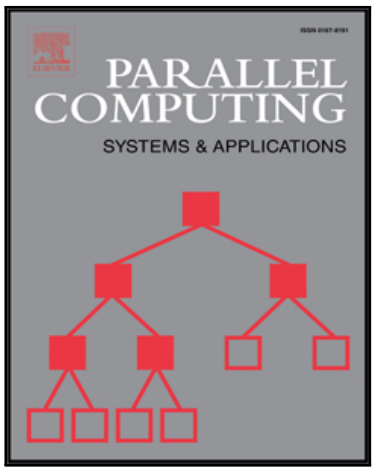

Enabling Scalable and Accurate Clustering of Distributed Ligand Geometries on Supercomputers

Boyu Zhang, Trilce Estrada, Pietro Cicotti, Pavan Balaji, Michela Taufer

PII: S0167-8191(17)30026-1
DOI: 10.1016/j.parco.2017.02.005
Reference: PARCO 2367

To appear in: *Parallel Computing*

Received date: 15 March 2016
Revised date: 26 November 2016
Accepted date: 28 February 2017

Please cite this article as: Boyu Zhang, Trilce Estrada, Pietro Cicotti, Pavan Balaji, Michela Taufer, Enabling Scalable and Accurate Clustering of Distributed Ligand Geometries on Supercomputers, *Parallel Computing* (2017), doi: 10.1016/j.parco.2017.02.005

Highlights

- Scalable method to cluster molecules from docking simulations on distributed systems
- Projections and interpolations into 3-D and 6-D capture molecular geometries
- Our approach scales up to 2,048 processing cores and 2 TB input data
- Our approach is more accurate than energy-based and centralized clustering methods