**LEADERSHIP-CLASS COMPUTING FACILITY (LCCF)** **$93,000,000**

**Appropriated and Requested MREFC Funds for the Leadership-Class Computing Facility**[1]

(Dollars in Millions)

| FY 2024 Request | FY 2025 Estimate | FY 2026 Estimate | FY 2027 Estimate |
|---|---|---|---|
| $93.00 | $247.00 | $147.00 | $33.00 |

[1] Internal NSF cost analysis indicates the Total Project Cost would range from $520.0 million to $620.0 million depending on the acquisition strategy used for the data center. The acquisition strategy will be determined during the Final Design Phase and prior to award, if authorized. Future budget requests will be modified accordingly.

## Brief Description

Computer simulation, together with artificial intelligence (AI) methods and data analytics, is critical to enabling transformational science and engineering (S&E) research. From understanding the origin and evolution of our universe to exploring atomic-scale biomolecular processes, computational methods are now an integral part of almost all curiosity-driven, use-inspired, and translational S&E research. The LCCF, led by the Texas Advanced Computing Center (TACC) at the University of Texas at Austin, is envisioned as a distributed facility that will provide unique computational and data analytics capabilities, as well as critical software and services, for the nation's S&E research community to enable discoveries that would not be possible otherwise. Furthermore, the project will deploy a comprehensive range of education and outreach activities that will expand and nurture our nation's future S&E workforce in data and computational science.

The LCCF will integrate multiple systems located at geographically distributed data processing and discovery sites. The LCCF processing capabilities will be anchored by a computing system called *Horizon,* which will substantially increase the computational and data analytics capacity of the current NSF leadership-class computing system, *Frontera*. The LCCF will include four Distributed Science Centers (DSCs), which will be located at the Atlanta University Center Consortium (AUCC), the National Center for Supercomputing Applications (NCSA) located at the University of Illinois Urbana-Champaign, the Pittsburgh Supercomputing Center (PSC), and the San Diego Supercomputer Center (SDSC). These DSCs are designed to leverage expertise across the broader cyberinfrastructure (CI) ecosystem and provide critical edge computing services to the LCCF user community. The partnership with AUCC will enable engagement with four Historically Black Colleges and Universities (HBCUs) to provide workforce pathways for HBCU students into leadership computing and computational data science. The partnership with NCSA will explore and provide new processor technologies for Artificial Intelligence (AI). The partnership with PSC will provide data intensive computing and data mirrors for published archives. Finally, the partnership with SDSC will provide testbeds focused on supporting Machine Learning (ML) and instrument data analytics in scientific workflows, and methods to democratize access to LCCF.

In addition to supporting large-scale simulations, the LCCF will also support new usage modes such as interactive computing required by scientific tasks that involve human-in-the-loop processing, as well as urgent computing for emergency response scenarios that will need immediate access to

computing resources and real-time data. The LCCF will also support the full scientific data lifecycle, which is critical to the modern S&E discovery processes and to unleashing the potential of rapid advancements in Machine Learning and AI. The extensive LCCF Education and Public Outreach (EPO) effort has the goals of expanding the learning and workforce pipeline and broadening participation in research computing, especially from underrepresented groups, to advance Diversity, Equity, and Inclusion. A key component of the LCCF EPO activities will be a Visitor Center that will include virtual and physical exhibits, along with multiple pilot activities that will be designed to engage students, teachers, and the broader public across the nation.

**Baseline History**

In Fiscal Year 2017, NSF released solicitation NSF 17-558 (*Towards a Leadership-Class Computing Facility - Phase 1*)[1] to support the acquisition of a Phase 1 leadership-class computing system and to initiate the planning process for a future Phase 2 leadership-class computing facility. After a rigorous merit review of the submitted proposals, NSF requested that the National Science Board (NSB) authorize an award to TACC. In July 2018, NSB authorized the acquisition of the *Frontera* system, the first acquisition in the two-phased process, and funding to advance the planning of a Phase 2 LCCF. As noted in NSF 17-558, and in response to the recommendations set forth in the report *Future Directions for NSF Advanced Computing Infrastructure to Support U.S. Science and Engineering in 2017-2020*[2], the goal of the LCCF Phase 2 planning is the design of a major facility that will support all S&E research that requires the largest and most computationally intensive capabilities to enable discoveries that would otherwise not be possible.

The Major Facilities Design Stage, as defined in the NSF Research Infrastructure Guide[3], consists of three phases – Conceptual Design, Preliminary Design, and Final Design. In July 2019, the LCCF began the Conceptual Design Phase following approval by the NSF Director to formally enter the Design Stage. The Conceptual Design Review (CDR) was conducted in FY 2020 by an external panel of experts to evaluate the proposed site-independent design of the LCCF. A successful outcome from the CDR resulted in the project entering the Preliminary Design Phase in September 2020.

Following extensive Preliminary Design Phase planning activities, the project successfully completed the Preliminary Design Review (PDR) in January 2022. At PDR, the project was judged by an external panel of experts to have made appropriate progress in clearly articulating a bottom-up cost estimate and a near-final project definition, as well as sufficiently mature risk analysis to allow determination of the risk-adjusted Total Project Cost (TPC) and construction duration for establishing a budget request to Congress. A successful outcome from the PDR resulted in the project entering the Final Design Phase in August 2022.

[1] *NSF 17-558, Towards a Leadership-Class Computing Facility - Phase 1*, www.nsf.gov/pubs/2017/nsf17558/nsf17558.htm

[2] Future Directions for NSF Advanced Computing Infrastructure to Support U.S. Science and Engineering in 2017-2020, www.nationalacademies.org/our-work/future-directions-for-nsf-advanced-computing-infrastructure-to-support-us-science-in-2017-2020

[3] NSF Research Infrastructure Guide, www.nsf.gov/pubs/2021/nsf21107/nsf21107.pdf

**Project Status**

The LCCF project is currently in the Final Design Phase. In this phase, the project will finalize the details in the project execution plan to ensure it is ready to enter the Construction Stage. The project's current activities include refining the technical requirements, completing the implementation of the Project Management Control System, resolving negotiations for all external contracts, and incorporating any events, conditions, or risks previously unforeseen at the PDR into the final plan.

**Meeting Intellectual Community Needs**

The scientific requirements for the LCCF are defined by the needs of the S&E community. These requirements include the need to expand computation and data analytics capabilities, as well as the size and diversity of the workforce capable of using these capabilities. LCCF requirements also include the need to be agile and adapt to rapidly changing technology, new application formulations, and requirements, as well as new paradigms of computation-based research. Furthermore, the LCCF will need to enhance the broader high-performance computing ecosystem by providing unique resources and expertise at our nation's Institutions of Higher Education that complement other investments in the federal government and in industry.

The LCCF scientific requirements were determined through a broad set of engagements with the S&E community. These included: distillation from a wide variety of scientific community reports; analysis of usage patterns on leadership computing systems across government; feedback from project-organized planning meetings with the research community; discussions under non-disclosure agreements with industry vendors concerning technology roadmaps; technical evaluation of advanced system prototypes; and direct conversations and interviews with scientists and engineers.

During the Preliminary Design Phase, the LCCF augmented its science requirements gathering activities by selecting twenty-one Characteristic Science Application (CSA) teams as partners with the project. The CSA partners represent applications across a broad range of S&E domains and were selected to enable the development of a suite of science drivers that will be used to verify and validate the facility's effectiveness. The CSA partners will also provide requirements to inform design decisions and enable the project to acquire experience and expertise in coding/performance-tuning enhancements in preparation for facility operations.

**Governance Structure and Partnerships**

<u>NSF Governance Structure</u>
NSF oversight of LCCF is provided by a Program Officer in the CISE Office of Advanced Cyberinfrastructure (OAC), working cooperatively with other NSF staff through an Integrated Project Team (IPT). The LCCF IPT includes members from OAC leadership, Large Facilities Office (LFO), Office of the General Counsel (OGC), Budget Division, Division of Acquisition and Cooperative Support, and Office of the Director. In addition, LFO staff provides advice to OAC and assists with agency oversight and assurance. CISE leadership and NSF's Chief Officer for Research Facilities provide additional high-level guidance and oversight to the project.

<u>External Governance Structure</u>
The LCCF Senior Management Team consists of the LCCF Project Director (PD), the Project Manager

(PM), and the Deputy Project Manager (DPM). This group is responsible for the day-to-day management of the LCCF project. The PD serves as the primary interface for the project with external oversight from NSF. The Project Management office is headed by the PM and DPM, who report to the PD. The PM has line responsibility for operations, reporting, and process management within the project, with support from the DPM. The PD is advised by two groups: the Technology Advisory Board and the Science Advisory Board. Each board consists of members who are leaders in their fields and represent cyberinfrastructure providers and experts (the Technology board) and stakeholders from the community of scientists who will be the ultimate customers of the LCCF (the Science board).

Partnerships and Other Funding Sources

The LCCF project includes partnerships with over twenty-six academic institutions that will contribute to the design, validation, and eventual operations of the LCCF. Academic partners include various institution types, including Minority Serving Institutions and Historically Black Colleges and Universities. The project also has extensive industry partnerships in various high-tech economic sectors, such as semiconductor design, computing hardware, software, and data center colocation.

## Cost and Schedule

Pending NSF and NSB authorization for a Construction Stage award after successful completion of a future Final Design Review (FDR), and appropriation of sufficient funds by Congress, LCCF construction will begin in FY 2024. The schedule proposed at PDR has a duration of 34 months, providing facility acceptance in FY 2027.

**Total Funding Requirements for LCCF**

(Dollars in Millions)

| | Prior Years | FY 2022 Actual | FY 2023 Estimate | FY 2024 Request | ESTIMATES[1] FY 2025 | FY 2026 | FY 2027 | FY 2028 | FY 2029 |
|---|---|---|---|---|---|---|---|---|---|
| *R&RA:* | | | | | | | | | |
| Development & Design | $12.50 | $3.50 | - | - | - | - | - | - | - |
| Operations & Maintenance | | - | - | - | | | 40.00 | 40.00 | 40.00 |
| Subtotal, R&RA | $12.50 | $3.50 | - | - | - | - | $40.00 | $40.00 | $40.00 |
| *MREFC: Implementation*[2] | - | - | - | 93.00 | 247.00 | 147.00 | 33.00 | | |
| Subtotal, MREFC | - | - | - | $93.00 | $247.00 | $147.00 | $33.00 | - | - |
| **TOTAL REQUIREMENTS** | **$12.50** | **$3.50** | **-** | **$93.00** | **$247.00** | **$147.00** | **$73.00** | **$40.00** | **$40.00** |

[1] Outyear numbers are for planning purposes only.

[2] Internal NSF cost analysis indicates the Total Project Cost would range from $520.0 million to $620.0 million depending on the acquisition strategy used for the data center. The acquisition strategy will be determined during the Final Design Phase and prior to award, if authorized. Future budget requests will be modified accordingly.

## Future Operations Costs

Contingent on the availability of funds, NSF plans to support LCCF operations and maintenance for an initial five years with a possibility of renewal for another five years. The current estimate is that LCCF operations will cost $40.0 million annually. NSF support for LCCF beyond 10 years, and possible re-competition of the operations and maintenance award, will be informed by a planned future study by the National Academies of Sciences, Engineering, and Medicine or other similar reports by community-based bodies and will follow internal NSF procedures for renewal, competition, and disposition.

**Reviews**

Technical Reviews

The LCCF project has been technically reviewed multiple times during the Design Stage by external committees of experts. These include stage gate reviews at CDR and PDR in FYs 2020 and 2022, respectively. The LCCF CSA program was also reviewed by a panel of external experts in FY 2021. The LCCF FDR is planned for Spring 2023.

Management, Cost, and Schedule Reviews

In accordance with NSF's No Cost Overrun Policy, the agency has developed a risk adjusted TPC estimate post-PDR (based on known risks) to inform the budget request to Congress. Using the TACC proposal as the basis for the cost analysis, BFA employed a series of independent cost estimate reviews to inform NSF's post-PDR cost analysis. These included an independent cost estimate in accordance with U.S. Government Accountability Office good practices, conducted by an external contractor for appropriate project components, BFA's internal Cost Analysis and Pre-award Branch for other elements of cost, and LFO for assessments of budget contingency and schedule. This information was reconciled against the original proposed budget from TACC in determining the recommended value for potential budget inclusion in a future budget request.

**Risks**

Technical

The LCCF project deploys state-of-the-art technologies and services for the Nation's S&E research community. Several risks related to the readiness and reliability of the future deployed systems and services are included in the LCCF risk register. The LCCF CSA partnerships will be key in monitoring and managing these risks as the CSA teams will be the early users of the facility and will help identify and remedy issues as they arise. Cost and schedule contingencies have also been factored into the plan to mitigate these risks. In addition, a scope management plan provides options to mitigate cost or schedule overruns by eliminating project scope if necessary. LCCF's scope management plan is arranged so that any decision to descope can be made as late as possible in the project and with minimal impact on the science goals.

Site

The LCCF project proposes to use a colocation datacenter operated by Switch 10 miles from the TACC campus as the preferred hosting option for the core *Horizon* system. Alternatively, a new datacenter would be constructed on the TACC campus. Risks and opportunities related to both site options are being analyzed for impact during the Construction and Operations Stages of the project. A final decision on the site location will be determined at FDR.

Environmental Health and Safety

The project includes several health and safety-related risks in its risk register. These risks include weather and health-related events that may have a system-wide impact on the LCCF construction schedule. Contingencies have been factored into the construction schedule to mitigate these risks if necessary. Furthermore, an assessment for compliance with National Environmental Policy Act regulations has been completed by NSF OGC.

Partnership Risk

The project has established partnerships with several technology vendors. Risks of withdrawal by partners due to unforeseen business reasons are included in the risk register and are carefully monitored and managed by the project. Project assessment of these partnership risks is that the probability of their occurrence is very low. However, mitigation plans have been developed.

System Integration Risk

An experienced technology integrator for the core *Horizon* system has been identified, and overall system integration risk is low. However, several other risks related to delays in datacenter access and technology roadmap change have been identified by the project that could impact the system integration timeline and schedule. Cost and schedule contingencies have been factored into the plan to mitigate these risks.