综 述

复杂事件处理技术及其应用综述

江连峰，赵佳宝

（南京大学工程管理学院，南京 210093）

摘 要：近年来，处理更高级别的事件（通常被称为业务事件）的需求迅速扩大，复杂事件处理（Complex Event Processing,CEP）的发展满足了这些新的市场需求。CEP 能结合来自多个数据源的数据来推断更为复杂的事件或模式。CEP 的目标是从事件流中找出有意义的事件（如机会或威胁），并尽快做出反应。文本介绍了 CEP 的研究现状，包括通用的参考架构，代表性的 CEP 系统及其事件处理语言，以及最常见的应用场景。

关键字： 复杂事件处理；事件处理语言；CEP 应用

中图分类号：TP391.7 **文献标识码：**A DOI：10.3969/j.issn.1003-6970.2014.02.074

本文著录格式：[1] 江连峰，赵佳宝 . 复杂事件处理技术及其应用综述 [J]. 软件 ,2014,35(2)：188-192

The Summary of the Complex Event Processing Technology and its Application

JIANG Lian-feng, ZHAO Jia-bao

(School of Management and engineering, Nanjing University, Nanjing 210000*,China)*

【Abstract】In recent years, the demand of processing higher-level events (often referred to as business events) rapidly expands. The development of Complex Event Processing (CEP) meets needs of these new markets. CEP can combine data from multiple data sources to infer more complex events or patterns. The goal of CEP is to identify meaningful events (such as opportunities or threats) from the event streams and react as quickly as possible.This article introduces the research status of CEP, including a common reference framework, representative CEP systems with their event processing languages and the most common application scenarios.

【Key words】Complex Event Processing; Event Processing Languages; CEP Applications

0 引言

复杂事件处理（Complex Event Processing,CEP）是一门新兴的学科，其根源于离散时间仿真、主动数据库、网络管理和中间件技术。David Luckman 所著的书 The Power of Event 极大地推动了 CEP 的发展 [1]。

Luckman 把事件定义成系统中任意一个活动的发生，其特点有显着性（发生在系统感兴趣的特定领域），瞬时性（发生在一个特定的时间点）和原子性（发生或不发生）。数据与事件的发生紧密联系在一起，一些数据是所有事件共同的（如发生的时间），而另一些则仅特定于某些事件。与事件相关的数据项被称为属性。

事件之间通过时间顺序、因果关系、聚合关系相互联系，事件之间的这些联系被定义成事件模式。CEP 采用事件分层结构，定义了一组规则把低层事件逐步聚合成高层事件，同时事件越来越接近应用层而远离 IT 硬件层，即高层事件 A 可以由一组低层事件 $\{B_i\}$ 抽象而成，此时事件 A 被定义成复杂事件（complex event）。

CEP 主要的任务是事件模式匹配，即在事件流中识别应用程序域中的重要事件模式。监视事件流的事件处理代理（Event Processing Agents, EPA）执行模式匹配和事件处理。从本质上讲，EPAs 过滤，分离，聚合，转换和充实事件。此外，根据事件之间的相关性从简单事件合成复杂事件。最后，事件传播到下游系统（如应用程序或仪表板）来触发相应的事件处理程序。

1 CEP 系统参考架构

文献 [2] 提供了图 1 所示的 CEP 通用参考架构，是当前 CEP 架构的一般抽象，其专注于 CEP 系统的共同要素和它们之间的相互关系，避免依赖于具体的技术。设计并执行具体的 CEP 系统需要许多额外的详细设计决策和技术选择。该架构描述了四个不同的实体：

(1) 事件源：产生的事件通常是原子事件，但也有可能是高层次的事件。

(2) 事件建模器：定义了事件模式。这些事件模式在事件处理媒介中处理。

(3) 事件处理媒介：用于处理事件的平台（可能是分布式的），提供通道访问来自事件源的事件数据，并把处理结果输出给事件消费者。事件处理媒介为事件的选择、聚合、层次结构内的分类、更高层次事件的抽象提供技术架构。事件处理媒介可能是单一的 EPA（单独的事件处理组件），也可能是一个包含了一组 EPA 的分布式事件处理网络。

(4) 事件消费者：接收来自事件处理媒介的复杂事件并通过

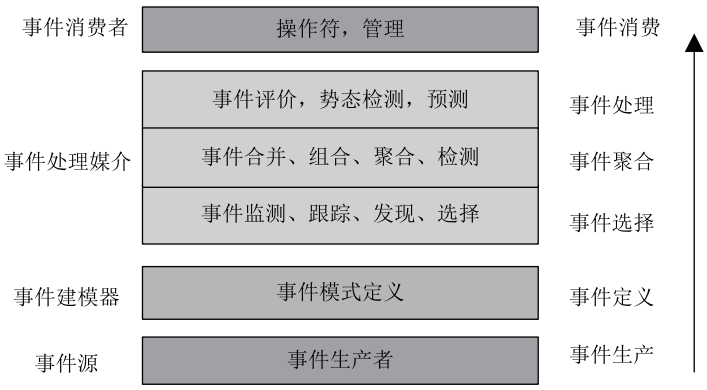

图 1 CEP 通用参考架构

反应规则对其响应。事件消费者可能会产生新的事件而作为事件源。

目前 CEP 解决方案供应商提出了几个具体的 CEP 的参考架构。一些参考架构与具体实施技术关联性不强，而另一些基于具体的解决方案和技术之上。特定领域的参考模型和元模型也已被提出[3,4]。实际架构的某些技术选择，不是由要求本身所决定，而是由当时可选择的技术所决定。

2 CEP 事件查询

数据库查询是对有限的一组数据的一次性且临时的查询，而事件查询是对（概念上）无限的事件流连续且长期的查询。由于数据连续到达，事件查询需要使用滑动窗口处理无限的数据流，这意味着事件有一定的租用期，过期则被删除，越抽象的事件，租用期越长。

通常，事件处理语言描述事件模式，并由相应的规则引擎来执行事件查询。事件处理语言的要求可以描述成以下四个方面（也称为事件查询的四个维度）：

(1) 数据提取：事件包含的数据决定是否及如何对该事件作出响应。提取事件中的数据作为事件对象的属性值，用于查询过程中的条件测试、创建新的事件对象或者触发相应的反应。

(2) 事件合成：从若干低层事件对象中提取属性值，用于聚合成更高层的复杂事件。

(3) 时间（因果）关系：事件查询往往涉及时间约束，即事件的发生必须在一个特定的时间间隔内或按一个特定的顺序。根据 Luckman 提出的因果时间原理，事件间的因果关系暗示了事件间的时间关系，所以复杂事件描述语言满足时间关系就能满足因果关系。

(4) 事件积累：事件查询的非单调特征，如事件否定（事件未发生）和一定时间跨度内的事件数据聚合对无限的事件流是无意义的，因为它们只能在事件流结束时才能确定。因此，这类查询只能针对特定时间窗内的事件。

3 典型 CEP 系统和事件处理语言

以下列举一些代表性的 CEP 系统与其事件处理语言。

3.1 PAPIDE-EPL

Luckman 提出了 PAPIDE-EPL[1]。PAPIDE-EPL 是一种强类型语言，它可以申明事件类型，然后对事件类型进行匹对并从事件对象中提取有用信息。PAPIDE-EPL 的基本类型包括标准整型、字符串、布尔型、数组和记录。事件有两种类型的属性——标准属性（如类型名、来源和创建时间）和用户自定义属性。匹配规则包括针对事件属性的关系运算和控制流结构（如循环和条件）。

3.2 SASE

加州大学伯克利分校的 Eugene Wu 等人设计了 SASE 系统，该系统主要针对 RFID 应用场景[6]。SASE 在主动数据库的事件处理语言的基础上进行改进来满足 RFID 技术在监控领域的应用。这些扩展包括事件否定，参数化谓词和滑动窗口。SASE 基于非确定性有限自动机模型并采用基于查询计划的方法予以实现，这与基于固定数据结构的事件系统（如有限自动机，树，Petri 网）形成了鲜明的对比。

SASE 解决了事件流处理中的两个突出的问题：滑动窗口和中间结果集较大。当使用的窗口很大时，使用一种辅助数据结构——活跃实例堆栈(active instance stacks，AIS)来促进序列构造。同时向下推进谓词和窗口到序列算子来减小中间结果集的大小。

SASE 语言的结构如下：

EVENT <event pattern>

[WHERE <qualification>]

[WITHIN <window>]

<event pattern> 代表需要查询的事件模式；<qualification> 过滤事件对象的属性；<window> 代表时间约束。

SASE 存在的缺陷：

1. 只能将原子事件转化为复杂事件，不能把复杂事件转化为更为复杂的事件。

2. 相关查询是基于原子事件是有序的假设。

3. 不支持聚合功能，如 count(),avg() 等，这在其后续的系统 SASE+ 中得到了改进。

3.3 NFA^b

相对于正则表达式匹配，事件流的模式匹配存在两个挑战：1）匹配语言更加丰富，因为事件匹配必须支持定序、选择相关事件和克林闭包（从输入流中提取数量有限但无界的具有特定

属性的事件）；2）运行更加高效，因为输入事件的数量在进行匹配前是不知道的[7]。由此，Agrawal 提出了用于事件流的查询评估模型 NFA^b（带有匹配缓冲的非确定性有限自动机）和查询评估框架，并利用共享存储和处理过程来提高运行时的效率。

一个 NFA^b 自动机由 5 个部分组成：一组状态，一组有向边，一组标记这些有向边的公式，开始状态和最终状态。这些都是非常类似于传统的正则表达式，除了转换根据的是匹配公式而不是输入标记的类型。Agrawal 提供了一个算法，将一条模式查询转化成 NFA^b 模型以便进一步的优化，同时可以很好地比较不同语言之间的表现力。Agrawal 还建立了一个运行时引擎，用来执行 NFA^b 查询。

3.4 Cayuga

Cayuga 由 Demers 等人在康奈尔大学建立的一个处理复杂事件的高性能系统[8,9]。Cayuga 是 NFA^b 形式的另一实现，但是它侧重于多个事件流的优化。Cayuga 的事件查询语言基于事件代数理论，并通过自定义的自动机模型和内部命名方法来高效地探测复杂事件。其格式与 SQL 和 SASE 很相似，格式如下：

SELECT <attributes>

FROM <algebra expression>

PUBLISH <output stream>

其中 <attributes> 是输出流模式的属性，<algebra expression> 是 Cayuga 的事件模式，PUBLISH 分句对输出流命名（用于分层处理和订阅）。

Cayuga 的一个重要特征是能够适用于各种事件到达比率和事件查询数量，后者的扩展性尤为重要，因为对于消费类应用，数以百万计的用户可能同时对系统进行查询。

Cayuga 查询语言的操作符有明确定义的形式化语义，这使 Cayuga 能够实现查询重写优化。同时所有的操作符都是可组合的，这允许 Cayuga 从简单的子模式建立复杂的模式。Cayuga 采用索引和垃圾收集等技术，使其以非常高的速率处理数据流。由于 Cayuga 由发布 / 订阅系统发展而来，所以它能在多个模式中寻找共同的子表达式。

与 SASE 系统相比，Cayuga 系统支持简单的重复操作和时序限制，但不支持否定操作，而且先前参与复杂事件匹配的基本事件不再参与后面的匹配。另外，Cayuga 系统不支持修正乱序，可能会导致参与匹配的有效乱序事件被删除，这样部分复杂事件就不会被正确检测。

3.5 XChange^{EQ}

Bry 和 Eckert 在文献[10]首次将关系代数引入到复杂事件的匹配与过滤中，从而对复杂事件的检测就转化为具有时间条件（事件按固有的时间顺序发生）的关系代数表达式的增量评估问题。关系代数为事件查询提供了清晰的理论基础，尤其是通过等价定律实现了查询重写，从而可获得更有效的查询计划。它还可以充分利用已有数据库方面的研究，如操作符的实现，内存数据库（因为事件处理通常在内存中进行），以及分布式数据库（实现事件查询的分布式评估）。据此，Bry 和 Eckert 提出了事件处理语言 XChange^{EQ}[15]。XChange^{EQ} 是第一个能处理 XML 的形式事件的事件处理语言。它特别适合于建立在 Web 服务标准之上的商业应用领域。

XChange^{EQ} 采用 Tarskian 模型理论和不动点理论提供语义支持，虽然这已用于规则语言，但并没有被应用到事件查询。模型理论存在一个问题，即许多模型同时适对于一个给定的程序，这可以通过将程序定义为一个基于模型理论的不动点方程的解来获得唯一的模型。

XChange^{EQ} 的存在一个限制：它不支持事件实例的选择和事件实例消耗。可以说，这两点对业务级别事件不太重要，因为事件中所包含的数据可达到类似的效果。

3.6 SARI

Josef Schiefer 等人设计了一个实时商务智能系统 SARI (Sense And Respond Infrastructure) 来对业务流程自动地感知，解释，预测及响应，以此来减少业务决策所花费的时间[11]。SARI 能够持续不断地从不同的事件源接受事件，并将这些实时事件转化成绩效指标和智能商务行动。SARI 处理过程主要有五个步骤 感知、解释、分析、决定和响应。

SARI 系统的事件处理语言为 EAExpression，它是一种嵌入式的语言。文献[12]详细的介绍了 EAExpression 的事件模型、语法和语义、语言结构、事件处理映射和事件驱动规则。

4 CEP 典型应用

4.1 金融

金融领域与交易，风险管理及合规管理等有关的数据大幅增长，此外另类交易系统和暗池的发展增加了市场数据的来源。CEP 可以有效地解决市场数据源和数据量大幅上涨所带来的问题。CEP 已经是全球主要金融服务公司收入和竞争优势的来源。

在技术和架构方面，金融企业可以利用 CEP 的三个层次[13]。在第一层次，低延迟系统对毫秒级或更小的市场数据进行操作，这是整个架构中最复杂、资源最密集的层次。第二层次，业务数据被用于各种操作流程，包括风险管理和合规操作。这些操作流程可以在几秒内完成。第三层次，用于满足其直接向监管机构或交易存储库提交交易报告的要求，其操作可以在几分钟甚至几小时内完成。

CEP 在金融领域的应用包括：(1) 大量市场数据的实时处理（标准化，清洗，精简）；(2) 算法交易机会的相关分析；(3) 低延迟的订单路由和执行；(4) 实时定价和风险管理。(5) 欺诈和合

规应用；(6) 服务水平监测；

4.2 RFID

RFID 应用通常有两种类型：对象跟踪和实时监控，这两者都需要从 EPC 数据中提取隐含语义将原始标签事件转化成业务事件。CEP 可用于任何一种基于消息的分布式系统的信息提取和分析，同时 CEP 能快速地处理有时间限制的高容量 RFID 数据流，故使用 CEP 处理 RFID 数据已成为一种普遍的选择。

RFID 系统主要由四部分组成：标签，阅读器，中间件（或边缘服务器）和应用程序。RFID 中间件有利于物理层和企业应用层之间的数据和信息的沟通。目前的 RFID 中间件仅输出与业务流程不直接相关的 RFID 原始事件。文献 [15] 提出为了使 RFID 系统响应更快，RFID 中间件应采用 CEP 来关联业务逻辑并产生业务事件，而不是仅仅提供原始标签信息。

由于 RFID 事件流中的数据错误经常发生，基于 RFID 的 CEP 包括两方面的内容，数据清理和事件检测。数据清理机制包括时间平滑、空间平滑、完整性约束和事件缓冲区缓存。而 RFID 事件包括高度的时间性和存在非自发的事件，传统的事件检测系统来很难支持 RFID 事件检测。文献 [16] 开发了基于 CEP 的 RFID 复杂事件检测的算法（RCEDA）来处理 RFID 事件，并引入伪事件处理非自发事件。

CEP 在 RFID 领域的典型应用场景：

(1) 企业信息系统

RFID 技术使企业信息系统能获得实时准确且详细的数据，而 CEP 可以处理数量庞大的数据，并发现这些数据背后的业务信息且能自动响应信息，这就弥补了原始数据和管理人员可操作的信息之间的信息差，同时支持企业自动响应可操作的信息。文献 [17] 提出了针对企业信息系统的事件处理架构、事件的元模型和 CEP 规则，利用事件实例的分类和分区提高事件检测效率。

(2) 实时定位系统

实时定位系统（RTLS）使用 RFID 实时定位来识别带有标签的运动物体，并根据在其位置信息和其他信息做出适当的处理。 文献 [18] 提出了基于时间自动机理论的 RFID 复杂事件处理的方法 GEEP，有效地解决了 RTLS 中使用 RFID 技术和 CEP 技术时遇到的问题。

(3) 物流

文献 [19] 利用 CEP 建立的第四方物流供应商（4PLP）监测模型，有效地确认运输延迟责任方和采取前瞻式行动避免违约。文献 [20] 在港口作业集装箱堆场业务流程的自动化系统中使用了 CEP。

(4) 医疗保健

RFID 技术和嵌入式设备如生理传感器和环境传感器已经应用于医疗保健领域，因此，医院需要处理不同来源的大量实时数据，且这些数据的格式与传入速率不同。CEP 可以提供很好的解决方案。文献 [21] 提出了一个用于手术治疗和处理紧急情况的基于 RFID 的 CEP 框架，并且已经实现了一个用于手术治疗的原型系统。文献 [22] 提出了一个在电子医疗网络中用于个人健康监测系统数据集成的 CEP 架构。该架构基于可戴式无线传感器，智能手机和数据流挖掘技术，并实现了个人健康监测系统和医疗服务提供商系统之间的以事件驱动的互操作。

4.3 感知网

文献 [23] 提出了基于事件驱动架构的传感器网络的通用软件架构，适应不同的应用场景。相应地，并提出了一个粗粒度的事件层次结构和事件处理网络，且将 CEP 作为事件处理模型。将该架构应用到智能交通管理系统的决策支持系统大大地提高了智能交通管理决策支持系统的实时性。

文献 [24] 提出了一个用于无线传感器网络环境下的目标跟踪和入侵检测的 CEP 解决方案。

4.4 Web 服务适配器

Web 服务接口之间存在差异导致它们无法互相通信，其差异可以归类为签名不兼容和协议不兼容，现有技术能解决其中一种不兼容，但很少能同时解决两者。文献 [25] 介绍了基于 CEP 的解决方案能同时解决 Web 服务接口之间的签名和协议不兼容。

5 结论

CEP 的主要思想如下：首先从大量数据中抽取原始事件；其次根据特定的规则，使用事件关联和事件聚合操作生成复合的业务事件；再对原始或复合事件进行事件处理，以获得它们的时间、因果、层次和其他的语义关系；最后响应可操作的业务信息。

事件处理语言支持事件模式（事件定义、选择和消费政策）的规范以及事件处理的规则。

CEP 通常用于处理大量事件流数据且根据事件的变化做出低延时反应的应用场景。由于 CEP 应用如此广泛，必须根据不同的应用场景正确配置与定制 CEP 系统。不同的应用领域，可能需要特殊的事件处理方法，如采用分布式还是集中式模式。

CEP 进一步的研究课题包括处理不确定性的事件（例如，用概率论方法），检测一个预先未知的复杂事件（例如，事件流的数据挖掘）等。

参考文献

[1]David L.The Power of Events: An Introduction to Complex Event Processing in Distributed Enterprise Systems [M]. US: Addison-Wesley Professional, 2002.

[2]Paschke A, Vincent P. A reference architecture for Event Processing[C]. Proceedings of the Third ACM International Conference on Distributed

Event-Based Systems. ACM, 2009: 25.

[3] Ammon R, Silberbauer C, Wolff C. Domain specific reference models for event patterns–for faster developing of business activity monitoring applications[C].VIP Symposia on Internet related research with elements of M+ I+ T+. 2007, 16.

[4]Blanco R, Wang J, Alencar P. A metamodel for distributed event based systems[C]. Proceedings of the second international conference on Distributed event-based systems. ACM, 2008: 221-232.

[5]Robins D. Complex event processing[C].Second International Workshop on Education Technology and Computer Science. Wuhan. 2010.

[6]Wu E, Diao Y, Rizvi S. High-performance complex event processing over streams[C]. Proceedings of the 2006 ACM SIGMOD international conference on Management of data. ACM, 2006: 407-418.

[7]Agrawal J, Diao Y, Gyllstrom D, et al. Efficient pattern matching over event streams[C]. Proceedings of the 2008 ACM SIGMOD international conference on Management of data. ACM, 2008: 147-160.

[8]Demers A J, Gehrke J, Panda B, et al. Cayuga: A General Purpose Event Monitoring System[C]. CIDR. 2007, 7: 412-422.

[9]Brenna L, Demers A, Gehrke J, et al. Cayuga: a high-performance event processing engine[C]. Proceedings of the 2007 ACM SIGMOD international conference on Management of data. ACM, 2007: 1100-1102.

[10]Bry F, Eckert M. Rule-based composite event queries: The language xchangeeq and its semantics [M]. Web Reasoning and Rule Systems. Springer Berlin Heidelberg, 2007: 16-30.

[11] Schiefer J, Seufert A. Management and Controlling of Time-Sensitive Business Processes with Sense & Respond[C]. Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on. IEEE, 2005, 1: 77-82.

[12] Rozsnyai S, Obweger H, Schiefer J. Event Access expressions: A business user language for analyzing event streams[C]. Advanced Information Networking and Applications (AINA), 2011 IEEE International Conference on. IEEE, 2011: 191-199.

[13]Jaswal A. Complex Event Processing in Asia-Pacific-Addressing Risk Management and HFT Requirements [OL],[2011]. http://www.celent.com/reports/complex-event-processing-asia-pacific.

[14]Adi A, Botzer D, Nechushtai G, et al. Complex event processing for financial services[C]. Services Computing Workshops, 2006. SCW'06. IEEE. IEEE, 2006: 7-12.

[15]Hu W, Ye W, Huang Y, et al. Complex event processing in RFID middleware: A three layer perspective[C].Convergence and Hybrid Information Technology, 2008. ICCIT'08. Third International Conference on. IEEE, 2008, 1: 1121-1125.

[16]Wang F, Liu S, Liu P. Complex RFID event processing [J]. The VLDB Journal-The International Journal on Very Large Data Bases, 2009, 18(4): 913-931.

[17]Zang C, Fan Y, Liu R. Architecture, implementation and application of complex event processing in enterprise information systems based on RFID [J]. Information Systems Frontiers, 2008, 10(5): 543-553.

[18]Liu Y, Zhang H, Wang Y. RFID Complex Event Processing: Applications in Real-Time Locating System [J]. International Journal of Intelligence Science, 2012, 2(24): 160-165.

[19]Roth M, Donath S. Applying Complex Event Processing towards Monitoring of Multi-party Contracts and Services for Logistics–A Discussion[C]. Business Process Management Workshops. Springer Berlin Heidelberg, 2012: 458-463.

[20]Kim Y, Yoo J W, Park N. RFID Based Business Process Automation for Harbor Operations in Container Depot[C]. Proceedings of Industrial & Manufacturing Engineering Graduate Research Symposium at Wayne State University. 2006: 213-226.

[21]Yao W, Chu C H, Li Z. Leveraging complex event processing for smart hospitals using RFID [J]. Journal of Network and Computer Applications, 2011, 34(3): 799-810.

[22]Mouttham A, Peyton L, Eze B, et al. Event-driven data integration for personal health monitoring[J]. Journal of Emerging Technologies in Web Intelligence, 2009, 1(2): 110-118.

[23]Dunkel J. On complex event processing for sensor networks[C]. Autonomous Decentralized Systems, 2009. ISADS'09. International Symposium on. IEEE, 2009: 1-6.

[24]Bhargavi R, Vaidehi V, Bhuvaneswari P T V, et al. Complex Event Processing for object tracking and intrusion detection in Wireless Sensor Networks[C].Control Automation Robotics & Vision (ICARCV), 2010 11th International Conference on. IEEE, 2010: 848-853.

[25]Taher Y, Parkin M, Papazoglou M, et al. Adaptation of web service interactions using complex event processing patterns [J]. Service-Oriented Computing, 2011: 601-609.