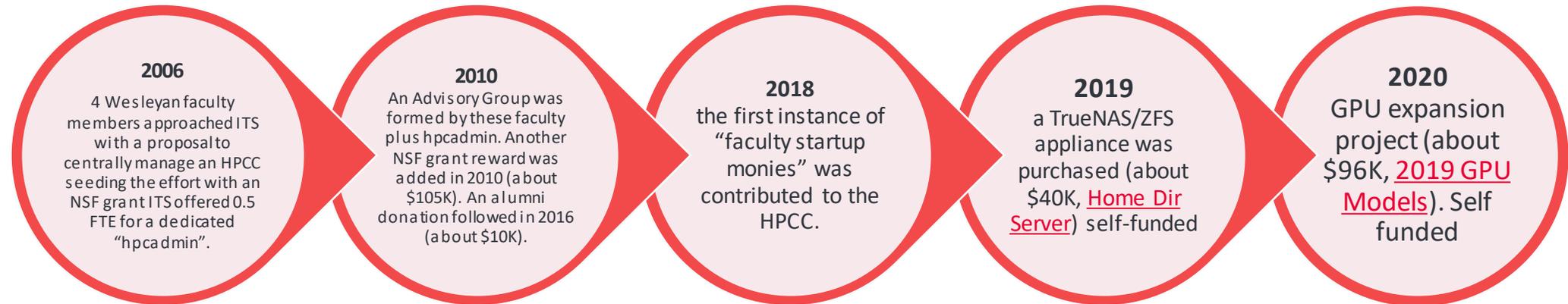# Wesleyan University

Eastern Regional Network
Broadening the Reach Workshop
December 11, 2020

Karen Warren
Deputy CIO
Wesleyan University

History of HPC and Research

2006

4 Wesleyan faculty members approached ITS with a proposal to centrally manage an HPCC seeding the effort with an NSF grant ITS offered 0.5 FTE for a dedicated "hpcadmin".

2010

An Advisory Group was formed by these faculty plus hpcadmin. Another NSF grant reward was added in 2010 (about $105K). An alumni donation followed in 2016 (about $10K).

2018

the first instance of "faculty startup monies" was contributed to the HPCC.

2019

a TrueNAS/ZFS appliance was purchased (about $40K, Home Dir Server) self-funded

2020

GPU expansion project (about $96K, 2019 GPU Models). Self funded

2006 cluster - ~$190K, two racks full of Dell PE1950, a total of 256 physical cpu cores on Infiniband

2020 cluster – 1,450 physical cpu cores (all Xeon), 72 gpus (20x K20, 4x GTX2018Ti, 48x RTX2080S), 520 gb of gpu memory and 8,560 gb of cpu memory. Provided by about 120 compute nodes and login nodes.
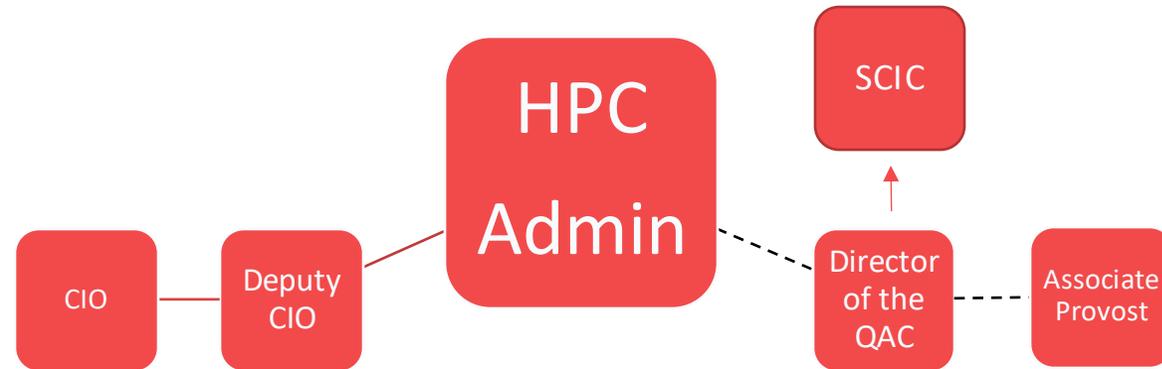
Wesleyan University

Structure

The Wesleyan HPCC is part of the Scientific Computing and Informatics Center (SCIC).

The SCIC project leader is appointed by the Director of the Quantitative Analysis Center QAC. The Director of the QAC reports directly to the Associate Provost.

The hpc admin has a direct report with the ITS Deputy CIO and an indirect report with the QAC Director.

The QAC has an Apprenticeship Program in which students are trained in Linux and several programming languages of their choice and other options (like SQL or GIS). From this pool of students, the hope is some become the QAC and SCIC help desk and tutors.

Wesleyan University

Funding

After an 8 year run of the HPCC, and a drying up of grant opportunities at NSF, it was decided to explore self-funding so the HPCC effort could continue without external dependencies on funds. A *report was compiled of the HPCC progress* including topics such as Publications, Citations, Honors Theses, Growth in Jobs Submitted, Pattern of Pending Jobs, and a General Inventory.

Academic Affairs would annually contribute $25K *if* the HPCC user base raised $15K annually in contributions. These funds would "roll over". That would amount to $160K in 4 years, enough for a hardware refresh or new hardware acquisition. Finance also contributed $10K annually for maintenance such as failed disks, network switches, etc … but these funds do not "roll over".

Wesleyan University

Funding Contribution Schedule

In order for the HPCC user base to raise $15K annually, CPU and GPU hourly usage monitoring was deployed (using scripts parsing the lsb.acct file). A dictionary is maintained listing PIs with their associated members (student majors, lab students, grads, phd candidates, collaborators, etc).
Each PI then quarterly contributes to the user fund based on a scheme supposedly yielding $15K annually.

Contribution Scheme for 01 July 2019 onwards
Hours (K) - Rate ($/CPU Hour)

- 0-5 = Free
- >5-25 = 0.03
- >25-125 = 0.006
- >125-625 = 0.0012
- >625-3125 = 0.00024
- >3125 = 0.000048

A cpu usage of 3,125,000 hours/year would cost $ 2,400.00
A gpu hour of usage is 3x the cpu hourly rate.

Priority Policy

1. *Contributions, of any kind and from any source, immediately become a community wide resource.*

2. *Priority access is granted for 3 years starting at the date of deployment (user access).*

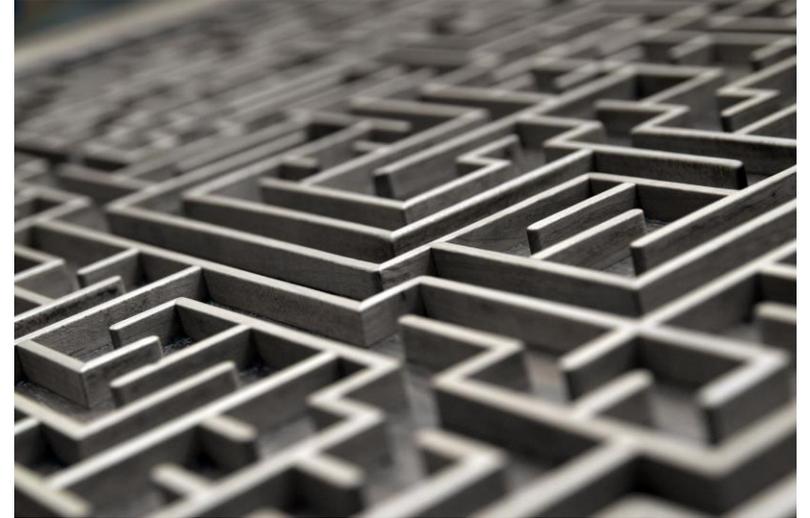3. *Only applies to newly purchased resources which should be under warranty in the priority period.*

The first principle implies that all users have access to the new resource(s) immediately when deployed. Root privilege is for hpcadmin only, sudo privilege may be used if/when necessary to achieve some purpose. The hpcadmin will maintain the new resource(s) while configuration(s) of the new resource(s) will be done by consent of all parties involved. Final approval by the Advisory Group initiates deployment activities.

The second principle grants priority access to certain resource(s) for a limited time to a limited group. The same PI/users relationship will be used as is used in the CPU/GPU Usage Contribution scheme. Priority access specifically means: If during the priority period the priority members' jobs go into pending mode for more than 24 hours the hpcadmin will clear compute nodes of running jobs and force those pending jobs to run. This by now is an automated process.

The main objective is to build an HPCC community resource for all users with no (permanent) special treatment of any subgroup.

Wesleyan University

Challenges

- **Network Infrastructure:** Transferring data generated in labs (instruments in labs, or internet downloaded and locally manipulated) to the HPCC and copying results back is time consuming. Twice Wesleyan has tried to get grants to deploy a 10G network between several science buildings to speed up transfers.

- **Storage:** There is a lot of content being transferred to and from the HPCC storage platform (available to all compute nodes) and the science labs. Multiple copies float around in many locations, some with poor or no backup schemes. Sharing content, either between labs or with other universities, is problematic. A solution where one central point of managed access is provided for all (researchers, collaborators, compute nodes) would go a long way to facilitate research (e.g. Globus https://www.globus.org). A possible funding angle might be collaboration with other small liberal arts colleges, each contributing their expertise in different areas (physics, engineering, social media, deep learning maybe).

- **Programming support:** The apprenticeship program helps and provides continuity. Faculty mostly rely on their own skill set for complex work. The university graduate body is small compared to their aspirations. How to fund this is not clear to me. Resources are available at large universities (like via XSEDE), but that requires quarterly allocation requests and proposal submissions, which is seen as a hassle.

Wesleyan University

Sources, contacts and Information

History of the HPC
https://dokuwiki.wesleyan.edu/doku.php?id=cluster:189

Scientific Computing and Informatics Center
https://www.wesleyan.edu/scic/

Quantitative Analysis Center (QAC), Manolis Kaparakis, Director
https://www.wesleyan.edu/qac/

Apprenticeship Program
https://www.wesleyan.edu/qac/apprenticeship/index.html

Henk Meij, HPCC Administrator and Unix Systems Manager
hmeij@wesleyan.edu

Karen Warren, Deputy CIO
kwarren@wesleyan.edu