

Systems of Systems Test and Evaluation Challenges

Dr. Judith Dahmann
MITRE
McLean, Virginia USA
jdahmann at mitre.org

George Rebovich
MITRE
McLean, Virginia USA
grebovic at mitre.org

Dr. Jo Ann Lane
University of Southern California
Los Angeles, California USA
jolane at usc.edu

Ralph Lowry
Modern Technology Solutions, Inc.
Alexandria, VA USA
ralph.lowry at mtsi-va.com

Abstract - *A growing number of military capabilities are achieved through a system of system approach and this trend is likely to continue in the foreseeable future. Systems of systems differ from traditional systems in ways that require tailoring of systems engineering processes to successfully deliver their capabilities. This paper describes the distinct characteristics of systems of systems that impact their test and evaluation, discusses their unique challenges, and suggests strategies for managing them. The recommendations are drawn from the experiences of active system of system engineering practitioners.*

Keywords: System of systems engineering, test and evaluation, test techniques.

1 Background and Introduction

The United States (US) Department of Defense (DoD) recognizes the importance of systems of systems (SoS) in meeting user capability needs. The *DoD Guide for Systems Engineering of Systems of Systems* [1] (SoS SEG) defines SoS as a "collection of systems, each capable of independent operation, that interoperate together to achieve additional desired capabilities". SoSs differ from traditional systems in ways that require tailoring of systems engineering (SE) processes. The distinctive characteristics of SoS have implications for the application of test and evaluation (T&E) to SoS.

T&E for SoS has traditionally been addressed in the US DoD from the perspective of testing individual systems in an operationally realistic environment as well as certifying system inoperability. [2] Here the focus is on T&E of the individual system in the larger context with attention on the issues of how to cost-effectively create the test environment.

More recently the US DoD has begun to shift from a system perspective to a capability perspective where the value to users is the collective effect of a set of systems rather than any one system. This leads to questions about how to define capabilities and how to integrate and test at the SoS level [1, 3, 4, 5].

This paper looks at SoS and T&E from the perspective of systems engineering and addresses the questions: What are the critical characteristics of SoS that affect T&E? What are the T&E implications for SoS? The answers to these questions draw on the experiences of SE practitioners currently working in SoS, including those used as the basis for the SoS SEG [1]. This paper reviews the characteristics of SoS as they impact T&E, and how aspects of T&E are addressed by the practice of SoS SE. Finally it discusses the implications for T&E of SoS, including specific challenges and the strategies currently employed to address them.

The focus of this paper is on ‘*acknowledged* SoS’. *Acknowledged* SoS have recognized objectives, a designated manager, and resources for the SoS; however, the constituent systems retain their independent ownership, objectives, funding, development and sustainment approaches. Changes to the constituent systems are agreed collaboratively between the SoS and the systems. Many SoS in the DoD today exhibit the characteristics of *acknowledged* SoS since the DoD has adopted a de facto strategy to maintain and leverage currently fielded systems to meet new and emerging needs wherever possible. For budgetary reasons, this is likely to continue well into the future.

This paper also applies to “mission-level” SoS where multiple platforms, information technology (IT) and other systems are brought together to meet larger mission level capability objectives. Other SoS, such as platform-level SoS (integration of separately developed systems on a submarine for instance) and IT-based SoS (where information across an SoS is managed as an enterprise asset), share some of the issues addressed here, but they have their own specific considerations, as well.

2 SoS Characteristics that Impact T&E

SoS present unique development challenges. These result from several factors, including broad based mission level SoS capability objectives, lack of control by the SoS over the constituent systems, and the dependence of SoS

capability on systems which address single-system user needs as well as those of the SoS. Often, SoS are not formal programs of record but depend on changes made through individual system acquisition programs or operations and maintenance of fielded systems. As a result, the questions addressed here are not simply how do we implement T&E for SoS, but what does it mean to test and evaluate the SoS in the absence of formal acquisition direction.

constituent systems which evolve in response to their user needs, technical direction, funding, and management control independent of the SoS. SoS evolution then is based on cooperation and leveraging of its constituents, which are each addressing the needs of their original users, the subject SoS, and possibly other SoS. This leads to several key challenges. First, deliveries of system upgrades to meet SoS needs are done asynchronously and are often bundled with other changes to the system in response to needs beyond those of the SoS. Second, interactions among the systems may lead to unintended effects or emergent behavior; the larger the number and the greater the variability of the systems, the greater the likelihood of emergent behavior.

Table 1 contrasts the characteristics of systems and acknowledged SoS and highlights key implications for SoS T&E. The differences between systems and SoS shown in the table are largely a result of the independence of the

Table 1. Comparing Systems and Acknowledged Systems of Systems

Aspect [1]	System [1]	Acknowledged System of Systems [1]	SoS T&E Implications
Management & Oversight			
Stakeholder Involvement	Clearer set of stakeholders and aligned objectives	Stakeholders at both system level and SoS levels (including the system owners), with competing interests and priorities; in some cases, the system stakeholder has no vested interest in the SoS; all stakeholders may not be recognized.	Validation criteria more difficult to establish
Governance	Aligned PM and funding	Added levels of complexity due to management and funding for both the SoS and individual systems; SoS does not have authority over all the systems.	Can not explicitly impose SoS conditions on system T&E
Operational Environment			
Operational Focus	Designed and developed to meet operational objectives	Called upon to meet a set of operational objectives using systems whose objectives may or may not align with the SoS objectives.	System level operational objectives may not have clear analog in SoS conditions that need T&E
Implementation			
Acquisition	Aligned to ACAT Milestones, documented requirements, SE	Added complexity due to multiple system lifecycles across acquisition programs, involving legacy systems, systems under development, new developments, and technology insertion; Typically have stated capability objectives upfront which may need to be translated into formal requirements.	Depends on testing of constituent systems to SoS requirements as well as SoS level testing
Test & Evaluation	Test and evaluation of the system is generally possible	Testing is more challenging due to the difficulty of synchronizing across multiple systems' life cycles; given the complexity of all the moving parts and potential for unintended consequences	Difficult to bring multiple systems together for T&E in synchrony with capability evolution.
Engineering & Design Considerations			
Boundaries and Interfaces	Focuses on boundaries and interfaces for the single system	Focus on identifying the systems that contribute to the SoS objectives and enabling the flow of data, control and functionality across the SoS while balancing needs of the systems.	Additional test points needed to confirm behavior
Performance & Behavior	Performance of the system to meet specified objectives	Performance across the SoS that satisfies SoS user capability needs while balancing needs of the systems	Increased subjectivity in assessing behavior, given challenges of system alignment.

3 SoS T&E Challenges and Strategies

The seven core elements of SoS SE described in the SoS SEG [1] are illustrated in Figure 1. The four indicated by dashed outlines are critical to T&E of SoS. This section walks through SoS T&E challenges and draws on current SoS SE efforts for examples.

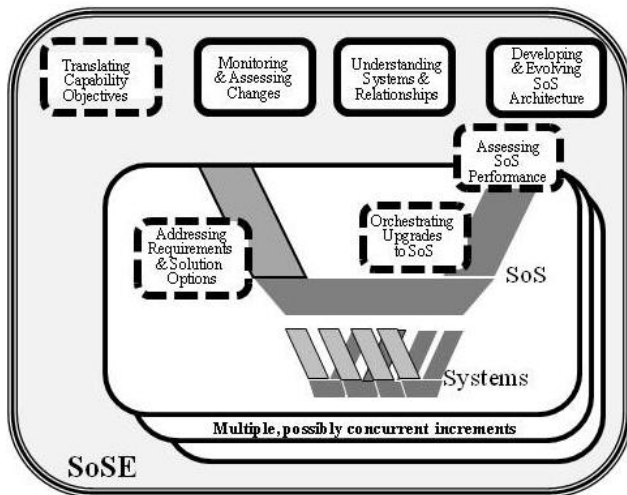


Figure 1. SoS SE core elements critical to T&E.

3.1 Level of SoS capability objectives

SoS capability objectives are often stated at a high level, particularly when an SoS is initially recognized. The objectives establish the capability context for the SoS, which grounds assessment of the current SoS performance. In many cases, SoS do not have ‘requirements’ per se; they have capability objectives or goals that provide the starting point for specific requirements which drive changes to the constituent systems in increments of SoS evolution.

For example, the Ballistic Missile Defense System (BMDS), capability objective is to defend against all ranges of enemy ballistic missiles in all phases of flight. [6] This defines the top level mission objectives and provides the foundation for identifying systems to support BMDS, for developing the BMDS architecture, and for recommending changes or additions to systems to enable the capabilities.

Similarly, the overall capability objective of the Enterprise Distributed Common Ground Station is to achieve Joint and Coalition intelligence, surveillance, and reconnaissance (ISR) mission interoperability through a multi-intelligence, multi-source collaboration strategy and by integrating ISR assets and information into command and control structures with linkages to national intelligence capabilities. [7] This includes the ability of a Joint Commander to flexibly tailor and employ ISR capabilities from any or all sources to support military operations.

Almost all SoS have these types of top-level objectives which guides the rest of the SoS actions. From a T&E perspective, the important point is that in SoS, capability objectives are not specific ‘requirements’ or even key performance parameters. Capability objectives are a starting point for developing a statement of expectations at the SoS level and require further specification and elaboration to conduct T&E.

3.2 Requirements specified at system level

Improvements in SoS performance accrue from additions to or changes in constituent systems, which collectively address the top-level capability objectives. SoS-level analysis identifies options for improvement. Alternatives are evaluated with the system SE teams that culminate in agreements with constituent systems owners on changes to be made to the systems to support the SoS. In the SoS SE core element “assessing requirements and solution options” [1], typically increments of SoS improvement are planned by the SoS and system managers and their SE teams. For each increment, requirements for changes in systems are specified, as well as an anticipated overall SoS performance effect. Defining specifications for system-level changes is generally straightforward. Defining specifications for the SoS capability that results from the cumulative system changes can be exceedingly difficult.

As a result, for most SoS, requirements are specified at the level of the system for each upgrade cycle and they provide the basis for assessing system-level performance. Consequently, T&E of system changes is typically done by the systems as part of their own processes; for changes introduced to benefit the SoS, T&E at the system level may not be able to demonstrate the intended SoS capability.

For example, BMDS [6] had adopted a ‘block’ process in which changes in systems are made for an increment of capability improvement in each block. The systems changes are documented and included in the baselines of the individual systems. This general approach is common across SoS. What vary are the ways the system requirements are specified and the formality of agreements between the SoS and systems. The key point for SoS T&E is that the requirements are typically specified at the system level, not the SoS level.

3.3 Implementation and test of SoS changes

Systems implement SoS changes as part of their own development processes and system level T&E validates the implementation of those system requirements. A major source of SoS T&E challenges is that SoS upgrades are the product of changes in systems which can and do operate independently from one another and of the SoS. In the core SoS SE element, Orchestrating Upgrades to SoS, the SoS SE team works with the systems’ SE teams to plan, fund

and track changes in systems which will contribute to SoS capability objectives. The type of SoS oversight employed depends on the nature of the changes and can range from simply getting reports from system level T&E to actively participating in system T&E design and implementation.

There are significant challenges in creating an SoS-wide test environment to assess a mission level SoS capability. In most cases, SoS integration and test does not comprehensively address the broad SoS capability objectives. Instead, it addresses one or more mission threads that are the focus of system-level changes. The costs of conducting a SoS wide test can be prohibitive when it includes assembling all participating systems, developing scenarios, and data collection and analysis across the SoS. The nature of SoS makes defining boundaries difficult. Systems whose influences are difficult to anticipate can impact system performance and testing.

In some cases the nature of the SoS objectives will drive the need for SoS-wide testing. The Naval Integrated Fire Control – Counter Air SoS [1] is one example. In the BMDS [6], system level tests are augmented with SoS testing by adding SoS collateral test events to constituent system testing. This can take considerable coordination but it takes advantage of already scheduled tests. The complexity may be mitigated via T&E of a subset of systems before fielding the entire SoS increment, although possibly with increased risk of T&E validity.

3.4 Asynchronous constituent system processes

Typically, constituent system development and testing are asynchronous and independent of the SoS. This challenges straightforward SoS level T&E. Integration and testing of the systems is an SoS team responsibility, worked in collaboration with the constituent systems' SE teams. While it is desirable to coordinate the development plans of the systems and synchronize the delivery of upgrades, as a practical matter it is often difficult or impossible. Even when synchronous developments are planned, asynchronous deliveries may result from delays in individual system development schedules, particularly when there are many systems or their developments are complex.

SoS constituent systems generally make changes as part of a development increment that will be ready to field when successfully tested and evaluated at the system level. However, other systems in the same SoS increment may not be ready to test at the same time, thwarting end-to-end SoS level testing. As autonomous entities, the individual systems expect to field their systems based on their results, independent of the larger impact on SoS capability. Delaying a system from fielding until all systems in an increment are ready to test is impractical and undesirable in most cases. Systems owners are understandably reluctant to defer delivery pending additional SoS integration tests,

and even more reluctant to stop system delivery if SoS T&E uncovers problems when the single system test was successful. The management independence of the systems means that the systems are not constrained by SoS level testing. Consequently, contingency plans should be prepared for this situation.

These issues were addressed in the SoS SEG. Referring to the issues raised by the independence and asynchronous development schedules of constituent systems, the SoS SEG states that SoS have addressed this conundrum in different ways. For example, "... a number of SoS initiatives have adopted a 'bus stop,' spin, or block-with-wave type of development approach in which there are regular time-based SoS 'delivery' points, that systems target for their changes. Integration, test, and evaluation are done for each drop. If systems miss a delivery point because of technical or programmatic issues, they have another opportunity at the next point (there will be another bus coming to pick up passengers in 3 months, for instance). The impact of missing the scheduled bus can be evaluated and addressed based on the specifics of the development cycles. By providing this type of SoS battle rhythm, discipline can be inserted into the inherently asynchronous SoS environment. In a complex SoS environment, multiple iterations of incremental development may be under way concurrently (e.g., MDA concurrent blocks in the development of the BMDS; NSA roadmap)." [1, pg. 68-69]

However, the SoS SEG also points out that there are downsides: "Approaches such as this may have a negative impact on certification testing, especially if the item is related to interoperability and/or safety issues (such as Air Worthiness Release). When synchronization is critical, considerations such as this may require large sections of the SoS, or the entire SoS, to be tested together before any of the pieces are fielded." [1, pg. 69]

The impact of the asynchronous constituent systems development processes leads to the same type of issues for testing.

3.5 SoS performance assessment

SoS capability objectives provide a foundation for identifying systems supporting an SoS, developing an SoS architecture, and recommending changes or additions to systems to meet the capabilities. They also provide the basis for defining and measuring top-level SoS performance.

In SoS, addressing SoS level performance is typically tied to 'end to end' SoS functionality often portrayed as mission threads which capture the activities implemented across the SoS to meet SoS objectives. These cross-cutting sets of activities collectively constitute the SoS behavior. Hence when looking at SoS performance, it is important to measure behavior of the individual systems in the context of the end to end behaviors supporting SoS capabilities.

This implies a need to generate metrics defining the end-to-end SoS capabilities that provide a 'benchmark' for

SoS development. Developing these metrics and collecting data to assess the state of the SoS is accomplished as part of the SoS SE core element “assessing the extent to which SoS performance meets capability objectives over time”. [1] This element provides the capability metrics for the SoS. They may be collected from a variety of settings. They provide input to the SoS SE on the performance of the SoS under a variety of conditions, and serve as a source of information about new or emerging conditions that affect the SoS. Hence, assessing SoS performance is an ongoing activity, which goes beyond assessment of specific changes in elements of the SoS (e.g. changes in constituent systems to meet SoS needs, and system changes driven by factors unrelated to the SoS).

T&E objectives, particularly key performance parameters, are the basis for making a fielding decision. Because SoS are typically comprised of a mix of fielded systems and new developments, there may not be a discrete ‘SoS’ fielding decision; instead the various systems are deployed as they are ready, at some point reaching the threshold that enables the new SoS capability. Consequently, SoS metrics, discussed above, provide a ‘benchmark’ for SoS development which should show an improvement over time in meeting capability objectives.

In some circumstances, the SoS capability objectives can be effectively modeled in simulation environments which can be used to identify appropriate changes at the system levels. The fidelity of the simulation provides the validation of the system changes needed to enable SoS-level capability. In cases in which the system changes are identified by SoS-level simulation, the fidelity of the simulation may also be used for SoS T&E.

Given the nature of SoS T&E challenges, it might be surmised that modeling and simulation (M&S) is used extensively through the SoS SE process, particularly in Assessing Performance to Capability Objectives. Interviews with SoS SE practitioners in developing the SoS SEG indicated that while the potential of M&S was widely appreciated, its use was limited. A follow-on survey [8] confirmed this initial finding and illuminated several inhibiting factors, including: shortage of staff skilled in applying M&S to SoS SE, insufficient fidelity, flexibility and adaptability of tools, difficulty in obtaining data to populate the M&S tools, and lack of funding.

4 Strategic Approaches to SoS T&E

Given these SoS T&E challenges, there are several opportunities to provide users with better information on expected systems and SoS performance.

4.1 SoS SE as the Framework for SoS T&E

To effectively conduct T&E, there needs to be a clear understanding of objectives and requirements of the ‘test

item’. For an SoS, where the value is accrued from the collective behavior of the SoS toward user capabilities, it is critical that systems engineering be conducted at the SoS level to develop the capability objectives and develop metrics to address performance of the SoS capabilities over time. These SoS objectives and metrics serve as the basis for requirements on the constituent systems and for setting and evaluating SoS capability test objectives and methods.

The effective application of SE at the SoS level, does not remove the challenges of SoS T&E, but it does provide a structured framework to address these challenges. As the SoS SE team develops approaches to addressing the asynchronous development paths of the constituent systems, they can consider extensions of these approaches to support SoS T&E. For example, in some cases SoS SE teams employ periodic recurring test events to address changes in the constituent systems using an extension of the ‘bus stop’ development approach. [1] These periodic SoS regression tests ensure changes in the constituent systems have not impacted the SoS performance.

In other ways the SoS SE actions can help mitigate T&E challenges. As is discussed in the SoS SEG [1], SoS architectures which shelter the SoS from changes in the systems tend to be more robust over time. They may also facilitate more partitioned testing, reduce the number of active participants needed to test changes in systems and assess the impact on the SoS.

4.2 Evidence-based approach

In many cases, the idea that the SoS can be ‘fully’ tested before deployment is simply not realistic. It may be more appropriate to view SoS T&E as an evidence-based approach to addressing risk. The SoS SE team identifies issues critical to success of each increment of SoS development, as well as places where changes in the increment might adversely impact user missions, and then focuses pre-deployment T&E on them. Risks are assessed using evidence from a range of sources, including live test. The evidence is based on activity at the SoS level, as well as roll-ups of activity at the constituent system level. The activity might include explicit verification testing, results of models and simulations, use of linked integration facilities, and results of system level operational test and evaluation. Analytical models of the SoS behavior are used to assess system level performance in operational scenarios, validate requirements allocations to systems, and otherwise provide an analytical framework for SoS level verification. The models may also be used to develop expectations for SoS performance. Typically, operational conditions are developed with end user input and sometimes guided by the design of experiments to explore a broad range of conditions to identify and assess risks. Finally, these risks are factored into SoS and system development plans. If T&E results indicate that the changes will have a negative impact, they can be discarded

or postponed without jeopardizing the delivery of the other system updates. The results are then used to provide feedback to end users in the form of ‘capabilities and limitations’ as done by the Navy Battle Group Assessment process [9], instead of test acceptance criteria for SoS ‘deployment’.

SoS SE teams employ a range of venues to assess SoS performance over time. SoS end-user metrics can assess the results of systems changes on SoS capability performance over a range of opportunities, both planned and opportunistic. Performance data from the latter can support periodic assessments of evolving capability and provide valuable insight to developers and users, including the identification of unexpected behavior.

4.3 Feedback Process

Because of the difficulty in assessing SoS performance before fielding, it is prudent to establish a robust process for post-fielding feedback. Once deployed, continuing "T&E" of the SoS can identify operational problems and be the basis for future improvements. This continual evaluation can be facilitated by instrumenting systems to collect data to provide feedback on incipient failure warnings, and unique operational conditions. This provides a vital link to emerging operational needs of the SoS. In addition to instrumenting systems for post-fielding data collection, consideration should be given to embedding a member of the SoS SE or management team with the SoS operational organization. Well-developed, continually exercised feedback mechanisms between operational and acquisition/ development communities are an enabler of “building the system right” and continuing to do so throughout the multiple increments of SoS evolution.

5 Conclusions

This paper has reviewed characteristics of SoS and the challenges they pose for SoS T&E. Typically SoS evolve via constituent systems incorporating changes to their development plans to meet SoS needs. As a result, SoS capabilities are developed and tested as part of system development activities. As autonomous entities, constituent system owners expect to field their systems based on their own T&E results, independent of the impact on SoS capability. Deferring system upgrades until all constituents in an increment are ready to test successfully is impractical and undesirable in most cases. Postponing the fielding of a successfully tested system because of problems in testing the SoS may not be practical. Since most SoS are comprised of already fielded systems, there may not be a discrete fielding decision. Full SoS level testing can be costly and it can be very difficult to create test environments which realistically represent the expected results in an operation environment because of the size and complexity of many SoS environments. These form core impediments to mapping traditional T&E to SoS.

The paper describes several higher level strategies that SoS SE teams are employing to achieve “right-sized and effective SoS-level testing. These focus on SoS SE establishing a framework for SoS T&E, evidence-based approaches, assessment of the SoS over time, and extending testing to include continual feedback processes. While this may not be optimal from the SoS T&E perspective, it fits well with the US DoD business model for managing the evolution of SoS capabilities, leveraging existing testing activities at the constituent system level, identifying opportunities for testing at the SoS level, and in some cases, building in SoS assessment capabilities into the SoS itself.

Acknowledgements

The authors would like to thank John Palmer, The Boeing Company, for his insights into the issues presented in this paper.

References

- [1] Department of Defense, *Systems Engineering Guide for System of Systems*, version 1.0. 2008.
- [2] Bjorkman, Eileen et al., “Testing in a Joint Environment 2004-2008,” *ITEA Journal* 2009; 30:39-44.
- [3] Columbi, John et al. “Interoperability Test and Evaluation: A Systems of Systems Field Study,” *Crosstalk*, November 2008.
- [4] Conley, Stephen, “Test and Evaluation Strategies for Network-Enabled Systems,” *ITEA Journal* 2009;30-111-116.
- [5] Valerdi, Ricardo. “A Prescriptive Adaptive Test Framework (PATFrame) for Unmanned and Manned Autonomous Systems: A Collaboration Between MIT, USC, UT Arlington, and Softstar Systems”, Presentation at USC Center for Systems and Software Engineering Annual Research Review, March 10, 2010.
- [6] Eccles, D. “Engineering and integrating the Ballistic Missile Defense System”, *Crosslink*, Aerospace Corporation, Spring 2008.
- [7] Meiners, Kevin. “Net-Centric ISR”, Presentation at NDIA SE Conference, October, 2005.
- [8] Lane, JoAnn, et al. “Modeling and Simulation Support for the SE of SoS”, Presentation at NDIA SE Conference, San Diego, CA, October 2009.
- [9] McConnell, Jeffrey. “The Navy Distributed Engineering Plant -- Value Added for the Fleet” Presented at the Engineering for the Total Ship Symposium, Gaithersburg, MD, February 2002.