A SPATIOTEMPORAL DATA AGGREGATION TECHNIQUE FOR PERFORMANCE ANALYSIS OF LARGE-SCALE EXECUTION TRACES

IEEE Cluster 2014 - 25th September 2014

Damien Dosimont [1 2], Robin Lamarche-Perrin [3],
Lucas M. Schnorr [4], Guillaume Huard [2 1], Jean-Marc Vincent [2 1]

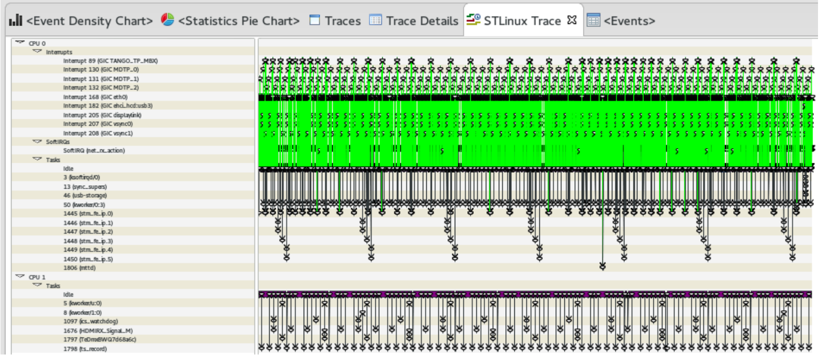
[1] Inria, [2] Univ. Grenoble Alpes, LIG, CNRS, F-38000 Grenoble, France
first.last@imag.fr
[3] MPI for Mathematics in the Sciences, 04103 Leipzig, Germany
robin.lamarche-perrin@mis.mpg.de
[4] Informatics Institute, UFRGS, Porto Alegre
schnorr@inf.ufrgs.br

INTRODUCTION

TRACE VISUALIZATION PROBLEMATIC

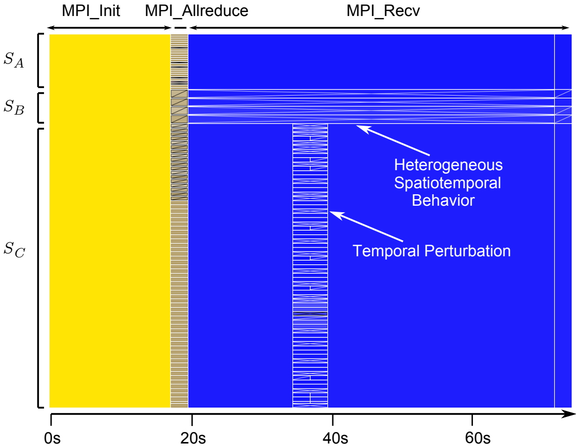
- **Trace contents**:
  - **SPACE** = application structure:
    - **hardware** components: *clusters, machines, cores, etc.*
    - **software** components: *processes, threads, etc.*
  - **TIME** = timestamped events:
    - *function calls, communications, CPU load, malloc, etc.*
- Traces can be **HUGE**
  $\rightarrow$ **scalability issues** of space-time representations

PROBLEMATIC VISUALIZATION



OBJECTIVE: SPATIOTEMPORAL OVERVIEW...

- Overcoming these issues thanks to **data aggregation**
- Showing **meaningful information**



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

TABLE OF CONTENTS

1. Introduction

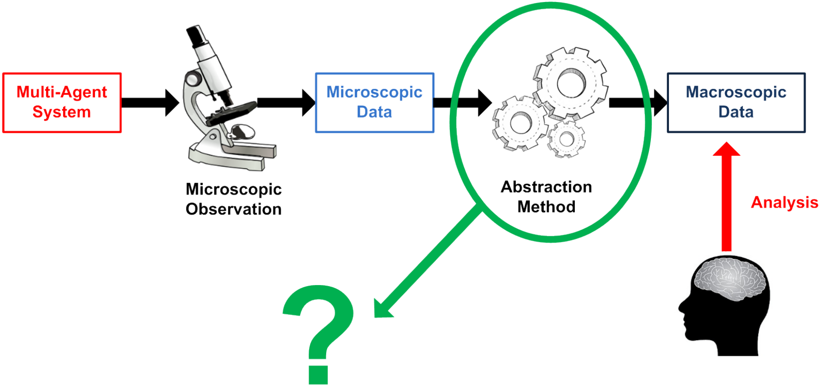
2. Previous works

3. Spatiotemporal data aggregation

4. Conclusion

PREVIOUS WORKS

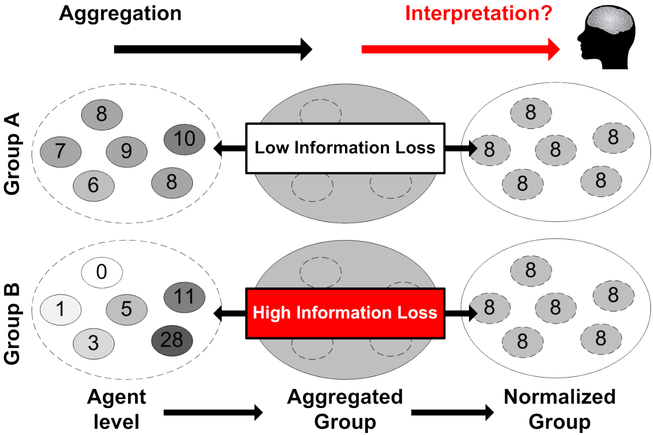
ADAPTING AN AGGREGATION METHODOLOGY



DATA AGGREGATION METHODOLOGY

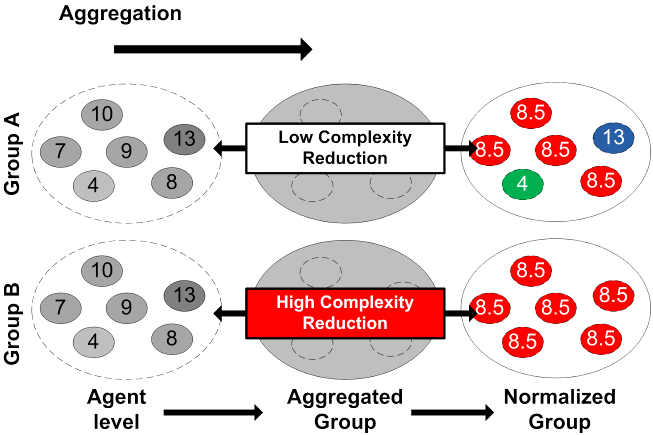
- A1. Choose a **model** and a **metric**
- A2. Choose on **which dimension(s)** aggregate
- A3. Define the **operands**
- A4. **Constrain** the aggregation : $\rightarrow$ partitions $\mathcal{P}$ allowed
- A5. Define the **operator**
- A6. Define the **trigger -** the aggregation condition
- A7. Build the **algorithm** satisfying A1-A6

INFORMATION LOSS: KL DIVERGENCE



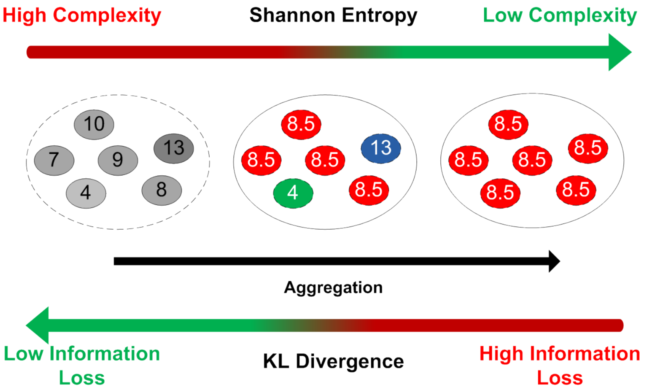
$$\mathrm{loss}_E = \sum_{e \in E} \rho_e \log_2 \left(\frac{\rho_e}{\rho_E} \right)$$

COMPLEXITY REDUCTION: SHANNON ENTROPY



$$\mathrm{gain}_E = \rho_E \log_2 \rho_E - \sum_{e \in E} \rho_e \log_2 \rho_e$$

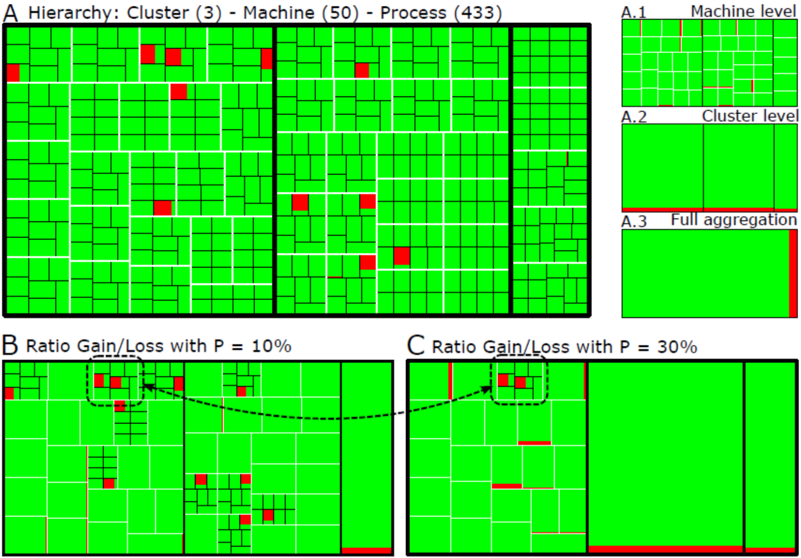
TRADE-OFF: PIC



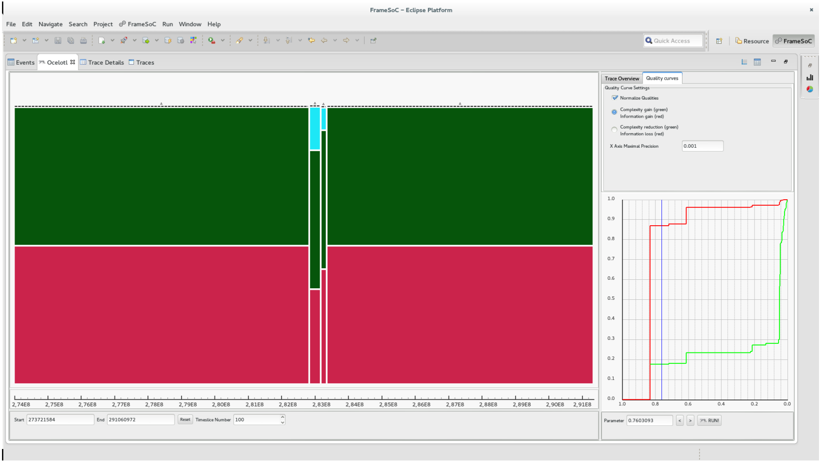
$$\mathrm{pIC}_E = p\,\mathrm{gain}_E - (1-p)\,\mathrm{loss}_E \qquad \mathrm{pIC}_{\mathcal{P}} = \sum_{E \in \mathcal{P}} \mathrm{pIC}_E$$

- For a given p : choose $\mathcal{P}$ with the highest pIC
- Aggregate in priority most homogeneous values

VIVA: SPATIAL AGGREGATION



OCELOTL: TEMPORAL AGGREGATION

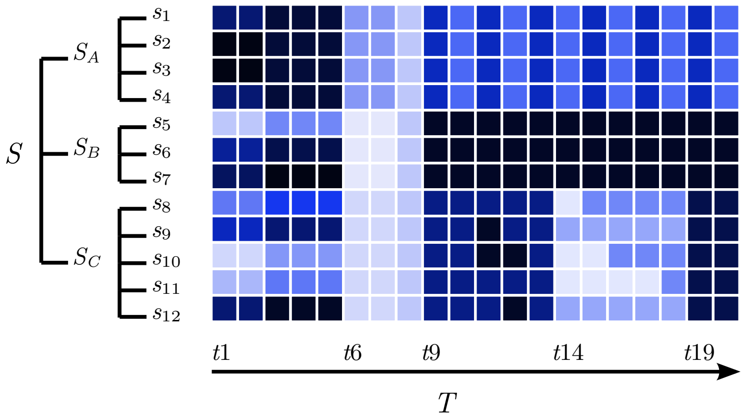


SPATIOTEMPORAL CRITERIA

- M1. Spatiotemporal representation
- M2. Aggregation coherence

SPATIOTEMPORAL DATA AGGREGATION

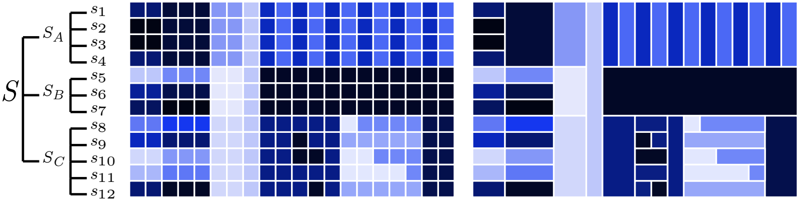
A.1 MICROSCOPIC MODEL



$$|X| = 2,\ \rho_x(s,t) = d_x(s,t)/d(t) \in [0,1],\ \rho_1(s,t) = 1 - \rho_2(s,t)$$

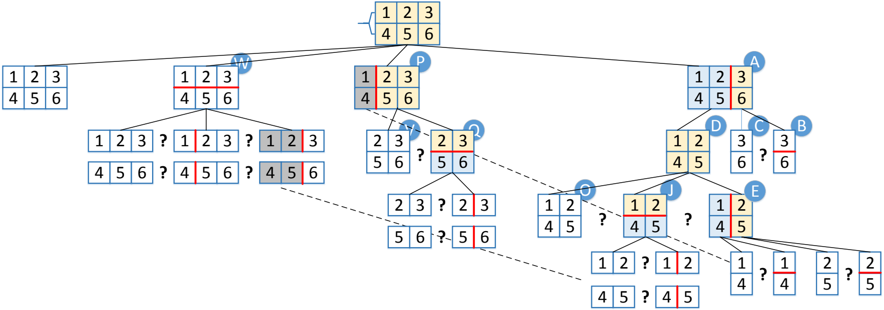
A2-A5

- A2. We aggregate simultaneously on T and S
- A3. Operands: $(s, t) \in S \times T$
- A4. Constraint: $\mathcal{A}(S \times T) = \mathcal{H}(S) \times \mathcal{I}(T)$
  Aggregation result is a partition $\mathcal{P}(S \times T) \in \boldsymbol{\mathcal{A}}(S \times T)$
- A5. Operator: $+$
- A6. Trigger: maximize pIC of the partition $\mathcal{P}(S \times T)$

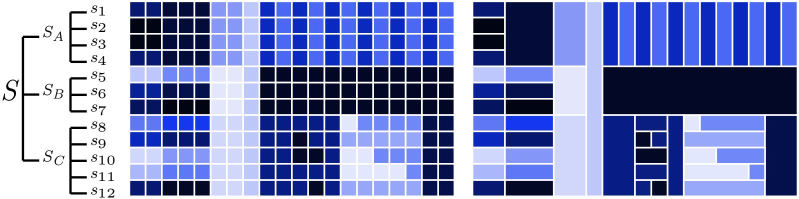


BEST CUT ALGORITHM

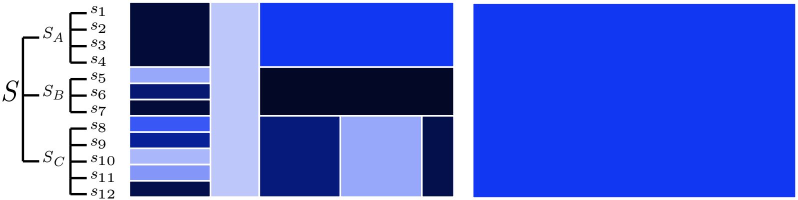
- Compute the partition with the highest pIC :
  - Cut an area : time, space (or no cut)
  - Best cut: the partition $\mathcal{P}$ where $\sum_{E \in \mathcal{P}} \mathrm{pIC}_E$ is max
  - Recursively cut and evaluate the partitions of $E_1, E_2 \in \mathcal{P}$
  - Useless recomputation is avoided



INFLUENCE OF P

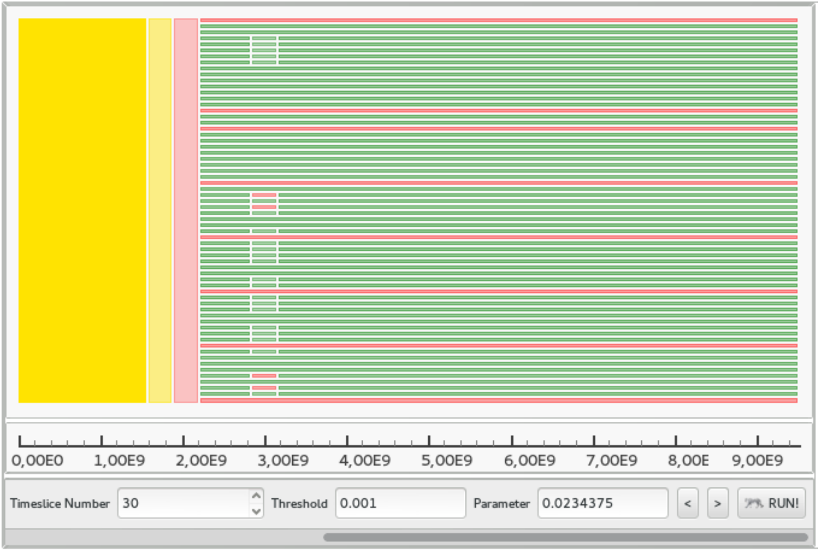


$p = 0$ $\quad$ $p = a,\ 0 < a < 1$

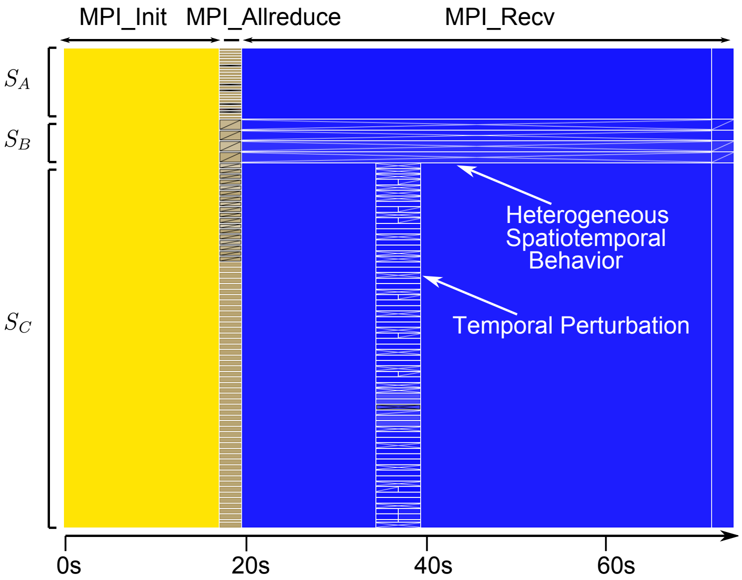


$p = b,\ 0 < a < b < 1$ $\quad$ $p = 1$

CG CLASS C, 64 PROCESSES ON G5K RENNES



LU CLASS C, 700 PROCESSES ON G5K NANCY



PERFORMANCES

	Case A	Case B	Case C	Case D
Application	CG, class C	CG, class C	LU, class C	LU, class B
Processes	64	512	700	900
Site	Rennes	Grenoble	Nancy	Rennes
Clusters (nodes)	parapide(8)	adonis(9), edel(24), genepi(31)	graphene(26), graphite(4), griffon(67)	paradent(38), parapide(21), parapluie(18)
Event number	3,838,144	49,149,440	218,457,456	177,376,729
Trace size	136.9 MB	1.8 GB	8.3 GB	6.7 GB

Ocelotl computation times (30 time slices)				
Trace reading + Microscopic description	5 s	31 s	222 s	174 s
Aggregation	$<$1s	$<$1s	2s	2s

CONCLUSION

CONCLUSION

- Visualization based on spatiotemporal data aggregation
  - Solves screen, computing and analyst capability limitations
  - Gives meaningful information about homogeneity (phases, perturbations)
  - Two use cases show its relevancy
- Future work :
  - Improve visualization and interaction to get more details
  - Extend methodology and design new algorithms ($\mathcal{H}(S) \times \mathcal{H}(S) \times \mathcal{I}(T)$, surface, etc.)

LINKS

Ocelotl:

http://github.com/dosimont/ocelotl

Framesoc:

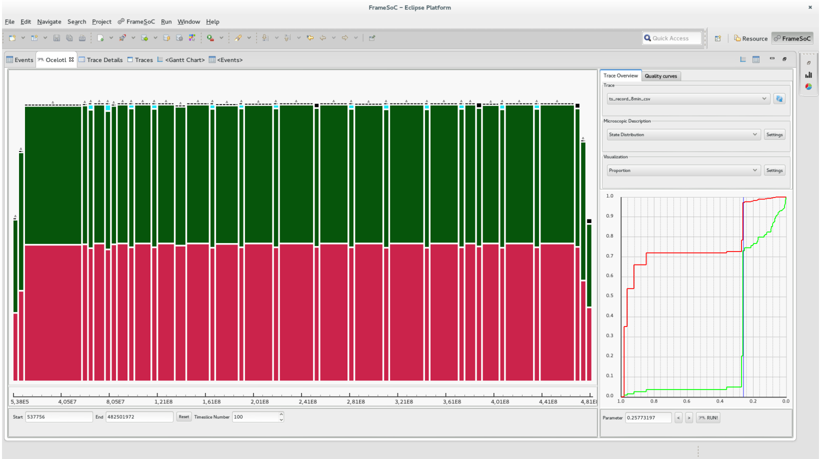
http://github.com/generoso/framesoc

Viva:

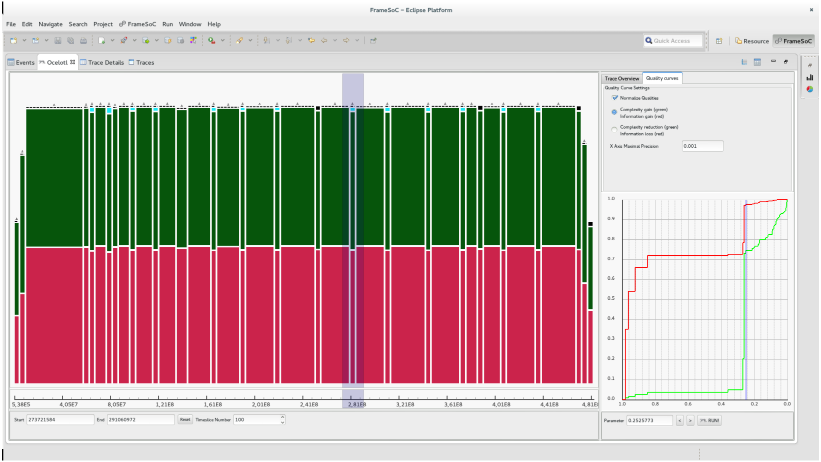
http://github.com/schnorr/viva

THANK YOU FOR YOUR ATTENTION

OCELOTL: TEMPORAL AGGREGATION (1)



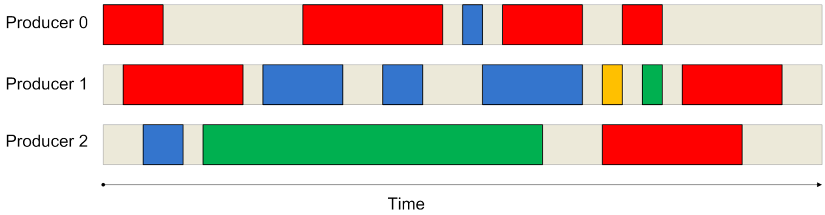
OCELOTL: TEMPORAL AGGREGATION (2)



ELMQVIST-FEKETE CRITERIA

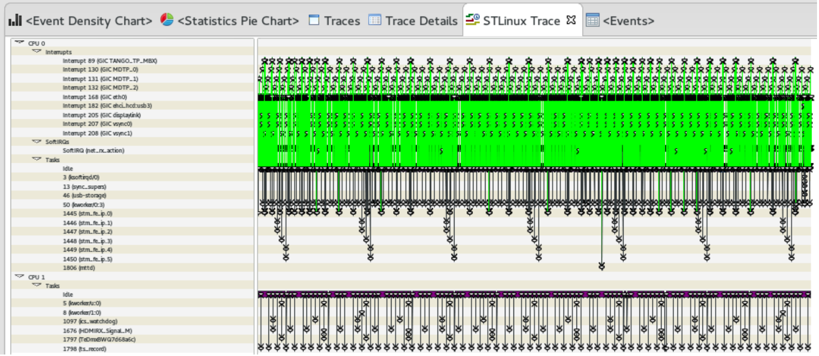
- **Shneiderman** : **overview**, zoom and filter, then get details on demand
- **Elmqvist & Fekete**: guidelines to design an **overview** visualization based on hierarchical aggregation
  - G1. Entity Budget
  - G2. Visual Summary
  - G3. Visual Simplicity
  - G4. *Discriminability*
  - G5. Fidelity
  - G6. *Interpretability*

VISUALIZATIONS NOT FULFILLING THESE CRITERIA (1)



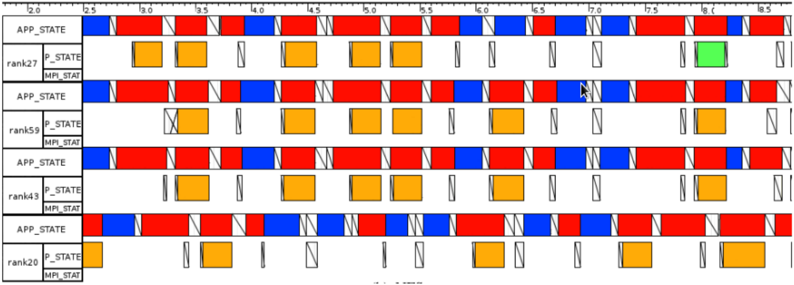
Example of Gantt chart - space-time diagram

VISUALIZATIONS NOT FULFILLING THESE CRITERIA (2)



KPTrace: $\overline{\text{G1}}$ (time), $\overline{\text{G2}}$, $\overline{\text{G4}}$, $\overline{\text{G5}}$

VISUALIZATIONS NOT FULFILLING THESE CRITERIA (2)



Pajé: $\overline{G1}$ (space), $\overline{G2}$