



TEMPORAL AGGREGATION

For Voluminous Trace Visualization

AEP 10 - 12th June 2014

Damien Dosimont
Guillaume Huard
Jean-Marc Vincent

Inria, Grenoble University
first.last@imag.fr



How an Internet Provider Loses a Customer

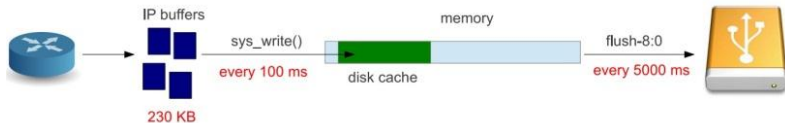
RECORDING A STREAMING WITH A SET-TOP BOX

- A customer wants to see the last episode of Game of Throne on his TV
- Unfortunately, the episode is released the morning
- So, he decides to record the streaming on a hard drive thanks to his set-top box
- Eventually, he watches the episode just after the work
- But...

THE VIDEO

Play

ST MICROELECTRONICS' TS RECORD USE CASE



- HD video streaming : **big quantity of data** transmitted through the **network**
- Data stored in **IP buffers**, waiting to be sent to the disk
- **sys_write()** function send the data to the disk every 100 ms
- The **kernel flushes** the disk cache every 5000 ms

Trace Visualization Issues

HOW TO DEBUG OUR APPLICATION?

Tracing

Collect **information** about the **application behavior**

→ **Structure :**

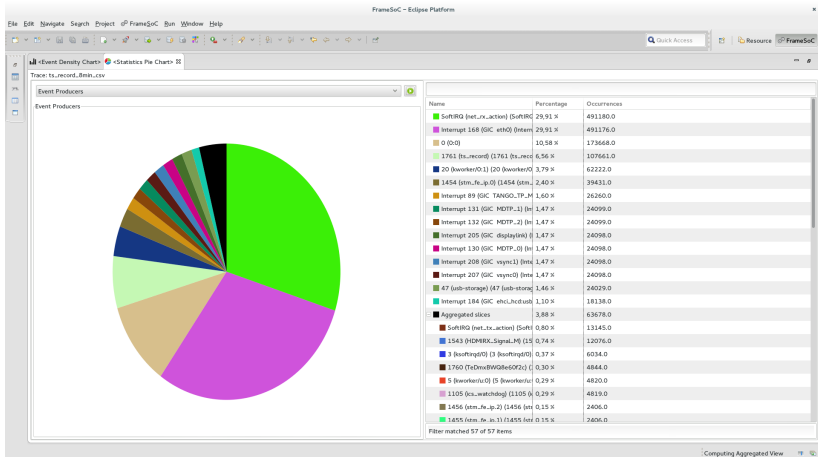
- **hardware** components (machines, cores, dedicated hardware)
- **software** components (processes, threads, system, middlewares)

→ **Timestamped events** : function calls, synchronization, communications, CPU load, memory utilization etc.

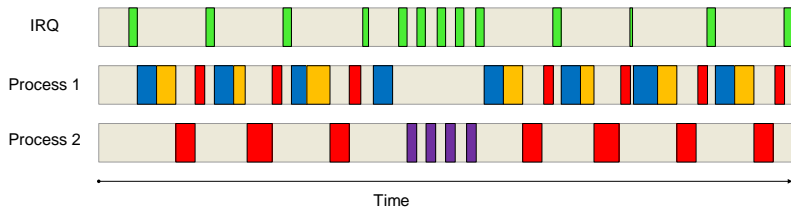
Visualization

Statistics, Gantt chart, call graphs etc.

STATISTICS



EXAMPLE OF GANTT CHART



VISUALIZATION ISSUES

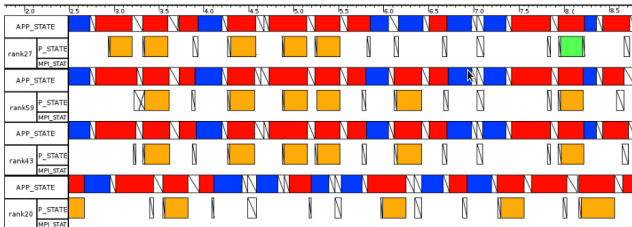
- **Statistics** are not able to show temporal perturbations
- Traces are **huge** (we have worked with traces up to 200 million events)
- Traditional temporal visualization does not respect **Shneiderman methodology**
Overview first, zoom and filter, then, get details-on-demand
- Some tools propose **aggregation techniques** but they do not convey **meaning** to the user

The screenshot displays the STLinux Trace tool interface. At the top, there are navigation tabs: '<Event Density Chart>', '<Statistics Pie Chart>', 'Traces', 'Trace Details', 'STLinux Trace', and '<Events>'. The 'STLinux Trace' tab is currently selected.

The main area is divided into two sections, one for CPU 0 and one for CPU 1. Each section contains a list of tasks on the left and a corresponding event density chart on the right.

CPU 0 Tasks:

- initrcs
- Intmrgt 89 (SIC TMS320C79_M80)
- Intmrgt 130 (SIC MCFIF-1)
- Intmrgt 131 (SIC MCFIF-1)
- Intmrgt 132 (SIC MCFIF-1)
- Intmrgt 133 (SIC MCFIF-1)
- Intmrgt 134 (SIC MCFIF-1)
- Intmrgt 135 (SIC MCFIF-1)
- Intmrgt 136 (SIC MCFIF-1)
- Intmrgt 137 (SIC MCFIF-1)
- Intmrgt 138 (SIC MCFIF-1)
- Intmrgt 139 (SIC MCFIF-1)
- Intmrgt 140 (SIC MCFIF-1)
- Intmrgt 141 (SIC MCFIF-1)
- Intmrgt 142 (SIC MCFIF-1)
- Intmrgt 143 (SIC MCFIF-1)
- Intmrgt 144 (SIC MCFIF-1)
- Intmrgt 145 (SIC MCFIF-1)
- Intmrgt 146 (SIC MCFIF-1)
- Intmrgt 147 (SIC MCFIF-1)
- Intmrgt 148 (SIC MCFIF-1)
- Intmrgt 149 (SIC MCFIF-1)
- Intmrgt 150 (SIC MCFIF-1)
- Intmrgt 151 (SIC MCFIF-1)
- Intmrgt 152 (SIC MCFIF-1)
- Intmrgt 153 (SIC MCFIF-1)
- Intmrgt 154 (SIC MCFIF-1)
- Intmrgt 155 (SIC MCFIF-1)
- Intmrgt 156 (SIC MCFIF-1)
- Intmrgt 157 (SIC MCFIF-1)
- Intmrgt 158 (SIC MCFIF-1)
- Intmrgt 159 (SIC MCFIF-1)
- Intmrgt 160 (SIC MCFIF-1)
- Intmrgt 161 (SIC MCFIF-1)
- Intmrgt 162 (SIC MCFIF-1)
- Intmrgt 163 (SIC MCFIF-1)
- Intmrgt 164 (SIC MCFIF-1)
- Intmrgt 165 (SIC MCFIF-1)
- Intmrgt 166 (SIC MCFIF-1)
- Intmrgt 167 (SIC MCFIF-1)
- Intmrgt 168 (SIC MCFIF-1)
- Intmrgt 169 (SIC MCFIF-1)
- Intmrgt 170 (SIC MCFIF-1)
- Intmrgt 171 (SIC MCFIF-1)
- Intmrgt 172 (SIC MCFIF-1)
- Intmrgt 173 (SIC MCFIF-1)
- Intmrgt 174 (SIC MCFIF-1)
- Intmrgt 175 (SIC MCFIF-1)
- Intmrgt 176 (SIC MCFIF-1)
- Intmrgt 177 (SIC MCFIF-1)
- Intmrgt 178 (SIC MCFIF-1)
- Intmrgt 179 (SIC MCFIF-1)
- Intmrgt 180 (SIC MCFIF-1)
- Intmrgt 181 (SIC MCFIF-1)
- Intmrgt 182 (SIC MCFIF-1)
- Intmrgt 183 (SIC MCFIF-1)
- Intmrgt 184 (SIC MCFIF-1)
- Intmrgt 185 (SIC MCFIF-1)
- Intmrgt 186 (SIC MCFIF-1)
- Intmrgt 187 (SIC MCFIF-1)
- Intmrgt 188 (SIC MCFIF-1)
- Intmrgt 189 (SIC MCFIF-1)
- Intmrgt 190 (SIC MCFIF-1)
- Intmrgt 191 (SIC MCFIF-1)
- Intmrgt 192 (SIC MCFIF-1)
- Intmrgt 193 (SIC MCFIF-1)
- Intmrgt 194 (SIC MCFIF-1)
- Intmrgt 195 (SIC MCFIF-1)
- Intmrgt 196 (SIC MCFIF-1)
- Intmrgt 197 (SIC MCFIF-1)
- Intmrgt 198 (SIC MCFIF-1)
- Intmrgt 199 (SIC MCFIF-1)
- Intmrgt 200 (SIC MCFIF-1)
- Intmrgt 201 (SIC MCFIF-1)
- Intmrgt 202 (SIC MCFIF-1)
- Intmrgt 203 (SIC MCFIF-1)
- Intmrgt 204 (SIC MCFIF-1)
- Intmrgt 205 (SIC MCFIF-1)
- Intmrgt 206 (SIC MCFIF-1)
- Intmrgt 207 (SIC MCFIF-1)
- Intmrgt 208 (SIC MCFIF-1)
- Intmrgt 209 (SIC MCFIF-1)
- Intmrgt 210 (SIC MCFIF-1)
- Intmrgt 211 (SIC MCFIF-1)
- Intmrgt 212 (SIC MCFIF-1)
- Intmrgt 213 (SIC MCFIF-1)
- Intmrgt 214 (SIC MCFIF-1)
- Intmrgt 215 (SIC MCFIF-1)
- Intmrgt 216 (SIC MCFIF-1)
- Intmrgt 217 (SIC MCFIF-1)
- Intmrgt 218 (SIC MCFIF-1)
- Intmrgt 219 (SIC MCFIF-1)
- Intmrgt 220 (SIC MCFIF-1)
- Intmrgt 221 (SIC MCFIF-1)
- Intmrgt 222 (SIC MCFIF-1)
- Intmrgt 223 (SIC MCFIF-1)
- Intmrgt 224 (SIC MCFIF-1)
- Intmrgt 225 (SIC MCFIF-1)
- Intmrgt 226 (SIC MCFIF-1)
- Intmrgt 227 (SIC MCFIF-1)
- Intmrgt 228 (SIC MCFIF-1)
- Intmrgt 229 (SIC MCFIF-1)
- Intmrgt 230 (SIC MCFIF-1)
- Intmrgt 231 (SIC MCFIF-1)
- Intmrgt 232 (SIC MCFIF-1)
- Intmrgt 233 (SIC MCFIF-1)
- Intmrgt 234 (SIC MCFIF-1)
- Intmrgt 235 (SIC MCFIF-1)
- Intmrgt 236 (SIC MCFIF-1)
- Intmrgt 237 (SIC MCFIF-1)
- Intmrgt 238 (SIC MCFIF-1)
- Intmrgt 239 (SIC MCFIF-1)
- Intmrgt 240 (SIC MCFIF-1)
- Intmrgt 241 (SIC MCFIF-1)
- Intmrgt 242 (SIC MCFIF-1)
- Intmrgt 243 (SIC MCFIF-1)
- Intmrgt 244 (SIC MCFIF-1)
- Intmrgt 245 (SIC MCFIF-1)
- Intmrgt 246 (SIC MCFIF-1)
- Intmrgt 247 (SIC MCFIF-1)
- Intmrgt 248 (SIC MCFIF-1)
- Intmrgt 249 (SIC MCFIF-1)
- Intmrgt 250 (SIC MCFIF-1)
- Intmrgt 251 (SIC MCFIF-1)
- Intmrgt 252 (SIC MCFIF-1)
- Intmrgt 253 (SIC MCFIF-1)
- Intmrgt 254 (SIC MCFIF-1)
- Intmrgt 255 (SIC MCFIF-1)
- Intmrgt 256 (SIC MCFIF-1)
- Intmrgt 257 (SIC MCFIF-1)
- Intmrgt 258 (SIC MCFIF-1)
- Intmrgt 259 (SIC MCFIF-1)
- Intmrgt 260 (SIC MCFIF-1)
- Intmrgt 261 (SIC MCFIF-1)
- Intmrgt 262 (SIC MCFIF-1)
- Intmrgt 263 (SIC MCFIF-1)
- Intmrgt 264 (SIC MCFIF-1)
- Intmrgt 265 (SIC MCFIF-1)
- Intmrgt 266 (SIC MCFIF-1)
- Intmrgt 267 (SIC MCFIF-1)
- Intmrgt 268 (SIC MCFIF-1)
- Intmrgt 269 (SIC MCFIF-1)
- Intmrgt 270 (SIC MCFIF-1)
- Intmrgt 271 (SIC MCFIF-1)
- Intmrgt 272 (SIC MCFIF-1)
- Intmrgt 273 (SIC MCFIF-1)
- Intmrgt 274 (SIC MCFIF-1)
- Intmrgt 275 (SIC MCFIF-1)
- Intmrgt 276 (SIC MCFIF-1)
- Intmrgt 277 (SIC MCFIF-1)
- Intmrgt 278 (SIC MCFIF-1)
- Intmrgt 279 (SIC MCFIF-1)
- Intmrgt 280 (SIC MCFIF-1)
- Intmrgt 281 (SIC MCFIF-1)
- Intmrgt 282 (SIC MCFIF-1)
- Intmrgt 283 (SIC MCFIF-1)
- Intmrgt 284 (SIC MCFIF-1)
- Intmrgt 285 (SIC MCFIF-1)
- Intmrgt 286 (SIC MCFIF-1)
- Intmrgt 287 (SIC MCFIF-1)
- Intmrgt 288 (SIC MCFIF-1)
- Intmrgt 289 (SIC MCFIF-1)
- Intmrgt 290 (SIC MCFIF-1)
- Intmrgt 291 (SIC MCFIF-1)
- Intmrgt 292 (SIC MCFIF-1)
- Intmrgt 293 (SIC MCFIF-1)
- Intmrgt 294 (SIC MCFIF-1)
- Intmrgt 295 (SIC MCFIF-1)
- Intmrgt 296 (SIC MCFIF-1)
- Intmrgt 297 (SIC MCFIF-1)
- Intmrgt 298 (SIC MCFIF-1)
- Intmrgt 299 (SIC MCFIF-1)
- Intmrgt 300 (SIC MCFIF-1)
- Intmrgt 301 (SIC MCFIF-1)
- Intmrgt 302 (SIC MCFIF-1)
- Intmrgt 303 (SIC MCFIF-1)
- Intmrgt 304 (SIC MCFIF-1)
- Intmrgt 305 (SIC MCFIF-1)
- Intmrgt 306 (SIC MCFIF-1)
- Intmrgt

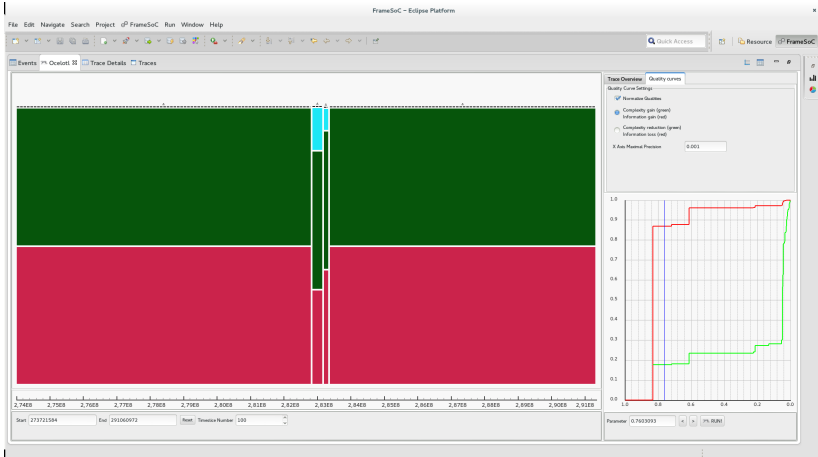


OUR PROPOSAL: OCELOTL

A temporal aggregation technique that provides an overview over time

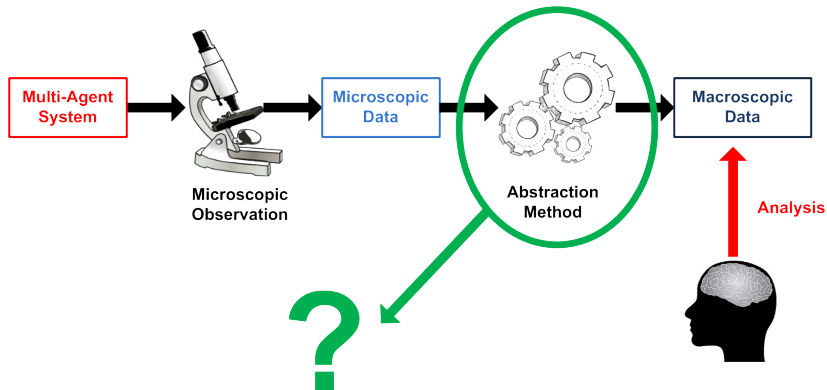
- We use a **meaningful aggregation** algorithm that gathers temporal parts of the trace where the behavior is similar
- The user chooses the **aggregation strength**
- He is aware of the **information loss**
- He can **interact**
- We manage up to 200 million events

EXAMPLE OF OCELOT VISUALIZATION

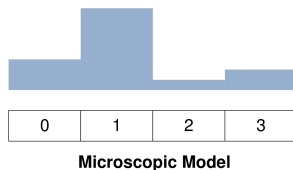


Temporal Aggregation Technique

LAMARCHE-PERRIN: BUILD A MACROSCOPIC MODEL

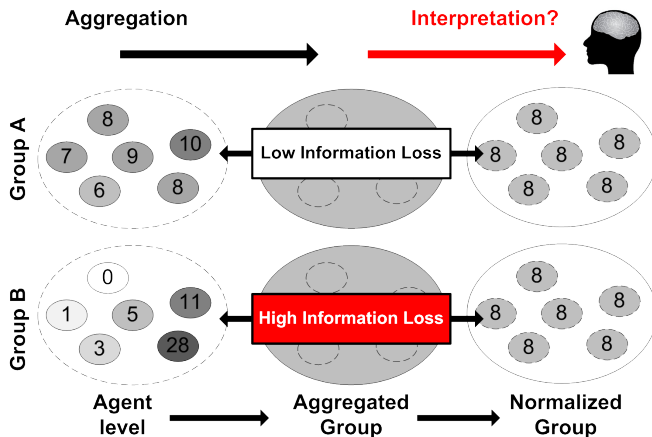


AGGREGATION METHODOLOGY



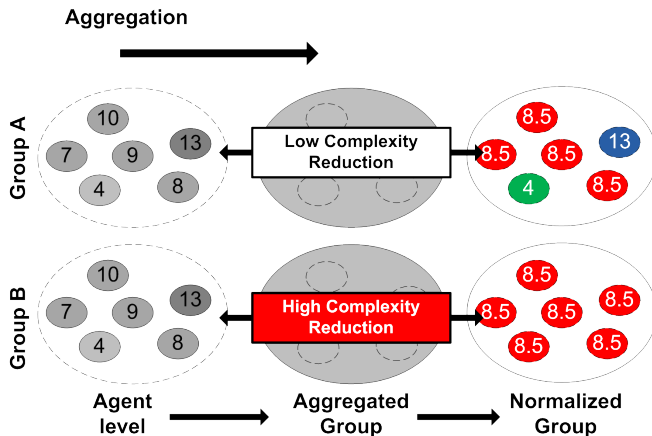
- Determine **the dimension(s)** to aggregate:
time
- Define a **microscopic model**:
time slices with associated values
- **Constrain** the aggregation:
only consecutive parts
- Choose a **criterion** that **triggers** the aggregation:
information and complexity
- Build an **algorithm** to provide a partition of the system

INFORMATION LOSS: KL DIVERGENCE



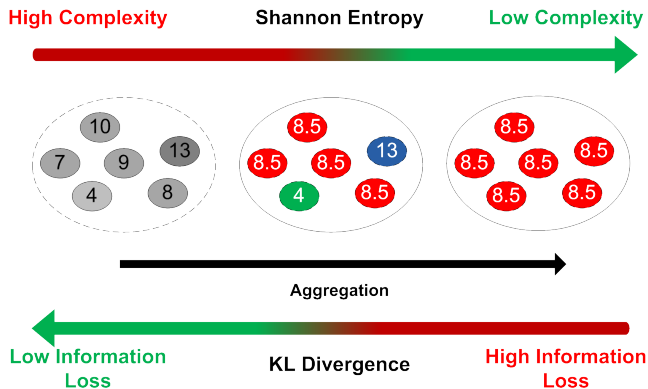
$$\text{loss}_X = \sum_{x \in X} \rho_x \log_2 \left(\frac{\rho_x}{\rho_X} \right)$$

COMPLEXITY REDUCTION: SHANNON ENTROPY



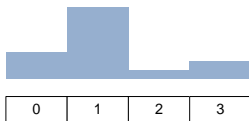
$$\text{gain}_X = \rho_X \log_2 - \rho_X \sum_{x \in X} \rho_x \log_2 \rho_x$$

COMPROMISE: PIC



$$pIC_X = p \text{ gain}_X - (1 - p) \text{ loss}_X$$

AGGREGATION ALGORITHM



Microscopic Model

0123			
012	123		
01	12	23	
0	1	2	3

Quality Measures X2 (Loss, Gain)

0123			
012	123		
01	12	23	
0	1	2	3

pIC for a given p

Best Cuts: select the highest pIC

$$\text{Best}_0 = 0$$

$$\text{Best}_{01} = \text{Best}_0 \ 1 \text{ or } 01$$

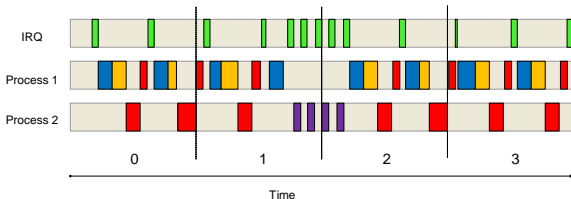
$$\text{Best}_{012} = \text{Best}_{01} \ 2 \text{ or } 012 \text{ or } \text{Best}_0 \ 12$$

$$\text{Best}_{0123} = \text{Best}_{012} \ 3 \text{ or } 0123 \text{ or } \text{Best}_0 \ 123 \text{ or } \text{Best}_{01} \ 23$$

Result:

$$\text{Best}_{01} \ 23 = \text{Best}_0 \ 1 \ 23 = 0 \ 1 \ 23$$

BUILD THE TRACE MICROSCOPIC MODEL



IRQ	0	0	0	0
Process 1	1	2.1	1	3
Process 2	4.1	2	4.1	4

IRQ	2	4.9	3	2.4
Process 1	0	0	0	0
Process 2	0	0	0	0

And so on...

COMPUTE THE QUALITIES

$$\begin{array}{|c|c|c|c|} \hline 0123 \\ \hline 012 & 123 & & \\ \hline 01 & 12 & 23 & \\ \hline 0 & 1 & 2 & 3 \\ \hline \end{array} = \begin{array}{|c|c|c|c|} \hline 0123 \\ \hline 012 & 123 & & \\ \hline 01 & 12 & 23 & \\ \hline 0 & 1 & 2 & 3 \\ \hline \end{array} + \begin{array}{|c|c|c|c|} \hline 0123 \\ \hline 012 & 123 & & \\ \hline 01 & 12 & 23 & \\ \hline 0 & 1 & 2 & 3 \\ \hline \end{array} + \dots$$

Red State Quality Measures

IRQ Red State Quality Measures

Process 1 Red State Quality Measures

$$\begin{array}{cccc} & & + & \\ \begin{array}{|c|} \hline 0123 \\ \hline \end{array} & & & \\ \begin{array}{|c|c|} \hline 012 & 123 \\ \hline \end{array} & & & \\ \begin{array}{|c|c|c|} \hline 01 & 12 & 23 \\ \hline \end{array} & & & \\ \begin{array}{|c|c|c|c|} \hline 0 & 1 & 2 & 3 \\ \hline \end{array} & = & \begin{array}{|c|} \hline 0123 \\ \hline \end{array} & + & \begin{array}{|c|c|} \hline 012 & 123 \\ \hline \end{array} & + & \begin{array}{|c|c|c|} \hline 01 & 12 & 23 \\ \hline \end{array} & + & \dots \end{array}$$

Green State Quality Measures

IRQ Green State Quality Measures

Process 1 Green State Quality Measures

+

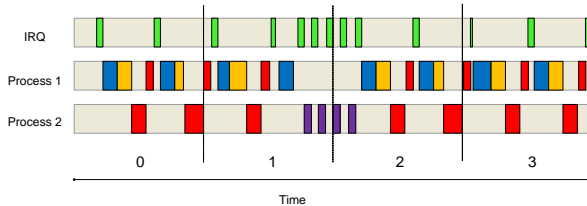
...

—

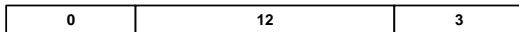
0123			
012	123		
01	12	23	
0	1	2	3

Quality Measures

REPRESENT THE AGGREGATES



Temporal Aggregation



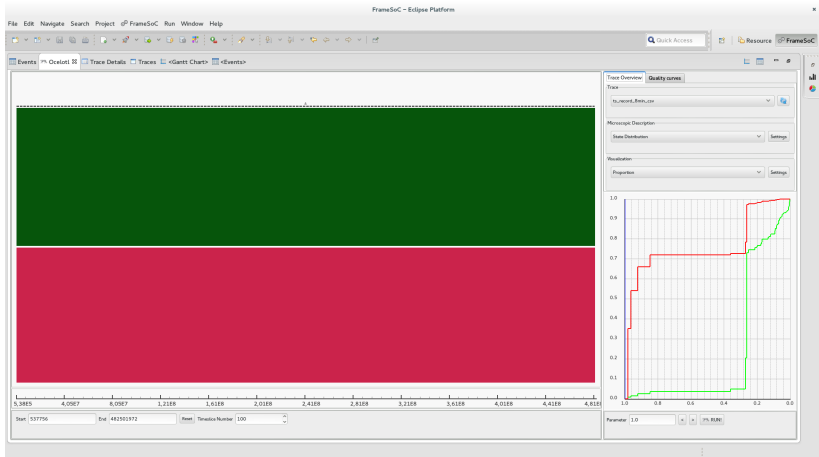
Spatial Aggregation



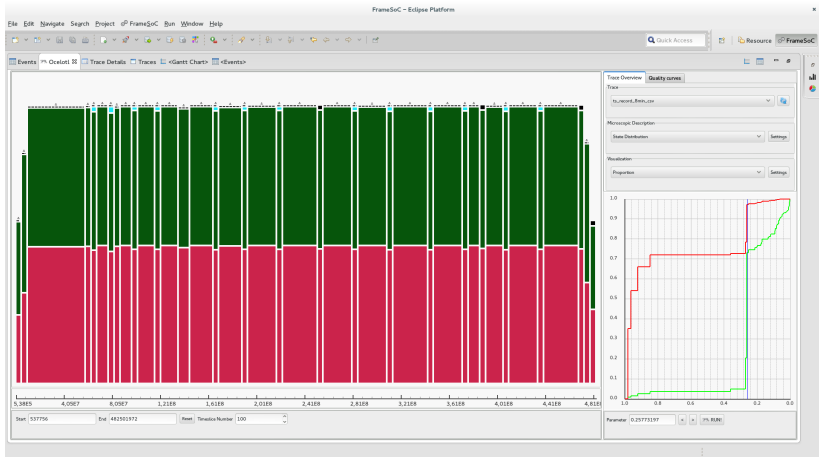
TS Record Analysis

25

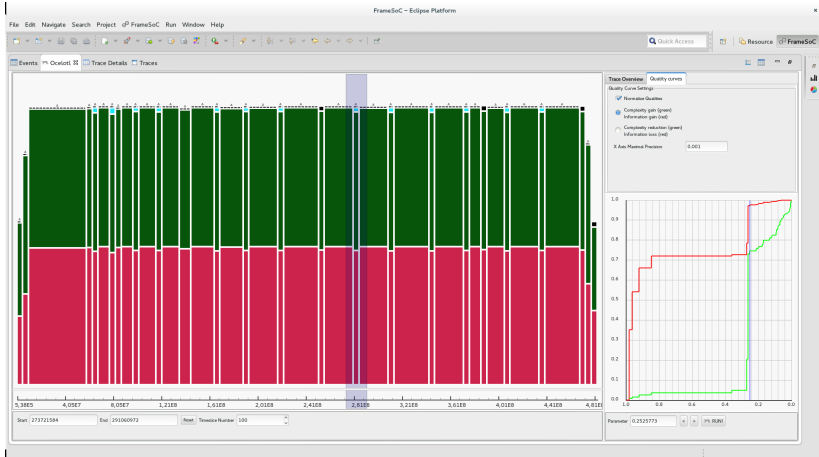
MACROSCOPIC DESCRIPTION



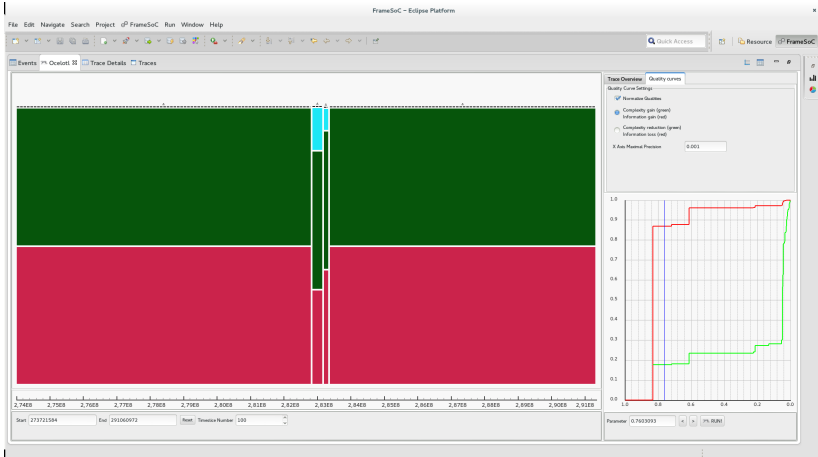
MESOSCOPIC DESCRIPTION



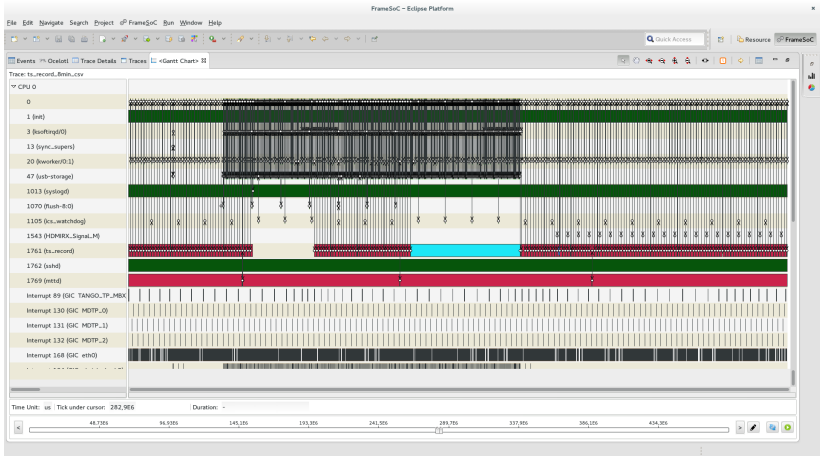
FOCUS ON AN AREA



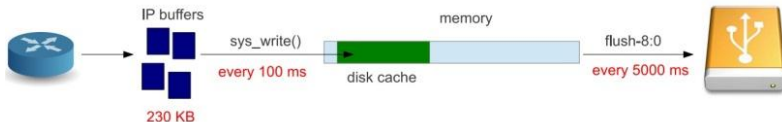
REPEAT THE AGGREGATION PROCESS



FIND THE BUG WITH THE GANTT CHART



ANALYSIS CONCLUSION



→ During the flush, the cache is **locked**, and `sys_write()` is **blocked**

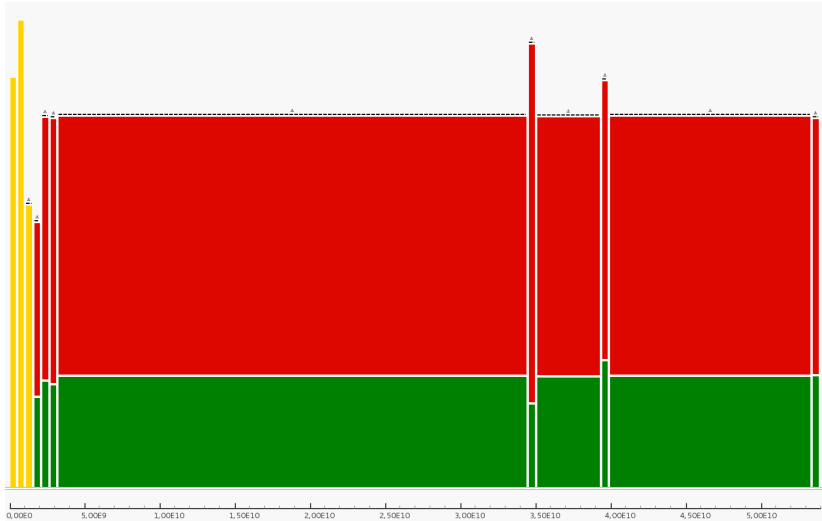
Problem

Writing on a USB disk is **slow**, IP buffers are often **overflowed**

Solution

Linux configuration: Wake up the flush according to the **data quantity**

OCELOTL APPLIED ON A PARALLEL APPLICATION TRACE

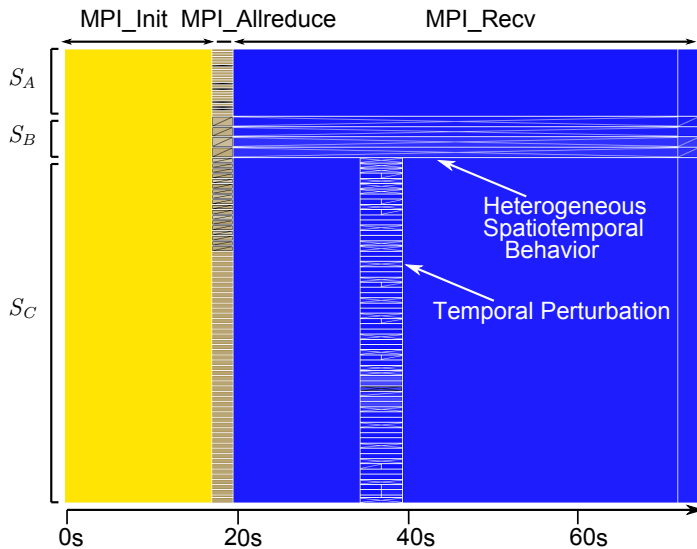


Grid'5000 Rennes, 64 Cores, NASBP CG, class C

CONCLUSION

- **Shneiderman** methodology respected
- Experiences show the ability to understand the **temporal behavior**
- We found why the episode of Game of Throne was not recorded correctly ;-)
- Able to work with **200 million events** (theoretically, no limit)
- Satisfying algorithm and interaction performance

CURRENT WORK : SPATIO TEMPORAL AGGREGATION



Grid'5000 Nancy, 700 Cores, NASBP LU, class C