HiCOMB 2015 Keynote Talk

Genomic Imperfections: Diagnosing the Causes of Rare Disease

Ramesh Hariharan, CTO, Strand Life Sciences

ABSTRACT

Two brothers who die early in life, a family where several individuals lose central vision in their 40s, siblings whose hearts fail in the 30s, an infant with critical bone marrow failure, a baby with cancer of the eye, or 5% of us who have trouble differentiating red from green, are all affected by variations in genomic sequences. Using these examples, this talk will describe how high-performance computing applied to Next Generation Sequencing data helps identify the root genomic cause in such cases.

HiCOMB 2015 Invited Talk 1

Emerging Architectural Frameworks for Microbiome Applications

Ananth Kalyanaraman,
Washington State University, Pullman, USA

ABSTRACT

Microbial communities inhabit various environmental habitats as pervasive and as diverse ranging from a handful of soil to ocean floors, acid drainage, and human gut. The study and detailed characterization of these communities is the primary goal in *microbiome* efforts. In this talk I will present novel parallel frameworks and the role of emerging parallel architectures in shaping the future of such studies. More specifically, the techniques to be described have their roots in massive scale sequence and graphtheoretic analyses.

HiCOMB 2015 Invited Talk 2

Enabling In-situ Analysis of Ligand Geometries in
Drug Design Simulations on Supercomputers

Michela Taufer (Uinv. Of Delaware), Trilce Estrada (Univ. of New Mexico), Pietro Cicotti (San Diego Supercomputer Center), and
Pavan Balaji (Argonne National Laboratory), USA

ABSTRACT

We present an efficient and accurate clustering method for the in-situ analysis of protein-ligand docking datasets on large distributed-memory systems. For each ligand conformation in the dataset, our clustering algorithm first extracts relevant geometrical properties and transforms the properties into a single metadata point in the N-dimensional (N-D) space. Then, it performs a N-dimensional (N-D) clustering on the metadata to search for predominant clusters. Our method avoids the need to move ligand conformations among nodes because it extracts relevant data properties locally and concurrently. By doing so, we transform the analysis problem (e.g., clustering or classification) into a search for property aggregates. Our analysis shows that when using small computer systems of up to 64 nodes, the performance is not sensitive to data content and distribution. When using larger computer systems of up to 256 nodes the scalability of simulations with strong convergence toward specific geometries is less sensitive to overheads due to the shuffling of metadata information. We also demonstrate that our method of metadata extraction captures the geometrical properties of ligand conformations more effectively and clusters and predicts near-native ligand conformations more accurately than do traditional methods, including the hierarchical clustering and energy-based scoring methods.

/15 $31.00 © 2015 IEEE
DOI 10.1109/IPDPSW.2015.174