As modern mobile SAM systems and precision tactical ballistic missiles proliferate, military constraints on access to bases ashore and on traditional methods of projecting power from those bases will be added to the political constraints which already exist

Demands on Today’s Navy in the New Security Environment

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 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. One key to meeting this challenge will be improved sensors and weapons for existing and planned naval platforms, as well as better data networks linking those platforms together.

This section will be organized around discussions of strike warfare, undersea warfare, and anti-air warfare. The evolution of the demands in each warfare area since the Cold War will be described, and projections will be made of how those demands are likely to change further in the transition from the near to the more distant term security environment. The section ends with shorter discussions of the evolving role of space and of new challenges in countering weapons of mass destruction (WMD).

This discussion will lay the groundwork for the next section, which will look at opportunities within those warfare areas for new sensors, weapons, and networks on or supporting the Navy's main platform communities – aviation, submarine, and surface. The specific systems discussed will be those which best help the Navy to meet its near term demands, while at the same time preparing it for a more distant security environment where access constraints of all types will be more serious.

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

The U.S. Navy emerged from World War II victorious in two ASW campaigns. In the Battle of the Atlantic, U.S. and allied antisubmarine forces beat back the challenge posed to their sea lines of communication by Doenitz's U-boats, while in the Pacific, a prosubmarine campaign was waged by American submarines that cut the sea lines of communication within Japan's Greater East Asia Co-Prosperity

Sphere. In the years immediately after the war, the U.S. Navy confronted a major challenge to its undersea warfare dominance. German submarine development, driven by the rigors of waging the Battle of the Atlantic against the Allies' increasingly potent ASW forces, had leapt forward during the course of WWII. By the end of the war, using snorkels, greater battery capacity, and better hull forms, the Kriegsmarine had deployed Type XXI submarines with vastly improved offensive performance while submerged. These developments came too late to influence the outcome of the war, but they were a harbinger of things to come, since their designs also fell into the hands of the Soviet Union.

Soviet submarines based on these 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 community of ASW as its primary Cold War mission. Although this focus on ASW predated the introduction of nuclear power, its full potential was realized in the early 1960s when quiet nuclear submarines were developed that could hear their louder Soviet counterparts at much greater ranges than they themselves could be heard. This acoustic superiority lasted almost through to the end of the Cold War.

Submarines were certainly never the only ASW instrument during the Cold War. Maritime patrol aircraft also played a key role as undersea surveillance systems became fully operational in the early 1960s. Patrol aircraft offered speed that submarines lacked, making them particularly useful in the initial localization of a contact 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 and short range ASW weapons 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 battle groups 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 current trends in America’s external security environment may confront the U.S. Navy with 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 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 current diesels like the Russian Kilo.

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 the radar horizon without the need for the noisy and battery-draining approach run necessary for a traditional, torpedo-armed, diesel-electric boat.[20] 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 asset-intensive and protracted 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.

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 were 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 Baltic and the Yellow Sea 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 merge, as the submarine is 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 of joint forces.

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. But the Navy’s AAW forces also face the brand new challenges of countering both conventional and WMD-armed, land attack ballistic missiles and projecting defense ashore against them. Today’s antiship cruise missile threat is the descendant of the Kamikaze threat and represents the primary above-the-waterline access constraint on naval surface combatants. Ballistic missiles do not pose such a threat to ships at sea, but the threat they pose to targets ashore may only be countered in an access-constrained environment from the sea. Thus, the

Navy will need to defend itself at sea, and project a defense for joint forces ashore from the sea.

Antiship Missile Defense During the Cold War

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-range 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, 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 integrated 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 antiship missiles (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. Its 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 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. In the not too distant future, 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.

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 is the specter of supersonic, sea skimming ASCM attack in the littoral from truck-mounted launchers ashore, fast boats, or non-nuclear submarines that are largely immune to, or which evade a Battle Group’s traditional outer defenses, and 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 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 a 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 intermediate 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 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.

Tactical Ballistic Missile Defense After the Cold War

Alongside ASCM defense lies the all new AAW challenge of tactical ballistic missile defense (TBMD). Tactical or theater ballistic missiles are attractive to lesser powers because they provide a method of launching long range fires against a major power such as the United States where the barriers to entry created by scale economies are much lower than they are for combat aviation. This is because the first missile that a regional power deploys gives it an initial capability, whereas combat aviation requires a whole system of systems before it can provide a credible capability against a major power. Thus, a country such as Iraq could spend many billions on modern Soviet and French fighters and not have one of its aircraft penetrate Saudi airspace during Desert Storm, while the best evidence indicates that few if any Iraqi SCUD missiles were shot down after being launched, and few if any mobile SCUD launchers were destroyed in their launch areas.

TBMs can be used as indiscriminate terror weapons whether they are armed with weapons of mass destruction or with conventional high explosive warheads. More ambitiously, with the advent of satellite-based navigation systems like GLONASS and GPS, conventional TBMs can be used with relative precision against high value military targets if they are provided a maneuverable payload with INS/GPS guidance. These two potential TBM missions pose brand new access challenges to U.S. forces.

First, opposing TBMs, and especially TBMs armed with WMD, create a political problem if and when they cause potential allies of the United States to weigh the advantages and disadvantages of balancing regional threats with U.S.military support. In these cases, the potential ally will need to be convinced that military cooperation with the United States against a regional, missile-armed threat will enhance its security rather than decrease it by making it a potential target of missile attacks.

During the Cold War, the United States assured allies such as Germany and Japan of the value of their close ties to the United States by extending or projecting its nuclear deterrent forces to cover them, promising for example to use nuclear weapons first if such use was deemed necessary to turn back a conventional attack, and promising to treat a nuclear attack against an ally as if it were a nuclear attack against the United States. In return for these promises, and the repeated and very expensive efforts mounted to preserve their credibility, U.S. forces were granted extensive peacetime access to the bases needed to mount a credible defense of its allies’ territory and prevent Soviet expansion.

The analogue to extending or projecting deterrence in today’s security environment will depend largely on the U.S.’s ability to extend or project a credible defensive umbrella over allied territory. Such an umbrella need not be impermeable to have the desired political effect, which is to demonstrate U.S. resolve to protect its potential allies from threats against which they might otherwise be naked. Thus, TBMD will be an access enabler because it will reduce the likelihood that potential allies will be blackmailed into appeasing regional aggressors rather than balancing against them by allying with the United States.

Opposing TBMs will also pose direct military challenges to U.S. forces when they become capable of attacking specific military targets with high accuracy. This will enable conventional missile attacks against soft, fixed, aboveground targets. Unhardened air bases of the type that expeditionary air forces must often use will be vulnerable to such attacks, as will ports where military and commercial sealift must debark. The emergence of such a conventional missile threat will depend largely on whether potential opponents develop and deploy INS/GPS guidance for the already ubiquitous TBMs whose range and payload bump up against or exceed existing Missile Technology Control Regime (MTCR) limits of 300 kilometer range and 500 kilogram payload.

In both cases, the need to project TBMD ashore is a brand new, post Cold War challenge. Despite the fact that the Soviet Union deployed a large TBM force during the Cold War, defenses against that threat were never considered necessary for two reasons. First, because of the deep and prolonged cooperation between the U.S. and its main allies, elaborate and very expensive measures to harden overseas air bases against conventional, chemical, or even nuclear attack were possible and were implemented. At the same time, prior to GPS, conventional TBM guidance was limited to all-inertial systems which could not give the accuracy needed for precision attacks against such bases. Second, the geographic scale of the main fronts of the Cold War allowed the United States and its Allies to use strategic depth to protect the more vulnerable nodal points of its logistics infrastructure from conventional missile attack. Thus, for example, many (but not all) major NATO ports of debarkation lay outside the range of TBM systems such as the SS-21.

By contrast, even in Saudi Arabia during Desert Storm, which intentionally built a surplus of expensive hardened airbases during the Cold War, many allied air units were forced to operate from unhardened bases within range of Iraqi Scud missiles, and both main Saudi ports of debarkation were within range of Iraqi Scuds as well. Because Iraq only fired conventional Scuds, and because those missiles had primitive guidance, they could not be aimed accurately at such inviting targets. This threat will almost certainly continue to evolve in ways that greatly constrain the ability of land-based forces to operate without fear of attack at their operational and logistics bases unless those bases are provided a credible defense.

Strike Warfare

Over the course of the Cold War, and into today’s security environment, strike warfare operations mounted by aircraft have evolved into a mature system. In that system, individual platforms have become much more lethal because of precision weapons, but the cost of penetrating modern defenses with manned platforms has also risen sharply.

Alongside traditional combat aviation are newer standoff precision weapons such as Tomahawk cruise missiles, and TBMs such as Land Attack Standard Missile (LASM) or Army Tactical Missile System (ATACMS). These systems are substantially less mature in their development than combat aviation, but therefore also face increasing returns on investment. Thus, they will get both cheaper and more capable with time.

The new security environment will demand that strike warfare assets become both more lethal, particularly against moving or mobile targets in addition to fixed targets, and less vulnerable to opposing air defenses, particularly un-cooperative ones such as those encountered in Allied Force, which seek only to survive and remain a threat in being, diverting strike capabilities to the task of defense suppression. In cases where an opponent possesses WMD-armed ballistic missiles, there will also be a demand for platforms that can strike with surprise and en masse, in order to give political leaders the option to attack all of the opponent's WMD weapons and infrastructure at the outset of a conflict. This will be a particular challenge for the Navy, whose carriers will provide the best access for tactical aviation in both crises and early in regional conflicts, but whose deckspace is finite, amplifying the negative effects of any diversion of its air wings away from true strike operations.

In facing this unique challenge, naval aviation will also possess some unique advantages, the main one being the fact that the Navy's other major strike warfare assets in the surface and submarine communities are aggressively pursuing the increasing returns on investment available from further stand off precision weapon development. Together, the air, surface, and submarine communities face significant opportunities for combined arms solutions to problems like finding and attacking mobile targets, or quickly destroying rather than merely suppressing a non-cooperative air defense system.

Strike Warfare During the Cold War

Methods of performing the strike warfare mission during the Cold War varied largely according to changes in the offense-defense relationship between combat aircraft and air defenses, because during much of that period, aircraft were the dominant strike platform. Changes in this relationship affected both the Air Force and naval aviation.

In the beginning, aircraft were designed to simply fly over enemy defenses, using a combination of speed and altitude. This trend reached it's apotheosis with aircraft like the B-70, which was designed to exceed Mach 3 at 60-70,000 feet. In the Navy, the progression from Savage (AJ-1), to Skywarrior (A-3), to Vigilante (A-5) in heavy attack squadrons illustrates the same trend. This approach was rendered obsolete in the early 1960s by the SAM which, by using a rocket motor, finally eliminated for good the high altitude sanctuary that aircraft designers had pursued since the dawn of the air age.

There were two main responses to the SAM. One led to the adoption of ballistic missiles, which restored to the offense the advantage in height and speed, albeit in a platform that was limited to delivering nuclear weapons because of its relative inaccuracy compared to aircraft. The second led to the adoption of low level penetration tactics by aircraft. These relied on the fact that terrain obstructions masked a low level penetrator from surface radars, and that background clutter masked it from airborne radars looking down at it. The classic example of an aircraft designed for this mission was the F-111, which sought survival in fast, terrain following flight. This is also the tactic that allowed B-52s and A-6s to remain effective as lone penetrators beyond the early 1960s. It was adopted for both nuclear and conventional air operations, and became threatened with obsolescence in those two mission areas for different reasons.

The air war in Vietnam, as well as the Israeli experience in the Yom Kippur war, demonstrated that low altitude attacks were not well suited to conventional operations. Aircraft flying low and fast could not find and bomb targets with great precision. Nuclear weapons could compensate for this imprecision, but in a conventional war, pilots were forced to climb to find the target and then dive to deliver weapons more precisely on it. Against unattrited terminal air defenses, which included both SAMs and dense antiaircraft artillery (AAA) barrages, these tactics led to significant losses and still did not provide the precision necessary to deliver unguided iron bombs accurately enough to destroy important targets like bridges or hardened bunkers.

This was less of a problem in nuclear operations, because nuclear weapons could destroy even the hardest targets within a lethal radius of hundreds of feet. On the other hand, nuclear operations against the Soviet Union required passing through an air defense system that included an enormous fleet of manned interceptors. Low flying bombers depended on terrain clutter to hide them from airborne radars, but by the early 1970s, the U.S. was using doppler signal processing to allow such radars to distinguish fixed from moving targets in the their field of view when looking downward. Look down/shoot down radars, once deployed by Soviet air defense forces, would eliminate the low altitude sanctuary.

The responses to these two separate challenges were quite different. For conventional operations, medium altitude tactics were adopted. These tactics depended on two innovations. The first was the creation of forces dedicated to suppressing enemy SAMs, while the second was the creation of precision guidance techniques that greatly increased the accuracy with which weapons could be delivered from medium altitude. SAM suppression tactics varied by service and country, but in all variations used some combination of radar homing weapons, jamming, and deception to kill or confuse SAM radars, thus creating a medium altitude sanctuary against ground-based air defense systems. From medium altitude, strike aircraft could locate their targets and guide new precision weapons to them, either semi-actively using a laser beam to designate the target, an approach favored by the Air Force, or by command using a data link to steer the weapon based on the readout provided by a terminal seeker in its nose, the method initially preferred by naval aviation.

This defense suppression tactic was not available to the long range bombers of the Air Force's Strategic Air Command, since its aircraft could not operate as part of a massive strike package containing fighters, Wild Weasels firing antiradiation missiles, and various jamming and other electronic warfare aircraft.[21] One answer was the B-1, which essentially sought to preserve the low altitude tactic by combining speed with a very sophisticated electronic countermeasures (ECM) suite. Its cancellation in the late 1970s led to both standoff weapons and stealth aircraft. The standoff tactic kept the launching aircraft out of range of opposing air defenses, relying for penetration on long range cruise missiles. The small size and terrain-hugging flight of these missiles made them hard to detect and even harder to kill, and they could be launched in numbers sufficient to saturate opposing defenses. Perhaps most important, a new type of guidance system enabled these missiles to fly long distances at very low altitude with precision equal to manned bombers.

Stealth, on the other hand, sought to restore to the aircraft the ability to penetrate defenses by eluding them. Technologically, this means designing aircraft which either absorb radar energy or reflect it away from its transmitter, hence the unusual shapes of aircraft like the F-117 and the B-2. When first deployed, stealth allowed a lone aircraft to penetrate unattrited air defenses at medium altitude and subsonic speed as long as it avoided daylight operations when visual detections were possible.

Strike Warfare During Desert Storm

Systems representing every stage in this evolution participated in Desert Storm. Stealth aircraft carrying laser guided bombs (LGBs) and conventional cruise missiles with terminal seekers launched from Navy ships and submarines were the only weapons aimed at targets inside the ring of terminal defenses surrounding metropolitan Baghdad. American war planners sent only F-117s and Tomahawks against these targets both because they were the most heavily defended, and because they were in areas where collateral damage was least acceptable. Other well defended targets in Iraq were attacked by large, medium altitude strike packages in which escorts outnumbered bomb droppers by as much as 3 to 1. When able to use precision weapons, mostly LGBs, the strike packages were very effective, but there were relatively few LGB-capable aircraft available. Strike packages using traditional iron bombs were much less effective. In neither case did aircraft in these packages suffer significant losses. The low altitude tactic remained the preferred penetration method of the Royal Air Force, which like other European members of NATO had never fully embraced the strike package method because of its great cost. As a result, its Tornados experienced higher, though still historically low, loss rates.

Very rapidly, these combined operations destroyed or suppressed the Iraqi air defense system to such a degree that a medium altitude sanctuary over Iraq for essentially any aircraft was created within days. This allowed B-52s and, on occasion, even AWACS and tanker aircraft to operate safely in opposing airspace with only limited fighter and defense suppression escorts.

Strike Warfare in Allied Force

The Desert Storm experience confirmed both the value of precision weapons and the increasing expense of delivering them against well defended targets using manned aircraft that must overfly the target. However, it only hinted at the promise of precision weapons, since percentage wise so few were actually used, and of those used, the overwhelming majority were laser guided gravity bombs delivered by aircraft. Thus, other than the Navy’s Tomahawk cruise missile, which played a major role early in the conflict, other uses of ballistic missiles and cruise missiles, both surface and air-launched, were extremely limited. Also, Desert Storm demonstrated the limitations of any laser or IR-guided weapon when used through weather. Also, the various means of delivering precision weapons were tested along only one axis, that being their ability to penetrate defenses. Other potential challenges to precision weapon delivery were absent due to the immediate and wide availability of local bases ashore. Also, the Desert Storm defense suppression experience was with a cooperative opponent, i.e. one that sought at least initially to complete SAM engagements against Allied aircraft even if that made engagement radars and batteries more vulnerable to destruction by antiradiation missiles. Finally, the moving or mobile target problem in Desert Storm presented itself to the Allies in a relatively benign geographic and operational environment. Thus, the desert terrain was flat and relatively featureless, giving unrestricted, relatively clutter free views of the battlefield to allied sensors, while operationally, mobile targets appeared in “weapons free” environments where the opponent had to concentrate in order to be effective, and in which collateral damage was generally not a concern.

By contrast, the Allied Force experience produced a very different set of lessons. First, precision weapons, and specifically laser-guided bombs, were widely used. Because of their wide use, and because periods of cloud free weather were significantly rarer in the temperate

European climate, Allied air operations encountered many periods when LGBs could not be used effectively. On the other hand, Allied Force also saw the first, limited use of INS/GPS guided weapons immune to the effects of weather, but limited to attacks against fixed targets whose location is known. Second, unlike the Iraqis, the Serbs operated their air defense system in a way designed to preserve it as a threat in being. Thus, Allied air planners never faced the relatively benign "air supremacy" phase that they experienced in Desert Storm after largely destroying rather than merely suppressing Iraqi SAM batteries. And third, allied air planners in Allied Force faced a very different mobile target problem than they faced in Desert Storm, one in which geography limited the view of the battlefield for standoff sensors like JSTARS, in which mobile Serb ground units were intermingled with civilians, and in which those ground units were never really forced to concentrate and move en masse in order to attack or defend territory from opposing ground forces.

From LGBs to INS/GPS. 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.[22] Post-Vietnam development of forward looking infrared (FLIR) technology allowed night/clear weather LGB operations on aircraft which combined a laser designator and a FLIR. This second generation capability was not demonstrated on a large scale until Desert Storm, and even then, a relatively small percentage of the total force in that conflict was so equipped.[23] The wide deployment since Desert Storm of FLIR/laser illumination pods in both the Air Force and the Navy has greatly increased the percentage of the force with such night/clear weather precision strike capabilities against fixed targets, as demonstrated in more recent operations over Iraq, Bosnia, and more recently, Serbia and Kosovo.

LGBs allow clear weather, precision strikes from medium to high altitudes, but operations from those altitudes frequently encounter cloud cover. This constraint prevents all weather use of LGBs and therefore reduces LGB-based precision strike capabilities to the extent that cloud cover over the target is common. Even over the deserts of Iraq and Kuwait, this constraint proved troublesome, and it proved crippling at times in the more cloudy, climate typical of Serbia and Kosovo, a characteristic obtaining throughout the temperate zones of the world, including all of the Asian littoral.

The solution to this problem will be weapons that integrate GPS and inertial navigation systems (INS). Integrated GPS/INS provides an all weather, through the cloud, weapon guidance capability that is compact, relatively cheap, and which can be made robust against countermeasures. As with second generation LGBs during Desert Storm, weapons reliant only on INS/GPS were first introduced amidst great acclaim in Allied

Force, but only on a limited scale, mostly in the form of some 600 joint direct attack munitions (JDAMs) dropped over 78 days by 6 B-2s.[24]

INS/GPS guidance will revolutionize precision strike against fixed targets because, compared to laser guidance, it will make the accuracy of precision weapons completely independent of weather, range as well as altitude of delivery, and perhaps most important, a man in the loop to identify and lase the target. In principal, this should mean that all strikes against fixed targets will eventually be conducted with standoff weapons of sufficient range to put their launch platforms out of range of surviving enemy defenses.[25] This does not mean, however, that all precision strikes against fixed targets will be made with 700 mile range weapons like Tomahawk at $500,000 apiece. Non-stealthy aircraft dropping cheap gravity bombs like JDAM will still be needed to destroy a large percentage of the total target set, but will need better defense suppression support to fully exploit their unique advantages, and thereby avoid the need to expend at least two or three HARMS at $250,000 apiece on each strike package sortie.[26]

That INS/GPS will enable a more robust, standoff, precision strike capability against fixed targets is fortunate, because as we shall see in the next section, enemy defenses will likely become much more effective than they already are. But these defenses will also complicate any solution to the mobile target problem by making it difficult to deploy survivable sensor networks within line-of-sight of the mobile targets that need to be found and identified. In the traditional approach to attacking both fixed and mobile targets, a man in the loop within line-of-sight of both the target and its defenses combines the target location and weapon aiming/guiding functions. INS/GPS weapons can eliminate the need for a human to guide the weapon, allowing standoff operations, but a network of sensors within line-of-sight of the targets must be developed if the human is to be replaced in locating the target with GPS-quality precision. No such network exists today.

From the SAM-2 to the SAM-6 to the SAM 10. Smaller countries which anticipate conflict with the United States generally do not plan on mounting a preclusive defense of their own air space. Instead, they depend largely on radar-guided SAMs, man-portable IR SAMs, and antiaircraft artillery (AAA) and assign these systems the operational goal of imposing costs rather than providing a preclusive defense. These costs can be measured in three ways: directly, as a function of opposing aircraft shot down; or indirectly, either as a function of opposing strike assets diverted to defense suppression missions, or of strike missions flown at altitudes which limit their effectiveness. There is also an operational tradeoff between air defense tactics designed to maximize direct costs and those designed to maximize indirect costs, and this tradeoff is manifest in the different approaches taken by Iraq in Desert Storm and Serbia in Allied Force.

In pressing home their SAM engagements, Iraq only succeeded in shooting down one allied aircraft that enjoyed a direct Wild Weasel escort, but their SAM units suffered enormous attrition from HARM attacks.[27] Within a week, these losses caused Iraqi radar-guided SAM activity to drop off precipitously, and allied aircraft were able to operate freely without HARM escorts at medium altitudes throughout much of Iraqi air space. Thus, by initially seeking to maximize the number of allied aircraft shot down, the Iraqis also rapidly expended their radar-guided SAM force.

The Serbs, faced with a similar operational challenge, chose instead to maximize the indirect costs incurred by allied air operations. In more than two months of operations, they only shot down three allied aircraft, but their radar-guided SAMs also managed to survive the war in large numbers. In particular, 19 out of 22 of their most modern, mobile SAM-6 batteries survived, even though they were used throughout the war, having fired at least 266 missiles.[28] The Serb strategy appears to have been to preserve the threat of its most potent, ground based air defenses in order to force the allies to continue allocating the full panoply of defense suppression assets needed to suppress them on each strike mission. They did this by repeatedly refusing to press home SAM engagements, in many cases wasting their missiles, but making it quite clear that they were still extent and operational. They also repeatedly moved their SAM batteries after such engagements had revealed their position. Given the relatively low numbers and high value of allied defense suppression assets, and given the continuing demand for them, their availability put an upper bound on the rate at which the air war could be prosecuted, a ceiling which was much lower than would have been the case if Serb radar-guided SAMs had been destroyed at the outset. This was one of the big indirect costs incurred by the allies.

Thus, by husbanding their SAMs, the Serbs were able to limit the intensity of NATO air operations to that which could be supported by their limited defense suppression assets. Yet the Allies expended HARMs at roughly the same rate as they did in Desert Storm, but with much less effect.[29]

These problems will get much worse if and when Allied air forces encounter more modern, mobile SAM systems such as the Russian SAM-10. SAM-10 missiles provide the greater than 100 km range of strategic SAMs like the SAM-2, with the mobility of shorter range systems like the SAM-6. Furthermore, its phased array main engagement radar has both a much higher power-aperture product and a much more agile beam than its mechanically scanned predecessors. This gives the radar a much longer detection range against even low radar cross section targets such as the stealthy F-117 and the B-2, and allows it to more quickly acquire and track multiple targets.

Systems such as the SAM-10 will greatly increase the indirect costs of defense suppression if traditional methods are maintained. This is because the greater effective detection range of the system will prevent even stealthy aircraft from attacking it without a weapon with a significantly greater range than the HARM. In this scenario, Wild Weasel aircraft would have to become stealthy, and their antiradiation weapons would need greater range than today's HARM, and higher speeds, all in a package small enough to be carried internally by an F-22 or a JSF. This dramatic increase in the cost of individual Wild Weasel platforms would at best buy an equal capability against new systems such as the SAM-10 as its predecessors provided against systems like the SAM-6. Yet by refusing to press home its engagements, a modestly sized force of SAM-10s could still extract indirect costs, forcing U.S. forces to limit their operations to the level that could be supported by still scarce and now much more expensive Wild Weasel platforms.

The unattractiveness of this scenario has led to consideration of an alternative approach to defense suppression. In it, SAM engagement radar locations are instantaneously and more precisely determined using multiple rather than single platform geolocation techniques. These techniques allow detection of even the briefest signals with a precision sufficient to target the emitter with an INS/GPS-guided standoff weapon rather than an antiradiation missile. Such an approach is attractive both because it deals with the advancing threat of opposing air defense systems, and with the tactic of using them in ways that emphasize indirect rather than direct costs.

In the near term, this is one of the areas where programs like TENCAP can produce significant leverage by using national assets in space to help form networks of multiple sensors within line-of-sight of the relevant targets.

From Fixed to Mobile Targets. Fixed targets are often found and identified using traditional intelligence methods, often well in advance of a conflict. Even when the value of certain fixed targets, such as command posts and WMD storage sites, varies significantly over the course of a conflict, their positions are still usually known with precision in advance of the conflict, even if the time when it is optimal to strike them is not.

By contrast, though mobile targets can sometimes be preemptively struck at their bases using pre-conflict intelligence, usually they must be found and identified while in the field. Rather than traditional intelligence methods, this creates the demand for surveillance and reconnaissance capabilities with continuous, wide area coverage that can search for and detect potential targets, classify them as real targets, and locate them in both time and space with accuracies compatible with the accuracy, lethal radius, and time late of the weapons that will be used to attack them. As any practitioner of ASW will understand immediately, it is the ability to find and identify these targets in a “noisy” environment with an acceptable false alarm rate that will be the most difficult challenge. In different ways, and under different circumstances, both the failure of the great SCUD hunt of Desert Storm, and the success of Serb ethnic cleansing activities in Kosovo during Allied Force, demonstrate that this challenge is a long way from being met.

Other examples from the recent past might seem to vitiate this point. For example, in two cases during the Gulf War, the Battle of Khafji and the later Iraqi retreat from Kuwait City toward Basra, allied air forces pummeled Iraqi ground forces from the air, in the former case stopping an attack, and in the latter case turning a retreat into a rout. The difference is that in both these cases the false target problem was moot. At Khafji, this was because the battle took place in a low or zero noise environment. In other words, detection equaled classification because the only vehicles in the area were Iraqi military vehicles. Later, at the so-called Highway of Death, where civilian and military targets were intermingled, the false target problem was initially ignored, although this did not last for more than 48 hours, when the decision was made to cease operations at least in part out of humanitarian concerns.[30]

The Desert Storm SCUD hunt and the effort to slow Serb ethnic cleansing operations in Kosovo during Allied Force were different because the false target problem was real. In the SCUD hunt, one of the main problems was that allied aircraft had difficulty distinguishing SCUD launchers from trucks and other vehicles. There were many more of the latter than the former, and both used the same road networks. Thus, many fuel trucks were attacked and destroyed, but few if any SCUD launchers. Another problem was that the best sensors were not always available for the SCUD hunt because it occurred relatively deep in Iraqi airspace. This often prevented assets like JSTARS and Rivet Joint from participating. In Kosovo, roughly the same problem of distinguishing military from civilian vehicles was exacerbated further by an extreme aversion to civilian casualties and collateral damage, political concerns that were entirely absent from the SCUD hunt, which was dominated by the overwhelming political imperative of keeping Israel out of the war by reducing, or of at least appearing to reduce, the SCUD threat. Also, there is considerable evidence that the Serbs were more aggressive than the Iraqis in their use of decoys, making an already noisy environment even noisier, and diverting Allied weapons from their real targets.

An obvious first step toward addressing these problems is to find ways of providing continuous, theater wide, synthetic aperture (SAR) and moving target indicator (MTI) radar surveillance coverage in support of future conflicts. Such a step would not by any means be sufficient as a solution to the mobile target problem, but it is almost certainly a necessary step toward one.

SAR and MTI are two different radar techniques that can be combined in a single surveillance platform.[31] 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 natural aperture of the platform's radar antenna. With SAR, a radar gains an imaging capability with resolutions approaching 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 combined, a SAR/MTI radar can detect and track moving vehicles over a wide area using the MTI mode, or provide high resolution, precisely located images of a series of spots within that area. SAR/MTI radars can not yet interleave these two different modes rapidly enough such that a target detected using the MTI mode can be imaged and more precisely located using the SAR mode as soon as it stops moving, and then picked back up on MTI once it starts moving again. When and if this capability is developed, it will in theory allow continuous, all weather tracking of high value mobile targets within the coverage area of the radar. In practice, this capability will be dependent both on the density of SAR/MTI coverage over the battlefield, and on the skill of the network's human operators.

Two major technical challenges will dominate SAR/MTI sensor development. One concerns the precision with which it can locate targets, and the other concerns the degree to which it can identify and classify the targets it detects. Mobile targets pose particular challenges in both areas. Current MTI radars cannot by themselves provide targeting quality geolocation information for moving targets, and their classification capabilities are limited to relatively gross distinctions like that between tracked and wheeled vehicles. There will certainly be great technical progress in both these areas, and it is well beyond the scope of this report to speculate about the details of what will result, but it is possible now to outline how different outcomes in SAR/MTI sensor development might broadly effect other elements of the mobile target problem.

At one extreme, perhaps in the near term, one can imagine SAR/MTI radars providing a cueing function to other platforms which would classify and attack the target. The latter two functions might be combined on one platform, probably a combat aircraft, which would result in the simplest, most evolutionary architecture, albeit one that would remain constrained by the future defense suppression challenge. At the opposite extreme, one could imagine SAR/MTI radars, probably linked with other sensors using different phenomenologies, in a global or theater-wide network which could find, classify, track, and target a variety of mobile targets of interest. This targeting information could be used to launch standoff weapons from platforms deployed outside the range of enemy defenses. The simplest and cheapest of these weapons might be capable only of quickly attacking a set of GPS coordinates provided by the network. This would be useful against a mobile target temporarily at rest, such as a SAM-10 radar. More complicated would be a weapon able to receive continuous GPS targeting updates from the network in flight, providing a closed loop between the sensor and the weapon, and making it possible to attack targets that had moved after the weapon was launched. Most complicated would be a fire and forget weapon with a terminal seeker able autonomously to reacquire and attack a moving target designated by the network using automatic target recognition (ATR) algorithms.

In all of these architectures a common theme is that targets are found, and in many cases identified and precisely located, by sensors separated from the weapon delivery platform. This separation implies an important additional characteristic about future strike operations against mobile targets. They will, in today's jargon, be net-centric, meaning that the strike platforms participating in such operations will be dependent on their connectivity to offboard sensors via a network for their effectiveness. In many cases, elements of these networks will need eventually to migrate into space.

The Navy and Space During and After the Cold War

During the Cold War, space was a sanctuary used for intelligence or military purposes by both superpowers, mostly for remote sensing, communication, and navigation. The main value of space to naval warfighters was and is that it provides an elevated perch from which to send and/or receive signals from earth. With some important exceptions, the Cold War saw space-based remote sensors focused on the intelligence and early warning function, rather than on supporting conventional military and naval operations. Space was also widely used for satellite communications by conventional military forces, and the Navy was the prime developer and by far the widest and deepest user at the operational and even tactical levels. The Navy also developed and was the prime user of Transit, a satellite navigation system designed in the late 1950s to support SSBN operations, which constituted the first purely military (as opposed to intelligence community) use of space.

Remote sensing, communications, navigation, and timing will remain the primary uses of space by the intelligence and military communities in the future security environment, but little else about space in the pre and post Cold War security environments will be the same. The two main changes will be the need for much greater military exploitation of space, particularly in remote sensing, and the possibility or even likelihood that space will not remain a sanctuary. There will therefore be increased technical and budgetary tradeoffs to be resolved between military and intelligence community requirements, and between steps taken to exploit access to space and steps taken to assure access to that medium and the assets deployed within it. The need to resolve these technical and budgetary tradeoffs will spur debates, some of which have already begun, over the organizational structures needed to make the necessary decisions.

The Navy and Space During the Cold War

The Navy was in many respects the dominant user of space during the Cold War.[32] As noted above, the first purely military use of space was the Transit navigation satellite system, which was used from the early 1960s onward as a means for nuclear submarines, particularly SSBNs, to provide periodic, precision updates to their ships inertial navigation systems while at sea. The Navy also developed the very precise clocks that have become the heart of the Global Positioning System.

The Navy was also the most aggressive service in its use of space-based sensors in direct support of its operations, developing and operating dedicated systems such as Classic Wizard, and gaining access to other national assets through its highly successful TENCAP program. Cold War Navy TENCAP programs exploited national assets to help detect, identify, locate, track, and target ships, and to detect and track aircraft in flight. And the Navy has been the most successful developer and aggressive user of satellite communication systems, particularly at UHF starting in the 1960s, and more recently at EHF.

Space-based sensors, satellite communications, shore and ship-based command centers, and deployed naval platforms formed some of the first U.S. sensor-to-shooter networks in the 1970s when the Navy began focusing on the over-the-horizon targeting requirements for Harpoon and, later, the anti-ship version of Tomahawk. Similar networks using different sensor phenomenologies were formed to support the outer air battle against Soviet Naval Aviation.

Neither the United States nor the Soviet Union chose to pursue anti-satellite (ASAT) technologies seriously during the Cold War, though both sides did develop and test limited systems capable of attacking satellites in low earth orbit. Thus, the Navy and the other services were largely free of the need to focus on assuring their access to space, or to denying their opponent access to space. Certainly, one of the reasons that space remained a sanctuary during the Cold War was because space simply never became necessary on a wide enough scale to the operations of either side's general purpose forces. Yet it was nevertheless of great importance to identify and track the opponent's satellites even if that information was not to be used to support ASATs. In support of this surveillance requirement, the Navy also developed and continues to operate a major element of the United States' space tracking system, which was used through most of the Cold War both for intelligence purposes, and as a means of warning deployed forces of the imminent arrival overhead of a Soviet satellite.

The Navy and Space After the Cold War

The future security environment will likely force all three of the military services to exploit space more vigorously, and technology will both enable that evolution and threaten it. The drive to exploit space more intensively will come from the need to identify and precisely locate significant targets, to allow platforms to stand off from the defenses deployed by an opponent, and to permit operations in dispersed fashion so as not to present fixed or concentrated targets for that opponent to attack. The technology to exploit space, particularly in remote sensing, is growing rapidly, both in the military and the commercial sectors. For example, just as it was discovered during the Cold War that early warning satellites could detect and track aircraft, modern early warning satellite technology has already demonstrated an ability to detect the flash of a general purpose bomb exploding on the ground, potentially enabling real time assessment from space of the effects of conventional strikes on the ground. At the same time, the sensors, communication circuits, and navigation transmitters placed in space to enable this evolution will be subject to a variety of soft kill satellite countermeasures that the technology already exists to support, as well as the future threat of ASATs.[33]

Among the organizational and technical tradeoffs that will need resolution are the following questions. What will be the balance between continuing intelligence operations in space, where the protection of collection sources and methods will remain a paramount priority, and growing military operations in space, where the main priority is the timely delivery of the product in usable formats to the user in the field? What will be the balance between efforts devoted to exploiting space and efforts devoted to assuring access to it while denying access to others? Will the DOD continue to allow the individual military services to generate requirements for space systems? Will all military space systems be developed as joint, common user systems? What will be the balance between dedicated military satellite development by DOD and the use of growing commercial capabilities in space? Will DOD-commercial partnerships in space system development be possible, particularly for broadband communications systems? What will be the balance in low earth orbit between large, multipurpose satellites launched by large boosters in small numbers providing intermittent coverage, and networks of many more small satellites launched by much smaller boosters providing continuous or near continuous coverage? What will be the balance between the data rate and jam resistance built into the RF circuits forming the uplinks, downlinks, and crosslinks of future space networks?

The range and depth of the uncertainties captured by these questions show how fluid and undefined the future of space is in the future security environment. This combination of great potential and organizational and technical uncertainty resembles in some ways the situation regarding aviation during the interwar period. This analogy has already been used by some who argue in favor of a new, independent space service which would presumably be a more fervent advocate for space, just as an independent Air Force was perceived by many to have been the key to a more aggressive development of air power. Such analogies can carry great weight in the political arena and it is important that they be fully explored before serving as a guide to the future. There are actually at least three models of air power development that are potentially appropriate to the future development of space.

The most commonly cited analogy is to the Army's view of air power in the 1930s. Here, ironically, space advocates argue that today's Air Force's view of space is like yesterday's Army's view of air power. In this view, neither service was or is willing to make major investments in the exploitation and control of a new medium of operation. This explains the advocacy for either a central military space advocate and manager, or should that step prove inadequate, the further step of creating an independent space service.

Supporters of an independent space service do not discuss a second possible analogy to air power development, that being the Air Force's view of its purpose once it became independent. That view focused on the control of the air as a means toward conducting independent, strategic air bombardment operations to the exclusion of operations supporting ground and naval forces. In this view, it may be correct to assume that an independent space service would invest more in space, but it may be incorrect to assume that those investments would be focused on the support of conventional ground, naval, and air operations. Rather, it is possible that an independent space service would focus more on space control than on space exploitation, and take the same attitude to systems designed to support the other services as the U.S. Air Force took to tactical air forces supporting the Army in the 1950s.

There is still a third analogy between aviation and space development that is almost never discussed, and that is the analogy between the development of naval aviation in the interwar period and the Navy’s approach to space during the Cold War. In this view, space is a medium that is no different from the air, surface, and undersea mediums. As the Navy pursues its responsibilities, it needs access to any medium that might support its ultimate objectives. Thus, unlike the Army of the 1930s, the Navy was more willing to experiment with air power, and therefore did not lose its air arm to the independent Air Force. And unlike the Air Force of the 1950s, the Navy never developed a narrow doctrine for air power which insisted on its independence from surface forces. Instead, naval aviation remained a member of a combined arms team.

In comparing these alternative approaches to space in the new security environment, one conclusion seems clear, which is that it would be a mistake for DOD to centralize all development of military space systems in one location. This does not mean that new organizations devoted to space should not be created, nor does it mean that there are not important efficiencies to be gained through the centralized management of the procurement and operation of military space systems, but it does mean that neither a new, independent space service nor the Air Force should be given an absolute monopoly of control over the development of military space systems. The best analogy for such a move to centralize all space development activity in one organization at a time of both great technological ferment and budgetary limitation would be the British decision after World War I to concentrate all aviation activity, including naval aviation, in the newly independent Royal Air Force. This move seriously hampered the development of British naval aviation during the interwar period, with grave consequences for the Royal Navy during World War II.[34]