UMBC High Performance Computing Facility (HPCF)

Complete information at the HPCF webpage hpcf.umbc.edu:

- Research including public relations: list of projects, publications, outcomes, posters
- Resources for users: system description with schematics and photos, tutorials and sample code, link to consulting support, supporting material, account request form
- Point of contact: Matthias K. Gobbert, *gobbert@umbc.edu*, 410-455-2404

- **HPCF is the community-based, interdisciplinary core facility for scientific computing and research on parallel algorithms at UMBC.**
- The current machine in HPCF is the distributed-memory cluster **taki** with over 200 nodes. **Taki** comprises three clusters: (1) The CPU cluster contains 42 compute nodes with 18-core Intel Skylake CPUs and 384 GB memory per node, connected by a EDR InfiniBand interconnect; 49 compute nodes with 8-core Intel Ivy Bridge CPUs and 64 GB memory per node and 82 compute nodes with 4-core Intel Nehalem CPUs and 24 GB memory per node, both connected by a QDR InfiniBand interconnect. (2) The GPU cluster contains 18 CPU/GPU nodes with two NVIDIA K20 GPUs and 1 node with four NVIDIA Tesla V100 GPUs connected by NVLink. (3) The Big Data cluster contains 8 nodes with 18-core Intel Skylake CPUs, 384 GB memory, and 48 TB disk space per node, connected by a 10 Gb/s Ethernet network. All nodes are connected via InfiniBand to a central storage of more than 750 TB.
- The HPCF initiative originated in 2008 with an MRI proposal to the National Science Foundation (NSF) by over 20 faculty from more than 10 departments and research centers from all three colleges at UMBC. It built on the experience with a partnership between the Department of Mathematics and Statistics and the Division of Information Technology (DoIT) in jointly operating the 33-node cluster **kali** (purchased in 2003 with funds from an NSF SCREMS grant and UMBC) that had also been used by researchers from several other departments, notably in the College of Natural and Mathematical Sciences. The initial machine **hpc** in HPCF with 35 nodes was funded jointly by several participating faculty, the administration, and DoIT in 2008 and extended the partnership with DoIT to the entire UMBC community. Its complete replacement and significant expansion to the 86-node **tara** in 2009 was supported by NSF grants from the MRI and SCREMS programs, plus funding from individual researchers. The dramatic expansion to **maya** with over 300 nodes in 2013 used a second NSF MRI grant, UMBC funds, individual researchers' funds, and a gift from NASA, the first-ever significant gift of computing equipment to UMBC.

- **Unique features of HPCF:**
  - Integration with education: Math 447/627 Introduction to Parallel Computing, computational chemistry classes, computer science / information systems classes. *REU Site: Interdisciplinary Program in High Performance Computing* (hpcreu.umbc.edu) funded by NSF, NSA, and DOD as summer program 2010-2017. Resource for other grants, e.g., NSF-funded Biology-Mathematics UBM@UMBC program (ubm.umbc.edu) and NSF-funded CyberTraining (cybertraining.umbc.edu).
  - Individual user support available by two RAs funded by UMBC and via consulting through the *Center for Interdisciplinary Research and Consulting* (circ.umbc.edu).
  - HPCF Governance Committee comprised of members of the user community. Regular meetings of the user support team with DoIT staff on system management.
  - Scientific and parallel computing research across all colleges visible on campus (more than 400 users over time; over 35 faculty research groups)
  - Interdisciplinary research opportunities among departments and with research centers (JCET, CUERE, IRC, etc.)
  - Over 250 publications, including over 100 papers in peer-reviewed journals (including Nature, Science, and other top-tier journals in their fields), 25 refereed conference papers, and 30 theses.

How does HPCF stack up?

- **Research-group (or department) owned computer:**
    - Size of system: up to 10s nodes
    - Network: basic Ethernet
    - Storage: few TB (not backed up, no redundancy)
    - System administration: by users (graduate student)
    - User support: none
- **College-based central computer:**
    - Size of system: 100 nodes
    - Network: varies
    - Storage: few TB (some redundancy, maybe backed up)
    - System administration: professional (part-time)
    - User support: varies, usually none
- **HPCF (after 2018 expansion):**
    - Size of system: over 200 nodes, including 20 with GPUs and 8 Big Data nodes
    - Network: high performance (low latency, wide bandwidth) InfiniBand (QDR or better)
    - Storage: over 750 TB with redundancy, including 10 TB with backup
    - System administration: professional (redundant full-time)
    - User support: by local ticket and in-person locally available via consulting approach
- **National supercomputer centers (e.g., NSF funded XSEDE members):**
    - Size of system: thousands of nodes, hundreds with GPUs
    - Network: high performance nowadays
    - Storage: significant, at least 100s TB (redundancy, maybe mirrored, no full backup)
    - System administration: professional (several full-time)
    - User support: professional and/or student-worker remotely by e-mail or forum only
- **National labs (e.g., Department of Defense, Department of Energy):**
    - Size of system: thousands of nodes or specialized system
    - Network: high performance
    - Storage: very significant
    - System administration: professional (multiple full-time)
    - User support: irrelevant, since system not open to public