

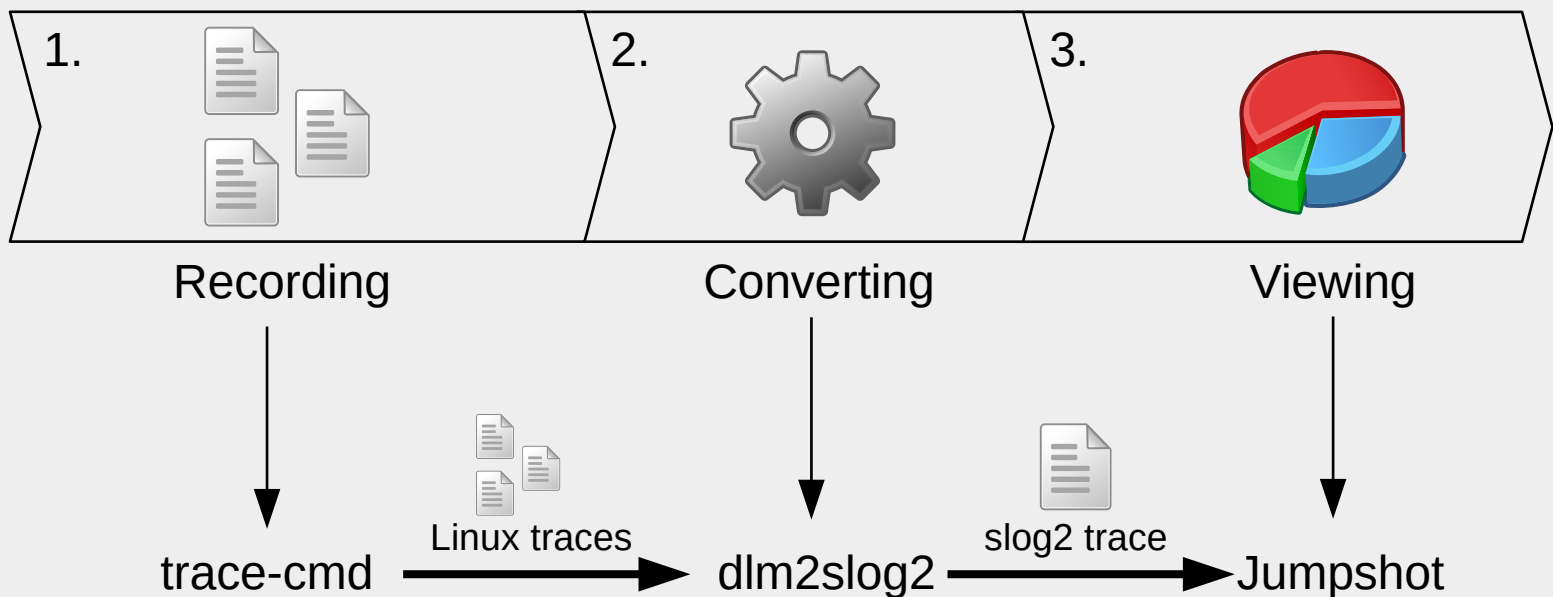
bring network and time together using Linux tracing

Alexander Aring

What to expect from this talk?

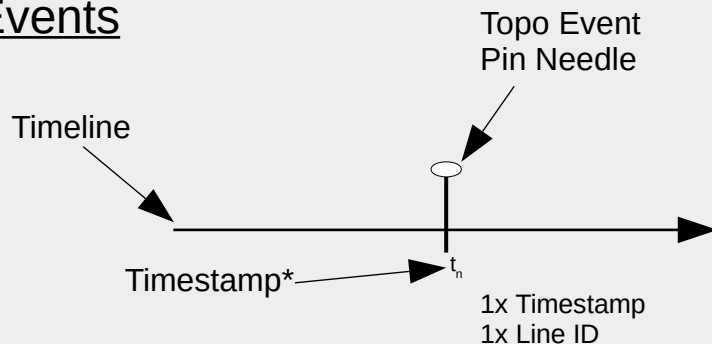
- Every Timeline shown is time synchronized
- Software to Visualize DLM (net protocol)
 - trace-cmd (Upstream User Space Tool, C)
 - slog2sdk (Tracing Framework for MPI Apps, Java)
 - dlm2slog2 (DLM Trace Converter “bridge”)
- How slog2sdk visualize traces for “MPI”
- DLM is an Application Example here!

Our Pipeline Steps

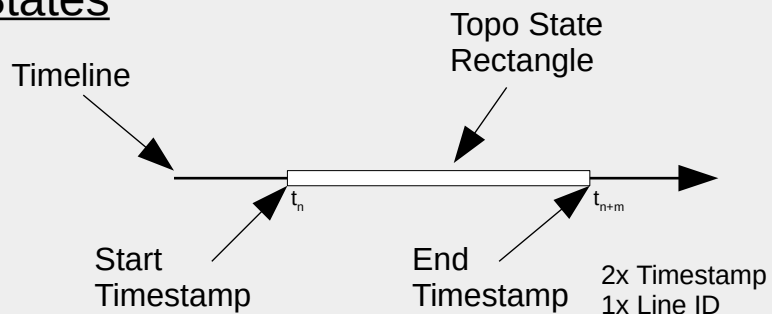


Visualizing Traces (slog2)

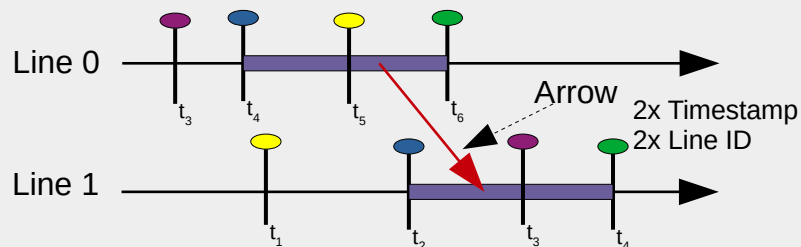
Events



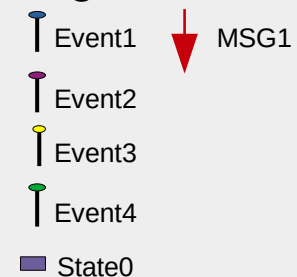
States



Example (GANTT)

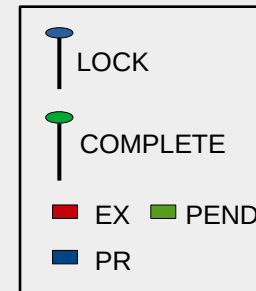
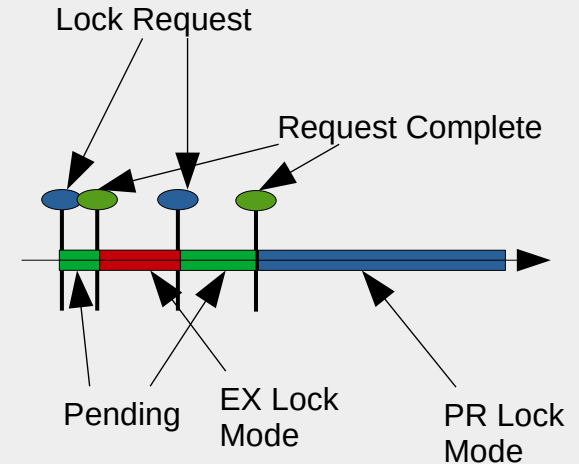


Legend

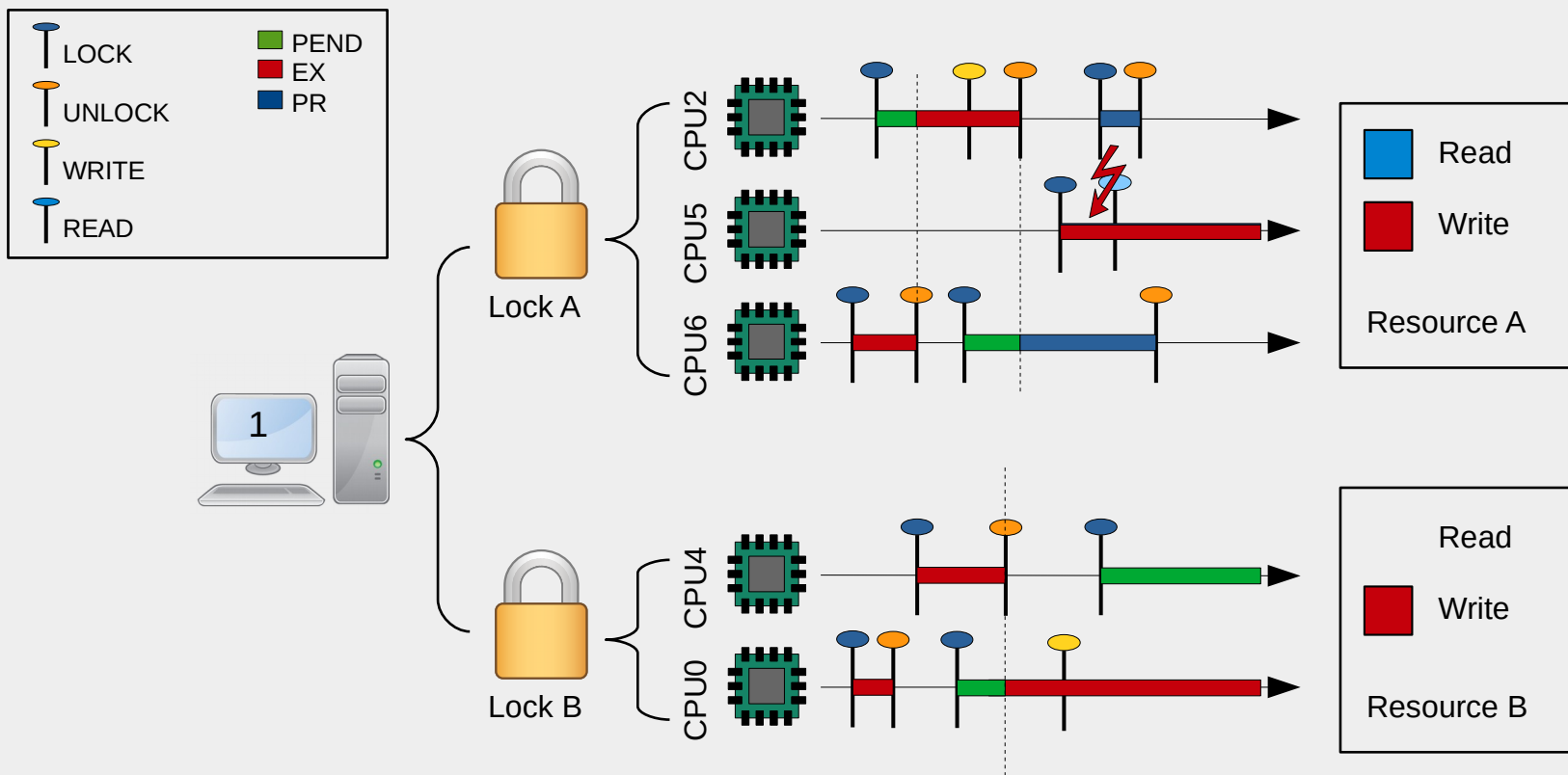


Linux Distributed Lock Manager

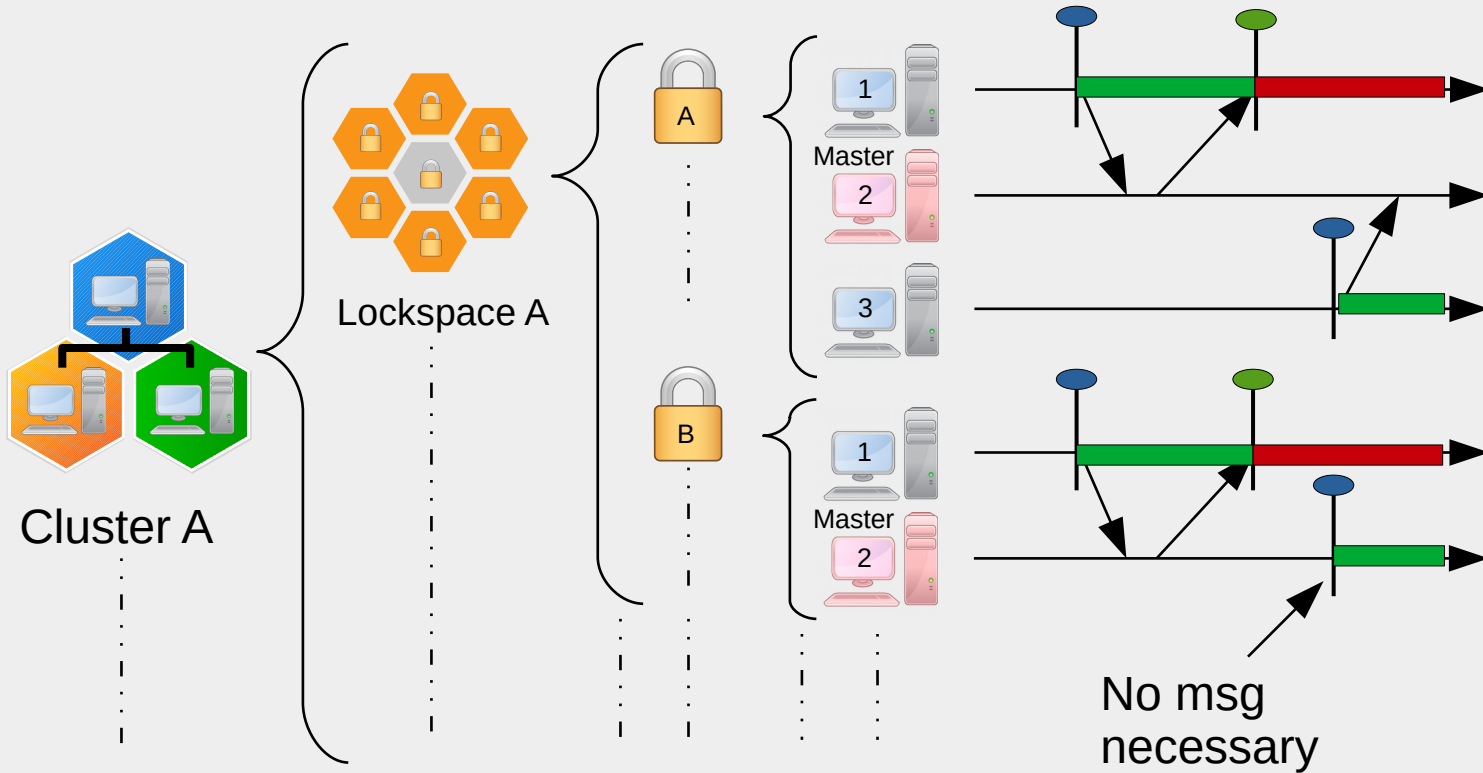
- Networking Protocol
- Requires Cluster Manager
- Asynchronous API
- Locks have Lock Modes
- Principle Lock Masters



Traced Local Locking

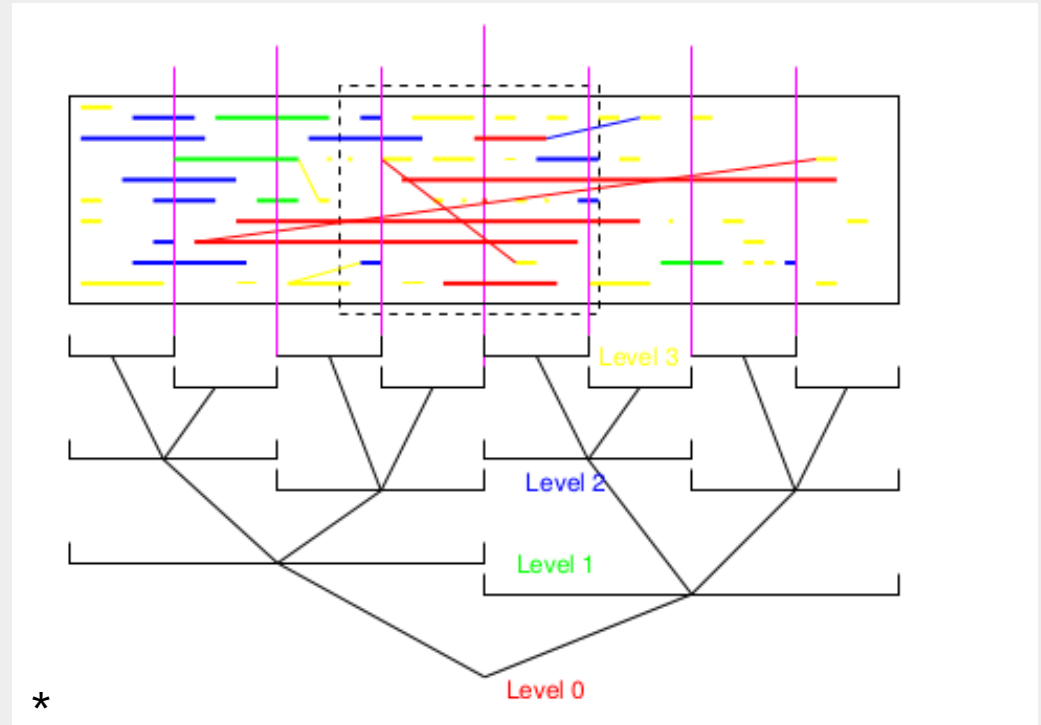


DLM Linux-Cluster Hierachy



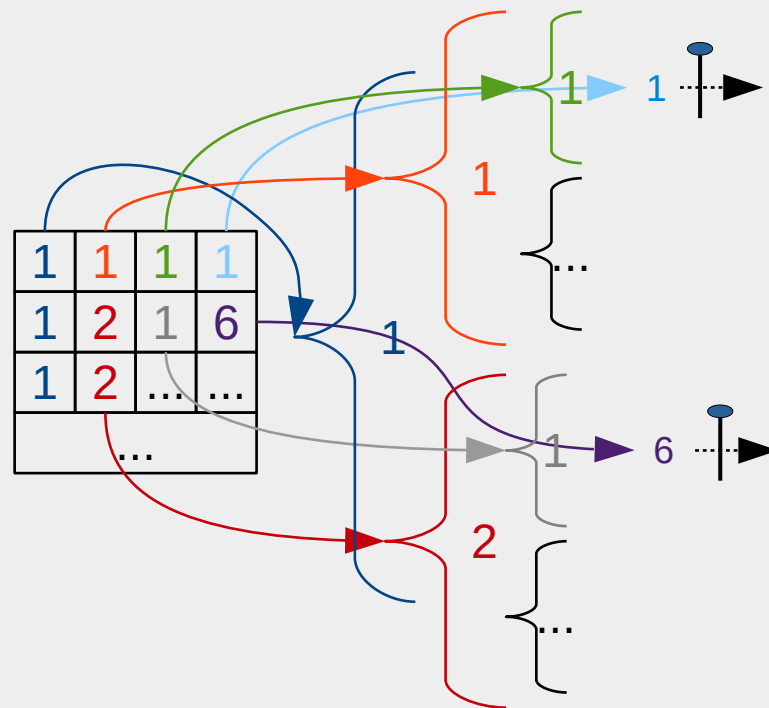
SLOG2 Timeline Scaling (Kshark?)

- YA Fileformat?
- Large Records
- Level Of Detail
- Partial-Read
- Zooming
- X-Axis Scaling
- Y-Axis?



SLOG2 YCoordMap

- Multidimensional Array
- Columns for Y-Axis
- Column to LineID*
- DLM Cluster Hierarchy



dIm2slog2

1) Analyze

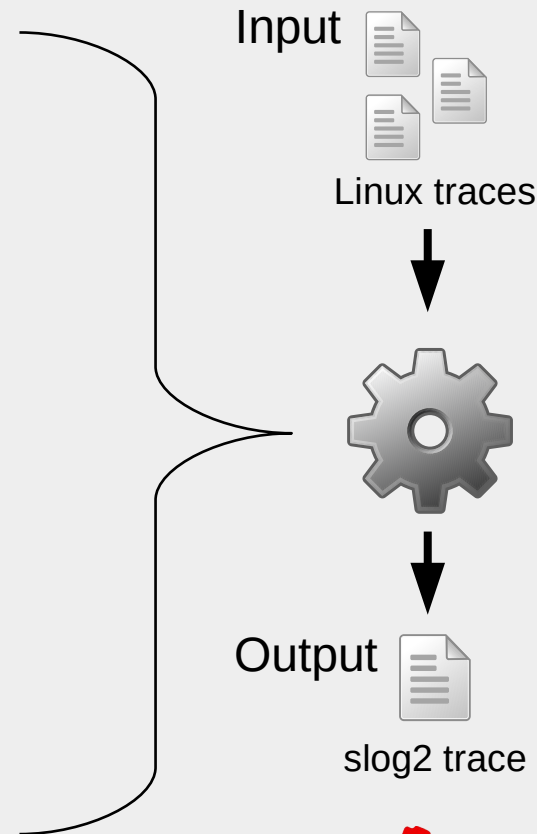
- Java-Bindings (Parsing)
- Overview about Locks

2) User Filtering

- Provide Way Scale Y-Axis

3) Generate

- Events and States
- Generates YCoordMap



dlm2slog2



Menu About

Add Node

- Nodeid: 1	Select Linux Trace Filepath	./examples/dlm_traces/trace-5.dat
- Nodeid: 2	Select Linux Trace Filepath	./examples/dlm_traces/trace-4.dat
- Nodeid: 3	Select Linux Trace Filepath	./examples/dlm_traces/trace-3.dat

Analyze Resources

Available Resources				Filtered Resources			
#Events	Lockspace ID	Resource Name		#Events	Lockspace ID	Resource Name	
315	0x08e4f48a18	3	1d9d9f	1551	0x08e4f48a18	2	450ea7
315	0x08e4f48a18	3	805d	1063	0x08e4f48a18	3	449eb3
308	0x08e4f48a18	3	11	370	0x08e4f48a18	5	451252
266	0x08e4f48a18	2	45107a	364	0x08e4f48a18	3	4037
263	0x08e4f48a18	3	91120f	327	0x08e4f48a18	3	1f9b37
246	0x08e4f48a18	5	451121	299	0x08e4f48a18	5	45107a
182	0x08e4f48a18	2	451252	222	0x08e4f48a18	3	6f3af7
180	0x08e4f48a18	5	45108e				
134	0x08e4f48a18	2	451121				
123	0x08e4f48a18	5	45107e				
111	0x08e4f48a18	5	45107b				
105	0x08e4f48a18	5	45107c				
84	0x08e4f48a18	2	45107c				
81	0x08e4f48a18	5	451088				
78	0x08e4f48a18	2	45107e				
71	0x08e4f48a18	2	45107b				
70	0x08e4f48a18	5	200557				
67	0x08e4f48a18	5	2005cf				
67	0x08e4f48a18	5	917b5f				
67	0x08e4f48a18	5	917b5e				
66	0x08e4f48a18	5	917b6a				
66	0x08e4f48a18	5	2005db				
62	0x08e4f48a18	5	45107d				
61	0x08e4f48a18	2	917b6a				
61	0x08e4f48a18	2	4513e8				
60	0x08e4f48a18	5	4514e3				
60	0x08e4f48a18	5	45112b				
58	0x08e4f48a18	2	26c46b				
58	0x08e4f48a18	2	451088				
58	0x08e4f48a18	2	200557				
58	0x08e4f48a18	2	917b5e				

Select Slog2 Filepath: ./example.slog2

Generate and Save

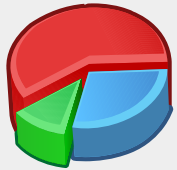
100%

Linux Traces

Available

Filtered

To Slog2



Jumpshot

Legend

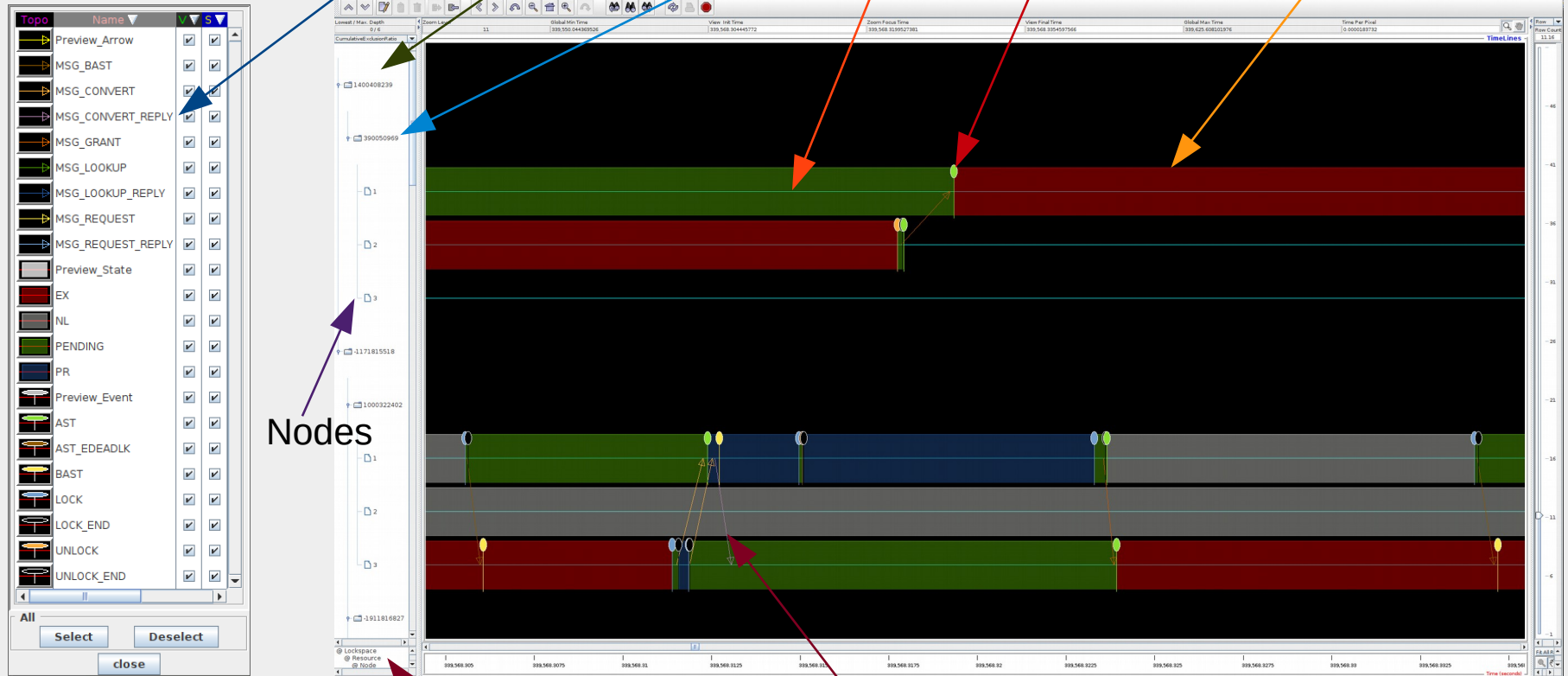
Lockspace

Lock

Timelines

Events

States

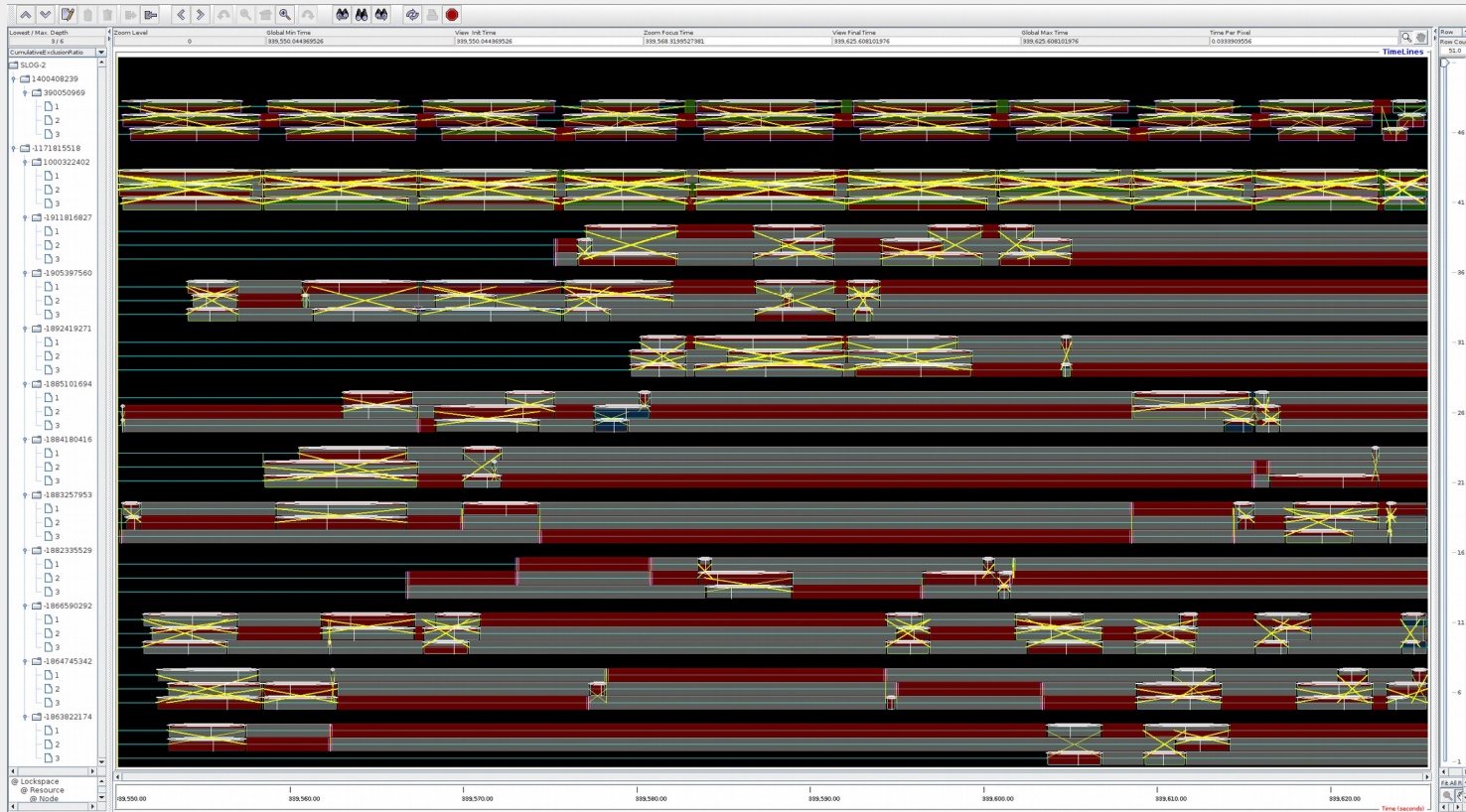


Nodes

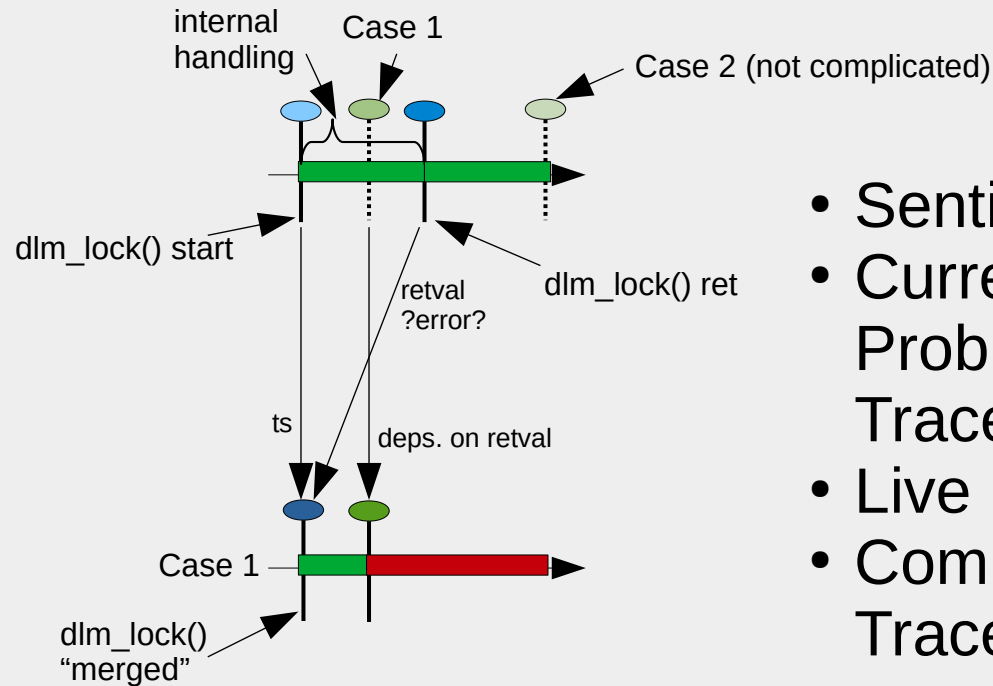
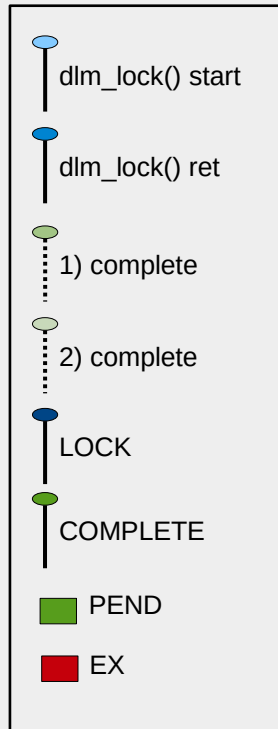
YCoordMap

Arrows

Jumpshot Non-zoomed Preview



Problem: Tracing async Function

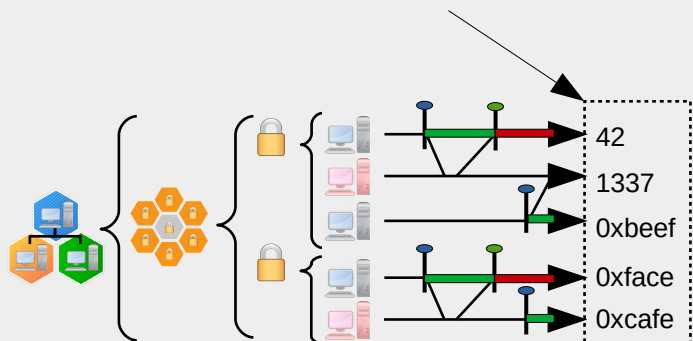


- Sentinals?
- Currently no Problem because Tracefile
- Live Recording?
- Common Field: Traceevent handle?

Idea: Traceevent Classification?

Remember Slide 7?

Let's talk about LineIDs!



- Linux Traffic Control
- User defined fields for each Timeline
- Different than Trace Filter?
- Timeline as separate Buffer?
- Done on which Layer?
- Current solution could end in memory issues

Future Work

- Continuous Kernel Integration
 - DLM Locktorture Test
 - Check for Lock incompatible Modes
- Runtime Kernel Optimization
 - Predigt and Switch Lock Master
 - Networking Queues with skb mark?

End

Thanks

<https://gitlab.com/netcoder/dlm2slog2/-/wikis/home>