GIG PERFORMANCE ASSESSMENT FRAMEWORK

Tony Modelfino, George Case
Stratogis Networks LLC

Tony Desimone
OSD Networks &
Information Integration
and

Subramaniam Kandaswamy
Johns Hopkins Applied
Physics Laboratory

ABSTRACT

This paper discusses the GIG Performance Assessment Framework (PAF) that was developed by the Office of the Secretary of Defense for Networks & Information Integration (OSD NII) to evaluate E2E application and service performance across the GIG, particularly to the tactical edge. The paper describes the Use Case based strategy that was developed to define GIG operational scenarios, and the simulation models that were developed to predict end-to-end performance. In addition, the paper details the Performance Evaluation Tool (PET) that was developed to allow rapid assessment and parametric analysis of GIG performance. Also, the function of the GIG Performance Working Group and its subgroups are described. This paper also contains a brief summary of evaluation comments gathered from share holders and hands-on testing experience with the PET. These feedback data is currently being used to identify performance shortfalls and to evaluate the effectiveness of enterprise network solutions, being developed as part of the Enterprise Wide GIG System Engineering effort.

INTRODUCTION

The Performance Assessment Framework (PAF) was developed by the Office of the Secretary of Defense for Networks & Information Integration (OSD NII) for evaluating Global Information Grid (GIG) end-to-end (E2E) performance and ensuring E2E performance will meet end-user expectations and needs. The PAF was developed to support the development of the NetCentric Implementation Document (NCID). The PAF enables the identification of GIG performance shortcomings and provides a methodology and facility for evaluating the effectiveness of end-to-end solutions. The PAF goal is to present E2E performance in metrics that end-users can readily understand and evaluate, such as service application availability and response time. This is in sharp contrast to the packet level performance metrics typically used to categorize GIG transport segment performance.

The PAF is required for a number of reasons. First, individual GIG transport development programs are, and should be, focused on the performance of a single network and typically do not evaluate E2E performance across multiple networks. Therefore, segment engineering design decisions may be made to optimize intra-segment performance, without recognizing the impact on E2E performance over multiple segments. Second, GIG application and services development programs typically do not consider the full range of transport network performance, particularly tactical edge networks, when developing new end user applications. As a result, tactical users may experience degraded application performance due to low bandwidth, high delay and high packet loss that often plague tactical edge networks. Finally, GIG component development programs seldom consider interactions between all the layers of the data plane and control protocol stack. For example, GIG transport programs do not validate segment performance based on E2E application performance, but rather via segment level packet performance. This can result in misleading E2E performance and capabilities estimates.

Originally, the PAF was envisioned to be a strategy for assigning portions of an E2E performance target to individual GIG segments and sub-segments. It became apparent that this top down allocation strategy could not be successful for a number of reasons. First, there was no definitive acceptable E2E performance threshold for GIG applications and services. Second, it was impossible to allocate portions of E2E performance to GIG segments since segments do not specify performance using these metrics. For example, transport segments define and assess segment performance using packet delay and loss characteristics not message delay. Similarly, services and application programs specify performance as measured at the Local Area Network (LAN), or possibly DISN core interface, not at the tactical edge. Even if such an allocation were possible, a strategy for determining the most practical and cost effective allocation between segments or programs does not exist. Third, GIG segment performance, particularly transport performance is constrained by inherent physical and/or technical limitations which cannot be improved; satellite propagation delay and rain attenuation being examples of this. There was considerable concern that PAF segment and sub-segment allocations might not be physically or technically achievable given these inherent constraints. Finally, even if all these issues could be resolved, or mitigated, it might not be practical to define and impose a large number of new requirements on existing development programs. Fortunately, the objective of the PAF process is not to optimize E2E GIG performance, but rather to insure acceptable performance across the GIG while minimizing the number of new segment requirements and maintaining intra-domain performance.

The PAF addresses these problems by defining a comprehensive set of E2E use cases which span the full spectrum of GIG user types, applications and networks. In addition, the PAF defines a set of operating conditions and a segment performance categorization strategy that is consistent with GIG segments' approach to evaluating and specifying performance. Finally, the PAF developed a modeling and simulation strategy and a set of tools for assessing E2E performance. This included a packet level simulation model that was developed to predict message level transfer time across the full range of GIG transport networks. The results from this model were integrated into a Performance Evaluation Tool (PET), which estimates performance for thousands of GIG use cases and provides the capability to rapidly assess the impact of segment performance or architecture changes on E2E use case performance. The PAF modeled E2E performance under a wide range of transport operating conditions to assess the impacts of operating load and other environmental conditions on E2E performance.

The PAF is meant to be used by segment developers, planners and operators, and end-users as the GIG capability evolves. Segment developers can use the process to assess the impacts of other GIG segments on their segment performance. Additionally, segment developers can evaluate the impacts of performance changes to their segment on E2E performance. GIG planners and operators can use the process to evaluate the impacts of service architecture decisions on E2E performance, particularly for tactical users. For end-users, the process provides an estimate of application performance and its impact on mission performance and mission effectiveness. In summary, the PAF is not meant to be a single pass process, but rather an iterative process with multiple feedback loops. These feedback loops insure that use cases are representative of critical DoD communication requirements; that GIG component performance assumptions are reflective of actual component performance; that performance shortcomings are real and warrant correction; and that solutions consider the impacts on all GIG developers, operators and users in a clear and transparent process. Since the PAF is an evolving process, a Performance Working Group (PWG) was initiated to bring users, developers and operators together to insure that the PAF was accurately representing and assessing segment performance and users needs. The following sections describe the development of GIG Use Cases, the E2E GIG modeling strategy, the tools developed as part of the PAF process and the workings of the GIG PWG.

GIG USE CASES

The first step in developing the PAF was to identify GIG user and mission requirements. This was initially accomplished by considering the Netcentric Operating Environment Joint Integrating Concept (NCOE JIC), NetCentric Enterprise Services (NCES), Joint Mission Areas and multiple transport and application program DODAF operational views and Information Exchange Requirements (IER). These user and mission requirements clearly show that the GIG is not a homogenous network, but rather a collection of networks with widely different performance capabilities and limitations. Similarly, GIG users span a broad spectrum including strategic, tactical, and business functions. These users can access the GIG from locations as varied as fixed sites in CONUS to an HMMWV (HumVee) in theater. In addition, these users will employ a wide range of applications that will place different demands on the GIG.

Ideally, application performance should not depend on the user type, location or network, but in practice it does. GIG networks and access technologies have inherent bandwidth, latency and loss characteristics that will affect application performance. The PAF recognized that these limitations made it impossible to define a single application performance objective for all GIG users across all networks. Instead, performance must be assessed for each user type, access technology and network connectivity. Toward this end, a set of GIG use cases were developed. Each use case was composed of a GIG user or users, GIG ingress/egress access technology, a GIG transport network or networks connecting users and a GIG service or application.

Initially, the PAF attempted to identify only the most stressing GIG use cases in an effort to minimize the number of scenarios that had to be analyzed. It became apparent that this strategy was destined to fail given that performance requirements varied significantly across the full range of GIG users. Ultimately, the PAF chose to evaluate a broad range of use cases. Currently, over 5000 use cases have been identified and analyzed for a variety of service architectures. A typical use case is shown in Figure 1. In this example, the commander in the COTM vehicle queries a portal located at a CJTF for available imagery. The COTM terminal connects to the beyond line-of-site portal using a satellite network such as WGS or TSAT. The wireless network interconnects with the satellite at a

SATCOM PoP while the portal server has a high speed connection to a teleport. The portal relays the search query to a CONUS datacenter using a satellite connection through a teleport which is connected to the DISN-core (GIG-BE). The request is relayed to a CONUS datacenter such as the Defense Enterprise Computing Center (DECC). The datacenter authenticates the user, performs a federated search and returns the results to the portal which then relays the results to the commander in the COTM vehicle. The commander then downloads the imagery from the closest content delivery server, which in this case is the COTH command post located in theater. This service oriented architecture allows for greater data dissemination and improved decision making, but also requires additional user-to-service and service-to-service communications. The objective of the PET is to identify these service communications, define the transport paths and determine the E2E transfer time for each constituent message and the overall E2E service response time and availability.

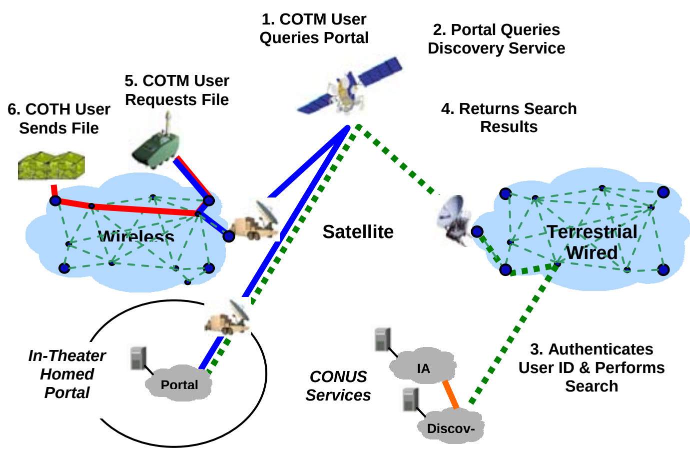

Figure 1 – Typical GIG Use Case

GIG users are defined as the senders and receivers of information and can be either people or machines. The current version of the PAF defines eleven user types based on the users' operational mode and technology employed for accessing the GIG. These user types are listed in Table 1. Each user type has distinctly different operational capabilities and performance.

GIG connectivity is defined as a network or series of networks that connect sender and receiver user types. The PAF currently includes six GIG wired and wireless network types. The current PAF network types and examples of the networks that have been integrated into the PET are shown in Table 2. This delineation was selected because the categories have distinctly different bandwidth, packet loss, packet delay and availability performance. The PAF recognized that there may be significant capability variation between network types within a given class; for example, Wideband Gap Filler Satellite (WGS), Advanced Extremely High Frequency (AEHF) and TSAT are all considered satellite networks yet there are significant performance and operational differences between these architectures. As the PAF matures, additional transport networks will be added to the process and to the PET.

Table 1 GIG User Types

User Type	
Dismounted	Aircraft, Tactical
COTM	Aircraft, C2
COTP	ISR Aircraft
COTH	Ship
Fixed – CONUS	Submarine
Fixed – OCONUS	

Table 2 GIG Network Types

Network Type	PET Component
Wireless Ad Hoc	JTRS SRW
Wireless PTP	JTRS WNW
SATCOM (1 Hop)	TSAT (1 Hop)
SATCOM (2 Hop/ Cross Link)	TSAT (Cross Link)
Intra-Theater Wired	Intra-Theater GIG-BE
Inter-Theater Wired	Inter-Theater GIG-BE

The PAF defined the access capabilities for each user type to each GIG network type. Access metrics included bandwidth, ingress/egress delay and packet loss, and availability. In addition, the PAF categorized the performance of each network type as a function of IP service class. The PAF considered five IP service classes that were consistent with the NCID T300 Service Class segmentation. Network performance metrics included packet loss, packet delay (minimum, mean, standard deviation, 90th and 95th percentile delay) and transport segment availability. Finally, the PAF characterized the performance of interconnection nodes that are used to connect networks such as gateways, wireless POPs and teleports. Major GIG net-centric transport programs such as TSAT and JTRS provided network and user performance data based on analysis, test or program requirements, as available.

The PAF created 26 composite networks composed of the six GIG network elements listed above. In principle, there are many more possible network combinations that could be built from the six network types. These 26 were selected because they represent highly likely GIG connectivity configurations. User-to-user connectivity was defined using these 26 possible composite networks. In total, 380 user-to-user connection paths were defined in the PAF and included in the PET. Both the framework and the PET software can be readily adapted to support additional connection paths that may be identified during PAF development.

GIG user services are applications or sets of applications that users execute over the GIG. The PAF currently includes over 30 different applications ranging from legacy applications to Netcentric service oriented applications such as collaboration and discovery. These are summarized in Table 3. The legacy services represent the dominant services, from a bandwidth perspective, on the GIG today based on Secret Internet Protocol Router Network (SIPRNET) and Non-classified Internet Protocol Router Network (NIPRNET) network monitoring that was conducted by OSD NII. These services also represent a significant portion of traffic identified for future tactical and satellite network programs such as WIN-T and TSAT. Netcentric services are currently modeled after NCES services. The PAF recognizes that additional services will be used over the GIG, particularly specialized Community of Interest (COI) services, and the framework can be upgraded to include those additional services as they are identified.

Table 3 Example of GIG Applications in PET

Legacy Applications	SOA Applications
VoIP	Discovery
VTC	Collaboration
Sensor Streaming	Mediation & Messaging
HTTP	Security Services
FTP	Content Delivery
Email	Net Management

The PAF defines which user types utilize each application. Each application or service is decomposed into a series of building block message exchanges that resemble a service or application DODAF OV-6C event sequence diagram. Each message exchange has a sender and receiver which may be another user type or a service. In addition, each message is defined by a message size, transport protocol, and service class. The following section describes the methodology for assessing the performance for each message.

GIG use cases can be complicated, involving multiple user-user, user-service and service-service communications as shown in the previous example. A single use case can involve multiple nodes and multiple networks. In addition, the number of GIG use cases grows combinatorially as more user types, networks and applications are added to the analysis. The current PET combination of GIG user types, composite networks, applications/services, and service architecture has the potential to generate over 1 million use cases if every combination was considered. Fortunately, reasonable operational assumptions reduce the number of use cases significantly. The objective of the PET is to automatically generate these use cases from a small number of user inputs. The PET manages a process that allows the user to rapidly select these parameters; generate use cases; estimate end-to-end performance and evaluate that performance relative to end-user requirements. PET presents the results in a fashion that allows the user to identify those messaging or processing events most responsible for a performance shortfall.

GIG E2E PERFORMANCE MODELING

The objective of the PAF network modeling is to estimate E2E application performance for each GIG use case based upon the performance of the individual GIG components involved. The GIG network modeling strategy needed to strike a balance between accuracy and calculation complexity. Given the wide range of GIG applications and network types, many of which are still in development, it became apparent that a single E2E GIG model that included all network and application features did not currently exist. In addition, while various GIG segments had developed a range of modeling and simulation tools to assess segment level performance, it would not be feasible to integrate these tools in the short term. Therefore, the PAF decided to develop both short term and long term network performance modeling strategies.

The long term modeling strategy seeks to develop an integrated E2E model that combines program supported detailed GIG segment models. The success of this approach hinges on selecting a standardized core simulation model upon which to integrate each of the segment models and to develop a standardized set of model interfaces that would enable easy integration of the segments models. Most GIG segment models are built using an OPNET core simulation model which makes an integrated model possible. Unfortunately, segment models, particularly for future GIG netcentric transport and services programs have not been developed with standardized interfaces. The PAF thus decided to implement a short term network modeling strategy that was able to predict E2E use case performance given the current state of GIG segment models.

The short term modeling strategy simulates GIG network connectivity as an IP cloud whose performance is defined by the IP packet delay, loss and network availability of its constituent segment networks. The segment delay characteristics are assumed to have an offset Gamma probability distribution (PDF). The decision to use an offset gamma delay distribution was arrived at after considerable analysis of simulated and monitored network delay performance provided by GIG programs and measurements taken for commercial IP networks. The gamma distribution has the heavy tail characteristics typical of network congestion particularly wireless bandwidth-on-demand networks.

The delay distribution for each GIG transport network was generated by service class using the minimum delay, the average delay and delay variance that the GIG network segments provided. A standard gamma distribution would generate packet delays that vary randomly from packet to packet; thus, one packet might experience a long delay while the next packet may experience a significantly shorter delay. This behavior is not generally consistent with measured and simulated packet delay performance for messages comprised of multiple TCP transmission segments. In fact, E2E delay for packets associated with a common flow tend to be highly correlated as these packets typically follow the same network path and experience similar network delay. The PAF delay model incorporates this behavior into the end-to-end delay model by correlating packet delay for individual messages. The correlation decreases as the difference in packet generation times increases. This delay modeling strategy produces a heavy tailed packet delay distribution behavior. In the PAF, an abstracted model that captures this performance behavior is referred to as the 'IP cloud model'. Comparison of end-to-end packet delay correlation using a complete network simulation model (a model explicitly includes all routers and links between nodes) has shown that this abstracted modeling approach produces results that are consistent with detailed end-to-end model message delay results, as will be discussed next.

In an effort to assess the accuracy of the IP cloud model, we simulated E2E performance for a 55 terminal, bandwidth-on-demand (BoD) satellite architecture employing IP QoS and supporting three real-time (VoIP, VTC and sensor streaming) and two data (FTP and HTTP) applications. The satellite transponder was operated at a 70% average load. Table 4 compares the mean, 80th and 90th percentile E2E message transfer times from the IP cloud model to the full end-to-end model. The E2E model included all the datalink layer and BoD protocols while the IP cloud model simply used the mean packet loss and minimum, mean and standard deviation packet delay results from the E2E model. The results show excellent agreement (within +/-10%) in all metrics for 500 and 5000 KB imagery message transfer times. The results for 5 and 50 KB C2 messages show excellent agreement for the average and 80th percentile metrics, however the cloud model tends to overestimate 90th percentile transfer performance relative to this architecture. Given that the ultimate objective of the PAF is to identify performance shortfalls, this level of accuracy appears to be adequate in the near term.

Table 4 Comparison of E2E Transfer Times (seconds)

Message Size (KB)	Performance Metric	COTP Terminals	
		Cloud Model	E2E Model
5	Average	4.2	3.7
50		7.2	6.7
500		15	15
5000		45	43
5	80th Percentile	4.5	4.3
50		8.1	7.8
500		17	19
5000		48	44
5	90th Percentile	9.4	5.2
50		15.0	9.1
500		23	26
5000		59	61

A typical GIG service or application may be composed of multiple messaging events. Figure 2 shows the event sequence (ES) diagram for a user authentication service. The ES diagram shows 6 messaging and 3 processing events. This service requires a total of 33 messages (10 user-to-service and 23 service-to-service) to identify and authenticate the user. Most of the messages involve a single small packet such as a TCP handshake which includes three 40 byte messages. While the service-to-service messages are transported over a high speed WAN (GIG-BE) or a LAN, the user-to-service messages can experience significant delay and packet loss depending on the path.

The PAF identified the messaging characteristics for each GIG service based on program OV-6C event sequence diagrams. Each messaging event was then decomposed into a series of building block components which are then further decomposed into standard networking protocols. Examples of building blocks and standard networking protocols used in the PAF are shown in Table 5. The standard protocols define a series of messages whose E2E transfer time is computed based on the bandwidth, delay, loss and availability of the composite network that connect the sender and receiver of each message. The total application response time and service availability is determined by combining the performance for each constituent message, taking into account that some messaging events occur serially while others are executed in parallel. The following section describes the Performance Evaluation Tool that was developed to facilitate the calculation of E2E performance for a large number of GIG use cases.

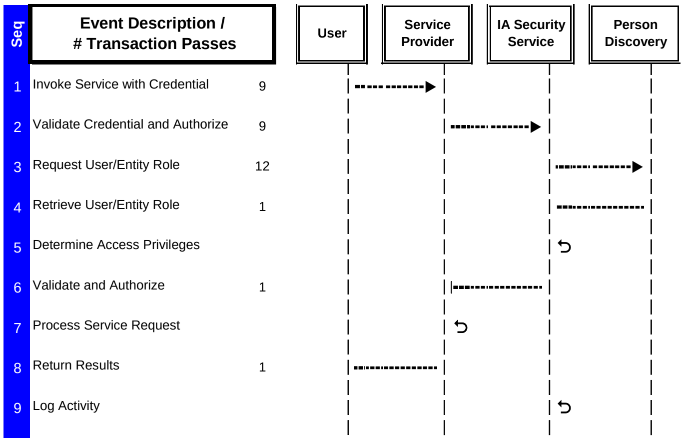

Figure 2 – Typical GIG Use Case

Table 5 – Sample Building Block Protocols

Layer 2 Building Blocks	Layer 1 Protocols
Discover	TCP Connection Setup
Messaging Setup	TLS Handshake
Request User/Entity Role	Directory Lookup (LDAP)
Service Request	DNS Query
File Transfer	HTTP Request/Response
Collaboration Set Up	Session Initiation Protocol (SIP)
	SMTP Setup
	UDP Logging
	SOAP

PERFORMANCE EVALUATION TOOL

The PAF recognized that the large number of GIG use cases coupled with the wide array of GIG segment design and performance parameters required an automated process to generate and assess use case performance. The PAF developed an interactive Performance Evaluation Tool (PET) to serve this purpose. The PET was built in EXCEL to enable distribution beyond the GIG modeling and simulation community to the larger GIG system engineering and program engineering community. The PET has a graphical user interface (GUI) that enables the analysis of a single GIG use case or the full set of GIG use cases. The GUI also enables the user to add additional GIG users, networks or applications and automatically generates new use cases. The PET includes the performance for each GIG segment and allows the user to parametrically vary segment performance to assess the impact on E2E performance. Segment performance includes parameters such as network delay, packet loss, node availability, server response/processing time, etc.

The PET includes a large database of message delay performance results that were generated using the OPNET-based packet simulation IP cloud model. The database contains E2E message transfer time performance as a function of network ingress/egress load, message size, service class, E2E delay, E2E packet loss and minimum connection bandwidth. In addition, the database includes performance for standard TCP (Windows 2000 default settings) and TCP with Performance Enhancement Proxies (PEP). TCP PEP is integrated into a number of DoD tactical networks as a means for improving performance on high latency and loss satellite and wireless links. The TCP PEP used for the PET incorporated a modified congestion avoidance algorithm that took advantage of the explicit congestion notification (ECN) flag in the IP header. This ECN-based TCP PEP approach is consistent with a HAIPE-based black core network provided GIG routers are ECN enabled. Other TCP enhancement approaches are being evaluated for future PET upgrades. The database was generated by simulating multiple nodes for thousands of simulated seconds. A single simulation generated over 100,000 messages and 10 million IP packets. The results of the simulations were processed to generate a statistical distribution of packet and message transfer time and the PET performance database includes mean, 75^{th}, 90^{th}, 95^{th} and 98^{th} percentile E2E performance. The PET user selects the percentile performance metric of interest as well as the desired operating loads and TCP implementation.

The PET tool includes output post–processing features that can be used to identify performance shortfalls based on specified minimum performance requirements for each GIG service/application. The tool includes target application performance thresholds derived from the NCOE JIC, program message/application speed of service requirements and industry standard performance thresholds. The PET compares the performance of each use case to the specified threshold and allows the tool user to drill into any use case to isolate the cause of poor performance.

The PAF built the PET to serve a wide range of purposes and the software is being made available to the GIG user, developer and operator communities. GIG end-users can use the tool to estimate E2E performance relative to end-user mission requirements. Both transport and services GIG segment developers can use the tool to investigate the sensitivity of E2E performance to segment level performance and to overall service architecture. The tool currently supports CONUS and global fixed site centralized service architectures, and in-theater and portal based decentralized service architectures. This allows the causes for performance shortfalls to be readily identified and solutions evaluated. Finally, GIG operators can use the tool to assess the impact of operating load and Service Level Agreement provisioning on E2E performance.

PERFORMANCE WORKING GROUP

The Performance Working Group (PWG) was instituted in the Spring of 2006 as a mechanism for refining and updating the PAF. The team coordinated with transport, services, and infrastructure developers and included members from DISA, Joint Staff (J6), Joint Forces Command (JFCOM), Space and Naval Warfare Systems Command (SPAWAR), Department of Defense for Acquisition, Technology, and Logistics (ATL) and Naval Research Laboratory (NRL) personnel. The PWG met on a monthly basis to review the network and mission modeling strategies; refine GIG use cases; agree upon GIG operating assumptions; develop a strategy for categorizing and obtaining GIG segment performance; and reviewing GIG E2E performance results. PWG members also met regularly with NCID working groups (including the QoS, Services and Computing and Infrastructure working groups) to insure the consistency of PAF models with the NCID compliance requirements. The results of the PWG were captured in the GIG Performance Assessment White Paper V3.0 which was released in October 2006.

The PWG was composed of four sub-groups: Mission Modeling, Network Modeling, Use Case Development and Application Event Definition. The PWG formed a Mission Modeling (MM) sub-group to identify documented joint mission areas, mission threads, and mission measures-of-effectiveness (MOEs) and devising a realistic mission modeling strategy that characterized appropriate message traffic input files for the network model. The message traffic characteristics were developed from architecture products that reflect doctrinal force structure and operational context. Mission thread(s) included identification and mapping of mission messaging requirements to PAF user types and applications/services. Also a network modeling (NM) sub-group was formed to identify a short-term and long-term modeling and simulation strategy for predicting GIG end-to-end application performance for the wide range of GIG use cases. The NM sub-group reviewed a series of potential GIG network models and down selected the IP cloud model as the short-term approach for assessing E2E performance. As part of this effort, the NM defined the strategy for defining GIG transport segment performance and the performance parameters that were required for each GIG transport segment. Finally, the NM sub-group defined the operating conditions (load, environmental, traffic models) for assessing GIG performance.

The Use Case Development (UCD) sub-group was formed to evaluate these use cases for accuracy and completeness. The UCD sub-group identified additional user types and network connectivity paths that were missing from the original Use Cases. In addition, the UCD sub-group identified new GIG services and service architectures. Finally, the Application Event (AE) development subgroup was formed to identify the messaging components of current and future Netcentric services and applications. This effort included identifying end nodes, service architectures, application protocols and message characteristics for each service.

PET EVALUATION AND FUTURE EFFORTS

The PET software has been used and evaluated by a number of DoD organizations including: major programs, network operators, and modeling and simulation groups. The evaluations served to provide performance information for the model; evaluate the accuracy of the models; identify interface changes that improve ease-of-use; and identify upgrades and improvements. The PET model and interfaces have been revised to reflect many of these suggestions. In addition, the PET development team is evaluating the upgrades and improvements that were suggested. These include: developing a of mission modeling capability to link service/application performance thresholds to mission requirements; upgrading the model to include a broader range of operating loads and background traffic; incorporating additional E2E performance metrics; incorporating additional TCP performance enhancement proxies into the model; and developing a web-based version of the model to improve configuration control, facilitate ease-of-use and broaden PET distribution in the DoD community. In addition, the team is evaluating approaches to upgrade the framework and models to include additional operating conditions such jamming, On-The-Move blockage, and Denial of Service attack) and operating loads.

The PET team is initiating a Pilot effort to obtain segment performance for additional transport, service and applications programs. This data will be added to the PET and the use cases will be expanded to include these programs; thereby providing a better understanding of E2E GIG performance for a broader range of operating environments.