Multicore scheduling

Claudio Grandi

INFN Bologna

This is a brainstorming…

… just a fewslides with a list of items / questions

Some of them have already activities associated

Some may be irrelevant

Not proposing solutions… leave them to the audience!

Hopefully not too much CMS oriented ;-)

Not covering multi-threaded/multi-process/parallel **programming**

Sites p.o.v.

- Technology
  - Just Clouds or already on Grids?
  - Whole VMs or fraction of (real/virtual) nodes?
    - assuming no whole real nodes… is this negotiable?
  - Do I need to communicate with clients on slots? How? One-way or bi-directional? Job/Machine Features WG, see also [1]
- Resource management
  - How to avoid dead time while draining nodes? (How to avoid having to drain nodes?) – Fixed vs. variable size slots
    - …may also affectthe memory/core ratio when buying hardware…
  - How to keep 100% resource occupation? In other words: How to provide fair shares with long slot allocation?
  - How do I chose the allocation (slot) to be terminated? See [1]
- Accounting/ Monitoring
  - CPT vs. N*WCT? Draining time? Wasted killed slots time?
  - How to identify inefficiencies, e.g. (partially) idle slots?

VO p.o.v.

- Discovery
  - What kind of resources does an infrastructure (site) offer? Where can I find a resource that suits my needs?
    - What's new: number of cores, cost (e.g. on public clouds)
    - What's old but never done: architecture, memory, disk, …
- Request
  - If a site offers more than one configuration, how do I specify what I need? (already done? Grids, Clouds, …)
- Allocation
  - In the pilot/VM: how do I know what resources have been allocated? How long? May I extend?
- Control
  - How long may I keep running? How can I release resources? All at once or as soon as I free them?
  - How do I (efficiently) mix single and multi-core jobs on a multi-core resource? See [2]

Event services

See [3]

Event services require a completely different approach

- Long lived master + many relatively short lived workers
- Good network connection between master and workers?
  - Is memory sharable? At least on the same host?
- Similar to classic parallel processing (MPI, …)

Ideas from CHEP…

1. I. Sfiligoi "Minimizing draining waste through extending the lifetime of pilot jobs in Grid environments”
2. A. Pérez-Calero Yzquierdo "CMS Multicore Scheduling Strategy"
3. S. Campana "Evolution of the ATLAS Distributed Computing system during the LHC Long Shutdown"