

Smart Infrastructures: computational resources to burn

Technology Horizons Program | SR 1042

1 | INTRODUCTION

Over the next 15–20 years we will overcome limits in availability of our computational resources. While today, high-performance computing applications are mostly limited to capital-intensive industries like petroleum exploration, aircraft and automotive design, and pharmaceuticals, over time these capabilities will migrate to mass markets and eventually into the hands of consumers. By 2010 sensor networks will be pervasive—according to report from Ernst & Young, there will be nearly 10,000 telemetric devices for every person on earth, and Ray Kurzweil postulates that by 2019 a \$1,000 computing device (in 1999 dollars) will have raw computational power, *although not necessarily the intelligence*, of a human brain. In this world of abundant computing, our interactions with computers will no longer be constrained to laptops, desktops, and handhelds. High-powered computing capability will be embedded in our physical environment, in living things, medicine, walls, furniture, garments, tools, utensils, and toys. We'll be able to interact with information in place as naturally as we interact now with physical things, which will become increasingly less passive, and more active. In short, the computation revolution will have huge impacts on daily life, workplaces, and in many industries.

In this memo, *Smart Infrastructures: Computational Resources to Burn* (SR-1042), we review the fundamental technologies driving exponential growth in computing resources and some of the likely applications in areas such as entertainment, gaming, health, and communications.

TABLE OF CONTENTS

1 Introduction

2 Technology Enablers

- Multi-core processors
- Metacomputing grids
- Nanoscale processors
- Molecular (DNA) processors
- Quantum-effects computing

3 Future Applications of Abundant Computing

- Interfaces
- Space and Place
- Immersive Media
- R&D
- Networks
- Youth
- Health



124 University Avenue, 2nd Floor
Palo Alto, CA 94301
www.iftf.org

2 | TECHNOLOGY ENABLERS

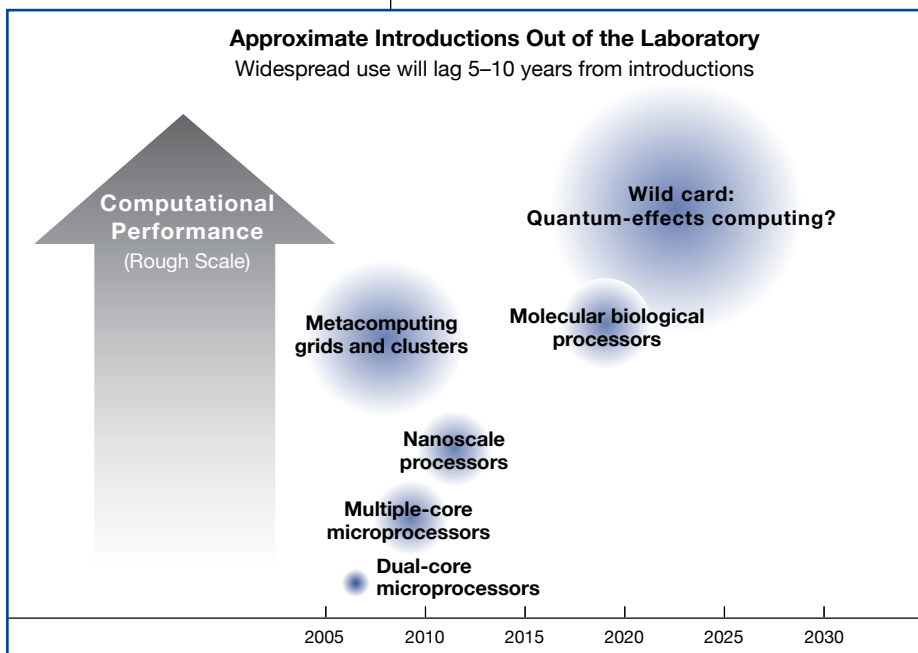
Many technologies will be harnessed to deliver abundant computing within the next decade and beyond—multi-core processor chips, grid computing, and perhaps even biological and quantum-effects based computers. However, when performing properly for the end user, these technologies will present extraordinary performance by today’s standards. Except for a small subset of high-latency network applications for the developing world, we believe that both local and grid approaches will deliver very high computing capability to the individual user.

Five key technological developments will enable this transformation:

- Multi-core processors
- Metacomputing grids
- Nanoscale processors
- Molecular (DNA) processors
- Quantum-effects computing

These technologies will emerge from labs over the course of the next ten to 30+ years (in the case of quantum computing) and will likely be in widespread use within 5–10 years of their release.

Timeline for the Release of Enabling Technologies



Source: Institute for the Future



Processor chips will grow from single-processor cores to multiple-processor cores on a chip.

To better manage the excessive power consumption and resulting heat from bigger processor chips, designers are selling chips with dual and eventually multiple processor cores.

Forecast

Today, dual-processor core chips are available from companies like AMD and Intel. And both have announced intentions to expand the number of processors on a die over the next ten years, but have been reluctant to state exactly how many processors may ultimately reside on a single chip.

Implications

As processors on a chip become redundant, processing net-useful power may increase at a rate beyond Gordon Moore's famous forecast of a doubling of the number of transistors every 18 months, and prices for computation will continue to drop correspondingly. And all this cheap, extra computing power will enable a range of powerful applications from immersive media to high-resolution telepresence, personal data-mining, complex modeling, and simulations.

Innovators

Intellisys is not waiting to build a chip with many processor cores on a chip. Its SEAForth™-24A Embedded Array Processor features 24 core processors on a single chip capable of combined sustained 24 billion operations per second. Intellisys is targeting software-defined radios for wireless communications, audio-signal processing, and remote data-collection applications for the chip.

RESOURCES

Intel Prototype May Herald a New Age of Processing

<http://www.nytimes.com/2007/02/12/technology/12chip.html>

The Landscape of Parallel Computing Research: A View from Berkeley

http://view.eecs.berkeley.edu/wiki/Main_Page

GLOSSARY

micron-scale: a measurement in units of microns, or one-millionth of a meter

submicron-scale: a measurement in units of less than a micron

nanoscale: a measurement in units of nanometers, or one-billionth of a meter

nanometric: refers to measurements in nanoscales

molecular scale: a measurement in terms of molecules, usual in nanometers

quantum dot: a semiconductor crystal with a diameter of a few nanometers; also describes an atom so confined and isolated that the removal or addition of a single electron can be detected

Individual computers will become linked together over ubiquitous networks into metacomputer grids.

Grid computing is one name for a whole family of research programs developing very-high performance computing programs linking computer systems across broadband networks. Autonomic computing, adaptable computing, cluster computing, on-demand computing, utility computing, and agile IT are all intended to provide extreme computational power that is accessible anytime, anywhere. Extra computing power is provided by tapping spare cycles on other computers in the network. This means that supercomputing power is available without the huge costs of supercomputers, and that CPU cycles that would otherwise be wasted are put to good use.

Forecast

According to a DARPA study released in July 2005, actual utilization of available online, high-performance computing is a fraction of available resources, because of the difficulty and expense of programming new applications. But given the relentless progress of multi-core and nanoscale processor design, the demand will increase irresistibly for programmers to learn how to program massively parallel and threaded applications. So, by 2015 the programming obstacles should be largely solved, and large-scale supercomputing services will be widely utilized over broadband terrestrial and wireless networks.

Implications

On-demand supercomputing will be increasingly used for pervasive computing, sensor nets, speech recognition, language translation, image recognition, online games, and ubiquitous media. Additionally, industries will increasingly benefit from capabilities to casually use very-high resolution simulations, real-time interactive graphic models, and huge numerical models.

Innovators

- IBM, HP, Sun, Oracle, SAS, and others are offering grid service extensions to existing ASP hosted web services offering variations with brand names like IBM On-Demand Computing and HP Utility Computing. But the uptake on these has been modest, for the aforementioned difficulty and expense of programming new applications.



- Google is building the one of the largest and fastest growing cluster based services. Already, Google is offering users free storage of videos, unlimited e-mail storage, and renders a very-high resolution image of the Earth on its professional version of Google Earth using its massively distributed, load-balanced network of thousands of generic Linux computers. Yahoo!, Microsoft, and AOL are likewise offering services utilizing massively distributed resources.
- The Globus open-source software toolkit was developed for grid computing by an international community of researchers in universities, national laboratories, and corporations. Globus is responsible for the reference implementations and de facto standards for grid computing, including reference software and practices.

RESOURCES

Getting Up to Speed: The Future of Supercomputing

S.L. Graham, M. Snir, and C.A. Patterson, eds, The National Academies Press, Washington, D.C. 2004.

http://www7.nationalacademies.org/cstb/pub_supercomp.html

Grid Computing "Overhyped"

Ingrid Marson, ZDNet UK, May 19, 2005

<http://news.zdnet.co.uk/communications/networks/0,39020345,39199143,00.htm>

Grid Computing: The Current State and Future Trends

A. Roxburgh, K. Pawlikowski, and D.C. McNickle, University of Canterbury, Christchurch, New Zealand.

http://www.cosc.canterbury.ac.nz/research/reports/TechReps/2004/tr_0401.pdf

Google and Akamai: Cult of Secrecy vs. Kingdom of Openness

http://corp.mooter.com/MIT_-_Apr_2004_475.aspx?pageid=475

High Performance Computing Software Survey

E. Joseph, A. Snell, C.G. Willard, S. Tichenor, D. Shaffer, S. Conway, Sponsored by: Defense Advanced Research Projects Agency, Council on Competitiveness, IDC, Framingham Mass. July 2005.

http://www.compete.org/hpc/hpc_software_survey.asp

National e-Science Centre

<http://www.nesc.ac.uk/>

Survey of U.S. HPC Industrial Users

E. Joseph, C. Willard, A. Snell, Sponsored by: Defense Advanced Research Projects Agency, Council on Competitiveness, IDC, Framingham Mass. July 2004.

http://www.compete.org/pdf/HPC_Users_Survey.pdf

Globus Open-Source Software Toolkit

http://www.globus.org/grid_software/

Processor power will increase by reductions in gate sizes from micron-scale to nanoscale.

Given the continuing progress in reducing scale and increasing performance, computer processors which have gates or junctions around one micron in size will be much tinier, measured in nanometers instead of microns, (a thousand times smaller). Processors with gate sizes approaching the size of a single molecule will be increasingly powerful and use less power for each computer instruction. Over time all processors will function at a nanometric scale. Instead of micron-scale processors, these much tinier nanoscale processors will be widely adopted for general computing in most parts of the world.

Molecular scale, or nanoscale, computing will likely be achieved in a variety of ways: electrically, mechanically, chemically, and eventually using quantum effects at significantly higher speeds, smaller scale, and at lower costs than micron-scale transistors. Transistors—or something like transistors—will still be employed for essentially the same computational functions—1s and 0s for memory storage and logic operations. First-generation nanoscale transistors will perhaps be carbon nanotube-based. Later on, quantum dots and other forms of quantum computation, offering even more massive improvements in performance, miniaturization, and cost reduction may come into use, although the technologies for processing information on the individual dots and how to read from the dots remains to be determined.

Forecast

Revenue-producing multiple core processors based at near-nanoscale will begin by about 2015, but much later for quantum computation using carbon nanotubes.

Implications

In the long-term future, it is likely that computers will be extremely application specific. That is, each molecular component employed will address only one specific problem. In short, there need not be one form of computation in 2030 (as there is today with the domination of the transistor paradigm of the IC industry). Rather, nanotubes, quantum computing, and molecular computation might each serve their own purposes relative to their material properties and associated advantages.



Innovators

- Research is underway in nanocomputing at universities and private labs working on mechanical nanocomputers, electronic nanocomputers, chemical molecular computers, and quantum computers. Progress is continual with fairly frequent, promising break throughs. Researchers at Oxford and Cambridge are working on all aspects of quantum computing including heuristics and algorithms.
- The U.S. National Institute of Standards (NIST), along with colleagues from New Zealand and Germany, are working on nanoscale quantum computing hardware focusing on a test suite implementing very difficult mathematics.

RESOURCES

Quantik

Lists continuing updates and new developments in quantum technologies.
www.quantiki.org

Semiclassical Quantum, Fourier Transform in a Scalable System

www.google.com/url?sa=U&start=3&q=http://www.boulder.nist.gov/timefreq/

Small Wonders, Endless Frontiers

A Review of the National Nanotechnology Initiative, Committee for the Review of the National Nanotechnology Initiative Division on Engineering and Physical Sciences, National Research Council. National Academy Press, Washington D.C. 2005.

Molecular-scale, DNA based processes will be developed for computation.

DNA computing is the use of biological molecules for computational devices. Computational problems are encoded into DNA-like strings, which are then mixed into a test tube. The nucleotide matching properties guide the association of these strings into DNA strands, which provide the solution to a particular problem.

Forecast

The first practical DNA computing devices will be demonstrated in 2015—devices that might be offered later as production computational resources.

Implications

A DNA computer will eventually be a cheap and powerful massively parallel problem-solving machine and potentially capable of combinatorial optimization, molecular nanomemory with fast associative search, AI problem solving, medical diagnosis, drug discovery, cryptography, and biocomputing applications like processing of DNA labeled with digital data, and for genetic-sequence comparisons.

DNA computing's impact will be most profound in the life sciences, medicine, health, and agriculture. DNA biocomputers could give humanity capabilities to manipulate both the behavior and physical forms of living cells at microscopic levels. As a result, a whole spectrum of new man-made life forms and medicines becomes possible, perhaps bringing the eradication of many diseases and the creation of engineered super organisms—plants, animals, and humans.

Research efforts in DNA computing are benefiting greatly from research to fulfill demands and applications of genomic sciences and nanotechnologies. Rapid progress may be achieved due to the massive economic challenges of global health and food production coupled with continuously growing demand for computing power.

Software development that takes advantage of massively parallel processing capabilities is the gating factor for practical implementation of DNA and molecular computing. But given the focus on DNA computing in universities in the United Kingdom, Israel, the United States, and elsewhere, some reasonable programming tools may be available by the time “bio-ware” (software for biological computers) is developed.



Innovators

- Israeli scientists have devised an experimental DNA computer that can perform 330 trillion operations per second, more than 100,000 times the speed of the fastest PC.
- In 2002, researchers from the Weizman Institute of Science in Rehovot, Israel demonstrated a programmable molecular computing machine composed of enzymes and DNA molecules instead of silicon microchips.
- In 2003 the same Israeli team demonstrated a new device, the single DNA molecule that provides the computer with the input data also provides all the necessary fuel.

RESOURCES

Computer Made from DNA and Enzymes

Stefan Lovgren for National Geographic News, February 24, 2003.

http://news.nationalgeographic.com/news/2003/02/0224_030224_DNAcomputer.html

Molecular Computing: An Overview

Byoung-Tak Zhang, Biointelligence Laboratory, School of Computer Science and Engineering, Seoul National University March 13, 2002

http://bi.snu.ac.kr/Courses/g-ai02/materials/DNAC_tutorial.pdf

Tiny Computing Machine Fueled by DNA

<http://news.nanoapex.com/modules.php?name=News&file=article&sid=3135>

Quantum phenomena may be harnessed for computation.

Simon Bone and Matias Castro from the Imperial College in London provide a concise explanation of quantum computing in their work: “A Brief History of Quantum Computing.”

In the classical model of a computer, the most fundamental building block, the bit, can only exist in one of two distinct states, a 0 or a 1. In a quantum computer, the rules are changed. Not only can a “quantum bit,” usually referred to as a “qubit,” exist in the classical 0 and 1 states, it can also be in a coherent superposition of both. When a qubit is in this state it can be thought of as existing in two universes, as a 0 in one universe and as a 1 in the other. An operation on such a qubit effectively acts on both values at the same time.

Despite the simplicity of this explanation, even many computer science experts have a hard time understanding the underlying principles, and resulting implications.

Forecast

Quantum computing is not a foregone conclusion, as enormously difficult theoretical breakthroughs are required to make it a reality. However, if Moore’s Law continues to hold, one bit of information will be encoded into a single atom by the year 2017, and by around 2012 quantum effects in nanoscale processes in normal chip architectures could become very important, functional phenomena perhaps even rendering quantum computation necessary. And according to researchers at the University of California, Berkeley, “revenue-producing products will likely be available using carbon nanotubes in 2020 for quantum computation in 2040–2060.” Implementation of quantum computing will make computation extremely fast, potentially trillions of times faster—and more secure using encryption techniques that are unbreakable.



Implications

Because of the almost unimaginable scale of possible performance of quantum computing in the number of instructions that can be executed simultaneously and in parallel, a whole range of computationally intensive tasks previously impossible, from image understanding and real-time speech recognition to unbreakable codes and extremely compact data and media compression, could become common.

Innovators

- The Centres for Quantum Computing at Oxford and Cambridge is lead by David Deutsch (Oxford University) who in 1985 described how the quantum Turing machine might be built, in principle, and how the “superposition” of 0s and 1s simultaneously led to quantum parallelism.

RESOURCES

[A Brief History of Quantum Computing](#)

Simon Bone and Matias Castro, Imperial College in London, 1997.

http://www.doc.ic.ac.uk/~nd/surprise_97/journal/vol4/spb3/#1.1%20Quantum%20computer%20basics

[Feynman, Einstein, and Quantum Computing](#)

<http://webcast.cern.ch/Projects/WebLectureArchive/cern/lectures/academ/2001/quantum/1/rawdata/hey1.ppt>

[Simulating Physics with Computers](#)

Richard Feynman, *International Journal of Theoretical Physics*, 1982.

<http://citeseer.ist.psu.edu/feynman82simulating.html>

[The Centres for Quantum Computation at Oxford and Cambridge](#)

<http://www.qubit.org/>

CHALLENGES AHEAD

Almost all new computer architectures are inherently parallel, and so will require the development and adoption of easy-to-use parallel programming tools. Parallel programming, programming for hundreds or thousands of concurrent independent processes or “threads,” will necessarily become the dominant paradigm for software development for most computers excepting some microscale processors used for mobile or embedded devices.

The system software used on most contemporary supercomputers is a crude variant of UNIX; most commonly, with limited programs written in Fortran, C, and C++ augmented with a few language or library extensions for parallelism and application libraries like the Titanium extensions to Java for higher performance computing. Even though some early software tools exist, programming supercomputer software supporting massive parallelism requires special expertise, for optimizing performance at multiple layered levels below the operating system. Modern mainstream application programmers deeply knowledgeable about their application domains have little or no practical knowledge of using massive parallel processes to generate better results or user experiences, and have no high-level tools to complete their goals, while masking complexity of lower level, massively parallel software and hardware processes. Indeed, applications designers have not yet been educated in how to think about discrete computational tasks to take advantage of enormous computing machinery that will be widely available starting around 2015. But things are changing. Already institutions like the Centre for Parallel Computing at Massey University in New Zealand and the University of California at Berkeley have launched pilot programs to begin teaching design of massively parallel applications.

Over the next 10–20 years research laboratories, and software companies will therefore necessarily develop new tools for programmers to harness the power of ever increasing computer hardware, and educational institutions will increase instruction in effective uses of parallel processing for application design. By 2015, computation will be based on many machine models, supported by virtual computers of using multiple operating systems selected by applications on an as-needed basis to resolve task-specific requirements.

Applications designers will have to be capable of determining the utility and cost of a solution depending on expert judgments on factors other than time taken—for instance, on accuracy or trustworthiness. According to a recent National Academy of Sciences study, “Determining the trade-off among these factors is a critical task. The calculation depends on many things—the algorithms that are used, the hardware and software platforms, the software that realizes the application and that communicates the results to users ... the design of the algorithms, the computing platform, and the software environment governs performance and sometimes the feasibility of getting a solution.” All of these are dependent on human expertise to orchestrate the overall effectiveness of implementation.



A higher level of abstraction and/or a more restricted model of parallelism are essential to be able to comprehend the behavior of a large parallel code, debug it, and tune it. It is not possible to understand the behavior of 10,000 concurrent threads that may interact in unexpected ways.”

Whether multi-core, microscale, nanoscale, or distributed grids, properly designed applications for massively parallel computing architectures will be able to perform bigger and more complex tasks, on larger and large data sets, and to perform ordinary computation faster and more accurately.

To date, in addition to secret government signal-sensing and cryptographic applications, petroleum, automotive, and aircraft companies and pharmaceutical and biomedical startups are driving demand for massively parallel software. And increasingly media companies like Industrial Light and Magic, Dreamworks, and Pixar Productions are using massively parallel processes to render movie graphics. Over time these photorealistic capabilities will migrate to interactive entertainment like games, and other new high-resolution media.

The National Academy of Sciences study concludes “Advances in algorithms and in software technology at all levels are essential to further progress in solving applications problems using supercomputing. Supercomputing software, algorithms, and hardware are closely bound. As architectures change, new software solutions are needed. If architectural choices are made without considering software and algorithms, the resulting system may be unsatisfactory. Because a supercomputing system is a kind of ecosystem, significant changes are both disruptive and expensive. Attention must therefore be paid to all aspects of the ecosystem and to their interactions when developing future generations of supercomputers. Educated and skilled people are an important part of the supercomputing ecosystem. Supercomputing experts need a mix of specialized knowledge in the applications with which they work and in the various supercomputing technologies.”

RESOURCES

Compute-Intensive, Highly Parallel Applications and Uses

Intel Technology Journal Volume 09 Issue 02 Published May 19, 2005.
<http://www.intel.com/technology/itj/2005/volume09issue02/foreword.htm>

MPI: A Message-Passing Interface Standard

<http://www.mpi-forum.org/docs/mpi-11-html/mpi-report.html>

Parallel Computer Architectures

CS 258, David E. Culler, Computer Science Division, UC Berkeley.
<http://www.cs.berkeley.edu/~culler/cs258-s99/>

Parallel Programming Models and Paradigms

Grid Computing and Distributed Systems (GRIDS) Lab, University of Melbourne, Australia
<http://www.buyya.com/cluster/v2chap1.pdf>

3 | FUTURE APPLICATIONS OF ABUNDANT COMPUTING

Here we present seven possible future uses for abundant computing for the next decade. This is not intended to be a comprehensive inventory of all the future impacts of this computing revolution, but rather to provide ideas that provoke thinking and insight into the role that abundant computing could play in future organizations, markets, and the lives of consumers.

These potential applications share three key assumptions from our technology assessment.

1. Applications will be platform and technology agnostic.

Many technologies will compete to deliver abundant computing in the world ten years from now—multi-CPU processor chips, grid computers, and embedded constellations of smart sensor networks. However, when performing properly for the end user, these technologies will all present essentially the same performance to users. Except for a small subset of latency-tolerant applications that will be designed to operate on the fragile networks of the developing world, we believe that both local and grid approaches will deliver similar levels of computing capability to the individual user.

2. The development of new software techniques will be the gating factor on applications.

Abundant computing will require drastically new approaches to programming and software development. A growing gap will develop between the capabilities of advanced computers and their everyday use. Over time, a growing portion of this excess computational capability will go unused.

3. Pattern matching and recognition will be the core capability around which use scenarios will be built.

While realistic graphics rendering and realistic modeling of the physical world will drive the early wave of applications, over time the proliferation of network-accessible data and environmental sensors will shape the future of abundant-computing applications. Brute-force approaches to pattern matching and recognition will drive the development of computationally intensive context-awareness: a world of computers sensing and making sense through grinding combinatorials.



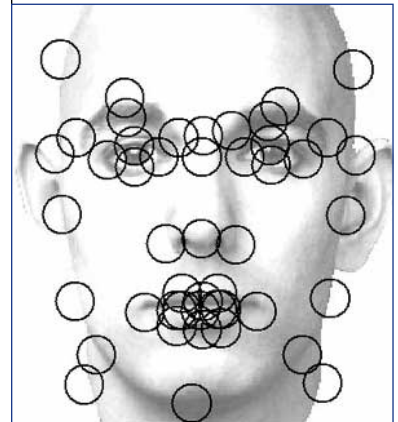
1. INTERFACES | Natural Computing

Today's computers devote the lion's share of their processing power to rendering and sensing the graphical user interface (GUI). However, the limits of this form of human-computer interaction are becoming increasingly clear as the demands of mobility, information visualization, and the convergence of communications and computing erode the desktop PC's dominance.

Personal and embedded computers of the future will leverage abundant-computing capabilities to provide more natural interfaces for computing. Computationally intensive facial recognition and speech-recognition processes will be embedded in many systems, so that computers will identify us and understand our intentions and instructions. Gestures may replace tactile input devices such as keyboards and the mouse.

One potential wild card will be the emergence of real-time, machine language translation across many media. Already, Minnesota-based SpeechGear, Inc. provides software that translates text from digital photographs, allowing tourists take photographs of street signs and receive real-time translation. In the next decade, this technology strand will merge with improved speech recognition, leveraging the supercomputing powers enabled by abundant computing.

Facial Recognition Is Computationally-Intensive



Source: Worcester Polytechnic Institute

Real-Time Multi-Media Language Translation



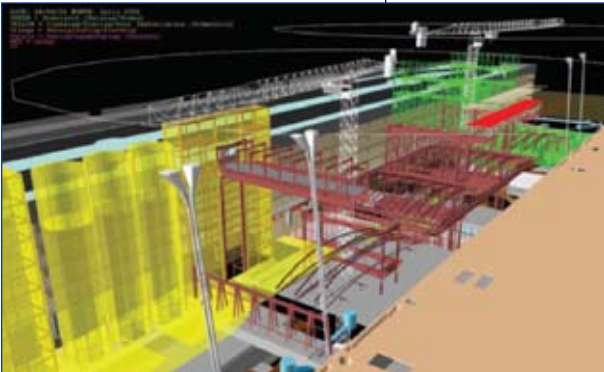
Source: SpeechGear, Inc.

2. SPACE AND PLACE | Intelligent Environments That Get Smarter over Time

Abundant computation will transform the way we design and manage the cities, homes, and public spaces that we inhabit every day. Today, we think of these places as fixed, static, and dumb. However, Paul Seletsky, Director of Digital Design at global architecture firm SOM, believes that in the next decade, all new buildings will begin and end their lives as a computer model. This model will guide the design process, simulate building operation, and be used to integrate sensing and operations during the building's working life.

Software will in many ways become “personality” of buildings, homes, and public spaces—and its capabilities and shortcomings will define the experience of living, working, and playing within them as much as the physical structure itself. Buildings will summon and manage people, robots, and resources to sustain themselves and their inhabitants. They will employ agent-based models of occupants' activities to predict everything from lunchtime congestion on elevators to evacuation patterns during emergencies. As they age, intelligent environments will grow and develop with the human communities that inhabit them.

4D Model for London Heathrow's Terminal 5



ARUP's 4D model for London Heathrow's Terminal 5 provides a hint of the future of modeling in building life-cycles. By adding time as a new dimension to CAD, this design technology is opening the door for software to develop new intelligence at the building level.

Source: <http://www.arup.com/DOWNLOADBANK/download498.pdf>



3. IMMERSIVE MEDIA | Entertainment, Education, And Communication

The pursuit of ever-more realistic computer graphics has long been a primary driver of demand for computational horsepower. In the future, abundant computing will be pushed to its limits by developers of deep, immersive gaming experiences that blur the boundary between real and virtual lives.

Abundant computing will enable new levels of immersive reality that allow people to play together in photorealistic, shared online game spaces. Full-body interaction with computers through motion detection and gaze sensing will render these experiences increasingly life-like. As immersive media spreads, it will have widespread social and economic impacts. The popularity of *World of Warcraft*, the most successful massively-multiplayer online game (MMOG) today, is an early indicator of how new forms of social interaction and cooperative problem solving could change the nature of team collaboration in the workplace. Immersive media will transform the nature of education as simulation and graphic visualization replaces written and oral forms of literacy among young people. Instead of just reading or writing about new ideas, children will be able to experience anything and employ immersive media as rapid prototyping tools for interactive communications.

Familiarity with immersive media in education and entertainment will also push telepresence beyond the glass ceiling that currently limits its use as a substitute for physical presence in business communications. Immersive media will pave the way for a future workplace dominated by multi-sensory, multi-channel telepresence.

Augmented Perception Is Built on Precision Positioning of High-Resolution Graphic Objects in Real Space



Source: Steven K. Feiner, Columbia University

An Artist's Photoshop Mashup of a Speculative 3D Television



Source: http://www.worth1000.com/entries/114500/114781AFHE_w.jpg

McLaren's Designs Are Developed Using a Mini Grid-Computing System That Creates a Virtual Supercomputer



Companies like McLaren International, the designer of Formula One race cars, are creating mini grid-computing networks of dozens of workstations that allow for supercomputing strength in even the smallest of engineering design enterprises.

Source: http://www.deskeng.com/articles/02/aug/feature1/main_dig.htm

4. R&D | Democratized Supercomputing Transforms the Global Innovation Landscape

For much of the last 50 years, massive computational resources were the exclusive province of advanced supercomputing and high-energy physics research centers in the world's superpowers. "Big iron"—as the fastest supercomputers were affectionately called—were expensive, finicky devices that required experts to build and operate them.

In the 1990s, experiments in distributed computing aggregated massive networks of personal computers to tackle complex problems like modeling protein folding and factoring very large prime numbers. However, in the next decade, multi-core processors and grid computing will dramatically democratize access to supercomputing power for research and development.

The impact will be felt most powerfully in fields such as life sciences and social sciences where major theoretical strides in modeling and simulation techniques are demanding massive computational power. Desktop and on-demand supercomputing could also usher in a wave of grassroots innovation in biotechnology, economic and financial forecasting, and molecular design. Grid computing will provide access to unprecedented levels of computational power for the developing world, unleashing an innovation revolution.



5. NETWORKS | Social Awareness

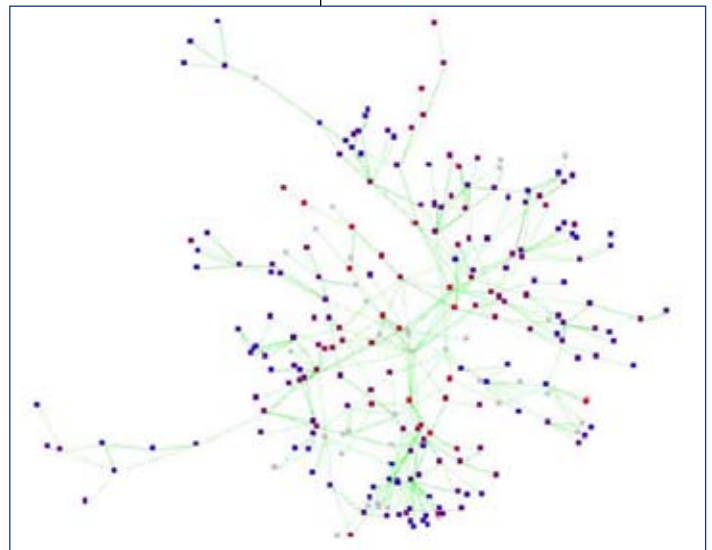
Google's success over the last five years has largely come from its combination of simple but clever algorithms like PageRank and massive distributed computing power to bring context to the mass of information on the Web. PageRank is, at its heart, a way of measuring social context by counting the number of links to a page that have been inserted by other Web authors.

Abundant computing will unleash a Google in every computer, and turn the mining of contextual data and patterns from social networks into an everyday application. The next generation of data mining will not be about information and documentation, but about people and their associations and interests.

This use of abundant computing to derive understanding of social networks and social context from information on the Web will have widespread applications:

- Organizations will mine their own social structure. In 2001, researchers at HP Labs showed how computer analysis of e-mail logs could be used to map important hidden social connections and networks that transcended formal organizational structures.
- Governments will mine social data about their own citizens. The National Security Agency has reportedly created a database of every phone call every made in the United States to be used to find patterns of social interaction to aid anti-terror investigations. The Chinese government is building a socially omniscient, state-censored Internet that relies heavily on pattern-matching that will become more sophisticated over time.
- Consumers will mine the social web in a variety of contexts. The web is accumulating tagged data about social networks and the people in them, which will be mined and searched by a variety of applications targeted for consumer use.

Social E-Mail Networks in HP Labs



Source: Joshua R Tyler, Dennis M. Wilkinson, and Bernardo A. Huberman. "E-mail as Spectroscopy: Automated Discovery of Community Structure Within Organizations" HP Labs (Palo Alto, CA).

6. YOUTH | Thinking Companions

Abundant computation will have a dramatic impact on today's children, who will be exposed to lifelike intelligence from a young age, both in education and play. We saw a glimpse of this future in Steven Spielberg's 2001 film *AI*. In that film, David (played by Haley Joel Osment) was accompanied and helped by an animatronic robot companion named Teddy. Teddy offered advice, warnings, and emotional support in a non-threatening way and in every way operated as a thinking companion.

Such thinking companions for children are likely to take many derivative forms of devices and services we see on the market today: robots that follow or are carried, virtual pets that are accessed through networks and immersive media, and glyphic media that interact with real-world objects and spaces. Thinking companions may also serve as avatars or simulacra of deceased friends and family members returning as virtual ancestors to teach and guide the child. Increasingly, abundant computing will offer the potential for thinking companions to grow and evolve with their masters throughout childhood and even into adult life.

Teddy Served as a Thinking Companion



Source: http://www.geocities.com/ravager_of_worlds/TeddyAI2.jpg



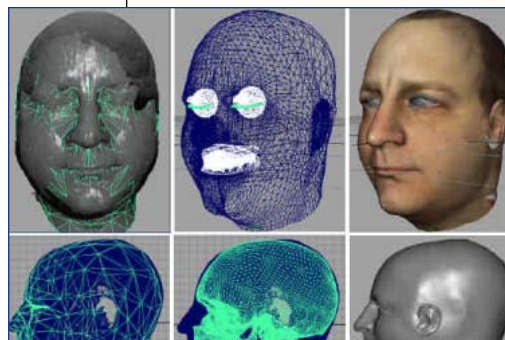
7. HEALTH | Personal Health Simulations

Abundant computation will enable consumers to run powerful simulations that help them make better, more-informed decisions about their everyday personal lives. From deciding whether to drive the car or take the train to work, to whether or not to bring an umbrella, these personal simulations will integrate distributed sensory inputs with computationally intensive modeling of the real world.

Health care will quickly emerge as a dominant application of abundant computing to personal simulation. Pharmacies and other point-of-sale and primary treatment centers will seed simulations with data from an intensive scan, using various CAT/PET/MRI equipment (themselves driven down in price by cheaper computing) to build a model of the user's body. Various implants would then provide continuous, real-time feedback to the model on stress, heart rate, blood chemistry, and so on. Other relevant data from health histories, personal genetic code, and even calendars would also be incorporated.

The health simulation would provide direct feedback to influence the user's behavior. For instance, it might provide statistical projections on the likelihood of heart disease or other negative events to discourage unhealthy behavior. It could even provide real-time feedback during a workout to help the user optimize the effectiveness of her workout.


Simulations Are Already Being Used by Plastic Surgeons



Source: http://www.digitalelite.net/Pages/Papers/DEE_SPIE2004_MRIPlasticSurgery.pdf

POST-SCRIPT: GEOGRAPHIC DIFFUSION AND CULTURAL NUANCES

The future of abundant computing is not likely to develop in a uniform fashion around the world. Local cultural practices, the state of broadband infrastructure, and consumer trends in devices will all heavily influence the form of abundant computing around the world in ten years. Specifically, different paradigms for abundant computing will dominate among major world regions.



In the United States, the large installed base of desktop personal computers, and the relatively slower pace of broadband infrastructure connecting the home will drive more computation to the PC. Entertainment and telepresence applications will dominate.

In Europe, scientific applications including visualizations and simulations are likely to drive adoption of abundant computation, most notably computational grids and clusters among academic and enterprise users.

In Asia, rapid deployment and continuous upgrading of wireless broadband infrastructures and the popularity and rapid product cycle of mobile devices will favor an on-demand, grid-based approach that delivers massive computation to small, portable devices. A general lack of resistance to pervasive, integrated sensing of consumers will enable applications that are likely to be socially unacceptable outside this region.

In Africa and much of the developing world, the lack of PCs and growing base of mobile phones will also dictate a grid-based approach to delivering abundant computing. However, poor communications infrastructure will require more latency-tolerant systems.

ABOUT THE ...

Technology Horizons Program

The Technology Horizons Program combines a deep understanding of technology and societal forces to identify and evaluate discontinuities and innovations in the next three to ten years. We help organizations develop insights and strategic tools to better position themselves for the future. Our approach to technology forecasting is unique—we put humans in the middle of our forecasts. Understanding humans as consumers, workers, householders, and community members allows IFTF to help companies look beyond technical feasibility to identify the value in new technologies, forecast adoption and diffusion patterns, and discover new market opportunities and threats.

The Institute for the Future

The Institute for the Future is an independent, nonprofit strategic research group with nearly 40 years of forecasting experience. The core of our work is identifying emerging trends and discontinuities that will transform global society and the global marketplace. We provide our members with insights into business strategy, design process, innovation, and social dilemmas. Our research generates the foresight needed to create insights that lead to action. Our research spans a broad territory of deeply transformative trends, from health and health care to technology, the workplace, and human identity. The Institute for the Future is located in Palo Alto, California.

©2007 Institute for the Future. All rights reserved. Reproduction is prohibited without written permission.

TECHNOLOGY HORIZONS PROGRAM
MAY 2007
SR-1042

Authors:

Michael Liebhold and
Anthony Townsend

Editor:

Maureen Davis

Art director:

Jean Hagan

Graphic designers:

Robin Bogott and
Yeshai Ling