Evaluating Trace Aggregation for Performance Visualization of Large Distributed Systems

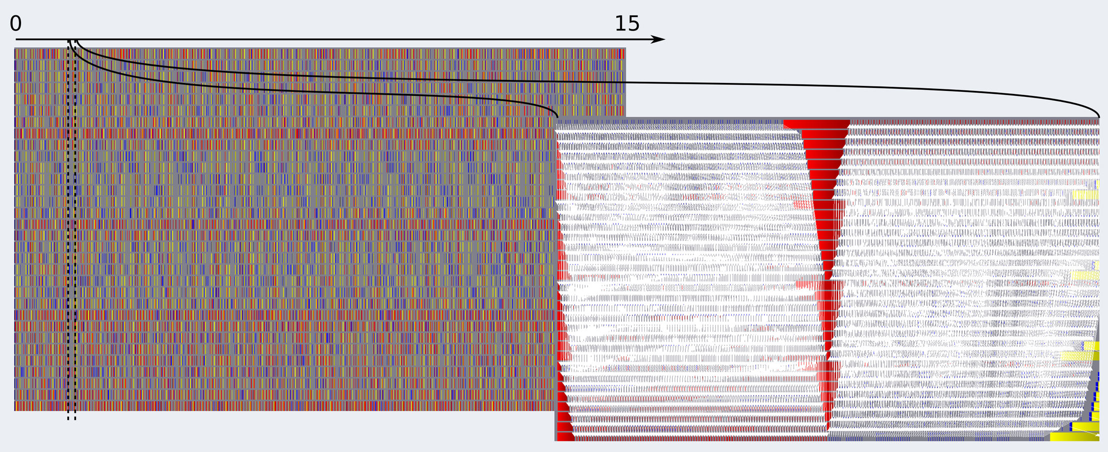
Robin Lamarche-Perrin, Lucas Mello Schnorr, *Jean-Marc Vincent*, and Yves Demazeau

Grenoble Informatics Laboratory

Large Scale HPC Application Traces Analysis Problem

- Performance analysis through visualization techniques usually suffers from severe semantic limitations due to the size of parallel applications (visualization of too much data).
- Macroscopic visualization (with an applicative semantic) are either inexistent, or ad hoc or rely on inadequate viewers' graphical capabilities (anti-aliasing).
- Performance anomalies are usually located on a small portion of the visualization and hard to isolate.
- Our objective:

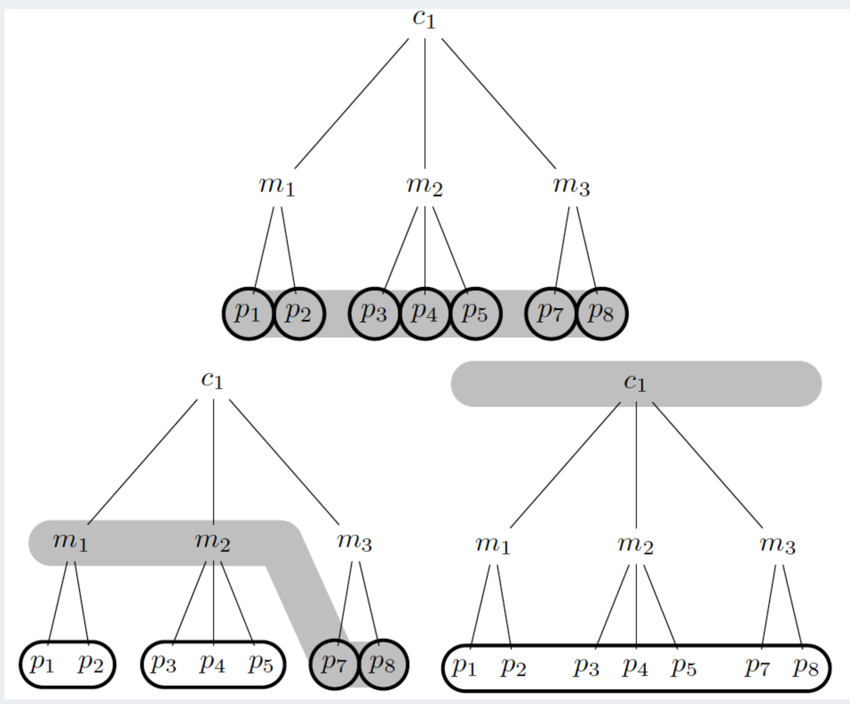
Evaluate the quality of aggregated representations – using measures from information theory – to build consistent multiresolution representations of large execution traces.



Typically cryptic Gantt representation of a LU factorisation on a standard 32 nodes cluster

Information Theory Based Approach : Information Loss vs Complexity Reduction

Multi-level representation



Possible aggregations in a hierarchy : disaggregated, partially or completely aggregated

Which one is the more meaningful ?

Parametrized Information Criterion

$$\mathrm{pIC}(\mathcal{A}) = p \times \mathrm{gain}(\mathcal{A}) - (1-p) \times \mathrm{loss}(\mathcal{A})$$

Shannon **entropy** measure of complexity

$$\mathrm{gain}(A) = (v(A)\log_2 v(A)) - \sum_{e \in A} (v(e)\log_2 v(e))$$

Kullback-Leibler **divergence** estimates the information loss

$$\mathrm{loss}(A) = \sum_{e \in A} v(e) \times \log_2 \left(\frac{v(e)}{v(A)} \times |A|\right)$$

- The *sum property* of quality measures enables independent computation of aggregates
- Hierarchical organization allows a recursive evaluation of branches
- Efficiently implemented through dynamic programming

Optimal aggregation algorithm

Require: A tree $\mathcal{T}$ with $Gain$ and loss labels on nodes.
Require: A gain/loss ratio parameter $p \in [0,1]$.
Ensure: An aggregation within $\mathcal{T}$ that maximizes pIC.

procedure FINDBESTAGGREGATION($\mathcal{T}$,p)
  $rootAgg \leftarrow \{root\}$
  if $\mathcal{T}$ is a leaf **then return** $rootAgg$
  for all S direct subtree of $\mathcal{T}$ **do**
    $aux \leftarrow$ FINDBESTAGGREGATION(S, p)
    $childAgg \leftarrow$ UNION$(childAgg, aux)$
  if pIC of $rootAgg > pIC$ of $childAgg$ **then return** $rootAgg$
  else return $childAgg$

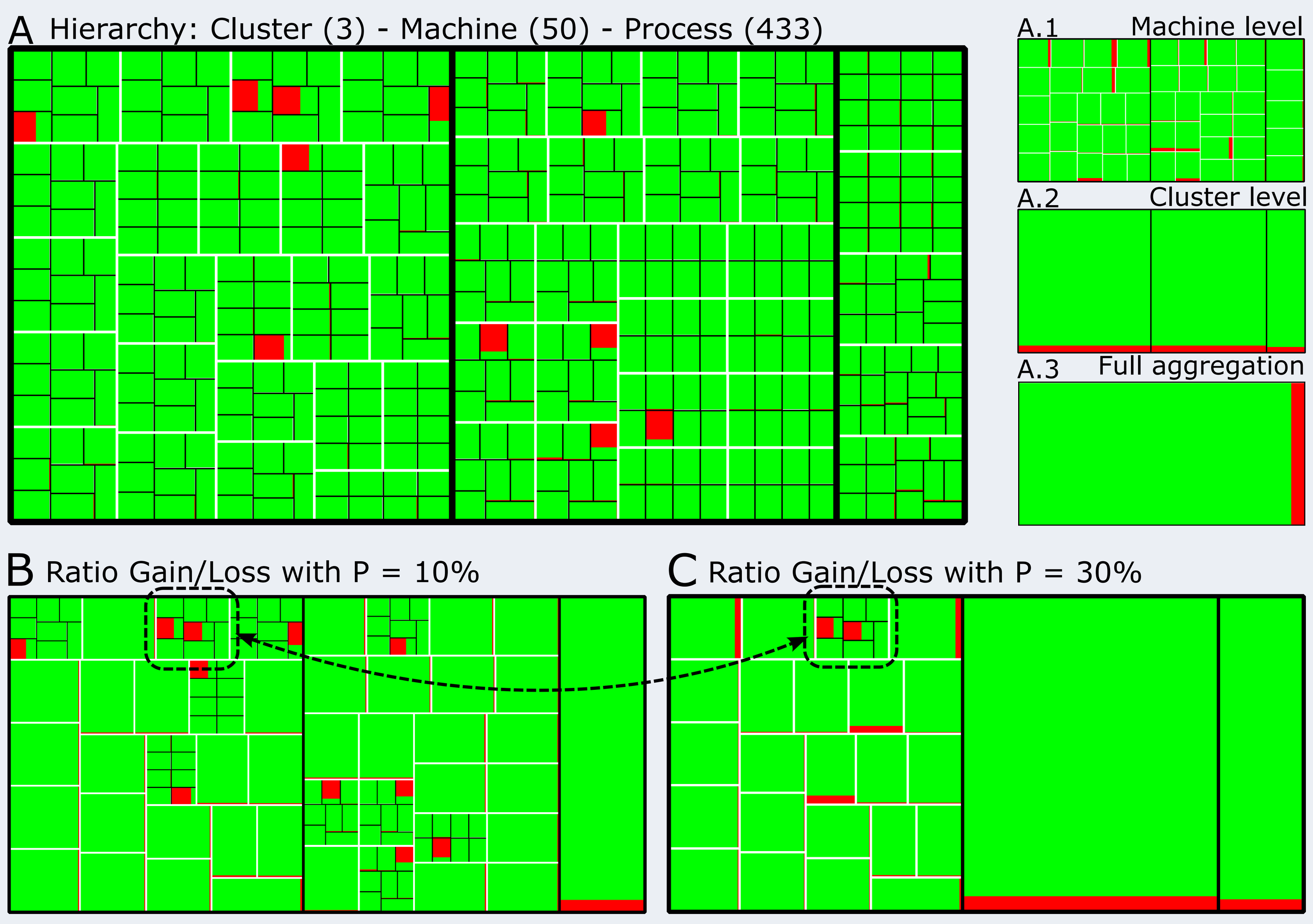
- Classical tree traversal

$\Rightarrow$ **Time complexity** $\mathcal{O}(n)$

Work Stealing Evaluation

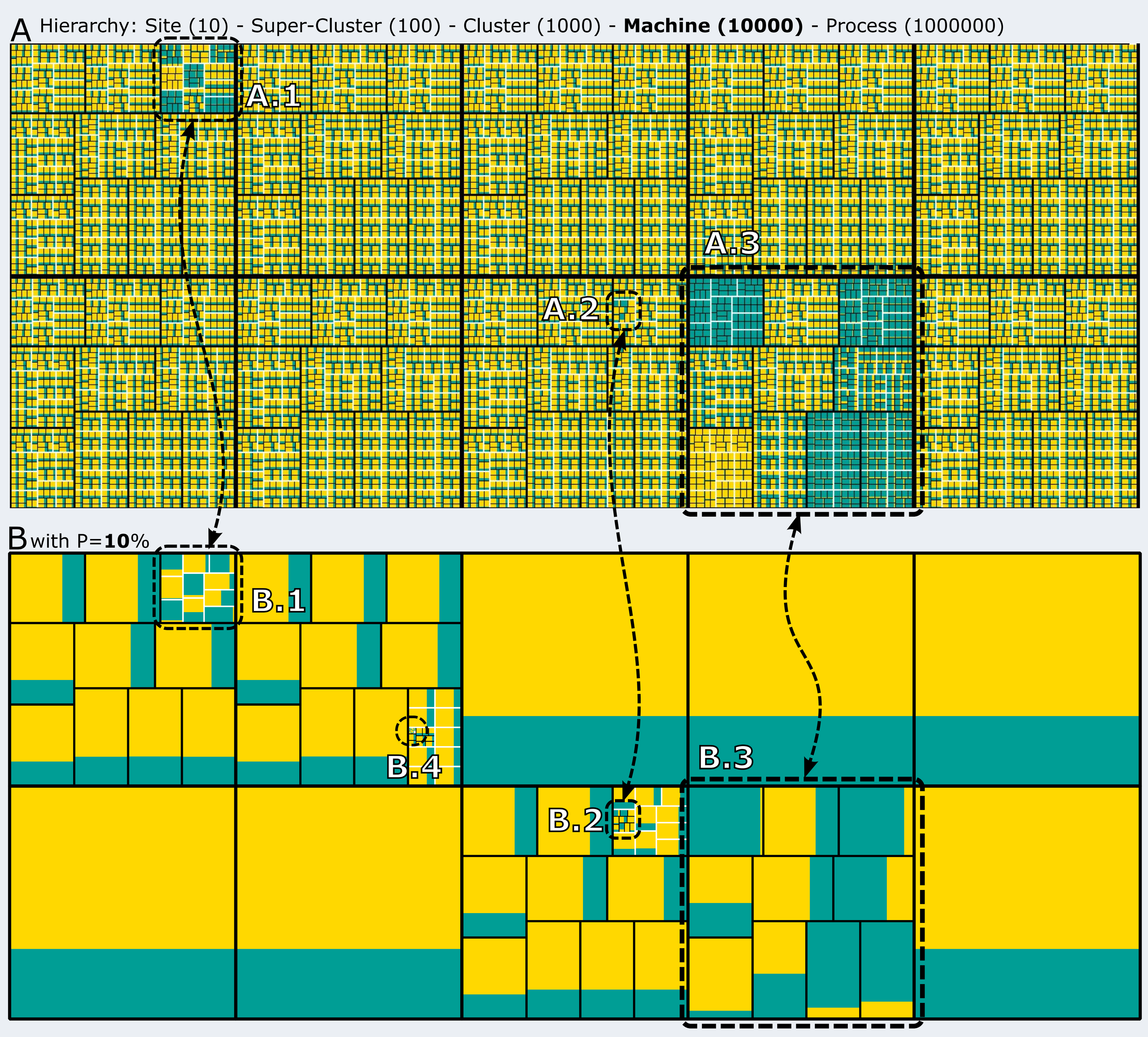
- Work stealing activity of a task-based parallel application.
- KAAPI framework on Grid'5000 platform
- Two states :

run (green) state, process executing tasks,
steal (red) state, process trying to steal work.



Anomaly identification at process or machine level

Scalable and Interactive Visualization of a Million of Processes



Synthetic workload with anomaly injections:
$\Rightarrow$ all anomalies are recovered by the multi-scale representation

Some References (long version [2])

[1] L. M. Schnorr, G. Huard, and P. O. A. Navaux, "A hierarchical aggregation model to achieve visualization scalability in the analysis of parallel applications," *Parallel Computing*, vol. 38, no. 3, pp. 91–110, 2012.

[2] R. Lamarche-Perrin, L. M. Schnorr, J.-M. Vincent, and Y. Demazeau, "Evaluating Trace Aggregation Through Entropy Measures for Optimal Performance Visualization of Large Distributed Systems," Laboratoire Informatique de Grenoble, France, Research Report RR-LIG-037, 2012.

[3] R. Lamarche-Perrin, Y. Demazeau, and J.-M. Vincent, "How to Build the Best Macroscopic Description of your Multi-agent System?" in *Advances in Practical Applications of Agents and Multiagent Systems*, ser. LNCS/LNAI, Y. Demazeau and T. Ishida, Eds., vol. 7879. Springer-Verlag Berlin, Heidelberg, 2013, pp. 157–168.

[4] I. Csiszár, "Axiomatic Characterizations of Information Measures," *Entropy*, vol. 10, no. 3, pp. 261–273, 2008.

[5] L. Mello Schnorr and A. Legrand, "Visualizing More Performance Data Than What Fits on Your Screen," in *Tools for High Performance Computing*. Springer-Verlag, 2013.

Future Works

- Generalization of the approach/algorithm to
  - networks of processes (spacial aggregation)
  - series of states (temporal aggregation)
  - sets of causally-related events (spatiotemporal aggregation).
- Combination with other complexity measures
  - Minimal Description Length,
  - theoretical framework (Kolmogorov Complexity)
- Application to other scientific domains:
  - embedded systems,
  - multi-agent systems,
  - geography, social sciences...

http://www.liglab.fr/ Robin.Lamarche-Perrin@mis.mpg.de, schnorr@inf.ufrgs.br, {Jean-Marc.Vincent,Yves.Demazeau}@imag.fr