research

CHECKLIST REPORT 2019

A Step-by-Step Guide to Adopting a Hybrid Integration Platform

By David Loshin

Sponsored by:

actian™ activate your data

tdwi TRANSFORMING DATA WITH INTELLIGENCE™

JUNE 2019

TDWI CHECKLIST REPORT

A Step-by-Step Guide to Adopting a Hybrid Integration Platform

By David Loshin

TABLE OF CONTENTS

tdwi

Transforming Data With Intelligence™

555 S. Renton Village Place, Ste. 700
Renton, WA 98057-3295

T 425.277.9126
F 425.687.2842
E info@tdwi.org

tdwi.org

FOREWORD

There are three key trends that have motivated a rapid yet radical transition in how organizations manage their information environments, namely:

- **AS-A-SERVICE PLATFORMS**, such as software-as-a-service (SaaS) and platform-as-a-service (PaaS) environments, that allow organizations to transition core business and data management functionality to external service providers
- **CLOUD COMPUTING**, which provides a broad array of low-cost hosted computing, storage, and application services
- **MODERNIZATION**, the desire to consider ways to re-engineer existing business applications to better meet both current and anticipated business requirements

Together, these trends have triggered waves of data and process migration out of the traditional on-premises data center toward a variety of hosted and cloud-based environments. Yet this migration cannot take place in one fell swoop.

Modernization must be planned and incrementally executed, resulting in the growing complexity of what could be called a hybrid environment. These hybrid environments conceptually incorporate data and system services across a variety of platforms including on-premises, a variety of hosted cloud environments, as well as a growing pool of operational systems or Internet of Things (IoT) devices that stream data.

A *hybrid integration platform* (HIP) is a suite of capabilities that address the challenges of integration across a variety of platforms, sources, and applications.

Gartner specifies that a true HIP should support:

- "**PERSONAS** (constituents): Integration specialists, ad hoc integrators, citizen integrators and digital integrators
- **INTEGRATION DOMAINS**: Application, data, B2B and process
- **ENDPOINTS**: On-premises devices, the cloud, mobile devices and IoT devices
- **DEPLOYMENT MODELS**: Cloud (potentially across multiple environments), on-premises, hybrid (cloud and on-premises) and embedded in IoT devices [as well as XaaS platforms and applications]"[1]

Although you need the right tools and technology for a HIP, there are some important concepts to focus on if you plan to embrace this kind of environment. This checklist report highlights some of these concepts and provides a step-by-step guide for adopting a hybrid integration platform.

[1] "Use a Hybrid Integration Approach to Empower Digital Transformation," Rob van der Meulen, accessed 04/23/2019 via https://www.gartner.com/smarterwithgartner/use-a-hybrid-integration-approach-to-empower-digital-transformation/

1 UNDERSTAND THE DIFFERENT DEPLOYMENT MODELS

Evolving enterprise systems span a variety of system types. For example, a manufacturer may need interoperability among three instances of different deployment models: an on-premises enterprise resource planning (ERP) system, a horde of factory floor sensors and devices continuously streaming measurements and operating data, and a cloud-based predictive analytics engine monitoring for events indicating the need for part replacement.

Hybrid integration encompasses the ability to seamlessly interoperate among and integrate a variety of deployment models within the extended information enterprise. One key aspect of adopting a hybrid integration platform is acknowledging and understanding the need for data and functional integration across these different deployment models, including:

- Conventional on-premises data management environments, including on-premises file systems, relational databases, data warehouses, specialized cross-functional systems (such as ERP, CRM, or HCM systems), and legacy data processing systems
- Conventional technologies deployed to the cloud, such as hosted relational databases, data lakes, or data warehouses
- Emerging data management technologies suited to cloud platforms, such as cloud-hosted NoSQL, in-memory/HTAP, or highly scalable data management technologies
- Hybrid cloud environments, such as virtual data lakes whose contents are distributed among multiple cloud hosts
- Embedded services, such as containerized applications and services or edge computing systems
- Systems executing on mobile or hand-held devices
- Streaming information environments, such as IoT/machine-generated data or data produced by mobile applications

As more API- and services-based applications are developed (especially ones that bridge on-premises, cloud, and mobile computing and IoT devices), there will be an increasing need to assess the application and data integration landscape and understand the need for integration across deployment environments. This implies the need for technologies that not only identify the deployment models that need to interoperate and then classify and catalog the different sources and platforms but also modulate and facilitate both data interchange and transformation to enable digital transformation.

2 UNDERSTAND INTEGRATION PATTERNS

Understanding the deployment models is necessary for implementing a HIP, but it is certainly not sufficient. You must also be aware of the methods for ensuring interoperability across a variety of operational environments and scenarios/use cases characterized as integration patterns, including (but not limited to):

- **APPLICATION INTEGRATION (A2A).** This pattern centers on communication, exchanging data, and sharing operational state among two or more different applications to allow them to interoperate.
- **DATA INTEGRATION.** This pattern focuses on blending and merging data from two or more sources originating across a wide variety of data management environments.
- **PROCESS INTEGRATION.** This integration pattern bridges communication and interoperability across a collection of processes.
- **OPERATIONAL B2B INTEGRATION.** The intent of this pattern is coordination of business processes whose successful execution requires actions performed by different business entities. It could mean integration of business processes with different business units within the enterprise or among different entities outside the enterprise.
- **"AS-A-SERVICE" INTEGRATION.** This pattern is used when organizations engage service providers to perform some aspect of their business operations, such as a hosted CRM, sales automation, or marketing automation SaaS application.
- **E-COMMERCE INTEGRATION.** In this integration pattern, the transactions and data streams originating from your company's e-commerce platform are synchronized among the front-end storefront, transaction processing on the back end, and order fulfillment partner systems.

The objective of these integration patterns is to ensure that information and operational directives are conveyed in a timely manner across the different components of an enterprise system, regardless of the deployment model. In essence, any of these integration patterns may be aligned across multiple deployment models. An example is operational e-commerce integration for mobile apps on a smartphone that push data to a centralized cloud-based data store, with summarized data streamed to an on-premises data warehouse hosting an executive dashboard.

Understand your organization's integration patterns. Survey your operational environment, identify the different integration scenarios, and determine the different end points. This intelligence task will help you determine where a hybrid integration platform can simplify business process interoperability across hybrid environments.

3 EMPOWER ANYONE TO INTEGRATE—KNOW THE NEEDS OF DIFFERENT PERSONAS

Adopting a hybrid integration platform requires more than just understanding the hybrid cloud models and associated integration patterns. You must also understand the different types of users or personas that are expected to either be developing new integration processes or configuring existing integration templates. These personas represent a range of technical skills and awareness of integration processes.

- **CITIZEN ANALYSTS/INTEGRATORS** have limited technical skills, high data awareness, and acute familiarity with advanced analytics within their knowledge domains
- **BUSINESS ANALYSTS** have some degree of technical skills coupled with moderate knowledge of data within their business contexts
- **DATA INTEGRATION PROFESSIONALS AND SPECIALISTS** are skilled in the use of integration tools and technology and have experience applying those skills across multiple applications and business functions
- **AD HOC INTEGRATORS OR APPLICATION DEVELOPERS** are tasked with incorporating integration patterns within application solution stacks
- **EMBEDDED SERVICES DEVELOPERS** are able to embed integration via defined development platforms, SDKs, and automated services

Each of these personas has distinct integration needs, ranging from configuration-based analytical integrations to tightly controlled low-latency mission-critical programmatic integration patterns.

Sitting at one end of the spectrum, citizen analysts and business analysts are likely to rely on no-code, self-service templates and wizards to help integrate information from across the hybrid environment to generate reports and visualizations that drive business decisions.

Data and application integration specialists (who are creating and modifying integration procedures and managing trading partner integration) and application developers (assembling development environments supporting different personas) sit in the middle of the spectrum—they may use intuitive integrated development environments (IDEs) for designing integrations and mapping purposes.

At the other end of the spectrum are the embedded system developers looking to deploy embedded integration patterns directly into interoperable applications (XaaS) whose operations cross the different components of the hybrid environment.

Empowering all of these personas means synthesizing the hybrid integration techniques to support their needs across a variety of application paradigms. Knowing and understanding the needs of the different personas is the first step in selecting the right technologies for your HIP.

4 INTEGRATE ANYTHING—PROVIDE BROAD UNIVERSAL CONNECTIVITY

One of the biggest integration challenges is the broadening variety of data sources and systems that compose the emerging hybrid enterprise. Some of these data sources are ones that most people are accustomed to from their on-premises environments, including:

- Conventional relational database management systems (RDBMSs) supporting transaction or operational processing
- Data warehouses and data marts supporting reporting and analytics use cases
- Data sets that have been extracted from production databases
- Flat files (such as those produced from desktop applications such as Excel)

In a hierarchical hybrid organization that touches both on-premises and multiple cloud environments, there are additional dimensions that need to be addressed for connectivity. Hybrid integration must embrace a wider variety of sources (deployed both on premises and in the cloud), including:

- Data produced in semistructured data formats such as XML and JSON
- Comprehensive enterprise applications (such as ERP, HCM, CRM systems) that provide interfaces for data interchange
- Data provisioned via SaaS and PaaS systems such as sales force automation or marketing automation services
- API- and data services-based applications, which are increasing in popularity among modern application developers
- Internet of Things (IoT) devices continually broadcasting automatically generated data streams

Only considering fundamental sources for data integration unnecessarily narrows the view of integration. We must also examine the contexts for connectivity and integration at the business level. *Process integration* provides a dimension that is layered on top of data integration and focuses on how compartmentalized workflows work together for process execution. At a higher level, there are a variety of logical vertical business integration scenarios that not only cross platform boundaries within an organization but also establish interoperation across organizational borders.

Examples include coordination for retail transactions (between e-commerce sites, fulfillment companies, and delivery providers), healthcare (coordinating among providers and payers), supply chain (involving inventory, warehousing, and shipping providers), and many others.

Provide broad universal connectivity. Provide network and connection capabilities to a wide range of data sources, using different formats, originating from different locations, and interacting using different exchange protocols.

5 INTEGRATE ANYWHERE—SUPPORT INTEGRATION ACROSS A HYBRID ENTERPRISE ENVIRONMENT

This checklist has been discussing a hybrid integration platform. However, when we use the term "hybrid" to describe an environment, what are the practical implications? As can be surmised, there is significant complexity in integration patterns that cross the barrier between on-premises and cloud-based platforms. The complexity is further compounded when considering the additional characteristics of what might be called the "extended information enterprise," which is composed of a network of operational platforms, possibly including:

- On-premises environments consisting of traditional "behind the firewall" systems
- Virtual private cloud systems in which the organization brings its own applications (BYOL) or product licenses and deploys them on virtual cloud resources
- Cloud SaaS/multicloud including services platforms whose capabilities and data must be integrated into the enterprise (such as Salesforce.com, Marketo, and other SaaS automation systems)
- Hosted cloud services, such as storage, compute, database, and even machine learning and artificial intelligence services that can be configured using Amazon Web Services, Google Cloud Platform, or other cloud hosts
- Embedded systems including other types of "as-a-service" (XaaS) systems, machine-generated data sources, edge computing systems, and containerized systems
- iPaaS or integration platforms as-a-service that are frameworks to facilitate integration among on-premises and multicloud environments

Cross-platform integration grows in difficulty as the number of sources and end points grows. If you want to enable "anywhere integration," look for technologies that have multiple deployment options coordinated through an integration manager that will support integration tasks that cross these different contexts and frameworks.

6 INTEGRATE ANYTIME—SIMPLIFY INTEGRATION WITH THE RIGHT ORCHESTRATION

As the number of parties with an interest in integration increases, there is a risk that the complexity of the environment will overwhelm the ability to both manage and maintain the growing collection of integration pipelines. In a traditional on-premises computing environment, there will be numerous application end users, many of whom will have a proactive voice in dictating the data preparation, workflow configuration, and pipeline integration steps for producing a report or analysis.

Recognize that this growing pool of integration workflows and the need to support all personas with a variety of integration patterns and sources across a range of deployment options can become unmanageable in a hybrid environment.

Simplify the management of these hybrid integration pipelines through *orchestration*, which is the ability to automate the creation of, establish metrics for, manage, schedule, and monitor the collected set of integration workflows at any time. This pool of workflows includes:

- Scheduled periodic batch integration workflows, such as conventional extraction, transformation, and loading of data into an on-premises data warehouse
- On-demand batch integration workflows, such as analyst interactions that pull data from a range of sources through configured workflow pipelines to create on-demand reports
- Event-driven integration workflows, such as data ingestion and transformations that are triggered automatically by cloud-based serverless computing frameworks as data sets are pushed to cloud-hosted storage environments
- API-driven integration workflows, such as trading partner business operations that invoke hosted data services
- Embedded process integration workflows such as those incorporated into edge analytics systems that ingest IoT data streams, apply embedded analytical models, and forward data streams to a centralized cloud analytics environment

Automate integration pipeline orchestration. Seek products and technologies for hybrid integration that are capable of supporting "anytime" integration orchestration for all types of workflows across the different computing platforms.

7 FUTURE-PROOF YOUR INTEGRATIONS—ENSURE PORTABILITY ACROSS THE ON-PREMISES/CLOUD BOUNDARY

The need for hybrid integration is a byproduct of the rapid evolution of technologies, as the breadth of platforms and systems widens. However, be aware that the rate of technology evolution is not slowing. Rather, the accelerating pace of innovation will continue to rapidly eclipse development efforts.

This means that applications developed for a particular platform or bound to a particular technology are at risk of becoming obsolete much faster than in the past. This issue becomes particularly acute as virtual environments span a variety of deployment domains.

Future-proof your integration workflows by making their integration applications portable. As your business grows and changes, make sure that developers can effortlessly move their integrations from on premises to cloud or vice versa, keeping their integrations intact. Leverage a combination of approaches to portability that support both on-premises and cloud environments, such as:

- Reusable code templates that are mapped to documented integration patterns
- Standardized representations that simplify development of data access and integration no matter where the data is sourced
- Serverless computing models that can automate triggered code execution on cloud platforms
- Containerization of applications and workflows so they can be portable and deployed in a variety of environments
- Platform virtualization that provides a simplified mechanism for implementing and deploying containerized applications in a virtual environment

The common thread across all of these suggestions is *portability*—the ability to develop and encapsulate reusable components in ways that can be easily deployed in many environments. Emerging approaches to virtualization and containerization provide an insurance policy against technical obsolescence.

Look for a HIP product that fundamentally embraces portability to extend the shelf life of developed integration workflows.

AFTERWORD

Application modernization is the pathway to embracing an extended information enterprise that spans the different components of a hybrid environment. Although cloud migration may be the holy grail for many organizations, we are bound to operate within this hybrid environment of on-premises and cloud platforms for the near and medium terms. That means that hybrid integration is going to be a critical competency for successful modernization.

That being said, simplify your approach for integration by adopting a hybrid integration platform. Look for technologies and products that combine these capabilities and features:

- Easy to use
- Rapid time to value
- Extensible
- Portable and supports encapsulation, APIs, and containerization
- Simplifies the cross-platform migration process
- Supports a multicloud hybrid environment
- Supports a wide range of integration personas
- Provides universal connectivity to a broad range of data and application sources
- Supports multiple deployment options, including on-premises, cloud-based, and embedded processes
- Support both batch and real-time integration

By considering the suggestions in this checklist, your organization will be prepared to simplify the orchestration of integration across a variety of implementation environments.

ABOUT OUR SPONSOR

Actian, the hybrid data management, analytics, and integration company, delivers data as a competitive advantage to thousands of organizations worldwide. Through the deployment of innovative, enterprise-class, hybrid data products, fully managed cloud services, mobile and IoT edge data management, and industry solutions, Actian ensures that business-critical systems can analyze, transact, and connect at their very best—both on premises and in the cloud. Thousands of forward-thinking organizations around the globe trust Actian to help them solve their toughest data challenges to transform how they run their businesses, today and in the future.

For more, visit http://www.actian.com.

ABOUT TDWI RESEARCH

TDWI Research provides research and advice for BI professionals worldwide. TDWI Research focuses exclusively on analytics and data management issues and teams up with industry practitioners to deliver both broad and deep understanding of the business and technical issues surrounding the deployment of business intelligence and data management solutions. TDWI Research offers reports, commentary, and inquiry services via a worldwide membership program and provides custom research, benchmarking, and strategic planning services to user and vendor organizations.

ABOUT THE AUTHOR

David Loshin, president of Knowledge Integrity, Inc, (www.knowledge-integrity.com), is a recognized thought leader, TDWI affiliate analyst, and expert consultant in the areas of data management and business intelligence. David is a prolific author on topics related to business intelligence best practices. He has written numerous books and papers on data management, including *Big Data Analytics: From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL, and Graph* and *The Practitioner's Guide to Data Quality Improvement*, with additional content provided at www.dataqualitybook.com. David is a frequent invited speaker at conferences, online seminars, and sponsored websites and channels including TechTarget and The Bloor Group. His best-selling book, *Master Data Management*, has been endorsed by many data management industry leaders.

David can be reached at loshin@knowledge-integrity.com.

ABOUT TDWI CHECKLIST REPORTS

TDWI Checklist Reports provide an overview of success factors for a specific project in business intelligence, data warehousing, analytics, or a related data management discipline. Companies may use this overview to get organized before beginning a project or to identify goals and areas of improvement for current projects.