

int.eu.grid MPI

Kęstutis Paulikas, KTU
2007-09-24

EGEE MPI palaikymas

- **Nepakankamas ir nelankstus palaikymas iš WMS**
- **Dokumentacijos trūkumas**
- **Techninės kliūtys nedidelės, tačiau labai apsunkina naudojimą**
- **Daug klasterių turi MPI lokaliai, bet MPI neprienamamas per GRID**
- **Sprendžia EGEE TCG WG, kartu su int.eu.grid**

Rekomendacijos

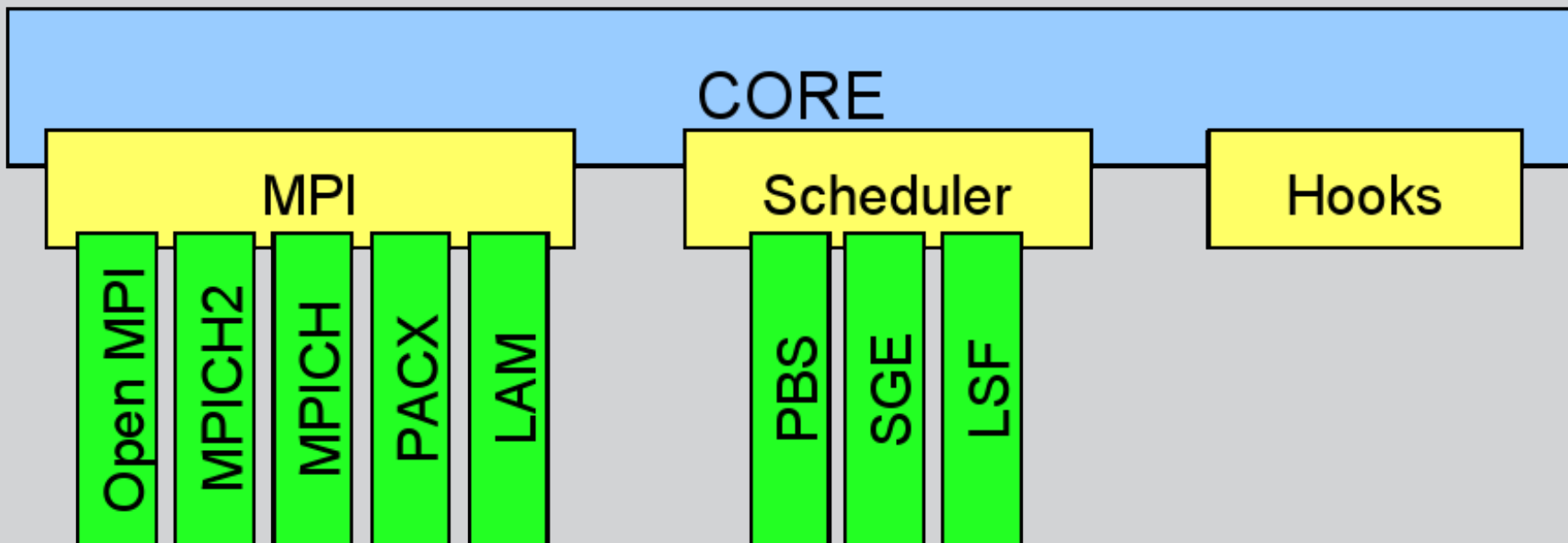
- **ne naudoti specialaus užduoties tipo, o paprastoje užduotyje nurodyti kelių mazgų reikia; MPI aplinką nusistato pats vartotojas**
- **informaciniai tag'ai instaliuotų versijų paskelbimui; aplinkos kintamieji WN, kad užduotys galėtų atrasti MPI**
- **"mpi-start" skriptas vartotojo darbui supaprastinti**

... Rekomendacijos

- **GlueHostApplicationSoftwareRunTimeEnvironment**
- **Aplinkos kintamieji: MPI_MPICH_PATH, MPI_MPICH_VERSION, MPI_SHARED_HOME**
- **"mpi-start" instaliuotas ant CE ir WN**
- **"mpirun" pakeisti tusciu skriptu.**

"mpi-start"

- shell skriptas
- prieš iškviestas "mpi-start", vartotojas nustato aplinkos kintamuosius



"mpi-start" pavyzdys

```
JobType = "MPICH";  
NodeNumber = 8;  
Executable = "mpi-start-wrapper.sh";  
Arguments = "mpi-test OPENMPI";  
InputSandbox = {"mpi-start-wrapper.sh", "mpi-hooks.sh", "mpi-  
test.c"};  
Requirements = Member("OPENMPI",  
other.GlueHostApplicationSoftwareRunTimeEnvironment);
```

JDL

```
# Setup for mpi-start.  
export I2G_MPI_APP=$MY_EXECUTABLE  
export I2G_MPI_TYPE=$MPI_FLAVOUR  
export I2G_MPI_PRE_RUN_HOOK=mpi-hooks.sh  
export I2G_MPI_POST_RUN_HOOK=mpi-hooks.sh  
# Invoke mpi-start.  
$I2G_MPI_START
```

mpi-start-wrapper.sh

```
pre_run_hook () {  
mpicc -o ${I2G_MPI_APP} ${I2G_MPI_APP}.c  
}
```

mpi-hooks.sh

pavyzdys BalticGRID'ui

int.eu.grid planai

- **Testas, SAM sensorius: reikalinga atskira užduotis.**
- **Turi tikrinti, ar paskelbta konfigūracija veikia ("mpi-start", MPI, NFS)**
- **konfigūracija realizuota, kaip Quattor šablonai, YAIM funkcija sukurta**
- **patchas LCG-RB, deramasi su JRA1 dėl gLite WMS modifikacijos**
- **<http://www.grid.ie/mpi/wiki>**
- **project-eu-egee-tcg-mpi@cern.ch**

Klasterių konf.

shared failų sistema

pbs (ne lcgpbs) job menedžeris

instaliuoti MPI irr mpi-start

paskelbti informacija per GLUE RTE

nustatyti MPI aplinkų kintamuosius

- pakeisti "mpirun" tuščiu skriptu

- pakeisti GLUE "torque" į "pbs"

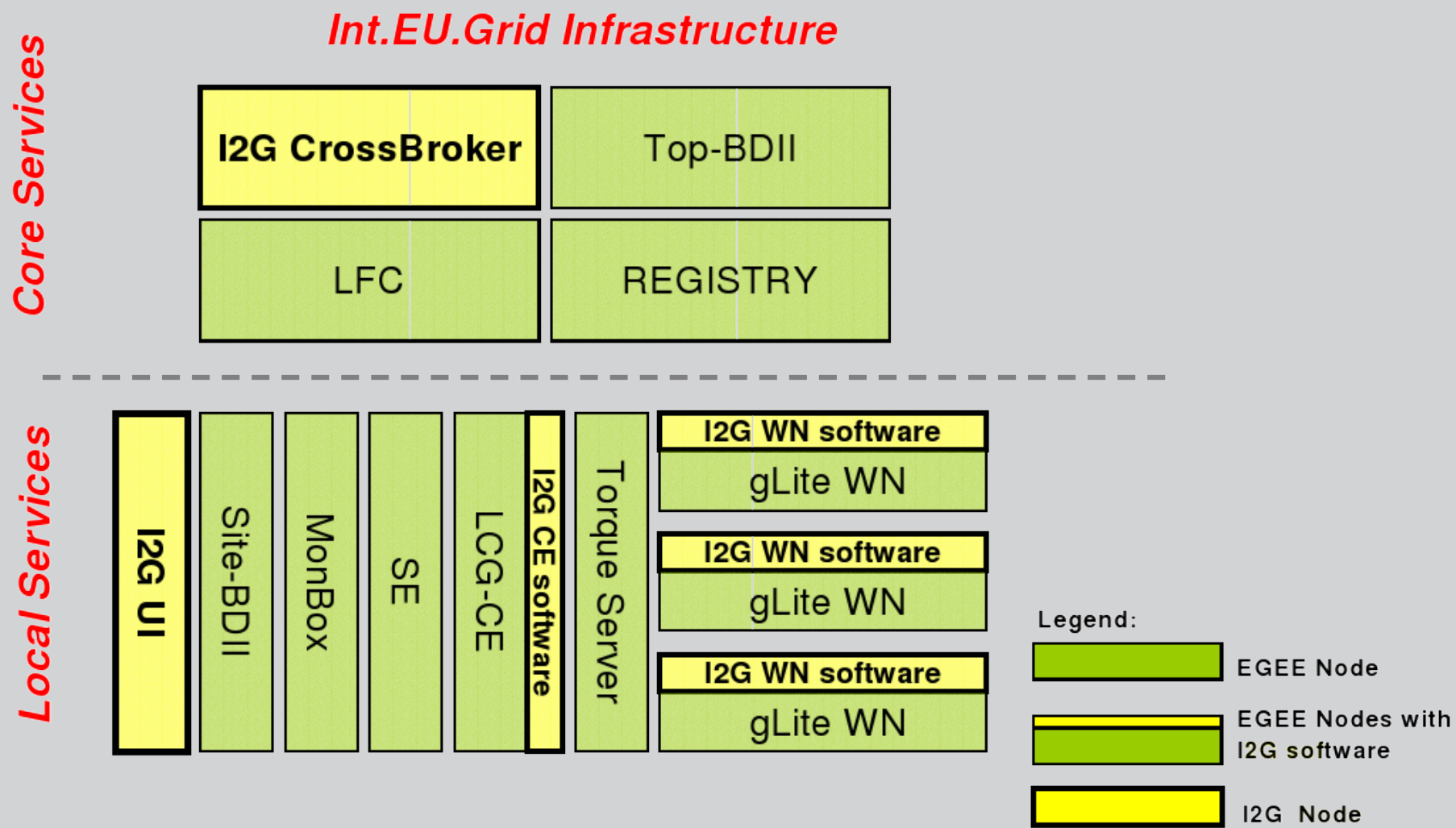
int.eu.grid

- **interaktyvios užduotys**
- **lygiagrečios užduotys (iter-cluster)**
- **vizualizacija (video srautai)**

Programinė įranga:

- **gLite**
- **CrossBroker (LCG-RB + parallel + video)**
- **UI su video palaikymu.**

I2G / EGEE klasteris



CrossBroker

- lygiagrečių (OpenMPI, PACX-MPI, MPICH-P4, MPICH-G2, plain) ir interaktyvių užduočių palaikymas
- JobType: "Paralel"
- SubJobType: "openmpi", "pacx-mpi", "mpich", "mpich-g2", "plain"
- Nėra garantijos, kad inter-site MPI startuos vienu metu.
 - naudot tik didelius saitus
 - išsiskirt resursus ir laukt, kol startuos

Pavyzdys parallel

```
Type = "Job";
VirtualOrganisation = "imain";
JobType = "Parallel";
SubJobType = "pacx-mpi";
NodeNumber = 5;
Executable = "test-app";
Arguments = "-v";
InputSandbox = {"test-app", "inputfile"};
OutputSanbox = {"std.out", "std.err"};
StdErr = "std.err";
StdOutput = "std.out";
Rank = other.GlueHostBenchmarkSI00 ;
Requirements =
    other.GlueCEStateStatus == "Production";
```

Pavyzdys interactive

```
Type = "Job";
VirtualOrganisation = "imain";
JobType = "Parallel";
SubJobType = "openmpi";
NodeNumber = 11;
Interactive = TRUE;
InteractiveAgent = "glogin";
InteractiveAgentArguments = "-r -p 195.168.105.65:23433";
Executable = "test-app";
InputSandbox = {"test-app", "inputfile"};
OutputSanbox = {"std.out", "std.err"};
StdErr = "std.err";
StdOutput = "std.out";
Rank = other.GlueHostBenchmarkSI00 ;
Requirements =
    other.GlueCEStateStatus == "Production";
```

glogin

- <http://www.gup.uni-linz.ac.at/glogin/>
- interaktyvus prisijungimas per Globus API
- portų forwardinimas, VPN palaikymas
- autentifikavimas pagal sertifikatą
- funkcijos panašios į ssh

- glogin turi būti abiejose pusėse (pvz.: UI ir CE)
- jungiamasi į CE, bet ne į WN

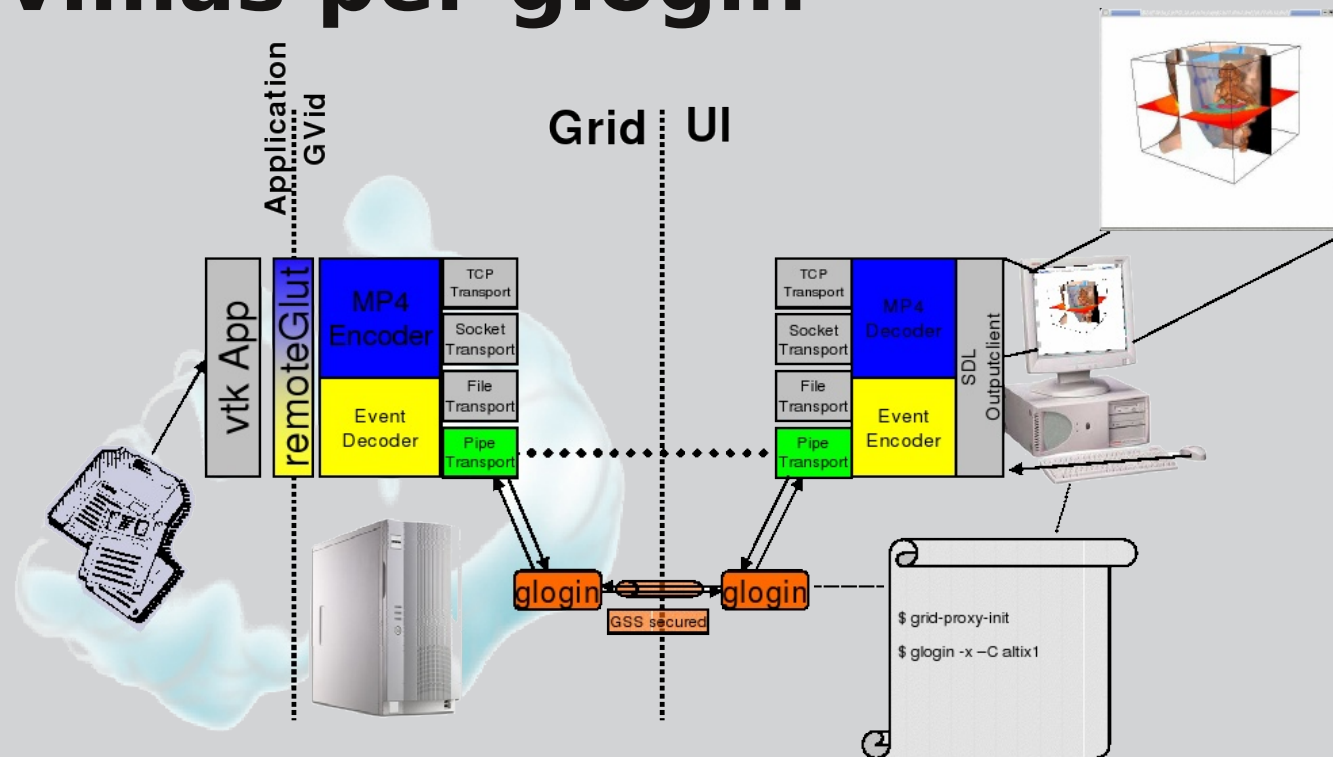
i2clogin

- <https://savannah.fzk.de/autobuild/module-i2g-glogin.html>
- perdarytas glogin
- jungiasi iš WN i UI

- veikia ant daugumos BalticGRID saitų, siunčiant kartu su užduotim
- turi būti iš išorės prieinamas portas ant UI

gvid

naudojamas video codec'as video
informacijos spaudimui
intaraktyvumo palaikymas
komunikavimas per glogin



MPI su MDesktop

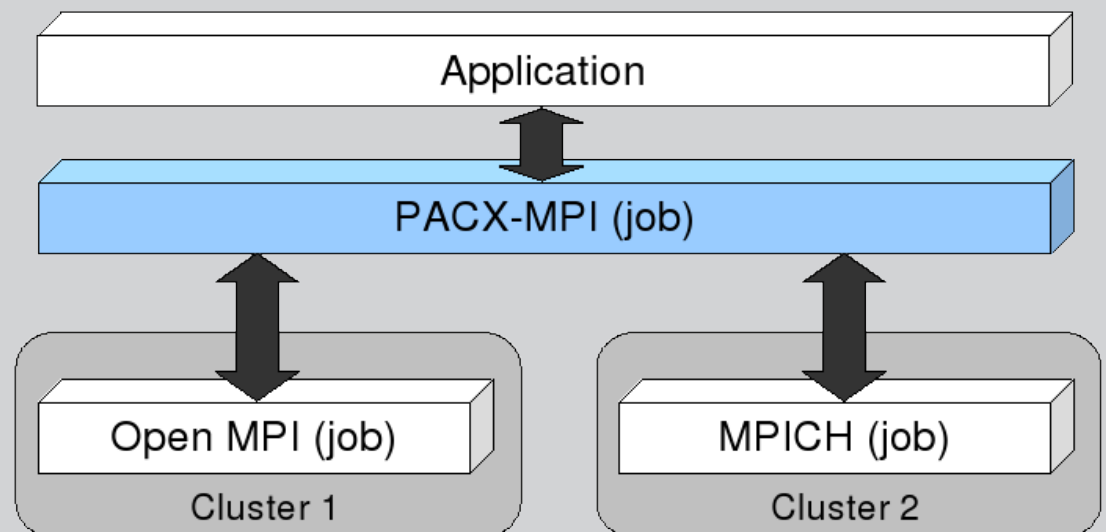
- Palaikomos int.eu.grid savybės (MPI)
- komunikavimui su interaktyviom užduotim naudojamas glogin/gvid

OpenMPI

- **pirma versija 2005m.**
- **tikslai: pilnas MPI-2, švariai perrašytas, pilnas thread palaikymas, išvengt projektų šakojimosi, komponentinė architektūra, pranešimų fragmentacija ir apjungimas, heterogeninių aplinkų palaikymas**
- **paleidžiant prisitaiko prie aplinkos**
- **"ompi_info" komanda informacijai.**

PACX-MPI

- skirtas programos vykdymui lygiagrečių kompiuterių tinkle, sukurta 1995m.
- norint naudoti reikia tik perkompiluoti programą
- naudoja vietoje instaliuota MPI realizacija komunikavimui klasterio viduje



PACX architektūra

- startuoja 2 papildomus komunikavimos procesus kiekviename klasteryje: **out_daemon (rank 0)**, **in_daemon (rank 1)**

