



Royal United Services Institute
for Defence and Security Studies

Occasional Paper

The Future of NATO's Air and Missile Defence

Sidharth Kaushal, Archer Macy and Alexandra Stickings



The Future of NATO's Air and Missile Defence

Sidharth Kaushal, Archer Macy and Alexandra Stickings

RUSI Occasional Paper, July 2021



Royal United Services Institute
for Defence and Security Studies

190 years of independent thinking on defence and security

The Royal United Services Institute (RUSI) is the world's oldest and the UK's leading defence and security think tank. Its mission is to inform, influence and enhance public debate on a safer and more stable world. RUSI is a research-led institute, producing independent, practical and innovative analysis to address today's complex challenges.

Since its foundation in 1831, RUSI has relied on its members to support its activities. Together with revenue from research, publications and conferences, RUSI has sustained its political independence for 190 years.

The views expressed in this publication are those of the authors, and do not reflect the views of RUSI or any other institution.

Published in 2021 by the Royal United Services Institute for Defence and Security Studies.



This work is licensed under a Creative Commons Attribution – Non-Commercial – No-Derivatives 4.0 International Licence. For more information, see <<http://creativecommons.org/licenses/by-nc-nd/4.0/>>.

RUSI Occasional Paper, July 2021. ISSN 2397-0286 (Online).

Royal United Services Institute
for Defence and Security Studies
Whitehall
London SW1A 2ET
United Kingdom
+44 (0)20 7747 2600
www.rusi.org
RUSI is a registered charity (No. 210639)

Contents

Executive Summary	v
Introduction	1
I. The Evolving Geopolitical Context of the Missile Threat	5
Russia	5
The Middle East and North Africa	7
The Threat Landscape: Adversary Capabilities and Intentions	7
II. Responding to the Challenge: Priorities for NATO's Air Defence	13
Air Defence in its Operational Context	13
Gaps in NATO's Current System	16
The Path Forward: Offence–Defence Integration	19
SEAT and SEAD	21
Critical Planning Decisions	23
Critical Operational Decisions	25
III. Next Steps for NATO	29
Practical Ramifications of Reframing IAMD	30
Key Lines of Effort	31
Conclusion	33
About the Authors	35

Executive Summary

SINCE THE RUSSIAN invasion of Crimea in 2014, NATO has attempted to retool itself to gradually prepare for great power competition. The Alliance has undertaken a number of steps consistent with the NATO Readiness Initiative to generate the capacity to respond to escalation on its eastern flank.¹ NATO's forthcoming future warfighting Capstone Concept will also emphasise countering threats emanating from Russia.² However, this paper argues that NATO's approach to air and missile defence has not kept pace with this broader transformation. Given the importance of long-range precision strike assets to Russia's strategic and operational frameworks, this represents a gap in the Alliance's posture and thus its credibility. This paper identifies the shortcomings that would prevent NATO from mounting a credible defence against the multi-tiered salvos that Russia can generate. The paper also highlights the changes NATO needs to make to its air defence system so that it can respond to the twin challenges of a peer competitor and sub-peer threats.

Key Points

- NATO's air and missile defence systems currently lack both the capacity and the capability to meet the air threat challenges of great power competition. This has ramifications for NATO's ability to project power and provide for the safety of its populations and critical infrastructure.
- In order to meet the evolving threat environment, the Alliance's priority must be to conceptually reframe its understanding of integrated air and missile defence (IAMD).
- NATO should not view air defence in binary terms – defended or undefended. The purpose of IAMD is to protect a select subset of critical targets for long enough for the air threat to be negated by other means.³
- The efficiency of air defence will be drastically improved through effective offence–defence integration. The suppression of enemy air threat capabilities, such as launchers, surface vessels and bombers, should be elevated from a subcomponent of offensive counter-air in NATO doctrinal parlance to assume a place equal to suppression of enemy air defences (SEAD) as a joint strike objective for Allied operations. Without this, the air defences on which SEAD missions depend will be overwhelmed.
- There should be closer coordination between the NATO Air and Missile Defence Committee and the Civil Emergency Planning Committee to better integrate active and passive defence mechanisms.

1. NATO, 'Defence Ministers Agree to NATO Readiness Initiative', press release, 8 June 2018, <https://www.nato.int/cps/en/natohq/news_155348.htm>, accessed 21 April 2021.

2. Tim Sweijs et al., 'The NATO Warfighting Capstone Concept: Key Insights from the Global Expert Symposium Summer 2020', Hague Centre for Strategic Studies, 2020, pp. 1–5.

3. For the purposes of this paper, all missile threats will be classed under the rubric of the air threat.

- In the long run, the requirements of a genuinely integrated system can only be met and afforded through functional specialisation at a national level and resource pooling for the Alliance.
- As such, this paper proposes a series of steps in which key air defence assets are committed to NATO for employment as tensions increase, in a similar manner to the European Phased Adaptive Approach, which will support a NATO-level integrated air and missile defence system.⁴ The expansion of shared funding beyond command and control for certain NATO-provisioned assets will also need to be considered.

4. Of course, not all components of the European Phased Adaptive Approach are under peacetime NATO command – the guided-missile destroyers at Naval Station Rota are one exception. These assets are capable of rapidly transitioning to NATO command, however.

Introduction

NATO IS RECALIBRATING to meet the challenges of great power competition while still preparing to manage sub-peer threats. In doing so, the Alliance will need to surmount several capability and capacity challenges.¹ This paper focuses on one aspect of the Alliance's transformation – the recalibration of NATO's air and missile defences. As others have suggested, developing a credible and integrated air and missile defence capability will be key to maintaining deterrence.² Given that Russia intends to use precision strike assets to target both the logistical and political seams within the Alliance, this recommendation is understandable.³ However, this paper contends that NATO's first task should be to examine whether it is conceptually prepared for the air threat that it will face.

This paper aims to contribute to the nascent discussion on NATO's future air and missile defence system by discussing the shortfalls in its current approach to integrated air and missile defence (IAMD), as well as the key lines of effort needed to remedy these deficits. Chapter I describes the contours of the contemporary air threat. Chapter II identifies the conceptual and structural shortfalls in NATO's current air defence system. Based on this analysis, Chapter III outlines the priorities that should guide the evolution of the Alliance's air defence system.

While the concept of great power competition is central to this paper, it does not suppose the disappearance of other commitments. However, great power competition will likely act as an organising factor around which other challenges are situated. For example, when civil wars become internationalised the challenge posed by sub-peer competitors can increase if they are provided with increased access to strike capabilities. The effect of broader systemic competition on sub-peer competitors is evident in Iran's support to the Houthi missile arsenal and Russia's provision of the Bastion-P system to the Syrian government.⁴

-
1. With regard to some of the other challenges that the Alliance will need to surmount, see, for example, Heinrich Brauss, Ben Hodges and Julian Lindley-French, 'Moving Mountains for Europe's Defence', Center for European Policy Analysis (CEPA), March 2021.
 2. Ben Hodges et al., 'One Flank, One Threat, One Presence: A Strategy for Europe's Eastern Flank', CEPA, May 2020.
 3. There exists a consensus in Russian doctrinal documents and military journals that this will be achieved by inflicting 'assigned damage' on NATO members who Russia deems to have limited will to join a conflict. Russian planning assumes that the asymmetry of commitment between NATO members can be exploited by demonstrative strikes which illustrate the costs of conflict without inflicting irreversible damage. See Michael Kofman, Anya Fink and Jeffrey Edmonds, 'Russian Strategy for Escalation Management: Evolution of Key Concepts', CNA, April 2020, p. 18.
 4. See Fabian Hinz, 'Meet the Quds-1', *Arms Control Wonk*, 15 September 2019, <<https://www.armscontrolwonk.com/archive/1208062/meet-the-quds-1/>>, accessed 21 April 2021; Sebastien

The challenge posed is twofold. First, competitors across the spectrum can use the threat posed by strike assets to constrain the Alliance politically, for example, by slowing down its response to a crisis through the threat of attacks on its infrastructure.⁵ Even if it is not the case that Russia can control escalation through limited coercion, the belief that this can be achieved can lead to risk-acceptant behaviour. Indeed, the aggressor's belief that it can dictate the tempo of escalation has historically been central to deterrence failures.⁶

Second, if the political conditions for NATO mobilisation are met, the military infrastructure which the Alliance depends on to mobilise, such as airfields, ports and command and control (C2) nodes, can still be immobilised by a peer competitor, further slowing the Alliance's response cycles. This could significantly undercut Alliance deterrence in a scenario in which a competitor aims to present NATO with a *fait accompli*.

The central contention of this paper is that NATO IAMD⁷ is neither conceptually nor functionally integrated with the broader campaign-level concept of operations necessary to prevail in a significant conflict with a peer competitor.

IAMD, as currently envisioned in a NATO context, is out of step with the re-emerging requirements of great power competition⁸ in two ways:

1. The current NATO approach to IAMD, built to provide security against a limited air threat from Iran, will be strained beyond its capacity against a great power. The contemporary missile defence threat is presumed to be presented primarily by limited numbers of

Roblin, 'Why Russia's Bastion-P Truck-Launched Anti-Ship Missiles Could Be a Real Killer', *National Interest*, 8 February 2020.

5. The assumption that calibrated pressure will slow down Alliance decision-making is a core part of Russian planning. See, for example, Dave Johnson, 'Russia's Conventional Precision Strike Capabilities, Regional Crises and Nuclear Thresholds', Livermore Papers on Global Security No. 3, Lawrence Livermore National Laboratory, February 2018, pp. 20–40. However, this might not be the case. Indeed, such attempts at coercion may backfire, although this is difficult to ascertain in advance. The low levels of enthusiasm for supporting a NATO ally evinced by the publics of many NATO member states suggests that the Russian planning assumption is, at least, plausible. On public sentiment, see T X Hammes, 'The Melians' Revenge', Atlantic Council, June 2019, p. 4.
6. Alexander George and Richard Smoke, *Deterrence in American Foreign Policy: Theory and Practice* (New York, NY: Columbia University Press, 1974), pp. 519–26.
7. This paper uses the term integrated air and missile defence (IAMD) to mean defence against all manner of damaging objects arriving in the atmosphere, including ballistic, strike and cruise missiles, and UAVs. This includes rockets, artillery and mortars, but generally the challenge of defeating them is left to shorter-range air defence systems associated with manoeuvre forces.
8. On the subject of NATO's renewed emphasis on great power competition, see Jens Stoltenberg, speech at the Centre for Strategic Studies at Victoria University of Wellington, New Zealand, 5 August 2019, <<https://www.nato.int/cps/en/natohq/168242.htm>>, accessed 8 April 2021.

medium- to long-range regional ballistic missiles.⁹ While IAMD missions such as cruise missile defence, counter-UAV and defence against tactical ballistic missiles are discussed by NATO, in practice the development of these systems occurs at a national level, in parallel to overall European territorial defence via NATO ballistic missile defence (BMD).¹⁰ Against a Russian military that can flexibly target both Alliance members' territory and military assets with the same suite of ballistic and air-breathing missiles, this distinction between NATO-owned, upper-tier BMD assets, such as the SM-3, and nationally-owned air defence systems is invalidated. It creates a stove-piped and federated system entirely unsuited to the strategic context within which the air and missile threat environment facing NATO is nested.

2. A doctrinal framework that views air defence as a task to be accomplished parallel to missions such as offensive strike raises the risk of saturation – there will never be enough defence for every point. To deliver enough air defence at the right points to enable manoeuvre, the air threat must be attrited even as active defences operate. Commanders must determine their air defence priorities in relation to what can be accomplished by offensive means – with active air defences being concentrated against those vectors that cannot be suppressed left of launch. This requires that suppression of the enemy air threat (SEAT) be elevated to a core joint airpower role, to which each service can contribute, in a manner comparable to suppression of enemy air defences (SEAD). SEAT should be planned as a core joint defensive mission alongside active IAMD, with defence defined in terms of effects, not tactics.¹¹ Current NATO planning, which places missions comparable to SEAT under the offensive counter-air mission, has the effect of diluting the effort placed on these tasks, which compete with a range of other missions within offensive counter-air, and organisationally stovepiping them from the work of air defenders.¹²

9. NATO, 'Brussels Summit Declaration', 11 July 2018 (updated 27 August 2018), <https://www.nato.int/cps/en/natohq/official_texts_156624.htm?selectedLocale=uk>, accessed 7 April 2021.

10. Luc Dini, 'Air and Missile Defence in Europe: Building a Consensus', Friends of Europe, 1 April 2015, <<https://www.friendsofeurope.org/insights/air-and-missile-defence-in-europe-building-a-consensus/>>, accessed 9 April 2021.

11. This would be comparable to the Russian concept of an aerospace deflection operation or 'Strategic Operation for Repelling Aerospace Aggression' (SORASA). See Dmitry Adamsky, 'Moscow's Aerospace Theory of Victory: Western Assumptions and Russian Reality', CNA, February 2021, pp. 1–5.

12. Insofar as the threats posed to air bases, even those with active defences, can significantly attrit the forces needed for suppression of enemy air defences (SEAD). Modelling conducted by Daryl Press suggests that a Russian air and missile campaign against NATO airbases, while incapable of defeating NATO air forces on the ground, can effectively prevent NATO from generating the sortie rates needed for a meaningful SEAD campaign for up to 30 days. See Daryl Press, presentation at RUSI Space and Missile Defence Conference, 27 February 2020. The suppression of enemy air assets is considered a subcomponent of offensive counter-air but exists as a subordinate mission within this rubric alongside the maintenance of sorties against airborne assets and focuses primarily on aircraft as opposed to other missile launch platforms and the infrastructure on which they depend. There is, by contrast, no equivalent to joint-SEAD. On NATO's rubric for the conduct

To generate the efficiencies needed to contend with the contemporary threat environment, this paper argues that three adaptations are needed:

1. The introduction of SEAT into NATO's doctrinal lexicon and the conceptual integration of active defence and strike.¹³
2. The elevation of NATO's role in IAMD from merely providing a C2 framework for nationally-held assets to exercising effective administrative and operational control over an IAMD system generated and operated in a manner similar to the assets devoted to the US European Phased Adaptive Approach (EPAA).¹⁴
3. Coordination between air defence and ongoing NATO efforts to generate civil resilience. This creates efficiencies because not every target needs to be defended as some can withstand a certain amount of damage and remain functional or can be repaired.¹⁵

The political and organisational challenges that the initiatives proposed by this paper entail should not be understated. However, if NATO is to be defended from air threats, it must now begin the arduous process of meeting these goals.

This paper's analysis derives from a review of the secondary literature and discussions held at the RUSI Missile Defence Conferences. The professional experience of one of the authors as a senior official responsible for delivering IAMD in both a US and NATO context and for creating NATO's missile defence system has also informed this paper to a significant degree.

of offensive and defensive air operations, see NATO, 'Allied Joint Doctrine for Air and Space Operations', AJP-3.3, April 2016, paras 1–9. This document treats SEAD as an independent joint function co-equal to offensive and defensive counter-air.

13. Conceptual integration does not always necessitate technological integration, but it can be a pre-condition for it. For a discussion of its technological aspects, see Tom Karako and Wes Rumbaugh, *Distributed Defense: New Operational Concepts for Air and Missile Defence* (Washington, DC: Center for Strategic and International Studies (CSIS), 2018), pp. 10–20.
14. See Jaganath Sankaran, 'The United States European Phased Adaptive Approach Missile Defence System: Defending Against Iranian Missile Threats Without Diluting the Russian Deterrent', RAND, 2015; NATO, 'NATO Ballistic Missile Defence', July 2016, <https://www.nato.int/nato_static_fl2014/assets/pdf/pdf_2016_07/20160630_1607-factsheet-bmd-en.pdf>, accessed 8 April 2021.
15. Notably, this does not necessarily require the subordination of all assets to a single authority, as with military support to civil disaster response. Literature on cross-government integration suggests that three factors are critical: awareness that cross-boundary interests exist; channels of communication; and development of a shared professional language. This enables rapid integration of effort without hierarchical subordination of all functions under a single command. See, for example, Louise K Comfort, *The Dynamics of Risk: Changing Technologies and Collective Action in Seismic Events* (Princeton, NJ: Princeton University Press, 2019), p. 108.

I. The Evolving Geopolitical Context of the Missile Threat

THE RESURGENCE OF Russian opportunism and the ongoing instability in northern Africa, the Arabian Peninsula and Western Asia, typified by the civil wars in Syria and Libya, have collectively created a range of potentialities that could see Europe threatened from the air by the missiles and strike assets proliferating in these conflict zones.

Russia

Competition with Russia will likely represent the most potent challenge in this environment. To be sure, this does not necessarily presage a return to the era in which there was a realistic threat of a full-scale Soviet assault on Europe. Contemporary Russia lacks both the capabilities and, in all likelihood, the will for such a gamble. Instead, the objectives that will guide the current era of competition are establishing positions of relative competitive advantage in the regions that form the geographical flanks of both Russia and the Alliance.¹⁶

A predominant position in regions such as the Baltic and High North, the Black Sea and the Eastern Mediterranean would potentially confer significant strategic advantages on Russia.¹⁷ For example, Russia's naval and sea denial facilities in Crimea and Tartus raise the prospect of Russia enjoying a pivotal position along one of the world's vital maritime arteries.¹⁸

Similarly, the ability to threaten the Baltic states – whether Russia wishes to invade them or not – confers it with a degree of leverage over the Alliance. Even if Russia sees little value in annexing one of the Baltic states, maintaining a credible capacity to do so and underscoring this reality through limited coercive steps could be seen as a means of providing it with a veto over the West's freedom of action more broadly.¹⁹ This is comparable to how both Stalin and Khrushchev used pressure on West Berlin as a means of exerting influence on the West's freedom of action

16. Janusz Bugajski and Margarita Assenova, *Eurasian Disunion: Russia's Vulnerable Flanks* (Washington, DC: Jamestown Foundation, 2016); Sidharth Kaushal, 'Persistent Engagement and Strategic Raiding: Options for the UK's Future Carrier Strike Capability', *RUSI Occasional Papers* (November 2020), pp. 10–30.

17. See, for example, Bugajski and Assenova, *Eurasian Disunion*.

18. Jonathan Altman, 'Russian A2AD in the Eastern Mediterranean: A Growing Risk', *Naval War College Review* (Vol. 69, No. 1, 2016); Dmitry Gorenburg, 'Russia's Naval Strategy in the Eastern Mediterranean', Marshall Center, July 2019.

19. Hodges et al., 'One Flank, One Threat, One Presence'.

in areas as far away as Cuba, where the Soviet Union enjoyed little local material power.²⁰ Khrushchev's colourful description of West Berlin as a pressure point on which he could 'squeeze every time I want to hear them [the West] scream' comes to mind.²¹

The strategic challenge posed by Russia is not necessarily limited to the threat of large-scale aggression akin to an armoured thrust through the Fulda gap. Instead, what Russian policymakers want – and have wanted since the 1990s – has been a veto over the West's security decision-making.²²

Within this evolving strategic context, there are three scenarios in which Russia's air and missile effectiveness across Europe would be salient:

1. Peacetime emplacement of missiles by Russia to alter the behaviour of rivals in crucial regions. This might mirror Chinese activities in the South China Sea and is arguably already occurring in regions such as the Arctic and Eastern Mediterranean.²³ While not a kinetic act, the presence of strike assets can have a latent effect on the behaviour of parties in a contested zone.
2. A limited European contingency such as a relatively bloodless land grab by Russia. For example, a limited invasion of one of the Baltic states or a maritime clash in the Black Sea. In this context, the posturing and limited use of strike assets acts as a deterrent to an Alliance-wide response to Russian actions.
3. A full-scale invasion of one of the Baltic states or an attack on NATO in the High North. While relatively unlikely, such a contingency cannot be ruled out.

20. On the ways in which considerations regarding Cuba were often linked to West Berlin in US policy discussions, see John Lewis Gaddis, *Strategies of Containment: A Critical Appraisal of US National Security Policy During the Cold War* (Oxford: Oxford University Press, 2005).

21. Warren Cohen, *The New Cambridge History of American Foreign Relations* (Cambridge: Cambridge University Press, 2013), pp. 116–40.

22. Andrei Tsygankov, *Russia and the West from Alexander to Putin: Honour in International Relations* (Cambridge: Cambridge University Press, 2012), pp. 136–50, 230–50.

23. On the impact of Chinese missiles in the South China Sea, see Gregory Poling, 'The Conventional Wisdom on China's South China Sea Islands Is Dangerously Wrong', *War on the Rocks*, 10 January 2020. On the Arctic, see Matthew Melino and Heather Conley, 'The Ice Curtain: Russia's Arctic Military Presence', CSIS, 2020. Measuring the latent effect of emplacing strike assets on adversary behaviour is fraught with difficulties, but there is a general consensus that these assets do typically alter the day-to-day behaviour of regional contestants. For more on this, see Mark Gunzinger, Bryan Clark and Jesse Sloman, 'Winning in the Gray Zone: Using Electromagnetic Warfare to Regain Escalation Dominance', Center for Strategic and Budgetary Assessments, 5 October 2017.

The Middle East and North Africa

NATO will continue to face sub-peer challenges on its southern flank. If Iran obtained nuclear weapons, the original missions for which the EPAA was intended would become more important. Moreover, as Iran's successful efforts to abet the Yemeni Houthis in producing cruise missiles and UAVs has illustrated, non-state proxies can also represent a source of air and missile threats. Quasi-deniable support to proxies can allow a major power to tie down the resources of its rivals in a cost-effective way.²⁴

While there are currently no proxies in regions adjacent to NATO territory, such as North Africa, that constitute an air threat, the rapid increase in Houthi capabilities in the last decade shows how quickly an air threat can materialise. European territory has also been threatened from the air before, during Libyan strikes on the Italian island of Lampedusa in 1986.²⁵ An increased emphasis on great power competition does not necessarily entail the exclusion of other threats.

The Threat Landscape: Adversary Capabilities and Intentions

Russia and the Eastern Flank

Modern Russian military planners have developed a framework for understanding military competition that is more flexible than Soviet-era thinking, which tended to assume that conflict would necessarily be total in its nature.²⁶ By contrast, contemporary thinking envisions a spectrum of contingencies ranging from non-kinetic competition through to limited conflict, local wars, regional wars and full-scale conflict.²⁷ These conflicts are differentiated by both the intensity of the fighting and the number of actors involved – both of which are assumed to covary on the assumption that the higher the intensity of a conflict between Russia and one of its neighbours, the more likely Alliance commitments will come into play, which in turn incentivises further escalation.²⁸ It is noteworthy that both conventional and nuclear precision strike assets – grouped together under the rubric of strategic deterrence forces – have a key role in every stage of conflict, with the active or passive nature of the role depending on the stage

24. Ian Williams and Shaan Sheikh, 'The Missile War in Yemen', CSIS, 9 June 2020.

25. Alessandro Marrone and Karolina Muti (eds), *Europe's Missile Defence and Italy: Capabilities and Cooperation* (Rome: Istituto Affari Internazionali, 2021), p. 55.

26. On Soviet thinking, see Jack Snyder, 'The Soviet Strategic Culture: Implications for Limited Nuclear Options', RAND, September 1977.

27. Charles Bartles, 'Book Review of "Preparation and Conduct of Military Actions in Local Wars and Armed Conflicts" by Sergei Batushin', NATO Defence College, November 2018.

28. Johnson, 'Russia's Conventional Precision Strike Capabilities, Regional Crises, and Nuclear Thresholds', pp. 16–20.

to which a dispute has evolved.²⁹ For example, in a regional conflict conventional strikes can coincide with the posturing of tactical nuclear assets to deter retaliation in kind.³⁰

Conventional precision strike assets were first added to the list of tools serving strategic deterrent roles in the 2010 revision of Russia's military doctrine, and the relative weight accorded to them as a means of influencing the scope and tempo of a conflict has grown in subsequent doctrinal literature.³¹

The Russian mechanism of non-nuclear force deterrence is an integral part of the mechanism of force strategic deterrence. It is predicated on the concept of achieving outsized strategic effects by striking specific high-value targets. This entails the identification of criteria for the destruction of potential targets which are central to the functioning of a target society. In theory, threatening these targets and inflicting well-calibrated damage on them will make it possible to contain a specific stage of conflict by underscoring the risks of intervention in a local war involving Russia and one of its neighbours.³²

In the early stages of a conflict, the task of strategic deterrence forces is to identify the vitally important objects of a target country and either threaten to or actually inflict 'assigned damage'.³³ The term 'assigned damage' means limited damage intended not to cause mass casualties but rather to demonstrate a capacity to do so.³⁴ As a conflict escalates, the objective shifts from gradually escalating assigned damage to the infliction of unacceptable damage sufficient to compel an opponent or coalition of opponents to desist. At this stage of a conflict, the potential exists for the use of non-strategic nuclear weapons in a demonstrative or limited role both as a coercive measure and as a means of dissuading massed conventional or nuclear strikes on the

29. Kofman, Fink and Edmonds, 'Russian Strategy for Escalation Management'.

30. On the 'aggressive sanctuarization' of Russian assets, see Johnson, 'Russia's Conventional Precision Strike Capabilities, Regional Crises, and Nuclear Thresholds'.

31. Russian Federation, 'The Military Doctrine of the Russian Federation', 5 February 2010, <https://carnegieendowment.org/files/2010russia_military_doctrine.pdf>, accessed 22 April 2021; Ministry of Foreign Affairs of the Russian Federation, 'Basic Principles of State Policy of the Russian Federation on Nuclear Deterrence', 8 June 2020, <https://www.mid.ru/en/foreign_policy/international_safety/disarmament/-/asset_publisher/rpOfiUBmANaH/content/id/4152094>, accessed 22 April 2021; Nikolai Sokov, 'Russia Clarifies its Nuclear Deterrence Policy', Vienna Center for Disarmament and Non-Proliferation, 3 June 2020.

32. Y A Pechatnov, 'Metodicheskii podhod k opredelniuy sderzhivayushchego ushcherba s uchetom subyektivnykh osobennostei ego vospriyatiya veroyatnym protivnikom' ['Methodological Approach to Determining Deterrent Damage Considering Subjective Specialties of its Perception by Likely Adversary'], *Vooruzheniye I Ekonomika [Armaments and Economics]* (Vol. 3, No. 15, 2011).

33. Michael B Petersen, 'Strategic Deterrence, Critical Infrastructure and the Aspiration-Modernization Gap in the Russian Navy', in Jeffrey Mankoff (ed.), 'Improvisation and Adaptability in the Russian Military', CSIS, April 2020, pp. 29-32.

34. Johnson, 'Russia's Conventional Precision Strike Capabilities, Regional Crises, and Nuclear Thresholds', pp. 23-25.

Russian homeland.³⁵ A recent example of the use of missiles as a signalling device is the alleged use of a Russian Iskander-M against Azerbaijani forces in the recent Nagorno-Karabakh conflict. This was likely a means of signalling that the limits of Russia's patience were being reached.³⁶ The incorporation of strike assets into Russia's strategic framework is described in Table 1.

Table 1: The Role of Strike Assets Across the Spectrum of Conflict in Russian Military Thought

Conflict Phase	Pre-Conflict	Local War	Regional War	Global War
Role of strategic deterrence forces	<ul style="list-style-type: none"> • Posturing of conventional strike assets. • Limited demonstrative salvos using conventional strike assets. 	<ul style="list-style-type: none"> • Large-scale strikes on target state. • Limited strikes by conventional forces on potential interveners to inflict assigned damage and dissuade intervention. • Posturing of non-strategic nuclear weapons to deter counter-escalation. 	<ul style="list-style-type: none"> • Strikes to inflict unacceptable damage, possibly including limited non-strategic nuclear weapons. • Posturing of strategic nuclear forces to deter retaliation or further escalation. 	<ul style="list-style-type: none"> • Large-scale use of both conventional and nuclear assets.

Source: Michael Kofman, Anya Fink and Jeffrey Edmonds, 'Russian Strategy for Escalation Management: Evolution of Key Concepts', CNA, April 2020, pp. 8–18.

In addition to their role in strategic deterrence forces, many Russian strike assets are also dual-hatted as components of general purpose forces, which are tasked with impacting the balance of military force more so than political will. As such, assets such as the Iskander missile system or Kalibr salvos may also be used to target logistical nodes, C2 centres and key bases, or to break up enemy formations, as they did during the 2008 war with Georgia.³⁷ Because of the large number of potential targets and the limitations of Russia's arsenal of precision strike assets, targets are to be selected based on the aggregate system effects that their destruction would produce.³⁸ While there is relatively little work detailing precisely how system effects would be generated, it would be expected (based on a systems-centric view) that non-substitutable

35. Kofman, Fink and Edmonds, 'Russian Strategy for Escalation Management'.

36. *Foreign Policy News*, 'Russia Allegedly Used Iskander-M Ballistic Missile Against Azerbaijan', 5 April 2021.

37. Johnson, 'Russia's Conventional Precision Strike Capabilities, Regional Crises, and Nuclear Thresholds'; Lester Grau and Charles Bartles, *The Russian Way of War* (Washington, DC: Foreign Military Studies Office, 2016).

38. A V Skrypnik, 'Metodicheskii apparat ranzhiovaniya kriticheskii vaznykh obyektov protivnika v tselyakh resheniya zadachi silovogo strategicheskogo sderzhivaniya' ['Methodological Approach to Ranking

points of fragility, such as aerial refuelling tankers critical to sustaining a tempo of operations, fifth-generation fighters without which an IADS cannot be easily penetrated, C2 nodes, and key logistical points of failure such as railway bottlenecks would all be targeted.³⁹

While Russia's capacity to produce modern strike assets is limited, the pace of production appears to be increasing.⁴⁰ Moreover, Russia retains a large stockpile of legacy Soviet missiles such as the KH-55 and P-700, which, while of lower quality, add mass to the force and can at a minimum act as decoys for more sophisticated missiles. As of 2019, it was estimated that Russia fielded 1,300 standoff strike assets, including both cruise and ballistic missiles, which could strike European targets. This includes sea- and air-launched cruise missiles, as well as theatre ballistic missiles, including the capabilities of the 12 Iskander Brigades fielding the 9M723K1 quasi-ballistic missile.⁴¹ Russian theatre ballistic capabilities also include the air-launched KH-47M2 Kinzhal ballistic missile.

The Southern Flank

Traditionally, it was assumed that the threat posed by sub-peer competitors was from WMD-tipped ballistic missiles, as the inaccuracy of less sophisticated ballistic missiles made the use of any other form of payload ineffective.⁴² The ability of such weapons to insure conventionally weak states, such as Iran and North Korea, against a conventional attack has made procuring them an appealing option for these countries.⁴³ This challenge has not necessarily diminished, and may re-emerge depending on the fate of the Joint Comprehensive Plan of Action framework.

However, sub-peer states increasingly have capable conventional strike options, which may threaten European physical safety or political interests. Over the last decade, improvements in precision guidance have seen the circular error probable of Iranian ballistic missiles decline significantly.⁴⁴ This could render these missiles capable of executing limited but accurate attacks on military bases or critical infrastructure in parts of Europe from the territory of allies, such as Syria. Iran also fields an increasingly capable suite of cruise missiles such as the Ra'ad, Soumar

Critically-Important Objects of the Adversary to Resolve the Task of Strategic Deterrence Through the Use of Force'], *Vooruzheniye I Ekonomika [Armaments and Economics]* (No. 3, 2011), pp. 101–07.

39. *Ibid.*

40. *Defense World*, 'Sixty Kalibr Cruise Missiles Supplied to Russian Army in May-June', 26 July 2017.

41. Jonas Kjellen and Nils Dahlqvist, 'Russia's Armed Forces in 2019', in Fredrik Westerlund and Susanne Oxenstierna (eds), *Russian Military Capability in a Ten-Year Perspective – 2019* (Stockholm: FOI, 2019), p. 35.

42. Sidharth Kaushal, 'A Critical Enabler for Power Projection: Options for a UK Missile Defence Capability in an Age of Escalation Control', *RUSI Occasional Papers* (May 2019).

43. Vipin Narang, *Nuclear Strategy in the Modern Era: Regional Powers and International Conflict* (Princeton, NJ: Princeton University Press, 2014).

44. Ralph Savelsberg, "'Massive Improvement" in Accuracy of Iranian Missiles Over Scud-B', *Breaking Defense*, 15 January 2020.

and Ya-Ali.⁴⁵ The presence of missile factories in Syria operated by Iran's Islamic Revolutionary Guard Corps and the previous transfer of missiles such as the Zelzal-2 to Hezbollah raises the prospect of Iranian-made precision strike assets within range of parts of European territory.⁴⁶ While Hezbollah lacks the long-range strike assets to hit the European mainland, it can affect sea denial around its coastlines using the Yakhont anti-ship cruise missile.⁴⁷ Additionally, its rocket and missile arsenal is capable of hitting military infrastructure, such as RAF Akrotiri in Cyprus.

To be sure, these missiles do not represent an imminent threat to most of Europe, and are targeted mainly at regional actors, such as Israel.⁴⁸ However, this might change should either the policies of European states or Iran itself change. In many ways, it is more useful for an analysis to focus on capabilities than policy, given the relative fluidity of the latter.⁴⁹ For example, as recently as a decade ago, the idea of a Russian cruise missile threat to Europe was deemed so implausible as not to warrant a NATO response.

A range of North African actors also field credible missile capabilities. For example, Algeria fields the Russian Iskander system and Egypt is considering purchasing the supersonic BrahMos cruise missile.⁵⁰ While not currently a threat to Europe, if these states 'failed', non-state actors could acquire these assets, which could then pose a risk to NATO countries. Although this is unlikely, both states are signposted as being at relative risk in the global fragility index, which means that this eventuality cannot be ruled out.⁵¹

45. Shahryar Pasandideh, 'Under the Radar, Iran's Cruise Missile Capabilities Advance', *War on the Rocks*, 25 September 2019.

46. Britain Israel Communications and Research Centre, 'Hezbollah's Precision Missile Project', February 2019.

47. The group recently obtained these missiles during the Syrian civil war. Wargames suggest this could pose significant challenges in a regional clash. See Paul Bracken, *The Second Nuclear Age: Strategy, Danger and the New Power Politics* (New York, NY: Times Books, 2012); Barak Ravid, 'Hezbollah Moved Advanced Anti-Ship Missiles into Lebanon, U.S Officials Say', *Haaretz*, 10 January 2018, <<https://www.haaretz.com/.premium-u-s-hezbollah-moving-missiles-1.5307607>>, accessed 22 June 2021.

48. Although countries such as Turkey represent an exception. See NATO Supreme Headquarters Allied Powers Europe (SHAPE), 'NATO Patriot Mission in Turkey', <<https://shape.nato.int/ongoingoperations/nato-patriot-mission-in-turkey->>, accessed 22 April 2021.

49. On the question of basing policy on capabilities versus intentions, see John Mearsheimer, *The Tragedy of Great Power Politics* (New York, NY: W W Norton, 2001), pp. 15–20.

50. *Army Recognition*, 'Russian Iskander-E Ballistic Missiles Delivered to Algeria', 12 March 2018, <https://www.armyrecognition.com/march_2018_global_defence_security_army_news_industry/russian_iskander-e_ballistic_missiles_delivered_to_algeria.html>, accessed 8 April 2021; *Al-Monitor*, 'Egypt Considers Purchase of Indian Missile System', 2 February 2021, <<https://www.al-monitor.com/originals/2021/02/egypt-india-russia-made-brahmos-missile-system-weapons.html>>, accessed 23 June 2021.

51. Fragile States Index, 'Country Data', <<https://fragilestatesindex.org/country-data/>>, accessed 8 April 2021.

The threat environment on the southern flank is currently more benign than on the eastern flank. However, the possibility that this might change makes preparation to meet a limited challenge from this vector necessary.⁵² In effect, NATO IAMD needs to meet the requirements of a 'one-and-a-half' conflict posture – a significant threat on its eastern flank and a more limited southern threat.

52. Historically, this has informed the policy of European states in relation to Iran's missile programme. See, for example, Cornelius Adebahar, 'Europe Needs a Regional Strategy on Iran', Carnegie Europe, 13 May 2020.

II. Responding to the Challenge: Priorities for NATO's Air Defence

IN ORDER TO adapt to the strategic and operational challenges described in Chapter I, NATO will need to conceptually reframe its approach to air defence in Europe. As opposed to being a stand-alone task, air defence will need to be nested within a concept of operations that explicitly links it to a campaign plan through which a joint force ends the strategic threat and in which active air defence is only part of the solution.

Air Defence in its Operational Context

The fundamental purpose of air defence is the protection of critical assets, own and partner forces and populations from damage caused by objects arriving in the atmosphere, either ballistically or aerodynamically, and to do so for long enough to negate⁵³ the air threat by other means.

Making this purpose clear ensures that the discussion, as well as strategic, operational and tactical planning does not get muddled by the distinctions between 'air defence', 'ballistic missile defence', 'cruise missile defence', and equivalent differentiations.⁵⁴ While those differences are important from certain technical perspectives, the threat they present when combined is what must be considered, planned for and defeated. Thus, IAMD must seamlessly address all the different threat categories.

Furthermore, any reasonable assessment of the balance of inventory between air threat attack weapons and air defence weapons will usually lie in favour of the former. This is because attackers can concentrate forces against a specific target more rapidly than defenders and because strike assets typically tend to be less expensive and less complex, mainly when aimed at non-mobile targets. Cruise missiles such as the Kalibr, for example, are estimated to cost Russia \$980,000 each, compared to the \$5-million unit cost of the upgraded PAC-3 interceptor. This cost asymmetry will be compounded by the fact that defenders will need more interceptors than the attacker has threats to execute a two-shot doctrine.⁵⁵ The Kalibr is, moreover, a relatively new system and older missiles such as the KH-55 stockpiled from the Soviet-era cost even less

53. 'Negating' aims to reduce the enemy air threat to as low as possible effectiveness, ideally to zero, but at a minimum to the point where any remaining damage capability can be tolerated without impeding the defender's campaign objectives.

54. Based on the professional experience of one of the authors as director of the Pentagon's Joint Integrated Air and Missile Defence Organization.

55. See, for example, Roger McDermott, 'Russia's Futuristic Military Plagued by Old Problems', Jamestown Foundation, 29 July 2016'. On the PAC-3, see Jen Judson, 'House Panel Wants Cheaper Patriot Missile', *Defense News*, 5 June 2019. This claim is also based on the professional experience of one of the authors.

– merely the price of maintenance. When one considers that UAVs can be used as decoys, as Iranian attackers did at Abqaiq and Khurais,⁵⁶ the average unit cost of air threats diminishes further. For example, modelling in an Asia-Pacific context suggests that a salvo of just under 60 cruise missiles could result in the destruction of all the hangars, fuel tanks and parking areas of a major air base, such as Andersen AFB in Guam.⁵⁷ Moreover, the risk to critical assets posed by a robust peer competitor arsenal remains real even when one assumes a generous single shot per kill (SSPK) of 0.8 probability for air defence systems.⁵⁸ There are meaningful differences between the European theatre and the Asia-Pacific – most notably Europe's strategic depth and the greater number of airfields by comparison to countries like Japan.⁵⁹ However, a force of 1,300–1,500 standoff assets could still likely overwhelm air defences at key political and military pressure points. Russia could eliminate NATO Combined Air Operations centres such as Rammstein and Udem and disrupt activity at key ports with relative ease. It could also strike critical national infrastructure such as power stations and chemical plants, the loss of which could produce cascading economic effects.⁶⁰ Moreover, the SSPK of air defences is almost certainly much lower than 0.8, especially for supersonic and hypersonic missiles like the P-800 and Zircon which leave air defenders with limited warning times.⁶¹

Thus, it is impractical for the defenders to simply defeat incoming missiles, as they will run out of defence capability before the enemy runs out of threat capability. 'Playing catch' is not a viable approach, as illustrated by the recent Saudi campaign against Houthi attacks, in which a reasonable rate of intercept by Saudi defences nonetheless failed to prevent the Houthis from exacting a significant toll on Saudi national infrastructure.⁶²

As such, the requirements of air defence must always be determined not in absolute terms but relative to the avenues available for suppressing or negating the air threat using both strike and passive defences, as well as the requirements of a more comprehensive campaign plan. Planning and execution must account for this.⁶³

56. Uzi Rubin, speech at RUSI Space and Missile Defence Conference, 27 February 2020.

57. Eric Heginbotham et al., *The US–China Military Scorecard* (Santa Monica, CA: RAND, 2015), p. 133.

58. Thomas Shugart and Javier Gonzales, 'First Strike: The Chinese Missile Threat to US Bases in Asia', conference transcript, Center for a New American Security, 28 June 2017, p. 12.

59. Press, presentation at RUSI Missile Defence Conference.

60. Filipe Batista e Silva et al., 'HARCI-EU, a Harmonized Gridded Dataset of Critical Infrastructures in Europe for Large-Scale Risk Assessments', *Scientific Data* (Vol. 6, No. 126, 2019).

61. Kyle Mizokami, 'Russia's New Hypersonic Zircon Travels at Nearly Two Miles a Second', *Popular Mechanics*, 26 December 2018. While optimised for an anti-ship role, the P-800 has a demonstrated ability to play a land attack role as well, see Tom Parfitt, 'Russia Fires New Cruise Missile at Syrian Rebels', *The Times*, 17 November 2016.

62. Jack Watling and Sidharth Kaushal, 'The Democratisation of Precision Strike in the Nagorno-Karabakh Conflict', *RUSI Commentary*, 22 October 2020; Williams and Sheikh, 'The Missile War in Yemen'.

63. Against a sufficiently large adversary salvo, even efficient air defences can only hold out for so long. For a discussion of this challenge in the parallel east Asian context, see Shugart and Gonzales, 'First Strike'. Russian planning tends to already take this temporal factor into account.

The temporal dimension is key here, given that both active and passive defences can only be expected to hold for long enough for other capabilities to terminate the threat before air defence capacity is exhausted, given the likely rates at which interceptors protecting any given area will be expended. Moreover, while some capabilities, such as runways, can be reconstituted relatively quickly, it is difficult and time consuming to replace others, such as fifth-generation aircraft or critical powerplants.

This raises the complex question of which aspect of NATO's scheme of manoeuvre is most vulnerable.⁶⁴ If the primary risk is the disruption of political will in key Alliance members, then the defence of strategic targets is most key, and the task of air defenders should be defined relative to an assessment of how long political will can last. If the primary risk is the destruction of critical supporting infrastructure in an operational echelon, then demand signals from component commanders should determine the requirement.⁶⁵ Russia will use the same assets to strike both categories of targets, so it is critical to know which risk should be prioritised at different points during a campaign.⁶⁶

Literature on decision-making in democracies suggests that public support for conflict is difficult to secure but, once it has been, it tends to hold even under coercive adversary pressure.⁶⁷ Therefore, strategic defence should receive priority in the periods when NATO is deliberating its response to a provocation, while the defence of operational rear areas should be prioritised once public support has crystallised and formations are on the move, given the vulnerability of key supply lines to disruption.⁶⁸ In summary, it is the purpose of air defence to prevent an effect on an aspect of the campaign, be that the erosion of political will or a reduction in battlefield capabilities, and which effect takes precedence at a given point can only be deduced once air defence is situated within a campaign plan.

See, for example, Michael Kofman, 'It's Time to Talk About A2AD: Rethinking the Russian Military Challenge', *War on the Rocks*, 5 September 2019.

64. The scheme of manoeuvre is the operational plan that underpins the operational-level coordination at the joint level. See NATO, 'Allied Joint Doctrine for the Planning of Operations', AJP-5, 24 May 2019, para. 4-50.
65. On the distinct ways in which strike assets can be used, see Robert Pape, *Bombing to Win: Airpower and Coercion in War* (Ithaca, NY: Cornell University Press, 1996). Component commanders currently feed into the air defence critical asset list through the Joint Force Command, but not ballistic missile defence, which is coordinated by SHAPE. However, there is no system for temporally managing political and military demands in the same theatre, because territorial defence and theatre defence are segmented in expeditionary operations. See NATO, 'Allied Joint Doctrine for Air and Space Operations', AJP-3-3, April 2016, para. 1-9.
66. Kofman, Fink and Edmonds, 'Russian Strategy for Escalation Management', p. 19.
67. See, for example, Dan Reiter and Allen Stam, *Democracies at War* (Princeton, NJ: Princeton University Press, 2002).
68. For example, see Ben Hodges, Tony Lawrence and Ray Wojick, 'Until Something Moves: Reinforcing the Baltic Region in Crisis and War', CEPA, April 2020.

The only scenario that might represent a partial exception to this rule – in which air defence alone can achieve an operational or strategic aim – is the prevention of limited engagement.⁶⁹ In this scenario, a small number of threats are launched against Europe as a political statement, and whose successful defeat would visibly deny the aggressor's goal. Even in this case, however, certain air threats, such as supersonic and hypersonic cruise missiles, may require a combination of active defence and dedicated strike, because even moderate numbers of these missiles can overwhelm air defences due to their high speeds.

Thus, it is crucial that active air defence is not treated as a stand-alone subcomponent of an air campaign, as is the case in recent planning.⁷⁰ To be effective, planning for the defensive counter-air mission needs to be carried out on the assumption that air defences can provide only finite temporal windows for other forces to act.

Gaps in NATO's Current System

A number of conceptual and practical gaps impede the creation of a defensive system that is compatible with the requirements of a multi-domain threat.

First, discussions of the requirements that militaries will face tend to assume air and missile defence availability as a consistent enabler.⁷¹ In a NATO context, however, this cannot be assumed. Competing demands on different assets will strip at least some parts of the battlefield of support for some of the time. Again, the temporal dimension is key.

Second, air defence and strike are highly – and in this paper's view artificially – stove-piped from each other in NATO doctrine.⁷² While some attention is paid to the potential for units to responsively strike the sources of ballistic threats to their own operations, this understates the contribution such assets can make.⁷³ Long-range fires could, for example, ease the burden on air defenders tasked with defending strategic targets by neutralising launchers or holding adversary airbases at risk. Similarly, naval vessels could significantly exacerbate Russia's missile delivery system bottlenecks by targeting the littoral vessels that serve as maritime launch platforms. If planned in tandem with strike, active defence could then be massed against vectors that cannot be suppressed for political or practical reasons. For example, if it was deemed viable to suppress

69. The 2019 Abqaiq and Khurais attacks are a recent example. The scope and selected targets of the Iranian raid were carefully chosen to demonstrate capability while limiting damage. While the effect on Saudi oil production was measurable, the political message was of far greater importance. See Rubin, speech at RUSI Space and Missile Defence Conference.

70. For example, see NATO, 'Allied Joint Doctrine for Air and Space Operations', para. 1-9. Air defence, under the rubric of defensive counter-air, operates in tandem with offensive counter-air and SEAD.

71. See, for example, Wilson Blythe, 'Warfighting at the Division and Corps Level', presentation at HQ US NATO, Sembach, Germany, 29 March 2021. Cited with permission from the author.

72. NATO, 'Allied Joint Doctrine for Air and Space Operations', para. 1-9.

73. For a discussion of responsive targeting of threats at a corps level, see Blythe, 'Warfighting at the Division and Corps Level'.

ship-based launchers in the High North but not ground-based launchers in Kaliningrad, active defences could be massed against threats emanating from this vector.

Third, the stated aim of convergence of capabilities and effects in the context of multi-domain integration across the campaign sequence – extending beyond synchronisation – tends to be excluded from IAMD discussions, which tend to focus on interoperability.⁷⁴ While interoperability is necessary for convergence, it is not sufficient.⁷⁵ Convergence entails the interaction of system components in a way that erodes conceptual boundaries between asset classes and allows for functional agnosticism.⁷⁶ For example, in a strike context, an F-35 may cue in a ground-based High Mobility Artillery Rocket System asset or conduct a strike itself – its role as an intelligence, surveillance, reconnaissance, targeting and ground attack asset are context dependent.⁷⁷ A similar convergence could in principle be envisioned when countering the air threat. In this scenario, the F-35 could track missiles for a ground-based system, attempt an intercept itself, or leave the defensive task to a system, such as Patriot, while conducting a strike on the launcher.

Practical concerns also matter. Existing air and missile defence systems in Europe have to cover significant territory with densely distributed significant potential targets to meet the aim of Alliance interoperability. Systems as heterogeneous as Sky Sabre, Patriot, Aegis (both afloat and ashore) and NASAMs will likely have significant difficulties talking to one another and to assets across the joint force in the near term.⁷⁸ For example, while the SAMP/T has taken part in joint exercises alongside Patriot systems, the ability of the two systems to speak to one another is open to question.⁷⁹ Similarly, while exercises such as Formidable Shield have tested NATO C2 structures, multi-country maritime IAMD systems have some way to go before they are truly networked – that is, able to communicate seamlessly with each other as opposed to being coordinated through a Maritime Air Operations Centre.⁸⁰ This shortfall is likely to be underscored by adversary developments that will strain both capacity and capability. On the capability front, the ‘birth to death’ tracking of targets such as hypersonic glide vehicles will require enablers on a geographic scale that can only be provided at a NATO level, and which will not be provided if individual countries generate self-contained national defence systems. There is precedent for Alliance-level procurement of key enablers, with both tankers and AWACS procured at a NATO level.⁸¹

74. For example, see Vivienne Machi, ‘NATO Missile Defence Systems Strive for Interoperability’, *National Defense*, 31 July 2017.

75. Jack Watling and Sean MacFarland, ‘The Future of the NATO Corps’, *RUSI Occasional Papers* (January 2021), p. 18.

76. *Ibid.*

77. Shawn Snow, ‘Marines Connect F-35 to HIMARS for the First Time’, *Marine Corps Times*, 5 October 2018.

78. Brad Hicks, presentation at RUSI Space and Missile Defence Conference, 26 February 2020.

79. Francis Mahon, ‘NATO Allies Must Work Closely Together, But Do Their Missile Defence Systems?’, *Reuters*, 2 October 2014.

80. Hicks, presentation at RUSI Space and Missile Defence Conference.

81. NATO, ‘AWACS: NATO’s Eyes in the Sky’, 16 March 2021; NATO, ‘First Aircraft of NATO’s Multi-Role Tanker Fleet Lands at Eindhoven Airbase’, 30 June 2020.

Moreover, large peer competitor offensive arsenals will likely provide the Alliance with capacity and capability challenges that cannot be affordably met at a national level. The development of broad volume and wide velocity detection systems capable of tracking targets ranging from cruise missiles to hypersonics and ballistic systems and cueing shooters across domains will require a resilient network of enablers in the air, stratosphere and outer space.⁸² National efforts at creating self-contained air defence systems will impede the creation of such an Alliance network by channelling resources that could be pooled to develop or purchase much more limited individual networks.⁸³ Functional specialisation, in which individual countries contribute towards the production of well-delineated parts of an integrated solution such as radar or satellites, will inevitably be necessary for future European territorial air defence. While political concerns and defence industrial opposition may act as an impediment, the first steps towards such an evolution might be taken by expanding the remit of NATO missile defence to include all aspects of air defence and allowing countries the option to pool individual specialised resources in a NATO system, either through situation-based commitments comparable to the EPAA model or joint procurement, as is the case with AWACS and tankers.

It is, of course, likely that individual countries will still seek to develop nationally-owned capabilities, but this can co-exist with a NATO-owned system. For example, Germany subordinates one of its air defence missile groups to Dutch command while it retains control over the other three.⁸⁴ As such, pooling resources for a truly NATO-owned system, which would enhance functional specialisation, can co-exist with nationally-owned assets, although it would imply a rebalancing of priorities within countries. Given that working through NATO for BMD is already the preference of countries such as the UK, this could create avenues by which a critical mass of willing countries could shift NATO policy by creating structural incentives for others to buy into a shared integrated system.

A final consideration worth signposting is the relationship between NATO and the EU. Ongoing discussions about European strategic autonomy represents a potential opportunity to achieve specialisation but could also become another potential risk to effective IAMD because it could duplicate the integration of the IAMD command structure and capabilities, as well as strategies for asset protection, choice of defended assets, counter-strike target allocations and

82. On a space-based sensor layer, see Ian Williams, 'Achilles Heel: Adding Resilience to NATO's Missile Defence System', CSIS, 5 August 2019. Stratospheric assets, such as JLENS, and aerial assets, such as UAVs, will likely be critical to tracking increasingly fast, low-flying missiles such as the Zircon. See Tom Karako and Ian Williams, 'JLENS Future Bleak, But the Need for Capability Remains', CSIS, April 2016. On UAVs and other air-breathing assets in an air defence role, see Sidharth Kaushal and Alexandra Stickings, 'Time Sensitive Targets, Space and Missile Defence', RUSI Conference Report, 22 April 2020.

83. Based on professional experience of one of the authors as a senior official responsible for delivering NATO missile defence.

84. Justyna Gotkowska, 'Rebuilding Germany's Air Defence Capabilities: On the Eve of Crucial Decisions', Center For Eastern Studies, 17 May 2018.

overall campaign objectives and timelines.⁸⁵ Projects such as TWISTER, emphasising a locally developed endo-atmospheric interceptor and a space-based sensor system, could enhance NATO defence but risk wasting resources duplicating interceptor capabilities that already exist or are in development in the form of US assets like THAAD and the SM-6 IB.⁸⁶ They also raise the question of how – if at all – a future European sensor layer will be integrated with NATO C4ISR. This represents a challenge given that one of the critical elements of the defence of Europe is NATO's Article 5, and the explicit support of non-EU defence entities such as the US, Canada, and (now) the UK, which are not part of the various proffered expansions of the EU's Common Defence and Security Policy structure, is enabled and engendered by that commitment.

Moreover, less ambitious projects emphasising functional specialisation in areas like counter-rocket, artillery and mortar (C-RAM), counter-UAV, or cost-effective methods to achieve cruise missile defence could allow individual members operating through a Permanent Structured Cooperation framework to reinforce NATO's deterrent posture by targeting the gaps within it.

The first task that must be accomplished in resolving the approach to air defence among these actors is to agree on the methods for strategy planning, operational coordination and provision of capabilities. This will involve rationalising and integrating decisions, plans and cooperation. Against great power adversaries who can launch massed salvos and have the attacker's luxury of choosing the time and place of an engagement, the absence of this rationalisation process will increase the risk that inevitably results in isolated air defence sub-systems being defeated in detail.

The Path Forward: Offence–Defence Integration

As currently structured, NATO's air defence system would struggle to cope with limited Russian salvos in the context of a 'local war', or unacknowledged or proxy attacks against major European symbols such as government buildings, transportation centres or seaports. It would not be able to withstand a massive Russian attack across the eastern approaches.

As a primarily south-facing system built around BMD against an Iranian launch, NATO BMD has relatively little capacity to defend every target which might be deemed a political pressure point by Russia in a limited contingency conflict. The effective use of ballistic missiles against key points in NATO's air defence network could, in turn, enable salvos of air-breathing threats. If a conflict should escalate, it is unclear whether NATO can mitigate massed salvos.⁸⁷

85. On European strategic autonomy, see Vincenzo Camporini et al., 'European Preference Strategic Autonomy and the European Defence Fund', ARES Report, November 2017.

86. On TWISTER, see Sebastian Sprenger, 'Germany Joins Nascent European Push to Shoot Down Hypersonic Missiles', *Defense News*, 30 November 2020; on the SM-6 IB, see Joseph Trevithick, 'Navy SM-6 Will Attempt to Shoot Down a Mock Hypersonic Missile', *The Drive*, 14 April 2021.

87. Hodges et al., 'One Flank, One Threat, One Presence'; on non-strategic nuclear weapons, see Kofman, Fink and Edmonds, 'Russian Strategy for Escalation Management'; and Kaushal, 'A Critical Enabler for Power Projection'.

As previously discussed, air threat defensive measures need to account for how long adversary effects must be denied for other capabilities to negate the threat and ensure the progress of the Alliance's campaign to end the conflict on favourable terms. This analysis leads to three interrelated tasks:

1. Identifying particular assets that need to be defended well before a conflict, based on their importance to national will or effective manoeuvre.
2. Efficiently allocating active and passive defences to avoid creating redundancies. It may be the case that many assets can be defended through deception or resilience while others cannot, so a shared awareness between the military and civil entities responsible for passive defence and air defenders can create efficiencies. NATO initiatives, including the setting of baseline requirements for civil resilience, are attempting to achieve this via a series of other threat responses, including against unconventional attack.⁸⁸ Leveraging the national-level capabilities created by these initiatives will require air defenders and the services and agencies responsible for resilience to share an operating picture and a professional lexicon – something that involves pre-planning and exercising, but not necessarily shared control of all assets.⁸⁹
3. Offence–defence integration, to unify strike and defensive assets under a single campaign plan that delineates how strike and defensive forces can support each other and create a mutual understanding and employment of capabilities and limitations to generate operational efficiencies that will provide sufficient protection 'long enough to negate the air threat by other means'.

Some aspects of these tasks, in particular offence–defence integration, pose political challenges. However, this was true of other major conceptual shifts within the Alliance, for example, with the adoption of Follow-on Forces Attack to harmonise NATO planning with AirLand Battle.⁹⁰ When a clear strategic case could be made that Warsaw Pact preponderance allowed few alternatives, however, the approach was adopted despite misgivings regarding escalation. This example shows how flexibility on which parts of a campaign each country would specialise in helped secure buy-in. Allowing certain governments to exempt themselves from politically sensitive deep-strike capabilities and instead specialise in shorter-ranged systems or ISR enablers was crucial to generating unity.⁹¹ Functional specialisation, then, may go some way to facilitating militarily necessary but politically sensitive adaptations, as it allows countries to specialise in ways consistent with their political constraints.

To be of value in a campaign against a great power, air defence must be nested within a concept of operations that specifies which effects are to be provided by offences and which enabled by defences. This integration of purpose will make it possible to delineate commanders'

88. NATO, 'Resilience and Article 3', 11 June 2021.

89. On integration without centralised control, see Comfort, *The Dynamics of Risk*, pp. 50–60.

90. US Congress, Office of Technology Assessment, 'New Technology For NATO: Implementing Follow on Forces Attack', June 1987, pp. 110–20.

91. *Ibid.*

responsibilities within the different phases of the overall campaign – given that these responsibilities can only be established in relation to both each other and an overarching operational plan.⁹² Within offence–defence integration, the allocation of purpose is best described as ‘the use of offensive forces to attack the missile threat throughout the kill chain and defensive forces to intercept targets in any phase of flight’.⁹³ This will allow defenders to prioritise key defence requirements for critical assets and likely targets and provide the conceptual capstone needed to integrate active defences and passive defences more effectively with other capabilities – in time, space and purpose.

SEAT and SEAD

Effective offence–defence integration will require the incorporation of SEAT into NATO’s lexicon.⁹⁴ At present, the emphasis of strike assets appears to be on SEAD, with air defence largely operating in parallel.⁹⁵ The SEAT concept would be distinguishable from traditional offensive counter-air by a specific focus on threats against own key assets, and as a critical enabling mission with an emphasis on creating temporal windows of opportunity in tandem with air defences. This stands in contrast to current NATO planning for offensive counter-air, which views targeting launch platforms as a subcomponent of a much wider set of objectives to be achieved primarily through airpower.⁹⁶ Instead of defining core airpower missions as defensive or offensive based on the tactics involved (strike versus interception, for example), they should be categorised based on their effect – negating an enemy air threat or establishing air control. Within this context, SEAT would amount to a joint defensive measure planned alongside active defence. By contrast, SEAD and strike missions aimed at establishing air control – rather than negating an air threat – would fall under the rubric of offensive actions.

92. This process of integration and handing-off of responsibilities between heterogenous forces has already been conceptualised with respect to campaign-level objectives such as joint-SEAD. See, for example, James R Brungess, *Setting the Context: Suppression of Enemy Air Defences in an Uncertain World* (Maxwell, AL: Air University Press, 1994). The integration of special forces into the Gulf War SCUD hunts, along with the emplacement of missile defences in vulnerable areas, represented an embryonic and ad hoc attempt at bringing together offensive and defensive assets to counter an air threat, although it fell short of integration. See, for example, Mark E Kippit, *Crossbow and the Gulf War Counter-Scud Efforts: Lessons from History* (Maxwell, AL: Air University Press, 2003) pp. 10–15.

93. Brian Green, ‘Offence–Defense Integration for Missile Defeat’, CSIS, 6 July 2020, p. 11.

94. The SEAT concept is known as offensive-counter air in US doctrine, but its tenets are not reflected in the same manner within NATO. While NATO does have a concept of offensive-counter air, it encompasses everything from air-to-air engagement to SEAD, meaning that, when subordinated to this concept, SEAT cannot be carried out as a properly resourced joint activity. See NATO, ‘Allied Joint Doctrine for Air and Space Operations’, para. 3-1. It is also worth noting that US doctrine does not address in detail the temporal nature of integrating threat reduction with the duration of air defence capability. See US Joint Chiefs of Staff, ‘Joint Publication 3-01: Countering Air and Missile Threats’, 21 April 2017.

95. Blythe, ‘Warfighting at the Division and Corps Level.’

96. NATO, ‘Allied Joint Doctrine for Air and Space Operations’, para. 3-1.

Introducing SEAT as a parallel mission with the temporal requirement of effectively accomplishing it before air defences are overwhelmed would alert the component commands that hold strike assets of their crucial role in facilitating an effective air defence. In turn, this will be an enabler for tasks such as SEAD, insofar as these missions will require the infrastructure of power projection to be protected against attack. Countries such as China, which have always had to worry about a superior air threat, already have precisely such a framework in mind as a priority that will orient the use of its air force, navy and strategic rocket forces at the outset of a campaign.⁹⁷ In many ways NATO needs to catch up conceptually.

As such, prior to SEAD, multi-domain operations should begin with a priority focus on SEAT. While every launcher from which an air threat can emerge can hardly be suppressed, many can be attrited to enable the more effective operation of active and passive defences. For instance, Russian surface vessels are highly vulnerable and generally operate in limited areas.⁹⁸ Strategic bombers can potentially be held at risk in their airbases, given their size and the difficulty in hiding them.⁹⁹ Subsurface assets are more difficult to suppress, though their suppression will likely form part of an anti-submarine warfare campaign in any case.

The importance of SEAT as a complement to passive defence will likely be necessary, albeit in a tailored form, in at least some more limited contingencies as well. As the speeds of missiles increase, the number of missiles needed to overwhelm defences decreases given the short warning times afforded to defenders and their limited engagement times.¹⁰⁰ Thus, a limited declaratory strike by Russia involving, for example, the hypersonic Zircon missile might most effectively be defended against by combining active defence with a more limited form of SEAT – forcing a given vessel to move before it can launch all its missiles through a ‘shot across the bows’, for example. Alternatively, against a sub-peer proxy, suppression of the launchers responsible for a strike functions as a combination of denial and punishment – an approach central to Israel’s operations against Iranian proxies in Syria.¹⁰¹

There is another latent strategic advantage to SEAT in limited contingencies. Work on deterrence suggests that initiators tend to act when they believe they can control escalation – for example,

97. The Chinese analogue to SEAT is the Joint Firepower Strike Campaign. See Mark Cozad, ‘PLA Joint Training and Implications for Future Expeditionary Capabilities’, RAND, 2016, p. 5; Shugart and Gonzalez, ‘First Strike’, pp. 1–5.

98. Kaushal, ‘Persistent Engagement and Strategic Raiding’, p. 35.

99. This has been the case for a long time. For example, see RAND, ‘Vulnerability of US Strategic Air Power to an Enemy Surprise Attack in 1956’, 1953. While early discussions revolved around suppression by nuclear attacks, conventional weapons could target strategic bombers under current circumstances.

100. On the warning times afforded by the Zircon, see Kyle Mizokami, ‘Russia is Putting Hypersonic Missiles on its Battlecruisers’, *Popular Mechanics*, 16 April 2016.

101. Sebastien Roblin, ‘Israel Demolishes Iranian Guard Corps HQ in Syria Following Rocket Attacks’, *National Interest*, 24 November 2019.

the Berlin Blockade of 1948 and the Taiwan Strait Crisis in 1958.¹⁰² A SEAT framework that raises the possibility in the opponent's mind of large-scale offensive strikes and perhaps even pre-emption (even if this is not the intended effect) forces opponents to consider that limited actions can lead to uncontrolled escalation. Thus, while seemingly destabilising, these risks can lend themselves to strategic stability by dispelling initiator notions that escalation can be carefully managed through assigned damage.¹⁰³

While the 1991 Gulf War illustrated the difficulties of suppressing mobile ballistic missile launchers, advances in technology – most notably the ability of systems, such as the Space-Based Infrared System (SBIRS),¹⁰⁴ to identify heat signatures after launch and potentially pass their locations on to high-velocity and quick-reaction strike assets – could allow these systems to be neutralised more responsively than was previously possible.¹⁰⁵

Critical Planning Decisions

The likely practical limitations on interceptor inventory mean that not every square inch of agreed-upon Allied territory can be defended against all threats. Therefore, it becomes necessary to determine what is sufficiently critical to require active air defence.

This brutal decision to deem certain assets less critical than others and leave them undefended, commonly encountered in BMD planning, will be encountered in IAMD planning due to the fact that increasing threat capabilities and enemy threat inventories will drive equivalent planning necessities.

The decision calculus is dependent on the difference between the list of assets for which defence is desired – the critical asset list (CAL) – and the list of assets which in time, place, purpose and inventory can realistically be defended – the defended asset list (DAL).¹⁰⁶ An agreement of what will be on each list needs to be achieved before the first shot is fired. The time available for identifying key points of failure, weighting of trade-offs, and subsequent decision-making will be circumscribed once a conflict begins. At a minimum, the criteria and methodology by which inclusion in the DAL is decided must be predetermined, should the exclusion of certain Allies' territory in peacetime prove fractious.

102. See George and Smoke, *Deterrence in American Foreign Policy*, chapters 1–4, 8.

103. This seemingly paradoxical idea is best captured by Glenn Snyder's stability–instability paradox. See Glenn Snyder, *Deterrence and Defence* (Princeton, NJ: Princeton University Press, 1961), pp. 20–30.

104. The SBIRS is a satellite constellation, now operated by the US Space Force, that provides early missile warning using infrared surveillance. For more information, see Lockheed Martin, 'Space Based Infrared Surveillance SBIRS', <<https://www.lockheedmartin.com/en-us/products/sbirs.html>>, accessed 14 April 2021.

105. Jerome Dunn, chief architect of counter-hypersonics at Northrop Grumman, presentation to RUSI, 27 February 2020.

106. Department of the Army, 'ADRP 3-37 Protection', July 2019, p. 2-10.

In NATO IAMD operations, this DAL must be agreed by all Allies for each major planning scenario. This will be necessary for the analytical work needed to generate a CAL and DAL, as the criteria and methodology for generating these lists cannot be put in place during a crisis. This requires NATO members to resolve sensitive discussions before a conflict. In such political discussions, a prior strategy serves as a focal point by which the criteria for selecting targets for defence prioritisation are critical to breaking logjams. The key decisions that will set the agenda for subsequent planning must be made before actual conflict, and thus in an environment of anticipation rather than knowledge of the real threat. Once a CAL and DAL have been determined, a decision regarding which resources demand active defences and which will require passive defences needs to be made. Securing assets that will not receive the coverage of active defences will depend on other techniques, such as civil and military resilience in which NATO work is ongoing. As such, this paper proposes that the NATO Air and Missile Defence Committee set up a schedule of regular meetings with the Civil Emergency Planning Committee.¹⁰⁷

The timescales involved mean that these decision-making assignments and processes, from initial threat-level determination to real-time engagement decision authorities, will have to be agreed by all partners before the conflict starts.¹⁰⁸ The earlier decisions will mean agreeing to allocate and release resources at certain threat levels and Indications and Warning events. In contrast, the later decisions will mean assigning a single decision system the authority to conduct the defence – with the concomitant loss of national authority that must ensue. A system coming under Alliance control as conflict potential escalates – like the Aegis Ashore sites – is critical.

The next evolutionary step will be rationalisation. SEAT missions must enable an active defence system that more effectively rationalises limited resources. This requires a keen understanding of how long it will take other military operations to convince the aggressor to stop its behaviour. Moreover, it will require the ruthless prioritisation of the critical capabilities and nodes necessary to maintain Western political cohesion and enable military operations.¹⁰⁹ This will occur in a

107. NATO, 'Civil Emergency Planning Committee (CEPC)', 15 November 2011.

108. Professional experience of one of the authors in planning air defence.

109. This is based on the assumption that a NATO response to aggression will take time to mobilise and that for it to mobilise population cohesion and critical mobilisation infrastructure need to be maintained. The deterrence of allied intervention by the fear of counter-value attacks was a factor in Chamberlain's appeasement of Hitler. See Gerald Lee, "'I See Dead People': Air Raid Phobia and Britain's Behaviour in the Munich Crisis', *Security Studies* (Vol. 13, No. 2, 2003), pp. 230–72. This paper bases its assessments on the potential political risk of countervalue attacks on the already low levels of population support within NATO countries for taking significant risks to meet Alliance commitments. See polling data in Hammes, 'The Melians' Revenge', p. 2. This paper's assumptions regarding the fragility of infrastructure on which a response could be mobilised is based on the fact that it is vulnerable to being overwhelmed by the needs of mobilisation even without an enemy attack. See Hodges, Wojick and Lawrence, 'Until Something Moves'. It stands to reason that already overstretched infrastructure would be further hampered by missile attacks. Airbases are less vulnerable, but subject to disruption sufficient to ensure that a NATO SEAD campaign takes longer to begin than would otherwise be the case and that NATO potentially fails to establish air superiority over a theatre such as the Baltics before the

context where the number of preferred defended targets, civilian and military, will invariably exceed defensive capacity – making it critical for political disagreements to be harmonised well ahead of a crisis.

Once engaged, the decision process focuses on demand: how the Alliance will allocate or restrict the use of interceptors and strike assets, depending on the threat level, the intensity of the enemy salvos, the remaining interceptor and strike inventory, and the damage already incurred or avoided. This continues as inventory reduces and threats continue:¹¹⁰ how strike and offensive assets are allocated between damaging or negating the enemy's air threat capability (such as launchers, airfields and missile storage areas) and other targets equally vital to achieving conflict termination.¹¹¹

Critical Operational Decisions

Active air defence expenditure rates can only be altered by two factors under the control of the IAMD commander: shot doctrine; and modifications to the DAL, which can imply decisions to put more previously defended assets at risk. Therefore, prior planning and agreement among Allies will need to provide the IAMD commander with sufficient guidance to make those decisions in a way that is understood and agreed by the partners.

The challenge becomes the interaction of the various offensive and defensive capabilities available to the Multi-Domain Battle Commander as they balance objectives, progress, assets and remaining inventories, and allocate the remaining assets among functional commanders conducting other missions such as SEAD, SEAT and fires supporting manoeuvre forces.¹¹² In the past, this has been the purview of Fires Coordinators and Air Warfare Coordinators, depending on service terminologies and warfare area focus, but now it will be necessary to integrate information, situational awareness and C2 across all fires assets to achieve the Commander's Intent and allow for the agile adjudication of disputes regarding resource allocation. While integrated systems can be disrupted by non-kinetic means such as electronic warfare and

window of political opportunity in which Russian gains could be stalled or rolled back closes. See Press, presentation at RUSI Space and Missile Defence Conference.

110. The risk of inventory exhaustion by large adversary arsenals has been recognised as acute. See Mark Gunzinger and Bryan Clark, 'Winning the Salvo Competition: Rebalancing America's Air and Missile Defenses', Center for Strategic and Budgetary Assessments (CSBA), 20 May 2016.

111. The subordination of missions to a central concept and overarching mission is crucial to executing operational art. See, for example, Shimon Naveh, *In Pursuit of Military Excellence: The Evolution of Operational Theory* (Abingdon: Routledge, 1997). This paper assumes the central concept in a clash involving NATO is compelling an opponent to cease operations.

112. This is based on the assumption that long-range strike assets will pose a challenge to all the assets described. See Johnson, 'Russia's Conventional Precision Strike Capabilities'.

cyber attacks, there exists a number of means to improve their resilience by multiplying communications pathways within the force – meaning that it is possible to reap the benefits of integration while mitigating costs.¹¹³ Moreover, one of Russia's primary means for disrupting C2 at theatre ranges is precision strike, which is precisely the capability such a system would be set up to negate.

Three impediments hinder this evolution. The first, explained earlier in this paper, is that nationally owned assets may not always be entirely compatible with one another. Their development at a national level may create duplications of effort,¹¹⁴ creating a federated system which potentially wastes resources on redundant efforts. For example, the existence of multiple different surface-to-air missile and lower-tier BMD systems, such as Patriot, NASAMS and Aster, creates a duplication of systems where groupings of individual countries could specialise in developing answers to niche challenges such as cruise missile defence, C-RAM or counter-UAV on a multilateral basis either informally, as with the Aster project, or through frameworks like Permanent Structured Cooperation.

Second is the lack of command integration and communications between strike and air and missile defenders that currently exists at both a national and an Alliance level. On today's battlefield, with its advanced technology, expanded data and decision-making capability and the complexity of threats and capabilities, the independent planning and execution of offence and defence will increase the chances of failure.

Third, the management of enablers is highly fragmented both at a national and an Alliance-wide level. Enablers in domains such as space, the cyber and electromagnetic spectrum and the air domain (except for the NATO-operated E3A AWACS) are held by organisations and services within countries, and their development and coordination are not managed at a NATO level.¹¹⁵ There is no NATO equivalent to the Chinese Strategic Support Force, which unifies the activities that will underpin the 'informatisation' of the PLA under a single organisation.¹¹⁶ While not suggesting an identical solution, this paper suggests that NATO will need a response to the challenge of sharing information across the Alliance to enable the unification of strike and active defence.

113. For a fuller discussion of this, see Bryan Clark, Dan Patt and Harrison Schramm, 'Mosaic Warfare: Exploiting Artificial Intelligence and Autonomous Systems to Implement Decision Centric Operations', CSBA, 2020, pp. 15–25.

114. Mahon, 'NATO's Allies Work Closely Together, But Do Their Missile Defences?'

115. Alexandra Stickings, 'Space as an Operational Domain: What Next for NATO?', *RUSI Newsbrief*, October 2020; Matthew Willis and Panagiotis Stathopoulos, 'Cyber-Electromagnetic Domain: The Necessity of Integrating the Electromagnetic Spectrum's Disciplines Under a Single Domain of Operations', Joint Air Power Competence Center, 2020.

116. On the strategic support force, see Elsa Kania and John Costello, 'Seizing the Commanding Heights: The PLA Strategic Support Force in Chinese Military Power', *Journal of Strategic Studies* (Vol. 44, No. 2, 2021), pp. 218–64; Stickings, 'Space as an Operational Domain'.

Some progress has been made in the space domain, which raises the possibility of better integrating Alliance member states' space assets and the information they provide. The declaration of space as an operational domain by NATO in December 2019 showed the recognition by the Alliance of the importance of space and raised questions as to how integration and the sharing of information would work in practice.¹¹⁷ NATO does not operate its own space assets, and among the member states there is a vast range of space capabilities, with the US still providing the majority. Decisions will need to be made about how the information from space assets is coordinated to fully enable many types of NATO operation, including air defence.

117. Stickings, 'Space as an Operational Domain'.

III. Next Steps for NATO

THREE ESSENTIAL PRECEPTS of system integration should guide NATO efforts: functional rationalisation; prioritisation; and creating efficiencies through the synergy of complementary effects.

Functional rationalisation involves the efficient allocation of resources to ends. For example, within the context of a missile defeat campaign, the optimal means of precluding certain effects may be civil resilience, while achieving others may require active defence or left of launch pre-emption. An integrated operational system would specify the roles of each system subcomponent in the service of a given campaign-level objective.

Prioritisation between the necessities and roles of the various subsystems may differ depending on the objective in question. For example, in a limited exchange short of war in the vein of Abqaiq and Khurais, civil resilience efforts and active defence might take priority, while strike assets might be postured for signalling purposes but not used. If the aim were to negate the missile threat in wartime, by contrast, a different configuration of roles would likely be optimal.

Synergies between system subcomponents can increase the efficiency of each subsystem. As mentioned earlier, the IAMD and resilience communities could more effectively allocate resources if they drew on each other's institutional knowledge. Equally, strike and IAMD systems would benefit from technical integration and institutional knowledge sharing, and alignment of their different assignments to achieve specific effects that directly support the parallel objectives of negating the air threat and achieving conflict termination on favourable terms.

An effects-based understanding of IAMD across NATO should replace current local conceptions and priorities, nesting the targets of IAMD within the desired effects, which, in turn, are situated within a campaign-level objective. For example, defence against counter-value threats should be understood less in terms of absolutely precluding threats than in terms of limiting the erosion of political will until the culmination of a campaign – meaning that a certain level of damage short of the threshold which erodes campaign-level objectives will have to be tolerated.

Similarly, force protection should emphasise protecting well-delineated points of fragility within NATO's overall military system for a specified period. Prioritising which points of failure must be defended against strike capabilities will be vital.¹¹⁸ In the absence of a drumbeat of exercises integrating offence, defence and manoeuvre, this process of prioritisation will prove difficult.

118. This is based on the assumption that interceptor arsenals will prove insufficient to protect everything and that salvos will eventually overwhelm active defences. For an in-depth discussion of this issue, see Gunzinger and Clark, 'Winning the Salvo Competition'.

Integration must provide as clear an understanding as possible of how a SEAT strike campaign can be employed to end or sufficiently attrit an air threat by, for example, attacks on launchers, and of which assets represent points of fragility in this effort, for example, tanker aircraft, missile launchers or relatively scarce fifth-generation fighters. Conversely, understanding how long offensive strike and manoeuvre elements need to negate the threat sets the requirement for air and missile defenders. The current separation of command functions between strike and missile defence leads to a degree of conceptual stovepiping, resulting in both a diffusion of efforts and seams that the enemy can exploit.

Practical Ramifications of Reframing IAMD

An integrated effects-based approach to IAMD could enable synergies at both the strategic and operational levels to facilitate a more effective system.

At the strategic level, IAMD should be more closely integrated with resilience initiatives already underway within Europe. Rather than selecting targets solely based on how significant they are, fragility and system effects should also be criteria for defence. A fragile target lacks robustness, redundancy and substitutability. A given asset, such as a power station, might be critical in terms of its importance and robustness – if repairs can be conducted rapidly. The attacks on Abqaiq and Khurais are an excellent illustration of this dynamic – the targets in question were critical but also robust, at least against a limited salvo, and were repaired within a few days.¹¹⁹ Resilience initiatives have already gone some way towards classifying targets based on their fragility and the likelihood of disruption producing network effects within European critical infrastructure systems such as power and transportation.¹²⁰ Integrating NATO's IAMD organisations with those of resilience initiatives could allow for the more efficient allocation of both active and passive defences held by civilian agencies (such as fire brigades, runway repair units), avoiding overlaps and redundancies.

At the level of integrating strike and IAMD, the replacement of the current formulation must be enabled, in which independent strike and defensive commanders exist in the air, sea and land domains, with joint strike and defensive commanders under each component command linked across domains by liaison elements. This would create the organisational basis for technical integration between IAMD and strike, which could enable a more agile and responsive approach.

Finally, in the future, information operations could be integrated into this framework. For example, the attacker's task can be complicated by targeting the assets needed to enable reconnaissance and assess battle damage. For example, Russia's constellation of Liana and Persona electronic intelligence and earth observation satellites can, in principle, be disrupted by

119. Frank Verrastro, 'Attack on Saudi Oil Infrastructure: We May Have Dodged a Bullet for Now', CSIS, 18 September 2019.

120. Batista e Silva et al. 'HARCI-EU, a Harmonized Gridded Dataset of Critical Infrastructures in Europe for Large-Scale Risk Assessments'.

a range of counter-space capabilities such as dazzling, cyber attacks and jamming.¹²¹ However, in the absence of accurate ISR and battle damage assessment, attackers can still target fixed assets but would need to do so with larger salvos, thereby obviating the operational aim of achieving outsized effects with limited strikes.

Key Lines of Effort

The end state that this paper proposes for NATO IAMD is a model that is akin to – but more expansive than – the EPAA. This would entail emulating the EPAA’s approach to emphasising the generation of capabilities in line with a well-defined Alliance sequentially increasing readiness levels, and NATO provision of key assets. In this context, individual nations contribute towards a shared budget for enablers like C2 which sustains a system permanently held under Allied HQ Ramstein, Germany and serves as a template for future efforts towards countering the air threat to NATO.¹²² Cognisant of the political hurdles between where NATO currently is and the desired end state, this paper proposes a series of intermediate steps that could serve as the springboard from which a whole of NATO air defence system could be created.

1. NATO should conduct an integrated review of its air defence capabilities. This should be performed by a team from Supreme Allied Commander Europe and Supreme Allied Commander Transformation, with the guidance of the Military Committee, to identify the gaps between existing systems and organisations and aspirations for comprehensive air defence of Europe.¹²³ The result should include details for the following requirements:
 - o The agreed specific contributions of which NATO air defence capabilities will be provided by which members.
 - o The agreed single structure of command and decision authority will prepare for, exercise and conduct NATO air defence across the spectrum of conflict, from attempted coercion to conquest.
 - o The agreed structure for coordination, preparation and execution of civilian resilience capabilities and efforts as part of the air defence of the NATO members’ homelands.

121. The vast majority of Allied counterspace capabilities belong to the US. Russia also possesses a range of these capabilities that could be targeted against Allied space-based ISR and IAMD assets. While there are broader questions as to the utility of these capabilities and the extent to which their use could lead to escalation, either in space or other domains, their existence means that the risks to space assets cannot be ignored. Decisions on the number of space assets and the extent to which they can be relied upon must be considered within broader ISR/IAMD architectures, with redundancies created in terrestrial and/or airborne platforms should access to space assets be denied.

122. Sankaran, ‘The United States’ European Phased Adaptive Approach Missile Defence System’.

123. For an example of a missile defence review, see US Department of Defense, ‘2019 US Missile Defence Review’, 2019.

2. Cultivate links between NATO's IAMD system and its resilience efforts. As a preliminary step, a representative from the Air and Missile Defence Committee might, for example, sit on NATO's Civil Emergency Planning Committee and vice versa. Additionally, individual countries might take steps to prepare civil defences for the risks of missile attack and examine areas where they can offset the limitations of active defences (training highway repair crews to rapidly repair cratered runways, for example). The Scandinavian model of Total Defence and Taiwan's overall defence concept, wherein military exercises involve the active participation of a range of private and civilian actors to test societal resilience and capacity for involvement in defensive measures, might represent a template of sorts.¹²⁴
3. Develop a structure of integrated operating plans, similar to the way that nuclear integrated plans such as the US Single Integrated Operational Plan eliminated redundancies and delineated their roles in various contingencies, in which Supreme Headquarters Allied Powers Europe would develop a conventional cross-service integrated strike and defensive architecture for an Alliance-level conventional air defence deep battle. As a starting point, one might envision integrating defensive exercises such as Formidable Shield with exercises involving strike assets.
4. In the operational domain, doctrinal and procedural concurrence and preparation for offence–defence integration will be needed to determine how air defences can best enable strike and manoeuvre. Without this integration at both the strategic and operational levels, active defences will simply be overwhelmed before meaningful effects can be delivered. In addition, adjudicating when resources can best serve the operational ends of different functional missions, such as supporting the close-in battle, SEAT and SEAD, and deducing which enablers are shared by assets meeting each challenge, will be crucial to rationalisation.
5. At an Alliance level, the scale implied by an integrated effort would likely see national contributions shift from the production of systems to system subcomponents. However, functional specialisation at a national level could allow the problem of scale to be circumvented by funnelling scarce resources at narrow subcomponents of a campaign-level challenge.

124. See, for example, Lee Hsi-min and Eric Lee, 'Taiwan's Overall Defence Concept, Explained', *The Diplomat*, 3 November 2020; J K Wither, 'Back to the Future? Nordic Total Defence Concepts', *Defence Studies* (Vol. 20, No. 1, 2020), pp. 61–81.

Conclusion

THE RETURN OF great power competition underscores that many recent NATO capabilities developed or evolved in a less threatening international environment. As the 21st century progresses, the Alliance will need to shake off the legacy of political bargains made in a safer era which separated a limited NATO BMD system from nationally-owned air defence assets.

NATO's federated air defence system, with its gaps between BMD and IAMD, is a prime example of such a legacy system. However, as the century progresses, it will become increasingly apparent that air defence is no longer a stand-alone asset capable of defeating all threats to civilian populations and fielded forces – as turn of the millennium policy envisioned.¹²⁵

Instead, NATO's defensive system must – if it is not to lose all utility – go through a process of rationalisation and integration, resulting in an Alliance-wide capability that addresses the spectrum of potential antagonists' threats.

As a first step towards this end-state, the Alliance can attempt to meet the first enabling challenge of offence–defence integration by introducing concepts such as SEAT into the Alliance's military lexicon. Conceptual shifts tend to produce organisational adaptation, which in turn sets the conditions for technical integration.¹²⁶ Any effort which focuses on technology as the first IAMD integration step risks getting this order wrong.¹²⁷

This will then facilitate the second key imperative of integrating the work of air defenders with resilience efforts and other forms of passive defence and creating benchmark reviews to guide the development of national strike and air defence capabilities. While not binding, these targets can serve to orient national investments.

In due course, this process of determining air defence needs relative to broader Alliance-level plans may habituate countries towards accepting a more integrated model and thus create the preconditions for an advanced EPAA-like model. This would enable the pooling of resources and functional specialisation that is so necessary for effective air defence.

125. For a summary of NATO policy in this era up to 2011, see NATO Public Diplomacy Division, 'Missile Defence Fact Sheet', 26 June 2011, <https://www.nato.int/nato_static_fl2014/assets/pdf/pdf_2011_06/20110608_Factsheet-Missile_Defence.pdf>, accessed 19 June 2021.

126. On the importance of conceptual foundations and critical task focus to technological adaptation, see Michael C Horowitz, *The Diffusion of Military Power: Causes and Consequences for International Politics* (Princeton, NJ: Princeton University Press, 2010), pp. 30–40.

127. *Ibid.*

The end state towards which the Alliance should build is a system that demonstrates three salient characteristics:

1. Conceptual integration of active defence with strike and passive defence.
2. Intra-Alliance integration of assets, peacetime C2 and procurement under a single NATO system.
3. Functional specialisation at the national level.

The purpose of this paper has been to lay out the strategic and operational rationale for this evolution and the steps that should be taken towards realising it.

About the Authors

Sidharth Kaushal is a Research Fellow in Sea Power and Missile Defence at RUSI.

Archer Macy is the former Director of the US Joint Integrated Air and Missile Defense Organization, served as the US representative to the NATO Air and Missile Defense Committee and was one of the co-chairs of the 2010 Ballistic Missile Defense review.

Alexandra Stickings is an Associate Fellow at RUSI. She was previously a Research Fellow for Space Policy and Security in the Military Sciences team.