**Equipment**

The Genome Informatics Facility (GIF) has bought into a shared resource named Condo that has a total of 2784 total processor cores and 25,516 GB of RAM that is housed and maintained by the High Performance Computing facility at ISU. This shared resource has 174 nodes. Each node has 128 GB of RAM except for two, which have 1,000 and 2,000 GB of RAM, respectively. GIF can use 315 thousand cpu hours every month on this machine. GIF also purchased 110 Terabytes of redundantly backed-up RAID-6 storage space, 10-20 Terabytes of which is available to conduct these analyses. This machine is where the bulk of the data analysis will be performed (genome assembly, annotation, alignment, comparative genomics etc). If we find the computational resources at ISU are insufficient for the assembly due to limitations in RAM or for the annotation due to limitations in processors then as an XSEDE campus champion for Iowa State, Andrew Severin will obtain a research allocation of computational resources on the NSF eXtreme Science and Engineering Discovery Environment (XSEDE). The resources at XSEDE include Bridges, which has 16 nodes with 12,000 GB of RAM, four nodes with 3,000 GB of RAM for memory intensive programs and many hundreds of nodes with 128 GB of RAM for compute intensive programs.

GIF collaborates with researchIT in the College of Liberal Arts and Sciences to host secure, redundantly backed-up, virtual machines (VM). These VMs host the GIF website, DokuWiki, and JBrowse. DokuWiki is used as a secure online notebook to record the commands and programs used in the genome assembly, annotation and comparisons. The wiki page is made available to all members of the research team to provide feedback and enhance coordination, collaboration, and communication between members of our team. A local installation of the Genome Browser (JBrowse) is used to visualize the next generation sequencing data to ascertain the quality of the assembly and annotation. A secure JBrowse instance will be created for members on the project. The final assemblies and annotations will be hosted by ISUGIF.

All data is redundantly backed up using RAID-6 storage to prevent loss of data. Data will be provided to other members of the project team through password-protected access to folders containing the data, GIF has an archive server with 132 TB of storage. This machine serves as tertiary backup of all raw data and scripts that generate data analyses. This secondary site backup ensures the safety and integrity of the raw data and analyses in case of catastrophic failure at the primary site of data analysis in the High Performance Computing Facility at Iowa State University.