PLCC Cluster Manual

About

This is a short manual for *PLCC cluster*, a collection of scripts which allows you to update the PTGL v3 server with fresh PDB data.

PLCC cluster runs the normal *plcc* java software for many PDB files on a computer cluster or on several machines in a LAN.

Versions

Available versions

Using PLCC cluster, you can run an update of the PTGL in different technical setups:
1. SINGLE: on a single machine (not recommended, most likely takes several weeks)
2. LAN: on some machines connected via LAN, which have a shared network drive (but no queue management software or any other cluster stuff installed)
3. CLUSTER: on a computer cluster which runs the openPBS queue. For this, two sub versions exits:
    1. CLUSTER_MPI: using openMPI. One job contains many (like 200) PDB files, and all cores of the node which is assigned the job work on them until all are done. The job is done when all PDB files are done.
    2. CLUSTER_NOMPI: without openMPI. Each job contains only a single PDB file, and MPI is not used.

Which version to use

Here is a rough comparison of the versions:

	System setup	Job management	Speed
SINGLE	Very easy	None	Very poor
LAN	Easy	None	Medium to OK
CLUSTER_MPI	Very complicated, MPI-specific tech problems	Full queue support, wall time, etc	Very good
CLUSTER_NOMPI	Complicated	Full queue support, wall time, etc	Very good

As of April 2015, we use the cluster version without openMPI (3.2 in the list above) to update the PTGL, and this is the recommended way to do it.