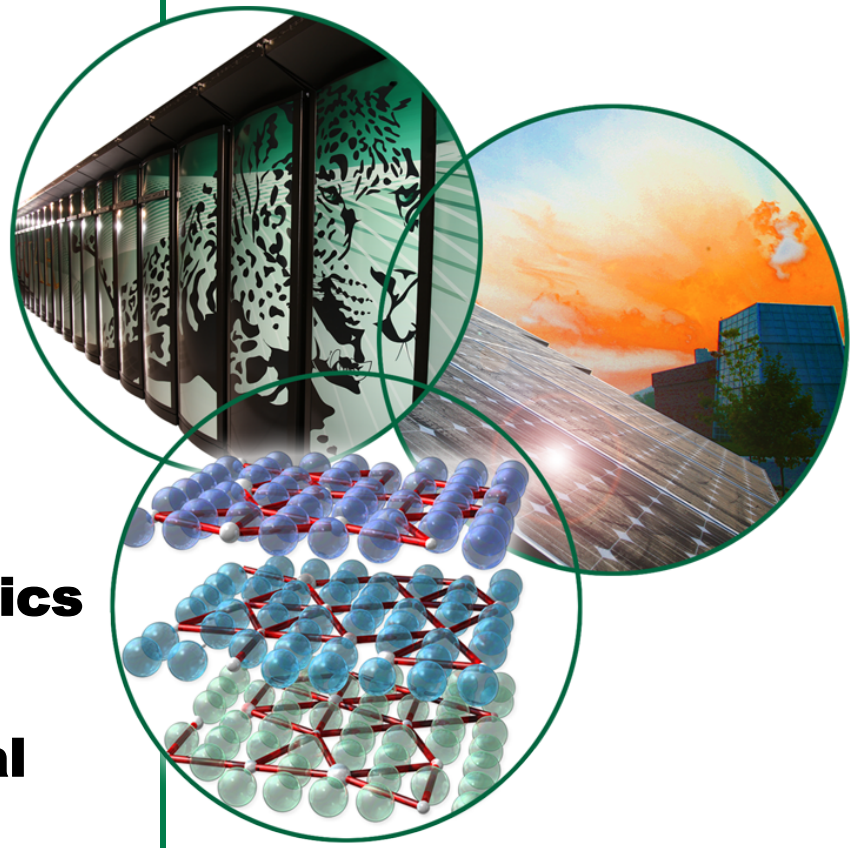
RAVEN: RAS data Analysis through Visually Enhanced Navigation

Byung-Hoon Park, Junseong Heo, Guruprsad Kora, and Al Geist

Computer Science and Mathematics Division, ORNL

National Center for Computational Science (NICS)



U.S. DEPARTMENT OF ENERGY

OAK RIDGE NATIONAL LABORATORY

MANAGED BY UT-BATTELLE FOR THE DEPARTMENT OF ENERGY

Motivation

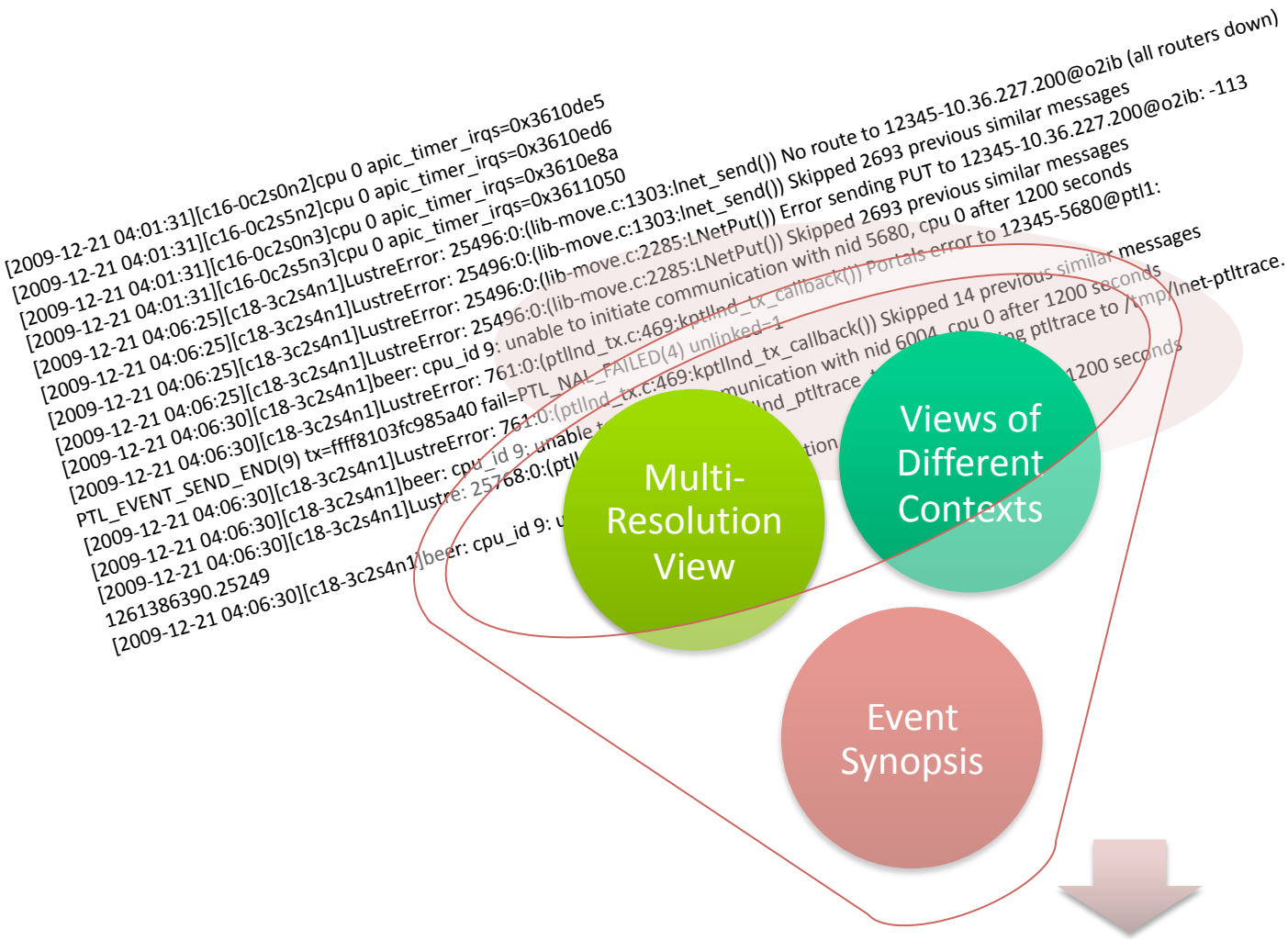
- RAS Logs are often the only resource for obtaining clues for system failure (or system abnormality).
- Ever-increasing volume of Logs is beyond the capacity of manual analysis by humans.
- Console logs are mixture of free outputs from many open source packages.
- Rendering unstructured RAS logs in intuitive and human friendly format is greatly desired.

er_irqs=0x3610de5
er_irqs=0x3610ed6
mer_irqs=0x3610e8a
mer_irqs=0x3611050
: 25496:0:(lib-move.c:1303:lnet_send()) No route to 12345-10.36.227.200@o2ib (all routers down)
: 25496:0:(lib-move.c:1303:lnet_send()) Skipped 2693 previous similar messages
: 25496:0:(lib-move.c:2285:LNetPut()) Error sending PUT to 12345-10.36.227.200@o2ib: -113
: 25496:0:(lib-move.c:2285:LNetPut()) Skipped 2693 previous similar messages
[2009-12-21 04:06:25][c18-3c2s4n1]LustreError: 25496:0:(lib-move.c:2285:LNetPut()) Portals error to 12345-5680@ptl1:
[2009-12-21 04:06:25][c18-3c2s4n1]beer: cpu_id 9: unable to initiate communication with nid 5680, cpu 0 after 1200 seconds
[2009-12-21 04:06:30][c18-3c2s4n1]LustreError: 761:0:(ptllnd_tx.c:469:kptllnd_tx_callback()) Skipped 14 previous similar messages
[2009-12-21 04:06:30][c18-3c2s4n1]LustreError: 761:0:(ptllnd_tx.c:469:kptllnd_tx_callback()) PTL_EVENT_SEND_END(9) tx=ffff8103fc985a40 fail=PTL_NAL_FAILED(4) unlinked=1
[2009-12-21 04:06:30][c18-3c2s4n1]beer: cpu_id 9: unable to initiate communication with nid 6004, cpu 0 after 1200 seconds
[2009-12-21 04:06:30][c18-3c2s4n1]Lustre: 25768:0:(ptllnd_ptltrace_to_file()) dumping ptltrace to /tmp/lnet-ptltrace.
1261386390.25249
[2009-12-21 04:06:30][c18-3c2s4n1]beer: cpu_id 9: unable to initiate communication with nid 6371, cpu 0 after 1200 seconds

Properties of RAS Logs

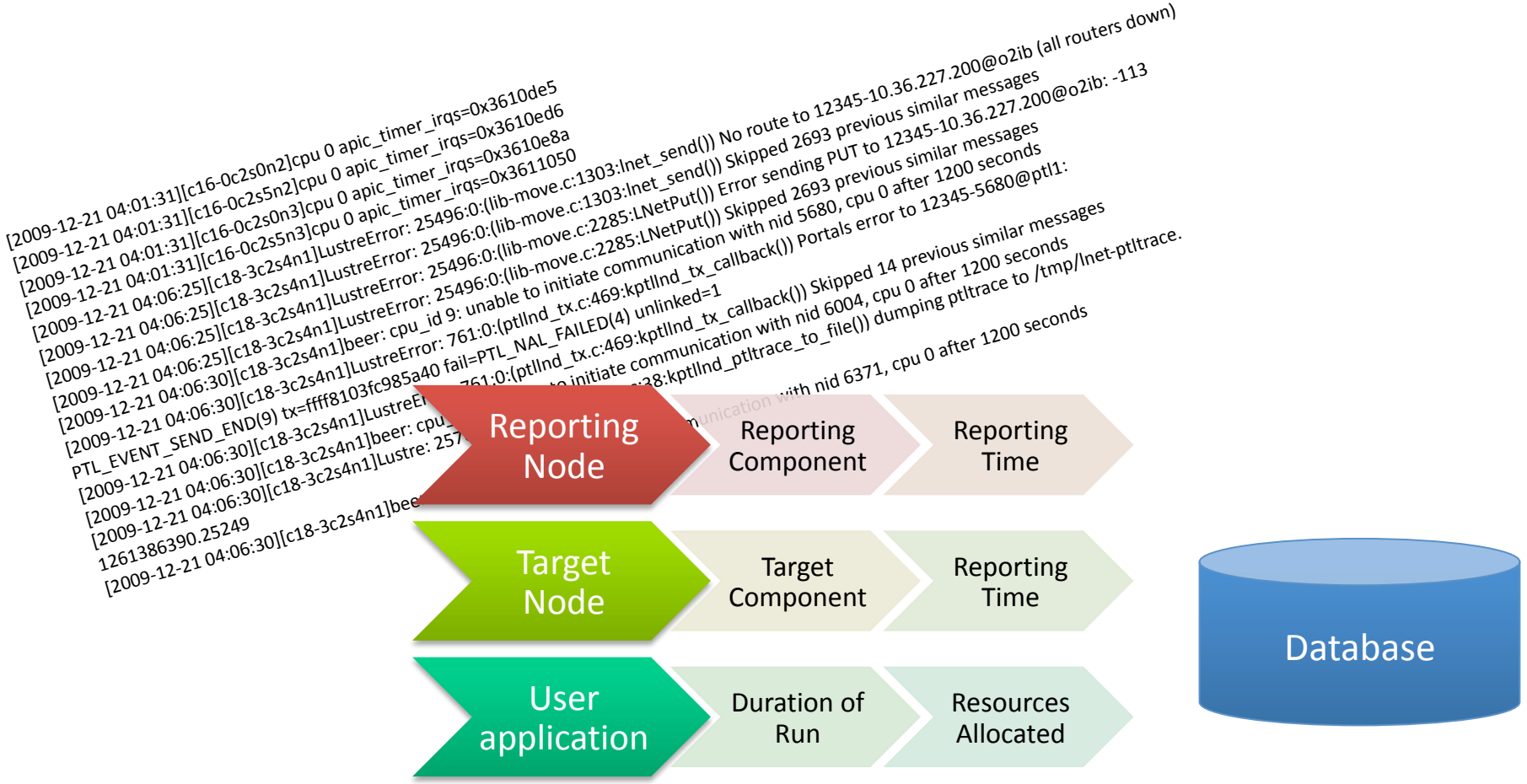
Irregular	• Event occurrences are irregular, often come in bulk.
Redundancy	• The same root cause triggers different observations of many components.
Implicit Correlation	• Some event types are inherently correlated. • Can be used to confirm the situation.
Implicit Context	• Many Messages embed pair wise relations. • Messages are clustered based on the user application.

Better Use of RAS Data

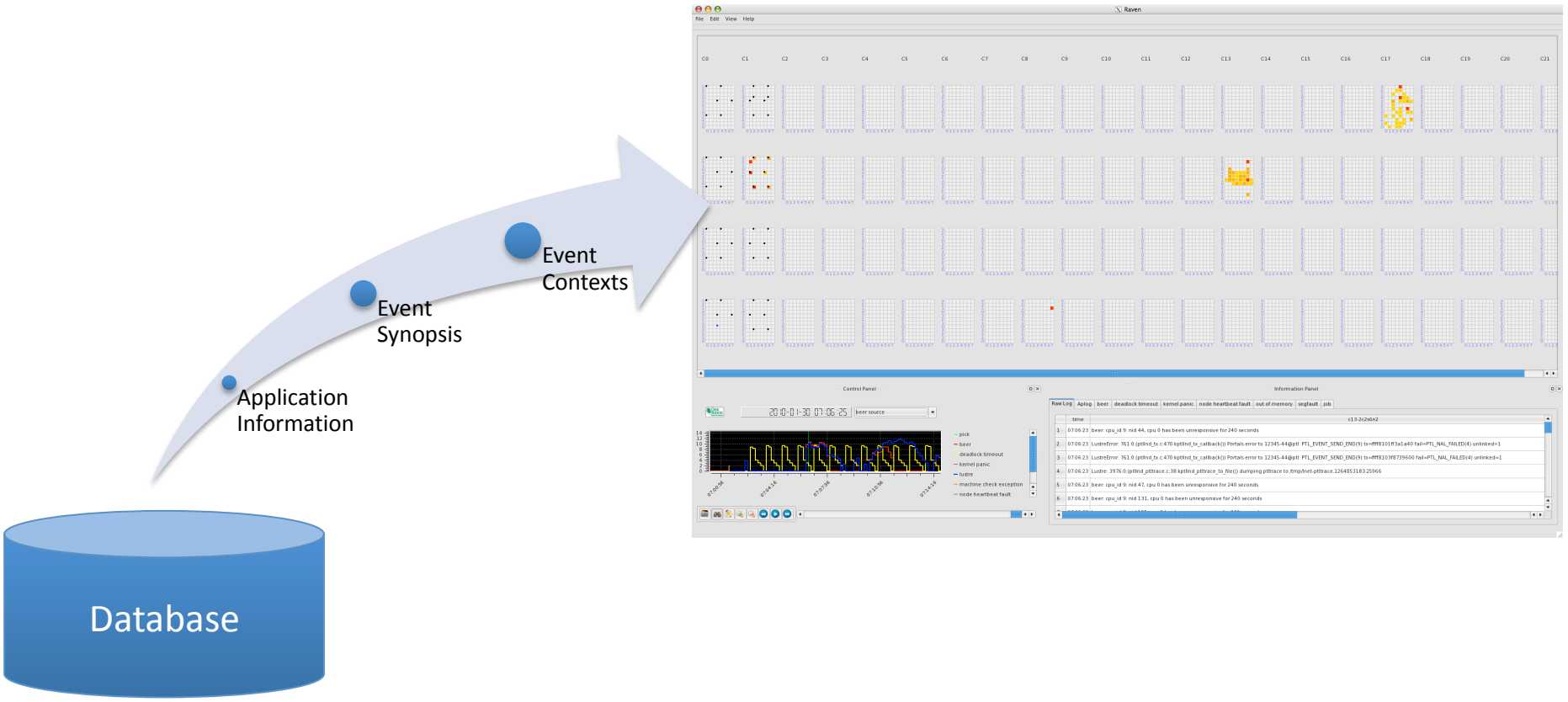


Context Driven Temporal and Spatial Analysis

Abstraction of RAS Data



RAVEN:
RAS data Analysis through Visually Enhanced Navigation



RAVEN At a Glance

Raven is a system independent tool

- Database Driven (MySQL)
- Configurable through XML files.

RAVEN is GUI tool.

- Frontend Interface is written with Qt4.6
- Graph Plotting is written with Qwt5.2

RAVEN has been ported to

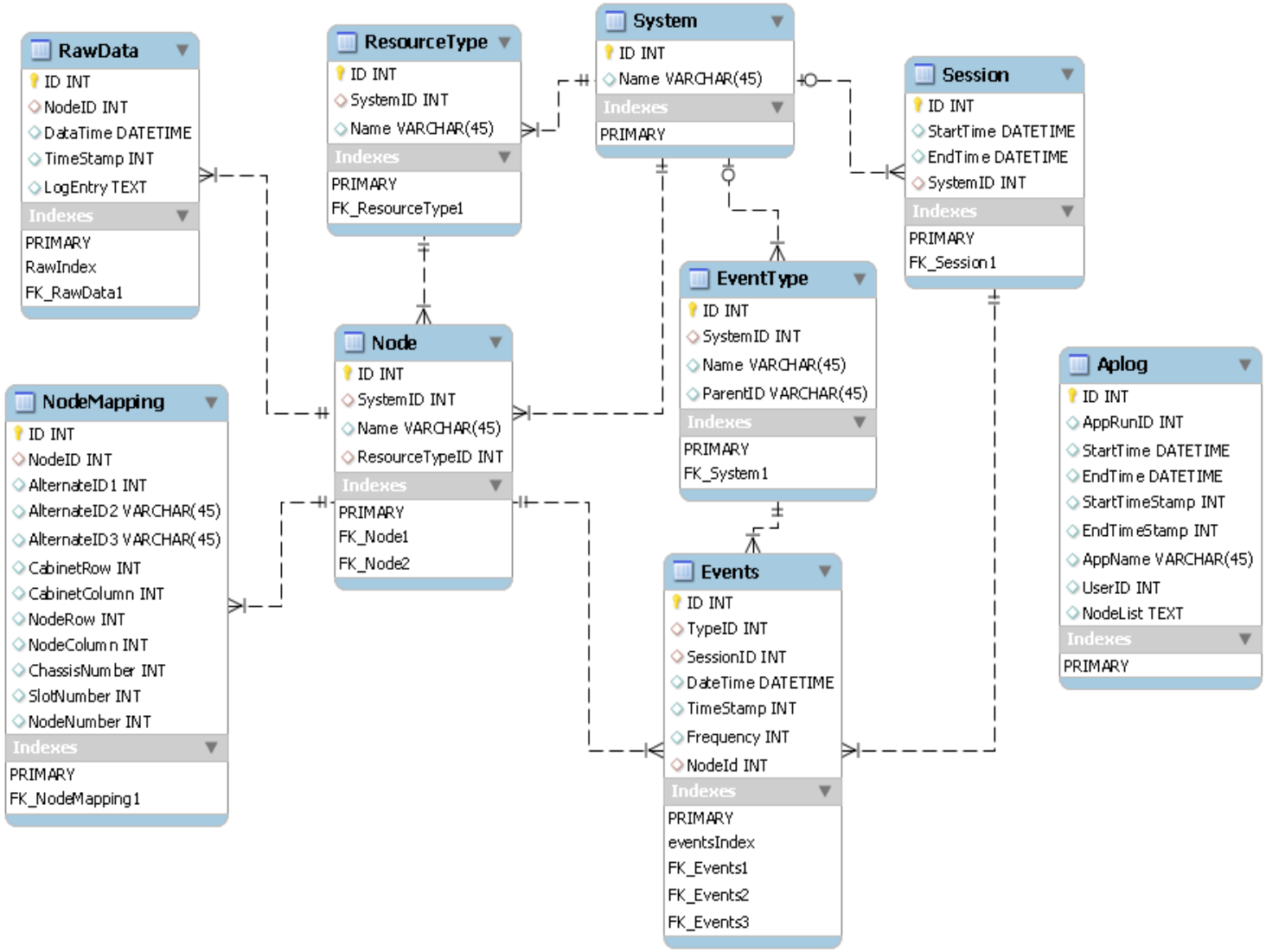
- Cray XT5 (Jaguar, Kraken)
- IBM BG/P at ANL (Underway)

RAVEN RAS Events for Cray XT5

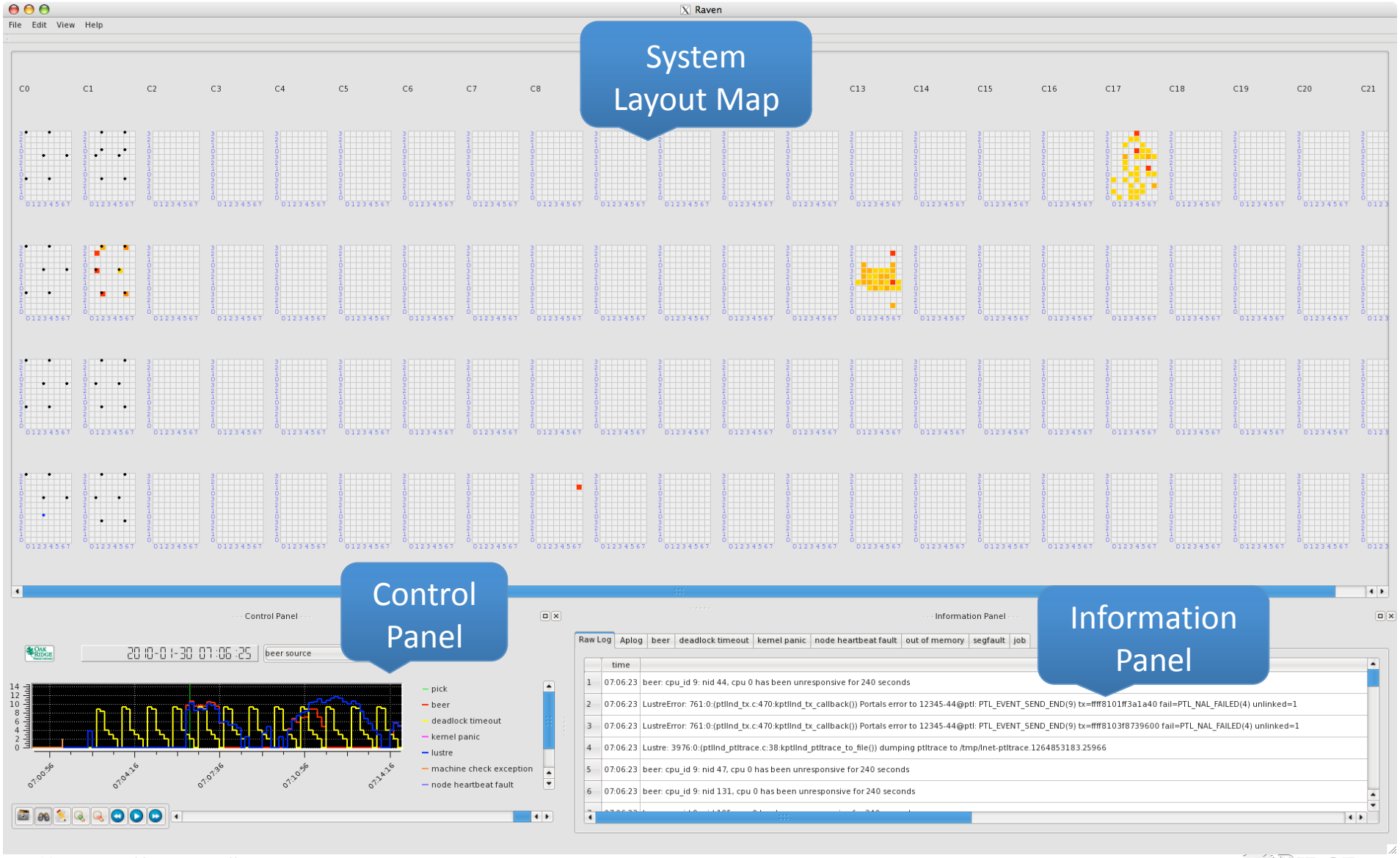
Console	• Lustre, BEER, Segfault, OOM, MCE, Kernel Panic, etc.
Netwatch	• Link Inactive, Deadlock Timeout, uPacket Squash (off), etc.
Consumer	• Node Heartbeat Failure, Seastar heartbeat failure, Node voltage fault, Verty health check fault, etc.
Apsched	• Application Name, User ID, Duration of Run, Allocated Nodes, ResID, etc.

RAVEN Backend Database Design

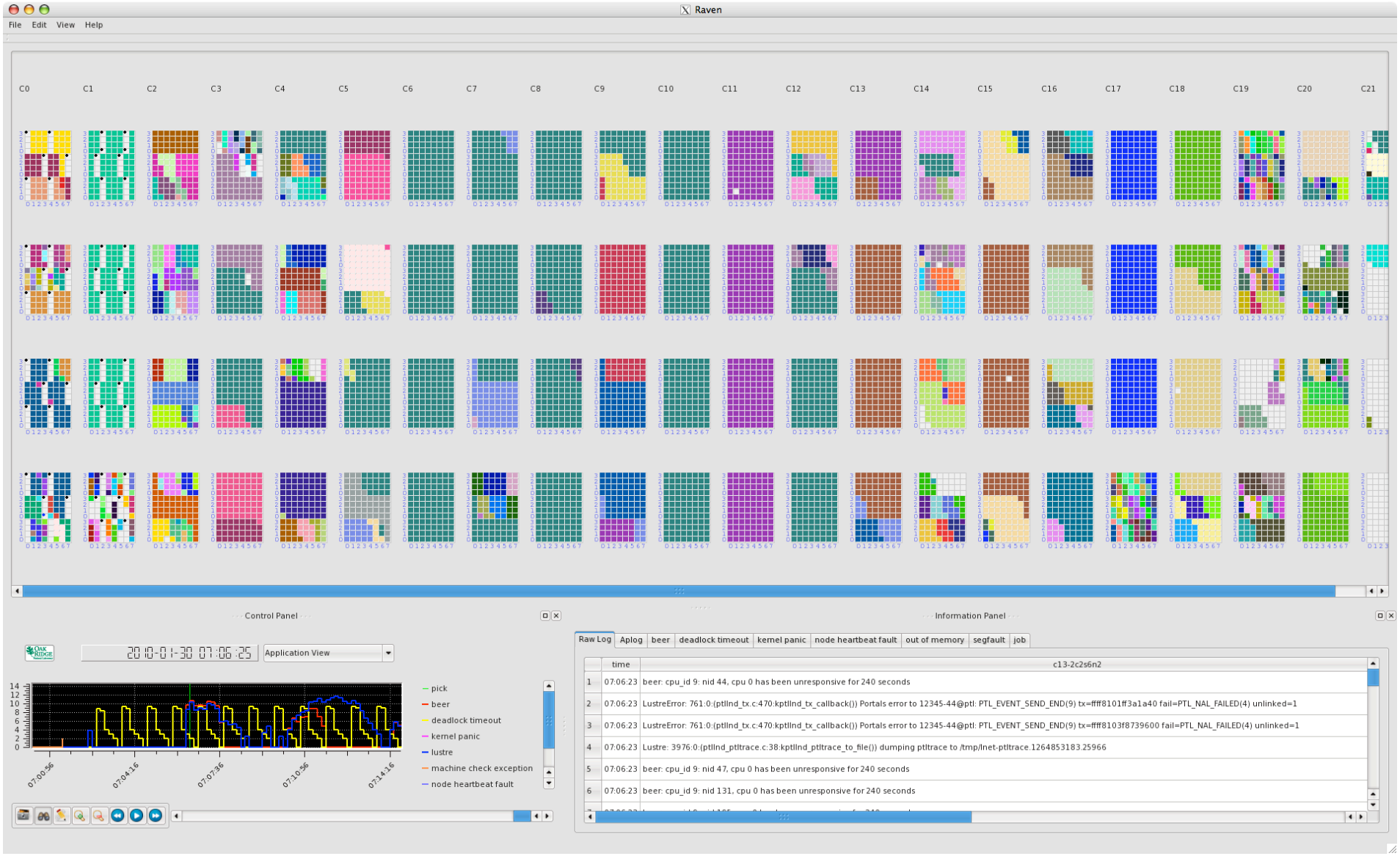
RAVEN RAS Log Backend Design



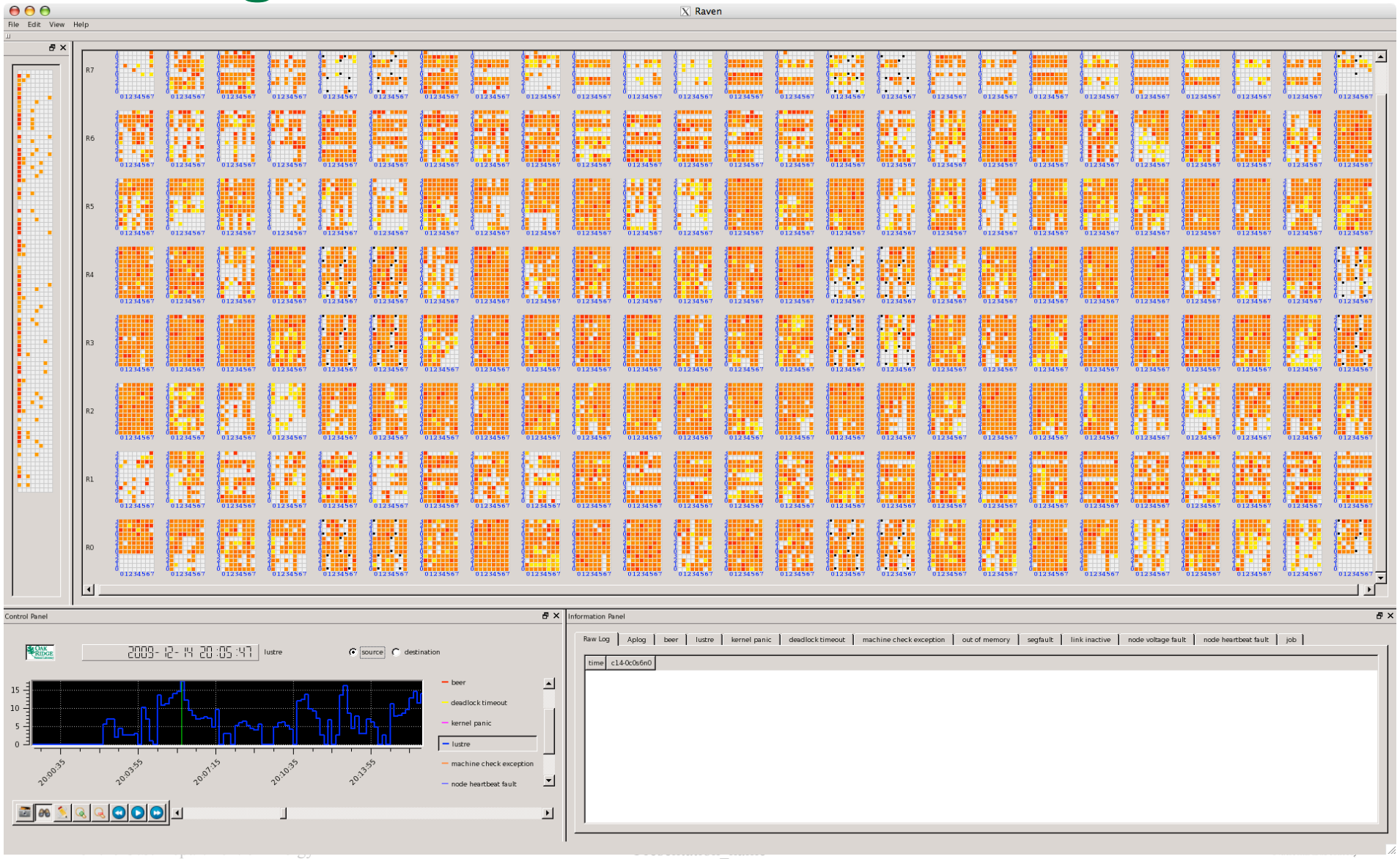
RAVEN Frontend



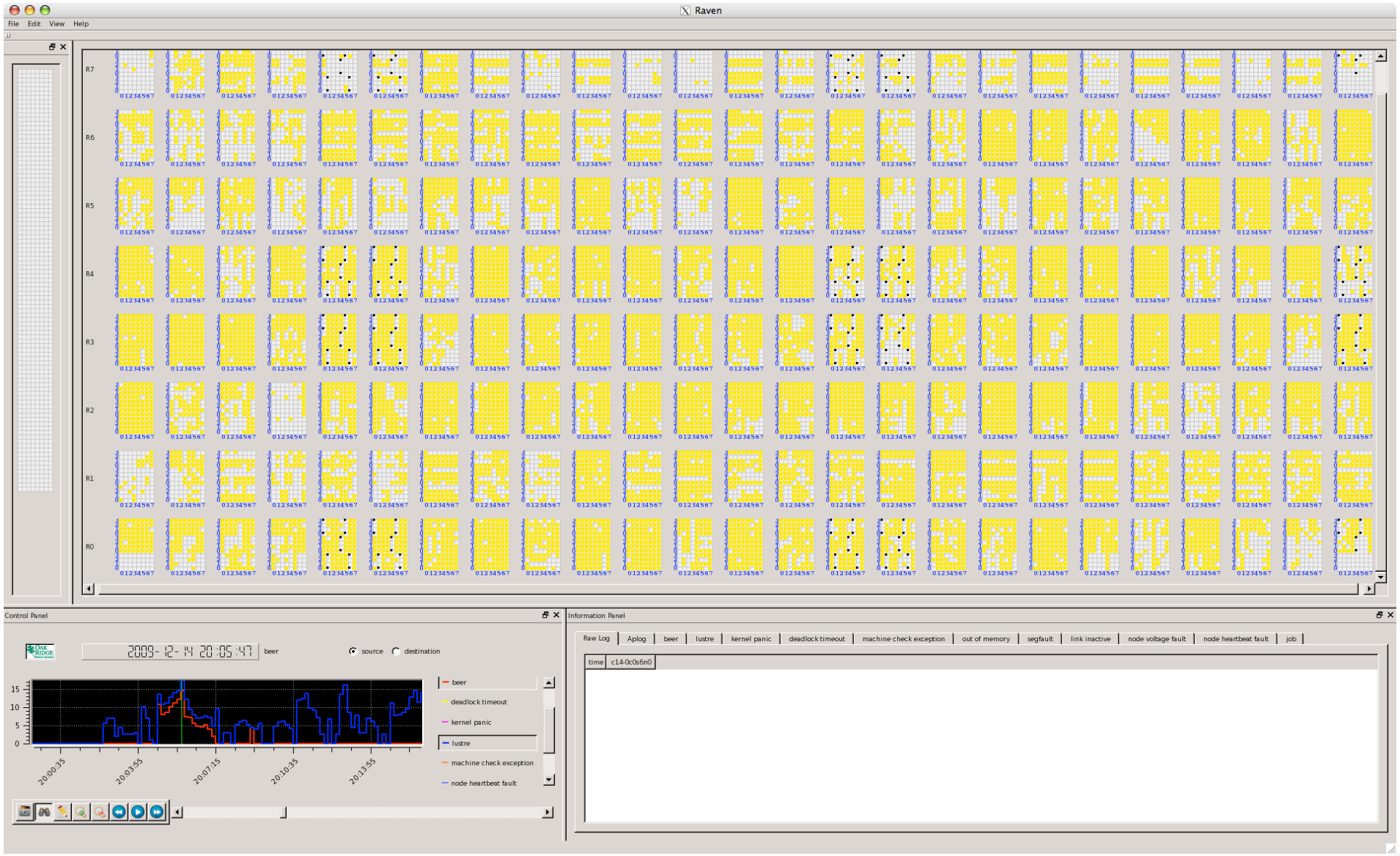
Application Displacement



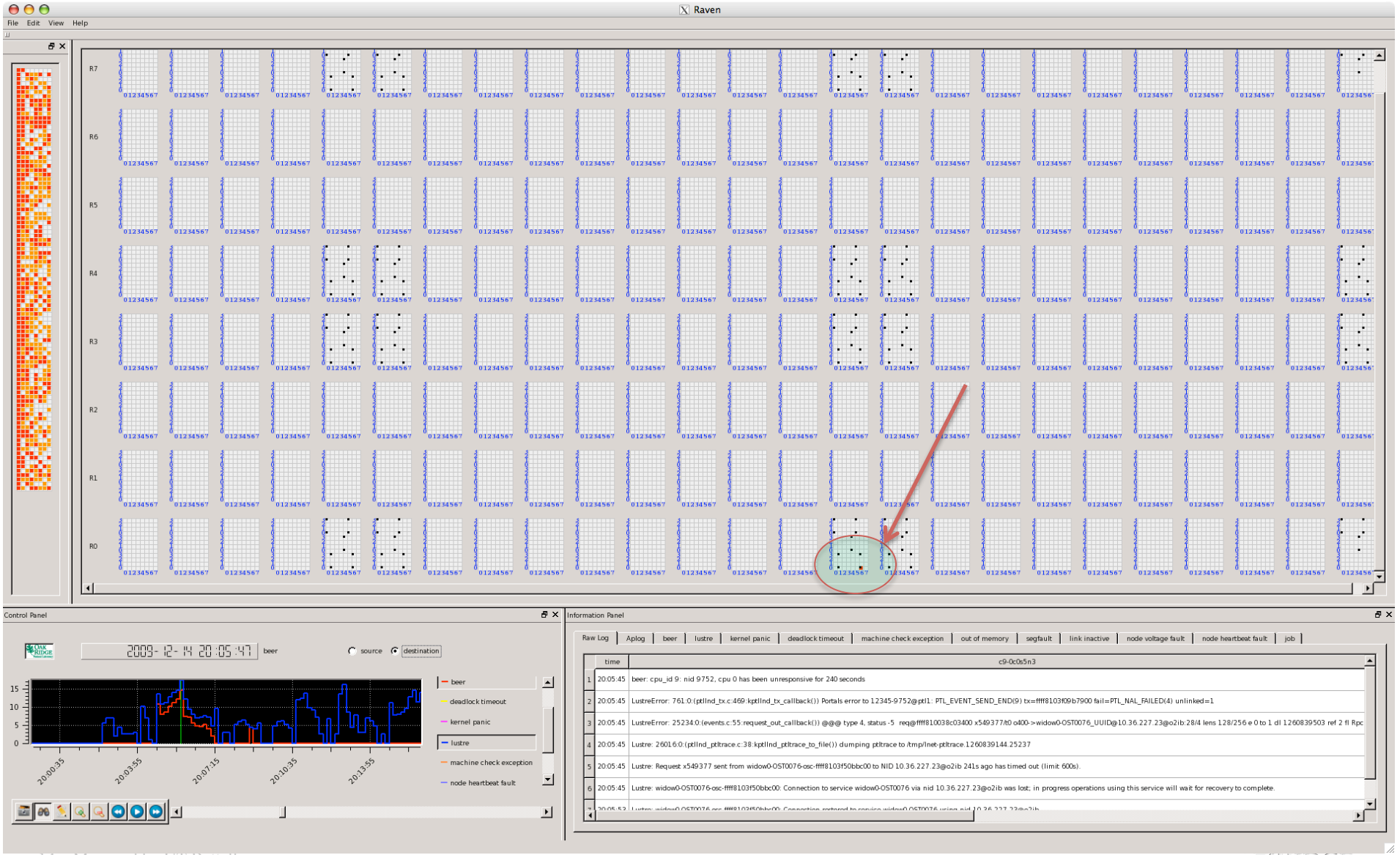
Case Study: A Flood of Lustre Messages



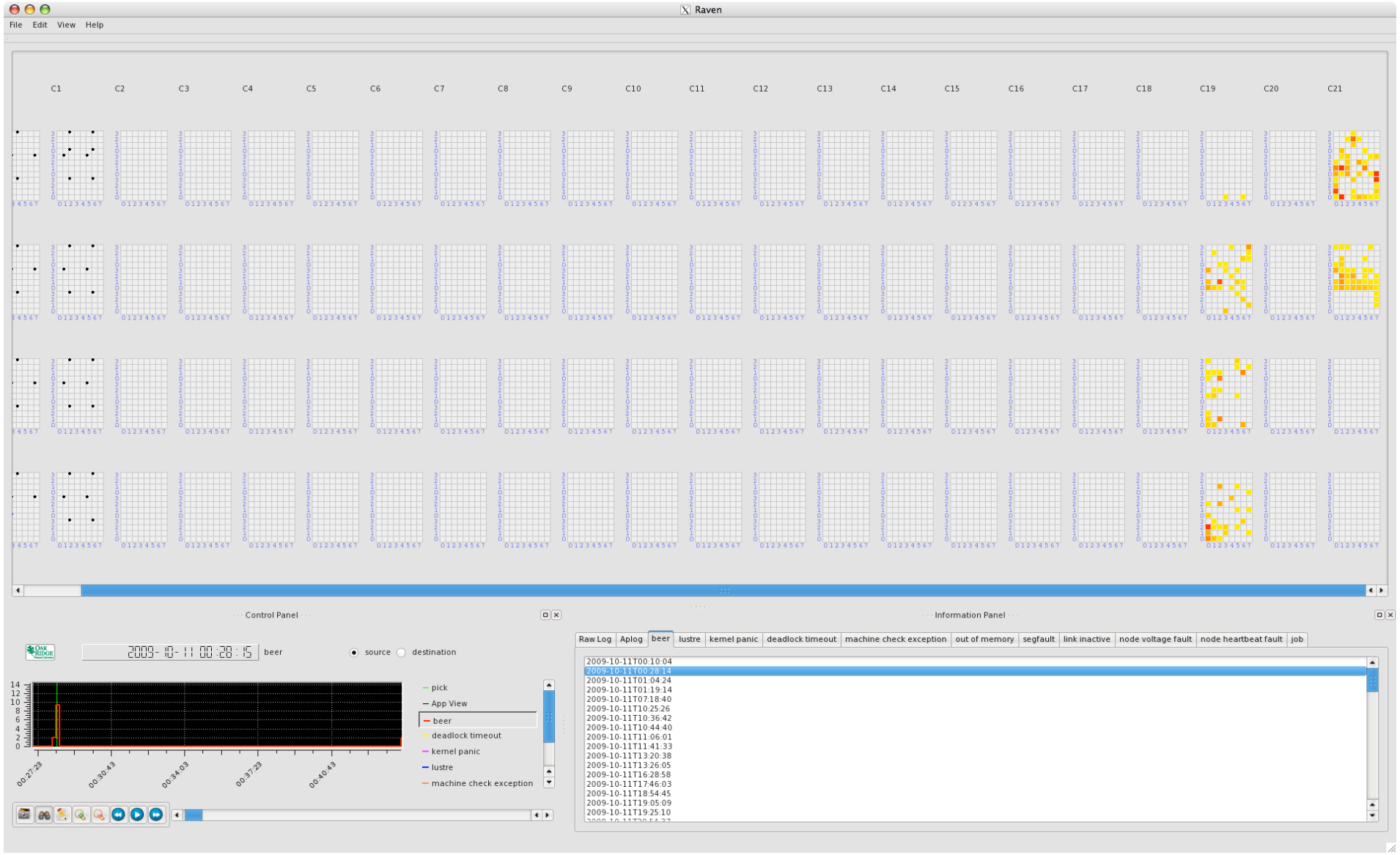
And a Flood of BEER Messages Together



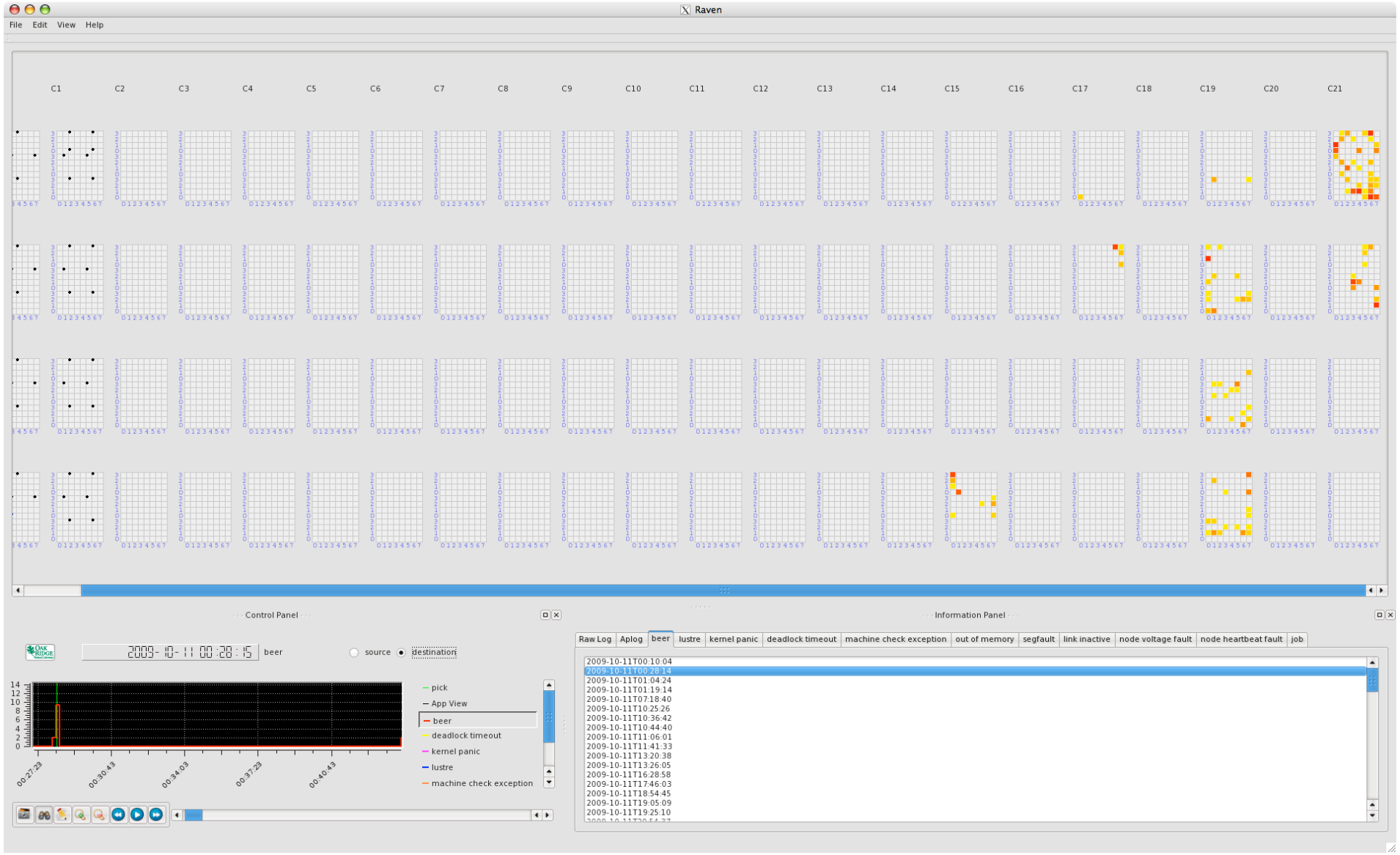
Pointing to a Single Router Node



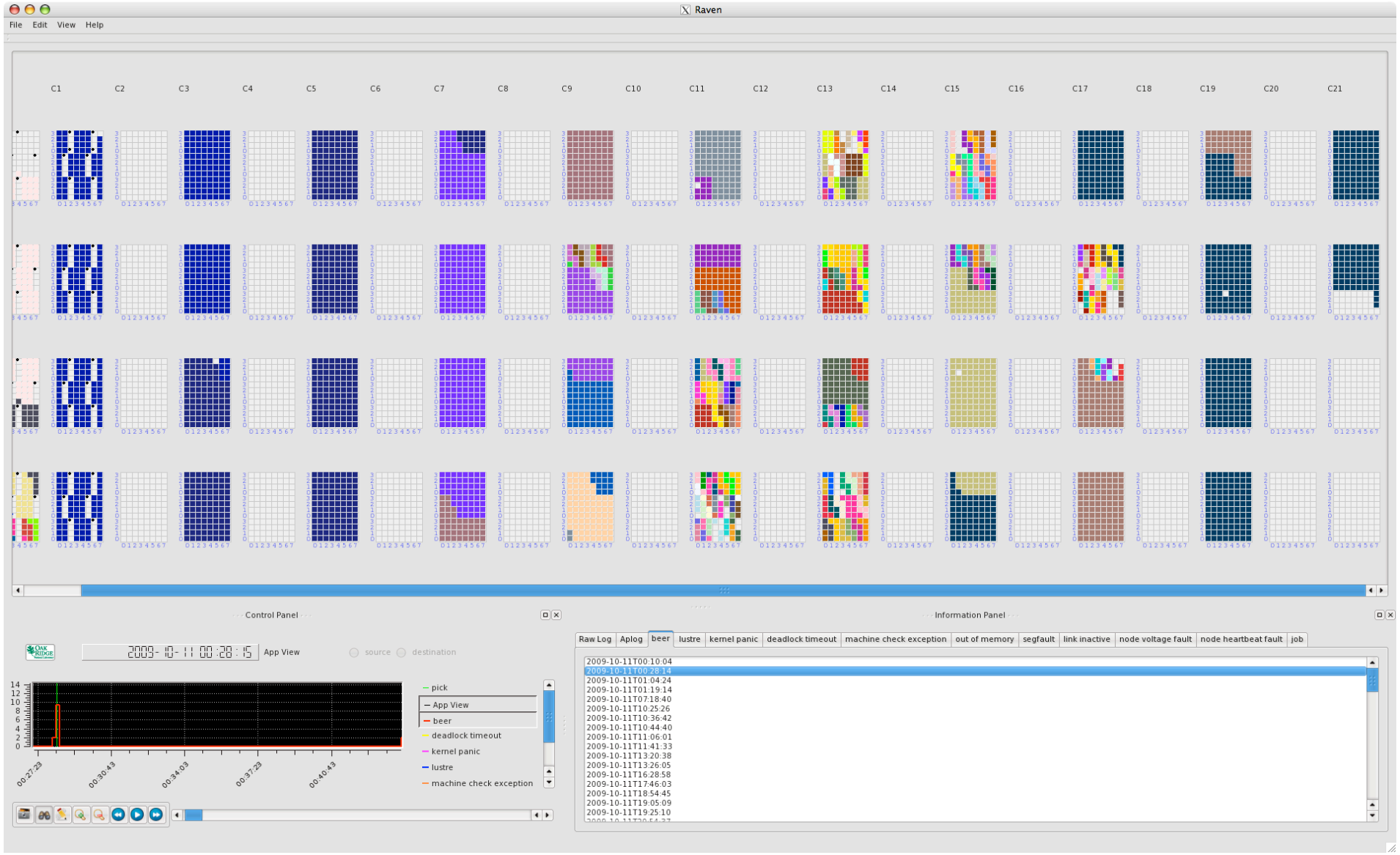
Case Study: BEER Messages Between Nodes of the Same Application



BEER Messages Between Nodes of the Same Application



BEER Messages Between Nodes of the Same Application



Next Step

- Provide live feed of RAS Events into RAVEN
- Incorporate a richer set of RAS event types
- Embed a intelligent module that sifts abnormal signatures from live feed of logs
  - Cluster of nodes showing abnormal behavior.
  - Time interval showing abnormal behavior.
  - Job showing abnormal behavior.
  - Etc.
- A Web-based RAVEN is under development.

Acknowledgement

- Raghul Gunasekaran (ORNL)
- David Dillow (ORNL)
- Galen Shipman (ORNL)
- Don Maxwell (NCCS, ORNL)
- Jeff Beckleheimer (Cray Inc.)
- Rick Mohr (NICS, UTK)