Cabot Partners Group, Inc. 100 Woodcrest Lane, Danbury CT 06810, www.cabotpartners.com

WHITE PAPER

A Total Cost of Ownership Study (TCO) Comparing the IBM Blue Gene/P with Other Cluster Systems for High Performance Computing

Sponsored by IBM

Srini Chari, Ph.D., MBA

November, 2008

mailto:chari@cabotpartners.com

Cabot Partners

Optimizing Innovation

Introduction

This paper is a realistic TCO analysis of the fastest supercomputers today, in a bid to support well-informed business and financial decisions when evaluating and deciding on various supercomputing systems available in the marketplace.

High performance computing (HPC) has penetrated beyond government and academic research into aerospace, automotive, petroleum, financial services, life sciences and nuclear industries with immense commercial gains for the these industries. Innovations resulting from competitive pressures and the insatiable need to process mammoth data into actionable knowledge and insights have fuelled the widespread adoption of supercomputers across these diverse industries. These in turn have fuelled the invention and development of faster and more efficient supercomputers for HPC systems and solutions across these industries.

The market is concentrated with a few major players such as HP, SUN, DELL, and IBM. There are other participants such as Cray, SGI, Verari, and many more with much less market share but important in certain segments. The offerings span a broad range of architectures ranging from x86 clusters to ultrascalable systems such as IBM's Blue Gene/P. IBM leads the market in very large scale supercomputing offerings for unprecedented scalability, portfolio breadth, increased energy-efficiency, capital costs, compute density and operational efficiency.

The major attribute of a supercomputer is its processing prowess usually measured in total or peak teraflops. However, in recent years and with escalating energy costs, the metric of gigaflops/watt has become increasingly important. The performance of the highest ranking supercomputer in the Top 500 supercomputer list went from around 140 teraflops in 2005 to an impressive and previously unimaginable 1 petaflop in 2008. And today, the horizon of multi-petaflop computing poses the question: what are the true trends in the total costs that are incurred when making investments in supercomputers and what is the value derived as we move forward?

Several factors in addition to acquisition costs drive the total cost of ownership (TCO). These include significant and generic factors, viz. energy costs, floor space costs, infrastructure hardware costs, and people costs. Other factors such as costs for applications porting and migration, retraining of the IT staff, software licensing, unplanned outages and solution deployment are also typically incurred.

In this paper we evaluate the Blue Gene/P solution from IBM with other comparable HPC platforms from the perspective of Total Cost of Ownership over a three-year period. In this study, to maintain relevance and objectivity across diverse industry scenarios, we consider only the generic factors: energy costs, floor space and data center costs, hardware acquisition costs, and people costs.

The platform differentiators for the Blue Gene/P are its higher energy efficiency, unparalleled scalability and smaller footprint (high package density), coupled with a standardized parallel programming model and software tools that permit the migration of a broad number of HPC applications with minimal parallel algorithmic invention. However, relative to standardized commodity x86 Cluster architectures which are the mainstream platforms for HPC, the Blue Gene/P needs continuing focused investments in application migration and optimization. IBM and a broad range of application partners continue to make these investments, and a repertoire of optimized HPC applications both ISV and in-house proprietary applications continues to grow. The Blue Gene has been used in many industrial applications ranging from financial services to the life sciences delivering unsurpassed performance, scalability, energy-efficiency, and substantial reduction in data-center footprint and costs.

Trends and Challenges in HPC

As HPC continues to grow faster than the overall server industry, HPC solution providers have geared to engineer breakthroughs in performance, scalability, price/performance, space and energy costs, and software for management and applications enablement.

The anticipated energy costs in large HPC data centers in order to power and cool hardware infrastructures are likely to increase more rapidly during the next decade unless economical approaches to energy production are developed in the near-term. The IT industry is defining additional metrics such as gigaflops/watt that rate the HPC solution providers. The Top Green 500 list[1] is becoming as important as the Top 500[2] list of supercomputers as HPC solution providers compete for bragging rights.

In recent years, HPC solution providers have made significant innovations in "green" and next generation data centers to reduce the Total Cost of Ownership, reflecting not just capital costs but operational and maintenance costs as well. More than an economic pain point or a social responsibility, bringing sound environmental principles to bear in operating a data center can become a competitive advantage and a source of operational stability and increased reliability. It is from this perspective that we compare prominent HPC systems available today.

HPC Systems Analyzed

This study evaluates the IBM Blue Gene/P with comparable systems from Sun and x86 based Clusters over a three-year period. The study estimates the Total Cost of Ownership of four commercial high performance computing systems- the Intel® based dual core Cluster, the Intel based quad core Cluster, the Sun® Constellation, and the Blue Gene/P. These systems have been carefully selected from the current list of Top500 supercomputers to provide objectivity. The TCO is calculated assuming a data center is designed and constructed around the computing performance desired.

---

[1] The Green 500 List, www.green500.org/Lists.html

[2] The Top 500 list at www.top500.org

For the sake of objectivity, people costs pertaining to training and application-environment migration differ across industries and were not accounted for in the study. Only generic factors consistent across all the industries have been considered. We have also not considered potential costs incurred due to failures and unplanned outages.

The following supercomputer models were chosen for the TCO study from the Top 500 list of supercomputers, because they are benchmark representatives of the top ten supercomputing systems in the globe today. The key attributes for the systems at performance of 13.9 Tflops are itemized below:

Installation	**IBM Blue Gene/P Systems**	**Intel-based Dual-Core Cluster Systems**	**Intel-based Quad-Core Cluster Systems**	**Sun Constellation Systems**
No. of Server Racks	1	14	7	2
Floor Space (ksf)	0.32	1.36	0.8	0.4
Electricity Cost (M $/year)	0.054	0.4	0.2	0.1
Annualized IT-related Capital Costs (M $/year)	0.7	1.3	1	0.9
Annualized Site-related Capital Cost (M $/year)	0.6	1.2	0.9	0.7
Annualized Operating Expenses (IT- and site-related in M $/year)	1.96	2.6	2.5	2.4
Total Annualized Cost (M $/year)	3.2	5.4	4.6	4

Factors That Fuel the TCO

The TCO is comprised of several significant factors, such as electricity cost that accounts for the annual energy cost of wattage being consumed, floor space cost that accounts for the density, architectural costs associated with infrastructure including networks and storage, capital hardware acquisition cost, and the people cost was restricted only to maintaining the system at the customer site in the data center.

The comparative study and its analysis were based on the knowledge that an Intel Cluster has standardized components on which far greater number of applications can run which lowers expected deployment and customization costs. Hence factors like application enablement and migration costs, operating costs and the training costs for all systems were not considered in the study. Software licensing costs vary across providers and industries, and hence were not considered.

Energy Consumption: With technological and performance advancements in processor technology and design R&D which enable high density packaging of processing cores on a server, the processor energy consumption could rise. Higher number of transistors translates into increased operational power costs and higher levels of heat generated per chip.

System failure rates increase with increase in temperature and also with larger footprints and hence adequate cooling is essential for efficient functioning. In addition to the direct cooling costs involved, cooling systems could occupy additional space on racks. Each rack cannot be fully populated with only server nodes, and more racks would be needed for a particular performance level.

As the energy requirements in large data centers rise, additional UPS and backup power capacities are needed for the operation and cooling of the data center.

Total Floor Area: In the last 5 years the performance of systems in data centers has increased exponentially. Advanced networking technologies and high speed Infiniband switches have enabled clustering of a large number of nodes. Most equipment layouts are in a single row of rack-mounted servers, forming aisles between them. Network switches and storage devices, placed alongside the racks, are often as big as the racks themselves. This has caused a significant strain on the infrastructure of data centers that were built for hardware with much less capability than what is being shipped today. The much higher rack power levels have caused customers to spread out their server products in order to cool them in the current facilities, using up valuable and expensive raised floor space.

The electrically active floor area of a data center is estimated to be only about 40% of the total floor area of the data center. Chillers, fans, pumps, service aisles between racks and other electrically inactive components make up the remaining.

IT Acquisition Costs**:** The IT-related capital cost, is Capital Recovery Factor times the capital costs incurred for a 3 year life. It accounts for the total IT-related capital cost incurred from investment in total number of filled racks, internal routers and switches, and rack management hardware. At a low level, the total IT-related capital cost for the investment in total number of filled racks includes the acquisition cost for servers, disk and tape storage, and networking.

Other Facilities Costs: Other facilities costs include interest during construction estimated based on total infrastructure and other facility capital costs at a fixed rate of interest, land costs, architectural and engineering fees, and inert gas fire suppression.

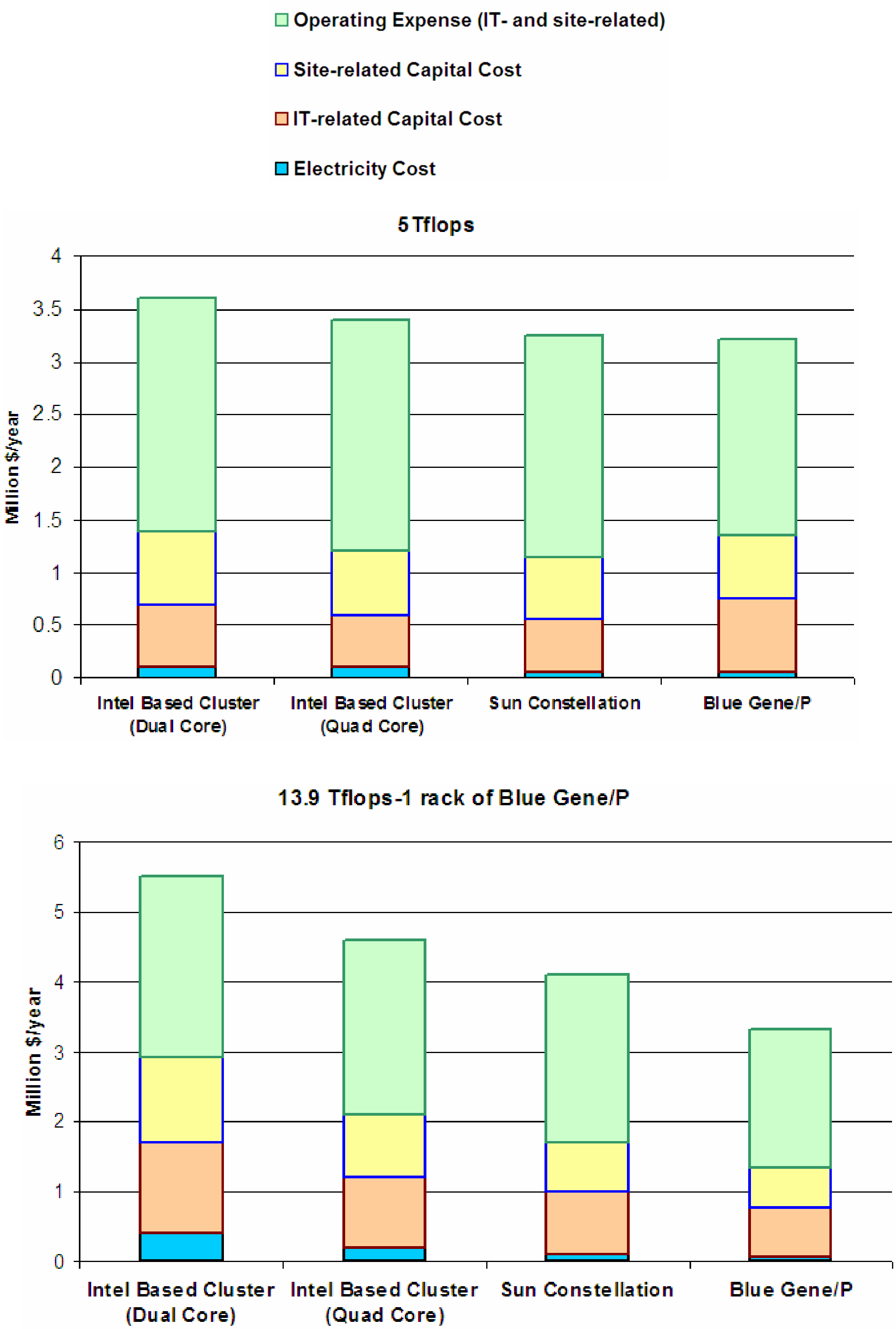
Land costs are based at $100,000 per acre and architectural and engineering fees are estimated as 5% of kW related infrastructure costs plus other facility costs (electrically active).

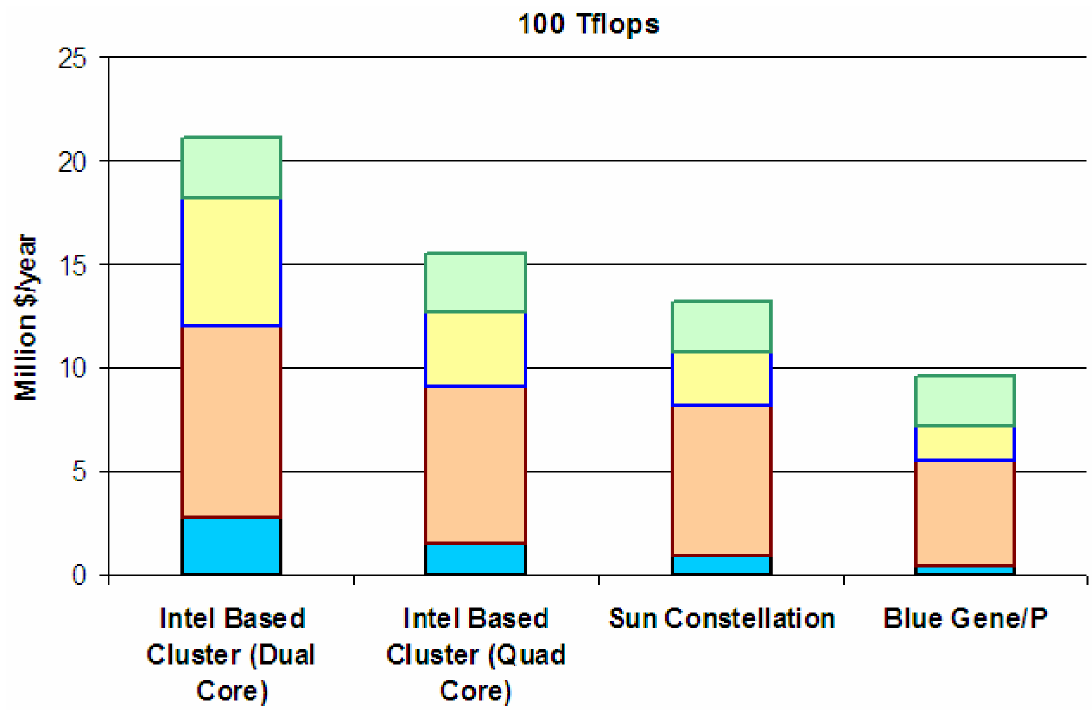
Interpretation of the Results

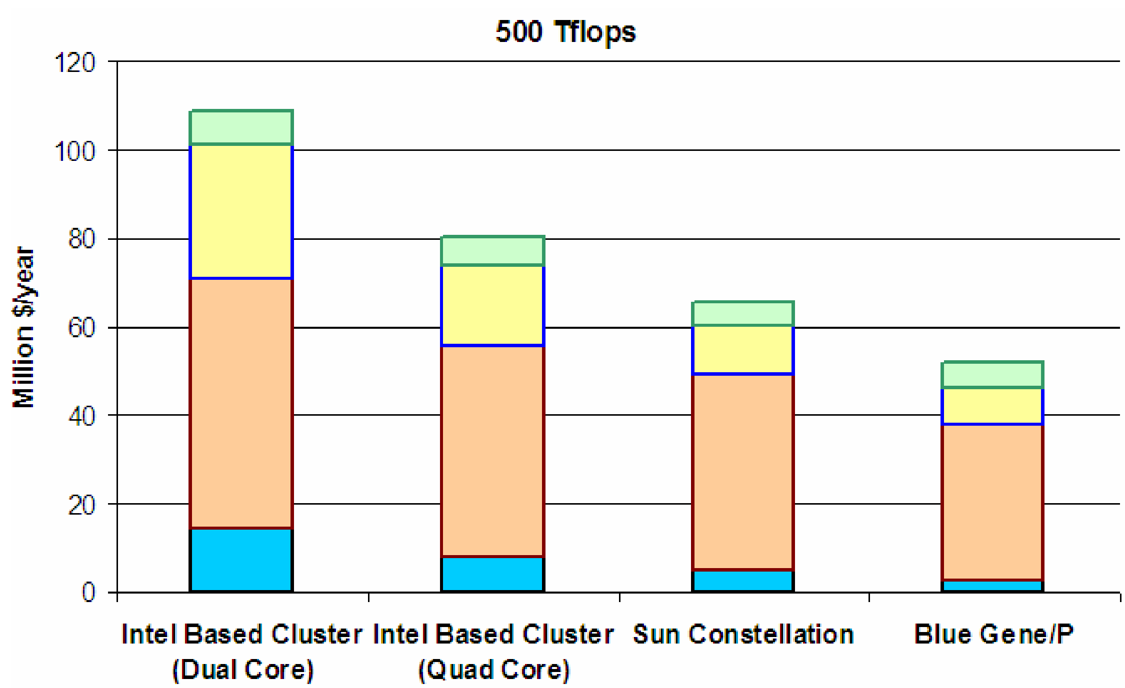
This TCO study includes the energy costs for the entire data center, the IT capital costs and infrastructure capital costs annualized over a 3 year life and a 15 year life for land and other fixed property, and the annualized operating expenses. The analysis accounts for the maintenance cost of $100,000 annually for the second and third year of operation of a Blue Gene/P. No additional maintenance costs are assumed for other systems. We studied systems at 5, 13.9, 100 and 500 teraflops, and the following interpretation is formed from an arithmetic mean of the relevant annual costs incurred at the performance ranges for 100 and 500 teraflops.

Energy Costs: Energy costs observations were derived from the total costs incurred by the components (servers, disk storage, tape storage and networking) for IT load, cooling and auxiliaries. Overall, the Blue Gene/P is the most economical system for the entire performance range from 5- 500 Tflops and is the most energy-efficient system as we scale up. As the performance increases, the gap between the increased energy cost for the Blue Gene/P and other systems widens. Also the Intel-based Quad-Core Cluster is found to be more energy-efficient than the Dual-Core Cluster. Servers contributed to the most energy consumption for all the systems for the range of 5-500 Tflops. The contributors to IT load power consumption are servers and disk storage for Blue Gene, and servers and networking for Intel quad-core Cluster. Server power consumption for IT loads was small in the case of Sun Constellation system for 5-13.9 Tflops. Networking energy costs were the least in case of the Blue Gene system.

IT Capital Costs: The IT capital costs include acquisition expenses incurred on racks, external hardwired connections, internal routers and switches, storage and rack management hardware. The IT capital costs for a Blue Gene/P rose significantly only over higher performance ranges from 100- 500 Tflops, and overall average IT-related capital costs of a Blue Gene/P was found to be 38.6% less than the Intel based dual core cluster, 27% less than a quad core Intel based cluster, and 22% less than the Sun Constellation system.


Operating Expense (IT- and site-related)
Site-related Capital Cost
IT-related Capital Cost
Electricity Cost
5 Tflops
Million $/year
0
0.5
1
1.5
2
2.5
3
3.5
4
Intel Based Cluster (Dual Core)
Intel Based Cluster (Quad Core)
Sun Constellation
Blue Gene/P
13.9 Tflops-1 rack of Blue Gene/P
Million $/year
0
1
2
3
4
5
6
Intel Based Cluster (Dual Core)
Intel Based Cluster (Quad Core)
Sun Constellation
Blue Gene/P


100 Tflops
Million $/year
25
20
15
10
5
0
Intel Based Cluster (Dual Core)
Intel Based Cluster (Quad Core)
Sun Constellation
Blue Gene/P


500 Tflops
Million $/year
120
100
80
60
40
20
0
Intel Based Cluster (Dual Core)
Intel Based Cluster (Quad Core)
Sun Constellation
Blue Gene/P

Facilities Costs: Facilities cost, an equally critical consideration for clients, demonstrate that the Blue Gene is the least expensive at $4.95M compared to Intel's dual-core clusters at $18.4M, and $11M for quad-core, and Sun's Constellation at $6.95M. For each of the components, the kW related infrastructural cost is the power consumption times the power related cost (in dollars). Significant cost escalations occurred over higher performance ranges from 100- 500 Tflops.

Operating Expenses: The IT- and the site-related operating expenses account for the total operating expenses incurred in investing in each of the four systems under study. Operating expenses demonstrated a marked rise for higher performance ranges from 100- 500 Tflops. The Blue Gene/P demonstrated the cost advantage of a 22.26% reduction compared to the Intel-based Dual-Core Cluster, a 10.44% reduction against the Intel-based Quad-Core Cluster, and a minor 6.55% increase when compared to the Sun Constellation.

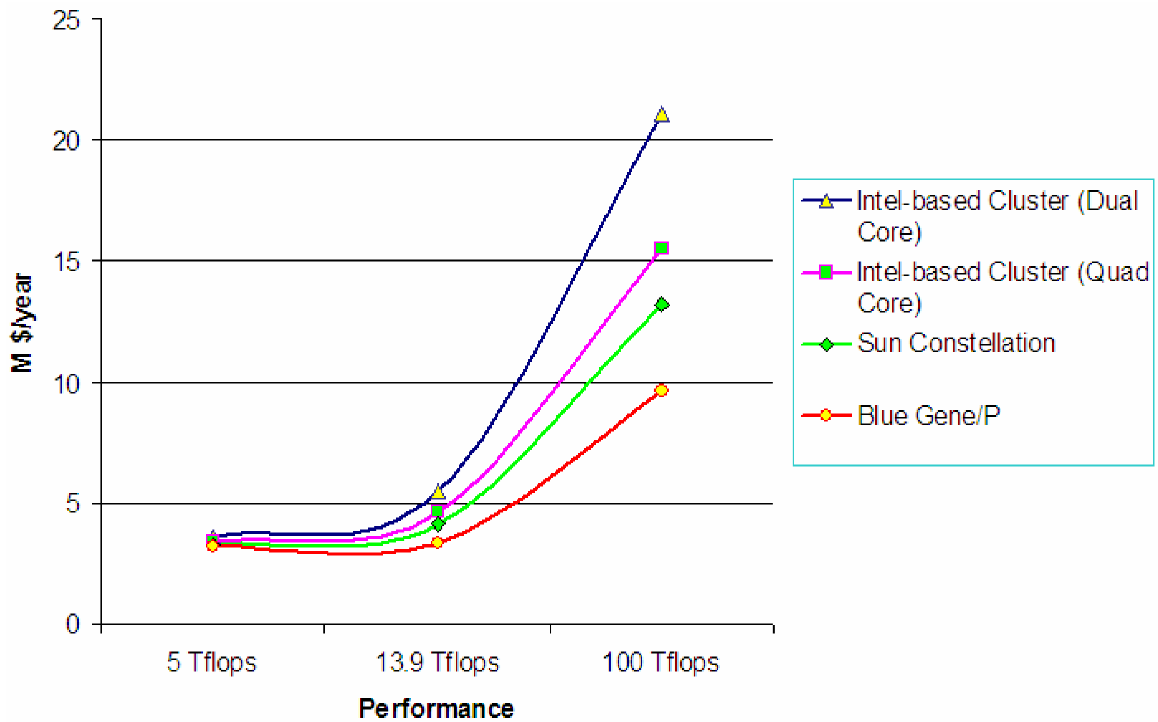
Total Costs: The summary of all the four costs for all the four systems at all performance levels undertaken in the study finds the Blue Gene/P to be the most cost-effective system of choice whose average, of all expenses at all performance ranges, is 52.62% less than the Intel-based dual-core cluster, 35.86% less than the Intel quad-core cluster and 21.88% less than the Sun cluster.

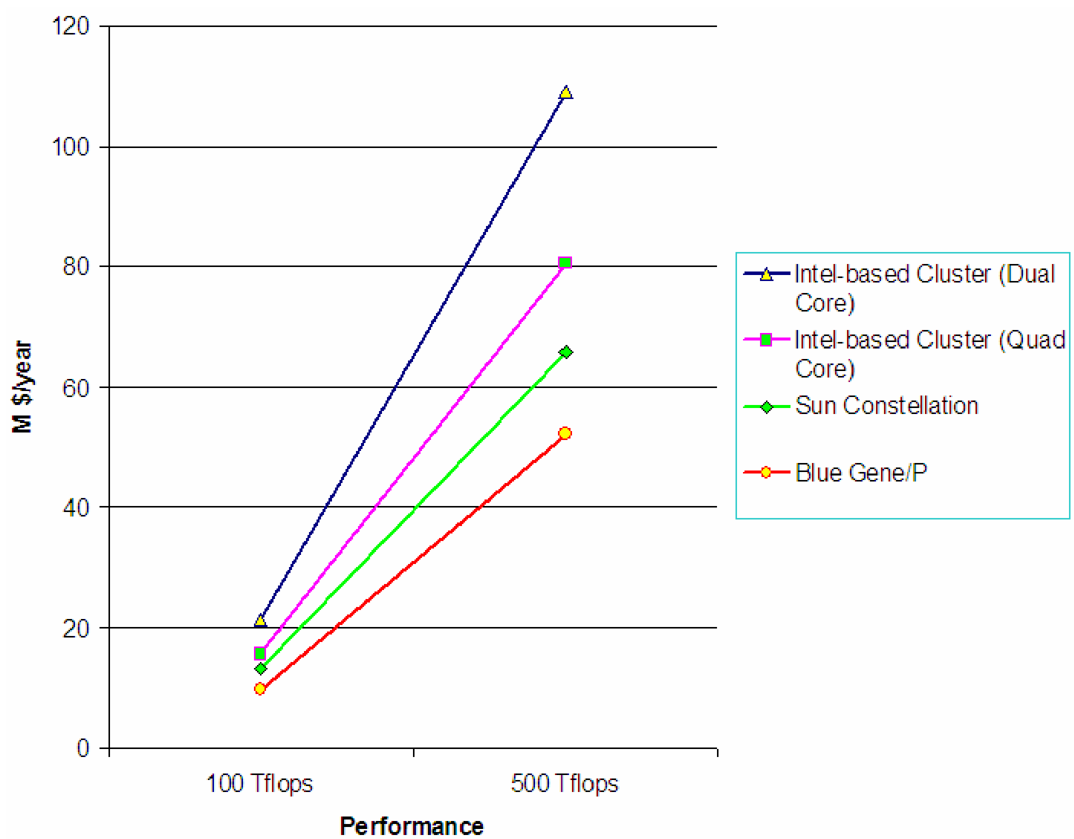
Other Considerations: Intel's cluster architecture is the most prevalent offering in the HPC market for a considerable span of time. This architecture has a thriving ecosystem of applications, software tools, and other HPC components. Hence the costs (not considered in this analysis) of application migration, training, and deployment for this architecture could be significantly less than other architectures.

To run and support many applications on an IBM Blue Gene, a fair amount of skills and performance optimization would be required. The good news (for the Blue Gene) is that over the last 5 years, the portfolio of optimized applications available on the Blue Gene continues to grow rapidly in several HPC disciplines.

Energy costs and cooling costs have been a major buzz in the IT industry but it forms a relatively small component (3%-7%) in the study on Total Cost of Ownership. However, energy considerations could be the limiting factor for the growth of HPC data center capability in urban and semi-urban locations. These facilities restrictions could make the Blue Gene/P a very attractive choice in these HPC data centers.

Next, we examine the TCO for the various systems over a broad range of performance levels.


25
20
15
10
5
0
M $/year
5 Tflops
13.9 Tflops
100 Tflops
Performance
Intel-based Cluster (Dual Core)
Intel-based Cluster (Quad Core)
Sun Constellation
Blue Gene/P


120
100
80
60
40
20
0
M $/year
100 Tflops
500 Tflops
Performance
Intel-based Cluster (Dual Core)
Intel-based Cluster (Quad Core)
Sun Constellation
Blue Gene/P

At lower performance levels, the cost of standard clusters is probably more attractive especially if application rework and deployment and training costs are considered. But when scaling up to higher performance systems beyond 50 teraflops, the Blue Gene/P easily emerged as the winning system of choice with a widening cost advantage.

At higher performance levels, the Blue Gene/P's total costs are lower reflecting its fundamental advantages of an ultrascalable architecture, technology design, and energy optimization for supercomputing.

The Bottom line

The Blue Gene/P system from IBM is attractive because of the total performance it offers, and its total cost that is far lower than the other systems in the study, over the three-year period.

Even then, the appetite for HPC is growing by a factor of 10 (from 1 core in 1986, 10 cores in 1992, 100 cores in 1998, 1,000 cores in 2004, and a realistic projection of 10,000 cores by 2010), hence the biggest challenge is to make the systems more efficient from the perspective of energy, floor space, operating expense and deployment costs. The Blue Gene/P has the appropriate architecture to meet these future requirements. Furthermore, the lower energy consumption and smaller foot prints of Blue Gene/P systems especially at larger performance levels could significantly increase system reliability and hence reduce downtime costs.

Intel-based Clusters have one major current advantage in that more HPC applications are supported hence application migration and deployment costs could be much smaller especially for environments that do not require large scalability. For HPC environments that require large performance scalability and economical operation, the Blue Gene/P is an excellent platform and the costs of application migration and training could be well worth the investment.

For More Information

To learn more about IBM's high performance computing product portfolio and the IBM Blue Gene/P, contact your IBM representative. More information on the IBM Blue Gene/P can be had at the website: http://www-03.ibm.com/systems/deepcomputing/bluegene/
The cost drivers of TCO are quantified based on the model provided by the Uptime Institute, Inc.: http://uptimeinstitute.org/content/view/57/81/

Assumptions:

1. Fraction of racks allocated to different categories based on experts' experience and judgment for high performance computing in financial applications.
2. Racks are standard (6.5 feet high with 42 Us per rack).
3. Energy use per U taken from selective review of market/technology data. Server power and costs per watt assumes IBM X-3550 1U system.
4. Total electricity use is the sum of IT, cooling, and auxiliaries. Cooling and auxiliaries together are equal to the IT load.
5. Total electricity consumption is calculated using the total power, a power load factor of 95%, and 8766 hours/year (average over leap and non-leap years).
6. US average industrial price of electricity is assumed to be 0.09 $/kWh consistent with year 2008.
7. Watts per thousand 2007 dollars of IT costs taken from selective review of market and technology data. Server number calculated assuming IBM x3550 1U server as described in next note.
8. Cost per filled U taken from selective review of market and technology data. Server street cost calculated assuming IBM x3550 1U server with 8 GB RAM, two dual core 2.66 GHz Intel processors (19.2 GFLOPS max/server).
9. Total IT costs are the product of the number of filled Us and the cost per filled U.
10. Interest during construction estimated based on total infrastructure and other facility capital costs assuming a 7% real interest rate for one year.
11. Capital costs with three year life include all IT equipment costs (which include internal routers and switches).
12. Capital costs are annualized with the capital recovery factor calculated using the appropriate lifetime (three or 15 years) and a 7% real discount rate.
13. Total operating expenses include electricity costs, network fees, and other operating expenses.
14. All costs are in year 2007 dollars.