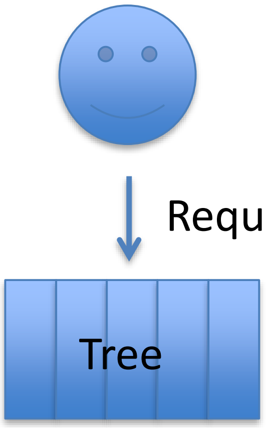
Parallel Unzipping Update

Zhe Zhang

Comparison among Three Modes of Parallel Unzipping

- Enabling Thread Building Blocks(TBB) in TTtree::GetEntry() and each task works on a branch.
- Disabling TBB in TTree::GetEntry() and unzipping baskets in cache using pthread.
- Disabling TBB in TTree::GetEntry() and unzipping baskets in cache using TBB.

TBB in GetEntry()

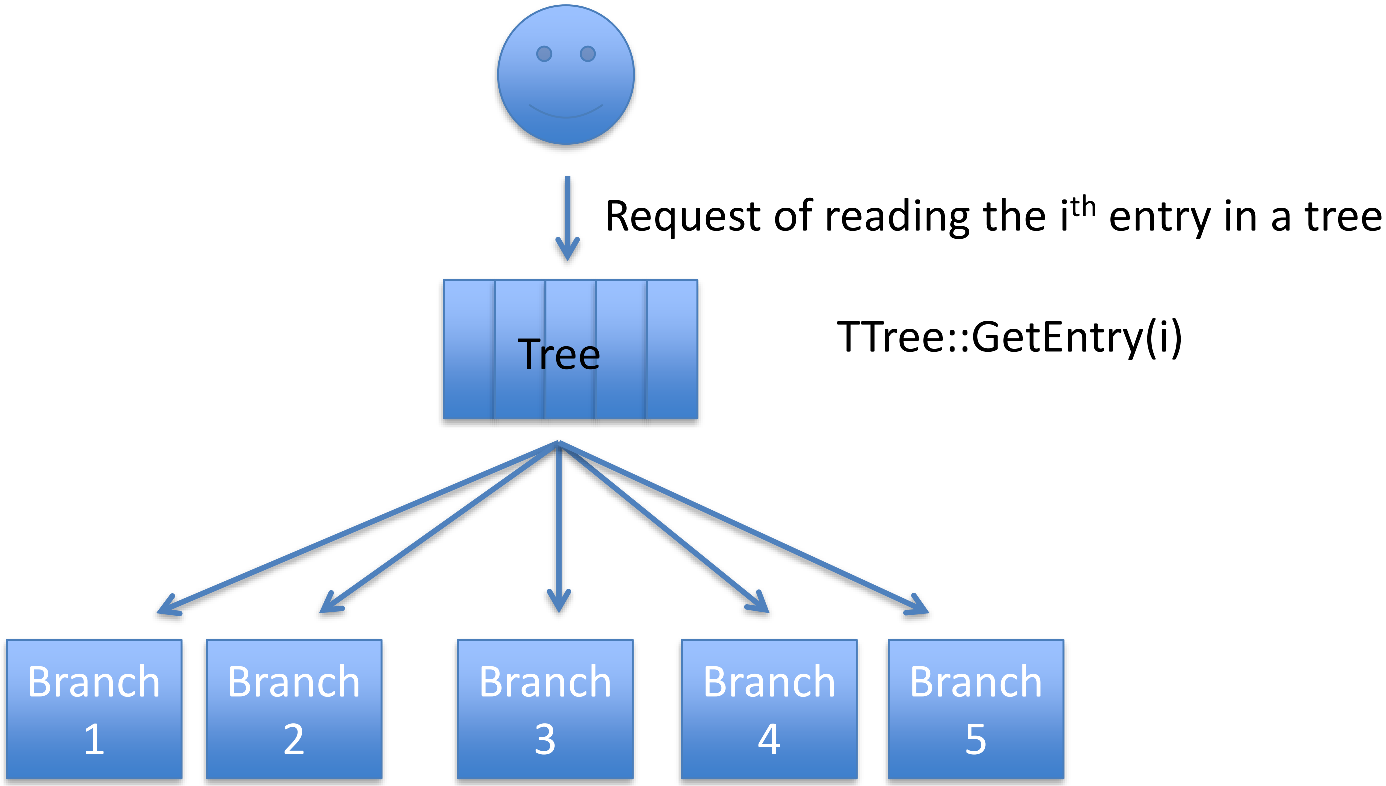


Request of reading the ith entry in a tree

Tree

TTree::GetEntry(i)

TBB in GetEntry()



Request of reading the i^{th} entry in a tree

TTree::GetEntry(i)

Tree

Branch 1

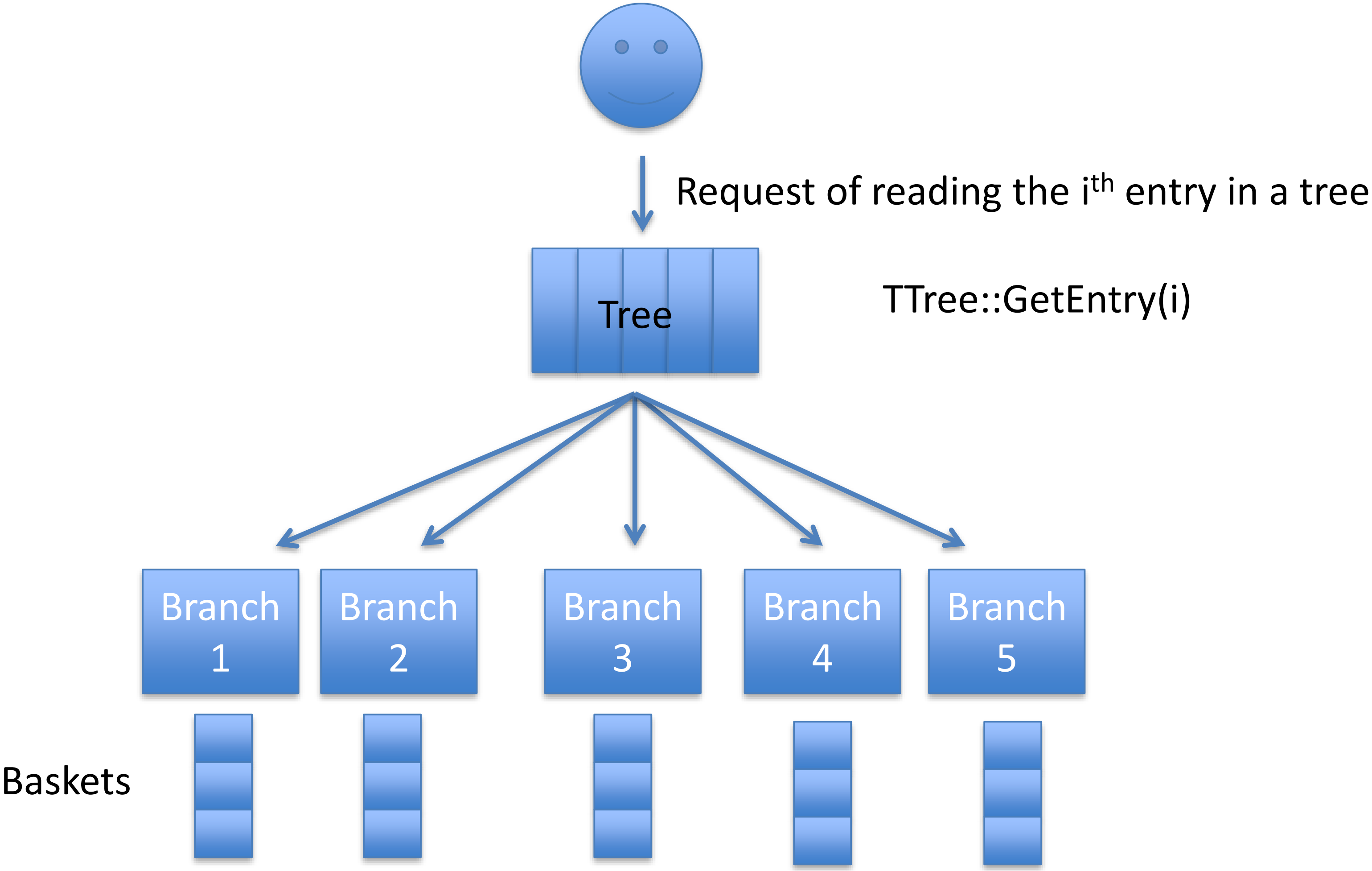
Branch 2

Branch 3

Branch 4

Branch 5

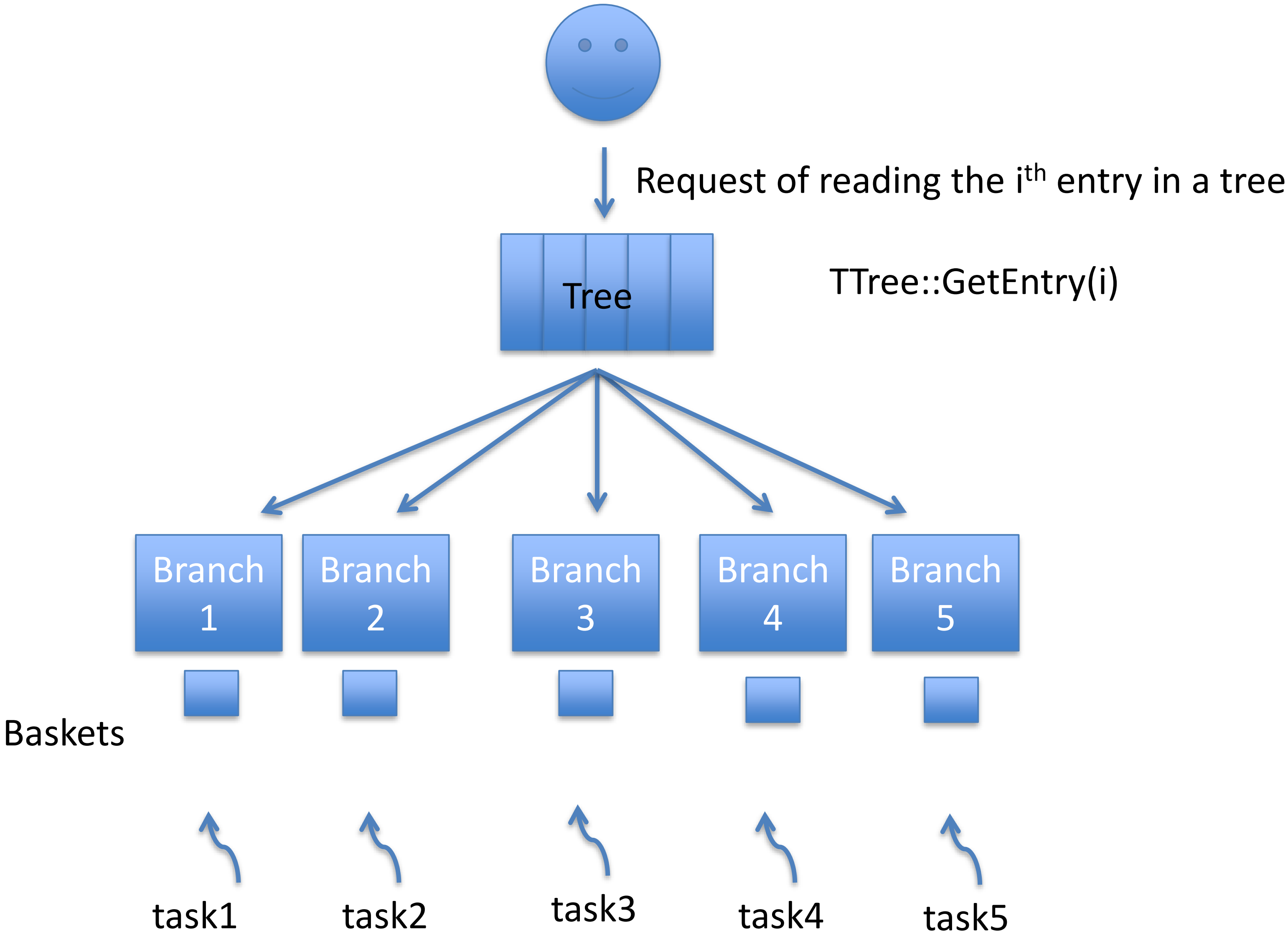
TBB in GetEntry()



Request of reading the ith entry in a tree

TTree::GetEntry(i)

TBB in GetEntry()



Request of reading the ith entry in a tree

TTree::GetEntry(i)

Tree

Branch 1

Branch 2

Branch 3

Branch 4

Branch 5

Baskets

task1

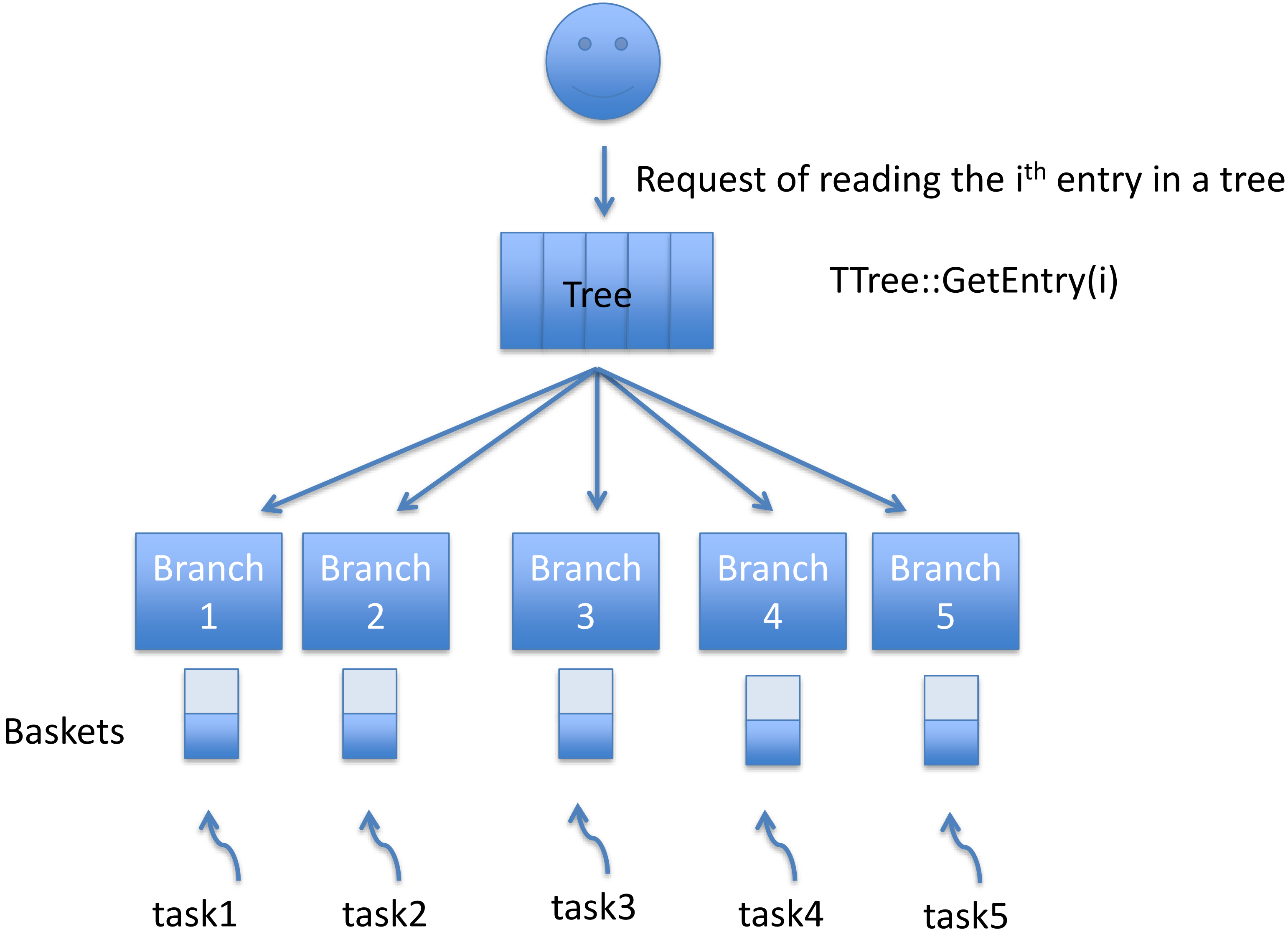
task2

task3

task4

task5

TBB in GetEntry()



Request of reading the ith entry in a tree

Tree

TTree::GetEntry(i)

Branch 1

Branch 2

Branch 3

Branch 4

Branch 5

Baskets

task1

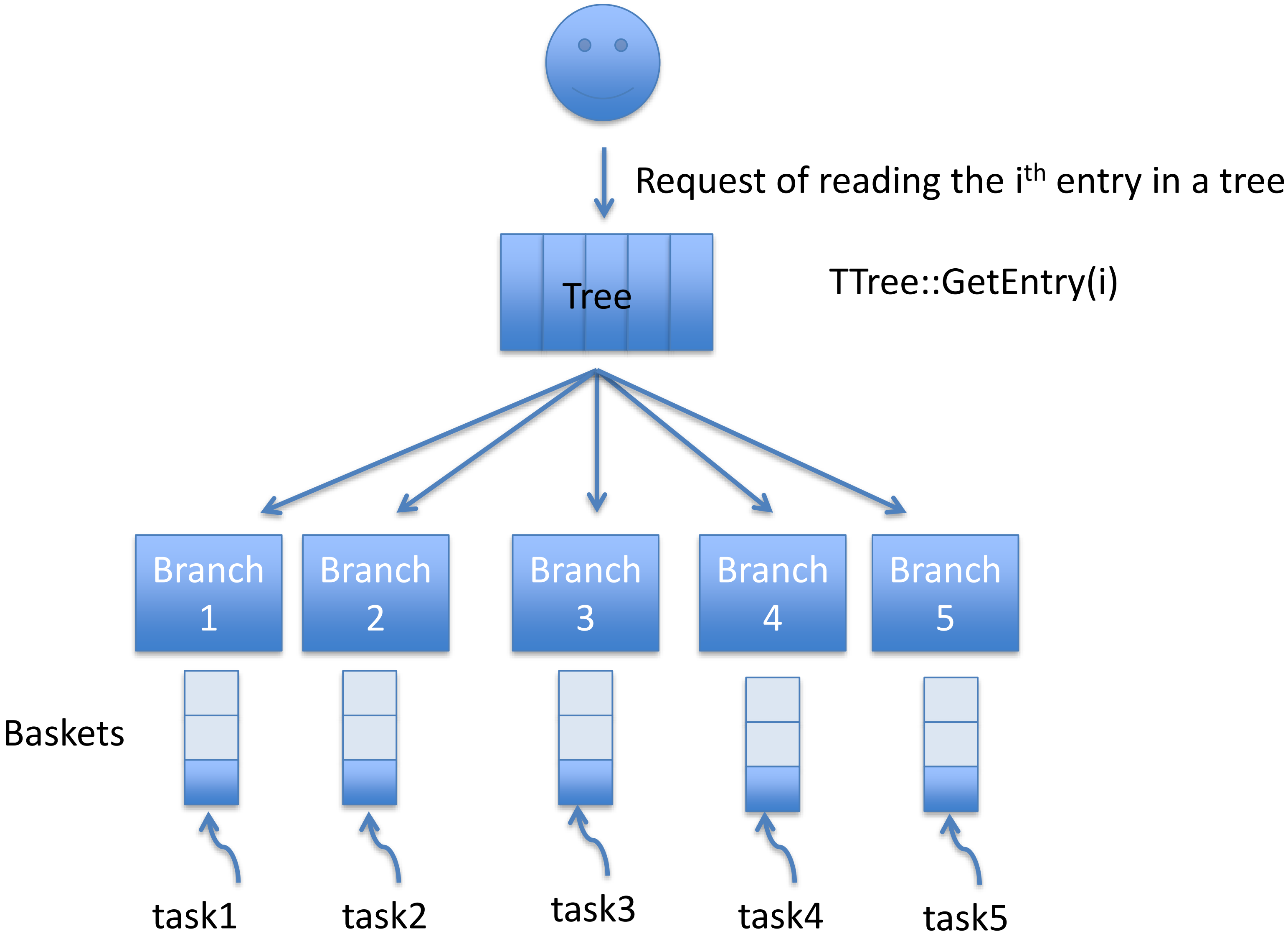
task2

task3

task4

task5

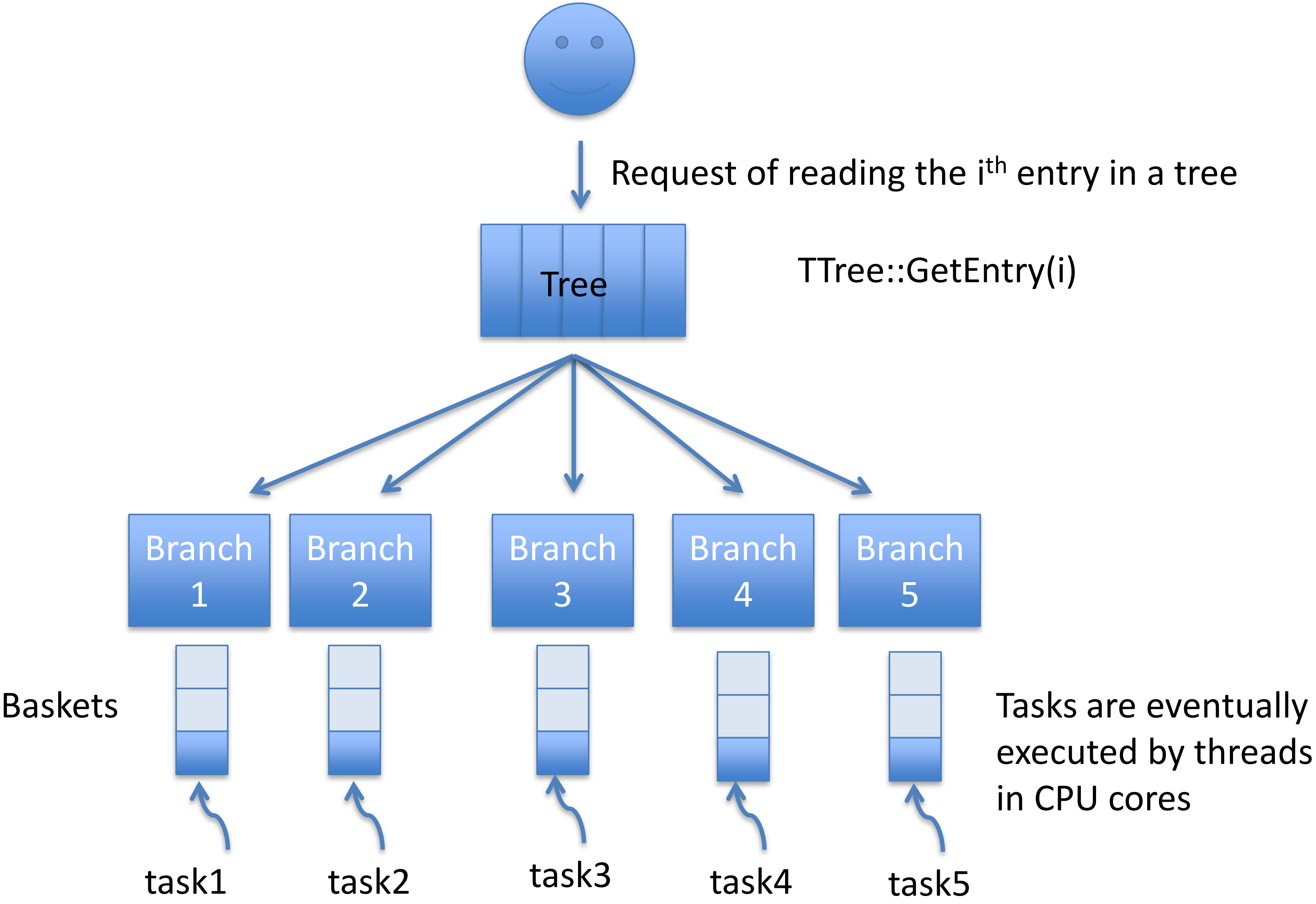
TBB in GetEntry()



Request of reading the i^{th} entry in a tree

TTree::GetEntry(i)

TBB in GetEntry()



Request of reading the i^{th} entry in a tree

Tree

TTree::GetEntry(i)

Branch 1

Branch 2

Branch 3

Branch 4

Branch 5

Baskets

Tasks are eventually executed by threads in CPU cores

task1

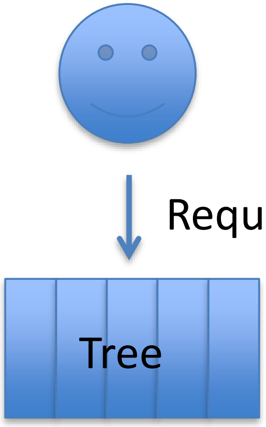
task2

task3

task4

task5

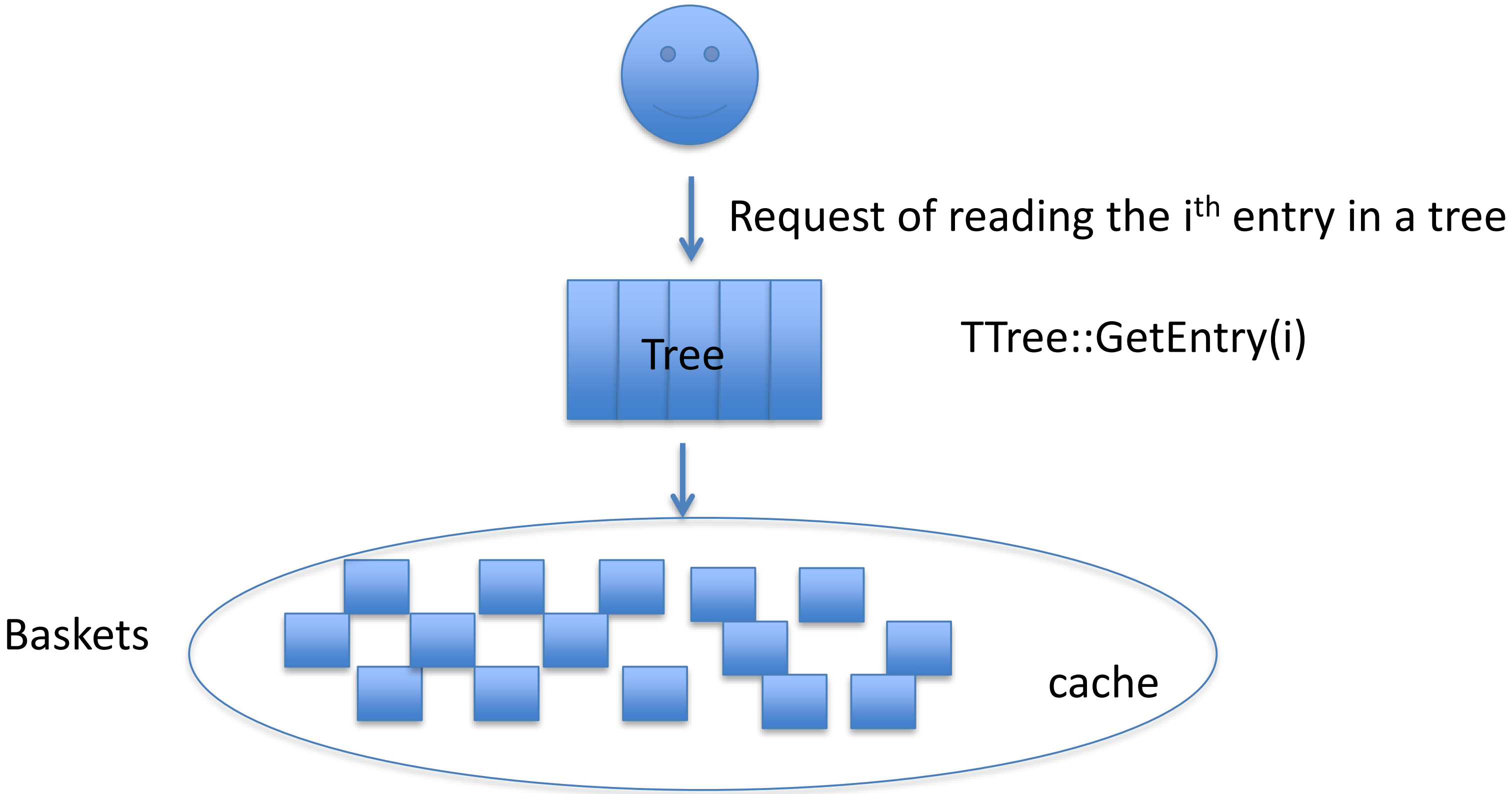
Unzipping Baskets using Pthread



Request of reading the ith entry in a tree

TTree::GetEntry(i)

Unzipping Baskets using Pthread



Request of reading the ith entry in a tree

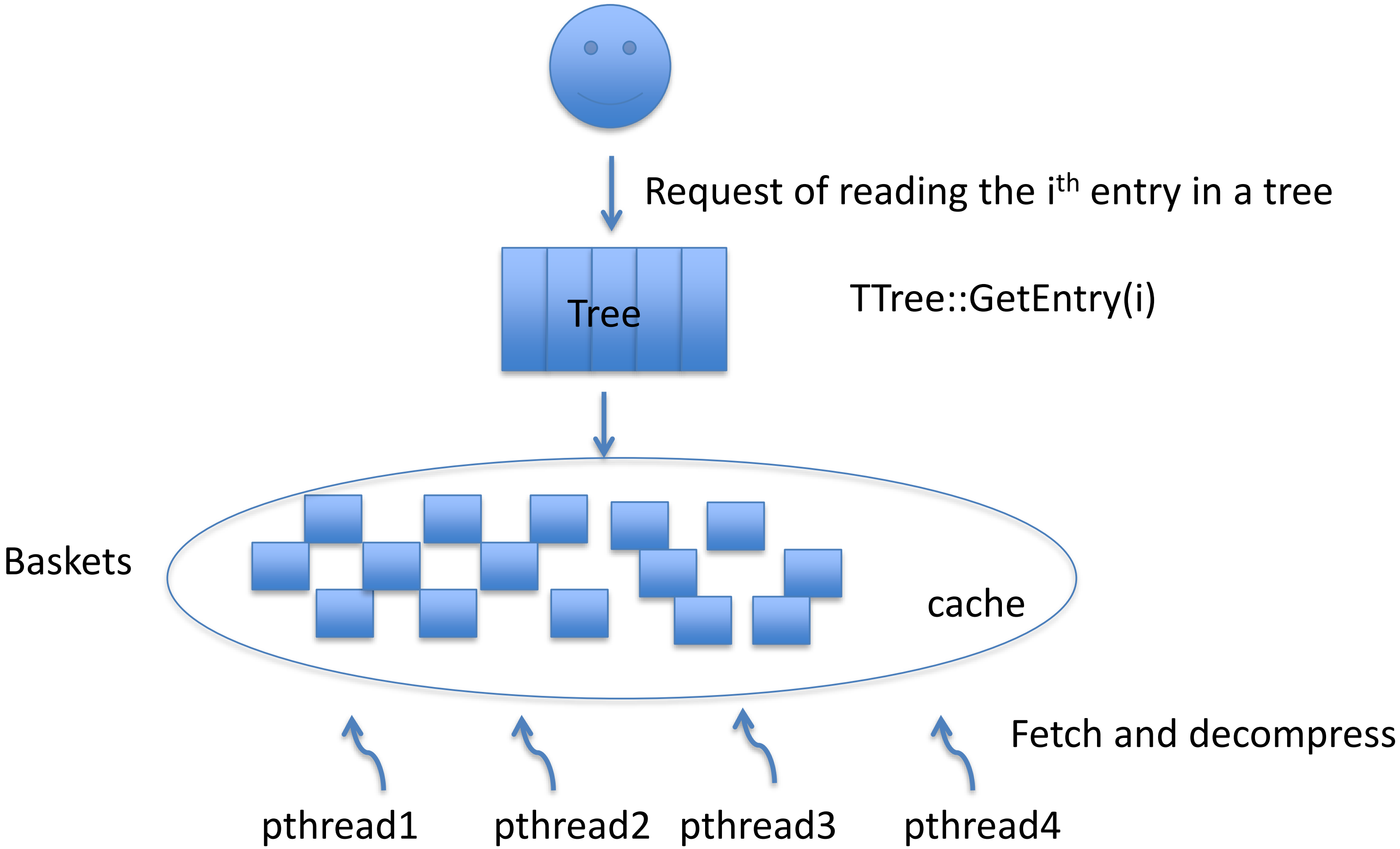
Tree

TTree::GetEntry(i)

Baskets

cache

Unzipping Baskets using Pthread



Request of reading the ith entry in a tree

Tree

TTree::GetEntry(i)

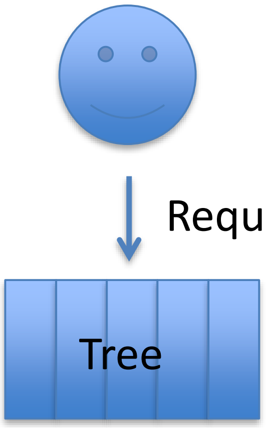
Baskets

cache

Fetch and decompress

pthread1 pthread2 pthread3 pthread4

Unzipping Baskets using TBB

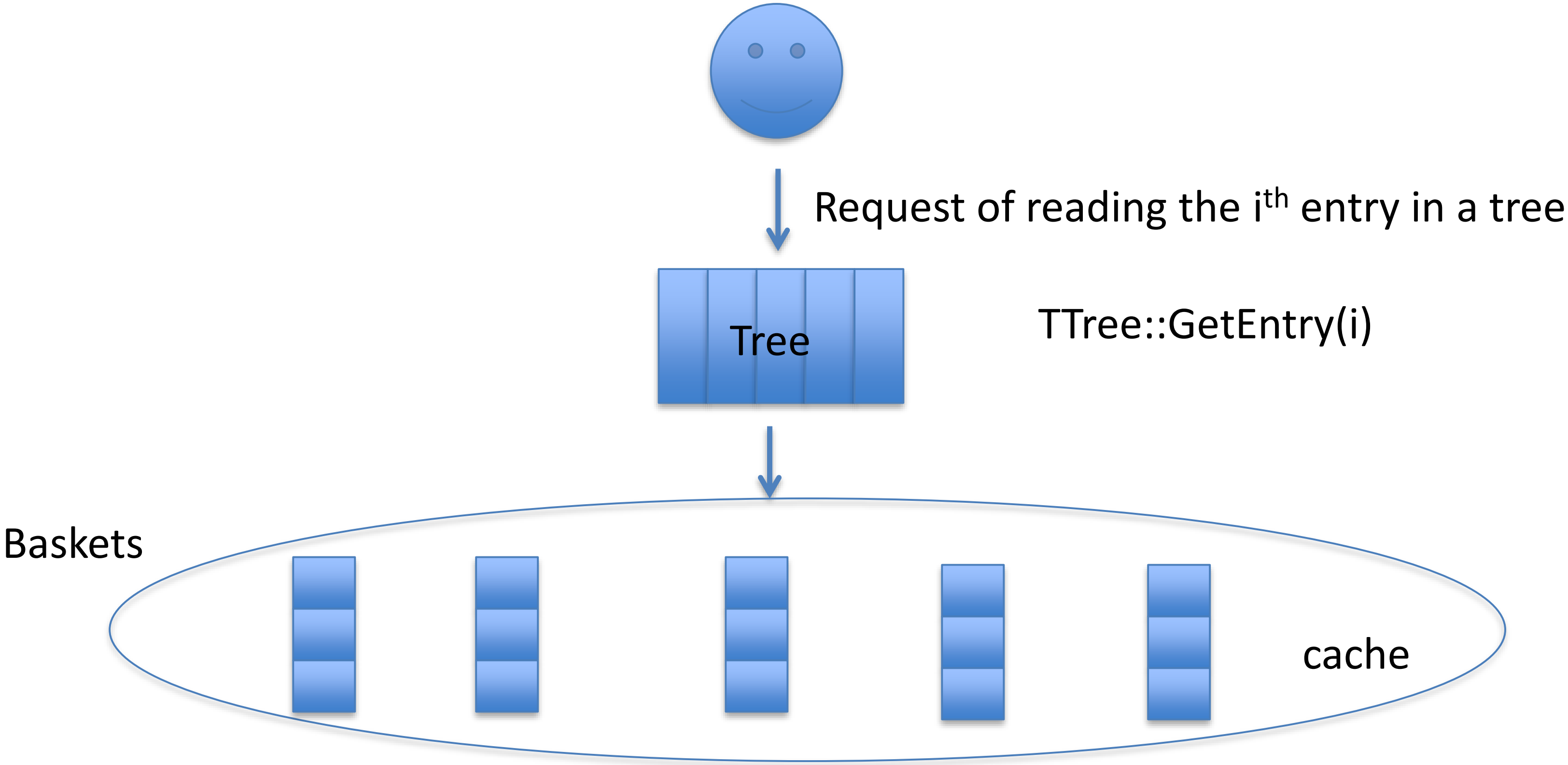


Tree

Request of reading the ith entry in a tree

TTree::GetEntry(i)

Unzipping Baskets using TBB



Request of reading the i^{th} entry in a tree

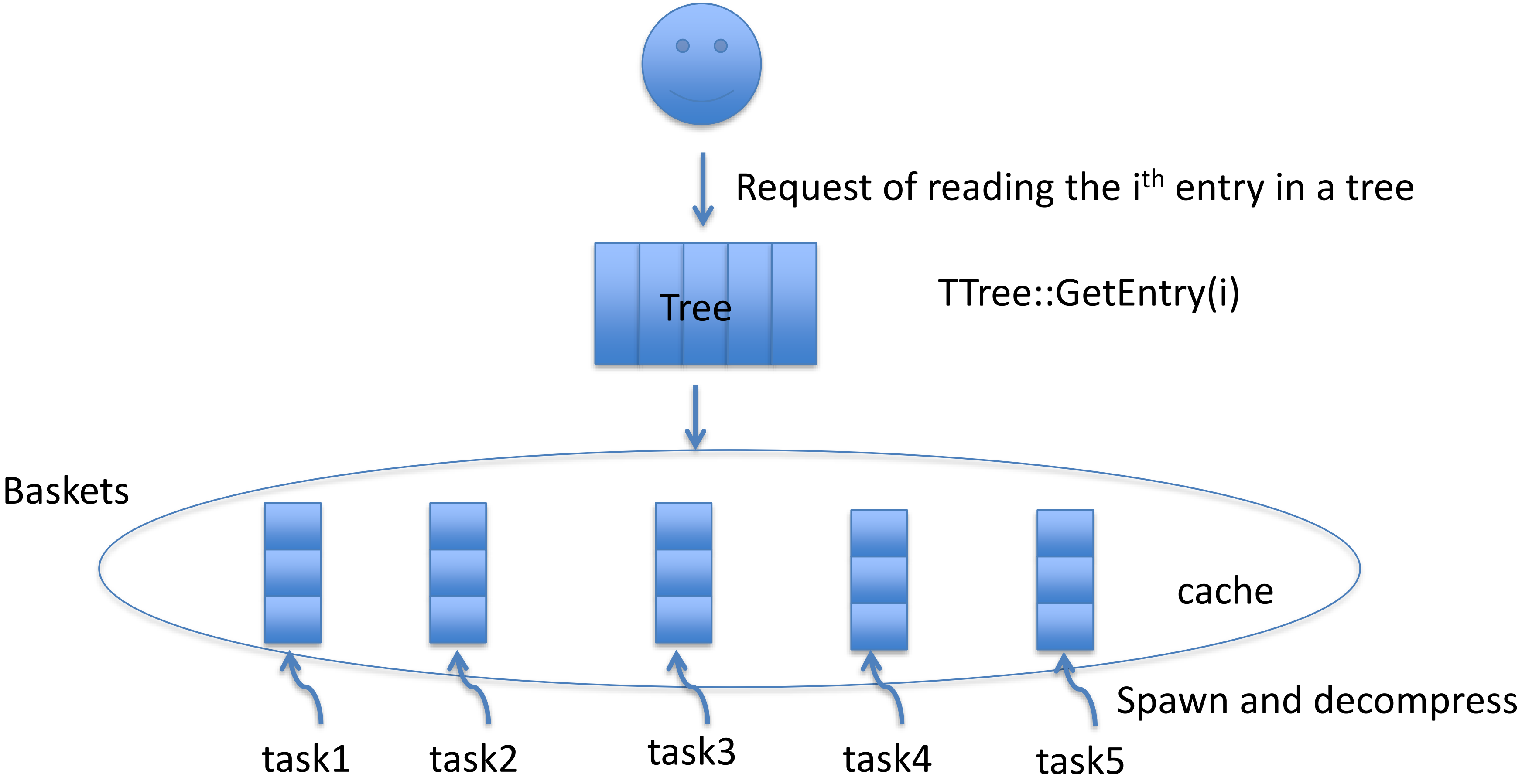
Tree

TTree::GetEntry(i)

Baskets

cache

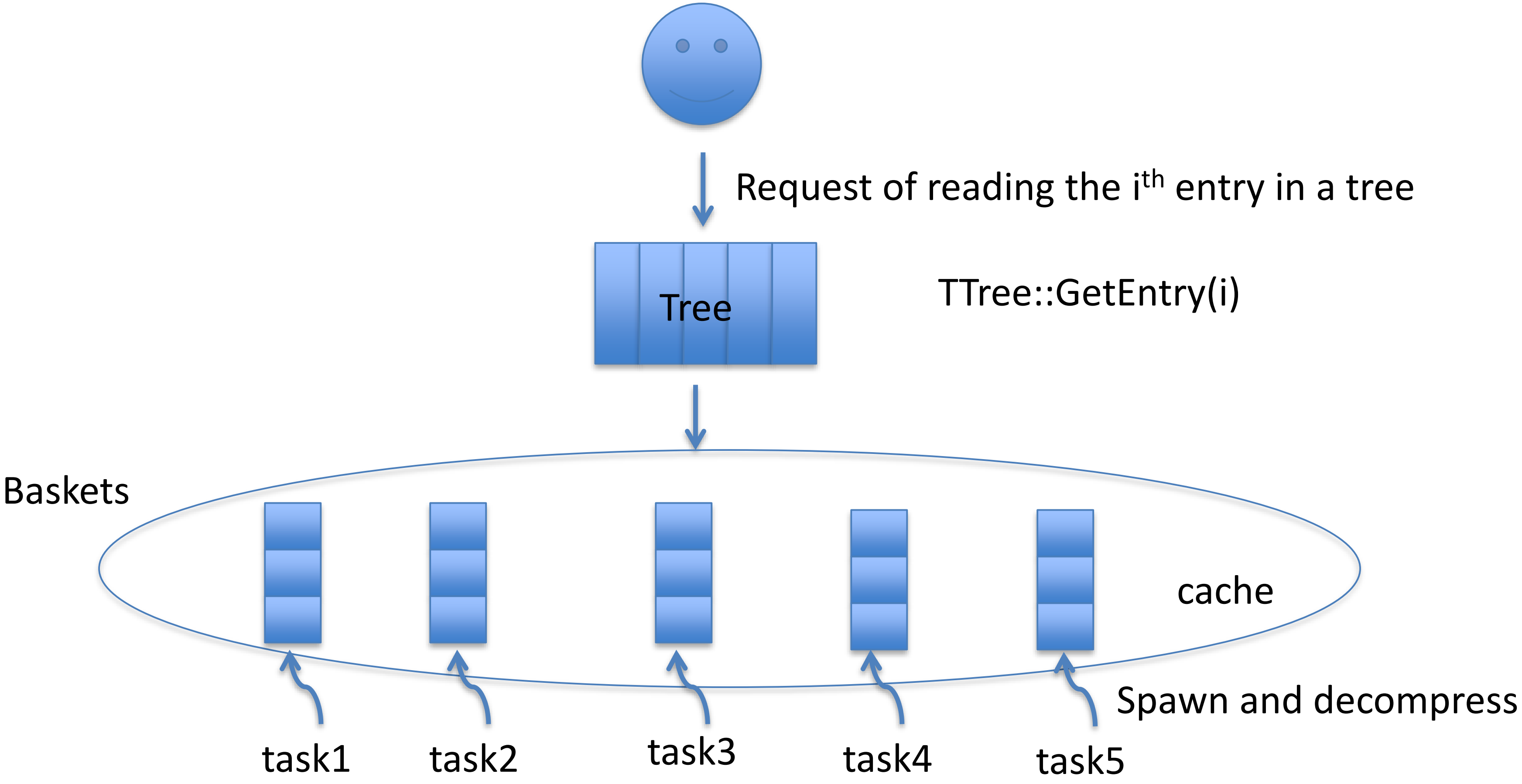
Unzipping Baskets using TBB



Request of reading the ith entry in a tree

TTree::GetEntry(i)

Unzipping Baskets using TBB



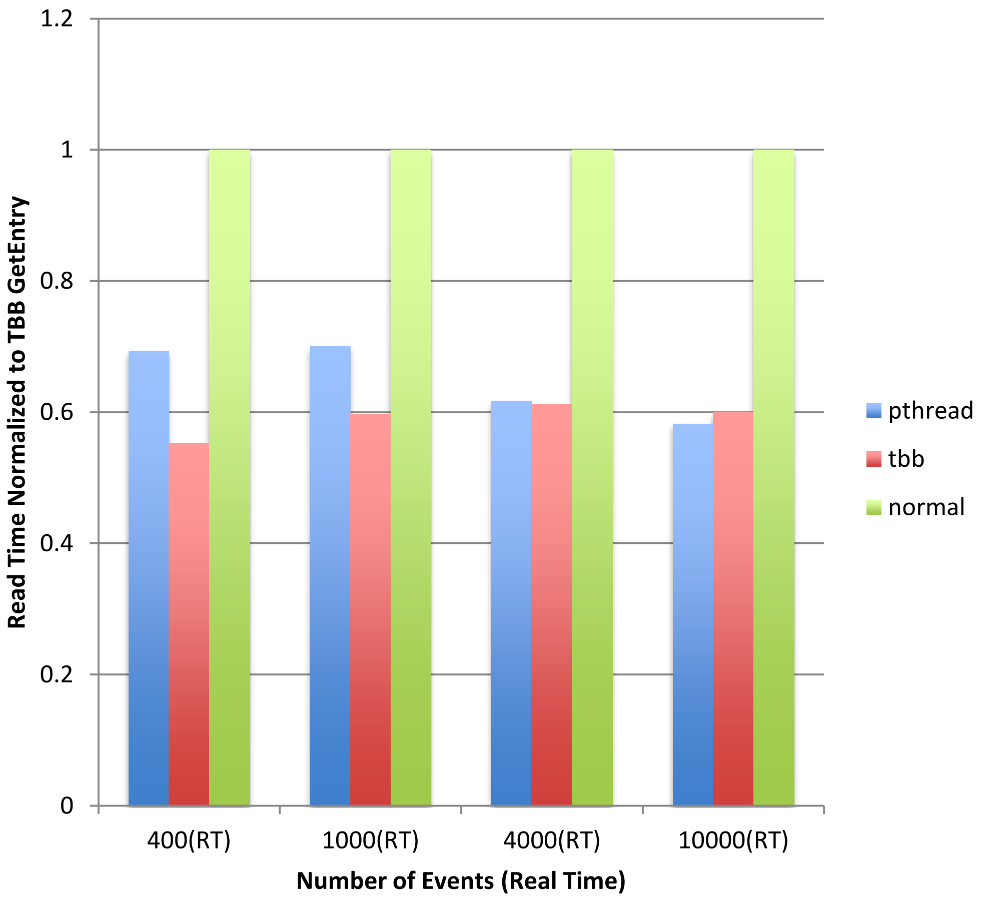
Request of reading the i^{th} entry in a tree

TTree::GetEntry(i)

Spawn and decompress

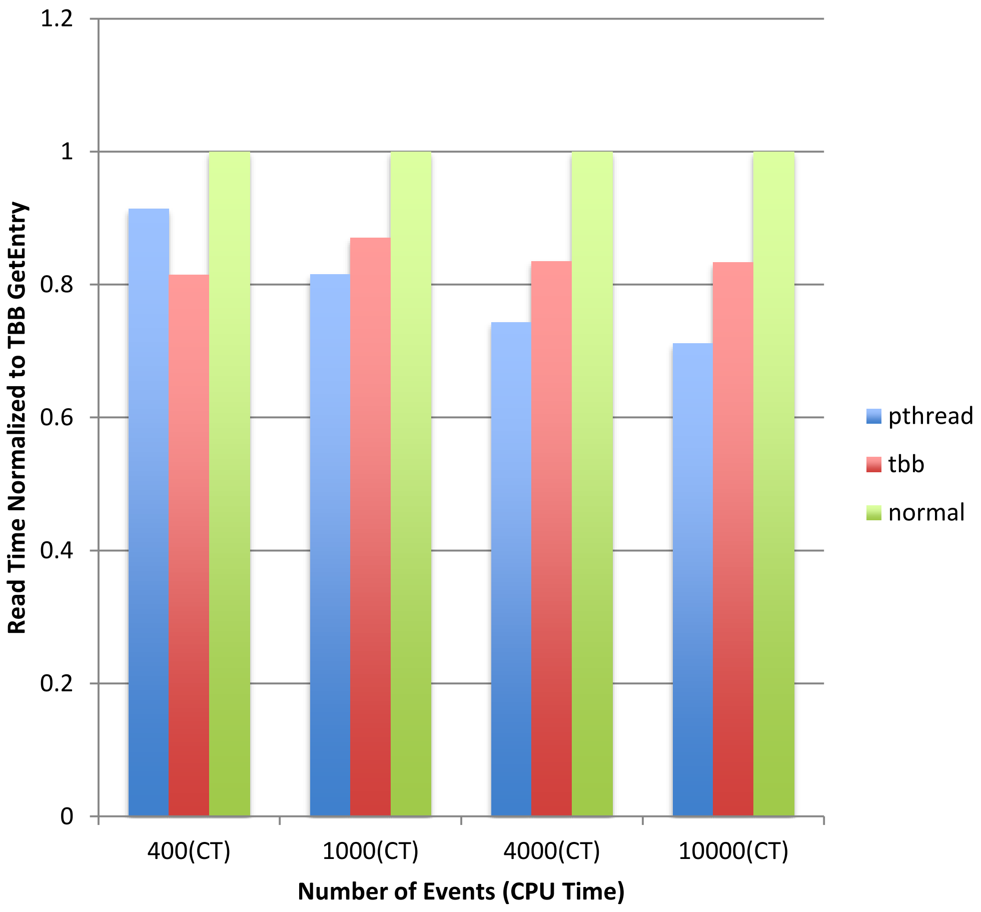
Tasks are eventually executed by threads in CPU cores

Read Time in Real Time



- TBB GetEntry() is the slowest.
- TBB Unzipping is faster than Pthread in most cases because it avoid mutex and locks.
- Pthread is faster when reading 10,000 events.

Read Time in CPU Time

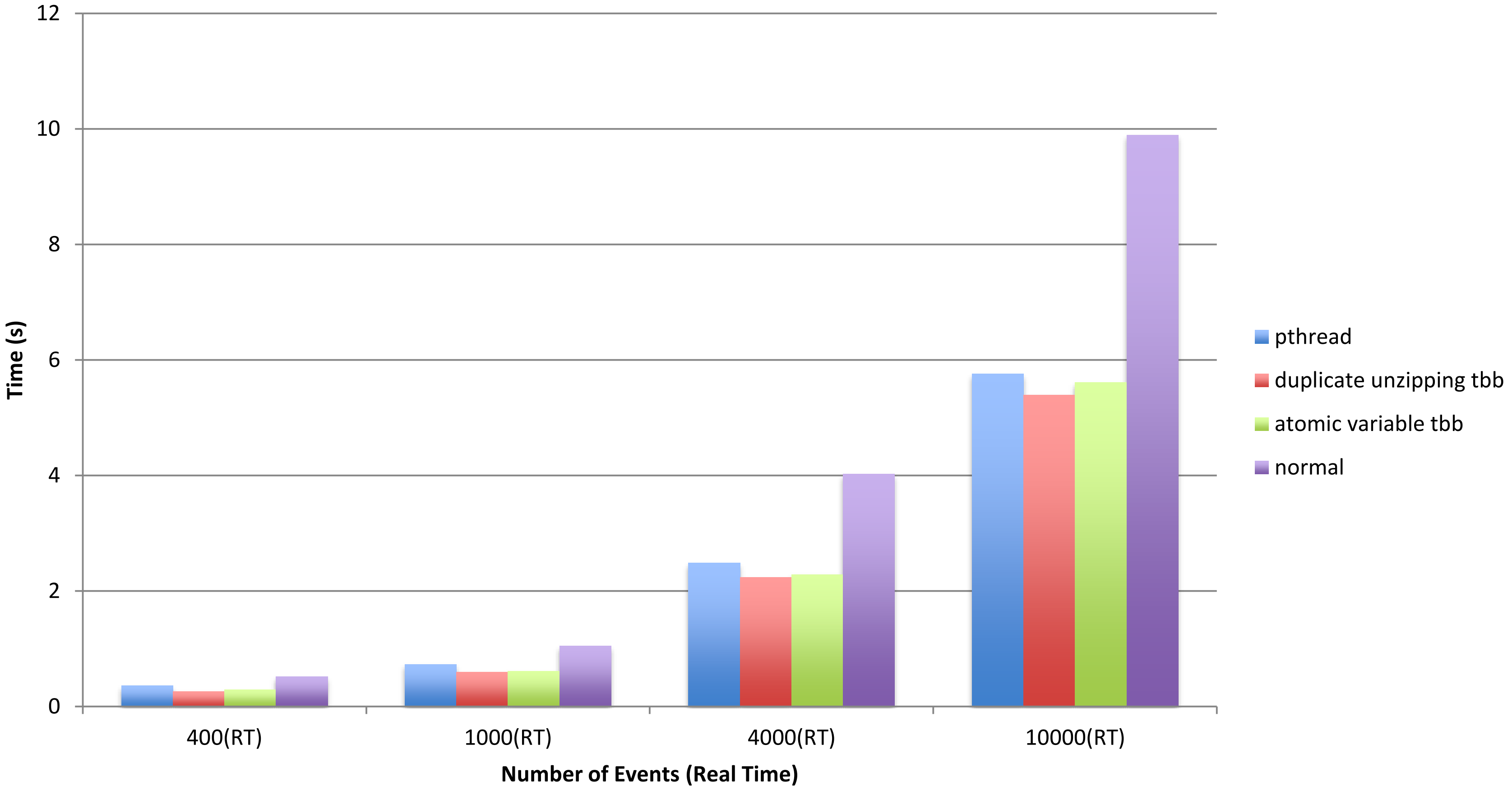


- TBB unzipping spends more time in CPU because of duplicate decompressing.

Two TBB Implementations

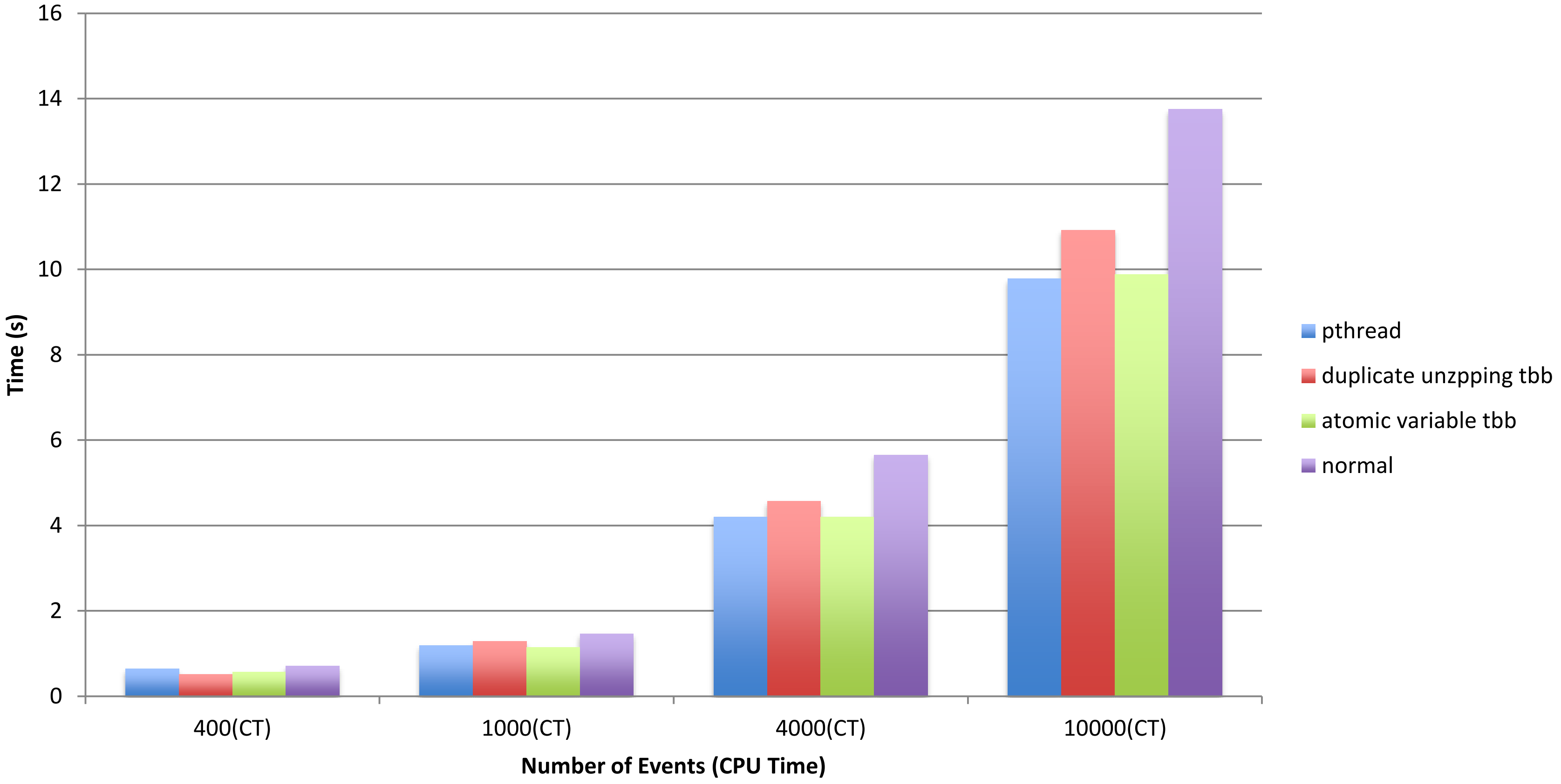
- Duplicate unzipping:
  - allowing duplicate unzipping same baskets when cache misses.
- Atomic variable:
  - Using atomic variables and compare and swap instructions to synchronize basket unzipping.

Read Time in Real Time



- TBB that allows duplicate unzipping runs faster than the implementation with atomic variables.

Read Time in CPU Time



- But duplicate unzipping costs more CPU cycles than the implementation with atomic variables.

Questions?