JOINT

U.S. Defense Science Board
UK Defence Scientific Advisory Council

TASK FORCE

on

DEFENSE CRITICAL TECHNOLOGIES

March 2006

Office of the Under Secretary of Defense
For Acquisition, Technology,
and Logistics
Washington, D.C.
20301-3140

Defence Scientific Advisory Council,
Level 1, Zone J,
MoD Main Building, Whitehall,
London,SW1A 2HB,
United Kingdom

This report is a product of the Defense Science Board (DSB) and the Defence Scientific Advisory Council (DSAC).

The DSB is a Federal Advisory Committee established to provide independent advice to the Secretary of Defense. Statements, opinions, conclusions and recommendations in this report do not necessarily represent the official position of the Department of Defense.

DSAC is an advisory Non-Departmental Public Body established to provide independent advice to the Secretary of State for Defence in the fields of science, engineering and technology. The report does not necessarily represent the views of the Ministry of Defence or its agencies.

This report is unclassified.

Defence Scientific Advisory Council (DSAC)
MoD Main Building (01.J.06)
Whitehall
London
SW1A 2HB

Defense Science Board OUSD (AT&L)
3140 Defense Pentagon
Room 3D865
Washington
DC 20301-3140

MEMORANDUM TO UNDER SECRETARY OF DEFENSE FOR ACQUISITION, TECHNOLOGY AND LOGISTICS
SCIENCE AND TECHNOLOGY DIRECTOR, MINISTRY OF DEFENCE

SUBJECT: Report of the Defense Science Board and Defence Scientific Advisory Council Joint working party on Critical Technologies.

We are pleased to forward the final report of the Defense Science Board (DSB) and Defence Scientific Advisory Council (DSAC) joint working party on Critical Technologies. This is the first collaborative science board effort between the U.S. Department of Defense and the United Kingdom's Ministry of Defence. In this report, the joint working party examines five major transformational technology areas that are critical to meeting the defence needs of the United States and the United Kingdom. These technology areas are:

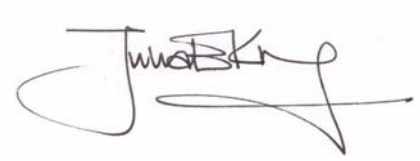
- Advanced Command Environments
- Persistent Surveillance
- Power Sources/Management For Small, Distributed Networked Sensors
- High Performance Computing
- Defence Critical Electronic Components

The working party concluded that the U.S. and UK lead in critical technologies is under threat and that commercial off-the-shelf technology is insufficient to meet U.S. and UK defence needs in these areas. The working party finds that government investment is essential to ensure that technological development translates into military capabilities and that it is in the interests of the United States and the United Kingdom to collaborate on selected critical technologies in order to help the DoD and MoD retain or increase their competitive advantage over potential adversaries.

In addition, the report presents the observations and lessons learned from the collaborative joint working party process. The working party co-chairs recommend that, based on the positive experience of the members, the DSAC and the DSB collaborate further on joint studies exploring in-depth, focused areas where U.S. and UK perspectives differ.

We endorse all the recommendations of the working party and encourage you to read their report.

William Schneider, Jr.
Chair, Defense Science Board



Julia E. King
Chair, Defence Scientific Advisory Council

OFFICE OF THE SECRETARY OF DEFENSE
3140 DEFENSE PENTAGON
WASHINGTON, DC 20301-3140

DEFENSE SCIENCE
BOARD

MEMORANDUM FOR CHAIRS, DEFENSE SCIENCE BOARD AND DEFENCE SCIENTIFIC ADVISORY COUNCIL

SUBJECT: Report of the Defense Science Board and Defence Scientific Advisory Council Joint Task Force on Critical Technologies.

The attached report is the result of a joint U.S. and UK working party on Critical Technologies. The study examined five transformational technology areas that are critical to meeting the defense needs of the United States and the United Kingdom. These technology areas included two broad, high-level topics (Advanced Command Environments and Persistent Surveillance) and three focused technology topics (Power Sources/Management for Small Distributed Networked Sensors, High Performance Computing and Defence Critical Electronic Components).

The working party was organized into five separate panels to examine each of these technology areas. The panels found common ground on the two broad technology areas, while the panels focusing on the three specific technology topics agreed to pursue their topics relatively independently and then integrated and harmonized their results at the end of the study. The following are the working party's principal findings and recommendations.

Commercial Technology Cannot Ensure Military Capability

All panels found opportunitites for government investment and strategic development that could translate new technologies into military capabilities. The panel findings suggest that commercial off-the-shelf (COTS) technologies will be insufficient to meet unique military needs and that the DoD and MoD should invest to accelerate technological development or to address technology gaps where there is no current commercial market. The panel findings also suggest that the availability of COTS technologies to adversaries further increases the need for DoD and MoD to invest in critical, defense-niche technologies in order to assure competitive advantage over potential adversaries.

U.S. and UK Lead in Critical Technologies is Under Threat

A combination of the U.S. consolidation of the defense contractor base, the migration off-shore of some critical manufacturing and design capabilities, and the reduction in the number of engineers with experience in critical areas all contribute to the erosion of the U.S. and UK lead in key technologies. The working party findings indicate a need for the DoD and MoD to not only assure their lead in critical technologies, but to reduce the acquisition time for intelligence, command and control, and weapons systems, in order to be prepared to deal with COTS-equipped adversaries.

U.S. and UK Collaboration for Delivering Military Capability

The working party concludes that it is in the interest of the United States and the United Kingdom to collaborate on selected, critical technologies. The working party finds that the value of multi-national collaboration comes from engagement of individuals that bring different perspectives on mission, on technology, on the industrial base available to a nation, and on the assumptions that flow from the scale of the technology efforts contemplated. The working party co-chairs recommend that, based on the positive experience of the members, the DSAC and DSB collaborate further on selected, joint studies where both organizations bring a useful perspective to the other. Further, co-chairs emphasize the importance of early in-person meetings between U.S. and UK members. Topics should be explored in depth involving larger panels in order to increase the diversity and richness of perspectives contributed.

We wish to thank the U.S. and UK members for their dedication and hard work in addressing these complex issues.

Hon. Anita Jones
Task Force Co-Chair

Admiral William Studeman, USN (Ret)
Task Force Co-Chair

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This joint study by scientific advisory boards that advise the Department of Defense (DoD) and the Ministry of Defence (MoD) is an effort to deepen the cooperation of the two organisations as they both pursue a strategy of investing in and developing technology to achieve military advantage. The Defense Science Board (DSB) and the Defence Science Advisory Council (DSAC) undertook this collaborative study to explore transformational technologies that are critical to meeting national defence needs. The joint working party was led by three co-chairs: Dr. Anita Jones and Admiral William O. Studeman, USN (Ret) to lead the U.S. team, and Dr. Julia King to lead the UK team.

The joint tasking statement, the terms of reference, appears in Appendix A. Members of the U.S. team are listed in Appendix B. It is the policy of the MoD not to publish the names of their advisors. The MoD re-considered its policy in the light of this joint study and concluded that the policy should be sustained. Consequently, the names of UK members do not appear in Appendix B. They are accomplished scientists and technologists, with academic and industrial experience in areas related to defence.

Because convening a joint working party is a new mode of cooperation for the two organisations, the chairs decided to focus on both broad and narrowly defined technology areas. They selected two broad mission areas (Advanced Command Environments and Persistent Surveillance) and three focused technology areas (Power Sources/Management for Small Distributed Networked Sensors, Defence Critical Electronic Components and High Performance Computing). The working party was organized into five topic panels with four to five members in each country drawing on a wider circle of outside experts as needed. The U.S. and UK members of each topic panel conducted separate but collaborative studies. Recommendations from the joint panels appear at the end of the topic chapters.

A key working party objective was to identify lessons learned from the collaborative process that can be applied to any future joint studies. Our conclusions on cooperation are:

- We recommend future collaboration between the Defense Science Board and the Defence Science Advisory Council. It is fruitful.
  - Topics should be narrowly defined and limited in scope.
  - Collaboration requires face to face meetings. While technology facilitated exchange is helpful, achieving collective understanding, and collaborating on complex topics necessitates personal interchange. In particular, face to face kick off meetings between the U.S. and UK teams would be very valuable.
  - Differences in style and approach need to be worked out early in the collaboration. Face to face meetings will aid in reaching consensus on how to manage the joint working processes.
- We recommend that the MoD and DoD consider ways to mitigate the shortage of U.S. and UK nationals opting to take undergraduate and higher degrees in science, engineering and technology and recruit more young scientists and engineers into defence research. In addition to national programs,
  - Establish collaborative programmes involving opportunities to train at universities in both the United States and the United Kingdom.
  - Create opportunities to work in U.S. and UK laboratories to show the importance of, the excitement and challenges offered by, working in defence research areas.

Overall the experience for the members of both boards was positive. Our militaries have unique missions that are not sufficiently served by commercially available technology. Cooperative and complementary technology development serves both nations. As the United Kingdom and the United States increasingly join in coalition operations with each other, as well as, with other nations, coalition considerations need to be considered early in the exploitation of technology. Joint DSB and DSAC studies could aid in this consideration.

The key recommendations of each of the sub-groups working in the five areas are summarised in the following table. Detailed recommendations can be found at the end of each chapter.

Chapter topic	Recommendations and key points concerning
Power Sources (Chapter 2)	▪ Develop system and power source metrics to meet military and intelligence requirements. ▪ Leverage advances from commercial and industry sectors in areas where rapid change will happen without investment from the defence and intelligence communities. ▪ Focus on areas where there is little commercial investment, but where concepts, material and design tools will be critical to the successful deployment of distributed sensor networks. ▪ Develop remotely read, unpowered nodes as a potential solution to the power problem.
High Performance Computing (Chapter 3)	▪ Fund the Defense Advanced Research Projects Agency's High Productivity Computing System programme robustly. ▪ Initiate a UK HPC programme to perform research on the most demanding military applications to improve performance using commodity clusters. ▪ Invest in research critical applications and technologies.
Defence Critical Electronic Components (Chapter 4)	▪ Maintain U.S. leadership in semi-conductor technologies critical to national defence: o Develop computer aided design tools, o Maintain U.S. lead in dual-use technologies, o Expand and continue trusted foundry initiative, and o Develop joint DoD-MoD technology in areas that may provide new capabilities for defence systems but have limited commercial use. ▪ Re-evaluate the U.S. Export Administration Regulations. ▪ Initiate studies to understand strengths, weaknesses, and vulnerabilities of COTS based systems to counter COTS-equipped adversaries. ▪ Forge coalitions with other sectors to find a common set of requirements to meet reliability, openness, and quality needs for COTS parts. ▪ DoD and MoD conduct longitudinal analysis of the emergence of novel electronics to determine whether or not the "discovery engine" has slowed down.
Advanced Command Environments (Chapter 5)	▪ Develop a cooperative U.S./UK programme to collaborate on physical design aspects, internal functionality and tools, and other human factors related to optimizing future command/decision environments: o Develop a trial link between U.S. and UK facilities to host an initial set of experiments and

	o Sponsor a conference with a call for papers to address a broad range of topics related to Advance Command Environments.
Persistent Surveillance (Chapter 6)	▪ Advance integrated sensing. ▪ Further horizontal knowledge integration. ▪ Establish U.S. Persistent Surveillance effort/office or DABINETT counterpart. ▪ U.S. should review the UK DABINETT model/approach as potential way to go forward. ▪ Devise a coordinated UK/U.S. approach, where combined approach has advantage

This page intentionally left blank.

CHAPTER 1. INTRODUCTORY OVERVIEW

ORIGINS

Both the Department of Defense (DoD) and the Ministry of Defence (MoD) pursue a strategy of investing in and developing technology, and then applying both commercial and military-unique technology to achieve military advantage. The armed forces of the United States and the United Kingdom have a long history of operating closely with one another in the field, as well as in the development and application of technology. Supporting this strategy of exploiting technology rapidly, both the DoD and the MoD enlist the advice of their respective technical advisory boards: the Defense Science Board (DSB) and the Defence Science Advisory Council (DSAC). This report is the product of the two organisations working together.

In an effort to deepen cooperation in areas of overlapping mutual interest, the DoD and MoD requested the DSB and the DSAC to undertake a collaborative study exploring transformational technologies that are critical to meeting national defence needs. The joint working party appointed three co-chairs to lead the study: Dr. Anita Jones and Admiral William O. Studeman, USN (Ret) to lead the U.S. team, and Dr. Julia King to lead the UK team.

OBJECTIVES AND CRITERIA

The objective of the working party was to identify technologies that would meet DoD and MoD requirements in several key areas. The working party considered the technologies identified in the terms of reference as a starting point from which to select topics for study. The working party used a broad definition of the term "technology" to encompass specific technologies, components, and processes.

For each of the areas, U.S. and UK chairs applied specific criteria in determining whether or not a technology should be included in the

study. First, the working party considered whether the technology area requires defence funding to meet military requirements, and whether it will produce a national security advantage. Second, the working party selected technology areas that are of specific interest to the United States and the United Kingdom. Third, the technology area had to inherently improve inter-operability and information sharing capabilities. Fourth, in order to maximize opportunity for UK and U.S. working party collaboration, the technology area had to avoid complex security and industry proprietary issues. Finally, the technology area had to be generally applicable at the nexus of the threats posed by weapons of mass destruction and terrorism.

This resulted in the inclusion of narrowly focused topics, such as Power, and broad ones such as Advanced Command Environments. Through collaborative discussion, the U.S. and UK chairs narrowed these areas to three specific technologies and two broader areas in which technology enables military advantage, but must be embedded into a system to extract value. Different missions and specific opportunities led to the selection of each topic. The two broad areas are:

- Advanced Command Environments, and
- Persistent Surveillance.

Specific technology areas are:

- Power Sources and Management for Small Distributed Networked Sensors,
- High Performance Computing, and
- Defence Critical Electronic Components.

The chairs assembled a panel of experts for each area. The panels explored the current state of each technology area, including its commercial and defence industry status and applications. U.S. members are listed in Appendix B. It is the policy of the MoD not to publish the names of their advisors. The MoD reconsidered its policy in the light of this joint study and concluded that the policy should be

sustained. Consequently, the names of the UK members are not listed in Appendix B.

The working party identified the key opportunities each technology area represents, and assessed what courses of action the U.S. and UK science and technology communities can take both separately and together in order to realize these opportunities. Finally, the working party examined where differences in U.S. and UK approaches to a technology either inhibit collaboration or provide an opportunity for complementary research.

The working party also reviewed various lists of disruptive, emerging technologies, and possible applications to exploit them. The working party drew on recent work conducted by the UK members on identifying emerging, disruptive technologies that may rapidly alter our current status quo. Several chapters of this report contain excerpts of specific technology lists related to the chapter topic.

The working party members were chosen for expertise in specific topic areas. Consequently, we did not feel that we had expertise to build a comprehensive list of disruptive, emerging technologies for all of national security. Instead, we focused on the specific topics chosen for study. The chapter on Electronic Components provides an interesting comparison. Table 1 provides a list of top 15 electronic technologies excerpted from the UK Defence Critical Technologies List. Three levels of priority are ascribed to the technologies. In contrast, Figure 2 gives a list of technologies (without priority) that were viewed as most critical for future space surveillance, as viewed by the National Reconnaissance Office Space Research and Development Industrial Base Study. The two lists indicate that experts will have different views. Also, the comparative lists show that if one views technology through the lens of a specific mission, and possibly specific system architecture in which the technology will be exploited, that the lists may differ.

So, while the terms of reference requested that the working party develop a methodology to identify unique defence technologies and to apply the methodology to develop a list of defence critical technology, we instead focused on just a few technology areas, used

existing technology lists to bring in the expert judgement of others, and explored the specific areas that were selected for study.

Lastly, because convening a joint working party was a new mode of cooperation, a key working party objective was to identify lessons learned from the collaborative process that can be applied to any future joint studies.

PROCESS

The working party was organized into five topic panels with four-five members in each country drawing on a wider circle of outside experts as needed. The U.S. and UK members of each topic panel conducted separate but collaborative studies. The panels met independently but cooperated throughout the study, sharing research and insights during video (which were poor quality, unclassified and frustrating) and tele-conferences to produce an integrated working party report. The Advanced Command Environments and Persistent Surveillance panels conducted visit exchanges for face to face collaboration and joint briefings.

The working party also held three plenary sessions during which the U.S. and UK working party members were connected via video-teleconference and exchanged updates on their progress to date.

The joint working party was directed to produce an unclassified report. While the working party agreed to try to overcome security classification restrictions on collaboration wherever possible, the U.S. and UK working parties agreed to independently examine in greater depth any area involving security issues. The U.S. and UK members also agreed to identify areas of opportunity for collaboration where classification issues presented an obstacle.

The working party obtained an International Traffic in Arms Regulations (ITAR) exception for the study, which granted DoD advisors and working party members the authority to share ITAR-controlled information with the United Kingdom for the purposes of the study. The U.S. working party executive secretary and the DoD

liaisons assigned to each panel were designated as exchange points of contact for all documents provided.

MAINTAINING LISTS OF CRITICAL, DISRUPTIVE AND EMERGING TECHNOLOGIES

The terms of reference for this study asked that we define a methodology for coming up with a list of critical/disruptive/ emerging technologies, and to define such a list. Along the way, we reinterpreted this task after discovering that both MoD UK and the DoD maintain respective lists of future technologies of high interest. We did review the DSAC prepared list applicable to the UK Research Acquisition Organisation, who are developing the MoD UK Research Programme. Likewise, we reviewed a similar Office of the Secretary of Defense list which is more limited to programs targeted for funding consideration. The Defense Advanced Research Projects Agency (DARPA) also has active programs defined and agreed in DARPA. The recent U.S. Intelligence Quadrennial Review known as the QICR Challenge has also recommended that the Assistant of the Director National Intelligence for Science and Technology, via the National Intelligence Science and Technology Counsel (which has extensive DoD participation), maintain such a list updated annually, working in cooperation with the National Intelligence Office for Science and Technology. It would be good if the Office, Director National Intelligence list, when developed, be cross-walked with the DoD list. In any case, methodologically, it is recommended that both the DoD and the MoD maintain such lists to be reviewed and updated annually, and that as much as possible, the United States and the United Kingdom compare and share their respective lists. The respective intelligence organisations can be involved where appropriate.

For the purposes of this report, we have elected to compile lists of critical technologies that applied only to the technology topics addressed by this report. They are contained in each of the technical chapters. Relatedly, we have deliberately not included the various DoD and MoD lists in this report.

STRUCTURE

In some cases, the U.S. and UK teams adopted different approaches to the study, shaped by differing perspectives. The DSAC members on the whole concentrated on individual technologies while the DSB, in addition, focused on mission, organisation, and system integration of technology. The DSB was also more concerned with U.S. technology bases, i.e. the laboratories and industries that develop military-unique technology and build systems that incorporate that technology. As a result, some chapters in the report represent a harmonization of different approaches.

The chapters contain the integrated findings and recommendations of the U.S. and UK panels. Chapter 2 discusses Power Sources and Management and argues that optimizing system performance, rather than that of individual components, is essential to the successful development of power efficient distributed networks. The chapter also highlights how power sources will continue to dominate the size and weight of systems and limit their lifetime, and points to the need for exploring less sophisticated but lower cost, smaller size, and higher reliability nodes that may solve the power problem.

In chapter 3, the U.S. and UK panels address different aspects of High Performance Computing. The U.S. group focused on defence priorities in very high scale or integrated High Performance Computing, while the UK panel explored technologies such as grid and cluster computing and applications. The panel findings call for MoD and DoD collaboration to initiate a UK High Performance Computing programme to complement existing DARPA activities.

In chapter 4, Defence Critical Electronic Components, the U.S. panel focused on industrial and political issues while the UK group adopted a bottom-up approach to analyze known military capability requirements. In addition to calling for increased UK and U.S. collaboration, the panel recommends an overhaul of export control regulations, which the panel finds is currently ineffective in denying

semiconductor technology to potential adversaries and in some cases encourages the development of foreign sources of critical technology.

Chapter 5 is the joint report of the Advanced Command Environments panel, which emphasizes the need to integrate the human factor into the development of complex technology and information systems. The U.S. and UK Advanced Command Environments panels initially differed in their approach; the U.S. team was interested in enabling technologies such as visualization, displays, and ergonomics, while the UK team sought to develop a framework to help determine whether a technology was worthwhile. The chapter illustrates how collaboration between the U.S. and UK panels led to a convergence in perspectives and a common understanding of the problem.

In chapter 6 on Persistent Surveillance, the U.S. and UK groups followed two distinct but complementary threads. The United States sought to understand how to better exploit outputs from persistent surveillance technologies, while the United Kingdom instead focused on identifying technologies that require specific defence investment or that could benefit from UK/U.S. collaboration. Chapter 6 argues for increased UK and U.S. collaboration to establish common standards and interoperability, link high-level Network Enabled Capability and Network Centric Warfare activity, and explore the benefits, opportunities, and challenges of aligning capabilities to drive improvements in persistent surveillance.

By definition, the use of small teams for each of these topical assignments means that the technology addressal of these topics was "thin" when compared to normal DSB and DSAC output. We adjudged that it was more important to explore the processes of working jointly together on topics, and we picked diverse topics to challenge those processes. Observations about these processes are contained in the "Conclusion" chapter of this report.

CHAPTER 2. POWER SOURCES/MANAGEMENT FOR SMALL, DISTRIBUTED NETWORKED SENSORS

INTRODUCTION

Distributed sensor systems on the ground, under water, and in the air have been used by the military for many years. The key applications are in intelligence gathering and to better understand and measure the battlefield – this includes the detection and monitoring of personnel, military vehicles, weapons, and communications. Emerging technologies will allow small, low-cost networked sensors to autonomously coordinate amongst themselves to achieve a larger sensing task. While initial applications for these new sensor systems are in the commercial market (for example, power and equipment monitoring; climate control; structural, seismic, and environmental monitoring; and inventory management and tracking), these technologies will revolutionize information gathering and processing by the military and intelligence communities across a range of terrains including urban, farm/rural, jungle, mountain, and desert. As a measure of their impact, the market for small, autonomous distributed sensor networks (also called "motes" or "smart dust") is estimated to be $50 billion in ten years, dominated by civilian uses.

Distributed in irregular patterns across remote and often hostile environments, sensor networks create daunting engineering challenges for sensor system designers, builders, and military users. Each node, which consists of a sensor, processing electronics, communications, and a power source in an environmental package, must be small, lightweight, inexpensive, low-power, and, because of the projected size of the network, low-cost. The system architecture provides the overarching control strategy. In order for these systems to be most effective for the military, sensor networks must self-organize, be robust and provide high information assurance despite individual node failures, intermittent connectivity, and tampering. In addition, support for lengthy

mission lifetimes constrains power consumption to miserly rates when not in an energy conserving dormancy.

Figure 1 shows a schematic drawing of the components in a typical node (left) and photographs of currently available (centre) and emerging (right) sensor nodes. What is obvious from these pictures is that despite the tremendous advances in sensors, control/processing electronics, and communications systems that have occurred over the last few decades (and continue to occur), the power source completely dominates the size and weight of the individual node. In addition, packaged electronics are far more robust than the power source to the environmental extremes experienced by the military (high/low temperature, water/humidity, shock and vibration, dust/dirt, etc.). Thus, there is a growing consensus that advances in power source technology and low-power circuit design cannot, by themselves, meet the energy needs of future sensor systems and that entirely new architectures and protocols must be developed. "Node-centric" power issues (which are constrained by the laws of physics, chemistry, and thermodynamics) include low-power sensing, low-power electronics for data processing and storage, communications (both transmit and receive), and the power source itself. Addressing these issues will increase the lifetime of an individual sensor node and therefore enhance network longevity. Once these hardware constraints are better understood, one can then explore higher level systems and software issues (such as the development of advanced architectures, protocols, and algorithms); as well as, the key technical system synergies and trade-offs between hardware and software to ensure that the network maintains its high level of functionality while still conserving energy.

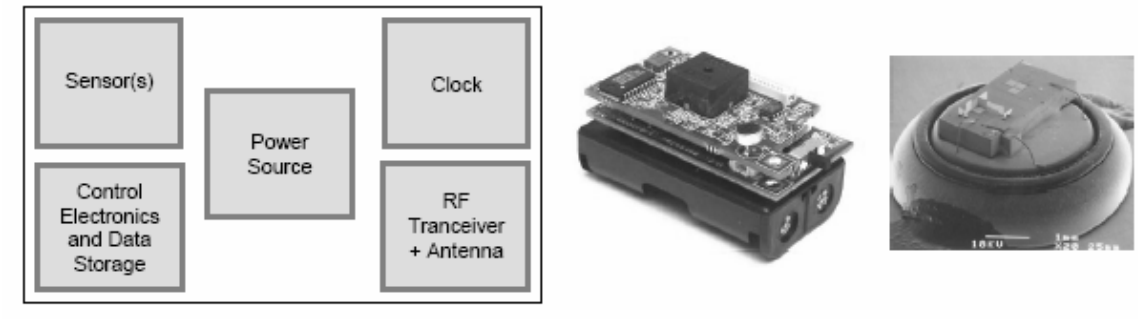


Figure 1. *The power source bottleneck.* [1]

LOW-POWER SENSING

The commercial market and numerous defence programmes are developing a vast array of compact, light-weight, low-power sensor systems for incorporation into distributed networks that are relevant to the needs of the defence and intelligence communities. These include sensors for explosives; chemical, biological and nuclear weapons; thermal, motion, pressure and magnetic/metal sensors for tracking military vehicles and weapons; audio and imaging technologies (including infrared, visible and ultraviolet) for detecting and monitoring both personnel and vehicles; and radio frequency (RF) and audio sensors for monitoring communications. Today's sensors represent only a small fraction of the power requirements of a distributed network system. Depending on the type of sensor employed and its use profile, power consumption can be either extremely low-level, continuous (nanoWatt – milliWatt) or require bursts of power (>Watt). Representative examples of the energy use and lifetime for a number of sensors applicable to the military and intelligence communities are shown in Table 1.

1. A typical sensor node (left) consists of a sensor (or group of sensors), control electronics/signal processor/data storage, clock and an RF transceiver with an antenna. Such systems are available for purchase today (centre) and operate on two "AA" batteries. Nodes emerging from research laboratories are far smaller and operate at far lower power, in this case a "coin cell" (right). Despite substantial advances in technology, the node size and weight is still dominated by the power source.

Table 1. Energy use and lifetime of several currently available sensors. Note that by switching from a AA alkaline battery to a comparably sized, commercially available Li-CFx cell will increase the sample size by a factor of approximately five.

Sensor	Energy/sample (3V)	Samples per AA alkaline battery
Microphone	1.5 nanoJoule (nJ)	20 trillion
Temperature	30 nJ	1 trillion
Accelerometer	1.5 microJoule (μJ)	20 billion
Passive infrared (IR)	75 μJ	5 billion
Magnetometer	300 μJ	100 million
Pressure	300 μJ	100 million
CMOS Imager	1 milliJoule (mJ)	30 million
Gas (electrochemical)		Unlimited (generates power)

Figure 2 displays representative examples of packaged sensor systems available on the commercial market today. Maturation of these technologies in terms of size, weight, power draw, reliability, and especially cost is already occurring rapidly, driven primarily by the needs of the commercial market place (e.g., high resolution cameras in cell phones, accelerometers and gyroscopes in wireless joysticks and mice, microspectrophotometers and microelectrochemical cells for glucose monitoring by diabetics, etc.). Many of these sensors are extremely robust and driven by the needs of the implantable medical device and automotive industries (currently the single largest consumer of Micro Electro-Mechanical Systems (MEMS) accelerometers for airbags and an emerging player in the passive infrared market). These latter sensors must meet environmental standards which in many cases are more stringent than those of the military. While some of the technologies noted above require additional maturation and environmental hardening before fielding, this is already occurring in the commercial/industrial, biomedical and defence sectors. Therefore, we believe that no supplementary (incremental) investments in low-power sensors are needed at this time. In addition, the academic, small business, and research laboratory communities are developing a vast array of even lower power sensors with higher sensitivities based on polymer electronics, nanotechnology (e.g.,

carbon and silicon nanotubes) and biomimetics (e.g., electronic noses).

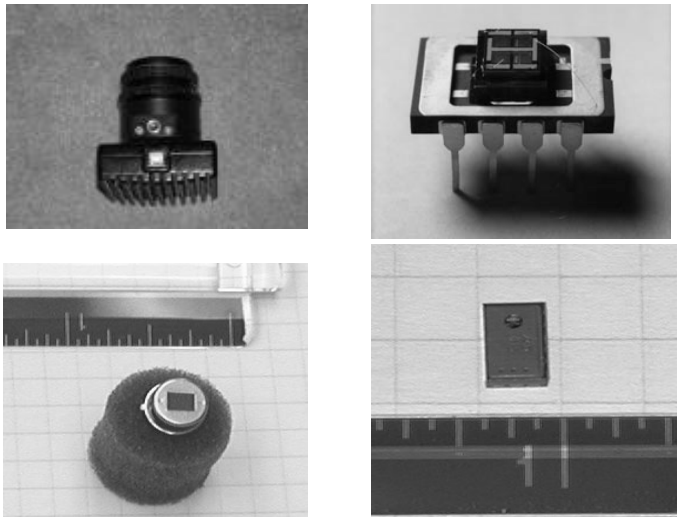


Figure 2. *Commercially available packaged sensors including complete visible camera, gas sensor, infrared imager and accelerometer.*

The military and intelligence communities face two critical issues with respect to the development and exploitation of emerging low-power sensors in distributed sensor networks, however. The first issue involves the speed with which these sensors will continue to improve in functionality and decrease in size, weight, cost, and power and the limited ability of the military procurement system and traditional defence contractors to rapidly incorporate them into state-of-the art systems. The second issue is perhaps more ominous: the availability of state-of-the art sensors to our potential adversaries given that most of these are produced in large quantities for the commercial market in overseas factories.

LOW-POWER ELECTRONICS

Today's sensor nodes all require some form of electronics for data collection, manipulation, storage, etc. Power consumption of the key electronic components for sensor nodes (e.g., clock,

microprocessors, memory, etc.) will continue to fall at a pace driven by Moore's Law. As feature sizes shrink, operating voltages drop and new device architectures are developed. Today feature sizes of commercial devices are routinely below 100 nanometre (demonstrations at ~20 nanometre) and supply voltages less than 200 millivolt are possible. Like sensors, much of the need for low-power, reliable electronics is driven by the commercial, portable communications, and entertainment markets (e.g., cell phones, MP3 players, laptop computers, etc.). Biomedical applications are having an increasing impact in this area (e.g., cochlear and retinal implants). Consumers are demanding increased performance, functionality, and run time and suppliers are delivering. Depending on the use profile, power needs for low-level, continuous (quiescent) operation can be in the microWatt (μW) to milliWatt (mW) range requiring bursts of mW's to W's during processing intensive periods. Note that many microprocessors already shut down part of their system when not in use to conserve power.

Nothing comes for free, however, and there are trade-offs, for example, between smaller feature size (smaller chip size, lower die cost) and increased leakage current (more power draw) or faster clock speed (more operations per second) and increased power usage. If one is willing to drop the clock speed substantially, one can build a 16 MHz general purpose processor that uses less than 1 mW during operation and only a few μW in standby. The key to the efficiency of these low-power systems is to only turn on the portion of the circuit that is being used at any given time (this is driven by system architecture, see below). For low duty cycle operations, it is the standby power that determines the lifetime of the system and thus low leakage current is absolutely critical. Decreased system flexibility can also yield a substantial improvement in power efficiency. For example, there is a two-order of magnitude trade-off between the power efficiency of a dedicated verses a general purpose microcontroller. More energy efficient data storage hardware (e.g., static random access memory [SRAM] verses dynamic random access memory [DRAM]) can also minimize power requirements. All of these issues are being addressed in commercial and emerging commercial systems. In addition, much

of the environmental tolerance (e.g., extreme temperatures, humidity, dust, mud, etc. operation) and extremely high reliability, of critical importance to defence, is now driven by automotive, biomedical, and computer/communications intensive "road warrior" needs.

The military and intelligence communities face the same two critical issues noted above: the speed with which power efficient electronics will continue to improve (and the limited ability of the military procurement system and traditional defence contractors to rapidly incorporate them into state-of-the art systems), and the availability of these systems to our potential adversaries given that most of these are produced overseas.

COMMUNICATIONS

Transceivers are the single largest consumer of power in a distributed network system. The total power consumption of both the transmitter and the receiver is critically dependent on the system specifics including the stand-by, wake-up and transmit/receive power; operating frequency; clock synchronization (the more accurate the system clock, the higher the use fidelity between the transmitter and the receiver – see discussion of system architecture below); the use profile (duty cycle, typically ≤1% and approximately linear in power usage); the system architecture/control algorithms; the node spacing, placement and location (e.g., on soil/ground cover, in buildings or trees, etc.), and the extent of on-board signal processing verses the quantity of data transmitted for remote processing. A qualitative view of these latter two trade-offs is shown in Figure 3 and is a key driver in the design of distributed sensor systems.

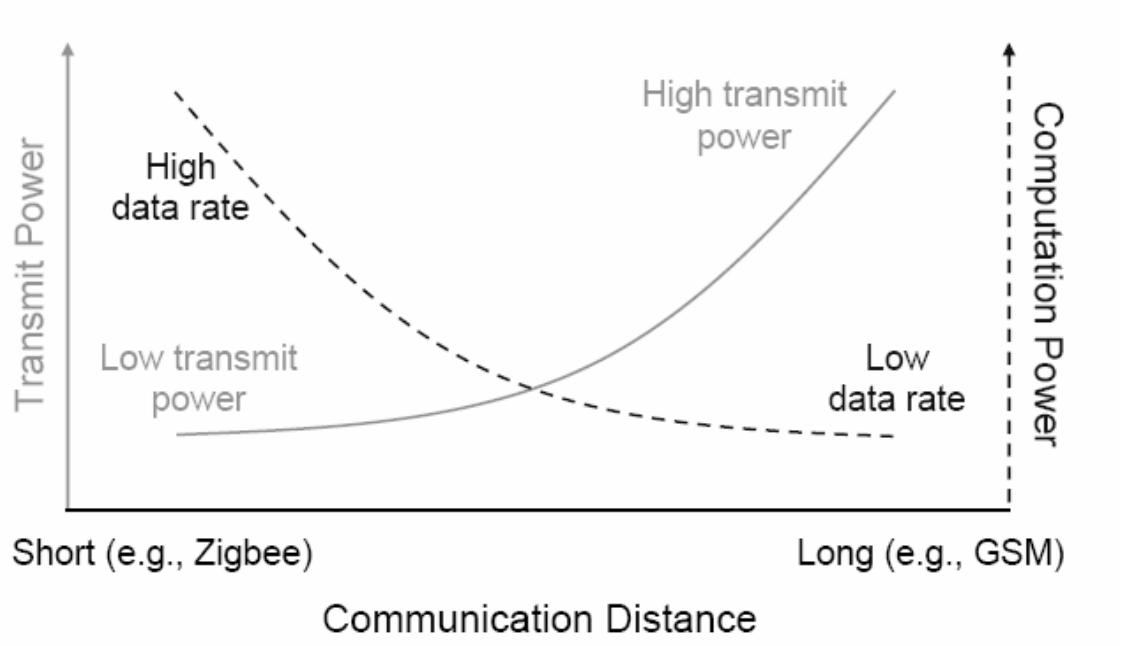


Figure 3. *Schematic view of the trade-offs between the power requirements for transmit power (solid) and on-node processing (dashed) and as a function of the distance between nodes. Zigbee and Global System for Mobile communications (GSM) are standard communications protocols used by sensor nodes and cell phones, respectively.*

The commercial market is making great strides toward minimizing power consumption in transceivers. This is driven largely by the use of BlueTooth, 802.11, 802.15.4 (Zigbee), etc. protocols in portable consumer electronics and commercial sensor systems. Today's commercially available low-power transmitters require only ~20 milliAmp (mA) of current to transmit 250 kbps (thousands of bits per second) and next generation commercial systems will require only 4 mA. Today's laboratory-based systems are pushing the envelope even further and use less than 0.4 mA to transmit 50 kbps. Because of privacy issues, encryption and crosstalk/interference are also being addressed by commercial industry. Like sensors and low-power electronics, the automotive industry is driving these devices to have broad environmental tolerance (required for remote entry, wireless headsets, etc.). Finally, while many of these systems are designed to operate in electronically noisy environments (e.g., industrial settings), they are

not built to detect or tolerate active jamming – a key problem for the military given the need to operate at extremely low transmit power to prolong power source life.

While improvements in sensors, electronics, and radios are being driven rapidly by the needs of the commercial market, antenna enhancement has lagged far behind. This is especially true for antennas of specific interest to the defence and intelligence communities. Most commercial transceiver systems, whether for consumer, commercial or industrial use, operate at least one metre from the ground. While antenna size is important for compactness, it does not have to be covert. This is not the case for distributed sensor systems for the military – where most distributed sensor systems will be on the ground, inside buildings or even underwater and should not be readily detected. Antenna performance and therefore transmission distance, drops dramatically as one approaches the ground due to anomalous reflections, turbulence in the atmosphere, diffraction effects from objects, and the proximity of a lossy (dielectric) medium. In free space, the power required to transmit a signal with an omni-directional antenna, increases as distance squared (r^2) while near the ground or underwater it can be as high as r^4.

The curves in Figure 4 show that the performance of a vertical antenna drops dramatically as the height above a gravel surface is reduced from 120 inches to 7 inches and finally to 4 inches. It is expected that the performance will degrade even further as the distance decreases to near 0 inches and the antenna shrinks in size or changes in direction from vertical to near-horizontal (crucial for covert operation). In addition, the condition of the ground is critically important to how well electromagnetic waves will propagate (e.g., conductivity – asphalt verses soil, the presence of obstacles such as rocks and vegetation – surface roughness, etc.). Systematic tests of these variables have not been performed to date and are critical to the design of sensor networks. Finally, because of size constraints, one cannot use high-gain antennas on individual sensor nodes. This further limits the performance of both the node and the system.

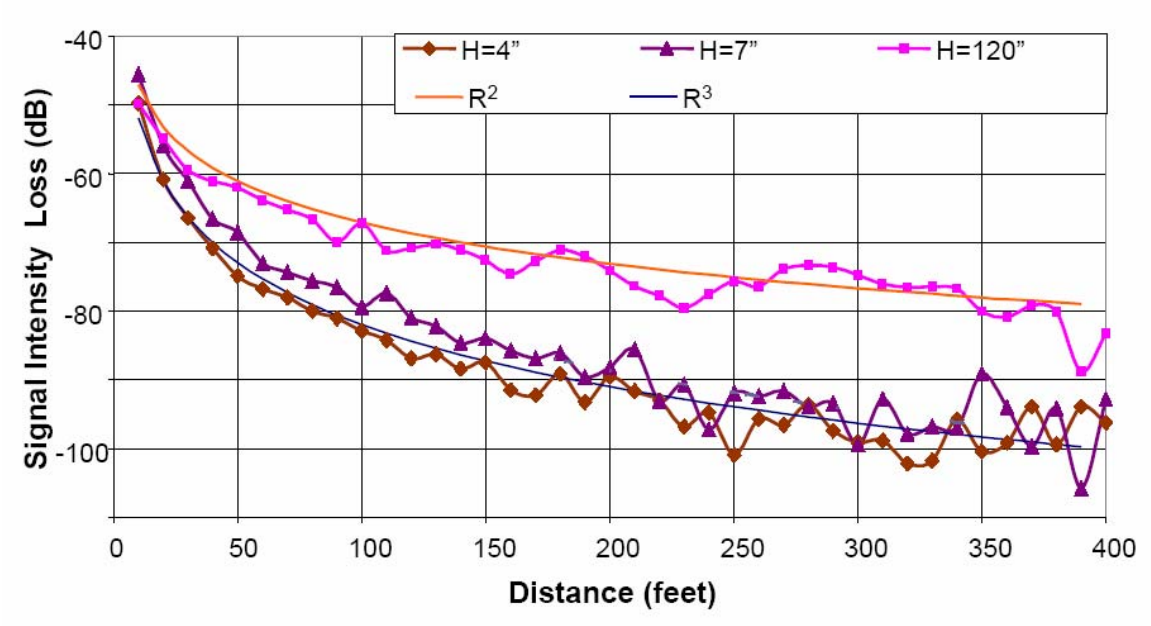


Figure 4. *Performance (signal intensity loss verses distance between transmitter and receiver) of a vertical antenna degrades rapidly as it is brought closer to a gravel surface (120 inches: pink squares, 7 inches: purple triangles, 4 inches: brown diamonds). The smooth curves show signal intensity falling off as r^2 (orange) and r^3 (blue).*

Given the importance of antennas to the performance of distributed sensor systems, it is critical for the military to develop a fundamental understanding of electromagnetic propagation on and near the ground (<10 centimetre [cm]) as a function of operating frequency, bandwidth, ground cover environment (e.g., soil, sand, asphalt, cement, grass, bushes, trees, etc.), weather, etc. Once a sound experimental and theoretical basis is established, one then needs to create design tools and build and test high performance, compact (stealthy) antennas and antenna systems (e.g., with, for example, micro load coils to raise the "effective" height of the antenna) specific for near-ground operation in a range of environments critical to many defence and intelligence community applications.

POWER SOURCES

Numerous small power sources (e.g., batteries) and energy harvesting systems (e.g., solar cells) exist and have been used in commercial distributed sensor networks for a variety of applications (e.g., automated irrigation and fire alert systems). Table 2 lists several current and emerging power sources which may be applicable to distributed sensor networks. Commercial systems do not require stealth, can be readily serviced, and under certain circumstances tap into the existing electrical grid for power. Commercial technology can and should be used for military applications having similar operational characteristics. However, the military has unique missions that will require more demanding attributes such as stealth, long-term operation without human intervention, and harsh operating environments. Stealth and long-term operation clearly require high energy density (energy/unit volume) power sources. In addition, the dimensions of the power source may be a critical design feature (e.g., a very thin power source may be desirable from both systems architecture and manufacturing standpoints). Whatever the power source, they must be able to operate over a broad range of environmental conditions (high and low temperature, high and low humidity, dust/dirt/mud, etc.) and should not emit detectable signatures.

Table 2. Comparison of various potential power sources for wireless sensor networks. Values shown are actual demonstrated numbers except in two cases which have been italicized. For systems where the source of power may be intermittent, secondary storage (e.g., a rechargeable battery or supercapacitor) may be required. Supercapacitors may also be used in systems where short pulses of high power are required.

Power Source	P/cm^3 ($\mu W/cm^3$)	E/cm^3 (J/cm^3)	P/cm^3/yr ($\mu W/cm^3/Yr$)	Secondary Storage Needed	Voltage Regulation	Commercially Available
Primary Battery	-	2880	90	No	No	Yes
Secondary Battery	-	1080	34	-	No	Yes
Micro-Fuel Cell	-	3500	110	Maybe	Maybe	No
Supercapacitor	-	50-100	1.6-3.2	Maybe	Yes	Yes
Heat engine	-	3346	106	Yes	Yes	No
Radioactive (^{63}Ni)	0.52	1640	0.52	Yes	Maybe	No
Solar (outside)	15000 *	-	-	Usually	Maybe	Yes
Solar (inside)	10 *	-	-	Usually	Maybe	Yes
Temperature	40 *†	-	-	Usually	Maybe	Soon
Human Power	330	-	-	Yes	Yes	No
Air flow	*380* ††	-	-	Yes	Yes	No
Pressure Variation	*17* †††	-	-	Yes	Yes	No
Vibrations	200	-	-	Yes	Yes	No

* Denotes sources whose fundamental metric is power per **square** centimetre rather than power per **cubic** centimetre.

† Demonstrated from a 5 ºC temperature differential.

†† Assumes air velocity of 5 m/s and 5 % conversion efficiency.

††† Based on a 1 cm^3 closed volume of helium undergoing a 10 ºC temperature change once per day.

While there is generally a trade-off between energy density and power density, the key driver for most systems is energy. Thus, batteries are the obvious choice for most applications as they are hermetically sealed, can operate over wide temperature ranges (depending on the electrolyte and the cell chemistry), are robust, will have little or no signature at the power rating envisioned for distributed sensor networks and are readily available at low cost. In contrast to batteries, supercapacitors (also known as electrochemical double layer capacitors) offer very high power density (>10 Watts/gram [W/g]) with limited energy storage. These may be useful in hybrid systems where high power communications are routinely required. Only after a thorough system analysis reveals that batteries will not meet the energy requirement of the mission should other alternatives be pursued. Energy harvesting or the conversion of high-energy content fuels to electricity are alternatives to batteries, but add system complexity, limit the conditions over which the system may operate, may decrease reliability, and certainly increase cost. In some instances one could envision a distribution of sensors, which would not require any power at all. For example, randomly distributed sensors that have the ability to change state when exposed to a triggering signal could be "read" periodically by passing vehicles (see below). This would not be as sophisticated as a network of nodes that communicate with each other but could nonetheless be a valuable asset in battlefield management (e.g., land mine detection). Such systems could be built using technologies and components from the rapidly growing radio frequency identification market.

Energy harvesting approaches offer the advantage of very long duration missions without the need for changing batteries or refueling. Even for small power loads the integrated energy over time can well exceed the energy content of any known or anticipated battery chemistry. There are many sources of ambient energy and means to convert this energy to electrical power. Some of these are listed in Table 2 above. For distributed sensor networks the amount of energy harvested is likely to be small due to the source energy content and the small footprint of the "harvester." It may be necessary to harvest and store energy over some period of time in order to enable power draws, e.g., periodic radio

transmissions, which exceed the capabilities of the energy harvester. This will necessitate a hybrid approach: the use of a rechargeable battery and/or supercapacitor to store the harvested energy for later use. The required power loads and duty cycles will determine the design of the hybrid components. Reliability will be determined by the integrity of the energy harvester (e.g., a solar cell could become obscured by debris and rendered useless) and the energy storage device (e.g., long term battery performance, which can be affected by cycling, depth of discharge, temperature extremes, or self-discharge).

Energy conversion of fuels to electricity using purely thermal, mechanical or electrochemical means is a very attractive option due to the high energy content of many fuels (e.g., hydrogen: 33 Watt hours/gram [Wh/g], diesel/jet fuel: ~13 Wh/g, methanol: 5.6 Wh/g). If air is used as the oxidant, it does not have to be carried and thus there is no volume or weight penalty associated with it, however if the system gets flooded with water or is contaminated (dust, mud, chemicals, etc.) the power source could be compromised or cease to function, perhaps permanently. The same is true for air (oxygen) "breathing" batteries (e.g., Zn-air, Al-air, Li-air).

For very short missions the fuel volume relative to that of the energy conversion device is insignificant, so fuel conversion efficiency is not important. It is unlikely, however, that for short mission scenarios envisioned, this approach would be better than batteries or supercapacitors, which are much simpler to implement and less susceptible to environmental contamination. Energy conversion efficiency is critical for long missions, as the amount of fuel required will determine the power source system weight and volume and the energy conversion device will be only a small fraction of the total. Thus, small fuel cells may play a role here. While the fuel energy content is a compelling incentive for developing these small systems, many subsystem components, e.g., insulation, shielding, air and fuel management systems, thermal management, will contribute to reducing the overall system energy density for a given mission. Variable load profiles will add

complexity and likely reduce overall system efficiency, energy, density, and reliability.

Radioactive sources provide a steady source of heat and/or nuclear particles (typically alphas and betas) which may be converted directly to electricity and exploited for compact, long-term energy conversion devices. Systems based on direct thermal conversion with lifetimes greater than 20 years have been developed for the space program; direct electrical conversion of high energy electrons (e.g., betavoltaics) in a compact package is emerging as a possible new high energy density power source. Proliferation is not an issue here as the amount of material is small (for example, most of today's home smoke alarms contain a small amount of radioactive material). While the energy density of a nuclear source is orders of magnitude higher than that of a chemical fuel, the conversion efficiency is still quite low leading to high energy density, but low power density systems. Proper shielding will also be required for safe handling and stealth. In addition, the source must not degrade any of the components of the power source or any other system components (e.g., electronics). Note that nuclear sources produce a constant output and cannot be throttled or shut off. Thus, they must be used as part of a hybrid system incorporating a rechargeable battery or supercapacitor if large excursions in power demand (such as transmitting or receiving data) are required by the application.

For all of these power sources proper metrics need to be developed for the power source (power, energy, cycle life, efficiency, etc.) in the context of the load profile for the anticipated mission under the expected environmental conditions. This will ensure good trade studies that will lead to the optimum solution for a given application and mission. Small nuclear sources, very high energy density batteries, and energy conversion devices should be evaluated for potential research and development funding to improve performance in distributed sensor networks. In addition, the state-of-the-art for existing energy harvesting technologies or concepts should be assessed in the context of distributed sensor networks in order to determine what, if any, energy shortfalls exist. Finally, the military should exploit the use of very low-cost,

unpowered sensor systems (built around architectures developed for radio frequency identification tags, see Figure 5) for military and intelligence applications. It may be much simpler and more cost effective, and reliable to deploy a suite of unsophisticated sensors, each reporting on different agents or signals, than one highly sophisticated sensor that attempts to do everything.

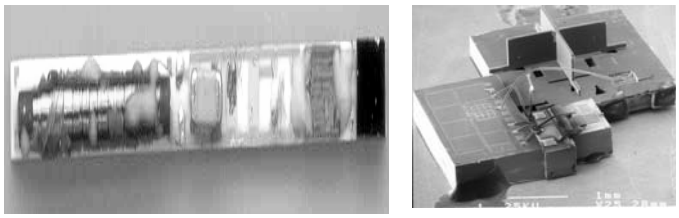


Figure 5. *Potential for compact, unpowered (or extremely low power) distributed sensors based on passive radio frequency identification (left) and optical corner-cube retroreflectors (right). The node receives its power to respond from an interrogating pulse (e.g., RF or light) from, for example, an overhead asset or passing vehicle.*

SYSTEM ARCHITECTURE

Optimized network architecture and the operating protocols and algorithms used to drive the system are critical to minimizing power usage while maximizing performance and robustness. This is a rapidly evolving field driven by significant investments from both the military and civilian parts of the government (work performed by both small companies and university performers), venture capitalists (through small companies), and large commercial/military suppliers. Typical sensor architectures include star, cluster tree, mesh, and hybrid. Examples are shown in Figure 6. A star architecture (Figure 6[a]) uses a central node to mediate all communication. It is a very simple system (exemplified by 802.11) and derives its power savings via time, division, and demand multiple access (TDMA). Unfortunately, it has a single point of failure and no redundancy. A cluster tree network (Figure 6[b]) uses a branching tree structure where each cluster head controls sub-nodes. This extends the range of the system and provides

power savings through scheduling sub-networks. Unfortunately, the individual routes are longer and there is still no redundancy. A mesh network (Figure 6[c]) uses every node as a relay or routing point which provides short routes, redundancy, and easy deployment. The increased listening times makes power conservation difficult, however. A hybrid network (Figure 6[d]) uses elements from star and mesh architectures to provide short routes, ultra-low power leaf nodes and easy deployment at the expense of increased complexity. While certain types of networks may be ideal for specific situations, the keys for the military are to ensure the reliability of the information and to make the system adaptable to the addition or loss of new nodes and robust to changing conditions.

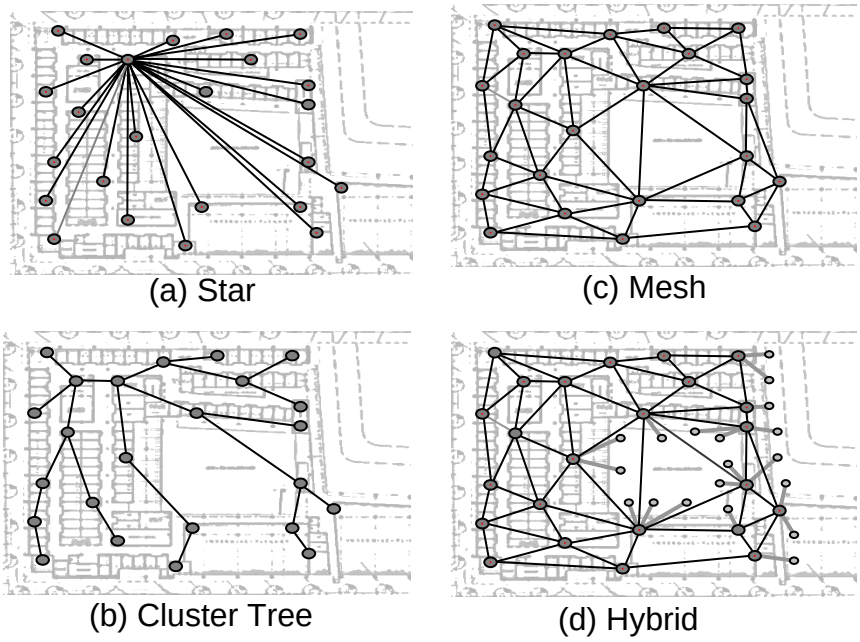


Figure 6. *Representative distributed sensor network architectures each with its own advantages and disadvantages (see text).*

The choice of system architecture is dependent on many variables. Most optimized network architectures use a hybrid system involving a mixture of very low-power sensor nodes (of order mW with 1 kilobyte compute power) combined with star,

cluster and/or mesh nodes which function as higher level network controllers (mW – W, 1 megabyte – 1 gigabyte compute power). Due to the rapid drop in microprocessor/memory power requirements relative to that of transceivers, more signal processing is occurring on the local and star nodes verses transmitting data to a central location for processing. System trades involving, for example, shorter distances, but multi-hop verses longer distance single hop, optimized duty cycle, two-way handshaking, etc. are already being performed, but more work clearly must be done given the specific needs of the military. In contrast to the commercial sensor network market, it is critical for the military to operate with randomly placed sensors, some of which may be compromised. Such control structures have been developed and tested.

While the system architecture determines the distance over which an individual signal must travel and the inherent redundancy of the network, the operating protocols determine how "alert" the network is. The average system power requirements can be determined by summing up the fraction of time the system is asleep (typically >95%) times the sleep power plus the fraction of time the system spends in wake-up mode times its power requirements (very short, but may be power intense) plus the fraction of time and the amount of power required to do "real" work – sense, compute, transmit/receive, etc. ($P_{ave} = f_{sleep}*P_{sleep} + f_{wakeup}*P_{wakeup} + f_{work}*P_{work}$). Numerous methods are used to control the relative amount of time in each state depending on the requirements of the system. These include synchronous wake-up where the entire system turns on simultaneously for a short amount of time to determine if it has detected anything and then shuts down again. This requires a high quality system clock and is not particularly energy efficient since the system will turn on when there is nothing to report. Alternatively, "sentries" in the system can provide alerts as soon as something is detected and turn on the entire network. While this may minimize transmit power requirements, this requires receivers to listen more often and thus increases their power usage.

Once the overarching system network architecture is decided on, node addressing protocols to control transceiver function must be established. Most use standard communications protocols (many of which have been developed for the cell phone industry, e.g., GSM, Code Division Multiple Access and TDMA). There are significant development efforts in both universities and companies to write efficient, yet flexible, operating systems to control both the system and each individual node. These include the development of both open source (e.g., TinyOS) and proprietary operating systems. High-speed algorithms optimized for specific applications are then written. Since the military operates in harsh environments with the need for high reliability, it is critical that the network contain some redundancy, fault tolerance and a low probability of detection and interception. While the requirements are not as strict for most civilian applications, encryption and error correction are already part of many of these systems. The defence and intelligence communities can certainly leverage the vast array of work going on in this field and can steer research into appropriate directions.

SYSTEM SYNERGIES AND TRADE-OFFS

While developers of distributed, networked sensor systems understand that one must optimize the entire network, not a single node or single function, many of the key hardware components are being created in isolation (e.g., sensors for numerous stand-alone and networked applications, electronics for a broad range of consumer needs, standardized communications driven primarily by networked consumer devices, general purpose power sources, etc.). Today, poor system design results in a sensor node dominated by the size and weight of the power source (Figure 1) rendering advances in component miniaturization essentially irrelevant. Thus, it is imperative that a total system design approach include power generation, power conditioning, energy storage and management, etc. and be carried out from the earliest stages of development. Most distributed sensor network developers are focusing on the use of low-cost, general purpose components for a broad market and thus coordination and optimization mainly occurs through software (system architecture). More specialized integrated electronic and

communications systems (systems on a chip or systems in a package) are emerging from university and government research laboratories and can operate at much less power than conventional designs, as well as, provide for smaller size and higher reliability. An example of such a system is shown in Figure 7. One can go much further, however and use multifunctional approaches to packaging where the power source components are fully integrated with the sensor, package, antenna, etc. (for example, using printed batteries, fuel cells or solar cells). This can further reduce the weight, volume, and footprint of a node and potentially increase its reliability.

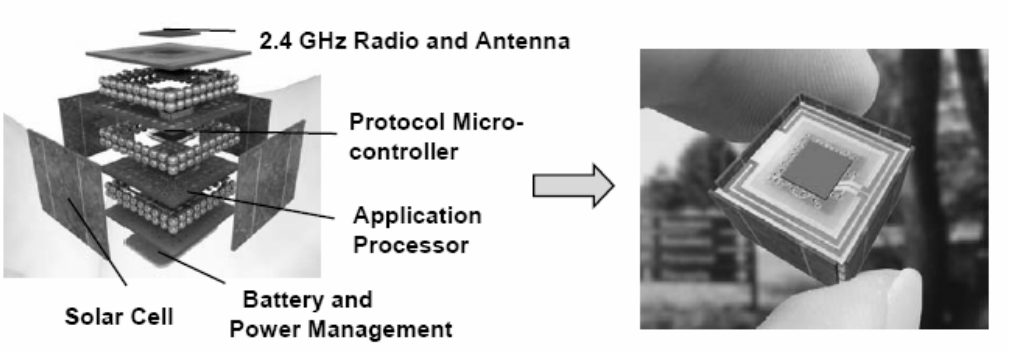


Figure 7. *Sensor node with integrated power source (solar cell), electronics, radio and antenna emerging from today's university research laboratories.*

This systems approach is critical to the successful development and deployment of low-cost, power efficient distributed networks. Nevertheless, hardware is only one small piece of the solution (a two- to five-fold improvement) – software holds far more promise to affect the longevity of a distributed network system (possibly one-two orders of magnitude). Thus, the general conclusion that we reach is that research on application specific integrated systems (power generation, power conditioning, energy storage and management in conjunction with sensors, control electronics, signal processing, transceivers, etc.) with highly efficient software to control system load demands that minimize energy use while

optimizing performance, reliability, lifetime, etc. for the military should be emphasized. This is critical to developing robust systems that cannot be detected or compromised by our adversaries.

RECOMMENDATIONS

Despite huge investments in distributed sensor systems from the commercial market, the military and intelligence communities have an important role to play to ensure that the resulting systems meet their critical needs. Key recommendations are to:

1. Develop critical system and power source metrics so that proper trade studies can be made in the context of military and intelligence community requirements. Optimizing the performance of the system (rather than that of the individual components) is critical to the successful development and deployment of power efficient distributed networks.
2. Enhance and leverage the rapid advances emerging from the commercial and industrial sector including application specific integrated systems (power generation (e.g., very high energy density batteries and small nuclear sources), power conditioning, energy storage and management in conjunction with sensors, control electronics, signal processing, transceivers, etc.), and highly efficient software to control system load demands that minimize energy use while optimizing performance, reliability, and lifetime. Most of these systems will continue to improve rapidly in functionality and decrease in size, weight, cost, power, etc. without additional investment by the defence and intelligence communities. The resulting products, however, must be tailored to meet specific defence needs (extremely harsh environments, robustness/redundancy, and low probability of detection/low probability of intercept, jamming).
3. Focus on areas where there is minimum commercial investment, in particular, antennas that must be covert and operate very close to the ground in a variety of terrains and

in harsh operating conditions. A more thorough understanding of electromagnetic propagation in these specific environments and the development of new modelling and design tools and advanced materials, and concepts are critical to the successful deployment of distributed sensor networks.

4. Given the rapid, continuous advances in sensors and electronics, for long-term missions, the power source will dominate the size and weight of the system and limit its lifetime. The development of remotely read, unpowered (or minimally powered) nodes which may be less sophisticated (but have lower-cost, smaller size, higher reliability, etc.) could solve the power problem and should be explored. Such systems may also limit detection and jamming by our adversaries.

SUMMARY

Based on these recommendations, there are a number of critical defence-specific technologies that must be developed to ensure our ability to field effective distributed sensor networks.

- Reliable, very high energy density power sources that can operate in the extremely harsh environmental conditions critical to the defense and intelligence communities. The optimum solution may involve hybrid systems (e.g., energy scavenging or continuous low-power systems for quiescent operation coupled with high pulse power delivery for communications).
- Highly efficient software, architectures, and system design tools to optimize and control system loads for minimum energy use and optimum performance, reliability, and lifetime.
- High efficiency, covert antennas that are optimized for near-ground use. This will require an improved understanding of electromagnetic propagation within

centimetres of the ground, as well as new materials and antenna design concepts.

- Remotely read, unpowered (or minimally powered) sensor systems (components, architectures, software, etc.). These could be built on technologies developed for radio frequency identification tags, but will require the development of systems architectures to transfer multiple bits of information from ultra-low power sensors with minimal signal processing ability and at large stand-off distances.
- Technologies to allow low-power sensor nodes and distributed sensor systems to operate and communicate reliably in electrically noisy and jamming environments.

Beyond technology, the military and intelligence communities face two critical issues with respect to the development and exploitation of emerging low-power distributed sensor networks. The first issue involves the speed with which these systems will continue to improve in functionality and decrease in size, weight, cost and power, and the limited ability of the military procurement system and traditional defence contractors to rapidly incorporate them into state-of-the art systems. The second issue is perhaps more ominous: the availability of these systems to our potential adversaries given that most of these are produced in large quantities for the commercial market in overseas factorie

CHAPTER 3. HIGH PERFORMANCE COMPUTING

INTRODUCTION

Military applications have been a leading reason to develop high performance computing (HPC) – both hardware and software – throughout the evolution of modern computers. HPC continues to be critical to cryptanalysis and intelligence analysis, as well as to the design of military platforms. Computational fluid dynamics codes perform aerodynamic modelling and hydrodynamic modelling for air craft, ship, missiles, and nuclear weapon design. The national security communities of both the United Kingdom and the United States make extensive use of such HPC technology, and the United Kingdom benefits from the U.S.' investment and leadership in both hardware and software.

Since the onset of the earliest computers like ENIAC, there have been a stressing set of military problems that demanded ever more powerful computation. That remains the case today. Even the highest performance computers are inadequate to solve a variety of challenging military and intelligence problems. New problems routinely arise. This chapter will discuss two stressing problems that can not be solved at an acceptable level today, but which could be solved through the development of new high performance computers: knowledge discovery and image/video processing.

Today's high performance computers all utilize processing elements that execute in parallel. Architectures, and of course the speed of various components, determine the delivered performance of a computer. This report will discuss three architectures: very high performance computers, cluster computers, and grid computers.

From the beginning of the study, the U.S. and UK working parties took complementary approaches. The UK working party emphasized cluster and grid computing. The U.S. working party instead focused on very high performance computing. There are several reasons for the difference. The first is that the United States has dominated high

performance computing until the recent past when Asian nations began challenging U.S. leadership by building state-of-the-art high performance computers, such as Japan's Earth Simulator. There is great concern that lack of investment is eroding U.S. leadership in this field; as well as, negatively impacting our ability to meet defence mission requirements. A second reason for the U.S. focus is that technology that is developed in the context of high performance computers "flows down" to help advance mass market computers. Thus, if the United States does not aggressively pursue very high performance computer technology, then innovation in mass market computers will slow. The U.S. and UK militaries gain advantage from the assured access that they have to the upper end of mass market computation, which includes cluster and grid computing, as well as, very high performance machines and software. In addition, the military benefits from the economies of scale that derive from U.S. leadership in the medium and low end of the computer market.

As a result of these different perspectives, the U.S. and UK working parties undertook complementary studies. Although the two studies were conducted largely independently, there were frequent email exchanges and teleconferences enabling the two groups to develop the common understanding of the issues and to produce this integrated report.

DRIVING APPLICATIONS

Both the United States and United Kingdom have applications that cannot be solved with sufficient speed or with sufficient precision, e.g. cryptanalysis and design. To underscore the need for advancing high performance computation, we describe two problems that in the last several years have become critical to both nations. They cannot be adequately solved with today's computers and software. *Knowledge discovery and integration* can make a large contribution both to intelligence analyses, for example, in locating terrorists before they strike and in preparation of the battlefield. Effective, rapid *image and video processing* has become more important as sensors proliferate. For example, the majority of video collected in

London must be processed automatically if the data is to be processed at all.

Knowledge Discovery and Knowledge Integration

Knowledge discovery is the analytic or search problem of finding selected items of data with a huge reservoir of disparately formed data items and recognizing a relationship between them. In the vernacular, this is called "connecting the dots." Given a relationship of potential interest, the analyst seeks to extend it or corroborate it by relating further selected data items. Somewhere in this process, data is converted to knowledge, and in some cases actionable knowledge. In practice analytic knowledge is encoded in a range of products within an analyst's workflow: analytic reports, working notes, stored query results etc. It is often these higher level products that need to be queried to answer the "is there anything known about…?" question, rather than the raw data. Sometimes these activities are referred to as knowledge discovery; the term knowledge integration is also used.

For this application the search queries are both manually inserted and automatically generated by the knowledge discovery system. The analysis of query results is a continuous function of the system. Knowledge discovery occurs not just when data that is "out of the norm" emerges, but where new patterns between data, or properties of the data, become apparent. Discovery of some new correlations may very often generate new "queries" that either search for other occurrences of the same pattern, or build on the pattern to formulate larger patterns.

Knowledge discovery is representative of a large subset of applications that are particularly key to both national security and different in character from traditional uses of HPC in modelling the physical world, in at least the following ways:

- Mathematics used, graph theory, combinatorics, pattern recognition, logic,
- Algorithms used, discrete algorithms, fuzzy matching, clustering, inferencing, and

- Data structures used, linked lists, semantic nets, dynamic data typing.

As a consequence, there are extreme demands placed on computational resources, such as: large uniformly-addressable memory, random access, high bandwidth, low latency for small data updates, potential for very high amounts of random concurrency but balanced by increasing need to propagate the effects, and implications of changes through widely scattered data items.

While knowledge discovery has synergies with a few commercial applications, including bioinformatics (protein and gene databases, drug discovery, etc.), information retrieval and web search, a recent report from the U.S. National Academies states, "The scale of this knowledge discovery problem is significantly larger than the largest commercial data mining operations."[2]

Knowledge discovery also differs dramatically from modern "database" applications. Much of the knowledge discovery is internal and driven not by programmer inputs but by software that observes potential correlations that surface from convolving large amounts of data from different sources against each other. There is also a need for collaboration across organisational boundaries, bringing multi-disciplinary skills to a problem or to allow the federation of geographically dispersed data sources; this is where the term "knowledge integration" is most obviously relevant.

Neither massive grid nor cluster computing, both of which are cheaper per cycle than very HPC, will scale effectively for large instances of this application. In physical modelling using partial differential equations, spatial locality often makes it possible to decompose problems and make efficient use of machines whose bandwidth and latency scaling is inferior to their compute power scaling. This describes the common HPC systems, and is, for example, the reason that the TOP500 benchmark favours such

2. Getting Up to Speed – The Future of Supercomputing, Susan L. Graham, Marc Snir, and Cynthia A. Patterson, Editors, Committee on the Future of Supercomputing, National Research Council (NRC), November 2004.

systems. In contrast, the knowledge discovery application lacks physical locality and emphasize random access, fine-grained global operations, and pointer-walking graph analysis. "The knowledge discovery problem requires the understanding of extremely large graph networks with a dynamic collection of vertices and edges."[3] Walking from pointer to pointer in large graphs generates large amounts of traffic between computer nodes. When many pointers need to be simultaneously de-referenced, the computation becomes bandwidth limited, and the low-bandwidth of cluster or grid computers makes them ineffective. If only a few pointers need to be de-referenced, the computation becomes latency limited and the high latency of these machines leads to the same end result: inadequately used hardware and slowly produced results.

The knowledge discovery process is heavily dependent on metadata. Metadata is "data about data": information that provides a summary description of the content of the data item. A familiar example is an index of television programme – a compact textual description of channels and their planned programme – just enough information to characterize the content. An organized approach to metadata, for example a common data directory, is a primary enabler for knowledge integration. Metadata from different sources and describing different types of data (e.g. text, video, map co-ordinates) can be combined to allow concepts or generic types in the data to be contrasted and compared. The initial production of metadata from raw signals (e.g. voice recognition and image processing) is often extremely compute-intensive.

Image and Video Processing

A wealth of raw intelligence data is collected in the form of still image sequences, such as reconnaissance photos and videos from surveillance cameras (e.g., at airports, borders, and secure locations). Today, much of this data is discarded without exploitation due to limited human resources. An emerging class of image and video

3. [NRC p. 56, EAGLE (Ref: http://crewman.uta.edu/psi/download/Cook_Holder/Graph_Based_Anomaly_19May_2004.pdf)

analysis applications holds the potential to make use of this discarded data. Image analysis applications can provide direct intelligence by interpreting the raw images and video – identifying objects and people of interest and, in video, tracking their movements. Image analysis can also provide indirect intelligence by providing annotations that can be used as input to the knowledge discovery process – for example, producing descriptions of the people who were observed at particular locations to store in a database from which further inference can be drawn.

The field of image analysis has advanced greatly from the early failures of automatic target recognition. However, in many ways this field is still in its infancy. It can potentially benefit enormously from recent advances in statistical machine learning, for example by training a programme on a corpus of images and annotations until the programme learns to make the same annotations as an expert image analyst. Statistical methods are also being applied to discriminate "natural" movement – e.g., waves breaking and tree branches blowing – from "artificial" movement – e.g., a boat moving or person walking.

Image analysis is computationally demanding, but quite amenable to parallel solutions. Parallelism exists at the high level of separate video streams and images and at the lower level of separate pixels (millions), objects (tens to hundreds), and templates (hundreds to millions) that can be processed in parallel. It is critical that the computational needs of these emerging image and video analysis applications be quantified and that future high-performance computers, hardware and software, provide the capabilities needed to meet rising mission requirements.

A modern integrated defence knowledge system utilizes a *federation* of systems that each serves the needs of their local organisations, with an integration tier that allows the consolidation of this information when needed. Federated systems may use their own data formats, but provide an information service to the integration tier using standardized metadata. Security features are necessary to allow each authority to retain control over the release of its own data.

Because access to interim and final analytic products ("knowledge") may prove more important than access to raw data, the security protocol needs to support sharing of knowledge products as well as raw data.

Image and video analysis algorithms require: irregularly-structured, linked, and dynamically changing data, of varying types e.g. text, images and databases; increasing multi-channel, high-bandwidth, continuous, real-time input/output; 24x7 availability; and controlled cooperation between disparate organisations, including managing and optimizing their operation.

A key characteristic of these two critical and stressing applications is that they often demand flexible, on-demand programming, in stark contrast to the long software development cycles for the more long-established, stable, HPC applications. This, in turn, introduces demands on the supporting technologies, especially software development methods.

ARCHITECTURES

The following section discusses the three architectural categories of high to very high performance computing. The fundamental difference is the distance (in processor cycles) between the processing components and the memory from which data is accessed.

Grid

This architecture is well suited to situations in which different organisations each own part of the relevant data and want to protect their resources and data according to their own security standards. The knowledge discovery problem often involves multiple organisations that have these types of relationships.

Some in the defence community downplay grids because of their "open" architecture; however, a grid computer can provide controlled collaboration, because the architecture is well matched to situations in which multiple organisations are working on the same or related problems and sharing data. Grid computers, each with

their own memory, are geographically separated and often the transfer of data is achieved by exchanging messages across the networks that connect them.

Two technical problems need to be solved to make grid computing more applicable. First, the security mechanisms of today inadequately support the multi-level security required across multiple, indeed national, administration domains. Second, for some applications it is necessary to co-schedule tasks to run concurrently on multiple nodes in the grid – across multiple computers that may be in different administrative domains.

Most girds in use today are classified as less than high performance computer systems. However, a grid with high bandwidth communication, large memories, and many fast (though not necessarily high performance nodes) can be considered a high performance computer.

Clusters

A cluster computer consists of an interconnection of high-end commercial microprocessors, each acting as one node of the cluster. The node interconnection network can either be a commodity interconnect, like Ethernet, or a specialized low latency network.

Cluster computer architectures are typically designed to provide cost effective computation, not necessarily optimized for single application performance. This is commonly referred to as "capacity computing." However, there are cluster computers among the fastest 500 computers today. The challenges faced by cluster designers are exactly the same as those that have to be addressed by more integrated HPC machines, including limits on scaling and the cost of electricity that can be the dominant lifetime cost.

Limited enhancements to commercial processor elements can make a significant improvement to the effectiveness of these systems. For example, custom-made communications hardware that integrates optical input/output with protocols implemented in silicon has the potential to deliver much better bandwidth and latency than

standard commercial interfaces. The commercial market alone is unlikely to produce key components such as integrated low-latency interconnects due to cost. Because multiple processors can fetch from the same memory rapidly, it is possible to use multiple processors in concert on the same task. For example, in processing a video stream, multiple processors can perform portions of the analysis in parallel.

Novel Architectures

Novel hardware architectures involve either custom processor designs, or custom design that improves the architecture in the power/performance trade-off space.

The use of simpler processors makes it possible to obtain over an order of magnitude improvement in power/performance, as well as significant savings in silicon area, if it is feasible to achieve slightly more parallelism, especially on chip. Today, a current embedded computer (Central Processing Unit [CPU]) (e.g. an ARM 11) exhibits a factor of 10 improvement in power/performance over a high-end microprocessor (e.g. a Pentium 4).

The embedded processor saves die area by omitting certain functions required for general purpose computing (e.g., virtual memory management). Such functions may not be important for data-intensive or cryptographic applications, and useful savings (in complexity and power consumption) can be made if instruction sets can be simplified.

Simpler instruction sets, again reducing die area, can be designed if the custom processor need only execute one algorithm or one class of algorithms. For example, pattern-matching algorithms implemented using Field Programmable Gate Arrays (FPGAs) deliver around ten times the performance of a high-end microprocessor. Although the processor is limited to executing a single algorithm, it can support applications ranging from molecular matching to facial recognition. Thus, a key research problem is finding a programming model (like pattern matching) that can host other classes of useful applications.

Research into novel architectures is relatively inexpensive compared to "classical" HPC and has the potential to deliver machines that provide improved power and performance for very important military applications.

RECENT U.S. HPC ACTIVITIES AND STUDIES

In recent years, many studies of high performance computing have been conducted in the United States. Their conclusions are in substantial agreement, as is summarised below. This section highlights their key recommendations because the U.S. working party believes that these studies chart the correct course for the United States in high performance computation innovation.

Current HPC work in the United States was strongly influenced by two studies conducted in 2000 and 2001. A previous Defense Science Board (2000) Task Force on Supercomputing Needs made the following recommendations:

- For the short term, support the development of the Cray SV2,
- For the medium term, develop an integrated system based on commercial off-the-shelf (COTS) microprocessors and a new high-bandwidth memory system, and
- For the long term, invest in research on critical technologies.

Funding was provided by the DoD and the National Security Agency for the Cray SV2, and the National Science Foundation funded the acquisition of high performance computers and the construction of the Teragrid. However, no long-term research programme was initiated on the critical technologies for the future.

Four additional studies, listed below, were initiated to analyze the state of HPC and to make recommendations. Summaries of the studies are in Appendix D.

- Information Science and Technology (ISAT) (2001) – Technology Gaps and Bottlenecks,
- Integrated high-end computing (IHEC) (2002) – High Performance Computing and National Security,
- National Research Council (NRC) (2004) – Getting Up To Speed, The Future of Supercomputing, and
- HECRTF (2004) – Federal High-End Computing Revitalization Task Force.

We draw selectively and substantially from them in the remainder of this section.

A consequence of the industry focus on the desktop and commercial markets is missed technology opportunities and the lack of development of novel computer architectures capable of delivering the computational power needed for defence applications. Two figures from the study of the Defense Advanced Research Projects Agency (DARPA) Information Science and Technology Study Group eloquently quantifies this situation.

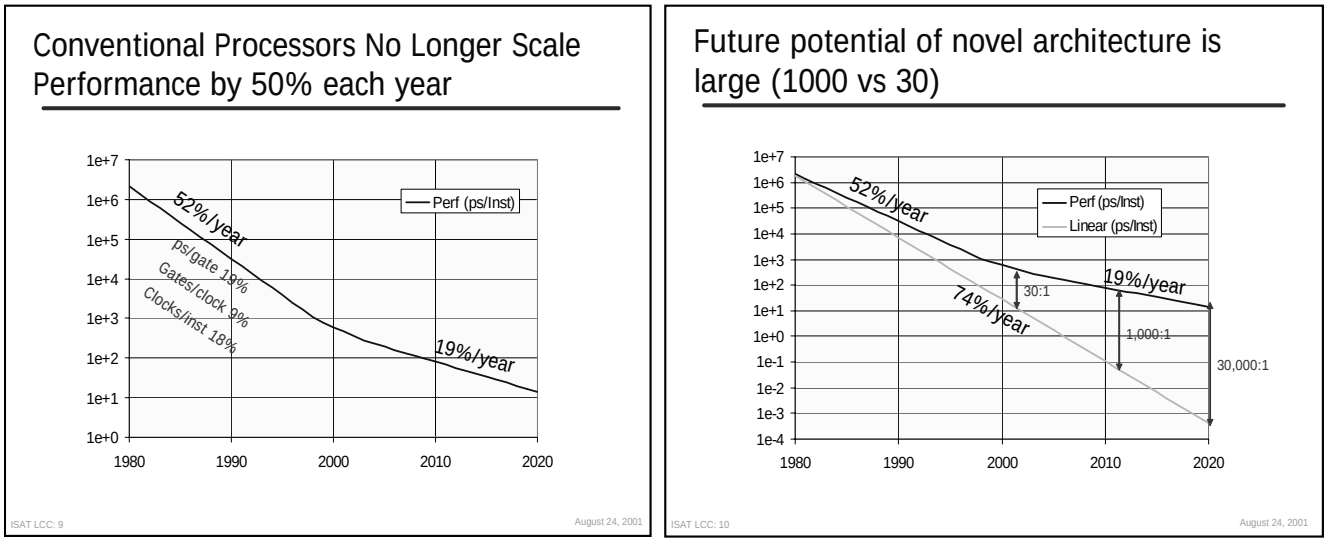


Figure 1

Figure 1 shows that the increases in computer performance experienced in the last 20 years (52%/year) will decrease to 19% per year in the next 20 years, measured in picoseconds/instruction verses year. This is due to the fact that modern processor designs have nearly exhausted the benefits of pipelining. The study reports that "conventional architectures are struggling to sustain even one instruction per cycle. Without further innovations, performance improvements will at best only match the rate of improvement due to further process technology innovations, which is projected to continue at 19% per year."

Novel processor architectures are needed to bridge the gap between performance gains in process technology (19% per year) and the historical gains (process and architecture) of 52% per year. This differential of untapped performance potential, compounded to 2020, results in an untapped performance factor of 30,000. "This quantity represents a tremendous opportunity for novel architectures to help bridge the performance gap and to enable future computer systems to solve increasingly complex and important problems."[4] While maintaining the historical gains may not be technically possible, substantial potential exists and will surely not be realized unless the government makes an investment in long-term research.

Multiple studies, such as the recently completed National Research Council study, conclude that "the supercomputing needs of the government will not be satisfied by systems developed to meet the demands of the broader commercial market."[5] The government must bear primary responsibility for ensuring that it has access to the custom systems that it requires. While leveraging developments in the commercial computing marketplace will satisfy many needs, the government must routinely plan for developing what the commercial marketplace will not, and it must budget the necessary funds.

While instability in long-term funding continues, the government loses opportunities to gain important advances in applications using high performance computing and in its supporting technology.

4. The Last Classical Computer ISAT STUDY, August 24, 2001
5. Getting Up To Speed – The Future of Supercomputing, NRC, November, 2004

DARPA has established the High Productivity Computing Systems (HPCS) programme. This is the only significant U.S. government-sponsored advanced development HPC activity at this time. HPCS is not a research programme. HPCS was initiated in 2002 in response to concerns that commercial systems were not adequate for meeting some very critical aspects of the defence mission. A goal of the HPCS programme is to create a new generation of systems that double in productivity (or value) every 18 months, rather than merely a doubling in unachieved, peak performance.

While primarily a DARPA programme, HPCS has received significant support from other U.S. defence agencies such as National Security Agency and the National Reconnaissance Office, as well as the Department of Energy, the National Aeronautics and Space Administration and the National Science Foundation.

In summary, all these studies have made substantially similar findings and recommendations. The high performance needs for national security will not be satisfied by systems designed for the broader commercial marketplace. A long-term programme funding the development of HPC systems is required to ensure that the DoD and MoD mission agencies can meet their requirements. This programme must fund both near-term acquisitions, alternative architectures, and long-term research; the existing DARPA activities need to be expanded accordingly.

SUPPORTING TECHNOLOGIES

Due to the challenges of building large-scale HPC machines it would be easy to focus exclusively on hardware research; however it is important to address software development tools and other supporting technologies to enable the full benefit to be obtained from these machines. We highlight three of the most important technologies.

Programming Technologies

It is quite difficult to write parallel application programmes. Coherent memory architectures provide perhaps the only existing programming model that effectively decouples the application from the machine-level parallelism. However, there may be no effective way of maintaining coherence over a large distributed system, because of the communication implications. As a result, attributes of the machine-level parallelism must be accounted for in programming, modelling, testing, and scheduling. The software challenges that are unique to HPC include:

- HPC parallel programming systems (languages, compilers, and development environments) that enable effective development of programmes for integrated and distributed HPC,
- Development aids, e.g. tools to partition and predict the performance of algorithms under different distribution strategies, and
- Key applications and libraries (especially numerical methods, image and signal processing, graph processing, and knowledge discovery/management) targeted to all classes of HPC.

These software tools, needed solely by HPC, are not likely to be supported by commercial development in the foreseeable future, if ever. Investments in this area must be a continuing effort of technology refresh, not a one-time development of a new technology.

Co-Scheduling and Collaboration

The vision of multi-disciplinary applications collaborating in a single business workflow implies that it is possible to coordinate the execution of these applications. This might involve two or more large jobs scheduled on two high performance computers controlled by different organisations. Efficient interchange between them must be carefully synchronized to implement the desired workflow. This approach is called co-scheduling.

The notion of a service-oriented architecture is growing in popularity as a way of avoiding some of these issues: organisations expose "services" to provide a lightweight, low-coupled means of interaction between collaborating institutions. The use of a service-oriented architecture may reduce the need for co-scheduling, when services can be used asynchronously and on demand. For some needs, e.g. data mining, and visualization this may be a more satisfactory way of combining applications than co-scheduling.

However, some components of some distributed applications require co-scheduling. Significant development is needed in the theory and practice of federated co-scheduling.

Security

The essence of security for the applications described above is to support collaboration between organisations, in such a way that each organisation can retain control of and protect its own resources. As with scheduling, there are two different interaction models for security: job oriented and service oriented. The job-oriented view is that collaboration involves launching a remote task, and therefore this approach relies on typical operating system mechanisms: user accounts, process sandboxing, group management, and file systems. For grid computers mechanisms such as the Grid Security Infrastructure[6] provide user identity propagation, but support for other facilities is limited.

The service-oriented view is that collaboration involves connection to a service offered by the remote system. At present the protocols of choice are web-services, but standardization has not occurred. Different commercial interest groups who do not see individual gain from standardization have stalled the development of serious collaborative applications using this technology.

6. http://www.globus.org/security/overview.html

RECOMMENDATIONS

1. Fund DARPA's HPCS programme robustly

DARPA should continue to fund its High Productivity Computer Systems programme fully, overlapping each phase to ensure continuity.

Support the Third Phase of the Current HPCS Program

DARPA should continue its HPCS programme through its third phase. This requires funding prototype development by at least *two* HPCS vendors.

Make HPCS a recurring programme

Within a recurring HPCS programme there should be multiple "waves," each lasting seven to eight years. The objectives of each wave will vary, and should be determined by analysis of the evolving mission requirements as well as the pace of commercial technology development. The ending of one wave should overlap with the beginning of its successor. This will ensure continuity of the science and technology. The funding level of each HPCS wave should be approximately $1 billion over its seven to eight-year lifetime.

The performance objective, circa, 2025, is to develop an *ExaFlop* scale HPC system. The work should be undertaken so as to encourage the "flow down" of relevant technologies to mid range HPC systems such as clusters.

2. Initiate a UK HPC Programme

The UK Ministry of Defence and the Engineering and Physical Sciences Research Council should create a programme to perform research for the most demanding military applications with the objective of improved performance using commodity clusters. Specifically, the United Kingdom should complement U.S. activities by:

- Promoting the development of special purpose communications hardware integrating optical input/output with protocols implemented in silicon that have the potential of much higher bandwidth and lower latency than commercial offerings,
- Performing research into novel architectures for special-purpose processors for associative pattern matching, which have the potential to offer at least an order of magnitude improvement in cost-benefit of conventional processors for specialized problems, e.g. image recognition (the work should include production of prototypes and evaluation of their potential on real applications; in some cases this can be achieved with an outlay of around £1M),
- Developing specialized programming aids, tools, and techniques for all classes of HPC machines with special emphasis on tools to plan the partitioning of important algorithms, such as computational fluid dynamics, and
- Defining a clear road map for security mechanisms including standards for web-services security to enable grid computing to be used widely on defence applications, and implementing key security mechanisms which are not being addressed by industry and academia, e.g. to progress work on web-services security.

3. Invest in Research on Critical Applications and Technologies

Both DARPA and the MoD Engineering and Physical Sciences Research Council programmes should address key hardware and software technologies including programming aids and tools for knowledge discovery. The newly reconstituted HPCS programme should be expanded to encompass all aspects of DoD high performance computing; including hardware and software research, prototyping, deployment, and emerging applications. Research in the United States and United Kingdom should address:

- Exploration and prototyping of advanced high-end computing models that emphasize very large memory and its contents as connected "information," not discrete arrays of simple data types,
- Programming tools suitable for HPC systems, addressing parallel programming, both centralised and distributed,
- System software for HPC computers, e.g. to support service oriented architectures and web services,
- Improved knowledge discovery algorithms that can run largely unattended,
- New inference engines capable of translating relatively open-ended high-level queries into efficient search procedures that understand the knowledge sets and their structure as currently available to the system,
- Support tools for rapid, high-productivity, flexible programming and re-tasking of such systems, and
- Appropriate metrics and one or more "open" test bed applications that permit the research community to explore and evaluate alternatives without revealing national security information.

As far as is practicable the United States and United Kingdom should collaborate on such activities.

CHAPTER 4. DEFENCE CRITICAL ELECTRONIC COMPONENTS

INTRODUCTION

The defence industry's leadership role in the electronic component industry has been diminishing for two reasons. First, the growth in commercial demand for electronics results in commercial markets that dwarf defence markets. Second, defence communities have chosen to develop high performance, reliable, complex systems that can only be afforded in small quantities (e.g., satellites and stealth aircraft). It is the ability to embed microelectronics in all parts of a platform, including weapons, that critically enables the design of such complex systems, and thus electronics impacts all aspects of system design.

It is the combination of improvements in targeting, geolocation and navigation precision that enables our militaries to not only hit the target on the first attempt, but to hit in the right spot. Today, a B-2A bomber uses the Global Positioning System (GPS), intelligence imagery systems, Milstar communications systems, and the precision GPS guided joint direct attack munitions weapons to provide the destructive power of an entire air wing in World War II. And, it does so in all weather conditions. Electronics are a major contributor to the net increase of 50,000 in targets per sortie from World War II to Operation Iraqi Freedom.

The Electronics Components Panel assessed the current and projected status of critical defence related electronic technologies. We focused on four principal areas: defence critical component technology, silicon-based component technologies and manufacturing resources, non-silicon based component technologies and manufacturing resources, and available design talent and design tools.

ELECTRONIC COMPONENT METHODOLOGY AND SUPPLIER ASSESSMENT

While federal funding of defence enabling technologies (radio frequency (RF), electro-optical and infrared (IR) devices) continues, the explosion of commercial markets in the last decade, driven by a great consumer appetite for advanced electronic technologies, has fundamentally changed the availability, adaptability and access for leading edge complementary metal-oxide semiconductor (CMOS) component technology for defence applications. While the worldwide semiconductor demand increased by a factor of three to four over the past 15 years, the DoD market share decreased as shown in Figure 1.

Simultaneously, in the United States the DoD contractor base restructured and consolidated, as reported by the Defense Science Board (DSB).[7] These remaining few large contractors are now defined by the large complex system-of-systems they design, integrate, and produce. They sell these highly effective systems in much fewer numbers and rely on small quantities of very specialized electronic components. This consolidation, driven by the reduced demand for defence equipment, also had an impact on the electronic component suppliers' ability to support the unique and often stressing requirements of the defence systems.

As a result, the DoD and the MoD are not only left in the position of being followers in some technologies that are critical to the military, they often have no choice but to rely upon overseas suppliers as well. This trend affects a number of electronic components and has motivated the recent studies by the Defense Science Board and the Advisory Group on Electron Devices.[8] [9] [10]

7. Defense Science Board Report " Vertical Integration and Supplier Decisions", May 1997
8. Defense Science Board Report "DSB Task Force on High Performance Microchip Supply", February 2005.
9. Advisory Group on Electron Devices, "Special Technology Area Review on Commercial Off-the-Shelf (COTS) Electronic Components," February 1999.
10. Advisory Group on Electron Devices, "Special Technology Area Review on Field Programmable Gate Arrays for Military Applications", September 2004.

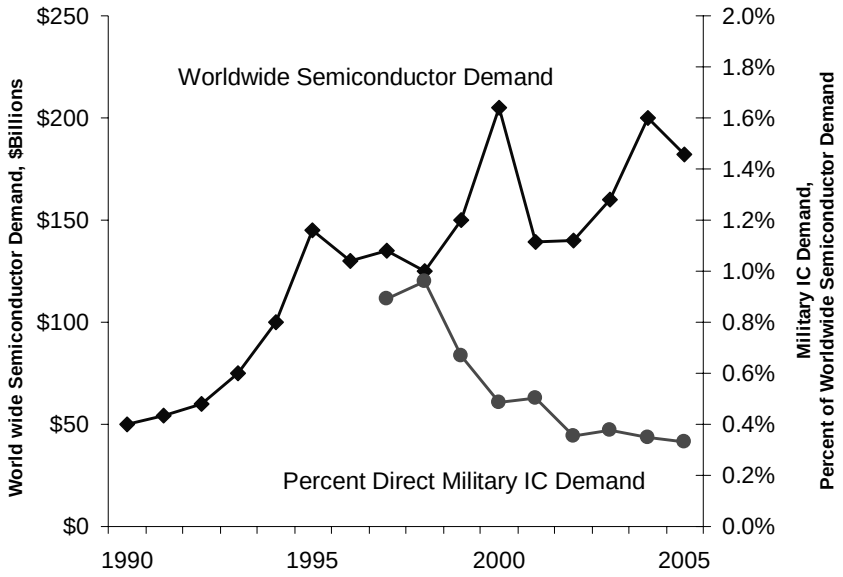


Figure 1. *Reduced DoD Semiconductor Market Share in the last Decade.*

ELECTRONIC COMPONENT CRITICAL TECHNOLOGY ASSESSMENT

Defence critical component technology

Electronic systems play a critical role in nearly all areas of defence applications. Sensors, communications, electronic warfare, command and information systems, intelligence systems, avionics, vehicle electronics, information assurance techniques, weapons, and virtually all logistical and weapons platforms rely heavily on integrated electronics and associated software. The commercial world is also experiencing an ever-growing dependency on electronics in areas such as entertainment, communications, transportation, power generation, lighting, medical systems, and security. As an example, approximately 30% of the value of a modern premium automobile is now the cost of its electronic systems.

This huge investment in commercial electronics brings about both opportunity and threat for defence technology. Opportunity, because much of it can be re-used to meet defence needs (given sufficient

environmental and reliability standards), but also threat due to the accessibility of this technology to military and terrorist adversaries. However, access to commercial electronic component technology does not automatically confer an advantage on an enemy; often systems knowledge, software technology, and operational concepts have a critical part to play. The Export Administration Regulations (EAR) managed by the Department of Commerce attempts to provide the United States with control on the export of commercial dual use electronic components identified on the Commerce Control List that could possibly bring a military benefit to other nations; but often the global market place limits the effectiveness of the regulations and promotes the off-shore migration of the electronic industry. U.S. industry representative indicated that these regulations often have a negative impact on their research and development investments and business strategy. This frequently results in industry moving research, development, and production off-shore.

The relevance of the capability of electronics within military systems can be considered in three categories:

1. Applications where the electronics does not play a vital role in giving the equipment or system a competitive advantage. For example, an intercom system within an armoured fighting vehicle, a laptop computer running logistics software, or an aircraft landing gear control module. In all of these examples the electronics are important, but having access to better components does not bring a greater military benefit. This category was not considered further.
2. Applications where the electronics play a vital part in determining the performance of the military equipment, but where leading edge commercial components provide a sufficient capability. Examples include the use of commercial Field Programmable Gate Arrays (FPGAs) in the latest military radios, digital signal processors (DSPs) in missile seekers and large screens used in tactical displays. In all of these applications, commercial technology provides a "good enough" capability and investing in defence specific variants would not be cost-effective.

3. Applications where commercial components do not meet military requirements, and where military-specific components are required to create the capability. Examples include thermal imaging arrays, analogue to digital converters for some specialized areas of electronic warfare, and radio frequency components for military-band radars.

One determinant of the quality of an electronic component is the quality of its associated design tools. Digital Application Specific Integrated Circuits (ASICs) can be highly integrated microcircuits and could soon reach a density of one billion gates. The design tools used to synthesise a specification into a usable circuit are highly complex software suites. Similarly, the performance of mixed signal ASICs (containing both analogue and digital circuits) and analogue ASICs are to a great extent dependent on the quality of their associated computer design tools. The evolution of the design tools has experienced difficulty in keeping up with the rapid evolution in microelectronic components as reflected in Moore's Law.

The availability of electronic components also depends upon the capability of the industrial base used to supply them, and the means by which governments can influence this supply base. Generally, government-funded, defence-specific electronic components (category three above) will be manufactured on-shore in "trusted foundries," whose security can be controlled. This supply base is reasonably assured.

Commercial components used in defence-critical applications (category two above) can present more problems. Vendors may choose to discontinue components due to commercial pressures such as obsolescence. "Off-shore" (non U.S. or UK) vendors may not be trusted to supply in times of tension. "Off-shore" vendors may have an opportunity to insert unwanted functions (trapdoors) into ASIC designs or to reverse-engineer their purpose. The required environmental envelope of a military application may be outside of commercial specifications, and the components may not be guaranteed to perform beyond such a specification.

While the DSB Electronic Components panel concentrated their analysis primarily on industrial and political issues to enable a more robust and innovative industrial base, the DSAC panel focused their efforts on a "bottom up" analysis of known military capability requirements. The UK results were shared with the U.S. counterparts and vice versa to allow the two efforts to reach a common conclusion. The studies are entirely complementary.

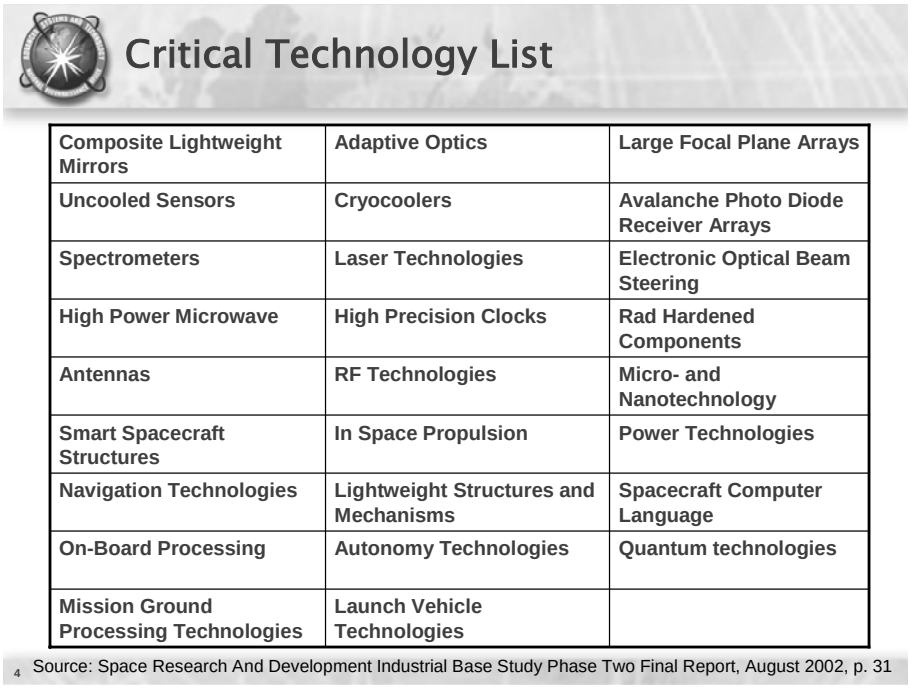
Table 1 below lists the UK's top 15 electronic technologies and their priority. The specific technologies on the list were found to be very similar to the U.S. needs although the individual priorities vary due to differing needs and national industrial strengths.

Table 1. UK Defence Critical Technology list and its associated priority category.

Defence Critical Technology Title	Priority*
Analogue to Digital Converter (ADC)	1
Optically Sampled ADC	2
16-Bit Fixed-Point Precision DSP Implementations	2
32-Bit Floating-Point High Precision Arithmetic DSP Implementations	1
RF Components for Adaptive Array Radar	1
EW and Communications Systems Power Amplifier	1
Components for Antennas - Ultra-Broad Band Manpack Electronic Countermeasures /Communications Antenna	1
Superconducting Filters for EW and Communications	3
RF through Optics	3
Components for Burst Illumination and 3D Imaging	1
Advanced Thermal Imaging Detectors	1
Ultra Fast Photon Counting Technology: Single Chip Photon Counting	2
Fibres for High-Power Transmission and Fibre Lasers	1
Sub-mm Wave (TeraHertz) Detectors and Sources	2
Software tools for design, capture and simulation of components.	2

*Priority ranked as 1 (highest) to 3 (lowest).

Often a technology list is best viewed through the lens of a specific mission or system. It is through the specific system design that the true leverage and value as well as the specification can be calculated. Thus, the priorities are really mission specific and need to be viewed in that context. Figure 2 lists the key technologies as seen by the U.S. intelligence's remote surveillance mission area. It can be seen that many technologies are similar, but also there are others that are unique.



Critical Technology List

Composite Lightweight Mirrors	Adaptive Optics	Large Focal Plane Arrays
Uncooled Sensors	Cryocoolers	Avalanche Photo Diode Receiver Arrays
Spectrometers	Laser Technologies	Electronic Optical Beam Steering
High Power Microwave	High Precision Clocks	Rad Hardened Components
Antennas	RF Technologies	Micro- and Nanotechnology
Smart Spacecraft Structures	In Space Propulsion	Power Technologies
Navigation Technologies	Lightweight Structures and Mechanisms	Spacecraft Computer Language
On-Board Processing	Autonomy Technologies	Quantum technologies
Mission Ground Processing Technologies	Launch Vehicle Technologies	

[4] Source: Space Research And Development Industrial Base Study Phase Two Final Report, August 2002, p. 31

Figure 2. *National Reconnaissance Office (NRO) key technologies required for the future of remote surveillance.*

In considering the criticality of evolving and emerging electronic components it is important to understand the system architecture into which the components are being incorporated. For example, conventional receivers typically have used several stages of down-conversion to translate the RF input signal down to a centre frequency that is low enough to be sampled by a realizable analogue-to-digital (A/D) converter. Each down-conversion added complexity

and sometimes undesired effects that limited overall system performance. As A/D converters become available with both high dynamic range and high sample rates, it is now feasible in many cases, to sample the RF signal directly. This reduces the number of down-conversion stages, and also eliminates the complexity and distortions introduced by the multiple stages. Although this architecture may require a costly state-of-the-art A/D converter, the receiver simplicity and high fidelity performance make it attractive for many applications.

Other notable emerging technologies will impact system design. An example is Micro Electro-Mechanical Systems (MEMS). The advantage of MEMS technologies do not necessarily lay in new functionality or better performing elements, but rather in the miniaturization and higher level of integration.

MEMS is now enabling an even more transformational or disruptive capability, that of an entire sub-system on a single chip. Each key component measuring centimetres has been re-engineered using various MEMS techniques to archive dramatically smaller components measuring fractions of a millimetre, yet still maintaining high performance. An entire GPS receiver can now be integrated into a wristwatch to enable precise knowledge of time and space. A complete networked radio/information system can be integrated into sun-glasses or a helmet-goggle system. This sub-system on a chip technology enabled by advanced ASICs and MEMS technologies show promise and could spur further miniaturization.

Silicon-based component technologies and manufacturing resources

Most information processing is realized through the combination of memory chips (DRAMs, SRAMs, etc.) which store data (including programmes), and programmable components, such as ASICs, application specific signal processors (ASSPs), programmable gate arrays (PGAs), CPUs, network processing units, and DSPs, which perform operations on the data. Of the two classes, the

programmable components have more intricate designs and are where the DoD derives most of its advantage.

Manufacturing

Stand-alone CPU and DSP parts are largely produced by integrated device manufacturers whereas PGAs, ASSPs, and ASICs are often manufactured at independent foundries.

Of the major U.S. commercial integrated device manufacturers and foundries, only IBM, Jazz Semiconductor and Peregrine Semiconductor retain domestic foundries to support leading edge defence applications. Table 2 shows the remaining manufacturers, the location of their U.S. based foundries and the technologies they support. As a result, the Department of Defense has launched a programme to develop a "trusted foundry" model, initially with IBM, with planned extensions to other foundry suppliers.[11] This DoD programme will provide an assured "chain of custody" for classified and unclassified integrated circuits, ensure that there will not be any reasonable threats related to disruption in supply, prevent intentional or unintentional modification or tampering, and protect them from unauthorized attempts at reverse engineering or evaluation of their possible vulnerabilities.

11. See for example, http://www.manufacturingnews.com/news/04/0203/art1.html, or R. Price, "Trusted ICs for Defense Applications", GOMACTech-04 Tutorial, Monterey, 2004.

Table 2. Major U.S. Commercial Integrated Device

Company	Business Model		Manufacturing Location	Available Technology						
	Integrated Device Manufacturer	Foundry Services		Bulk CMOS	CMOS/SOI	SiGe BiCMOS	Rad Hard CMOS	SiGe CMOS	III-V (GaAs, InP)	II-VI (HgCdTe)
AMD	•		TX, CA	•						
AMI	•		Pocatello, ID	•						
ATMEL	•		CO, TX	•	•	•	•	•		
BAE Systems	•	•	NH, VA	•			•			
Cirent Semi	•		Orlando, FL	•						
Cypress Semi	•		CA, TX, MN	•						
Freescale Semiconductor	•		AZ, TX	•	•	•			•	
Hewlett Packard	•		CA, CO	•		•			•	
Honeywell	•	•	Plymouth, MN	•	•	•	•			
HRL Laboratories	•		Malibu, CA						•	•
IBM	•	•	Fishkill,NY, Burlington, VT	•	•					
IDT	•		CA, OR	•						
Intel	•		NM, AZ, OR, CO, CA	•						
Jazz Semiconductor		•	CA	•		•		•		
LSI Logic	•		CA, CO, OR							
Maxim Semiconductor	•		CA	•		•		•		
Micron Technology	•		VA, ID	•						
Northrop Grumman	•		Linthicum, MD	•			•	•	•	
Peregrine Semiconductor	•	•	SanDiego, CA		•					
Texas Instruments	•		TX	•		•				
Vitesse Semiconductor	•		CO, CA	•		•			•	

Manufacturers and Foundries

While the initial Trusted Foundry effort is under development at IBM, DARPA is funding a parallel technical effort to extend the successful Metal-Oxide Semiconductor Implementation System (MOSIS) to include trusted foundry suppliers. The current and planned technology extensions are shown in Table 3 and will provide access to leading edge CMOS technology at modest volumes on a foundry basis.

Table 3. Trusted Foundry Technology Extensions to MOSIS Program

Technology (microns) / Supplier	other	0.5 – 0.35	0.25	0.18	0.13	0.090	0.065
IBM CMOS			6SF, 6RF	7SF, 7RF	8SF, 8RF, 8SFG	9SF	10SF
IBM SOI					9S2	9SFG	
IBM SiGe		5HP OR 5SO 5DM, 5AM 5HPE, 5PA	6HP, 6DM	7HP, 7WL	8HP, 9HP		
AMI CMOS	ABN (1.5u)	C5N, C3N					
TSMC		CL035	CL025, CM025	CL018 CM018			
Peregrine		SOI-SOS					
AMS	BiCMOS 0.8u (CMP)						
OMMIC	GaAs (CMP) 0.2u						
Vitesse	InP HBT						
JAZZ							

■ Available or planned from Trusted Foundry*
■ Available from both Trusted Foundry and MOSIS
■ Available from MOSIS
□ Presently unavailable

Note: MOSIS develops access to advanced processes as soon as a viable customer base exists to support the new process. TFMOSIS will be driven by the specific needs of the DoD in this regard

***** GOMACTECH 2004

The U.S. Integrated Device Manufacturers environment is somewhat healthier, as U.S. based companies continue to lead in the high volume CPU and DSP sectors. Although their manufacturing chains are globalized, a substantial fraction of the leading edge wafer

fabrication facilities associated with these parts remains in the United States. [12]

Component Design

There is evidence that leading edge application specific (ASIC/ASSP) design is migrating off-shore.[13] Perhaps more importantly, the non-recurring engineering costs associated with the development of an ASIC and ASSP have been rapidly increasing, to the point that ASIC/ASSP design is cost-effective only in cases where there is a compelling case that the resultant chip will be produced in sufficient volume to amortize the non-recurring engineering costs.[14] Thus, the use of leading edge ASICs by the DoD is becoming increasingly problematic.[15]

In contrast, U.S. based companies continue to lead in the design of application independent programmable parts (CPUs, digital signal processors, Network Processing Units, Programmable Gate Arrays). For example, Texas Instruments is a leader in DSPs; IBM and Freescale (formerly Motorola) are leaders in CPUs and Systems on a Chips with embedded processors; Xilinx and Altera are leaders in Field Programmable Gate Arrays; and Intel is a leader in Central Processing Units and Network Processing Units.

Placing emphasis on the use of "standard" programmable parts is attractive because these parts are manufactured in high volume, thus when the DoD uses them it is a beneficiary of the economies of large manufacturing scale. However, DoD challenges and opportunities related to various types of Programmable Gate Arrays are so

12. As discussed in the High Performance Microchip Supply standing, part packaging and test is almost exclusively performed off-shore and some IDMs are reliant on off-short mask suppliers.
13. This topic was extensively covered in the "High Performance Microchip Supply" study.
14. There may be additional low-volume cases where the value of integration (e.g., to reduce weight and power) make the non recurring engineering (NRE) justifiable. However, these become less common as NRE costs increase.
15. NRE costs for parts that are a few generations behind the leading edge tend to be considerably lower.

significant that they are addressed in a separate set of findings and recommendations.

Although U.S. based leadership does not in and of itself assure the trustworthiness of these parts or the continued scaling of their performance, it does put the DoD in a superior position to potential adversaries, whose systems rely on U.S. based designs and/or inferior parts. This advantage accrues not only to fielded weapon systems, but to all aspects of the defence community and of our national infrastructures. It would be a distinctly disadvantageous situation if all of our nation's PCs, servers, routers, DSPs, etc. were to some day be built using designs whose suppliers are based in an adversarial country.

The current advantageous position cannot be taken for granted. In the absence of a commitment to U.S. leadership in this space, there is reason to believe that the global consolidation underway in the semiconductor industry will lead to the off-shore migration of these design capabilities. There is some evidence that process has already begun. Of perhaps greater importance is the decline in the DoD's investments in the visionary university research that has sustained the pace of innovation in information processing techniques, i.e., in the design and application of these parts. These reductions are threatening the wellspring of innovation in computer architectures, algorithms, programming languages, distributed systems, etc. Without a continuing supply of such innovations, the ongoing translation of Moore's Law progress into defence systems, and thus our overall information superiority, is threatened. In particular, without dramatic improvements in the design and application of programmable devices, the DoD and MoD will not be able to extract benefit from their unparalleled leadership in many critical defence technologies, such as sensors.

Non-silicon based component technologies and manufacturing resources

In other technologies such as IR focal-plane-arrays (FPAs), radiation hardened electronics, compound semiconductors (III-V and II-VI) and Mixed-Signal ASICs, the United States currently maintains

a significant technological lead. By securing, expanding and controlling this lead, the DoD and the MoD will continue to have the opportunity to leverage technological innovation to maintain a significant edge in system performance.

DoD has traditionally funded leading-edge, low-volume electronic technologies (e.g., Microwave/Millimetre Wave Monolithic Integrated Circuit, IR FPAs, analogue-to-digital and digital-to-analogue converters, etc.). However, for those technologies where large commercial markets have developed (e.g., Gallium Arsenide (GaAs), high power amplifiers for wireless communication, medium-performance analogue to digital converters and digital to analogue converters, etc.) the DoD has found it difficult to maintain reliable onshore suppliers willing to design and/or manufacture for low-volume, specialty military applications.

The panel believes that these specialty electronic technologies can enable unique military capabilities in the future. Specific examples of these technologies include:

- Mercury-Cadmium-Telluride (HgCdTe) Focal Plane Arrays and associated Readout integrated circuits: Large-format imagers and longer wavelength (e.g. Very Long Wave IR) imagers can provide the ability to detect and image relatively cool threats in complex backgrounds. This capability will be a key to future air and missile defence systems. However, difficult material and manufacturing issues (and the resulting costs) make it unlikely that significant commercial markets will develop for these technologies.
- High-speed InP Mixed-Signal Circuits: DARPA's Technology for Frequency Agile Synthesized Transmitters programme has recently demonstrated complex digital logic devices with clock rates in excess of 150 GigaHertz. This technology will enable significant improvement in A/D converters, digital to analogue converters, direct digital synthesizers and high-bandwidth communications. The systems that

will be developed using these components (e.g., radars, signal intelligence receivers, digital radios, etc.) will be lower in cost, higher in performance, higher in reliability, and more adaptive due to the flexibility added once analogue subsystems are replaced with digital circuitry.

- Gallium Nitride (GaN): Newly emerging components in GaN will offer unique performance enhancements for military systems. Prototypes of high-power amplifiers, low-noise amplifiers and even digital circuits have been demonstrated in GaN. Among GaN's unique characteristics are: ability to operate at high temperature, ability to tolerate high terminal voltages, and a presumed inherent radiation tolerance. GaN devices have demonstrated very high power densities and competitive noise figures at frequencies well into the millimetre waves.
- Other emerging technologies: exotic electronic materials are currently being examined or developed by the DoD. Examples include the antimonide compounds, diamond materials, and even carbon-nanotube-based devices. Many such materials have special thermal or bandgap properties that make them potential critical enablers of future systems.

Although the panel was briefed on research and development efforts related to each of the specialty electronics areas listed above, we were not made aware of any coordinated efforts by the Office of the Secretary of Defense to ensure either the maintenance of a critical edge or low-volume design and production capability for these technologies. For these technologies, there is no equivalent of the government supported efforts in CMOS (i.e., MOSIS and the trusted foundry). We heard concerns from the developers and suppliers of specialty electronics technologies regarding the viability of their future design and low-volume production capability.

It is the opinion of the panel that DoD and the MoD should jointly and separately develop and manage a plan to not only maintain our

lead, but to expand it in these technologies. Keeping in mind the lessons learned from CMOS in the 1990's, the panel believes that the DoD and MoD should ensure that onshore design and low- to medium-volume microelectronic fabrication capabilities exist in these technologies.

Available design talent and design tools

The defence trend toward fewer but more capable systems means that there are fewer people in the defence work force with programme experience. Reduced personnel with appropriate experience will ultimately limit the ability to design and develop complex systems and system-of-systems. It is the case that this work force – in engineering to manufacturing – is more efficient due to the use of modern tools, such as automation tools for 3-D computer aided design, simulation and modelling to create a simulation based design capability and the automation of the factory floor with computer aided manufacturing processes.

This limited programme experience base gives more reason to develop advanced design tools that can make an individual more productive. Of particular value are tools to rapidly capture a design in a form that allows parametric functional modelling in order to perform "virtual" evaluation and design libraries containing modular, standard functional blocks that can be rapidly reused.

The solid state electronics which found their way into nearly every aspect of the communication satellites provide a good example of the increasing complexity the designers must manage. As electronic components became more capable, of lower weight and smaller, and more reliable, the satellite system's capability grew along with its complexity. This evolution in capability and complexity are shown in Figure 3. Standard cells and gate arrays used in satellites grew in complexity along with general purpose microprocessors. Since 1995, the introduction of ASICs and FPGAs enabled custom switched telephony (that connects arbitrary caller to arbitrary caller without the aid of any ground based systems) and

permitted the transfer of data at tens of gigabytes per second which made high bandwidth video applications like Direct TV possible.

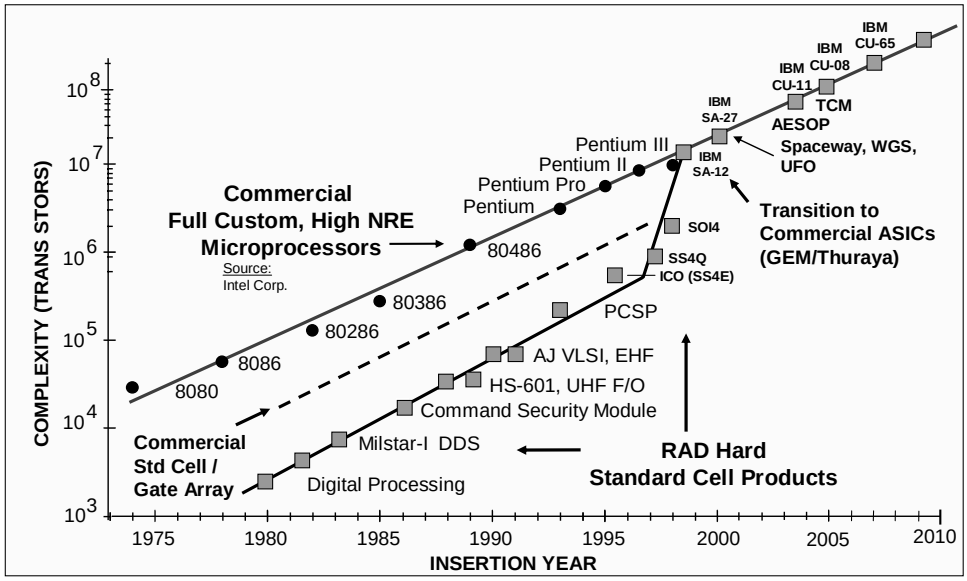


Figure 3. *The evolution of processing complexity over the last 30 years in Military and Commercial Communication Satellites.*

FPGAs are employed in large numbers as a standard way to manage the cost and design complexity of modern electronic systems. It is not uncommon to find subsystems using 100 FPGAs and tens of ASICs all programmed to accomplish different functions, and modern systems having 1000 FPGAs and 100 ASICs in total. If design tools and design engineers were of sufficient supply today, we would see entire systems designed primarily with ASICs due to their higher performance, but due to limitations in both tools and the work force, companies are forced to organize their work force around a few FPGAs and ASICs. This may limit the capability of future systems.

Radiation hardened microelectronics are required to ensure that key military systems can perform in the combined nuclear and natural space radiation space environments. Without radiation hardened microelectronics, U.S. military power – including

conventional military power – and battlefield satellite communications and data transfer could be disrupted. Typical total dose and dose rate radiation hardened requirements for defence systems are illustrated in the Figure 4. Radiation tolerant electronic components are sufficient for commercial communications satellites and many National Aeronautics and Space Administration missions. Requirements for upper radiation hardened electronics are almost exclusively dominated by defence systems. The COTS market for radiation hardened electronics is insufficient to warrant investments by the commercial sector and must be addressed by the DoD. Design talent and tools will only be available if funded by the defence users.

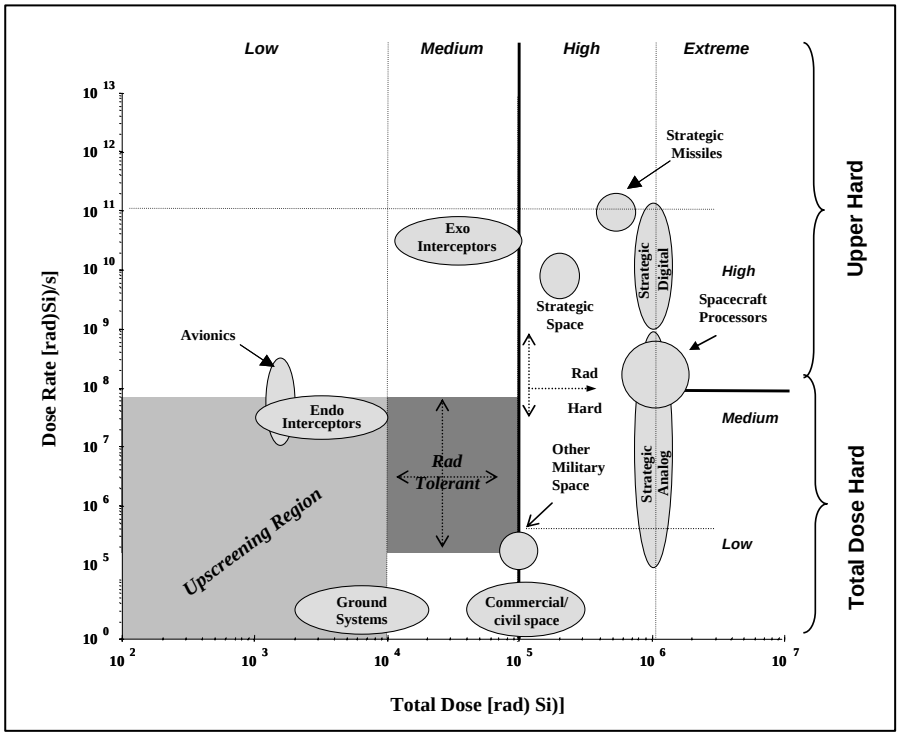


Figure 4. *Radiation Hardening map for Radiation Tolerant to Radiation Hard.*

ASSESSMENT OF ELECTRONIC COMPONENT CRITICAL TECHNOLOGY DEVELOPMENT AND TRANSITION PATHS

In the silicon-based semiconductor space, the industry has partnered with government and universities to establish the Semiconductor Research Corporation and Microelectronics Advanced Research Corporation (MARCO) research consortia. The former operates in a roadmap-driven fashion while the latter undertakes longer term and higher risk research. While successful, the MARCO programme has never been fully funded and the DoD contractor base has not been effectively integrated into it. More importantly, there is no equivalent partnership to stimulate university research in computer architecture or its application (software). Historically, DARPA has been a major agent driving innovation in these spaces. However, in recent years its role and influence over university research has substantially diminished. Arguably, this has resulted in less aggressive behaviour on the part of the university research community, i.e., a shift in the balance towards incremental research as opposed to the visionary, high-risk efforts that have enabled information superiority.

RECOMMENDATIONS

Finding: Maintaining U.S. Leadership in Semiconductor Technologies critical to National Defence

Electronic technologies play a critical role in defence systems; the DoD must continue to protect and support critical electronic technologies to maintain asymmetric advantage. These critical technologies are listed in Table 1 with high priority emphasis on FPGAs, compound semiconductors, electro-optical components; radiation hardened components, as well as high power and high frequency radio frequency components.

State-of-the art wafer manufacturing facilities (fabs) are being established in the Far East in order to support the rapid increase in consumption in the region. As a result, the United States no longer has the asymmetric technology advantage that it once had vis-à-vis

our adversaries and potential adversaries. The United States is not behind. But it is no longer as far ahead as it once was. One consequence of this is that state-of-the-art COTS technology is available to our adversaries.

Some observers' worry that the move of wafer fabrication (fab) to the Far East is the beginning of a trend, and that soon, all wafer manufacturing will move to the Far East, and the United States will not have an assured and trusted supply of state-of-the-art integrated circuits (ICs). However, it is the view of the DSB task force that, in the next 10-20 years, wafer manufacturing in the United States will remain sufficiently strong to support DoD needs in the event of supply disruption in the Far East. One key reason is that the cost of wafer fab in the Far East is not substantially cheaper than the cost of wafer fab in the United States. Wafer fab costs are dominated by capital depreciation, and low cost labour has minimal impact on wafer fab cost. These issues have been addressed by the recent Defence Science Board Task Force report on High Performance Microchip Supply (February 2005). This panel endorses the relevant recommendations made in that report.

Technologies that are critical to DoD needs and that have minimal commercial use require DoD support for the continued advancements needed to maintain superiority. They include:

- III-V components: The DoD need for high frequency and high power components requires III-V technology. Most of the commercial RF products currently use frequencies 5 GigaHertz and below. For these applications, Silicon CMOS, Silicon Bipolar CMOS and Silicon Germanium Bipolar CMOS are replacing III-V components.
- II-VI components: II-VI compounds, such as HgCdTe for night vision systems, remain almost exclusively the domain of DoD systems. There is almost no commercial utilization of II-VI technology.
- Radiation hard components: Commercial satellites require radiation tolerant components which defence

satellites require radiation hardened components. DoD requires availability of "Upper Rad Hard" components. DoD must continue the development of rad-hard technology.

- Technologies for which DoD systems require higher performance components to enhance total system performance beyond the commercial needs include:
  - A/D converters with performance beyond 14-bit at 500 MSPS,
  - Opto-electronic components and subsystems,
  - TeraHertz components,
  - High efficiency, space qualified solar cells, and
  - Readout ICs.

Recommendation: Design Tools. The DoD should develop Computer Aided Design Tools to enable the design of affordable low volume, high performance custom ASICS.

The cost of designing high-performance custom ASICs in 90 nanometre CMOS is on the order of $20M and increases with every technology generation. The high cost of reticles is often cited as the reason for the demise of low volume, custom ASICs in 90 nanometre CMOS, but, in fact, the design cost far exceeds the $1M cost of reticles. DoD systems requirements demand high performance, low volume custom ASICs.

Tool development should be directed to solve the specific problems that result in the increase in design cost for leading edge CMOS, particularly timing closure in physical design and statistical design to optimally design margin for parameter variation.

The tools should be integrated into the industry mainstream toolset for CMOS design. The development can be cost-shared with the commercial industry. The commercial Electronic Design Automation industry is currently developing such tools, but rate of investment is not sufficient to meet the challenge for DoD. While commercial industry will benefit from these tools, the DoD will

benefit disproportionately because of the importance of their contribution to the development of low volume, high performance custom ASICs.

Recommendation: Dual Use Technologies. Notwithstanding the erosion of the U.S. lead in technology as discussed in the DSB Task Force report on High Performance Microchip Supply, the United States has a continuing lead in the design of the following dual-use technologies:

- Fixed/Floating point DSPs and CPUs and their software,
- FPGAs (see finding on use of COTS by our adversaries) and their firmware, and
- Very high performance A/D converters.

Because these technologies are critical for the superior performance of DoD systems, the DoD needs to ensure that this lead is maintained, and that these technologies are exploited for DoD systems.

Recommendation: Trusted Foundries. The DoD should continue and expand the trusted foundry initiative. Additional foundries should be added. CMOS among the trusted foundries should be standardized in order to reduce the cost to DoD for developing new chips and to provide multiple sources for ICs.

DoD should also monitor the number of new U.S. based fabs and provide incentives for U.S. companies to locate fabs in the United States as required.

Recommendation: Military Unique Electronic Components. There are many technologies that have limited commercial use at present and may provide new capabilities for future defence systems. They include:

- Components for Adaptive Array Radar,

- Wide bandgap (power) devices for extreme environments Components for Antennas,
- Superconducting Filters for Electronic Warfare and Communications,
- RF through optics (photonics),
- Components for 3-D Imaging,
- Advanced Thermal Imaging Detectors,
- Single Chip Photon Counting,
- High-Power Fiber Lasers & Diodes,
- Sub-millimetre Wave (TeraHertz) tech,
- Retro-reflective tags,
- Novel materials for antennas, and
- Light-emitting polymers.

These technologies are candidates for joint DoD-MoD technology development efforts. These hedge-technology investments might best be developed in either a collaborative fashion or each country could focus on different pre-competitive technologies and share the research results in a fashion similar to that done by industry.

Recommendation: Talented Workforce. The United States and the United Kingdom face an increasing shortage of engineers, and this shortage affects results in a shortage for DoD and for MoD contractors. The shortage is particularly acute in the area of mixed-signal design. The number of American engineers graduating from U.S. universities is decreasing, and enrolment in U.S. engineering graduate schools is increasingly dominated by foreigners. The United States needs to take action to develop engineering manpower for DoD engineering needs. In 2005, the Director of Defense, Research and Engineering started to address this need for highly educated U.S. citizens in technical fields relevant to DoD by initiating the National Defense Education Program. In addition to vigorously supporting this program, the DoD should do the following to improve the workforce available to support defence electronics needs:

- Encourage development of government sponsored U.S. student industry/academic apprenticeships, and
- Advocate changes to our immigration quotas to enable and encourage immigration and retention of the best students who get their graduate degrees in U.S. engineering colleges and who want to remain in the United States.

Finding: Importance of Adaptability and Programmable Gate Arrays

The United States appears to have a unique advantage in the design of FPGAs and with the design of systems that exploit a mix of FPGAs and ASICs.[16] Recently, the defence primes have begun to adopt Programmable Gate Arrays (PGAs) technology to support a variety of digital signal processing needs. PGAs, especially FPGAs, are rapidly becoming the essential, high performance integrated circuit building block of choice for many commercial and defence systems. As their performance, complexity, cost, and capacity have improved, these devices have begun to challenge the use of Application Specific Integrated Circuits in many electronic systems. In some applications their ability to incorporate built-in core functionality such as those of microprocessors or DSPs have led to preferred system level solutions over traditional design approaches.

A comprehensive Special Technology Area Review (STAR) was recently conducted by the Advisory Group on Electron Devices. Among other things, this review found that the Department of Defense has been an early adopter of FPGA technology to reduce the design cost and development time of technology insertion. Furthermore, the STAR found that a very capable, but currently small, DoD design community is emerging.

16. Special Technology Area Review on Field Programmable Gate Arrays (FPGAs) for Military Applications. Report of the Department of Defense Advisory Group on Electron Devices, July 2005.

The STAR reviewed all major aspects of FPGAs for military use and found that immature FPGA technology has already been incorporated into military systems with significant impact on cost and schedule. The traditional role that FPGAs have played as "glue chips" in electronic systems has changed as their complexity and functionality has increased in recent years. DoD designers have chosen FPGAs over ASICs overwhelmingly because FPGAs have a large variety of intellectual property (IP) hardware cores (e.g. Power PC 405) including various high-speed input/output circuits up to 10 Gigabits per second (Gbps) Rapid input/output.

This panel endorses the findings and recommendations of this FPGA STAR. They include:

- FPGAs have found widespread use and acceptance in digital military electronics systems. The domain of applicability is growing at the expense of general purpose processor and ASIC solutions due to functionality, availability and development/prototyping advantages.
- Processing speed, power consumption/dissipation and functionality are the primary performance selection metrics for applications. Low non-recurring engineering cost, availability, and reduced development time are economic selection drivers.
- Reconfigurability to address evolving requirements and extend mission life remains a selection driver. This is especially important for long life, remote platforms, such as satellites.
- Application for space and harsh military environments is still problematic due to unknown reliability problems and/or lack of design information and qualification/radiation immunity testing.
- System design tools lag hardware in capability and ease of use serving to limit the realization of FPGA technology for military applications.

- The engineering cadre skilled in the design and application of FPGAs in military electronics system design is scarce.
- There is no standardization among FPGA hardware and design software. This restricts portability of designs among and across vendor's product lines.
- Short FPGA product life cycle is at odds with military legacy system logistic demands. A new paradigm is required in the government for planned upgrades.
- The Trusted Source concept is viewed as potentially restricting competition, reducing technology advances, increasing cost and reducing supply.
- Security concerns with FPGAs are not viewed as a severe problem due to domestic procurement of standard and programming control. Battery back-up and encryption storage methods are viewed as solutions to be implemented as required.
- Off-shore fabrication is not identified as a major concern except for the lack of visibility into processing changes at the manufacturer and their potential detrimental affect on performance/reliability. Periodic lot testing is suggested as a solution.

In addition, this panel endorses the recommendations of the DSB Task Force Report on High Performance Microchip Supply. We repeat their recommendation that the DoD support research to enable firmware integrity.

A targeted DoD programme in the area of firmware integrity would likely lead to the rapid development, dissemination, and adoption of improvements to these trust-related aspects of programmable parts. Today's standard parts, especially FPGAs, offer limited protection against the compromise of their firmware, i.e., the configuration software that is loaded into the parts prior to or during execution. The loading of low-level firmware (e.g., the BIOS) into CPUs can also have similar vulnerabilities. However, it is likely that suppliers of commercial parts would incorporate protective measures

if they were readily available. Thus, DoD investment in university research in this area could yield significant improvements in the trustworthiness of standard parts.

Recommendation: Adaptability and Programmable Gate Arrays. The DoD should consider the utility of the creation of a library of re-usable DoD-relevant firmware building blocks. In the commercial sector, one of the attractions of PGAs is the availability those building blocks either in open source or licensable form. This advantage could be extended to the defence sector through the creation of a web-based forum for the exchange of PGA firmware building blocks that can be re-used across DoD systems. The re-use of known good building blocks could reduce cost, development time, and risk. Furthermore, it could ameliorate the impact of the scarce pool of designers with expertise in the hardware implementation of key signal processing algorithms.

Finding: Commercial Electronics marketplace has made obsolete the U.S. Export Administration Regulations (EAR)

The U.S. EAR is based on the assumption that technology advantage resides to a great extent in the United States, and its export should be controlled. In electronics, while U.S. companies do have certain technology advantages, commercial companies have global markets for both development and sales. Forces in those markets, together with impediments created by the Export Administration Regulations, often promote off-shore migration of technology development and production capability of the electronic industry.

Where EAR identify technology capability thresholds, sometimes those thresholds are not maintained at appropriately discriminating levels. EAR restrictions can even encourage the development of foreign sources of technology and are completely ineffective in keeping critical technology from our potential adversaries where components of comparable performance are available elsewhere. These problems with the current Export Administration Regulations result in a loss in technology development investments by U.S. companies in addition to a loss of business and a high administrative

cost of compliance, sometimes with no redeeming protection of U.S. technology.

This issue of economic competitiveness is also relevant to ITAR regulations, which are focused on controlling the exports of items and technologies with direct military (and little or no civilian) application. With increased reliance on foreign military sales and technical cooperation with international partners, a more agile process that balances security and industrial concerns are needed in order to deliver military capability in a more timely and efficient manner.

Recommendation: The United States should re-evaluate export controls (EAR and ITAR) as it pertains to the critical microelectronics. New regulations should recognize modern market forces and allow the export of technology that is available in foreign markets. A time phased downgrading should allow older generations of technology to be exportable. This will help to balance the threat of the loss of technical leadership against the restriction of technology released to foreign markets. Ideally, new regulations would allow the United States to export technology that is just a bit better than the foreign competition.

Finding: Use of COTS by our adversaries

Modern COTS electronics technologies are enabling both nation-states and terrorists to obtain significant military capabilities at modest costs. Furthermore, the rapid pace of change in the commercial world allows such adversaries to quickly evolve and advance their capabilities.

In the past, the development of key military technologies (e.g., computers, night vision, spread-spectrum radio, etc.) was controlled by the military. Governments spent large sums of money developing and controlling these technologies. The huge consumer markets for advanced technologies have not only overtaken the government influence in the electronics business, but the magnitude of these markets has accelerated the rate of technological advance.

Video games, computers and consumer electronics drive the processor markets. Telecom drives the networking market. Cell phones drive the spread-spectrum and software-controlled radio businesses. Even DoD-funded technologies that were once controlled by DoD (such as GPS) had to yield to commercial market influences. Extremely powerful COTS technologies such as Wi-Fi, Bluetooth, public-key encryption, and the Internet continue to evolve and advance at a dizzying pace due to commercial market forces.

Very effective and militarily significant tools are at the disposal of our adversaries. The improvised explosive devices used today in Iraq are one example of this. Commercial satellite imagery is widely available and subject to resolution enhancement by COTS processors. Hand-held GPS units can provide precision target locations. The internet, cell phones, pagers, satellite telephones, etc. can be combined to provide a robust, global command and control network at an insignificant cost. Furthermore, if the past is any predictor of the future, very substantial advances in these kinds of capabilities at even lower costs can be anticipated.

Recommendation: DoD and MoD need to prepare for dealing with the COTS-equipped adversary. Studies should be initiated to understand the strengths, weaknesses, and vulnerabilities of COTS based systems that can serve key military capabilities, like communications, sensing, etc. These studies could lead to the development of counter-COTS systems to exploit vulnerabilities. Exploitable vulnerabilities that should be explored include intercept, detection, geolocation, infiltration, and spoofing. It is equally important that the DoD and the MoD reduce the acquisition cycle time for electronic components for its'systems in order to exploit the rapid pace of change for intelligence, command and control, and weapons systems. This could be accomplished by designing systems intended to be adaptable to a rapidly evolving threat based on leveraging the commercial market offerings.

Finding: Failure of COTS to meet DoD/MoD Quality & Reliability requirements

COTS integrated circuits have different design, testing and environmental requirements and are not suitable in many military applications without additional adaptation and/or qualification. As the DoD and the MoD have come to rely on these COTS parts, some critical systems have become vulnerable.

Some of the environmental challenges are relatively obvious, i.e., COTS ICs are explicitly not designed or tested for space applications. However, there are other cases where a COTS part might come close, but not quite be able, to meet military requirements. For example, COTS ICs typically are not intended for use at the extremes of the outdoor (artic and desert) temperature ranges to which military systems are sometimes subjected. Similarly, the performance of a COTS part might degrade over time in a manner that does not compromise its integrity during its specified commercial lifetime, yet may compromise longer-lived military systems.

In some of the above cases, it might be possible for a third party to qualify parts for the harsher environment through extended testing (e.g., for lifetime and temperature range), environmentally controlled packaging, etc. However, the opaqueness of COTS parts, i.e., the limited availability of design information, can mask fundamental limitations that may not be apparent through such "black box" testing.

One particular area of concern is with testing, which is a substantial fraction of the manufacturing cost of modern integrated circuits. Manufacturers rigorously test their products to reduce the substantial cost of allowing parts that do not meet their specifications to "escape" into the market resulting in customer disappointment, recalls, brand impairment, etc. Nonetheless, the degree of testing to which parts are subjected is the subject of a cost and benefit analysis that trades-off the incremental costs of additional testing against the risks and potential costs of escapes. A key assumption underlying these trade-offs is that most COTS ICs are not intended for use in "life or death" applications. Thus, when escapes occur, their remedy

typically involves short-term inconvenience rather than loss of life or the long term impairment of a strategic asset (such as a space vehicle). Escaped parts may produce dire consequences in military situations.

Recommendation: The DoD should forge coalitions with other sectors that face similar stringent requirements for COTS parts that meet higher quality and openness demands, e.g., the automotive industry, public utilities, healthcare, etc. These broader coalitions should find a common set of requirements that meet their reliability, openness, and quality needs. Of these, the automotive industry may be the most advanced in obtaining COTS-like parts that meet quality standards considerably above those of consumer electronics.

Finding: Where are the new innovations and technologies?

One final concern that arose during the panel's deliberations is the belief that the list of critical technologies is surprisingly familiar with few, if any, being new additions that would not have been on a list compiled five or ten years ago. Of those that are new, many are technologies that can be substituted for existing components – as opposed to those that create revolutionary new capabilities. A failure to discover and develop revolutionary new technologies is of particular concern, given the belief that new systems concepts often arise in response to the development of new component technologies.

Recommendation: The panel recommends that the DoD and MoD conduct a longitudinal analysis of the emergence of novel electronics to determine whether or not the electronics "discovery engine" has slowed and, if so, identify the root causes, including the impact of organisational and process changes. For example, one process hypothesis to be investigated is whether the early coupling of research to DoD acquisition programme offices is impeding the development of new technologies, i.e., if a key criterion for investment in a new component technology is the support of a pre-existing systems programme office then one should not be surprised to see a dearth of revolutionary new technologies of the sort that in

the past would have led to the creation of new capabilities. This type of linkage might be particularly acute in the current environment, in which few new systems are under development.

SUMMARY

The DSB and DSAC Electronic Components panels have studied a number of specific defence critical electronic components with a view to examining the role they play in determining the performance of defence systems, and the key issues that are facing our defence industries in the light of ever increasing globalization. The critical security of supply chain and the robustness and innovativeness of our industrial base were reviewed and recommendations made to better focus and enable their long term health as well as the security of our respective nations.

CHAPTER 5. ADVANCED COMMAND ENVIRONMENTS

INTRODUCTION

The reason for our interest in Advanced Command Environments (ACE) is to enhance the command function and increase military effectiveness. Command is at the heart of military capability and is essentially, a human activity. We therefore take a human-centric view of ACE.

The U.S. and UK panels encountered the ACE problem first hand in the course of this study. Information technology and current state-of-the-art commercial tools fell far short of enabling collaboration at a distance, even for these modest sized panels. It was not until the UK and U.S. panels decided to physically meet and visit some of the key military locations to hear, firsthand, the warfighter's view, that consensus on the goals of the study emerged. The U.S. team came to agree with the UK team that without a framework for analysis and metrics for assessing a particular technology, there would not be much chance for sound evaluation.

DISCUSSION

Definition of Advanced Command Environments

We believe it is appropriate to use a broad definition of a command environment if we are to ensure that all components are successfully integrated and together support superior human command performance. The ACE must therefore take into account physical layout, equipment, information, processes, organisation, team membership, doctrine, culture, and any other relevant contextual factors.

An advanced command environments has component elements including the physical surroundings of the members of the command staff, the information space that the ACE supports, and the physical

and social connectivity provided to material assets and to other members of the command staff and its supporting organisations. In addition, the environment is so integrated into command staff activity, that it must support the formal and informal processes being executed by that staff. This holds true for all levels of command.

Physical Surroundings

The physical elements in the near vicinity of command staff members form their surroundings, architecturally and ergonomically. Physical objects, light, sound, and the way that the material objects affect humans all play a part in the command team's ability to be aware, understand, remain focused, carry out physical and mental tasks, and not suffer undue fatigue over the watch time.

Information Space

The command environment supports the presentation of a virtual information space or volume. Just as the personnel occupy a physical location, they operate within this information space. The most critical aspect of the information space is whether each member of the command staff receives, digests, and can use the information elements needed, while being able to ignore information that is irrelevant or of insufficient priority for the task being carried out by the human. The job of the environment is to permit humans to operate at high productivity levels.

Technological systems within the command environment present information, graphically display it in myriad forms, on demand, and automatically. Other systems process extant information to change its form and to derive new information. Technology for processing and displaying information has advanced to the stage that command staffs demand what they describe as situational awareness presentations, which are the result of accumulating and amalgamating raw and interpreted information from many sources into a holistic view. The advanced command environment should present situational awareness descriptions in a fashion that best

serves the specific commander and staff, with their own knowledge, insight, and judgement.

Information can be presented in myriad representations. One of the challenges of designing an advanced command environment is to determine how to adapt information from one format to another – without loss of accuracy – and to amalgamate information from many sources; possibly deriving new information that some may view as new information or even knowledge. Because the information fed into a command environment is typically created independently from the command environment, integration of both representations and even of information itself is an immense challenge. Feeds may be technically and semantically incompatible.

The command environment must support an understanding of situational awareness that appears to the command staff to be stable over time. Information needs to be timely, if feeds permit. Information that will be presented repeatedly needs to be kept in data bases accessible to command environment systems. Hence, there is substantial background task of maintaining both historic and reference data, as well as the information in the instantaneous situational picture.

Connectivity

The command environment provides connectivity from the command staff members to assets that they control on air, land, and sea, as well as, network connectivity to remote members of the staff, or to supporting and superior organisations. Network bandwidth depends upon the locations of the communicating entities. Over time bandwidth improves as communication technologies improve. In the near future new capabilities such as Transformation Communications, Future Combat Systems, and the Joint Tactical Radio System should enhance bandwidth. Today, in practice, information feeds are gated by the ability of the command environment to receive and to send using the network assets deployed as well as competing demands on backbones. Latency, the time for a signal to travel from end to end, and therefore interactive

responsiveness is limited fundamentally by the speed of light, and practically by less than idea communications.

Support of Command Staff Processes

The advanced environment is more integral to the activity of the command staff than environments of the past. As a result, the environment must effectively support the command staff activity and the human value it brings to operations. This entails support for its procedures, often step by step. Consequently, the environment must scale to fit location, communication capability, size of the command team, and its specific objectives. For coalition partners in a command environment, the handling of multiple levels of security on a "need to share" versus the current "need to know" is required both for current data and for the background intelligence needed for interpretation. The environment must allow a command thread, from initiation of a command action through to its execution, to form, perform, and disperse on an ad hoc basis. In summary, the environment needs to support organisational as well as individual agility.

Our assessment of ACE must be focused on the output, or end-product. Our measure of "goodness" is the effectiveness in delivering desired military effects in the desired manner. Concepts and doctrine, types of operation, and conditions of operation (e.g. in coalition, with global visibility) will determine the what, when, and how of the ACE outputs.

The idea of a human-centric view of the architecture is a key theme in our findings. Without the ability for human beings to interact around diverse perspectives of the same situation, decision-making, effects-based operations and collaborative evaluation are not easily achieved. This was illustrated by the panel's own experience of collaboration. The social (human) architecture enabled us to reach consensus while an architecture based on technology (video teleconference, etc.) did not. Furthermore, this approach allowed us to consider the issue of organisational agility. To paraphrase a

concept paper prepared by the Joint Doctrine and Concepts Centre[17], by *organisational agility* we mean the ability of the humans in a command organisation to think creatively, adapt, improvise, and respond to the unexpected. It is what humans do best. One of the dangers of modern command systems is that they can allow or even lead a commander to "over-control" the environment which may prevent necessary improvisation at the lower levels of command. Organisational agility is characterised by four attributes: responsiveness, robustness, flexibility, and adaptability. If an advanced command environment is to support and enable organisational agility, then that environment must be designed around the human participants, answering to their needs and amplifying their strengths and capabilities.

Our focus remains on the development and use of technologies to enhance command, and the key to doing this effectively is to consider command environments to be human-centric systems. Since the United States and the United Kingdom can expect to be part of the same military "system" (projecting a coalition military capability) in the future, the development and application of a human-centric command environment should be an area of technical collaboration between us.

TECHNOLOGIES ASSOCIATED WITH ACE

The critical technologies other than human-centric methods that are associated with ACE are in various stages of development in a wide variety of commercial and government markets. Three technologies will drive the command environment as key enabling elements. The first is large scale media management as practiced by major video broadcast news agencies such as FOX, NBC or Sky News. The second is remote collaboration which is emerging in global enterprise management in both the government and the private sector. The third is in visualization technology which has its roots in the entertainment and media sector.

17. The UK Joint High Level Operational Concept: An Analysis of the Components of the UK Defence Capability Framework, para 202

Large scale media management encompasses the end-to-end flow of information from the creation, development, retention, dissemination, and destruction of each bit in a distributed system. The required technologies for end-to-end media management ($E2EM^2$) in ACE are network management for converged information architectures. Multiple protocol label switching will allow information to flow through the environment as voice, video or data all in the Internet protocol format while accommodating the widely different data rates, display capabilities or security of end devices or users. Very large database technologies, such as tuple storing, will allow the rigor of a structured database to have some of the advantages of Google's capability. Databases will be measured in exabytes not petabytes but will still have the required control for command. This will enable search and access using XML and will create the ability to horizontally integrate traditional islands of automation without converting relational databases. $E2EM^2$ will also benefit from new storage technologies such as laser optical tape and metadata generation technologies.

Collaboration technology is accelerating in the large multinational corporations. Collaborative network environments that run on top of $E2EM^2$ systems allow for teams to develop a joint understanding of the opportunities and constraints that are before them. Pharmaceutical research companies are moving quickly to deploy corporate wide collaboration technology. Key elements of this technology are advanced video teleconferencing; human interface devices such as video tables, large displays, eye-limited resolution work stations, gesture recognition tools, virtual environments, attention cueing, graphical relations analysis tools, Bayseian-based decision tools, and chat schemes. These technologies are efforts to improve the human to human connection.

Visualization technology is rooted in the entertainment sector. The use of special effects to convey a complex situation is becoming more important in the art of story telling. In the ACE, the $E2EM^2$ will perform the complicated information tasks but the visualization layer will allow the command team to do what humans do best which is to understand the complex. A critical step between information and

knowledge is visualization. Key advances in this area are very large seamless displays. Synthetically fused images from heterogeneous sources create an almost omnipresent experience for the commander. Visual simulations of real world models are commonplace in the massive gaming industry. Holographic and other three dimensional techniques are in development to give better situational awareness and improve the common operating picture.

THE HUMAN AT THE CENTRE OF THE ENVIRONMENT

The challenges of harnessing new information technologies to design and operate advanced command environments are great. While the human must be at the centre of the environment, the technologies permit changes in process and organisation that can yield new and useful capabilities. The environment must enhance and not hinder the human beings who are providing flexibility, rationality, accountability, and creativity. Ingenuity will still come from people, not from automated systems.

Three major defence challenges exist when developing advanced command environments:

1. Technology must not erode or limit valuable human capability. Yet, for some highly sophisticated, technical system-of-systems, the human appears to be the limiting factor. Key concerns include:
    - Constraints on people imposed by technology that may erode the value humans bring to military capability (e.g. information management demands can constrain flexibility),
    - Demands on people to deal with a greater complexity and range of functions,
    - Demands on people to respond more rapidly than they are comfortable with,

- The diversity of demands of future operations and command, including the need to carry out very different tasks,
- Insufficient integration between teams, horizontally and vertically, to create a seamless networked capability,
- Difficulties in team interactions involving a wider range of cultures, processes, situations (e.g. more remote team working, other government departments, coalition) and the need to rapidly integrate and collaborate with many new people in coalition command centres,
- Insufficient, and diminishing, numbers of people and lack of appropriate skills, and
- Greater technological and information overhead.

2. New command-related concepts are being developed which rely on a greater understanding of human interaction and performance issues. These concepts need to be evaluated in the new military environment. Some of the concepts are:
   - Distributed decision making,
   - Agile mission groups (forces that can be configured on demand to address a wide range of effects based operations; fighting, peacekeeping, humanitarian, etc,
   - Increased reach-back,
   - Command "by exception,"
   - Command in a distributed environment,
   - Effects based operations/comprehensive approach, and
   - Large scale modelling, simulation, and war-gaming.

3. The way in which existing command concepts, such as mission command, can operate in the new military environment of Effects Based Planning, is not understood. Current military planning has a constrained set of options that is quickly expanding to include the use of military systems for relief efforts, counter-terror, policing, and administration of local governments. Given the sensitivity of human behaviour to environmental influences, we cannot take for granted that current effective human performance will continue unaffected as we introduce new technology and information.

An advanced command environment is intended to support transformed command processes. Compared to processes of the past, tomorrow's command processes will involve faster decision cycles, orders of magnitude more data available to be digested by the command staff, and more diverse sources of data as well as human perspectives. The desire to build advanced command environments breaks new ground. Today's experts have insufficient understanding of how human beings interact, reason, reach consensus, and make decisions in a fast-paced, information-rich, collaborative environment. The relative immaturity of knowledge about human beings is a formidable impediment. As we have said before, the command environments must be human-centric. Therefore, the issues of how humans perform and behave must be considered early in the design. It is very difficult to articulate requirements for systems when fundamental knowledge of how to design information systems to support human processes is not well understood.

There has been insufficient investment in research to achieve an understanding of how humans perform in a system-of-systems military context, where technology, information, organisation, culture, process, and people are integrated. More specifically, new command concepts have not been well-articulated and thoroughly tested. So, it is not surprising that there is difficulty in articulating the requirements for command environments that can deliver the underpinning to support those concepts.

Understanding and applying human science to ACE (and other military systems) is critical to the exploitation of technology and to

the delivery of overall military capability. Human science is the means for transforming the components of a command environment into effective military command.

THE SCOPE OF "HUMAN DISCIPLINES"

Many disciplines contribute to better understanding human behaviour in a command environment. Many disciplines can contribute knowledge to be employed in creating more effective systems and environments. These disciplines include psychology, ergonomics, human factors, cognitive science, neuroscience, biology, archaeology, philosophy, history, literature, sociology, anthropology and more. Some are "hard science" in the sense that they use rigorous methods such as laboratory-based experimentation and mathematical analysis. Others are "softer" using methods such as field studies and ethnography. In each case there must be a methodical and controlled approach to knowledge.

We shall use the term *human knowledge base* to describe the body of knowledge that offers relevant observations and understanding that relate to human performance in the command environment. The skills and methods required to apply human knowledge base is a discipline in itself. We shall use terms such as *human factors integration, human systems integration,* and *human factors engineering* to refer to the application of such knowledge. The first term is in common use in the United Kingdom, and the latter two in the United States.

Human behaviour processes, especially those used in complex environments are not well understood. Some experts contribute perspectives as they describe "system thinking," "complex adaptive systems" and "socio-technical systems." It is understood that humans are strongly affected by influences external to the immediate system in which they are working. Influences include cultural attitudes, expectations, motivations, and values. In some cases they are not taken sufficiently into account. A system-of-systems approach will deliver a more reliable assessment of overall performance by factoring in the broader influences which determine human

behaviour. To do so necessarily calls upon knowledge from a broader range of human disciplines. Operational analysis techniques are being developed to carry out integration at this level, and the human factor will be central to this. The ACE panel finds that there is a critical need for the development of more advanced means of applying the human knowledge base to the design of future advanced command environments.

SPECIFIC AREAS TO BE ADDRESSED

The panel has identified that there is insufficient application of human factors in the development of military command capability. There are research findings in areas such as human decision-making, remote and co-located team-working, cultural determinants of behaviour, intuitive and analytical thinking, the use of information and knowledge in developing understanding. However the findings from such studies need to be applied in the context of military capabilities and used to inform doctrine, organisation, process, and technology application. The ACE panel suggests that there are several ways in which this can be done more effectively:

- Experimentation to understand human (military practitioner) requirements for ACE using new technologies in display, communications capability, visualization, special effects to replace iconography, and simulation tools,
- Experimentation to assess the impact on individuals and teams of potential ACE or components of ACE,
- Full utilization of both human discipline specialists and human factors engineers throughout the ACE lifecycle from capability analysis to termination. Human discipline specialists will have a deep knowledge of areas such as cognition, sociology, problem solving, decision making, and team building. Human factors engineers will have experience in the practice of incorporating knowledge about human behaviour into the systems engineering process,

- Modelling human cognition and behaviour to develop technologies to support own command behaviours and to help in the analysis of opponent behaviour,
- Development of metrics, language and taxonomy to assist in multi-disciplinary collaboration in a human-centric systems approach to command environments, and
- Early engagement of stakeholders, especially the future commanders, in the design process.

The explorations proposed in the list above can only be conducted with substantive military involvement. Indeed, the panel believes that outside contractors and universities are generally unable to do this work in a way that sufficiently involves the military. Consequently, we believe that a concentrated programme needs to be created by the two nations. In the next section we discuss two organisations within the MoD and the DoD that have demonstrated capability in this area.

JFCOM AND NITEWORKS

Two leading stakeholders in the area of developing advanced command techniques and environments are the U.S. Joint Forces Command (JFCOM) Joint Futures Laboratory (JFL) and the UK NITEworks (Network Integration Test and Experimentation Works).

The JFL conducts transformational experimentation locally or globally, through live or virtual means. The JFL facilities serve as the hub for the Headquarters of the Combined Joint Forces Command-Future, a community of joint and service headquarters linked through common Collaborative Information Environment tools, networks, modelling and simulation federations, and methodologies to integrate and synergize joint experimentation. The JFL and the Distributed Continuous Experimentation Environment conduct continuous, collaborative joint experimentation to accelerate solution sets to current warfighter challenges, develop potential solutions and opportunities for future warfighting, and inform development of the Joint Force Headquarters of the future. JFCOM also works with the

Defense Advanced Research Projects Agency (DARPA) to develop transformational technologies to enhance the capability of commanders and staffs to collaboratively plan and conduct effects-based campaigns.

NITEworks is an experimental environment that allows the UK Ministry of Defence to assess the benefits of Network Enabled Capability (NEC) and the options for its effective and timely delivery. NITEworks is a partnership between the Ministry of Defence and the defence industry rather than a traditional and more formal customer/supplier relationship. This unique arrangement allows NITEworks to draw on the widest possible range of specialist skills, information, and facilities in addressing problems set by the Ministry of Defence. In brief, NITEworks conducts activities such as warfighting experiments, including manipulation experiments, and empirical studies, tests and visualisations. Through NITEworks, various industry capabilities are accessible including: training simulators, test and evaluation facilities, the Battlespace Transformation Centre, and the Battlespace Management Environment.

The JFCOM Joint Futures Laboratory and NITEworks are not the only candidate organisations that are qualified to be involved in this kind of experimentation. Appendix E lists other organisations in the United States and abroad.

EXPLORING COMMAND ENVIRONMENT SUPPORT FOR ORGANIZATIONAL AGILITY

Anticipating our recommendation for joint experimentation between JFCOM JFL and NITEworks, this section explores a concrete example of the kind of exploration that would be valuable. It focuses on the dimensions and techniques that would enable a command environment support system.

Experimentation, the development of high level operational analysis, and the involvement of multiple stakeholders are already well in evidence (JFCOM and NITEworks), and these activities could

be built upon effectively. Table 1, below, provides an example of some of the concerns and potential solutions that could be assessed through the application of human science to ACE concepts.

Table 1. A U.S.-UK Common Goal

DELIVERING AGILITY THROUGH THE COMMAND ENVIRONMENT*(in the UK's Joint Doctrine & Concepts Centre definition this includes: flexibility, adaptability, responsiveness, robustness)* Both the United Kingdom and the United States wish to increase military agility at all levels and scales. Command will play a key part in achieving this.	
Obstructions to agility	**Potential sources for enhancing agility**
• Compartmentalisation of process and information for security reasons • Necessity to integrate new people with different cultures for the purposes of effects based and coalition operations • Separation of decision makers and warfighters through reach back • Static structures, stove-piped information environments	• Re-configurability of ACE to look inwards or outwards, to work as a whole or splinter into groups and individuals, to integrate back into a team. This will support different command styles and doctrine, the ability to deal with both micro and macro threats, and the movement of people, tasks and information across different environments • Concept of operations to allow simultaneous collaboration & compartmentalisation
• Lack of exploitation of lessons learned • Long design and modification cycles for command environments	• Rapid processes for responding to lessons learned (organisational learning); technology to assist in rapid learning; self-learning technology (e.g. modelling) • Involvement of all stakeholders (including suppliers) in early activities to understand requirements
• "Locked-in" thinking arising from expectations, dependency on process, rule-based behaviour and inappropriate simplification of complex situations	• Visualisation and thinking aids to help challenge assumptions, encourage lateral thinking, support broad and deep understanding of a situation • Cultural diversity
• Over-dependency on technology; impoverishment of human understanding through de-skilling and removal from detail; misunderstanding of application functionality and mode status (e.g. modelling tools) • Erosion of human perception & understanding – through reducing the communication bandwidth (using IT not face to face) and trying to pump more information through this reduced bandwidth.	• More effective modelling, simulation, and visualisation to match cognitive styles of commanders, analysts and planners. • Tools, organisation & processes to assist in rapid thinking (e.g. brainstorming tools) • Policy for face to face communication
• Effects of accountability and global visibility	• Training, organisation, tools for providing audit trail of discussions and decisions

RECOMMENDATIONS

1. Collaborative Experimentation

The United States and the United Kingdom should develop a cooperative programme to collaborate on the physical design aspects, internal functionality and tools, and other human factors aspects related to optimizing future command/decision environments. The purpose of such efforts would be to spotlight the importance of optimizing decision environments at different levels of command, and to formalize opportunities for commands to draw from the technologies and human factors work already being done in battle and futures labs working in these areas across the respective governments and in industry. As a first step, we recommend a trial link between U.S. and UK facilities, such as NITEworks and the JFCOM JFL, and possibly other UK/U.S. military organisations (see Appendix E for additional organisations). This would host an initial set of experiments in the area of advanced command environments to assess the effectiveness of alternative architectures and components and supporting future command modalities. Specific issues to be explored are discussed in earlier sections.

This experimentation should develop a model for more expanded collaboration on this topic, including:

- Provision of a comprehensive method, including a definition of the scope of the operational context we need to address in order to assess command functions,
- Establishment of measures and metrics for the output of ACE and the performance of elements within them such as human command performance or the extent to which a technology supports or hinders specific human performance,
- Develop specifications and prototype requirements for ACE and their components in order to help define future research needs and to

frame the findings in terminology meaningful to both the militaries and supply industries, and

- Test and evaluate the most relevant technologies such as:
  - Display technologies to support partially shared information environments,
  - Special effects as a visual tool to enhance situational awareness,
  - Technologies to enhance remote collaboration,
  - Large scale media management systems,
  - Human-in-the-loop modelling (using expert judgement to temper complex computation),
  - Tools and methods for enhancing lateral and intuitive thinking,
  - Training and education techniques for cultural understanding,
  - System interfaces to intuitive, emotional and social dimensions of human cognition,
  - Generic, agile planning aids, and
  - Large scale modelling and simulation for war-gaming.

An important goal should be to identify which aspects of the command environment are generic, and which are context-dependent (e.g. dependent on particular scenarios). This can be achieved by using three or more vignettes or scenarios representing different extremes of key variables.

A baseline should be established with existing conditions (e.g. a single commander operating hierarchically with existing command information systems using current doctrine for group decision making and course of action analysis).

2. Conference

The programme should also sponsor a conference with a call for papers to address the broad range of topics including behavioural modelling of specific populations, individual person modelling, methods for characterising individual members of a hypothetical command team, ergonomics, colour, light, sound, etc. These ideas should be evaluated and either dismissed or adopted as part of a broader command environment toolset after experimentation.

It will be important to have key high-level stakeholder buy-in from the start and a commitment and mechanism to act on the outcomes. There will be important issues about how to provide resources for such a study, what timetable is feasible given existing programmes, who would have ownership of such an over-arching initiative, and what needs to be in place before such experimentation can take place.

SUMMARY

As part of the terms of reference for the working party, each panel was tasked to develop a critical, disruptive and enabling technology list for their topic. The ACE panel submits the following list of technologies with potential to meet the requirements for advanced command environments listed above:

- Advanced Displays, Visualizations and the Use of Special Effects,
- Advanced and Collaborative Networks, Collaboration Technologies and Network Protection,
- Advanced Computing and Communications,
- Advanced Knowledge Management; End-to-end Media and Info Management,
- Virtual Training,
- Advanced Modelling, Simulation, Gaming, Demonstration and Experimentation,

- Human Performance, Psychological Factors and Behaviours,
- Autonomous and Self-Organizing Networks,
- Analysis Tools and Technologies,
- COTS (Including Wireless) Advances and Vulnerability Mitigation,
- Energy Sources and Management,
- High Capacity Communications,
- Advanced Encryption, Public Key Infrastructure and Identity Concepts,
- Advanced Information Assurance,
- Advanced Collaboration Concepts,
- Smart Robots/Micromachines,
- Advanced Materials and Designs, and
- Compact Climate and Instrument Controls.

The panel believes that there is great value in joint experimentation between the two nations' militaries. The result of this initial activity should then be evaluated to determine if a long term relationship can be achieved.

CHAPTER 6. PERSISTENT SURVEILLANCE

INTRODUCTION

This report is the U.S. and UK contribution to a joint UK/U.S. study of Persistent Surveillance (PS) for defence applications. The two teams met both separately and together and, as a result followed both individual and complementary threads. The UK team focused on the underlying technologies and the programme view required to achieving PS, while the U.S. group considered the definitional, organisational and architectural/systems issues in depth. This reflects the different approaches of the United States and the United Kingdom, i.e. the United States has extensive investment in PS technologies but requires effort on how to better exploit the outputs from these. The UK MoD has recently initiated a major coordinated programme in PS (DABINETT), while requiring a clearer vision of which emerging technologies require defence specific investment. The U.S. DoD and intelligence communities have undertaken a series of activities to discuss and analyze the PS attributes and definitions. Although U.S. members have not investigated the full extent of U.S. DoD and intelligence community activities in this area, we are not aware of a formal U.S. programme defined for persistent surveillance.

Despite near-universal enthusiasm for next-generation "persistent surveillance assets" there is remarkably little informed discussion of what constitutes persistence and why we want it. We believe that increasingly targets of interest will require us to go beyond the current norm of point or sweeping type episodic surveillance. Enhancing the ability to track "lower signature" objects of interest (including persons and related entities) and better link them to specific threat activities will require a focused effort across the full spectrum of collection, analysis, mission and information management, and customer interaction. Persistence is a relative construct and a sensing system may be said to be sufficiently persistent if we can capture and recreate with sufficient fidelity the

relevant temporal characteristics of the target. Thus, persistent surveillance is most important if and only if, we are interested in some temporal characteristic(s) of the "target." Specifically, we are interested in analyzing the "activity" of the target. To appreciate the texture of an activity we must observe systematically, i.e. "sample" the target (entities as well as objects) at a rate dictated by the activity, not by the sensor, sensor system, or platform(s). The role of surveillance – in the troika of Intelligence, Surveillance and Reconnaissance (ISR) – is to deny the adversary any opportunity to act unobserved. Just as intelligence aspires to omniscience, surveillance aspires to be omni-present with virtual dwelling, staring, resolution, and other performance attributes that make such persistence more powerful and differentiated from current and past ISR. Current trends also require that such PS also be able to support a wide variety of weapons delivery scenarios integral to the conduct of such surveillance.

The context for PS is that adversaries increasingly are non-state players with many individual actors and entities, globally distributed with low signatures. Existing UK and U.S. collection capabilities are platform based. Increasingly as we align the information from different systems and phenomenologies, there will be significant implications for mission management, information management, processing, analysis, and archiving and network enablement. Interestingly, it is via these features of PS that backtracking of targets (or from events) (particularly those targets "hiding in plain sight") in order to facilitate retrack (or for forensic benefits) onto other like or related targets is facilitated. As we achieve improvements in PS, it is envisaged that new interfaces will be required that will leverage traditional defence, as well as homeland-security and law enforcement capabilities.

This PS study utilised the knowledge of subject experts plus solicited input from relevant experts when any capability gap was apparent. We must sense differently and more effectively and we must process and exploit that sensing with greater skill and with much improved results. This chapter summarises the key findings and the supporting evidence. The panel took the following approach:

twenty-nine defence critical technologies for PS were identified (not defined in this report, but available on request). These were categorised into Front-end Sensors, Back-end Capabilities and Interconnectivity Fabric. Three of these technology areas (identified below) were selected as "high priority." The panel assumed the following caveats: space-based sensors are addressed in existing collaborations and were therefore not considered in this study. However, the output from these sensors was considered and we believe that capable space sources are critical to our evolution of a comprehensive PS programme.

This chapter focuses less on the technologies of PS per se, and more on how to think about PS and for defining a way forward in PS. The technologies for PS are many and varied in the context of all domain legacy and "to-be," penetrating and standoff, front-end (tasking, mission control and management, sensors and platforms) verses back-ends (processing, exploitation, analysis[story finding], reporting [storytelling], and target development), non-weaponized and weaponized ISR, etc.

DISCUSSION

The subject of the panel (PS) is an emerging concept within U.S. and UK defence operations, and this is reflected in the numerous definitions available. The main approach adopted by the UK team was to divide the area into three categories i.e. Front-end Sensors, Back-end Capabilities, and the Interconnectivity Fabric. The aim was to identify a subset of key technologies from these three categories that would directly benefit from focused defence investment, architecture development, system-of-systems engineering of PS integration, and enhanced collaborative efforts between the United States and the United Kingdom (and other international, coalition or allied partners). There is also a broader set of issues related to enabling coalition PS, including security management (biasing toward information sharing while simultaneously protecting sources and methods) and integration into a network enabled information environment co-shared by both operational and intelligence actors.

A significant aspect of the shift towards persons, entity, and activity surveillance is the increased need for data forensics, modelling, and in-depth research based analysis of the subject behaviour. Improved analysis and modelling may lead to improvements in our ability to accurately predict activities. Such modelling and analysis needs to be taken into account as the United States and the United Kingdom define the future of PS architectures. Said another way, "*its not just what the sensor brings to you, it's also what you bring to the sensor.*"

The following sections first expand on the concept of PS and its application, followed by an outline of the analysis for each of the three key technologies identified in the summary.

Defining "Persistent Surveillance"

Today, there are many definitions of PS, most of which we have analyzed. For the purposes of this report, we have come up with our own definition, to wit:

> The systematic and integrated management of collection, processing, and customer collaboration for assured monitoring of all classes of threat entities, activities and environments in physical, aural or cyber space with sufficient frequency, accuracy, resolution, precision, spectral diversity, spatial extent, spatial and sensing diversity and other enhanced temporal and other performance attributes in order to obtain the desired adversary information, even in the presence of deception.

In this definition, there is a desired objective ability to observe with sufficient frequency and precision that the target will not be able to move, change, or function without notice.

Understanding and quantifying the temporal characteristics of an intelligence target requires persistence. Therefore, we are interested in persistent intelligence collection, because there are a wide variety of defence and intelligence problems for which the temporal dimensions are important. Among the relevant parameters of persistent systems are:

- Temporal parameters—e.g., sample rate, sampling duration, or "epoch" (time on target), and response time (time to target),
- Geo-Spatial parameters—e.g., area covered (field of regard), and area of coverage (field of view or, ground-sampled area) and overall resolution,
- Spectral parameters—e.g., frequency bands available, and sample frequency (spectral region of view), and

- Adaptability—e.g. actively monitors and adjusts the sensing and processing to insure the capabilities are optimized for the situation and entities.

Of pertinence to the summary recommendation is that persistent collection systems beg for automated processing—we recognize the impact of all the man-hours needed for staring at a target for long periods, with nothing happening most of the time – without machine enabled exploitation. Even if we were willing to spare the manpower, the "vigilance" literature guarantees that when the extremely rare event for which we are watching happens, we frequently miss it. Recognize that the humans need to be in the loop to drive the automated processing and react to the discovered information.

PS in Network Enabled Capabilities and Network Centric Warfare

From the dialogue between the U.S. and UK teams it was agreed that while U.S. Network Centric Warfare (NCW) and UK Network Enabled Capabilities (NEC) are different they converge in the area of the Common Operating Picture (COP) derived from PS. In addition, commitments for future demands on collaborative operations between the United Kingdom and the United States require a COP. It was thought by our joint review of PS that the NCW and NEC concepts and programmes will become better aligned, with the NCW programmes leaning more towards the acceptance of the UK view that PS can only be enabled by the web character of the networks, and future trends in establishing the various layers of the global grid.

It was agreed that *space sensors would not be considered as part of this study* for future collaborations since there are currently other Memorandums of Understanding and collaborations in this area. However, information and data that is collected from space sensors were considered in this study as part of the back-end capability and interconnectivity fabric considerations.

IDENTIFIED TECHNOLOGIES

The UK PS panel has identified their own list of technologies that it considered were defence critical technologies (not included in this report). These technologies were categorised into three broad areas. These were 1) Front-end Sensors; 2) Interconnectivity Fabric; and 3) Back-end Capabilities. Of these technologies, it was decided that three in particular should be proposed as a top priority for further MoD and DoD consideration. It should be noted that all three were analysed in the context of the Three-Block War Scenario as a grounding exercise. These three are detailed below.

Integrated Sensing – Physical and Computer

Definition: The processes and technologies for "persistently" monitoring the full spectrum of Information and Communication Technologies (ICT)/ISR activity within a selected area or within a range of selected networks and nodes of interest. Includes target development, collection management, and sensor application and management aspects of PS, as well as, PS connections to weapons, and weapons battle damage assessment/combat effectiveness assessment.

Why this technology is considered defence critical?

Unlike other technologies, adversaries will likely be aware of many of our ICT/ISR technologies and they might have the means to exploit these PS modes for an asymmetric advantage or subject them to denial and deception. There is a need to integrate across the sensing domain with traditional geographic-based sensors as part of the COP. As well as monitoring the enemy, sensing can be used to monitor friendly forces. It has specific application in the Three-Block War, which assumes an implicit asymmetric threat.

How will operations change using this technology?

Operations will indeed change within a PS enabled framework. In addition to the physical world improvements driven by PS,

cyberspace – both Signal Intelligence and Information Warfare –will be an important element of the PS framework.

Observation: To scope, conceive, design, and model a joint U.S.-UK study team.

Horizontal (Knowledge) Integration

Definition:To achieve a shared universal COP, comprehensive and timely situational awareness and integration of tasking, collection, analysis, and customer coordination is needed. All the attributes which were defined during the U.S. joint Under Secretary of Defense for Intelligence and Assistant Deputy of Central Intelligence (Collection) horizontal integration efforts are involved.

Why this technology is considered defence critical?

Ideally different users want their own smart pool of information that is extracted from all the data gathered from sensors. This is because different users require different interpretations of the data depending on their requirements.

How will operations change using this technology?

This technology provides greater detail and confidence, increased precision and can reduce fratricides.

Observation: The United States and the United Kingdom need to establish common standards and interoperability e.g. on sensor formats and meta-data. We need to move towards "service-oriented" warfare, and there is a need to link high level NEC/NCW activity. Both human and technical aspects must be considered to enable this. Establish goals to drive this activity.

Software Agents

Definition: Any software or tools which are part of making PS work as a differentiated form of ISR. Agents are part of a computing and IT environment, which is capable of flexible autonomous action

in a dynamic, unpredictable, and complex target environment. They enable information fusion, knowledge management, and back-end integration.

Why this technology is considered defence critical?

Software agents are the key elements for large scale PS. They are applicable at each level of PS.

- Apply to both the back and the front end,
- Need for collaboration on putting together the infrastructure that enables software agents to be exploited,
- Software agents need to be developed for military purposes to make them more:
  - Robust,
  - Fault-tolerant,
  - Address speed and tempo of defence arena, and
  - Scalable.

This technology is also cross-cutting across the other critical technology panels. For example, it can be applied to manage distributed processes in HPC, and intelligent management of resources in large-scale sensor networks, including power management. They have also already been adapted to support advanced command and control interfaces.

How will operations change using this technology?

There is a strong military advantage in bringing all the software agent streams together from across the Atlantic. A specific example for the application of software agents in persistent surveillance would be in long dwell monitoring including covert surveillance. Agents will increase the rate at which information can be matched and reduce the time taken to back-trawl through data. Efficiency in the human loop can also be improved.

Observation: Given that the United Kingdom has a PS programme (DABINETT), there is an opportunity for a joint UK-U.S. study for defence requirements. Activity in this area could be improved with increased emphasis. This effort could align activities and better leverage respective strengths. A joint U.S.-UK study would avoid any wasted effort and potentially save ourselves from running down the same dead end streets.

RECOMMENDATIONS

The UK and U.S. Persistent Surveillance panels therefore make the following specific recommendations.

1. Establish a U.S. Persistent Surveillance Effort, Office or Programme Counterpart to DABINETT

The U.S. Office Director National Intelligence (ODNI) and the Secretary of Defense should consider creating a programme or focused PS effort or office to define, build and integrate U.S. PS capabilities. The establishment of such a focused effort jointly at the level of the ODNI and DoD Under Secretary of Defense for Intelligence would provide a natural office/staff for working PS on a broadly based context and for transatlantic collaboration and coordination of U.S. and UK (as well as other) PS related programmes. This U.S. PS office would then be the natural interface point between the U.S. efforts and the UK DABINETT programme.

The United States should review the UK DABINETT as a potential framework for establishing its counterpart. The United Kingdom has put a considerable amount of sophisticated staff energy into framing their DABINETT focus on PS, and the United States should close with the MoD, subject to their concurrence, to use the UK thinking as a potential reference model.

2. Establish a High Level Joint UK/U.S. Coordination Group for Addressing Persistent Surveillance

The United States and United Kingdom should consider establishing a high level Joint Coordination Group on PS to work the TransAtlantic aspects of PS, including doctrine, concepts, sensors, architecture, analysis, processing, information management, interoperability, network enablement, customer support, security, export control, acquisition, systems engineering, etc. issues.

SUMMARY

The working party was asked to identify critical defence technologies. The Persistent Surveillance panel submits the following list of critical technologies in this area:

- Advanced Sensors and Sensor/Collection Management,
- New Concepts for Tagging, Tracking and Locating,
- High Performance and Quantum Computing,
- Advanced High Capacity Communications,
- Information Management and Analysis Technologies,
- Alignment of Surveillance with Weapons Applications,
- Stealth, Counter-Stealth and Signature Reduction/Management,
- Low Power Consumption,
- Counter Improvised Explosive Devices Attributes/Explosive Detection,
- Autonomous Self-Organising Networks,
- Advanced Materials, Engines, Energy Management, Modelling and Simulation,
- Space and Satellite Remote Sensing and Cost Reduction,
- Virtual Training,
- Data and Network Protection and Survivability,

- Nanotechnologies, Robotics and Micro-machines and related control,
- Antennas/Apertures,
- Data Fusion and Correlation; Networks; Software Agents,
- Processing and Exploitation; Geolocation/Navigation,
- Responsive Space, Air and Ground Systems and related Target Access Means,
- Hypersonics, EMP, HPM and IO Protection/Defences, chemical, biological, radiological and nuclear explosive/WMD Detectors,
- Pointing and Tracking,
- Data Warehousing and Storage,
- Collaboration Technologies,
- Many of the same technologies in ACE,
- Persistent Surveillance Total Systems Architecture,
- Development of a common data dictionary of sensor data formats and associated meta-data formats for full interoperability, and
- Context filters and compression.

CHAPTER 7. CONCLUSIONS

THE VALUE OF WORKING JOINTLY

This joint DSB/DSAB effort, in its own small way, echoes the value and power of U.S and U.K. collaboration that has existed over modern time. Our history together, given our respective world views, resource bases, assessment of mutual critical interests, application of policies, technologies and operations, and talents of our peoples proves again that there is great synergy and benefit to be derived from working together on important problems. Separated by an ocean and perhaps by a common language, this project, involving collaboration on five major and differing major defence technology areas, reaffirmed the benefits and importance of working together.

This study was ambitious in terms of its breadth and scope of topics. We picked both traditional and new topics, some of which had relationships to one another. We depended on small teams to carry the major burden of detailed problem definition, collaboration, analysis and the formation of conclusions. Working the five task areas by five different teams using five different work styles resulted in differing approaches, problems and even frustrations related to bridging the distance, maintaining momentum and in the mutual collaboration processes, aided by video teleconference, teleconference, electronic mail as well as some face to face meetings. The observations and lessons learned from these differing approaches are discussed below.

PROCESS

Our five topics of study included two broad, high level topics and three focused technology topics. As each of the panels proceeded, there was cross-fertilization on processes and procedures, but sometimes less on technical content. In the case of the three focused technology topics, the UK and U.S. panels operated independently – to a great degree.

The chairs concluded that we selected more parallel efforts than was optimal. The individual panels were small, and in some cases would have benefited by being larger. The value of multi-national collaboration comes from engagement of individuals who bring different perspectives on mission, on technology, on the industrial base available to a nation, and on the assumptions that flow from the scale of the technology efforts contemplated. Larger panels would have brought a richer perspective, and would have provided for richer technology detail.

Each panel grappled early on with finding common ground between the members from the two nations. And in the end it was only for the broad topic panels that common ground was established. In the three focused technology areas, the panels from the two nations did not meet face to face and agreed to pursue their topics independently, and then to integrate and harmonize their results at the end of the study. The process is reflected in the final written product.

One impediment for the U.S. members could have been the export control – International Traffic in Arms Regulation (ITAR). However, the lawyers at DoD provided a release from the constraints of technology information release to the members of this working party. Although, most of the U.S. members of the working party were, for the most part, from U.S. industry and still felt constrained by their previous corporate ITAR histories related to these specific subjects. Therefore, they did not make optimum use of forward leaning posture which the ITAR staff provided for this project.

In the two broader – closer to the mission – topics, the panels directly grappled with finding common ground. This was complicated by differences in approach to the mission by the two nations' militaries and intelligence communities. Different views of mission lead to different views on technology needs. One panel, the Advanced Command Environments panel, made mutual exchange visits. The entire panel held meetings in both countries, receiving briefings from military organizations, including from the U.S. Joint Forces Command and the U.K. NITEworks. Panel members

interactively engaged with briefers and with each other. Based on those meetings the Advanced Command Environments panel did find a common ground. This experience is discussed in the opening of their chapter. Once they had come to better understand each other's perspectives, they developed a joint product.

The U.S. members of the Persistent Surveillance panel visited the UK MoD headquarters and took briefings, but did not visit individual organization locations in the UK. Face to face meetings in which they could vigorously debate issues such as "how persistent surveillance differs from traditional surveillance" formed a basis for their consensus.

One reason why the face to face meetings are important is that the U.S. and UK styles of interaction are different. U.S. members were more animated and probing in style of interaction – brash, some may say – than the UK members were more reserved – more actively seeking consensus, some may say. But, when the individuals from the two nations had the opportunity to meet face to face, they forged a compatible view and the final product benefited from the face to face visits.

The working parties utilized teleconference, video-conferencing, virtual networked working areas and electronic mail exchange. The technology supported information sharing and general discussion, albeit awkwardly and unreliably when dealing with live images. However, the technology support available to the working party was not – by itself – effective in creating genuinely integrated teams. The tools that we used were not secure. Future DSB/DSAC collaboration, at least for some topics, may require secure interchange support tools.

There was another element that influenced the interaction of members from both countries. The DSAC traditionally focuses more on technology. The DSB takes the view that it must focus on the needs of clients within the DoD, their objectives and their policy and budget context. Therefore, a DSB working party routinely focuses on the objectives and constraints of the specific sponsors of the working party members of the DoD leadership who pose the issue being addressed. The DSB is concerned with the health of the industrial

base, lessons learned from military operations, policy constraints, the characteristics of future missions, and resource allocation. Technology, alone, is rarely the focus of a DSB study.

In a visit to the United States, the DSAC leadership found this wider scope interesting and one reason for the creation of this joint working party is that the DSAC leadership wanted to experience working on a problem using the approach of the DSB. So, one objective for this study was to expose DSAC and DSB members to each other's style in the context of grappling with real and specific problems.

The working party certainly succeeded in exposing panel members with two styles of operation, particularly the panel members that dealt with the two broad topics. It remains to be seen if the DSAC will adapt its operation in any way as a result of the experience with this working party.

RECOMMENDATIONS

Recommendations from the five specific areas of study are included at the end of each chapter and summarised in the table in the Executive Summary. The three co-chairs have two further recommendations relating to the benefits of continuing the relationship which has now been established between DSAC and DSB through future collaboration, and relating to common concerns on the state of the "pipeline" for future defence scientists and engineers in the United States and the United Kingdom.

Recommendation 1

Based on our positive experience, the three co-chairs of this working party recommend that the DSAC and DSB collaborate further on joint studies. The topic of each future study should be constrained, explored in depth and defined to be in areas in which there are different perspectives in the two nations along dimensions important to the chosen topic. Early working party meetings should

be in person and should involve briefings by appropriate organizations from both sides.

Recommendation 2

During working party discussions, plenary sessions, and teleconferences between the co-chairs, the issue of the shortage of U.S. and UK nationals opting to take undergraduate and higher degrees in science, engineering and technology was raised repeatedly. At a time when potential adversaries have access to high levels of technology (often at low cost), and when the rapidly growing economies of the east, such as China and India, are developing advanced and highly competitive industries in areas such as information and communication technologies and electronics, the United States and the United Kingdom are both experiencing a decline in the number of citizen graduates able to support our defence research programmes and defence-critical industries. We recommend that the MoD and DoD jointly consider ways to mitigate this problem and recruit more of our brightest young scientists and engineers into defence research. Collaborative programmes involving opportunities to train at universities in both the United States and the United Kingdom, and to work at U.S. and UK laboratories is one way to emphasise both the importance of, and the excitement and challenges offered by, working in this area in the 21st Century.

This page intentionally left blank.

APPENDIX A. TERMS OF REFERENCE

This page intentionally left blank.

April 21, 2004

MEMORANDUM FOR CHAIRPERSONS, U.S. DEFENSE SCIENCE BOARD
UK DEFENCE SCIENTIFIC ADVISORY COUNCIL

SUBJECT: Terms of Reference —Identifying and Sustaining U.S. Department of Defense/UK Ministry of Defence Defense Critical Technologies (Study)

Technology plays a vital role in the success of U.S. and UK armed forces. The development of useful technology arises from different sources: Government may itself develop unique technology for which there is no evident commercial use; Government may provide the initial impetus to enable industry to take over future development; or Government may choose to adapt technology developed solely in the market place. None of these sources, alone, is capable of satisfying the need for defense technology.

Governments can not afford to duplicate market-driven technology development, while the marketplace does not always develop technology which fulfills U.S. DoD or UK MoD niche needs or in the required time frame. For example, there is a requirement for a small number of radiation hardened integrated circuit chips for some missions but there is no commercial market to fulfill this need. High Performance Computing (HPC) represents another technology in which military needs are not fully met. As industry commits to massively parallel machines there exist several computing domains that are resistant to this technological approach. Quantum computing may hold the key to future military HPC needs but this unproven technology is still several years, if not decades away.

The Study will develop a methodology to identify unique defense technologies as well as commercially developed technologies needing augmentation to fulfill defense niche areas, and then apply the methodology to develop a list of defense critical technologies. The Study should focus its effort on high leverage, differentiated and transformational technologies. The Study may then use this list of defense critical technologies to further assess the tools available to the U.S. DoD or UK MoD to develop its critical technology needs. Some of the considerations the Study should examine include mechanisms to develop niches in pre-existing technologies, foster new technology until the commercial marketplace takes over, or develop technology without any expectation of commercial development; the analysis should include a review of the applicable acquisition/business case. Finally, the Study should consider the impact of technology development in other countries and the implications that this may have on Anglo-U.S. unique needs.

The Study will specifically address U.S. DoD and UK MoD technological needs in the following areas: power systems; HPC; materials, including energetic, structural and functional; advanced micro- and opto-electronics; communication systems; security and information assurance; vaccines and pharmaceuticals; and human factors. The

Study should assess relevant technologies and the means of transferring them to the defense arena using the above methodology.

This Study will operate under an exchange of letters. The Defense Science Board and Defence Scientific Advisory Council will work in parallel, comparing interim findings and working together to produce a unique UNCLASSIFIED report.

The UK part of the Study will be sponsored by the UK MoD Science and Technology Director and chaired by Dr Julia King. The Executive Secretary and UK Point of Contact will be Dr. Alexander Churchill.

The U.S. part of the Study will be co-sponsored by me as the acting Under Secretary of Defense (Acquisition, Technology and Logistics) and the Director, Defense Research and Engineering. Admiral William Studeman and Dr. Anita Jones will serve as Study Co-Chairpersons. Mr. John Grosh will serve as the Executive Secretary and Commander David Waugh will serve as the Defense Science Board Secretariat Representative.

Michael W. Wynne
Acting Under Secretary of Defense
(Acquisition, Technology and Logisitics)

Mike Markin
Science & Technology Director

APPENDIX B. WORKING PARTY U.S. MEMBERSHIP

CO-CHAIRS

Dr. Anita Jones	University of Virginia
ADM Bill Studeman, USN (Ret)	Private Consultant

WORKING PARTY MEMBERS

ADVANCED COMMAND ENVIRONMENTS PANEL	
Mr. Rocky Rocconova (Chair)	Northrop Grumman Technology Ventures
Mr. Chuck Benson	RTA, Inc.
Dr. Wayne Zachary	CHI Systems
Mr. David Jakubek	DoD Liaison

POWER MANAGEMENT FOR SMALL, DISTRIBUTED NETWORKED SENSORS PANEL	
Dr. Larry Dubois (Chair)	SRI International
Dr. Robert Nowak	Private Consultant
Dr. Brad Ringeisen	DoD Liaison

HIGH PERFORMANCE COMPUTING PANEL	
Mr. Steve Wallach (Chair)	Centerpoint Ventures
Professor Bill Dally	Stanford University
Dr. John Gilbert	University of California, Santa Barbara
Dr. Peter Kogge	University of Notre Dame
Dr. Bob Lucas	University of Southern California, Information Sciences Institute
Mr. John Grosh	DoD Liaison

ELECTRONIC COMPONENTS PANEL	
Dr. David Whelan (Chair)	The Boeing Company
Dr. Dennis Buss	Texas Instruments
Dr. Matt Ganz	HRL Laboratories, LLC

Mr. Zach Lemnios	MIT/LL
Dr. David Tennenhouse	Intel
Dr. Chuck Byvik	DoD Liaison

PERSISTENT SURVEILLANCE PANEL

Mr. Jeff Harris (Chair)	Lockheed Martin
Dr. Edward Gerry	The Boeing Company
Mr. Lee Hammarstrom	Pennsylvania State University
Mr. Rich Haver	Northrop Grumman
Mr. Leo Hazelwood	SAIC
Dr. Joe Markowitz	Private Consultant
Mr. Robert Gold	DoD Liaison

EXECUTIVE SECRETARY

Mr. John Grosh	ODUSD (S&T)

DSB REPRESENTATIVE

CDR Cliff Phillips, USN
LtCol Dave Robertson, USAF

STAFF

Dr. Evelyn Dahm	Strategic Analysis, Inc
Ms. Julie Evans	Strategic Analysis, Inc

APPENDIX C. BRIEFINGS RECEIVED BY THE PANELS

U.S. ELECTRONIC COMPONENTS PANEL	
Dr. Charles Holland and Mr. Sonny Maynard, OSD	Trusted Foundry Discussion
Ms. Joan Pierre, DTRA	RDD
Dr. John Zopler, DARPA MTO	DARPA Component Technology Briefing
Mr. Warren Snapp, Boeing	DARPA Design Technology Programs Address Emerging Challenges to DoD System on Chips
Dr. Matt Goodman	FPGA
Dr. Pete Rustan, NRO	Discussion of Critical Electronic Components
COL Tim Gibson, USA, DARPA, ATO	Multi-level Coalition Functionality
Mr. Rick Thompson, BAE Systems	Future Advances in High Performance Silicon Essential for DoD to Maintain Technological Superiority
Mr. Roger Van Art, Jazz Semiconductor	Commercial Pure Play Semiconductor Foundry Business Models Serving Aerospace & Defence
Mr. Joe Jensen, HRL	Mixed-signal electronics
Dr. David Chow, HRL	High-performance RF electronics
Mr. Randy Isaac, IBM	Leveraging Commercial Silicon Foundries for Government Needs
Dr. Paul Monticciolo, MIT/LL	High-Performance FPGAs in DoD Systems
Dr. Sonny Maynard	Discussion
Dr. George Valley, Aerospace Corporation	Photonic Analogue-to-Digital Converters: Fundamental and Practical Limits
Lt Col Chris Warwick, USAF, OUSD (IP)	Defence Industrial Base

Dr. Barry Gilbert, Mayo Clinic	Can the U.S. Military and agency communities execute their long-term mission without (on shore) access to advanced electronic technologies?
Dr. Eliot Cohen, EBCO Technology Advising, Inc.	Accelerating the Insertion Of Electronic Component Technology Into DoD Systems
Mr. Bob Walden, Phenomenon Consulting	Analogue-to-Digital Conversion in the Early 21st Century
U.S. POWER MANAGEMENT AND SMALL NETWORK PANEL	
Mr. Larry Schuette, NRL	Near Earth RF Propagation
Mr. Dave Watters, SRI International	Wireless Sensor Tags: Smart Passive Devices for Remote Monitoring
Dr. Robert Poor, Ember	Ember Architecture
Dr. Vjay Raghavan, DARPA	Ad Hoc Wireless Sensor Networks: Applications and Challenges
Dr. Clark Nguyen, DARPA/MTO	Microscale power generation, and Radio isotope micropower sources
Dr. Daniel Radack, DARPA/MTO	Thoughts on low power dissipation in electronics
Mr. Jim Smith, In-Q-Tel	Private Company Landscape of Power Sources
Mr. Tim McVey	FBI Perspective
Dr. William Kaiser, Dr. Mani Srivastava and Dr. Deborah Estrin, Center for Embedded Networked Sensing	Challenges in Energy-aware Embedded Networked Sensors
Dr. Paul Wright	MicroScale Systems supported entirely by Vibrational Energy Scavenging
Dr. Bob Brodersen, UC Berkeley	Low Power Architectures
Mr. George Methlie, In-Q-Tel	Power Sources for Distributed Systems
Dr. David Culler, UC Berkeley	Power issues in Wireless Sensor Nets
Dr. Jan Rabaey, Berkeley Wireless Research Center (BWRC);Gigascale Systems Research Center (GSRC); Department of EECS, University of California, Berkeley	Ultra low power wireless communications - status, challenges and opportunities

U.S. HPC PANEL	
Mr. Robert Graybill, DARPA	High Productivity Computing Systems Programme- Technology Update
Visit to National Security Agency	NSA/CES HPC Arsenal- Tom Page CA Needs for HPC- John McNamara HPC for CA and Beyond- Mike Merrill SPDs- Baron Mills CES Perspective and ECI Briefing- Dave Muzzy Superconductivity study- John Pinkston HPC Research and Vendor Interaction- Candace Culhane LUPS verses GUPS: What does CES need in a Supercomputer?- Boyd Livingston Emerging Areas for HPC at NSA- George Cotter Turmoil- Mike McGlynn/Jeff Fritz KSP- Steve Pritchard or John Walker
Visit to NRO	Organisational Overview HPC related to R&D
U.S. ADVANCED COMMAND ENVIRONMENTS PANEL	
Mr. Richard Lee, Office of the Deputy Under Secretary of Defense Advanced Systems & Concepts	Advanced Concept Technology Demonstrations (ACTDs) Showcasing Emerging Technology for Contributions to the GIG
Mr. Don Diggs Director, C2 Policy, OASD(NII)	The Unified Command Structure (UCS) A Net-Centric Approach to Information Integration & Decision and Global C2 Services
JOINT ACE PANEL MEETING	
JDCC	Briefings from Mr. Julian Starkey, Wg CDR T. Harris
Farnborough Dstl	Briefings from Dr. Graham Mathieson, Human Systems, Dstl
NITEworks	Tour of the Battlespace Management Environment (Alastair Prickett). Briefings from Wg Cdr Mike Oldham and Mr. Christopher Morris
JFCOM visit	Hosted by General Thomas Matthews and

	Mr. David Morris Command Overview/Transformation/Multinational briefing Joint Center for Operational Analysis Lessons Learned briefing Standing Joint Forces Headquarters Core Element Tour of Joint Advanced Training Technology Lab Combined Joint Task Force Headquarters of the Future U.S. JFCOM DARPA Integrated Battle Command Joint Intelligence Operations Center – Iraq Joint Futures Lab tour
Visit to VASCIC, Northrop Grumman	Hosted by Mr. W. Dennis Gallimore Demo of FLEXLAB- a full scale mock-up of the O3 deck (command space) of a Nimitz class carrier and if there is time a briefing from CVN 21, SPAWAR

APPENDIX D. RECENT U.S. HPC STUDIES

Four studies were initiated to analyze the state of HPC and make recommendations. They are briefly described below.

Information Science and Technology (ISAT) STUDY (2001) – Technology Gaps and Bottlenecks

A consequence of the industry focus on the desktop and commercial markets is missed technology opportunities and the lack of development of novel computer architectures. The ISAT study of 2001, eloquently describes this (picoseconds/instruction verses year).

In the ISAT study it was shown that the increases in computer performance experienced in the last 20 years (52%/year) will decrease to 19% per year in the future 20 years. This study states:

"Modern designs have nearly exhausted the benefits of pipelining, and conventional architectures are struggling to sustain even one instruction per cycle. Without further innovations, performance improvements will at best only match the rate of improvement due to further process technology innovations, which is projected to continue at 19% per year."

With an increase of only 19% a year (due to process technology innovations), the potential performance gain due to novel architectures (as in HPCS) is greatly enhanced. Again the study concludes:

"Until now, the differential between the 74% (When accounting for increased transistor counts and faster transistor switching speeds, the capability of microprocessor-scale integrated circuits has been improving at 74%/year) and 52% rates has resulted in only a factor of 30 of untapped performance potential. However, with only 19% per year projected in the future, the differential is expected to increase to a factor of 30,000 by 2020. This quantity represents a tremendous opportunity for novel architectures to help bridge the performance gap and to enable future computer systems to solve increasingly complex and important problems."

Integrated High-end Computing (IHEC) Study (2002) – High Performance Computing and National Security

The IHEC study summarised in the report: (Ref: "High Performance Computing for the National Security Community", July, 2002)

For the working groups involved with this report, the situation is clear. The mix of research, development, and engineering programs lack balance and coordination and is far below the critical mass required to sustain a robust technology/industrial base in high-end supercomputing. Requirements identified as critical by the national security user community (such as improved memory subsystem performance and more productive programming environments) will not be addressed.

The impact is that the national security community will be unable to solve critical computational problems required to maintain our technology lead for select but important classes of problem, examined in the course of the study, which included:

- Weapons Development Program
- Comprehensive Air Vehicle Design
- Army Future Combat Systems
- Stealthy Ship Design
- Nuclear Stockpile Stewardship
- Cryptanalysis
- Global Ocean Modelling and Operational Fleet Weather Forecasting
- Biological Sciences
- Intelligence Support
- Threat Systems M&S
- Signals & Image Proc
- Nuclear Effects
- Future Critical Problems
- Missile Defence

NRC STUDY (2004) - Getting Up To Speed, The Future of Supercomputing

The recently completed NRC study (Ref: "Getting Up To Speed, The Future of Supercomputing", NRC, November, 2004) made the following observations.

Conclusion: The supercomputing needs of the government will not be satisfied by systems developed to meet the demands of the broader commercial market. The government has the primary responsibility for creating and maintaining the supercomputing technology and suppliers that will meet its specialized needs.

Conclusion: Government must bear primary responsibility for maintaining the flow of resources that guarantees access to the custom systems it needs. While an appropriate strategy will leverage developments in the commercial computing marketplace, the government must routinely plan for developing what the commercial marketplace will not, and it must budget the necessary funds.

Conclusion: The government has lost opportunities for important advances in applications using supercomputing, in supercomputing technology, and in ensuring an adequate supply of supercomputing ecosystems in the future. Instability in long-term funding and uncertainly in politics has been the main contributors to this loss."

FINDINGS

Whilst technology is advancing all the time, the results of these studies remain valid and relevant.

The Defence Advanced Research Projects Agency (DARPA) has established the High Productivity Computing Systems (HPCS) program. This is the only significant US government-sponsored advanced development HPC activity at this time. HPCS is not a research program. HPCS was initiated in 2002 in response to concerns that COTS systems were not adequate for meeting some very critical aspects of the Defence mission. A primary focus of the HPCS program is on the "Last Dimension of HPC, User & System Efficiency & Productivity". The goal is to create a new generation of systems

that double in productivity (or value) every 18-months, rather than merely a doubling in unachieved, peak performance.

While primarily a DARPA program, HPCS has received significant support from other Defence agencies such as National Security Agency (NSA) and the National Reconnaissance Office, as well as the Department of the Energy, the National Aeronautics and Space Administration (NASA) and the NSF. While DARPA is leading, HPCS is truly a collaboration of all U.S. government agencies with a major stake in HPC.

All these studies, from the DSB (2000) Study to the recent NRC study have made similar findings and recommendations. The high performance needs of DoD mission agencies will not be satisfied by systems designed for the broader commercial marketplace. A long-term program funding the development of HPC systems is required to ensure that the DoD mission agencies can meet their requirements. This program must fund both near-term acquisitions, and long-term research; the existing DARPA activities need to be expanded accordingly.

High-End Computing Revitalization Task Force (HECRTF) Study (2004)

The HECRTF report makes these recommendations:

- **Make high-end computing readily available to federal agencies that need it to fulfil their missions.**
- ***The overarching conclusion of the HECRTF is that action to revitalize high-end computing in the U.S. is needed now.*** **The federal government's historical success in motivating HEC R&D, the oversubscription of current HEC resources, the scarcity of alternative architectures for delivering high performance to applications, and the lack of current incentives for industry to engage in HEC architecture research all argue strongly that the Federal government should move to revitalize HEC R&D.**

APPENDIX E. ADVANCED COMMAND ENVIRONMENT EXPERIMENTATION AND FACILITIES

U.S. ADVANCED COMMAND ENVIRONMENT EXPERIMENTATION AND FACILITIES

Combined Air Operations Centre-Experimental, Langley VA (Air Force Command and Control [C2] Intelligence Surveillance and Reconnaissance [ISR] Center)

The C2 ISR Centre was established to support experimentation with processes, procedures, and systems associated with the USAF Air and Space Operations Center. It was intended to facilitate the acquisition of fielded capabilities through a rapid spiral process, resulting in "leave behinds" for operations. Combined air operations centre experimentation is conducted under the supervision of the Air Force Experimentation Office using the facilities and infrastructure at Hurlburt Field, Air Force Base and Nellis Air Force Base. http://www.hanscom.af.mil/hansconian/Articles/2001Arts/06082001-5.htm

Unit of Action Maneuver Battle Lab, Ft Knox KY (Army Battlelabs)

This facility conducts experimentation, studies and analysis for the development of the Army's Future Combat System and Unit of Action. http://www.knox.army.mil/center/uambl/index.htm

Futures Lab Fort Leavenworth, KS (U.S. Army Battle Command Battle Lab)

Battle Command Battle Lab -Leavenworth's Futures Laboratory is the focal point for the experimentation and demonstration of emerging Battle Command capabilities in support of Army Transformation Objectives. Experimentation and demonstration events span echelons

from battalion through Army service component command. The Futures Lab is responsible for the maintenance, configuration, and set up of Models, Simulations, Communications and Networks to support the Battle Command Battle Lab –Leavenworth and TRADOC experimentation objectives.
http://www.defense.gov/news/Jun2001/n06212001_200106214.html

Integrated Command Environment (Naval Surface Warfare Center Dahlgren VA)

Integrated Command Environment developed a series of radically different options to influence command decision-making environments in future warships.
http://www.navyleague.org/sea_power/sep_01_12.php

Joint Expeditionary Force Experiment (U.S. Air Force Experiment)

Joint Expeditionary Force Experiment (JEFX) is a large-scale Air Force experiment designed to assist the U.S. Air Force in preparing for the challenges of the 21st Century Expeditionary Air and Space Force operations. The experiment is an operational innovation activity that attempts to anticipate and model a future command and control system. The JEFX Enterprise consists of Electronic Systems Command at Hanscom AFB, 505th Command and Control Wing at Hurlburt AFB, and the Air Force Command and Control & Intelligence, Surveillance and Reconnaissance Center and Air Combat Command/SC at Langley AFB. http://afeo.langley.af.mil/

Tidewater Node of the FORCEnet (Fn) Composeable Environment (FnCE) (Space and Naval Warfare Systems Command [SPAWAR] SC-CH Norfolk, VA)

A composeable Fn portal interconnects voice, video and data with the entire SPAWAR claimancy in a very user-friendly manner. The portal provides an advanced collaborative engineering environment for the SPAWAR Enterprise labs to deliver higher quality & more expeditious solutions to the warfighter. Design Objectives: provide

state-of-the-art information presentation systems & IT systems to replicate various system design options & to evaluate mission effectiveness/customer satisfaction. Designed with advanced human factor features to elicit optimum performance throughout the engineering process including requirements analysis, systems engineering, experimentation, acquisition & fielding. Provide an environment to expeditiously exchange systems information with key developers and fleet customers (S2C) so that the Navy can maintain information superiority to generate transformational combat effectiveness (Fn objective). https://mirage.norfolk.navy.mil/fnce/

Decision Architectures Research Environment (U.S. Army Research Laboratory)

This facility was created to facilitate the integration and testing of decision support technologies on Army Research Laboratory (ARL) platforms. The architecture consists of different layers to include communications, data transmission, and data presentation. These layers are united through communication protocols, enabling interplay of multiple components, and providing the "plug and play" capabilities. After integration, the technologies will become a part of ongoing ARL programmes where experiments with soldiers are conducted and the value of the technologies are assessed. http://www.arl.army.mil/main/researchopportunities/alliances/advanced_decision_architectures_2005a.cfm

MULTINATIONAL ADVANCED COMMAND ENVIRONMENT EXPERIMENTATION AND FACILITIES

Combined Federated Battle Laboratories Initiative

The Combined Federated Battle Laboratories (CFBL) Initiative links scientists and defence teams from the United Kingdom, United States, Canada, Australia, New Zealand and NATO. The CFBL network facilitates international research into command and control and information exchange between military forces - nationally or internationally. The output of this research can then be demonstrated in a representative battlespace environment for assessment by military

users. From the front line forces through the command chain to the Commander-in-Chief and across coalition boundaries, this research will support strategic, theatre, operational and tactical information requirements with the aim of enhancing command and control of national and coalition forces.
http://www.dstl.gov.uk/pr/press/pr2002/24-04-02.htm

Multinational Experimentation (North Atlantic Treaty Organization Allied Command Transformation)

North Atlantic Treaty Organization Allied Command Transformation conducts multinational experimentation examining command and control and coalition operations concepts.
http://www.act.nato.int/

International Technology Alliance (UK/U.S. collaboration)

A bilateral UK/U.S. collaboration opportunity is currently being established. It's called the International Technology Alliance (www.usukita.com) and draws on the UK Defence Technology Centres and the U.S. Collaborative Technology Alliance (run by the Army Research Laboratory). It involves government, academia and industry and currently consortiums are being assembled to bid for this.
http://www.usukita.com/program_info.html

APPENDIX F. ACRONYMS

ACE	Advanced Command Environments
A/D	Analogue-to-Digital
ADC	Analogue to Digital Converter
ARL	U.S. Army's Research Laboratory
ASSPs	Application Specific Signal Processor
ASICs	Application Specific Integrated Circuit
CFBL	Combined Federated Battle Laboratories
CM	Centimetre
CMOS	Complementary Metal-Oxide Semiconductor
CPU	Central Processing Unit
COP	Common Operating Picture
COTS	Commercial Off-the-Shelf
DARPA	Defense Advanced Research Projects Agency
DoD	Department of Defense
DRAM	Dynamic Random Access Memory
DSAC	Defence Science Advisory Council
DSB	Defense Science Board
DSPs	Digital Signal Processors
EAR	Export Administration Regulations
$E2EM^2$	End to End Media Management
Fn	FORCEnet
FPA	Focal Plane Array
FPGA	Field Programmable Gate Arrays
GaN	Gallium Nitride
GaAs	Gallium Arsenide
GB	Gigabytes
GPS	Global Positioning System
Gbps	Giga bits per second
GSM	Global System for Mobile Communications
HECRTF	High-End Computing Revitalization Task force
HgCdTe	Mercury-Cadmium-Telluride

HPC	High Performance Computing
HPCS	High Productivity Computing System
IC	Integrated Circuits
ICT	Information and Communications Technologies
IHEC	Integrated High-end Computing
IP	Intellectual Property
IR	Infrared
ISAT	Information Science and Technology
ISR	Intelligence, Surveillance and Reconnaissance
ITAR	International Traffic in Arms Regulation
JEFX	Joint Expeditionary Force Experiment
JFL	Joint Futures Laboratory
JFCOM	Joint Forces Command
KB	Kilobytes
KBPS	Thousands of Bits Per Second
MARCO	Microelectronics Advanced Research Corporation
MEMS	Micro Electro-Mechanical Systems
mA	MilliAmp
mJ	MilliJoule
mW	MilliWatt
MoD	Ministry of Defence
MOSIS	Metal-Oxide Semiconductor Implementation System
NCW	Network Centric Warfare
NEC	Network Enabled Capability
NITEworks	Network Integration Test and Experimentation Works
nJ	NanoJoule
NRC	National Research Council
NRO	National Reconnaissance Office
NSA	National Security Agency
ODNI	Office, Director National Intelligence
PGA	Programmable Gate Arrays
PS	Persistent Surveillance
RF	Radio Frequency
RFID	Radio Frequency Identification

SPAWAR	Space and Naval Warfare Systems Command
SRAM	Static Random Access Memory
STAR	Special Technology Area Review
TDMA	Time Division/Demand Multiple Access
µJ	MicroJoule
µW	MicroWatt
W/g	Watts/gram
Wh/g	Watt hours/gram