ITG STATUS REPORT

November 2006

BUSINESS CASE FOR IBM SYSTEM BLUE GENE SOLUTION: REDUCING THE COSTS OF COMMERCIAL SUPERCOMPUTING

Challenges

At the end of 2005, there were around 20 supercomputers worldwide capable of delivering more than 20 teraflops of Rpeak performance. On current trends, by the end of the decade thousands will be in use.

Technology advances, accompanied by dramatic declines in the cost of processing power, are enabling organizations to develop new generations of breakthrough applications. The potential is emerging for far-reaching transformations not only in government and academic research, but also in industries where competitiveness is affected by research and development (R&D) investment.

For many companies, conventional strategies for leadership through R&D will erode as large-scale supercomputing becomes the norm among Fortune 1000 corporations as well as smaller players seeking to challenge them through aggressive innovation. Competitive pressures will force escalating supercomputer investments.

The dynamics of this process will vary between industries. Transformations, however, may be rapid, and companies need to plan for them now. A great deal of work will be required in application conception and development. But this should not be the only focus. Critical infrastructure challenges are also posed.

Commodity microprocessors and Linux cluster architectures offer the potential for unprecedented levels of hardware price/performance. They also, however, increase data center space, power and cooling requirements. Costs for these may more than offset the advantages of inexpensive hardware, and may create reliability and manageability problems that impair system effectiveness.

It is from this perspective that the potential role of the IBM System Blue Gene solution should be viewed.

Infrastructure Costs

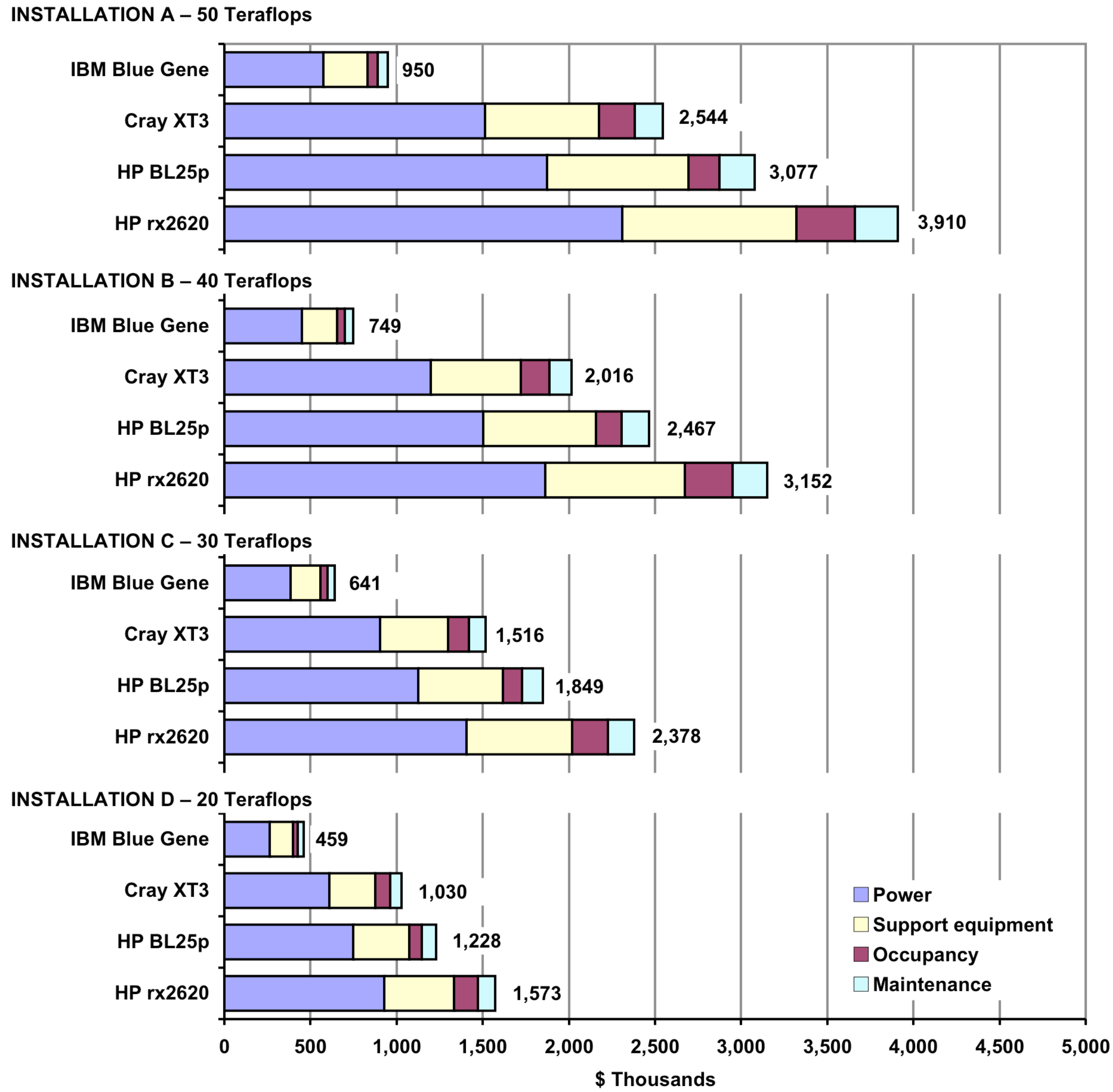
This report compares infrastructure costs – including costs of data center floor space, power and support equipment – for commercial supercomputing systems.

Four platforms are compared: IBM Blue Gene systems; Cray XT3 systems and Hewlett-Packard (HP) BL25p blade clusters based on AMD Opteron processors; and HP Cluster Platform 6000 systems built around Intel Itanium 2-based rx2620 servers. All platforms are configured with current technology dual-core processors. Calculations are based on four projected supercomputer installations delivering between 20 and 50 teraflops peak performance.

If existing data centers are employed, three-year infrastructure costs for IBM Blue Gene systems average 60.6 percent less than those for Cray XT3 configurations; 67.5 percent less than those for HP BL25p blade clusters; and 74.6 percent less than those for HP Cluster Platform 6000 configurations delivering comparable performance.

ITG STATUS REPORTS contain analyses of important topics and future trends in information technology and are published periodically by the International Technology Group.

Results are summarized below.



If existing data centers are employed, three-year infrastructure costs for IBM Blue Gene range from $1.59 million to $0.57 million less than those for Cray XT3 configurations; from $2.13 million to $0.77 million less than those for HP BL25p configurations; and from $2.96 million to $1.11 million less than those for HP Cluster Platform 6000 configurations.

If new data center facilities must be constructed, infrastructure costs for IBM Blue Gene range from $2.16 million to $0.79 million less than those for comparable Cray XT3 configurations; from $2.61 million to $0.93 million less than those for comparable HP BL25p configurations; and from $4.03 million to $1.54 million less than those for comparable HP Cluster Platform 6000 configurations.

Floor space and power values are for processor complexes and external Gigabit Ethernet switches, along with support equipment including uninterruptible power supplies (UPS), power distribution systems (PDS), cooling systems and backup diesel generator capacity.

Calculations do not include hardware, software or maintenance costs for supercomputer systems and storage. Configurations, vendor discounts, financing arrangements and other factors that affect these may vary widely between organizations.

Calculations are prorated; e.g., if an installation requires 3.5 UPS units, floor space is calculated based on the footprint of the unit multiplied by 3.5. Equipment purchase and maintenance costs are calculated similarly. The same approach is applied to other resources.

Cost breakdowns for the four installations, along with the basis of calculations, are presented in more detail at the end of this report.

Cost Drivers

The computer industry has a long history of cost surprises. In the 1980s, for example, it was widely believed that use of PCs and local area networks (LANs) offered the potential for dramatic reductions in IT costs. More recently, comparable claims have been advanced for commodity Intel processor-based Windows and Linux servers.

Such expectations have not been met. Despite apparently dramatic improvements in price/performance, IT costs have continued to increase.

The key lesson learned has been that price/performance measurements based solely on hardware purchase costs do not accurately reflect overall cost structures. Other factors may more than outweigh the advantages of inexpensive hardware. It thus becomes vital to understand the full range of IT cost drivers.

User experiences with supercomputer systems have highlighted the significance of two of these factors:

1. ***Power and heat***. Higher levels of microprocessor performance are achieved by greater design densities. As more transistors are embedded, and clock speeds increase, the amount of power required – and the heat generated per chip – increase.

   Users are faced not only with escalating electricity costs, but also with the need to cool increasingly dense system packages. Ripple effects are created across the entire data center cost structure. Higher system power consumption means that greater UPS and backup generator capacities are required, and increased heat output requires additional cooling equipment. These, in turn, add to power loads.

   Users, for example, are beginning to experience problems with the amount of heat generated by dense blade packages. In supercomputing as well as commercial applications, an increasing number of organizations are unable to fully populate blade cabinets for this reason, or are obliged to invest in new cooling technologies, or both.

   Failure rates increase as temperatures rise. If cooling is inadequate or suboptimal, organizations face more frequent, and potentially more serious disruptions to system operations over time.

Under any scenario, users can expect continued increases in electricity prices. According to the Energy Information Administration, between July 2005 and July 2006 (the most recent month for which data was available when this report was prepared), average U.S. industrial prices per kilowatt hour increased by approximately 11 percent. The rate of growth is accelerating over time.

Other prices are also rising. In August 2006, for example, Emerson Network Power announced an average increase of 5.5 percent for its market-leading Liebert brand of data center power and cooling products. The company cited rising costs for lead, copper, steel and other raw materials. The following month, citing higher raw materials costs as well as more stringent emissions standards, diesel generator supplier Cummins announced price increases of up to 10 percent.

Further price increases for these and other types of critical data center support equipment can be expected in the future.

2. ***Floor space***. The basic measure of floor space in a raised floor environment is cabinet size. Many supercomputer installations already require dozens to hundreds of cabinets.

   There is again a ripple effect. Allowance must also be made for service clearances, main aisles and other inactive areas for computers as well as support equipment. One commonly used rule of thumb is that electrically active equipment occupies only 30 percent of data center space.

   Unless there is a great deal of surplus space within existing facilities, organizations thus face potentially major increases in occupancy costs, or the need to invest in new facilities. There are, moreover, signs that with the latest generation of supercomputers, the limitations of conventional raised floor cooling techniques are starting to be reached. Even users who do not face space constraints may be faced with major new data center outlays.

   Recent estimates of average data center construction costs in the United States are in the range of $400 to $500 per square foot, and this figure is increasing over time. Organizations operating in comparatively expensive urban areas may experience significantly higher costs.

Some attention is being paid to these issues now. There is often, however, a tendency to assume that because the potential benefits of supercomputer applications are so great, or because supercomputer investment is mandated for competitive reasons, infrastructure costs should be, at best, a subsidiary factor in decision-making.

Such attitudes, in the past, were common among users of other types of server. They are less prevalent today. Server consolidation and infrastructure cost reduction initiatives have become the industry norm.

There is no obvious reason to repeat these mistakes. Experience has shown that, without proper planning, inefficiencies may be embedded into infrastructures in a manner that causes sustained long-term cost inflation. Such inefficiencies may also impair system availability and performance. In both areas, experiences with small-scale systems may prove a poor indicator of the challenges organizations will face as workloads expand far beyond early levels.

One conclusion emerges clearly. Among organizations planning to use the new generation of large-scale supercomputers, close attention should be paid to infrastructure issues. The right choices must be made at an early stage of the decision-making process.

Not least of these choices will be which system to employ.

Platform Differentiators

All of the supercomputer platforms covered by this report enable high levels of computational performance. From the perspective of infrastructure costs, however, the architecture and technology of the Blue Gene system is significantly differentiated.

This is notably the case in two areas:

1. ***Power and heat***. The 700 MHz dual-core PowerPC processor employed in Blue Gene systems is characterized by a low bandwidth design and high levels of on-chip integration, which result in consumption of approximately 12 watts per processor. IBM refers to a dual-core Blue Gene processor as a "node."

   In comparison, 2.6 GHz dual-core Opteron processors are rated at 95 watts, and 1.6 GHz Itanium 2 dual-core processors at 104 watts. There are corresponding differences in system-level power consumption, and in heat output and cooling requirements.

2. ***Packaging density.*** High levels of on-chip integration, as well as an exceptionally space-optimized overall design mean that up to 1,024 Blue Gene processors (2,048 cores) may be incorporated in a single 36-inch wide by 36-inch deep cabinet.

   In comparison, XT3 systems are built around a 22.5-inch wide by 56.75-inch deep cabinet containing up to 96 Processor Elements (192 cores). The Cluster Platform 6000 is built around EIA cabinets that typically contain 16 to 20 rx2620 units (64 to 80 cores using two dual-core processors per unit). Because of heat generation issues, it is unclear what the actual density of large BL25p cluster configurations would be in practice.

A key point should be emphasized. These Blue Gene characteristics are not the result of simple engineering enhancements. They reflect fundamental parameters of architecture and technology design. Blue Gene infrastructure costs are lower because the system is, to a much greater degree than for the other platforms, designed and optimized to minimize these.

Certain characteristics of Blue Gene reliability, availability and serviceability (RAS) design also suggest that the system will be able to maintain high levels of uptime. Soldered-down memory and fewer moving parts than commodity server-based designs reduce the number of potential points of failure.

Blue Gene systems also benefit from a tapered duct structure that conveys cooled air more efficiently to higher cabinet levels than conventional techniques. Some users have found that, with raised floor cooling, failure rates for high-density systems tend to be higher toward the top of cabinets because air circulated through floor tiles has less effect here. This problem can be addressed by mounting cooling units on top of cabinets. The Blue Gene approach is, however, a great deal less expensive.

Additional Information

This ITG Status Report is based upon the preliminary results and methodology for an upcoming series of Industry Management Briefs to be released by the International Technology Group. For copies of these Industry Management Briefs, please email requests to info-itg@pacbell.net.

Detailed Data

Installation Values: IBM Blue Gene Systems

INSTALLATION	A	B	C	D
Teraflops (1)	50	40	30	20
Compute processors (2)	9,216	7,168	6,144	4,096
Number of cabinets (3)	9	7	6	4
FLOOR SPACE (sq. ft)				
System (4)	280	232	208	160
Support equipment (5)	271	213	181	124
Total	**551**	**445**	**389**	**284**
POWER CONSUMPTION (kWh)				
System (6)	274	215	184	126
Support equipment (6)	123	97	83	56
Total	**397**	**312**	**267**	**182**
SUPPORT EQUIPMENT COSTS ($000)				
UPS/PDS (7)	112	88	75	54
Cooling (8)	71	56	48	39
Backup generation (9)	75	59	50	41
Total	**258**	**203**	**173**	**134**
THREE-YEAR OPERATING COSTS ($000)				
Power (10)	574	451	386	263
Occupancy (11)	58	47	41	30
Support eqpt. maintenance	60	48	41	32
Total	**692**	**546**	**468**	**325**
EXISTING FACILITY – OVERALL INFRASTRUCTURE COSTS ($000)				
	950	**749**	**641**	**459**
NEW FACILITY COSTS ($000)				
Floor space construction (12)	220	178	156	114
NEW FACILITY – OVERALL INFRASTRUCTURE COSTS ($000)				
	1,170	**927**	**797**	**573**

(1) Rpeak.
(2) PowerPC 700 MHz dual core.
(3) Rounded to nearest cabinet.
(4) Based on manufacturer cabinet specifications and recommended service clearances, plus cable ducts and 60" aisles where appropriate. Includes system and external Gigabit Ethernet switches.
(5) Based on manufacturer specifications and recommended service clearances for equipment described below, plus aisles and other inactive areas where appropriate.
(6) Based on manufacturer specifications and industry norms. Includes computer room heat load.
(7) Based on Liebert 610 1000 kVA UPS & power distribution equipment.
(8) Based on Liebert Deluxe System/3 30-ton water-cooled units.
(9) Based on Cummins 2250 kW units.
(10) Based on $0.055 per kWh. Assumes 24/365 operations.
(11) Based on $35 per square foot per year.
(12) Based on $400 per square foot.

Installation Values: Cray XT3 Systems

INSTALLATION	A	B	C	D
Teraflops (1)	50	40	30	20
Compute processors (2)	4,888	3,854	2,914	1,974
Number of cabinets (3)	52	41	31	21
FLOOR SPACE (sq. ft)				
System (4)	1,256	1,007	727	545
Support equipment (5)	713	564	426	289
Total	**1,969**	**1,571**	**1,153**	**834**
POWER CONSUMPTION (kWh)				
System (6)	720	571	430	290
Support equipment (6)	326	258	195	132
Total	**1,046**	**829**	**625**	**422**
SUPPORT EQUIPMENT COSTS ($000)				
UPS/PDS (7)	287	227	171	115
Cooling (8)	183	145	109	74
Backup generation (9)	192	152	115	77
Total	**662**	**524**	**395**	**266**
THREE-YEAR OPERATING COSTS ($000)				
Power (10)	1,512	1,198	903	610
Occupancy (11)	207	165	121	88
Support eqpt. maintenance	163	129	97	66
Total	**1,882**	**1,492**	**1,121**	**764**
EXISTING FACILITY – OVERALL INFRASTRUCTURE COSTS ($000)				
	2,544	**2,016**	**1,516**	**1,030**
NEW FACILITY COSTS ($000)				
Floor space construction (12)	788	628	461	333
NEW FACILITY – OVERALL INFRASTRUCTURE COSTS ($000)				
	3,332	**2,644**	**1,977**	**1,363**

(1) Rpeak.
(2) AMD Opteron 2.6 GHz dual core.
(3) Rounded to nearest cabinet.
(4) Based on manufacturer cabinet specifications and recommended service clearances, plus cable ducts and 60" aisles where appropriate. Includes system and external Gigabit Ethernet switches.
(5) Based on manufacturer specifications and recommended service clearances for equipment described below, plus aisles and other inactive areas where appropriate.
(6) Based on manufacturer specifications and industry norms. Includes computer room heat load.
(7) Based on Liebert 610 1000 kVA UPS & power distribution equipment.
(8) Based on Liebert Deluxe System/3 30-ton water-cooled units.
(9) Based on Cummins 2250 kW units.
(10) Based on $0.055 per kWh. Assumes 24/365 operations.
(11) Based on $35 per square foot per year.
(12) Based on $400 per square foot.

Installation Values: Hewlett-Packard Blade Cluster BL25p

INSTALLATION	A	B	C	D
Teraflops (1)	50	40	30	20
Compute processors (2)	4,800	3,840	2,880	1,920
Number of cabinets (3)	50	40	30	20
FLOOR SPACE (sq. ft)				
System (4)	861	680	513	346
Support equipment (5)	884	709	532	352
Total	**1,745**	**1,389**	**1,045**	**698**
POWER CONSUMPTION (kWh)				
System (6)	892	715	536	356
Support equipment (6)	404	324	243	161
Total	**1,296**	**1,039**	**779**	**517**
SUPPORT EQUIPMENT COSTS ($000)				
UPS/PDS (7)	355	285	213	142
Cooling (8)	226	181	136	90
Backup generation (9)	238	191	143	95
Total	**819**	**657**	**492**	**327**
THREE-YEAR OPERATING COSTS ($000)				
Power (10)	1,873	1,502	1,126	747
Occupancy (11)	183	146	110	73
Support eqpt. maintenance	202	162	121	81
Total	**2,258**	**1,810**	**1,357**	**901**
EXISTING FACILITY – OVERALL INFRASTRUCTURE COSTS ($000)				
	3,077	**2,467**	**1,849**	**1,228**
NEW FACILITY COSTS ($000)				
Floor space construction (12)	698	556	418	279
NEW FACILITY – OVERALL INFRASTRUCTURE COSTS ($000)				
	3,775	**3,023**	**2,267**	**1,507**

(1) Rpeak.
(2) AMD Opteron 2.6 GHz dual core.
(3) Rounded to nearest cabinet.
(4) Based on manufacturer cabinet specifications and recommended service clearances, plus cable ducts and 60" aisles where appropriate. Includes system and external Gigabit Ethernet switches.
(5) Based on manufacturer specifications and recommended service clearances for equipment described below, plus aisles and other inactive areas where appropriate.
(6) Based on manufacturer specifications and industry norms. Includes computer room heat load.
(7) Based on Liebert 610 1000 kVA UPS & power distribution equipment.
(8) Based on Liebert Deluxe System/3 30-ton water-cooled units.
(9) Based on Cummins 2250 kW units.
(10) Based on $0.055 per kWh. Assumes 24/365 operations.
(11) Based on $35 per square foot per year.
(12) Based on $400 per square foot.

Installation Values: Hewlett-Packard Cluster Platform 6000 Systems

INSTALLATION	A	B	C	D
Teraflops (1)	50	40	30	20
Compute processors (2)	3,852	3,096	2,340	1,548
Number of cabinets (3)	107	86	65	43
FLOOR SPACE (sq. ft)				
System (4)	2,142	1,744	1,330	900
Support equipment (5)	1,092	880	663	438
Total	**3,234**	**2,624**	**1,993**	**1,338**
POWER CONSUMPTION (kWh)				
System (6)	1,100	886	669	441
Support equipment (6)	499	402	303	200
Total	**1,599**	**1,288**	**972**	**641**
SUPPORT EQUIPMENT COSTS ($000)				
UPS/PDS (7)	438	353	266	176
Cooling (8)	279	224	169	112
Backup generation (9)	293	236	178	118
Total	**1,010**	**813**	**613**	**406**
THREE-YEAR OPERATING COSTS ($000)				
Power (10)	2,311	1,862	1,405	927
Occupancy (11)	340	276	209	140
Support eqpt. maintenance	249	201	151	100
Total	**2,900**	**2,339**	**1,765**	**1,167**
EXISTING FACILITY – OVERALL INFRASTRUCTURE COSTS ($000)				
	3,910	**3,152**	**2,378**	**1,573**
NEW FACILITY COSTS ($000)				
Floor space construction (12)	1,294	1,050	797	535
NEW FACILITY – OVERALL INFRASTRUCTURE COSTS ($000)				
	5,204	**4,202**	**3,175**	**2,108**

(1) Rpeak.
(2) Integrity rx2620 2 x Itanium 2 1.6 GHz dual core.
(3) Rounded to nearest cabinet.
(4) Based on manufacturer cabinet specifications and recommended service clearances, plus 60" aisles where appropriate. Includes system and external Gigabit Ethernet switches.
(5) Based on manufacturer specifications and recommended service clearances for equipment described below, plus aisles and other inactive areas where appropriate.
(6) Based on manufacturer specifications and industry norms. Includes computer room heat load.
(7) Based on Liebert 610 1000 kVA UPS & power distribution equipment.
(8) Based on Liebert Deluxe System/3 30-ton water-cooled units.
(9) Based on Cummins 2250 kW units.
(10) Based on $0.055 per kWh. Assumes 24/365 operations.
(11) Based on $35 per square foot per year.
(12) Based on $400 per square foot.

International Technology Group

4546 El Camino Real, Suite 230
Los Altos, California 94022-1069
Telephone: (650) 949-8410
Facsimile: (650) 949-8415
Email: info-itg@pacbell.net