Version 27.0: Spring '13

Integration Patterns and Practices

Last updated: March 15, 2013

Table of Contents

INTRODUCTION

Chapter 1

Integration Patterns Overview

When you implement Salesforce, you frequently need to integrate it with other applications. Although each integration scenario is unique, there are common requirements and issues that developers must resolve.

This document describes strategies (in the form of patterns) for these common integration scenarios. Each pattern describes the design and approach for a particular scenario rather than a specific implementation. In this document you'll find:

- A number of patterns that address key "archetype" integration scenarios
- A selection matrix to help you determine which pattern best fits your scenario
- Integration tips and best practices

Purpose and Scope

This document is for designers and architects who need to integrate the Force.com platform with their on-premises applications. This content is a distillation of many successful implementations by salesforce.com architects and partners.

Read the pattern summary and selection matrix if you're considering large-scale adoption of Salesforce-based applications (or the Force.com or Database.com tool sets) so that you can become familiar with the integration capabilities and options available. Architects and developers should consider these pattern details and best practices during the design and implementation phase of a Salesforce integration project.

If implemented properly, these patterns enable you to get to production as fast as possible and have the most stable, scalable, and maintenance-free set of applications possible. Salesforce.com's own consulting architects use these patterns as reference points during architectural reviews and are actively engaged in maintaining and improving them.

As with all patterns, they cover the majority of integration scenarios, but not all. While Salesforce allows for user interface (UI) integration—mashups, for example—such integration is outside the scope of this document. If you feel that your requirements are outside the bounds of what these patterns describe, please speak with your salesforce.com representative.

Pattern Template

Each integration pattern follows a consistent structure. This provides consistency in the information provided in each pattern and also makes it easier to compare patterns.

Name

The pattern identifier that also indicates the type of integration contained in the pattern.

Context

The overall integration scenario that the pattern addresses. Context provides information about what users are trying to accomplish and how the application will behave to support the requirements.

Problem

The scenario or problem (expressed as a question) that the pattern is designed to solve. When reviewing the patterns, read this section to quickly understand if the pattern is appropriate for your integration scenario.

Forces

The constraints and circumstances that make the stated scenario difficult to solve.

Solution

The recommended way to solve the integration scenario.

Sketch

A UML sequence diagram that shows you how the solution addresses the scenario.

Results

Explains the details of how to apply the solution to your integration scenario and how it resolves the forces associated with that scenario. This section also contains new challenges that can arise as a result of applying the pattern.

Sidebars

Additional sections related to the pattern that contain key technical issues, variations of the pattern, pattern-specific concerns, and so on.

Example

An end-to-end scenario that describes how the design pattern is used in a real-world Salesforce scenario. The example explains the integration goals and how to implement the pattern to achieve those goals.

Pattern Summary

The following table lists the integration patterns contained in this document.

List of Patterns

Pattern	Scenario
UI Update Based on Data Changes	The Salesforce user interface must be automatically updated as a result of changes to Salesforce data.

Pattern	Scenario
Remote Process Invocation—Request and Reply	Salesforce invokes a process on a remote system, waits for completion of that process, and then tracks state based on the response from the remote system.
Remote Process Invocation—Fire and Forget	Salesforce invokes a process in a remote system but doesn't wait for completion of the process. Instead, the remote process receives and acknowledges the request and then hands off control back to Salesforce.
Batch Data Synchronization	Data stored in Force.com should be created or refreshed to reflect updates from an external system, and when changes from Force.com should be sent to an external system. Updates in either direction are done in a batch manner.
Remote Call-In	Data stored in Force.com is created, retrieved, updated, or deleted by a remote system.

Pattern Approach

The integration patterns in this document are classified into two categories:

- Data Integration—These patterns address the requirement to synchronize data that resides in two or more systems so that both systems always contain timely and meaningful data. Data integration is often the simplest type of integration to implement, but requires proper information management techniques to make the solution sustainable and cost-effective. Such techniques often include aspects of Master Data Management (MDM), data governance, mastering, de-duplication, data flow design, and others.
- Process Integration—The patterns in this category address the need for one business process to leverage the work steps in another business process. When you implement a solution for this type of integration, the triggering application crosses the process boundary of the target application. Usually, these patterns also include both orchestration (where the triggering application is the central "controller") and choreography (where applications are multi-participants and there is no central "controller"). These types of integrations can often require complex design, testing, and exception handling requirements. Also, such composite applications are typically more demanding on the underlying systems because they often support long-running transactions, and the ability to report on and/or manage process state.

Choosing the best integration strategy for your system is not trivial. There are many aspects to take into consideration and many tools that can be used, with some tools being more appropriate than others for certain tasks. Each pattern addresses specific critical areas including the capabilities of each of the systems, volume of data, failure handling, and transactionality.

Pattern Selection Guide

The selection matrix table lists the patterns, along with key aspects, to help you determine the pattern that best fits your integration requirements. The patterns in the matrix are categorized using the following dimensions:

Aspect	Description
Source/Target	Specifies the requestor of the integration transaction along with the target(s) that provide the information. Note that the technical capabilities of the source and target systems, coupled with the type and timing of the integration, may require an additional middleware or integration solution. See the details of each pattern for additional details.

Aspect	Description
Type	Specifies the style of integration: Process or Data. • Process—Process-based integrations can be defined as "the integration of the functional flow of processing between the applications."[1] These integrations typically involve a higher level of abstraction and complexity, especially with regard to transactionality and rollback. • Data—Data integrations can be defined as "the integration of the information used by applications."[2] These integrations can range from a simple table insert or upsert, to complex data updates requiring referential integrity and complex translations.
Timing	Specifies the blocking (or non-blocking) nature of the integration. • Synchronous—Blocking or "near-real-time" requests can be defined as "request/response operation, and the result of the process is returned to the caller immediately via this operation."[3] • Asynchronous—Non-blocking, queue, or message-based requests are "invoked by a one-way operation and the results (and any faults) are returned by invoking other one-way operations."[4] The caller therefore makes the request and continues, without waiting for a response.

Note: An integration can require an external middleware or integration solution (for example, Enterprise Service Bus) depending on which aspects apply to your integration scenario.

Pattern Selection Matrix

The following table lists the patterns, along with key aspects, to help you determine the pattern that best fits your integration requirements.

Source/Target	Type		Timing		Key Pattern(s) to Consider
	Process Integration	Data Integration	Synchronous	Asynchronous	
Salesforce –> System (s)	X		X		Remote Process Invocation—Request and Reply
				X	Remote Process Invocation—Fire and Forget
		X	X		Remote Process Invocation—Request and Reply
				X	UI Update Based on Data Changes

[1] "Back-end Integration," IBM Corporation, last accessed May 18, 2012, *http://www.ibm.com/developerworks/patterns/application/*.
[2] Ibid.
[3] "Synchronous and asynchronous processes," IBM Corporation, last accessed May 18, 2012, *http://publib.boulder.ibm.com/infocenter/adiehelp/v5r1m1/index.jsp?topic=%2Fcom.ibm.etools.ctc.flow.doc%2Fconcepts%2Fcsynchf.html*.
[4] Ibid.

Source/Target	Type		Timing		Key Pattern(s) to Consider
	Process Integration	Data Integration	Synchronous	Asynchronous	
System –> Salesforce	X		X		Remote Call-In
				X	Remote Call-In
		X	X		Remote Call-In
				X	Batch Data Synchronization

DESIGN PATTERN CATALOG

Chapter 2

UI Update Based on Data Changes

Context

You use Salesforce to manage customer cases. A customer service rep is on the phone with a customer working on a case. The customer makes a payment, and the customer service rep needs to see a real-time update in Salesforce from the payment processing application, indicating that the customer has successfully paid the order's outstanding amount.

Problem

When an event occurs in Salesforce, how can the user be notified in the Salesforce user interface without having to refresh their screen and potentially losing work?

Forces

There are various forces to consider when applying solutions based on this pattern:

- Does the data being acted on need to be stored in Salesforce?
- Can a custom user interface layer be built for viewing this data?
- Will the user have access for invoking the custom user interface?

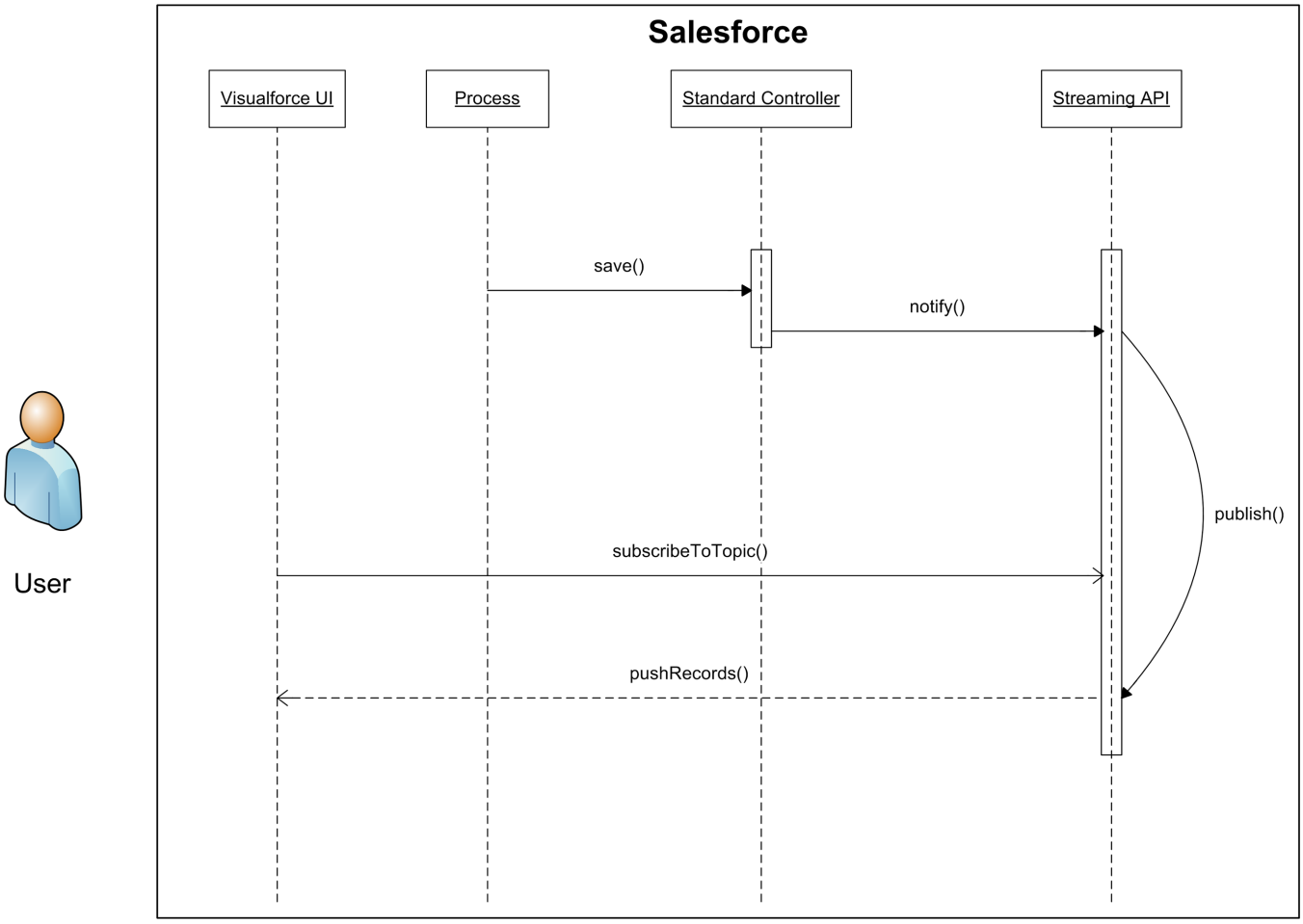
Solution

The recommended solution to this integration problem is to use the Salesforce Streaming API. This solution is comprised of the following components:

- A PushTopic with a query definition that allows you to:
  - ◊ Specify what events trigger an update
  - ◊ Select what data to include in the notification
- A JavaScript-based implementation of the *Bayeux* protocol (currently *CometD*) that can be used by the user interface.
- A Visualforce page.
- A JavaScript library included as a static resource.

Sketch

The following diagram illustrates how Streaming API can be implemented to stream notifications to the Salesforce user interface. These notifications are triggered by record changes in Salesforce.



Results

Benefits

The application of the solution related to this pattern has the following benefits:

- Eliminates the need for writing custom polling mechanisms
- Eliminates the need for a user-initiated feedback loop

Unsupported Requirements

The solution has the following limitations:

- Delivery of notifications isn't guaranteed
- Order of notifications isn't guaranteed
- Notifications aren't generated from record changes made by Bulk API.

Security

Standard Salesforce organization-level security is adhered to. It's recommended you use the HTTPS protocol to connect to Streaming API.

Sidebars

The optimal solution involves creating a custom user interface in Salesforce. It's imperative that you account for an appropriate user interface container that can be used for rendering the custom user interface. Supported browsers are listed in the Streaming API *documentation*.

Example

A telecommunications company uses Salesforce to manage customer cases. The customer service managers want to be notified automatically when a case is successfully closed by one of their customer service reps.

Implementing the solution prescribed by this pattern, the customer should:

- Create a PushTopic that sends a notification when a case is saved with a Status of "Closed" and Resolution of "Successful."
- Create a custom user interface available to customer service managers. This user interface subscribes to the PushTopic channel.
- Implement logic in the custom user interface that shows alerts generated by that manager's customer service reps.

Chapter 3

Remote Process Invocation—Request and Reply

Context

You use Salesforce to track leads, manage your pipeline, create opportunities, and capture order details that convert leads to customers. However, Salesforce isn't the system that contains or processes orders. After the order details are captured in Salesforce, an order needs to be created in the remote system, and the remote system manages the order through to its conclusion.

When you implement this pattern, Salesforce makes a call to the remote system to create the order and then waits for successful completion of that call. To signify successful completion of the call, the remote system synchronously replies with the order status and order number. As part of the same transaction, Salesforce updates the order number and status internally. The order number is used as a foreign key for any subsequent updates to the remote system.

Problem

When an event occurs in Salesforce, how do you initiate a process in a remote system, pass the required information to that process, receive a response from the remote system, and then use that response data to make updates within Salesforce?

Forces

There are various forces to consider when applying solutions based on this pattern:

- Does the call to the remote system require Salesforce to wait for a response before continuing processing? In other words, is the call to the remote system a synchronous request-reply or an asynchronous request?
- If the call to the remote system is synchronous, does the response need to be processed by Salesforce as part of the same transaction as the initial call?
- Is the message size relatively small or large?
- Is the integration based on the occurrence of a specific event, such as a button click in the Salesforce user interface or DML-based events?

Solution

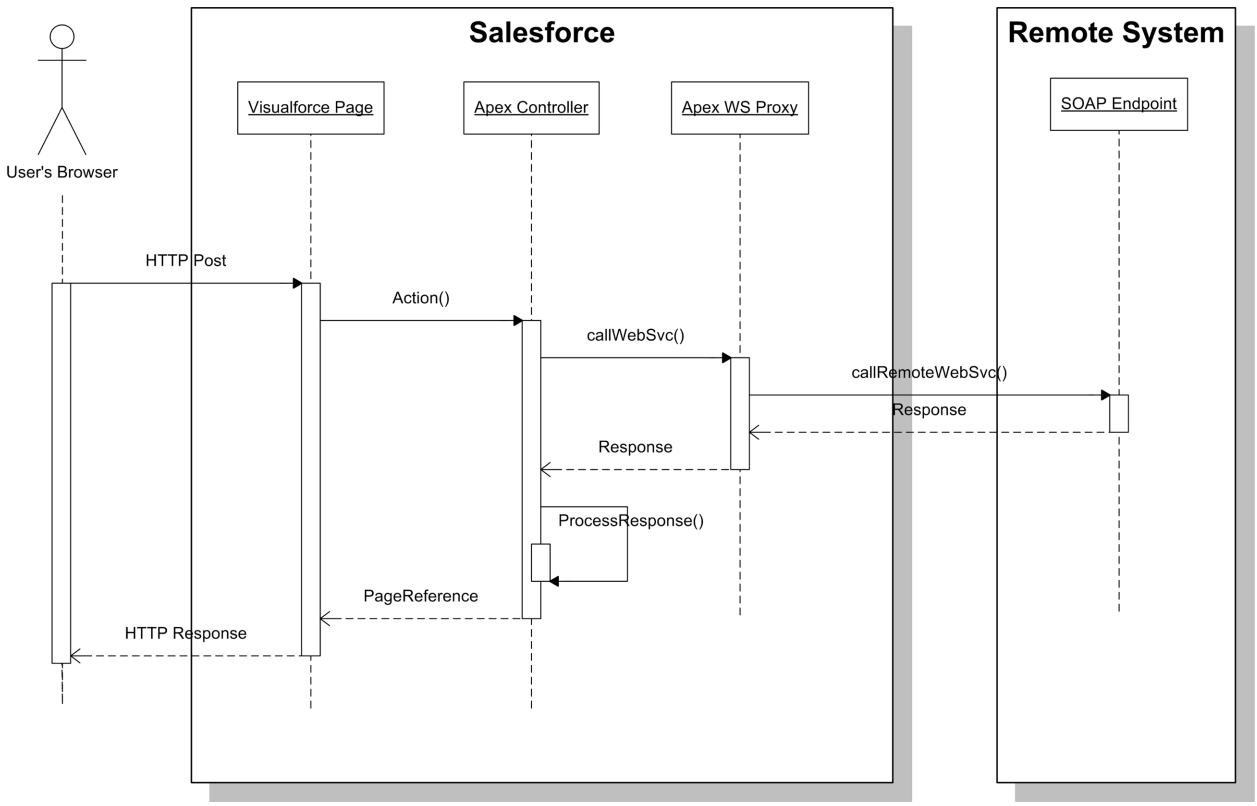
The following table contains various solutions to this integration problem.

Solution	Fit	Comments
A custom Visualforce page or button that initiates an Apex SOAP callout in a synchronous manner.	Best	Salesforce provides the ability to consume a WSDL and generate a resulting proxy Apex class. This class provides the necessary logic to call the remote service. A user-initiated action on a Visualforce page then calls an Apex controller action that then executes this proxy Apex class to perform the remote call. Visualforce pages require customization of the Salesforce application.

Solution	Fit	Comments
A custom Visualforce page or button that initiates an Apex HTTP callout in a synchronous manner.	Best	Salesforce provides the ability to invoke HTTP services using standard GET, POST, PUT, and DELETE methods. A number of HTTP classes can be used to integrate with RESTful services, although it's also possible to integrate to SOAP-based services by manually constructing the SOAP message. The latter is not recommended as it's possible for Salesforce to consume WSDLs to generate proxy classes. A user-initiated action on a Visualforce page then calls an Apex controller action that then executes this proxy Apex class to perform the remote call. Visualforce pages require customization of the Salesforce application.
A trigger that's invoked from Salesforce data changes performs an Apex SOAP or HTTP callout in a synchronous manner.	Suboptimal	You can use Apex triggers to perform automation based on record data changes. An Apex proxy class can be executed as the result of a DML operation by using an Apex trigger. However, all calls made from within the trigger context must be executed asynchronously from the initiating event. Therefore, this isn't a recommended solution for this integration problem. This solution is better suited for the Remote Process Invocation—Fire and Forget pattern.
A batch Apex job that performs an Apex SOAP or HTTP callout in a synchronous manner.	Suboptimal	Calls to a remote system can be performed from a batch job. This allows for batch remote process execution and for processing of the response from the remote system in Salesforce. However, there are limits to the number of calls for a given batch context. For more information, see Governor Limits. A given batch run can execute multiple transaction contexts (usually in intervals of 200 records). The governor limits are reset per transaction context.

Sketch

The following diagram illustrates a synchronous remote process invocation using Apex calls, where the state needs to be tracked by Salesforce.



In this scenario:

1. The user initiates an action on the Visualforce page (for example, clicks a button).
2. The browser performs an HTTP POST that in turn performs an action on the corresponding Apex controller.
3. The controller calls a previously-generated Apex Web service proxy class.
4. The proxy class performs the actual call to the remote Web service.
5. The response from the remote system is returned to the Apex controller, which then processes the response, updates any data in Salesforce as required, and re-renders the page.

In cases where subsequent state needs to be tracked, the remote system returns a unique identifier that's stored on the Salesforce record.

Results

The application of the solutions related to this pattern allows for event-initiated remote process invocations, where the result of the transaction needs to be handled by the invoking process in Salesforce.

Calling Mechanisms

The calling mechanism depends on the solution chosen to implement this pattern.

Calling mechanism	Description
Visualforce and Apex controllers	Used when the remote process is to be triggered, as part of an end-to-end process involving the user interface, and the resulting state must be displayed to the end-user and/or updated in a Salesforce record. For example, the submission of a credit card payment to an external payment gateway, where the payment results are immediately returned and displayed to the user.

Calling mechanism	Description
	Integration that's triggered from user interface events usually requires the creation of custom Visualforce pages.
Apex triggers	Used primarily for invocation of remote processes using Apex callouts from DML-initiated events. For more information about this calling mechanism, see pattern Remote Process Invocation—Fire and Forget.
Apex batch classes	Used for invocation of remote processes in batch. For more information about this calling mechanism, see pattern Remote Process Invocation—Fire and Forget.

Error Handling and Recovery

An error handling and recovery strategy must be considered as part of the overall solution.

- *Error handling*—When an error occurs (exceptions or error codes are returned to the caller), error handling is managed by the caller. For example, an error message displayed on the end-user's page or logged to a table requiring further action.
- *Recovery*—Changes aren't committed to Salesforce until a successful response is received by the caller. For example, the order status won't be updated in the database until a response that indicates success is received. If necessary, the caller can retry the operation.

Idempotent Design Considerations

Idempotent capabilities guarantee that repeated invocations are safe and will have no negative effect. If idempotency isn't implemented, then repeated invocations of the same message can have different results, potentially resulting in data integrity issues, for example, creation of duplicate records, duplicate processing of transactions, and so on.

It's important to ensure that the remote procedure being called is idempotent. It's almost impossible to guarantee that Salesforce only makes the call once (especially if the call is triggered from a user interface event). Even if Salesforce makes a single call, there's no guarantee that other processes (for example, middleware) will do the same.

The most typical method of building an idempotent receiver is for it to track duplicates based on unique message identifiers sent by the consumer. Apex Web service or REST calls should be customized to send a unique message ID.

In addition, operations that create records in the remote system should always check for duplicates before inserting. You can do this by passing a unique record ID from Salesforce, and if the record exists in the remote system, the record should be updated. In most systems, this is termed an upsert operation.

Security Considerations

Any call to a remote system must maintain the confidentiality, integrity, and availability of the request. The following are security considerations specific to using Apex SOAP and HTTP calls in this pattern:

- One-way SSL is enabled by default, but two-way SSL is supported with both self-signed and CA-signed certificates to maintain authenticity of both the client and server.
- WS-Security is not currently supported by Salesforce.
- Where necessary, consider using one-way hashes or digital signatures using the Apex Crypto class methods to ensure request integrity.
- Implement the appropriate firewall mechanisms to protect the remote system.

Sidebars

Timeliness

Timeliness is of significant importance in this pattern. In most cases:

- The request is typically invoked from the user interface, therefore, the process shouldn't keep the user waiting.

- Salesforce has a configurable timeout of up to 60 seconds for calls from Apex.
- Completion of the remote process should be executed in a timely manner to conclude within the Salesforce timeout limit and/or within user expectations.

Data Volumes

This pattern should be used primarily for small volume, real-time activities. This is due to the relatively small timeout values and maximum size of the request or response for the Apex call solution. Therefore, this pattern shouldn't be used in batch processing activities where the data payload is contained in the message.

Endpoint Capability and Standards Support

The capability and standards support for the endpoint depends on the solution that you choose.

Solution	Endpoint considerations
Apex SOAP callouts	The endpoint must be capable of receiving a Web service call via HTTP. The endpoint must be accessible over the public Internet by Salesforce. This solution requires that the remote system be compatible with the standards supported by Salesforce. At the time of writing, the Web service standards supported by Salesforce for Apex SOAP callouts are: • WSDL 1.1 • SOAP 1.1 • WSI-Basic Profile 1.1 • HTTP
Apex HTTP callouts	The endpoint must be capable of receiving HTTP calls. The endpoint must be accessible over the public Internet by Salesforce. Apex HTTP callouts can be used to call RESTful services using the standard GET, POST, PUT, and DELETE methods.

State Management

When integrating systems, keys are important for on-going state tracking, for example, if a record gets created in the remote system, in order to support ongoing updates to that record. There are two options:

- Salesforce stores the remote system's primary or unique surrogate key for the remote record.
- The remote system stores the Salesforce unique record ID or some other unique surrogate key.

There are specific considerations for handling integration keys, depending on which system contains the master record (Salesforce or the remote system), as shown in the following table.

Master system	Description
Salesforce	In this scenario, the remote system should store either the Salesforce RecordId or some other unique surrogate key from the record.
Remote system	In this scenario, the call to the remote process should return the unique key from the application and Salesforce stores that key value in a unique record field.

Complex Integration Scenarios

In certain cases, the solution prescribed by this pattern can require the implementation of several complex integration scenarios, usually best served by using middleware or having Salesforce call a composite service. These scenarios include:

- Orchestration of business processes and rules involving complex flow logic
- Aggregation of calls and their results across calls to multiple systems
- Transformation of both inbound and outbound messages
- Maintaining transactional integrity across calls to multiple systems

Governor Limits

Due to the multi-tenant nature of the Salesforce platform, there are limits to Apex callouts.

- Only 10 callouts can be made in a given execution context
- A maximum of 60 seconds invocation time for a given callout and 120 seconds of invocation time for all callouts in a given execution context
- A maximum message size of 3 MB for a given callout request or response

Middleware Capabilities

The following table highlights the desirable properties of a middleware system that participates in this pattern.

Property	Mandatory	Desirable	Not required
Event handling		X	
Protocol conversion		X	
Translation and transformation		X	
Queuing and buffering		X	
Synchronous transport protocols	X		
Asynchronous transport protocols			X
Mediation routing		X	
Process choreography and service orchestration		X	
Quality of service (encryption, signing, reliable delivery, transaction management)		X	
Routing			X
Extract, transform, and load			X

Example

A utility company uses Salesforce and has a separate system that contains customer billing information. They want to display the billing history for a customer account without having to store that data in Salesforce. They have an existing Web service that can return a list of bills and their details for a given account number, but cannot otherwise display this data in a browser.

This requirement can be accomplished with the following:

- Salesforce consumes the billing history service WSDL from an Apex proxy class.
- Create a Visualforce page and custom controller to execute this Apex proxy class with the account number as the unique identifier.
- The custom controller then parses the return values from the Apex callout and the Visualforce page and subsequently renders the bill to the user.

This example demonstrates the following:

- The state of the customer is tracked with an account number stored on the Salesforce account object.
- Subsequent processing of the reply message by the caller.

Chapter 4

Remote Process Invocation—Fire and Forget

Context

You use Salesforce to track leads, manage your pipeline, create opportunities, and capture order details that convert leads to customers. However, Salesforce isn't the system that holds or processes orders. After the order details are captured in Salesforce, an order needs to be created in the remote system, then the remote system manages the order through to its conclusion.

When you implement this pattern, Salesforce makes a call to the remote system to create the order, but doesn't wait for successful completion of that call. The remote system can optionally update Salesforce with the new order number and status in a separate transaction.

Problem

When an event occurs in Salesforce, how do you initiate a process in a remote system and pass the required information to that process without waiting for a response from the remote system?

Forces

There are various forces to consider when applying solutions based on this pattern:

- Does the call to the remote system require Salesforce to wait for a response before continuing processing? In other words, is the call to the remote system synchronous request-reply or asynchronous?
- If the call to the remote system is synchronous, does the response need to be processed by Salesforce as part of the same transaction as the call?
- Is the message size relatively small?
- Is the integration based on the occurrence of a specific event such as a button click in the Salesforce user interface or DML-based events?
- Is guaranteed message delivery from Salesforce to the remote system a requirement?
- Is the remote system able to participate in a contract-first integration where Salesforce specifies the contract? In some solution variants (for example, outbound messaging), Salesforce specifies a contract that must be implemented by the remote system endpoint.
- Are declarative configuration methods preferred over custom Apex development? In this case, solutions such as outbound messaging are preferred over Apex callouts.

Solution

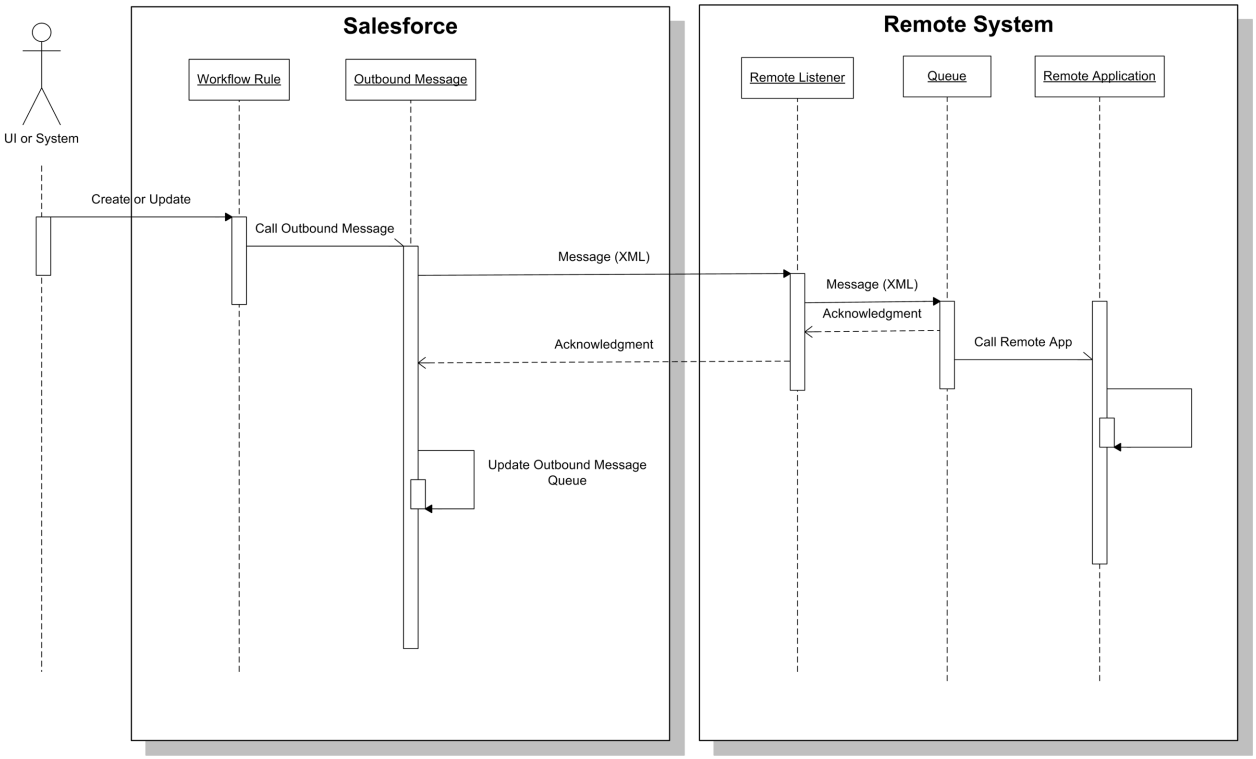
The following table contains various solutions to this integration problem.

Solution	Fit	Comments
Workflow-driven outbound messaging	Best	No customization is required in Salesforce to implement outbound messaging; therefore, it's the recommended

Solution	Fit	Comments
		solution for this type of integration, where the remote process is invoked from an insert or update event. Salesforce provides a workflow-driven outbound messaging capability that allows the sending of SOAP messages to remote systems triggered by an insert or update operation in Salesforce. These messages are sent asynchronously and are independent of the Salesforce user interface. The outbound message is sent to a specific remote endpoint. The remote service must be able to participate in a contract-first integration where Salesforce provides the contract. On receipt of the message, the remote service must respond with a positive acknowledgment, otherwise Salesforce attempts to retry sending the message, thus providing a form of guaranteed delivery. When using middleware, this becomes a "first-mile" guarantee of delivery.
Custom Visualforce page that initiates an Apex SOAP or HTTP asynchronous callout	Good	This solution is typically used in user interface-based scenarios, but does require customization. In addition, the solution must handle guaranteed delivery of the message in the code. Similar to the solution for the Remote Process Invocation—Request and Reply pattern solution that specifies using a Visualforce page, together with an Apex callout. The difference is that in this pattern, Salesforce doesn't wait for the request to complete before handing off control to the user. After receiving the message, the remote system responds and indicates receipt of the message, then asynchronously processes the message. The remote system hands control back to Salesforce before it begins to process the message; therefore, Salesforce doesn't have to wait for processing to complete.
A trigger that's invoked from Salesforce data changes performs an Apex SOAP or HTTP asynchronous callout	Suboptimal	You can use Apex triggers to perform automation based on record data changes. An Apex proxy class can be executed as the result of a DML operation by using an Apex trigger. However, all calls made from within the trigger context must be executed asynchronously.
A batchApex job that performs an Apex SOAP or HTTP asynchronous callout	Suboptimal	Calls to a remote system can be performed from a batch job. This allows for batch remote process execution and for processing of the response from the remote system in Salesforce. However, there are limits to the number of calls for a given batch context. For more information about limits, see the *Salesforce Limits Quick Reference Guide*.

Sketch

The following diagram illustrates a call from Salesforce to a remote system, where the call is triggered by the create or update operations on a record in Salesforce.



In this scenario:

1. A DML update or insert occurs on a given set of records in Salesforce.
2. A Salesforce workflow rule triggers, based on a given set of conditions.
3. This workflow rule invokes a pre-configured outbound message that sends a SOAP-based message to a remote listener.
4. The remote listener receives the SOAP message, places the message on a local queue, and returns a positive acknowledgment to Salesforce as part of the same transaction context.
5. The queueing application forwards the message to the remote application for processing.
6. Salesforce receives the acknowledgment and completes the request, but doesn't wait for the remote application to process the message.
7. Salesforce waits for an acknowledgment from the remote system for up to 10 seconds. After 10 seconds, Salesforce retries sending the outbound message request for up to 24 hours.

In the case where the remote system needs to perform operations against Salesforce, an optional call back operation can be implemented. The outbound message sends a SessionId that can be used in the call back to authenticate and authorize a subsequent API or Web service call into Salesforce.

Results

The application of the solutions related to this pattern allows for:

- User interface-initiated remote process invocations, where the result of the transaction might or might not be displayed to the end user
- DML event-initiated remote process invocations, where the result of the transaction might or might not be processed by the calling process

Calling Mechanisms

The calling mechanism depends on the solution chosen to implement this pattern.

Calling mechanism	Description
Visualforce and Apex controllers	Used to invoke a remote process asynchronously, using an Apex callout.
Workflow Rules	Used only for the outbound messaging solution. Create and update DML events trigger the Salesforce workflow rules, which can then send a message to a remote system.
Apex triggers	Used primarily for invocation of remote processes, using Apex callouts from DML-initiated events.
Apex batch classes	Used for invocation of remote processes in batch mode.

Error Handling and Recovery

An error handling and recovery strategy must be considered as part of the overall solution. The best method depends on the solution you choose.

Solution	Error handling and recovery strategy
Apex callouts	• *Error handling*—The remote system hands off invocation of the end process so the callout only needs to handle exceptions in the initial invocation of the remote service. (For example, a timeout event is triggered if no positive acknowledgment is received from the remote callout.). The remote system must handle any subsequent errors once the initial invocation is handed off for asynchronous processing. • *Recovery*—Recovery is more complex in this scenario. A custom retry mechanism needs to be created if quality of service requirements dictate it.
Outbound Messaging	• *Error handling*—Because this pattern is asynchronous, the remote system needs to handle any error-handling. In the case of outbound messaging, Salesforce initiates a retry operation if no positive acknowledgment is received within the timeout period, for up to 24 hours. Error handling needs to be performed in the remote service because the message is effectively handed off to the remote system in a “fire-and-forget” manner. • *Recovery*—Because this pattern is asynchronous, the system needs to initiate any retries based on the service’s quality of service requirements. In the case of outbound messaging, Salesforce initiates retries if no positive acknowledgment is received from the outbound listener within the timeout period, for up to 24 hours. The retries escalate at increasing intervals. Administrators must monitor this queue for any messages exceeding the 24 hour delivery period and retry manually, if required. In the case of custom Apex callouts, a custom retry mechanism needs to be built if the quality of service requirements warrant it.

Idempotent Design Considerations

The idempotent design considerations in the Remote Process Invocation—State Tracking pattern also apply to this pattern.

Idempotency is particularly important with outbound messaging because it’s asynchronous and retries are initiated if no positive acknowledgement is received. Therefore, the remote service must be able to handle messages from Salesforce in an idempotent fashion.

Outbound messaging sends a unique ID per message and this ID remains the same for any retries. The remote system can track duplicate messages based on this unique ID. The unique record ID for each record being updated is also sent, and can be used to prevent duplicate record creation.

Security Considerations

Any call to a remote system must maintain the confidentiality, integrity, and availability of the request. Different security considerations apply, depending on the solution you choose.

Solution	Security considerations
Apex callouts	Any call to a remote system must maintain the confidentiality, integrity, and availability of the request. The following are security considerations specific to using Apex SOAP and HTTP calls in this pattern: • One-way SSL is enabled by default, but two-way SSL is supported with both self-signed and CA-signed certificates to maintain authenticity of both the client and server. • WS-Security is not currently supported by Salesforce when generating the Apex proxy class. • Where necessary, consider using one-way hashes or digital signatures using the Apex Crypto class methods to ensure the integrity of the request. • Implement the appropriate firewall mechanisms to protect the remote system.
Outbound Messaging	For outbound messaging, one-way SSL is enabled by default. However, two-way SSL can be used together with the Salesforce outbound messaging certificate. The following are some additional security considerations: • Remote integration servers should whitelist Salesforce server IP ranges. • The remote system should be appropriately protected by firewalls.

Sidebars

Timeliness

Timeliness is less of a factor with the fire-and-forget pattern because control is handed back to the client either immediately or after positive acknowledgment of a successful hand off in the remote system. With Salesforce outbound messaging, the acknowledgment must occur within 24 hours or the message expires.

Data Volumes

Data volume considerations depend on which solution you choose and communication type.

Solution	Communication type	Limits
Apex callouts from a batch or a trigger	Asynchronous or batch	It's possible to perform Apex callouts from an Apex batch class, but there are a limited number of callouts (10) per batch execution (a single batch run is comprised of multiple small batch executions). Therefore, callouts in batch need to be able to support multiple records in a given request rather than being treated as atomic transactions. In addition, asynchronous Apex callouts must be invoked through triggers that fire due to DML record changes. DML operations in Salesforce can occur for more than one record and a single trigger invocation can consist of up to 200 records. Therefore, any DML-initiated or batch-initiated

Solution	Communication type	Limits
		integration can't be implemented on a single record per message basis given the previous limitations.
Outbound messaging	Asynchronous	A single outbound message can contain up to 100 records. Therefore, it's important that the remote listener can handle more than a single record in its implementation.

Endpoint Capability and Standards Support

The capability and standards support for the endpoint depends on the solution that you choose.

Solution	Endpoint considerations
Apex SOAP callouts	The endpoint must be capable of processing a Web service call via HTTP. The endpoint must be accessible over the public Internet by Salesforce. This solution requires that the remote system be compatible with the standards supported by Salesforce. At the time of writing, the Web service standards supported by Salesforce for Apex SOAP callouts are: • WSDL 1.1 • SOAP 1.1 • WSI-Basic Profile 1.1 • HTTP
Apex HTTP callouts	The endpoint must be capable of receiving HTTP calls and be accessible over the public internet by Salesforce. Apex HTTP callouts can be used to call RESTful services using the standard GET, POST, PUT, and DELETE methods.
Outbound message	• The endpoint must be capable of implementing a listener that can receive SOAP messages in predefined format sent from Salesforce. • The remote listener must participate in a contract-first implementation, where the contract is supplied by Salesforce. • Each outbound message has its own predefined WSDL.

State Management

When integrating systems, unique record identifiers are important for on-going state tracking. For example, if a record is created in the remote system, in order to support ongoing updates to that record, there are two options:

- Salesforce stores the remote system's primary or unique surrogate key for the remote record.
- The remote system stores the Salesforce unique record ID or some other unique surrogate key.

The following table contains considerations for state management in this pattern.

Master system	Description
Salesforce	In this scenario, the remote system should store either the Salesforce RecordId or some other unique surrogate key in the Salesforce record.

Master system	Description
Remote system	In this scenario, Salesforce must store a reference to the unique identifier in the remote system. Because the process is asynchronous, storing this unique identifier can't be done as part of the original transaction. Salesforce must provide a unique ID in the call to the remote process. The remote system must then call back to Salesforce to update the record in Salesforce with the remote system's unique identifier, using the Salesforce unique ID. This implies either specific state handling in the remote application, to store the Salesforce unique identifier for that transaction to use for the call back when processing is complete, or the Salesforce unique identifier is stored on the remote system's record.

Complex Integration Scenarios

Each solution in this pattern has different considerations for complex integration scenarios, including transformation, process orchestration, and so on.

Solution	Considerations
Apex callouts	In certain cases, solutions prescribed by this pattern might require the implementation of several complex integration scenarios, usually best served by using middleware or having Salesforce call a composite service. These scenarios include: • Orchestration of business processes and rules involving complex flow logic • Aggregation of calls and their results across calls to multiple systems • Transformation of both inbound and outbound messages • Maintaining transactional integrity across calls to multiple systems
Outbound messaging	Given the static, declarative nature of the outbound message, no complex integration scenarios, such as aggregation, orchestration, or transformation, can be performed in Salesforce. These types of operations must be handled by the remote system or middleware.

Governor Limits

Due to the multi-tenant nature of the Salesforce platform, there are limits to outbound callouts. Limits depend on the type of outbound call and the timing of the call.

Solution	Limits
Apex callouts	A number of governor limits apply when invoking either Apex SOAP or HTTP callouts from Salesforce. • Only 10 callouts can be made in a given execution context. • A maximum invocation time of 10 seconds for a given callout and 120 seconds for all callouts in a given execution context. • A maximum message size of 3 MB for any callout request or response. In addition, there's also a maximum of 10 asynchronous requests per transaction context and a maximum number of asynchronous requests per 24 hour window per license count. Because this event is executed via DML and can occur in batch, this approach should be implemented with extreme care based on the expected transaction volumes.

Solution	Limits
Outbound messaging	No governor limits apply.

Reliable Messaging

Reliable messaging attempts to resolve the issue of guaranteeing the delivery of a message to a remote system, where the individual components themselves might be unreliable. The method of ensuring receipt of a message by the remote system depends on the solution you choose.

Solution	Reliable messaging considerations
Apex callouts	Salesforce doesn't provide explicit support for any reliable messaging protocols (for example, WS-ReliableMessaging). We recommend that the remote endpoint receiving the Salesforce message implement a reliable messaging system (based on JMS or MQ, for example) to ensure full end-to-end guaranteed delivery to the remote system that ultimately processes the message. However, this doesn't ensure guaranteed delivery from Salesforce to the remote endpoint that it calls. Guaranteed delivery must be handled through customizations to Salesforce. Specific techniques, such as processing a positive acknowledgment from the remote endpoint in addition to custom retry logic, needs to be implemented.
Outbound messaging	Outbound messaging provides a form of reliable messaging because the process retries for up to 24 hours if no positive acknowledgment is received from the remote system. However, this only guarantees delivery to the point of the remote listener. In most implementations, the remote listener calls another remote service. Ideally, the invocation of this remote service should be via a reliable messaging system (based on JMS or MQ, for example) to ensure full end-to-end guaranteed delivery. The positive acknowledgement to the Salesforce outbound message should occur after the remote listener has successfully placed its own message on its local queue.

Middleware Capabilities

The following table highlights the desirable properties of a middleware system that participates in this pattern.

Property	Mandatory	Desirable	Not required
Event handling		X	
Protocol conversion		X	
Translation and transformation		X	
Queuing and buffering	X		
Synchronous transport protocols			X
Asynchronous transport protocols	X		
Mediation routing		X	
Process choreography and service orchestration		X	

Property	Mandatory	Desirable	Not required
Quality of service (encryption, signing, reliable delivery, transaction management)	X		
Routing			X
Extract, transform, and load			X

Solution Variant — Outbound Messaging and Message Sequencing

Salesforce outbound messaging can't inherently guarantee the sequence of delivery for its messages. This is because a single message can be retrieved over a 24 hour period. There are multiple methods for handling message sequencing in the remote system.

- Salesforce sends a unique message ID for each instance of an outbound message. The remote system discards any messages that have a duplicate message ID.
- Salesforce sends only the RecordId. The remote system then makes a call back to Salesforce (for example, via SOAP API or REST API) to obtain the necessary data to process the request.

Solution Variant — Outbound Messaging and Deletes

Salesforce workflow rules can't track deletion of a record, it can only track the insert or update of a record. Therefore, it's not possible to directly initiate an outbound message from the deletion of a record. You can do this indirectly with the following process:

1. Create a custom object to store key information from the deleted records.
2. Create an Apex trigger, fired by the deletion of the base record, to store information such as the unique identifier in the custom object.
3. Implement a workflow rule to initiate an outbound message based on the creation of the custom object record.

It's important that state tracking is enabled either by storing the remote system's unique identifier in Salesforce or storing the Salesforce unique identifier in the remote system.

Example

A telecommunications company wishes to use Salesforce as a front end for creating new accounts using the Lead to Opportunity process. The creation of an order is initiated in Salesforce once the opportunity is closed and won, but the back end ERP system will be the data master. The order must be subsequently saved to the Salesforce opportunity record and the opportunity status changed to indicate that the order was created.

The following constraints apply:

- The ERP system is capable of participating in a contract-first integration, where its service must implement a Salesforce WSDL interface.
- There should be no custom development in Salesforce.
- The user doesn't need to be immediately notified of the order number after the opportunity converts to an order.

This example is best implemented using Salesforce outbound messaging, but does require the implementation of a proxy service by the remote system.

On the Salesforce side:

- Create a workflow rule to initiate the outbound message (for example, when the opportunity status changes to "Close-Won").
- Create an outbound message that sends only the opportunity RecordId and a SessionId for a subsequent call back.

On the remote system side:

- Create a proxy service that can implement the Salesforce outbound message WSDL interface.

- The service will receive one or more notifications indicating that the opportunity is to be converted to an order.
- The service transforms and places the message on a local message queue and on notification of receipt, replies with a positive acknowledgment back to the Salesforce outbound message.
- The local message queue forwards the message to the backend ERP system so the order can be created.
- After the order is successfully created, a separate thread calls back to Salesforce using the SessionId as the authentication token to update the opportunity with the order number and status. This call back can be done using previously documented pattern solutions, such as the Salesforce SOAP API, REST API, or an Apex Web service.

This example demonstrates the following:

- Implementation of a remote process invoked asynchronously
- End-to-end guaranteed delivery
- Subsequent call back to Salesforce to update the state of the record

Chapter 5

Batch Data Synchronization

Context

You're moving your CRM implementation to Salesforce and want to:

- Extract and transform accounts, contacts, and opportunities from the current CRM system and load the data into Salesforce (initial data import).
- Extract, transform, and load customer billing data into Salesforce from a remote system on a weekly basis (ongoing).
- Extract customer activity information from Salesforce and import it into an on-premises data warehouse on a weekly basis (ongoing).

Problem

How do you import data into Salesforce and export data out of Salesforce, taking into consideration that these imports and exports can interfere with end-user operations during business hours, and involve large amounts of data?

Forces

There are various forces to consider when applying solutions based on this pattern:

- Should the data be stored in Salesforce? If not, there are other integration options an architect can and should consider (mashups, for example).
- If the data should be stored in Salesforce, should the data be refreshed in response to an event in the remote system?
- Should the data be refreshed on a scheduled basis?
- Does the data support primary business processes?
- Are there analytics (reporting) requirements that are impacted by the availability of this data in Salesforce?

Solution

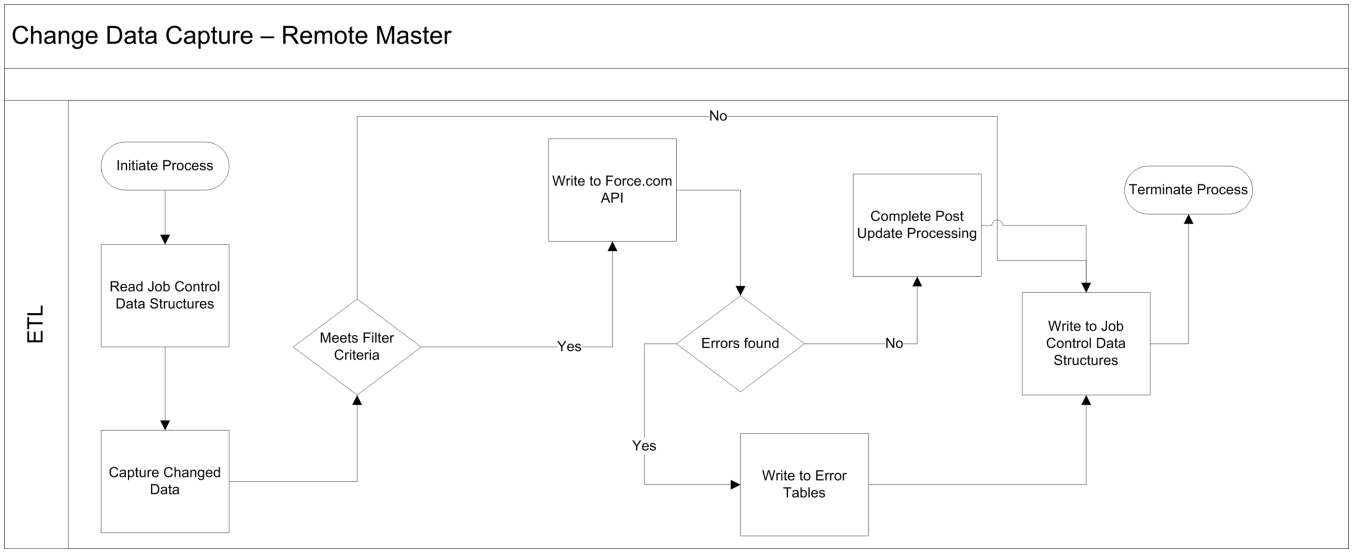
The following table contains various solutions to this integration problem.

Solution	Fit	Data master	Comments
Change data capture	Best	Remote system	Leverage a third-party ETL tool that allows you to run change data capture against source data. The tool reacts to changes in the source data set, transforms the data, and then calls Salesforce Bulk API to issue DML statements. This can also be implemented using the Salesforce SOAP API.

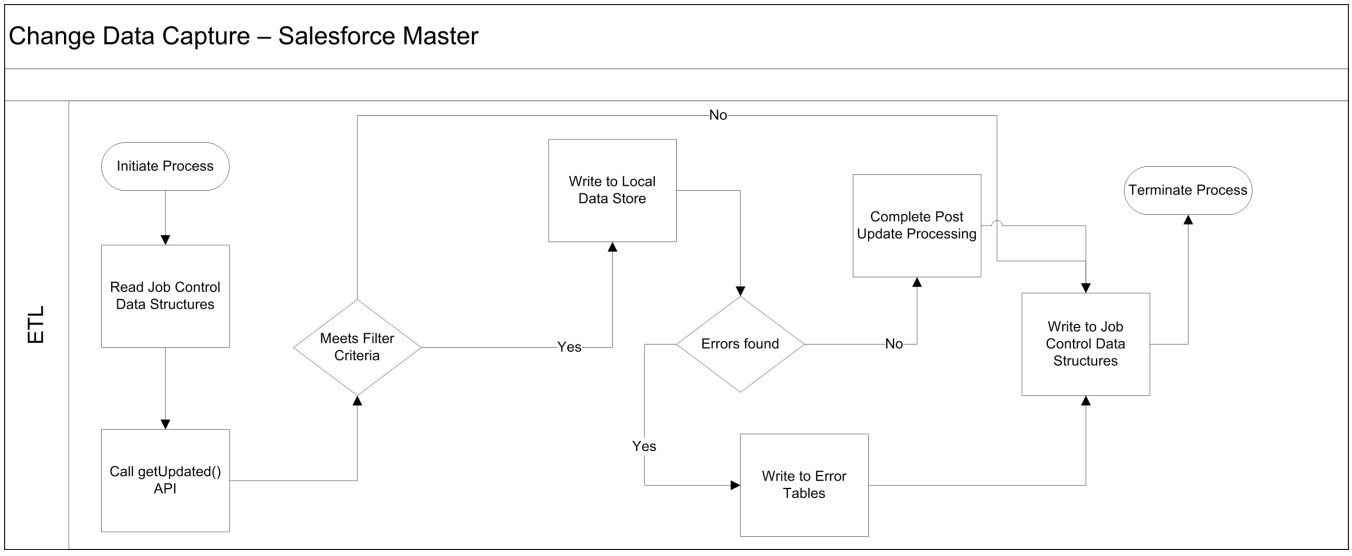
Solution	Fit	Data master	Comments
Change data capture	Best	Salesforce	Leverage a third-party ETL tool that allows you to run change data capture against ERP and Salesforce data sets. In this solution, Salesforce is the data source, and you can use time/status information on individual rows to query the data and filter the target result set. This can be implemented by using SOQL together with SOAP API and the `query()` method, or by the using SOAP API and the `getUpdated()` method.
Remote call-in	Suboptimal	Remote system	It's possible for a remote system to call into Salesforce by using one of the APIs and perform updates to data as they occur. However, this causes considerable on-going traffic between the two systems. It also requires that greater emphasis be placed on error handling and locking because this pattern has the potential for causing continual updates, which has the potential to impact performance for end-users.
Remote process invocation	Suboptimal	Salesforce	It's possible for Salesforce to call into a remote system and perform updates to data as they occur. However, this causes considerable on-going traffic between the two systems. It also requires that greater emphasis be placed on error handling and locking because this pattern has the potential for causing continual updates, which has the potential to impact performance for end-users.

Sketch

The following diagram illustrates the sequence of events in this pattern, where the remote system is the data master.



The following diagram illustrates the sequence of events in this pattern, where Salesforce is the data master.



Results

You can integrate data that's sourced externally with Salesforce under the following scenarios:

- External system is the data master—Salesforce is a consumer of data provided by a single source system or multiple systems. In this scenario, it's common to have a data warehouse or data mart that aggregates the data before the data is imported into Salesforce.
- Salesforce is the data master—Salesforce is the system of record for certain entities.

In a typical Salesforce integration scenario, the implementation team does one of the following:

- Implement change data capture on the source data set.
- Implement a set of supporting database structures, known as control tables, in an intermediate, on-premises database.

The ETL tool is then used to create programs that will:

1. Read a control table to determine the last run time of the job and extract any other control values needed.
2. Use the above control values as filters and query the source data set.
3. Apply predefined processing rules, including validation, enrichment, and so on.
4. Use available connectors/transformation capabilities of the ETL tool to create the destination data set.
5. Write the data set to Salesforce objects.
6. If processing is successful, update the control values in the control table.
7. If processing fails, update the control tables with values that enable a restart and exit.

For an ETL tool to gain maximum benefit from data synchronization capabilities, consider the following:

- Chain and sequence the ETL jobs to provide a cohesive process.
- Use primary keys from both systems to match incoming data.
- Use specific API methods to extract only updated data.
- Group the imported data to avoid locking. For example, if you're importing contact data, be sure that all the contacts for a single account are grouped together. Otherwise, the account record locks when importing the first contact record, causing subsequent contact records for that account to fail.
- Any post-import processing, such as triggers, should only process data selectively.
- If your scenario involves large data volumes, follow the best practices in the white paper *Best Practices for Deployments with Large Data Volumes*.

Error Handling and Recovery

An error handling and recovery strategy must be considered as part of the overall solution. The best method depends on the solution you choose.

Error location	Error handling and recovery strategy
Read from Salesforce	• *Error handling*—If an error occurs during a read operation, implement a retry for errors that aren't infrastructure-related. In the event of repeated failure, standard processing using control tables/error tables should be implemented in the context of an ETL operation to: ◊ Log the error ◊ Retry the read operation ◊ Terminate if unsuccessful ◊ Send a notification • *Recovery*—Restart the ETL process to recover from a failed read operation. If the operation succeeds but there are failed records, an immediate restart or subsequent execution of the job should address the issue. In this case, a delayed restart might be a better solution because it allows time to triage and correct data that might be causing the errors.
Write to Salesforce	• *Error handling*—Errors that occur during a write operation can result from a combination of factors in the application. The API calls return a result set that consists of the information listed below. This information should be used to retry the write operation (if necessary). ◊ Record identifying information ◊ Success/failure notification ◊ A collection of errors for each record • *Recovery*—Restart the ETL process to recover from a failed read operation. If the operation succeeds but there are failed records, an immediate restart or subsequent execution of the job should address the issue. In this case, a delayed restart might be a better solution because it allows time to triage and correct data that might be causing the errors.
External master system	Errors should be handled in accordance with the best practices of the master system.

Security Considerations

Any call to a remote system must maintain the confidentiality, integrity, and availability of the request. Different security considerations apply, depending on the solution you choose.

- A Force.com license is required to allow authenticated API access to the Salesforce API.
- We recommend that standard encryption be used to keep password access secure.
- Use the HTTPS protocol when making calls to the Salesforce APIs. You can also proxy traffic to the Salesforce APIs through an on-premises security solution, if necessary.

Sidebars

Timeliness

Timeliness isn't of significant importance in this pattern. However, care must be taken to design the interfaces so that all of the batch processes complete in a designated batch window.

As with all batch-oriented operations, we strongly recommend that you take care to insulate the source and target systems during batch processing windows. Loading batches during business hours might result in some contention, resulting in either a user's update failing, or more significantly, a batch load (or partial batch load) failing.

For organizations that have global operations, it might not be feasible to run all batch processes at the same time because the system might continually be in use. Data segmentation techniques using record types and other filtering criteria can be used to avoid data contention in these cases.

State Management

You can implement state management by using surrogate keys between the two systems. If you need any type of transaction management across Salesforce entities, we recommend that you use the Remote Call-In pattern using Apex.

Standard optimistic record locking occurs on the platform, and any updates made using the API require the user, who is editing the record, to refresh the record and initiate their transaction. In the context of the Salesforce API, optimistic locking refers to a process where:

- Salesforce doesn't maintain the state of a record being edited by a specific user.
- Upon read, it records the time when the data was extracted.
- If the user updates the record and saves it, Salesforce checks to see if another user has updated the record in the interim.
- If the record has been updated, the system notifies the user that an update was made and the user should retrieve the latest version of the record before proceeding with their updates.

Middleware Capabilities

The most effective external technologies used to implement this pattern are traditional ETL tools. It's important that the middleware tools chosen support the Salesforce Bulk API.

It's helpful, but not critical, that the middleware tools support the `getUpdated()` function. This function provides the closest implementation to standard change data capture capability on the Salesforce platform.

Example

A utility company uses a mainframe-based batch process that assigns prospects to individual sales reps and teams. This information needs to be imported into Salesforce on a nightly basis.

The customer has decided to implement change data capture on the source tables using a commercially available ETL tool.

The solution works as follows:

- A cron-like scheduler executes a batch job that assigns prospects to users and teams.
- After the batch job runs and updates the data, the ETL tool recognizes these changes using change data capture. The ETL tool collates the changes from the data store.
- The ETL connector uses the Salesforce SOAP API to load the changes into Salesforce.

Chapter 6

Remote Call-In

Context

You use Salesforce to track leads, manage your pipeline, create opportunities, and capture order details that convert leads to customers. However, Salesforce isn't the system that contains or processes orders. Orders are managed by an external (remote) system that needs to update the order status in Salesforce as the order passes through its processing stages.

Problem

How does a remote system connect and authenticate with Salesforce and update existing records?

Forces

There are various forces to consider when applying solutions based on this pattern:

- Does the call to Salesforce require the remote process to wait for a response before continuing processing? Remote calls to Salesforce are always synchronous request-reply, although the remote process can discard the response if it's not needed to simulate an asynchronous call.
- What is the format of the message (for example, SOAP or REST, or both over HTTP)?
- Is the message size relatively small or large?
- In the case of a SOAP-capable remote system, is the remote system able to participate in a contract-first approach, where Salesforce dictates the contract? This is required where our standard Web service APIs are used, for which a predefined WSDL is supplied.
- Is transaction processing required?
- What is the extent to which you are tolerant of customization in the Salesforce application?

Solution

The following table contains various solutions to this integration problem.

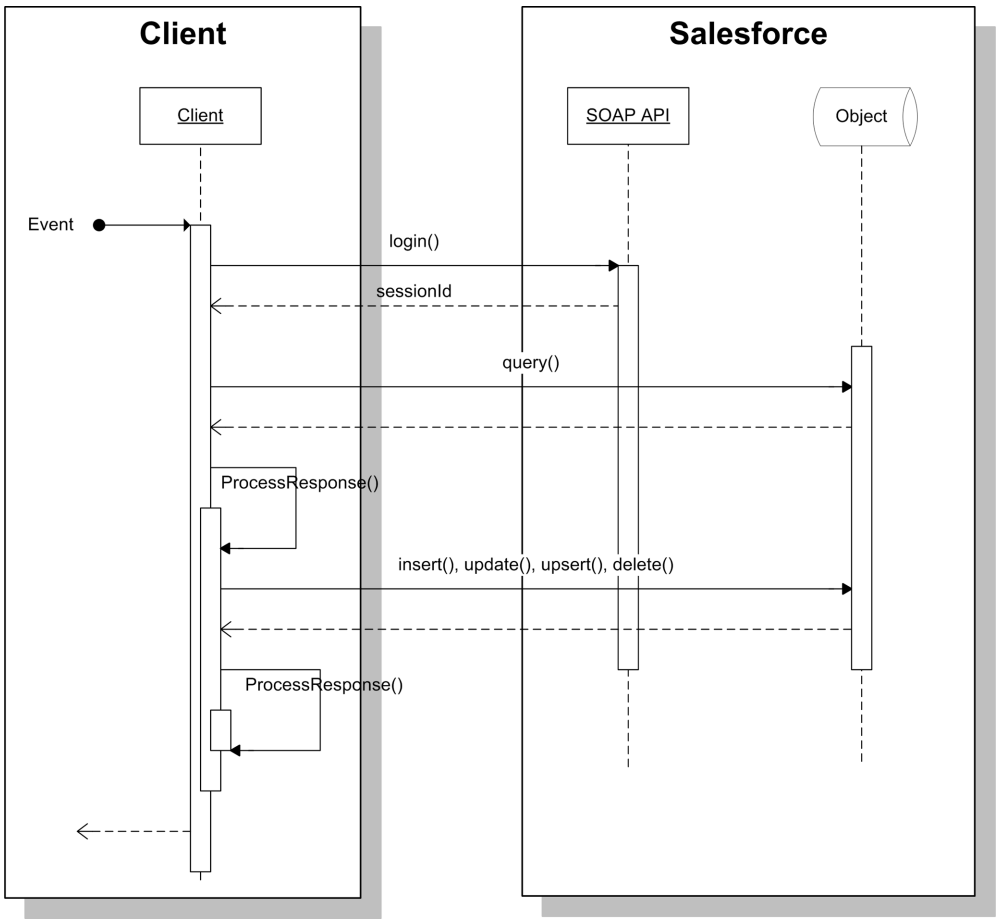
Solution	Fit	Comments
SOAP API	Best	• *Accessibility*—Salesforce provides a SOAP API that remote systems can use to: ◊ Query data in your organization ◊ Create, update, and delete data ◊ Obtain metadata about your organization ◊ Run utilities to perform administrative tasks • *Synchronous API*—After the API call is made, the remote client application waits until it receives a

Solution	Fit	Comments
		response from the service. Asynchronous calls to Salesforce aren't supported. • *Generated WSDL*—Salesforce provides two WSDLs for remote systems: ◊ Enterprise WSDL—Provides a strongly-typed WSDL that's specific to a Salesforce organization. ◊ Partner WSDL—Contains a loosely-typed WSDL that's not specific to a Salesforce organization. • *Security*—The client executing SOAP API must have a valid login and obtain a session to perform any API calls. The API respects object-level and field-level security configured in the application based on the logged in user's profile. • *Transaction/Commit Behavior*—By default, each API call allows for partial success if some records are marked with errors. This can be changed to an "all or nothing" behavior where all the results are rolled back if any error occurs. It's not possible to span a transaction across multiple API calls. To overcome this limitation, it's possible for a single API call to affect multiple objects. • *Bulk Data*—For bulk data operations (more than 500,000 records), use the REST-based Bulk API.
REST API	Best	• *Accessibility*—Salesforce provides a REST API that remote systems can use to: ◊ Query data in your organization ◊ Create, update, and delete data ◊ Obtain metadata about your organization ◊ Run utilities to perform administrative tasks • *Synchronous API*—After the API call is made, the remote client application waits until it receives a response from the service. Asynchronous calls to Salesforce aren't supported. • *REST API vs. SOAP API*—REST exposes resources (entities/objects) as URIs and uses HTTP verbs to define CRUD operations on these resources. Unlike SOAP, the REST API requires no predefined contract, utilizes XML and JSON for responses, and has loose typing. REST API is lightweight and provides a simple method for interacting with Salesforce. Its advantages include ease of integration and development, and it's an excellent choice for use with mobile applications and Web 2.0 projects. • *Security*—The client executing REST API must have a valid login and obtain a session to perform any API calls. The API respects object-level and field-level

Solution	Fit	Comments
		security configured in the application based on the logged in user's profile. • *Transaction/Commit Behavior*—By default, each record is treated as a separate transaction and committed separately. Failure of one record change doesn't cause rollback of other record changes. This can be altered to an "all or nothing" behavior. • *Bulk Data*—For bulk data operations (more than 500,000 records), use the REST-based Bulk API.
Apex Web services	Suboptimal	Apex class methods can be exposed as Web service methods to external applications. This is an alternative to SOAP API, and is typically only used where the following additional requirements must be met: • Full transactional support is required (for example, create an account, contact, and opportunity all in one transaction). • Custom logic must be applied on the Salesforce side before committing. The benefit of using an Apex Web service must be weighed against the additional code that needs to be maintained in Salesforce.
Apex REST service	Suboptimal	An Apex class can be exposed as REST resources mapped to specific URIs with an HTTP verb defined against it (for example, POST or GET). Unlike SOAP, there is no need for the client to consume a service definition/contract (WSDL) and generate client stubs. The remote system requires only the ability to form an HTTP request and process the returned results (XML or JSON).
Bulk API	Optimal for bulk operations	Bulk API is based on REST principles, and is optimized for loading or deleting large sets of data. It has the same accessibility and security behavior as REST API. Bulk API allows the client application to query, insert, update, upsert, or delete a large number of records asynchronously by submitting a number of batches, which are processed in the background by Salesforce. In contrast, SOAP API is optimized for real-time client applications that update small numbers of records at a time. Although SOAP API can also be used for processing large numbers of records, when the data sets contain hundreds of thousands to millions of records, it becomes less practical. This is due to its relatively high overhead and lower performance characteristics.

Sketch

The following diagram illustrates the sequence of events when you implement this pattern using SOAP API. The sequence of events is the same when using REST API.



Results

The application of the solutions related to this pattern allows for:

- Remote system to call the Salesforce APIs to query the database and execute single-object operations (create, update, delete, and so on).
- Remote system to call custom-built Salesforce APIs (services) that can support multi-object transactional operations and custom pre/post processing logic.

Calling Mechanisms

The calling mechanism depends on the solution chosen to implement this pattern.

Calling mechanism	Description
SOAP API	The remote system uses the Salesforce Enterprise or Partner WSDL to generate client stubs that are in turn used to invoke the standard SOAP API.
REST API	The remote system has to authenticate before accessing any Apex REST service. The remote system can use OAuth 2.0 or username/password authentication. In either case, the client

Calling mechanism	Description
	must set the authorization HTTP header with the appropriate value (an OAuth access token or a session ID can be acquired via a login call to SOAP API). The remote system then generates REST invocations (HTTP requests) with the appropriate verbs and processes the results returned (JSON and XML data formats are supported).
Apex Web service	The remote system consumes the custom Apex Web Service WSDL to generate client stubs that are in turn used to invoke the custom Apex Web service.
Apex REST service	As per REST API, the resource URI and applicable verbs are defined using the `@RestResource`, `@HttpGet`, and `@HttpPost` annotations.
Bulk API	Bulk API is a REST-based API, so the same calling mechanisms as REST API apply.

Error Handling and Recovery

An error handling and recovery strategy must be considered as part of the overall solution.

- *Error handling*—All the remote call-in methods, standard or custom APIs, will require the remote system to handle any subsequent errors, such as timeouts and the management of retries. Middleware can be used to provide the logic for error handling and recovery.
- *Recovery*—A custom retry mechanism needs to be created if quality of service requirements dictate it. In this case, it's important to ensure idempotent design characteristics.

Idempotent Design Considerations

Idempotent capabilities guarantee that repeated invocations are safe and will have no negative effect. If idempotency isn't implemented, then repeated invocations of the same message can have different results, potentially resulting in data integrity issues, for example, creation of duplicate records, duplicate processing of transactions, and so on.

The remote system must manage multiple (duplicate) calls, in the case of errors or timeouts, to avoid duplicate inserts and redundant updates (especially if downstream triggers and workflow rules fire). While it's possible to manage some of these situations within Salesforce (particularly in the case of custom SOAP and REST services), we recommend that the remote system (or middleware) manages error handling and idempotent design.

Security Considerations

Different security considerations apply, depending on the pattern solution chosen. In all cases, the platform uses the logged-in user's access rights (for example, profile settings, sharing rules, permission sets, and so on). Additionally, profile IP restrictions can be used to restrict access to the API for a specific IP address range.

Solution	Security considerations
SOAP API	Salesforce supports Secure Sockets Layer v3 (SSL) and the Transport Layer Security (TLS) protocols. Ciphers must have a key length of at least 128 bits. The remote system must log in using valid credentials in order to obtain a session ID. If the remote system already has a valid session ID, then it can call the API without an explicit login. This is used in call-back patterns covered earlier in this document, where a preceding Salesforce outbound message contained a session ID and record ID for subsequent use. We recommend that clients that call SOAP API cache and reuse the session ID to maximize performance, rather than obtaining a new session ID for each call.
REST API	We recommend that the remote system establish an OAuth trust for authorization. REST calls can then be made on specific resources using HTTP verbs. It's also possible to make

Solution	Security considerations
	REST calls with a valid session ID that might have been obtained by other means (for example, retrieved by calling SOAP API or provided via an outbound message). We recommend that clients that call the REST API cache and reuse the session ID to maximize performance, rather than obtaining a new session ID for each call.
Apex Web service	The same security considerations as SOAP API apply.
Apex REST service	The same security considerations as REST API apply.
Bulk API	The same security considerations as REST API apply.

Sidebars

Timeliness

SOAP API and Apex Web service API are synchronous. The following timeouts apply:

- Session timeout — The session will time out if there's no activity based on the Salesforce organization's session timeout setting.
- Query timeout — Each SOQL query has an individual timeout limit of 120 seconds.

Data Volumes

Data volume considerations depend on which solution and communication type you choose.

Solution	Communication type	Limits
SOAP API or REST API	Synchronous	• *Login*—The login request size is limited to 10 KB or less. • *Create, Update, Delete*—The remote system can create, update, or delete up to 200 records at a time. Multiple calls can be made to process more than a total of 200 records, but each request is limited to 200 records in size. • Query Results Size — By default, the number of rows returned in the query result object (batch size), returned in a `query()` or `queryMore()` call is set to 500. Where the number of rows to be returned exceeds the batch size, use the `queryMore()` API call to iterate through multiple batches. The maximum batch size is 2,000 records. However, additional rules might apply, so for more information, see *Salesforce Limits Quick Reference Guide*.
Bulk API	Synchronous	Bulk API is optimized for importing and exporting large sets of data asynchronously. Bulk API is synchronous when submitting the batch request and associated data. The actual processing of the data occurs asynchronously in the background. For more information about API and batch processing limits, see *Bulk API Limits*. • Up to 2,000 batches can be submitted per rolling 24–hour period. • A batch can contain a maximum of 10,000 records.

Endpoint Capability and Standards Support

The capability and standards support for the endpoint depends on the solution that you choose.

Solution	Endpoint considerations
SOAP API	The remote system must be capable of implementing a client that can call the Salesforce SOAP API, based on a message format predefined by salesforce.com. The remote system (client) must participate in a contract-first implementation where the contract is supplied by Salesforce (for example, Enterprise or Partner WSDL).
REST API	The remote system must be capable of implementing a REST client that invokes Salesforce—defined REST services, and processes the XML or JSON results.
Apex Web service	The remote system must be capable of implementing a client that can invoke SOAP messages of a predefined format, as defined by salesforce.com. The remote system must participate in a code-first implementation, where the contract is supplied by Salesforce after the Apex Web service is implemented. Each Apex Web service has its own WSDL.
Apex REST service	The same endpoint considerations as REST API apply.

State Management

When integrating systems, keys are important for on-going state tracking, for example, if a record gets created in the remote system, in order to support ongoing updates to that record. There are two options:

- Salesforce stores the remote system's primary or unique surrogate key for the remote record.
- The remote system stores the Salesforce unique record ID or some other unique surrogate key.

There are specific considerations for handling integration keys in this synchronous pattern.

Master system	Description
Salesforce	In this scenario, the remote system should store either the Salesforce RecordId or some other unique surrogate key from the record.
Remote system	In this scenario, Salesforce must store a reference to the unique identifier in the remote system. Because the process is synchronous, the key can be provided as part of the same transaction using external ID fields.

Complex Integration Scenarios

Each solution in this pattern has different considerations, when handling complex integration scenarios such as transformation and process orchestration.

Solution	Considerations
SOAP API or REST API	SOAP API and REST API provide for simple transactions on objects. Complex integration scenarios, such as aggregation, orchestration, and transformation, can't be performed in Salesforce. These scenarios will need to be handled by the remote system or middleware, with middleware as the preferred method.
Apex Web service or Apex REST service	Custom Web services can provide for cross-object functionality, custom logic, and more complex transaction support. This solution should be used with care, and you should always

Solution	Considerations
	consider the suitability of middleware for any transformation, orchestration, and error handling logic.

Governor Limits

Due to the multi-tenant nature of the Salesforce platform, there are limits when using the APIs.

Solution	Limits
SOAP API, REST API, and custom Apex APIs	• *API request limits*—Salesforce applies a limit on the number of API calls per 24–hour period. The limit is based on the Salesforce edition type and number of licenses. For example, Unlimited Edition provides 5,000 API requests per Salesforce or Force.com license per 24 hours. For more information, see *Salesforce Limits Quick Reference Guide* • *API query cursor limits*—A user can have up to 10 query cursors open at a time. Otherwise, the oldest of the 10 cursors is released. If the remote application attempts to open the released query cursor, an error results. For example, if sharing integration user credentials, the maximum query cursors need to be considered. Middleware may need to execute requests across multiple users in a "round robin" fashion. • *Call limits*—See Data Volumes sidebar for create, update, and query limits.
Bulk API	See Data Volumes sidebar for more information.

Reliable Messaging

Reliable messaging attempts to resolve the issue of guaranteeing the delivery of a message to a remote system where the individual components themselves might be unreliable. The Salesforce SOAP API and REST API are synchronous and don't provide explicit support for any reliable messaging protocols, per se (for example, WS-ReliableMessaging).

We recommend that the remote system implement a reliable messaging system to ensure that error and timeout scenarios are successfully managed.

Middleware Capabilities

The following table highlights the desirable properties of a middleware system that participates in this pattern:

Property	Mandatory	Desirable	Not required
Event handling		X	
Protocol conversion		X	
Translation and transformation		X	
Queuing and buffering	X		
Synchronous transport protocols	X		
Asynchronous transport protocols			X
Mediation routing		X	
Process choreography and service orchestration		X	
Quality of service (encryption, signing, reliable delivery, and transaction management)	X		

Property	Mandatory	Desirable	Not required
Routing			X
Extract, transform, and load		X (for bulk/batches)	

Example

A printing supplies and services company uses Salesforce as a front-end to create and manage accounts and opportunities. Opportunities on existing accounts are updated with printing usage statistics from the on-premises Printer Management System (PMS), which regularly monitors printers on client sites. Upon creation of an opportunity, an outbound message is sent to the PMS to register the new opportunity. The PMS stores the Salesforce ID (Salesforce is the opportunity record master).

The following constraints apply:

- The PMS is capable of participating in a contract-first integration, where Salesforce provides the contract and the PMS acts as a client (consumer) of the Salesforce service (defined via the Enterprise or Partner WSDL).
- There should be no custom development in Salesforce.

This example is best implemented using the Salesforce SOAP API or REST API.

In Salesforce:

- Download the Enterprise or Partner WSDL and provide it to the remote system.

In the remote system:

- Create a client stub from the Enterprise or Partner WSDL.
- Log in to the API using the integration user's credentials (or the opportunity owner who created the record, assuming the session ID is provided in the initial outbound message).
- Call the update operation on the Salesforce record ID provided in the outbound message and pass in the relevant field updates (usage statistics).

This example demonstrates the following:

- Implementation of a Salesforce synchronous API client (consumer)

  .
- A callback to Salesforce to update a record (aligned with previously covered request/reply outbound patterns).

APPENDICES

Appendix A

Resources—External

1. Hohpe, Gregor, and Bobby Woolf. *Enterprise Integration Patterns*. Boston: Addison-Wesley Professional, 2003.
2. Microsoft Corporation. *Integration Patterns (Patterns & Practices)*. Redmond: Microsoft Press, 2004.
3. IBM Corporation. *Application Integration Patterns*. IBM Corporation, 2004.
4. "Synchronous and asynchronous processes," IBM Corporation, last accessed May 18, 2012, *http://publib.boulder.ibm.com/infocenter/adiehelp/v5r1m1/index.jsp?topic=%2Fcom.ibm.etools.ctc.flow.doc%2Fconcepts%2Fcsynchf.html*.

Appendix B

Resources—Salesforce

Developer Documentation

- *SOAP API Developer's Guide*
- *REST API Developer's Guide*
- *Salesforce Streaming API Developer's Guide*
- *Bulk API Developer's Guide*
- *Apex Code Developer's Guide*
- *Salesforce Object Query Language (SOQL) Reference*
- *Salesforce Object Search Language (SOSL) Reference*
- *Salesforce Limits Quick Reference Guide*

White Papers, Presentations, Websites

- *Best Practices for Deployments with Large Data Volumes*
- *Integration Success on Demand with Salesforce.com*
- *Force.com Connect: Five Paths to Integration Success*
- *Consulting Resource Center: Large Data Volumes*

Index