Our research targets the application of data analytics techniques for extracting insight from collected performance data. Algorithms similar to the ones used for (big) data analysis can be applied to performance data to extract structure, correlations and information from the raw data.

Objectives

The maxim of the Performance Analysis Tools team is "Detail and Intelligence". With this aim, our group develops the following topics of the Performance Analytics area:

- Clustering, or cluster analysis, is applied to detect different trends in the application computation regions with minimum user intervention. This detection provides an unique insight of the application behaviour that serves as a starting point from which to perform different types of analyses around the applications' computation structure.
- Tracking correlates performance information from different executions or time intervals. The most relevant code regions and their behaviour are automatically identified and tracked across multiple experiments, enabling very diverse parametric and evolutionary studies.
- Folding combines both instrumentation and sampling for trace-based performance analysis tools. The
folding mechanism takes advantage of long execution runs and low frequency sampling to finely
detail the evolution of the user code without penalising the application or distorting the performance
results. The reports provided by the folding mechanism are extremely useful to understand the
behaviour of a region of code at a very low level.

- On-line analysis aims to maximise the amount of useful information presented to the analyst, while
keeping the trace size reasonably small. Clustering and time analysis techniques are applied to
intelligently select the most relevant data as it is being collected.

For more information, visit us at tools.bsc.es/research

Barcelona Supercomputing Center - Centro Nacional de Supercomputación