Informatica Best Practice Guide for Salesforce Wave Integration:

Building a Global View of Top Customers

Company Background

Many companies are investing in top customer programs to accelerate revenue and growth rates, by shifting their approach to their top 20, 30, or 50 customers—from a transactional sales relationship to that of strategic partners with shared success. Such a strategic approach to the customer, however, requires a complete view of the relationship, and for complex enterprises, sophisticated operating models and technology infrastructures mean that getting a comprehensive single view of the relationship with each of their top customers is difficult. This best practice guide will describe one firm's approach to the integration requirements, data issues, and technical approaches to creating a single view of each of its top customers.

Cephalos[1] is a global financial services firm that works with clients of many sizes across a number of global industries – to provide not only investment and credit services, but also a set of strategic and consultative advisory services to increase the success & growth driven by the additional capital.

Cephalos manages global operations along three regional lines: Americas, EMEA, and APAC; each region uses separate systems for managing customer interactions and data. For example, each region has an independent CRM strategy and technology, separate customer master databases, and separate billing and finance systems. In addition, divisions within each region operate in the areas of leasing, inventory finance, bill of trade, and specialty industry finance divisions for capital intensive industries such as Energy and Aviation. These divisions also demonstrate variations in business process and data about customers.

For a global organization such as Cephalos, a distributed approach provides the front office with flexibility and agility. It allows the business units (BUs) to be run in a decentralized fashion. It also allows BUs to address regulatory concerns and regional business practices. Additionally, this strategy lets the BUs address local needs in an agile fashion and fit the offerings to the local market. However, it increases the difficulty, risk, and expense of making major system changes.

The Goal of Global Customer View and the Challenges of Large Enterprise Operating Models

With five divisions and three regions, Cephalos has massive potential in terms of the information it has on hand to better serve its customers and grow the relationships with its top customers. To that end, it established a team of global relationship managers for its top 30 accounts, with the following goals:

- Provide clients with a single global point of contact and consistent level of service
- Enhance the customer experience by accelerating Cephalos's responsiveness
- Support client growth in key countries through existing platforms or partnerships
- Create a best-in-class purchase platform
- Provide superior liquidity, risk transfer, and cash conversion
- Utilize best-in-class operations for its AR/factoring offering

With these goals in mind, Cephalos embarked on a multi-phased project to unlock information stored in its regional silos. Elements of the project included:

- Requirements and options analysis
- Phase 1 proof of concept
- Phase 2 deployment of the Top 30 Program

Summary: Best Practices for Automating Cloud Integration for Phase 2 Deployment of the Top 30 Program

Before beginning implementation, Cephalos needs to take a holistic view of what applications and databases have relevant data necessary to include with each Top 30 Account in Salesforce. The main objective is to create 360-degree dashboards for each top customer. These data-driven dashboards require continuously integrating, cleaning, and relating large volumes of data from multiple sources (such as ERP system and cloud apps) in a batch integration (automated ETL) approach. Before batch integration can be set up for production data, the company needs to prepare prototypes of the required data sets with business rules ensuring that the data is clean and formatted and in the views needed for analysis in Salesforce Wave. Cephalos uses Informatica® Rev, with its spreadsheet-style interface, for Data Preparation and allows business users to easily prepare data sets for dashboards and analysis. Additionally, Data Quality and Profiling rules need to be applied to each Account field to standardize conflicting fields and values across different Cephalos subsidiaries.

Once those rules are applied, Cephalos needs to use Hierarchy Management together with Deduplication to establish the "Golden Record" (or single source of truth) for each Account as well as the master org from which Analytics Cloud will draw information. Address validation and enrichment enable greater localization for the various geographies, which will help Cephalos better target product offers. Informatica Cloud's Designer collapses these several transformations and functions onto a Design Canvas, which makes it easier to create data sets in Analytics Cloud. The time saved from creating these data sets cascades forward to the lenses that are created for the customer satisfaction scores.

Introduction – Phase 2 Deployment of Top 30 Customer Program

Cephalos uses Salesforce across its five divisions and three regions to scale its operations and achieve growth in new markets. The next phase of growth relies on a more focused approach involving its global Top 30 customers and uncovering areas of opportunity to expand Cephalos's product offerings within each account. To do so, Cephalos needs to provide analytical capabilities to its VP of Sales and VP of Sales Operations within each division. These stakeholders need 360-degree dashboards of each top customer showing total revenue spend, length of relationship, most profitable geographies, and product mix.

For the Phase 2 overall objective, Salesforce Analytics Cloud needs to provide a global customer data view that consolidates views of key Opportunity performance for Cephalos with clients across disparate industries using automated batch integration (also often referred to as "automated ETL"). Automating the batch integration and data stewardship of the analytics process requires a complementary cloud integration solution that aligns with the speed to market and ease of use of Wave. Cephalos activated the Informatica Cloud solution portfolio to automate the data consolidation and stewardship process at scale as Phase 2 requires. The Informatica Cloud portfolio includes batch Data Integration, real time Application Integration, Cloud Master Data Management (Cloud Customer 360) and Cloud Test Data Management (Cloud TDM) which together constitute a set of end-to-end integration and data management services that enable Salesforce Wave, the Analytics Cloud.

The steps to achieving a global customer data view include:

1. Data Preparation to prototype dashboards and uncover issues with the data

    a. Extract a sample of data to analyze

    b. Uncover issues such as inconsistencies and duplicates

    c. Prepare the data to fit into the right format for the dashboards

    d. Track these changes to be shared for the next steps

2. Integration of the data from multiple systems into Analytics Cloud

    a. Extracting product and revenue data from ERP systems

    b. Associating this data with each of the Top 30 accounts

    c. Ensuring that a timeline of revenue sold into each account is visible

3. Data Quality: Profiling, Cleansing, Enrichment and De-duplication

    a. Data Profiling helps determine the quality of the data in the source systems. At this stage, inconsistencies, fragmentation, and any issues with standardization can be highlighted.

    b. Cleansing helps to rectify issues that profiling has uncovered. Cleansing includes the validation of address and contact details as well as the opportunity to use purchased data lists to enrich the data.

    c. De-duplication involves identifying potential duplicate Accounts, Contacts, or Leads and then automating merging them or facilitating a safe and wizard-based manual merge.

4. Multiple Org Coordination

    a. Each Salesforce org will have its own business processes and details. Coordinating these environments ensures that the right details are making up the master record as well as ensuring there is clarity on which BU/Org is the source of which Account/Contact/Lead record.

5. Hierarchies and The Master Record

    a. Finally, once the data has been cleansed and de-duplicated, the next stage is mastering the record to create a Golden Record in the Cloud Analytics org. This also includes highlighting necessary connections between subsidiary accounts that relate to Cephalos' Company and/or Contacts.

1. Data Preparation

The initial steps before starting the automated ETL process are to analyze and prepare the data. A business analyst can use Informatica Rev and its Excel-like interface to analyze a sample of the combined data from the multiple sources for initial prototypes of the reports and dashboards. This step helps ensure that the business is involved upfront in how the data will appear in the Analytics dashboards. At this stage, Cephalos can perform trial runs of different views such as the financial history or wins and losses to see if the results look correct in the Analytics org.

This data preparation task also allows for uncovering and profiling issues within the data.

i. Data Profiling

Profiling is the first step to identifying the irregularities in the data that could potentially cause issues with reporting later. Key aspects of this step include:

1. Identifying issues of standardization—Are addresses and other details in the same format?

2. Identifying fragmentations in the data—Are the records missing crucial information that can be enriched through purchased data lists or through combining data from other systems?

3. Is the data valid and up to date? Which source system has the most up-to-date information on the record? Are contact and address details correct?

Because this best practice guide focuses on the automated ETL process, the focus here on profiling and cleansing concerns issues that can be resolved through automated cleansing.

Once the profiling is complete, cleansing, validating, and enriching the data can begin. The following screenshot (Figure 1) shows Informatica Rev in action.

Figure 1. Informatica Rev

2. Integrating Data from Multiple Sources to Provide 360-degree Customer Views

The first step toward building the Analytics Cloud capabilities is to identify all applications, databases, and other data sources that would contain the data necessary to build the data sets, lenses, and dashboards to serve the two primary stakeholders—the VP of Sales and the VP of Sales Ops—so that they have a 360-degree view of the customer. Once the data sources have been identified, pre-built connectors from Informatica Cloud parse the necessary APIs transparently and push the data into an Analytics Cloud data set. Informatica Cloud Data Integration has more than 100 native connectors to applications and databases such as Marketo, Sales Cloud, Service Cloud, Marketing Cloud, Salesforce Analytics Cloud, Workday, NetSuite, SAP, Oracle E-Business, SQL Server, MySQL, and Hadoop to support batch and real-time integration use cases.

Cephalos has a mix of recurring and fixed revenue financial products whose information is stored in various ERP systems. Given the size of the company, different business units use different ERP systems. The corporate ERP system runs on SAP; some locations use Oracle E-Business while others have turned to a SaaS solution such as NetSuite. These ERP systems contain master identification numbers for every product SKU and its associated list price and cost, every customer account and its subsidiaries, as well as every product ever sold to each customer.

The overall objectives of integrating data into Analytics Cloud include:

1. Associating revenue for each product line with each of the Top 30 customers

2. Merging data from various disparate ERP systems across all of Cephalos's subsidiaries into a single data store

3. Achieving profitability information to best align product offers to the Top 30 customers

To achieve these overall business objectives, this ERP data needs to be pushed into a series of tables in an Analytics Cloud data set and associated with the appropriate Salesforce Account. Informatica Cloud Data Integration can aggregate the ERP data and associate it with the appropriate Account when moving it into the data set. This data set furnishes the initial data necessary to form the 360-degree dashboards of each customer showing total revenue spend, length of relationship, most profitable geographies, and product mix. Informatica Cloud Data Integration has a batch engine that can move millions of rows of data rapidly. Informatica Cloud also has a library of advanced data transformations that can be applied quickly and easily.

Once the ETL automation process has been completed, the next step is managing the data that has been transferred into the Salesforce Analytics org.

3. Data Quality

The challenge facing Cephalos when consolidating multiple CRM systems and different BUs is that each data source will have different business processes and thus each Account or Contact in each system may have different attributes and potentially conflicting or duplicated information. To maintain a single dashboard that will bring together data from the multiple CRM and ERP systems, the differences first need to be identified and then standardized.

i. Cleansing, Validation, and Enrichment

After data integration, cleansing and enrichment are necessary to minimize disruption to the business units.

Informatica Cloud's Contact and Address Validation solution ensures that each Account/Contact has the correct contact details according to a database of global postal addresses and a domain and first-name email address validation. Ensuring that the correct address and contact details are in place is crucial to creating the relationships that will eventually make up a top Account with the various subsidiaries that relate to it. This step will also help with uncovering any geographic patterns with regards to opportunity management and sales to help ensure that resources are used effectively.

Enrichment from third-party data sources such as data.com can ensure that fields are populated to enable optimal matching at the de-duplication level.

ii. De-duplication

Step two of data quality involves the identification and removal of duplicates that can occur in the individual systems. With multiple BUs and regions feeding into the one global source, duplicates across the systems are likely.

Informatica Cloud's Master Data Management solution uses field matching to compare records and score them based on the likelihood that they are duplicates. This step factors in such elements as Address, Zipcode, Phone numbers, and Company name. Because some fields in an Account may be unique to each BU, in this process it's best to compare the common fields that they all share.

Records with a high matching score will then be identified as duplicates and a mass auto merge can be run in the initial implementation (as shown below in Figure 2). This process will then act as a preventative measure when new records are synchronized into the global org matching them against existing records before creating a new account.

Figure 2. Dedupe and Cleanse with Informatica Cloud Customer 360

4. Coordination of Multiple Orgs

Many businesses including Cephalos have multiple instances of Salesforce that each exist in their own right. For instance, different product teams or business units may choose to maintain these siloed instances to maintain their unique business processes or buyer.

There are different approaches to managing multiple orgs; the approach Cephalos chose was to use a global reporting hub (the Salesforce Analytics Cloud org; as shown in Figure 3).
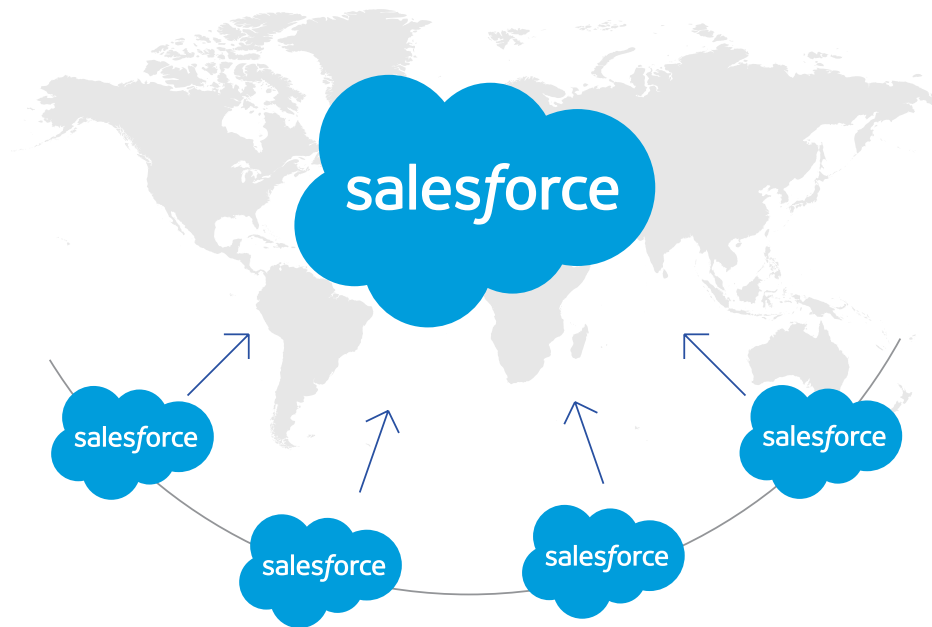

Figure 3. Managing Multiple Salesforce orgs

The global reporting hub approach is a periodical upload of data from the other CRM systems into one global org, which has Salesforce Wave Analytics sitting on top. With this approach, the global org acts as a master hub for all the customer data.

Using Informatica Cloud integration to upsert[2] the data from the multiple orgs means the dashboards will always be working with the most current version of the truth. When not all of the data needs to be migrated to the Salesforce Analytics org, Cephalos can merely reference the data via a lookup so that the dashboards in the primary org always have fresh data.

[2] definition of "Upsert": Insert data where there is none, update data to the most recent value where data is already present.

5. Hierarchies and the Master Record

Benefit: Single trusted view of the customer. Sales have the details they need.
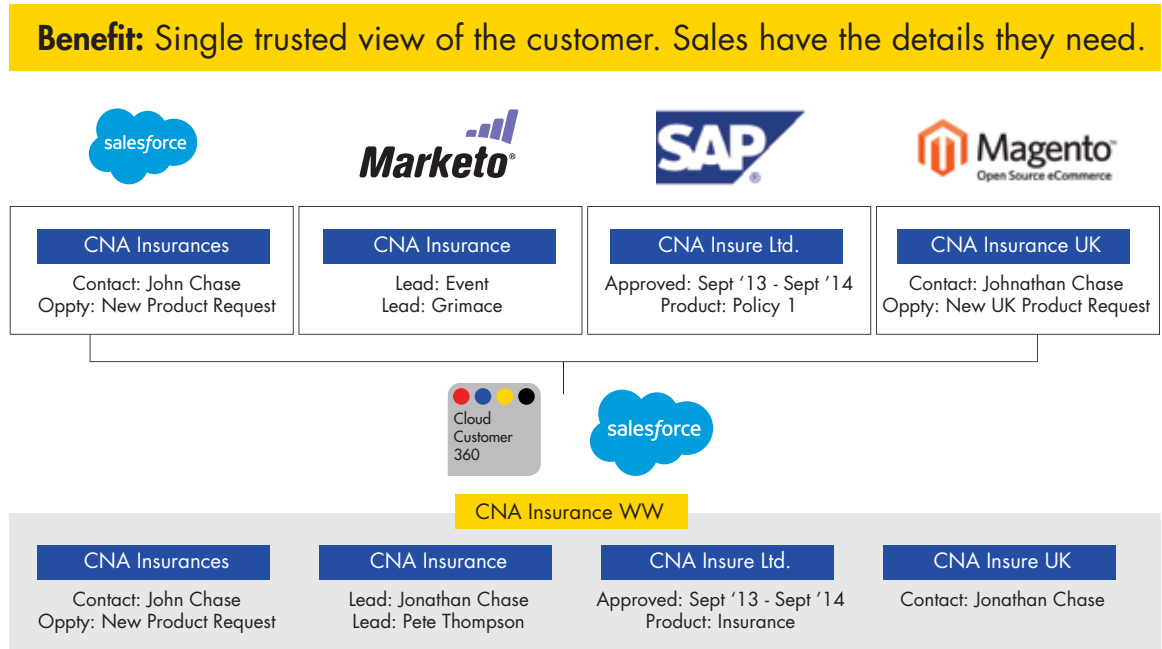

Figure 4. Managing customer hierarchies with Informatica Cloud Customer 360

New fields can be created to track historical transactions to ensure that key information about wins and losses and when they're occurring can be tracked over time. Once cleansing of the data has occurred and duplicates have been identified and merged, the next stage is to create the master accounts.

With 30 key accounts as the core focus, it's easy to forget that each of these accounts may have multiple subsidiaries, locations, or simply multiple versions of the truth. Aligning with each BU's key products, interactions, buyers, and possibly locations (an activity done at the profiling stage) meant that Cephalos could decide at its field level, using Informatica's Cloud Master Data Management (Customer 360) solution, if there were conflicts between CRM systems or other back-office systems on certain key fields such as Product, Address, Finances, and Contacts. Cephalos could also select which system should be the master "golden record" based on business rules that prioritize the most appropriate system. For example, for tracking conversion ratio trends, the marketing data may be the best to follow along with Lead conversion. For historical financial references, using ERP system data is often typical because these systems tend to be more fixed than CRM data. Informatica Cloud Customer 360 is then able to automate these rules on an ongoing basis, ensuring that the right data makes it to the live Account record.

When multiple versions of an Account are required because it represents a different sister company/subsidiary of an account, then the Informatica Cloud Master Data Management (Cloud Customer 360) solution can automate a hierarchical connection between the accounts to help streamline reporting. This connection created through automated hierarchies will then carry over when building the dashboard and reports to enable white space analysis and rolling up total revenues across BUs and regions by a global account (see Figure 5).

Figure 5. Multi-dimensional account hierarchies

Once the data has been prepped, integrated, cleansed, and then mastered, Cephalos can ensure that it has the visibility into its global top 30 accounts and know that this view is trusted and up to date. These cleansing and synchronizing processes will then continue to run automatically in the background with no disruption to the source systems so that getting a current view of the business is a fast and error-free process.

About Informatica

Informatica Corporation (Nasdaq:INFA) is the world's number one independent provider of data integration software. Organizations around the world rely on Informatica to realize their information potential and drive top business imperatives. Informatica Vibe, the industry's first and only embeddable virtual data machine (VDM), powers the unique "Map Once. Deploy Anywhere." capabilities of the Informatica Platform. Worldwide, over 5,500 enterprises depend on Informatica to fully leverage their information assets from devices to mobile to social to big data residing on-premise, in the Cloud and across social networks. For more information, call +1 650-385-5000 (1-800-653-3871 in the U.S.), or visit www.informatica.com.

For a free 30 day trial of Informatica Cloud, visit www.informaticacloud.com.

ınformatıca

Put potential to work.™

IN09_0215_02834