# Hunter College Bioinformatics Cluster

Silicon Mechanics proposes to provide a fully-integrated, high-performance, scalable, centrally managed compute cluster with fast scratch storage.

This document outlines the management software, network, compute and storage resources that will constitute the cluster. It also includes discussion of options with regard to competing CPU architectures.
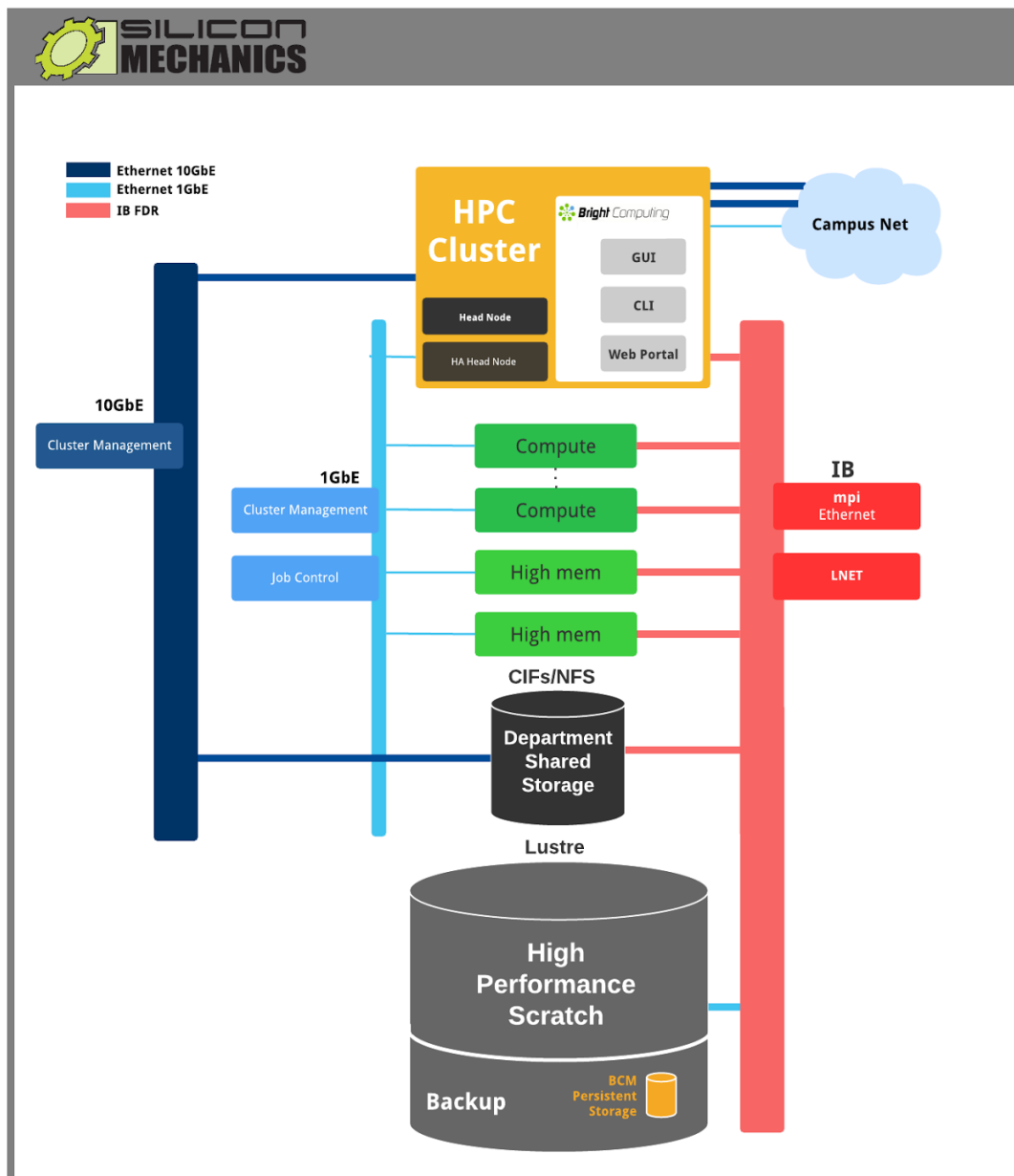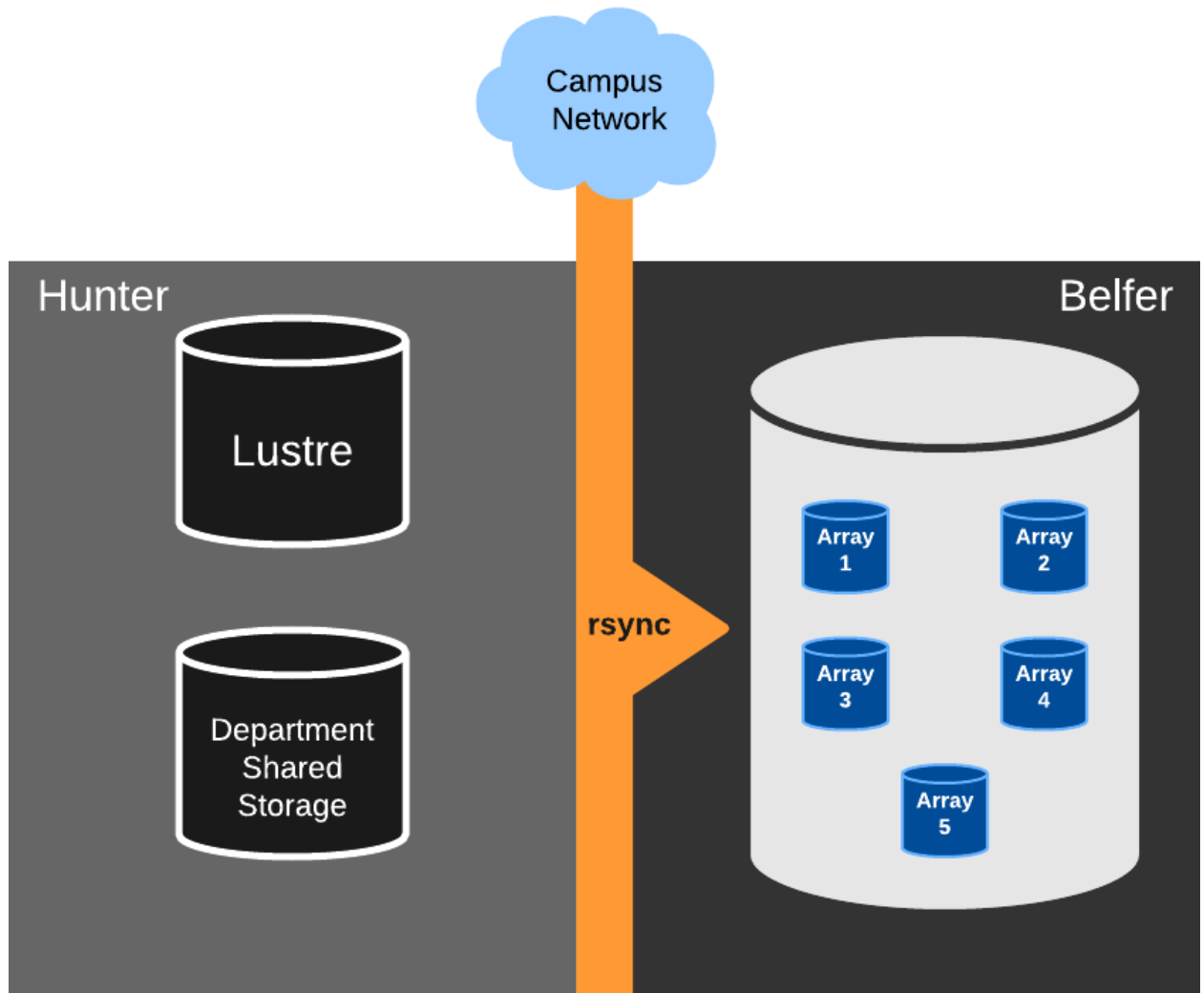
Storage
- Shared Scratch
  - As much storage as possible up to 5PB
  - data workflow
    - Sequencers and propitiatory on-prem data generating sources (expect CIFS)
      - Source [CIFS / 10GbE] -> Dept Shared Storage Node [LNET] -> Scratch
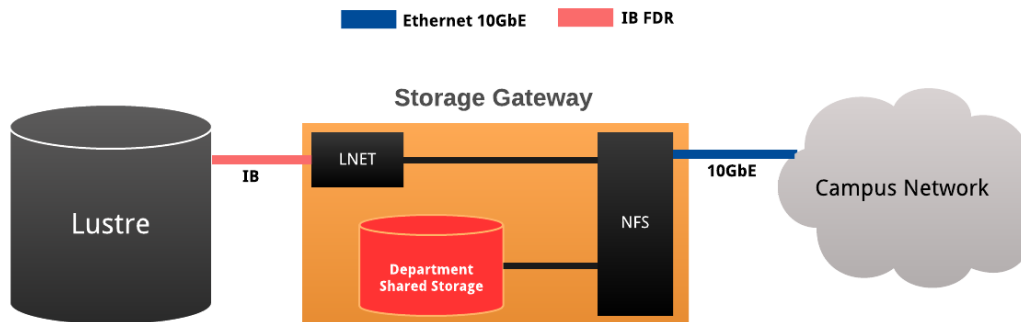
    - Ingest rate
      - 10GbE
      - 1 x sequencer (Win7)
    - Select data is configured to be backed up to NAS located in Belfer
  - Computation generates results data
    - egress to permanent storage
    - backup strategy
      - Rsync over 10GbE campus network from Hunter College to Belfer.
      - Select files/folders will be backup
  - Backup
    - The backup will be at Belfer (0.5 miles from Hunter). A 10gig point to point connection between the 2 building. Select files/folders on the Lustre appliance will be backed up to the backup solution.
    - 180TB of name plate storage (~160TB after file system creation) spread over 5 independent arrays. Each array has 2 parity disks and a hot spare for extra data protection.
    - Requirements
      - 8 Rack Units
      - 4 power outlets
      - watts: 1111
      - va: 1140
      - btu/h: 3791

- ○ amps:110V: 10.4
- ○ amps:208V: 5.5

- ●
  Local Scratch
    - ○ Per node may not be needed now that fast shared scratch is deployed
- ● Dept Shared Storage
    - ○ 5TB home dirs - not much scaling
    - ○ Backup strategy
   - ■ backup up onto supplied Network Attached Backup Storage located in Belfer

Compute
- ● Compute Node
    - ○ Standard job run
    - ○ Single Queue
        - ■ Queue Name: TBD
- ● Mid Memory
    - ○ Single Node
    - ○ Single Queue
        - ■ Queue Name: TBD
- ● High Memory
    - ○ Single Node

- ○ Single Queue
  - ■ Queue Name: TBD

Networking
- Segmented network
  - ○ 1GbE Ethernet
    - ■ management (ssh, ipmi)
    - ■ provisioning (pxe)
    - ■ job control
    - ■ some NFS
  - ○ 10GbE VLAN
    - ■ **Requires 3** ports on the campus network
    - ■ Head, Dept Shared Storage, Campus Net
      - ● NFS, CIFS
        - ○ Bridge to LNET
    - ■ Head node and Shared Storage Node are directly connected to campus network
    - ■ Backup data over campus network to Belfer.
  - ○ Infiniband
    - ■ MPI
      - ● Cluster jobs
    - ■ LNET
      - ● NFS/CIFS share to users on Campus Net
- Protocols
  - ○ MPI (HPC application communication)
  - ○ LNET
    - ■ shared high performance cluster storage (scratch)
    - ■ Bright Cluster Manager persistent data

Proposed Cluster Physical Layout
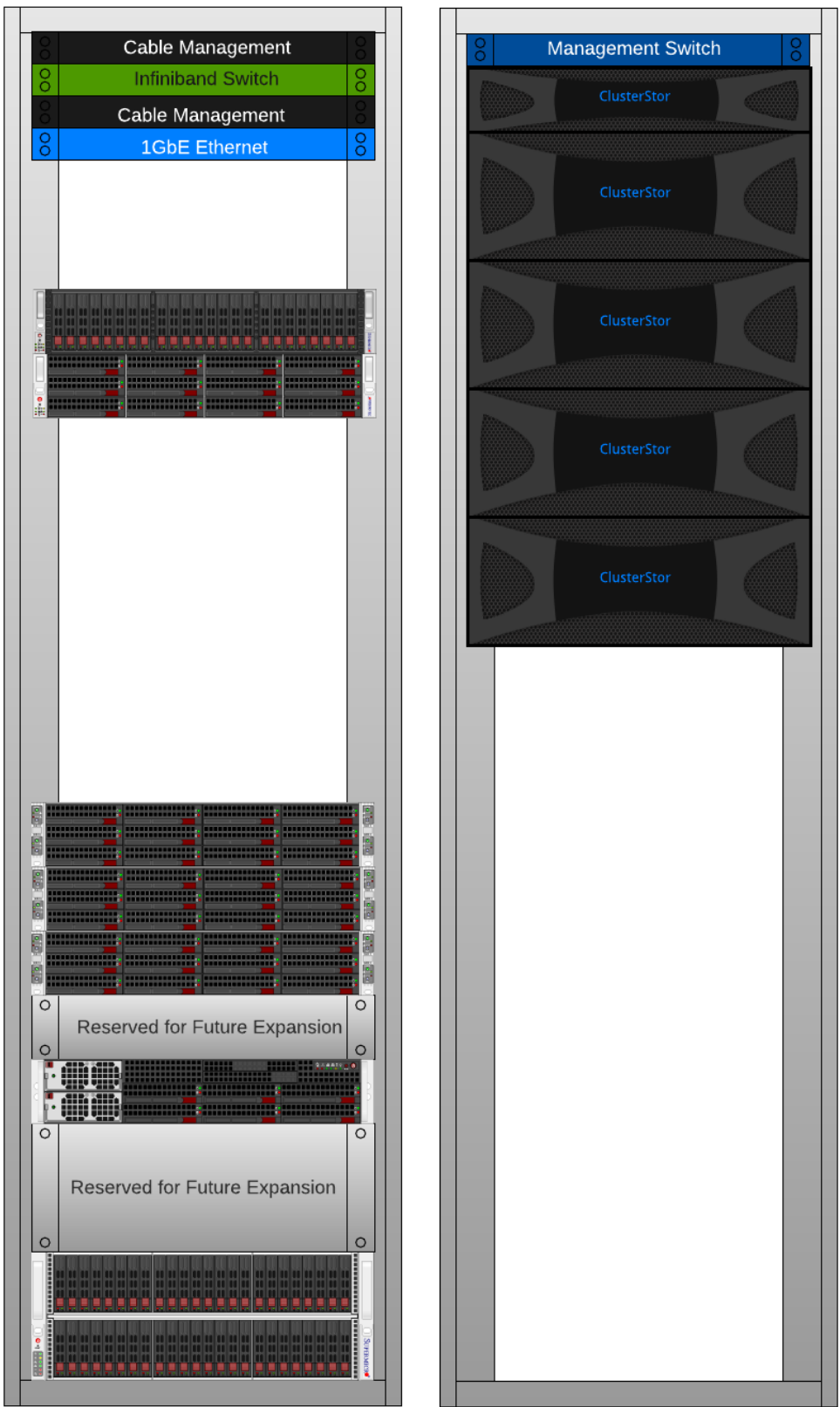
Cluster Middleware and Management

- GUI (or CLI) access to all cluster management functionality
- OS to be used is CentOS 6.5
- Workload manager is Grid Engine
- Provisioning and Image Management
    - Multiple images can be supported
    - RPM packages can be added and images can be changed without re-booting
- Monitoring and Cluster Management
    - Head Node manages Compute Nodes' IPMI services
    - User authentication
    - Wide range of cluster performance metrics
- Scalable and Available
    - Supports future GPU implementation
    - Supports dual Head Nodes with failover (Included)