VA Enterprise Design Patterns:

2. Enterprise Architecture

2.2. End-to-End Application Performance Management (APM)

Office of Technology Strategies (TS)
Architecture, Strategy, and Design (ASD)
Office of Information and Technology (OI&T)

Version 2.0

Date Issued: December 2015

THIS PAGE INTENTIONALLY LEFT BLANK FOR PRINTING PURPOSES

APPROVAL COORDINATION

TIMOTHY L MCGRAIL 111224

Digitally signed by TIMOTHY L MCGRAIL 111224
DN: dc=gov, dc=va, o=internal, ou=people, 0.9.2342.19200300.100.1.1=tim.mcgrail@va.gov, cn=TIMOTHY L MCGRAIL 111224
Date: 2015.12.29 13:54:50 -05'00'

Date:

Tim McGrail
Senior Program Analyst
ASD Technology Strategies

PAUL A. TIBBITS 116858

Digitally signed by PAUL A. TIBBITS 116858
DN: dc=gov, dc=va, o=internal, ou=people, 0.9.2342.19200300.100.1.1=paul.tibbits@va.gov, cn=PAUL A. TIBBITS 116858
Reason: I am approving this document.
Date: 2015.12.29 15:00:32 -05'00'

Date:

Paul A. Tibbits, M.D.
DCIO Architecture, Strategy, and Design

REVISION HISTORY

Version	Date	Organization	Notes
1.0	9/25/14	ASD TS	Initial version approved staffed for formal approval by DCIO ASD
2.0	12/4/15	ASD TS	Update accounting for lessons learned and re-alignment to current template for Enterprise Design Patterns

REVISION HISTORY APPROVALS

Version	Date	Approver	Role
1.0	9/25/14	Joseph Brooks	APM Design Pattern Lead
2.0	12/29/15	Tim McGrail	ASD TS Design Pattern Final Reviewer

TABLE OF CONTENTS

FIGURES

TABLES

1 INTRODUCTION

Information Technology (IT) projects within the Department of Veterans Affairs (VA) developed or acquired solutions in a stove-piped fashion, resulting in redundant functionality including application performance management (APM). This duplication has resulted in an increased total cost of ownership (TCO) and APM complexity. APM tools were not deployed end-to-end, hindering VA's ability to evaluate application health consistently and to identify problems proactively.

This Enterprise Design Pattern provides guidance to projects in applying end-to-end APM capabilities provided by Enterprise Shared Services (ESS). This document guides projects in the use of standard APM capabilities provided by VA regional data centers. Additional coordination with OI&T Service Delivery and Engineering (SDE) Enterprise Operations (EO) on capacity and operations planning is required prior to deployment.

1.1 Business Need

VA business owners derive a number of benefits from APM capabilities within the agency's IT infrastructure investments. APM provides business owners the following "justification themes," according to *APM Best Practices* (see Appendix D):

Table 1 – APM Justification Themes for Business Owners

Justification Theme	Benefits
Availability vs. Performance Monitoring	• Enhances visibility into the behaviors of distributed systems and how to correlate and resolve various incidents • Reduces the time to first alert for a performance incident • Provides performance monitoring capability across transport protocols and Java/.NET platforms
Resolving Application Incidents and Outages	• Enables efficient tracking and resolving performance issues • Provides separate responses for availability and degradation incidents • Allows more effective use of the monitoring tool infrastructure through active capacity reporting and planning
Improving Application Software Quality	• Decreases overall time-to-market for new software systems • Confirms accuracy and utility of load testing during development • Improves production experience based on a consistent set of key performance indicators (KPI)
Pre-production Readiness	• Validates of low overhead of agent and transaction definitions

and Deployment	• Supports definition of the monitoring dashboards and reporting
Managing Service Level Agreements (SLA)	• Enhances relationships with business owners • Enables reliable transactions that are defined and focused • Provides accurate and rapid performance and capacity forecasting
Enhancing the Value of the Monitoring Tool Investment	• Decreases time-to-market schedule • Allows for optimal use of existing and proposed monitoring technology • Helps evolve skill sets and competencies of technical staff
Proactive Monitoring	• Achieves proactive management by catching performance problems during quality assurance (QA) and user acceptance testing (UAT) (DevOps) • Enhances triage of performance problems • Enhances overall software quality from the operations perspective
Trending and Analysis	• Increases use of the monitoring environment • Establishes comprehensive capacity management planning practices • Establishes more capable triage technical practices
Single-View of Service Performance (Dashboards)	• Gives real-time view of business service performance • Provides visibility into application component interactions and the end-user experience

1.2 Approach

End-to-end APM is currently available at regional data centers to monitor all operational systems and services, including ESS (Appendix E). All new applications are required to integrate the APM capabilities provided by the VA data centers. The current approach applies to solutions deployed at VA's data centers and will accommodate VA-approved external cloud service providers as the VA Cloud Strategy is deployed.

2 CURRENT CAPABILITIES AND LIMITATIONS

2.1 Traditional APM Approach

The following figure depicts VA's monitoring approach focused on specific domains. Example domains include message queues (MQ), operating systems, and Java Virtual Machines (JVM). This approach does not provide the full visibility into an entire business transaction using all domains, resulting in monitoring inefficiencies and a longer mean time to repair (MTTR).

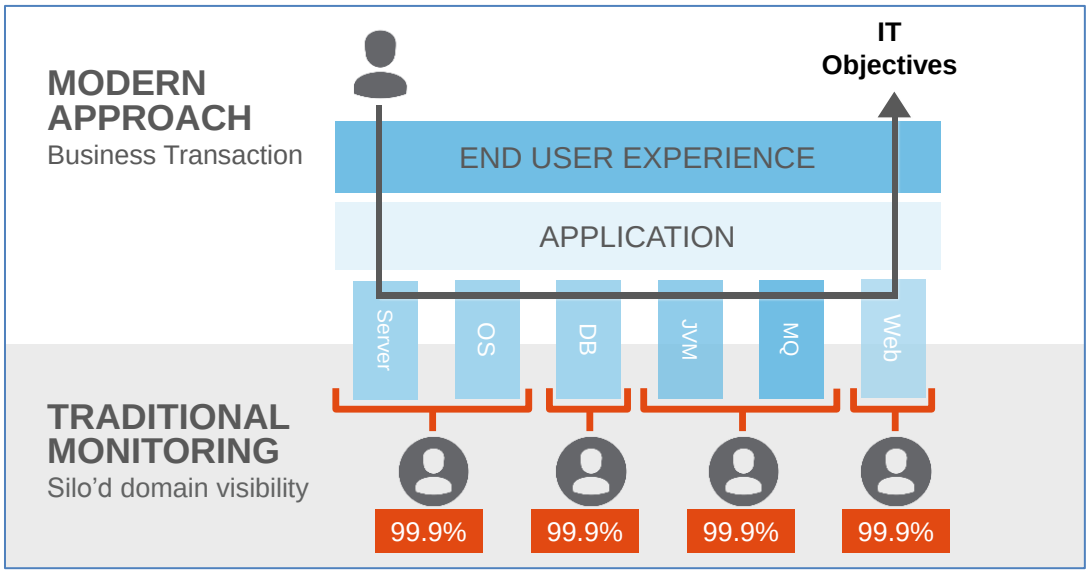

Figure 1 – Traditional Monitoring Approach Providing Visibility to Specific Domains of a Typical Business Transaction

Industry best practices recommend for end-to-end APM to leverage a top-down approach focusing on the complete application stack. VA's regional data centers (e.g., Austin Information Technology Center, Hines Information Technology Center, and Philadelphia Information Technology Center, etc.) offer the full spectrum of APM capabilities to monitor application health.

2.2 Current APM Deployments

The current end-to-end APM tools in VA data centers deliver a holistic view of all user transactions, helping IT stakeholders understand the health, availability, service impact, and end-user experience of critical applications. APM enables projects to diagnose and resolve problems proactively while optimizing the performance of mission critical services. APM supports prioritization of incidents based on service impact and quickly pinpoints problems across disparate technology silos.

2.3 Common Technical Capabilities

The following figure describes APM products deployed by SDE Enterprise Operations (EO):

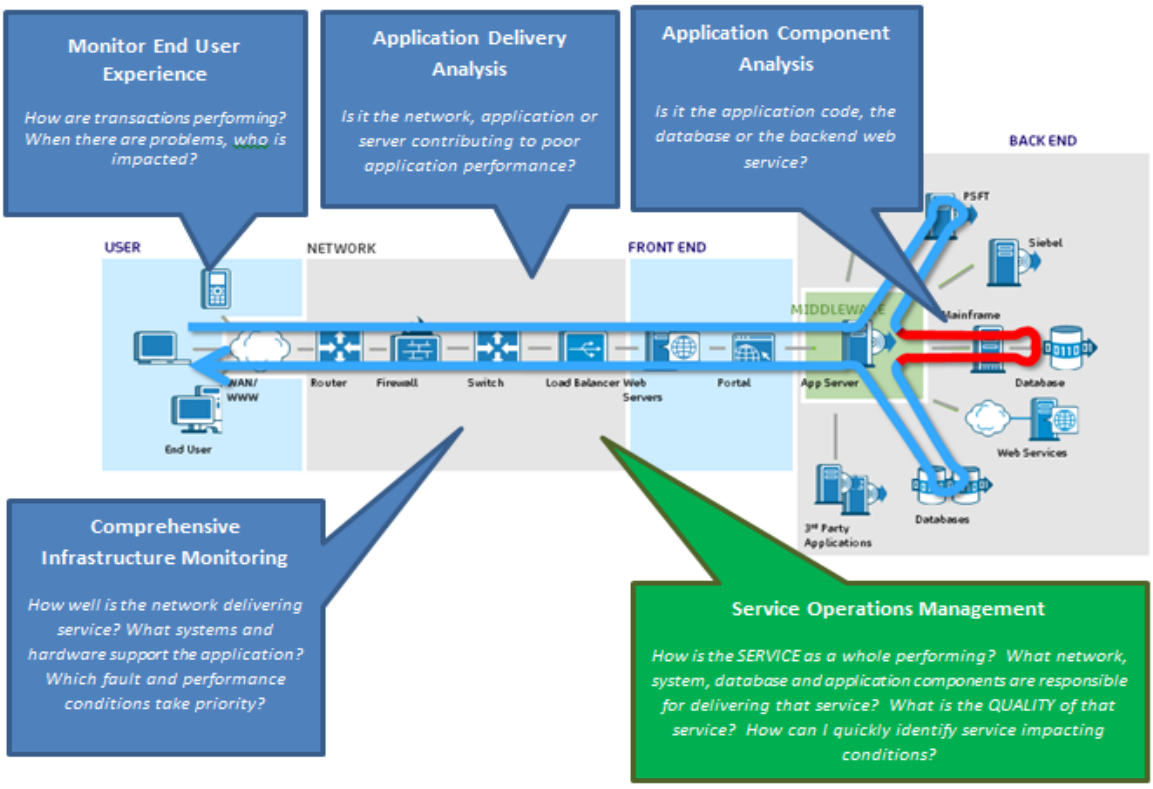

Figure 2 – APM Capabilities from the End User to Back-end Services and Databases

The following APM capabilities are available as enterprise infrastructure services:

End-user Experience Monitoring – Ensures consistent end-user experience and high service levels meeting business objectives by monitoring all end-user transactions (including web-based and non-web-based services) on a 24x7 basis with minimum overhead. APM measures end-user transaction performance to ensure applications are delivering against service level agreements (SLA) using application-specific Key Performance Indicators (KPI).

Application Behavior Analytics – Discovers anomalous application behavior automatically and proactively alerts IT operators of potential problems that could disrupt performance. The instrumentation tools provided by EO automatically mine the vast repository of rich data created by APM and, within hours of setup, can start identifying anomalous behavior in components, providing a view of potential issues between related components.

Smart Triage – Reduces downtime and optimizes the performance of services by proactively identifying, diagnosing and resolving performance problems before they impact end users. The EO-provided APM tools map all transactions to the dependent infrastructure in real-time for a single view of application health, business process flow, and the entire transaction path to

quickly triage issues, help eliminate problem resolution guesswork and accelerate mean time to repair.

Rapid Root-cause Diagnosis – Improves IT productivity and controls costs by diagnosing problems occurring within the application and infrastructure. End-user experience monitoring capabilities integrate with behavior analytics and deep-dive problem diagnosis features to understand performance issues in context, pinpoint failures, and accelerate problem resolution.

Business-centric Management – Assure high-value transactions receive the highest service levels by understanding problems in business context to identify critical transactions that may be at risk, prioritize problem resolution efforts, dispatch the right resources, and fix the problem affecting key functionality or end users. APM provides application health metrics understood by non-application experts and easily communicated to business users.

2.4 Current Limitations

The current APM capabilities focus on the VA regional data centers and emphasize web-based applications. VA will also require the monitoring of cloud services and mobile applications in the future. PD and SDE will require upfront coordination prior to Milestone 1 to conduct capacity planning and establishing KPIs prior to deployment. This necessitates a "DevOps" mindset involving close collaboration between development and operations staff, especially as VA shifts to a continuous integration and deployment paradigm.

3 FUTURE CAPABILITIES

3.1 APM for Mobile and Cloud Services

The future-state operational vision consists of end-to-end APM covering on-premise, cloud and hybrid environments, and support for DevOps practices for building, testing, and deploying applications. The concept diagram below (source: IBM) depicts the types of services to be provided by APM capabilities in various hosting environments.

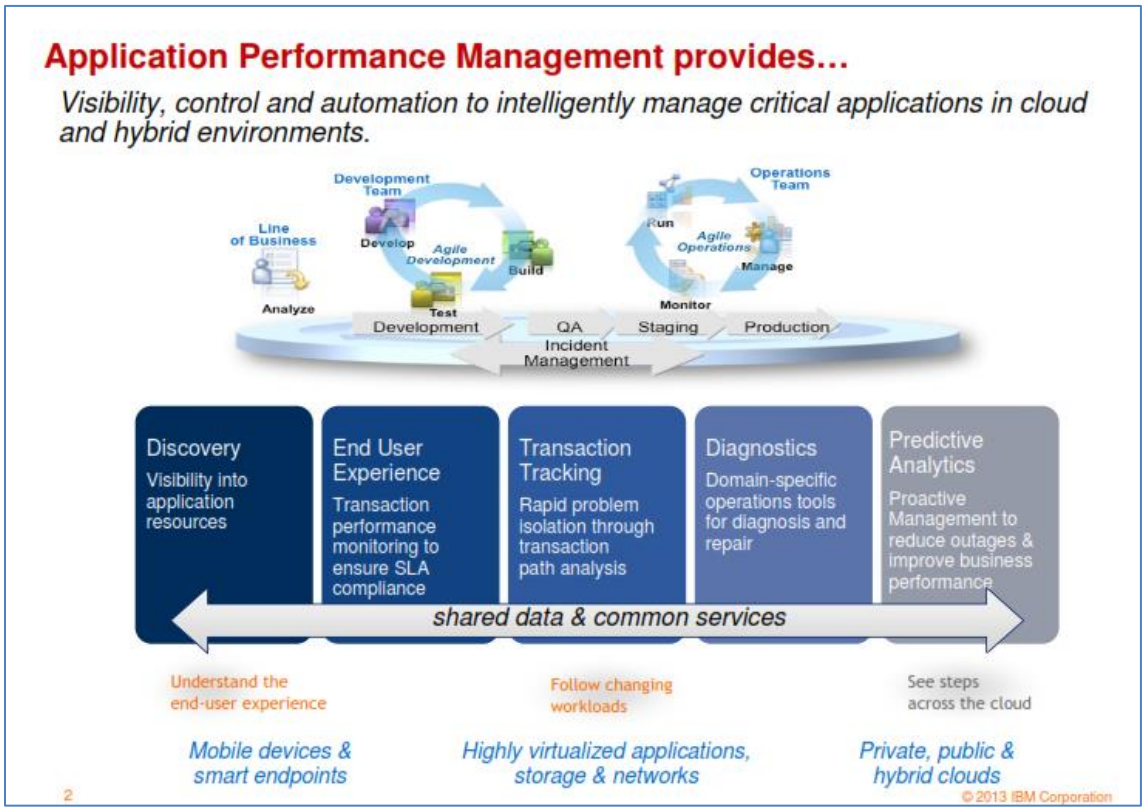

Figure 3 – End-to-end VA APM Capabilities and Transaction Visibility Conceptual Overview (Based on Materials Provided by IBM)

APM includes visibility into cloud environments, as discussed in detail in the Cloud Computing Enterprise Design Patterns. APM integrates predictive analytics capabilities to enhance proactive monitoring and trouble resolution. These capabilities support mobile applications and will integrate with enterprise mobile analytics capabilities, as explained in the Mobility Enterprise Design Patterns. Specifically, end-user experience monitoring supports mobile analytics, and APM achieves this through the following functions:

- Deploying a mobile performance agent on top of end-user monitoring capabilities, which may require adding a library and recompiling the code to perform APM for the application
- Agent piggybacking on other user's service calls through an application programming interface (API)
- Generating crash analytics to create a snapshot of device crash statistics

Deploying APM for cloud services will include health endpoint monitoring. This is typically the combination of two factors: the checks (if any) performed by the application or service in response to the request to the health verification endpoint, and analysis of the result by the tool or framework that is performing the health verification check. The response code indicates

the status of the application and, optionally, any components or services it uses. Additional best practices include:

- Checking storage or a database for availability and response time
- Checking other resources or services located within the application, or located elsewhere but used by the application

3.2 Proactive Planning for APM

Proactive planning for APM enables a DevOps paradigm that includes collaboration with both developers and operations staff. Projects will accomplish this through the following tasks:

- Establish KPIs with EO and conduct evaluations in a pre-production test environment
- Coordinate infrastructure support and conduct operations support planning prior to Milestone 0 to establish monitoring KPIs
- Develop a monitoring plan with known KPIs prior to Milestone 1

Projects require a common set of KPIs to monitor, control and track relative to indicating poor performance or equipment outage. Below is a representative list of KPIs:

- Message queue length
- Transaction or message throughput rate
- Transaction response time
- Database query response time
- Event management states
- Garbage collection behavior
- File I/O abnormalities
- Percentage of free storage space available
- Percentage of network retransmissions
- Network round trip time
- Network connection time
- SNMP connection failure (indicates complete equipment unavailability)
- Memory management patterns (e.g., JVM Heap)

All new applications require load testing in pre-production environments. APM must be available in these environments to measure expected performance and identify potential issues. Projects will work with EO to identify which of these KPIs needs to monitor during the testing phase to mitigate pre-production performance risks. Appendix D contains technical references on APM, and Appendix F includes pain points identified by EO to guide capacity and operations planning.

3.3 Alignment to TRM

All APM products used in regional data centers are commercial-off-the-shelf (COTS) and cataloged in the TRM, and new APM products require TRM approval. APM aligns to the following technology categories: Application Management, Monitoring, and Network Performance Optimization.

4 USE CASES

The business process below shows a prescriptive flow for how end-to-end APM works within the VA enterprise for three scenarios:

1. End user experience monitoring
2. Network performance monitoring
3. Back-end infrastructure monitoring

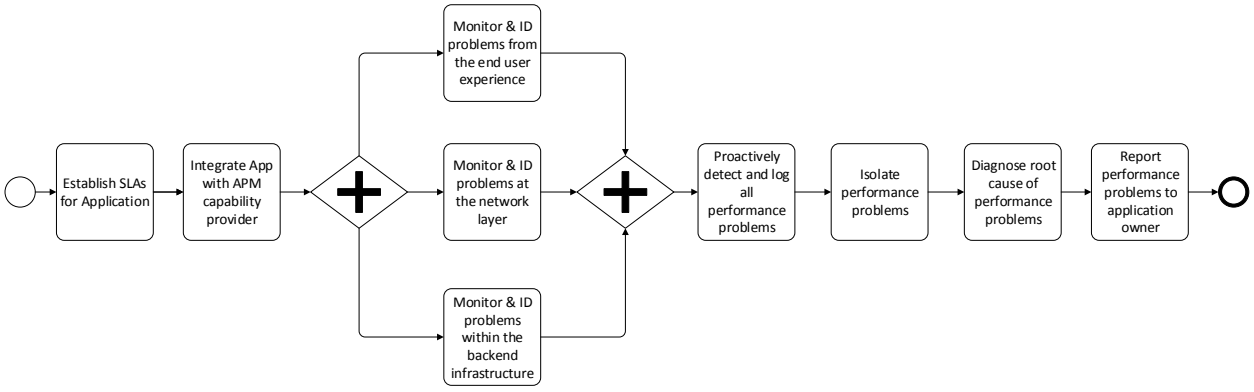

Figure 4 – Process for APM with User Experience, Network, and Infrastructure Monitoring

The basic flow of events between application owner and infrastructure actors (e.g., APM capability provider) is as follows:

1. Application owner establishes appropriate KPIs for the application in pre-production, including SLAs between service consumers and providers
2. Application owner deploys application into the VA IT infrastructure production environment and integrates with APM capability provider
3. APM capability provider monitors all business transactions traversing the entire VA IT infrastructure:
    a. Monitor and identify problems associated with the application layer (e.g., end-user experience)
    b. Monitor and identify problems associated with application delivery over the network

c. Monitor and identify problems associated with the backend infrastructure (e.g., application servers, web services, or databases)

4. APM capability provider proactively detects and logs all performance problems in each part of the infrastructure (Parts 3a-c)
5. APM capability provider isolates performance problems detected in Step 4
6. APM capability provider diagnoses root cause of performance problems in Parts 3a-c
7. APM capability provider reports performance problems to application owner

Appendix A. Document Scope

A.1 Scope

The purpose of this document is to provide guidance on using end-to-end APM capabilities that support Enterprise Shared Services (ESS). Specifically, this document guides projects to use the standard set of APM capabilities provided by VA data centers and to coordinate with OI&T Service Delivery and Engineering (SDE) Enterprise Operations (EO) early in the development lifecycle to ensure proactive performance monitoring. This applies to all new applications requiring integration into VA's enterprise IT. The guidance in this document applies to both COTS software (including open-source) acquisitions and applications developed by VA Product Development.

This document focuses on APM capabilities provided by VA's regional data centers. It provides high-level guidance for establishing application performance metrics (e.g., CPU usage, memory trends, input, and output operations). The following content is out of scope for this document but addressed in related Enterprise Design Patterns:

- Mobile analytics (covered by Mobility Design Patterns)
- Log management for auditing and compliance (covered by Privacy and Security Design Patterns)
- Network traffic monitoring (covered by Privacy and Security Design Patterns)
- Cloud service monitoring (covered by Cloud Computing Design Patterns)
- Vulnerability scanning and incident management processes (covered by IT Service Management Design Patterns)

A.2 Document Development and Maintenance

This Design Pattern was developed collaboratively with participation from VA's Office of Information and Technology (OI&T), Product Development (PD), Office of Information Security (OIS), Architecture, Strategy and Design (ASD), Service Delivery and Engineering (SDE), and industry. This document contains a revision history and revision approval log to track all

changes. Updates will be coordinated with the Government lead for this document, which will facilitate stakeholder coordination and subsequent re-approval depending on the significance of the change.

Appendix B. DEFINITIONS

Table 2 – Definitions

Key Term	Definition
Enterprise Shared Service (ESS)	A SOA service that is visible across the enterprise and accessed by users across the enterprise, subject to appropriate security and privacy restrictions.
Service	A mechanism to enable access to one or more capabilities, where the access is provided using a prescribed interface and is exercised consistent with constraints and policies as specified by the service description.
Service Oriented Architecture (SOA)	A paradigm for organizing and utilizing distributed capabilities that may be under the control of different ownership domains. It provides a uniform means to offer, discover, interact with, and use capabilities to produce desired effects consistent with measurable preconditions and expectations.
Service-Level Agreement (SLA)	An agreement between two parties regarding a particular service. They contain quantitative measurements that: • Represent a desired and mutually agreed state of a service • Provide additional boundaries of a service scope (in addition to the agreement itself) • Describe agreed and guaranteed minimal service performance
Key Performance Indicator (KPI)	Performance metrics targeting service provider organization tactical and strategic objectives. These metrics are used to measure: • Efficiency and effectiveness of a service • Service operation status Not all metrics automatically become KPIs. KPIs must be bound to the organization or service goals and must drive continuous improvement and efficiency.

Appendix C. ACRONYMS

Table 3 – Acronyms

Acronym	Description
APM	Application Performance Management
ASD	Architecture, Strategy and Design
BPEL	Business Process Execution Language
BAM	Business Activity Monitoring
BPM	Business Process Monitoring
CoE	Center of Excellence
COTS	Commercial Off-the-Shelf
eMI	Enterprise Messaging Infrastructure
EO	Enterprise Operations
ETSP	Enterprise Technology Strategic Plan
IPT	Integrated Project Team
JVM	Java Virtual Machine
KPI	Key Performance Indicator
MTTI	Mean Time to Identify
MTTR	Mean Time to Repair
PD	Product Development
SDE	Service Delivery and Engineering
SLA	Service Level Agreement
SNMP	Simple Network Management Protocol
TCO	Total Cost of Ownership

Appendix D. References, Standards, and Policies

Table 4 – References, Standards, and Policies

#	Issuing Agency	Applicable Reference/Standard	Purpose
1	VA OIS	VA 6500 Handbook	Directive from OIS that establishes an information security program in VA, which applies to all applications subject to APM.
2	VA ASD	VA Enterprise SOA Design Pattern	Provide a reference for the use of end-to-end application performance monitoring as part of the integration with SOA support infrastructure services. These documents standardize and constrain the solution architecture of all healthcare applications in the VA.
3	VA ASD	ESS Strategy Document and Directive	Provides the overarching strategy for developing, deploying, and managing ESS throughout the VA
4	VA ASD	VA Enterprise Technology Strategic Plan (ETSP)	Provides long-term IT vision for systems management capabilities that include APM
5	VA ASD	OIT Infrastructure Architecture	Provides a list of instrumentation/monitoring products to be used (based on business/technical requirements) for the monitoring, proactive detection, triage and diagnosis of performance problems in VA's Data Centers.

Additional technical references are as follows:

1. *APM Best Practices: Realizing Application Performance Management* by Michael J. Sydor, 2010, CA Press, ISBN-10: 1430231416
2. MSDN Cloud Design Patterns: Health Endpoint Monitoring https://msdn.microsoft.com/en-us/library/dn589789.aspx
3. MSDN Application Architecture Guide v2, Chapter 17 Cross-cutting Concerns: https://msdn.microsoft.com/en-us/library/ee658105.aspx

Appendix E. ESS INTEGRATION WITH APM

Alignment to ESS Architecture Construct

APM capabilities monitor the performance of Enterprise Shared Services (ESS) using the approved IT infrastructure hosted by VA's data centers. APM is a platform capability that constitutes the service-oriented architecture (SOA) support infrastructure "backplane," and it does not represent a specific business service, per the following ESS architecture layer construct:

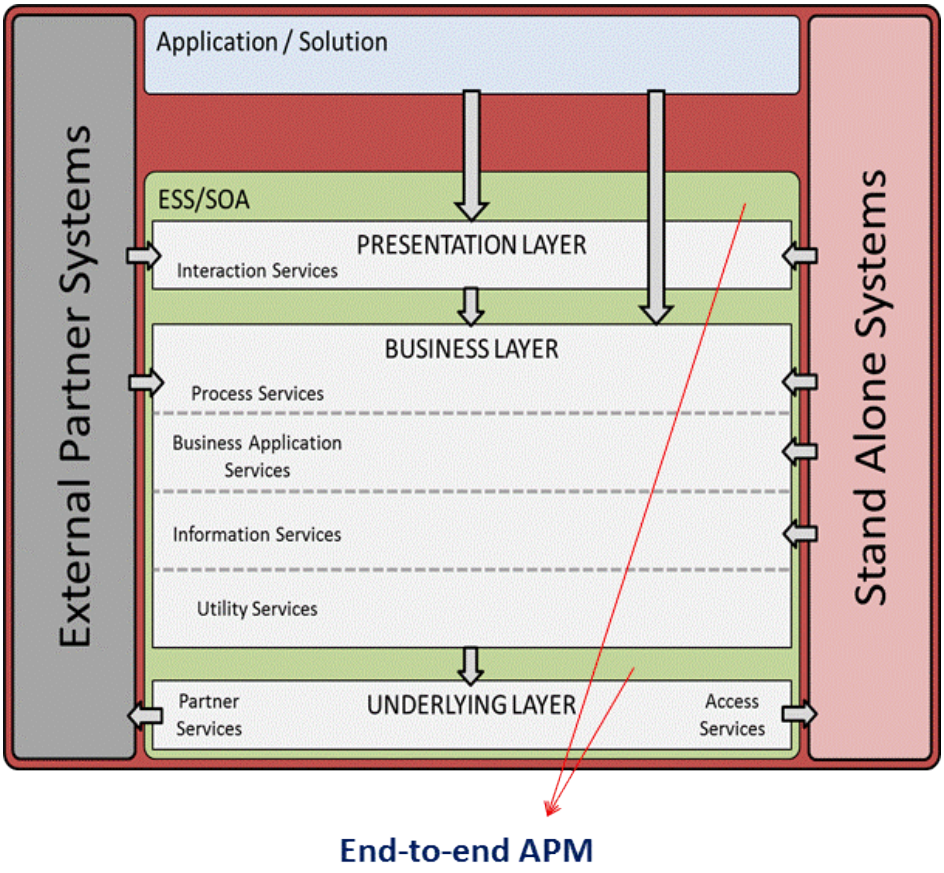

Figure 5 – APM within the ESS Architecture Construct

APM monitors both front-end and back-end performance associated with common utility services shared across numerous applications meeting diverse business requirements. Per the ESS Strategy document, new applications consuming ESS coordinate with the ESS Center of Excellence (CoE) and follow applicable architecture guidelines provided by the CoE to ensure proper integration with ESS. APM is a crosscutting concern and referenced in platform architecture models. These models support service-specific architecture models for ESS in

alignment with business capabilities and drivers. APM also integrates with existing Business Process Monitoring (BPM) and Business Activity Monitoring (BAM) functions in the Enterprise Messaging Infrastructure (eMI), monitoring workflows and service orchestrations. APM aligns to the Open Group SOA Reference Architecture (Quality of Service Layer) by enabling the following SOA management functionality:

- IT Systems Monitoring and Management: This category of capabilities provides monitoring and management of IT infrastructure and systems. This includes the ability to monitor and capture metrics and status of IT systems and infrastructure.
- Application and SOA Monitoring and Management: This category of capabilities provides monitoring and management of software services and applications. This includes the ability to capture metrics and to monitor and manage application and solution status.
- Business Activity Monitoring and Management: This category of capabilities provides monitoring and management of business activities and business processes. It provides the ability to analyze this event information, both in real-time/near real-time, as well as stored (warehoused) events, and to review and assess business activities in the form of event information and determine responses or issues alerts/notifications.

ESS Monitoring Approach

APM unifies end-user experience and network performance monitoring through a single appliance that provides a single source of truth on how network behavior affects the end-user experience, making it faster and easier to identify, diagnose, and resolve transaction problems caused by the network. APM provides application-aware infrastructure monitoring for any TCP-based application without desktop or server agents to deliver a consistent and common set of response-time metrics, mitigate risks from planned changes and unexpected events, and resolve problems faster. By providing the TCP-level view of applications running over the network and from tier-to-tier within the data center, it enables rapid troubleshooting of network and performance bottlenecks and provides insight into the duration, frequency, pervasiveness, and severity of problems. Automatic, intelligent baselines establish an understanding of normal performance, such that when deviations are detected, diagnostic data can be gathered that helps further enable faster resolution of performance problems. All of this information is accessible from a single, flexible APM dashboard for rapid troubleshooting and triage.

The eMI contains products that combine information from WebSphere Service Registry and Repository (WSRR), observations and business process execution language (BPEL) business process definitions. This allows users to:

- Reconcile services in WSRR with those monitored in target systems by the APM solution
- Topology views show relationships between service operations and BPEL business processes for impact analysis
- Forwards status information to WSRR to allow selection of services based on performance and other metrics

Appendix F. Identified Current Pain Points in Application Performance

SDE EO has identified key pain points in performance for new VA applications. These relate to the load and capacity testing capabilities, which enable measurement of these items prior to production. Mitigation of these pain points would remove 90% of the performance issues that operations encounter according to EO. In general, degraded performance on a new application is due to the application and not the infrastructure.

The following are pain points identified by SDE EO that can greatly reduce poor performance in production. While pain points on memory management and throughput refer to the Java stack, they also relate to .NET and other platforms.

- Java Heap issues are one of the primary problems of new systems in production because of inadequate load testing of the applications. EO has tools that observe Heap behavior and recommend the optimum Java Heap settings to try to minimize the pain of all projects using the "out of the box" heap settings that work in development but not in production.
- Stuck threads typically do not manifest in pre-production without rigorous load testing.
- Poorly written SQL queries are, by far, the single biggest application performance issue. Developers often do not recognize what database indexing is required to optimize the query return. EO also has many projects using Hibernate to generate queries but the projects do not understand how to optimize the queries produced in Hibernate.
- Production database size often causes application performance problems because testing is conducting against small test databases instead of the real-sized production-like database. Therefore, performance is great in pre-production but lackluster in production.