**PDL Computing Systems**

### Introduction
The PDL provides a variety of computing systems for research and teaching
- Historically, most work was accomplished with dedicated HW
- Trend is toward specialized clusters, often built from virtual machines
- This poster provides a survey of our major systems
- Most of these systems are housed in the Data Center Observatory http://www.pdl.cmu.edu/DCO/index.html

### OpenCirrus
An open cloud-computing research testbed designed to support research into the design, provisioning, and management of services at a global, multi-datacenter scale
- Tashi testbed (big data management in VM-based cluster)
- Exploring performance benefits of data duplication in HDFS
- Using hadoop with panFS
- Evaluation of a MANET attack
- Lazybase, studying high-throughput continuously-updated analytical databases
- Rabbit filesystem with elastic capabilities
- Do-not-track effectiveness research
- https://opencirrus.org/

**Software:**
- Linux (Debian), KVM
**Hardware:**
- Virtual Machines are the primary resource
- Physical machine stats: 108 nodes, 864 cores, 1.7 TB RAM, 90 TB disk

### OpenCloud
A Hadoop cluster used for both research & teaching purposes
**Primary use cases:**
- Big data/science research (Astro-DISC, Quake, etc.)
**Software:**
- Hadoop, HDFS, Accumulo
**Hardware:**
- 113 nodes, 841 cores, 462 TB disk, 1.7 TiB RAM, 10 GbE network

### Small Specialized Clusters
Other machine clusters with long-term dedicated use cases
- Biometrics
  - 4 8-way machines
- FingerPointing/problem diagnosis
  - 38 machines
- LTI project servers/Oracle Grid Engine
  - 18 machines
**Software:**
- Purpose-specific choice of software

### vCloud
General-purpose VM cluster
**Status and features:**
- Offers groups "virtual data centers"
- Ongoing effort to engage external users
- Graphical instrumentation platform
**Software:**
- VMware’s vCloud, Hadoop On Demand
**Hardware:**
- 128 nodes, 1,248 cores, 3.375 TB RAM, 10 GbE network
- 120 TB NetApp filer with 10GbE and 8Gb FCP

### PRObE
- A computer facility dedicated to systems research
- Allows hands-on operation of very large computing resources
- Researchers have complete control of the hardware
  - Can inject both hardware and software failures while monitoring the system
  - Any operating system can be deployed
- Supports research in many systems related fields
  - As far as we know, no other system at this scale in the world provides this ability
  - CMU hosts two instances of the PRObE system
**MARMOT**
- 128-node PRObE staging cluster
- 1 GbE and 4x DDR Infiniband
- 256 cores/2.0 TiB RAM total
- Disk: 128 TB across 128 spindles
- 40 GbE backbone connection; 4x 10 GbE Panasas IB Routers
**SUSITNA**
- 34-node PRObE testbed for high core count research
- 40 GbE and InfiniBand 4x FDR10
- 2176 Cores total (4xAMD Opteron 6272 per node)
- 4.25 TiB RAM total (128 GiB per node)
- Disk: 238 TB across 102 spindles
- nVidia Kepler K20 GPU

### General Purpose Machines
Pool of physical machines available for short-term projects
**Use cases:**
- Flexible according to user needs
- Short term projects
**Software:**
- User is free to image machine to suit their needs
**Hardware:**
- 144 machines (1-2 processor cores, 2-4 GiB RAM, 250 GB+ SATA HDDs, GbE interconnect)