Complex-Event Processing Poised for Growth

Neal Leavitt

Companies today often face a flood of data generated by different information systems. This data includes the rise and fall of stock prices during the day, the movement of tagged luggage through an airport's baggage-handling system, the usage of a credit card, and ongoing activity on a network.

Businesses have been doing near-real-time event analysis processing for decades, noted University of California, Berkeley, professor Michael Franklin, cofounder and chief technical officer of event-processing vendor Truviso. "But until recently," he explained, "most of these systems had to be either hand-coded or were specific to a particular [application type]."

General event-analysis software typically doesn't function in real time. Instead, organizations must spend time analyzing past complex events only to yield static results relevant to just the data studied up to that point.

This makes it challenging for them to use the information to make on-the-spot decisions in matters such as stock trading, fraud detection, inventory tracking, network-performance monitoring, and airline-baggage handling.

"Basically, the faster a company analyzes its data streams and responds to the analysis results, the more competitive it is," said Christoph Heinz, head of marketing and sales for event-processing vendor RTM Realtime Monitoring.

"Traditional databases are not optimized for real-time analysis of continuous streams and complex events," he noted.

Furthermore, added Chris Bradley, chief technical officer of event-processing vendor Agent Logic, "The analysis challenge is exacerbated by combining databases with RSS feeds, message queues, and real-time sensors."

To solve the problem, businesses are increasingly turning to complex-event processing technology.

With CEP, said Franklin, "What's new is the availability of general-purpose platforms that provide on-the-fly analysis of information across a spectrum of potential applications."

As Figure 1 shows, CEP systems collect data from numerous sources about raw events within a company's operations on an ongoing basis and use algorithms and rules to determine in real time the interconnected trends and patterns that combine them into *complex events*. They then send the findings to the appropriate business user.

"A series of queries could look for a pattern of raw events that represents an opportunity, problem, or threat, and pass that alert to a downstream application or person," said John Morrell, vice president of product marketing for CEP vendor Aleri.

Major benefits include the increased ability of companies to make quick use of marketplace and other information they collect to enable faster response to ongoing events, said Heinz. For example, investment banks are using CEP for real-time liquidity management.

Also, credit card companies use the technology to analyze multiple customer transactions and recognize complex patterns that indicate fraudulent usage, thereby allowing them to close accounts quickly, noted Mary Knox, research director for banking and investment services for Gartner Inc., a market-research firm.

California Institute of Technology professor Kanianthra M. Chandy said branches of the US military and the US Department of Homeland Security are using CEP to analyze engineering, geological, space, and other types of data.

Nonetheless, the CEP market is still just emerging, said Stanford University Professor Emeritus David Luckham, an events-processing pioneer. Moreover, he noted, CEP faces a number of obstacles to commercial success.

PROCESSING CEP

In 1995, Luckham was working on the Rapide research project, studying discrete-event systems for simulating the behavior of proposed hardware designs in order to find errors and weaknesses.

0018-9162/09/$25.00 © 2009 IEEE Published by the IEEE Computer Society

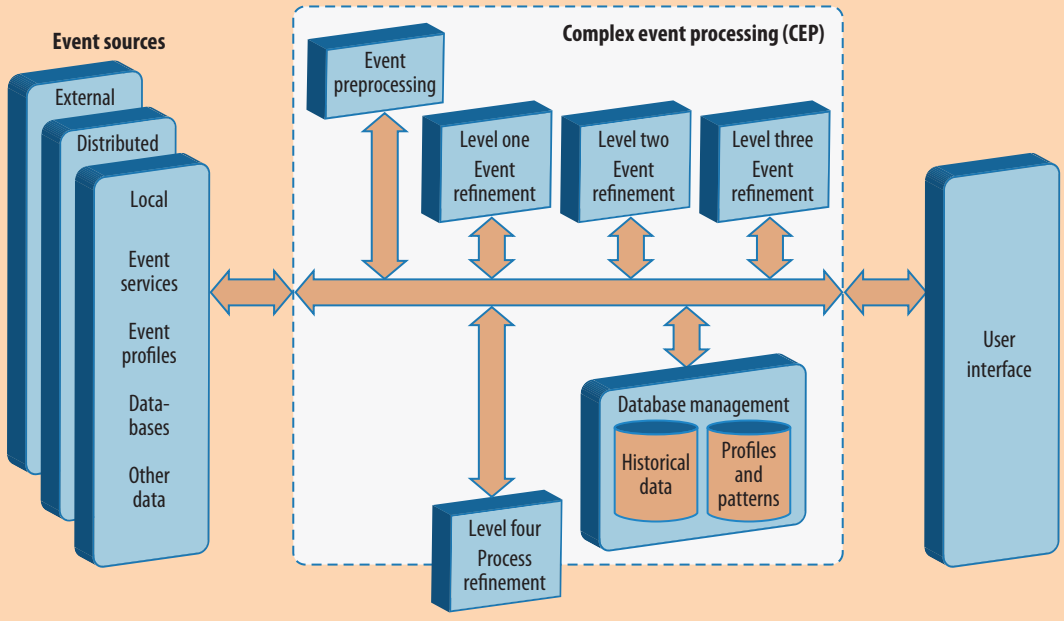

Figure 1. Complex-event processing systems take information from multiple sources of raw events and process them in numerous ways before determining if they represent complex events. If so, the system sends the information to the user.

He subsequently coined the term "complex-event processing" when he applied the Rapide toolset to the simulation and analysis of software systems. His 1998 proof-of-concept experiments demonstrated the use of CEP to monitor the performance of and detect errors in distributed-messaging systems.

Concurrent CEP research projects included Caltech's Infosphere, Cambridge University's Apama, and Active Middleware Technology at the IBM Research Laboratory in Haifa, Israel. Products using the technology began appearing in 2001.

Two technical advancements have helped drive demand for CEP systems. Data sources, including the increasing number of inexpensive sensors, have become able to continuously transfer their information to central analysis facilities, noted Gartner vice president of architecture and middleware Roy Schulte.

In addition, said Realtime Monitoring's Heinz, IT infrastructures have become more modular and flexible, making it easier to add CEP capabilities to existing systems.

Also driving the technology's popularity is the growing importance of event-driven software architectures; business-process management (BPM); and service-oriented architectures (SOAs), infrastructures in which users can invoke certain functions within a system as services.

CEP techniques could help improve BPM activities. And CEP could act as one of the services provided via an SOA, noted Luckham.

CEP products

Essentially, there are two types of CEP toolsets. Stand-alone software engines link to multiple applications. Other CEP services can be bundled into large products such as a database or enterprise-service-bus software. Typically, ESB software sits between business applications and services, and enables communication among them.

CEP products on the market now take different approaches. For example, the Aleri Streaming Platform uses high-level modeling tools that let clients implement business logic while processing events.

The Aleri product and StreamBase Systems' CEP toolset use SQL extensions for processing events, which is good for programmers used to working with the language. Systems can run continuous SQL queries over streams of raw events, noted Heinz.

Agent Logic's RulePoint uses a stand-alone, thin-client, Web-based platform that lets any user in an organization create and work with CEP rules. This, said the company's Bradley, delivers CEP benefits to nonprogrammers, while providing collaborative capabilities across an enterprise's business units.

Realtime Monitoring's RTM Analyzer is a library-based platform that has multiple modules with complementary functionality. For example, there is a module for SQL queries over event streams, a failover module for redundant execution of CEP logic, and a module for advanced statistical stream analysis. Based on customer

needs, the RTM Analyzer configures and connects the required modules, delivering a tailored, extensible CEP system.

TIBCO's popular BusinessEvents Enterprise Suite works with inference-based rules that are applied to incoming event streams to find patterns and determine causality.

How CEP works

A CEP system is a software infrastructure that typically consists of a layer of adapters that can connect directly to data sources and pipe information to the analysis engine.

In some cases, CEP products connect via adapters to the ESB, which in turn connects to data sources with its own adapters, Heinz noted. The adapters let CEP tools work with data from heterogeneous sources.

Users can specify the business logic they use to filter and analyze events via programming interfaces.

A CEP server evaluates the business logic, then runs event-stream-analysis algorithms and transfers the results continuously to users. "This can be done using a Web-based dashboard or a real-time notification such as e-mail or [short-message-service communications], or by publishing data to an upstream system for further analysis," said Baden Hughes, architect for CEP vendor Event Zero.

CEP products analyze events in many ways, including via pattern matching, inference, graph analysis, and models that support implicit and explicit event causality, said Opher Etzion, who chairs the Event Process Technical Society (EPTS) Steering Committee and is also an event-processing scientist at IBM's Haifa Research Lab.

The basic pattern-matching algorithm works by looking for a potential sequence among inbound events, said Marco Seirio, a spokesperson for CEP vendor RuleCore. The algorithm then determines whether a pattern exists, what it is, and which events belong to it, noted Etzion.

To help with the process, CEP tools filter out irrelevant information and enrich events with external data, he noted. They then aggregate and correlate information from the various events to identify the complex events.

The CEP engine optimizes the execution of the technology's complex data flow, Aleri's Morrell noted.

Vendors build CEP systems so that they initiate processing in response to inbound events, which enables them to function in real time. Traditional event-processing systems wait for a user action to initiate processing.

CEP promises to help organizations detect complex patterns in activities and recognize opportunities and threats.

Event data, like information in a relational database, can be organized as a table. CEP vendors, noted Gartner's Schulte, have thus used SQL as a starting point for designing event-processing languages. They borrow some of the SQL syntax and then add new capabilities specifically for processing events.

According to Morrell, CEP reports should include information on the raw events from which the complex event was derived. Delivering the raw events, as well as additional data, can provide important context for the recipient user or application, as well as the audit trails that government regulation or corporate policies require in some cases.

Current types of applications

CEP technology could be used for numerous applications. For example, it could analyze events within a network, power grid, database, or other large system to determine whether it is the target of an attack, is performing optimally, or is experiencing problems.

Optimization analysis could help activities such as inventory tracking, hazardous-materials handling, patient flow in hospitals, complex airline operations such as equipment usage, the detection of threats to national safety, and the scheduling of trucking fleets.

Financial-trading institutions could monitor worldwide stock, commodity, and other markets to recognize potential opportunities or problems with current holdings. CEP could also inspect e-commerce activities for anomalies and signs of fraud, such as credit-card theft.

Data from simulations could be run through CEP systems to analyze proposed business processes and other activities to determine whether they would be inefficient, cause legal problems, or violate customers' service-level agreements.

The technology could examine internal corporate activities to determine whether they conform to government regulations and corporate policies.

Similarly, regulators could use CEP to look through organizations' business activities to determine whether they are violating laws, such as by committing insider stock trading.

OBSTACLES TO CLEAR

Currently, Gartner's Schulte noted, typical CEP systems cost, on average, between $100,000 and $250,000. This is one reason, he said, why these products have been limited largely to demanding, high-value applications like those used for stock trading.

As CEP systems' prices drop, he added, the technology will be used in additional applications.

A major challenge to widespread CEP adoption is a lack of standards for the event-pattern and rule languages, according to IBM's Etzion.

Currently, vendors use proprietary languages. For example, RuleCore uses a custom pattern-detection language called the RuleCore Markup Language.

Standards would be required to enable the use of multiple event-processing engines, as well as to enable migration among different products, Etzion explained. In addition, added RuleCore's Seirio, customers like to spend training dollars on technology they can reuse in multiple contexts, not just with specific vendors' tools.

Said Luckham, "Education in and understanding of CEP techniques and how they can be applied are still ongoing, among both the vendor and customer communities."

"The real impediment to adoption is a lack of education," said Gartner's Schulte. "Most business analysts and software engineers and architects are not familiar with CEP technology yet."

Other hurdles, he said, include a lack of standardized benchmarks, as well as pattern and rule libraries. This would enable the reuse of existing patterns and rules, and eliminate the need for users to write all of their own code for each application.

"Better tools for managing [CEP] once it is deployed would help," added Schulte.

There are not many use cases available yet, noted Gartner's Knox, so best practices are still evolving.

CEP systems must continue being able to handle the ever-increasing volume and complexity of financial, stock-trading, and other types of data they encounter without undue latency, she said.

CEP systems also must better be able to access and incorporate contextual information to put their findings in perspective, she added. For example, if a system detects a rise or fall in a stock price, it should be able to identify whether the change is a regular, periodic event or represents a fleeting market opportunity or risk.

Recipients of CEP-system alerts, particularly those using legacy systems, will have to be able to receive and keep up with the increasing speed and volume of notices.

According to Aleri's Morrell, business-intelligence customers will soon use CEP to implement continuous intelligence, driving faster, more responsive actions to markets, customers, and operations.

"CEP as an individual commercial tool has a limited lifetime, probably another eight years or so, until the CEP techniques and range of applications are better understood," said Luckham. "Then, it will be gradually absorbed into more comprehensive technology suites."

The Aite Group, a research and advisory firm, estimates that revenues generated by CEP products will more than double from $180 million in 2008 to $370 million this year and increase to $460 million in 2010.

According to IBM's Etzion, "The need for CEP is obvious. It represents one of the next big competitive advantages."

It is moving from being a competitive differentiator to being a competitive requirement for many companies, particularly financial services firms, according to Gartner's Knox.

"You'll be able to use the technology to turn your business into a proactive, predictive enterprise with the ability to detect patterns of opportunity or, on the flip side, patterns of threat," said TIBCO senior product marketing manager Alan Lundberg. "Either could serve to make bottom-line contributions of time or money."

Neal Leavitt *is president of Leavitt Communications (www.leavcom.com), a Fallbrook, California-based international marketing communications company with affiliate offices in Brazil, France, Germany, Hong Kong, India, and the UK. He writes frequently on technology topics and can be reached at neal@leavcom.com.*

Editor: Lee Garber, *Computer*, l.garber@computer.org