Modeling Customer Relationships

A Flexible, Integrated Architecture
Enables Customer-Centric Marketing

SEQUENT

Abstract

This white paper draws upon the lessons learned by Sequent Computer Systems in implementing large-scale technology platforms to support Customer Relationship Management (CRM) strategies within major organizations worldwide. It is assumed that the reader is familiar with the strategic direction of most customer-focused organizations and understands the cyclical and repeatable nature of a technology-driven marketing strategy. Interested readers are likely to be concerned with how they might model customer relationships in a way which will support the transition from present product-focused views of marketing intelligence to more useful (and profitable)

customer-focused views. This paper provides the data architect or modeler with a generic template for modeling customer data. This approach is not product or technology specific but does provide a flexible data architecture for integrating the various technology components that use data to drive marketing. This paper also highlights commonly faced problems that occur when modeling customer data.

Editor: David Puckey
Sequent Computer Systems
UK Professional Services

Introduction

The strategic importance of managing customer relationships both drives and is driven by technology. In particular, this applies to data and the increasingly sophisticated and useful ways in which data is used to model relationships and to drive contact strategies. At the core of any technology enabler for CRM is the customer database. The customer database represents the data hub that integrates the various statistical modeling, campaign management, contact history and response tracking components of the marketing campaign lifecycle. This is true whether the database is used for the execution of marketing strategies (e.g., generates mailing lists), or whether it exists purely as an analysis engine that passes contact strategies and information to a separate customer interaction platform for execution (e.g., customer call centers).

The technology layer and its integration with emerging business processes is therefore key to the successful implementation of a data-driven Customer Relationship Management strategy. This paper describes, in a generic way, an approach that Sequent has successfully used to ensure that the core data structures in the CRM technology layer support the integration of the various components of the marketing process and reduce the time required to design and execute a campaign. This approach enables the creation of a complete model of customer relationships over time.

The Evolution of the Customer Database

Current approaches to the design of the customer database fall broadly into two camps. The first—the flat earth view of the world—hails from the glory days of target marketing in the late 1980s. This approach, which is popular with list providers and bureau operations, utilizes the concept of the customer file or list. Such a list tends to offer a current snapshot of the customer or prospect base and is often the product of much tortuous cleansing, de-duplication and point-in-time segmentation. This approach makes it difficult to analyze the ups-and-downs of an organization's relationship with a customer over time due to its current snapshot view of the customer and prospect base. Further, it typically lacks the transaction-level detail and promotional history needed to model customer behavior.

Evolution of Customer Database

	Campaign Management	Customer Relationship Management
contact horizon	one-shot	sequence
output	offer	information
systems	mail / phone	touchpoint
execution	manual	scheduled
departments	marketing	front-office
data types	purchases	contacts
update interval	monthly	daily
reaction time	billing cycle	transaction
goal	reduce waste	add revenue

Source: Raab and Associates

Figure 1: Customer databases evolve with integration of technology and business processes

The second approach has evolved from the data warehousing movement and Sequent's experience in helping hundreds of organizations design and implement data warehouses. Sequent has developed a mature methodology for delivering rapid business benefit by integrating sophisticated analytical tools with subject-oriented and time-consistent central databases. Such systems typically concentrate on the delivery of business intelligence and are generally not designed to plug directly into an organization's day-to-day operations. However, the modeling techniques employed by Sequent for the delivery of successful data warehousing projects represent a radical shift in emphasis from both flat earth views of data and the microscopic views of data used in online transaction processing (OLTP) systems. Sequent's dimensional view of data provides the optimum combination of analysis of facts over time and high system performance when dealing with large data volumes.

Sequent's approach to successfully delivering large-scale technology platforms to support CRM strategies uses the best attributes from both of the previous approaches. This approach incorporates the maximum degree of analytical flexibility for the marketer and marketing analyst with the efficient scoring, segmentation and extraction of data to execute marketing campaigns or contact strategies. It also places the customer or prospect at the center of the model and seeks to model all facets of a relationship with that customer over the known lifetime of the relationship. Sequent's approach to the design of customer databases is not list based and is not designed to simply support ad-hoc, point-in-time marketing solutions. Rather, the objective is to give the marketer true insight into the variability of his relationship with a customer or customer segment over time and to deliver seamless integration with the widest possible choice of campaign management and statistical modeling tools available.

Elements of a Customer Relationship Management Database

There are a number of required features of a CRM database that the architect must integrate in order to support the marketing lifecycle. These are (in no particular order):

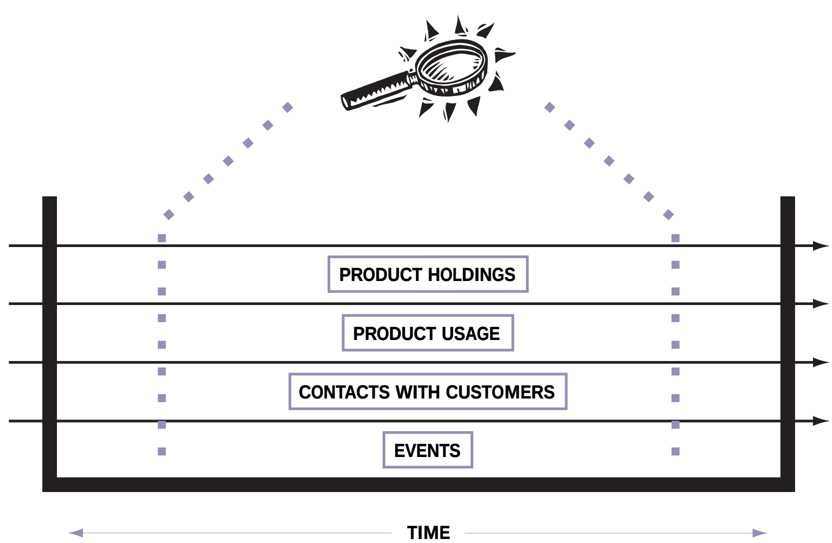

Figure 2: Typical facets of a customer relationship that need to be tracked over time

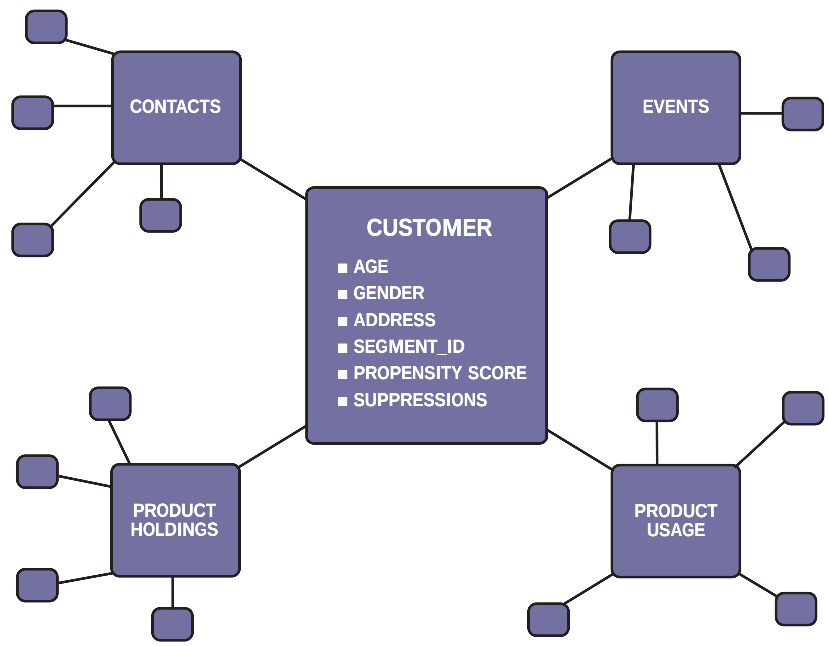

Figure 3: Customer focus is key. Each facet of a relationship may be treated as an island of analysis, linked centrally to the customer

- Customer or prospect focus
- All relevant facets of the relationship over time
- Integration of external prospect lists
- Integration of external data classifications
- Integration of external data enrichment
- Ability to directly score the database and segment the database many times
- Ability to evaluate different campaigns and treatment strategies over time and across millions of transactions and customers
- Campaign management, prioritization, etc.
- Ability to predict future customer behavior based on past behavior

It is not possible to achieve all of the above features using either a flat file approach or a standard data warehousing approach alone.

Taking a Lifetime View of the Customer

In order to fully realize a CRM strategy, the marketer must have information that enables him to take a lifetime view of the relationship. A relationship is most usefully defined as the starting point at which the organization has an initial interaction with a prospect. This relationship then needs to be tracked as the prospect is encouraged to climb the loyalty ladder from prospect to customer and eventually to highly valued customer. The marketer needs to see and understand past events, contacts and purchase information in order to assess the current and future profitability of the relationship. The commonly used marketing analysis of recency, frequency and monetary value of transactions indicates some of the facets of the relationship that should be tracked.

In Sequent's experience of facilitating client workshops to establish the business requirements for a CRM solution, four relationship facets appear common to most organizations. These facets are:

Product Holding–What products has a customer purchased and what products do they currently hold ?

Product Usage–How has the customer used that product? For example, can an increase in credit card usage be attributed to some prior interaction with the customer or some promotional activity?

Contacts–What has the organization's interaction with the customer been over time and what were the outcomes?

Events–What other events have occurred, either within the life of the customer (e.g., marriage) or externally to the relationship (e.g., competitor activity)?

Each of these facets may be treated by the modeler as an island of analysis linked centrally to an individual customer at a given point in time. For example, the marketer may take a point-in-time view of the relationship, a view over time or make prescient predictions for the future. Information about these four facets of a customer relationship enable the marketer to answer questions such as: *How many customers have bought product X? How many customers display a repeatable purchasing pattern? How often have I contacted this customer and when? Who are my most profitable customers? What events or contacts occurred prior to customer defection?*

The approach taken by Sequent to support this kind of questioning is to place a customer table at the center of the model and to surround it with satellite dimensional schema (star schema) representing each facet of the relationship to be modeled. Modeling the facets of the relationship dimensionally allows who, what, when, where style analysis. For example: *Which segment bought which products and what contacts preceded which purchase? Where do the contacts live, and how do they like to be addressed?*

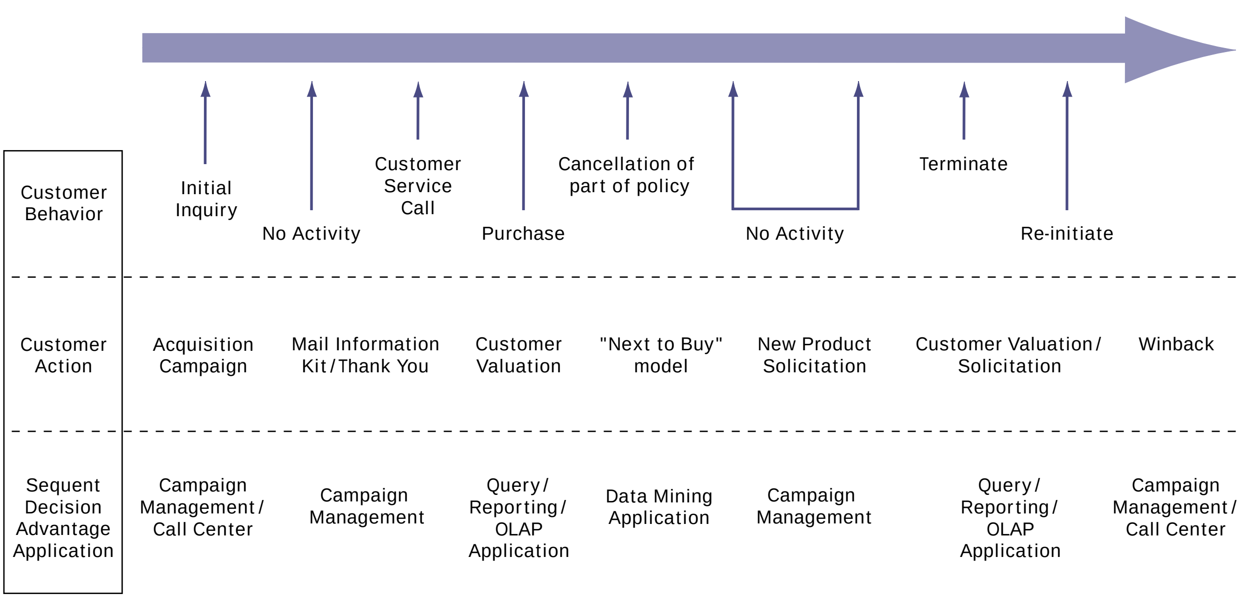

Figure 4: Example—The Customer-Centric Model at an Insurance Company

The customer-centric nature of the model also lends itself well to the prudent de-normalization of often-used facts, such as disposable income estimates, onto the customer table and helps facilitate the efficient extraction of contact lists and integration with statistical modeling tools, such as SAS or Unica. The customer-centric model also supports very well the iterative nature of the marketer's questioning, such as: *How many customers hold product Y? Which of those customers are profitable? Which of those customers did I contact last week and which of them complained about the contact?* It is also possible to assess what behavioral changes are exhibited as a result of identifiable interactions with the customer. Once the marketer has exhausted his questioning, which helps refine the contact list names, addresses and salutations may be simply extracted from the customer table using the relevant keys. Current suppression indicators and propensity scores may also be stored against the central customer record, allowing the possible automation of standard hygiene filtering.

The Customer-Centric Model at an Insurance Company

To see how this model might work, take the example of an insurance business. The firm's relationship with Mr. Jones begins when he makes an initial inquiry about health insurance via the organization's call center. This initial inquiry is the result of a press advertising campaign that reached Mr. Jones; this fact is recorded.

In response to his interest in the company's health insurance offering, the insurance business sends Mr. Jones an information pack. This step is also captured and recorded in the database. At this point, Mr. Jones does not have a product holding, but his name and address and contact records exist within the database.

Mr. Jones does not respond to the receipt of the information pack, and after three months the marketer plans a campaign targeted at Mr. Jones and all the other Mr. Joneses who have interacted with the organization but not purchased any products in the last three months.

In this case, a query can be run against the database asking, *Who has contacted us in the last three months with a contact type of inquiry?* This query will generate a list of keys into the customer or prospect table, which, without further refinement, could be used to generate a contact list. However, it is more likely that the marketer's questioning will continue further—*How many of these customers or prospects were sent an information pack?* The result set from this query will be matched against the result set from the last query to further refine the list of keys. This process may be further refined by asking, *How many people in this list do not have a product holding?* Once the marketer has completed his refinement of the list, it is a simple, and highly performant, exercise to take the resulting list of keys and extract the name, address, salutation data, etc. from the central customer or prospect table and perform further filtering based on suppressions on the customer table or assigning customers to campaign cells for different treatments based on segmentation keys on the customer record. Once the contact list is finalized, the customer keys are used to populate the contact table and to record the fact of the outbound contact. By storing all of this data in a centralized relational database management system (RDBMS), it is a relatively simple matter to make this data available to sophisticated campaign management tools and statistical modeling tools. These tools interface easily with an open RDBMS, such as Oracle, and almost without exception, such tools feature native connectivity options.

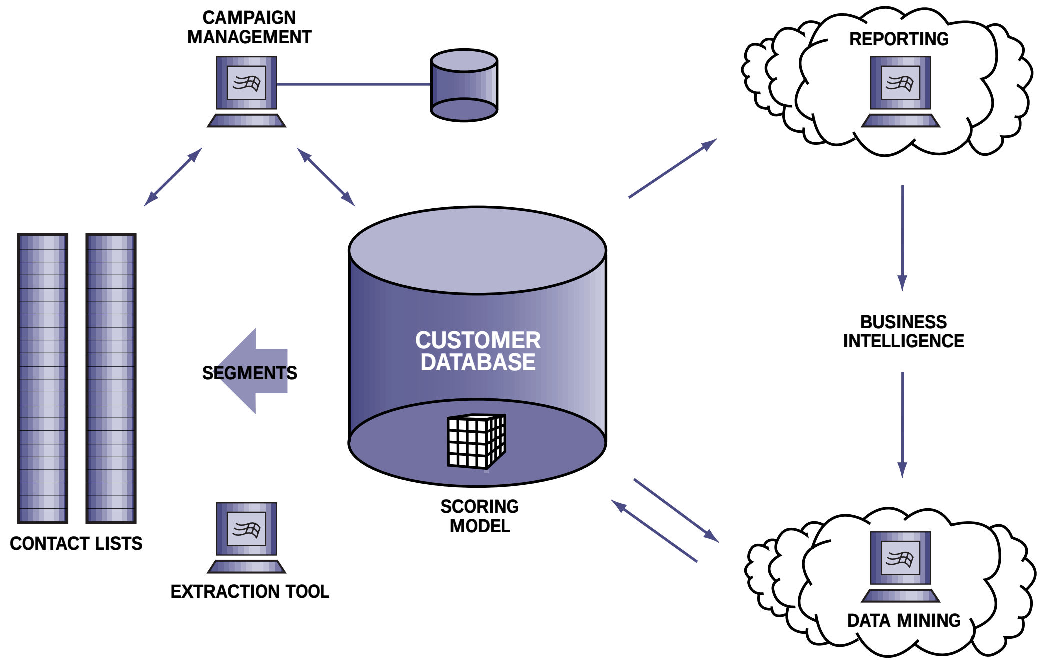

Figure 5: An integrated architecture reduces the marketing cycle

Those readers familiar with the processing dynamics of most RDBMS will immediately spot a major dependency of this model—the various software components deployed to support the marketing lifecycle must allow the generation of interim result sets. This is absolutely crucial in order to support the marketer's analytical processes as he constantly shrinks and expands potential target lists, possibly to generate the required list size to match a budget allocation. Already, a number of tools vendors are acutely in-tune with the mindset and thought processes of the modern marketer.

Integrated Infrastructure Supports Marketing Process

In the past, database marketing solutions often focused on individual user communities participating in the overall marketing process. While this focus has managed to hit the sweet spots of these often isolated communities (e.g., marketing analysis or campaign management) and to temporarily satisfy parochial needs, it has left a troublesome legacy for the integrator of the technology layer who seeks to accelerate the marketing cycle, empower the marketer and reduce the marketing department's dependency on highly skilled and expensive (and often obstructive) database experts. Such function-focused solutions have ensured that the walls that block the implementation of a virtuous circle of continuous improvement in the marketing process remain solid. The proliferation of file formats, APIs and unnecessary processing layers needed to integrate these elements have delivered a full employment charter for those who wrangle with the complexity of the technology layer at the expense of marketing responsiveness and creativity.

Sequent's solution to such technical anarchy is to focus firmly on a technological infrastructure that supports and integrates the overall marketing process, and underpins the progressive development of a relationship management strategy. The use of a centralized relational database and open systems to manage customer data, contact history and relationship history allows the easy integration, at the data level, of the various technologies deployed at different stages in the marketing process. Analysts' models may be stored alongside the actual data, and scoring and segmentation keys can be made directly available to campaign management and campaign scheduling software. The automation of routine communications is simplified and database triggers can be utilized to make marketing more event driven.

Typical Data Modeling Challenges

This section details some of the data modeling challenges, which, in Sequent's experience, are common across a number of industries and organizations.

Householding

The grouping of individuals by household or relationship patterns is often a difficult process in product-focused legacy systems. These systems often have great difficulty in even identifying the individual responsible for purchasing a given product. The benefits of groupings for the relationship marketer are many:

- Avoidance of unnecessary duplicate contacts per household
- Understanding loyalty patterns among relationship groups
- Identification of cross-sell and up-sell opportunities (e.g., family policies, etc.)
- Identification of significant life events (coming of age, birth, marriage, etc.)
- Analysis of geodemographic data by household

Multiple households can be problematic for both the marketer and the system designer. Individual customers may have multiple addresses, each of which is related to the customer via the product holding. For example, Mr. Jones has a main residence in the city and a weekend retreat by the coast. Mr. Jones has a household insurance policy for each address. An insurance marketer may wish to sell Mr. Jones a life insurance policy. However, for the modeler, a household is just a simple grouping of individuals. Specific business questions must be answered in order to track the household movements of individuals. The difficulty is in the actual identification of a household—particularly in high-density urban residential areas or areas with a highly transient population.

There are several approaches to handling customer householding, de-duping and geocoding challenges. These include:

- Service Bureau operations
- Integrating specialized software tools to perform this function on a regular basis (this also requires process integration for proper and effective handling)

A number of marketing data processing bureau services perform household identification, based on, for example, electoral register information, etc. However, such matching is never 100 percent accurate.

Products Held by Groups of People

Certain types of products, for example joint bank accounts, introduce a many-to-many relationship between product holdings and persons. This fact, if modeled literally, can cause performance problems in the database and confuse campaign management and extraction tools seeking to identify a single prospect. This is particularly true in cases where organizations are transitioning to a customer-focused marketing strategy yet still require the ability to market in the interim period based on product holding attributes. This situation is common in large businesses that cannot possibly switch from a product to a customer focus overnight. The only answer to this problem is a business one. Identifying a primary marketing contact for a product holding can simplify the problem in some cases.

Person Matching

Another key challenge for the designer of a CRM database is the identification of individuals. Often, seemingly multiple individuals on the database are in fact the same person, albeit at a different point-in-time, or with a different product holding, or at a different address. Organizations with multiple operational systems serving multiple customer touch points often find that the non-uniformity of input validation across these systems leads to situations where Mr. John Jones, Mr. J. Jones and Mr. J. B. Jones at the same address could perhaps be one, two or three actual people. This problem is further exacerbated when external prospect lists are brought into the database. Once again, the modeler can incorporate a simple grouping of people within the database design but the problem is identifying the actual grouping. In some cases, Sequent has allowed a “degree of confidence” value to be assigned to the grouping record to provide the marketer with a coefficient that validates assumptions. The business rules for deriving this coefficient clearly evolve over time, and can result in the creation of specific profiling questions targeted to specific customers during interactions.

As with householding, some marketing data providers can perform unique person identification based on postal lists, real estate listings, electoral rolls, and other data. This identification activity can be cumbersome as it involves exporting and re-importing data periodically. If the grouping of seemingly multiple individuals into one is handled as a grouping table, the impact on, for example, referential integrity within the database can be minimized. However, this kind of group can also make the model more complex—with a possible impact on performance.

Unfortunately, there are no magic cures for the problem of person matching, and the database modeler should be wary of the purveyors of such cures.

Classing and Banding

A number of marketing database designs use fields such as “date of birth” or “age” on the customer record. Though there is a clear use for such fields, marketers rarely wish to contact people who are, for example, 51 or 23 years of age. Usually, the marketer wants to target people aged between 25 and 35 or those who are past retirement age. Such targeting calls for some sort of banding of customers to reduce wasted processing and simplify the process for the marketer.

Age is not the only candidate attribute for banding. The modeler should seek to understand other candidates and include these in the model.

Regularly Used Measures

Initially, and over time, the modeler of the customer database should seek to identify those frequently asked marketing questions, such as: *Who earns more than $20,000? Who has made more than four insurance claims in the last period, etc.?* It makes little sense to have multiple marketing campaign designers all scanning the product usage table over and over again. This can be avoided by denormalizing regularly used measures directly onto the customer or prospect record.

Suppressions

Most organizations are able to identify a number of standard reasons for suppressing marketing communications. Suppressions can range from blanket "do not communicate at all" indicators to "do not market a specific product" to this individual. These suppressions should be held directly on the customer or prospect record to enable swift and easy filtering of targets.

Summary

While both flat file and standard data warehousing approaches to the customer database will allow analysis of customers and the selection of target lists, neither approach will, on its own, support the management of customer relationships over time. Likewise, neither will integrate all components of the marketing process in the most efficient way.

The template presented in this paper may form the basis of the data architect or analyst's initial attempts to define data structures, which will support both of the above objectives. This template reflects the work Sequent has done with a number of major organizations to support their database marketing activities and to drive the strategic implementation of Customer Relationship Management at both the business and the systems levels.

CRM is an emerging strategy and as such requires a fresh approach to systems design, along with the flexibility to accommodate unexpected change. Many piecemeal or point solutions in the market fail to take an integrated view of the entire marketing lifecycle and focus only on data structures to support their own specific components of that lifecycle. As CRM matures as an operational reality, it is imperative that organizations have an integrated view of business processes and data. Failure to take an integrated view of requirements will lead to significant effort and cost reengineering the organization's marketing databases—sometimes comprising many terabytes of data.

Corporate headquarters:
American headquarters:
Sequent Computer Systems, Inc.
15450 SW Koll Parkway
Beaverton, Oregon 97006-6063
(503) 626-5700 or (800) 257-9044
www.sequent.com

European headquarters:
Sequent Computer Systems, Ltd.
Sequent House
Unit 3, Weybridge Business Park
Addlestone Road
Weybridge, Surrey KT15 2UF
England
+44 (0) 1932 851111

Asia/Pacific headquarters:
Sequent Computer Systems
(Singapore) Pte Ltd.
80 Robinson Road, #18-03
Singapore 068898
+65 223 5455

With offices in:
Australia, Austria, Czech Republic, France, Germany, Hong Kong, India, Indonesia, Italy, Japan, Korea, Malaysia, The Netherlands, New Zealand, Philippines, Poland, Russia, Singapore, United Kingdom, and United States.

With distribution partners in:
Bahrain, Brunei, Croatia, Czech Republic, Egypt, Greece, Hong Kong, Indonesia, Japan, Korea, Kuwait, Malaysia, Mexico, Oman, People's Republic of China, Philippines, Russia, Saudi Arabia, Slovenia, South Africa, Sri Lanka, Taiwan, Thailand, United Arab Emirates, and Yugoslavia / Serbia.

CP-1340 12/98

SEQUENT