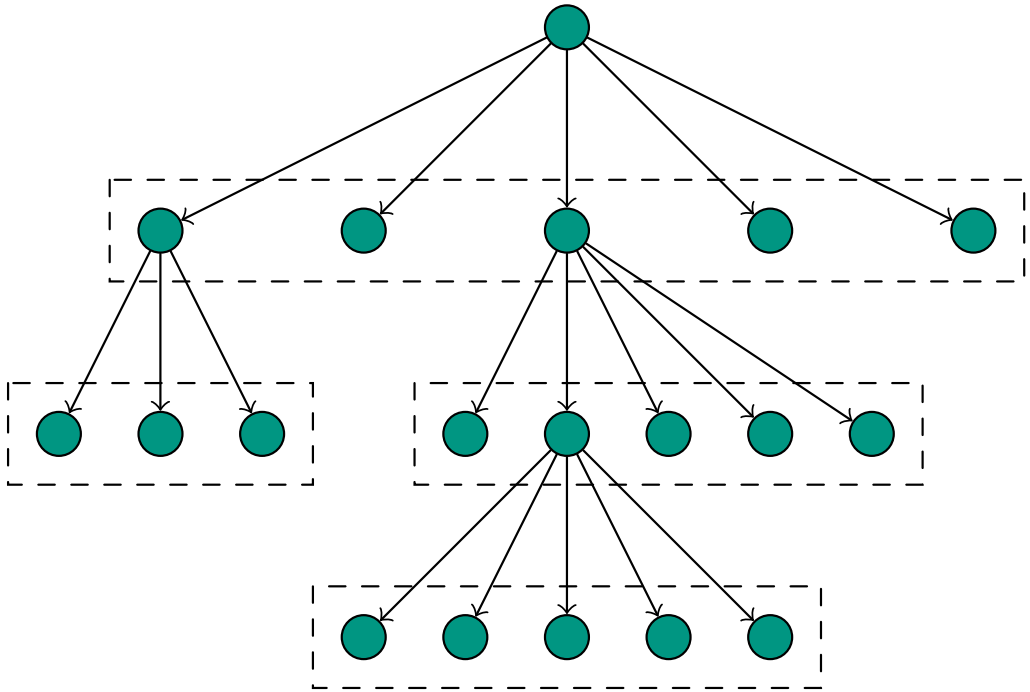
Relaxed

Alistarh et. al
The SprayList: a scalable relaxed priority queue (2014)

Parallel EM

Beckmann et. al.
Parallel multiway merge for the sequence heap (2009)

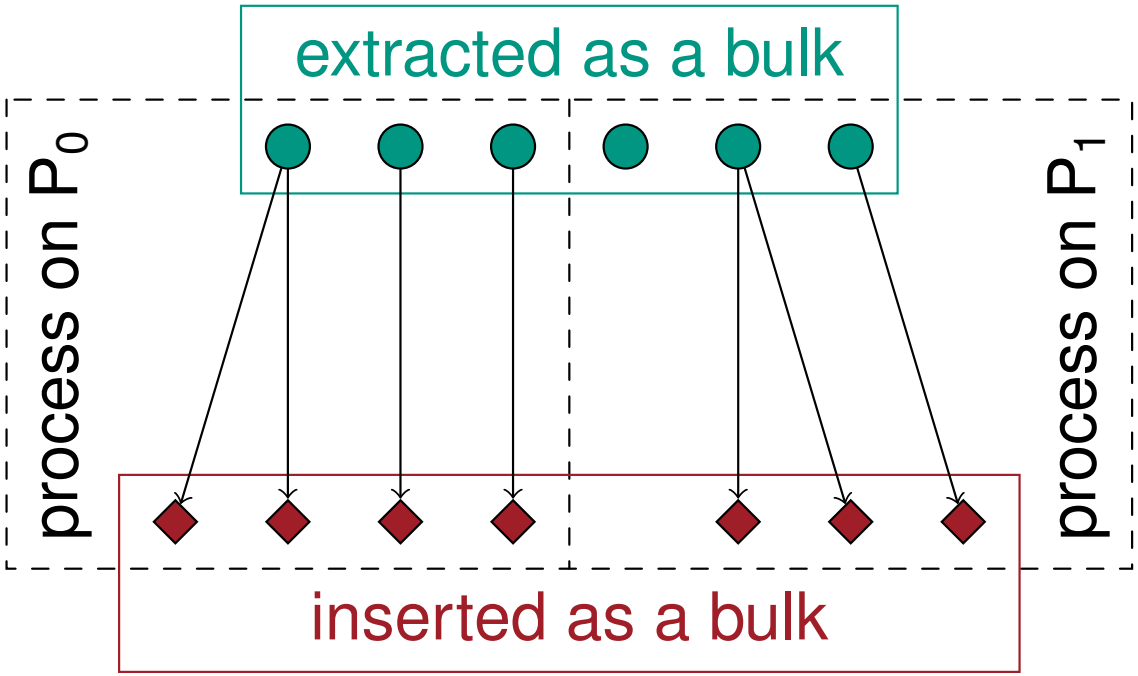
Bulk Push

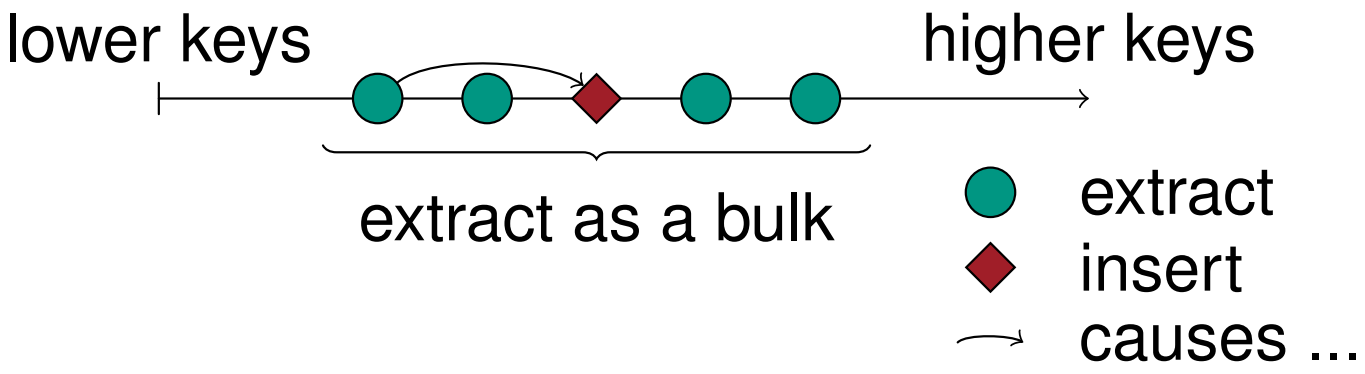


bulk push

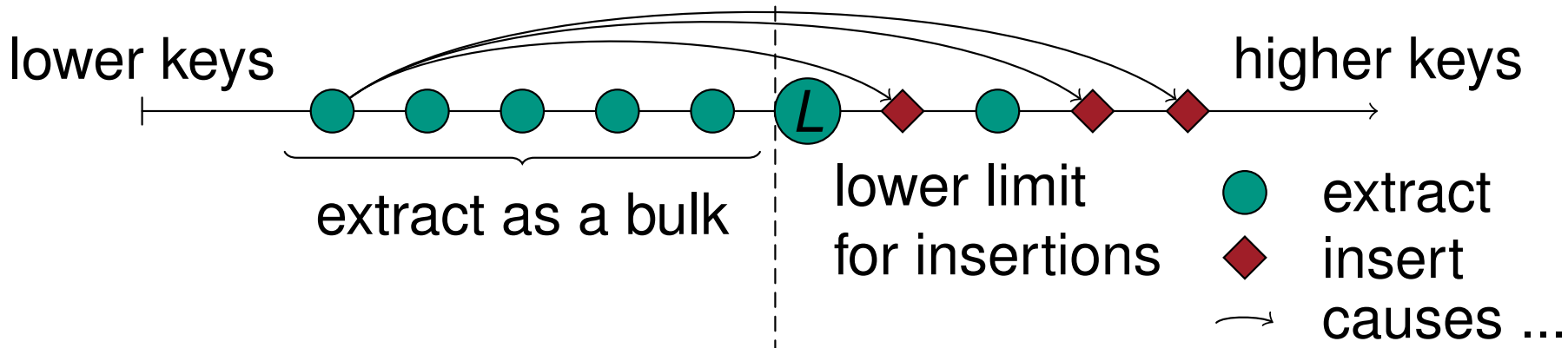
```
ppq.bulk_push_begin(
    approx_bulk_size);
#pragma omp parallel for
for (...) {
  ppq.bulk_push(value,
      thread_id);
}
ppq.bulk_push_end();
```

Bulk Pop


extracted as a bulk
process on P_0
process on P_1
inserted as a bulk

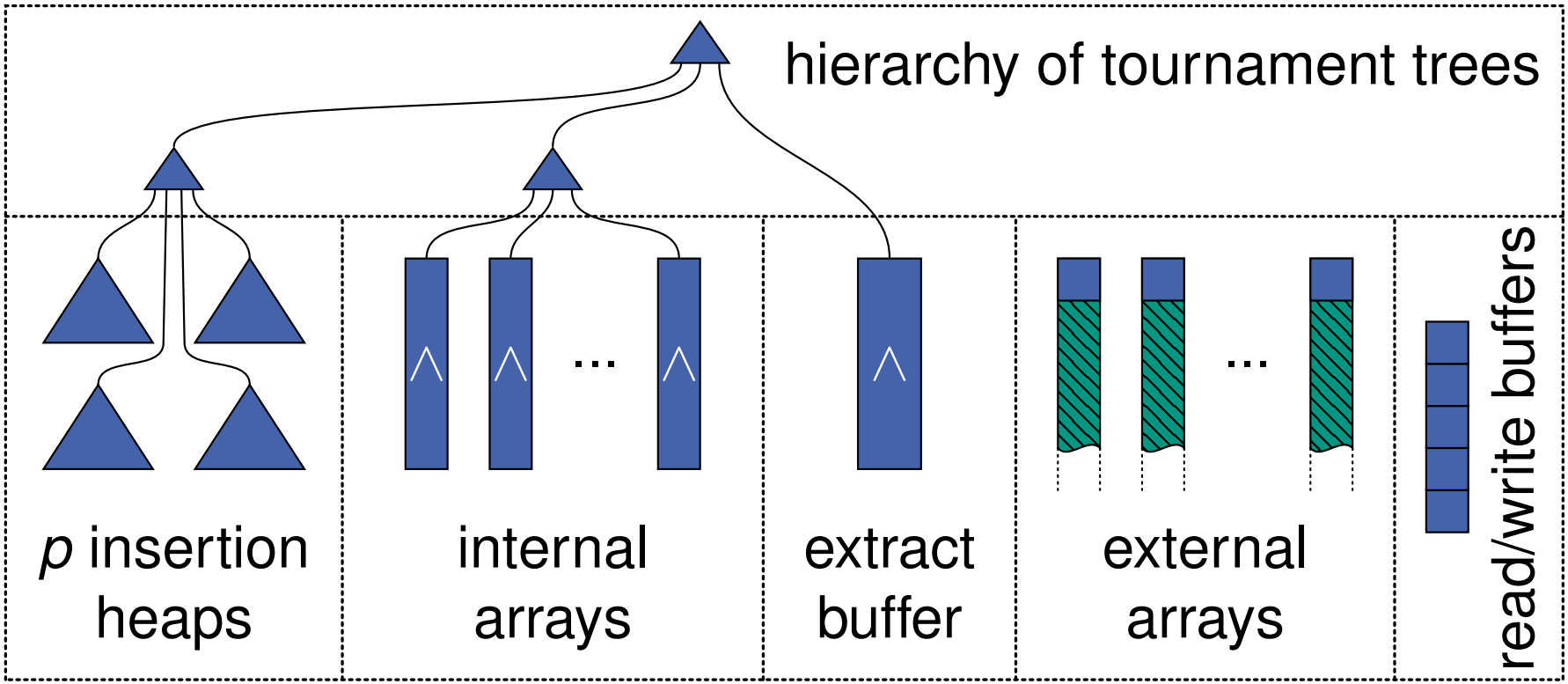

lower keys
higher keys
extract as a bulk
extract
insert
causes ...

Bulk Limit

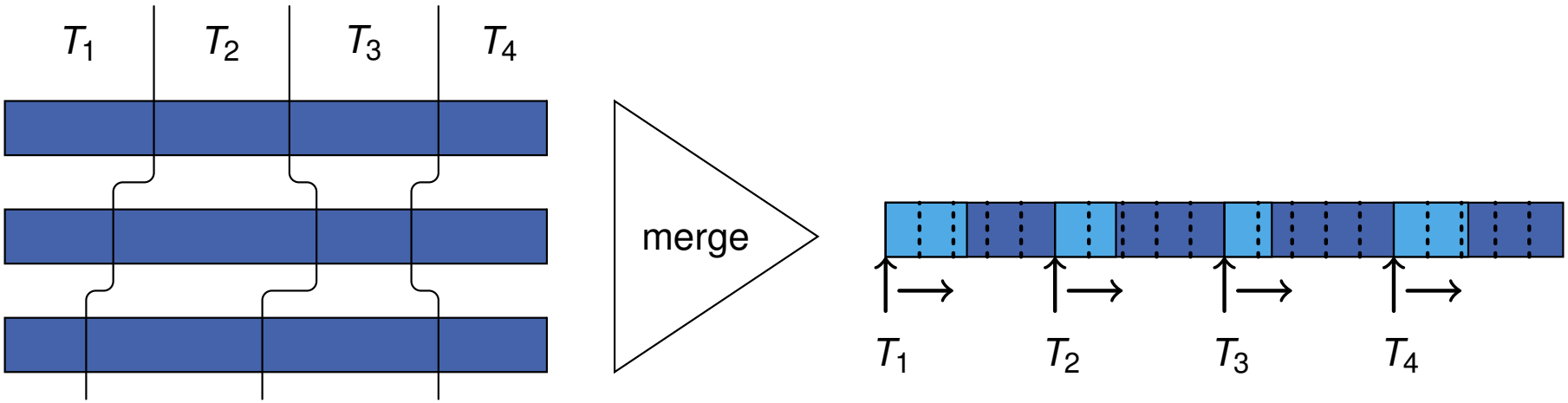



```
vector<item> work;
while (!ppq.empty()) {
    ppq.bulk_pop_limit(work, L);
    ppq.bulk_push_begin(approx_bulk_size);
    #pragma omp parallel for
    for (i = 0; i < work.size(); ++i)
        // process work[i], maybe bulk_push
    ppq.bulk_push_end();
}
```

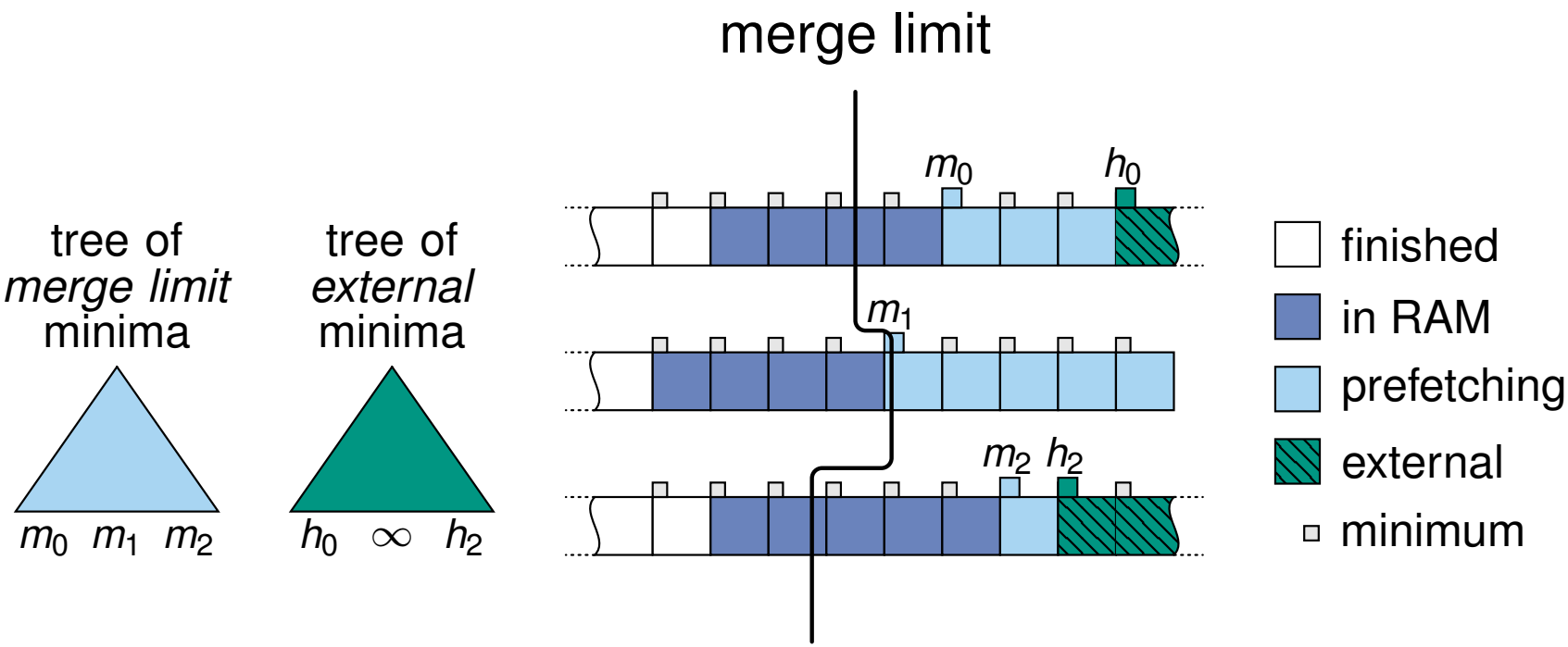
Components


hierarchy of tournament trees
p insertion heaps
internal arrays
extract buffer
external arrays
read/write buffers

Parallel Multiway Merge



Prefetch Prediction



Experiments

Platforms

A-Rotational	B-SSD
Intel Xeon X5550 2.66*GHz* 2 sockets, 8 + 8 HT cores 48*GB* RAM 8/17*GB/s* (1/8 cores) 6x Seagate SV35.5 1*TB* 85–170*MB/s* read/write parallel: 740*MB/s* R/W	Intel Xeon E5-2650 v2 2.6*GHz* 2 sockets, 16 + 16 HT cores 128*GB* RAM 12/16*GB/s* (1/16 cores) 4x Samsung SSD 840 Evo 1*TB* 512/475*MB/s* read/write parallel: 2.0/1.6*GB/s* R/W

Competitors

Bulk-Parallel Priority Queue (PPQ)

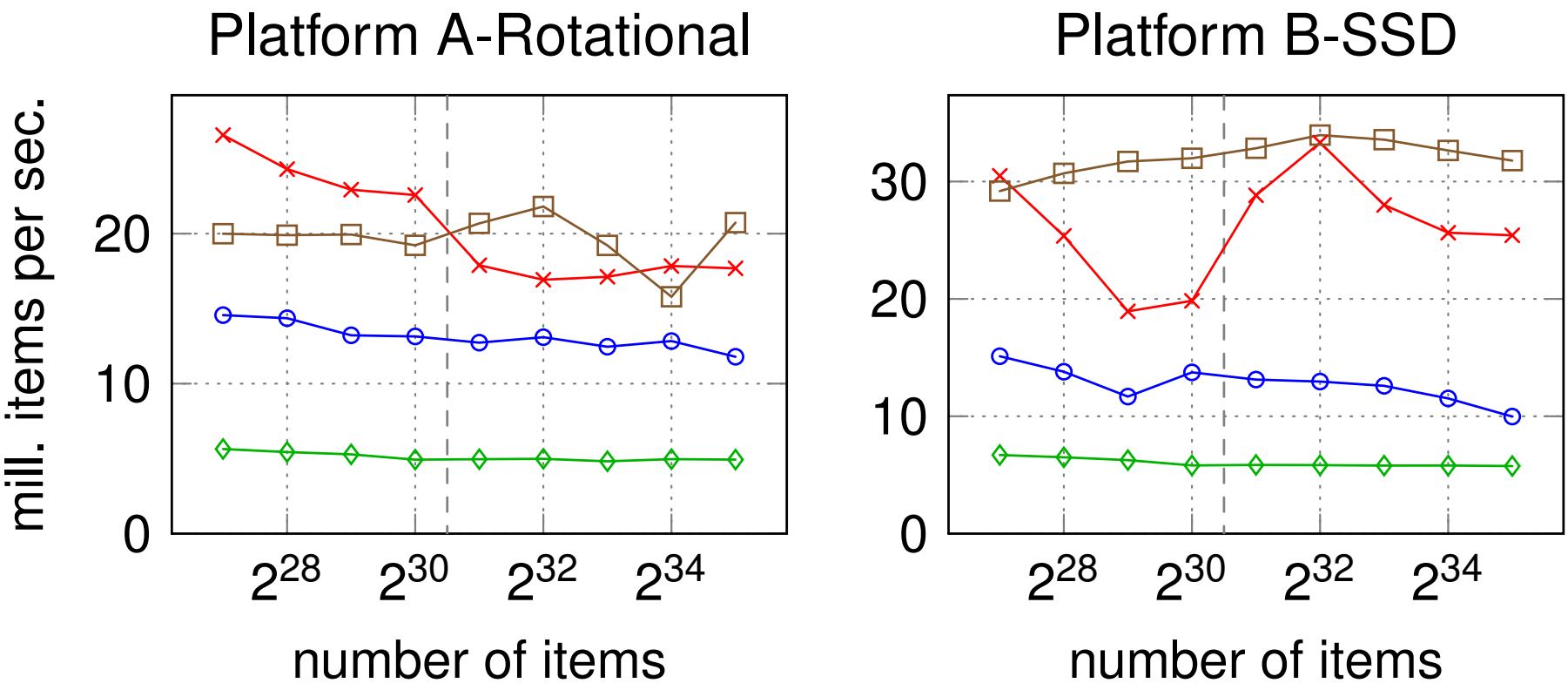
STXXL Priority Queue (SPQ-S)

STXXL Priority Queue + Parallel Multiway Merge (SPQ-P)

STXXL Sorter

Insert-All-Delete-All

push	15	151	2	67	...	|	97	13	83	...	|	...
pop	1	2	4	4	...	|	137	138	140	...	|	...



bulk size = 640k

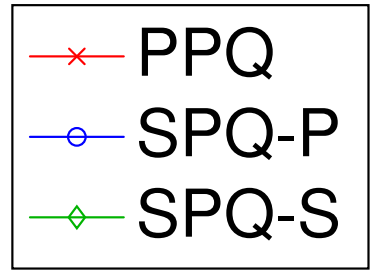
8-byte items

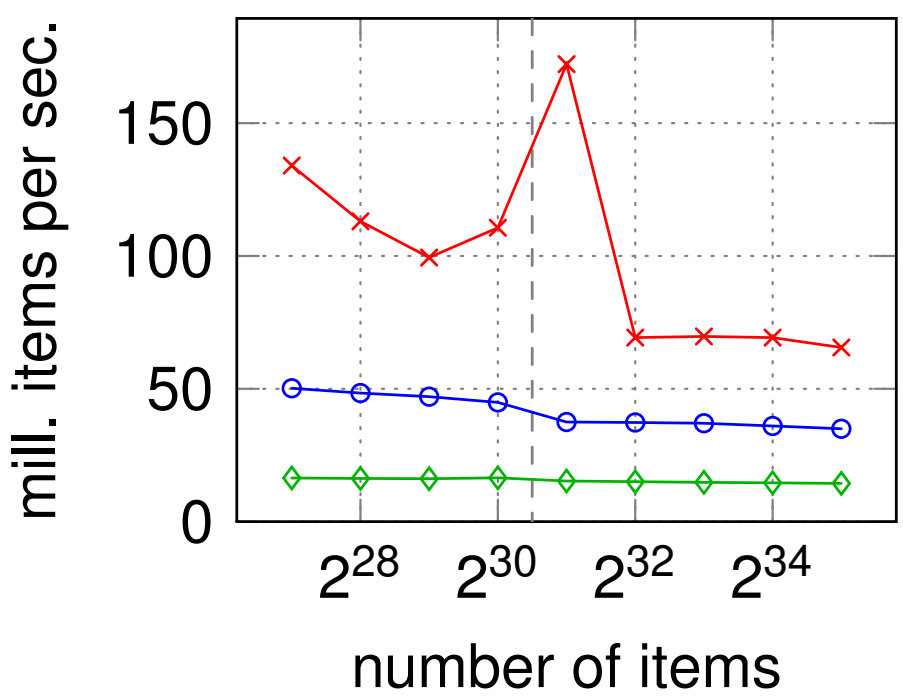
Random Bulk Rewrite

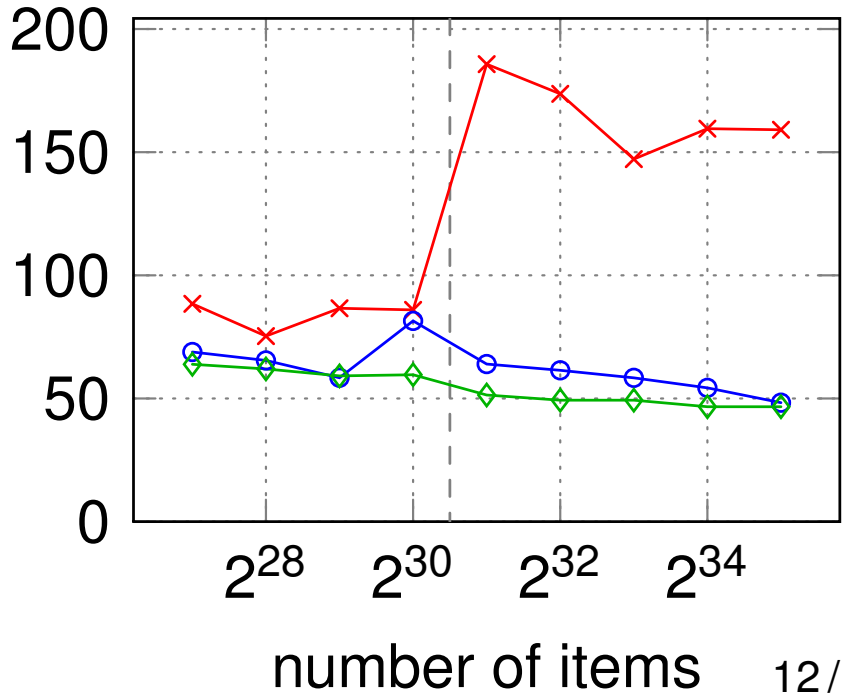
prefill	1	2	3	4	5	...	n
pop	1	2	3	4	5	...	b_0
push	n+1	n+2	n+3	...			n+b_0
...							
pop	n−b_k	...	n−3	n−2	n−1	n	
push	2n−b_k	...	2n−3	2n−2	2n−1	2n	

$b_i \in [0,640k]$

8-byte items





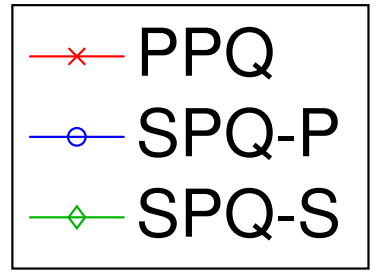


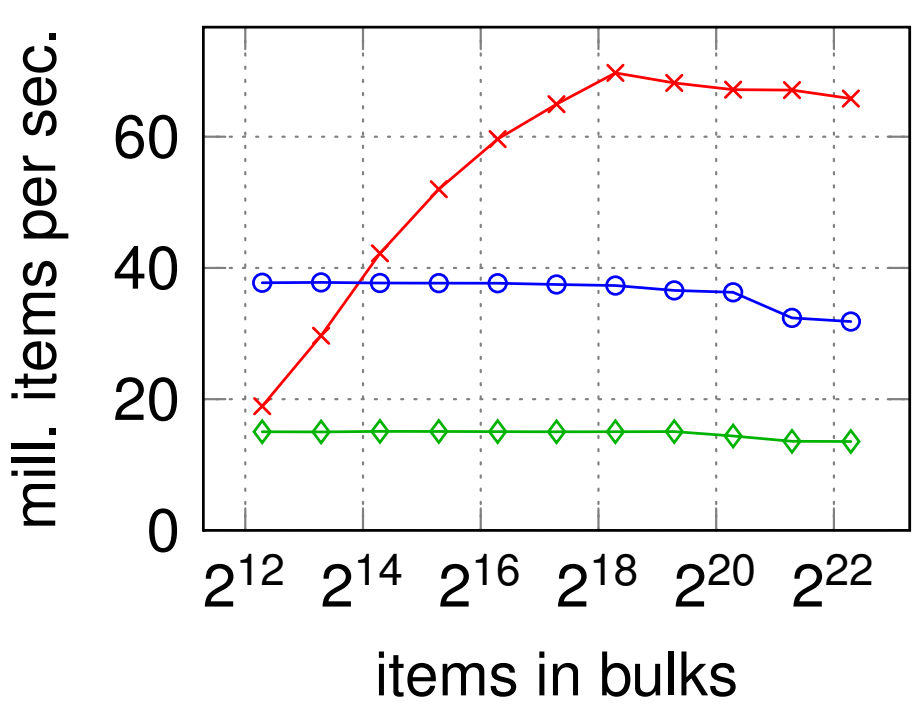
Variate Bulk Size

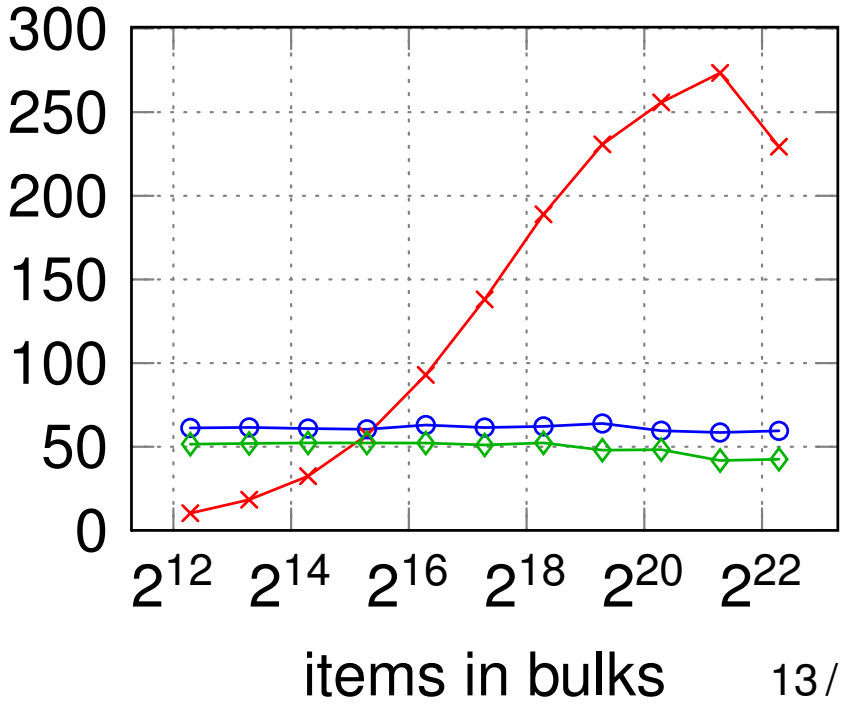
prefill	1	2	3	4	5	...	n
pop	1	2	3	4	5	...	b
push	n+1	n+2	n+3	...	n+b		
...							
pop	n-b	...	n-3	n-2	n-1	n	
push	2n-b	...	2n-3	2n-2	2n-1	2n	

V = 32 GB

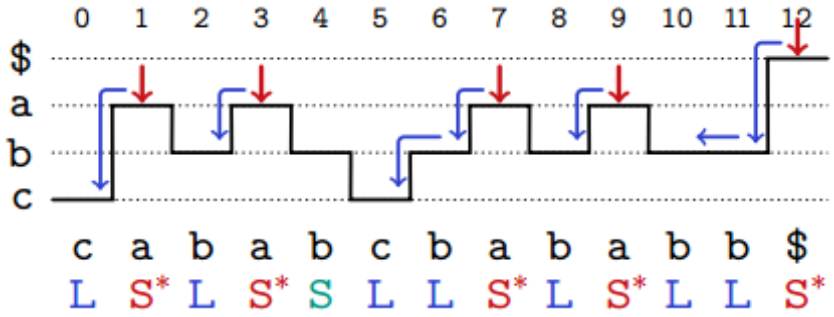
8-byte items







Future Work



- Integration in eSAIS (external suffix and LCP array construction)
- Overlap Application and Priority Queue work
- Experiments with other internal memory PQs