**Computer Science in Modern Biology Workshop**

**August 2020**

**LESSON PLAN: Introduction to High Performance Computing (HPC)**

Instructor: Jens Mueller

Requested TAs: TBD

**Logistics:**

Software needed: NX Nomachine from nomachine.com, WinSCP from winscp.net (Windows), filezilla from filezilla-project.org (MAC) MobaXTerm from <https://mobaxterm.mobatek.net/> (Windows)

What needs to be done in advance?

* Install packages on the students personal (laptop) computers
* Have Redhawk cluster student accounts created 2 weeks prior to workshop
* Have premade scripts and datasets on cluster ready for Exercises
* Have problem sets for students to deal with for Day 2
* Make cluster reservations for class 2 weeks prior to workshop

**Lesson plan, Day 1 (10am-12:30pm)**

**Intro – Overview of HPC:** **50 min**

* Historical background of high performance computing (HPC)
* Application areas of HPC
* Overview of current HPC systems in the world and in the US
* Overview of the Redhawk cluster

**Break 10 min**

**Accessing HPC systems 25 min**

* Shell access
* Desktop access via NX Nomachine and MobaXterm
* File system overview and file transfer

**Break 5 min**

**Interactive: Configuring access for users and connecting to Redhawk 60 min**

* Configuring NX, WinSCP, MobaXTerm, filezilla
* Configuring ssh keypairs (optional)
* Launching shell and accessing compute nodes

**Lesson plan, Day 2 (10am-12:30pm)**

**Intro to compute jobs: 40 min**

* Designing, configuring and submitting compute jobs (interactive, batch)
* Monitoring processes, resource usage and compute jobs
* Managing Input/Output (I/O)
* Overview of parallel compute jobs

**Break 10 min**

**Exercises I 35min**

* Design of a basic compute job (R, Matlab, Python …)
* Submission and monitoring of compute job
* Postprocessing

**Break 5 min**

**Exercises II 45 min**

* Design of parallel compute job (Matlab)
* Submission and monitoring of compute job
* Postprocessing

**Q&A: 15 min**