

www.bsc.es



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación

Introduction to Paraver

tools@bsc.es

SoHPC – Barcelona – June 2015

Humans are visual creatures

« Films or books?

- Two hours vs. days (months)

« Memorizing a deck of playing cards

- Each card translated to an image (person, action, location)

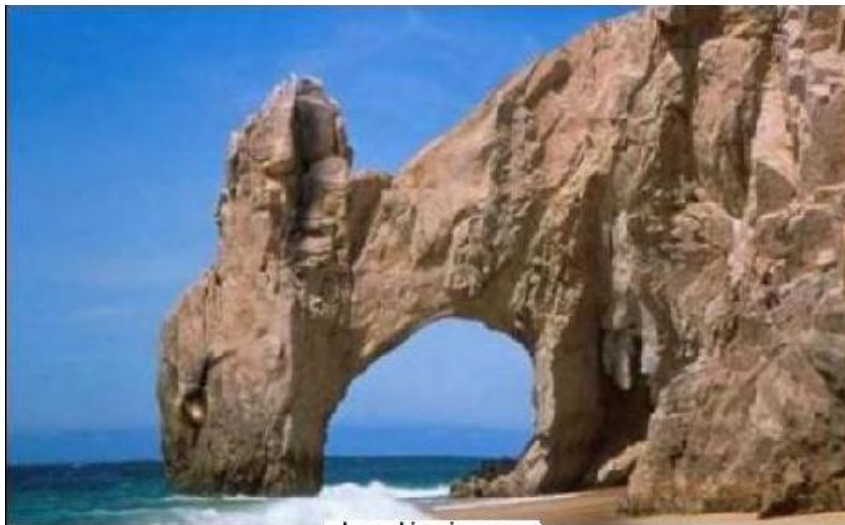
« Our brain loves pattern recognition

- What do you see on the pictures?

PROCESS

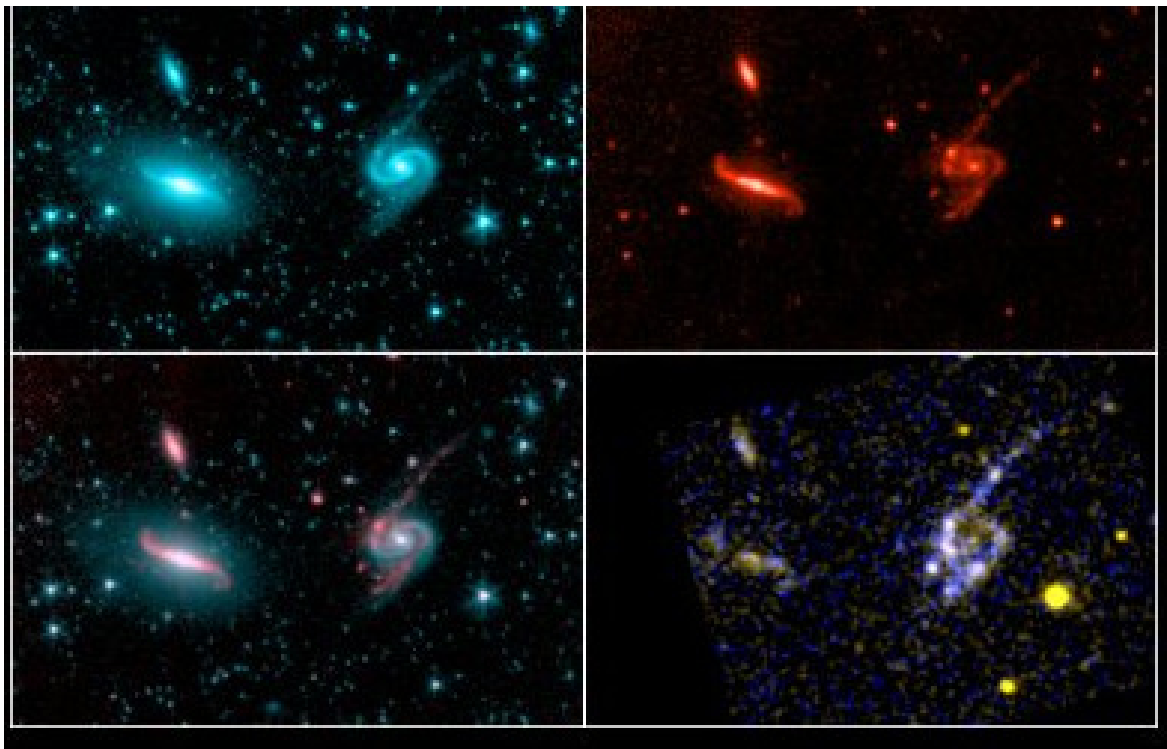
STORE

IDENTIFY



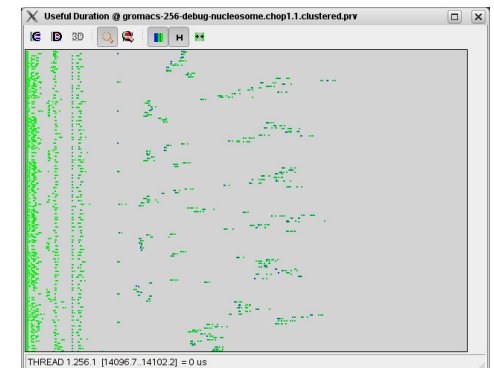
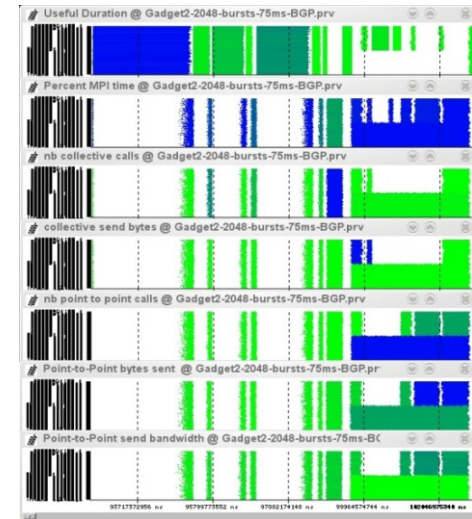
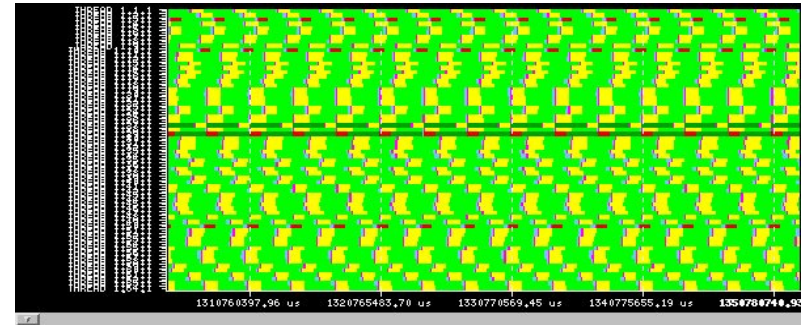
Multispectral imaging

- « Different looks at one reality
 - Different spectral bands (light sources and filters)
- « Highlight different aspects
 - Can combine into false colored but highly informative images



Our Tools

- « Since 1991
- « Based on traces
- « Open Source
 - <http://www.bsc.es/paraver>
- « Core tools:
 - Paraver (paramedir) – offline trace analysis
 - Dimemas – message passing simulator
 - Extrae – instrumentation
- « Focus
 - Detail, variability, flexibility
 - Behavioral structure vs. syntactic structure
 - Intelligence: Performance Analytics



www.bsc.es

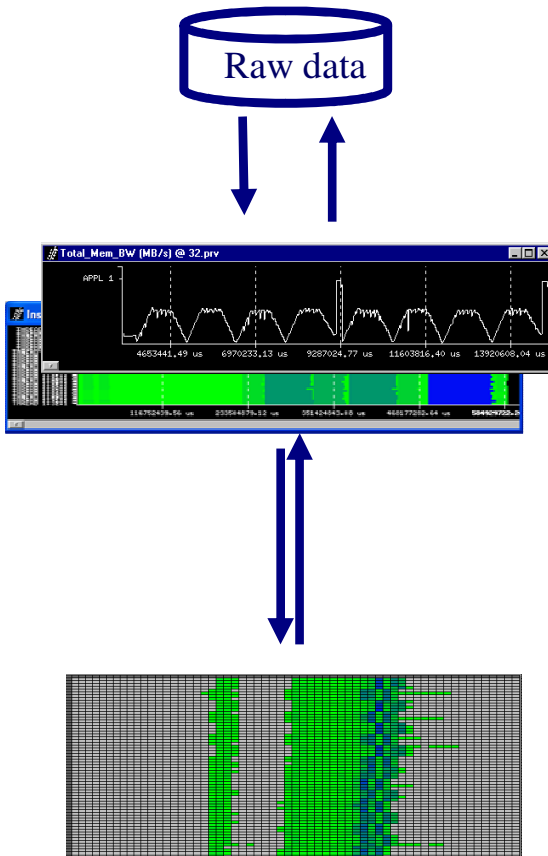


**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación

Paraver

Paraver – Performance data browser



Trace visualization/analysis

+ trace manipulation

Timelines

Goal = Flexibility

No semantics

Programmable

**2/3D tables
(Statistics)**

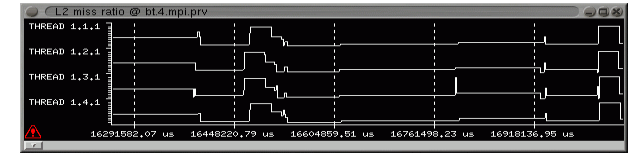
Comparative analyses

Multiple traces

Synchronize scales

Timelines

- Each window displays one view
 - Piecewise constant** function of time



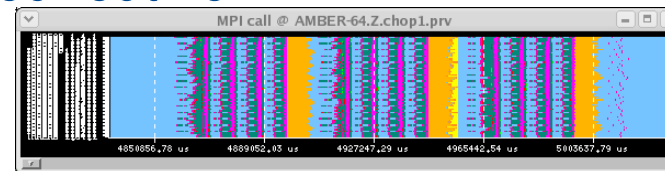
$$s(t) = S_i, i \in [t_i, t_{i+1})$$

- Types of functions

- Categorical
 - State, user function, outlined routine

$$S_i \in [0, n] \subset N, \quad n <$$

- Logical



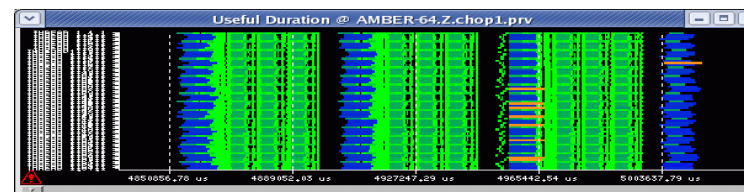
$$S_i \in \{0, 1\}$$

- In specific user function, In MPI call, In long MPI call

- Numerical

$$S_i \in R$$

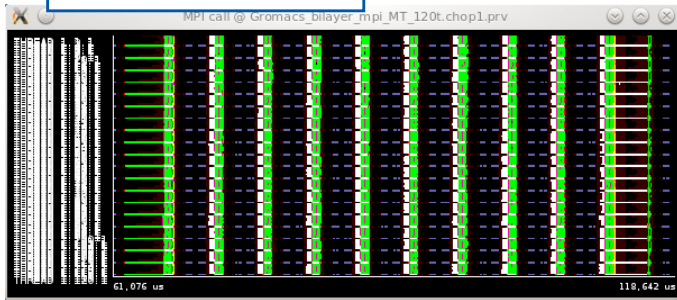
- IPC, L2 miss ratio, Duration of MPI call, duration of computation burst



Tables: Profiles, histograms, correlations

« From timelines to tables

MPI calls



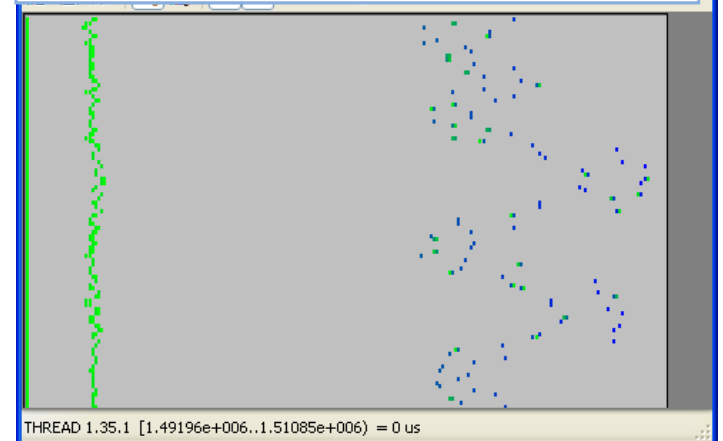
MPI calls profile

	Outside MPI	MPI_Send	MPI_Recv	MPI_Isend	MPI_Irecv	MPI_Waitall	MPI_Bcast	MPI_Reduce	MPI_Allr
THREAD 1.113.1	67.6081 %	0.0682 %	9.9182 %	2.5777 %	1.7698 %	5.1676 %	0.5934 %	0.1465 %	
THREAD 1.114.1	42.8434 %	-	20.5621 %	1.1947 %	1.0400 %	7.7056 %	-	-	
THREAD 1.115.1	68.6127 %	0.0707 %	9.6223 %	2.2589 %	2.0177 %	5.9825 %	0.5249 %	0.0297 %	
THREAD 1.116.1	74.6039 %	0.0531 %	9.6084 %	2.8813 %	2.5593 %	2.9286 %	0.5095 %	0.0483 %	
THREAD 1.117.1	74.3733 %	0.0691 %	9.7012 %	2.8517 %	2.5240 %				
THREAD 1.118.1	72.7770 %	0.0545 %	9.5489 %	2.8489 %	2.5353 %				
THREAD 1.119.1	66.7994 %	0.0682 %	10.0674 %	2.4206 %	1.9741 %				
THREAD 1.120.1	43.7224 %	-	20.5273 %	1.1912 %	1.0175 %				
Total	8,012.4546 %	7.3174 %	1,370.5276 %	288.6168 %	253.0137 %	54			
Average	66.7705 %	0.0690 %	11.4211 %	2.4051 %	2.1084 %				
Maximum	75.6821 %	0.4390 %	21.2505 %	2.9706 %	2.6369 %				
Minimum	40.5200 %	0.0129 %	8.8583 %	1.1489 %	1.0077 %				
StDev	11.3685 %	0.0474 %	4.0613 %	0.5984 %	0.5406 %				
Avg/Max	0.8822	0.1572	0.5374	0.8096	0.7998				

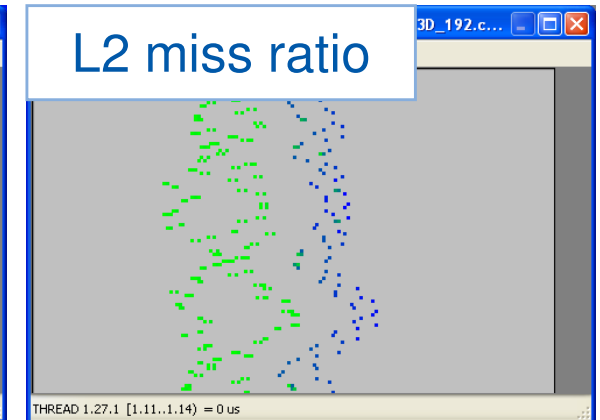
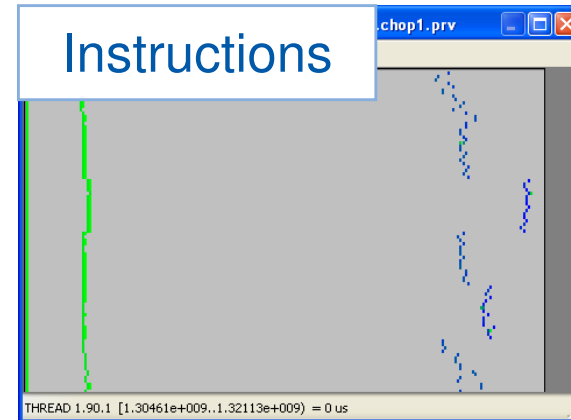
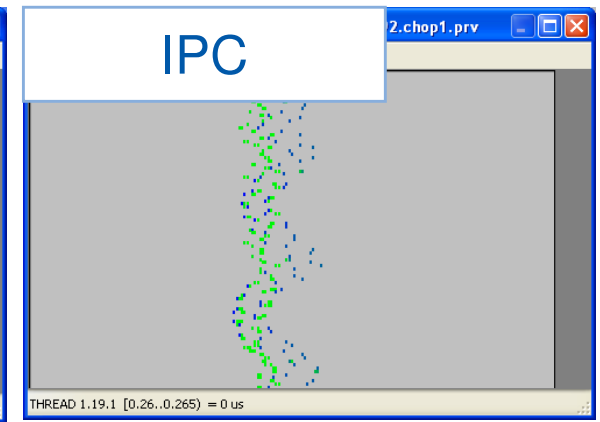
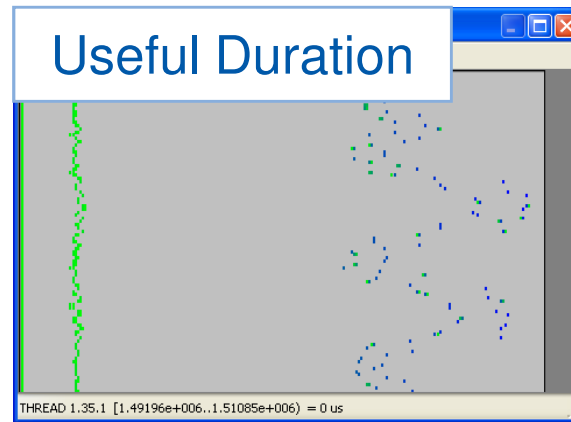
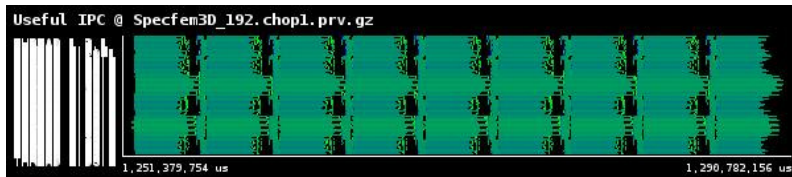
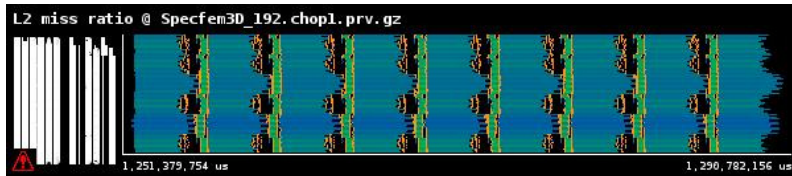
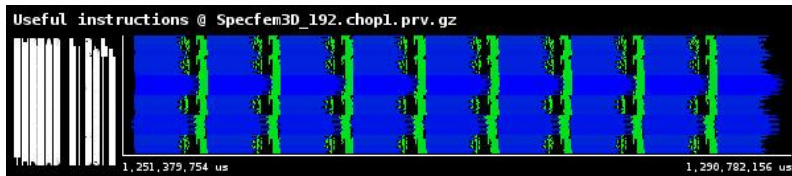
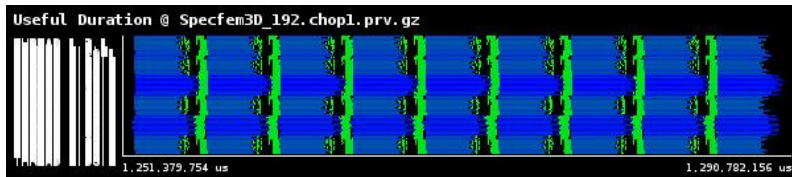
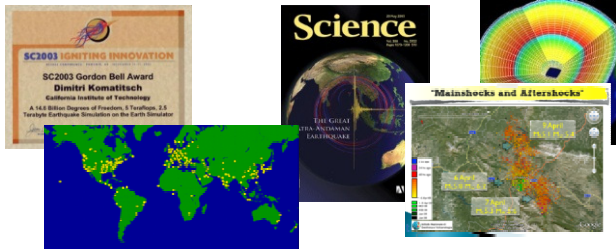
Useful Duration



Histogram Useful Duration

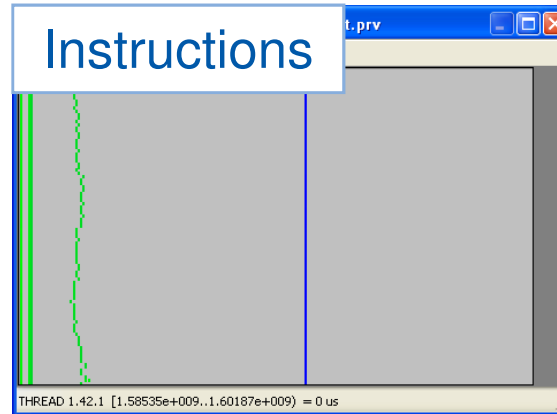
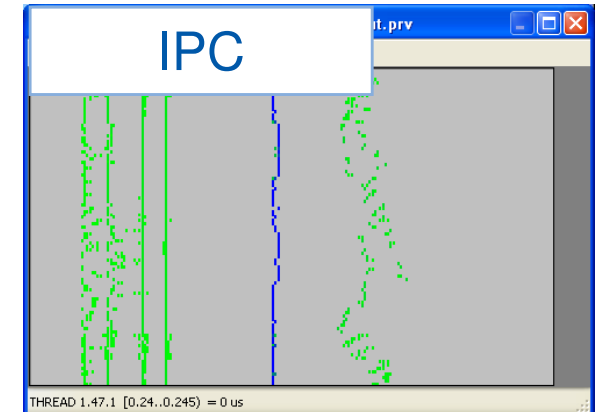
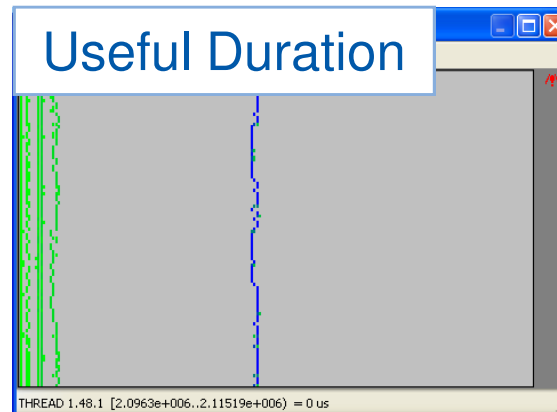


Analyzing variability through histograms and timelines



Analyzing variability through histograms and timelines

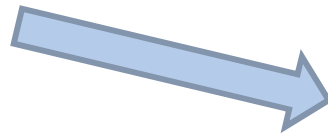
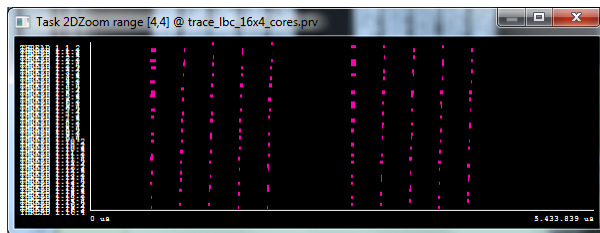
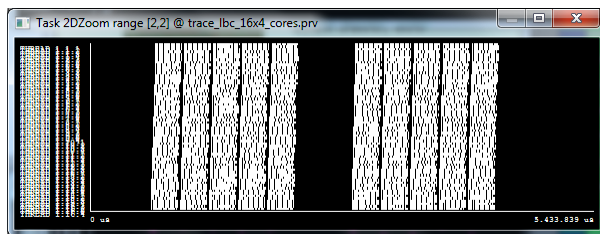
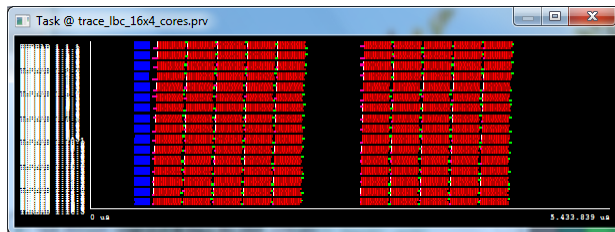
« By the way: six months later



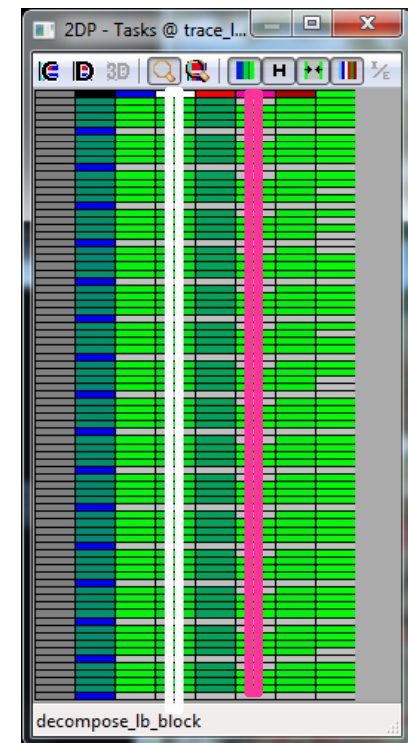
From tables to timelines

« Where in the timeline do the values in certain table columns appear?

– ie. want to see the time distribution of a given routine?

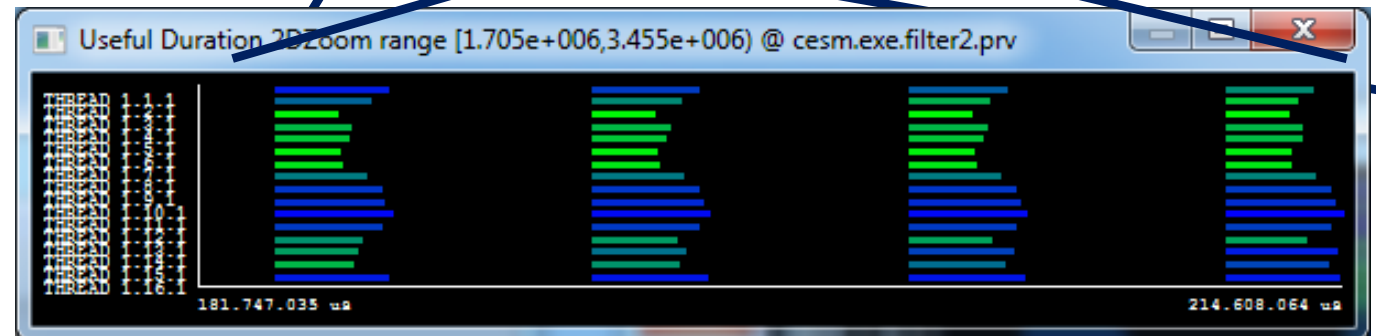
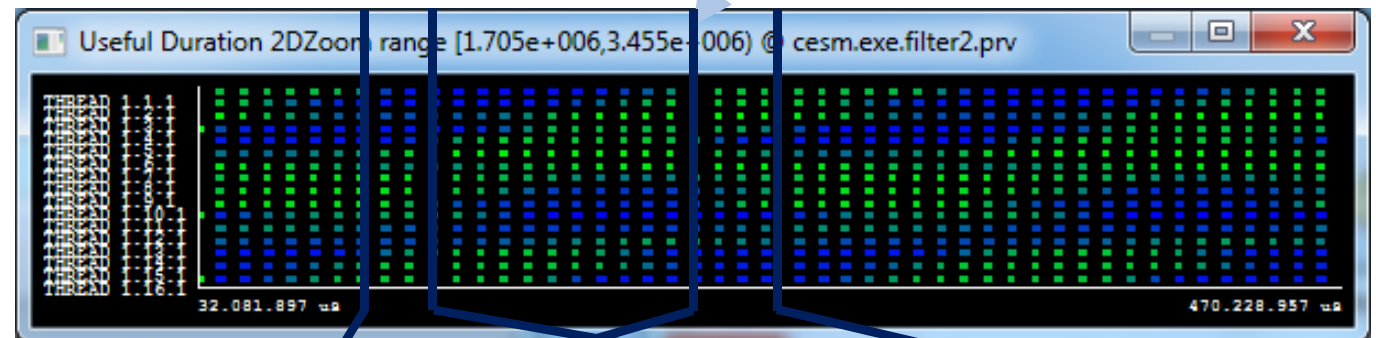
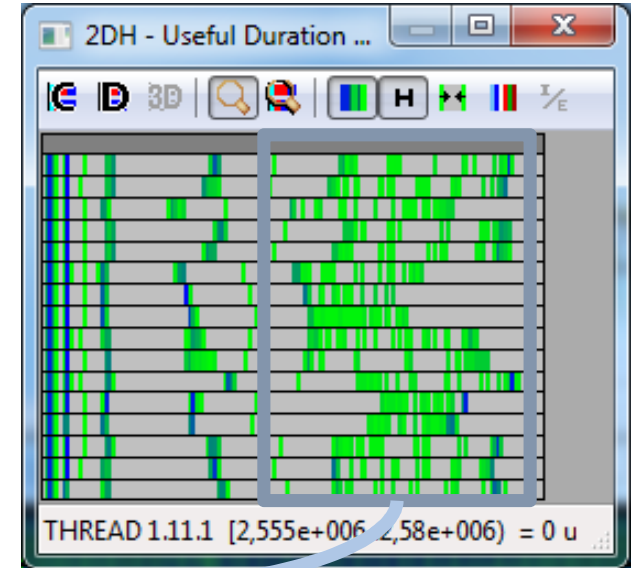


Only showing when a given value happens



Variability ... is everywhere

- « CESM: 16 processes, 2 simulated days
- « Histogram useful computation duration shows high variability
- « How is it distributed?
- « Dynamic imbalance
 - In space and time
 - Day and night.
 - Season ? ☺



Trace manipulation

« Data handling/summarization capability

– Filtering

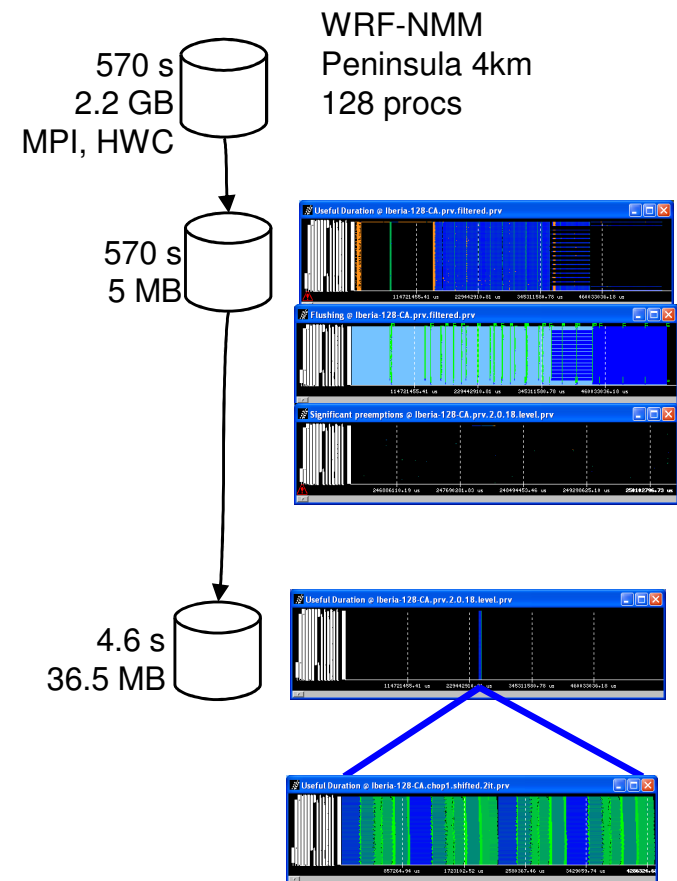
- Subset of records in original trace
- By duration, type, value,...
- Filtered trace IS a paraver trace and can be analysed with the same cfgs (as long as needed data kept)

– Cutting

- All records in a given time interval
- Only some processes

– Software counters

- Summarized values computed from those in the original trace emitted as new even types
- #MPI calls, total hardware count,...



« www.bsc.es/paraver

- downloads
 - Sources / Binaries
 - Linux / windows / MAC
- documentation
 - Training guides
 - Tutorial slides

« Getting started

- Start wxparaver
- Help → tutorials and follow instructions
- Follow training guides
 - Paraver introduction (MPI): Navigation and basic understanding of Paraver operation

Conclusions

- ⌘ We should take benefit of our visual brain
 - Effort = learn how to read views
- ⌘ Astrophysics enjoy their work :)
- ⌘ Look at your codes!

www.bsc.es/paraver



www.bsc.es



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación

Paraver Demos