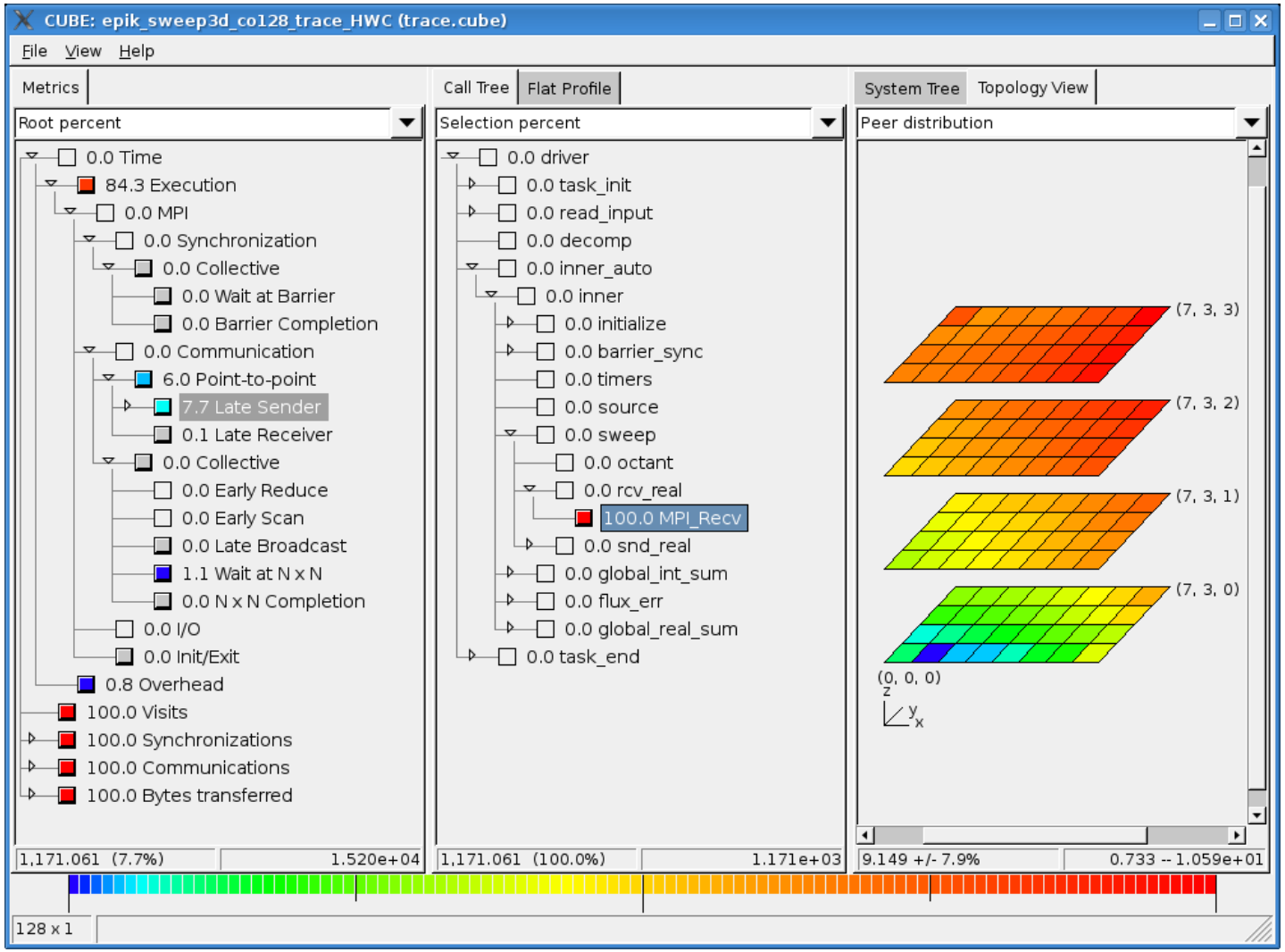
SCALASCA performance properties

"The metrics tour"

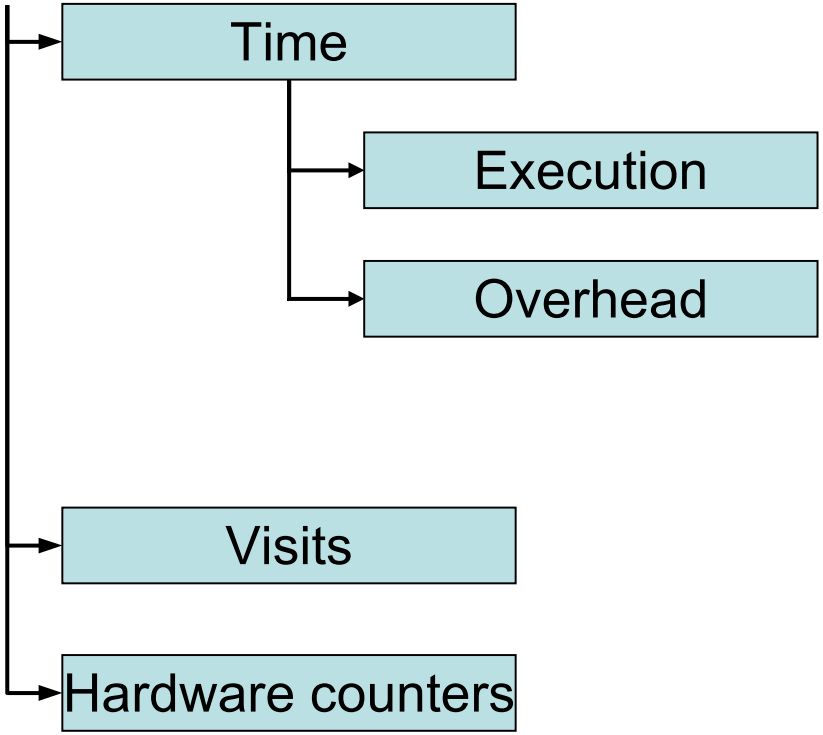
Markus Geimer
m.geimer@fz-juelich.de

SCALASCA analysis result



Generic metrics

Generic metrics



Total CPU allocation time

Execution time without overhead

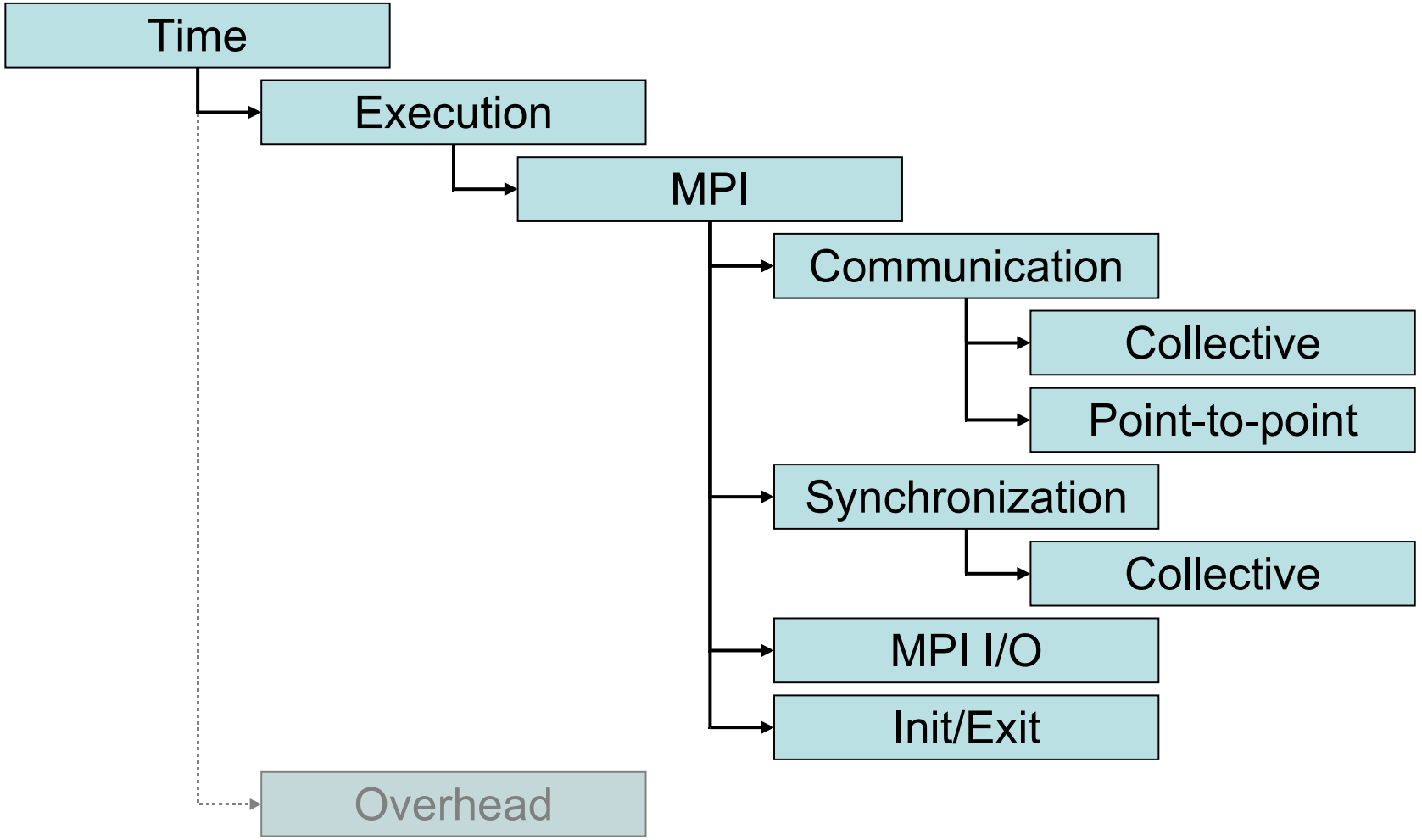
Time spent in tasks related to measurement (Does not include per-function perturbation!)

Number of times a function/region was executed

Aggregated counter values for each function/region

MPI-related metrics

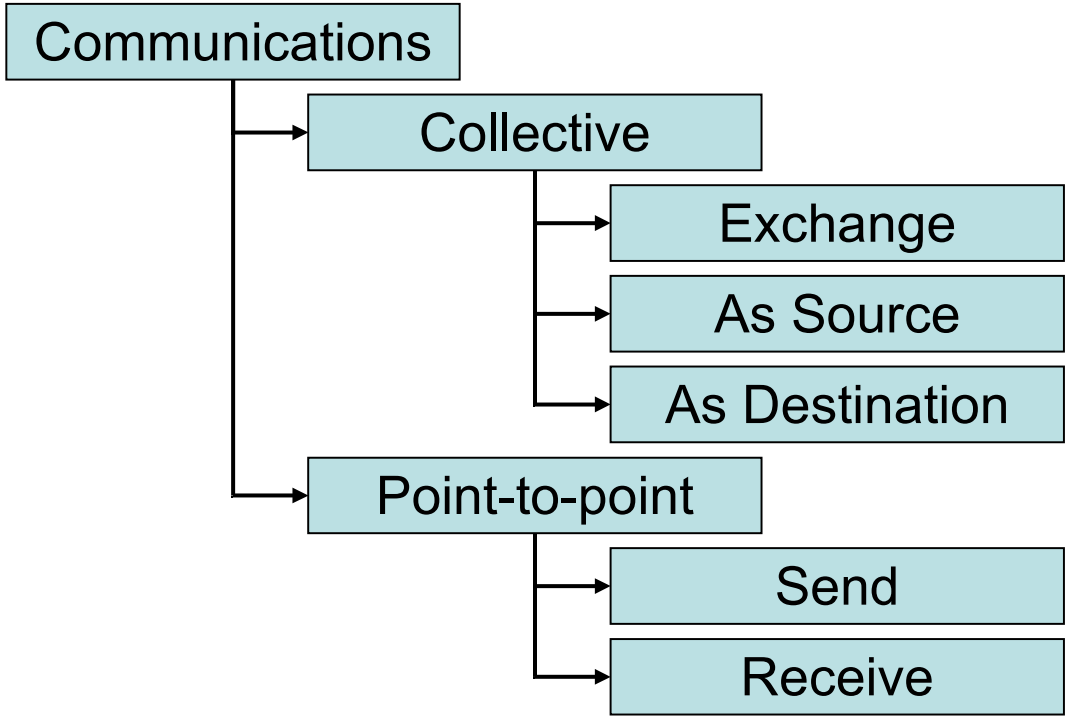
MPI Time hierarchy



MPI Time hierarchy details

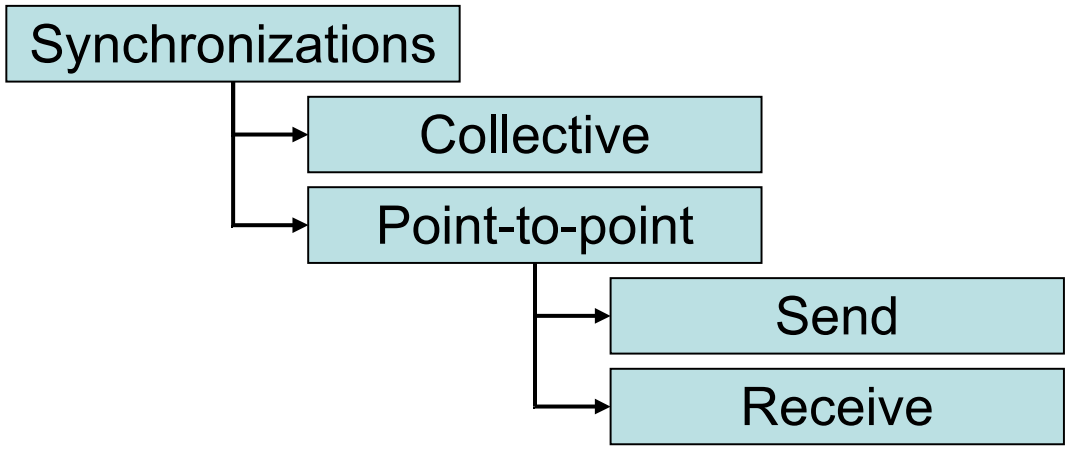
MPI	Time spent in pre-instrumented MPI functions
Communication	Time spent in MPI communication calls, subdivided into collective and point-to-point
Synchronization	Time spent in calls to `MPI_Barrier`
MPI I/O	Time spent in MPI I/O functions (not yet supported)
Init/Exit	Time spent in `MPI_Init` and `MPI_Finalize`

MPI Communications hierarchy



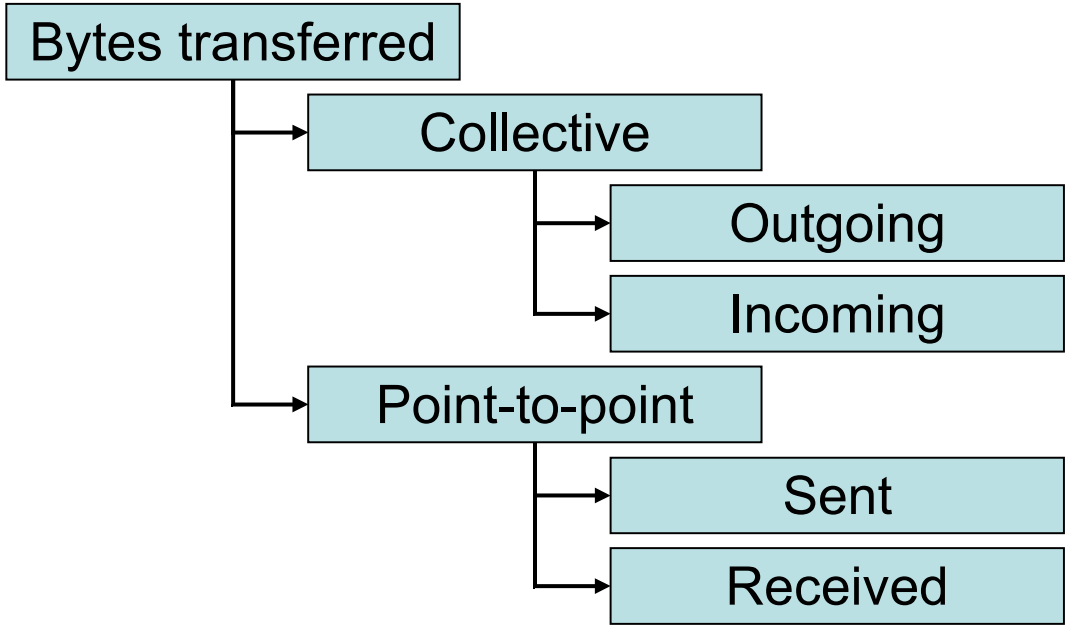
- Provides the number of calls to an MPI communication function of the corresponding class
- Zero-sized message transfers are considered synchronization!

MPI Synchronizations hierarchy



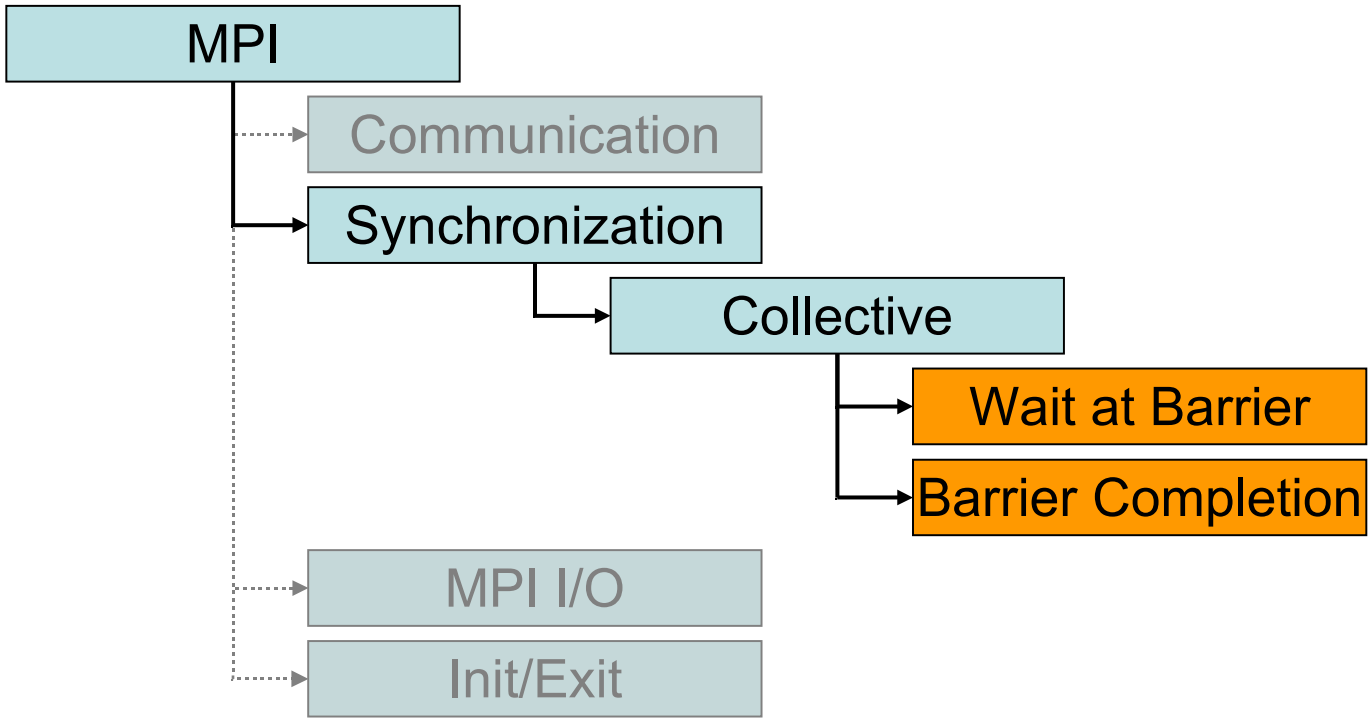
- Provides the number of calls to an MPI synchronization function of the corresponding class
- Synchronizations include zero-sized message transfers!

MPI Transfer hierarchy

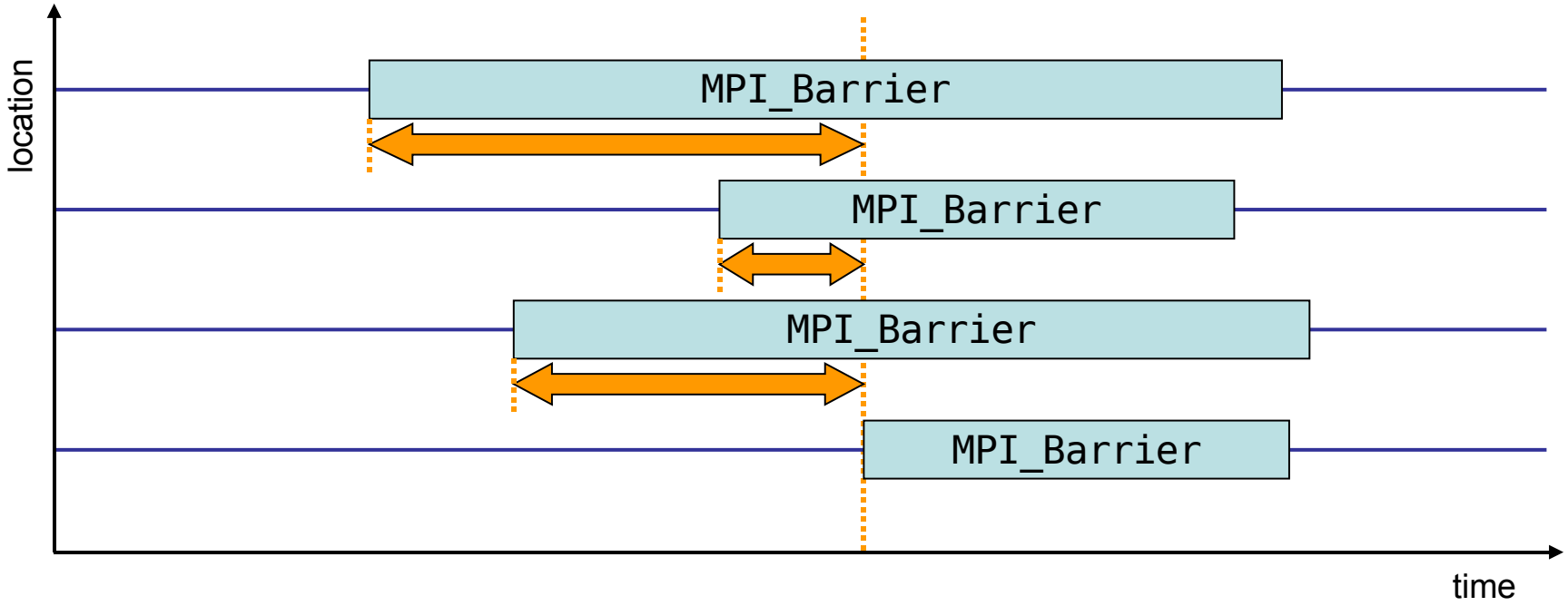


- Provides the number of bytes transferred by an MPI communication function of the corresponding class

MPI collective synchronization time

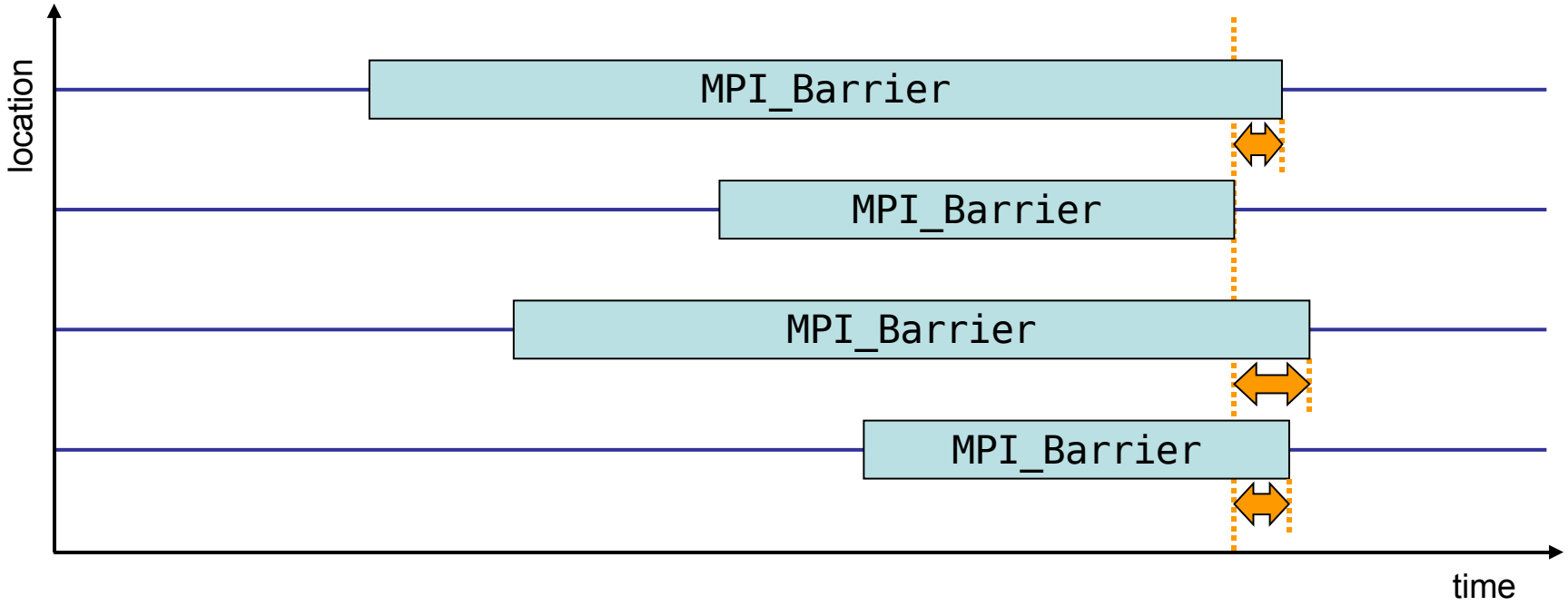


Wait at Barrier



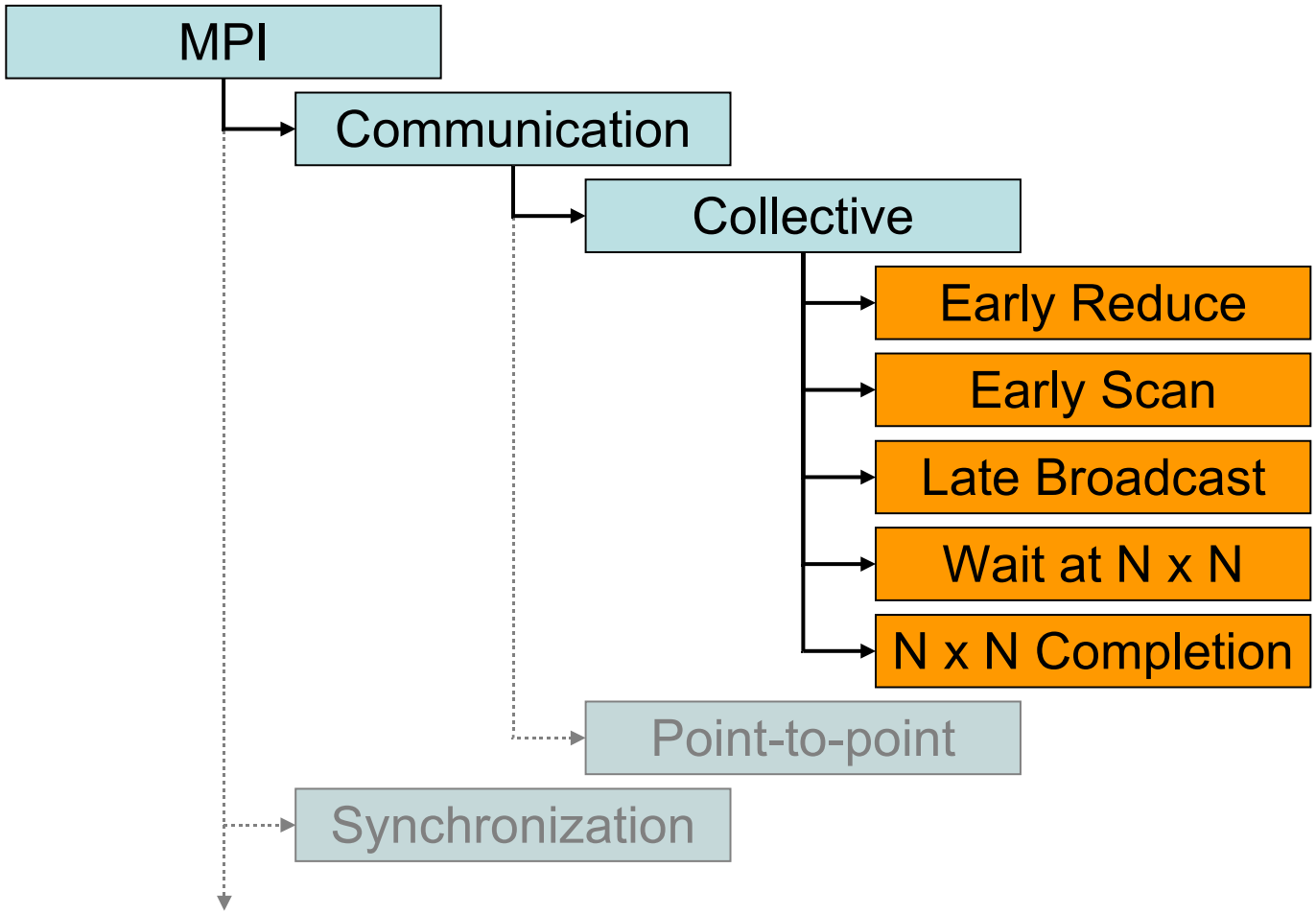
- Time spent waiting in front of a barrier call until the last process reaches the barrier operation
- Applies to: `MPI_Barrier`

Barrier Completion

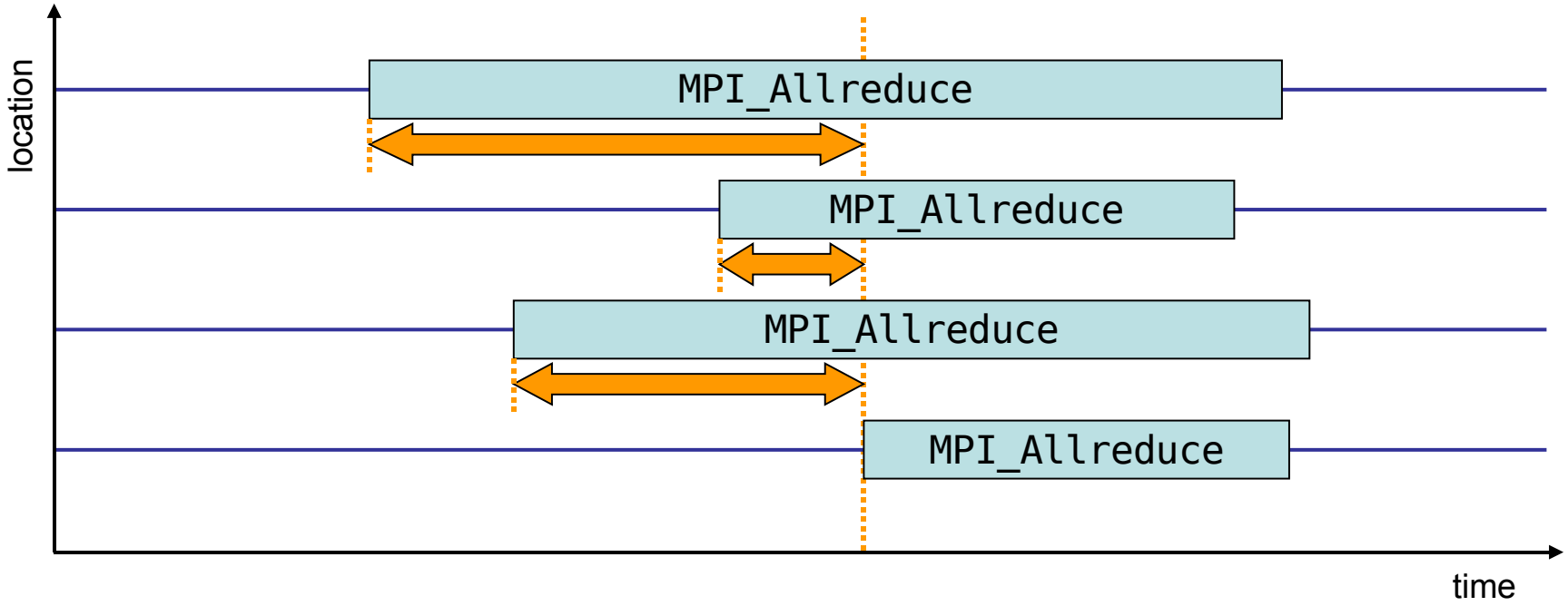


- Time spent in barrier after the first process has left the operation
- Applies to: `MPI_Barrier`

MPI collective communication time

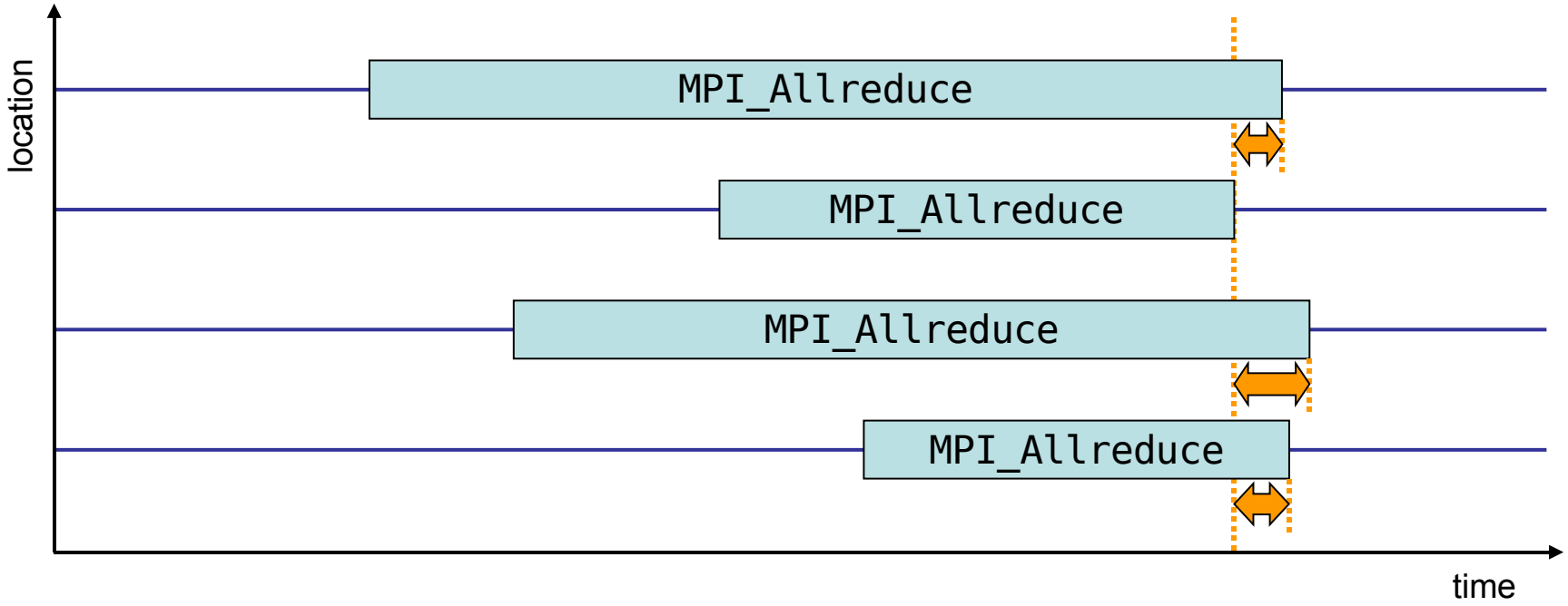


Wait at N x N



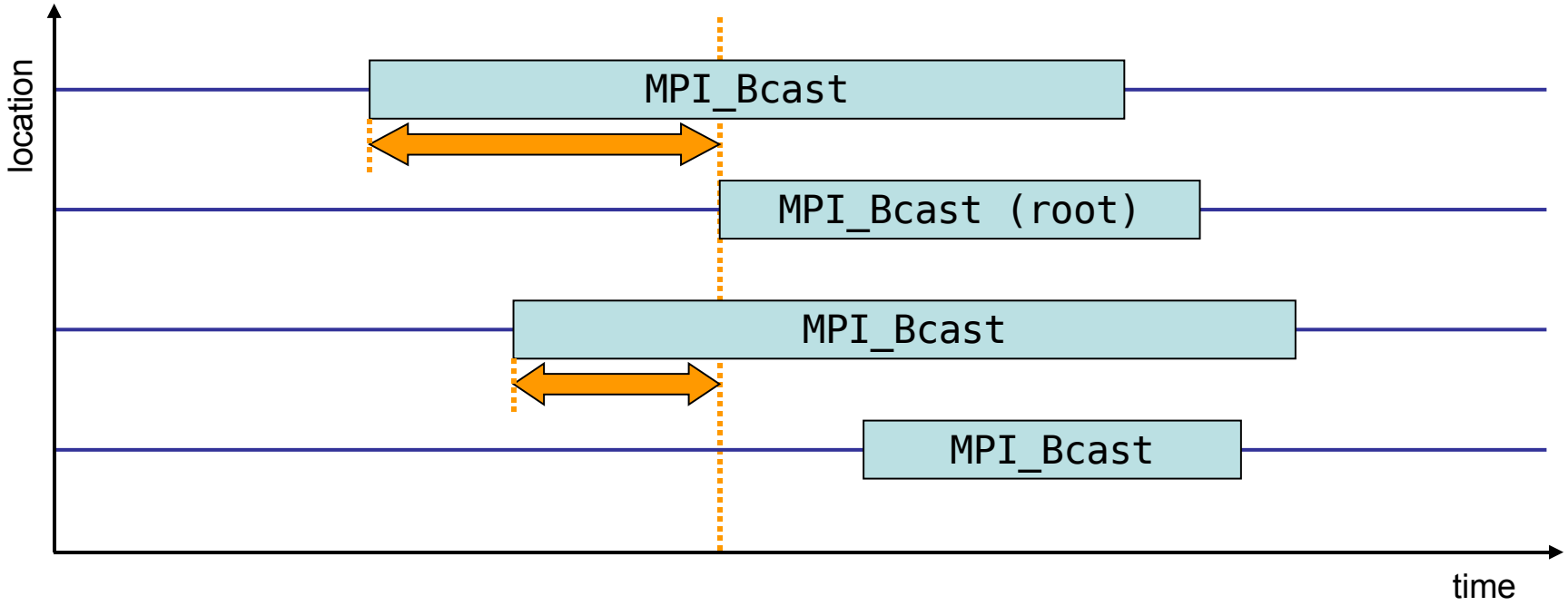
- Time spent waiting in front of a synchronizing collective operation call until the last process reaches the operation
- Applies to: `MPI_Allreduce`, `MPI_Alltoall`, `MPI_Alltoallv`, `MPI_Allgather`, `MPI_Allgatherv`, `MPI_Reduce_scatter`

N x N Completion



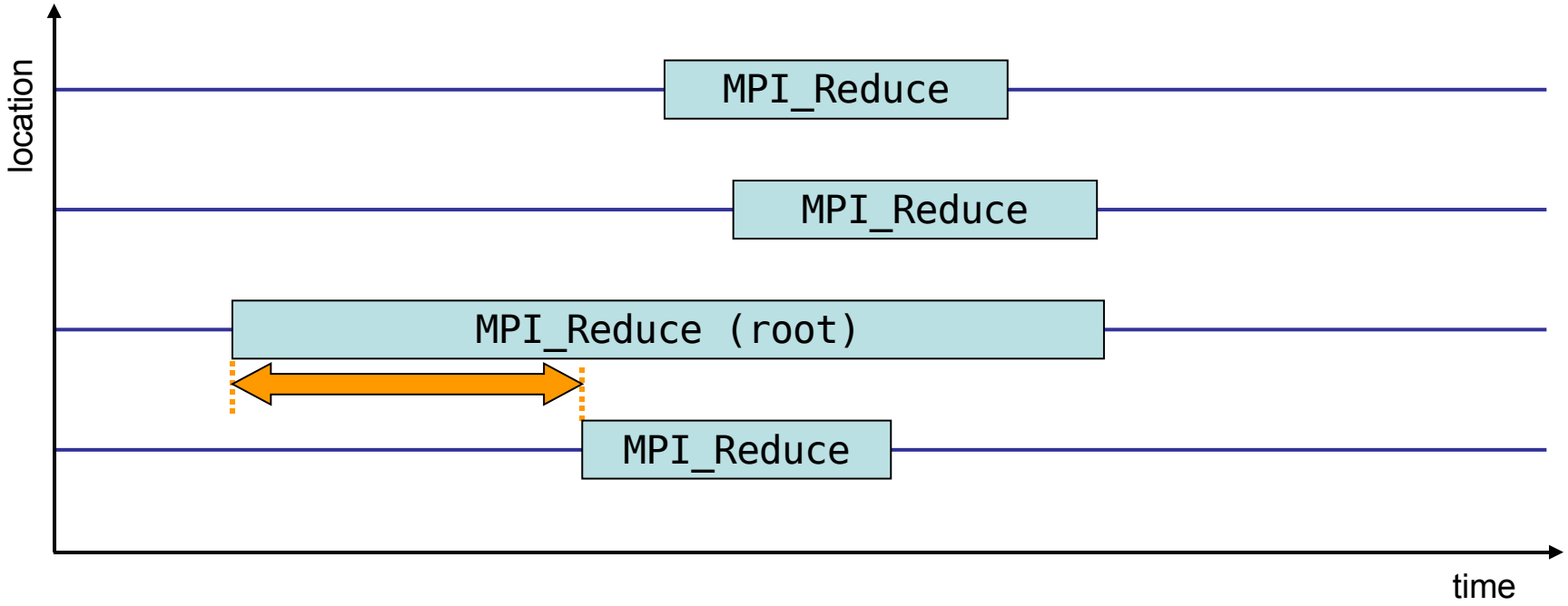
- Time spent in synchronizing collective operations after the first process has left the operation
- Applies to: `MPI_Allreduce`, `MPI_Alltoall`, `MPI_Alltoallv`, `MPI_Allgather`, `MPI_Allgatherv`, `MPI_Reduce_scatter`

Late Broadcast



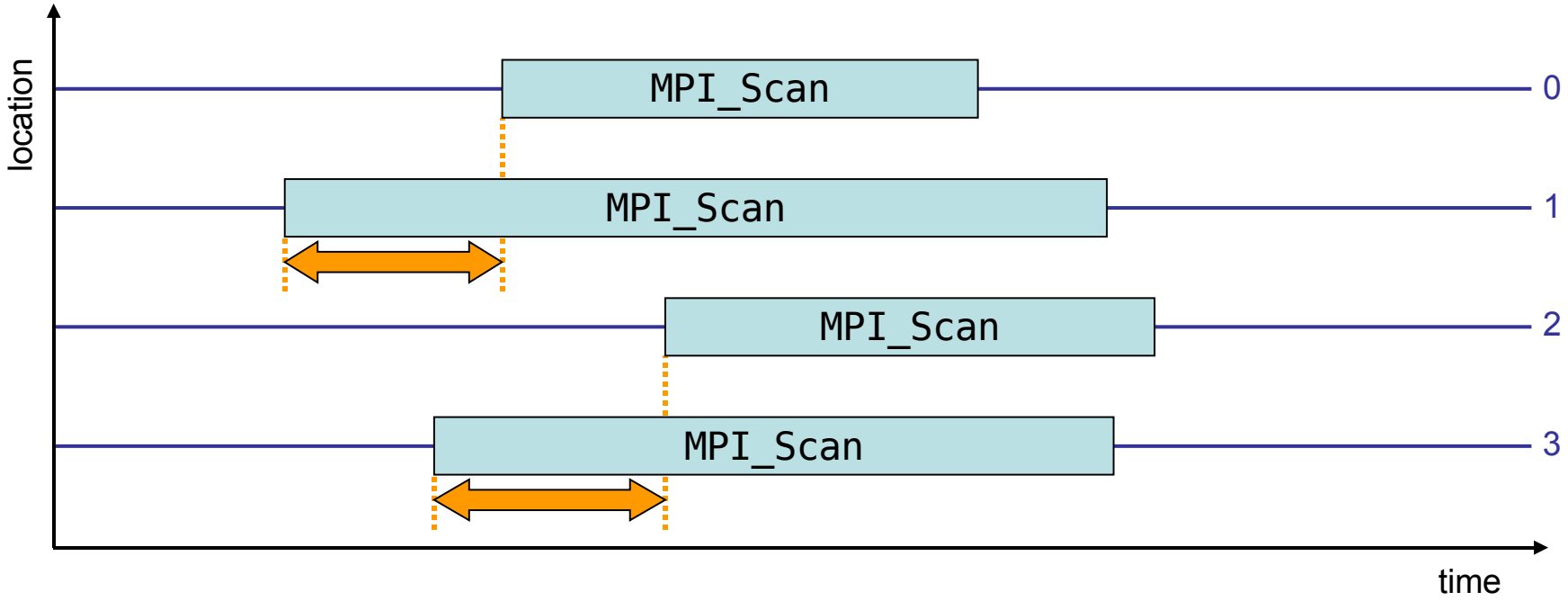
- Waiting times if the destination processes of a collective 1-to-N communication operation enter the operation earlier than the source process (root)
- Applies to: `MPI_Bcast`, `MPI_Scatter`, `MPI_Scatterv`

Early Reduce



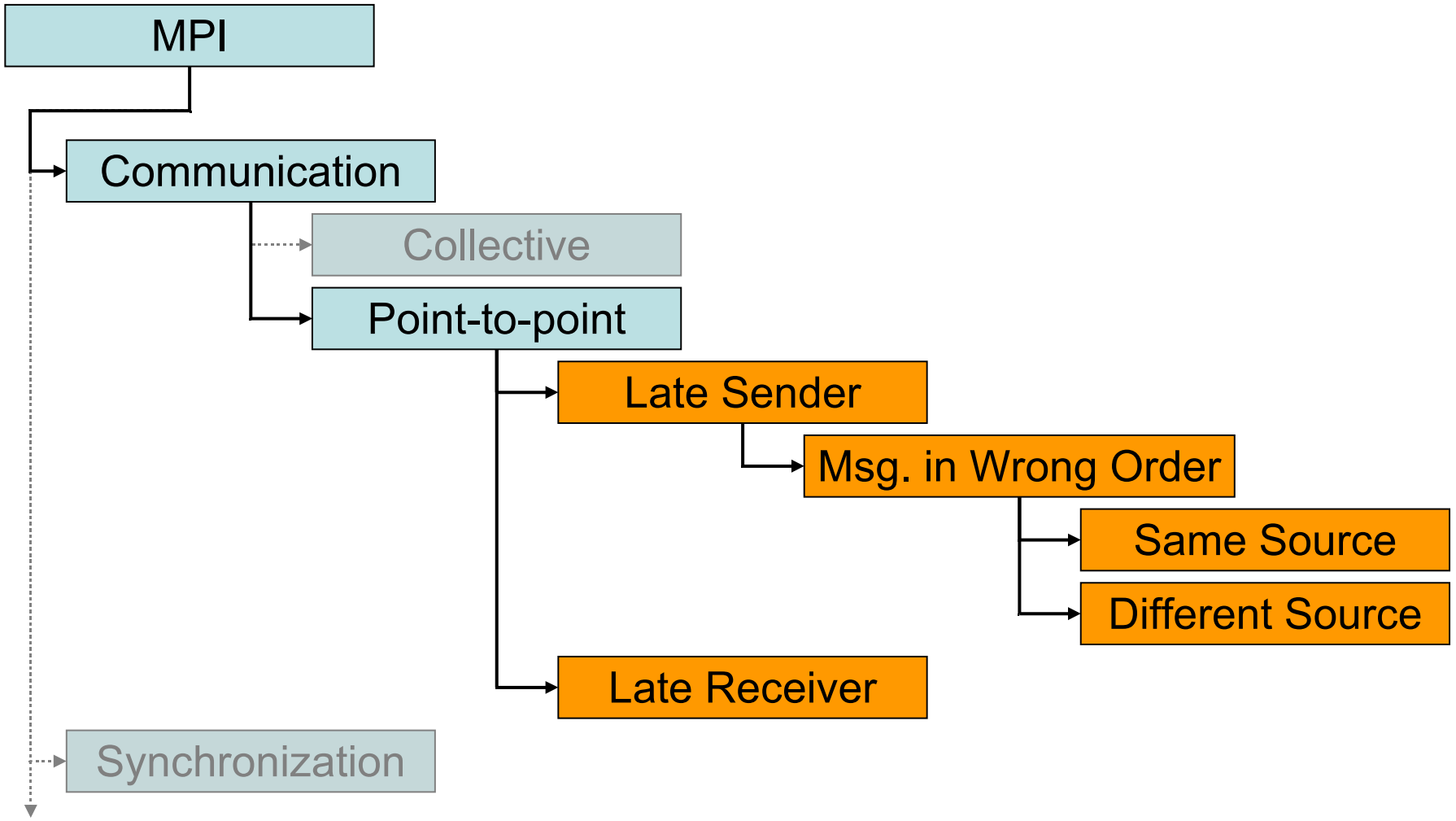
- Waiting time if the destination process (root) of a collective N-to-1 communication operation enters the operation earlier than its sending counterparts
- Applies to: `MPI_Reduce`, `MPI_Gather`, `MPI_Gatherv`

Early Scan

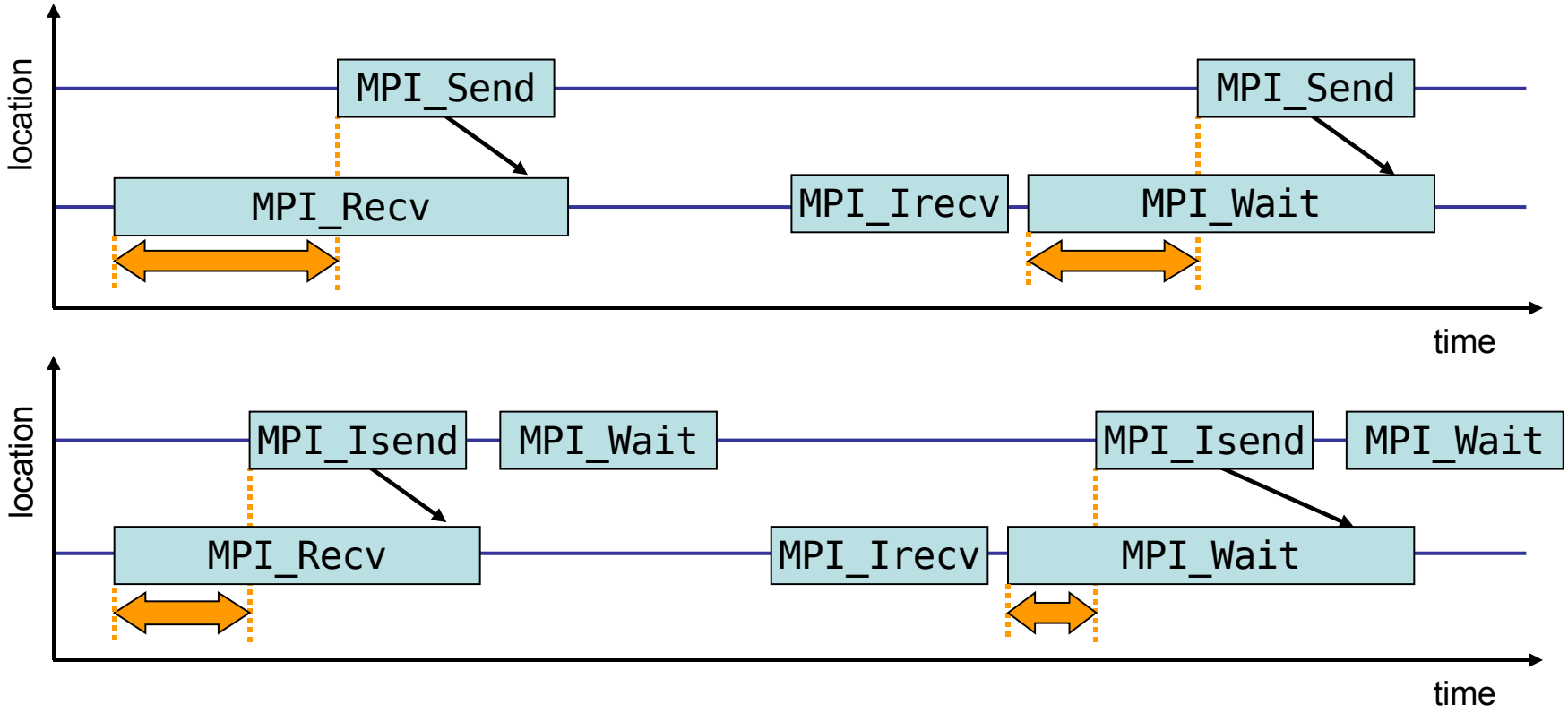


- Waiting time if process *n* enters a prefix reduction operation earlier than its sending counterparts (i.e., ranks 0..*n*-1)
- Applies to: `MPI_Scan`

MPI point-to-point communication time

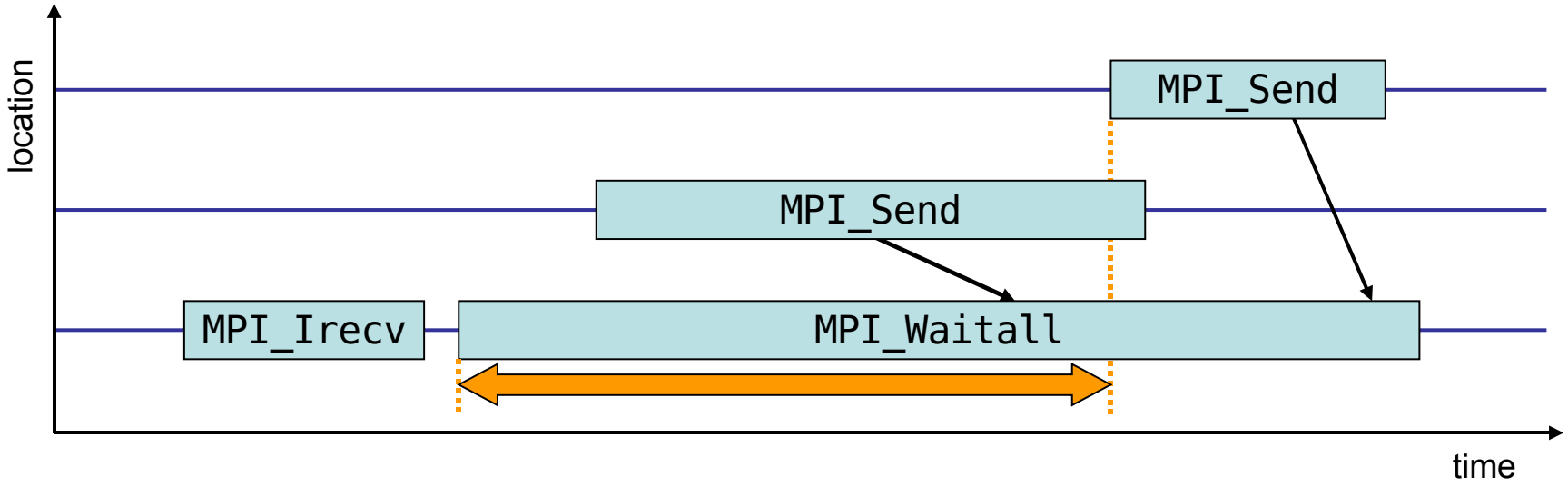


Late Sender



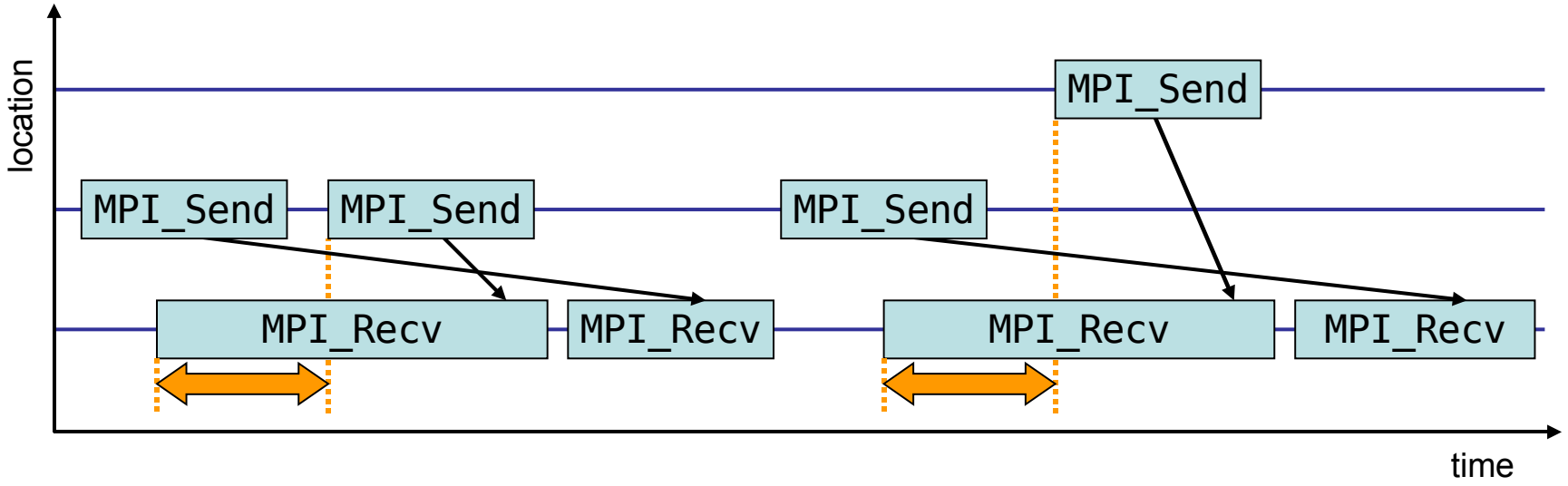
- Waiting time caused by a blocking receive operation posted earlier than the corresponding send operation
- Applies to blocking as well as non-blocking communication

Late Sender (II)



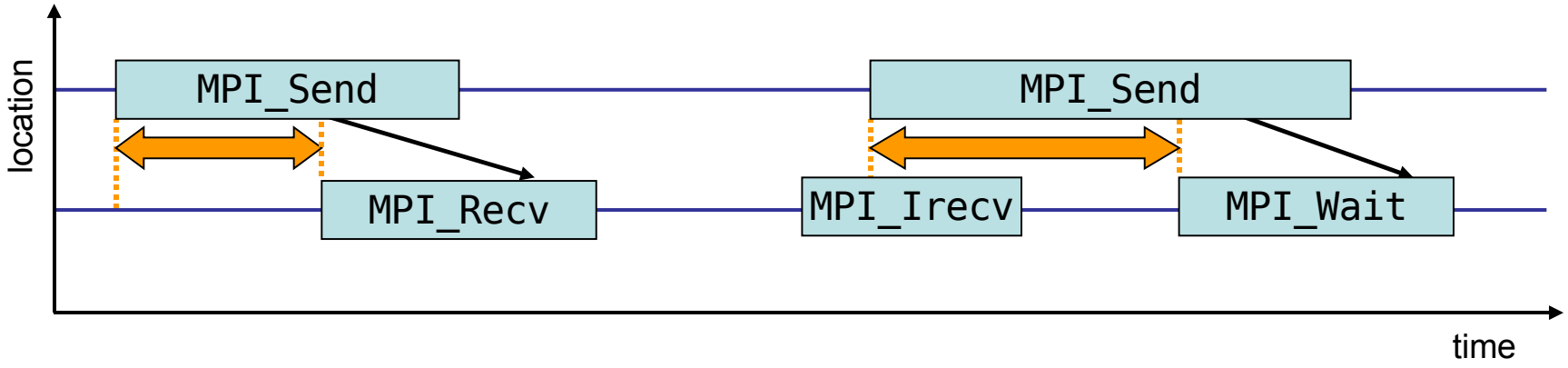
- While waiting for several messages, the maximum waiting time is accounted
- Applies to: `MPI_Waitall`, `MPI_Waitsome`

Late Sender, Messages in Wrong Order



- Refers to Late Sender situations which are caused by messages received in wrong order
- Comes in two flavours:
  - Messages sent from same source location
  - Messages sent from different source locations

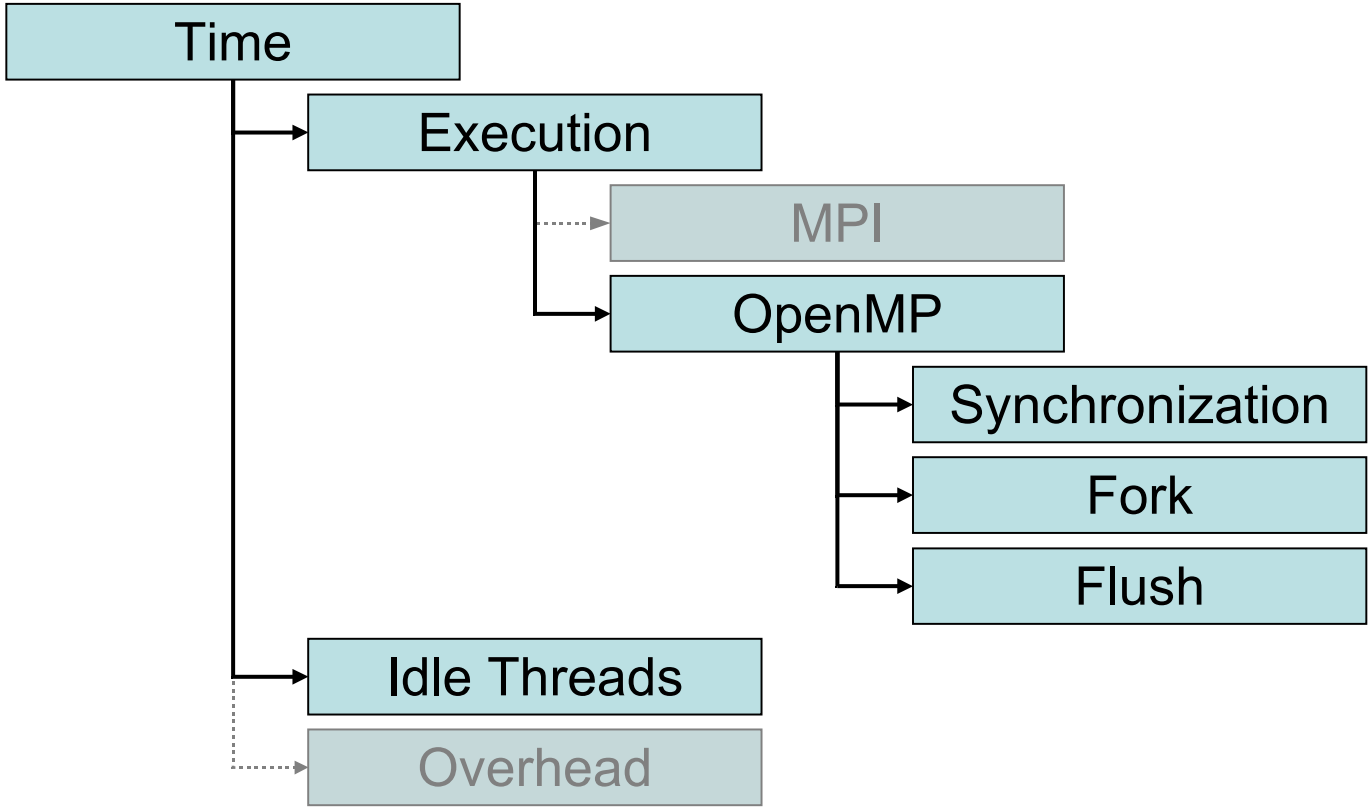
Late Receiver



- Waiting time caused by a blocking send operation posted earlier than the corresponding receive operation
- Calculated by receiver but waiting time attributed to sender
- Does currently not apply to non-blocking sends

OpenMP-related metrics

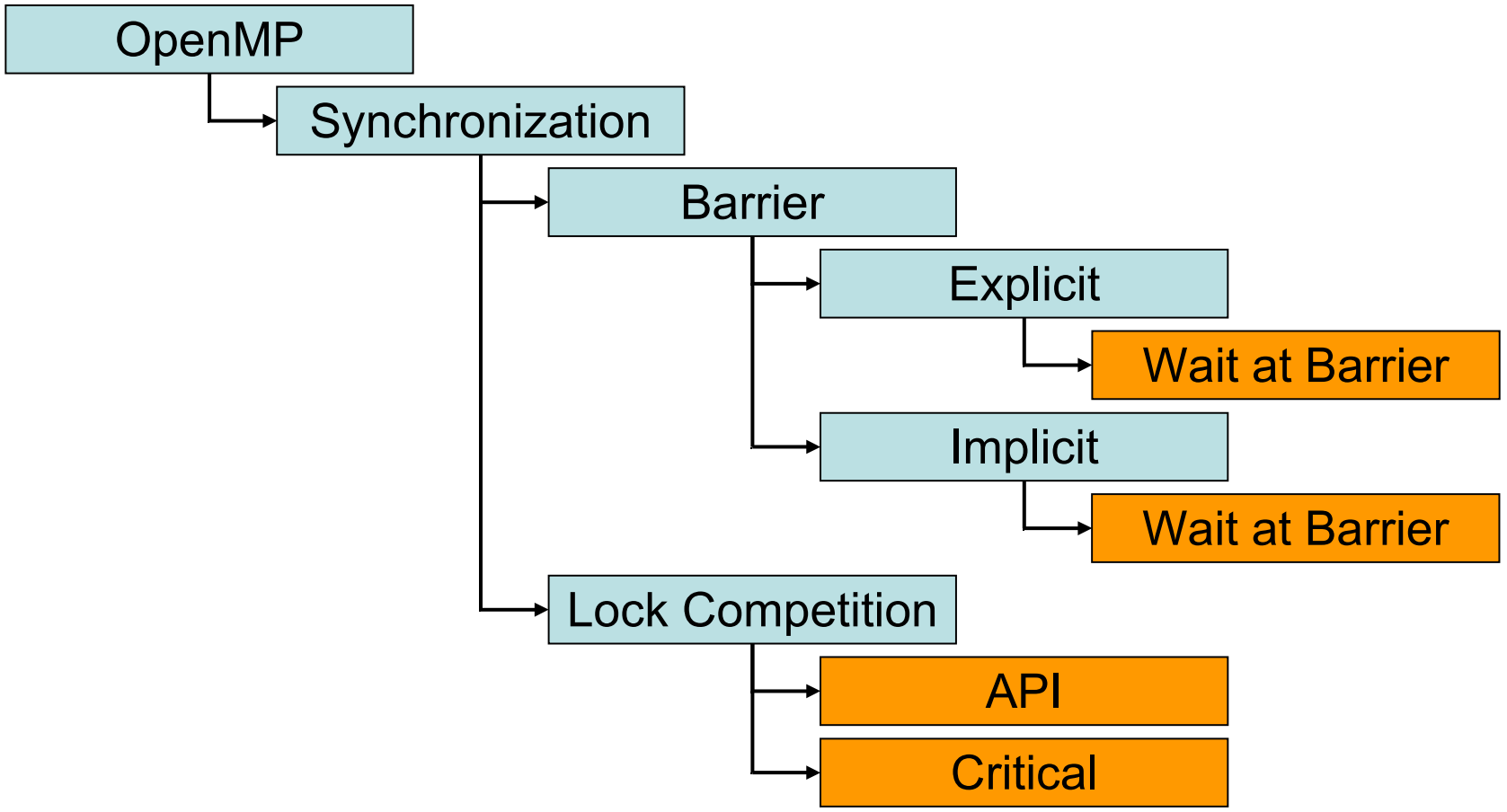
OpenMP Time hierarchy



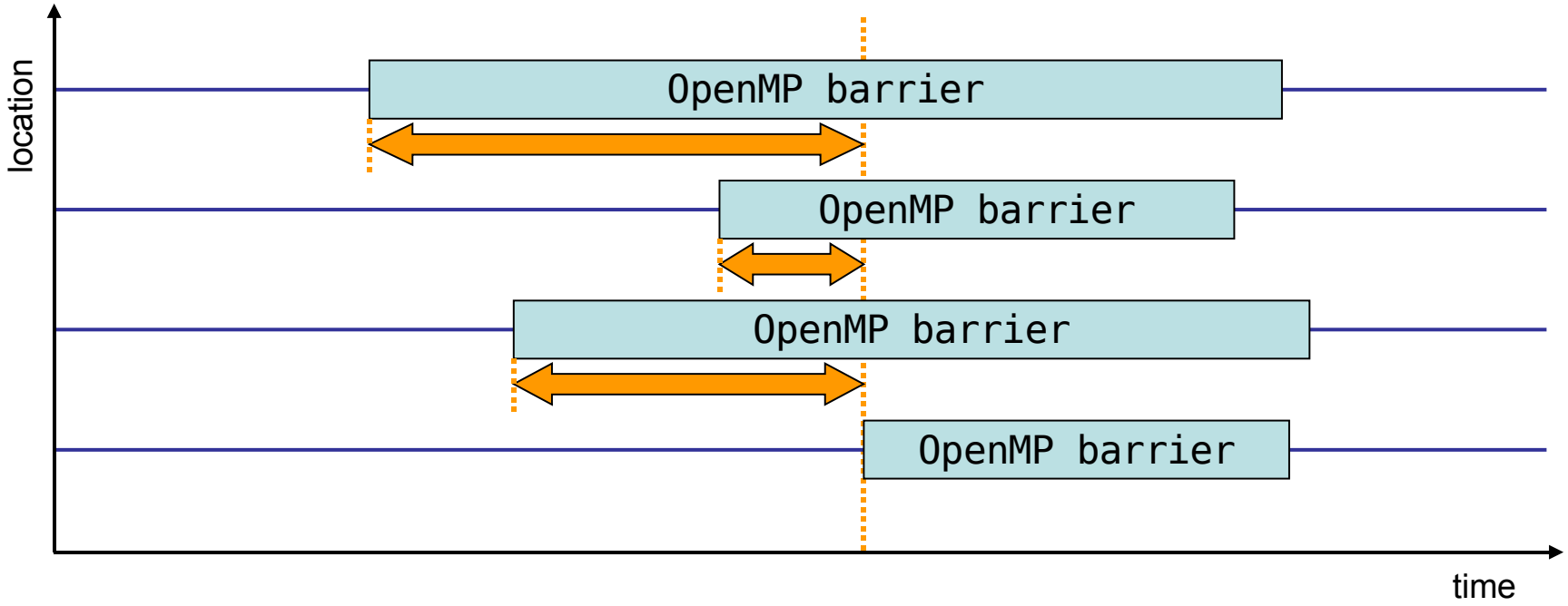
OpenMP Time hierarchy details

OpenMP	Time spent for OpenMP-related tasks
Synchronization	Time spent for synchronizing OpenMP threads
Fork	Time spent by master thread to create thread teams
Flush	Time spent in OpenMP flush directives
Idle Threads	Time spent idle on CPUs reserved for slave threads

OpenMP synchronization time

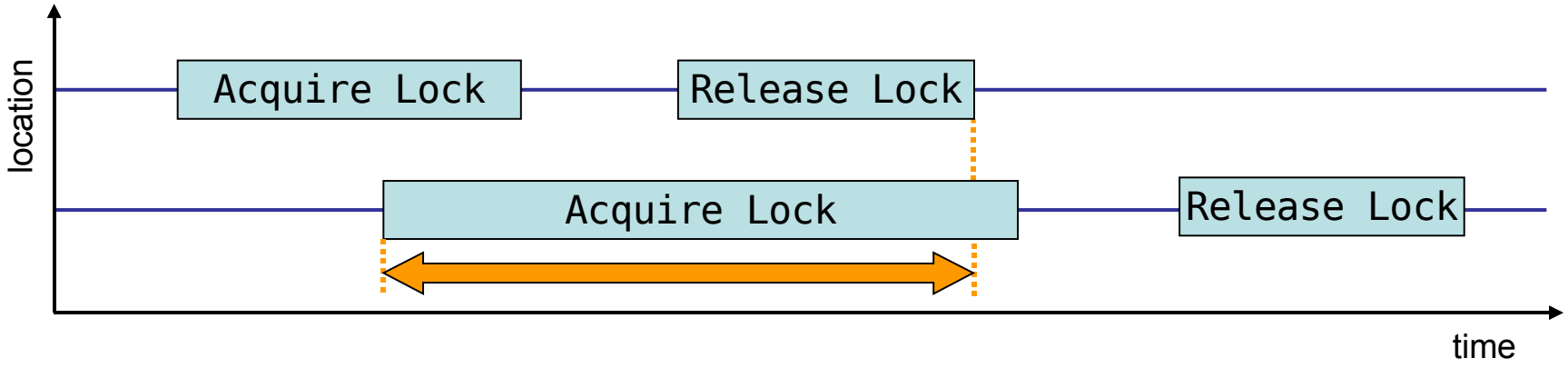


Wait at Barrier



- Time spent waiting in front of a barrier call until the last process reaches the barrier operation
- Applies to: Implicit/explicit barriers

Lock competition



- Time spent waiting for a lock that has been previously acquired by another thread
- Applies to: critical sections, OpenMP lock API