

Cray Phase 2 Migration

Technology update overview

- Replacing Ivybridge Xeon CPUs by next generation Broadwell Xeon CPUs
 - Broadwell CPUs are backwards compatible with Ivybridge executables
- Additional Broadwell nodes
- Interconnect (Aries) network not updated

	Phase 1 (Ivybridge)	Phase 2 (Broadwell)
CPU	24 cores (2 x 12 core) @ 2.7GHz	36 cores (2 x 18 core) @ 2.1GHz
Memory	64Gb (1866 MHz DDR3)	128Gb (2400 MHz DDR4)
Nodes	3,400	3,513 (+3% cf Phase 1)
Cores	84,096	130,212 (+55% cf Phase 1)
Tf _{sustain}	200	320 (+60% cf Phase 1)

- 6 Broadwell nodes \approx 8 Ivybridge nodes (4DVAR runtime)
- 2.5 Broadwell nodes \approx 8 Ivybridge nodes (memory capacity)

Migration timetable

- 21 January : Broadwell nodes added to cct
- 21-23 March : System sessions on cca & ccb
 - 1 new Broadwell cabinet (200 nodes) added to cca & ccb
 - 3 ccb cabinets (600 nodes) upgraded to Broadwell
 - 2800 parallel application nodes of Ivybridge
 - 700 parallel application nodes of Broadwell
- 30 March : ccb Broadwell queues open for internal users

7 weeks of testing
- 16 May : ccb down for 1 week, remaining nodes upgraded to Broadwell

3 weeks of testing
- 13 June : cca down for 1 week, all nodes upgraded to Ivybridge

4 weeks of testing

Migration – what's required? (1)

•Libraries

- Existing (Ivybridge) executables/libraries will work on Broadwell, but may not produce bit identical results
- Compiling with a later compiler release (cdt/15.11 onwards) creates libraries/executables that can run on both Broadwell and Ivybridge with bit reproducible result
 - BUT: Broadwell nodes have 36 cores, cf Ivybridge with 24 cores which will require a different task/thread decomposition to fit – for 4D VAR this will mean a loss of bit reproducibility anyway!
- With the latest compiler releases we can produce libraries/executables which run only on Broadwell, and gives a small (~5%) performance improvement
 - During the initial migration we don't anticipate using this feature, as we want the flexibility of running executables on either Ivybridge or Broadwell nodes
 - We will concentrate on performance improvements after the initial migration

Migration – what's required? (2)

- Scripts

- Minimal additional changes anticipated. Changes to Korn shell will require the same updates as were implemented in the lxc porting branch (widely used already). A small change to run_parallel is also required.

- prepIFS

- All the task/thread/memory allocations need to be updated to cater for the change in node size, and memory size.
 - Plan is to squeeze jobs to maximise memory usage (ie. run using the fewest cores possible)
 - Individual jobs may run slower
 - Overall throughput (and queue time) should decrease
- You will chose whether to run an experiment on Broadwell or Ivybridge nodes
 - Theoretically possible to change after launching, but all the task/thread decomposition information would need to be changed.

Which IFS cycles will IFS Section migrate?

Cycle	Libraries	prepIFS support
32r1 (ERA interim)	cdt/15.11	“an” only
36r4 (S4)	cdt/15.11	No
40r1 (OpenIFS)	cdt/15.11	“fc” only
41r1	cdt/15.11	All types
41r2	cdt/15.11	All types
42r1	cdt/15.11	All types
42r3	cdt/15.11	All types
43r1	cdt/15.11	All types