SYM-AM-26-111

EXCERPT FROM THE PROCEEDINGS

OF THE

TWENTY-THIRD ANNUAL ACQUISITION RESEARCH SYMPOSIUM AND INNOVATION SUMMIT

THURSDAY, MAY 7, 2026 SESSIONS

VOLUME II

"ACCELERATING WARFIGHTING CAPABILITIES"

Quantifying Return on Investment in Digital Technical Data Packages for Combat Systems: A Life Cycle Perspective

Published: April 30, 2026

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

Quantifying Return on Investment in Digital Technical Data Packages for Combat Systems: A Life Cycle Perspective

John G. Fiore—is Vice President of Strategic Growth at Kitty Hawk Technologies in Dahlgren, VA, where he leads innovation initiatives, identifies emerging opportunities, and builds mission focused partnerships that strengthen warfighting capability and community impact. A former member of the Senior Executive Service, he previously served as the Director for Acquisition and Sustainment at the Missile Defense Agency, overseeing enterprise wide acquisition and sustainment strategies. Before joining MDA, he was the Technical Director of the Naval Surface Warfare Center, Dahlgren Division, guiding a large technical workforce in delivering advanced capabilities to the Fleet.

Glenna Miller—is a premier technical talent who brings threat engineering to life in ways vital to national security. She has more than 35 years of experience across DoD technical, engineering, program management, and leadership roles, serving within government, the military, and as a contractor supporting Navy and Air Force programs. Her background includes serving as Deputy Threat Engineering Division and Threat Engineering Branch Head at NSWC/DD and leading the PEO IWS threat working group. Her expertise spans agile and traditional systems engineering, MBSE, S&TI analysis, and Surface Navy Warfare requirements. She now serves as Kitty Hawk's technical program manager.

Chris Petrin—is a systems engineer specializing in C UAS, remote sensing, and the application of digital engineering principles to accelerate effective designs and warfighter capabilities. At Kitty Hawk Technologies, he develops physics based and statistical flight models, advanced RCS characterization methods, and modular simulation architectures. Previously, at the Oklahoma Aerospace Institute for Research and Education, he led interdisciplinary teams advancing UAS/C UAS test and evaluation for DoD and DHS sponsors. He has designed RF measurement platforms, executed major C UAS test campaigns, and synthesized intelligence to shape test vignettes. Chris holds BS and MS degrees and is completing a PhD in Mechanical and Aerospace Engineering.

Abstract

Combat system acquisition is shifting toward digital engineering (DE) as a strategic necessity for accelerating warfighting capability. Traditional Technical Data Package (TDP) development relies on physical integration and live testing, often causing delays and high costs. Digital TDPs instead enable virtual performance exploration, executable mission threads, and early configuration control. Even in sustainment phases, they support faster updates and reduce risk through model-based validation.

Using ACAT I benchmarks, this paper quantifies DE return on investment (ROI) across the life cycle. One case draws on early Model-Based Test and Evaluation (MBT&E) applied to the AEGIS Ballistic Missile Defense baseline, reducing live-fire tests by 50%, avoiding about $222 million in costs, and producing a reusable modeling and simulation environment. A second case examines the Threat Digital Twin Advanced Technology Demonstration, showing how validated digital twins improve design performance, reduce test burden, and enable reuse across variants.

Results show DE delivers ROI at any life cycle stage, with the largest gains in Operations and Support, where most life cycle costs occur. Key benefits include fewer live tests, faster integration and validation, improved configuration management, and greater stakeholder confidence. DE transforms acquisition into a faster, more resilient force multiplier.

Introduction

Major Defense Acquisition Programs (MDAPs) operate under extraordinary technical complexity, decades-long service lives, and statutory oversight that magnify the consequences of engineering decisions made early in the life cycle. Yet despite more than a decade of policy emphasis on digital engineering (DE), the Department still lacks a quantitative, evidence-based understanding of where, when, and how DE produces measurable return on investment. This absence of analytic grounding is especially problematic for program managers, who are rarely provided discretionary resources to adopt new engineering practices or toolchains, particularly in legacy programs where technical debt is highest, documentation is fragmented, and budgets are most constrained. In this environment, DE is often mandated as a compliance requirement rather than justified as a beneficial life cycle investment, leaving programs without the evidence needed to defend or prioritize adoption. Policy has outpaced practice, creating a disconnect in which programs are expected to implement DE without the analytic basis needed to justify investment or the institutional support required to succeed.

The prevailing narrative in both policy and industry suggests that DE only produces meaningful return on investment (ROI) when adopted early, during Materiel Solution Analysis (MSA) or Technology Maturation and Risk Reduction (TMRR). While early adoption unquestionably provides the cleanest opportunity to establish a seamless digital thread, this assumption is rarely tested and often untrue. Many ACAT I programs begin DE adoption mid-stream or even during production and sustainment, yet still achieve measurable cost avoidance, risk reduction, and modernization agility.

This paper addresses that gap by examining ROI across two complementary scenarios. The first, and primary empirical case, draws directly on the author's experience as the former Deputy Technical Director and Chief Engineer for AEGIS BMD during the baseline development period (circa 2007–2014). It demonstrates how model-based test and evaluation principles reduced physical test volume, cost, and operational risk by more than $222 million, before such approaches were formalized in Department of Defense (DoD) policy. The second scenario examines the Threat Digital Twin ATD, a current program supported by Kitty Hawk Technologies, as a forward application of those proven principles, showing how the structural conditions for measurable life cycle ROI have been deliberately established in a mission area where no adequate physical test or training asset previously existed. Together, these scenarios provide both retrospective evidence and prospective framework demonstrating that DE generates significant ROI even when adopted mid-stream or late, and that the largest returns often emerge in O&S, where 60% to 70% of life cycle cost resides (GAO, 2024a).

Beyond illustrating ROI, this paper offers a practical framework for how programs can collect, curate, and structure the evidence needed to justify DE investment. By grounding the analysis in real program experience and aligning it with current DoD policy, the paper provides a defensible, actionable foundation for program managers seeking to adopt DE in environments where resources are limited, legacy constraints are significant, and the need for modernization is urgent.

Background

Defense Acquisition Context for ACAT I Programs

MDAPs, particularly those designated as ACAT I, operate under a uniquely demanding combination of technical complexity, statutory oversight, and decades-long service life. These programs progress through the DoD acquisition framework—MSA, TMRR, Engineering and Manufacturing Development (EMD), Production and Deployment (P&D), and Operations and Support (O&S)—yet the majority of life cycle cost and risk emerges only after Milestone C (entrance to the P&D phase). Across MDAPs, O&S consistently accounts for up to 70% of total life cycle cost, as documented in Selected Acquisition Reports (DoD, 2025). This dominance is driven by recurring maintenance and depot activity, continuous software sustainment and cybersecurity updates, parts obsolescence and diminishing manufacturing sources and material shortages (DMSMS) pressures, block upgrades required to pace evolving threats, and configuration drift across fielded units (DoD, 2025).

Traditional document-centric Technical Data Packages (TDPs) struggle to support these long-term demands; fragmented interface specifications, inconsistent configuration baselines, and limited traceability create brittleness that compounds over time. As a result, ACAT I programs face a structural challenge: the phases where most cost and risk accumulate are precisely the phases where traditional engineering artifacts are least effective. DE and model-based systems engineering (MBSE) directly target this gap by creating authoritative, persistent, and executable representations of the system—assets that remain useful throughout the life cycle rather than only during development.

Current DoD DE Strategy and Policy Environment

The DoD Digital Engineering Strategy (2018) remains the central policy anchor for DE adoption, articulating five goals: formalizing the development and use of models, establishing authoritative sources of truth, incorporating technological innovation, building supporting infrastructure, and transforming the culture and workforce (DoD, 2018). Importantly, these goals extend beyond early design and explicitly include sustainment, modernization, and test. Complementary policy reinforces this life cycle orientation, including DoDI 5000.97 on Digital Engineering for Defense Acquisition (DoD, 2023), DoDI 5000.61 on M&S Verification, Validation, and Accreditation (DoD, 2022), and 5000-series T&E policy permitting accredited M&S to replace or supplement physical testing like DoDI 5000.89 (DoD, 2020). A recent report from RAND further strengthens this policy support by noting that DE cost-benefit assessment is feasible at any life cycle stage and that systems-thinking and goal-based systems engineering are essential for realizing measurable returns (Whitehead et al., 2024).

These policies are aligned with larger trends across industries. The U.S. Government Accountability Office (GAO) surveyed 14 leading companies in product development and found extensive use of iterative design using modeling and simulation to create and maintain a digital thread along with validation using combinations of physical and digital prototypes (GAO, 2023). Digital twins, as an example of a key component of DE, are finding applications in manufacturing, aerospace, construction, agriculture, healthcare, retail, and mining (Attaran & Celik, 2023).

Together, these policy and analytic foundations support the central thesis of this paper: DE is a life cycle investment with quantifiable returns in sustainment, modernization, and test, not merely a front-end documentation replacement.

The Economics of DE Adoption Timing

A key finding from the RAND 2024 framework is that DE benefits are not limited to early adopters (Whitehead et al., 2024). While programs that begin DE during MSA or TMRR achieve the most seamless digital thread, meaningful ROI remains available at SRR, PDR, CDR, or even during production and O&S. Early adoption offers the highest theoretical return by enabling a full digital thread from concept and reducing integration risk entering EMD. Mid-stream adoption is the most common in real programs (GAO, 2025), and typically involves a digital retrofit of existing artifacts and yields improvements in integration, verification, and configuration control. Even late adoption during P&D or O&S may remain valuable when major block upgrades are planned, decades of service life remain, or obsolescence and software modernization pressures are high.

The economic logic is straightforward: when O&S represents 60% to 70% of life cycle cost, even modest efficiency gains can dwarf the cost of DE adoption. To illustrate this effect, Figure 1 shows the cumulative cost for the Illustrative System Life Cycle from the DoD O&S Cost-Estimating Guide (DoD, 2025). Two implementation cases are shown as well. In Case 1, an additional hypothetical 20% cost to implement DE is incurred during RDT&E, with variable percent cost savings in O&S. In Case 2, the hypothetical cost is incurred later to represent implementation during O&S, with benefits realized at the completion of the implementation. In Case 1 using the Illustrative System Life Cycle, only a 3.3% reduction in annual O&S costs is necessary to break even on the DE implementation investment. For Case 2, the break-even point becomes a 5.5% reduction in O&S costs. Both cases assume a worst case, that DE implementation yields no benefits until and only in O&S. Yet in both cases, modest O&S cost reductions can justify the implementation.

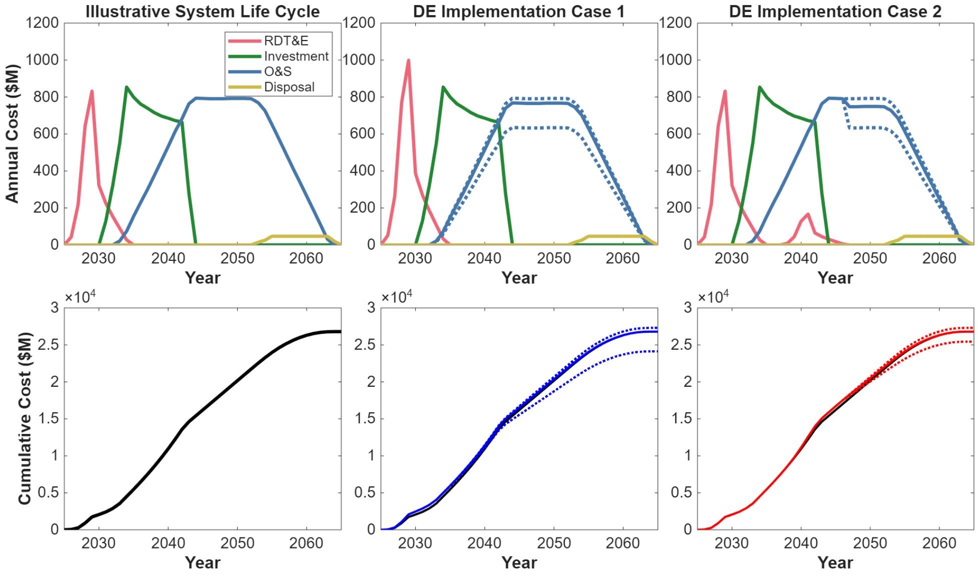
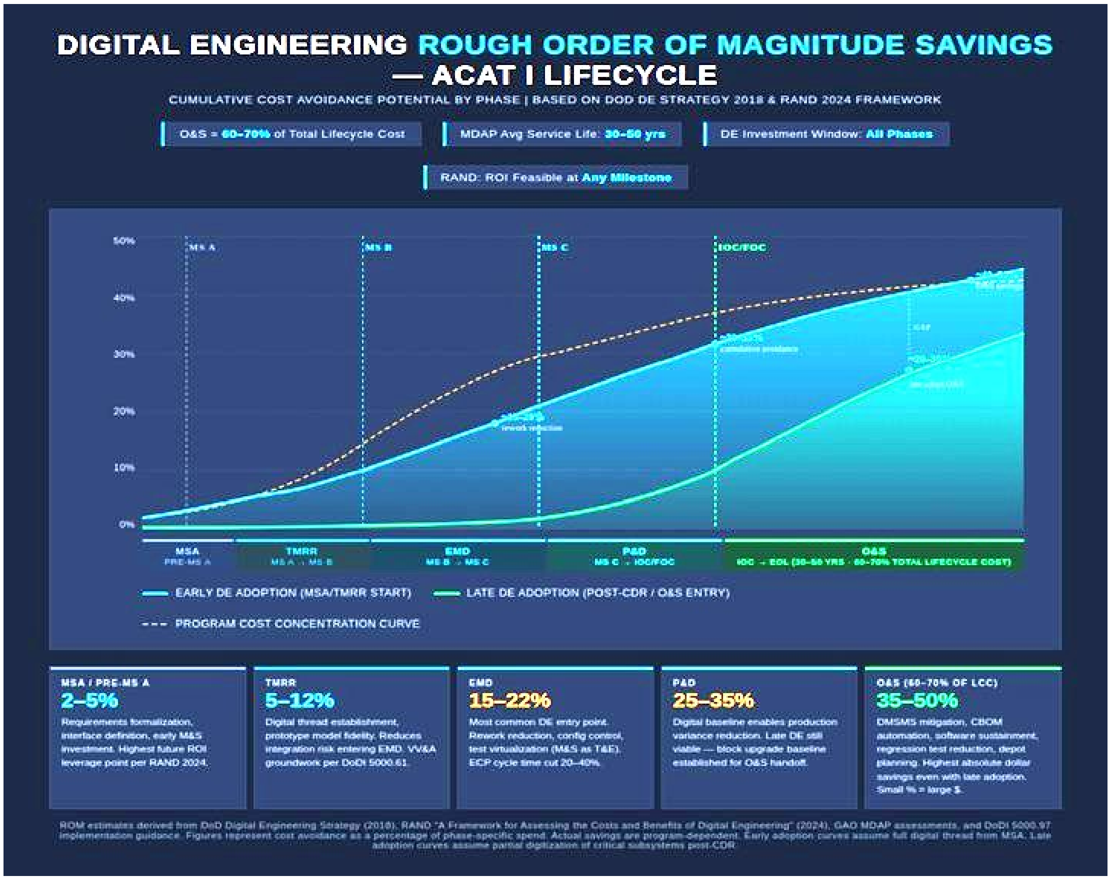


Figure 1. (Top) Annual Costs for the Illustrative System Life Cycle from DoD O&S Cost-Estimating Guide (DoD, 2025), and Two DE Implementation Cases. (Bottom) The Cumulative Cost Incurred for the Annual Costs Shown Directly Above. The dashed bands indicate 0% O&S savings (top dashed curve) and up to 20% O&S cost reduction (bottom dashed curve).

Figure 1 establishes the break-even threshold using DoD's own published life cycle cost curve—a conservative, single-source test of the economic logic. Figure 2 broadens that analysis into a phase-by-phase ROM estimate synthesized across four policy and analytic sources, showing where the acquisition life cycle DE value accumulates and at what order of magnitude.



DIGITAL ENGINEERING ROUGH ORDER OF MAGNITUDE SAVINGS — ACAT I LIFECYCLE

CUMULATIVE COST AVOIDANCE POTENTIAL BY PHASE | BASED ON DOD DE STRATEGY 2018 & RAND 2024 FRAMEWORK

O&S = **60–70%** of Total Lifecycle Cost

MDAP Avg Service Life: **30–50 yrs**

DE Investment Window: **All Phases**

RAND: ROI Feasible at **Any Milestone**

MSA / PRE-MS A
2–5%
Requirements formalization, interface definition, early M&S investment. Highest future ROI leverage point per RAND 2024.

TMRR
5–12%
Digital thread establishment, prototype model fidelity. Reduces integration risk entering EMD. VV&A groundwork per DoDI 5000.61.

EMD
15–22%
Most common DE entry point. Rework reduction, config control, test virtualization (M&S as T&E). ECP cycle time cut 20–40%.

P&D
25–35%
Digital baseline enables production variance reduction. Late DE still viable — block upgrade baseline established for O&S handoff.

O&S (60–70% OF LCC)
35–50%
DMSMS mitigation, CBOM automation, software sustainment, regression test reduction, depot planning. Highest absolute dollar savings even with late adoption. Small % = large $.

ROM estimates derived from DoD Digital Engineering Strategy (2018), RAND "A Framework for Assessing the Costs and Benefits of Digital Engineering" (2024), GAO MDAP assessments, and DoDI 5000.97 implementation guidance. Figures represent cost avoidance as a percentage of phase-specific spend. Actual savings are program-dependent. Early adoption curves assume full digital thread from MSA. Late adoption curves assume partial digitization of critical subsystems post-CDR.

Graphic developed with AI assistance (Claude, Anthropic, 2025) based on the DoD Digital Engineering Strategy (2018); RAND Corporation, "A Framework for Assessing the Costs and Benefits of Digital Engineering" (2024); GAO MDAP program assessments; and DoDI 5000.97. Cost-avoidance values are rough-order-of-magnitude estimates and are program-dependent.

Figure 2. DE Rough Order of Magnitude Savings—ACAT I Life Cycle

Figure 2 illustrates cumulative cost-avoidance potential across the ACAT I life cycle. Early DE adoption beginning at MSA can yield 40%–45% cumulative avoidance by full operational capability (FOC)—roughly 15 percentage points greater than late adoption at CDR or O&S—but meaningful returns are available at every phase. Even late adoption produces 28%–35% cumulative avoidance, a compelling return given the absolute dollar scale of O&S. These ranges are rough-order-of-magnitude estimates derived from the DoD DE Strategy (DoD, 2018), RAND's 2024 framework (Whitehead et al., 2024), GAO MDAP assessments (GAO, 2025), and DoDI 5000.97 (DoD, 2023), and represent potential for cumulative, not annual, cost avoidance.

Table 1. DE ROM Cost Avoidance by Acquisition Phase (ACAT I Programs)

Acquisition Phase	ROM Cost Avoidance	Primary DE/MBSE Value Drivers	Why Savings Are Anticipated
MSA / Pre-MS A	2%–5%	Requirements formalization; interface definition; early M&S investment	Early clarity reduces downstream rework; highest leverage point per RAND 2024
Technology Maturation and Risk Reduction (TMRR)	5%–12%	Digital thread establishment; prototype model fidelity; VV&A groundwork	Reduces integration risk entering EMD; anchors early model credibility
Engineering and Manufacturing Development (EMD / MS B)	15%–22%	Rework reduction; configuration control; test virtualization; M&S as T&E	Cuts ECP cycle time by 20%–40%; reduces physical test load
Production & Deployment (MS C)	25%–35%	Digital baseline for production; variance reduction; upgrade pathway establishment	Late DE still viable; stabilizes configuration for O&S handoff
Operations & Support (IOC/FOC)	35%–50%	DMSMS mitigation; configuration-controlled parts automation; software sustainment; regression test reduction; depot planning	Highest absolute dollar savings because O&S is 60–70% of life cycle cost

Source: Fiore, J. G., "Digital Engineering ROM Cost Avoidance Methodology," internal working paper, 2025, derived from DoD DE Strategy (2018), RAND (2024), GAO MDAP assessments, and DoDI 5000.97.

Taken together, the acquisition realities, policy foundations, and economic analysis presented in this section establish a compelling case for DE as a life cycle imperative. ACAT I programs face an environment where 60% to 70% of total cost accumulates in O&S, where configuration drift and obsolescence pressures intensify over decades, and where traditional document-centric artifacts cannot keep pace with continuous software sustainment, cybersecurity, and block upgrades. The scenarios that follow demonstrate through concrete evidence how these principles translate into real reductions in test cost, operational risk, and life cycle burden.

Scenario 1—MBT&E: AEGIS BMD Baseline

(Personal Technical and Programmatic Reflection, John G. Fiore)

As a senior technical leader during the AEGIS BMD baseline development period (circa 2007–2014), MBT&E principles offered a powerful means of reducing cost, schedule pressure, and operational risk in a highly complex ACAT I environment. Although the formal digital-engineering constructs we reference today, such as MBTEMP, iTEMS, authoritative digital baselines, and enterprise digital threads did not exist circa 2008, the engineering team applied many of the underlying principles well before they were formally named or institutionalized.

At that time, Developmental Test, Operational Test, and Live Fire Test and Evaluation imposed significant burdens on the program, particularly because the AEGIS BMD kill chain was so tightly coupled that a failure in any segment invalidated the entire event. The Navy's Operational Test Agency (OPTEVFOR) initially requested 12 live-fire flight tests to demonstrate baseline performance, driven by the variability of threat environments, radar sensitivity to clutter and geometry, compressed engagement timelines, and the probabilistic nature of endgame lethality. Each event carried substantial cost—roughly $20 million per missile, $6.5 million in range operations, $1 million in instrumentation and telemetry, and $6–$11 million in ship operations, while also exposing ships, sailors, civilians, and collection assets to operational risk. Despite this investment, physical testing inherently under-sampled the performance space, leaving many edge cases and emergent behaviors unexplored.

Based on the technical realities we faced, the team adopted a structured alternative built around a reduced set of anchor tests supported by a high-fidelity modeling and simulation (M&S) environment. Although terms like "anchor tests" and "runs for the record" would later be formalized in DoD policy, the logic was already clear: use a small number of well-chosen physical events to calibrate and validate the digital environment, then rely on that environment to explore the broader performance space. We selected six anchor tests, replacing the original 12, to span radar performance boundaries, threat kinematic extremes, environmental stressors, and weapon-system timing constraints.

Once the models were verified, validated, and accepted by OPTEVFOR and Director, Operational Test and Evaluation (DOT&E), they were authorized for formal evidence in the Integrated Master Test Plan (IMTP) for Monte Carlo exploration of thousands of engagement variations, for excursions inside and outside the contracted envelope, and for regression testing supporting software builds and block upgrades. Although DoDI 5000.61 and the 5000-series T&E policies that explicitly permit M&S substitution would come later, the AEGIS BMD team was already demonstrating the feasibility and rigor of this approach.

The planning methodology we used closely resembled what would eventually become the Model-Based Test and Evaluation Master Plan (MBTEMP). Mission engineering defined operational threads; system models mapped those threads to system behaviors; test objectives were traced to model elements; anchor tests calibrated the model; and the validated environment enabled expanded exploration. Similarly, although the Integrated Test and Evaluation Management System (iTEMS) had not yet been conceived, our disciplined management of test artifacts, data pedigree, schedule dependencies, and configuration control mirrored the capabilities iTEMS now provides. In retrospect, the AEGIS BMD baseline effort stands as an early pathfinder that demonstrated the operational value of digital-engineering principles years before the DoD formally codified them.

As shown in Table 2, this approach produced measurable reductions in actual test costs. The accredited M&S environment enabled test-point reduction by identifying redundant conditions, applying design-of-experiments techniques to maximize coverage, and focusing physical tests on high-value regions, reducing the requirement from twelve planned tests to six anchor tests, a 50% reduction. Once validated, the models became reusable assets for software regression, block upgrades, threat updates, and mission-level analysis, extending their value far beyond the baseline delivery.

Table 2. AEGIS BMD Flight Test Cost Comparison: Original 12-Test Plan vs. Revised 6-Anchor-Test Plan

Cost Element	Per-Event Cost	Original Plan (12 Tests)	Revised Plan (6 Tests)	Savings
Missile / Test Article	~$20M	~$240M	~$120M	$120M
Range Operations	~$6.5M	~$78M	~$39M	$39M
Instrumentation & Telemetry	~$1M	~$12M	~$6M	$6M
Ship Operations & Crew	~$8 -11M (avg. $9.5M)	~$114M	~$57M	$57M
Safety / Risk Exposure	Non-monetary but significant	12 high-risk events	6 high-risk events	50% reduction in operational risk
Scenario Coverage	Limited	12 physical scenarios	Thousands via M&S	Massive increase
Total Quantifiable Cost	—	**~$444M**	**~$222M**	**~$222M saved (≈50%)**

Summary of ROI

- **Direct cost avoidance: ~$222 million**
- Operational risk reduction: 50% fewer high-risk live-fire events
- Scenario coverage: Orders of magnitude increase through Monte Carlo simulation
- Long-term asset creation: Accredited M&S environment reused for upgrades, regression, and threat updates
- Life cycle ROI: Compounding savings across future block upgrades

The AEGIS BMD baseline effort demonstrates that model-based T&E is not merely a cost-avoidance mechanism, it is a force multiplier for safety, operational confidence, and long-term modernization agility. Reducing twelve planned flight tests to six anchor events produced more than $222 million in direct savings, but the deeper return came from reducing high-risk live-fire evolutions by half, expanding scenario coverage from a dozen physical shots to thousands of analytically defensible engagement variations, and generating a validated M&S environment that continues to pay dividends across software builds, block upgrades, and threat-model updates.

Most importantly, this work, executed years before the DoD formalized MBTEMP, iTEMS, or digital-engineering policy, showed that the principles of DE deliver measurable, repeatable ROI even in the absence of a fully mature enterprise environment. When authoritative models, telemetry-anchored validation, and mission-thread traceability are combined, programs achieve higher confidence at lower cost, with greater safety and dramatically improved test sufficiency. This is the essence of digital-engineering value, and it remains as relevant today as it was when this pathfinding effort first proved what was possible.

Scenario 2—Applying DE Principles Forward: The Threat Digital Twin ATD

Life Cycle Cost Reduction Through DE and Threat Digital Twins

Scenario 2 presents the Threat Digital Twin ATD—a current program supported by Kitty Hawk Technologies—as a forward application of the principles demonstrated in Scenario 1. The AEGIS BMD baseline establishes the empirical proof of concept; the Digital Twin ATD illustrates how those same principles are being embedded in a current program from inception, and outlines the evidence framework that will quantify returns as the program matures. Program-sensitive cost data appropriately remains within the program office and is not presented here.

The AEGIS BMD baseline demonstrated what disciplined application of model-based principles can achieve in a mature ACAT I environment. The natural question that follows is whether those same principles translate forward into current programs, emerging threat domains, and mission areas where no adequate physical test or training asset previously existed. The Threat Digital Twin Advanced Technology Demonstration (ATD), a current effort supported by Kitty Hawk Technologies, provides a compelling answer.

Unlike AEGIS BMD, the Threat Digital Twin ATD is not yet at a stage where life cycle cost savings can be fully quantified. What can be described, and what makes this case analytically valuable, is the degree to which the program has deliberately embedded the foundational DE principles that produced measurable ROI in the AEGIS BMD case: authoritative digital baselines, physics-informed modeling, configuration-controlled artifacts, and a digital thread linking design, test, and validation. The argument is not that savings have been measured yet. The argument is that the structural conditions for measurable ROI have been deliberately established, and that the evidence framework for capturing those returns is already in place.

The Threat Digital Twin is not a generic visualization tool. It is a validated, physics-based, time-space correlated representation of a complex separating aerial vehicle, incorporating AI/ML-enhanced physics-informed neural network (PINN)-based threat motion modeling, six-degrees-of-freedom (6DoF) accounting for steady and unsteady aerodynamics, dynamic time and space-correlated RF signatures, and digital SE processes ensuring threat validity. These features make it a high-fidelity, configuration-controlled asset reusable across design, manufacturing, test, and sustainment, precisely the type of digital artifact the AEGIS BMD experience showed produces compounding life cycle returns.

Where the ROI Potential Lies

The AEGIS BMD case showed that the largest returns emerged from three mechanisms: reduction in physical test burden, creation of a reusable accredited M&S environment, and elimination of repeated re-characterization costs across software builds and block upgrades. The Threat Digital Twin ATD is structurally positioned to realize all three.

First, in a mission area where no adequate physical test asset previously existed, the threat digital twin directly substitutes for physical test articles across a range of design, analysis, and training scenarios. The cost avoidance associated with not fabricating multiple physical threat surrogates, each carrying substantial manufacturing, instrumentation, and range costs, is real, even if not yet fully tallied.

Second, the PINN-based dynamics model, once validated against flight-test data, becomes a reusable asset. New threat variants, geometry changes, and separation timing modifications can be assessed within the digital environment without new physical fabrication and extensive testing. This mirrors exactly the reuse value the AEGIS BMD accredited M&S environment delivered across subsequent block upgrades and regression test cycles.

Third, the authoritative digital baseline prevents the configuration drift that drives reverse engineering and re-characterization costs in legacy programs (DoD, 2025). Every manufactured unit, every design update, and every test configuration references the same validated source of truth, eliminating the ambiguity that compounds into life cycle technical debt. Once validated, the threat digital twin becomes a reusable asset that can be updated, extended, or adapted to represent new threat variants without requiring new physical test articles.

Evidence Framework for Quantifying ROI

A credible ROI case for the Threat Digital Twin ATD will be built from the following evidence streams, which the program is positioned to collect as it matures:

- Fabrication and test cost comparisons between physical threat surrogates and digital twin-enabled alternatives
- Reduction in re-characterization events as the validated model replaces repeated physical testing
- Configuration variance tracking across manufactured units against the digital baseline
- Upgrade cycle timelines for new threat variants developed within the digital environment versus those requiring new physical articles

These are not speculative metrics. They are the same categories of evidence that, in retrospect, would have quantified the AEGIS BMD ROI had the program been tracking them systematically from the outset. The AEGIS BMD scenario shows that that programs should instrument their DE adoption from the beginning, not reconstruct the evidence afterward.

Positioning in the Broader Argument

The Threat Digital Twin ATD case reinforces the central thesis of this paper in a specific and important way. The AEGIS BMD baseline proved that DE principles deliver measurable ROI even when adopted without a fully mature Digital Engineering Environment (DEE). The Threat Digital Twin ATD demonstrates that current programs, informed by that experience and supported by more sophisticated toolchains, are applying those principles earlier, more deliberately, and with a clearer eye toward life cycle returns.

The trajectory from AEGIS BMD to the Threat Digital Twin ATD is not merely historical. It is the institutionalization of hard-won lessons into current practice, which is precisely what the DoD’s Digital Engineering Strategy calls for, and precisely what this paper argues the Department should systematically enable and resource. Importantly, none of these practices require full DEE maturity. The Threat Digital Twin ATD shows that even partial adoption, authoritative models, and a digital-threaded V&V workflow, establishes the structural conditions for measurable, repeatable ROI across the life cycle.

Cross-Cutting Analysis

A central insight emerging from both scenarios is that DE delivers measurable returns on investment across the entire acquisition life cycle, but the mechanisms, magnitude, and timing of those returns differ depending on when adoption occurs. Early adoption provides the greatest architectural freedom and the cleanest opportunity to establish a seamless digital thread from concept through sustainment. Yet the case examples presented here, particularly the AEGIS BMD baseline, demonstrate that mid-stream and late adoption often generate the most tangible, defensible, and operationally meaningful returns. This is because the costliest phases of the life cycle occur after Milestone C, when sustainment, obsolescence, regression testing, and configuration drift dominate program expenditures (DoD, 2025).

"It Is Never Too Late" to Adopt DE

One of the most persistent misconceptions in the DE community is the belief that DE must begin early, ideally at MSA or TMRR, to be effective. The empirical evidence from Scenario 1 directly contradicts this. The AEGIS BMD baseline adopted model-based principles at the end of EMD, at a point when the system was already transitioning into legacy sustainment. Despite this late entry, the program achieved a 50% reduction in physical flight tests, avoided hundreds of millions of dollars in cost, and created a reusable, accredited digital environment that supported future upgrades and regression testing. Perhaps most importantly, the digital environment expanded the quality and breadth of evidence available to decision-makers, far beyond what live testing alone could provide, by enabling thousands of analytically defensible engagement variations through Monte Carlo simulation.

This is not merely a case study; it is a proof point. Programs facing legacy technical debt, configuration drift, obsolescence pressures, and high sustainment costs stand to benefit the most from DE precisely because these are the challenges where traditional document-centric approaches fail. Scenario 2 reinforces this by demonstrating that current programs applying DE principles from inception are establishing the structural conditions for returns that will compound over the system's service life.

Governance and Metrics That Matter

Across both scenarios, the common denominator of successful DE adoption is not tool sophistication but governance discipline. DE succeeds when programs treat models and digital artifacts with the same rigor historically reserved for software—configuration control, formal verification and validation pipelines, and cross-functional governance structures that ensure digital artifacts remain authoritative and credible across life cycle phases.

Effective governance includes:

- Model configuration control with the same discipline applied to software baselines
- A formal VV&A pipeline for any model used to support decisions, test reduction, or certification
- A DE Steering Group integrating engineering, T&E, logistics, cybersecurity, and PMO functions
- Life cycle–aligned metrics, including:
    - Reduction in engineering change proposals (ECPs)
    - Reduction in physical test points
    - Decrease in sustainment downtime
    - Faster upgrade cycles
    - Avoided reverse engineering and re-baselining

Without governance, DE devolves into a collection of disconnected models. With governance, it becomes a strategic asset that compounds in value over time.

Risk Considerations

DE adoption is not without risk. Programs may over-promise benefits, underestimate the VV&A burden, or become dependent on proprietary toolchains that limit interoperability. Data pedigree issues can undermine model credibility, and workforce readiness remains a persistent challenge across the DoD. However, the AEGIS BMD example demonstrates that these risks are manageable even in the absence of modern DE infrastructure.

The most effective mitigation strategies include:

- **Incremental adoption**—focusing first on high-value models and authoritative baselines
- **Transparent assumptions**—rigorous documentation of model limitations
- **Telemetry-anchored calibration**—which builds credibility and reduces VV&A friction
- **Strict configuration control**—ensuring digital artifacts do not drift from physical reality
- **Cross-functional governance**—preventing DE from becoming siloed within engineering

Synthesis: What These Scenarios Tell Us About DE’s True Value

The cross-cutting insight from both scenarios is that DE is not a monolithic capability but a portfolio of practices that deliver value in different ways at different times. Early adoption maximizes architectural flexibility and reduces integration risk. Mid-stream adoption improves verification, configuration control, and test sufficiency. Late adoption reduces sustainment burden, obsolescence risk, and regression test cost. In every case, the value proposition is tied to life cycle economics: when 60%–70% of total cost resides in O&S, even modest improvements in sustainment efficiency dwarf the cost of DE adoption.

The scenarios also demonstrate that DE’s most powerful contributions extend beyond cost avoidance to improvements in safety, evidence quality, and operational confidence. Reducing live-fire events reduces risk to ships and sailors. Expanding scenario coverage improves decision quality. Creating authoritative digital baselines reduces rework and accelerates modernization. These benefits compound over time, creating a virtuous cycle in which each block upgrade becomes faster, cheaper, and more predictable than the last.

Conclusions and Recommendations

The analysis presented in this paper demonstrates that DE delivers measurable return on investment across both test and evaluation and life cycle management, and that these benefits do not depend on early adoption. Scenario 1 draws on the firsthand experience of John G. Fiore as Deputy Technical Director and Chief Engineer for AEGIS BMD, providing empirical evidence that model-anchored T&E reduced physical flight tests by 50% and generated cost avoidance of approximately $222 million, despite being implemented at the end of EMD and without a fully mature DE Environment. Scenario 2 draws on the technical work of Glenna Miller at Kitty Hawk Technologies, illustrating how the Threat Digital Twin ATD is applying those same principles from inception and establishing the structural conditions for measurable life cycle ROI through authoritative digital baselines, physics-informed modeling, and configuration-controlled digital threads. Together, these scenarios show that DE is not merely a design-phase enhancement but a life cycle capability that improves affordability, test sufficiency, and modernization agility regardless of when adoption begins.

The broader implication is difficult to ignore. Over the past decade, DoD policy has aggressively mandated DE adoption without first establishing a quantitative foundation for ROI or providing the resources required to implement it. Policy has outpaced evidence, and the burden of reconciling this gap has fallen disproportionately on programs least able to absorb it—particularly legacy programs where technical debt is highest and discretionary budgets are most constrained.

Correcting this imbalance requires a shift in how the Department conceptualizes and funds DE:

- **Treat DE as a program-level investment, not an unfunded mandate.** Resources must be aligned to the phases where the greatest returns are achievable.
- **Establish standardized ROI measurement frameworks** across both life cycle and T&E domains, enabling programs to quantify benefits in a defensible and repeatable manner.
- **Develop authoritative exemplars**—such as the AEGIS BMD baseline—that demonstrate how disciplined application of DE principles produces evidence-based value even without full DEE maturity.
- **Prioritize model credibility, configuration control, and data pedigree** as foundational elements of any DE effort, ensuring digital artifacts remain authoritative and reusable across decades of service life.
- **Support incremental adoption pathways** that allow programs to begin with high-value, low-risk digital practices rather than attempting enterprise-level maturity in a single leap.

Ultimately, DE is not a slogan, a compliance exercise, or a tool acquisition strategy. It is a disciplined engineering approach that reduces risk, accelerates modernization, and lowers life cycle cost when applied with rigor and supported with appropriate governance. The scenarios in this paper show what is possible when DE is implemented thoughtfully and pragmatically. The next step is institutionalizing these lessons so that every program, regardless of age, phase, or legacy burden, can realize the full life cycle value that DE promises.

Disclaimer: *The views expressed in Scenario 1 are solely those of John G. Fiore, based on his direct experience as the former Deputy Technical Director and Chief Engineer for AEGIS Ballistic Missile Defense. The technical content and analysis presented in Scenario 2 reflect the work and professional judgment of Glenna Miller in her capacity at Kitty Hawk Technologies. The views expressed in this paper do not represent the official position of the Department of Defense, the U.S. Navy, the Missile Defense Agency, or any other government agency or contractor.*

References

Attaran, M., & Celik, B. G. (2023). Digital twin: Benefits, use cases, challenges, and opportunities. *Decision Analytics Journal, 6*, 100165. doi:10.1016/j.dajour.2023.100165

DoD. (2018). *Department of Defense digital engineering strategy.* Office of the Deputy Assistant Secretary of Defense for Systems Engineering.

DoD. (2020). *Test and evaluation* (DoD Instruction 5000.89). Office of the Under Secretary of Defense for Research and Engineering.

DoD. (2022). *DoD modeling and simulation verification* (DoD Instruction 5000.61). Office of the Under Secretary of Defense for Research and Engineering.

DoD. (2023). *Digital engineering* (DoD Instruction 5000.97). Office of the Under Secretary of Defense for Research and Engineering.

DoD. (2025). *Operating and support cost-estimating guide.* Office of the Secretary of Defense.

GAO. (2023). *Leading practices: Iterative cycles enable rapid delivery of complex innovative products* (GAO-23-106222).

GAO. (2024a). *Weapon system sustainment: DOD identified operating and support cost growth but needs to improve the consistency and completeness of information to Congress* (GAO-24-107378).

GAO. (2025). *Weapon systems annual assessment: DOD leaders should ensure that newer programs are structured for speed and innovation* (GAO-25-107569).

GAO. (2025). *High-risk series: Critical actions needed to urgently address IT acquisition and management challenges (GAO-25-107852)*.

Whitehead, N. P., Light, T., Luna, A., & Mignano, J. (2024). *A framework for assessing the costs and benefits of digital engineering.* RAND.

Acquisition Research Program
Department of Acquisition, Finance, and Manpower
Naval Postgraduate School
555 Dyer Road, Ingersoll Hall
Monterey, CA 93943

www.acquisitionresearch.net