Grid Accommodation of Dynamic HPC Demand

Grant L. Stewart
LANL
Los Alamos, NM 87545
gstewart@lanl.gov

Gregory A. Koenig
Energy Efficiency HPC Working Group
koenig@acm.org

Jingjing Liu
LBNL
Berkeley, CA
jingjingliu@lbl.gov

Anders Clausen
U of Southern Denmark
Odense, Denmark DK-5230
ancla@mmmi.sdu.dk

Sonja Klingert
U of Mannheim
Mannheim, Germany 68131
klingert@uni-mannheim.de

Natalie Bates
Energy Efficiency HPC Working Group
natalie.jean.bates@gmail.com

ABSTRACT

Sudden and short-duration power swings in modern supercomputers can have a challenging impact on the voltage of the adjacent power grids. The coming age of exascale supercomputers is expected to bring platforms that are capable of power fluctuations of up to 20 MW in 15 milliseconds (ms) or less (one cycle at 60 Hz).

This paper presents a framework that explains these relationships between rapid power flow changes in high performance computing environments (HPC) and power grid characteristics, such as voltage and frequency fluctuations, and offers insights on system improvements that can mitigate large power disturbance issues.

Four case studies of HPC facilities illustrate their unique power-grid setting and identify some of the components that help with the mitigation of large power fluctuations. HPC managers were asked about preparing for larger peak dynamic power demand both on their site and within the broader context of their local electrical grid. The data presented was collected based on questionnaires and interviews with these sites. The responses can provide insight on how to address this growing concern and mitigate the local impact of HPC facilities through grid stiffness or resiliency measures.

CCS CONCEPTS

Hardware-Power & Energy-Energy-Distribution-Power Networks

1 Introduction

The Energy Efficient HPC Working Group, Grid Integration Team became aware of a growing concern that ever-larger, power consuming, supercomputers can affect utility power operations and adjacent utility customers. For example, Oak Ridge National Laboratory's recently installed Summit platform has a peak power demand of 13 MW. Next generation supercomputers are expected to have even higher peak demand of 30 MW or more. Steady state demand of this magnitude is manageable for a utility provider, but they can become concerned when the demand fluctuates rapidly.

Researchers at Leibniz Supercomputing Centre (Huber and Labrenz, 2019) measured SuperMUC, Phase 1 demand fluctuations of up to 1.9 MW and saw corresponding 9% voltage fluctuations during initial testing [1].

Researchers at Los Alamos National Laboratory (LANL)(Backhaus et.al., 2013[2]) studied the power signature of Cielo platform at LANL. Figure 1 shows power fluctuations through four different operational transitions.

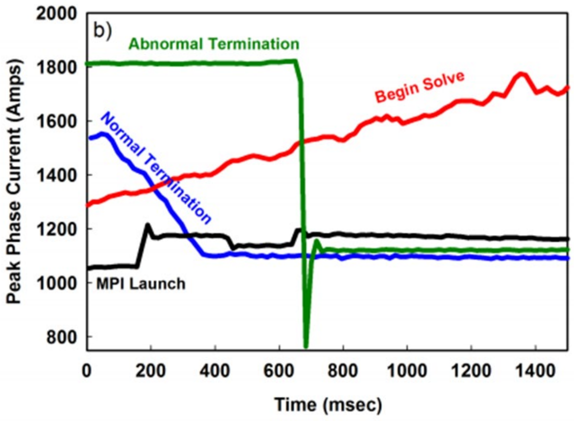


Figure 1. Real power logs and peak phase currents

For example, the "abnormal termination" event in the aforementioned study represented a 1.5 MW power fluctuation in less than one alternating current (AC) cycle or 16.67 msec. The authors measured an approximate 0.4% voltage fluctuation and they anticipated that voltage could fluctuate by as much as 4.0% with a 10 petaflops/sec computer.

HPC power fluctuations can occur over periods longer than 1 cycle, up to several seconds. Power meter capture of the shortest duration power fluctuations requires an advanced programmability for waveform capture triggered by a specified power flow change. Not all meters have this capability, and waveform capture is data intensive due to the short time step measurement required.

LA-UR-19-24333

2 Power Engineering for Supercomputers

HPC demand is highly variable and can challenge power quality on the adjacent grid circuits. To conserve power, newer computers use processor controls that ramp up and down the power demand during the execution of large run cycles. Memory intensive operations, the beginning or end of intensive CPU/GPU operations, system performance testing and cooling/fluid pumping component starts and stops can have highly variable electric demand. HPC power demand associated with motor starts is an inductive load, which can contribute to power quality concerns. Further, the memory testing operations for LANL's Trinity platform required peak-power demand greater than 120% of the supplier's nameplate demand and produced large power demand fluctuations.

With an adequate and robust grid connection and timely communication between the demand side and the utility provider, these demand characteristics may not introduce concerns such as voltage fluctuation for adjacent customers or the neighboring electric grid. As HPC demand grows, it is important to be aware of concerns that may arise and mitigation measures available to address them.

2.1 Grid Stability, Reliability & Stiffness

In North America, the North American Electric Reliability Corporation (NERC) is responsible for bulk power system (BPS) reliability and stability. In response to replacement of traditional synchronous generation fleet with distributed renewable generation, concerns have arisen for voltage stability [3]. The potential impacts to the BPS of rapidly growing supercomputer demand can be similar if they grow large enough. Grid stiffness or the systems short-circuit strength is a measure of the grids ability to withstand and rapidly recover from dynamic power fluctuation events.

2.2 Basics of Voltage Fluctuations

In an AC power system, resistance to current flow is impedance. Impedance includes resistance, and inductive and capacitive reactance. Real power, measured in megawatts (MWs), plus reactive power, measured in volt-ampere reactive (VARs), is apparent power, measured in mega-volt amperes (MVA). Resistance and the opposition to current flow caused by capacitors and inductors control real power flow.

- Resistance is a function of length, size and material of conductors. In a typical high-voltage AC circuit, the impact of resistance is small when compared to the impact of reactance. Lower voltage AC circuits experience more resistance because current flow is higher.
- Reactance causes current to be out-of-phase with voltage
  - Inductive reactance causes current to lag voltage, and is generated with the start of inductive motor loads
  - Capacitive reactance causes current to lead the voltage, and is generated more in underground circuits where cables are close together
  - Lagging or leading reactance is measured in VARs
  - VARs flow on a circuit supplies magnetizing current for inductive loads and charging current for capacitive loads, and the supply of VARs provides power-factor correction.

Power factor correction in a steady state circuit is achieved when the reactance from an inductive load (absorbing reactive power) is balanced with reactance from a capacitor (generating reactive power).

When power flow changes rapidly (fault, large motor start, imbalanced/automatic switching operation, or supercomputer rapid ramp-up or ramp-down), voltage can droop or spike. The magnitude of voltage fluctuation is greater on circuits with a low power factor. Devices sending power require adjustments to the changed impedance. These adjustments might include tap changer position on a power transformer, automatic or manual switching of capacitor banks, or exciter output on a synchronous generator controller.

Voltage droops or spikes in a power system can be an initiator of instability and typically considered disruptive to affected utility customers.

The following devices in a power grid can mitigate voltage fluctuations and improve stability [4].

- Generator with automatic voltage regulation (AVR)
- Static VAR Compensator (SVC)
- Static Synchronous Compensator (STATCOM)
- Regulating Transformer (automatic tap changing)
- Switched Shunt Capacitor

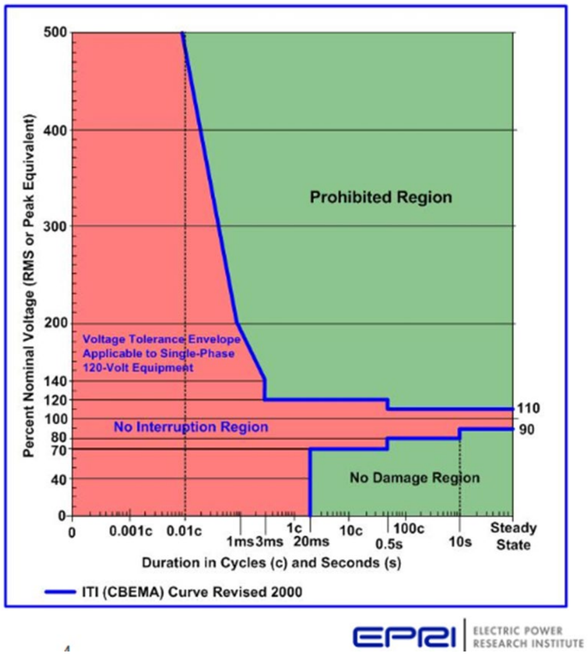
These devices have unique performance characteristics in a power supply system, and can be deployed in various combinations to provide the required voltage stability. Like tools and a tool chest, they all have a useful function. Utility power flow studies and modeling is required to anticipate HPC-driven voltage instability and optimally deploy voltage-stabilizing devices.

2.3 Dynamic Load Conditions

A large fault on a major transmission line is often followed by sectionalizer and recloser operations that cause power flow excursions for seconds or minutes after the initial fault. This is similar to the displacement of a spring system and the oscillations that follow until the system returns to equilibrium. Voltage fluctuations in a "stiff" grid will rapidly dampen and return to equilibrium. With a less-stiff grid, the same magnitude of disturbance can radiate outwardly through the grid and returning to a stable voltage can take longer. Similarly, a large and rapid demand change from a supercomputer in a less-stiff grid setting can cause voltage fluctuations to propagate outwardly on a distribution circuit.

Exascale computers are expected to have 30 to 40 MW peak demand and up to 20 MW of rapid dynamic load change over a period-of-time between 15 ms and several seconds.

Voltage disturbances can cause electrical contacts to open, relays to trip, damage to controllers, etc. Further, a supercomputer's intentional demand increase or decrease that causes a voltage fluctuation can have the unintentional cascading impact of shutting down other equipment in the circuit or adjacent circuits. The Information Technology Industry Council (ITIC) Curve below presented by Electric Power Research Institute (EPRI) indicates the magnitude and duration of disturbances that can cause harm in low-voltage systems.



3 Case Studies

Traction power networks for electrically-powered passenger trains (e.g., Bay Area Rapid Transit, Sound Transit, and Massachusetts Bay Transportation Authority) [5] and industrial arc furnaces for metals production are examples of large dynamic loads that have historically been served by power grids. Power fluctuations of more than 30 MW in less than 10 seconds are common. These systems have had to address voltage transient conditions similar to those potentially associated with next generation supercomputers. We can build empirically on successes from electric trains and arc furnaces as well as the grid context of existing HPC centers.

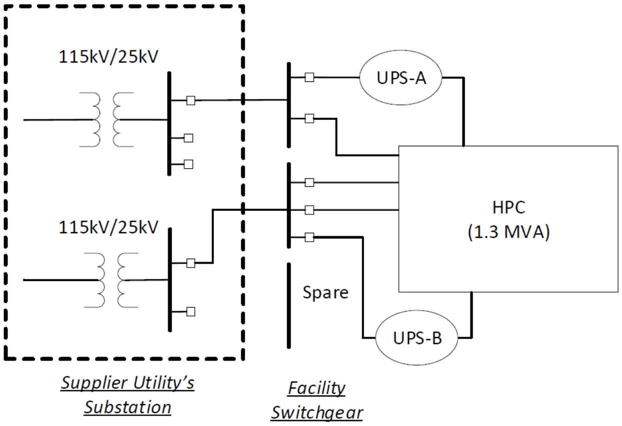
Four United States based HPC centers were examined as of this writing, and several more are scheduled to be in 2019 for this research. They are not named here to honor confidentiality.

Each case study consisted of a written questionnaire, follow up interview, and grid context review. Our inquiry for each case study included advanced metering, example power flow data for largest current machine, interactions with and power scheduling requirements of the utility service provider, dynamic power flow modeling, and plans for growth and development including power system changes to address power dynamics concerns. Not all of the information gleaned from the interviews is synthesized below.

Each site has a power system configuration that uniquely addresses grid stiffness and adaptations that both accommodate growing dynamic load and mitigate voltage disturbances.

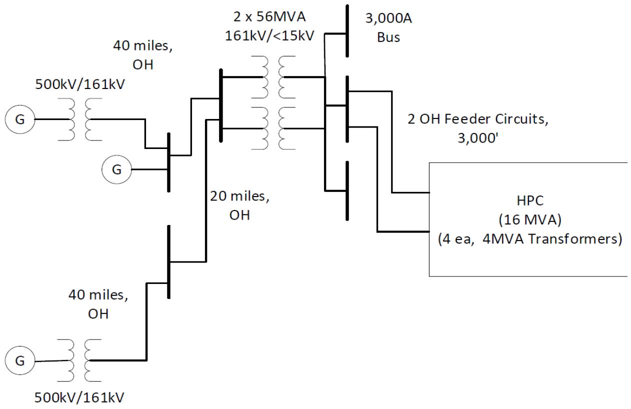
3.1 HPC-A (developing)

Grid setting: The power supply includes two 25 kV distribution circuits (less than one mile long). The HPC-A has access to two 115 kV/25 kV substations. It has both a primary and backup UPS that provide both emergency power and power quality support. This facility has an adjacent business with a 100-180MW array of gas turbine generators with automatic voltage regulation. HPC-A currently has a 1.3-MW peak demand with 1.0 MW fluctuations, and planned growth to a 14-MW peak power demand within 10 years. This HPC is smaller and developing; has not yet experienced issues with power dynamics, and is not required to schedule or limit demand with their utility service provider.



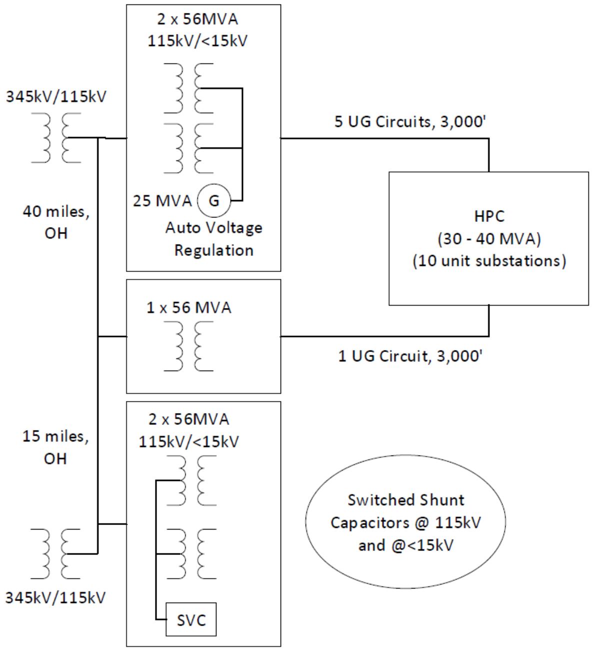
3.2 HPC-B (most stiff)

Grid setting: HPC-B has access to large generation facilities and two 500kV/161kV substations within 40-60 miles and three 161kV/13.8kV substations (< one mile from HPC). HPC-B has an available fault current at secondary transformers of 62,000 Amps. Series rated fuses address excessive fault current at the computer racks. Planning for HPC with 90MW peak demand, 30-40MW power fluctuation within 10 years, and two large machines operating (one operating, one in startup/testing). The utility service provider has so much available stiffness and capacity that the planned growth does not require major improvements or power flow restriction.



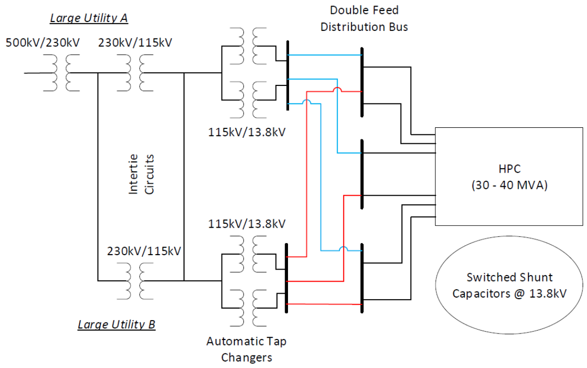
3.3 HPC-C (many VARs support components)

Grid setting: HPC-C has access to two 345kV/115kV substations within 15-40 miles, four on-site 115kV/13.8kV substations with automatic-tap-changers and switched shunt capacitors, an on-site 25MW generator with automatic-voltage-regulation, six dedicated underground feeder circuits, and a 100 MVar Static VAR Compensator. Planning for HPC with 60MW peak demand, up to 20MW power fluctuations, and two large machines operating (one operating, one in startup/testing). This HPC has several rotating UPS systems for power quality. The site is a transmission utility with a merchant-desk function for power scheduling.



3.4 HPC-D (high reliability & stiffness)

Grid setting: HPC-D has access to one 500kV/230kV/115kV substation with automatic tap changers for voltage regulation and one 230KV/115kV substation within 10 miles, served by two major utilities, and has double fed switchgear for reliability and dedicated distribution feeders. Planning for HPC with 65MW peak demand, up to 70% ramp down (25MW) in a fraction of a second, and two large machines operating (one operating, one in startup/testing). One of the serving utilities requires detailed power forecasting.



4 Conclusion

The power supply infrastructure necessary for large HPCs is expensive and takes a long time to development. In general, power infrastructure improvements must lead HPC systems by 5 to 10 years and they can be comparable in cost to the computing platform. Each site has a unique grid setting that is an artifact of its development over the years. Optimized planning for a supercomputer and its supporting power grid requires a clear understanding of grid stiffness characteristics and the magnitude and ramp rate of dynamic power fluctuations. An awareness through case studies of the challenges and adaptations planned in operating HPCs can hasten the development process for others. This research has just begun.

ACKNOWLEDGMENTS

Contributing EE HPC WG Team Members include Parks Fields, Michael Kercher, Aaron Anderson, Gary New, Jim Rogers, Rick Griffin, Anna Maria Bailey, Trent D'Hooge, Alynna Montoya-Wiuff, Patrick Segura and Eloy Romero.

REFERENCES

[1] Power Transients - SuperMUC Extreme Scaling Periods, Herbert Huber , Detlef Labrenz, EEHPC Workshop Presentation, Dallas 2019

[2] HPC Power Transient Testing, Scott Backhaus, Josip Loncaric, Cornell Wright, and Maura Miller, Los Alamos National Laboratory, 2013, LA-UR-13-21919.

[3] Short-Circuit Modeling and System Strength, White Paper, North American Electric Reliability Corporation NERC, February 2018

[4] Comparison of Different Types of Compensating Devices in Power Systems, Arun Pundir, et.al., International Research Journal of Engineering and Technology, November 2016 issue

[5] Analysis of Transit AC and DC Traction Power Networks, Richard N. Rauceo, et.al., American Public Transportation Association, 2014 Rail Conference Proceedings, Montreal, QC, Canada