Ocelotl User Guide

Youenn Corre
Damien Dosimont

Ocelotl User Guide
v. 1.1.8

Youenn Corre*
Damien Dosimont†

Project-Team MOAIS

Abstract: This guide describes the Ocelotl tool, in its version 1.1.8, available at the following github repository: `https://github.com/soctrace-inria/ocelotl`.

* INRIA, youenn.corre@inria.fr
† INRIA, damien.dosimont@imag.fr

Table of contents

1   Introduction

Ocelotl is a tool to explore, analyze and visualize execution traces through aggregation techniques. It is part of the SoC-Trace project, and as such, it is integrated in the Framesoc framework as an Eclipse plug-in.

2   Installation

Please refer to the installation procedure described on the Ocelotl wiki page: `https://github.com/soctrace-inria/ocelotl/wiki/User-Guide`.
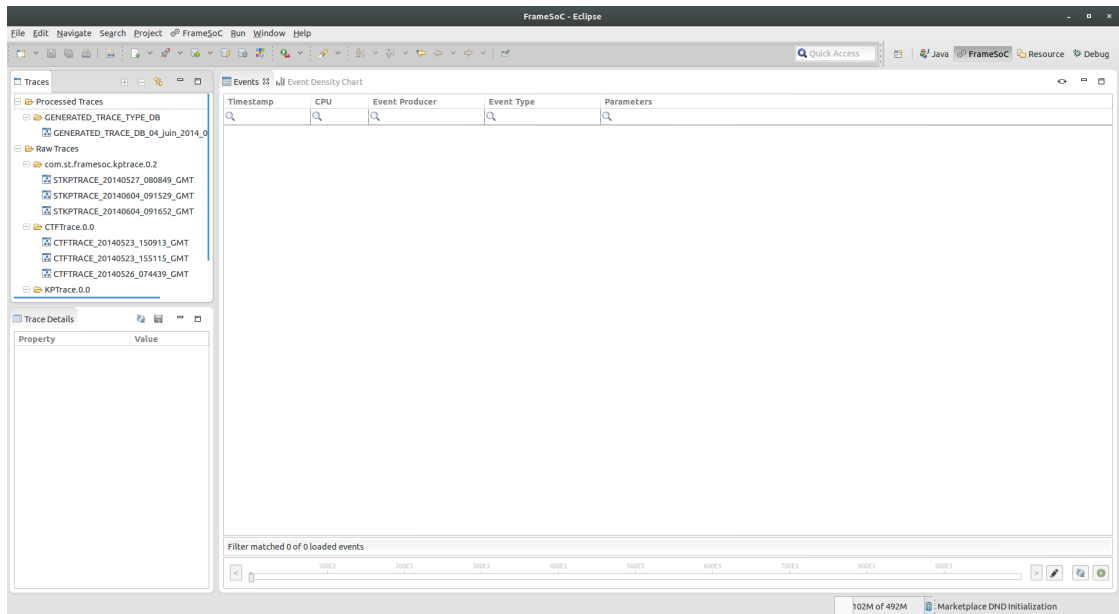
3   Launch Ocelotl

Figure 1: Framesoc Eclipse Workbench at startup

In order to launch Ocelotl:

1. - Go to the Framesoc menu, choose the *Trace Analysis* entry and click on *Launch Analysis Tool*.
    - Or click on the *Analyze Trace* icon.

2. Select the Ocelotl tool.

3. No argument is needed so just click *OK*.

The Ocelotl tool view should open, as illustrated in Fig. 4.
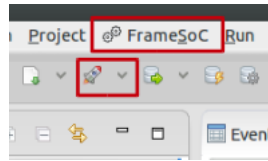
Figure 2: The Framesoc menu and the *Analyze Trace* icon
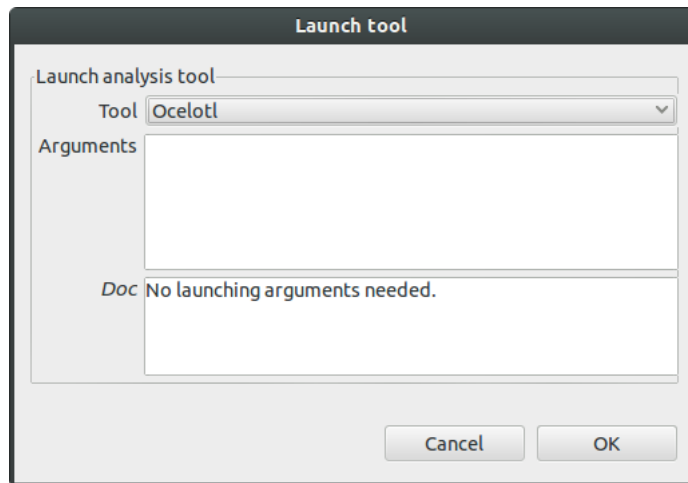
Figure 3: Ocelotl launching window
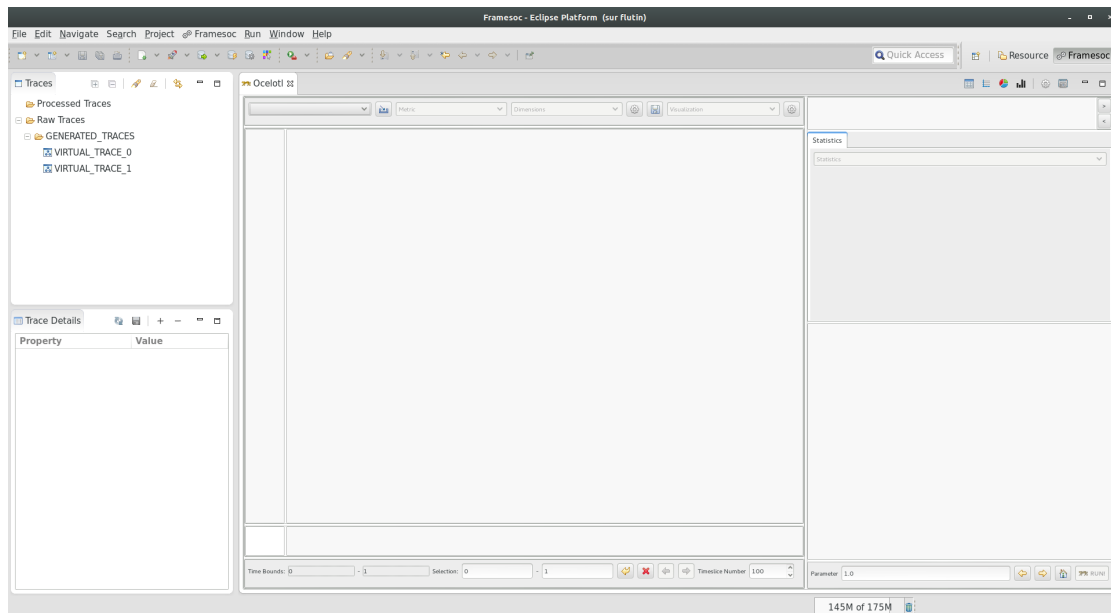
Figure 4: Ocelotl at startup

4   Ocelotl Features

In this section, we describe how to load a trace, launch the analysis and explore the results by modifying parameters.

4.1   Load a Trace

Figure 5: Trace configuration interface

4.1.1   Select a Trace

At the top of the Ocelotl window, the *Trace Overview* tool bar can be used to select a trace to analyze (element 1 in Fig. 5). It shows the traces already exported in Framesoc. Once a trace is selected, default settings compatible with the trace are automatically selected. However it is possible to change these settings, as described in the following subsections.

It is also possible to load a trace from the Framesoc Trace tree (the frame on the top left in Fig. 1): right-click on the trace to load and select *Show in Ocelotl* in the context menu. By using this method to load a trace, it is possible to launch multiple instances of Ocelotl, up to a maximum of five simultaneous instances[1].

4.1.2   Metrics

The *Metrics* combo (element 2 in Fig. 5) displays the metrics compatible with the current trace. At the moment, the available metrics are:

- *Event occurences* which performs an analysis based on the event occurrences. This type of analysis is compatible with all the traces.

- *State total duration* which performs an analysis based on the event duration. This analysis requires the trace to contain events of the type state.

- *State average duration* which performs an analysis based on the average duration of a state in a given time slice. This analysis requires the trace to contain events of the type state.

- *Variables* which performs an analysis based on the variation of the value of the event over time. This analysis requires the trace to contain events of the type variable.

---

[1] This limit can be set in the Framesoc configuration file. Please refer to the Framesoc user doc to configure Framesoc (`https://github.com/soctrace-inria/framesoc/blob/master/src/fr.inria.soctrace.maven.repository/archive/doc/framesoc_user_guide.pdf?raw=true`).

4.1.3 Dimensions

Once the type of metrics selected, it is possible to choose the aggregation operator (element 3 in Fig. 5). Three operators are currently available:

- The *Temporal Aggregation*, which performs an aggregation based only on the temporal dimension.

- The *Spatiotemporal Aggregation*, which performs an aggregation based on the temporal and the spatial dimension. The spatial dimension is based on the hierarchy of event producers.

It is possible to configure the analysis by clicking on the *Settings* button next to the metrics selection combo, which opens a setting window (cf. Fig. 6).
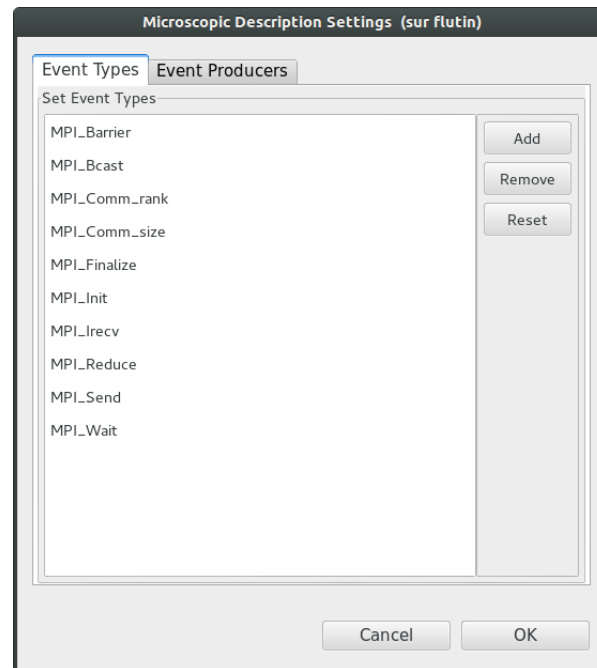
Figure 6: Aggregation Settings

In the settings, by default, all the event types and all the event producers are selected. The available settings are:

- in the *Event Types* tab, the filtering of the event types involved in the analysis;

- in the *Event Producers* tab, the filtering of the event producers (the resources that produce the events) involved in the analysis. The producers are shown with their hierarchy, and depending on whether the operator required a valid hierarchy or not[2], the deselection of a parent node can lead to the deselection of its children nodes. The *Add Result* button enables to load a set of event producers saved as an analysis result by another tool.

---

[2]The spatio-temporal operators take into account the hierarchy, while the temporal ones do not.

4.1.4 Visualization Settings

The *Visualization* combo (4 in Fig. 5) provides a list of available views, in accordance with the compatibility of the previously selected metrics and aggregation operators. The currently available choices are:

- *Partition*: is a visualization showing aggregates that correspond to an homogeneous behavior in the trace. The aggregates are separated by a blank space and all have a different colors. For a temporal aggregation, the available parameters in *Settings* are:

  - *Aggregate parts*: gather time slices that belong to the same aggregate. If not active, each time slice is separated, and their aggregation is shown thanks to their color.

  - *Show part number*: display the number of each aggregate part to help distinguishing them.

- *Temporal Stacked Bar Chart*: is a visualization showing the same aggregates as in *Parts*, but also the proportion of event occurrences (or the ratio of state duration, if state distribution was selected in the previous step) for each aggregate. In this visualization, each event/state is separated vertically in each aggregate and their proportion is shown by making the height of each event/state in an aggregate proportional to its occurrence/duration. The events/states that have a proportion too small to be visible, are aggregated and displayed on top of the aggregate as a dotted line. In the *Settings* window, you can customize the color of each event.

- *Mode*: is a visualization based on the value that appears most often[3]. It is available for both temporal and spatiotemporal aggregation operators. In the *Settings* window, you can customize the color of each event. In the settings of the spatiotemporal mode, you can also perform a filtering of the displayed event types and change the colors of the event producers. The difference with the event type filtering in the aggregation settings is that the aggregation process is performed with all the types selected in the aggregation settings, but the visualization takes into account only the types not filtered in its settings.

4.2 Quality Curves

Ocelotl displays curves representing the different aggregation levels, as shown in Fig. 7. The X axis is the aggregation ratio from 1 (fully aggregated) to 0 (no aggregation). The green curve shows the level of displayed complexity and the red is the level of displayed information. The vertical blue line shows the current value of the parameter.

4.3 General Settings

At the top right of the Ocelotl tool, a button (1 in Fig. 8) launches a window to configure the Ocelotl settings (Fig. 9). When clicking *OK*, the settings are saved for the current instance of Ocelotl. To permanently save the settings for all future instances of Ocelotl, click on the button *Save Settings*.

4.3.1 Aggregation

The tab *Aggregation* settings contains three subsections: the *Quality Curves* and *Parameter P* settings. The *Quality Curves* settings enable to change the appearance of the complexity and information curves provided by the aggregation algorithm:

---

[3]It is based on the notion of mode in statistics `http://en.wikipedia.org/wiki/Mode_(statistics)`
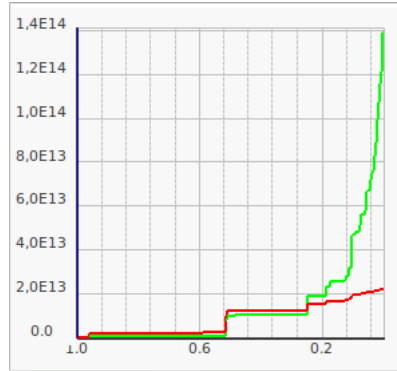
Figure 7: Quality curves in rising order.

Figure 8: Ocelotl Features: 1. General settings; 2. Take a snapshot

- *Normalize Qualities*: normalize the curves and scale them on the interval $[0:1]$ to improve their readability. **<u>Note:</u>** normalization is no longer taken into account in spatiotemporal operators since it provides incorrect results.

- You can choose between two representations of the curves: rising curves (complexity gain, information gain) or decreasing curves (complexity reduction, information loss).

- The *Threshold* parameter determines the quality curves precision. The more its value decreases, the more the precision increases and consequently more aggregation steps are found. However, setting a too small value may lead to bad performances, so we advice to keep the default value.

The *Parameter P* settings let you choose the strategy used to select the value of P that will be chosen as default after launching an analysis.

Enable visual aggregation prevents the drawing of cluttered representations by aggregating visually graphical objects that can not be represented properly.

4.3.2 Cache

To prevent time-consuming loading of traces from the database, we use a cache system to speed up the access to the trace raw data. Two operations performed by Ocelotl uses cache: the building of the microscopic model (*data cache*) and the computation of the parameter P values (*partition cache*). If the data cache is enabled in the settings, then the first time a trace is
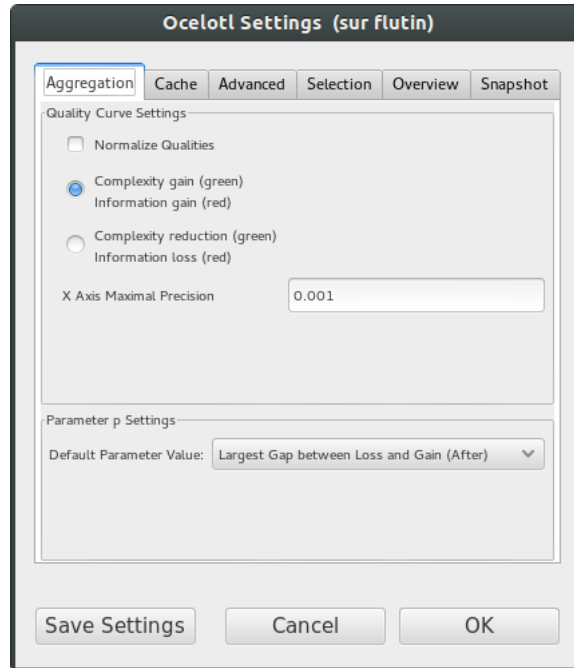
Figure 9: Ocelotl Settings

analyzed, a data cache file is generated. A data cache file is a plain text file that contains non-null values of the trace, and each cache is specific to an metrics operator, and a specific time region.

Several settings are available to configure the caches:

- *Empty cache*: delete all the cache files in the current cache directory.

- *Caches directory*: specify the directory where the cache files are stored. All the cache files specific to a given trace are stored in a subdirectory with a name composed of the name and the unique ID of the trace.

- *Maximum cache size*: specify the maximum cache size in megabytes. When the maximum size is reached, the cache files with the oldest accessed dates are deleted until enough space is available. Specifying a value of -1 means that there is no maximal limit to the cache size.

- *Trace abstraction cache enabled*: enable or disable the related cache.

- *Optimal time slice number*: the first time a trace is loaded, a cache file is generated — if the cache is enabled. This parameter specifies the number of time slices used for the generated cache. A higher number means a more accurate cache, but also a bigger cache file.

- *Data cache policy*: specify the rebuilding policy used to rebuild the data from the cache

file when a "clean" rebuilding is not possible[4]. The available choices are:

- ○ *Fast*: the dirty time slices are rebuilt by multiplying the values proportionally to the division of the cache time slice between the two new time slices (e.g. for a time slice which is split with 30% of its time in one time slice and 70% in the other, its values will be multiplied by 0.3 and 0.7 respectively). While this is a fast method, it may not be accurate in all cases (especially in cases with high ratio of dirty time slices).

- ○ *Accurate*: the dirty slices are rebuilt by performing a request to the database so that the events are rebuilt precisely. Depending on the event nature (punctual event, state, etc.), the rebuilding can be pretty slow.

- ○ *Automated*: try to automatically choose the best strategy between *Fast* and *Accurate*.

- ○ *Ask me*: show a dialog to the user at each run, from which he can select a strategy.

- • *Partition cache enabled*: enable or disable the cache used to generate the quality curves.

If an aggregation was performed using an approximate strategy, a warning is displayed in the status bar of Eclipse.

4.3.3 Advanced Settings

In the *Advanced Settings* tab, you can configure some more advanced settings[5]:

- • *Event Number Retrieved by Threads*: set how much events are loaded by thread for each iteration during the trace loading.

- • *Event Producers per Query*: enable to divide the query into several queries with a fixed number of event producers. This functionality is useful if the amount of free space on the drive is low, as the temp directory is used to store the query result during the trace reading. The default value is 0 which corresponds to fetching all producers in one query.

- • *Working Threads*: set the number of active threads during the trace reading.

- • *Leaf Aggregation*: it is possible to specify a maximum number of leaf event producers (i.e. event producers that have no child in the hierarchy). This is useful to bound the size and thus the memory footprint of the generated microscopic model, especially when there is a large number of event producers. If the maximum number of leaves is reached, then Ocelotl looks for the hierarchy level at which to perform the aggregation. This means that all the event producers that are above the level in the hierarchy will be merged with their parent in the hierarchy level selected for the aggregation. This pre-aggregation will of course affects the aggregation process, and thus the displayed results might differ from the ones where no event producer would be aggregated. If some event producers are aggregated, then a warning is displayed in the status bar of Eclipse. When performing a spatial zoom, if the number of leaves displayed is under the set maximal bound, then the event producers are no longer aggregated.

---

[4]A "clean" rebuild means that no time slice of the cache is split inside two time slices of the currently built microscopic description (such time slices are called "dirty" time slices). Typically it is clean when the number of time slices of the cache is dividable by the current number of slices (for example, if the cache was built with 10 000 time slices and the current rebuilt uses only 100 time slices, then one time slice of the new model will be built each from 10 time slices of the cache. But if the current rebuild is using 150 time slices, then some time slices of the cache would be split between two of the new time slices. The rebuild strategies are thus used for those "dirty" cases.

[5]**Note:** Modifying these values might decrease the trace reading performances.

4.3.4   Selection Settings

The *Selection Settings* can be used to change the color and transparency values of the selections performed on the diagram.

4.3.5   Overview Settings

The *Overview Settings* allow to enable or disable the overview, along with customizing the color and transparency settings of the selection displayed on the overview. It is also possible to enable a pre-aggregation of the leaf event producers, and specify the maximum number of leaves, in a fashion similar to what is done for the main view in the *Advanced Settings*.

4.3.6   Snapshot Settings

The *Snapshot Settings* allow to set the saving directory of the snapshots taken in Ocelotl, as well as the resolution of the snapshots. The latter feature can be useful since drawing is relaunched by taking into account the new resolution, meaning that it can overcome the limitation of the screen resolution, and thus avoid any visual aggregations.

4.4   Time Settings

In order to tune the time settings, several parameters can be modified in the bar under the graph display area (Fig. 10).

Figure 10: Time Settings

4.4.1   Timestamps

There are two types of timestamp values shown: the first one (1 in Fig. 10) shows the starting and ending timestamps of the currently displayed zone of the trace, the second (2) shows the bounding timestamps of the currently selected zone (or are identical to the first ones, if there is currently no selection). The first timestamps are not editable, but the second set of timestamps is, and allow to set the trace time bounds by modifying the values. The *Cancel Selection* button (4) cancels the current selection (both temporal and spatial) to the currently displayed timestamps and event producers. Alternatively, the selection can be canceled with a middle-click. The *Reset* button (3) also cancels the current selection and in addition resets the temporal values to the global starting and ending timestamp values of the trace, and also resets the event producer selection to all the event producers present in the trace.

4.4.2   Time Slices

The *Timeslice Number* field (6) corresponds to the number of time slices that will be used to compute the aggregation. It is advised to tune this parameter in order to find a granularity that is convenient to the analysis. However, since the aggregation algorithm complexity is dependent on this parameter, it is advised to increment it progressively to keep good performances.

4.5 Analysis Settings

To analyze the trace, several parameters allow to explore different levels of aggregation and visualization. The aggregation settings are located under the *Quality curves* display.
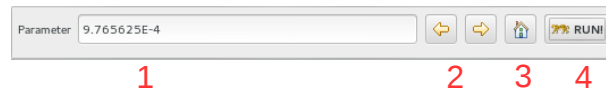

Figure 11: Aggregation Settings

4.5.1 Run an Analysis

Once the configuration is set, click on RUN! (4 in Fig. 11), or press the ENTER key on the keyboard. The aggregation is then computed. This operation can take a while. When it is finished, the visualization and the curves are shown, as illustrated in Fig. 13.

4.5.2 Change the Aggregation Strength

In order to find a relevant aggregation, you can tune the *Parameter* setting. This parameter allows to explore different aggregation levels, by changing the balance between the displayed information and its complexity. There are three ways to do this:

1. Click on the curves. The corresponding parameter value is retrieved, and used to show the associated aggregation.

2. Use the <- and -> buttons (2) (or use left and right arrows on the keyboard). It increments or decrements the parameter to the next interesting threshold.

3. Enter a value manually, in the *Parameter* field (1). The value must be between 0 and 1.

A value of 1 means that the aggregation strength is maximal: the whole trace is aggregated. On the opposite, a value of 0 means that each all heterogeneous consecutive time slices are disaggregated. An intermediate value provides a compromise between both extrema. The button (3) resets the parameter to its default value — i.e. the value displayed just after a run.

4.5.3 Zoom

Once you obtain an interesting aggregation level, you may want to focus on a particular time area. It is possible to zoom in on the trace by selecting an area. The selection is performed by clicking with the left button and dragging the mouse while maintaining the left button down. Once the mouse is released the selection will adjust to the nearest selected time slice. It is is also possible to select just one time slice through a simple left click. Then, click on RUN! or press ENTER to display the selected area. In order to zoom out, you must click on *Reset* and then on RUN!, or use the browsing history arrows (cf. *Zoom History* paragraph below).

Spatial Selection If a spatio-temporal aggregation operator was chosen, then it is possible to perform a spatial zoom also. If more than one producer are selected, then the spatial selection automatically adjusts to the parent node containing all the selected producers. A click on the RUN! button will zoom on the selected area. A simple click selects a single producer, (or a single producer node, if a producer is too small to be represented), and a double-click selects the whole aggregate.

Zoom History   During an exploration session of a trace, the zoom history is recorded, on both temporal and spatial dimensions (i.e. the selected time zones, and the selected event producers). It is possible to browse the history through the $<-$ and $->$ buttons in the bottom bar (cf. 5 in Fig. 10). The $<-$ button shows the previous zoom value, and the $->$ button shows the next zoom value. The history is lost when a different trace or a different aggregation operator is selected.

4.5.4   Switch to Another View

Figure 12: Framesoc tools buttons

It is possible to display the trace with the other visualization tools integrated in Framesoc: a Gantt chart, an event table, a statistics pie chart and a density chart. This is performed by clicking on the corresponding button in the top right corner, shown in Fig. 12. The newly opened view will show the events of the trace that are within the time bounds shown in the *Start* and *End* fields. You can focus on an area by selecting it with the mouse left button and then switching directly to another view, without loading the aggregation with the RUN! button.
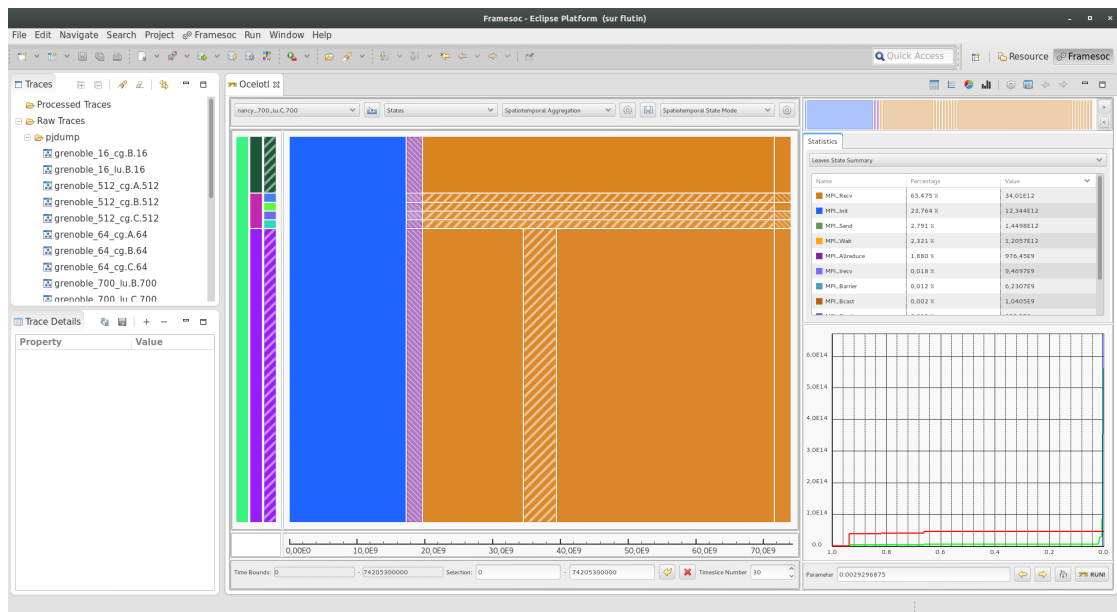
4.6   Graph Display

Figure 13: A graph showing an aggregated trace

Fig. 13 shows an aggregated trace, with the proportion view. At the bottom of the graph, a time axis shows the timestamps of the currently displayed events. Events of different types have

different colors. By letting the mouse over a part of the graph for a short time, a tooltip note shows the type of events represented by that part of the graph. For the event view, this note also displays the number of occurrences of the events divided by the duration of the aggregate[6]. For the state view, it is the state duration divided by the aggregate duration.

4.6.1 Spatiotemporal Aggregate

If a spatio-temporal aggregation is used, it is possible that the size of the window does not allow to correctly display the aggregates, especially if there is a great number of producers. If, in an aggregate, there are disparities only on the spatial level, then the aggregate is displayed with right to left stripes. If the disparities are also temporal, then the stripes go from left to right. Both of these cases can be seen in Fig. 13. In any case, to display the content of the aggregate, right-click on the aggregate and the content will be displayed in a new window. To close the newly created window, simply click on it.

4.6.2 Y Axis

Depending on the selected aggregation operator and visualization type, a Y axis displaying some additional information might appear on the left side of the graph display. For spatiotemporal aggregation, the Y axis displays the hierarchy of the event producers. This can be seen in Fig. 13. The hierarchy is shown as a set of rectangles, each representing a level of the hierarchy — from left to right, from the more general to the specific (typically root, cluster, machine, node). By hovering the mouse on a producer, the tooltip displays the name of the producer. If there is not enough place to correctly display the producers, they will be aggregated visually. By right-clicking on one of the producer, a hierarchy tree with the selected producer as root will be shown in a new window. Additionally, it can be used to select the event producer for the whole displayed time region, by simply clicking on the producer.

In the case of a proportion view with a temporal aggregation, the Y axis displays the proportion value. When available, the unit is displayed at the top of the Y axis.

4.7 Load and Save a Cache File

4.7.1 Load

An alternative way to load a trace, instead of using the combo box, is to load it from its cache file, by using the button to the right of the combo box (element 5 in Fig. 5). It opens a new window, allowing to select a cache file. If the selected cache file is valid and corresponds to a trace currently in the Framesoc database, then the trace is automatically loaded with the settings saved in the cache file.

4.7.2 Save

Once a run has been performed on a trace, a cache file is generated, in order to speed up the next analysis of the trace. It is then possible to save this cache to a specific location, in order to get a fast access later. This is done by pressing the button to the far right of the *Aggregation Operator* combo box (element 6 in Fig. 5).

Figure 14: An overview of a trace

4.8 Overview

In the top right corner, Ocelotl displays an overview of the current trace, which gives the context of the trace. When a zoomed view of a trace is displayed, the overview highlights the currently displayed view. When selecting a part of the view on the graph, the corresponding selection is highlighted in a different color on the overview.

The overview shows a temporal summary of the currently displayed trace. It is based on the mode operator described in subsection 4.1.4. The two small buttons on the right of the overview can be used to modify the aggregation level of the overview, and work in a similar fashion to the buttons 2 in Fig. 11. Alternatively, you can also use the keys u and i to browse between aggregation levels.

Note: Since the overview is computed in a different thread, there may be a delay between the display of the main graph and the apparition of the overview.

4.9 Statistics

Figure 15: An example of statistics

Ocelotl displays statistics about the currently selected data (as illustrated by Fig. 15). The data selection is performed by dragging the mouse in the display of the aggregated graph, or, to select a single time slice, by performing a single click on the desired time slice. The statistics display is updated when the mouse button is released. To reset the selection, click on the *Reset* button (3 in Fig. 10) which sets the statistics data back to the whole displayed region.

The combo box lists all the compatible statistics operators for the currently selected operators. For now, only one statistic operator is available. It summarizes the data of the currently selected time region:

---

[6]The time unit of the trace is displayed in the *Trace Details* tab on the left bottom of the screen.

- If the *Event occurence* metrics is selected, then all events are considered to be punctual events, then the statistics display shows for each type of event their number of occurrences.

- If the *State total duration* metrics is selected, then the statistics display shows the portion of time spent in this state during the selected time region. Note that the total percentage can be different of 100% since some event producers can be in no active state, leading to a total under 100%. It is also possible possible to have more than 100%, since some states can be embedded, meaning that an event producer can have more than one active state at a time.

The statistics can be sorted by clicking on the column names of the table.

4.10 Take a Snapshot

In order to save the currently displayed diagram, axes and curves, it is possible to save them as image files along with the current parameters of the trace in a text file. This is done by clicking the button at the top right of the display (2 in Fig. 8).

4.11 Known Issues

- Depending on the main window size or on the screen resolution, some components of the Ocelotl UI might not be displayed correctly or might be partially hidden. Workaround: Resize the window, put Ocelotl in full screen by double-clicking on the Ocelotl tab or use the resize handles of the Ocelotl widgets to resize them accordingly.

- The temporal selection between the main graph and the time axis view have a few pixels of difference.

- Resizing the window can take a while depending on the implementation of SWT. This phenomenon has been mainly encountered on Linux systems.

- Clicking repeatedly and quickly on the quality curves may cause a incoherence between the selected parameter p and the visualization that is printed (which corresponds to a previously computed aggregation). Click again on the curves to actualize the visualization.

For an exhaustive list of all the current known issues in Ocelotl, please refer to the issues section of our GitHub[7].

---

[7] https://github.com/soctrace-inria/ocelotl/issues