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.

Information Theory Based Approach: Information Loss vs Complexity Reduction

- Shannon entropy measure of complexity
  \[ \text{gain}(A) = v(A) \log_2 v(A) - \sum_{e \in A} v(e) \log_2 v(e) \]
- Kullback-Leibler divergence estimates the information loss
  \[ \text{loss}(A) = \sum_{e \in A} v(e) \times \log_2 \left( \frac{v(e)}{v(A)} \times |A| \right) \]
- The sum property of quality measures enables independent computation of aggregates
- Hierarchical organization allows a recursive evaluation of branches
- Efficiently implemented through dynamic programming

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.

Scalable and Interactive Visualization of a Million of Processes

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

Future Works

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

Some References (long version [2])