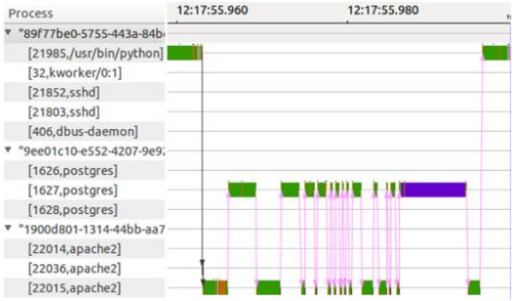
DISTRIBUTED COMPUTATION OF CRITICAL PATH

Pierre-Frédérick DENYS
Friday 11 June 2021

Agenda

- Introduction
- Actual architecture and challenges
- Proposed solution
- Related work : storage on disk (G. Bastien)
- Conclusion

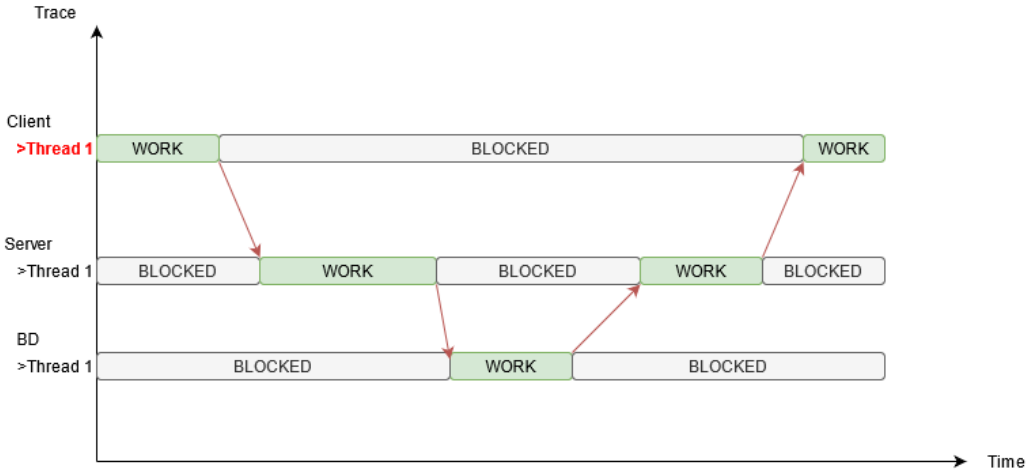
Introduction



- Need for large distributed systems tracing
- Critical path computation not optimized for this usage
- Critical path unavailable in Theia and Grafana plugin

Quick reminder about Critical Path

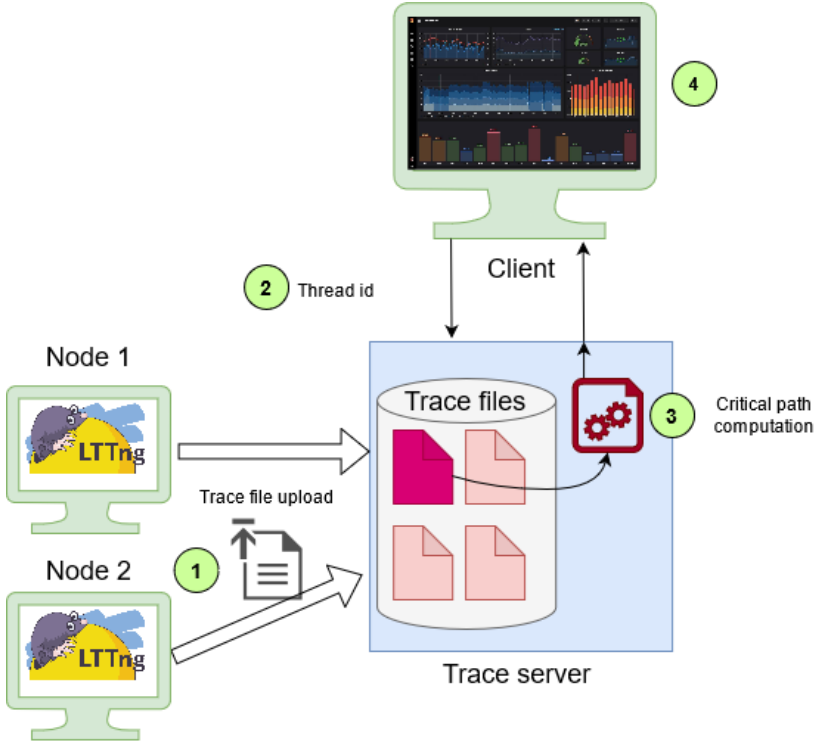
What is critical path



- The data structure as a two-dimensional doubly linked list, where horizontal edges are labeled with task states, and where vertical edges are signals between tasks (either a wake-up or a network packet)
- The active path of execution is the execution path where all blocking edges are substituted by their corresponding subtask

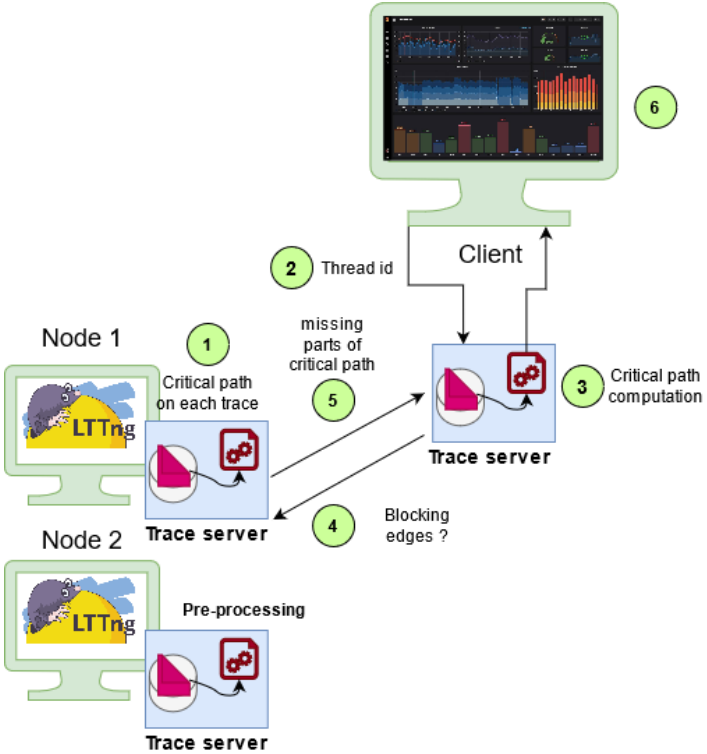
Part 1 : Actual architecture

Actual architecture in Trace Compass

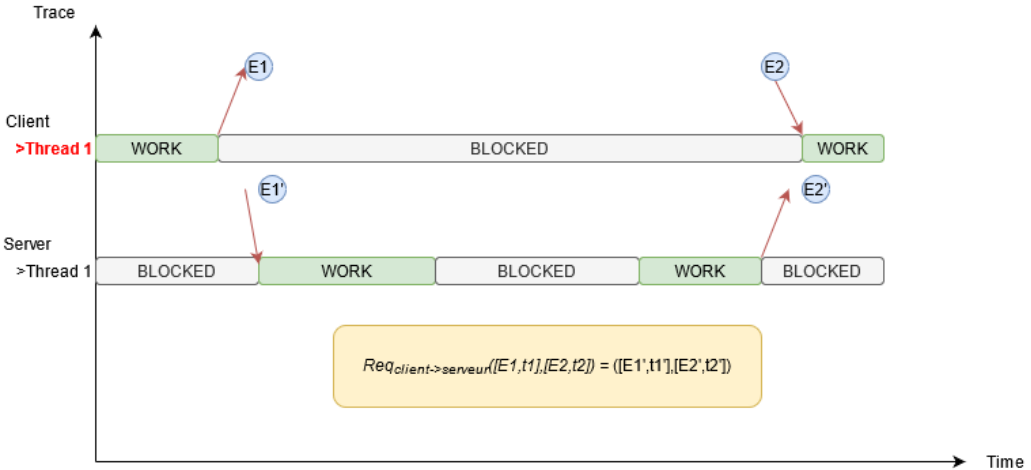


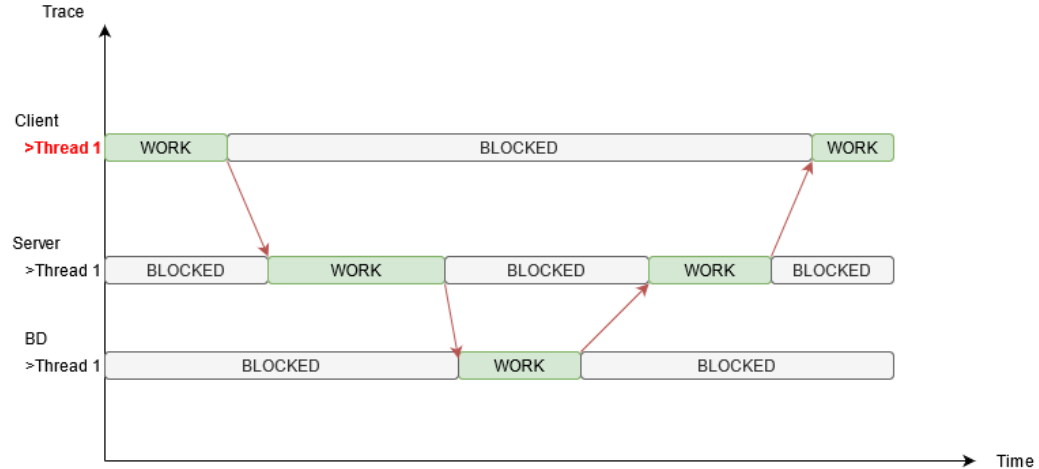
Part 2 : Parallelization of the architecture

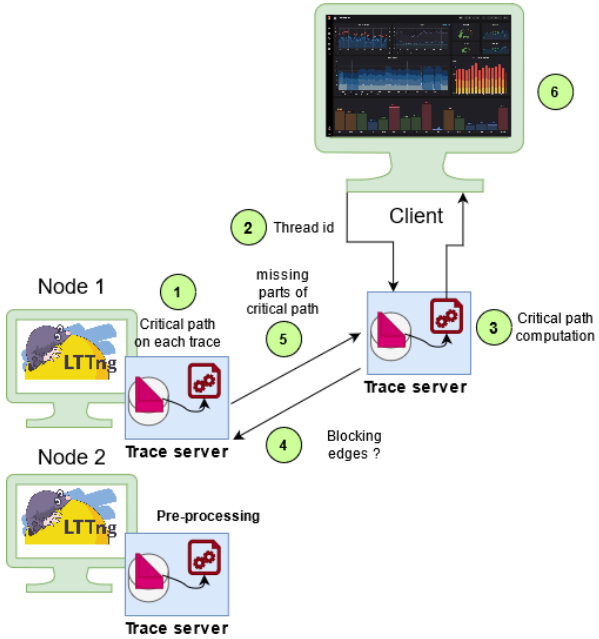
Parallelisation of the computation



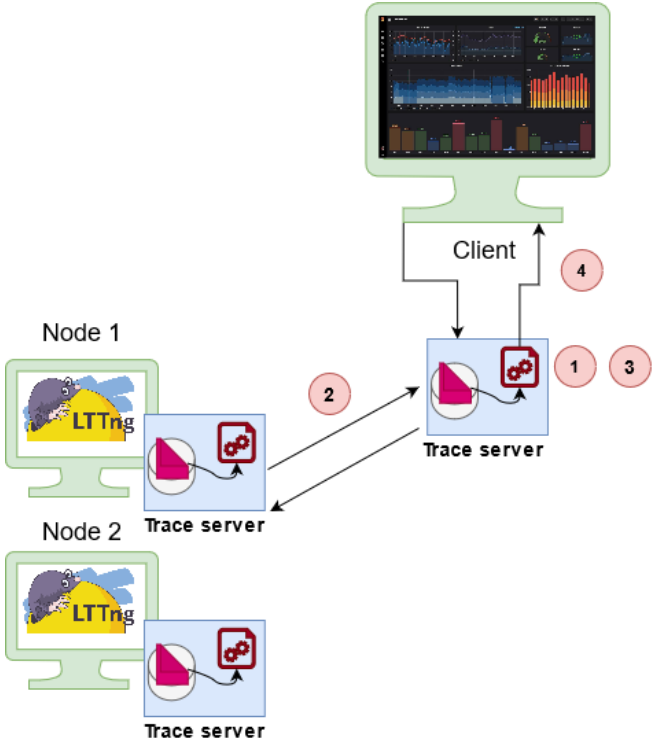
- Pre-processing of critical path on each node
- On client request, process the critical path of the trace, and ask only the missing parts of the path to other nodes
- Distributed processing, suitable for large number of nodes, less network load


Trace
E1
E2
Client
>Thread 1
WORK
BLOCKED
WORK
E1'
E2'
Server
>Thread 1
BLOCKED
WORK
BLOCKED
WORK
BLOCKED
Req client->serveur([E1,t1],[E2,t2]) = ([E1',t1'],[E2',t2'])
Time


Trace
Client
>Thread 1
WORK
BLOCKED
WORK
Server
>Thread 1
BLOCKED
WORK
BLOCKED
WORK
BLOCKED
BD
>Thread 1
BLOCKED
WORK
BLOCKED
Time


6
Client
2
Thread id
Node 1
1
Critical path on each trace
missing parts of critical path
5
3
Critical path computation
LTTng
Trace server
4
Blocking edges ?
Node 2
LTTng
Pre-processing
Trace server

Challenges



1. Current algorithm not suitable for parallel computation
2. Need a protocol to send and receive critical path elements between trace servers
3. Critical path elements are currently stored on memory, unlike traces and state history databases which reside on disk
4. Trace server protocol is not suitable for critical path, not possible to display it in Theia and Grafana

My work

Current :

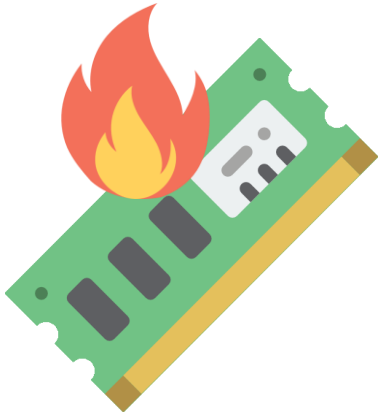
- Simplified case : improve algorithm to :
- ✅ Compute critical path of each trace(on a same trace server) independently
- ✅ Be able to process the full critical path from each part

Future :

- Define a communication protocol for critical path elements between nodes on the network
- Improve the Trace Server Protocol to access the critical path on Theia and Grafana

Part 3 : Storage of critical path on disk

Related work



- For large traces, OutOfMemoryException when building execution graph
  - Store the graph and critical path on disk
- First draft implementation : naive and simple, using available structures : state system for horizontal edges, segment store for vertical.
- Work done one by Geneviève Bastien

Conclusion

- Parallelisation of critical path computation
  - Needed for distributed systems traces analysis
- Integration of Critical path in Trace Server Protocol (for Theia and Grafana viewers)


Thank you for
listening !