EUROPEAN MIDDLEWARE INITIATIVE

Roadmap for parallel computing framework

Enol Fernández (CSIC)

gLite-MPI PT & MPI TF

JRA1.1.2 Tasks

Subtask	Name	Owner	Due
A11.1	define a proposal for a parallel execution framework within EMI	MPI TF	**M18**
A12.1	implementation of the proposal for a parallel execution framework within EMI	MPI TF	M32
A13.1	enable capabilities to support multi-core, multi-node execution in ARC	Arc CE	M36
A13.2	enable capabilities to support multi-core, multi-node execution in gLite	gLite JM	M36
A13.3	enable capabilities to support multi-core, multi-node execution in UNICORE	UNICORE *	M36
A13.4	enable capabilities to support cross-middleware multi-core, multi-node execution	MPI TF	M36

A11.1 Common Execution Framework

- EMI-ES will provide a common interface for *submitting* the jobs
- ***ParallelEnvironment*** in EMI-ES **is** the ***common framework*** for parallel jobs
- Modifications proposal by the MPI TF:
  - **Type** as a free string (e.g. MPI/OpenMP)
  - Change **ProcessesPerSlot** to **ProcessesPerHost**
  - **ThreadsPerProcess** with additional tag useSlotsPerHost="true"
  - Any extra features in the **option** element

ParallelEnvironment Implementation

- There is not a single parallelenvironment that can cover all kinds of parallel jobs.
- Each middleware stack provides its own mechanism for developing them:
  - ARC: based on shell scripts
  - gLite: ?
  - Unicore: based on XML templates

ParallelEnvironment Implementation

- mpi-start provides PE functionality for common MPI implementations & batch systems:
  - It is independent of the middleware
  - It just interacts with the batch system and the MPI implementation
  - Extensible and open to new parallel applications
  - Easy to configure by the sysadmins (shell scripts)

A12.1: Implementation of the proposal

- EMI-ES + mpi-start as basis for PE implementation
  - Already performed adaptation for ARC RuntimeEnvironments and UNICORE ParallelEnvironmnet.
- To be finished by M32, but should be fast once the EMI-ES implementations are ready

mpi-start roadmap

- Current version in EMI: 1.1.0
  - Open MPI, MPICH, MPICH2, LAM MPI
  - PBS/Torque, (S)GE, LSF, Slurm, Condor
  - Hybrid MPI/OpenMP support
  - Support for binding to core/socket/node
  - Processor/Memory affinity in Open MPI
- Towards EMI 2:
  - Processor/Memory affinity in MPICH2
  - Improved extensibility
  - Improved Slurm & Condor support
  - Bug fixing

A13.4

- A13.4: enable capabilities to support cross-middleware multi-core, multi-node execution (M36)
  - any ideas?
- mpi-start was used in int.eu.grid for multi-site execution using PACX-MPI
  - Requires a node with public IP
  - Main issue: co-scheduling/reservation

EUROPEAN MIDDLEWARE INITIATIVE

Parallel jobs consolidation

Enol Fernández (CSIC)

gLite-MPI PT & MPI TF

Common Execution Framework

- EMI-ES will provide a common interface for *submitting* the jobs
- ***ParallelEnvironment*** in EMI-ES **is** the ***common framework*** for parallel jobs
- Modifications proposal by the MPI TF:
  - **Type** as a free string (e.g. MPI/OpenMP)
  - Change **ProcessesPerSlot** to **ProcessesPerHost**
  - **ThreadsPerProcess** with additional tag useSlotsPerHost="true”
  - Any extra features in the **option** element
- Currently in discussion by MPI TF

ParallelEnvironment Implementation

- There is not a single parallelenvironment that can cover all kinds of parallel jobs.
- Each middleware stack provides its own mechanism for developing them.
- mpi-start *consolides* PE backend for common MPI vendors:
  - Will be adapted to each middleware mechanism

Consolidation Plan

- Consolidation plan created by MPI TF
- EMI-ES and ParallelEnviroment are already agreed
- No products phased out:
  - Each "CE" implements the EMI-ES
  - mpi-start provides backend implementation
- what are the main difficulties in the consolidation of your area?
  - Agreement on common attributes for PE

Consolidation Plan

- Common EMI cli/API, maintenance:
  - EMI-ES will be the common API, maintained by each CE
  - Backend maintained by gLite MPI (mpi-start)
- New products do not affect consolidation, no replacements.

EUROPEAN MIDDLEWARE INITIATIVE

Thank you

EMI is partially funded by the European Commission under Grant Agreement INFSO-RI-261611