

# Deployment of Virtual Clusters on a Commercial Cloud Platform for Molecular Docking & Fault Tolerance

*Derek Song  
NAIST, Nara, Japan*

*August 6, 2014*

# Expectations from Previous Week

- Organize the tremendous amount of ligand slices obtained from ZINC Database (In Progress)
- Modify and test Python codes to accurately split raw data of ligand slices (Completed)
- Compare test results from adjusting Amber Score parameters (Completed)

# Research Progress

- Adjusted Amber Scoring input file “dock.generic.amber.in” to test for most efficient time with accurate result
- 5-ligand test file test.mol2 was used

Default	26393s (~1.5 hrs/mol)	Result
Max_Ranked_Ligands=500	28499s (~1.6 hrs/mol)	same
Amber_Score_md_Steps=2500	15886s (~0.9 hr/mol)	very different
Amber_Score_md_Steps=3500	20494s (~1.1 hr/mol)	very different
Amber_Score_md_Steps=4000	25296s (~1.4 hr/mol)	very different

- Default was the most accurate and efficient

# Research Progress



- Downloaded Clean Drug-Like ligands from ZINC Database
- Unpacked and unzipped one slice for testing with python codes
- Python successfully spliced the slice into pieces containing 1000 molecules each
- Modified python codes further to count the number of molecules remaining in the last spliced slice
- Attempted to decompress all Clean Drug-Like ligand slices in my personal computer, but failed due size of files, time efficiency, and connectivity to internet
- Decompressing and organizing all ligand slices inside of a Virtual Machine

# Overall Plan of Action

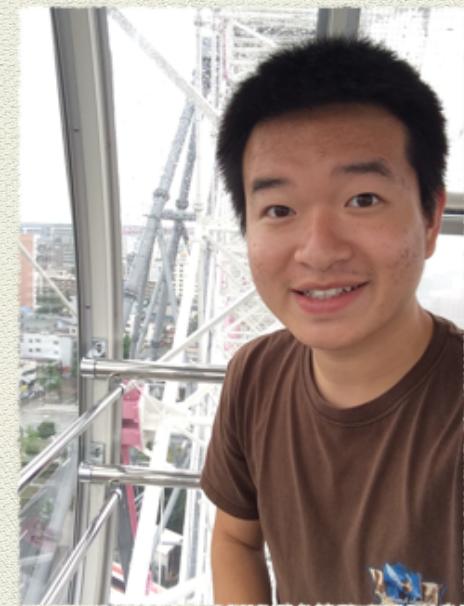
- After obtaining the raw ligand slices from ZINC Database online, I successfully spliced one slice into sub-slices with 1000 ligands inside of each. It was a very exciting news and a critical step to our project. Because of the size and amount of raw ligand slices are very large, I will be organizing them on a Virtual Machine. I also plan to look more into Hadoop and learn the process of running Dock with Hadoop.

# Research Expectations

- Go over the process of running Dock with Anthony
- Finish unpacking, unzipping, and organizing raw ligand slices on virtual machine
- Read more on Hadoop



*Tanabata @ Kyoto,  
Osaka Aquarium,  
Tenpozan Ferris Wheel,  
Tsutenkaku Tower @Osaka*





#UCSDPRIME2014  
*Osaka Aquarium, Osaka*

# Acknowledgments

- ◆ Mentors
  - ◆ **Dr. Jason Haga**, Advanced Industrial Science and Technology
  - ◆ **Dr. Kohei Ichikawa**, NARA Institute of Science and Technology
- ◆ UCSD PRIME
  - ◆ **Dr. Gabrielle Wienhausen**
  - ◆ **Teri Simas**
  - ◆ **Jim Galvin**
  - ◆ **Madhvi Acharya**
- ◆ Funding
  - ◆ Japanese Student Services Organization (JASSO), NAIST
  - ◆ **Dr. Abbie Celniker**, UCSD Alumna
- ◆ Special Thanks to **Karen Rodriguez, Kevin Lam, previous PRIME students, and graduate students in lab**

