Xtrack
User Guide

tools@bsc.es

December 11, 2015

Contents

Chapter 1

Xtrack User Guide

1.1 Introduction

Xtrack is a tool for comparing multiple experiments or different time intervals within the same experiment. Xtrack takes as input the output produced by the ClusteringSuite tool [1] as well as the output produced by the Extrae On-line module when configured to run automatic structure detection analysis. The input of the tool is a series of clustering scatter-plots. Each clustering scatter-plot (or frame) is the result of clustering the computations of one execution of a parallel application with respect to selected performance metrics. Then, the clusters in one frame represent the main performance trends of the computations of the program. Comparing multiple frames, we can study how the behavior of the computing regions change between experiments. If we are changing a given parameter between experiments, this is useful to study the impact of this parameter in the program performance.

The tool has two parts: the tracking algorithm and the GUI. The *tracking algorithm* takes as input the sequence of frames and performs a "who-is-who" correlation between the clusters that appear in all the frames. To do this, the tool applies several heuristics that look for different characteristics that can distinguish certain clusters from the others. Then, the results of the tracking algorithm can be visualized with the *Xtrack GUI*.

In this document we briefly introduce the user to both parts of the tool, and show how to use them and the different settings available.

1.2 The Tracking tool

The first step to perform this comparative analysis is to apply the tracking algorithm to the sequence of frames that results from the application of cluster analyses to different traces (or sub-traces). To do so, the tool takes as input the list of clustered traces and some optional arguments that define which heuristics will be used to do the tracking. A detailed description of each of the available heuristics can be found here [2]. By default, the tool tries to apply all the heuristics that are applicable with the information comprised in the trace with default settings.

SYNTAX

```
tracking [OPTIONS] [-l LIST | TRACE1 TRACE2 ...]
```

OPTIONS

```
-a MIN_SCORE
      Minimum SPMD score to use the alignment tracker
-c CALLER_LEVEL
      Enable the callers heuristic at the specified stack depth.
-d MAX_DISTANCE
      Maximum Epsilon distance to use the cross-classifier tracker
-m MIN_TIME_PCT
      Discard the clusters below the given duration time percentage.
-s DIM1,DIM2...
      Select the dimensions to scale with the number of tasks.
-o OUTPUT_PREFIX
      Set the prefix for all the output files.
-r    Enable the trace reconstruction with tracked clusters.
-t THRESHOLD
      Minimum likeliness percentage in order to match two clusters
      (special values: all | first).
-v[v] Run in verbose mode (-vv for extra debug messages).
-x CLUSTERING_CONFIG_XML
      Specify the clustering configuration to automatically
      cluster the traces.
```

As a result, the tool generates:

- Trajectory lines for all the clusters in the frames, that show how the behavior of the clusters change across experiments.
- Recolored frames: a new set of scatter-plots where the identifiers of the clusters and their colors have been changed to make them match across experiments for easy comparison.
- Tracked traces: the input trace is reconstructed, changing the clustering events so that the clusters identifiers and their colors are the same in all experiments for easy comparison.
- The **.xtrack* definition file. This file contains all the information about "who-is-who" between experiments, and is the input for the Xtrack GUI.

1.3 The Xtrack GUI

The Xtrack GUI takes as input the results of applying the tracking algorithm, and displays a graphical comparison of the different experiments. Figure 1.1 shows the main view of the tool.

1.3.1 Menu bar

The menu bar contains controls to manipulate the plotting areas 2 and 3. From left to right, the common controls in both bars are:

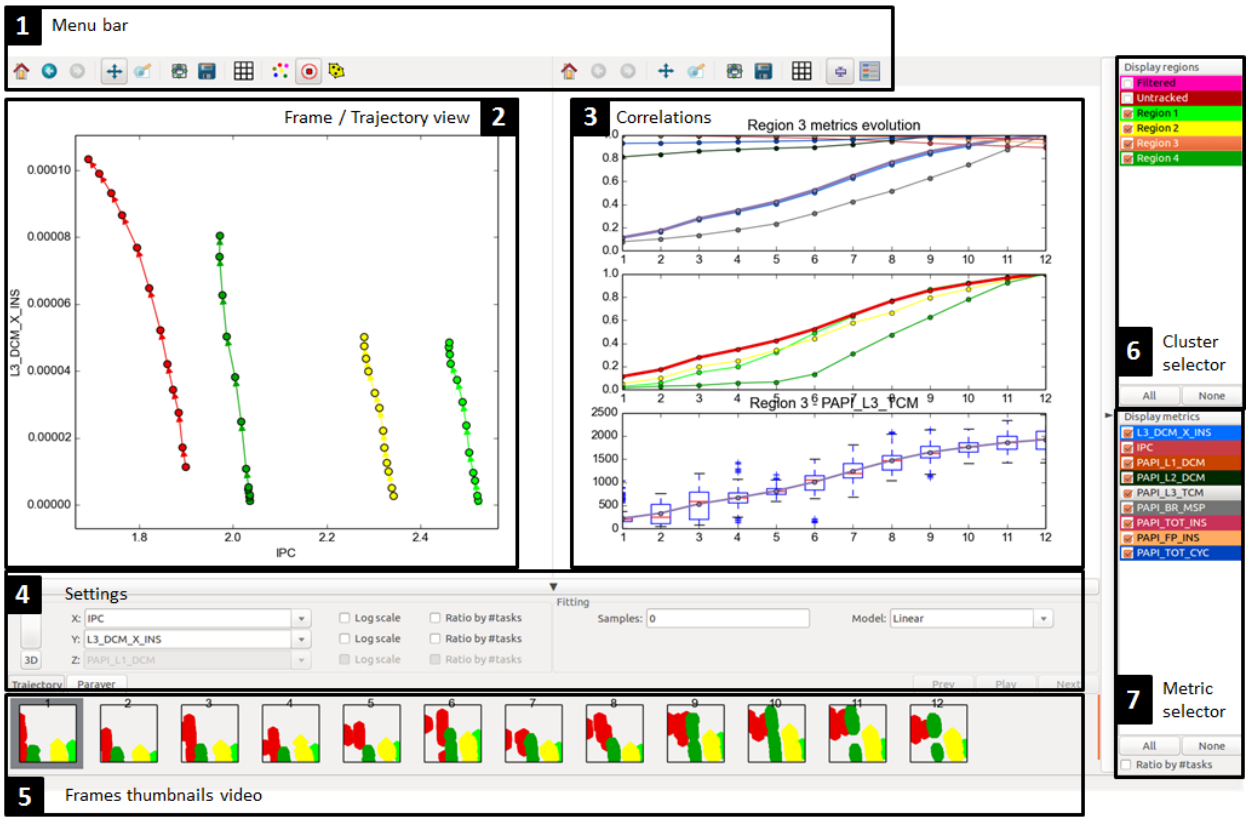

Figure 1.1: Xtrack GUI

- **Reset** the plot to the default view.
- **Undo** the last pan or zoom.
- **Redo** the last pan or zoom.
- **Pan** the graph towards any direction.
- **Zoom-in** dragging a box with the left mouse button, or **zoom-out** with the right mouse button.
- Adjust the plot **margins**.
- **Save** the visible plot.

The specific controls for the left menu bar (that controls plotting area 2) are:

- Show/hide a **grid** in the plot.
- Show/hide the **centroids** of the clusters.
- Show/hide the **perimeters** of the clusters.

And the specific controls for the right menu bar (that controls plotting area 3) are:

- Show/hide **boxplots** for dispersion plot in pane 3 (dispersion graph).
- Show/hide plots **legends**.

1.3.2 Frame/Trajectory view

This is the main view of the tool. When the *frame view* setting is selected in pane 4, it displays the cluster results for one single frame of the input sequence of experiments, the one that is selected in pane 5. When the *trajectory view* is selected in pane 4, it displays instead all the clusters from all the frames at the same time, and draws trajectory lines that show how the clusters are moving from one experiment to the next. In the image, the trajectory view is set by default.

1.3.3 Correlations

This pane shows correlations for different clusters and metrics across the sequence of experiments (X-axis). The top plot shows a correlation of all metrics that are checked in pane 7, for the cluster that is selected in pane 6. The middle plot shows a correlation of the selected metric in pane 7, for all the clusters that are checked in pane 6. Since these two plots can display different metrics or very different ranges for the same metric in the Y-axis, the value for this axis is normalized. The bottom plot shows the dispersion of the selected metric in pane 7 for the selected cluster in pane 6. In this case, the Y-axis shows real values.

1.3.4 Settings

The settings is divided in two parts. The *Axes* configuration (left) determine the metrics that are used to plot the performance space in pane 2. It is possible to select any performance counter that was used to cluster the trace or extrapolated, and also to select a third metric to change the view into a 3D plot. The *Log scale* checkbox can be ticked to draw each axis in logarithmic scale. The *Ratio by #tasks* can be selected to weight that axis metric with respect to the number of tasks that were used in the selected frame. This is useful when the metric is related to the number of instructions executed (e.g. total instructions, floating point instructions, etc.) for doing strong scaling or weak scaling comparisons, and studying code replication issues.

The *Fitting* settings (right) are used to do predictions based on the experiments we have. You can select a fitting model between linear, quadratic, cubic, logarithmic or log-linear (right), and a number of experiments to predict (left), and all the correlation plots in pane 3 will extend their X-axis to predict how the series will continue according to the selected model.

1.3.5 Frame thumbnails video

This pane shows a carousel of frames, where each frame is the clustering result of every single experiment. This representation provides a quick view on the changes across experiments. Also, this pane allows to select a single frame to be inspected in detail in pane 2.

1.3.6 Cluster selector

Ticking the checkbox of each cluster we can control whether that cluster has to be displayed/hidden in the plots in panes 2 and 3. Also selecting a single cluster from the list, changes the plots in pane 3 to display the metrics correlations for the selected cluster only.

1.3.7 Metric selector

Analogously, ticking the checkbox of each metric we can control whether that metric has to be displayed/hidden in the plots in pane 3. Also, selecting a single metric from the list, changes the plots in pane 3 to display the correlations of clusters for the selected metric only.

Bibliography

[1] BSC Tools. `http://www.bsc.es/paraver`.

[2] G. Llort, H. Servat, J. González, J. Giménez, and J. Labarta. Studying performance changes with tracking analysis. In *Proceedings of International Parallel Tools Workshop*, IPTW'14, Stuttgart, Germany, 2014.