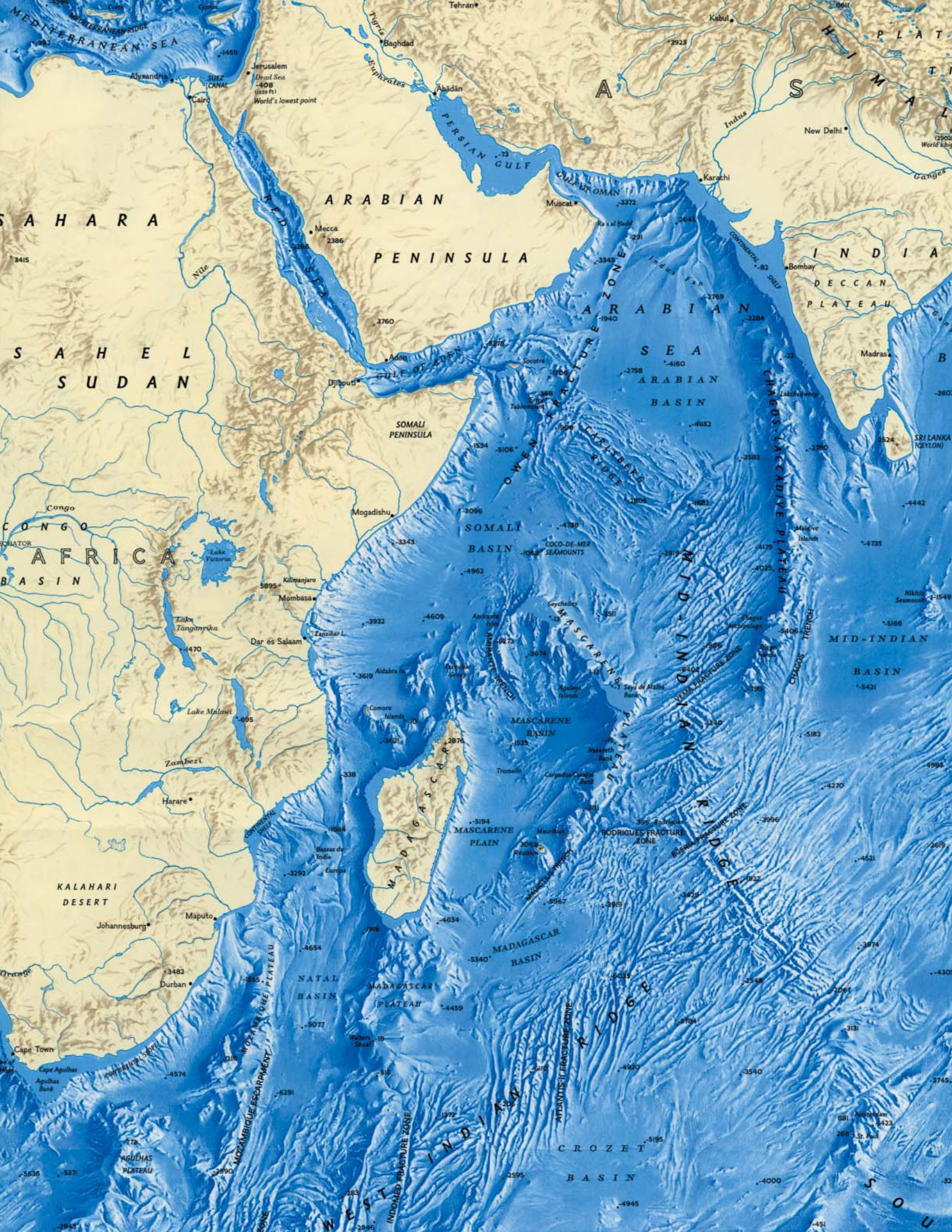
MIT Security Studies Program

The Future of Naval Aviation


MEDITERRANEAN SEA
MEDITERRANEAN RIDGE
Cyprus
Crete
Alexandria
SUEZ CANAL
Cairo
Jerusalem
Dead Sea
-408
(1339 ft)
World's lowest point
Tigris
Baghdad
Euphrates
Ahādān
PERSIAN GULF
Tehran
Kabul
Indus
New Delhi
Karachi
Ganges
Bombay
DECCAN PLATEAU
INDIA
ASIA
SAHARA
ARABIAN PENINSULA
RED SEA
Mecca
Nile
Muscat
GULF OF OMAN
Ra's al Hadd
CONTINENTAL SHELF
Indus Fan
ARABIAN SEA
ARABIAN BASIN
OWEN FRACTURE ZONE
CARLSBERG RIDGE
CHAGOS-LACCADIVE PLATEAU
Madras
SRI LANKA (CEYLON)
SAHEL
SUDAN
Adan
GULF OF ADEN
Djibouti
Socotra
Tablemount
SOMALI PENINSULA
Mogadishu
SOMALI BASIN
COCO-DE-MER SEAMOUNTS
Maldive Islands
Nikitin Seamount
Congo
CONGO BASIN
AFRICA
Lake Victoria
Kilimanjaro
Mombasa
Lake Tanganyika
Dar es Salaam
Zanzibar I.
Seychelles
MASCARENE RIDGE
MID-INDIAN BASIN
CHAGOS TRENCH
Chagos Archipelago
VEMA FRACTURE ZONE
Aldabra Is.
Comoro Islands
Lake Malawi
Zambezi
MASCARENE BASIN
Saya de Malha Bank
Nazareth Bank
Cargados Carajos Bank
Tromelin
Harare
MADAGASCAR
MASCARENE PLAIN
Mauritius
Réunion
RODRIGUES FRACTURE ZONE
Rodrigues
EGERIA FRACTURE ZONE
MAURITIUS TRENCH
Bassas da India
Europa
KALAHARI DESERT
Maputo
Johannesburg
MADAGASCAR BASIN
Orange
Durban
NATAL BASIN
MADAGASCAR PLATEAU
MOZAMBIQUE PLATEAU
Walters Shoal
Cape Town
Cape Agulhas
Agulhas Bank
CONTINENTAL SHELF
MOZAMBIQUE ESCARPMENT
AGULHAS PLATEAU
INDOMED FRACTURE ZONE
ATLANTIS II FRACTURE ZONE
WEST INDIAN RIDGE
CROZET BASIN
Amsterdam
St. Paul

The Future of Naval Aviation

Owen R. Cote Jr.

Associate Director

M.I.T. Security Studies Program

This report is a product of the M.I.T. Security Studies Program. It is the sixth in a series authored by Owen Cote and devoted to the subject of the U.S. Navy in the future security environment.

This series of reports is dedicated to the late Vice Admiral Levering Smith, U.S.N., the technical director of the Navy's fleet ballistic missile program during the development of Polaris, and from 1965 to 1977, its director.

The author would like to thank Admiral John Nathman, Vice Admiral Mark Fitzgerald, Rear Admiral Thomas Kilcline, and Malcolm Taylor for their support for this project, and Jeanne Abboud and Harlene Miller for their help in producing this report. It, along with other MIT/SSP reports, is available at http://web.mit.edu/ssp/

Contents

Executive Summary | 7

The Future Security Environment | 15

Sea Shield Past, Present, and Future | 35

The Revolution in Sea Strike | 55

Conclusion | 75

Notes | 78

When concentrated, multiple carriers are now able to produce effects ashore that would formerly have required many more wings of tactical aircraft operating from local bases in the region, while at the same time protecting the sea base and extending that shield ashore.

Executive Summary

TODAY, ALONGSIDE ITS ALL IMPORTANT OPERATIONS in direct support of the Global War on Terror, naval aviation also continues its now 60 year commitment to shaping the maritime and littoral environment through persistent forward presence. In the longer term, naval aviation is also adapting to a series of geopolitical revolutions which will dramatically increase the future demand for a secure sea base capable of projecting dominant power ashore in wartime against the full spectrum of possible opponents. It is adapting to these demands by exploiting technologies and operational practices developed in the last decade that will greatly increase its ability to surge and concentrate forces rapidly; protect the sea base from new air, surface, and undersea threats; and find, identify, locate, track, and strike mobile as well as fixed targets ashore, under all weather conditions, and in timely enough fashion to produce the desired effects.

Formal Alliances Provide Predictable Access, Informal Coalitions Do Not

The main source of new demands on naval aviation is the decline in secure and predictable access to overseas bases resulting from the shift in strategic focus away from the central front of Europe to the long Eurasian littoral, which extends from the Mediterranean to the Yellow Sea. Along this arc, the United States lacks and is unlikely to recreate the tight, long term alliance relationships which were a hallmark of the Cold War, from which flowed assured access to bases ashore. Instead, it faces a security environment in which ad hoc coalitions will form to solve specific problems; access to bases ashore will be episodic and unpredictable in advance of conflict, as was the case with Turkey in the run up to Iraqi Freedom; and the freedom to operate from such bases during conflicts can be suddenly withdrawn, as happened more recently in Uzbekistan.

Distributed Ground Forces Require Persistent, Distributed Air Support

Today's ground forces, operating in dispersed fashion, far from sustaining bases, in extremely austere environments, and in units ranging from section to brigade — rely on air forces for combat power more than they did during the Cold War. In particular, they depend on air forces for timely attacks against targets that emerge quickly and unpredictably in meeting engagements, and which must also be destroyed quickly. They also depend on air forces to detect, identify, and destroy larger concentrations of enemy forces moving to contact with elements of the distributed ground force. A sea base allows a supporting air force to operate in distributed and persistent fashion, while retaining the ability to concentrate quickly and bring dominant power to bear.

The Sea Shield Must Be Dominant If the Sea Base Is To Be Effective

In order for the sea base to play a central role in support of engaged forces ashore, it must be provided a dominant defensive shield, and must be able to project that shield ashore. A shield for the sea base preserves the access necessary for expeditionary maneuver warfare against both existing and emerging military threats. When anti-access threats are present, the sea base will need to be able to play offense and defense at the same time, and do so with absolute reliability, even when anti-access threats appear relatively low, and the temptation to trade sea shield capabilities for more sea strike will exist.

Adapting

Naval aviation is adapting to the demands of the new security environment across all its mission areas and against the full spectrum of threats. At the most aggregate level, naval aviation is developing a force structure and operating tempo which maximizes its contribution across the spectrum of conflict, whether measured in terms of time or threat level. In terms of time, this involves presence and shaping operations in peacetime, crisis response operations designed either to deter conflict or to maximize early arriving combat power should deterrence fail, and large scale surge operations whose purpose is the dominant application of sea-based power projection with an eye to the rapid and decisive defeat of the opponent. In terms of threat level, this involves the development and insertion of rapidly evolving technologies into sensors, weapons, and networks. The platforms which deploy these technologies will be capable of concentrating to dominate and defeat the high end threats, while also remaining prepared to operate multiple smaller force packages in dispersed fashion that retain robust shield and strike capabilities for use against more common mid- and low-level threats.

The Spectrum From Presence to Major Combat

For many years, the primary metric for assessing Naval Aviation capabilities has alternated between peacetime presence and wartime surge requirements. In the last decade, carrier forces have operated simultaneously along the entire spectrum from peacetime to major war. For example, since 1990, Naval Aviation has maintained a near continuous 1.0 presence in the Arabian Gulf, along with a forward-deployed carrier in Japan and a frequent presence in the Mediterranean. From this peacetime posture a forward deployed carrier has often been flexed in response to crises in places like the Arabian Gulf and the Balkans, trading presence in one theater for another. In other cases, pairs of carriers have been concentrated, as occurred off Taiwan for one month in 1996, and in the Indian Ocean for seven months in 1997-98. During both Operation Enduring Freedom (OEF) and Operation Iraqi freedom (OIF), as many as six carriers at one time were deployed to support combat operations, and as many as eight total carriers were employed over the course of the conflict.

Metrics based only on peacetime presence or wartime surge requirements fail to capture the complex requirements generated by a world in which presence, crisis response, and major surges are all necessary – sometimes simultaneously. This makes force planning more difficult, because the planner cannot focus on a single metric, such as needing five carriers to maintain one in the Indian Ocean, or twelve to be able to surge

eight. As it is more realistic to assume Naval Aviation will be asked to provide all three functions, as well as others not yet imagined, a carrier force that is able to sustain that more complex burden is needed.

The force that is providing presence or conducting the Global War on Terror today will often be the first to arrive in a crisis response or major combat scenario tomorrow. Combat power applied in the first days of a crisis response or major combat scenario is like the "golden hour" in combat medicine, because it produces relatively more deterrent or warfighting effect than a larger amount of combat power that arrives later. The Navy has historically avoided tailoring its deployed forces to the lower threat environment in which they often find themselves operating; instead, it deploys forces that are prepared for the full spectrum of combat. For example, when Iraq invaded Kuwait in August 1990, the Independence (CV-62) battle group was in the Indian Ocean, ready to respond, as was the Enterprise (CVN-65) battle group following the 11 September 2001 terrorist attacks on American soil.

Technology and the Spectrum of Threat

Naval Aviation must also ensure that its forces have the capabilities needed against the full spectrum of threats. This is a more technologically intensive process than force planning, and it has been the object of a decade-long Naval Aviation recapitalization strategy that is about to reach fruition. Advanced airborne early warning (AEW) aircraft; increased persistence and range for strike fighters; modernized airborne electronic attack (AEA) platforms; advanced surface, undersea, and mine warfare helicopters; and long range, persistent, land-based maritime patrol reconnaissance aircraft (MPRA) have been the major platform-related aspects of the carrier aviation modernization strategy.

The Value of Robust AEW

Any advanced, integrated air defense system relies on a fleet of airborne early warning (AEW) aircraft with powerful radars providing a persistent, high altitude view of the battlespace. Only upon this base can an outer area defense be erected that aspires to keep attackers well outside range of their targets, so they can be killed before they can launch their weapons. Without AEW, defenders cede enormous amounts of battlespace to their attackers, decrease warning time, eliminate multiple kill opportunities, and place primary reliance on difficult, close in, time critical engagements against arrows rather than archers. If one looks around the world on land, one can identify those few countries that actually seek to keep opponents out of their air space by their AEW aircraft. At sea, the presence or absence of carrier-based AEW is even more decisive. Airborne early warning also determines the ability of the sea base to project power ashore in the face of all but the most minimal anti-access threats. Thus, it is one of the main determinants of whether a sea base can play offense and defense at the same time.

No Substitute For Range in Carrier Aviation

The long range, sea-based strike fighter, with its ability to engage in multiple, simultaneous, and dispersed engagements 24 hours a day, is a key enabler of power projection ashore. Both OEF and OIF have proven the inestimable value of sea-based strike fighters as a distributed, timely source of fires over the battlefield, both against fleeting high

value targets, and in support of blue ground forces operating in dispersed fashion on non-linear battlefields. Adding range to today's relatively short-legged naval strike fighters through enhanced organic mission tanking and greater internal fuel capacity can expand the maneuver space of the sea base without compromising the reach or persistence of its main striking arm; increase the overland persistence and coverage of that striking arm from the same maneuver space; or some combination of the two. The greater the range extension, the more flexibility and capability result, allowing the commander at sea to maneuver and operate his sea base in such a way as to task-optimize his force along the full spectrum of conflict.

The Need for AEA Is Not Going Away

Airborne electronic attack (AEA) is a key enabler of strike operations by non-stealthy aircraft against even modest air defenses, and is central to any concept of air operations against more advanced air defenses. Although this is one of several possible future roles for UCAV, an AEA platform based on a long range strike fighter is necessary in the mid term. The range and persistence of AEA platforms must be equal or even better than the strike fighters they support.

Land-Based Maritime Patrol Aircraft

Long range, persistent, land-based maritime patrol aircraft provide the only way for a dominant naval power to maintain a continuous presence and surveillance throughout the vast ocean and littoral spaces over which it must exercise control. They often provide the most timely means of response, whether to a fleeting undersea acoustic contact, a report of a suspicious merchant ship, or an important signals intelligence collection opportunity. The maritime patrol fleet is evolving into a triad of more capable assets—the P-8A Multimission Maritime Aircraft, the Aerial Common Sensor, and the Broad Area Maritime Surveillance Unmanned Aerial Vehicle—that are an important element in Naval Aviation's recapitalization. Deployed in small expeditionary contingents at the strategic approaches to and operating locations along the Mediterranean-Indo-Pacific arc, the new maritime patrol triad will provide improved surface surveillance, antisubmarine warfare, antisurface warfare, and multisource intelligence collection.

Multimission Helicopters

Maritime patrol assets need to be complemented by a more distributed force capable of quickly responding to its cues; identifying potential contacts as friend, foe, or neutral; and either tracking the contacts or destroying them, depending on the circumstances. The specific sensors, and weapons used to acquire and prosecute contacts in surface warfare, undersea warfare, and organic mine warfare missions will vary, but because they will often be deployed and operated from the same platforms and in the same littoral battlespace, there is a tremendous premium on combining them on the same multimission helicopter when possible.

New Capabilities and Challenges

New sensors, networks, and weapons, along with new platforms, will introduce some radically new capabilities for naval aviation. For example, there are already emerging technologies that will enable through the weather attacks against mobile ground tar-

gets. There are also areas where the technology is less mature and where naval aviation faces significant challenges, such as in undersea warfare.

Eliminating the Weather Sanctuary for Mobile Targets

Radars suffer little or no interference from weather but, unlike in air-to-air combat, it has proven impossible to this point to use air-launched, radar-guided weapons against combat vehicles on the ground beyond ranges of five kilometers or so. That constraint has preserved a sanctuary from air attack for mobile targets when clouds lie between the attacker and the target, because other means of attacking mobile targets have historically relied on visual acquisition of the target. GPS-guided weapons solve this problem when the target's location is precisely known, but the vehicles which comprise an opposing ground force often move frequently and in unpredictable fashion, and combat aircraft do not currently have the ability to target GPS-guided weapons in real time.

In the near to mid term, strike fighters will become capable of targeting mobile targets temporarily at rest because the aircraft's own synthetic aperture radar processing will soon be capable of comparing an organically-generated SAR image with an onboard database of geo-registered SAR imagery automatically and in real time. Alternatively, toward the same end, it will soon be possible to use an organically generated SAR image as a scene-matching template for a weapon with an infrared terminal seeker, achieving the same goal of targeting through weather a mobile target temporarily at rest, while in this case, also reducing or eliminating reliance on GPS.

In the mid to longer term, one of several possible approaches to the all weather attack of targets that are actually moving will also be developed and deployed. One option will be to use bilateration or trilateration in a network of airborne radars to reduce the errors in azimuth intrinsic to single radars when they are tracking moving targets, and to use a data link to give a weapon in flight constant updates of the moving target's changing position. Another option will be to improve terminal seekers on weapons to the point where they can acquire, recognize, and home on a moving vehicle after the less accurate cueing provided by a single tracking radar.

Certainly, substantial limits will remain on what ground targets can be detected, identified, and attacked from the air, but with the weather sanctuary for combat vehicles greatly reduced, an opponent's freedom to concentrate and maneuver will also be greatly reduced, enabling the distributed, non-linear approach to battle that U.S. ground forces will increasingly adopt.

Providing a Dominant Defense of the Sea Base

In strike operations ashore, naval aviation will often be part of a joint team, but when defending the sea base from attack, the Navy will often be on its own. This is not a problem in much of the world because most countries have no capability to project power out to sea. Of those countries of concern that pose a threat to the sea base, many do so only with ground or small boat-launched, anti-ship missiles and mines, limiting both their reach and their effectiveness. Some more capable potential adversaries add missiles launched by aircraft or major surface combatants and non-nuclear submarines deploying mines, torpedoes, and, in a few cases, antiship missiles. At the highest end of the spectrum, the sea base may face an opponent with over-the-horizon sea surveillance

capabilities extending as far as 1000 miles for cueing antiship attacks by long range, terminally-guided, ballistic missiles and/or non-nuclear submarines with air independent propulsion (AIP).

Though formidable compared to the norm in today's world, these potential anti-access threats do not equal those the Navy faced down during the Cold War. On the other hand, because the Navy will be asked to play a much larger power projection role relative to land-based forces than was the case during the Cold War, the Navy's sea shield capabilities must dominate the high end anti-access threat. The path to dominance in this mission area varies between the air, surface, and sub-surface environment, and the consequences for naval aviation vary as well.

Shoot Archers Not Arrows

On the surface and in the air, surface combatants and aircraft will be detected long before they can target and launch their missiles against the sea base because of powerful, fully networked surveillance radars such as Advanced Hawkeye and the Broad Area Maritime Surveillance (BAMS) system. Even if missiles are successfully launched, strike fighters with AESA radars, as well as extended range Standard missiles, will give the sea base a greatly extended battlespace in which to engage them, allowing multiple intercept opportunities against each incoming missile. Technology will improve this picture further when it allows better long range combat identification and, in the event some missiles leak through, improved passive ships' self defense systems alongside the active systems already planned.

Make Opposing Submarines Pay For Their Inevitable Indiscretions

In the undersea environment the challenges are different. Here, sensor performance is limited, reducing detection ranges, and making wide area surveillance a more asset-intensive endeavor. Furthermore, unlike nuclear submarines, which usually produce a continuous acoustic signature, the best detection opportunities against non-nuclear submarines are both episodic and difficult to classify. On the other hand, there is a close correlation between the steps a submarine needs to take in order get into position to attack a target, and the operational indiscretions which provide the best detection opportunities for ASW forces. Therefore, contacts must be prosecuted and reliably classified as quickly as possible before they disappear back into the cluttered background as an unknown contact. This puts a premium on ASW platforms that can be deployed in numbers and distributed throughout the sea base, close a potential contact quickly, and deploy a menu of high quality acoustic and non-acoustic sensors to reacquire and identify the contact, classifying it as a false alarm, or trailing and/or attacking it.

Get Back In the Counter-Surveillance Business

Finally, in the longer term, against more formidable anti-access threats, the Navy will need to get back into the business of denying its opponent a reliable ocean surveillance capability. The worst potential threats to the sea base emanate from missile attacks launched from outside the sea base's defenses, but they therefore also demand that the weapon be launched from well beyond line-

of-sight to its target. This in turn requires effective over-the-horizon surveillance, classification, and tracking from other sources, and that these sources also provide cueing to weapons timely and accurate enough for their terminal seekers to reacquire and lock onto the correct target and complete the engagement. Alongside the added reach of such a centralized, networked, anti-access system come potential new vulnerabilities at each step of the engagement sequence, as the Soviet Navy learned during the Cold War.

The Force of the Future

The United States has long understood the value of single, full spectrum carrier battle groups forward deployed in peacetime. It has also understood the value of concentration in wartime when necessary; witness the multi-carrier operations practiced off the coast of North Vietnam and those envisaged by the Cold War maritime strategy. But it has been many years since concentrations of carrier aviation were in a position to be so unambiguously dominant in the power projection role, both in terms of its freedom to maneuver in the face of opposing defenses, and in terms of its ability to produce decisive effects ashore. If one combines the effects of the precision weapons revolution (multiple kills per sortie) with the high sortie rates possible on a large deck carrier, and then concentrates up to five or six of those carriers in one theater of operations, one is deploying a combat force analogous in capability to the Fast Carrier Task Forces employed to devastating effect in the last two years of the war in the Pacific. Such a force, operating in dispersed or concentrated fashion, will be the key to meeting the greater demands on naval aviation in the new security environment described above.

For example, a force of six carriers includes some 300 strike fighters, 30 E-2s, and 100 multimission helicopters, all operating in mutual support of both each other and expeditionary forces ashore. As President Bush noted at Annapolis recently, since Desert Storm, the number of targets ashore that a single carrier is capable of destroying in a single day has tripled. This means that a forward deployed carrier's capability to influence events during the "golden hours" early in a contingency have been greatly enhanced. It also means that when concentrated, multiple carriers are now able to produce effects ashore that would formerly have required many more wings of tactical aircraft operating from local bases in the region, while at the same time protecting the sea base and extending that shield ashore.

The relative value of Naval Aviation as a guarantor of national security and an instrument of national will has never been greater than in today's post-9/11 strategic landscape. From the Pacific campaigns of World War II through the complete battlespace dominance demonstrated during Operations Enduring Freedom and Iraqi Freedom, Naval Aviation continues to evolve as a decisive force through ongoing incorporation of advanced technology. The necessary investments must be made to sustain this preeminent force to ensure the United States' continued access and ability to shape world events.

Persistent surveillance, whether manned or unmanned, land or sea-based, is the foundation for success in all mission areas in the new security environment.

The Future Security Environment

THE LONG TERM security environment is obviously less predictable than the near term, and it is therefore more difficult to make specific assumptions about its likely characteristics. But assumptions can be made and trends identified in several areas. First, it is unlikely that the United States will have to make a major continental commitment in order to preserve a balance of power in Eurasia; second, it will have unpredictable access to bases along the Indo-Pacific littoral where it will most likely need to project power; third, it could face opponents with capabilities ranging across the full spectrum of conflict; fourth, new operational demands will result in more sea basing and a more distributed and persistent tactical air force; and fifth, technology trends will increasingly favor airborne as well as spaceborne sensor platforms, networking, persistence, decentralized execution, and robust, line-of-sight data links.

Major Continental Commitments on the Eurasian Land Mass Will Not Be Necessary

The key military aspect of the Cold War was the fact that the United States made a continental commitment to Western Europe's security. The grand strategic purpose of this commitment was to contain Soviet power, preventing it from uniting the resources of the Eurasian landmass. The military expression of this commitment was a large, continuing peacetime commitment of U.S. ground forces. It is extremely unlikely that any future threat to the Eurasian balance of power will require a new continental commitment. Rather, should there be such a challenge, it will likely arise in a maritime context. This is because nuclear weapons and political geography conspire to make significant landward expansion by Russia, China, India, or the European Union at the expense of each other all but unthinkable.

During the Cold War, Germany's division and its non-nuclear status made it unable to defend itself against Soviet attack alone. The unification of Germany and the collapse of the Soviet Union made Germany and Russia much more equal in basic power potential, and also established a number of medium-size buffer states between them. Thus, a unified Germany would be much less disadvantaged in a conventional military competition with Russia. And Germany's continuing non-nuclear status could evolve in three directions in the future, none of which would make major war between Germany and Russia likely.

First, current tensions aside, U.S. nuclear guarantees to Germany in the context of NATO might simply continue. Second, the further intensification of European unification might substitute for these guarantees via a European-based alternative. Third, and perhaps least likely, Germany might eventually develop nuclear weapons of its own, which would undoubtedly create tensions, but which would in some ways deepen rather than weaken the barriers to major landward expansionism, at least by one nuclear power at the expense of another.

The land border separating Russia and China has also acquired buffer states such as

Kazakhstan and Mongolia, so that the two larger countries abut only in China's upper Xinjiang province and more extensively along the border between Manchuria and the Russian maritime provinces. Both of these borders could become future sources of instability, but these instabilities should be constrained both by the fact that China and Russia are likely to remain major nuclear powers, and by the fact that the vulnerabilities along their land borders should tend to cancel each other out. That is, China is vulnerable to separatism in Xinjiang province, which is near the base of Russian land power, and Russia is vulnerable to separatism in its maritime provinces, which are near the base of Chinese land power.

Finally, India is likely to maintain rough parity between its nuclear forces and China's, and furthermore, the political geography of the subcontinent will continue to provide a powerful buffer against invasion along the entire Indo-Chinese land border. Central Asia and the Indo-Pakistani subcontinent are likely to be enormous sources of instability, but geography and nuclear weapons make it unlikely that that instability will provoke a major ground war between India and China, and absent such a war, events on the ground in this region are unlikely to cause major shifts in the overall Eurasian balance of power.

The most likely venue of great power competition and even war will, instead, have a more maritime focus. China and Japan is one obvious potential conflict dyad, and China and India is another. A triangular competition among all three powers over control of the energy flows from the Middle East and Central Asia is also possible. The medium powers that sit astride the key sea routes, such as Singapore, Malaysia, Indonesia, the Philippines, Korea, and of course, Taiwan, will all have stakes in the outcome of such a competition, and will all face competing pressures to balance or bandwagon against different perceived threats to their own interests. And unlike the newly independent medium powers in Central Europe, these powers are not part of any strong, transnational security organizations like NATO or the EU.

U.S. Alliance Relationships and Access to Overseas Bases Will Be Less Formal and More Unpredictable Than Those That Obtained During the Cold War

The main Cold War alliance relationships between the United States and NATO and Japan benefited from a basic agreement among the parties to each alliance on the threats that justified it, the tools needed to oppose those threats, and the essential equality of national interests and thermonuclear risks at stake for all its members. Although the United States dominated each alliance, it also committed itself to the most binding of security guarantees: the promise to use U.S. nuclear weapons, if necessary, to defend allied territory from attack, whether conventional or nuclear. In return for this commitment, U.S. allies granted unprecedented access to bases within their territory and allowed the United States to station hundreds of thousands of troops. The rights of access and operational activity granted by each host nation were codified in formal status-of-forces agreements and were therefore predictable and reliable enough to be assumed as a given in Cold War military planning.

Both alliances were a response to the Soviet threat, and both continue after its demise, but neither any longer provides the United States assured access to local bases near or along the long littoral from the Mediterranean to the Sea of Japan. There, a

better model for the alliance relationships that will provide such access, when it is granted, is the U.S.-Saudi relationship.

Originally formed early in the Cold War, the relationship grew in importance to both the United States and Saudi Arabia after the fall of the Shah appeared to eliminate Iran as a buffer between the Soviet Union and Persian Gulf oil. Yet the United States gained only limited access to Saudi bases before 1990 in support of its Rapid Deployment Force (RDF), mostly in the form of port visits and pre-positioning of ammunition and other supplies. Iraq's invasion of Kuwait resulted in a decision by the Saudi monarchy to allow U.S. forces unlimited access, but that decision was not made until four days after the invasion began, when Iraqi forces were already poised on the Saudi border.[i]

After the war, the Saudis allowed U.S. combat aircraft to remain deployed, but refused U.S. requests to pre-position a brigade set of heavy armor.[ii] During the decade or so of Southern Watch, those deployed air forces were put under strict operational restrictions, including a ban on flying any strike sorties from Saudi territory, even during crises such as Operation Desert Fox in December 1998. [iii]

Saudi reticence about granting the U.S. unrestricted access to its bases continued after 9/11. Though many support missions were flown out of Saudi bases during both Operations Enduring and Iraqi Freedom, few if any combat missions were. And of course, soon after the assumed completion of combat operations in Iraq, U.S. forces left Saudi Arabia altogether.

Many factors explained this Saudi schizophrenia about U.S. forces, even during the height of their cooperation. The Saudi regime is a Sunni feudal monarchy that sits across a narrow sea from Iran, a Shia fundamentalist theocracy; it is an Arab state that enjoys good relations with Israel's largest supporter; it is wealthy state with a small population that abuts several poorer states with large and growing populations. The United States could solve only some of the Saudis' security problems, while at the same time exacerbating others, and it was always difficult for the Saudi monarchy to determine the balance between these two effects of their military cooperation with the U.S. For example, there is no question that the Saudi regime's greatest domestic threat comes from fundamentalist Islamists, and the U.S. military presence certainly served as a rationalization, if not a cause, for claims by Bin Laden and others that the Saudi regime was failing in its sacred role of protecting the holy cities of Mecca and Medinah from the infidel.

Even before 9/11 and the two wars that have followed, both the 1997 *Report of the National Defense Panel* and the more recent Hart-Rudman Commission report *New World Coming* discussed why many of the uncertainties that characterized U.S.-Saudi military cooperation during the period described above will be endemic in the new security environment. For example, the latter noted that:

> In dealing with security crises, the 21st century will be characterized more by episodic "posses of the willing" than the traditional World War II–style alliance systems. The United States will increasingly find itself wishing to form coalitions but increasingly unable to find partners willing and able to carry out combined military operations.[iv]

When the alliances that produce base access are episodic and temporary, the access they produce will be as well.

Finally and perhaps most importantly, those who must consider granting access to U.S. forces in the future will do so without

the prospect of receiving the security guarantees against both nuclear and conventional attack that the United States gave its important Cold War allies. This will make it harder for them to determine whether giving U.S. forces access will increase or decrease their long-term security. For example, as the National Defense Panel argued, this might lead to limits on access for U.S. forces when potential allies face regional rivals armed with weapons of mass destruction.[v] Limits on access are also likely when potential regional allies face serious conventional threats.

This is not to argue that U.S. forces will gain no access to bases abroad. When faced with clear threats to their sovereignty, many states will ask for help, and when it is in the interests of the United States to respond, its forces will be given access. But this access will often come late, after a conflict has already begun; it will often be austere, in that few preparations will have been made in advance; and it will often be withdrawn or sharply limited after the particular conflict that generated it is resolved.

Future Military Opponents

One can imagine three types of military opponents for the United States in the new security environment; peer competitors in competition for energy resources with the U.S.; medium powers that seek or provide nuclear technology: and terrorist groups with global reach and their state sponsors.

Great Power Opponents

For lack of a plausible alternative, many focus on China as a future peer competitor of the United States. Certainly, it is possible to argue that China will be wealthy and technically advanced enough to play this role within decades, if not sooner. Any country able to sustain a peacetime defense budget of $150 billion a year or more, possessing a technical/military/industrial complex capable of designing, building, and operating modern, information technology-laden military systems, could cause a reprise of some aspects of the Cold War military competition between the U.S. and the Soviet Union. China does appear on a path that would put those capabilities within its reach, and there are certainly aspects of China's geopolitical position that might cause it to embrace that path.

The main point is that any future conflict with a peer competitor is likely to take the form of a struggle for control over the Indo-Pacific littoral because of the desire of some power to control and assure access to the seas and littoral chokepoints that will continue to link the world's primary users of energy to its main supply. Given that the United States Navy now provides that service to the world as a free good, it is ironic that the rise of a peer competitor would therefore be as likely to result from a retrenchment of U.S. maritime power in the Indo-Pacific region as from an overexpansion of that power.

In general, when maritime hegemony is used to assure rather than suppress free trade, it takes on the character of the trade that it assures; an exchange in which all achieve absolute gains larger than they would in the absence of such trade, but in which the largest traders may achieve larger relative gains than others. Thus, maritime hegemony, when it is used to assure free trade, is rarely itself the cause of great power conflict, because states generally ignore the balance of relative gains from trade unless they are already in a military competition with their potential trading partners. This is a fundamental difference between naval hegemony and military hegemony, defined as the effort to develop supremacy of land power relative

to one's neighbors, because such supremacy creates the threat of landward expansion. The threat of landward expansion by definition involves a threat to the basic sovereignty of the potential victim, and is therefore much more likely to provoke intense balancing behavior than is naval hegemony.

A more likely source of conflict between the U.S. and a peer competitor would arise if China and India became embroiled in a struggle for energy security that also involved the medium powers in the Indo-Pacific region. One potential flashpoint for such conflicts will be the unresolved status of island chains such as the Spratleys, sovereignty over which may provide access to considerable reserves of offshore oil and/or natural gas.

The goal of the United States would be to play the role of the balancer of last resort in these competitions, and the power that will determine the balance in these competitions will be seaborne. This will put a premium on forces that can independently survive in and gain control over contested sea and littoral battle spaces against all comers, and when necessary, can project power rapidly ashore. The requirements for power projection ashore will stop short of an independent ability to wrest control of significant land areas from another great power, and will be focused instead on the ability to deploy air forces and, when necessary, ground forces rapidly as an equalizer in land conflicts between medium powers and larger powers.

Thus, the United States should plan on maintaining its global naval hegemony without allied assistance, but it should only plan on fighting other great powers on land with the assistance of another medium power. In both cases, in the event that a peer competitor emerges, the battlefields of the longer-term security environment will be much more lethal because the asymmetry in wealth and technological prowess that favors the United States today will be gone or significantly reduced.

Medium Power Opponents

Alongside any great power competitions that might arise, or in their absence, there will be medium powers that aspire to regional expansion, seek nuclear weapons or provide them to others, and/or support terrorist groups with global reach. This type of conflict has been ubiquitous in the immediate post–Cold War era, and were today's "unipolar" moment to last forever, it would probably be the only type of conflict for which the U.S. military needed to prepare. Iraq, Iran, and North Korea all fall into this category.

The nature of the wars that the U.S. might fight against such powers vary widely. Wars could be fought to prevent or reverse territorial expansion by such powers, to stop the development of nuclear capabilities, or to eliminate support and/or sanctuary for terrorist groups. Wars to prevent or reverse territorial expansion would most clearly correspond to the major regional contingencies that dominated DOD force planning between the end of the Cold War and 9/11. Wars to stop a nascent nuclear program would start with relatively limited strikes against fixed targets, but would primarily be waged to deal with the target country's retaliatory response. For example, the primary military challenge of dealing with the current nuclear programs in Iran and North Korea is not to mount the attacks against their fissile material production facilities, but to deal with the potential military response to those attacks, whether that involves protecting shipping and oil industry infrastructure in the Persian Gulf from Iranian attack, or Seoul and its environs from North Korea attack. Wars to prevent support for terrorist groups will require successful regime

change, involving not only invasion but also the establishment of a new government, and these wars, should the U.S. choose to wage them, will most closely resemble the one the U.S. is engaged in today in Iraq.

The Global War on Terror

Operation Enduring Freedom sought not only to attack Al Qaeda, but also to eliminate the regime that provided it a base in Afghanistan and to prevent Afghanistan from being used for that purpose again. As such, Enduring Freedom probably represents the high end of the kind of wars that might be fought against Al Qaeda or groups like it. This has led some to note the paradox that the war on terror makes the U.S. as concerned about weak, failed states, whose very weakness creates a threat because it makes them candidates for use as a terrorist base, as it has historically been concerned with states with rising power.

Whether the war on terror includes another war like Enduring Freedom, it will almost certainly include operations like those being conducted today in Djibouti, where air, ground, and naval forces operating from an austere base conduct small scale, often covert anti-terror operations within the region broadly defined by the Horn of Africa.

Another military model for the future war on terror is represented by the Proliferation Security Initiative (PSI), which is a grouping of states willing in advance to allow their commercial ships to be boarded by other parties to the agreement in case there are suspicions of the transfer of illicit, WMD-related cargos. The PSI is a specific example of a much more general trend, which is that homeland defense in an age of terror is a question of maritime security as much as any other factor.

Operational Demands on U.S. Naval Aviation

Three operational demands on naval aviation will dominate the future security environment. The first two will be less dependent on the nature of the adversary, whereas the third, which is more of a tradeoff than a demand, will depend heavily on the nature of the opponent. The first is that U.S. air forces will increasingly need to deploy and sustain themselves from the sea, rather than from local bases on land. This demand derives not only from the changed nature of alliance relationships described above, but also from the revolution in vulnerability to fixed targets that is being caused by GPS. U.S. air forces are already exploiting this revolution to "solve" the fixed target problem, but its full consequences will only arrive when opposing forces start to solve the fixed target problem as well, which they inevitably will. Once this process starts, the key to successful projection of air power will be to avoid dependence on fixed bases near the opponent, and instead to deploy and operate from a secure, mobile sea base.

The second new demand will be that U.S. air forces adapt to the needs of U.S. ground forces on future battlefields. On those battlefields, U.S. ground forces will deploy with smaller, distributed force packages that, operating independently from each other, will use high speed and a flexible scheme of maneuver to quickly reach and threaten key objectives. In so doing, they will advance on external lines, both exposing their flanks and leaving pockets of opposing forces in their rear. U.S. air forces will have the responsibility of protecting the flanks and rear areas of these advancing forces from efforts by enemy ground forces to concentrate and maneuver in response.

To accomplish this task, U.S. air forces will need to be in many different places at

the same time, deploying networks of sensors and weapons able to quickly find, identify, track, locate, attack, and assess the damage of attacks against mobile targets when they emerge. To do this, U.S. air forces will need to supplement today's model of one large, centrally managed, theater wide network with a model that supports many distributed, self-contained networks, each able to complete the kill chain autonomously within its area of operation. In some low intensity cases, such networks will form around Expeditionary Strike Groups operating alone, but in any case where there are military threats the sea base will be formed around an Expeditionary Strike Force that includes a Carrier Strike Group or Groups.

The Shift in Emphasis From Local Land Bases to Sea Bases

Both political and technological trends limit the military's access to local bases in regions where it will most likely need to project power. Formal alliances dating from the Cold War, and the assured access to regional bases that they provided, have declined in relative importance compared to more informal coalitions formed after a conflict has already begun. In these coalitions, U.S. forces must generally negotiate access to regional bases on the fly, and the access that results is therefore less predictable in advance of a conflict than in the past.

At the same time, technological trends are making fixed targets more and more vulnerable to attack. In particular, when potential opponents follow the United States in marrying precision guidance technologies to cruise and ballistic missiles they will be able to hold fixed bases within a radius of up to a 1000 miles at risk of conventional attack by cruise and ballistic missiles against which defense is extremely difficult.

Political Constraints on Access to Bases Ashore. The political constraints on access to foreign bases affect land and sea-based forces differently. The main difference is not that one mode needs overseas bases while the other does not, but that land bases launch weapons while naval bases do not. This distinction makes naval bases less threatening to the host nation, not just because it makes those bases less of a target, though it may certainly have that effect as well, but because it separates the host nation politically from the actions taken by the U.S. naval forces that use those bases.

This is a function of the range and endurance of naval platforms, which go to sea for months at a time, and often operate literally around the world from their bases. Certainly, when ships go to sea, they are supported by an extensive train of supporting vessels, which replenish supplies of fuel oil, dry cargo, and, in the midst of a conflict, ammunition. But this umbilical connecting deployed navy combatants to the shore is largely indistinguishable from normal commercial activity, and is in fact often conducted by civilian-crewed ships that are essentially indistinguishable from their purely commercial counterparts. Thus, for example, when the U.S. Navy first began regular Indian Ocean battle group deployments in the late 1970s after the fall of the Shah, its oiler and stores ships were able to obtain needed supplies from a variety of countries along the Indian Ocean littoral, none of which were willing to provide any level of base access ashore to land-based American forces.

This distinction is reflected in actual legal arrangements. For example, everywhere the United States has access to overseas bases for land-based forces it has an accompanying set of agreements with the host nation about how those forces will or will not be used. Thus, in a recent example, after the Taiwan

straits crisis of 1997, the United States apparently sought to negotiate with Japan over future guarantees that bases on its territory which were not available in 1997, would be in the future under similar circumstances. No such negotiations were necessary regarding the use of Yokosuka and other U.S. Navy facilities in Japan, even though the primary U.S. military response to the 1997 crisis was naval, and involved forces homeported in Japan. Likewise, land-based forces have experienced significant constraints on the use of bases in the Persian Gulf region during Operations Enduring and Iraqi Freedom, while sea-based forces have enjoyed unlimited access to naval facilities in the same region.

Certainly, this asymmetry in favor of sea basing does not come without its costs, and sea-based forces may be less cost effective than analogous land-based forces in those cases where the latter have assured access to local, prepared bases in advance of a conflict. But as has become eminently clear, this level of access will be much less common in the future security environment than it was during the Cold War.

Access for land-based forces can be constrained in both time and scope. For example, a country might deny American strike aircraft based there permission to bomb a neighbor, but allow aircraft flying supporting missions to operate, as has been routine in the Persian Gulf region since Desert Storm. Similarly, a country that had allowed the United States to pre-position war material to support rapid deployment of land-based forces in a crisis will be likely to support such a deployment only under a particular set of circumstances, against a particular opponent. Also, a country might or might not allow a U.S. force to stage through its territory or its air space on the way to another country's bases, and again, there will be cases where a country will allow all forces to do so, and other cases where only support forces such as tankers and airlifters will be allowed to transit. Such access to enroute bases is necessary to any major deployment of land-based forces, both because Air Mobility Command airlifters must stage through them, and because air-refueling tankers must operate from them in order to refuel deploying aircraft. The same constraint applies to global bomber operations launched from the continental U.S., which are utterly dependent on access to enroute bases on foreign territory for use by KC-135s and KC-10s.

In addition to gradations in the scope of access granted U.S. forces, are gradations in the timing with which that access is granted. As noted in the first section, it took the Saudis four days after the Iraqis invaded Kuwait to decide that a massive deployment of American forces would be in their security interest. The speed of the deployment that followed was inestimably aided by the massive pre-positioning of U.S. equipment that the Saudis had agreed to during the Cold War. Thus, a country that already had a major military relationship with the United States, and which had allowed extensive pre-positioned stocks of equipment to be deployed on its territory in anticipation of a scenario like the one it faced in August 1990, still took days to decide how to respond. A country with a more tenuous relationship with the United States, or greater concerns about the projected military operation, might refuse such access, as Turkey did at the outset of Iraqi Freedom. Finally, access once granted can also be revoked, as happened abruptly in the Fall of 2005 in Uzbekistan.

Operations Enduring and Iraqi Freedom have provided myriad examples of these various gradations of access in both time and scope for land-based forces. The central points about this experience are two: the

access eventually gained has over time become extensive, but the process that produced this access began after the need for the access emerged, was protracted, and the results were utterly unpredictable from the vantage point of September 10, 2001.

Military Constraints on Access to Bases Ashore. The technical trends effecting the military security of overseas bases provide a case where the demands of the near and far term security environments appear to reinforce each other, even though the source of this similarity is different. This is because fixed bases within a radius of 500 to 1000 miles of an opponent are likely to become increasingly vulnerable in the near to mid term to conventional attack by GPS/INS-guided ballistic and cruise missiles. Over the longer term, this vulnerability to over the horizon attack will extend to large ships at sea, but in the near to mid term, such ships are likely to be more survivable than fixed land bases as long as they stand off over the horizon from an opponent. Satellites, and especially those in low orbits, will also become vulnerable in this time frame. Least likely to become vulnerable out through the far term, some 30 or more years from now, are undersea platforms. Fast, quiet nuclear submarines will remain the least vulnerable of all basing modes because antisubmarine warfare (ASW) is least affected by the technical trends that will potentially transform other warfare areas. Thus, ASW against modern nuclear submarines will remain both extremely demanding technically, very expensive, and still a largely fruitless endeavor.

The trends in the vulnerability of fixed bases within a theater will be driven in particular by the marriage of cheap, widely available GPS/INS guidance technology and conventional sub-munition payloads with existing, mobile, tactical ballistic missiles (TBMs) and cruise missiles. Cruise missiles will likely be chosen if an opponent wishes to extend his reach beyond about 600 km.[vii] Such cruise missiles might actually have more in common with small aircraft than with current cruise missiles like Tomahawk, and might derive their survivability less from high speed and low radar cross section than from their ability to blend in with a noisy background and deny an opponent positive identification.[viii] This discussion will focus on TBMs because they are already ubiquitous.

Before discussing the impact of GPS/INS, it is important to establish what is already true about existing mobile TBM capabilities. TBM defenses have already proven to be at the edge of the scientific-technical capabilities of the United States, and even assuming effective TBM defenses in the future, they will likely be on the losing end of the cost-exchange ratio between the attacker and the defender. Directed energy weapons may change this relationship in the distant future, but even directed energy weapons will be most effective against antiship missiles whose warheads must have terminal seekers, because it is more difficult to harden the latter against a laser or high power microwave than it is to harden an INS/GPS-guided warhead used in attacks against fixed targets.[ix] Heretofore, the TBM threat has been leavened by the fact that they have been very inaccurate. Armed with conventional warheads, their military effects have been both low and unpredictable, and armed with nuclear, chemical, or biological warheads, their use would bring the vastly superior American deterrent into play.

GPS/INS will lead to a quantum leap in conventional TBM capabilities because it will give them the accuracy needed to attack soft, fixed military targets with conventional submunition payloads. This is a capability that has already been deployed by the U.S. in

the form of the U.S. Army's Tactical Missile System (ATACMS). Single stage TBMs like the Chinese M series already provide ranges of 300-600 km (185-375 miles), payloads of 500 kg (1100 lbs), and conservatively estimated accuracies of 100-200 meters (330-660 feet). A 500 kg payload of simple submunitions like the M-77 grenade can destroy soft targets over a circular area of two million square feet centered at their point of release. Even with the relatively low accuracies assumed above, TBMs like this would wreak havoc on airfields and ports within their range. For the purpose of this discussion, I will focus on airfield vulnerabilities, and in particular on those vulnerabilities most relevant to the mid to far term security environment.

Airfields within range of opposing TBMs will be inherently vulnerable in the expeditionary environments typical of likely future air operations for four reasons. First, simply building the hardened shelters to protect five 72 aircraft wings of tactical fighters is a major, multi-billion dollar investment that few potential allies are likely to make on their own. Second, in a world where the location of future conflicts is less predictable than it was during the Cold War, the U.S. will not be able to invest in hardening airfields in all the potential areas where it may be called on to fight. Third, even in cases where the United States did make this investment during the Cold War, as in Saudi Arabia, it still proved too expensive to provide shelter to the thousands of U.S. personnel which must live and work on the base; hence the enormous tent cities which were erected on base, which would remain a lucrative target even at the hardest base. Finally, it is simply impossible to provide hardened shelters for the various large, high value support aircraft which are integral to any expeditionary air operation, such as AWACS, JSTARS, Rivet Joint, KC-135 and KC-10 tankers, and outsize airlifters like the C-5 and the C-17.[X] For all these reasons, it is likely at some point in the mid to long term future that large scale, expeditionary deployments of tactical combat aircraft and supporting assets will become limited by the need to avoid bases within 1000 miles of an opponent's territory.

This constraint will, at a minimum, reduce by half or more the sortie rate of a given size force of tactical aircraft, and increase the cost in terms of additional tanker support and bases for those tankers necessary to give that force the range to attack a given set of targets. Perhaps more important, there will be circumstances due to the political geography of a particular conflict in which the only bases potentially available will be within range of an opponent's missile force. This will lead to cases where land-based tactical aircraft will be denied access to a given theater altogether, or at least until the threat to their bases has been suppressed or eliminated by other forces.

The Special Case of Bombers and Overseas Bases. Land-based bombers use their much greater range than land-based tacair to reduce the number and/or the political salience of the overseas bases they use, and to increase the standoff range of those bases from the opponent. Politically, bombers can reduce their vulnerability to denied access simply by increasing the probability that an amenable ally can be found within range of the opponent. They can also reduce their vulnerability to being denied access by taking advantage of the fact that a country is more likely to allow tankers rather than combat aircraft to operate from its bases. Taken to its extreme, this latter tactic can involve 30-40 hour round trip missions for

the bomber and its crew between the United States and a target half way around the world. In these operations, the bomber never uses a foreign base, but the vast array of tankers that must pre-deployed along its route must do so intensively. Thus, it is almost never the case that long range bombers eliminate all dependence on access to foreign bases. What they do in almost all cases is lessen that dependence below the political threshold that can lead to access denial.

Added range is not a free good, which is why bombers are much larger and much more expensive than tactical fighters, and tend to be slower. Also, though their size gives them larger payloads, the added range over which they carry those payloads leads to reduced sortie rates. Finally, in contested air space, bombers cannot defend themselves against opposing air defenses, which precludes independent operations by B-52s and B-1s, and precludes daylight operations by B-2s. These tradeoffs between bombers and land-based tacair will become more important when military vulnerabilities at overseas bases are added to the political vulnerabilities that already exist. Assuming that there will be limits on the range of these threats for the foreseeable future, they will improve the relative advantages of long range aircraft compared to land-based tacair, because bases outside that range will not need to be defended and/or hardened.

Distributed Air to Support Distributed Ground Forces

> "…the long thin columns of vehicles penetrating through hostile territory were very weak, seemingly highly vulnerable to attacks on their flanks. Tactically, the columns were of course all flank and no "front." …(But) so long as the invasion columns kept up a high tempo of operations, their apparent tactical vulnerability was dominated by their operational advantage since the defender's intercepting and blocking actions would always be one step behind…The whole operation obviously rests on the ceaseless maintenance of momentum. Organizationally, this implies a very restricted deployment of heavier/slower artillery, the need to keep the supply tail light and fast moving will restrict the amount that can be deployed." [xi]

This description of the German Army's invasion of France in the Spring of 1940 applies almost verbatim to the U.S. Army and Marine Corps' initial operations against Iraq in Operation Iraqi Freedom. These operations will likely serve as a future model for expeditionary ground force operations.

In such operations, ground forces will likely strike from several locations at the same time, operating from expeditionary bases on external lines surrounding the opponent. Because of the expeditionary environment, and because of the need to maximize the speed and freedom of maneuver of the invading force, the size and firepower of the invading force will frequently not be sufficient to win a straight attrition battle with opposing forces. Rather, that force will seek to use speed and maneuver to confuse and paralyze the opponent's command system, and to quickly penetrate and descend on the objective from several directions.

Air forces must play a central role in such operations for two reasons. First, when the desired shock effect depends so directly on the maintenance of momentum, and when ground forces must be kept light to create that momentum, air support becomes a key source of fires enabling either the rapid elimination of

obstacles to the ground force's advance, or their avoidance and later reduction.

Second, such operations absolutely depend for their success on the creation of chaos and confusion among opposing forces. In cases when the opponent maintains its cohesion at the operational level, as for example did the Allied forces deployed in the Ardennes during the Winter of 1944, all of the potential tactical vulnerabilities described above become real, and the penetrating columns of the attacker are cut off from their supplies and attacked by opposing forces able to concentrate and maneuver against their weak, exposed flanks. A major if not dominant mission of future air forces will be to detect and destroy such concentrations down to the company level before they can strike. This requirement is one of the main drivers behind the need for air forces to develop mobile target kill capabilities against opposing ground forces.

The specific requirements for such a mobile target kill capability will be discussed below, but the main requirement is for air forces capable of being in many places at once, to match the distributed nature of the ground battlefield. Distributed air forces assure that targets can be struck quickly when the need for such strikes arises. In this respect, the mobile target problem will create demands for a very different force than will the fixed target problem. For example, in the latter case, the lack of time urgency and the desire for efficiency may call for large platforms with large weapon payloads that can attack their targets at leisure. Such a force would be inappropriate for the mobile target problem because at any one time, it would only be within minutes of striking one part of a distributed battlefield.

In addition, in cases where the opponent has significant air defenses, the demand for a distributed air force must be met from the very outset of the conflict with absolute reliability, given its' role in the combined arms battle. This is one of the main reasons why naval aviation is pursuing the cost-effective, all aspect stealth that F-35C will provide. It will enable distributed operations from the first day of any conflict without the prior need for a large defense suppression campaign, or the continuing need for dedicated defense suppression escorts.

Distributed operations of this type will likely dominate the future regardless of the opponent because they are a potential solution to problems that occur at both the low and high ends of the conflict spectrum. At the low end of the spectrum, distributed operations make it possible for the combatant commander to deploy a smaller, lighter, more sustainable force for the initial operation to repel or assault an opponent's forces in what will often be an austere environment against lesser powers. At the high end of the spectrum, it is possible if not likely that distributed operations will prove a necessity, both to reduce the ground forces' dependence on vulnerable fixed bases and staging areas that can be attacked by the opponent, and to avoid the need for breakthrough battles against that opponent's forces in order to envelop and destroy them, perhaps via the means of multiple, simultaneous vertical envelopments.

Concerns About Casualties and Collateral Damage

America's supposed aversion to casualties in post Cold War conflicts has been much discussed. Fear of casualties often measured in the thousands or even tens of thousands dominated the debate over whether to launch a ground war in Desert Storm. In the event, casualties during Desert Storm were orders of magnitude lower than expected, leaving the question of America's tolerance

for casualties open for debate.[xii] Then the events of early October 1993 in Mogadishu, Somalia seemed to resolve the debate.[xiii] The death of a small number of Rangers and Delta Force troopers abruptly led the United States to abandon that operation. A growing consensus developed that the United States could be stopped in its tracks with the deaths of a few of its soldiers, leading some to question the viability of its enormous but seemingly unusable military power.

The later experience in Kosovo certainly did not provide evidence that the United States was not casualty averse, and it also brought to the fore the related question of rules of engagement and whether its fears of causing collateral damage to civilians and civilian infrastructure had come to hamstring its operations to the point of impotence. For example, NATO air crews were ordered to remain above 15-20,000 feet throughout the entire conflict because it was only at that altitude that they remained immune from Serbian air defenses, and, of course, ground forces were foresworn from the outset. This restriction also reduced NATO's ability to distinguish between Serbian forces on the ground, Kosovar fighters, and fleeing refugees, making it all but impossible to stop or limit the ethnic cleansing being conducted by Serbian army and police units in Kosovo, driving NATO political and military leaders to adopt a gradual strategic bombing campaign designed to coerce Serbian compliance without causing excessive damage to the civilian infrastructure, which took months to succeed.

As of this writing, Iraqi Freedom has not provided unambiguous evidence on this question, but it is clear that for those who do oppose the war, U.S. casualties are a major cause of that opposition. The explanation for this aversion has more to do with the strength of the United States' position in the world, rather than the weakness of its leaders or its people. The U.S.'s basic security means that it rarely if ever has to fight wars of necessity against formidable opponents, and the general weakness of the enemies it does fight means that U.S. forces are rarely driven by military necessity to make attacks against targets where there is any but the smallest chance of collateral damage.

On the first point, the United States is the most secure country the world has ever seen:

> "..(which) leads to something of a paradox: Although solving many global problems requires active U.S. involvement, Americans do not see them as vital to their own interests and they are unwilling to expend much effort addressing them… Americans would like to coerce others to do what they want, but they aren't willing to risk much blood or treasure to make sure they do."[xiv]

In this view, America's aversion to casualties, and the degree to which political leaders will constrain how the military fights in order to reduce its exposure to casualties will depend on the stakes the United States has in the conflict. Because of the great overhang of American power in today's security environment, and because of its basic security, few if any conflicts are likely to engage its vital interests, and many conflicts, like Kosovo, the first Gulf War, and even the current war in Iraq, will be viewed by many Americans as wars of choice. Using the same logic, concerns regarding attacks that might cause collateral damage grow steep when it is difficult, against a very weak opponent, to argue that military necessity allows no alternative to launching those attacks.

This basic structural paradox sets the bar

extremely high for the U.S. military, because it must win while keeping its losses and collateral damage extremely low by historical standards. Certainly, the pressures in this regard will vary somewhat, depending on whether a conflict is a major contingency like Iraqi Freedom, or instead, a humanitarian intervention in Latin America or Central Africa. Yet in the absence of major war with a great power, there is little prospect that the U.S. military will see this bar lowered.

The main military consequences of this reality will be a growing demand for weapons which can stand off at a distance from enemy defenses and avoid direct fire engagements with their targets at short ranges, but also growing demands for very precise identification of targets before they are attacked and very precise results when attacks are approved. In many cases, such as in attacks from the air against high profile, fixed targets on the ground, precision weapons have or soon will solve these problems. In other cases, such as in attacks from the air against military vehicles or convoys outside of contact with friendly ground forces, the problem of combining precise identification and weapon effects with essential immunity from attack is far from solved. In still other cases, such as in today's urban counter-insurgency operations in Iraq, it is only possible to imagine solutions to this problem through the deep integration of air and ground forces down to the smallest units.

Another consequence of this problem is a tradeoff between concepts of operation which devolve the maximum degree of authority to execute operations to the engaged units, and those that concentrate decision making and execution authority at central operations centers. In an ideal world, distributed operations would lead to radically decentralized execution authority, thereby speeding up the decision cycle, maximizing the prospect that fleeting targets could be struck before they disappeared again, and reducing the overall operational risks associated with distributed operations. On the other hand, when political and military leaders have to weigh the political costs of either casualties or collateral damage in limited conflicts, they will remain reluctant to allow such decentralization, and concepts of operation will tend toward centralized control and execution.

The perceived cost of casualties and collateral damage would certainly change in wars of necessity when military necessity demanded it. But it is important to note that concepts of operation bred in today's security environment will have a tendency to "reify" themselves in the doctrine and force structure of tomorrow's forces. This issue is of particular importance to air forces, where both the desire and the ability to centralize execution is felt most strongly by military leaders prosecuting wars of choice against weak opponents, but where the consequences of building centralization into future doctrine and systems could be devastating in the event U.S. forces faced more formidable opponents.

Technology Trends in the New Security Environment

Longer term trends in technology are more difficult to predict, but some can be identified with more or less confidence. As in the previous discussion of operational demands, some trends in technology are general and apply across all or many mission areas, while others apply most strongly to a particular mission area. The following discussion will emphasize those trends that are likely to apply across all or most mission areas.

The conventional wisdom in the U.S. defense department regarding the future of

military technology appears to contain at least five assumptions or desires: that space-based sensors will play a larger role in future operations; that sensors, however based, will increasingly be linked to form networks, rather than used independently; that unmanned vehicles in the air, on the ground, and on and below the surface of the sea will increasingly takeover missions now assigned to manned platforms; that more and more battlefield information processing, decision making and control will be centralized in rear area operations centers; and that the bandwidth to enable networks, operate unmanned vehicles, and support rear area operations centers will be available. Certainly, all of these trends are real and all are likely to continue, but at the same time, each can also be taken too far.

Space-Based Sensors

There is a long history of the tactical exploitation by the defense department of so-called national capabilities, or space-based sensors. These efforts exploited assets that were already in orbit that were designed, procured, and operated by the intelligence community. Today, there is much discussion of space-based systems that would be developed and operated jointly by DOD and the intelligence community, and which would therefore be designed from the outset with DOD's as well as the intelligence community's needs in mind.

It is easy to show what the inherent limits of airborne sensor platforms are compared to spaceborne sensor platforms. Airborne sensors can only see out to the horizon, out to perhaps 200 miles at the most. Spaceborne sensors looking down at the earth from low earth orbit can see a patch of ground at least twice that area, and from geosynchronous orbit, satellites can "see" almost an entire hemisphere. Airborne sensor platforms are also easier to shoot down then satellites and must generally standoff some distance from hostile air space until opposing defenses are suppressed or destroyed.

By contrast, the limits of spaceborne sensors are more subtle and generally require a technical understanding that is largely absent from most public debates. For example, airborne sensors can dwell in a given area for an extended period of time, with its sensor performance limited only by line-of-sight constraints. Satellites, on the other hand, face a tradeoff between low orbits, where sensor performance is maximized but dwell time in any one area is measured only in minutes, and much higher synchronous orbits, where dwell time is maximized but where many sensor phenomenologies are ineffective because of the great distance to the earth's surface. For example, radar and optical imaging sensors cannot be deployed in synchronous orbit.

Therefore, to achieve the dwell time necessary for surveillance as opposed to reconnaissance, radar and optical imaging sensors in low earth orbit must be deployed in constellations of 20-40 satellites in order to ensure that one is over the area of interest at any time. By contrast, a single airborne sensor platform can provide continuous surveillance of a given area for many hours, and three or four can provide 24 hour surveillance indefinitely.

Airborne sensor platforms can also deploy larger antennas with more powerful and continuous power supplies. This is particularly relevant in the case of MTI radars, whose performance is very sensitive to power and aperture. Airborne sensor platforms also have significant advantages when used as SIGINT platforms against highly directional signals such as those generated by air defense engagement radars, whose main beams propagate horizontally rather than vertically.

Because of these various tradeoffs, space and airborne sensor platforms should complement each other, but as in many mission areas, it is difficult for many casual observers not to see two different platforms deploying the same types of sensors as duplicative. Add to this the fact that a satellite remains a vastly more expensive method of deploying a sensor than an airborne platform and it is easy to see how a push to transform space reconnaissance into space surveillance could have unintended and undesirable consequences for the Department of Defense if it comes at the expense of the ability to fund airborne surveillance platforms.

Sensor Networks

Whether deployed in space, in the air, on the ground, or under the sea, sensor performance can often be improved dramatically when the output of multiple sensors is compared and fuzed. For example, even the most advanced RF antennas have inherent limits in their resolution which, at longer ranges, produce errors in target location measured in 100s if not 1000s of feet in azimuth for MTI radars and ELINT receivers. These errors can be eliminated or dramatically reduced if several sensors are deployed within line-of-sight of the target and of each other and their output data-linked and processed. Two or three networked MTI radars using trilateration can precisely locate and track a moving target, and two or three networked ELINT receivers using some combination of TDOA and FDOA processing can precisely locate a hostile radar.

Furthermore, in many cases, networking can enable a significant reduction in the cost of individual sensors. For example, the best ELINT antennas with angular resolutions measured in single degrees or even fractions of a degree require large platforms like the RC-135, and are therefore very expensive. By contrast, the antennas used in a TDOA/FDOA network can be extremely small and cheap because the angular resolution of the nodes in the network become irrelevant to its performance. In this extreme case, a network comprised of several very simple nodes not only outperforms a single, very sophisticated platform, but is also much cheaper.

But many ignore the assumptions hidden in this comparison. First and most obviously, such networks fail deadly. That is, they require a certain number of nodes to function, and deprived of even one node they fail completely. There is no graceful degradation in a TDOA/FDOA ELINT network if its individual sensor nodes are not also designed to operate alone. Second, in the case where network designers anticipate this vulnerability, they can seek to deploy redundant sensor nodes to compensate for losses or malfunctions, but in this case they must also ensure that a high degree of self-organization and self-healing is built into the network. Otherwise, the additional, redundant sensor nodes can become a source of error rather than insurance. Finally, in denied or contested environments, it is often difficult to deploy and operate any sensors at all, never mind a redundant network.

An alternative approach to network design would look at networks as an opportunity to greatly expand the potential capabilities of individual sensor platforms that can function autonomously if necessary. Certainly, in the near to mid term, there is the need for considerable experimentation to determine both the strengths and weaknesses of networking, and a strategy that led to the deployment and use of networks that failed safe rather than deadly would have real benefits. There is also considerable promise in the use of networks where the nodes are so cheap as to be truly expendable. These will usually be networks of passive sensors and will therefore also provide the potential for

covert surveillance. Such networks would provide an ideal venue for experimentation with self-organizing and self-healing systems.

Unmanned Aerial Platforms

Taking the aircrew out of an airborne platform can clearly solve some important military problems. High endurance, high altitude UAVs like Global Hawk have amply demonstrated their worth as sensor platforms, and both the altitude and endurance of their operation are simultaneously the source of their relative value and the result of taking the aircrews out of the platforms and putting them on the ground. By the same token, a cruise missile like Tomahawk that can fly up to 1000 miles into contested air space eliminates the operational tradeoff between deep penetration attacks by platforms like the B-2 and the need to keep air crews within range of helicopter-based combat search and rescue (CSAR) assets whose range is limited to 2-300 miles.

The next step in unmanned aerial platforms is assumed by many to be some sort of unmanned combat air vehicle (UCAV) which would deploy both sensors and weapons into contested air space and return to be reused again. In many eyes, UCAVs are viewed as successors to manned strike fighters. This implies many things, among them that UCAVs can be provided the target recognition and situational awareness capabilities that air crews provide strike fighters.

For example, target recognition often involves the generation and interpretation of high resolution images. At some point in the future it may become possible to automate that process, but today and for a number of years target recognition will require people to interpret the images. Situational awareness allows many airborne vehicles to operate in the same air space without collision, and it also allows aircraft to respond immediately to a variety of hostile threats, either to avoid them or attack them. Air crews accomplish this function by constantly monitoring a variety of sensors and cues, including predominantly what they collect themselves visually on a continuous basis. As with target recognition, it is difficult to imagine automating this function.

In short, providing target recognition and situational awareness capabilities to UCAVs will likely still require crews, but those crews will be physically separated from the multitude of sensors that provide that capability. The bridge linking the crews to their sensors will be data links, some of the constraints of which will be discussed below. The one unambiguous advantage of separating air crews from their platforms is the increase in the latter's range and endurance that becomes possible. For example, in the case of carrier aviation, whatever payloads are deployed on carrier-based UCAVs, the ability to deploy those payloads on platforms with a 1500 mile radius and 12 hour endurance will be the factor that drives their adoption.

Centralized Air Operations Centers

Since the first Gulf War, combined air operations centers (CAOCs) and other rear area processing facilities have played an increasing role in air operations. The use of the word combined implies correctly what the function of these centers has been – to combine a series of normally separate streams of intelligence, surveillance, and reconnaissance assets into a fuzed picture of the battlefield to support decision making by high level leaders. A CAOC can also be described with equal accuracy as a means of centralizing decision making authority on the battlefield.

Specifically, CAOCs are a place to fuze the output of many independent sensors, whether spaceborne, airborne, on the surface or under the sea. When it is necessary or desirable to create networks of sensors, CAOCs are often a place to do the unique processing that gives those networks their power. And certainly, if UCAVs are deployed in large numbers, CAOCs or rear area locations like them will be one place to put the ground crews that will give those assets their target recognition and situational awareness capabilities.

Such centralization has two major consequences, one that is obvious and that will be discussed in the next section. The other consequence is more subtle. Operations by even modest-sized forces on cluttered battlefields generate the need for thousands of tactical decisions a second, most of which are obviously made on the spot in real time by small units and individuals who then also become the means of executing those decisions. If one seeks to take any significant percentage of this activity and physically separate the means of decision from the means of execution, and if one wishes to preserve if not accelerate speed and responsiveness, then both the number of independent, rear area decision makers and the bandwidth connecting them to their means of execution become fundamental determinants of the possible pace and scope of military operations.

Of course, this is ironic because CAOCs have for the last decade increasingly been the scene of successful efforts to speed up decision making and execution, but this progress has occurred during a phase when the main obstacle to fuzing or networking information was technical rather than bureaucratic. It also has occurred during conflicts in which U.S. forces were still designed and trained to operate without ISR fusion and networking if necessary, in which operations directly supported by the CAOC were a very small albeit important subset of the whole, and in which the forces deployed by the opponents were vastly inferior, particularly in qualitative terms.

Secure High Data Rate Communications

In a future where it is assumed that many tactically relevant sensors are deployed in space as well as in the air and on the surface, that sensors are often formed into networks, that sensors and weapons will deploy into contested air space on unmanned platforms, and that ISR information is combined and weapon release authority centralized in rear area operations centers, one capability above all others will be required—the ability to move vast amounts of data between many 100s if not 1000s of mobile platforms, reliably and over great distances. The solution so far to this emerging issue has been to lease bandwidth on commercial satellites.

But if future opponents have even modestly better electronic warfare capabilities than todays' opponents, never mind if a true peer competitor should emerge, than those high data rate communication links will have to be secured in the face of efforts to jam them. Today, in the entire Department of Defense, there is not a single communication system that comes close to achieving this objective.

The problem is that data rate and jam resistance directly compete for the same bandwidth. For example, UAVs like Global Hawk and Predator often use commercial satellite communication systems operating at Ku band (10-15 GHz) when they need to relay their sensor outputs over the horizon. These satellites are designed to maximize data rate and can support links with throughputs of 10s of MBs/sec. but are thor-

oughly vulnerable to simple noise jamming from anywhere within their uplink antenna's footprint, which for a geosynchronous orbit is always large. By contrast, Milstar, a DOD satellite system which operates at even higher Ka band frequencies (and therefore higher bandwidth), is optimized toward jam resistance and will only support data rates measured in 10s of KBs/sec; i.e. several orders of magnitude less. Hybrid waveforms at Ka band used on Milstar II that seek a compromise between jam resistance and data rate can support data rates of roughly 1.5 MBs/sec, or the equivalent of a T-1 line. At lower frequencies, such as the ubiquitous UHF SatCom systems which are most useful to tactical forces because they do not require expensive terminals and highly directional antennas, such as those used by Milstar, there is much less bandwidth available, and therefore both data rates and jam resistance are inherently low.

One option for addressing this problem within a specific area of operations is to rely more on networks of airborne relay platforms that are linked by line-of-sight communications. Such airborne networks have the potential to address the data rate/security tradeoff in at least three ways. First, the ground footprint of an airborne receive antenna is much smaller than for a satellite antenna, which means that the area from which an opposing jammer could introduce a spurious signal into the antenna is also greatly reduced, making it easier either to use array gain to suppress the spurious signal or to take offensive action against the jammer itself. Second, line-of-sight links lose much less of their signal strength to propagation losses and data rates are therefore inherently higher, providing the waveform designer with a higher base to trade from in seeking to build in some jam resistance. And finally, because airborne platforms are much cheaper than satellites, it is possible to create networks with many more nodes, all within line-of-sight of each other but arrayed along different azimuths from any one jammer. Azimuth diversity combined with packet switching could enable an airborne communications network that resembled the terrestrial internet in that it could experience a loss of the link between any two nodes and still allow them to communicate with each other.

If one looks around the world, one can identify those few countries with serious air defenses by the presence of AEW aircraft. At sea, the presence or absence of carrier-borne AEW is even more decisive.

Sea Shield Past, Present, and Future

BECAUSE THE OTHER SERVICES are likely to face political constraints on their access ashore early in future conflicts, the Navy will face greater demands on its power projection capabilities. But the Navy will also remain largely if not solely responsible for countering opposing access denial efforts at sea, both to ensure the security of its own base of operations, and to enable the safe entry and secure operation of joint, follow-on forces.

This section is organized in three parts. It discusses the evolution of the demands on the Navy's access assurance capabilities during and since the end of the Cold War; the current status of sensors, weapons, and networks in this mission area; and future opportunities for innovation in both technology and doctrine by naval aviation.

The Cold War Legacy

The pillars of today's and tomorrow's Sea Shield posture were laid during the Cold War in the undersea and antiair warfare mission areas. In these warfare areas, the Navy faced an opponent whose prime focus was the denial of access by sea-based power projection forces, and who chose what would now be called asymmetric means in the pursuit of that goal.

Undersea Warfare

Undersea warfare can be divided for our purposes into antisubmarine warfare and counter-mine warfare. Both warfare areas have experienced dramatic change since the end of the Cold War, but both remain important sources of sea denial leverage for future opponents. That is because modern, non-nuclear submarines and mines remain in some ways the ultimate conventional, asymmetric threats. They can do damage to major, high value naval platforms, yet they can only be countered by an effort whose cost greatly exceeds that necessary to generate the initial threat. Thus, they pose unique challenges in today's security environment because they remain one of the best ways to cause politically significant losses to American or allied ships despite the dramatic diminution in the overall level of the ASW and mine threat compared to the Cold War. This often makes the case for better ASW and mine warfare capabilities both important and difficult to make in today's budgetary environment.

ASW During the Cold War. After World War II, Soviet submarines based on captured German designs threatened to render obsolete much of the U.S. Navy's ASW posture, which had been focused on dealing with submarines that lost a substantial portion of their offensive capabilities when forced to submerge. At the same time, the Soviet Union, being a continental power, threatened to make the U.S. Navy's victorious submarine force irrelevant, since submarines were primarily useful as an anti-surface weapon against merchant shipping, and the Soviet Union could easily survive without merchant shipping.

Out of this challenge grew two initially separate innovations which, when brought

together, formed one of the cornerstones of the U.S. Navy's Cold War ASW posture.

The first innovation involved the exploitation of passive acoustics to detect and track submerged submarines, using the sounds they generated as a signature. Passive sonars significantly increased the range at which submerged submarines could be detected compared to active sonar, allowing for very wide area searches by ocean-wide sound surveillance systems, which in turn could be used to accurately cue ASW platforms to localize and prosecute the submarine contact. The second innovation began with the embrace by the U.S. Navy's submarine and maritime patrol communities of ASW as their primary Cold War missions, using passive acoustics as their primary method for search, classification, and localization.

Maritime patrol aircraft offered speed that submarines lacked, making them particularly useful in the initial localization of a contact provided by offboard surveillance systems, which could then be handed off to a platform with more endurance, such as a nuclear submarine. The surface warfare community remained dependent on active sonar until the late 1970s. Then, in response to the deployment of more capable Soviet submarine-launched antiship missiles, surface combatants also embraced passive acoustics and long range, shipborne ASW helicopters.

By the early 1980s, all of the Navy's platform communities were being used successfully in ASW operations against Soviet submarines, and increasingly these operations demanded a high degree of coordination as Soviet submarines became quieter. Earlier in the Cold War, when U.S. acoustic superiority was still unchallenged, each platform community's ASW operations had been relatively independent of each other. This independence reflected a natural division of labor based on the strengths and weaknesses of each ASW platform. Thus, submarines went forward into contested waters where other ASW platforms could not operate, maritime patrol aircraft used their speed to prosecute long range contacts generated by underwater surveillance systems, and surface combatants utilized their endurance to provide a local screen for battlegroups and convoys.

The key to success in these relatively uncoordinated operations was maintaining a high degree of acoustic superiority over Soviet submarines. Ironically, that superiority began waning in the 1980s, just as the Cold War was ending, in an echo of the end of World War II. This ending to what was the third battle of the Atlantic was fortunate, but the Navy will face new ASW challenges not unlike those it avoided when the Soviet Union collapsed, albeit on a smaller scale.

ASW After the Cold War. The threat to American acoustic superiority resulting from the first Soviet deployments of the Akula in the mid 1980s may recur in today's security environment with the increasingly wide proliferation of modern non-nuclear submarines. Deployed relatively close to their homes, in or near littoral waters through which the United States may need to project power from the sea, and where it is easier for a weaker Navy to obtain cueing information against U.S. ships, these submarines pose a potentially formidable threat. With a competent crew and the kind of advanced weapons that are now widely available in global arms markets, a modern non-nuclear submarine deployed in its own backyard might become a poor man's Akula. Of even more concern is the fact that modern weapons, such as wake homing torpedoes for example, tend to reduce the demands on submarine crews, making even less competent crews too dangerous to ignore.

Modern non-nuclear submarines are both better than those deployed by the Soviet Union during the Cold War, and more widely available as defense industries that served their home markets during the Cold War now use exports to stay alive. One reason that the submarines are better is because many decades of continual investment by countries like Germany and Sweden have finally paid off in the form of non-nuclear submarines with both rafted diesel propulsion plants that greatly reduce their acoustic signature when snorkeling, and with air independent propulsion (AIP) systems that make them more like true submarines rather than mere submersibles.

These submarines still do not provide anything like the mobility and endurance of a nuclear submarine, but they reduce the indiscretion rate of a traditional diesel-electric submarine when on a slow speed patrol. Such a submarine, patrolling in a limited area in or near its home waters, would need to expose its snorkeling mast much less frequently than do older versions of the Russian Kilo and would be less vulnerable when it did.

Such submarines will also be armed with better weapons and fire control systems. One particularly alarming development is the marriage made possible by the end of the Cold War of the air independent, non-nuclear submarine with the submarine-launched antiship missile. Armed with Harpoons or Exocets available from several western suppliers, or Russian missiles like the Novator 3M-54E, these platforms can launch fire and forget missiles from over a surface ship's radar horizon without the need for the noisy and battery-draining approach run necessary for a traditional, torpedo-armed, diesel-electric boat. Absent high quality over-the-horizon cueing, these attacks will be prone to homing on the wrong target in a cluttered environment, but will be very hard to defend against in those cases where the weapon homes on the right target. This threat circumvents the traditional ASW approach to dealing with very quiet diesel-electrics, i.e. to flood the ocean surface with radar and use speed to force the submarine to either run down its battery and expose itself in an attack run, or stay quiet and defensive.

There is also a political challenge associated with conflicts in which the United States is fighting over less than all out stakes. In such conflicts, there will be a very low tolerance for shipping losses, but the presence of an opposing submarine force will put great pressure on the Navy if it must rapidly project power and protect against those submarines at the same time.

Regarding casualties, even in a major regional contingency, the stakes for the United States are limited while those of its opponents are very high indeed. The opponent may be willing to run great risks and sustain high losses, while the U.S. is less willing to do so. Faced with the possibility or the reality of losses at sea, the Navy will need to mount a major effort to eliminate the threat of further losses. In order to be able to do this while still projecting its own power, the Navy will need to make ASW a less protracted and asset-intensive exercise.

A good analogy is to the great Scud hunt of Desert Storm. Thousands of sorties were diverted over several weeks from the air war during Desert Storm to hunt for SCUDs to little or no effect. From an ASW perspective, this experience is illuminating for both operational and political reasons.

Operationally, Scud hunting was like ASW using traditional methods against a very quiet target. A large area needed to be searched for objects that easily blended into the background and only intermittently exposed themselves. Thus radar was used to

flood SCUD operating areas, unattended ground sensors were also deployed, and aircraft were used to pounce on potential contacts. This was a protracted, extremely asset intensive endeavor, characterized by false alarms, high weapon expenditures, and low success rates. In short, a SCUD launcher was most likely to reveal itself by successfully launching its weapon, just as sinking ships are often the only reliable indication that there is a submarine in the neighborhood.

The political lessons of the SCUD hunt also apply to ASW. Before the war, the SCUD had rightly been dismissed as a serious military threat, but once they began landing in Israel, the political imperative to allocate scarce resources to at least appear to counter this threat rapidly overwhelmed these narrow military calculations. The same political pressures would be brought to bear on ASW forces facing active enemy submarines, but unlike the Iraqi Scuds, which were terror weapons without much military utility, submarines are a serious military threat as well a political one. Therefore, it will be important to avoid delays in containing the ASW threat, and an ensuing delay in the closure of Marine amphibians or Army sealift ships.

A delay of several weeks during the halting phase of a major contingency might not be a war stopper all by itself, but it is important to understand the consequences for current time phased force deployment list (TPFDL) timelines, which assume closure of millions of square feet of pre-positioned sealift within the first two weeks of the start of an MRC. This would transform a rapid deployment into a slow one, throw the deployment timelines of all the services askew, and open a window of indeterminate size at the outset of a conflict in which the enemy can operate unmolested except by those opposing forces already in theater, assuming they do not need an open sea line of communication to sustain themselves.

ASW will also be the primary tool for protecting merchant fleets if and when attacks against them are launched by an opponent seeking to coerce a third country by attacking it commercial shipping. The capability to assure the protection of its commercial shipping will probably be the single most important kind of security guarantee that the United States will be able to offer medium-sized powers along the Indo-Pacific littoral, because it is likely that submarine-based threats to that shipping will be the primary source of leverage deployed by any would be hegemon in the region. This argument applies with particular force to those countries dependent on external sources of energy that are shipped by sea.

There is also a doctrinal challenge the Navy faces as it attempts to increase its ability to project power from the sea. The Navy faces a new operating environment in which it is increasingly relevant and therefore in demand. Unlike in the post WWII era when the Navy was searching for a mission, it has been inundated with new missions in the post Cold War era, and these new missions compete with ASW for resources.

This has serious consequences for ASW because, as noted above, ASW is a multi-platform mission area performed by multi-mission platforms. As the Navy's strike warfare, anti-air warfare, missile defense, and amphibious warfare capabilities have grown in importance in the nation's military strategy, the Navy has shifted its focus away from an emphasis on blue water sea control toward power projection and land control in the littorals. Yet these missions must be performed by the same platforms that will perform ASW in the littorals - the air, surface, and submarine communities, all supported by the ocean surveillance community.

This "multi-mission pull" increasingly makes ASW compete with strike warfare

and theater air and missile defense for the same resources and training opportunities. This shift in orientation is occurring at a time when technology increasingly demands that ASW be a coordinated, "combined arms" exercise if it is to succeed. All elements of the Navy's ASW posture must be maintained to succeed in the fight against quiet submarines, but all three of the Navy's major platform communities also face pressures to improve the capabilities of their multimission platforms in other mission areas.

Mine Warfare During and After the Cold War. Counter-mine warfare in today's security environment shares much in common with ASW, but is also unique in several respects. Like modern non-nuclear submarines operating on battery, mines can not be detected at operationally significant ranges using passive sonar, and they "operate" in a shallow, cluttered environment in which their small size and ability to remain still while retaining operational effectiveness all conspire to make detection and classification with active sonar extremely difficult. Likewise, in their effects, they also pose the same kind of asymmetric threat in operations where the U.S. Navy and its allies must limit ship losses to very low levels.

Like submarine-launched torpedoes, mines attack ships under their waterline which makes them extremely lethal, but unlike submarines, mines lack mobility. Thus even more then submarines, mines are only effective when used in confined waters or chokepoints, and most mines also require relatively shallow water. Thus, mines have always had particular utility when used to limit passage to and from ports, to limit the operation of ships in shallow coastal waters or straits, and to frustrate or delay amphibious assaults.

All of these potential uses for mines have been of historic concern for the U.S. Navy, but during the Cold War its counter mine posture was determined largely by a small subset of this threat. First, traditional amphibious assaults was not considered likely in a major war with the Soviet Union, and though the Navy and the Marine Corps retained capabilities to clear mines in the approaches to a landing beach, the requirements in this mission area were set at the relatively low level expected in lesser contingencies. Second, the U.S. Navy's main operational focus during the Cold War lay in countering the Soviet Navy's expected attempts to contest control of the Atlantic and Pacific sea lines of communications (SLOCs). In this blue water environment, mines were a minor factor. Certainly there were ports at both ends of these SLOCs, and there were also shallow, enclosed seas like the Baltic and Yellow Seas which would have been contested, but here Allied navies bore the brunt of the counter-mine burden. The main exception to this division of labor lay in the need for the U.S. Navy to assure access to ports in the United States. For this purpose, the Navy developed and maintained a dedicated, U.S.-based Mine Countermeasure (MCM) force.

Desert Shield illustrated two weaknesses in this posture. First, early arriving naval forces lacked the organic MCM capabilities needed in the event of an aggressive Iraqi mine laying effort in the shallow waters of the Persian Gulf. In the event, a relatively small and incompetent Iraqi mine laying effort led to two major ship casualties. Second, even after dedicated MCM forces arrived in the Gulf after several months, these forces could not clear the extensive mine defenses the Iraqis had prepared along the Kuwaiti coastline with sufficient confidence to enable an amphibious assault.

This experience highlighted the new MCM challenges presented by the new

security environment. First, CONUS-based, dedicated MCM forces can not deploy fast enough to support a forward deployed Navy that must confidently operate in littoral waters early in a conflict, so those forward deployed forces must have organic MCM capabilities that at least allow them to find, identify, and evade mines that would otherwise limit its access. Second, a serious mining effort by a competent adversary using modern mines will demand MCM capabilities based on new technology not resident in existing MCM forces.

This challenge will be most serious in two specific scenarios where mines can extract the greatest leverage; in deterring amphibious assaults against prepared coastal defenses, and in delaying or interdicting the deployment and sustainment of land-based forces by mining the ports of debarkation to which their sealift must have timely and unimpeded access. In the second of these scenarios, the ASW and MCM challenges may merge, as the submarine may be the only mining platform available to a weaker power seeking to operate in an opponent's home waters. In both cases, the U.S. Navy's challenge is to enable power projection and sustainment by joint forces, and to protect commercial shipping routes.

Antiair Warfare

As with undersea warfare, elements of the U.S. Navy's current antiair warfare (AAW) posture can be traced back to its experience in World War II. Today's antiship cruise missile threat is the descendant of the Kamikaze threat and has traditionally represented the primary above-the-waterline access constraint for naval surface combatants. Nonnuclear, land-based ballistic missiles have not traditionally posed threats to ships at sea, but could in the future if more advanced adversaries can both deploy long range sea surveillance systems and maneuvering reentry vehicles with effective terminal sensors. GPS-guided ballistic and cruise missiles can already now pose threats to bases ashore. Thus, the Navy will both need to defend itself at sea against cruise missiles and, perhaps, ballistic missiles, and also project a defense against such weapons ashore.

Antiship Missile Defense During the Cold War. During World War II, the integrated air defenses contained within Carrier Task Forces became quite effective against Japanese dive bombers and torpedo bombers for two reasons. First, they projected the defense outward such that many Japanese aircraft never delivered their weapons, and second, their inner or terminal defenses greatly reduced the effectiveness of weapons that were delivered by deterring most Japanese pilots from flying the delivery profiles necessary to give the short-ranged and unguided antiship weapons of the day the accuracy needed to strike a maneuvering ship with reasonable probability.

During the last year of the war, two new AAW challenges presented themselves. First, the Navy's Carrier Task Forces switched from pursuing the by then defeated Japanese fleet to supporting amphibious assaults beyond the range of land-based, tactical aircraft. This fixed carrier operations in space and time, making their movements more confined and predictable, and therefore making them easier for opposing, land-based air forces to find. Furthermore, this limitation on the carriers' ability to use movement and deception to frustrate Japanese air attacks lasted for the weeks or months that it took to build up land-based aviation ashore.

Second, it was also at this point that the Japanese introduced the Kamikaze tactic. The challenge posed by Kamikaze aircraft

was that their pilots were no longer deterred by a Task Force's terminal defenses, making the platforms they were piloting into very intelligent missiles that were guided all the way to their targets. These aircraft had no better luck than their non-Kamikaze counterparts penetrating a task force's outer defenses, but those that did penetrate were much more lethal. Thus, Carrier Task Forces became easier to find because they were tethered to the shore for an extended period, and their terminal defenses were less effective against guided weapons that could not be deterred from pressing home their attacks.

During the Cold War, the evolution of the antiship missile threat went through three phases corresponding to the years when the Carrier Battle Group was expected to be a primary nuclear delivery platform against the Soviet Union (roughly 1948-1960), the years when Battle Groups were focused on projecting power in limited conflicts in the third world (roughly 1960-1975), and the years when Battle Groups refocused on operations against the Soviet Union, albeit in a primarily conventional rather than a nuclear role (roughly 1975-1990).

During the first phase, the Soviet Navy deployed radar-guided missiles in both air and submarine-launched versions that were designed to defend Soviet territory from carrier-based nuclear strikes. Launched from faster, higher flying, radar-equipped jet aircraft like the Badger, these air-launched missiles posed a day or night, all weather threat to the carriers which could not be countered by traditional air defense systems. Attacking jet aircraft approached the carrier too high and fast for reactive, deck-launched intercepts to be effective, while the tactic of having a continuous combat air patrol in the air above the carrier was infeasible using the Navy's early jet interceptors, which had low endurance and were not yet truly night/all weather platforms. Furthermore, antiaircraft guns were almost completely ineffective against antiship missiles with jet and later rocket motors.

Out of this threat grew several major innovations which have become keystones of any modern integrated air defense system. Carrier-based airborne warning and control aircraft with powerful radars were developed and deployed which greatly extended the outer ring of a Battle Group's defenses by providing much more warning of attack. Radar-guided surface-to-air missiles (SAMs) were developed and deployed. SAMs greatly increased the reach and effectiveness of an individual ship's defenses. Ships so equipped provided true night/all weather air defense capability, and with a family of missiles of varying size and range—the so-called 3-Ts: Terrier, Tartar, and Talos; these ships also contributed to both the outer and inner defenses of a Battle Group.

A less visible but equally important innovation of this period was the development and deployment of the Naval Tactical Data System (NTDS). NTDS was the first widely-used digital data link and it grew out of the need to integrate the Battle Group's air defense systems in a period when the speed and complexity of AAW operations had exceeded the capacity of voice radio links and yeomen with grease pencils writing backwards on glass tracking boards.

Thus began a classic measure/countermeasure race between Navy fleet air defense systems and Soviet antiship systems. Soviet ASMs grew faster and developed longer legs, forcing the Navy to further extend the outer rings of its Battle Groups' air defenses, and to improve its SAM-based inner rings. It was at this point that E-2 warning aircraft and F-4 interceptors armed with radar guided air-to-air missiles became the mainstay of the Battle

Group's outer ring of air defenses. The need to stand off from greater distances forced the Soviet Navy to improve its ocean surveillance and over-the-horizon targeting capabilities, which in turn led the Navy to place increasing emphasis on evading, spoofing, or destroying those systems.

This race abated somewhat during the Vietnam years when the Navy's Battle Groups were focused on power projection operations in Southeast Asia, but renewed with a vengeance during the third phase of Cold War AAW operations. The Navy emerged from the Vietnam years facing a Soviet Navy armed with a space-based ocean surveillance system that used radar and ELINT satellites to find and identify U.S. ships, and provide over-the-horizon targeting information to long range Soviet Naval Aviation (SNA) and nuclear powered cruise missile submarines (SSGNs). Launch platforms like the Backfire and the Oscar were armed with supersonic antiship missiles of 100-300 mile range. From this distance, SNA bombers and SSGNs sought to launch missiles from outside a Battle Group's outer defenses, thus saturating its inner defenses with multiple incoming missiles.

Out of this challenge grew the AAW posture designed to enable the forward Battle Group operations envisaged by the Maritime Strategy of the 1980s. E-2s and F-14s armed with long range Phoenix AAMs extended the Battle Group's outer ring. As important, aggressive efforts were mounted to provide strategic as well as tactical warning to the Battle Group of an impending SNA attack. Out of this particular initiative grew some of the first and most successful tactical exploitations of national capabilities (TENCAP), including a program which used missile early warning systems to detect and track the exhaust plumes of Soviet naval aviation aircraft in flight. Linked together by real time data links, these assets collectively extended the outer air battle hundreds of miles from the Battle Group, reestablishing a robust barrier that SNA needed to penetrate before it could launch its missiles.

At the same time, the Aegis weapon system was deployed during this period. Aegis vastly expanded the capabilities of the Navy's air defense cruisers to deal with antiship missiles that leaked through a Battle Group's outer ring. Its phased array radar could track hundreds rather than tens of targets simultaneously, and its target illuminators could guide up to 16 SAMs simultaneously, rather than one or two. Furthermore, because Soviet antiship missiles flew high altitude, arcing profiles in order to extend their range, Aegis could see them at great distances, and because of the speed with which Aegis could prosecute individual engagements, it could get off multiple shots against the same missile raid.

In addition to Aegis and the Outer Air Battle, the Navy aggressively pursued measures to counter Soviet ocean surveillance systems at the front end of the engagement cycle, as well as a panoply of close in systems designed to give each Battle Group combatant the ability to defend against antiship missiles in their terminal phase.

Soviet ocean surveillance systems, which by the 1970s included a substantial space-based component, provide an example of the kind of space capabilities that future adversaries might deploy. It's photo satellites, ELINT satellites, and radar satellites used technology that was quite advanced for the time, including systems designed to geolocate electronic emissions from space, and to use synthetic aperture techniques to distinguish between specific ship types. And the U.S. Navy's response to this system is also instructive, including a reporting system that

told ships when Soviet satellites were overhead, emission control tactics which denied ELINT satellites a signal to exploit, or false emitter tactics which put an emitter normally associated with a specific platform on a decoy platform.

One indication of the success of these countermeasures is the fact that the Soviets were never able to reduce their reliance on maritime patrol aircraft such as the Bear, which of course were quite vulnerable to a carrier's outer air defenses. It is important to keep this experience in mind for the future, because it demonstrates that the mere demonstration of space capability by a future opponent, even a very ambitious one like the Soviets deployed during the Cold War, will not necessarily translate into an effective ocean surveillance system

The Navy was also aggressive in improving terminal defenses during this period. In this category were systems like the Close In Weapons System (CIWS), a self-contained, radar-cued gatling gun designed to detect and attack incoming missiles automatically as they approached individual ships. Also, because Soviet antiship missiles were guided by small aperture radars in their terminal phase, decoys and jammers were deployed to either fool or blind those radars when they went active. In this context, the Navy also began to reduce the radar cross section of its ships, not to defeat Soviet surveillance efforts, but to enhance the effectiveness of decoys and jammers used against missile homing radars.

Antiship Missile Defense After the Cold War. In the new security environment, the AAW threat has changed in at least four basic ways. First, the days of large, saturation missile attacks launched at long range by platforms with an ocean-wide reach are over. In that sense, the antiship threat has declined dramatically. Second, on the other hand, the U.S. Navy aspires to a much more aggressive power projection posture than it did during the Cold War. For example, in today's security environment, in an analogue to what happened in the Pacific during WWII after the Japanese fleet was defeated, Battle Groups are expected to conduct protracted, high volume strike operations within 200 miles of an enemy coast. Surface combatants will be expected to provide naval surface fire support to engaged Marines ashore from just over the horizon of an enemy coastline. Third, for the foreseeable future, these operations will likely occur in crises or conflicts where there is a great asymmetry in the stakes in the outcome among the contestants favoring the United States' opponent. This will continue to make U.S. military and political leaders averse to human and material loss among its forces. And fourth, "export or die," post Cold War arms export markets will continue to provide potential U.S. opponents with modern sea skimming, antiship cruise missiles. A fifth change may involve the development of land-based ballistic missiles with an anti-ship capability.

This environment has already caused a fundamental shift in the Navy's AAW posture, and this posture will need to continue evolving to stay abreast of this threat. The essence of this threat today and in the near future is the specter of supersonic, sea skimming ASCM attacks in the littoral launched from truck-mounted launchers ashore, fast boats, or non-nuclear submarines which evade a Battle Group's ASW screen. Such attacks would give individual ship terminal defenses only minutes to detect and attack incoming missiles as they break the radar horizon at a distance of only 15-20 miles. This threat is already ubiquitous today in those operational scenarios where ships must

approach within line-of sight of a hostile coastline. Coming this close essentially solves the opponent's surveillance problem, and provides sufficient targeting information to launch truck-mounted, ASCMs down a bearing along which lies a U.S. surface combatant within 20-25 miles.

In order to extend this threat outward the 200-300 miles necessary to sharply limit Battle Group operations, the opponent will need to extend its view of the littoral battlespace by moving its surveillance assets upwards, and to extend the reach of its ASM platforms without thereby re-exposing them to a Battle Group's outer defenses. In assessing how potential opponents will grapple with this challenge, it is essential to be clear about the problems they will face.

The most important issue is the distinction between a wartime capability and one that functions effectively only in peacetime or a crisis. Wide area surveillance of the ocean surface requires putting sensors within relatively continuous line-of-sight of the area to be surveilled. In the case of any near term opponent, these sensors will need to be deployed in airspace that will be contested during a war. Certainly in the near term, the United States will win those contests when an opponent seeks to operate well outside its own airspace. Thus, it will be very difficult for some time for potential U.S. opponents to develop and deploy a robust, dedicated, ocean-wide or even littoral-wide surveillance system for use in wartime against U.S. naval forces.

Much more feasible is a system that seeks only to preserve the wartime reach of surveillance assets out to the "electronic horizon" of the littoral battlespace as viewed from the opponent's coastline. Depending on the range and elevation of the sensors used, the highly contested littoral battlespace in wartime would extend for at least 20-25 miles, and its outer limits would roughly correspond to the 200-300 mile radius limit for current, high volume carrier strike operations. Outside that radius, an opponent's view would be limited to peacetime or crisis operations in which vulnerable assets like long range patrol aircraft are able to operate because the rules of engagement do not allow U.S. attacks against them. This would enable an opponent to cue ASCM-equipped surface combatants with the speed and endurance to trail Battle Groups, providing a limited but potentially effective "first salvo" capability much like that pursued by otherwise vulnerable Soviet surface ships in the Mediterranean during the 1973 Yom Kippur War. But such a wide area system would not be effective against Battle Groups which survived or were not-exposed to the first salvo.

Inside a 200-300 mile radius, early in a conflict, Navy surface combatants will face the prospect of ASCM attacks launched from land, submarines, or small, fast boats, and cued by elevated, offboard sensors. The elevated offboard sensors, whether aircraft, UAVs, or aerostats, and their command, control, and processing facilities will be protected by modern, mobile SAMs able to reach some 50-100 miles outward from the opponent's coast, and at elevations of 50-60,000 feet, these sensors will have an horizon stretching some 200 miles. A further step upward in the opponent's anti-access capability will occur within 20-25 miles of its coast. Within this region of the littoral, an opponent's ASCM missiles will not need offboard cueing to be effective, and the opponent's ASCM launchers will be operating in a high clutter environment in which it will be much more difficult for the Battle Group to interdict or suppress these launchers before they launch their missiles. In this environment, extreme pressure will be placed on the inter-

mediate and terminal ASCM defenses of the ships comprising a Battle Group.

Thus, the near to mid term antiship missile defense challenge will likely resolve itself into three elements corresponding to the survivability of the opponent's surveillance capabilities: the opponent's peacetime surveillance system that gives extended reach but is vulnerable; it's extended littoral system which reaches out 200-300 miles and whose airborne sensors can survive as long as the modern, mobile SAMs that protect it remain unsuppressed; and its core wartime system which is limited to the 20-25 mile horizon from the opponent's own coastline.

It is important to note again that the most serious access challenge faced by the Navy in this area comes when it is playing the role of an enabling force for the other services. Thus, for example, Battle Groups standing off more than 300 hundred miles from an opponent's coast can still launch Tomahawk missiles and long range aircraft strikes essentially at will once an opponent's peacetime surveillance system has been destroyed, albeit at a lower sortie rate than when such operations are mounted over a shorter radius of operation. But naval combatants will have to close within 20-25 miles of a hostile shore to provide the naval fires that will enable ship to objective maneuver (STOM) by Marine Expeditionary Units (MEUs), and MEUs will often be the key to gaining access to the ports and airfields ashore that are necessary for reinforcing ground and air units.

Sensors, Weapons, and Networks for Gaining and Exploiting Access

The need to gain and exploit access in the new security environment will drive the Navy toward better sensors and weapons, and toward networks that link them together and process their output more effectively. There are both immediate opportunities in this regard, and opportunities which demand further development. This section will look at the immediate naval aviation opportunities in the Sea Shield mission area that are already being pursued.

Countering Submarines and Mines

The ASW and Mine Countermeasure problem in the littorals will always be difficult. But tremendous progress has been made in the ten years since the end of the Cold War on the main challenges in these areas. Compared to other warfare areas, ASW and MCM pose particular challenges in the areas of sensors and, to a slightly lesser extent, weapons. Networks are very important in ASW, but the networking technology needed is less demanding in many ways than the networking requirements in AAW. Networks are less important to MCM.

ASW Surveillance Sensors. The primary ASW challenge has always been wide area surveillance or search, and the main challenge initially posed by the new security environment in this mission area is a wide area search problem. Sound propagates better in deep water than in shallow water, and non-nuclear submarines can remain silent for extended periods when allowed to patrol small areas near their home ports at low speed. Using passive acoustics to search for such submarines is much more difficult than it was to search for relatively loud Soviet submarines operating in deep water during the Cold War. On the other hand, active sonars encounter serious problems with clutter in shallow water, much as early radars did when forced to look down at targets fly-

ing over land. And even in shallow water, the water column still remains relatively opaque to non-acoustic energy, limiting the role of RF and laser radars as long range sensors.

Three systems stand out as first steps toward regaining a wide area search capability in the littorals. The first is called the Advanced Deployable System (ADS); the second, derived from the Distant Thunder experiment, is Advanced Explosive Echo Ranging (AEER); and the third is the Low Frequency Active (LFA) variant of the Surveillance Towed Array Sensor System (SURTASS). ADS is a passive ocean bottom array that can be deployed by a surface ship, and whose output is currently collected and processed ashore via fiber-optic cable. AEER is primarily a signal processing adjunct to existing ASW combat systems, combined with coherent, air-droppable, active sound sources and a relatively simple data link that uses existing UHF radios on participating platforms. LFA is a very powerful, low frequency active sonar.

Unlike the Cold War Sound Surveillance System (SOSUS) arrays, which listened for low frequency, narrow band tonals propagating outward horizontally along the deep sound channel, nodes in an ADS array look upward along what is called the Reliable Acoustic Path (RAP). ADS is a derivative of the Cold War Fixed Distributed System (FDS) program, which was an attempt to repair the ASW barrier strategy by using many simple passive sensors in an upward looking array that used the reliable acoustic path (essentially the direct path) rather than the deep sound channel. Each sensor would cover a small cone of the ocean column, and fiber optic cable provided the bandwidth to network a vast array of these small sensors and bring their output ashore for processing. Future work on ADS will focus on deploying the arrays covertly via submarine, protecting the arrays from bottom trawlers, and using buoys to relay the array output directly to ships at sea.

AEER adds commercial off the shelf (COTS) processing to existing towed arrays on ships (and potentially, submarines) and air-deployed sonobuoys, and links the processors together using legacy radios with modems to form a network that can do bistatic or multistatic processing of the echoes from the air-dropped sound source. The essence of AEER is that it uses both spatial and temporal processing to extract a submarine's echo from the clutter and reverberation. Long wavelength towed arrays or directional sonobuoys use spatial processing to eliminate clutter and reverberation except on the azimuth of interest, and temporal processing allows reverberating echoes from the same object to be compared over time, thereby exploiting the fact that a submarine's echo loses less of its higher frequency spectrum in that time than do objects sitting on the bottom or floating on the surface.

One of the original concerns from Distant Thunder was that variations in bottom topography and content would interfere with AEER's temporal processing capability, but worldwide experiments have demonstrated excellent performance over a wide range of environments. Like all acoustic sensors, performance will vary in practice depending on many circumstances, but AEER will significant improve the detection ranges achieved using active sonar in the cluttered shallow water ASW environment. Another benefit of AEER is that it demonstrates long range performance under a wide variety of acoustic conditions, including the very common case in the littoral where sound is refracted away from the surface, a condition which drastically reduces the performance of a traditional, hull-mounted sonar. The main challenge facing AEER today is that its recognition differential is low when used by non-laboratory personnel.

Finally, both ADS and AEER are also great examples of the incredible power of networked sensors, and the relative ease of backfitting such a capability onto legacy platforms once the substantial initial challenge of developing the necessary signal processing algorithms is completed. Systems like AEER can be backfitted onto any towed array ship or submarine, and onto LAMPs helos and P-3s.

SURTASS was a late Cold War system that used towed passive acoustic arrays to supplement SOSUS coverage, or extend coverage into waters too shallow for a deep sound channel. SURTASS has been upgraded with better passive acoustic capabilities, and LFA is an additional update that adds an active, low frequency acoustic source, transforming SURTASS into an active sonar.

Specialized periscope or mast detection radars can also play an important role in the ASW search problem. Even during the Cold War, Soviet nuclear submarines regularly exposed a periscope when seeking a torpedo fire control solution against the fast ships of a Battle Group. And of course radar has an important role to play in preventing diesel submarines from snorkeling to recharge their batteries. Thus, a combination of speed, and radar deployed to search within the limiting lines of approach created by that speed, have always been an important ASW tactic against all submarines. Likewise, radar flooding in which a large area is flooded with RF energy so as to set off a submarine's radar warning alarm whenever it exposes a mast is also a traditional tactic against diesel submarines. But specialized mast detection radars like the APS-137 experience tremendous false alarm rates caused by both sea state and other floating objects and debris when their detection threshold is set low to maximize range.

The Automatic Radar Periscope Detection and Discrimination (ARPDD) program is developing the capability to process APS-137 returns in such a way as to allow very low detection thresholds (i.e. long range) and very low false alarm rates. Very impressive results have already been demonstrated in shipboard experiments, but unlike systems like AEER, ARPDD needs further development time to reduce the footprint of the massive processing capability it now requires before it can be deployed on surface ships, maritime patrol aircraft, or helicopters.

Maritime Patrol Aircraft and Helicopters. Long range, persistent, land-based maritime patrol aircraft provide the only way for a dominant naval power to maintain a continuous presence and surveillance throughout the vast ocean and littoral spaces over which it must exercise control. They often provide the most timely means of response, whether to a fleeting undersea acoustic contact, a report of a suspicious merchant ship, or an important signals intelligence collection opportunity.

In ASW, because sensor performance is limited, detection ranges are reduced, making wide area surveillance a more asset-intensive endeavor. Furthermore, all but the very quietest nuclear submarines, produce a continuous acoustic signature, whereas the best detection opportunities against non-nuclear submarines are both episodic and difficult to classify. On the other hand, there is a close correlation between the steps a submarine needs to take in order get into position to attack a target, and the operational indiscretions which provide the best detection opportunities for ASW forces. Therefore, contacts must be prosecuted and reliably classified as quickly as possible before they disappear back into the cluttered background as an unknown contact. This puts a premium on ASW platforms that can be deployed in numbers and distributed throughout the sea base, close a potential contact quickly, and deploy a

Among many roles, maritime patrol aircraft and multi-mission helicopters will provide the best means of making opposing submarines pay for their inevitable indiscretions.

menu of high quality acoustic and non-acoustic sensors to reacquire and identify the contact, classifying it as a false alarm, or trailing and/or attacking it.

Multimission maritime patrol aircraft and ship-based helicopters will play an increasingly important ASW role in the new security environment because they provide the best means of quickly responding to surveillance cues, especially when those cues do not include reliable classification of the target, as they often will not.

ASW Weapons. Torpedoes remain the primary ASW weapon in the littoral environment, although this environment also presents them with great challenges, particularly lightweight torpedoes, which are "fire and forget" weapons. Like all fire and forget weapons, the relatively small aperture and limited signal processing available to a lightweight torpedo's active seeker makes for problems in shallow water where there is a lot of clutter and the target is relatively small and moving slowly. The Mk. 50 modification to the Mk. 46 lightweight torpedo provides an initial response to this problem, and the more ambitious Mk. 54 a more robust response in a few years.

There is also an alternative ASW weapon opportunity that grows out of the intersection between MCM and ASW. One of the challenges in the organic MCM program is to do in stride mine neutralization and clearance from a helicopter, and the Rapid Airborne Mine Clearance System (RAMICS) program's approach to this problem may provide another ASW weapon opportunity as well. RAMICS is discussed in more detail below.

A Common ASW Operational Picture. One of the legacies of the formidable passive acoustic detection ranges possible in ASW during the Cold War is the tradition of relatively autonomous operation amongst the Navy's main ASW platforms. When the Soviet Navy finally deployed very quiet nuclear submarines near the end of the Cold War, the need for more coordination arose. Today, coordination is even more important, especially to give the ASW commander and all of his forces a wide area picture of the ASW battlefield. Such a picture would allow better utilization of multiple, often evanescent contacts against the same target produced by different sensors; it would give units knowledge of environmental conditions over a wide area, allowing them to better predict the performance of their sensors as they move about the battlefield; and it would identify resulting "holes" in ASW coverage where search assets could be concentrated efficiently.

Most of the individual pieces of work needed to accomplish this task are relatively simple, such as using common operational protocols when processing and communicating data, and using the same environmental models. But the task is complicated by the need to integrate these activities across many platforms.

MCM Sensors. As with ASW, sensor performance is central to success. And again, the beginning of the problem is always to detect and identify the mines in the first place. In the new security environment, this challenge is further complicated by the need to make such a mine hunting capability organic to the Navy's forward deployed Carrier Strike Groups, Expeditionary Strike Groups, and Submarines.

The key opportunities in this area being exploited today are very compact, imaging sonars and laser radars (LIDARS) able to detect and identify mines in the water column and on the bottom. Because these sensors can be made very

small, they can be deployed on or towed by smaller helicopters such as the CH-60; put on a surface ship-launched and controlled, semi-submersible vehicle; or even inside a torpedo-sized UUV launched and recovered from a submarine. Through the regular, peacetime employment of these sensors, the Navy can map the ocean bottom, particularly near key approaches or chokepoints. Doing so will facilitate the location of newly-placed mines, appearing as "deltas" from the peacetime picture, allowing forward deployed forces to rapidly focus on areas to avoid, or if they are critical, areas to clear. The unique advantage of the submarine-UUV combination is that this sensing can occur regularly without raising suspicion.

MCM Weapons. Once identified, mines need to be neutralized or destroyed. In many cases, the instruments that accomplish this purpose are not really weapons, but so called influence devices designed to create the signature needed to set off the mine in a way that does not destroy the mine sweeping platform. An influence sweep usually requires a platform that will not itself set off the mine, but which can tow a vehicle that will, hence the long tradition of relatively small, dedicated minesweeping ships with low magnetic and acoustic signatures. More recently, helicopters have been employed to tow influence sleds, but the size of the latter has required the towing services of heavy lift helicopters like the massive CH-53. Some of the same trends which will allow smaller MCM sensors will also allow smaller influence sleds, enabling an eventual transition to a CH-60 platform, and in turn allowing forward deployment on existing carriers, surface combatants, and amphibious ships.

In addition to influence sweeps, MCM forces also must have the ability to individually approach and remove or destroy all the mines it has found, because influence sweeps trade off speed for a reduced certainty that a minefield has been truly cleared. Here, one encounters perhaps the slowest and most labor intensive naval warfare area, in which today's dedicated MCM force utilizes explosive ordnance disposal (EOD) divers, marine mammal systems (MMS), and remotely operated underwater vehicles.

New approaches to this problem designed for use by organic MCM forces focus on helicopter-deployed systems. In the nearer term, a helicopter-delivered, remotely operated underwater vehicle will be deployed that can approach an already identified mine and explosively destroy both itself and the mine. In the longer term, the RAMICS system described above is being developed. RAMICS will combine a LIDAR and a Gatling gun firing supercavitating, 20mm projectiles. The LIDAR would be used to search for and identify mines, and the gun's projectiles would disable or neutralize it by penetrating the mine's shell and injecting a chemical initiator into it.

The MCM Network. Unlike sophisticated networks like AEER, and those that will be described below for AAW and strike warfare, the main network in MCM is human, and the center of this network is the dedicated MCM force. This is to say that even more than ASW, MCM success is not a science but an art that requires practice and extensive, detailed knowledge, and which is therefore extremely perishable. A dedicated MCM force is the home for this expertise, because it is the only place in the Navy where officers will do nothing but train for MCM, and where the intelligence on foreign mines will be sustained.

Also, the nature of the entire undersea warfare threat, and particularly the mine threat, is that its most challenging manifestations have primarily "purple" and "green" consequences. In other words, an aggressive, inshore mining campaign by an opponent will more directly impact the projection of Army and Marine Corps power than it will purely naval power, and even when the Navy does face a serious mine threat, it will usually arise when it is operating in direct support of the Marines, as in the NSFS mission. Combined with an aggressive organic MCM program, this might lead some to advocate the eventual dissolution of the dedicated MCM force for narrow budgetary purposes. A salutary warning of the likely consequences of such a decision is provided by the Air Force's decision after the Gulf War to retire its dedicated airborne electronic attack and air defense suppression assets in the belief that stealth would make such a dedicated force unnecessary.

Countering Missiles

Throughout the Cold War, the main AAW threat to U.S. Navy Battle Groups was the long range, air and submarined-launched, antiship missile. This threat presented itself at great distances from the Soviet homeland, and was supported by an ocean wide surveillance system. The seriousness of this threat provoked major attempts by the Navy to deal with it at every step in the engagement sequence. Efforts were mounted to defeat or fool the surveillance system, to attack the launch platforms before they could launch their weapons, to take multiple shots at the weapons themselves if they leaked through a battle group's outer defenses, and to defeat the weapon's seeker in the terminal phase with both active and passive countermeasures. All of these defensive measures required depth, and depth was naturally provided in this Cold War mission area by the great range at which Soviet sea denial operations against U.S. Battle Groups were mounted.

The main problem with the littoral AAW threat is that this depth is largely absent, both because the U.S. Navy seeks to close with its adversaries, and because those adversaries are generally constrained anyway to operations within the littoral battlespace. This means that an adversary's launch platforms will be buried in the clutter and noise of the littoral environment, either on land or in shallow inshore waters where it is easy for them to hide. It also means that the surveillance system that cues those launchers need not approach ocean-wide coverage, but rather must only aspire to cover a radius of several hundred miles outward from the coast. And finally, because ASCM weapon engagements will usually occur over an even shorter range within the contested littoral battlespace, the specific weapons used can be relatively short range, sea skimming missiles rather than the high arcing AS-6s and SS-N-19s of Cold War fame.

All of these factors conspire to radically compress an AAW engagement in space and time, reducing the role of the outer air battle, and reducing the number of shots available during the inner air battle. For the most serious sea skimming ASCM threats, launched from platforms that have successfully approached a Battle Group in the littoral clutter, the AAW engagement will begin when the attacking missile approaches the targeted ship's radar horizon—say 20 miles—and will be over, for better or worse, within one or two minutes.

Three interrelated steps are being taken to counter this threat. First, elevated sensors need to be developed which can eliminate or greatly reduce the clutter in the littoral environment which allows ASCM launchers to

hide, and which also prevents missile detection until the terminal phase of an engagement. Second, weapons need to be developed that can function in the same cluttered environment against small, fast targets. And third, these sensors and weapons need to be linked together in such a way as to allow an elevated sensor to provide the information needed for another platform to launch a defensive weapon against the incoming weapon from over the radar horizon.

Advanced Hawkeye Will Reduce Littoral Clutter. Central to the ASCM defense problem is a much better wide area picture of the littoral air space, particularly at the low altitudes relevant to the ASCM problem. The E-2 is the Navy's primary AAW surveillance system but it is not currently well equipped for this task. As a relatively low frequency, pure pulsed UHF radar, the existing E-2 APS-145 radar has tremendous difficulty detecting targets in the littoral for two basic reasons.

First, more than higher frequency, pulse doppler radars like that on the Air Force's E-3, the E-2 has trouble picking out low flying targets over land or among maritime clutter because it does not exploit Doppler signal processing. This was not a problem in a blue water environment because at UHF the sea surface is a mirror, but in the littoral or over land, clutter interferes with the ability to detect low flying targets. And even radars that do use Doppler signal processing have trouble with so-called low doppler targets. A low doppler target is one whose movement relative to the clutter background in the direction of the surveillance radar is low, either because the target is moving slowly in absolute terms, or because its direction of movement is perpendicular or nearly perpendicular to that of the radar's main beam. Historically, doppler signal processing in look down radars has been most effective against relatively high doppler targets, i.e. ones closing on the main beam of the radar at a relatively high rate. An ability to track low doppler targets in the littorals is critical because both surface ships and aircraft, as well as ASCMs, will often present themselves as low doppler targets.

Second, mechanically scanned UHF radars have inherently larger sidelobes than do higher frequency radars, which makes them more susceptible to both intentional jamming, and to inadvertent electromagnetic interference (EMI). EMI is particularly troublesome at the lower, roughly 400 MHz frequencies where the APS-145 operates because there are so many powerful commercial occupants near this band.

The radar on the Advanced Hawkeye will defeat these problems using two separate techniques. First, the APS-145 will be replaced by a digital, phased array radar called the ADS-18, whose 18 element array will allow electronic scanning over 160 degrees, and which will mechanically rotate to provide 360 degree coverage. The phased array antenna allows the radar to reduce its sidelobes electronically, significantly reducing the jamming and EMI problem. It also provides more gain in the main lobe, giving better detection ranges. Second, the ADS-18 will also allow temporal processing by providing three complete sets of measurements of the RF energy returning from a single spot, which will allow it to distinguish the moving target within the fixed clutter background of that spot because the target will move slightly during the interval between each of the three pulses.

ADS-18 will provide a quantum leap in the ability of the E-2 to detect ASCMs in the littoral environment, as well as a raft of other important targets. The next step is for the E-2 to provide its track information to shooters in the air and on surface ships in a way that maximizes their ability to shoot down the

missile. This can be done in three ways, roughly corresponding to degrees of both capability and risk, and the Cooperative Engagement Capability (CEC) is central to all three.

The Centrality of CEC on E-2. CEC is a very sophisticated data link that allows different platforms to share track information on targets with a speed and accuracy that allows one platform to shoot a weapon at a target that another is tracking. In practice, CEC enables both very accurate cueing, to provide warning to another platform that it is under attack by a target it cannot yet see, and to maximize that platform's radar energy management so that it can begin defending itself as soon as possible. More ambitiously, it allows for actual over the horizon engagements, where one platform launches a weapon that another guides to the target. In all cases, CEC extends the battlespace available to combat the ASCM threat, and this is particularly the case when CEC is combined with Advanced Hawkeye.

At a minimum, CEC can give warning to any ship with terminal ASCM defenses that it is going to come under attack from a very specific azimuth, allowing it to aim its ship self defense systems at that point on the horizon and to prepare to deploy decoys.

For ships with Standard missile or Sea Sparrow capability, CEC will provide cueing that allows search radars to focus their energy on the horizon, and will in some cases enable missile launch before the ASCM has broken the target ship's radar horizon.

Most ambitiously, when combined with the SM-2 Extended Range Antiaircraft Missile (ERAM), E-2/CEC will enable SM-2 intercepts out to 100 miles, even against low flying targets, at the very limits of the kinematic range of the interceptor. ERAM substitutes an active seeker based on the AMRAAM (but with a larger aperture) for the semi-active guidance of earlier SM-2s. This eliminates the need for an X-Band illuminator within line-of-sight of the target during the end game of the engagement. Instead, using track data provided by E-2, ERAM will allow engagements where not only is the intercept begun when the target is beyond line-of-sight of the launcher, but completed as well.

Active Electronically Scanned Antennas (AESA) and Overland Cruise Missile Defense. Just as cruise missiles pose serious threats to ships in the littoral, they also pose threats to targets ashore. Overland cruise missile defense presents all the problems described above, with the additional challenge that the endgame of the engagement is more challenging because small aperture AAMs have more difficulty locating and homing on cruise missiles against a ground clutter background than they do at sea. One element in the solution to this more challenging problem is the AESA radars that will soon be deployed on F-18E/F and, late in the decade, on F-35. Compared to mechanically scanned radars, AESA radars have much more capability against low cross section targets, both because they detect them earlier, and because they track them more accurately in azimuth and elevation. Earlier detection gives back battlespace, making for more favorable intercept geometries, while better tracking accuracy enables a fighter to guide AAMs like AMRAAM into smaller baskets within which their terminal seekers are more likely to acquire and home on the target.

Airborne electronic attack capabilities will grow, not decline in importance in the new security environment.

The Revolution in Sea Strike

Sensors, Weapons, and Networks for Dealing With Mobile Targets

FROM THE FIRST USE of military aviation during World War I until Desert Storm, it was common for air-dropped bombs to miss their targets by several thousand feet, and it therefore often took several hundred or even thousand strike sorties to destroy a single fixed target with conventional weapons. Precision weapons have quickly changed that equation so that now it is reasonable to expect a single strike sortie to destroy several targets. Furthermore, the leap from "clear weather" to "all weather" precision attack was essentially completed in the brief period between Desert Storm and Iraqi Freedom. This revolution in precision has solved the fixed target problem in the sense that no opponent of the U.S. can expect its fixed targets to survive long once they have been identified for attack. Of course, nothing about this revolution assures that the opponent's fixed targets will be discovered or correctly identified, not is it assured that effective attacks against fixed targets will have decisive effects, but the simple fact that once identified a fixed target can be quickly destroyed does represent a revolution in capability for U.S. air forces.

Both Enduring and Iraqi Freedom have also demonstrated the "solution" to another vexing problem for air forces; the ability for air forces to provide effective and timely support to ground forces engaged directly with opposing ground forces. Such close air support operations have always been bedeviled by the following kill chain requirement. The targets in question were small and mobile and could only be reliably identified by friendly ground forces in close contact with them; once targets were identified, it was difficult for friendly ground forces to mark them for supporting forces; and once targets were marked, it was difficult for supporting air forces to bring a weapon to bear that did not simultaneously threaten friendly and hostile forces. Fort McNair is only one of many monuments to the historic difficulty of closing this kill chain.

The keys to completing the solution to the close air support problem are three: the wide deployment among ground forces of both laser-based target markers/locators and ground controllers; and the provision of data links (rather than just voice links) between ground controllers and combat aircraft. The obstacles to this achievement are no longer technical, and the organizational/doctrinal obstacles that blocked progress in this mission area in the past seem to have faded. It is in this sense that one can say that the close air support problem is solved, and it is also in this sense that one can say that the mobile target problem is not solved.

Solving the mobile target problem will require the development and tight integration of new sensors, weapons, and networks for linking them together. These will be used to find, identify, track, locate, attack, and assess attacks against various types of mobile targets. The mobile targets of concern will include opposing armored units and their command posts, surface-to-surface and surface-to-air missile launching units, and leadership targets.

Sensors, Weapons, and Networks in the Littoral Battlespace

The littoral battlespace for strike operations against mobile/time critical targets will remain defined by the border between contested and uncontested air space, the maximum altitude at which combat operations occur, and the maximum range into contested air space that combat operations occur. Within that battlespace, naval strike forces will deploy platforms carrying sensors and weapons, both manned and unmanned. Some of these platforms will be multi-purpose, while others will have a single purpose; some will be autonomous, while others will be closely controlled from other platforms; some will need to send and receive large quantities of sensor data, while others will have less stringent need for high bandwidth connectivity; and finally, these platforms will vary in the degree to which they can survive independently in the face of opposing defenses.

The capabilities of sensors, and the platforms they are deployed on, have the largest impact on strike operations. The sensor platforms supporting future strike operations will primarily use radar, signals intelligence or SIGINT (defined here as the passive collection of either communications, or COMINT, or opposing radar emissions, or ELINT), and optics exploiting either visible light or infrared. They will generally be deployed on either satellites or airborne platforms.

Space-based sensors look down at their targets from low (~200 miles), medium (~11,000 miles), or high (~22,000 miles) orbits. Satellites in high orbits remain in the same position relative to the earth as it orbits, providing continuous coverage of the same quarter to a third of the earth's surface, meaning that three or four satellites can provide continuous global coverage. By contrast, satellites in low orbits move much faster relative to the earth's surface (90 minute versus 24 hour orbits). They see much less of the earth's surface at any one time, and it might take 12 hours for a single satellite to come within line-of-sight of the entire earth, but because they orbit so much closer to the earth, such satellites can deploy sensors whose power-aperture or resolution product is insufficient for deployment in high orbits.

Historically, satellites have been valuable sensor platforms, particularly for intelligence purposes, because they provide global coverage, at least intermittently, and because space was not seriously contested as a deployment medium during the Cold War. By contrast, airborne platforms must operate in or alongside contested air space in order to come within line-of-sight of their targets. In cases where enemy defenses have not been suppressed, and where airborne sensor platforms are vulnerable to attack, they must patrol outside contested airspace and look horizontally across the battlefield. The maximum theoretical detection range of almost any sensor under these circumstances is the distance to the horizon, or the distance of a tangent drawn between the platform and the earth's surface. For air breathing platforms, the maximum altitude of operation is about 60,000 feet, leading to a maximum detection range of about 250-300 miles.

This is significantly less field of view than even a satellite in low earth orbit, which can see upwards of 500 miles outward along either side of its ground track. But an airborne platform can orbit for many hours or even several days within line-of-sight of the same battlefield, whereas continuous coverage of the same battlefield from low earth orbit requires a constellation of many tens of satellites in order to ensure one will always be within line-of-sight. Alternatively, airborne platforms and satellites in high orbits both provide continuous coverage of the

same battlefield, but airborne platforms provide this coverage at distances that allow the use of sensors whose power-aperture product is inadequate for deployment in high orbiting satellites.

In reality, maximum and realistic detection ranges are almost always different. In the discussion that follows, I will look at the actual limits on the performance of sensor platforms, focusing on those deployed on airborne platforms. When the discussion shifts to how those sensors will be used in future sea strike operations, I will bring space-based sensor platforms back into the discussion.

As noted above, the maximum detection range of a sensor is determined by its distance from the horizon. In practice, maximum detection ranges often are considerably less because performance depends on the diameter of the sensor's antenna or aperture, and the power available to it – both of which consume weight and volume, which are always scarce on airborne platforms and extremely scarce on space-based platforms. Also, for a given wavelength, aperture size and design determines the angular resolution of the sensor, or the accuracy of the bearing to the target it provides. Again, for a given wavelength, passive sensors have longer range than active sensors, but on the other hand, active sensors can provide range to the target, whereas passive sensors provide only a bearing.

As a rule, SIGINT sensors are the only ones whose detection ranges will always extend out to the horizon, which on an airborne platform will be 200 or more miles. For radars and optical sensors, detection range will depend more heavily on aperture and power, and therefore on volume and weight. For example, JSTARS' radar has a maximum range of about 150 miles against a ground target, while Global Hawk's is more like 100 miles, and a current fighter's radar might have a maximum range of 50 miles. Passive optical sensors usually have shorter ranges than radars, with even the widest aperture airborne systems usually not exceeding 50 miles, and active optical sensors such as laser radars (LADARS) have the shortest ranges of all, with most airborne ladars limited to approximately 5 miles.

Measured in terms of resolution or accuracy, sensor performance is generally inverted, with very high frequency visible light sensors and LADARS providing the most accurate bearings and ranges, and the sharpest images, followed closely by IR sensors, with the performance of radars and SIGINT systems lagging behind in those performance metrics because of the much lower frequencies of the signals they exploit.

Finally, sensors vary according to how they are affected by weather, battlefield obscurants like smoke, ambient light levels, foliage, and physical obstructions like terrain and buildings. SIGINT sensors and radars are least affected by these factors, though the signals they collect can be blocked by major terrain obstructions, and radars do not penetrate foliage well. By contrast, optical sensors are dramatically affected by all or most of these factors. Neither IR or visible light sensors can see through cloud, and in addition, visible light systems are quickly blocked by obscurants such as smoke and become much less effective in low light conditions.

Within their detection ranges, and given the limits on their resolution, different sensors perform different functions. SIGINT sensors detect radio or radar transmitters and provide a line of bearing to them, with more sophisticated systems also providing an analysis of the signal that allows identification of the class of the emitter, and in some cases, the specific emitter itself. Such systems can also be used to develop an estimate of the location of the emitter by taking multi-

ple lines of bearing from several look angles, but these estimates do not provide a precise location (<100s of meters) because of limits on the angular resolution of the bearings.

Radars detect objects with sufficient reflectivity (radar cross section) to provide a detectable return. When airborne radars look down, terrain features provide a flood of returns that are difficult to distinguish from each other without specialized processing of the return signal. Today, in air-to-ground operations, the primary radar modes are synthetic aperture radar (SAR) and moving target indicator (MTI).

SAR uses the movement of the radar platform over time to create an artificially wide "aperture" or antenna that can be used to produce higher resolution images of a fixed target than could be produced using the real aperture of the platform's radar antenna. With SAR, a radar gains an imaging capability with resolutions that approach but do not equal those normally provided only at much higher optical wavelengths. By contrast, MTI exploits the relative movement of a moving target normal to the path of the radar platform. It does this by exploiting the fact that radar pulses reflected back from a target moving toward the radar have a higher, or doppler shifted, frequency than the pulses reflected from the stationary background around the target. With doppler signal processing, the radar can therefore be instructed to "see" only moving targets, and the background clutter can be filtered out.

When the two signal processing modes are combined, a SAR/MTI radar can detect and track moving vehicles over a wide area using the MTI mode, or provide high resolution images of a series of spots within that area. SAR/MTI radars do not yet interleave these two different modes rapidly enough such that a target detected using the MTI mode can be imaged using the SAR mode as soon as it stops moving, and then immediately picked back up on MTI once it starts moving again, but this capability will soon be deployed and will improve the ability of SAR/MTI radars to maintain continuous tracks of specific mobile targets.

Today, passive optical sensors are used primarily to collect very high resolution images or lower resolution video, and lasers are used primarily as range finders and target illuminators. In the not too distant future, LADARS will be deployed that will be able to "measure" targets very precisely in three dimensions, albeit at relatively short range.

One other metric related to sensors concerns varying demands on processing capabilities and downlinks. SIGINT sensors used in peacetime to collect and analyze signals require large amounts of signal processing. That processing can either be deployed on a large manned platform, such as an RC-135 or an EP-3, which is thereby able to perform its primary mission autonomously, or it can be separated from the sensor platform by a data link, allowing the use of smaller sensor platforms like U-2 or Global Hawk, which can fly higher and longer, but which must downlink their output to a command center with the requisite processing capabilities. The required data links must be wideband, with data rates measured in the multiple megabit/sec range, and can either be line-of-sight, to a command post within less than 200 miles or so, or via satellite, in which case they could link anywhere, albeit at lower data rates than the maximum available using line-of-sight links.

One of the main purposes behind peacetime SIGINT collection and analysis is to form libraries of the characteristics of communication and radar transmitters of interest. These libraries, or portions of them, can be loaded onto discs and carried on essentially any platform. During combat operations, SIGINT surveillance assets use these

libraries to generate immediate threat warnings of emitters within their field of view that require very little processing, and that can be broadcast on low bandwidth links. Today, large, sophisticated, intelligence platforms are generally used in this role, but not because of the need for their substantial processing or data link capabilities.

For example, assuming the provision of a digital threat library, a small UAV with a passive receiver, very little onboard processing capability, and a narrowband data link would be able to provide real time threat warnings of hostile radars within its field of view. These warnings would include a classification of the emitter type, the exact time of arrival and frequency of the intercepted signal, and a rough bearing to the emitter's location. Future networks of such UAVs might also provide immediate and precise location of threat emitters without significant additional processing or data link requirements, as I will discuss in more detail below. The point here is to note the dramatically different requirements for processing and data downlinks between tactical SIGINT surveillance operations that exploit already existing threat libraries, and the peacetime SIGINT collection operations that generate and maintain those libraries.

The requirements for processing and data downlinks for SAR/MTI radars are similar to those of peacetime SIGINT collectors. They can either be deployed on large manned platforms, where most of the processing is done onboard by human operators, such as on JSTARS, or on smaller manned or unmanned platforms such as U-2 and Global Hawk, which downlink their data continuously to ground-based processing centers. The processed radar output from a platform like JSTARS, consisting of SAR images and MTI tracks, can be transmitted using narrowband links, whereas the downlinks from the U-2 or Global Hawk radars are wideband, multi-megabit/second links.

Optical sensors generally do not require as much signal processing, but their requirements for data transmission vary according to the resolution desired, whether the imagery is still or video, and whether it needs to be transmitted in real time. Real time transmission of high resolution video requires enormous bandwidth, whereas still images and low resolution video can be transmitted in real time over narrowband links.

Quickly Detect, Identify, Track, Locate, and Assess Attacks on Mobile Targets

Successful attacks on mobile targets depend on networks of sensors that can quickly perform the tasks described above. Each step in this sequence creates different demands that will be reviewed below, but there are also some common challenges. First, mobile targets will gain significant operational sanctuary if sensor networks do not function in most weather conditions. This means either that sensors must be chosen that can operate through clouds, or that sensors must be deployed on platforms that operate beneath the weather. Second, sensor networks must either be designed to operate in the face of opposing defenses, or those defenses must be suppressed or destroyed, thereby creating sanctuaries from which sensor networks can safely operate. This tradeoff applies equally to the nodes of the network – the sensor platforms themselves – and to the data links that connect them.

The Effects of Weather and Defenses

As a rule of thumb, the altitude band between 15-20,000 feet is a good demarcation point

regarding both weather and opposing defenses. Above that altitude, sensor platforms will often be above significant cloud formations, but at the same time, they will also be at heights that force the opponent to use longer range air defense systems that must use radar for initial weapon cueing and guidance. Below that altitude, clouds will be less common, but at those lower altitudes, sensor platforms will face shorter range air defenses that can engage targets using only passive optical sensors, and are therefore essentially immune to suppression or destruction.

Of course, against a capable opponent whose longer range air defenses have not been suppressed or destroyed, there will be no operational sanctuary for airborne platforms above the battlefield. Under such circumstances, if sensor networks have to operate against unattrited defenses, sensor platforms will need to operate from distant horizontal standoff ranges or from space, or sensors will need to deploy on very stealthy platforms that can not be detected or targeted, or on platforms so cheap that redundant numbers can be deployed in a self forming and self healing network that can sustain losses and still function reliably.

Data links have different vulnerabilities compared to sensor platforms. Unlike radars, whose antennas are always designed to maximize transmission and reception efficiency in a specific direction, radio communication systems often use omnidirectional, or low gain antennas. Such antennas are smaller than high gain antennas and also eliminate the need for accurate pointing. Both of these characteristics make them useful for mobile platforms where weight and volume are at a premium. Even though such antennas are much less efficient, sufficient power can be generated except at the highest microwave frequencies to establish reliable one way links out to the horizon. Thus the ubiquity of military voice and data links at VHF and UHF using small, low gain antennas, particularly on aircraft and ground vehicles. The relatively large amount of power available at these lower frequencies also explains why satellite communications are possible at UHF with only slightly higher gain antennas that need only be pointed at the sky to transmit or receive, and can still be deployed on almost all platforms, albeit at much lower data rates than are available using line-of-sight links (50-100 kb/sec versus mb/sec).

The tradeoff that comes with depending on these simple VHF and UHF circuits is that they are inherently vulnerable to jamming by any transmitter within line-of-sight of the receiving antenna. VHF and UHF communication systems are not friendly to the two main methods for dealing with jamming – high gain antennas and spread spectrum or frequency agile waveforms – because both of these antijam methods are best implemented using much higher, microwave or millimeter wave frequencies.

High gain antennas with highly directional reception patterns "ignore" energy that is not in their main beam, greatly reducing the area from which an opposing jammer can insert spurious energy into the antenna. Special waveforms that rapidly vary transmission frequencies or use low power signals buried seemingly randomly in the background noise attack the jamming problem both by making communication signals covert and by making it difficult for the jammer to know what frequency to jam. Jam resistant communication systems depend on both these techniques, and both require the use of higher SHF or EHF frequencies because high gain antennas at UHF would be too large for mobile platforms, and because sophisticated, jam resistant waveforms are prodigious consumers of bandwidth that is scarce at UHF. And finally, when bandwidth

at the higher frequencies is used to buy jam resistance, data rates remain the same as on today's links – kb/sec using satellite links and mb/sec using line-of-sight links.

The issue of data link hardness and reliability is central because upon this question turns the design and viability of any concept of operation that includes a battlefield sensor network. Absent jam resistant data links, it will always be necessary to design in default modes where the network is unavailable and individual platforms must perform autonomously. In addition, assuming jam resistant data links are pursued, their development will need to be tightly integrated with the development of the sensor nodes in the network because of the impact that jam resistance has on data rates, particularly for satellite links. For example, the performance of unmanned or lightly manned sensor platforms like Global Hawk and U-2 that depend on multi mb/sec or even gb/sec satellite links if there are no ground-based command centers within line-of-sight will not degrade gracefully if forced to fall back on narrowband data links, whereas manned sensor platforms that do a lot of onboard processing of data could adapt much more easily to such an environment.

It is unlikely that jam resistant satellite communications will ever provide the same data rates available using line-of-sight links. The receiving antenna in a line-of-sight downlink will normally be able to shield its sidelobes from jamming signals, whereas receive antennas on high orbit satellites are much more exposed. This means that jam resistant satellite links will always depend on a high degree of spectrum spreading that comes at the expense of data rate. Today, when sensor platforms use satellite links, they use commercial links designed to maximize data rates that have no jam resistance.

A possible solution to this problem would be a laser satellite communication system, using laser uplinks and downlinks. Lasers suffer from significant propagation losses in the atmosphere, but so much bandwidth is available at optical wavelengths that it may still prove possible to provide extremely high data rates. At the same time, because of the extremely high frequencies involved, the beams produced are extremely narrow, which means that even a receive antenna in high orbit might be able to null signals emanating from outside a narrow cone surrounding the legitimate transmitter. Absent the successful development and deployment of such a system, satellite links will always present a harsher tradeoff between data rate and jam resistance than will line-of-sight links.

Assessing the potential physical vulnerability of space-based sensors and communication satellites is a special case. Heretofore, satellites have enjoyed a virtual sanctuary from attack. No war has ever been fought between a country with satellites and another with anti-satellites, never mind a war between two countries with both capabilities. Though both superpowers deployed direct ascent anti-satellite systems during the Cold War, neither chose to use them against the other in peacetime, and the absence of war between the superpowers leaves open the question of whether these systems would have been used in a war. In any case, neither side's anti-satellite systems were very extensive, nor did they have the capability to reach beyond low earth orbit.

Before attempting to judge the future relevance of this complicated issue, one needs to ask what role satellite-based sensors or communications systems are likely to play in the future in finding, identifying, tracking, locating, or assessing attacks against mobile targets.

Surveillance

Surveillance systems are characterized by large fields of view and persistence. Ideally, a surveillance system can continuously monitor the entire battlefield and pick out targets of potential interest with a low false alarm rate. But even the best surveillance systems gain this capability at the expense of the ability to identify, track, precisely locate, and/or assess attacks against those potential targets. Thus, historically at least, surveillance systems serve a cueing function for other sensors which complete the kill chain.

The surveillance challenge against an enemy's forces in the field largely boils down to the problem of detecting vehicles. This is obviously a difficult problem, both because military vehicles are hard to distinguish from other vehicles, and because all vehicles are difficult if not impossible to detect from a distance if they are not moving, emitting a signal, or launching a weapon.

Moving vehicles can be detected by MTI radars in all weather at ranges that depend on the altitude and aperture of the sensor. Future airborne surveillance platforms modeled on today's JSTARS, U-2, and Global Hawk will be able to detect targets out to 150 miles or more. On the other hand, faced with reasonably advanced air defenses, such platforms need to standoff some 100 miles until those defenses are destroyed.

The desire to escape this tradeoff is one reason why a space-based radar program is being pursued. In principal, a constellation of such satellites could be deployed that would provide continuous coverage of the earth within its orbital planes. Deployed in low orbit, such a constellation would need to include 20-40 satellites to ensure that one is always above the horizon. The expense of such a constellation has led to exploration of the alternative of deploying space-based radars in medium orbits, where many fewer satellites would provide continuous coverage, but where power/aperture products would need to be much greater. A third and technically more challenging alternative has therefore emerged envisioning a bistatic or multistatic system in which a satellite in medium orbit serves only as the transmitter in a network in which an airborne platform or platforms deployed within line-of-sight of the area of interest serve as the receiver.

The theoretical advantages of such an architecture are several. First, it would reduce the cost of achieving continuous coverage from the sanctuary of space without forcing the satellite designer into a power/aperture race that might not be winnable, particularly when using the MTI mode. Second, despite this concept's continuing dependence on airborne platforms, it would not be as vulnerable to opposing defenses as a purely airborne system because the airborne platforms in question would be passive receivers only, and therefore, in principal, highly stealthy and able to operate deep within rather than alongside contested air space. Third, if implemented as a multistatic system with several airborne platforms within line-of-sight of the same area of interest, such a concept would also enable the precise tracking of mobile targets by exploiting trilateration, or the reduction in angular resolution errors of single platforms by calculating the intersection of the error ellipses of two or three widely separated platforms tracking the same target.

One of the challenges with space-based radar is that it will be an expensive system, meaning that it will need to accommodate the requirements of both the Intelligence community and the Department of Defense, but these requirements will be difficult to reconcile. The Intelligence community's prime interest is in a system that provides global coverage in peacetime but is less

determined than DOD that that coverage be continuous, whereas DOD is most interested in continuity of coverage and might be willing to sacrifice the loss of peacetime access to deep inland areas that would result from adoption of a multistatic system.

Vehicles can also be detected when they use a radio or a radar. SIGINT systems designed to detect these signals have longer detection ranges than radars because the signals they collect are more powerful, not having suffered the attenuation of a two way trip in which the power of the signal falls as the fourth power of the distance traveled. Thus, for example, airborne SIGINT collectors such as RC-135, EP-3, and Guardrail can see further into contested air space than can airborne radars.

At the same time, the signals that SIGINT sensors collect are often highly directional. This applies with special force to ELINT sensors that collect radar signals. Since a radar's main beam sends out a much more powerful signal than do its sidelobes, and since that main beam is usually aimed at the horizon and scanned in both azimuth and elevation, the probability of detecting the radar's signal can vary significantly depending upon where the ELINT receiver is.

Thus, airborne ELINT sensors have a major advantage over satellite-based ELINT sensors when used in a tactical setting. First, and most generally, an air defense radar's main beam will rarely be aimed directly at a satellite because the air defense radar will be oriented horizontally to the horizon, whereas even a low orbit satellite will normally be looking directly down. This means that space-based ELINT systems must be capable of detecting a radar's sidelobes if they are to routinely and reliably detect it. Airborne ELINT systems have a double advantage in that it is much easier for them to place their sensors in the path of a radar's main beam, and, because they are closer, it is easier for them to detect a radar's sidelobes.

Like radars, SIGINT sensors provide a line of bearing to their targets which have errors in angular resolution of one, several, or tens of degrees, depending on the quality of the receiving antenna and the frequency of the signal, but unlike radars, a SIGINT receiver can not by itself provide a precise range to the target because it is passive. On the other hand, because most emitting targets are stationary while they emit, a SIGINT platform can calculate an estimated range to the target over time by taking multiple, separate bearings on the same signal. But in no case would the resulting position estimate have an error ellipse with a radius less than 100s of feet.

Finally, a vehicle can also expose itself to detection by launching a weapon, particularly a missile or a shell. Radars can sometimes detect such weapons soon after they take flight, and if they are flying anything approaching a ballistic trajectory, an estimate of their launch point can be quickly calculated. When a vehicle launches a weapon with a strong IR signature that lasts for more than a few seconds, it can often be detected and tracked by missile warning satellites in high orbits, both giving warning to those in the estimated impact area, and allowing a rough estimate of the launch point, and therefore of the launch vehicle's location if it has not moved. In principal, IR sensors could be deployed on airborne platforms and used for similar purposes.

In no case can an individual surveillance sensor complete the kill chain against a vehicle using the same methods it used to detect it. None of the sensor modes described above can precisely locate a vehicle; instead, they provide estimates of location with a radius of uncertainty measured in 100s or 1000s of feet. An MTI radar on

an airborne platform can provide a continuous track of a moving vehicle within its field of view, but a space-based system can only do so if its constellation is designed always to have a satellite in view. Also, MTI radars can not normally classify their targets, at least beyond such basic distinctions as between a tracked or a wheeled vehicle. This may change with the further development and deployment of MTI radars that can image moving vehicles when those vehicles are generating rotational movement relative to the radar, i.e. when rounding a turn on a road, but the images generated will not be high resolution. SIGINT sensors, on the other hand, will be able to classify their targets immediately upon detecting them because of the wide deployment of already-developed specific emitter identification capabilities. Finally, none of the sensor modes described can reliably assess the results of the attacks that result from their initial detections. Solutions to the rest of the kill chain will depend on some combination of networking among surveillance sensors and the introduction of imaging sensors.

Geolocation

Even during the latter part of the Cold War, target location errors measured in the thousands of feet were acceptable because an attacking aircraft only required knowledge of a target's position relative to it, not the target's absolute position. Manned aircraft used their navigation systems to fly to within line-of-sight of the estimated location of the target. As long as the accumulated error in estimated target location and in navigation system performance was less than the field of view of the aircrew, they had a good chance of acquiring the target visually or with their radar. Once acquired in this way, the target location error was irrelevant because the accuracy of the attack depended on how accurately the attacking aircraft could calculate its position and rate of closure relative to the target, not that target's absolute position.

This did not change with the early generation of precision weapons. Weapons like Paveway 1, Walleye, or Maverick did not need to know the absolute position of their targets. Rather, they went where they were told to go by the pilot, who still found the target using methods basically similar to those described above. It is the introduction of GPS-guided weapons that has driven the need for precise geolocation of targets, with errors measured in the 10s of feet rather than 100s or 1000s, and future weapons like small diameter bomb may create the demand for even greater precision.

There are two fundamental approaches to precise geolocation: one involves using networks to compensate for the inaccuracies of individual sensors operating autonomously, while the other compares images of a target collected by a single platform to a database of geo-registered imagery (or precise terrain data).

Networks can be used to compensate for the current inaccuracies of both SIGINT and MTI platforms. For example, in the case of ELINT, if three ELINT antennas are deployed within line-of-sight of a radar and data-linked together, they can be used to measure the time of arrival of a single radar pulse at three widely separate locations. With three receivers, there are three separate pairs of receivers and the output of each pair can be used to form a hyperbola of uncertainty along which the emitter lies. With three hyperbolae, there is only one point at which all three intersect and that point can be located with accuracy measured in 10s of feet. More ambitiously, if only two receivers are available, and at least one is moving relative to the emitter, both the time of arrival

and the precise frequency of the received signal can be measured. The difference in the signal's time of arrival at the two platforms can be used to form one hyperbola of uncertainty, while the difference in the doppler shift of the signal at the two platforms can be derived to form a curved line intersecting with the first hyperbola at only one point.

ELINT or COMINT networks using time difference of arrival or frequency difference of arrival (TDOA/FDOA) signal processing do not require high data rate links, but they depend on very low latencies, and the performance of these networks varies according to the geometry of the sensors relative to the emitter and the distance of the sensors from the emitter. The ideal geometry has the sensors relatively close to the emitter and essentially surrounding it.

One of the great strengths of these networks is that they do not require sophisticated, large aperture antennas. This means that there is great flexibility in designing the nodes of the network. For example, where groups of aircraft are already present on the battlefield for other reasons, it will be possible to turn them into a TDOA/FDOA network via Link 16 or its successor using their Radar Warning Receivers (RWR) as the network nodes. In addition, because the antennas can be so small, TDOA/FDOA nodes could be deployed on UAVs much smaller than those required to deploy radars. Thus, because the nodes are passive, and because they can be deployed on very small platforms, a small UAV-based ELINT network might prove to be the best means of targeting the radars of advanced air defense systems. Even more ambitious would be an ELINT network whose nodes were covertly placed unattended ground sensors.

Networks can also be used to compensate for MTI radar errors. If three widely separated radars are deployed within line-of-sight of a moving target and data-linked together, the error in target location caused by errors in azimuth resolution can be greatly reduced by fusing the error ellipses generated by each radar and finding their intersection. This allows moving targets to be tracked by MTI radars with an accuracy approaching that needed to target a GPS-guided weapon. As with the SIGINT networks described above, the area in which a network of MTI radars can precisely locate moving targets is limited by the area in which the coverage of all three radars overlap, and by the geometry of the radars relative to the targets within that area. Thus, in both cases, the performance of the network degrades if its nodes are forced to standoff from contested air space. For example, when standing off in line abreast formation networked to precisely locate moving targets, three widely separated MTI radars able by themselves to see 150 miles into contested air space might collectively cover an area roughly 50 miles deep, and their line abreast formation relative to the targets would also reduce the network's location accuracy.

This point illustrates one significant difference between SIGINT and MTI radar networks when used to locate ground targets in contested air space. MTI radars put much greater demands on their platforms for power and volume than do passive SIGINT sensors, which means that it should be much easier to create SIGINT and especially ELINT networks that use small, stealthy, and persistent platforms that can penetrate and operate within the coverage area of unattrited, advanced air defenses.

The other general method of obtaining precise target locations is to collect images of the potential target, compare them to an imagery database which is already geo-registered, and match the images collected to the database. Databases now exist that provide

precisely registered optical imagery and terrain elevation data to which can be compared images collected on the battlefield by cameras and SAR radars respectively. The term of art for this process is called mensuration.

Today, mensurated optical and SAR images are produced on the ground after airborne platforms like U-2 or Global Hawk or other national systems have downlinked raw, very wide bandwidth data to ground stations where the data is processed, or they are produced aboard larger, manned platforms like JSTARS and the Navy's experimental "Hairy Buffalo" aircraft, which can mensurate SAR images on board. Either approach today can require more than an hour to get targeting information from the sensor platform to the weapon platform.

Major efforts are already underway to reduce the time late associated with target mensuration in order to speed up the process of striking mobile targets while they are at rest. One initiative is to give strike fighters and bombers an organic capability to mensurate images found by their own sensors. Another more ambitious goal is to create airborne networks in which UAVs or UCAVs can downlink optical and SAR images by line-of-sight links to manned platforms capable of doing the mensuration. The manned aircraft could be command posts like the Air Force's MC2A or the Army-Navy Aerial Common Sensor or individual strike fighters like F-35. In either case, the objective will be to locate mobile targets precisely in real time so that once found and identified they can be struck before they move again.

Classification and Damage Assessment

Imaging sensors are one of several means of obtaining precise target locations, but they are often the only means of classifying or identifying targets before attack, and of assessing the results of an attack. Furthermore, depending on the rules of engagement and on the future trends in technology, it may be the case that many or most targets will continue to require imaging sensors for both classification and damage assessment, as they almost always do today.

Certainly, there are instances on the battlefield when classification and/or damage assessment can be gained by other means. For example, as was discussed above, SIGINT sensors can classify emitters without needing an image, though they can not perform damage assessment beyond noting that an emitter has gone off the air. Some kinds of mobile targets essentially classify themselves by moving en masse in response to battlefield events, such as an armored unit moving to (or retreating from) an engagement. Thus, in the case of MTI radars, targets or groups of targets can be sometimes classified simply by the fact that they are where they are when they are doing what they are. The same caveats apply to damage assessment, where, for example, there will always be occasions in which a target's destruction is easy to determine simply because of secondary explosions that can be seen from great distances.

But the fact remains that a significant number of mobile targets will likely remain which, by dint of their proximity to friendly troops or sensitive non-military facilities, their high value, or some other factor, will need to be precisely identified using high resolution imagery before being attacked, and an assessment of those attacks will often need to be performed using the same means.

Target classification and damage assessment are important for other reasons. Absent a process that leads to fairly rapid classification, surveillance sensors quickly become swamped because they detect many more

potential targets than they can track. In addition, absent reliable classification, false targets are generated, wasting valuable tracking time and causing the diversion of weapons and weapon platforms. Also, when reliable damage assessment is not available, multiple weapons must be assigned to high value targets to compensate for the possibility that one might malfunction. And in the particular case of attacks against opposing defenses, a lack of reliable damage assessment prevents U.S. forces from exploiting successful attacks. This is a particular problem in the case of attacks against air defense radars where, absent damage assessment, a successful attack is indistinguishable from a successful shut down by the opponent.

Beyond the common need for imagery, classification and damage assessment may require different degrees of resolution. Whereas 1 meter resolution might be sufficient to distinguish military from commercial vehicles, and perhaps even certain types of military vehicles from each other, 1 foot resolution might be required to distinguish friendly tanks from hostile tanks, or to distinguish different variants of the same basic prime mover, such as the missile TEL, radar, and command post vehicles within a SAM-10 battery.

At the same time, determining damage to a target can require even higher resolutions, measured in inches, especially when vehicles are attacked by penetrating weapons that leave only a small hole, or by sub-munitions which sand blast the outside of the vehicle but leave its basic structure intact. On the other hand, to avoid the need for such high resolutions, damage assessment can be performed probabilistically using much lower resolution images. For example, following the model anticipated for AARGM/Quickbolt, any GPS/INS weapon with a terminal seeker could be programmed immediately prior to detonation to relay its position, a health and status message, and an image of the target area via line-of-sight relay back to its launch platform. In the case of attacks against mobile targets, the purpose of the image would be to simply confirm or deny the presence of the vehicle in the target area, and its resolution could therefore be quite modest.

Today, imagery of relatively high resolution is routinely generated by optical, infrared, and SAR sensors, but the performance of these various sensors varies widely according to their maximum range and resolution; their performance at night, in bad weather, and in the presence of battlefield obscurants; the weight, volume, and power requirements they impose on their platforms; and their technical maturity.

Optical sensors have the best resolution, they can be given good detection ranges when provided modest aperture, and the technology of electro-optics is advancing rapidly, making the processing, storage, and transmission of optical images easier by the day. But optical sensors are shut down completely by weather and some important battlefield obscurants, and are greatly degraded at night. IR imaging systems are approaching optical systems in their range and resolution, which has had a dramatic impact on strike operations at night, but IIR systems are also shut down by weather. Thus, as tools of classification and damage assessment, EO/IR sensors must be deployable "under the weather" if they are to be routinely available. This reduces their field of view, increasing the number of platforms needed for wide area coverage, and also exposes those platforms to unsuppressed, short range air defenses.

These constraints already have led to efforts to improve the resolution of SAR radars, which are "through the weather" sys-

tems. Today, SAR imagery of 1 foot resolution and slant range of more than 100 miles is generated by a variety of platforms, and imagery with 4 inch resolution has been demonstrated, albeit from a slant range of only 25 miles. It is important to note that the resolution of SAR radar imagery is determined less by its angular resolution, which is not limited by its real aperture, as in other radar modes, but by its range resolution. Range resolution, in turn, is limited today by atmospheric distortion of individual radar pulses, the degree of distortion varying directly with slant range. Thus, with respect to high resolution SAR imagery, and unlike MTI radar, standoff platforms can not use aperture to compensate for distance, meaning that smaller penetrating platforms with smaller apertures will have an advantage in resolution.

Finally, work is being done to give MTI radars a better capability to identify moving targets. One technique exploits the excellent range resolution of these radars to form a crude image of the moving target, a technique that is also being developed for advanced air-to-air radars. A second technique inverts the SAR radar's normal routine by using the target's motion to exploit the doppler effect of an object rotating relative to the radar, and to again provide an image of a moving target, a technique long used by the Navy to identify ships at sea. In neither case are these techniques expected to produce truly high resolution images, but they will prove useful as a means of culling potential targets, reducing the number of tracks that need to be maintained before the target can be further classified using other means.

Weapons

For obvious reasons, the appropriate weapon for attacking a mobile target will depend on the capability of the both the supporting sensor network, and on opposing defenses. For example, during World War II, air forces gained a mobile target capability when they gained sufficient air superiority over the enemy's fighters to fly patrols over the opponent's army in the field. When that army was forced into large scale maneuver by friendly ground forces it filled up the local road network and aircraft flying very low and using short range cannon fire, bombs, and unguided rockets could relatively easily find and attack such columns. Deployed in the numbers and with the degree of air superiority achieved by the Allies in the campaign to liberate France, tactical air forces wrought havoc on the German Army at places like Mortain, where a German armored counterattack was stopped from the air, and the Argentan-Falaise gap, where a German Army was decimated as it sought to escape encirclement by allied armies after they had broken out of Normandy.

The wide deployment by the early 1970s of vehicle-mounted, radar-guided, 20 and 40mm AAA and hand held, heat seeking (infrared or IR), surface-to-air missiles (SAMs) greatly raised the cost of all low altitude attacks, whether against fixed or mobile targets, and drove attacking aircraft to higher altitudes. At these altitudes, smaller, mobile targets were much more difficult to find and identify, even when they were concentrated. Equally important, bombing accuracies from high altitudes had not improved much since World War II, and were still often measured in thousands of feet. At the same time, radar-guided SAMs were also eliminating the relative sanctuary heretofore provided at these higher altitudes from ground-based air defenses. An operational crisis resulted, experienced by both the American and Israeli Air Forces in their wars of this period, in which bombing effectiveness against both fixed and

mobile targets was low, and even high altitude operations faced serious opposition from ground-based defenses. This crisis led to major investments in technologies for suppressing radar-guided SAMs, and for increasing the precision of weapons dropped from medium to high altitudes.

Suppressing radar-guided SAMs would reestablish a high altitude sanctuary from ground-based defenses, while precision weapons would greatly increase the lethality and effectiveness of air operations from such altitudes. SAM suppression came to depend on specialized aircraft known as Wild Weasels, equipped with sophisticated, passive, direction-finding avionics which could identify and locate SAM radar emissions, and armed with high speed, antiradiation missiles (HARMs) which, once fired at an active radar, would either home on its emissions and destroy it, or force it to shut down, causing the SAM it was guiding to go ballistic and miss its target. The main method of increasing the lethality of bombing operations was the development of the laser-guided bomb (LGB). Aircraft equipped with a laser illuminator could drop bombs from high altitude that would home on laser energy reflected from the target. This greatly increased the accuracy of bombing attacks, now measured in tens of feet, and also made accuracy relatively insensitive to altitude, allowing effective operation from the high altitude sanctuary established by the suppression of opposing radar-guided SAMs. But this development did not address the problem of finding mobile targets from medium and high altitudes and was relevant mostly for attacks against fixed targets.

First generation LGBs were day/clear weather systems, and were used only in the latter part of Vietnam after the Air Force and the Navy experienced repeated failure in attacking high value fixed targets around Hanoi. Post-Vietnam development of laser-guided weapons was dominated by the fact that Europe, like Vietnam during the monsoon, usually had dense cloud cover, blocking the use of LGBs from medium altitude, and that Soviet radar-guided air defenses – whether mobile or fixed – were so dense that there would be no sanctuary at medium altitude. This led to the development of forward looking infrared (FLIR) sensors that would allow low altitude operation at night, putting the laser illuminator under the weather and the aircraft under the radar horizon, and also reducing to some degree the exposure to optically-guided short range air defenses. But flying low and fast at night, a fighter with a LANTIRN pod and LGBs would still have had only modest capabilities to find its own mobile targets. This second generation capability was not demonstrated on a large scale until Desert Storm, where it was used from medium altitude against a totally suppressed air defense system against both fixed and mobile targets.

The wide deployment after Desert Storm of FLIR/laser illumination pods in both the Air Force and the Navy greatly increased the percentage of the force with such night/clear weather precision strike capabilities against fixed targets and mobile targets with a clear IR contrast relative to their surroundings. But even over the deserts of Iraq and Kuwait, the need for clear weather proved troublesome, and it proved crippling at times in the cloudy climate typical of Serbia and Kosovo, a characteristic obtaining throughout the temperate zones of the world, including the entire Asian littoral.

The solution to this problem was weapons that integrate GPS and inertial navigation systems (INS). Integrated GPS/INS provides a through the cloud, weapon guidance capability that is compact, relatively cheap, and which can be made

robust against countermeasures. On the other hand, unlike LGBs, platforms carrying GPS-guided weapons are still generally unable to geolocate targets with their own sensors with the precision needed for an organic, closed loop targeting system. GPS weapons have played a large role in both Enduring and Iraqi Freedom, but when used against mobile targets the coordinates for their targets were provided by other platforms, often an individual on the ground within line-of-sight to the target.

Weapons vary according to the range and speed of their delivery. Standoff range is useful when opposing defenses are not suppressed, and speed of flight will determine the time lag between weapon launch and impact. GPS or laser-guided gravity bombs have ranges of at most 10-15 miles, well within the range of most radar-guided air defense systems. Glide bombs like JSOW can boost weapon range to about 40 miles, which may be far enough to protect stealth aircraft from a SAM-10, but in that case they must be carried internally. This is not a problem for B-2, but internal carriage becomes a real constraint with aircraft like the F-22, whose weapon bays are small. Hence the small diameter bomb program, which is a GPS-guided glide bomb with a maximum range of 40 miles, a 250 lb. payload, and terminal seeker. At 40 miles range, the time-to-target for a glide bomb is at least 8 minutes.

Current rocket-propelled standoff weapons like AGM-130 and 142 can deliver 500-1000 lb. payloads with less time-to-target over the same distance as a glide bomb, but these weapons are too large for internal carriage by any stealth aircraft, while at the same time they do not provide enough standoff for non-stealthy aircraft to use them against targets defended by double digit SAMs.

The HARM antiradiation missile is a special case in this category. It is faster than the other missiles, but carries a very small warhead. But like other air-launched rockets, it too is currently too large to fit internally on stealth fighters, and lacks the range to be used by non-stealthy aircraft against double digit SAMs. One element of the AARGM program is to develop a new motor for HARM which would increase both speed and range, while at the same time allowing internal carriage.

Air-launched jet-propelled cruise missiles like JASSM and SLAM are small enough to be carried externally by non-stealthy fighters and can reach out to about 150-200 miles carrying 500 lb. warheads. But at those ranges, which are necessary if the launch platforms are not stealthy, flying subsonically at less than 10 miles a minute, the weapon takes at least 15-20 minutes to reach the target.

Cruise or ballistic missiles can also be launched from over the horizon by surface ships or submarines, or from friendly territory. Standard 21" diameter and 20' long launchers can carry ballistic missiles that come in variants with ranges from 100-300 miles and payloads varying from 1000 to 400 lbs., or cruise missiles that can carry a 1000 lb. payload to a maximum range of 1000 miles.

Weapons also vary in other respects. Gravity bombs like Paveway and JDAM cost as little as $20,000 apiece, whereas the other weapons described above range in unit cost from $300,000 for the baseline JSOW upward. In principal, any air-launched gravity, glide, or rocket- propelled weapon could be targeted by its launch platform, whereas beyond about 40 mile range, offboard targeting usually becomes necessary. For targeting fixed targets or mobile targets at rest, weapons require either a laser spot to home on, a precise GPS coordinate, an image of the target and a terminal seeker with scene matching capability, or a terminal seeker and a data link back to the launcher that allows

the launcher's aircrew to do the scene matching. Traditionally, weapons designed to attack moving targets have used an optical seeker with a tracker and have relied on the pilot to designate the target and lock on the seeker before launch.

Future weapon developments will include new kinds of terminal seekers, smarter sub-munitions, and better scene matching or target recognition algorithms. Today's terminal seekers rely on scene matching and are usually passive, but future terminal seekers may use active millimeter wave or laser seekers more capable of recognizing targets themselves, potentially eliminating the need for a prior image of the target. Active seekers, if made small enough, can be used in smart sub-munitions which, if used against dense target arrays, could produce multiple kills from a single weapon, and could also significantly reduce the demands on targeting networks, particularly in regard to target geolocation accuracy.

At the same time, scene matching weapons could also be improved. Today, the terminal seekers on scene matching weapons usually collect IR or visible light images and compare them to templates generated using the same phenomenology. So an IR seeker needs an IR image as a template, or an optical seeker needs an EO image. In principal, the launch platform could generate these templates but only in clear weather or under the weather. Future scene matching weapons may retain their relatively simple, passive, EO/IR seekers, but use scene matching algorithms that allow target templates to be developed by either visible, IR, or SAR images. For example, this would allow a strike fighter to use its SAR to target a cheap gravity or glide weapon through the weather and in real time.

In addition to reducing the need for networks to produce precise target location information, smarter sub-munitions and better scene matchers would also reduce the current dependence on GPS for weapon guidance and possibly give weapons an equal capability against stationary or moving targets. In both cases, INS systems would put the weapon inside its terminal seeker's basket, with smart sub-munitions probably costing more, but also providing a much larger basket, whereas smarter scene matchers would likely remain very cheap but also dependent on their launch platform for a more accurate initial target location. Smart sub-munitions would be most compatible with standoff weapons for use in higher threat environments, while smart scene matchers would be most relevant when strike fighters have gained a medium altitude sanctuary.

Future strike fighter sensors will be both integrated and networked with other platforms.

When ready for operational use, carrier-based UCAVS will provide unprecedented range and endurance to carrier strike groups which value those attributes highly.

Conclusions

IN CONCLUSION, one way to assess the future of naval aviation is to look at its progression to this point. For example, consider the following words, written sometime in May or June 1944 by Bernard Brodie, known then primarily for his work as a naval historian.

> "The peculiar value of the airplane in modern naval war extends over two very diverse fields – reconnaissance and attack. Almost every aerial bomb that strikes home makes the front page of the news, but the accomplishments of the aerial observer are generally and understandably passed over in silence. Yet it may be doubted whether the reconnaissance value of aircraft over the seas adds up to much less than their attack value. The two cannot as a rule be separated, but great decisions have been made and large operations put in progress on the basis of intelligence gained from the air. Correct intelligence is the least publicized but also the most elemental prerequisite of successful warfare."[xvii]

One can see in these words the genesis of today's carrier-based strike fighters and land-based maritime patrol aircraft (as well as the long history of greater public attention to the former over the latter). One can also see by comparison to today one of the main changes in the structure of naval aviation that occurred in the intervening years – the emergence of the multi-mission helicopter. Helicopters combined the reconnaissance and strike function in a single platform that can be deployed on the smallest of naval combatants, making the surface community air capable, and allowing the widest possible dispersion of aircraft within the fleet.

Much about today's aircraft has obviously changed, and much will change in the future, but they will likely remain focused on reconnaissance and strike broadly defined, both at sea and ashore, where they will retain unique advantages in the altitude, speed, and maneuverability of their movements compared to platforms which are limited to maneuvering in two dimensions.

Some of the changes to expect in the future have already begun, as in the development and deployment of unmanned aircraft to complement and perhaps eventually replace manned aircraft. For the near to mid-term, it appears that this trend will be most pronounced in those applications where range, endurance and in some cases, altitude are the most valued attributes. By comparison, manned aircraft will likely remain above the battlefield for the foreseeable future in those applications where disparate, often ambiguous, sensor inputs must be quickly fuzed, assessed, and acted upon quickly and decisively in order to achieve the desired tactical outcome.

The most important and perhaps misunderstood change which is unfolding lies, on the one hand, in the increasing benefits that will flow from networking the sensors from multiple aircraft toward a common purpose, and on the other in the concomitant potential that the network itself will become a new source of vulnerability, either through the

insecurity of its links, or the indispensability of one of its nodes. The potential benefits to be derived from networking are overwhelming, as are those associated with unmanned platforms, but much discussion to this point is characterized by less focus on the potential vulnerabilities and constraints.

For example, many advocates of "net-centric warfare" possess a near mystical belief in both the power and future availability of bandwidth at radio frequencies on and over future battlefields. Certainly, much bandwidth there already is, and more there will be, but all RF bandwidth used for communication or data relay purposes can be optimized either to maximize data rates or minimize vulnerabilities to detection and jamming but not both. Today, many of the best examples of net-centricity that have been demonstrated depend on links with very high data rates. Networks dependent on such links would not be viable on many future battlefields, requiring that their nodes retain some capability for autonomous operation. But network designers anxious to reap the maximum benefits from a networked force, and infused with the assumption of infinite bandwidth, will be tempted to design network nodes that can not function autonomously.

These issues speak both to the question of whether to man airborne platforms or not, and whether to deploy multi-sensor platforms on which some degree of sensor fusion occurs onboard or to deploy highly specialized single sensor platforms whose output can only be fuzed with other sensors using a network. In the case of reconnaissance and surveillance aircraft, the trends toward unmanned operation and network dependence can and should be more pronounced than with strike aircraft, where both the operating environment and the nature of the mission will often require that sensors, weapons, and decision makers be collocated on the same platforms.

Another set of conclusions concerns likely future sensor capabilities in different environments. Four distinct environments are important: air (and space), sea surface, land, and undersea. Variations in terrain (or hydrography) will have significant effects in some environments (land and undersea) by reducing detection ranges. Variations in weather will be important wherever optical (as opposed to radar) sensors play an important role on the battlefield. Clutter, when it exists, will greatly complicate the ability of sensors to identify moving targets ashore and fixed targets undersea.

For these and other reasons, conclusions about future sensor capabilities for use against airborne and space-based platforms, or against fixed land targets should vary considerably from the capabilities for use against mobile land targets or stationary or slow moving undersea targets. In the former cases, detection ranges are relatively long and classification is often straightforward, whereas in the latter cases, detection ranges are almost always much shorter and classification much more difficult. In the former cases, radar is dominant, whereas in the latter cases a high degree of multi-spectral sensor fusion is often necessary.

As important as these largely technical trends will be a series of political and doctrinal trends. The most important political trend is the reduction in reliable and predictable access to overseas bases that accompanies the reduction in salience of formal alliances as compared to more informal coalitions. This trend is based on a basic structural change in the external security environment and is therefore likely to last for the foreseeable future. It is the main reason why sea basing has grown in such relative importance in both major combat opera-

tions such as Enduring Freedom and Iraqi Freedom, and in operations against global terror networks which often incubate and metastasize in exactly those regions of the world where one wouldn't want access to bases ashore even if it was available.

At the doctrinal level, the most important trends for naval aviation are the evolution toward distributed air-ground operations ashore, and toward the need for a dominant defense of the sea base in a littoral as opposed to a blue water environment. The former creates the demand for a persistent and distributed air force capable both of preventing major concentrations and movements of enemy forces on the ground, enabling the use of smaller and lighter friendly ground units, and of rapidly and precisely supporting friendly forces when they encounter and engage smaller pockets of enemy resistance that cannot be detected and attacked from afar.

The need for a dominant defense of the sea base in the cluttered littoral environment splits into two primary challenges: the challenge of recovering detection, classification, and engagement ranges against cruise missiles and their launchers operating ashore, on the sea surface, and in some cases undersea; and the challenge of adapting to the inherently reduced detection, classification, and engagement ranges against submarines (and mines).

As this report shows, naval aviation is progressing toward solutions to each of these challenges. Success in these areas will in turn enable the full exploitation of the sea base's capabilities to influence and control events both on the high seas and ashore, against the full spectrum of threats. It is necessary to sustain this progress in the coming years and decades because the future security environment will brook no alternative solutions to the major security challenges faced by the United States.

Notes

i *Conduct of the Persian Gulf War* (Washington, D.C.: U.S. Department of Defense, April 1992), p. 35.

ii Peter Grier, "Pentagon Speeds Forces to Hot Spots," Christian Science Monitor, October 18, 1994, p. 1.

iii Douglas Jehl, "Saudis Admit Restricting U.S. Warplanes in Iraq," New York Times, March 22, 1999, p. 6.

iv The U.S. Commission on National Security, New World Coming (Washington, D.C. August 1999) p. 7 <http://www.nssg.gov/Reports/>.

v Report of the National Defense Panel, December 1997, pp. 12–13 <www.dtic.mil/ndp/>.

vi On the United States being an offshore balancer of last resort, see Christopher Layne, "From Preponderance to Offshore Balancing: America's Future Grand Strategy," International Security, Vol. 22, No. 1 (Summer 1997), pp. 112–123. I am predicting that the United States will adopt that broad role for reasons quite different from those proposed by Layne, but his description of how such strategy might work in practice is useful. On Britain's historic role as an off shore balancer in Europe, see Daniel A. Baugh, "British Strategy during the First World War in the Context of Four Centuries: Blue-Water versus Continental Commitment," in Daniel Masterson, ed., Naval History: The Sixth Symposium of the U.S. Naval Academy (Wilmington, Del.: Scholarly Resources Inc., 1987), pp. 85–110.

vii At about 500-600 km range, missile designers must start considering staging if they wish to pre serve a reasonable payload. Though a two stage TBM is far from infeasible, it is a leap ahead in both technology and cost, and given the simplicity of the cruise missile alternative, the latter may be preferred.

viii For an excellent discussion of such TBMs and cruise missiles, from which this discussion of the vulnerability of fixed targets draws liberally, see John Stillion and David T. Orletsky, Airbase Vulnerability to Conventional Cruise-Missile and Ballistic-Missile Attacks: Technology, Scenarios, and U.S. Air Force Responses (Santa Monica, CA.: Project Air Force, RAND, 1999).

ix Again, in the long run, the U.S. Navy's recent decision to give its next surface combatant electric drive may prove fortuitous if this provides the power supply for terminal, directed energy defenses using solid state or free electron lasers if and when the latter become feasible. A surface ship is the perfect candidate for such a terminal defense because it is continuously moving, which means it can only be attacked by weapons with terminal seekers that a laser can burn. At the same time, it is among the largest and highest value moving targets, which means that unlike ground vehicles it has a large, organic supply of power which, on an electric drive ship, can be rapidly turned into electricity to power a laser.

x These are all points made by Stillion and Orletsky, Airbase Vulnerability, p. 31.

xi Edward N. Luttwak, 'The Operational Level of War," in *Strategy and History: Collected Essays*, Vol. 2 (Transaction Books, New Brunswick, N.J.: 1985) pp. 182-186.

xii For a summary of this debate, see Harvey M. Sapolsky and Jeremy Shapiro, "Casualties, Technology, and America's Future Wars," Parameters, Vol. XXVI, No. 2 (Summer 1996) pp.119-127.

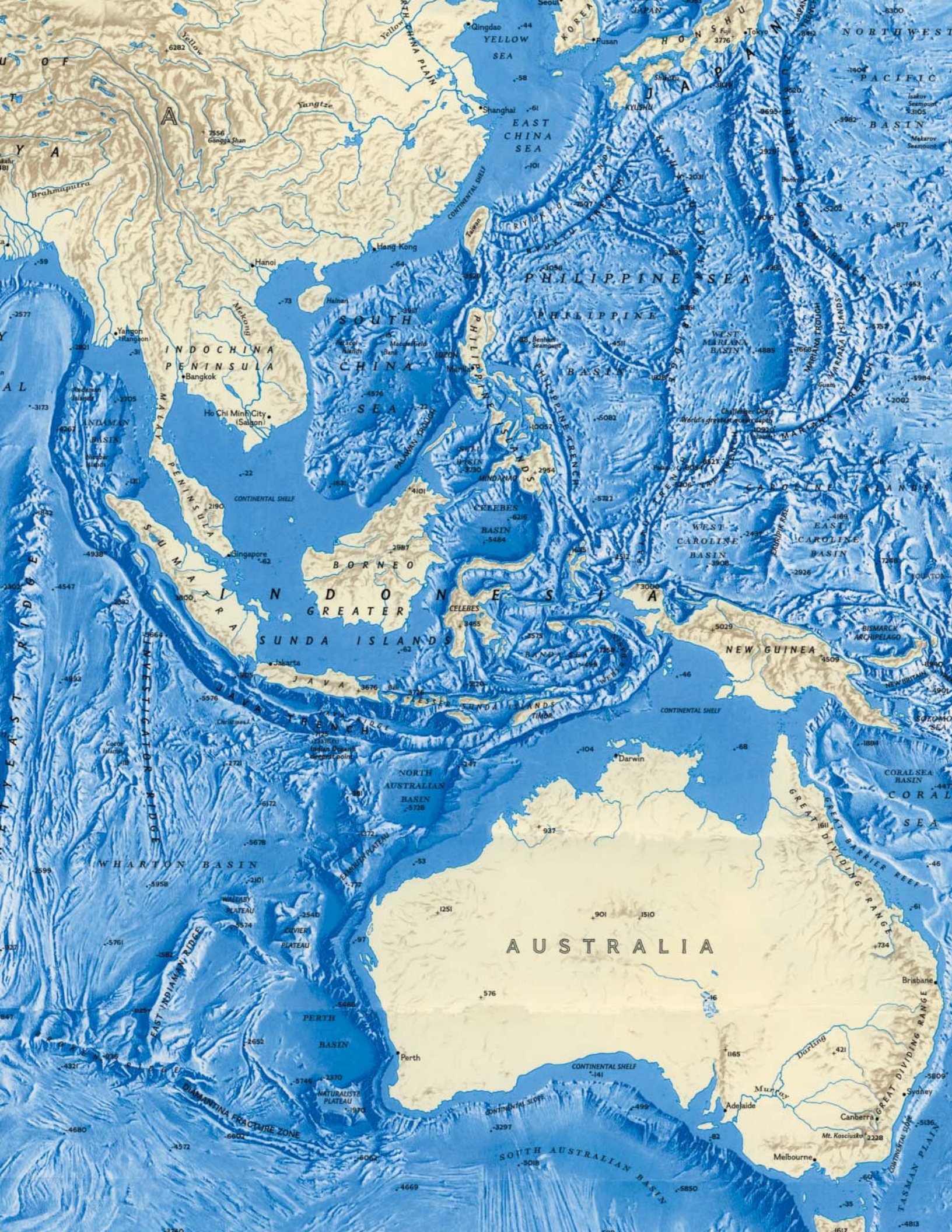
xiii The best account of those events is Mark Bowden's masterful Black Hawk Down: A Story of Modern War (New York: Atlantic Monthly Press, 1999).

xiv Stephen Walt, "Muscle-bound: The Limits of U.S. Power," Bulletin of Atomic Scientists, March April 1999, p.44.

xv Douglas Barrie, "India Begins To Receive Russian Missiles for Subs," Defense News, October 23, 2000, p. 60.

xvi For a vivid description of these early LGB operations, see Jeffrey Ethell and Alfred Price, *One Day In A Long War* (New York: Random House, 1989).

xvii Bernard Brodie, *A Guide to Naval Strategy,* Third Edition (Princeton, N.J.: Princeton University Press, 1944) pp. 211-212.


YELLOW SEA
EAST CHINA SEA
PHILIPPINE SEA
SOUTH CHINA SEA
INDOCHINA PENINSULA
GREATER SUNDA ISLANDS
BORNEO
CELEBES
NEW GUINEA
AUSTRALIA
WHARTON BASIN
NORTH AUSTRALIAN BASIN
PERTH BASIN
SOUTH AUSTRALIAN BASIN
CORAL SEA BASIN
CAROLINE ISLANDS
WEST CAROLINE BASIN
EAST CAROLINE BASIN
WEST MARIANA BASIN
MARIANA TRENCH
PHILIPPINE TRENCH
JAVA TRENCH
TIMOR TRENCH
BISMARCK ARCHIPELAGO
SOLOMON SEA
CORAL SEA
GREAT BARRIER REEF
GREAT DIVIDING RANGE
TASMAN PLAIN
NORTHWEST PACIFIC BASIN
Qingdao
Shanghai
Hang Kong
Hanoi
Yangon (Rangoon)
Bangkok
Ho Chi Minh City (Saigon)
Singapore
Jakarta
Darwin
Perth
Brisbane
Sydney
Canberra
Melbourne
Adelaide
Tokyo
Pusan
Seoul
Manila
Guam
Yangtze
Yellow
Mekong
Brahmaputra
Mt. Kosciusko 2228

Security Studies Program
Massachusetts Institute of Technology
E38-600
Cambridge, MA 02139

617.258.6531
http://web.mit.edu/ssp