

Tungregning (HPC)

Eirik Thorsnes

System Engineer
Parallab, BCCS

Oversikt

- Hvorfor trenger vi tungregning / HPC?
- Historie
- Hvordan løses HPC – arkitektur
- utfordringer for HPC
- Ny maskin – Cray XT4
- HPC – innkjøp og infrastruktur
- Spørsmål



Hvorfor trenger vi HPC?

- Løse svært store problemer (klima-forskning)
- Løse tidskritiske problemer (værmelding)
- Moore's lov: cpuer dobler tetthet hvert andre år
 - I løpet av 10 år, omtrent 100x mer ytelse per krone
 - En PC tar igjen en 100 CPU maskin etter 10 år
 - Alternativt kan en forsker som bruker en 100 CPU maskin ligge 10 år foran noen som bruker en enkelt PC



Historie

- Parallab har vært involvert i HPC siden 1985
- De tre siste maskiner ved Parallab:
 - 1996: SGI Origin 2000, 49.9 Gflops, 128 cores, 0.195 GHz MIPS R10000
 - 2002: IBM p690 Regatta, 499 Gflops, 96 cores, 1.3 GHz Power4 dual-core
 - 2008: Cray XT4, 51700 Gflops, 5552 cores, 2.3 GHz AMD Opteron quad-core



HPC arkitektur

- Klynger (cluster) – hyllevare maskiner
- Klynger med raskt nettverk (infiniband)
- SMP (symmetric multiprocessing)
- Klynger av SMP
- MPP (massively parallel processor)



Utfordringer for HPC

- Skalere software (algoritmer)
- Skalere OS (OS-jitter, latenstid)
- Stabilitet og feilrate for store systemer
- Infrastruktur
 - strøm
 - kjøling
 - plass



Ny maskin - Cray XT4

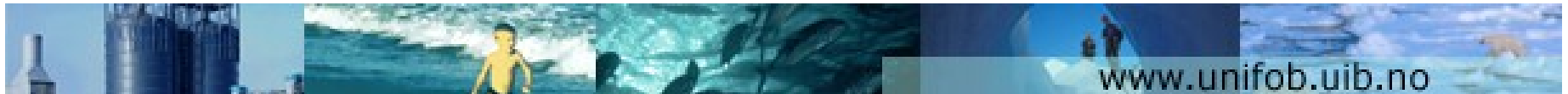
- 51700 Gflops (51.7 Tflops) – ca 30. plass top500
- 6064 GB, 800 MHz DDR2 RAM
- 5552 cores 2.3 GHz AMD Opteron quad-core
- 1388 compute noder
- 300 kW strømforbruk per time
- 254 TB ekstern disk (3 DDN S2A9550)
- 26 service nodes (boot, I/O, login)
- 1 PC for å styre det hele (SMW)



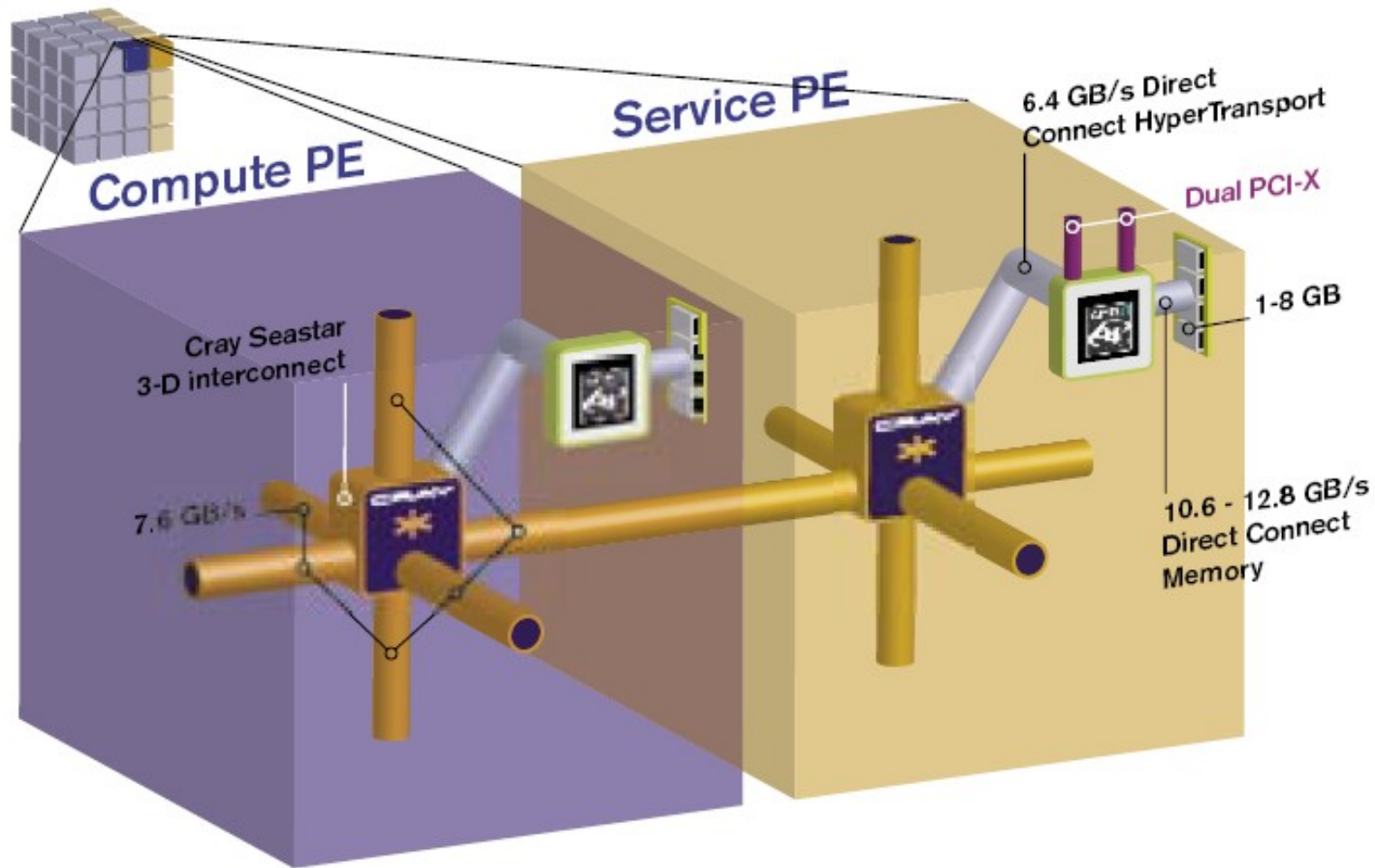


Cray XT4 – brukere

- Hvem kan bruke den nye maskinen?
 - I underkant av 50% av cpu-tid reguleres av NOTUR, alle kan søke dette nasjonale HPC prosjektet om regnetid på de ulike anleggene.
 - Havforskningsinstituttet har investert og har følgelig tilgang til sin andel av maskinen.
 - Forskere ved UiB og Unifob.
- Forbeholdt store, skalerbare problemer
 - mer enn 128 cores per jobb



Cray XT4 - arkitektur



Cray XT4 – CPU

- AMD Opteron quad-core
 - innebygd minnekontroller – lav latenstid
 - Hyper-transport – direkte forbindelse fra CPU til Cray Seastar2 (HSN) eller FC-kort for disk I/O
 - opptil 4 flop per klokkesyklus



Cray XT4 – Seastar2

- Høyhastighet, lav latenstid nettverk (HSN)
- Cray Seastar2 kommunikasjonsbrikke
 - Hyper-transport interface
 - PowerPC 440 CPU
 - mer enn 6 GB/s for hver av de 6 retningene
 - 3D torus topologi
 - avlaster CPU for kommunikasjon



Cray XT4 - OS

- SUSE (SLES 9) Linux på service noder (login)
- Compute Node Linux (CNL) på compute
 - Optimalisert linux kjerne med minimalt miljø
 - Redusert antall interrupts og systemkall for å unngå OS-jitter
- MPI og OpenMP (innen 1 node)
 - standard miljø og kompilatorer



Cray XT4 – HW

- Så få bevegelige deler som mulig
 - 1 stor vifte er eneste bevegelig del i compute del
- Avansert overvåkning og strømstyring
 - separat ethernet basert nettverk (CRMS)
 - skalerbar kontroll-struktur
- 3 DDN S2A9550 disk kontrollere, direkte koblet
 - dobbel paritet på RAID (8+2 tiers)
 - ECC sjekk for både lesing og skriving



HPC – innkjøp og infrastruktur

- HPC innkjøp er svært komplekse
 - første forberedelser for dette innkjøpet ble gjort allerede i 2005, maskinen ble installert i jan. 2008
 - avansert hardware gir høye krav til anbudspapirer og vurdering av tilbud
- Infrastruktur et økende problem for HPC
 - maskinrom 1.etg. HIB er økt med 85 m²
 - helt nytt strøm- og kjøleinfrastruktur
 - kapasitet øket med mer enn 500 kW



Spørsmål

"This new computer enables me to increase the problemsize and simulation time to a point where I get entirely new results, results which makes it possible to compete scientifically with large groups abroad."

(bruker-kommentar etter å ha brukt Cray XT4)

Mer info: <http://www.parallab.uib.no/resources/>

