

Online Prompt Reconstruction

Sridhara Dasu
University of Wisconsin

- PromptReco versus Reconstruction
 - Differences
 - Responsibilities
- PromptReco status
 - Reconstruction code
 - Infrastructure
 - Performance

Prompt Reconstruction

- Processes raw data (XTC) files saved from online data acquisition
- Runs TC-to-DIGI conversion, full reconstruction and monitoring/calibration based on the reconstructed data.
 - Integrates online and offline code
 - ElfApp = OepSequences + BearSequences + MonitorSequences + CalibrationSequences
- Stores raw + reconstructed + data-stream calibration data in Objectivity database

PR Software

- **OprManager**
 - Initiates PR processing
- **OprLoggingManager**
 - Receives events from Oep, saves them in intermediate TC file store and distributes to Opr processes
 - Ensures that all events are logged and distributed for processing.
- **OprDaemon**
 - One per processor node handles communication with LM and feeds the ElfApp
- **ElfApp**
 - Does the actual processing (a framework job)

Responsibilities

- Core Opr* code
 - Opr group
- ElfApp
 - Custom framework: Opr group
 - Subsystems for actual code: TC-to-DIGI, Reconstruction, Monitoring and Calibration
 - Oep group for integrating Oep sequences
 - Opr group for integrating Opr sequences
 - Reco group for reco sequences
 - Code and required TCL settings are needed.

PR Software Status

- Core code
 - Base functionality written and debugged
 - Manager programs are currently scripts
 - Can use more automation - Shifters shouldn't see PR.
 - Bells and whistles underway
 - Error recovery
 - Big worry
 - Need to fully integrate into error reporting scheme
 - Can run alternate, e.g., TC-to-DIGI only code on “killer” events.
- PR developers are stuck debugging non PR problems, leaving these holes

PR Software Status II

- **ElfApp**

- All sequences are now linked in
- Can run on cosmic/simulated/real events
- It has been a rough ride through a myriad releases and tags but we are probably close
 - Only 4 subsystem package tags in 8.1.0
- Thank you subsystem developers!
- Some tcl parameter settings need cleanup
- Some sequences don't actually do anything
- Some crashes are not quite resolved yet
- Subsystem experts can and should now run ElfApp to debug any problems
- Instructions for running underway.

Database Software

- **Unexpected Problems**

- Work around the lock and AMS problems
- Much work to avoid dead locks
- Improvements to Objectivity software to scale to large number of distributed processes accessing several of database files.
- It was important to pin point the exact problems in their software and ideas for their resolution to obtain “timely” fixes.
- Improved organization of the database server software and hardware.
- Thank you Adil, Andy, David & Jacek
- Thank you SCS

PR Test Results

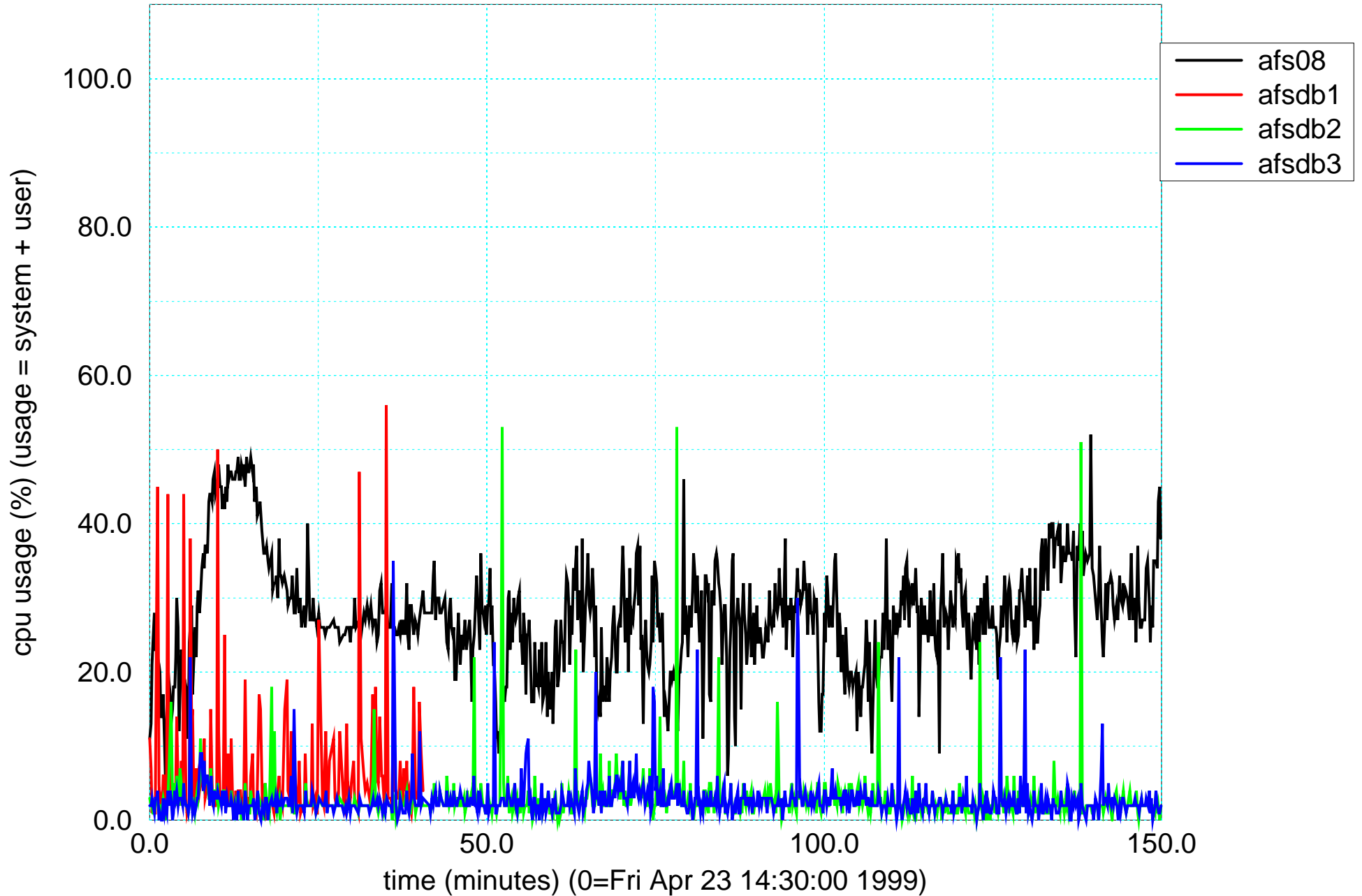
- Continuous testing of software, database and farm hardware since mid December
- The biggest run used 121 CPUs
 - 21.5 Hz peak rate
 - 13 Hz steady state
 - 5.5 Hz overall
 - Big start up time
 - Dominated by AFS access to the 200MB ElfApp from 121 machines!
 - Left a big site wide mess for SCS to cleanup
- Also ran a 104 CPU test
- Cannot repeat this often - until a solution is found

PR Scaling Test

# CPUs	Rate (Hz)	Startup latency(min)	# file descs	ools idle time(%)
1	0.18	3	108	87
2	0.35	4	130	76
4	0.69	4	162	65
8	1.38	7	235	45
14	2.38	9	354	47
31	4.82	21	633	41
33	~5.1	25	762	31
65	11.9	18	1040	
104	17.6	41		
121	21.5	55		

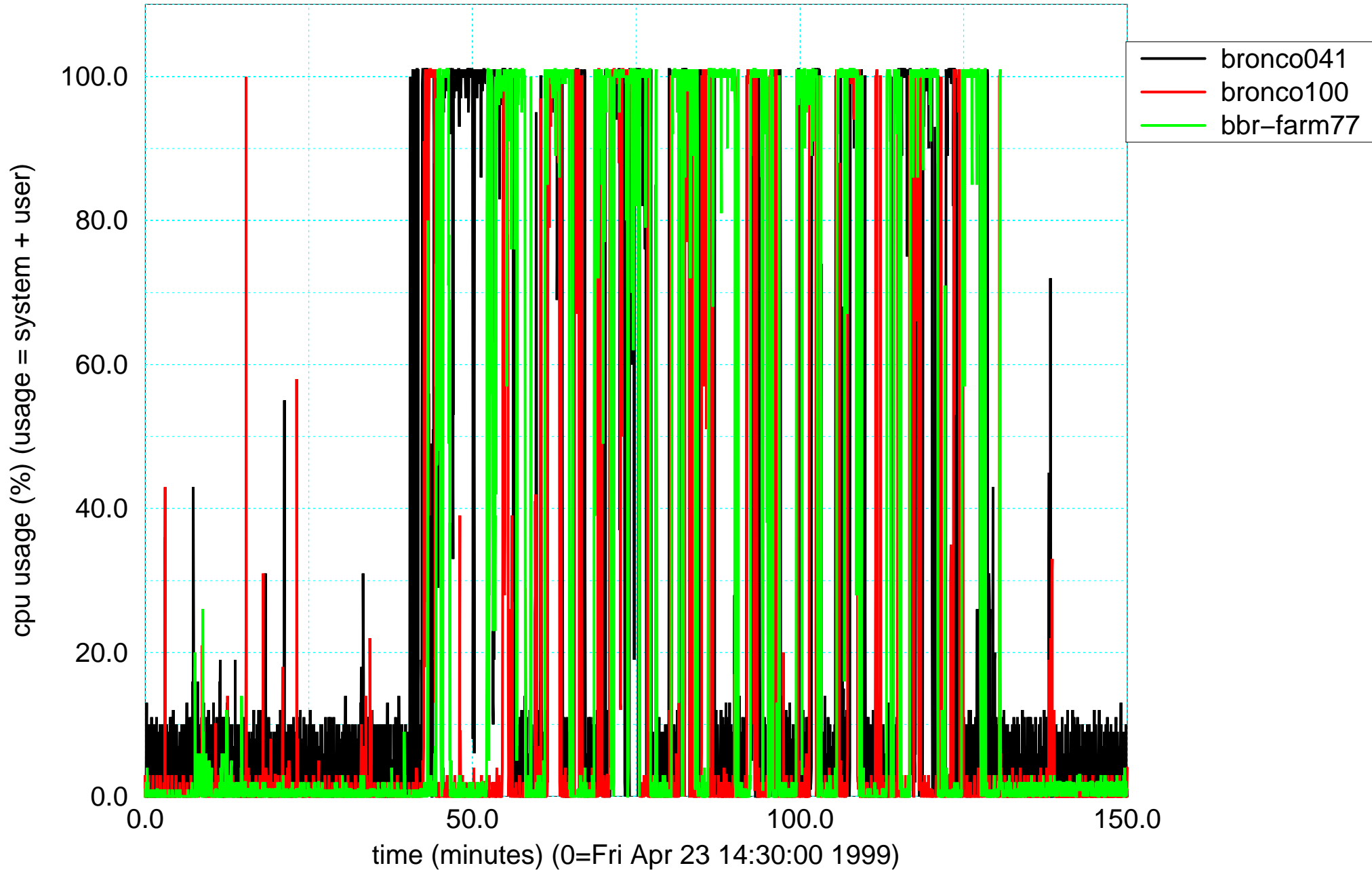
Run 1003: Interval 1, 120 nodes

From Fri Apr 23 14:30:00 1999 to Fri Apr 23 17:00:00 1999



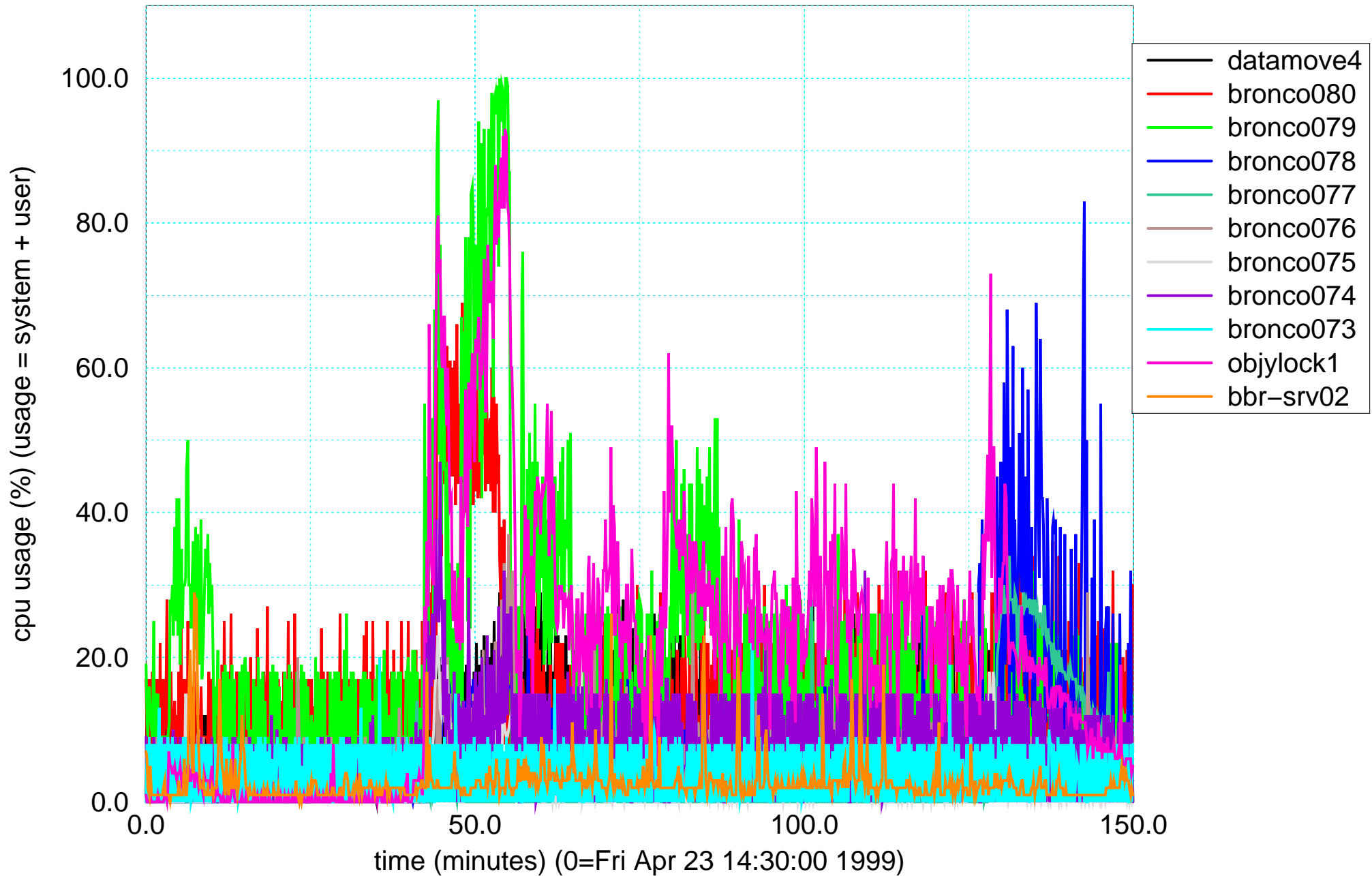
Run 1003: Interval 1, 120 nodes

From Fri Apr 23 14:30:00 1999 to Fri Apr 23 17:00:00 1999



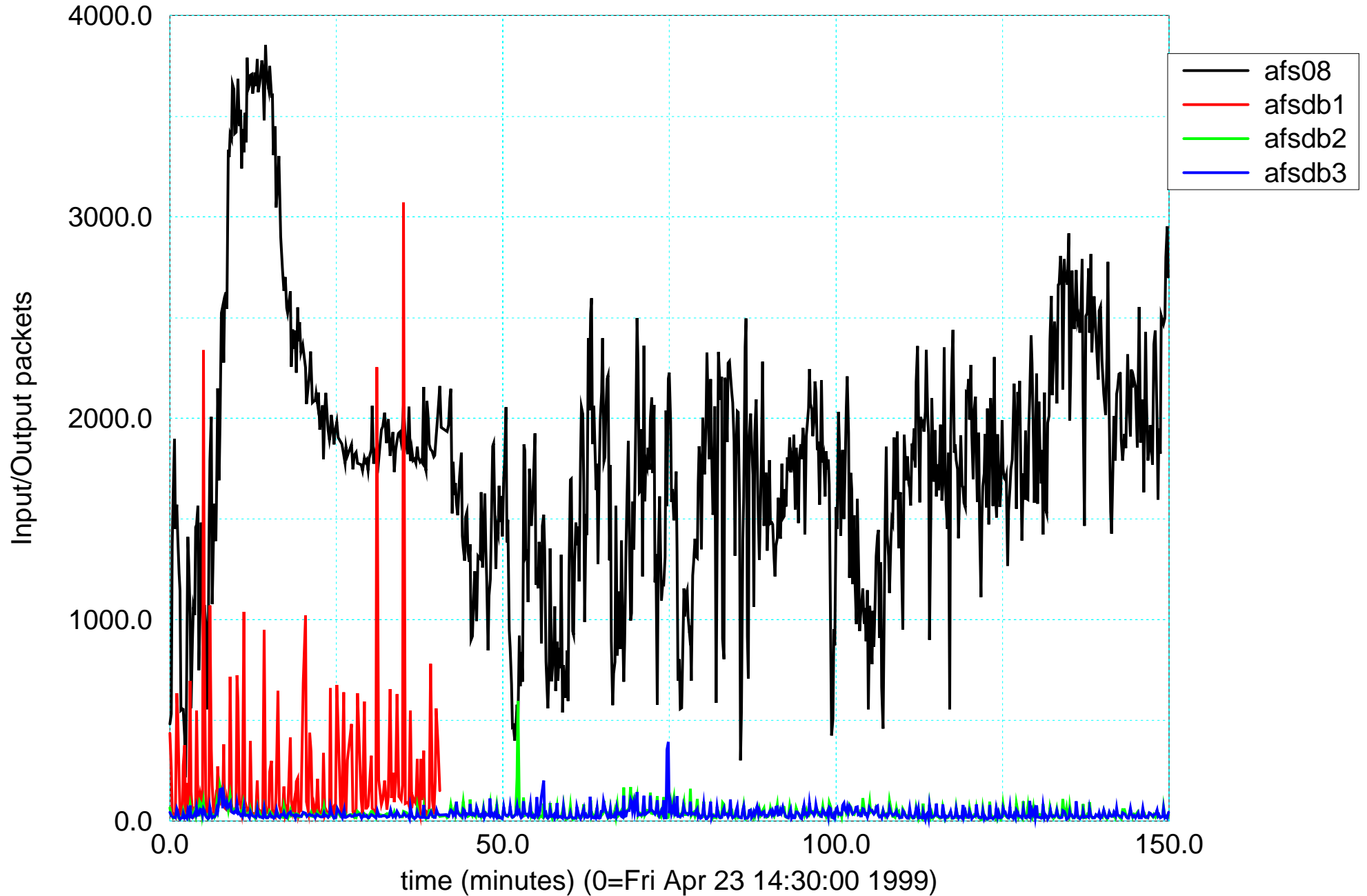
Run 1003: Interval 1, 120 nodes

From Fri Apr 23 14:30:00 1999 to Fri Apr 23 17:00:00 1999



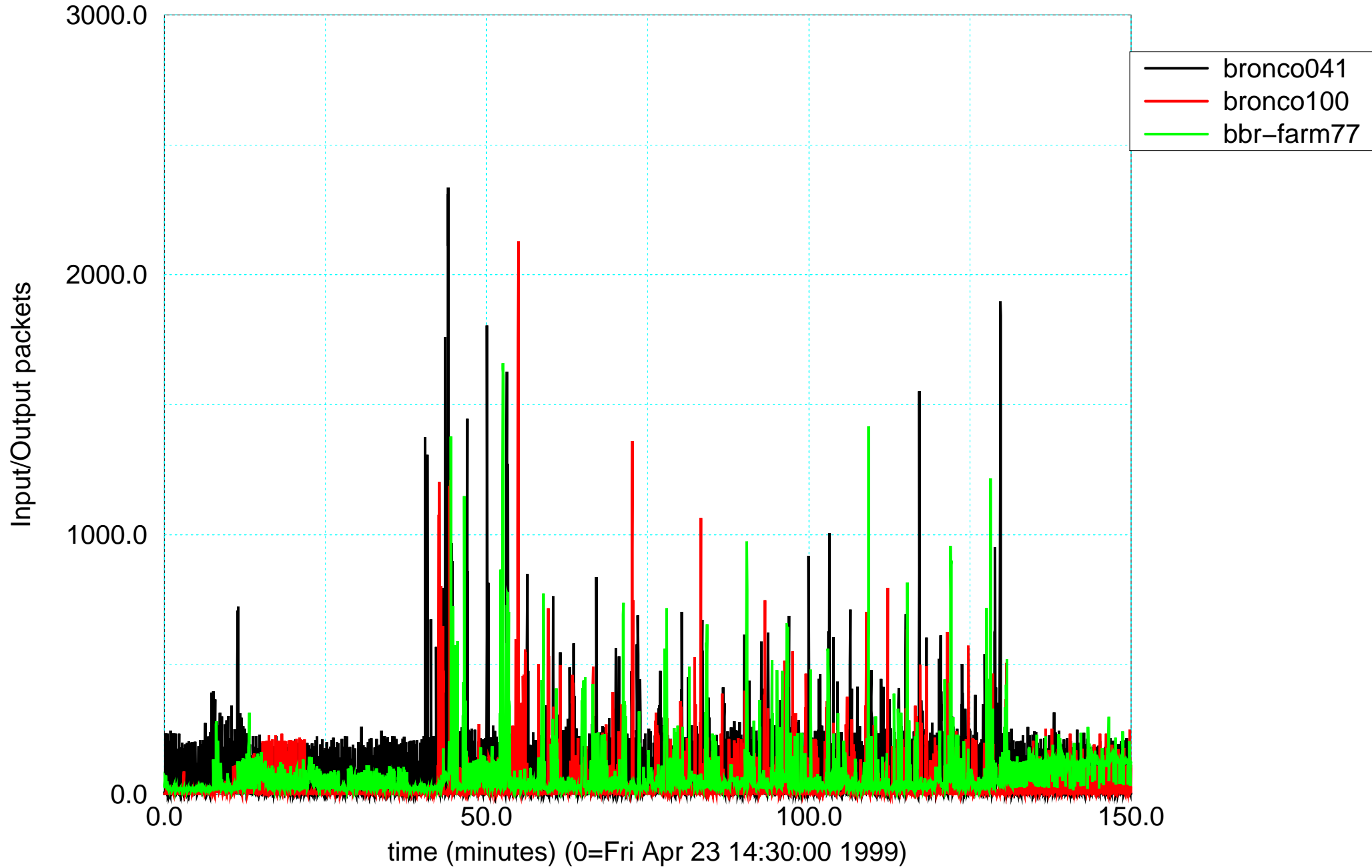
Run 1003: Interval 1, 120 nodes

From Fri Apr 23 14:30:00 1999 to Fri Apr 23 17:00:00 1999



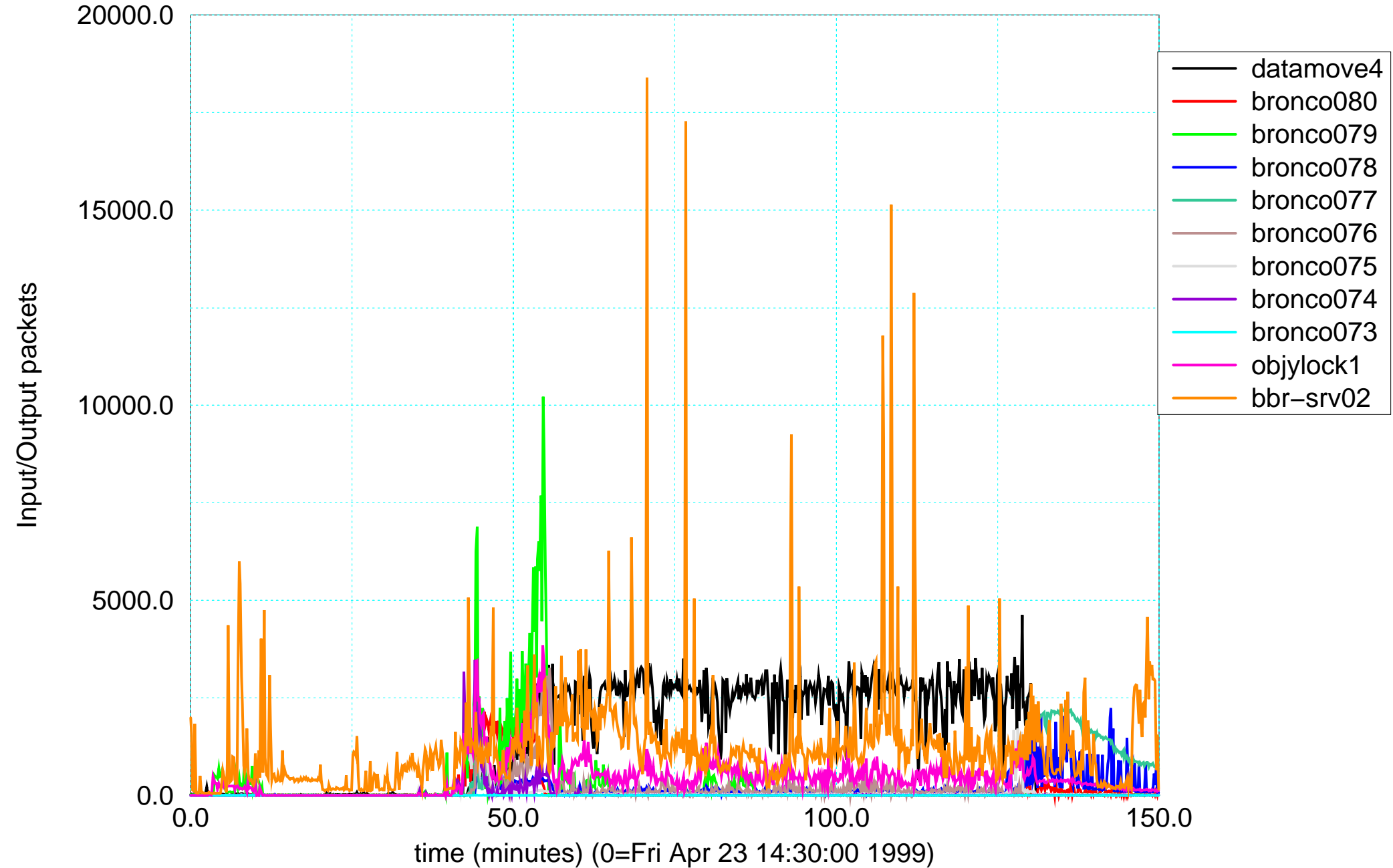
Run 1003: Interval 1, 120 nodes

From Fri Apr 23 14:30:00 1999 to Fri Apr 23 17:00:00 1999



Run 1003: Interval 1, 120 nodes

From Fri Apr 23 14:30:00 1999 to Fri Apr 23 17:00:00 1999



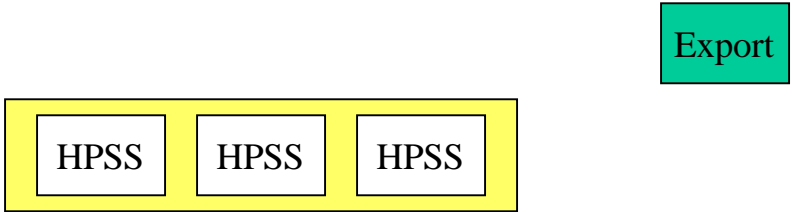
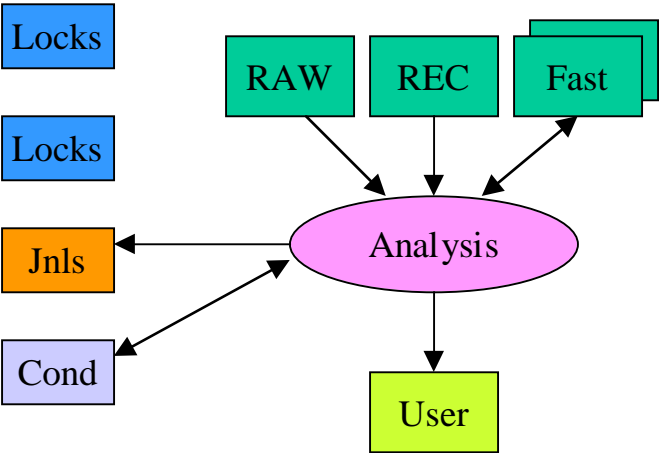
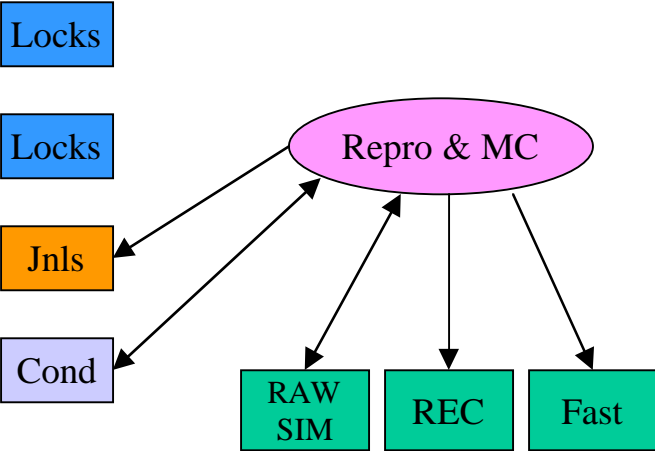
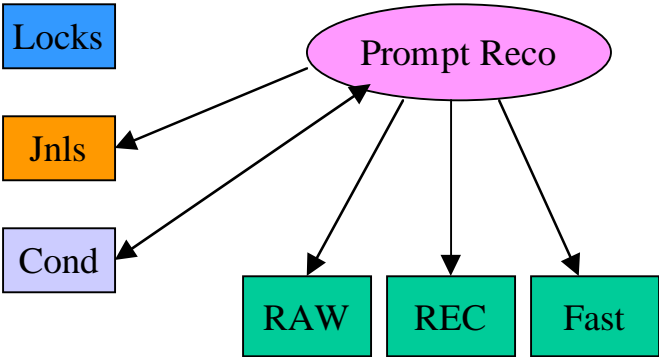
PR Operations / Farm

- PromptReco is unable to share hardware with other activities due to Database/Network reasons
- Farm scales to 600 nodes
 - 200 for PromptReco
 - 400 for Batch Analysis/Repro/Simulation
- 200 simultaneous Analysis Jobs
 - 100 interactive, 100 batch
 - Split between simulation and physics analysis

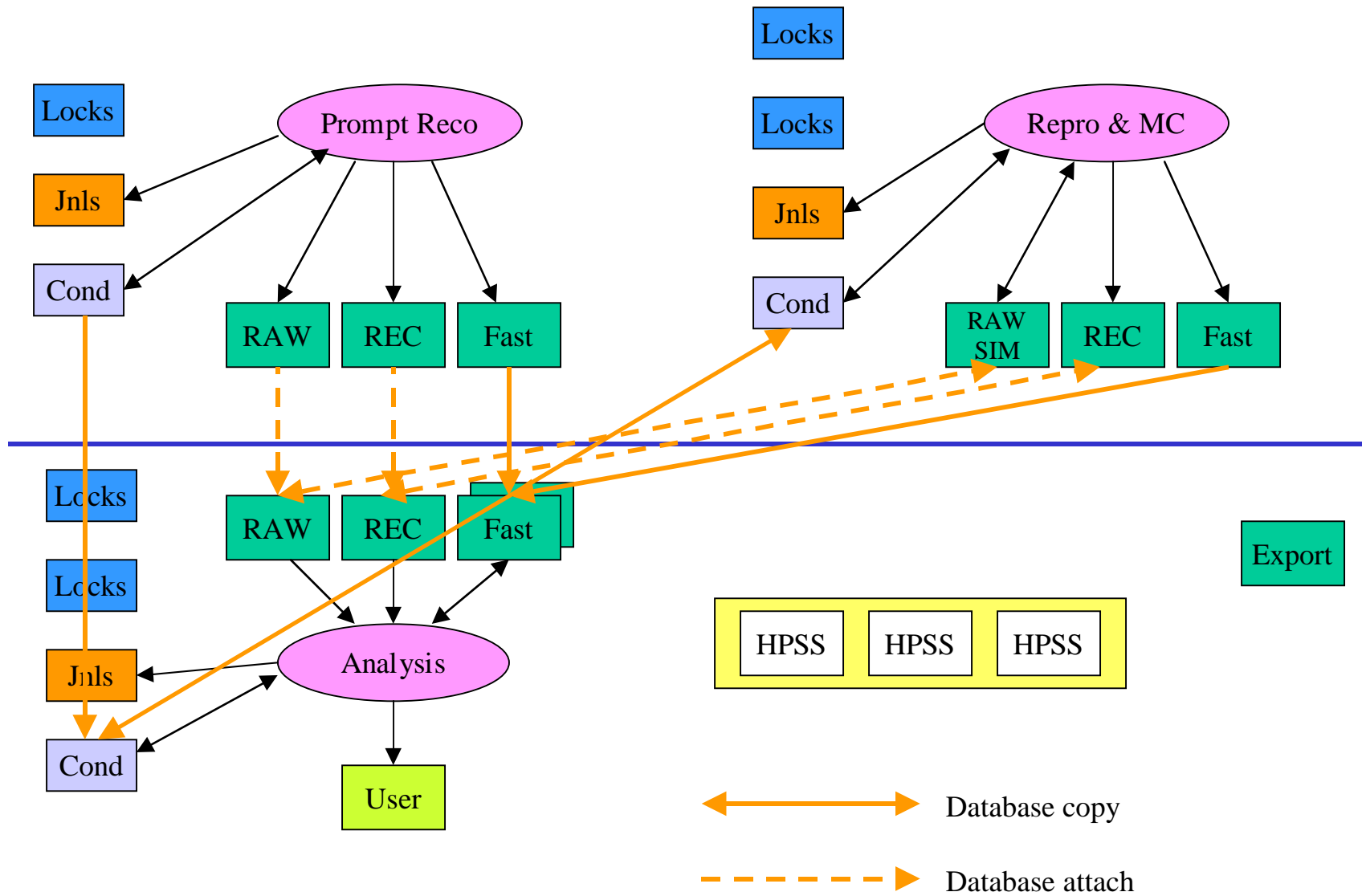
PR Operations / Databases

- Online/Prompt Reconstruction
- Analysis
 - Interactive & batch
 - Physics & Simulation
- Reprocessing
- Simulation Production
- Small scale activities
 - Test federations, QA, etc.

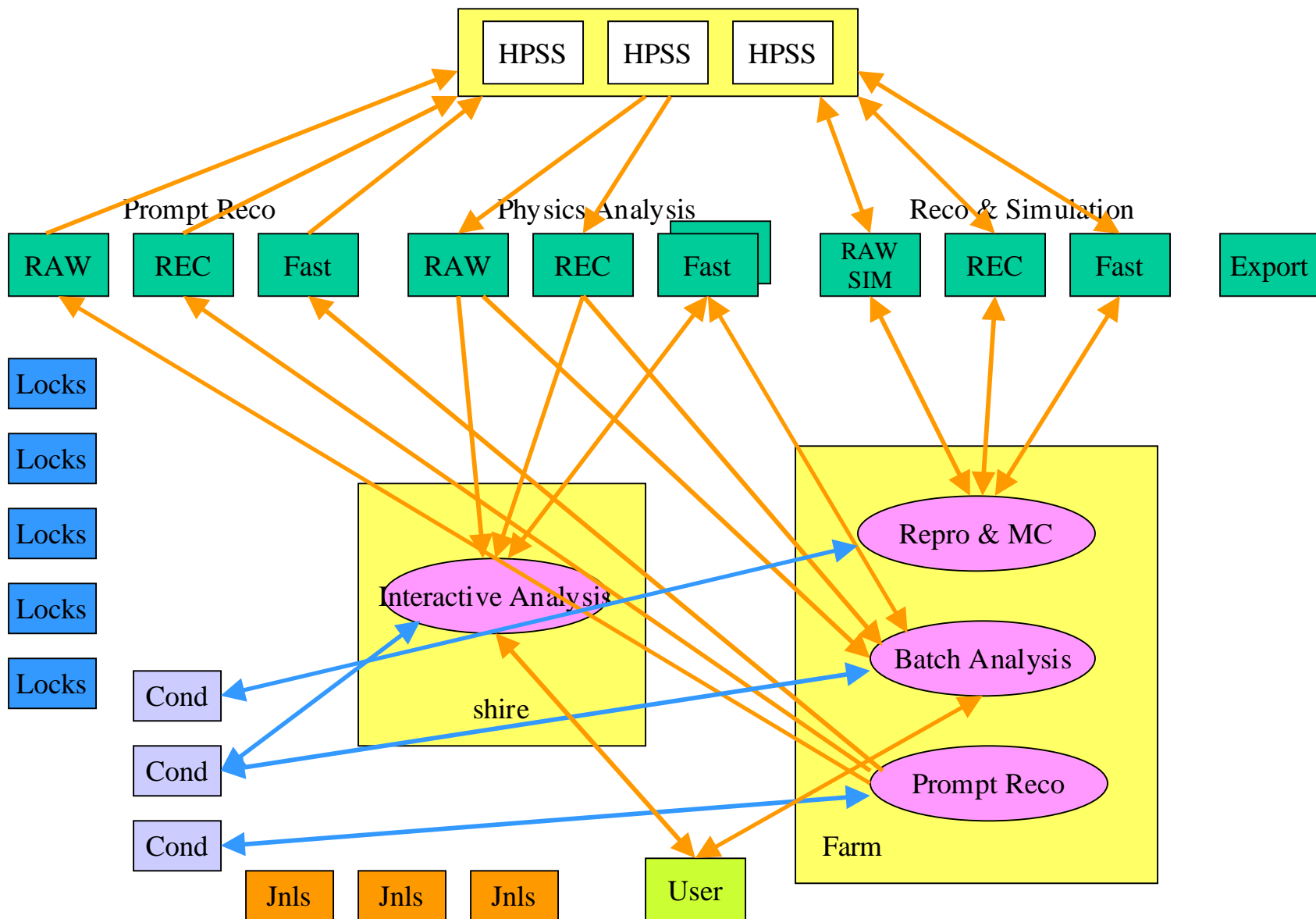
Parallel Activities and Distinct Databases



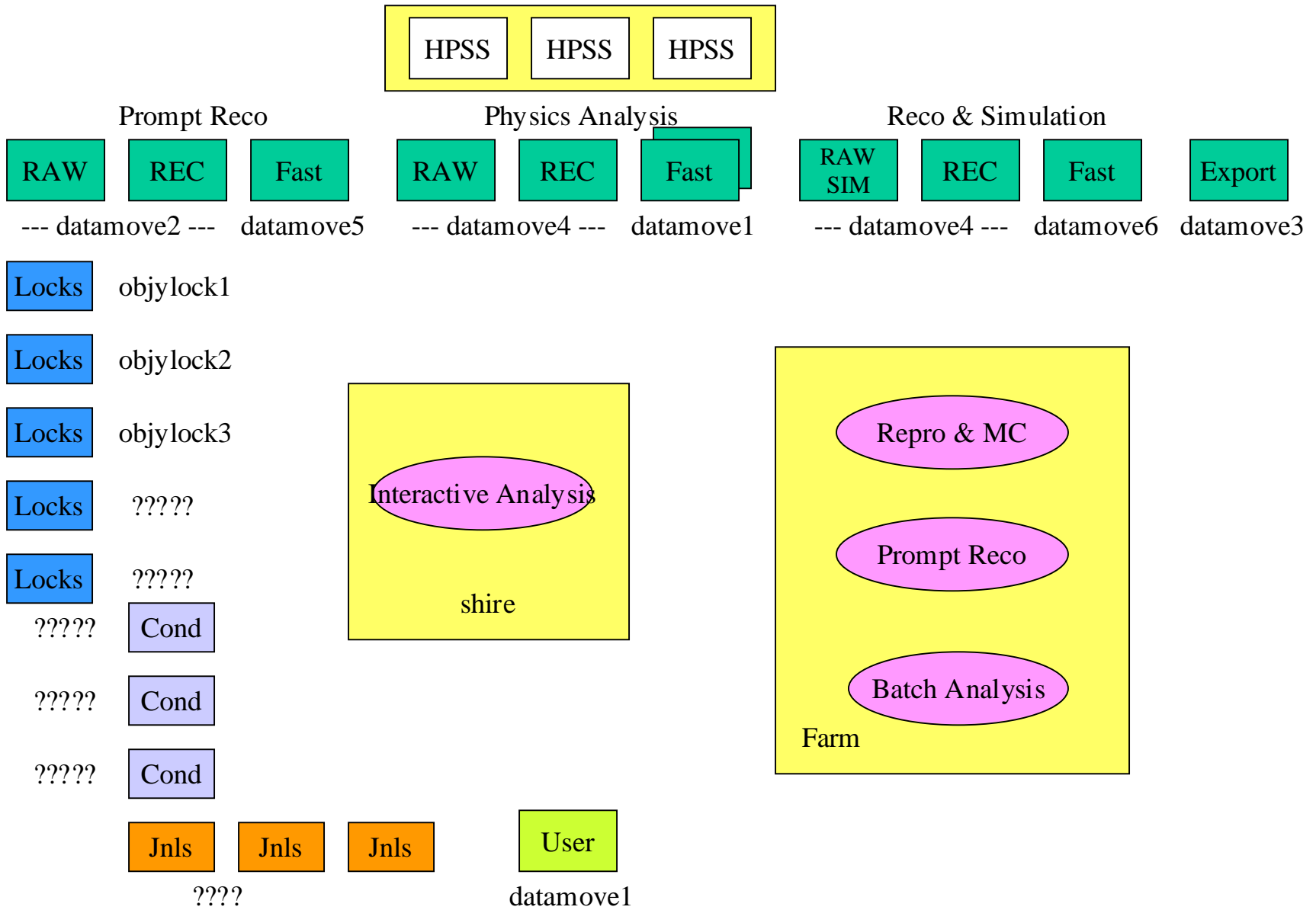
Activities and Data Distribution



Dataflows



Short term mappings (26 Apr?)



Glossary

RAW	4500 w/ 0.5TB
Fast	4500(s) with 1.5TB to 10TB
Locks	Ultra5
Cond	450 or Toaster?
Jnls	Ultra5, 250 or Toaster?
User	450 or Toaster (or done by FAST mover?)

- Should we be using Toaster instead of objyserv1?
(avoids file descriptor limitation)
- Need to backup Cond into HPSS
(at least one Cond server needs to be a 450)
- Need to understand the security implications of using Toaster
mountd -n
- Do we need the User server or can those just stay on FAST mover?
- Manage with 2 Cond servers?

Plans

- **Panic! Finish core code development.**
 - Particular concerns
 - Error recovery / Killer events
 - Monitoring/QA
- **Panic more! Scale to 200 CPUs & 100 Hz**
 - Get Objectivity fixes installed
 - Get server hardware and software setup
 - Get optimized ElfApp to run reliably
- **Panic even more!!! Operations setup**
 - Cleanup PR core scripts / core programs
 - Provide job monitoring
- **Relax - go on summer vacation**