Research Trends In High Performance Computing Application On Large Scale Power System Operation

Ranjit Amgai, Jian Shi, Sherif Abdelwahed, Yong Fu
Mississippi State University
ra372@msstate.edu

Keywords: High performance computing, Smart grid, Distributed Algorithms, Power systems

Abstract
Recent trends in smart grid technology and fast switching devices have escalated the need for utilizing high performance computing (HPC) in power system applications. Likewise, with the distributed generation and added intelligence, number of control variables are increasing that requires real-time solution. The need for comprehensive and detailed simulation of such complex interconnected system is of utmost importance for operation, control and monitoring actions. Research works relating to load flow, state estimation, contingency analysis, stability issues, unit commitment and cyber-security issues are discussed with focus in faster computation for efficient decision making. Considering the various aspects over reliable and sustainable power system, this work surveys current progression and suggests the future potentials of HPC application in power system.

1. INTRODUCTION
Inclusion of sophisticated controls and high speed switching devices along with the current trend in smart grid requires in-depth analysis of the underlying system dynamics. The computational demand for the modern integrated power system has increased in proportion to the escalated technology boost. Planning, simulation, optimization and real time controls demand high speed complex computation [1]. Emerging technologies in multiple architecture grid expansion, distributed generation, cyber-security factors, and data handling and visualization have added control and planning computation challenges. Utilizing high performance computing has been recognized as a potential solution since few decades back [2]. The power system has added significant intelligent functions since the analysis of its application and trends has been reported [3][4]. Analysis of the evolving HPC trends to specific application in whole power system delivers the insight of the potential research aspects in this field. Basic HPC architecture is shown in figure 1.

Comprehensive and data intensive real time simulation of the power system helps to provide the required operation and control signals. Supervisory control and data acquisition (SCADA) system, and more recently phasor measurement Unit (PMU) provide real time data for simulation. Over recent years there has been increased interest in the development of PMU based application development for monitoring power system dynamics [5]. The decision is then taken for specified objectives including dynamic state estimation, contingency analysis, economic dispatch, unit commitment, automatic generation control, and relay control. Complex mathematical equations and network theories need to be solved for dynamic study, consuming a significant amount of time. The demand of dynamic real time state estimation requires solution of differential algebraic equations (DAE). The solution of these problems by single processor computer generally takes order of minutes to simulate some second’s dynamics of large scale power system [6]. Appropriate tools to the above problem have been presented in HPC domain with the recent extension to grid computing and Graphics Processing unit(GPU) computing [7][8]. Massive interconnected networks are decomposed into modules and the parallel algorithms helps in concurrent execution to render the prompt results. Resource and work partitioning, parallel access control, system component interaction and memory management are some of the challenges [9] that designer has to face in utilizing HPC power system applications. Over the past decades, development in the HPC is witnessed towards faster processor speed, bigger memory and use of multiple processors to solve a problem. The utilization of HPC to the power grid application has benefited in computing sensor data for real-time grid monitoring and operation, data-driven models, visual analytic tools, stochastic real time simulations and secure sensor network infrastructure among many.
2. HARDWARE TRENDS

Moore’s law for governing the hardware performance trends has been reported [10] no more valid due to the thermal capacity limitation of the current chip technology. Empirical formulation of “Koomey’s law” [11] has defined the performance trends in terms of energy cost. Commercially available clock speed for desktop computers is around 3.4 GHz, whereas more focus is being given in deeper pipelining and cache management. Performance metrics has changed from single processor frequency measure to various HPC trends including multi-core computing, grid computing, GPU computing, and cloud computing.

The large scale HPC platform developed from Giga flop in 1980s, through Teraflop in 1994 to Petaflop in 2009. The architecture is expected to change drastically to reach the goal of Exaflop with the great challenge in power, concurrency and memory capacity [12]. InfiniBand up to 56Gb/s is already in the market [13] to address the future communication bottleneck challenges involved in HPC.

Accelerating power system algorithms requires tight coupling between the explicit parallel algorithms with the trending hardware platform. Shared memory access is still the great challenge in multi-core processor architecture. Shared memory architecture with hardware multithreading has been implemented recently in scalable shared memory supercomputing platform architecture like Cray XMT [14], cray MTA or Sun Niagara 2.

Grid computing has the potential to solve the complex problems with the use of geographically distributed resources. Dynamic communication is achieved from its open source tools in a secure and reliable way. Coordinated operation of the grid helps for parallel processing, virtual organization for collaboration, load balancing and management, and use of idle resources [7]. Similarly, several power system applications could have potential benefits to foster smart grid concepts. Grid computing concept is evolving with several projects such as NSF TeraGrid, EGEE, open science grid, and XSEDE.

With the recent release of CUDA™ 4.1, GPU computing has gained momentum for high performance in smaller scale. GPU are faster, cheaper, and more power efficient as compared to corresponding multi-core CPU for bulk concurrent data calculation. GPU computation is trending in many power system applications [15][16][8] and is likely to be closely integrated with smart grid concept. The hardware architecture is developing towards the combination of multi-core CPU along with GPU. Some works regarding cloud computing in power system application are seen using the Google app engine and azure platforms [17]. Features including ease of software use, real time interaction, and no added hardware costs are likely to boost the cloud computing future in smart grid technology.

3. HPC APPLICATIONS IN POWER SYSTEM

Some relevant aspects of the power system applications are analysed here with respect to utilization of HPC. The trending technology up to the end of year 2011 is discussed in terms of HPC technology used, computation time, speed-up, efficiency and other possible factors.

3.1. Load Flow (LF)

Load flow is the basic tool that provides steady state characteristics of the overall power system. Load flow computation needs to accelerated as several power system applications depends upon the load flow solution. After the concept of parallel algorithms for LF solution came in early 90’s [18] significant progress has been made. Decomposition method [19] for 43 bus system with 4 processors took 15 milliseconds [20] and splitting the work in four subsystems for steady state analysis. Several works [21][22] reported of using mapping radial distribution network into tree structure of the distribution system using forward/backward sweep. Work on [21] implemented 528 bus systems in 600 milliseconds with 13 parallel processors.

Well balanced factorization tree technique was implemented on [23] over 288 bus system. Implementation of 6 individual 32 bit transputers each with 4 Mb of cache took 255.36 ms with a cluster size 7. Speed-up near to 4 was demonstrated with the given hardware. With the boost in the hardware technology significant research interest was shown for distributed algorithms for large power system studies. LU decomposition and forward/backward substitution of Jacobean matrix or Y admittance matrix is no more required for parallel operation [24]. Field Programming Gate Array(FPGA) implementation of the optimal load flow was reported on [25][26], where first work implemented 7917 bus system to complete in 0.69 s using Jacobi method.

The use of cluster workstations further accelerated the power flow computation. Work on [27] used the cluster of 8 processors each of 1 GHz and 512 Mb of RAM each. Ethernet connection of 1 Gbit/s connects the workstation and runs on Linux using MPICH [28]. Parallel power flow solution to Newton’s method is claimed with speed-up of 4.22 for 39136 no of branches [27].

Convergence, computation time, infeasibility of the problem, decoupling, and efficiency are major attributes for comparison. Research work [29][30] has also focused on the efficient computation of loads flow on commonly available hardware platforms such as Intel quad core. Decoupled Newton Raphson three phase power flow model on [30] gave the speed-up of 1.08. Artificial Intelligence implementation has also been reported to solve partitional clustering power flow model with better convergence [31]. Recent work on [32][15] reported the use of GPU in CUDA architecture. GPU computing processor
with 128 processing units along with the 2.83 GHz host PC with 2G RAM is used in [15] for parallel power flow solutions with speed-up of 2.1. Some discrepancies are observed due to the data transfer issues between GPU and host PC.

### 3.2. Contingency Analysis (CA)

Depending upon the measurements obtained from the SCADA system, CA helps to assess the possible failures of the power system. Various credible contingencies along with the load flow need to be computed for post contingency analysis. To develop parallel model for contingency is not as complex as load flow. However, closer look needs to be done for computational load balancing. Necessary control actions depend upon the feasible contingency analysis. Reduction of the contingency list is one alternative [33] since long due to data intensive computation demand. Work on [34] divided the contingency into larger number of screening list and less number of full analysis list depending upon the ranks provided. The above work took 170.32 s for 1340 screening test and 100 full contingency test of 810 Buses using single program multiple data (SPMD) model. This work was later validated using MPI and PVM with better results in [35]. Performance analysis on work [36] was validated using 64 nodes with 512 Mb of shared memory, and 8 dual quad core nodes running at 2.33 GHz connected with InfiniBand. N-x contingency analysis is proposed here exploiting the multithreaded architecture with hybrid computing reducing the data transfer overhead. Various test cases are demonstrated ranging from 200 to 1700 selected contingencies for both dynamic and static load balancing. The current research [36][10][37] has demonstrated significant deduction in computation time using HPC platform.

Work on [38] has leveraged the use of GPU via the CUDA interface for contingency analysis in 300 bus system obtaining a speed factor of 4. The experiment was supported by dual quad core processors with 2048 nodes, where each node has a 3 GHz Intel processor with 16 GB RAM and 250 GB hard disk.

Apart from the grid computing, cloud computing in power system analysis is gaining its momentum in recent years. The work on [17] used InterPSS cloud edition over the Google App Engine (GAE) to provide the features including load flow, contingency analysis and open data model.

### 3.3. Simulation

Dynamic simulation of large scale power system has always been a computational challenge due to the need to capture the transient and sub transient dynamics in real time. Various abstractions have to be made to capture the dynamics of fast acting devices, reducing its fidelity. Optimal Control algorithms [39] for fast acting devices bring even more challenges. However, HPC is thriving to meet the real time computational requirement for grid operation. Distributed simulation is gaining momentum as the current computing trend is making its way towards grid and cloud computing. The performance of GPU based simulation [8] and development of the commercially available real time simulators for new device testing, protection, control, and operator training are stepping towards new era.

Real time simulation has evolved to assess the real time closed-loop testing for large scale power system study. Hardware in the loop study [40][41] has helped to understand the behaviour of the new device and validate the controller action. Commercially available simulators such as RTDS and OpalRT have reduced the simulation steps up to 50 us. Those simulators are generally based on of-the-shelf processors for processing and FPGA for input/output interaction. With the development of the new power system decomposition technologies hardware cores can be used to its optimum. Work on [42] has demonstrated the initial development of custom solver to accelerate the simulation of power system studies and is compared to the Simulink for its efficiency. Speed-up between 20x and 80x is reported for shipboard power system. The authors in [43] implemented Multi-Area Thévenin Equivalent algorithm (MATE) in 16 AMD Athlon™ 2.5 GHz PC units. Comparison with SuperLU DIST algorithm to solve the 14,327 bus WECC system gave up to 26 times faster solution.

Development of the average model and proper abstraction of the model specified to the particular objective helps to boost the simulation. Data parallelism and multi-threading has been exploited [16] to run the simulation on multiple GPU for scaling and high performance. Custom built solver application package [42] was reported to perform better as compared to commercially available tools in terms of CPU utilization.

### 3.4. State Estimation (SE)

State Estimation is one of the major function of Energy Management System (EMS) for real time monitoring and control of power system. Great effort is being made to perform SE near to the SCADA measurement cycles. This practice will have a vital role for monitoring and control action near collapse region when the system state is rapidly changing. HPC is being exploited to generate tools and techniques towards near-to-real time system state estimation [44] integrating with the information and communication technologies. Due to the decoupled nature of the state estimation problem multiple algorithms exist since long [2]. HPC methods are evolving, which includes weighted least square error methods, Conjugate Gradient algorithms to extended Kalman filtering approach. The PSO technique implemented [45] in 118 bus system is tested with the 8 processor system with consistent result reported speed-up of 4.8. The work on [46] used SuperLU [47] package on 128 1.5 GHz processors with 256 GB memory shared memory system which includes hardware multi-threading. Various comparisons gave the speed-up of 10, advocating more research towards HPC potentials on state estimation. Other
advancement in GPU and cluster implementation in solution of the complex mathematical vector computation [47][48] has also accelerated the power system application development such as SE, CA, LF among many. SE has also caught the pace of recent computation trends towards grid and cloud computing [49][50][17] for better vision of power system states.

3.5. Stability

GPU computation and grid computation have addressed the granularities involved in transient stability simulation [16][51] to great extent. Work on [16] utilized instantaneous relaxation based parallel processing technique to run with Tesla S1070 server equipped with 4 independent T10 GPUs each with 4GB of memory. Block diagonal structure of the Jacobian matrix is exploited to obtain 2-10 speed-ups for 9984 bus system as compared to sparse CPU-based solver. Frequency domain parallel eigenvalue search algorithm is implemented in [52] to deal with the small signal stability problem. This work achieved up to 2.32 speed-up and is claimed for high parallel efficiency. New parallel algorithms are developing with time to fully exploit available hardware potential. Butcher transformation is used in [53] along with the Runge-Kutta-Nyström method for transient stability simulation. Convergence and efficiency is compared with the conventional parallel-in- time Newton approach.

3.6. Unit Commitment (UC)

Present dynamic deregulated power market is incorporating more and more number of generating units, thus increasing the computation burden. Randomness included in the demand and with participation of renewable energy resources, unit commitment (UC) requires stochastic algorithms to be solved by high performance computer. Implementation of HPC techniques for forecasting, rapid information exchange, bidding and optimization is the need of the current trending power market. The work on [54] pursued MPI based parallel artificial bee colony (PABC) approach for UC problem involving 1000 generating units. Different HPC platforms are used to compare the parallel performance from shared memory and distributed memory models. Quality of solution and improvement of minute by minute dispatch was reported. A parallel UC algorithm considering AC optimal power flow (OPF) in [55] obtained acceleration up to 3.77. Test was done for 118 and 300 bus systems with six 1.7 GHz processors, 512 Mb of memory and 100 Mb/s network adaptor. Recently, GPU has found its application in economic dispatch problem using comprehensive learning particle swarm optimization (CLPSO) [56].

3.7. Cyber-Security

Technological complexity and interconnection between the computers, communications and the power grids have exponentially increased the cyber-security threat level in power system operation. Monitoring and detection of such security vulnerabilities is one challenge followed by the necessary control measures to be implanted. RICA [57], Monte-Carlo and model-based [58] algorithms provide security measures to help inform the cyber-security issues. Implementation of various statistical assessment algorithms for cyber-security on large system demands the need for HPC. The recent work reported on [59] the large amount of data is computed on the 180-teraflop Red Mesa6 HPC platform to simulate the cyber-attacks against generator, line, and bus protection as well as SCADA.

4. DISTRIBUTED PROCESSING AND DECOMPOSITION STRATEGIES

Cooperative utilization of multiple hardware components to perform a complex computation task requires well defined decomposition strategies [60][61]. Performance is optimized by implementing parallelism on different application levels of a device, such as architecture-level parallelism, core-level parallelism, thread/process-level parallelism or instruction-level parallelism. For thread-level parallelism, the most commonly used applications are SMT (simultaneous Multithreading) and CMP (Chip multiprocessor) [62][63].

SMT processors hold states of multiple threads in hardware, and only allows the executions of certain threads per cycle based on given instructions. The single processor could switch instantaneously from one thread to the other threads, thus making the overall process efficient. CMP processors on the other hand consist of a number of single processor cores on one chip, and every core can be utilized simultaneously by different processes or threads. With the support of parallel running of multiple application instances, the overall execution process time could be greatly minimized compared with traditional processors.

Within power system simulation, parallelism can normally be realized via methods including space-domain parallelism, time-domain parallelism, and time-space domain parallelism. Considering the complexity of reconstructing programs for parallelism and flexibility of each method, space-domain parallelism is more viable solution in platform with multi-core or multi-threaded processors. W-matrix algorithm [64] and BBDF algorithm [65][66] are two common efficient and widely used techniques in space-domain parallelism.

4.1. W matrix Algorithm

Power system problems often require the solution of sparse sets of linear equations of the form:

\[ \mathbf{A} \cdot \mathbf{x} = \mathbf{b} \]  

The matrix A can be factored as \( \mathbf{A} = \mathbf{L} \cdot \mathbf{D} \cdot \mathbf{U} \), which is also known as LDU decomposition, where D is a diagonal matrix.
and $L, U$ are unit triangular matrices with all the elements on the diagonals one. The solution to this problem includes three steps: forward substitution, diagonal scaling and back substitution. The computation of $L$ and $U$ is calculated by ordering of the rows and columns of $A$ to minimize the number of new non-zero terms. Define:

$$W_l = L^{-1} \quad W_u = U^{-1}$$  \hspace{1cm} (2)

$W_l$ is lower triangular while $W_u$ is upper triangular.

Then the solution to the original problem can be derived as:

$$x = W_u \cdot D^{-1} \cdot W_l \cdot b$$  \hspace{1cm} (3)

The matrix $L$ can be expressed as:

$$L = L_1 \cdot L_2 \ldots L_n$$  \hspace{1cm} (4)

Where each $L_i$ is an identity matrix except the $i$-th column. Then,

$$W_l = L_1^{-1} \cdots L_2^{-1} \cdot L_i^{-1} = W_l^n \cdots W_l^1$$  \hspace{1cm} (5)

where $W_l^j$ is $L_j$ with signs reversed on the off-diagonal elements.

Using this expanded form of $W$ in equation (3), the solution now can be expressed as:

$$x = W_l^1 \cdot W_l^2 \cdots W_l^n \cdot D^{-1} \cdot W_l^1 \cdots W_l^2 \cdot W_l^1 \cdot b$$  \hspace{1cm} (6)

All multiplications within each step are amenable to parallel processing.

4.2. BBDF Algorithm

Any large power system can be considered as comprising of a number of subsystems connected via networks. The partitioning of each subsystem can be based on either the regional characteristics of subsystems, or geography division [67]. However, simple implementation of each method would lead to more communication cost or load imbalance. Partitioning scheme utilizing hierarchical block bordered diagonal form of power network was proposed to greatly accelerate the numerical calculations. The strategy of dispatching computation tasks to multiple processors in parallel, focuses on the separation of the power system network and devices into several individual subsystems. Two principles should be considered simultaneously when partitioning the network [66]: 1) Minimization of the connections between different sub-systems. 2) Maintaining an equal computational load for every sub-system.

In a typical power system, the system dynamics can be expressed as

$$Y = AX$$  \hspace{1cm} (7)

where $A$ is the transmission matrix while $X$ and $Y$ are input and output system state vectors. In order to decompose the large system connected via networks, it is required that elements in each individual sub-system are interconnected strongly while at the same time, different sub-systems interconnections (also known as cut nodes) are loosely coupled [68].

In Fig 2 [68], the power system network admittance matrix is reformed as the Bordered Block Diagonal Form (BBDF), the border blocks ($C_1$ to $C_3$ and $B_1$ to $B_3$) connect the diagonal blocks ($A_1$ to $A_3$) with cut node block $A_g$, in this way the diagonal blocks are only coupled with each other via the cut node block loosely. The size of the cut node block has to be as small as possible to reduce the amount of information changing among each sub-system, and the size of each sub-system should be approximately the same to minimize the load unbalance. After the decomposition, the calculation of each sub-system shall be distributed to multi-processor computing devices. With limited interconnections between each processor, the computing efficiency could be greatly optimized.

5. ROADMAP FOR HPC IN POWER SYSTEMS

Emerging smart grid technologies are adding the number of control variables with more stringent time margin for computation. Development of the efficient near-to-real time operational tools to the demand of economic dispatch, optimal PF, stability, cyber-security, SE and CA analysis are the potential research path in HPC. The increasing mathematical complexity is the challenge faced by the HPC, commensurate with the increasing nodes, distributed generation, and integration of renewable...
energy. The proactive prediction towards self-healing, self-configuring and self-optimizing tools will be the answer for 21’st century demand in Energy Management System.

Power system application tools will be tightly coupled with HPC. It will follow the increasing trend of hardware performance along with the development of the parallel/distributed architecture and decoupled algorithms. The path for HPC will be towards Multi-core computation, Grid computation, GPU computation or Cloud computation determined by the custom goal.

Multi-core computation is the lowest level of HPC platform that an user can experience with the programming model like openMP. With the commercial availability of 8 core processors, single chip multi-core processor will be soon successfully exploited with multi level programming models. This approach could take the computing to the next level where desktop machine could handle massive computation capabilities without clusters. With the introduction of the CUDA, GPU computing has accelerated its pace since couple of years back in power system applications [8][32][38][15]. Large matrix handling capability, mapping with SIMD model is going to take the GPU computing deep into HPC market for years to come. Almost all of the attributes reviewed above facilitates from the GPU technology. Parallel core GPU, working in tandem with multi-core CPU development will soon be seen.

Emerging smart grid technology concept will be blended with the grid computing paradigm in the near future. Inclusion of the real time vision of the power system is possible with the grid computation for all kind of possible applications. More efficient decoupling algorithms will support the grid mission. Grid will soon be able to answer the challenges such as communication overhead and load balancing, whose effect can be seen in the application such as contingency analysis [33].

Cloud computing is in its preliminary stage with very handful of amount of work [1][17][69]. This technology will go far ahead supporting the smart grid concept. Standing above the grid computing cloud computing could reach to the end user to serve as the particular application. Scalability, security, reliability, interoperability and cost effectiveness will be fully exploited with this technology. In addition, privacy, cybersecurity, liability and billing issues are the new emerging challenges that needs to be addressed. Each house in the grid will have a necessary interface right in the personal computer. Hand-held devices will be soon be used for monitoring and control actions.

6. CONCLUSION

Over the past decades, tremendous development in the HPC is witnessed towards faster processor speed, bigger memory and use of multiple processors to solve a problem. The implementation of HPC to the power grid has benefited in computing sensor data for real-time grid monitoring and operation, data-driven models, visual analytic tools, stochastic real time simulations and secure sensor network infrastructure among many.

A wide scope of subjects covering some of the emerging technologies for improving the dynamic performance of electrical power systems in the new era have been analysed and reviewed in this paper. Knowledge of these trends will be beneficial for developing future power system applications. For the future Smart Grid power system, HPC clusters will play an important role in the real-time energy dispatch determinations and optimizations.

REFERENCES


