NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
UNITED STATES DEPARTMENT OF COMMERCE

High Performance Computing

Strategic Plan for FY2011-2015

October 16, 2008

Page Left Intentionally Blank

Table of Contents

Executive Summary

NOAA is recognized as a world leader in understanding and predicting the Earth's environment, through advanced modeling capabilities, climate research and real time weather products. The threat and growing concern of global climate change and its impact on national security, hurricanes, and other natural disasters has spurred growing public demand for climate and weather information with increased accuracy, shorter lead times and local detail of model simulations. "Members of the 110th Congress are introducing legislation related to global climate change at a faster pace than any previous Congress," according to the Pew Center on Global Climate Change.

To meet this demand, NOAA scientists require leadership-class, high performance computing (HPC) systems with petaflops-scale capabilities (a petaflop is one million-billion operations per second). NOAA estimates that this performance requirement is similar to that of the top 10 supercomputing environments in the world. However, NOAA's high performance computing capabilities have fallen to more than an order of magnitude behind other U.S. agencies with leadership-class systems, such as the Department of Energy and the National Science Foundation, who also work on aspects of climate and weather research.

Requirements

NOAA's HPC program must provide a broad continuum of information products based on computer models for weather, climate, and ecological predictions. NOAA currently has dozens of major models and hundreds of variants of models under development that support the full range of forecasting challenges. These models are maintained in the operational phase, producing millions of informational products daily for public use as well as for special needs, such as emergency management. Each model fills a niche in NOAA's mission and has unique requirements for input data, computational resources, run-time variability, and product delivery time.

New high performance computing hardware architectures require scientific applications to run across multiple processors, rather than a single processor, to achieve desired performance. Improvements in modeling techniques have led to environmental models that can utilize many thousands of computer processors, rather than a few hundred, which promises to dramatically increase both the accuracy and speed of environmental predictions. Test runs on some of the world's largest computers currently available have demonstrated that NOAA is ready to use these large scale systems.

In order to accommodate this new large-scale supercomputing approach to modeling, NOAA has determined that it requires a new, more flexible HPC architecture. This target architecture must span a vast array of technical and product delivery requirements to meet the needs of millions of diverse stakeholders, including NOAA scientists, academic researchers, private sector planners, Federal partners, and the general public, particularly when life and property are threatened.

Specifically, the future HPCC environment must meet the following requirements:

- ***Availability:*** Clock-driven (i.e., several times daily) product generation requires redundant, highly available (99.9%) HPC systems, networks, and data center facilities; calendar-driven products such as decadal climate models and environmental research have lesser availability requirements (96%) at a much lower unit cost
- ***Large-scale computing:*** Able to run massively parallel codes, multi-model ensembles, order of magnitude resolution increases, and dramatic increases in new observational data assimilation
- ***Life cycle management:*** Maximize the efficiency of the transition of research to operations and the return to research of operational improvements
- ***Software engineering:*** Adopt software engineering standards and practices that accommodate sharing of code internally and externally with other federal agencies and academic partners
- ***Software engineering staff:*** Streamline and optimize code for performance and enable adoption of new tools and technologies
- ***IT security:*** Advanced IT security to protect and support NOAA's computing from misuse and protect the integrity of NOAA's information products

The target HPC system architecture must also have the ability to support the rapid reprioritization of computing resources to address short-term, national needs for environmental security. In times of national emergency, such as major hurricanes or widespread wildfires, critical modeling efforts must be able to surge onto additional computational resources within NOAA's architecture to deliver results in near real-time. In addition, a longer term goal is to establish and coordinate an on-demand computing infrastructure with other Federal agencies to quickly respond to unforeseen events of national significance, such as the spread of a toxic plume. These capabilities will constitute an HPC *Environmental Security Architecture.*

Environmental Security Architecture

To meet these complex business and technical requirements, NOAA has designed a target HPC architecture for FY 2011-2015, comprised of four supercomputing subsystems:

1. A Operations System for the daily production of forecast model products (e.g., daily national weather forecast)
2. A backup identical to the Operations System that also supports testing and final integration of models into operations
3. A system of similar architecture to the Operations System, sized to support the research and development of applications bound for operations
4. A large-scale (preferably petaflops-scale) computer to support research and development (R&D) for environmental security, including long term climate and ecological information products delivery

To deliver the type and amount of information products needed and requested by NOAA's customers and stakeholders, the largest possible (petaflop-scale) computing is required. This would require substantial, sustained annual investments through 2015 to expand the scale of the

Environmental Security Architecture to accommodate high-resolution environmental models such as regional (multiple U.S. states) climate change prediction models. Thus all four of the supercomputer subsystems in this architecture would grow in size, but the fourth system would be expanded to tens of thousands of processors working together.

New world-class facilities will be required to house these computers, which would greatly exceed NOAA's present data center capacities. One option is the DoE Oak Ridge National Lab data center, which currently houses one of the world's top five computers. This type of data center is rare, due to the extremely high electrical power requirements. In addition, the data center must provide access to the fastest possible national research networks, including Internet2 or National Lambda Rail, to interconnect with NOAA's research infrastructure.

Software engineering capabilities will be another key component to the long term success of this new HPC architecture. NOAA will facilitate the development of unified modeling standards for NOAA's Environmental Modeling Program. More uniform utilization of internal software engineering standards and guidelines will enhance collaboration. Codes will be more robust, portable, and compiler-independent, which means an easier transition from research to operations and back to research. Additionally, effective use of standards will enable NOAA's codes to be more easily utilized by other federal agencies, universities, and the general public.

Lastly, acquiring these computational and facilities resources requires expert systems integration to manage all of NOAA's high performance computing, to meet the diverse technical and operational requirements of the sub-systems, and to minimize overall cost to NOAA.

Conclusion

NOAA must make substantive changes to its high performance computing capabilities in the next five years to meet the rapidly growing demand for its environmental modeling products. Recent developments in environmental modeling approaches offer a demonstrated path to enormous performance improvements given sufficient HPC resources, and require a different system architecture to achieve those improvements. NOAA requires significant and sustained investments to establish and maintain the target HPC architecture. This Environmental Security Architecture will have a profound effect on NOAA's ability to service the nation's needs for weather, climate, and ecological information and predictions.

Without substantially increasing the annual investments in HPC resources, NOAA's ability to address the growing national priorities will be limited. An ideal window of opportunity exists to remedy this situation during the FY 2011 to 2015 period, given the expiration of several long-term HPC contracts in 2011.

High Performance Computing Current State

NOAA HPC exists within the Environmental Modeling Program (EMP) to support the EMP outcomes and NOAA's mission outcomes. NOAA HPC assists the Environmental Modeling Program in delivering trusted, timely, accurate environmental assessments and predictions through next-generation models that are integrated, interoperable, mission driven and accessible.

NOAA HPC supports the following NOAA mission outcomes:

- Weather and Water: Improved severe weather warnings by providing longer lead times and reducing over warning through interdisciplinary modeling, and ability to expand scope of predictions
- Climate: Analysis, Understanding, and Prediction of the Earth System from weeks to centuries to support informed and reasoned decisions.
- Ecosystem: New Research and Development (R&D) architecture will make extensibility to Coastal and Ocean ecosystem modeling feasible.
- Commerce and Transportation: Extremely high resolution weather and coastal models critical to aviation, roadway transportation, and navigation.

To support the mission and functional outcomes NOAA HPC has a process to manage mission and functional requirements flowing from these outcomes. This process is linked to the supporting technical requirements which are coordinated to produce consistent actions that achieve NOAA's mission efficiently and effectively.

NOAA uses HPC for operational weather forecasts, for climate prediction and for the research that supports those forecasts and predictions. Operational activities are divided into those products that are produced on an hourly clock-driven cycle and those that are produced on a periodic calendar-driven cycle. The current total investment made each year by NOAA for HPC is $48M for both Operations and R&D.

On the weather operations side, NOAA has primary and backup computers in Gaithersburg, MD and Fairmont, WV. The primary computer runs the models from which the weather forecasts are made. Over 14.8 million products are created and disseminated each day to government agencies, commercial interests, and to the public. Since this is critical to the nation, NOAA also maintains a backup computer which can be quickly brought into production in the event of a failure of the operational system. When not used for operations, that backup computer is used for developing the next generation of forecast models to be run operationally including tuning them for maximum efficiency and ensuring that they provide forecast improvements. It serves as a critical link in the chain from research to development to operational weather forecasts.

NOAA also operates three systems for research and development in support of NOAA's mission. The largest of these, at NOAA's Geophysical Fluid Dynamics Laboratory (GFDL) on the campus of Princeton University, is primarily used for climate modeling. On this system the models for climate change are developed and runs of various climate scenarios are made for the Climate Change Science Program (CCSP). This system is critical to understanding the impacts of climate change and developing a national response. The results from this system were used by a large number of NOAA scientists and their collaborators in support of the Intergovernmental Panel on Climate Change (IPCC), and they shared in the Nobel Peace prize as a result. By some measures, the IPCC climate model developed and run at GFDL is the best in the world. GFDL also contributes research towards improving hurricane track and intensity prediction by coupling the hurricane prediction model with a full ocean model.

The research and development computer in Gaithersburg, MD is used for weather model research, exploring advanced methods of using satellite data in weather models, and developing improved climate forecasts on seasonal through inter-annual scales, such as El Niño forecasts. One third of the computational time on this supercomputer is devoted to the Climate Test Bed, improving the transition of short term climate forecast research to production.

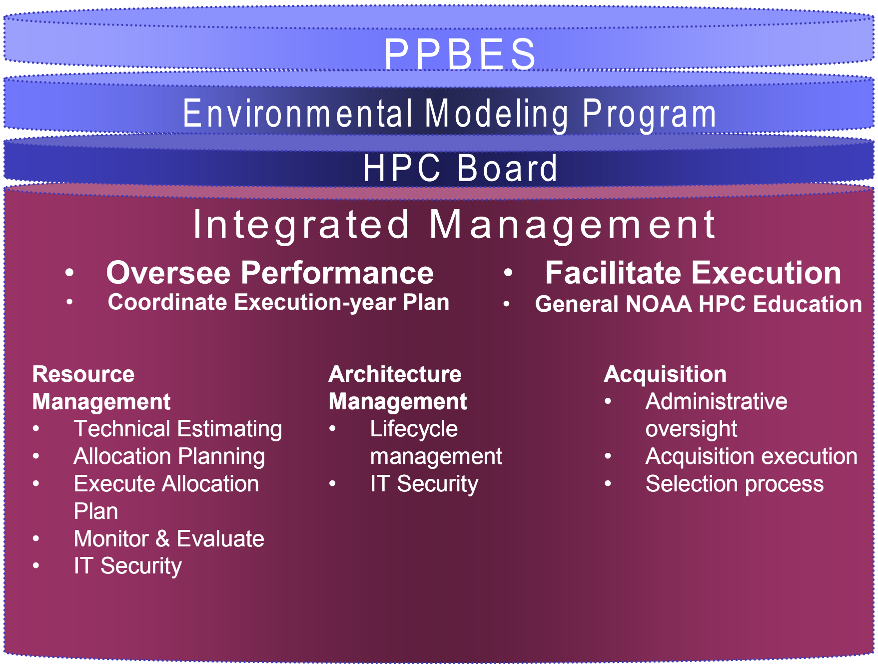
The research and development computer at NOAA's Earth System Research Laboratory in Boulder, Colorado serves many users. It serves as a portal for university scientists to work with NOAA in weather model development as part of the joint NOAA-NCAR (NSF-sponsored National Center for Atmospheric Research) Developmental Test bed Center (DTC). This resource helps NOAA leverage the investments of NSF in atmospheric research at universities around the country to improve the nation's weather forecasts.

Among the problems being addressed are improvements in flight-level forecasts for commercial aviation in support of the next generation air traffic control system being developed by the FAA. The system also is used to develop improved air quality models and to develop models that incorporate the changing chemistry and radiative properties of the atmosphere in severe smog events into weather models. Exploratory work is also going on in very high resolution hurricane models to lay the foundation for further improvements, not just in track forecasts, but also in the rapid intensification that can occur shortly before landfall increasing the destructive power of the hurricane.

NOAA also partners with other Federal Agencies to leverage the use of their HPC resources. In particular, collaborations with NASA, NSF, and DoE have been very beneficial to NOAA. NOAA has worked closely with NASA in developing an atmospheric model, capable of scaling to very high processor counts. This model was tested on one of the largest DoE computers and successfully scaled to over 8000 processors. NOAA is using the NSF-sponsored Ranger system at the Texas Advanced Computing Center (TACC) to support development of the next generation hurricane modeling system.

NOAA collaborates with other federal agencies in developing models that can be used and reused for the needs of the individual agency's missions. NOAA employs the use of community developed models to improve collaboration to accelerate the transition of these enhancements to operations. The recent emergence of software frameworks, such as the Earth System Modeling Framework (ESMF), has raised the possibility of having a NOAA-wide, and perhaps Nation-wide, common software architecture for complex numerical forecast systems.

The Environmental Modeling Program (EMP) is the central manager for all of NOAA's environmental modeling activities, and as such, is the main customer of the HPC resources. The EMP, in conjunction with the NOAA CIO, is responsible for implementing NOAA Administrative Order (NAO) 216-110. As directed by NAO 216-110, the EMP is responsible for managing NOAA's HPC through an HPC Board to meet NOAA's goals and sub goals. The HPC Board oversees the Integrated Management team, who are responsible for managing cost, schedule and performance on a day to day basis.



The Environmental Modeling Program (EMP) has utilized the PPBES process to ensure that the Mission goals identified in this plan are aligned with NOAA's strategic vision. PPBES is an extensive, iterative process that includes planning, detailed feasibility analysis and budgeting. Outcomes will be measured against corporate performance measures and execution metrics to assess the goals and outcomes relative to established targets.

Case for Change

The Environmental Modeling Program has set forth the following goals as top priorities for NOAA's HPC Program:

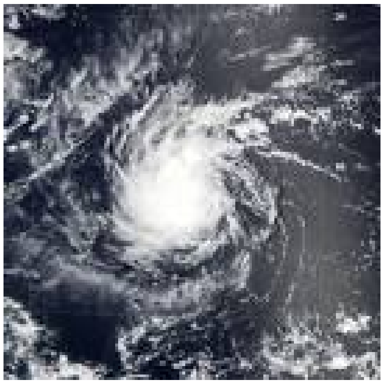
- Improve Hurricane, Tropical Cyclone, and Tropical Forecasts
- Understand and Predict Climate
- Improve and Deliver Timely Severe Weather Forecasting
- Accurate, Real-Time Weather information for Navigation
- Improve Air Quality Forecasting
- Ecological Forecasting
- Support Observing System Analysis

High Performance Computing Requirements:

Improve Hurricane, Tropical Cyclone, and Tropical Storm Forecasts

Mission Driver:

Stakeholders need an accurate and timely hurricane, tropical cyclone and tropical storm monitoring and forecast system that provides the necessary lead-time and storm intensity to make decisions regarding coastal evacuations and mobilization of resources. To meet this need NOAA's hurricane forecast system must reduce the error in 48-hour intensity forecasts for hurricane-strength storms by at least 10 kt (approximately one half of a Saffir-Simpson category) by 2011, with an emphasis on improved forecasting of rapid intensification and decay and re-intensification cycles.



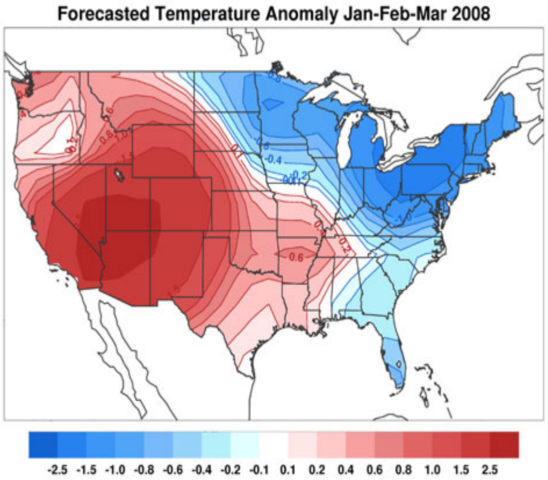
HPC Requirement:

To meet the needs of EMP, NOAA's HPCC program must provide the technical resources needed to increase the resolution of the hurricane model to 1km. Today's 9km hurricane model runs for approximately one hour across 80 processors. To attain the computational resources necessary to achieve this goal can be met by a sustained investment in HPC providing the model 100x more dedicated processors over a period 5x longer than currently available. With Moore's Law improvements in HPC, and system engineering enhancements to the model, in addition to the sustained investment, the 6000x additional computing needed to increase the model resolution can be achieved allowing a 1km hurricane modeling capability by 2015.

Understand and Predict Climate

Mission Driver:

Federal, regional, state, and local decision makers need credible climate information at finer scales to support strategies to mitigate and adapt to climate variability and change, including long-term resource management practices and public infrastructure decisions.



Forecasted Temperature Anomaly Jan-Feb-Mar 2008

The Nation's scientific community needs a dense and reliable network of climate observations, data records, models (with related computational and data storage capabilities), and analytical tools that continue to advance the understanding of climate change and its potential impacts.

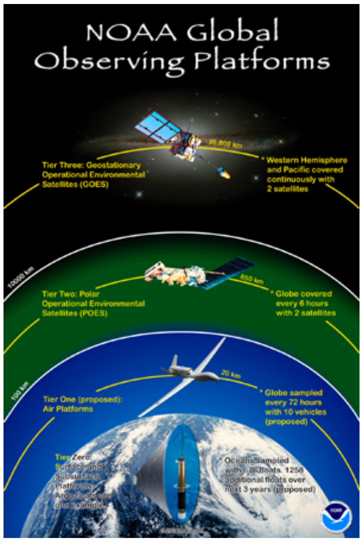
HPC Requirement:

Each doubling of a model's horizontal resolution requires a factor of 8x increase in computational power. Regional climate change information requires a quadrupling of the horizontal resolution, from 200kmin NOAA's 2004 climate models to 50 km, and twice as many vertical levels to resolve the stratosphere. **These models exist now but NOAA lacks the computing capability to verify, validate, and apply them.** To meet the needs of EMP for regional climate information, NOAA's HPCC program must provide resources **128x** over the current computational resources.

Additionally, the requirements for more comprehensive Earth System models that constrain uncertainty need an increase in computational power of at least a factor of **2x** for an interactive carbon cycle and **3x** for improved representation of aerosol and clouds in the models. Together these requirements need a factor of about **800x** overall improvement to computational capability just to get to reliable climate information at the regional scale.

Support for Observing System Analysis and Design

Mission Goals:



To conduct almost every facet of its operations, NOAA requires a capable and reliable observations infrastructure. The development of NOAA's observing system, its contributions to the Global Earth Observing System of Systems, the U.S. Integrated Ocean Observing System, and other satellite and in situ instruments, requires coherent decision-support tools for design, priorities, and investment considerations. Observing System Evaluations (OSE) and Observing System Simulation Experiments (OSSEs) are powerful ways to quantitatively assess the impact/value of a specified data stream, modeling methodology, or data assimilation technique on numerical environmental prediction results. OSEs/OSSEs provide guidance for optimizing NOAA's observing system architecture design, infrastructure investment prioritization, data quality, modeling, and data assimilation with respect to available technology, priorities, and resources. OSE's and OSSE's can measure the effectiveness of various configurations of the observing system, as well as assessing instrument/data quality, towards closing NOAA mission requirement gaps. OSEs evaluate the relative value of existing observation data streams, while OSSEs provide analogous insight for hypothetical/planned observing systems. Both provide assessments within the context of operational modeling and predictions, supporting prioritization of observing system investments and optimization of existing/future capabilities.

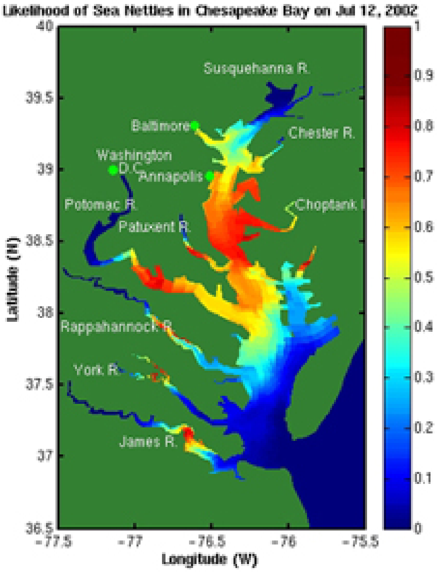
HPC Requirements:

OSEs/OSSEs employ sophisticated numerical models of governing dynamical, radiative, and some ecosystem processes, requiring notable HPC resources. Establishing an OSE/OSSE core capability requires HPC resources for access to a simulated set of environmental states, computing "reverse-engineered" synthetic observations, and assimilating the simulated observations into a forecast model.

The development, testing, implementation, and operation of an OSE/OSSE system requires notable increases in HPC capacity, comparable to the Operations system, or even larger, in order to support multiple simultaneous studies. Currently, it is not possible to run sufficiently large, high-resolution OSSEs on the Operations systems, nor on the NOAA R&D HPC systems.

Ecological Forecasting

Mission Goals:



Federal, regional, state, and local decision makers need credible ecological information at finer scales to support strategies to: identify and respond to harmful algal blooms and hypoxic events; to monitor and predict the impacts of pathogens and point and non-point source pollution in coastal habitats; and to understand, adapt to and mitigate the ecological impacts of climate variability and change, including long-term resource management practices and public infrastructure decisions.

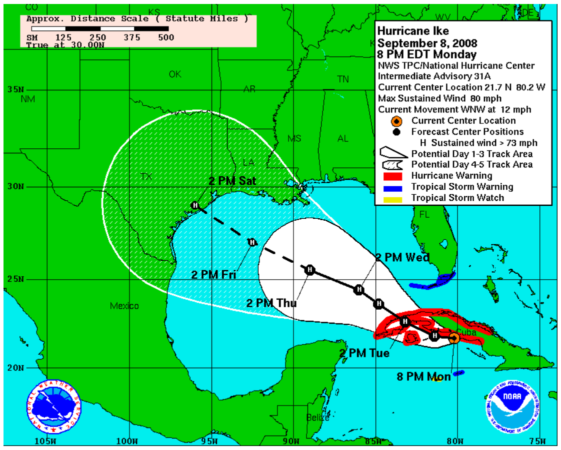
A dense and reliable network of ecological observations, data records, models (with related computational and data storage capabilities), and analytical tools are needed that will advance our understanding of our coastal marine ecosystems and enable monitoring and prediction of ecological events and change at a variety of time and space scales.

HPC Requirement:

To meet the needs of its stakeholders NOAA must develop a variety of types of ecological models, including: (1) Integrated models that directly influence physical modeling and processes in the ocean, such as sediment transport and chlorophyll (particle) models; (2) coupled models, one and two-way, with the physical modeling providing input, without or with reaction to coupled ecosystem model parameters; and (3) output user models that can be considered as downstream users of physical ocean models such as fish stock assessment models and habitat models. This variety of ecological models will require horizontal resolutions from 10's of km to 10's of meters at regional coastal domain scales in order to make useful predictions for regional response and planning.

Where we want to go

The HPC Board has developed a new target HPC system architecture to better address NOAA's needs for the foreseeable future. NOAA's existing HPC architecture limits the agency's ability to flexibly utilize computing and to address large scale environmental modeling challenges. In contrast, the new target architecture can scale to meet the future requirements defined by the EMP mission goals, changing IT technologies, and improved sharing and flexibility.



NOAA currently has dozens of major models and hundreds of variants of models under development. These models are maintained in the operational phase, producing millions of informational products daily for public use as well as for special needs like emergency managers. Operational models are generally associated with daily weather products, but models are also applied to long term ecological and climate predictions (e.g., decadal and centennial). Each model fills a niche in NOAA's mission, and has unique requirements for the ingestion of data, computational resources, run-time variability, and product delivery time.

The architecture of NOAA's HPC must meet the life-cycle requirements of these environmental models, including:

- Research into the physical and biogeochemical processes in the Earth System
- Development of numerical algorithms that describe these processes
- Integration of these algorithms into NOAA's environmental models
- Validation and verification of the improved models
- Transition into an operational capability
- Operational execution

The lifecycle phase of a model is an important architectural consideration. The R&D phase of a single model's life-cycle typically requires a larger fraction of an HPC system than during its operational execution. Research requires many runs for each case as algorithms are developed and debugged; physics chemistry packages tested and improved; parameterizations optimized and validated. In addition, the computing available to run a model will have improved through technological innovations (e.g., Moore's Law) by the time that improved model becomes operational. The further refinement and optimization of the mature code allows for more efficient processing.

The other salient feature of models that impacts the architecture is their expected product delivery cycle. For example, a more accurate hurricane model may take months to develop, but once it is operational, its product delivery cycle must be executed in mere minutes. In contrast, a centennial climate model may require years to develop, and final product delivery also occurs on the scale of years, as in the case of the periodic Assessment Reports of the IPCC.

NOAA's present architecture divides resources roughly equally between R&D and operations because the large suite of operational models for NOAA's daily forecast products drives a computing requirement that is comparable to that of current R&D. Strong demand for new long term and regional climate products, and higher resolution hurricane models is anticipated to dramatically increase the R&D component during the next several years as new models are developed. Operational products will also grow as the models with a shorter development cycle become mature.

The Operational and R&D computing systems share a number of characteristics and practices such as periodic technology refreshes, high levels of vendor support, and a robust set of security techniques and tools. However, the defining characteristics of an R&D HPC system are different from those of an Operational system. Specifically, R&D systems emphasize the capacity of the computing to explore new modeling capabilities; while an Operational HPC system must provide on-time delivery of products through highly redundant, serviceable, and stable computing systems. Consequently, the NOAA HPC Architecture separates the R&D and Operational computing systems, with the exception of shared storage. Even so, the overall Architecture must avoid barriers to the exchange of R&D and Operational results through the transition process and provide for seamless planning across the agency.

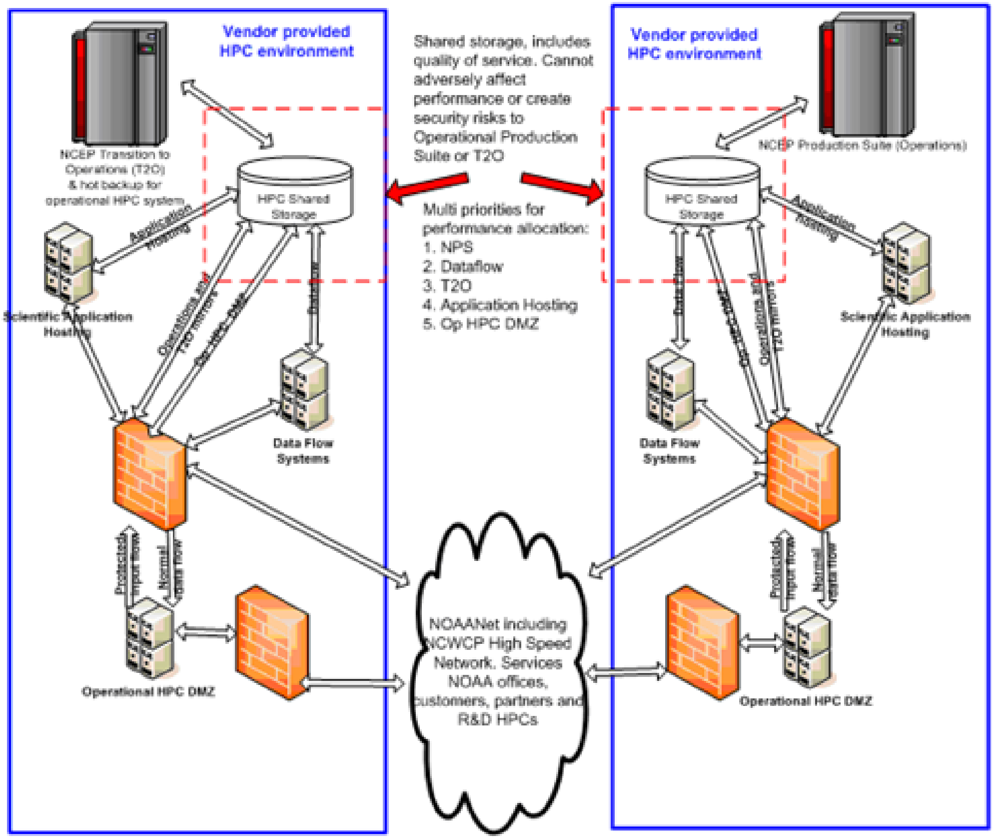
In the remainder of this section, the target architecture for NOAA HPC will be described. It is designed to embody the concepts outlined above while advancing NOAA's modeling capabilities.

Operational HPC Architecture

A critical component of the National Oceanic and Atmospheric Administration (NOAA) mission is the support of NOAA's operational high-performance computing (HPC) system. The NOAA operational HPC system creates and delivers environmental forecasts and guidance products with a high level of reliability. The customers for these products include the nation, military, academia, private enterprise, and other partners throughout the world.

The diagram below (figure 1) provides a conceptual architectural view of the operational HPC system.

Figure 1 - Conceptual view of next generation operational HPC architecture



The above diagram illustrates two HPC systems, symmetrically configured, hosted at two geographically diverse locations. This architecture has demonstrated that it can meet the stringent reliability and timeliness metrics required by NOAA.

This architecture also supports transition-to-operations (T2O) capabilities. Tasks associated with T20 include final development and testing of new modeling techniques and scientific advancement, external community review of all significant changes, and the migration of the changes into production operations. Similarly, the HPC architecture supports the flow of knowledge and data from the operations environment back into research (O2R).

In summary, the operational HPC architecture requires robust systems and processes to support critical performance capabilities that include:

- On-time delivery of forecast model products
- Hosting a diverse set of environmental models
- Transfer of new and enhanced science from research systems into operations, driving continuous improvements to operational forecast model products
- Feedback from operations back into research

The critical capabilities above require the following architectural components and practices:

- Highly redundant HPC systems, network and facilities
- Stable HPC technology, to support both reliability and scientific reproducibility
- High level of vendor support
- Robust HPC management processes, including configuration management, project management, HPC resource allocation management and user access management
- Technology transparency to facilitate software portability across modeling systems and vendor independence
- Seamless planning across all impacted organizations
- Risk reduction through robust testing
- Scheduled periodic system patching across all HPC environments
- Technology refresh every 2-3 years
- High level of security awareness and practices

NOAA has matured as a technology leader in managing operational HPC systems. This maturation has been a journey of continuous learning and improvements within both the operational and research HPC components and processes. Each architectural component has been developed, enhanced, tested, proven and integrated into an end-to-end HPC management system. This HPC management system provides the foundation for the capability to deliver operational model forecast products with scientific consistency and reliability, under any circumstances. The new NOAA HPC architecture retains this proven Operational configuration essentially unchanged, while seeking the programming and documentation necessary to work and collaborate across multiple platforms.

R&D HPC Architecture

NOAA’s R&D HPC system (R&D HPCS) provides four fundamental HPC functions:

1. Large-scale computing provides computing for development, testing, and production integrations of NOAA environmental models. The workload that runs on this subsystem is characterized by compute-intensive codes with I/O characterized by regular snapshots of diagnostic fields.
2. Analysis and interactive computing provides computing for the post-processing of data from production runs and the analysis of post-processed data, code development, and debugging. The workload that runs on this subsystem is characterized by data-intensive codes requiring high I/O bandwidth.
3. Data archiving provides long-term storage of post-processed model runs and analyses.
4. Networking links these subsystems together.

NOAA’s environmental modeling R&D uses these four functions to varying degrees.

The current configuration of the R&D HPCS is architected along organizational lines. Large-scale computing, analysis computing, and storage are located at sites in Princeton, NJ (collocated with OAR/GFDL), Boulder, CO (collocated with OAR/ESRL), and Gaithersburg, MD (collocated with NWS/NCEP’s Operational HPC system). These systems are now connected with a high-bandwidth network. This configuration reflects the historical development of HPC capability in NOAA over the last 50 years, coupled with the high cost and relative slowness of long distance networks; during this period, computing resources were acquired along organizational lines. The most recent procurement for the R&D HPCS has unified the acquisition and management of the system, yet the configuration of R&D resources has remained essentially unchanged since 1999.

Figure 2 R&D and Operational HPCS System Locations

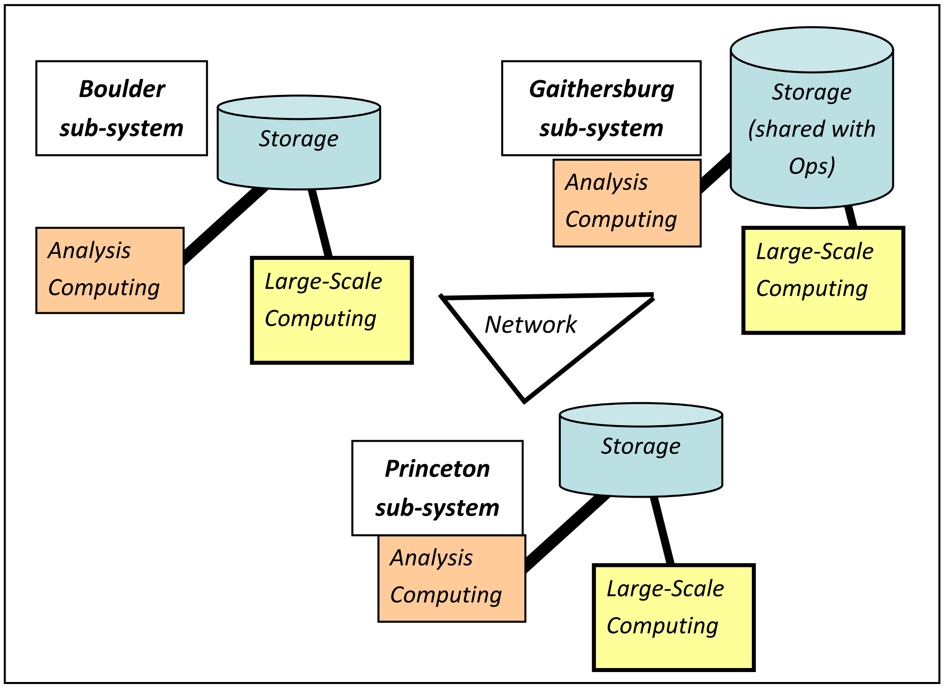


Figure 3 R&D HPCS System – Current Architecture

There are distinct advantages to the current architecture, including leveraging the experience in the organizations that use HPC to meet distinct NOAA mission goal objectives, and the development of robust software and management approaches used in the diverse set of systems NOAA now manages.

However there are also significant disadvantages to the current architecture that make the case for change. Physically separate large-scale computing systems practically limit a single job to the system size at a single site. This is inconsistent with NOAA using HPC to solve its largest computational problems (as defined in NAO 216-110). Separate large-scale computing systems provide barriers to moving NOAA's workload to available resources, thereby reducing the flexibility and utilization of these NOAA HPC systems. There is also some duplication of effort in providing support for multiple HPC systems.

The case for change leads to a new target architecture for NOAA R&D HPC that ameliorates these disadvantages yet retains the application of organizational expertise to NOAA's diverse set of computational challenges and facilitates the transition from research to Operations, and vice versa. The new target architecture consists of two large-scale computing systems with distributed analysis and storage sites local to the scientists using the output from NOAA's environmental models. All NOAA HPC systems are connected by a high-speed network.

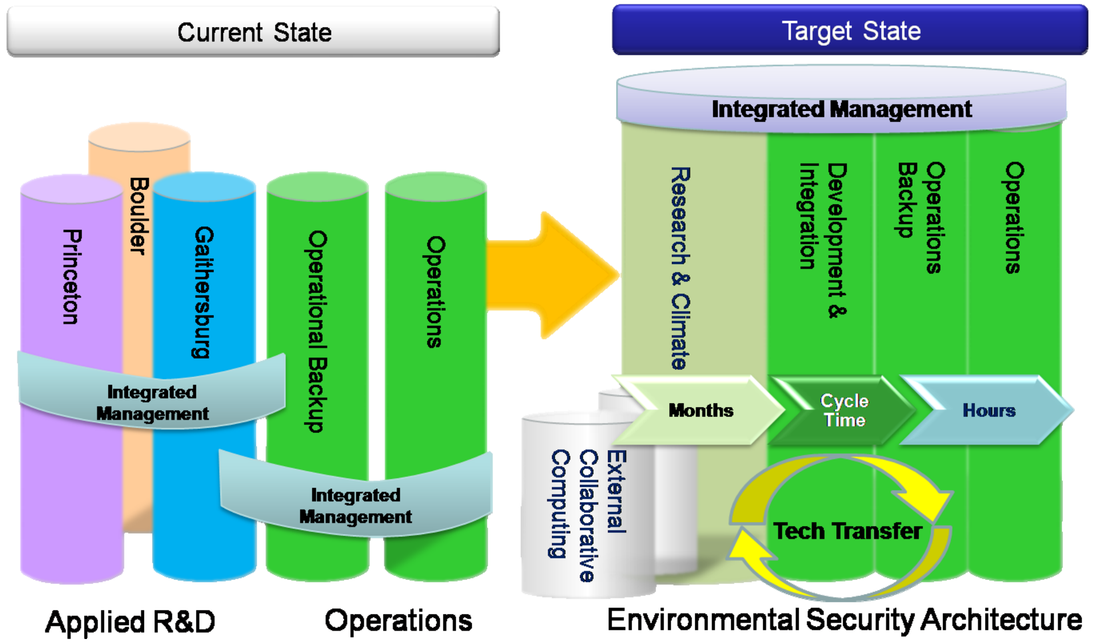


Figure 4 Target Architecture

In this new configuration, one of the large-scale R&D computing systems serves the needs of the broad NOAA research community, while the other serves as a research platform specifically to advance the modeling capabilities on the Operations system. This system is configured to be the same technology as the Operations system without incurring additional costs for the highest levels of availability. In moving from three to two large-scale computing systems, the job size will be maximized while still retaining some redundancy, NOAA realizes reduced support infrastructure, distributed analysis systems provide some redundancy in computational capability, and data and the systems that use them are near their users. Although overall wide-area networking requirements will increase, the majority of data movement will be localized to the sites hosting analysis and storage, which reduces networking requirements between the sites.

The target architecture will require additional investment in one or more facilities to house these more capable large-scale computing systems. The new architecture promotes a continuity of operations while allowing the exploration of novel, more efficient computing paradigms. Overall, this is expected accelerate advances in NOAA's environmental models.

Path to Desired State

Software

An aspect of the target architecture that impacts all NOAA modeling products is software engineering. In the coming years, the environmental modeling community will be faced with the challenge of writing efficiently coded model simulations to run across thousands of multi-core processors in programming environments that are still emerging. This challenge will require a cadre of engineering expertise to work closely with physical scientists to develop and optimize algorithms. NOAA must continue to enhance its software engineering discipline and expertise to achieve optimal code performance and scaling, maximize the efficiency of transitioning research to operations, and enable effective collaborative model development with partners both internal and external to NOAA.

To lead this effort NOAA must develop a cohesive, well trained, high performing software engineering staff. The software engineering staff will lead NOAA in developing unified modeling standards and frameworks that enable scientists to create and maintain robust, portable, and compiler-independent modeling software. These software characteristics will broaden the opportunity for partnerships and should strive to reduce the complexities of moving modeling software from research to operations and from operations to research. The engineering staff will also investigate promising novel architectures and approaches to utilizing their capabilities, which is important to position NOAA to move to new architectures when they are appropriate.

Along with standards, engineering expertise, and increased collaboration NOAA must ensure its code remains accessible to a variety of research partners. The coordinated use of software repositories within NOAA will enable researchers to share code while maintaining confidence in its integrity.

Communications

All of NOAA's HPC requires a robust and secure communications infrastructure. NOAA has been steadily building and integrating its HPC networks through the annual HPCC Program Office proposal process. These efforts have resulted in a rudimentary Gigabit backbone comprised of dark fiber local loops and Metropolitan Area Networks provided by commercial entities and that interconnect to the National Lambda Rail or Internet2 national research networks for interstate transport. Sites that currently host the R&D HPCS already have 10 Gigabit-per-second (Gb/s) capabilities, while the Operations sites have 1 or 2 Gb/s service, to be upgraded to 10 Gb/s by fall of 2008. Other major NOAA sites are gradually being connected at 1 Gb/s, including: Ann Arbor, MI; Ashville, NC; and Ford Island, HI.

Moving forward, this backbone will be strengthened by adding Layer 1 and Layer 2 services at 10 Gb/s between all the HPC sites in a cost-effective, modified star topology. Multiple routes will be added to achieve reliabilities similar to the primary computing or storage locations.

All routine, Layer 3 communications will be securely rerouted to the new Trusted Internet Connection (TIC) centers. However, Layers 1 and 2, which are inherently more secure are used for bulk data transfer, and are not routed internet connections, are exempt from the TIC and may continue to be directly interconnected between NOAA sites.

Finally, all large NOAA field offices will be connected via NOAAnet as it absorbs some of the existing infrastructure established by HPCC. Smaller field offices may be connected via secure virtual private network (VPN) hardware encryption routers. Individuals may connect via VPN software. All of these Layer 3 traffic streams will converge at the TIC sites, where they would enter the higher bandwidth HPC network.

Security

Security will continue to be a high priority for NOAA's HPC systems. The R&D system has already implemented two factor authentication (2FA) methods for most or all users at all sites. The NOAA-issued Common Access Card (CAC) will be central to the IT security strategy and fulfilling HSPD-12 compliance. Each NOAA employee or associate account holder will be issued a CAC card for physical building access. This same card holds unique identifier information (e.g., a fingerprint) that links it to a particular individual. Each card also contains an encoded PIN number selected by the user at the time of card issuance. Finally, each CAC card contains four electronic certificates for public key infrastructure (PKI) use in encryption. These allow digital signing of passwords, email, and other functions.

The HPC systems will capitalize on the CAC cards by mandating that all account holders, other than foreign nationals, log onto a NOAA network using their CAC card. This assures compliance with 2FA as the CAC represents "something you have", and the PIN number entered by the user represents "something you know."

Foreign nationals, who hold accounts due to research collaborations, such as via the Nobel Prize-winning Intergovernmental Panel on Climate Change (IPCC), will be issued a traditional security token. They will access the HPC system through a separate, more tightly controlled access point, with closer scrutiny on their account activities.

Surge Computing

NOAA is very interested in the potential of having a surge computing capability and exploring the manner in which it may be employed. Surge computing refers to the overflow of NOAA's critical Operations models onto other HPC systems in times of national emergency, such as a Category 5 hurricane, extensive wildfires generating interstate smoke plumes, or use of a weapon of mass destruction (WMD).

Surge computing would reprioritize the normal workload for the backup Operations system and the R2O system that shares the same computer architecture. This would potentially allow the use of all 3 systems for brief intervals to run larger ensembles of models, and/or multiple instances of models.

In addition to surging across its own internal HPC resources, NOAA is exploring the possible use of On Demand computing with other federal agencies. This concept extends the use of surge computing to partners outside of NOAA. Because of the complex and difficult nature of establishing and maintaining compatible architectures that could be employed on extremely short notice for episodic events, this remains a long term goal requiring protracted negotiations and cooperation among agency partners. Additional, potentially heavy, investments may be required in test and development subsystems.   A concept of operations will need to be crafted multilaterally.

Conclusion

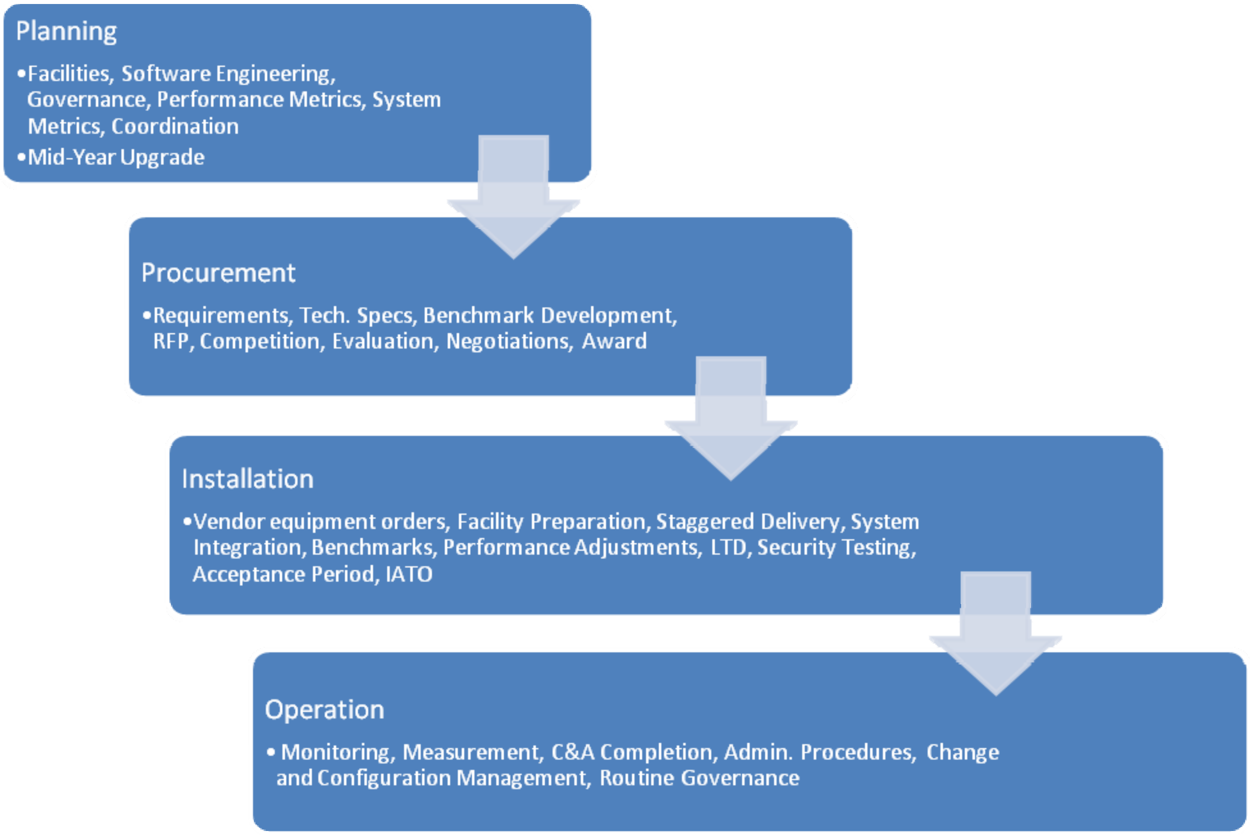
The threat and growing concern of global climate change and its impact on national security, hurricanes, and other natural disasters has spurred growing public demand for climate and weather information with increased accuracy, shorter lead times and local detail of model simulations. The public needs accurate and timely hurricane, tropical cyclone and tropical storm monitoring and forecasting that provides the necessary lead-time and storm intensity to make decisions regarding coastal evacuations and mobilization of resources. Federal, regional, state, and local decision makers need credible climate information at finer scales to support strategies to mitigate and adapt to climate variability and change, including long-term resource management practices and public infrastructure decisions.

Given the continuing growth in NOAA's HPC requirements coupled with the ever increasing complexities of supercomputing and software engineering, NOAA will have to make substantive changes in order to continue to meet mission requirements and achieve its organizational goals.

NOAA's operational HPC requirements will be met by the implementation of the target architecture which includes a primary HPC and a backup HPC connected by a robust communications and IT security infrastructure. NOAA's R&D requirements will be met by the use of a single large scale compute component, with long term storage, with an adequately sized operations-like component, and local compute and short-term storage resources. NOAA will continue to partner with other Federal Agencies to obtain additional compute cycles to help fill resource gaps. NOAA will also pursue a surge computing capability by partnering with other Agencies. A more focused and managed approach to software engineering will be pursued.

HPC requires a significant investment to acquire, operate, and manage and all present significant challenges to NOAA in moving forwards to the future. Funding will determine the pace at which the Environmental Security architecture is implemented and how soon the Nation will realize the benefits.

Appendix "A" Phased Approach



Planning

- Facilities, Software Engineering, Governance, Performance Metrics, System Metrics, Coordination
- Mid-Year Upgrade

Procurement

- Requirements, Tech. Specs, Benchmark Development, RFP, Competition, Evaluation, Negotiations, Award

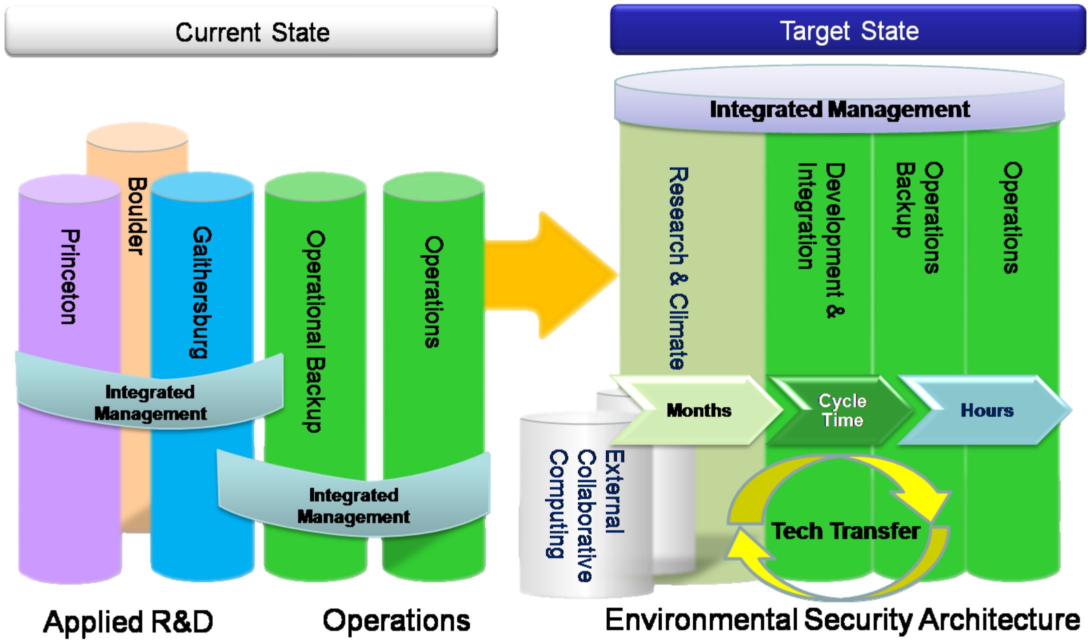
Installation

- Vendor equipment orders, Facility Preparation, Staggered Delivery, System Integration, Benchmarks, Performance Adjustments, LTD, Security Testing, Acceptance Period, IATO

Operation

- Monitoring, Measurement, C&A Completion, Admin. Procedures, Change and Configuration Management, Routine Governance

Appendix "B" Proposed Target Architecture



Appendix "C" HPC Facilities and Security

Facilities Introduction

As NOAA prepares to address the need for a future High Performance Computing (HPC) center to satisfy the computational and communication requirements to meet its mission in the 21st century, it needs to evaluate the requirements for its next generation high performance computing capability as well as its location. Major changes in the size, complexity, and resource requirements of next generation supercomputers need to be taken into account in designing and locating a new facility. In addition, increasingly strict security requirements based on the current global political realities need to integrate into the decision process.

Finally, consideration of current and projected NOAA resources, as well as those available through partnerships with other federal agencies, need to be carefully considered to provide NOAA with the best path forward that maximizes the technological capabilities available to meet the NOAA mission. NOAA's mission is highly dependent upon its ability to rapidly model and predict both atmospheric and oceanographic events. High Performance Computing is an essential tool for meeting this requirement and for keeping NOAA in the forefront of global climate and meteorological research. Beyond 2010, the supercomputing field will begin implementing petaflops and multiple petaflops computers. These computers can provide a substantial leap forward in the resolution and complexity of global climate models. To be the leader in global climate modeling, NOAA scientists must have leadership-class, supercomputing resources available to them.

The next generation of supercomputers is currently being designed. The industry as well as Defense Advanced Research Project Agency (DARPA) estimates show that electrical power requirements to both operate and cool the system will jump into the Megawatt range. This makes the facility a major power consumer in whatever area it is located and hence questions of electrical power cost, priority, and reliability become issues in locating the facility. It is imperative that the facility be designed to efficiently utilize energy, thereby minimizing its energy footprint. NOAA needs to carefully consider power related issues in both the design and location. Energy issues will be a driver in locating the facility.

In response to global terrorism, and cyber crime, the federal government is instituting increasingly strict requirements on Information Technology systems, especially when the facility has a major impact on an operational mission of the Federal Government. The center for NOAA's future HPC must meet these requirements. The location of the facility needs to minimize the risk from both natural and made-man vulnerabilities. This can be done by locating the facility in an area that has a low exposure to major natural catastrophic events. In addition, locating the facility on a limited access federal reservation away from major urban centers and outside of public view best minimizes the risk from man-made vulnerabilities.

With the increasing cost and complexity of next generation high performance computing, NOAA needs to analyze the best way to utilize existing resources across the Federal HPC community and take advantage of these resources.

Facilities Background Statement

The ability of NOAA to complete its mission is highly dependent on the ability to rapidly model and predict both atmospheric and oceanographic events. The purpose of this report is not to justify the need for new HPC capabilities but to set out the criteria that should be used in defining HPC requirements for the period 2010 to 2015 as part of that decision making process.

The selection criteria for new HPC capabilities are set out in the body of this report. They include issues such as the HPC capacity requirements, which must be 600% to 1,300% greater than current capabilities. Facility construction requirements are set out and include cooling, location, size and similar standards. Energy requirements, types, availability, reliability, and costs are estimated based on historical and currently available data.

In addition, communications capability requirements are examined in detail, with examples of three commercial vendor network options, other agencies' communications network options, and the potential for communications network collaboration with other agencies or with Government Owned Contractor Operated (GOCO) facilities such as National Laboratories.

Current and probable vulnerability issues are highlighted for consideration, including geological stability, historical wind (tornado/hurricane) risks, historical lightning risks, volcano risks, as well as flood and wildfire risks. Physical and cyber security risks are discussed, including potential for terrorism, crime, and cyber security vulnerabilities of utility providers through System Control and Data Acquisition (SCADA) and other Industrial Control Systems (ICS).

Finally, options for partnerships with existing Federal facilities are examined to highlight synergies and economies of scale.

While there is no perfect, risk free solution addressing all the selection criteria, certain options stand out above the less desirable options.

A summary matrix of requirements is included in the final section of this report.

Next Generation HPC Facility Criteria

Capacity

For NOAA to meet the future mission requirements of increased model resolution and complexity, it must migrate to a supercomputing capability that incorporates the best available technologies and operational parameters. Current projections for the FY2010-2015 time frame anticipate supercomputers in the 1-20 petaflops range, which represents a computing improvement of 600%-13,000% over current capabilities. Three example high performance computers that may expected in the years 2010 to 2015 are set out below in Table 1.

Model	IBM P7 IH 30 Frame Cluster[1]	Cray XT5 [2]	DARPA HPCS[2]
Petaflops (PF)	1	1	20
Power	4.3MW	6.5MW	15MW
Footprint (ft^2)	1,080	3,500	9,500
Cooling (Cooling Tons) [Size is for building machine will reside in]	Not specified	6,000	>6,000

Table 1: Expected Size of Leading Supercomputers in 2010-2015

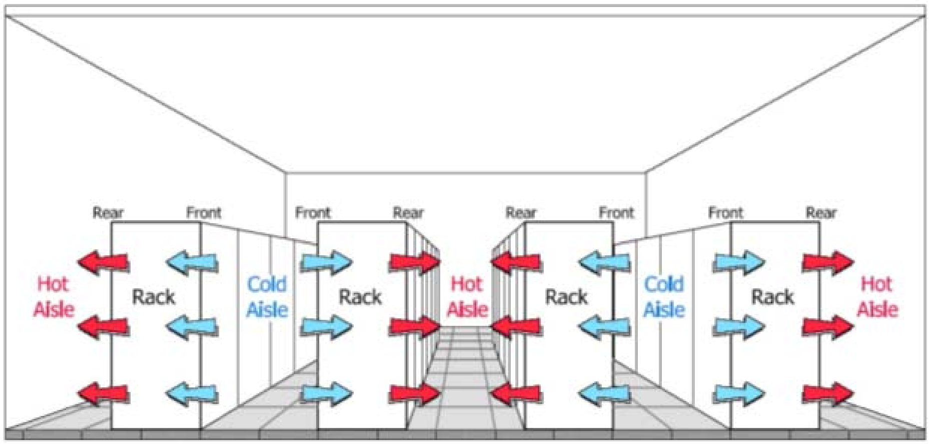
The data in Table 1 was provided by IBM and by Oak Ridge Laboratory (ORNL). ORNL expects to install the Cray XT5 and the DARPA HPCS in the FY2010-2015 time frame.

Other HPC Criteria:

Additional criteria that need to be considered in the design and specification of a new HPC building include:

- **Raised floor height** – Minimum raised floor height needs to be three feet.
- **Single fan per rack** – Use of a single fan per rack with top down cooling rather than traditional back to front cooling.

Figure 1 shows traditional front to back cooling is less efficient.



Source: ASHRAE (2004)

Figure 1: Traditional Front to Back Cooling

- **Tightly sealed computer rooms –** This is required for energy efficiency.

- **Primary transformer close to rack –** As petaflops computers have larger electrical power requirements, electrical power loss from friction in the lines between the transformers and the computers can be a significant cost factor. By minimizing this distance, enhanced efficiency can be achieved.

- **Chiller capacity expandable –** Growth capability needs to be considered for both the cooling and electrical systems. Experience at Oak Ridge National Laboratory is that in previous periods expected growth needs for cooling and electrical power greatly exceeded growth estimates. This disconnect required significant building modification and expansion.

- **Pipe Trenches –** Pipe trenches under a raised floor are now required for HPC systems.

- **Under-floor fire protection –** Under floor wet fire protection is needed in all raised floor areas.

- **Building automation –** Building needs HVAC, power distribution and chilled water system monitoring and automated control capability.

- **Operations staffing –** The facility needs 24/7 operators plus shift electricians and HVAC mechanics to ensure continuous operation.

- **Emergency response -** Due to the electrical power and cooling requirements for these systems, petaflops computing facilities have the potential for life threatening hazards to occur. As such, emergency response capabilities need to be situated in a reasonable proximity to the HPC facility such that timely response can occur. Further, the emergency response team needs to be trained for the specific hazards expected at the facility. Location on a limited access government reservation or substantial green space between the general public and the facility is recommended. The network system and facility infrastructure needs to be monitored 24/7.

- **Asset protection -** As the facility is the principle supercomputing facility for NOAA and the building is running as a "dark" facility with only control and maintenance staff, access needs to be very limited.

- **Reliable Power –** Such HPC requires zero point of failure in the power distribution system, meaning that HPC capabilities cannot be dependent on a single power generation facility. The HPC will need access to at least two external power generation facilities, each of which has lines going to closely located substations that connect to the HPC facility.

- **Estimated square footage for the HPC Center:**
  These estimates assume only operators and maintenance personnel are located in the building. The office space estimate includes conference room space. No visualization facility is included in the building. The raised floor space is estimated for a single supercomputer system and associated data storage. If more equipment is required, then additional raised floor space should be added.

Table 2: Sizing a Petaflops Facility

Case 1: 1 Petaflops Facility:		**Case 1: 20 Petaflops Facility:**	
Control Room	**1,000 sq. ft**	**Control Room**	**1,000 sq. ft**
Office Space	**2,500 sq. ft**	**Office Space**	**2,500 sq. ft**
HVAC / Electrical	**5,000 sq. ft**	**HVAC / Electrical**	**10,000 sq. ft**
Raised Floor Space	**10,000 sq. ft**	**Raised Floor Space**	**24,000 sq. ft**
Rest Room Facilities	**500 sq. ft**	**Rest Room Facilities**	**500 sq. ft**
Lobbies and Hallways	**2,000 sq. ft**	**Lobbies and Hallways**	**2,000 sq. ft**
Total	**21,000 sq. ft**	**Total**	**40,000 sq. ft**

From the above estimates, a Next Generation HPC facility will require a building in the range of 20,00 to 40,000 square feet as a minimum size. If growth is taken as a factor, then a building in the range of 30,000 to 50,000 square feet should be considered.

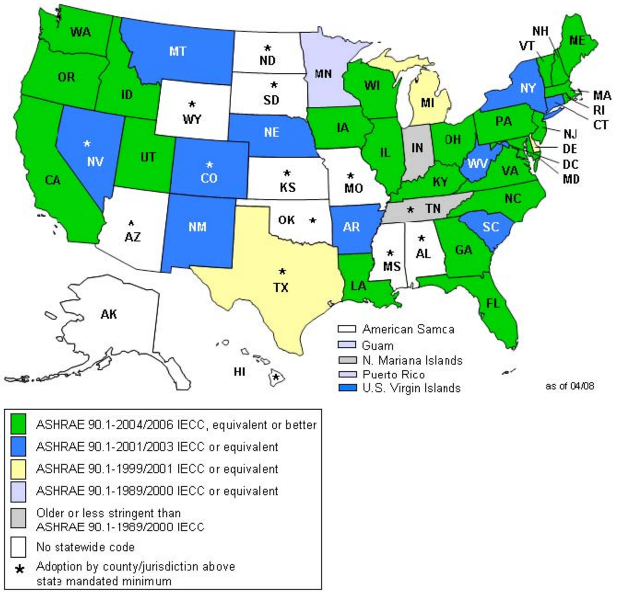
Location Factors

The determination of a location for the Next Generation HPC facility will be highly influenced by a number of external factors that vary across the United States but have a major impact on the success of the HPC facility. These factors include electrical power and its continuous availability, reliability, and cost; high-speed network infrastructure; site vulnerabilities, to include natural vulnerabilities, terrorism, and crime; and data flow between the satellite downlink site, HPC facility, and data storage site.

Building Energy Codes

Building energy efficiency codes have been standardized by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) and the International Code Council (ICC). The ASHRAE standard is 90.1 - Energy Standard for Buildings Except Low-Rise Residential Buildings and the ICC code is the International Energy Conservation Code (IECC). These codes are continually updated with versions bearing the year of the update. Under federal law, the states are free to adopt these codes (either the current version or an older version), modify the code to meet the state's requirements, or to adopt no code.

Figure 2: Current Status Building Codes



As energy efficiency is an extremely important requirement for the new HPC facility, locating the facility in states with less stringent codes does allow more flexibility for meeting specific NOAA design requirements.

Energy Requirements

Of all the factors that impact a location choice, energy may be the key location requirement for any facility interested in housing Petaflops-class computers. Given the need for a minimum of 6.5 MW of electrical power and a requirement for 24/7 operations in support of the National Weather Service production mission and international science interaction in research, the availability of power that is both reliable and cost effective will be a significant limitation on where the Next Generation HPC facility can be located.

Availability of Electricity

The Department of Energy completed a National Electric Transmission Congestion Study in 2006. Transmission congestion defines areas where there is insufficient capacity to meet demand at peak periods. Many of these areas are not building either new transmission lines or new local power plants to meet increasing demand requirements and the problem in these areas can only be expected to get worse. The Department of Energy found that three classes of congestion areas merit further Federal attention:

Critical Congestion Areas

These are areas of the country where it is critically important to remedy existing or growing congestion problems because the current and/or projected effects of the congestion are severe. Two such areas have been identified, each of which is large, densely populated, and economically vital to the Nation. They are:

- The Atlantic coastal area from metropolitan New York southward through Northern Virginia
- Southern California

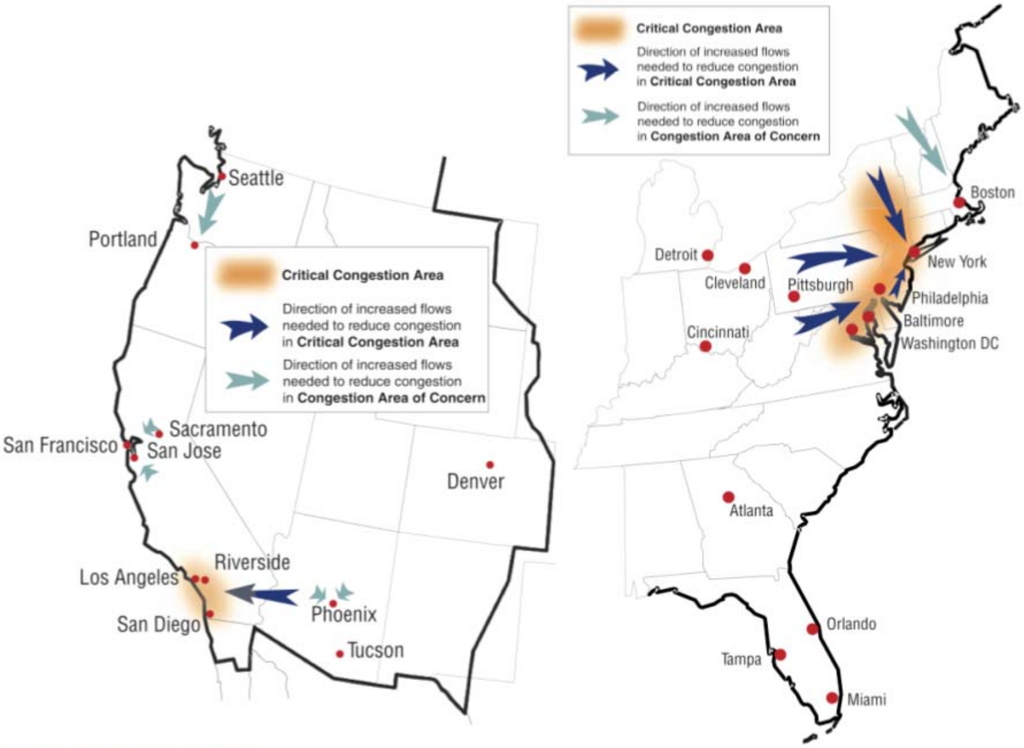
Congestion Areas of Concern

These are areas where a large-scale congestion problem exists or may be emerging, but more information and analysis appear to be needed to determine the magnitude of the problem and the likely relevance of transmission expansion and other solutions. Four Congestion Areas of Concern have been identified:

- New England
- Phoenix – Tucson area
- Seattle – Portland area
- San Francisco Bay area

These areas are shown in Figure 3. Currently, two of the NOAA HPC facilities are located in regions of Critical Congestion: Geophysical Fluid Dynamics Laboratory in Princeton, New Jersey and the National Center for Environmental Prediction's HPC Production facility at Gaithersburg, Maryland. In addition, the Washington Post published two articles in 2008 based on a study by the Maryland Public Service Commission that found that Maryland might face rolling blackouts as early as 2011 or 2012. Power could be shut down for perhaps an hour at a time in certain areas, such as on hot days when air conditioners strain the grid.

Figure 3: Areas of Critical Concern for Transmission Congestion



Source: (US DOE 2006b)

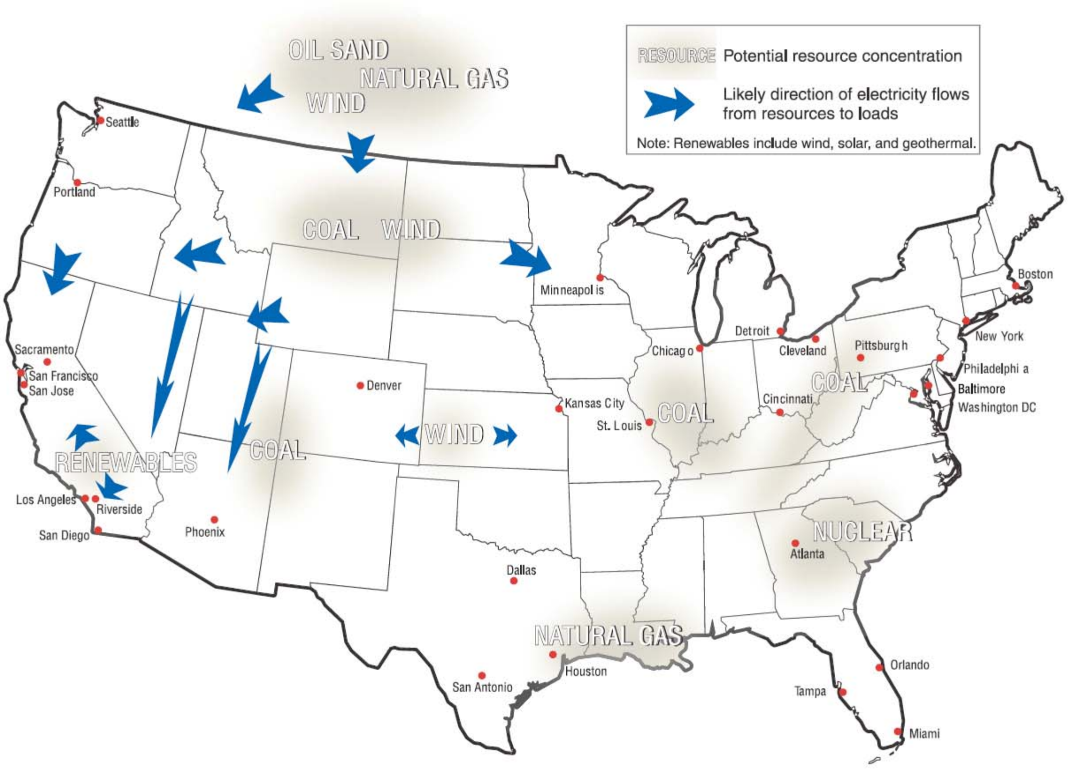
Conditional Congestion Areas

These are areas where there is some transmission congestion at present, but significant congestion would result if large amounts of new generation resources were to be developed without simultaneous development of associated transmission capacity. As shown in Figure 4, these areas have potential coal and nuclear generation capacities to serve distant interests. These areas are:

- Montana – Wyoming (coal and wind)
- Dakotas – Minnesota (wind)
- Kansas – Oklahoma (wind)
- Illinois, Indiana and Upper Appalachia (coal)

- The Southeast (nuclear)

Figure 4: Conditional Congestion Areas

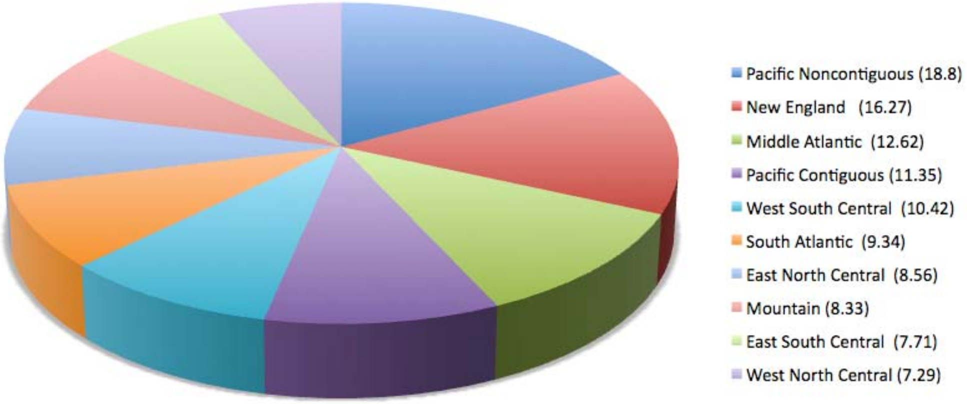


Areas listed as being in the Critical Congestion or Congestion areas of concern should not be considered for a facility that will require as much power and power availability requirements as the Next Generation HPC will demand. Areas listed under Conditional Congestion areas are acceptable as long as multiple power plants are located a relatively short distance from the facility, thereby minimizing potential future transmission problems. NOAA needs to have a power priority with the electric power provider in these regions.

Cost of Electricity

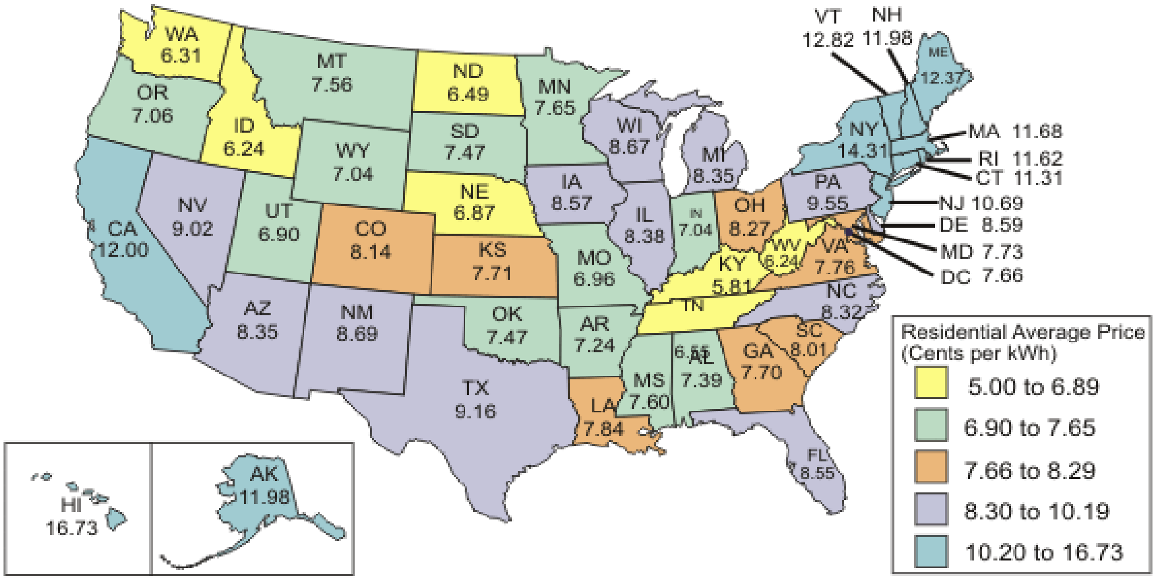
Electric Power cost measured in Kilowatt-Hours (KWh) varies widely across the United States. Within the Department of Energy, the Energy Information Administration (EIA) maintains information about electrical power cost for residential, commercial, and industrial rates on a nationwide average, regional average, and a statewide average. Figure 5 shows commercial electrical power rates for the 10 regions of the United States. Figure 6 shows residential electrical power rates per state for 2006.

Figure 5: Commercial Electric Power by Region (2007)



From Figure 5, the regions with the lowest rates are the West-North-Central Region (Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota), East-South-Central Region (Kentucky, Tennessee, Mississippi, and Alabama), and the Mountain Region (Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, and Wyoming). Electricity pricing varies across a region so not every state in a region—even a region with a low average cost— has a low electricity cost in their state.

Figure 6: Residential Electrical Power Cost by State (2006)



Source: Energy Information Administration, Form EIA-861, "Annual Electric Power Industry Report."

From Figure 6, the states in yellow and green are the lowest cost as defined in cents per kWh. While the map in Figure 6 is residential, commercial and industrial pricing is proportional so the state-to-state ratios are maintained.

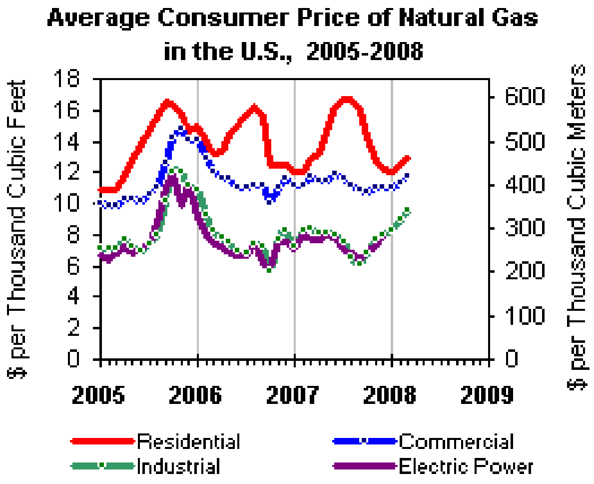
The energy source for electrical power generation is also an important parameter. Those states where a large percent of their electricity is generated from oil and gas plants are susceptible to fluctuations in the oil industry that can result in higher prices and potential shortages. For a reliable supply and stable pricing, states should be considered where a higher percent of electricity is generated from nuclear, coal, and hydroelectric.

From the above cost data, regions such as New England, the Middle Atlantic, Pacific Noncontiguous, and California represent costs that are too high to consider for economical operation of a Petaflops computing facility.

Alternate Energy Sources

Given the issues presented above, in an effort to both reduce cost and assure availability, a natural approach is to look at all types of electrical power supplies, including sources independent of the existing power grid in an effort to both reduce cost and ensure availability. A review of power alternatives might include Distributed Generation (DG) systems, solar, wind turbines, microturbines, fuel cells, gas turbines, and reciprocating engines. These technologies typically are designed to provide power in the 3kW to 10MW range depending on the technology used. Fuel cells typical can provide 10kW to 2MW; microturbines can produce 30kW to 250kW; gas turbines provide typically 500kW to 20MW and reciprocating engines can provide 100 kW to 3MW. Solar and wind are best in combination with other technologies as they are weather dependent and cannot provide 100% availability in all times.

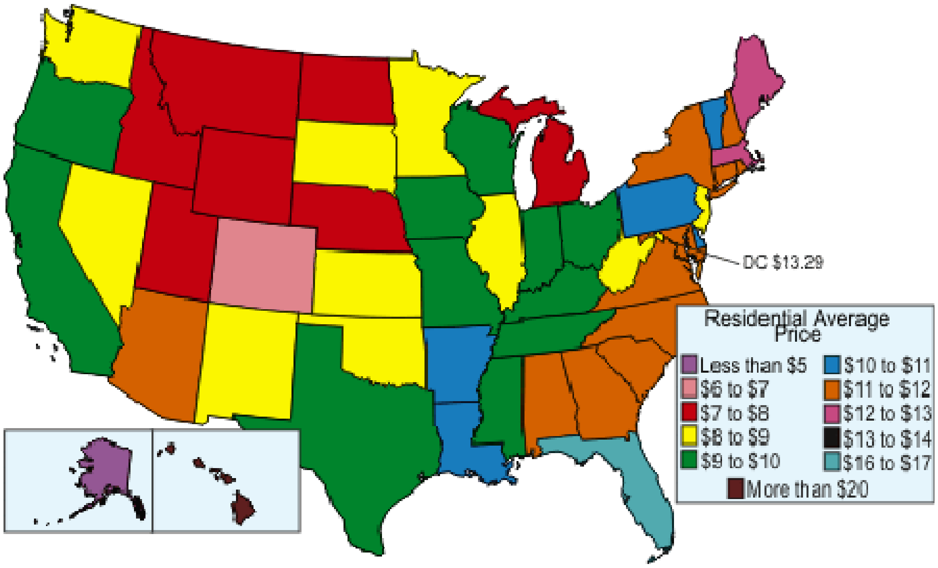
Figure 7: Natural Gas Pricing



However, an HPC center running a 1 to 20 Petaflops machine will require between 6MW and -20MW, meaning that the required power falls outside the range of Distributed Generation systems. Currently, only gas turbines are capable of generating 20MW of electrical power.

However, as the HPC initiative progresses, it would be wise to keep abreast of developments with natural gas fuel cells. They represent one of the most promising technologies from an energy efficiency and environmental perspective. Currently, this technology is being implemented at the NOAA Geophysical Fluid Dynamics Laboratory in Princeton, New Jersey. Data provided by the US Fuel Cell Council shows that there are eleven companies producing fuel cells and their output ranges between a 1 W portable system to a 125kW stationary system.

Figure 8: U.S. Natural Gas pricing by state (Dollars per Mcf)



In combination, current natural gas fuel cell systems could produce up to 2MW of power, which is well short of the requirement for petaflops systems. In monitoring advances in natural gas fuel cell technology, it will also be important to determine whether they develop any growth path for larger future systems as well.

In addition to being energy efficient, the cost of natural gas has not kept pace with the cost of petroleum-based energy products. Natural gas pricing (see Figures 7 and 8) has increased significantly over the past decade; there are areas of the country where it is plentiful and relatively inexpensive. Recent surges in the price of oil, and analyst predictions that it will go dramatically higher, make natural gas a comparative bargain despite the potential volatility of natural gas pricing during extended periods of cold weather. Figure 7 shows average natural gas prices between 2005 and 2008 and Figure 8 shows natural gas pricing by state for 2006.

Communications

The determination of a location for the Next Generation HPC facility needs to take into account the increasing quantity of data, both raw data and processed results, that will need to be transmitted in and from the HPC facility. Currently, NWS is using seven OC-3 lines to move data from the Wallops Island download site. This results in a minimum transmission bandwidth requirement at OC-24 or 1.244Gbps. However, for the Next Generation HPC facility, two key requirements must be taken into account.

First, model resolution must double at a minimum, which should double input data. Second, as the three existing HPC centers are consolidated into one center, the bandwidth of the new center must be capable of handling all the data flow of each former center plus projected growth. In addition, input from the satellite downlink site and output to the data storage center must be included. This results in a factor of eight increases in bandwidth requirements.

Given the current OC-24 bandwidth, the new center would need to have a baseline of OC-192 (9.952Gbps) or approximately 10Gbps. Again this would be a minimum for a petaflops center given that NOAA may have a mission in production climate modeling as well.

Commercial high-speed network backbones connect major urban centers and hence they have located the GigaPOP connections in these urban areas. Thus, the highest probability of obtaining 10Gbps and higher bandwidth is in or around these urban centers. As an example of where the high-speed fiber backbones exist, three major commercial carriers are included in this study. The Sprint MPLS network is shown in Figure 9; the Qwest network is shown in Figure 10; and the Internet2 (Level3 Communications) network is shown in Figure 11. Additionally, ESnet, a major government research network managed by the Department of Energy, is also shown in Figure 12.

The network diagrams in Figures 9 through 12 confirm the concentration of POPs in urban areas and show two regions that clearly lack high-speed networks: 1) New England north of Boston; and 2) the upper part of the West North Central region, specifically North and South Dakota, Montana, parts of Iowa, upper Minnesota, and upper Wyoming. The later has one of the lowest electric power rates but is clearly not able to support high-speed network requirements.

The major exception to the above observations results when a major federal agency such as the Department of Defense (DOD), Department of Energy (DOE), or the National Aeronautical and Space Agency (NASA) has a major facility located in a non-urban region that requires high-speed networks. ESnet (Figure 12) shows a number of examples, such as the DOE National Laboratories, where very high-speed connectivity is available in a non-urban setting. In these cases, both low energy prices and high-speed network requirements can be met. NOAA should strongly look at the option of partnering with another federal agency such as the DOE, NASA, and/or DOD.

Another option is the use of dark fiber—available through a downlink site at Wallops Island or the National Climate Data Center in Asheville, North Carolina—which would possibly allow higher bandwidth, lower latency, and access to direct optical fiber, non-Internet Protocol transmissions as well as access to many power companies that would allow high-speed connectivity to a GigaPOP. If one were to consider only the increasing quantity of transmitted data, locating the HPC facility closer to the transmission between facilities would appear logical. Yet, while this is especially important for raw data transmission, it is not the only criteria to be considered, as is noted in previous and following sections of this paper.

Figure 9: Sprint North American MPLS Backbone

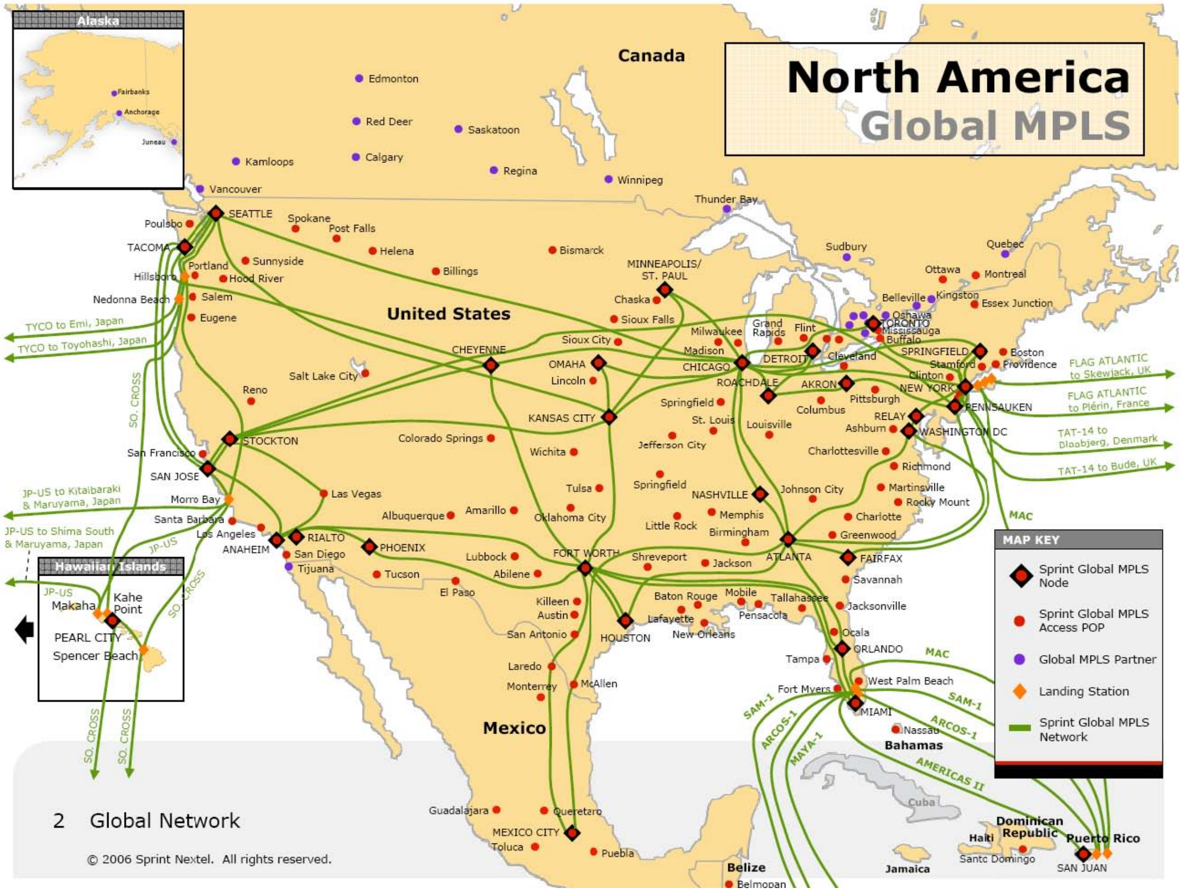


Figure 10: Qwest North American Fiber Backbone

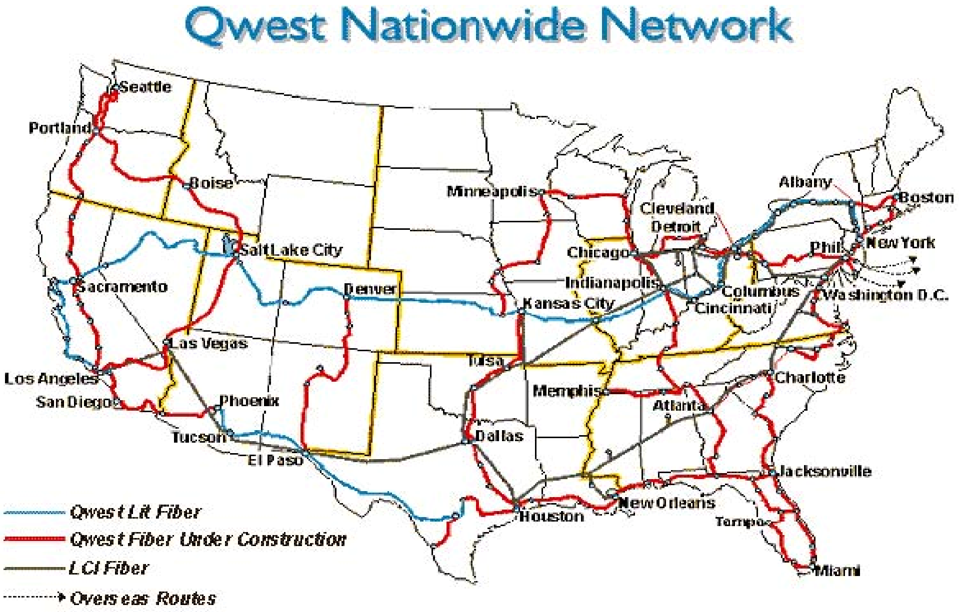


Figure 11: Internet 2 Backbone

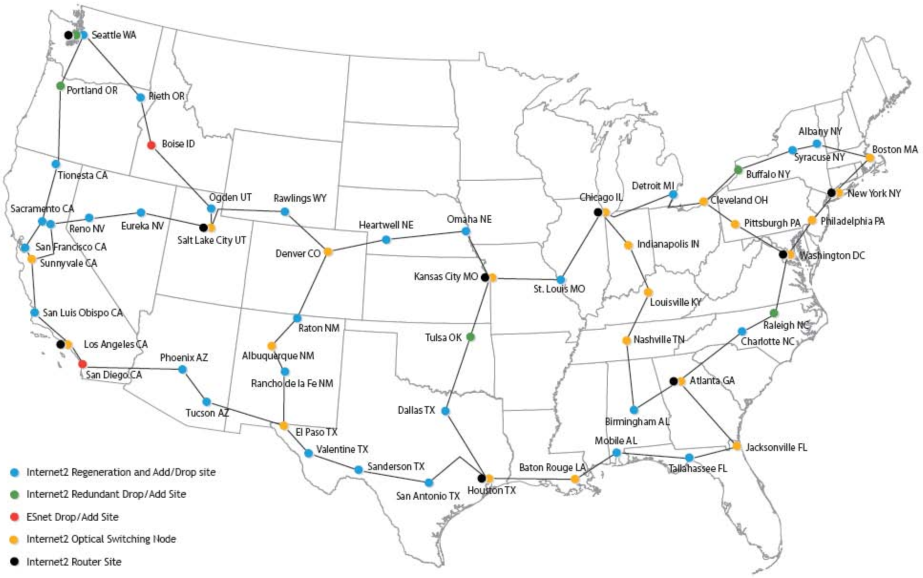
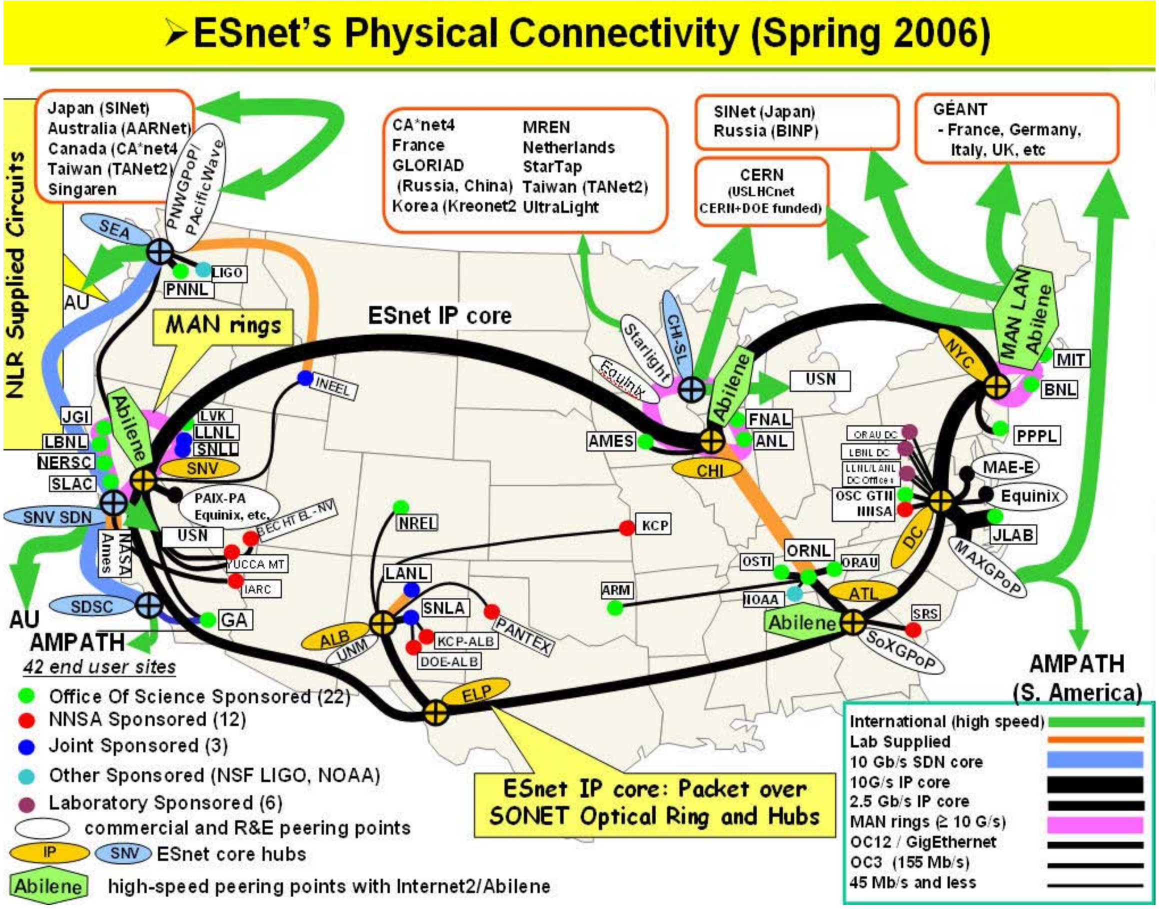


Figure 12: ESnet Backbone (Spring 2006)



Vulnerabilities

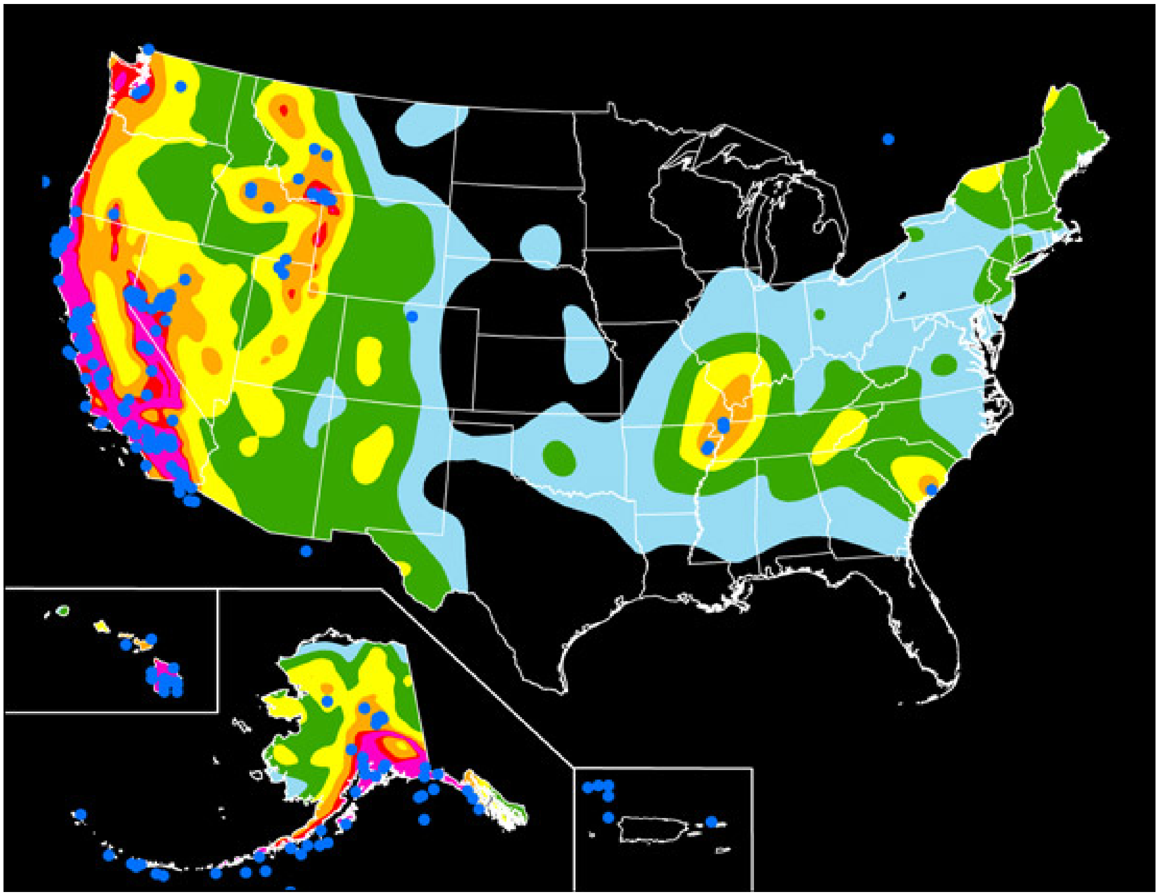
The location of a new facility must take into account both known natural vulnerabilities and vulnerabilities to human threats. Given the 24/7 operational nature of the facility, any event that would significantly impact operational status of the center should be considered and locations with a history of known catastrophic events should be avoided as a site for the facility.

Earthquakes

Earthquakes are a large-scale catastrophic event in parts of the United States. Figure 13 shows the history of strong shaking over the past fifty years. Areas of major earthquake concern are those areas which are colored orange, pink or red. These areas include the West Coast, Wasatch and Bitterroot mountains, and the New Madrid Fault zone.

Using earthquake risk as a factor, Charleston, South Carolina, as well as Southern Alaska and parts of Hawaii, are not optimum sites for location of the facility. Areas of high probability of earthquakes also have the highest probability of physical infrastructure damage such as damage to buildings, power lines and network communication lines, as well as damage to roads and bridges that would impact staff accessing the facility. In areas of moderate shaking, while the tremors may well be felt, they have a low probability of interfering with facility operation in an appropriately constructed facility.

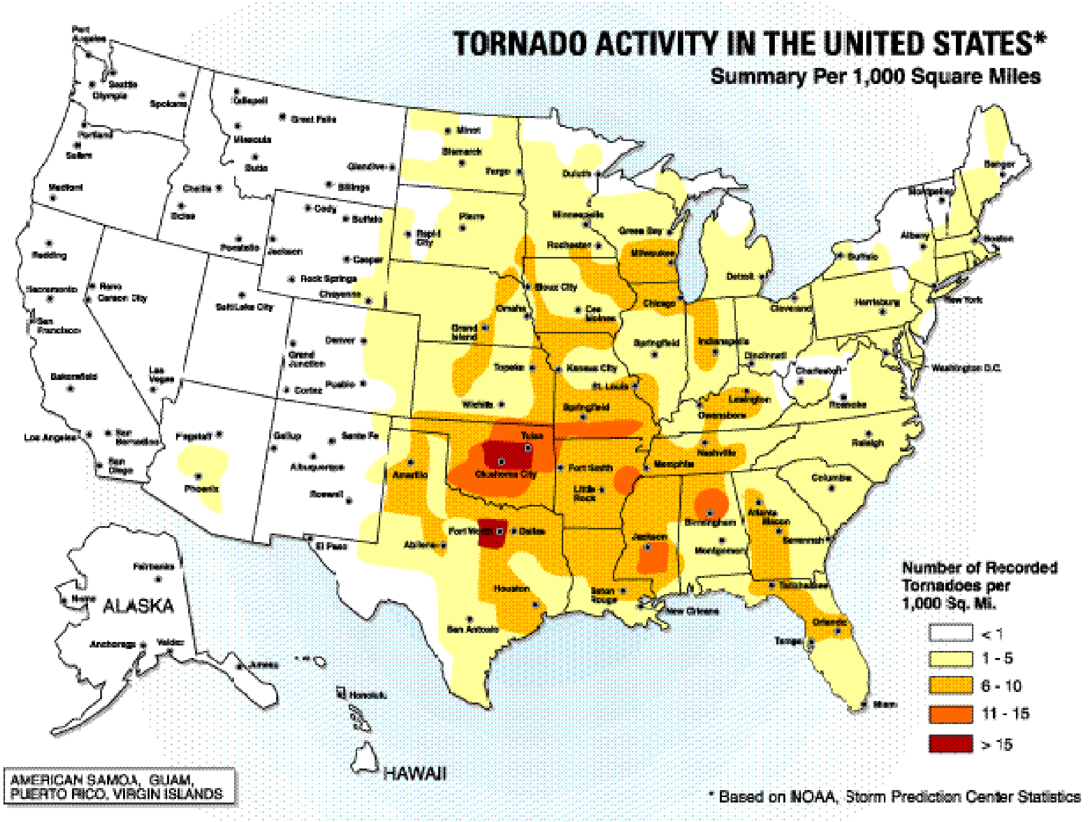
Figure 13: United States Map Showing Probability of Major Earthquake Threat



Tornadoes

Unlike earthquakes that can be felt uniformly over an extended area, tornadoes represent a much narrower field of impact. They can occur in most areas east of the Rocky Mountains. Figure 14 shows a summary of recorded tornadoes per 1000 square miles represented by areas that are shaded red, orange and dark yellow. This includes Oklahoma, Arkansas, Western Tennessee, Mississippi, Missouri, Northern Alabama, Northern Texas, and Northern Louisiana. While a direct hit on an HPC facility is statistically unlikely in all but the areas colored red or orange, the probability of disruption to power or communications may be of much higher probability. This would be especially true of power and communication lines that extend for many miles across tornado risk areas. Locating the HPC center in a region that is shaded light yellow or white minimizes the risk a tornado event impacting the center's operation.

Figure 14: Tornado Activity in the United States



* Based on NOAA, Storm Prediction Center Statistics

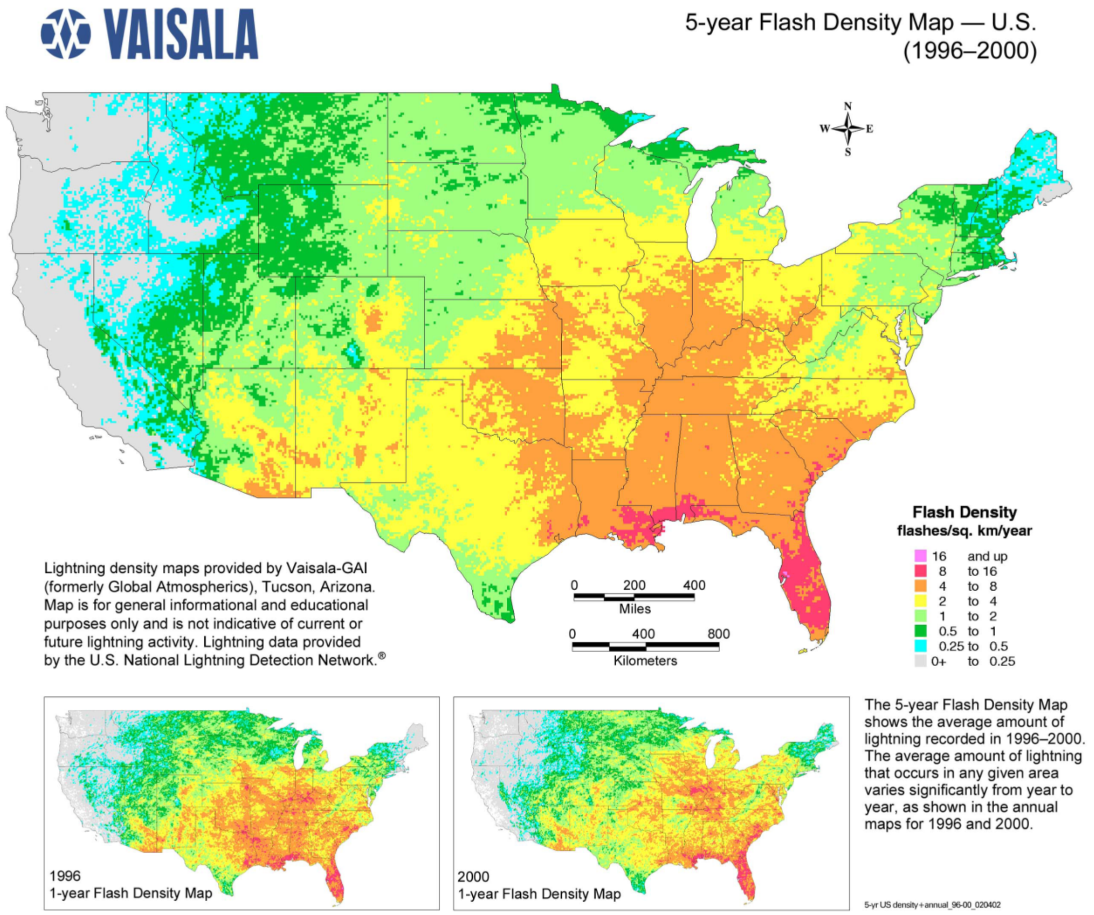
Figure I.1 The number of tornadoes recorded per 1,000 square miles

Lightning

Lightning is a leading cause of outages on power utility transmission and distribution systems. In the United States, lightning is estimated to cause more than $1 billion in damage and loss to utilities and their customers every year. Figure 15 shows flash density for the 5-year period of 1996 through 2003.

From the map in Figure 15, areas along the Gulf and up the Eastern Seaboard to North Carolina have the highest probability of lightning strikes. This is followed by the interior Southeast and up the Mississippi and Ohio rivers. Given that the Gulf coast, the State of Florida, and the Eastern Seaboard to at least North Carolina are also areas of extreme hurricane winds. In addition, the combination of wind and lightning strikes makes those highly inappropriate HPC locations.

Figure 15: United States Lightning Flash Density Map from 1996-2003



Lightning density maps provided by Vaisala-GAI (formerly Global Atmospherics), Tucson, Arizona. Map is for general informational and educational purposes only and is not indicative of current or future lightning activity. Lightning data provided by the U.S. National Lightning Detection Network.®

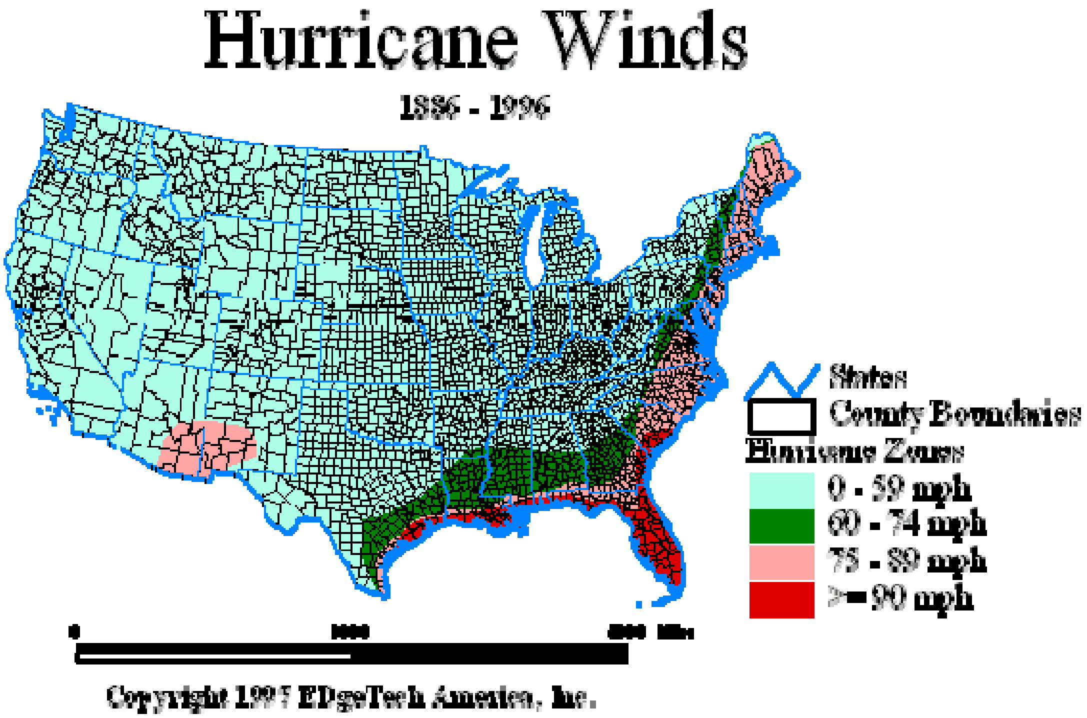
The 5-year Flash Density Map shows the average amount of lightning recorded in 1996–2000. The average amount of lightning that occurs in any given area varies significantly from year to year, as shown in the annual maps for 1996 and 2000.

Locations in the areas shaded orange and yellow are acceptable risk given compensating safeguards such as the Fault Analysis and Lightning Location System, developed by the Electrical Power Institute (EPRI). In addition, equipment manufactured by Vaisala can provide power companies with situational awareness tools that allow transmission switching decisions to be made that minimize power outages due to lightning. In addition, the use of backup generators also reduces or eliminates any impact on the HPC facility.

Hurricanes

Hurricanes are another large-scale catastrophic event that affects large areas. The location of the HPC facility should avoid areas where there is significant impact by hurricanes. Hurricane impact includes storm surge, flooding, and winds. Figure 16 shows hurricane wind levels across the continental United States. Locating the facility in areas where hurricane winds have historically been below 60 MPH eliminates threats from storm surge and significantly reduces risk of an operational impact from hurricane winds. The Gulf Coast from Texas to Florida, the entire state of Florida, and the entire Eastern Seaboard represent areas to avoid when determining the HPC site location.

Figure 16: Hurricane Winds from 1886-1996



Volcanoes

The eruption of volcanoes can be an extremely catastrophic event that can both completely destroy a facility and significantly interrupt power and communications to the HPC facility. Figure 17 shows that volcanoes are located in the Hawaiian Islands, Alaska, and the western part of the United States. Figure 18 highlights (solid red triangles) the locations of volcanoes that have been active in the past 2,000 years. All of these volcanoes are located in the Cascade Mountain Range of the Pacific Northwest and Northern California, with the exception of Long Valley Caldera located in the Serra Nevada Mountains of Central California. The areas most impacted from a site study perspective are locations in or near Seattle, Washington, Portland, Oregon, or Willamette Valley in Oregon.

Figure 17: United States Map Showing Areas Where Active Volcanoes are Located

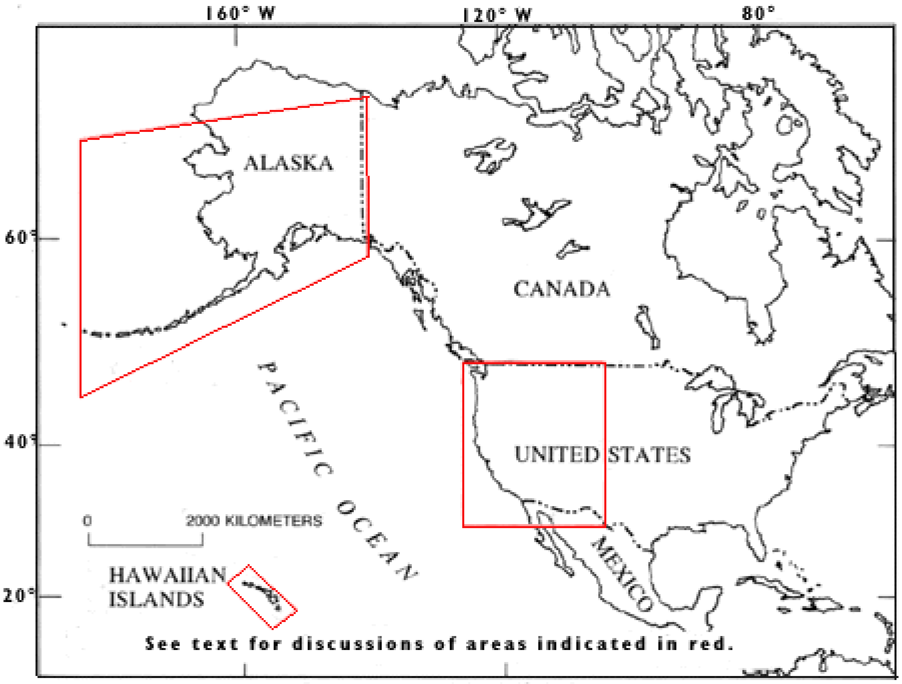


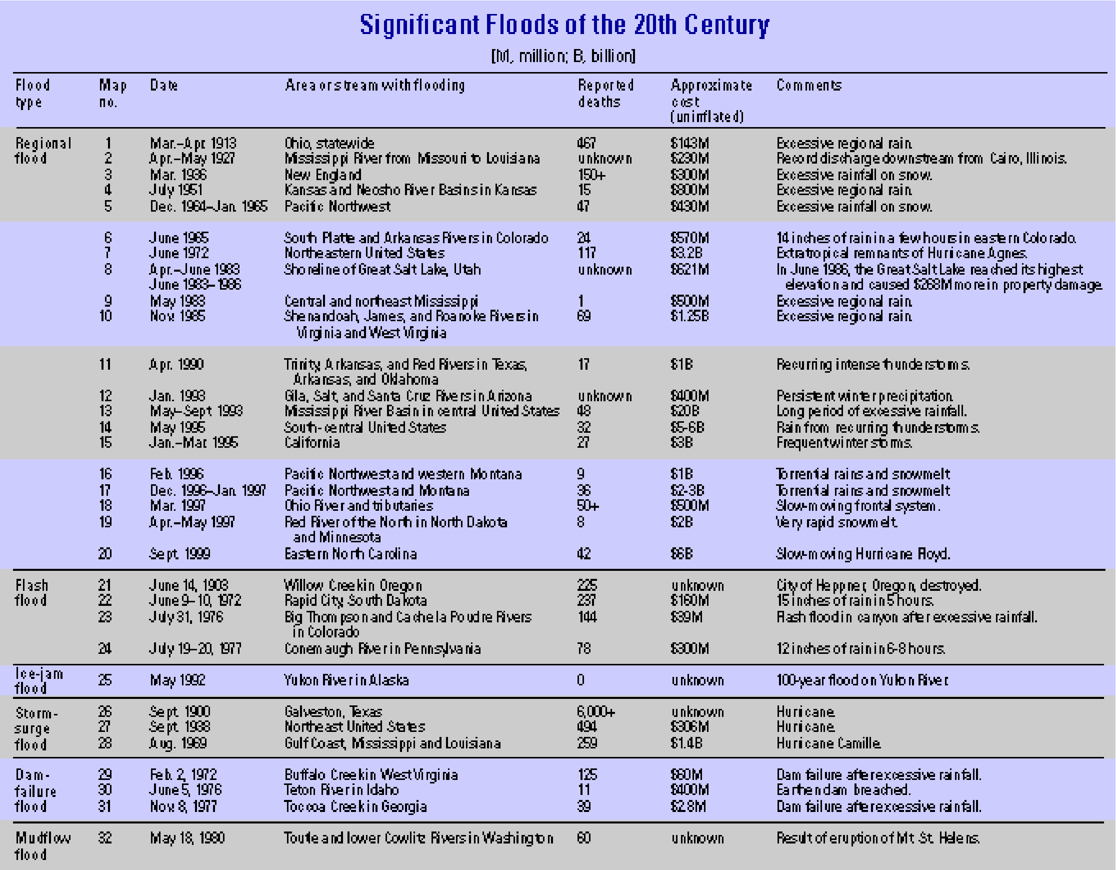
Figure 18: A Map of Active Volcanoes in the Continental Unites States

Topinka, USGS/CVO, 1999, Modified from: Brantley, 1994, Volcanoes of the United States: USGS General Interest Publication

Flooding

Major flooding events have occurred throughout the United States. Table 3 shows a list of significant flooding events in the twentieth century. Figure 19 shows a map of these floods. The principal areas affected by flooding are the Mississippi River and its connecting rivers, including the Missouri, Arkansas, and Ohio rivers as well as coastal areas where flooding is caused by hurricanes. Significant work has been done by the United States Government to control flooding in the twentieth century. However, as late as the 1990's, the rivers listed above have continued to have major flood events.

Table 3: Significant Floods in the 20th Century

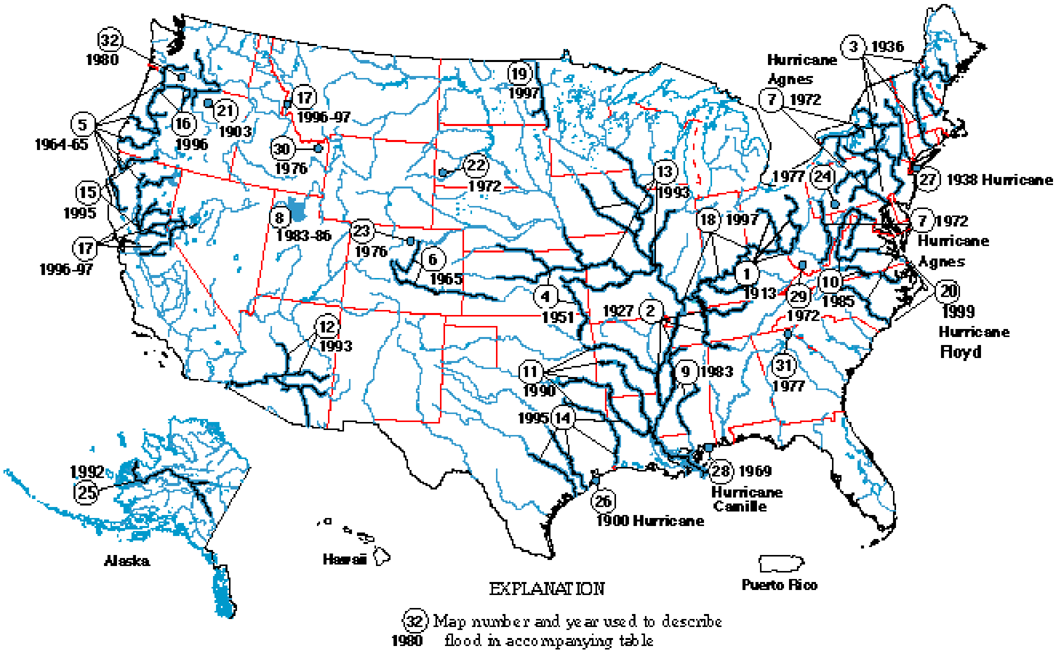


Significant Floods of the 20th Century

[M, million; B, billion]

Flood type	Map no.	Date	Area or stream with flooding	Reported deaths	Approximate cost (uninflated)	Comments
Regional flood	1	Mar.–Apr. 1913	Ohio, statewide	467	$143M	Excessive regional rain.
	2	Apr.–May 1927	Mississippi River from Missouri to Louisiana	unknown	$230M	Record discharge downstream from Cairo, Illinois.
	3	Mar. 1936	New England	150+	$300M	Excessive rainfall on snow.
	4	July 1951	Kansas and Neosho River Basins in Kansas	15	$800M	Excessive regional rain.
	5	Dec. 1964–Jan. 1965	Pacific Northwest	47	$430M	Excessive rainfall on snow.
	6	June 1965	South Platte and Arkansas Rivers in Colorado	24	$570M	14 inches of rain in a few hours in eastern Colorado.
	7	June 1972	Northeastern United States	117	$3.2B	Extratropical remnants of Hurricane Agnes.
	8	Apr.–June 1983 June 1983–1986	Shoreline of Great Salt Lake, Utah	unknown	$621M	In June 1986, the Great Salt Lake reached its highest elevation and caused $268M more in property damage.
	9	May 1983	Central and northeast Mississippi	1	$500M	Excessive regional rain.
	10	Nov. 1985	Shenandoah, James, and Roanoke Rivers in Virginia and West Virginia	69	$1.25B	Excessive regional rain.
	11	Apr. 1990	Trinity, Arkansas, and Red Rivers in Texas, Arkansas, and Oklahoma	17	$1B	Recurring intense thunderstorms.
	12	Jan. 1993	Gila, Salt, and Santa Cruz Rivers in Arizona	unknown	$400M	Persistent winter precipitation.
	13	May–Sept. 1993	Mississippi River Basin in central United States	48	$20B	Long period of excessive rainfall.
	14	May 1995	South-central United States	32	$5-6B	Rain from recurring thunderstorms.
	15	Jan.–Mar. 1995	California	27	$3B	Frequent winter storms.
	16	Feb. 1996	Pacific Northwest and western Montana	9	$1B	Torrential rains and snowmelt
	17	Dec. 1996–Jan. 1997	Pacific Northwest and Montana	36	$2-3B	Torrential rains and snowmelt
	18	Mar. 1997	Ohio River and tributaries	50+	$500M	Slow-moving frontal system.
	19	Apr.–May 1997	Red River of the North in North Dakota and Minnesota	8	$2B	Very rapid snowmelt.
	20	Sept. 1999	Eastern North Carolina	42	$6B	Slow-moving Hurricane Floyd.
Flash flood	21	June 14, 1903	Willow Creek in Oregon	225	unknown	City of Heppner, Oregon, destroyed.
	22	June 9–10, 1972	Rapid City, South Dakota	237	$160M	15 inches of rain in 5 hours.
	23	July 31, 1976	Big Thompson and Cache la Poudre Rivers in Colorado	144	$39M	Flash flood in canyon after excessive rainfall.
	24	July 19–20, 1977	Conemaugh River in Pennsylvania	78	$300M	12 inches of rain in 6-8 hours.
Ice-jam flood	25	May 1992	Yukon River in Alaska	0	unknown	100-year flood on Yukon River.
Storm-surge flood	26	Sept. 1900	Galveston, Texas	6,000+	unknown	Hurricane.
	27	Sept. 1938	Northeast United States	494	$306M	Hurricane.
	28	Aug. 1969	Gulf Coast, Mississippi and Louisiana	259	$1.4B	Hurricane Camille.
Dam-failure flood	29	Feb. 2, 1972	Buffalo Creek in West Virginia	125	$60M	Dam failure after excessive rainfall.
	30	June 5, 1976	Teton River in Idaho	11	$400M	Earthen dam breached.
	31	Nov. 8, 1977	Toccoa Creek in Georgia	39	$2.8M	Dam failure after excessive rainfall.
Mudflow flood	32	May 18, 1980	Toutle and lower Cowlitz Rivers in Washington	60	unknown	Result of eruption of Mt. St. Helens.

Figure 19: United States Map of Major Flood Events in the 20th Century



Wildfires

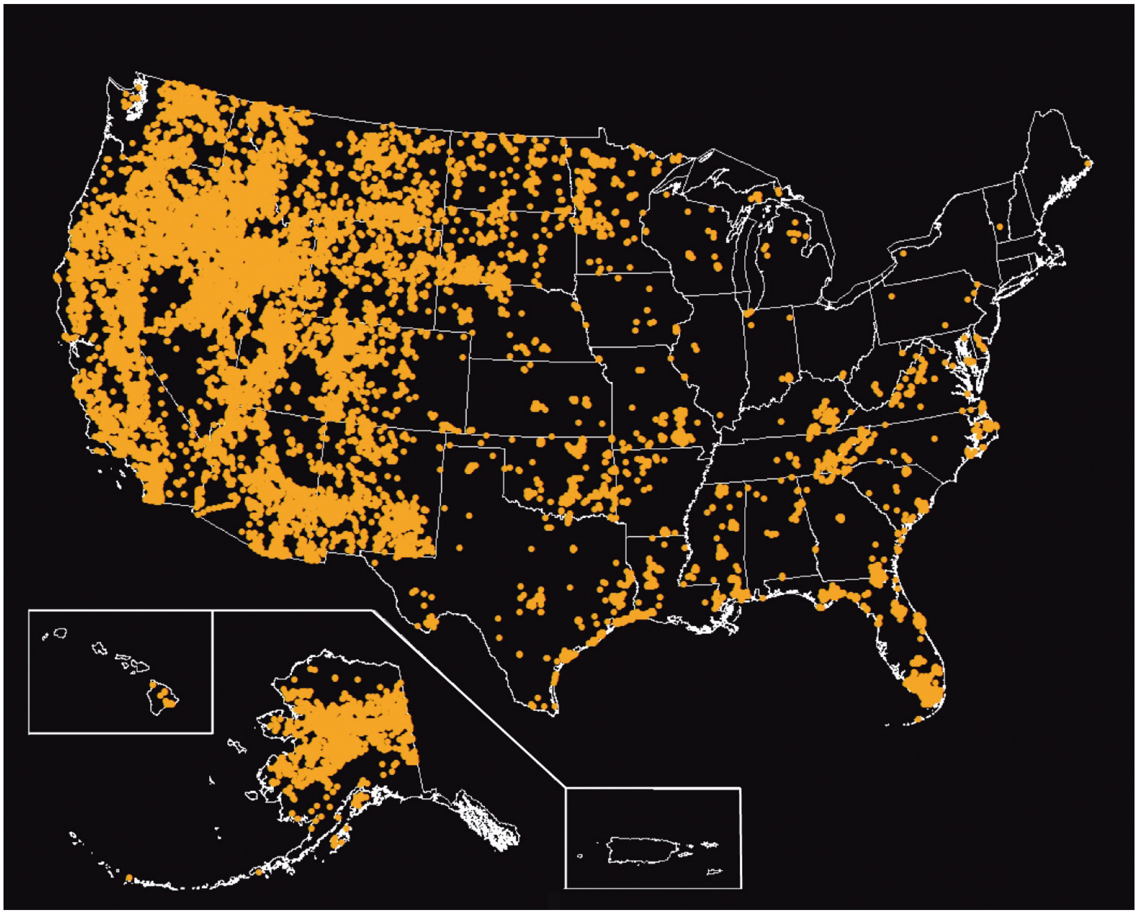
Wildfires are another event that can impact large areas. They can be especially destructive of power and communications lines. Figure 20 is a map of the United States showing areas that have experienced major wildfires during the period of 1980 to 2003. While the map shows that wildfires have occurred in all regions of the United States, the highest concentrations have been in Alaska and the Western third of the continental United States.

California has had significant impact from wildfires, an example of which was the Harris fire of October 2007. This fire damaged and disabled the Southwest Power Link, a 500,000-volt power line from Arizona to San Diego on October 21st, resulting in power outages to 335,000 Southern California Edison customers. The power outages were wide spread, affecting the areas of Ojai, Oxnard, Simi Valley, Santa Clarita, Thousand Oaks, Rialto, Fontana, San Benardino, Rancho Cucamonga, Mira Loma, Hesperia, Corona, Bloomington, Irvine, Calimesa, and Rubidoux. On October 23rd, the California Independent System Operator Corp had to declare an energy transmission emergency in Southern California, due to wildfires disabling 500,000-230,00- and 138,00-volt lines in San Diego. Over 29,992 residences lost power, due to lack of power from the power grid. The wildfire caused evacuations that displaced more than 500,000 people, resulting in the largest set of evacuations in the history of California.

Anyone considering potential HPC location sites along the Rocky Mountains or west of the Rocky Mountains should carefully analyze historic data on wildfires in the area and their impact.

Many areas of major wildfires such as Southern California are not prime choices for HPC location due to other issues such as earthquakes and electrical cost and availability.

Figure 20: Locations that Experienced Wildfires Affecting Areas Greater than 250 acres from 1980 to 2003.



Terrorist

Not all catastrophic events are natural – some are man-made. In the twenty-first century, terrorist attacks are an important concern for federal facilities. In conducting a location study, terrorism needs to be divided into two types of attacks - wide area attacks and point attacks on a specific, narrow group of buildings or people. In the latter case the very fact a federal facility exists can make it a target for a terrorist attack. In these cases physical security measures are the only way to minimize an attack.

It is for the wide area cases that strategic location is important. In such a case, an HPC facility would likely be a collateral damage victim and not the principal target of interest. The principal wide area attacks of interest are weapon of mass destruction (WMD) attacks aimed at large urban areas. The Department of Homeland Security has designated six urban areas as Tier 1 cities that are at risk. These urban areas are as follows:

- Bay Area (CA)
- Chicago (IL)
- Houston Area (TX)
- Los Angeles/Long Beach Area (CA)
- National Capitol Region (DC)
- New Your City/Northern Ney Jersey Area

For WMD attacks, the minimum safe distance from the urban center is approximately 300 miles[23,24]. By locating a facility outside of this distance to an urban area, the facility has the best chance of minimizing disruption of operation, which includes disruptions to both power and communications. Figure 21 shows the six Tier I urban areas with circles extending to 300 miles from the center of the urban area.

Further, it is important to note that on the North American continent large-scale wind movement is generally from the West to the East. Thus radiation, biological, and chemical plumes would follow weather patterns. It should be noted that the National Climate Data Center is outside of the minimum distance of all six urban areas, while the satellite downlink site is inside both New York and Washington DC's minimum distance. As the satellite downlink site is at the edge of both urban centers' minimum distances, physical damage is unlikely to occur and automated operations should continue unaffected by the event. Communication should not be impacted as packet switched communications were developed under just this type of scenario.

Figure 21: United States Map Showing the Minimum Safe Distance for Tier 1 Cities

However, if multiple urban areas are attacked, resulting in multiple GigaPoPs going offline, then communications could be impacted. If the location of the HPC facility was selected to be close enough to both the satellite downlink site and the data center, direct optical fiber non-standard protocol communication could be used, such as leasing fiber runs on power transmission links.

A secure and closed end system could also be developed that would not be affected by WMD attacks in the major urban areas listed above.

Crime

Crime is found at varying levels in all parts of the United States. As in the case of small-scale terrorist attacks, physical security is the best means of preventing an event from occurring. If the facility can be located on a larger federal reservation that has limited access, is physically not viewable by the public, and has a 24/7 protection force in place, as well as a 24/7 emergency response team, operational impact can be mitigated. The use of HSPD-12 physical access controls must be mandatory and will ensure that unauthorized persons cannot easily access the facility.

Cybersecurity

Cybersecurity will be a major driver in the requirements for a new HPC center. It will not have as major of an impact on location.

Analysis based on NIST Special Publication 800-63

- Access Control – Logical access control will be via HSPD-12 approved devices. See section on HSPD-12 below.

- Audit and Accountability – Given the large data sets and high performance nature of the system, both auditing and accountability techniques need to be predetermined to ensure they don’t have a negative impact on performance or operating costs.

- Configuration Management – Given the single purpose of high performance computing, the rules and strategies of configuration management should be put in place before initial startup.

- Contingency planning – The contingency planning needs to ensure that there are no single points of failure. The facility needs to have access to more than one GigaPOP in case the primary GigaPOP fails. Since most GigaPOPs are in urban areas, any natural or man-made catastrophic event could eliminate a single GigaPOP. Next, the facility needs to have the capability of obtaining power from more than one power generation station. At the power levels required for Petaflops computing, more redundancy is required than just having backup generators. Finally NOAA needs to have agreements in place to utilize other HPC facilities as a backup supercomputer site. Costs would be prohibitive to have a second petaflops or higher machine idle as a backup system. A cross federal government agreement may be a better solution.

- **Identification and Authentication –** All personnel accessing the system will need to have FIPS-201 Personal Identity Verification (PIV) in place. The PIV process has been developed by the National Institute for Standards and Technology (NIST). PIV outlines a standard procedure that all federal departments and agencies must follow to confirm the identities of its employees and contractors before issuance of a credential (identification badge). A PIV card is the generic name for a common identification card that is produced by the HSPD-12 system.

- **Maintenance –** Given the scarcity of petaflops computing maintenance experts, there would be a definite benefit to co-locating the HPC with other federal organization petaflops computing facilities such that maintenance personnel might well be stationed near the facility, with all PIV clearances in place and equipment in place so that nothing needs to leave the federal reservation.

- **Physical and Environmental Protection** Given that the HPC facility has limited staffing, considerable resource investment could be saved by locating the HPC facility at another government facility that can handle physical security and provide visitor control and HPSD-12 physical access control systems.

- **System and Service Acquisition –** As in the maintenance area discussed above, acquisition and life-cycle support could be greatly simplified if the facility is collocated with another government agency, such as the DOE National Laboratory that has existing strong vendor relationships and can provide economy of scale purchases of systems and components for NOAA.

- **System and Communication Protection –** The ability to have direct optical fiber (non-IP transmission) between the downlink site, the HPC facility, and the data storage facility helps improve both performance and security.

- **System and Information Integrity –** Data integrity is also enhanced by having direct optical fiber (non-IP transmission) between the downlink site, the HPC facility, and the data storage facility.

Trusted Internet Connections (TIC) Initiative

The Trusted Internet Connection Initiative, which needs to be implemented by June of 2008, will reduce the total number of federal external Internet connections to fifty. As long as the HPC facility's communication networks are within the federal space and do not need to transverse the external interfaces, this will not have a major impact on operations or location choice. It may have some impact on University and international users on conducting research using the supercomputers.

HSPD-12 Logical Access Control Systems

Under the HSPD-12 mandate, departments and agencies are required to have their logical access control systems ready to read and process the new smart card no later than October 27, 2009. Logical access control systems will require use of smart card readers. These systems will be required to authenticate and authorize an individual to access federally controlled information systems. Any new facility should be built with compliance in mind from the start. It is an unresolved issue how this will work on the research side where international agreements mandate access to the HPC systems.

SCADA Cybersecurity Concerns

As discussed above, as the HPC center begins to install one petaflops and higher computer systems, the electrical power requirements can be 5 MW or higher. Just having diesel engines for backup can no longer be considered sufficient for any outage that lasts more than a few hours as diesel fuel cost and probable availability will be an increasing problem in the future. As noted elsewhere herein, these power levels exceed the generation capacity of current fuel cells. Therefore the reliability of electric power is essential for HPC operation.

The United States electric power industry has a major issue in trying to secure Supervisory Control and Data Acquisition (SCADA) equipment and other Industrial Control Systems (ICS) from cyber attack. The GAO released a report in May 2008 on the Tennessee Valley Authority (TVA), which is a federally owned power utility. The GAO report is entitled “TVA Needs to Address Weaknesses in Control Systems and Networks[25].” While this report specifically addresses TVA, most of what is discussed is true, if not worse, at most of the power utilities in the United States. Since TVA is now mandated by Congress to fix these problems, the issues listed in the report are not an issue for picking a TVA supported site for the HPC facility in the 2011-2015 time period. However, regardless of where the facility is located, NOAA must actively be aware of the Cybersecurity posture of the power utility it is relying on for operation of its HPC facility. One advantage of TVA is, as a federal agency, NOAA should have no problem obtaining federal reviews of TVA Cybersecurity. Private power companies consider this information proprietary and it is difficult to obtain.

Future Security Concerns

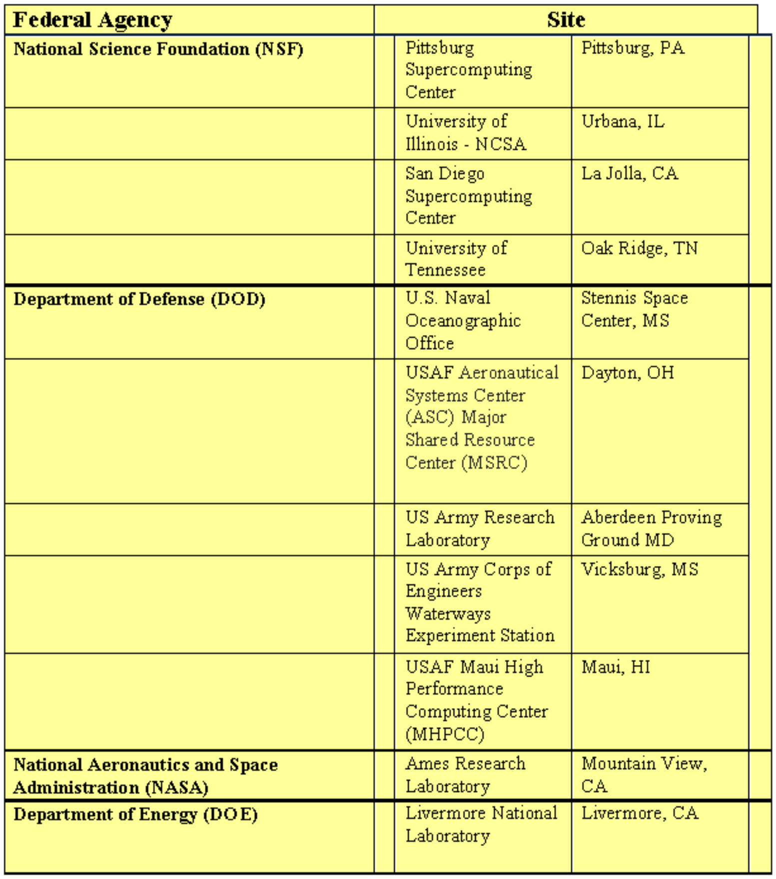
The biggest future security issue is the ability to do both research and operations in the same facility and/or with the same machine. The trend for the future is continued control and increased isolation of the .gov network. Research work will most likely be moved to Government Owned, Contractor Operated (GOCO) Federal Research Laboratories or universities that can maintain systems in other domains than the .gov domain.

Federal Partnerships

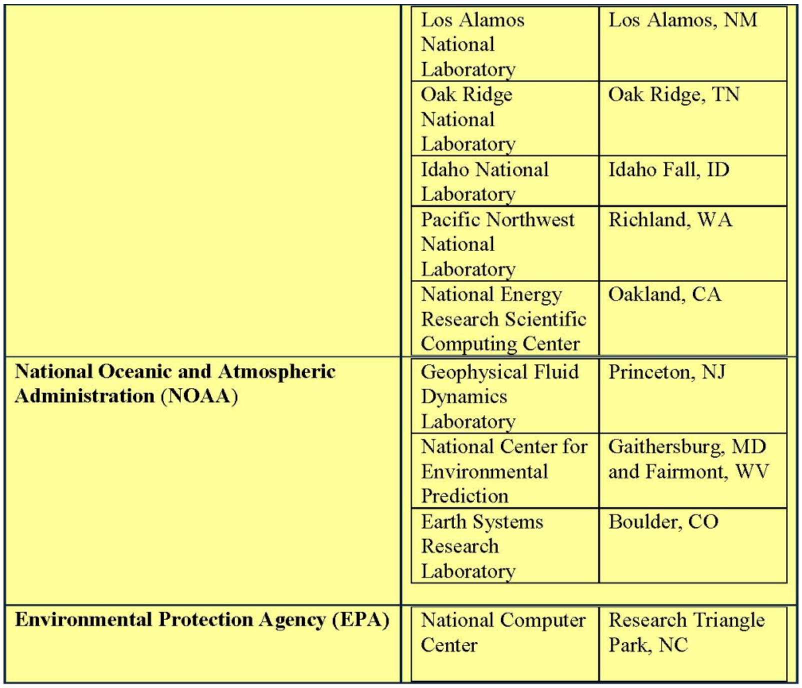
Federal High Performing Computing Locations

A number of federal agencies manage high performance computing facilities in the United States. Table 4 shows a list federal organization and the sites of their HPC facilities. NOAA's facilities are included for completeness.

Table 4: Federal High Performance Computing Facilities



Federal Agency	Site	
National Science Foundation (NSF)	Pittsburg Supercomputing Center	Pittsburg, PA
	University of Illinois - NCSA	Urbana, IL
	San Diego Supercomputing Center	La Jolla, CA
	University of Tennessee	Oak Ridge, TN
Department of Defense (DOD)	U.S. Naval Oceanographic Office	Stennis Space Center, MS
	USAF Aeronautical Systems Center (ASC) Major Shared Resource Center (MSRC)	Dayton, OH
	US Army Research Laboratory	Aberdeen Proving Ground MD
	US Army Corps of Engineers Waterways Experiment Station	Vicksburg, MS
	USAF Maui High Performance Computing Center (MHPCC)	Maui, HI
National Aeronautics and Space Administration (NASA)	Ames Research Laboratory	Mountain View, CA
Department of Energy (DOE)	Livermore National Laboratory	Livermore, CA



	Los Alamos National Laboratory	Los Alamos, NM
	Oak Ridge National Laboratory	Oak Ridge, TN
	Idaho National Laboratory	Idaho Fall, ID
	Pacific Northwest National Laboratory	Richland, WA
	National Energy Research Scientific Computing Center	Oakland, CA
National Oceanic and Atmospheric Administration (NOAA)	Geophysical Fluid Dynamics Laboratory	Princeton, NJ
	National Center for Environmental Prediction	Gaithersburg, MD and Fairmont, WV
	Earth Systems Research Laboratory	Boulder, CO
Environmental Protection Agency (EPA)	National Computer Center	Research Triangle Park, NC

In reviewing the locations for the federal non-NOAA HPC centers listed in Table 4 and in light of the previous sections on electrical power, cost, availability, network communications, natural vulnerabilities, and terrorism, one of the above locations, Oak Ridge National Laboratory (ORNL) in Oak Ridge, TN, has all, or nearly all, the requisite attributes to be an ideal location for a NOAA HPC center.

Oak Ridge National Laboratory

ORNL is a leading center for high performance computing in the United States and the Department of Energy's largest unclassified supercomputing center. The Laboratory has one of the largest supercomputing staffs in the country. ORNL is located outside of Knoxville, Tennessee and is outside the minimum safe distance of any Tier I city. The laboratory is located on the 33,750 acre, Department of Energy Oak Ridge Reservation, allowing a NOAA HPC building to be located there out of public sight and with full government physical security including HSPD-12 physical access control in place, as well as all emergency services. ORNL has been involved in a number of public-private partnerships for constructing leased buildings at the laboratory and has built two HPC facilities using this mechanism.

ORNL has direct access to the GigaPOP in both Atlanta and Chicago. It obtains its electrical power from the Tennessee Valley Authority and has one of the lowest power rates per kWh in the United States. As of May 2008, it was 6 cents per kWh. It is physically about 140 miles from ORNL to the NOAA Data Center in Asheville, North Carolina.

Federal Electric Power Utilities

The Tennessee Valley Authority is the only federally owned public electric power utility in the United States. TVA sells power to 158 local distributors that serve 8.7 million people and 650,000 businesses and industries in the seven-state TVA service area. It covers almost all of Tennessee and parts of Alabama, Georgia, Kentucky, Mississippi, North Carolina, and Virginia. TVA also sells power to 59 large industrial customers and federal installations. TVA's power system consists of a diverse mix of fuel sources, including fossil, nuclear, hydro, and renewables. TVA has 11 coal-fired and eight combustion-turbine plants; three nuclear plants; 29 hydroelectric dams and one pumped storage plant; and 16 solar power sites, one wind power site, and one methane gas site. TVA generates more electricity than any other public utility in the nation. Coal plants typically provide about 60 percent of TVA's power. The TVA transmission system has been 99.999 percent reliable for seven years in a row. TVA customers enjoy some of the lowest retail rates in the nation. As of May 2008, Oak Ridge National Laboratory was buying power for their supercomputers at 6 cents per kWh. TVA continues to expand generation facilities to meet rising demand while keeping rates affordable. In May 2007, TVA restarted Browns Ferry Nuclear Unit 1 making it the first U.S. nuclear unit brought online in the 21st century.

Summary

Of the other federal HPC centers, Oak Ridge National Laboratory is the best fit in minimizing cost and vulnerabilities while meeting requirements for high-speed communication availability. Table 5 compares both of these site as well as the National Climate Center against the criteria.

Table 5: Summary of Criteria for Principal Sites

Criteria	Earth Systems Research Laboratory in Boulder, Colorado	National Climate Data Center in Asheville, North Carolina	Oak Ridge National Laboratory (ORNL), Oak Ridge, Tennessee
Energy Costs			
Electricity Costs	9 cents per kWh	9 cents per kWh	6 cents per kWh
Priority/Reliability of Power	Would need to use natural gas fuel cells to minimize cost. Above 1 petaflop, power levels can exceed the power levels that fuel cells capable of providing.	Highest cost overall for both electricity and natural gas.	Site has transmission lines from four generation stations coming to Laboratory sub-station. Uses Federally owned power utility and has federally mandated priority.
Natural Gas Costs	$10 to $11	$16 to $17	$14 to $15
High Speed Networks Availability	Single Connection gigaPOP in Denver	Single Connection gigaPOP in Charlotte	Dual Connected gigaPOP in Chicago gigaPOP in Atlanta has Access to TVA dark fiber connects across the entire seven states TVA serves.
Vulnerabilities			
Earthquakes	Moderate	Moderate	Moderate
Tornadoes	NO	Low	Low
Lightning	Moderate	Moderate	Moderate
Hurricanes	No	No	No
Volcanoes	No	No	No
Flooding	Low	Low	Low
Wild Fires	Low	Low	Low

Terrorist	Outside Minimum Safe Distance In small city on multi-building campus with perimeter security and security force	Outside Minimum Safe Distance Downtown Urban setting	Outside Minimum Safe Distance Outside of small town on large federal reservation on multi-building campus with perimeter security and security force
Crime	Low	Low	Low
Federal Partnership	No	No	Multiple Federal agencies have HPC facilities at ORNL
Human Resources	Limited	Limited	The largest unclassified HPC facility and staff in the United States. Actively involved in Supercomputer research and design as well HPC software R&D