“From the perspective of system software, what are the biggest challenges facing users of BG?” Panel

Satoshi Matsuoka

Tokyo Institute of Technology / National Institute of Informatics

The 3 challenges, from the perspective of a SC center administrator

- Titech GSIC has the fastest SC in Japan today (TSUBAME Grid Cluster, 85 Teraflops peak)
- Based on that, the 3 challenges:
  1. Ease of Use — Most users run ISV or popular open source code --- the software availability problem
  2. **Ease of Use** — Ease of Porting a variety of Vector, SMP, and MPI code --- the small memory problem
  3. **Ease of Use** — Seamless Integration of the Computing Environment with the rest of the Campus Grid --- the ISA, the OS, compilers/libraries, tools, …

Most users run ISV or popular open source code --- the software availability problem

- What may be seemingly naïve requests from users have grave implications
  - "How do I run Gaussian/NASTRAN"?
  - "How do I run Matlab / Mathematica / SAS?"
  - "How do I run Word/Excel?"
- You may laugh but these are the majority of users we as a center support
  - Supporting just 10 users is not economically sustainable
  - High-end users must live off some leverage of the "commons"

みんなのスパコン
"Everybody's Supercomputer"

TITECH GRID

"Solitary" High-End

Any Freshman @ Titech can use the fastest supercomputer in Japan

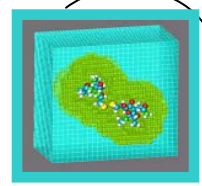
Huge Usage Gap

- Different command environment to familiar Windows, Linux
- Lack of seamless data sharing w/desktops
- Lack of Applications avai. on Clusters
- Lack of programming nevironments, tools
- Batch-only exec. Env.

Might as well use my desktop…

Only serving 20 users spending $10 mil in a 10,000 person organization

My desktop is now 1000 times more powerful



100TF, 1PB scale

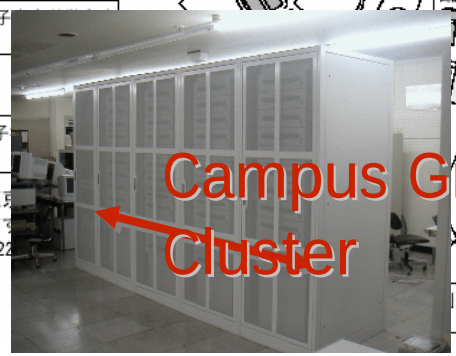
Seamless, Ubiquitous Campus Grid as a core infrastructure for day-to-day usage
➔ Incubation of power users of the future!

Titech Supercomputing Grid 2006

- **~13,000 CPUs, 90 TeraFlops, ~26 TeraBytes Mem, ~1.1 Petabytes Disk**
- **CPU Cores: x86: TSUBAME (~10600), Campus Grid Cluster (~1000), COE-LKR cluster (~260), WinCCS (~300) + ClearSpeed CSX600 (720 Chips)**

35km, 10Gbps

1.2km



Ease of Porting a variety of Vector, SMP, and MPI code --- the small memory (and other) problems

- Large memory SMP, fast balanced Integer/FP, fast I/O are ***catch-all***
- HW requirements of a code may be heterogeneous in those respects
  - In the same code & in coupled simulation scenarios
- Most systems too homogeneous
  - BG/L: lacking in most of above as performance/density tradeoff
  - Fat node SMP clusters: satisfies all above but often overkill
- Answer: tightly-coupled heterogeneous architecture (and the system SW support thereof)

Seamless Integration of the Computing Environment with the rest of the Campus Grid -- the ISA, the OS, compilers/libraries, tools, …

- Users are now seeing ubiquity of the environment as the norm
- Increasing intolerance towards high learning and usage costs
  - Quantifiable merits of using a shared SC versus moderate-sized personal clusters?
  - Users want seamless access so that the Grid is the natural extension of their desk/laptops and clusters
- Low power, high-density PC cluster versus BG/L?
  - Which will the users REALLY prefer?

MegaProto
17 Sockets/1U

Orion
DS-96s

Template

- How can we quantify that challenge?
- What can be done in the next year? In the next two years? Which of these are primarily software and which are strongly affected by hardware
- Are there opportunities for collaboration in solving some of these problems?