

# Report: hierarchical workstealing in XKA-API

XKA-API team

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Enabling HWS in XKA-API</b>	<b>4</b>
<b>3</b>	<b>HWS implementation in XKA-API</b>	<b>5</b>
<b>4</b>	<b>Benchmarks</b>	<b>7</b>
<b>5</b>	<b>TODOs</b>	<b>8</b>

# 1 Introduction

TODO: describe the need for HWS

## 2 Enabling HWS in XKA-API

### 2.1 Build configuration

No new configuration has been added to the build system. However, the runtime has to be compiled with *hwloc* and *numa* support for HWS to be enabled:

---

```
./configure --with-hwloc --with-numa
```

---

### 2.2 Environment variables

The steal request emission routine has to be specialized by setting the *KAAPI\_EMITSTEAL* environment variable to “hws”:

---

```
$> KAAPIHWSLEVEL=hws ./a.out
```

---

Specific memory hierarchy levels can be used by setting the *KAAPI\_HWS\_LEVELS* environment variable. It consists of a comma separated list of one or more of the following values:

- ALL: enables all the levels,
- NONE: disables all the levels,
- L3: enable the L3 cache level,
- NUMA: enable the numa level,
- SOCKET: enable the socket level,
- MACHINE: enable the machine level,
- FLAT: enable the flat level.

For instance:

---

```
$> KAAPIHWSLEVEL=hws KAAPIHWSLEVELS=FLAT,NUMA ./a.out
```

---

Not setting this variable enables the NUMA, SOCKET, MACHINE and FLAT memory levels.

### 3 HWS implementation in XKA-API

#### 3.1 Overview

The *src/hws* directory has been added to the runtime sourcecode. It implements:

- subsystem initialization (*kaapi\_hws\_initialize.c*),
- task pushing (*kaapi\_hws\_push\_task.c*),
- steal request emission (*kaapi\_hws\_emitsteal.c*),
- adaptive task handling (*kaapi\_hws\_adaptive.c*),
- initial splitter (*kaapi\_hws\_splitter.c*),
- workstealing queue implementation (*kaapi\_ws\_queue\_XXX.c*),
- related performance counters (*kaapi\_hws\_counters.c*),
- scheduler synchronization (*kaapi\_hws\_sched\_sync.c*).

#### 3.2 Hierarchy construction

TODO

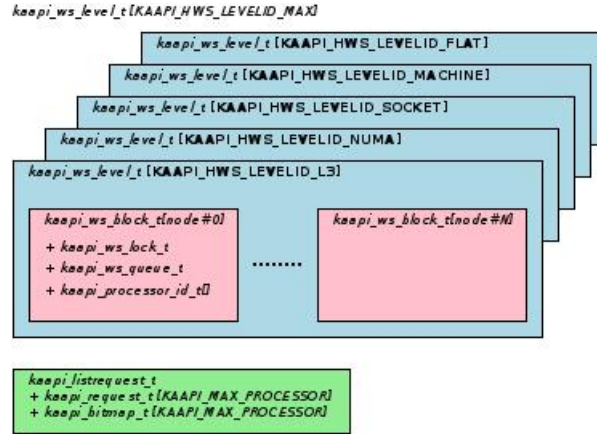


Figure 1: HWS data structure relations

#### 3.3 Steal request emission algorithm

The steal request emission entrypoint is the *kaapi\_hws\_emitsteal* routine. It is still in progress, the final implementation will depend on the performances achieved during the benchmarking process. Currently, the algorithm is as follow:

1. try to pop from the local queue, stop upon success.
2. steal in each parent level.
3. during the parent level iteration, try to steal in each child level.
4. if child level stealing fails, try to steal in parent level leaves.
5. fail the unanswered requests.

### 3.4 Workstealing queue virtualization

To provide better flexibility regarding workstealing decision, a new *queue* interface has been added. It contains the following methods:

- push: push a task in the queue,
- pop: pop a task from the local queue,
- steal: reply to a set of steal requests.

Currently, there is only a basic static lifo implementation.

### 3.5 Emistead routine virtualization

The *emistead* routine has been virtualized. The *KAAPI\_EMITSTEAL* environment variable controls the implementation used:

- hws: use the hierarchical workstealing implemented by the *kaapi\_hws\_sched\_emistead* routine,
- any other value: use the default *kaapi\_sched\_emistead* routine.

### 3.6 Workstealing blocks

TODO

## 4 Benchmarks

TODO

## 5 TODOs

The following list keeps track of previously discussed ideas not yet implemented:

- merge with current runtime. Especially, kprocs should be part of the hierarchy (ie. as leaves). The related queue methods should be specialized on the current stealprocessor and execframe implementation.
- the request should have a pointer for the data allocated area. If this pointer is set to NULL, this is left to the splitter to allocate the area and set the pointer accordingly.
- preemption of adaptive task is not yet supported. Implement an accumulate algorithm for testing.
- kaapi.c2c compiler integration.
- one bitmap per hierarchy level: avoid global bitmap contention by having one bitmap per memory level.