

前言

本白皮书将为您提供有关富士通 XML 型数据库引擎「Interstage Shunsaku(瞬索) Data Manager V6」的产品信息,并介绍其研究成果。

如果本白皮书能给您的信息系统构筑提供一些帮助,我们将感到不胜荣幸。

富士通株式会社

有关「Interstage Shunsaku(瞬索) Data Manager V6」的最新信息请参考以下网站。

<http://cn.fujitsu.com/it/products/Interstage/shunsaku.htm>

本白皮书所介绍的「Interstage Shunsaku(瞬索) Data Manager V6」的功能中包括了一部分仍处于计划阶段的内容。这些计划中的内容,根据商务环境和技术动向的变化,有可能在不予通知的情况下被中止。

关于注册商标

- Java 以及所有 Java 关联的商标及标识语均为美国 Sun Microsystems 公司的商标或其注册商标。
- Microsoft Excel 及 Microsoft Word 为美国微软公司的商品名。
- Oracle 为美国 ORACLE 公司的注册商标或商标。
- Solaris 操作环境为美国 Sun 公司的商标,并由 Sun 公司的关联公司 Sunsoft 公司提供授权。
- 本资料称在 Solaris 操作环境上进行运作的产品为「面向 Solaris OE 的产品」,「Solaris OE 对应」,以及「Solaris OE 版」。
- Interstage 及 Symfoware 为富士通株式会社的注册商标。
- UNIX 为 The Open Group 集团在美国及其他国家的注册商标或商标。
- Linux 为 Linus Torvalds 在美国及其他国家的注册商标或商标。
- 本资料所使用的图像以及公司名称,产品名称等专用名词均为各公司的商标或注册商标。
- 此外,网页中所记载的系统名称,产品名称等并未全部标记上 TM(Trademark:注册商标)或 R(Registered:已注册)等记号。

目录

前言	… 1
第 1 章 背景	
1.1 什么是信息的有效利用	… 3
1.2 在信息的有效利用方面所存在的问题	… 3
1.3 对变化应对自如的「XML 型」数据结构	… 4
1.4 仅靠 RDB 已不能适应市场的需求	… 4
1.5 富士通的研究成果	… 5
1.6 对 IT 技术进行最大限度的有效利用	… 6
第 2 章 Shunsaku(瞬索)的特长	
2.1 给信息系统带来革新的技术	… 7
2.2 Shunsaku(瞬索)的全貌	… 7
2.3 设计	… 7
2.4 开发	… 10
2.5 运用	… 11
2.6 应用程序开发	… 12
第 3 章 业务应用模式	
3.1 Shunsaku(瞬索)的应用模式(概要一览)	… 14
3.2 Shunsaku(瞬索)的应用模式(详细介绍)	… 14
第 4 章 今后的研究课题	
4.1 关于 Shunsaku(瞬索) V7 的研究	… 18
4.2 Shunsaku(瞬索) V7 的主要构成技术	… 18

第 1 章 背景

1.1 什么是信息的有效利用

信息是指数据 (data) 或情报 (information) 等人们在做出某项判断和采取某个行动时, 通过各种媒体获得的资料或知识。信息的有效利用是指“充分地利用各种信息情报, 高效率地找到解决问题的线索”。充分而有效地利用各种信息是使企业在激烈的竞争中立于不败之地, 获得成功的必要手段。

但是, 利用以往的信息系统, 信息的利用并没有能达到理想的效果。也就是说, 在以往系统中, 为了储存信息, 必须先设定一些规则, 然后根据这些规则对信息进行加工, 储存。所以, 信息的利用被局限在事先制定的规则范围内, 大量信息得不到充分的利用。可是, 当今社会已经进入了高度竞争的信息时代, 为了与其它企业拉开档次, 更多、更快、更有效地利用各种信息, 提高自身的竞争力, 采用超越以往这些束缚的信息利用系统成为了迫在眉睫的问题。

以往的信息系统中所制定的那些规则其实破坏了信息利用应有的意义。信息的利用原本是指人们为了某个特定的目的, 通过各种方式收集必要信息, 并对其运用运用的过程。可是运用以往的信息系统, 却形成了人们被自己制定的规则所束缚, 使大量信息得不到充分利用的局面。本应以人为主体的信息利用不幸陷入了以系统为主体的信息利用状态。

随着泛在时代 (Ubiquitous) 的到来, 我们必须重新正视信息有效利用的现状, 脱离以往条条框框的束缚, 建立起以人为主体的宽松自如的信息利用体系。

1.2 在信息的有效利用方面所存在的问题

随着商业环境的变化, 企业必须处理各种形式的数据。但是, 实际存在的问题是, 并非所

有的数据都能在业务和经营等方面得到有效利用。如果说得到有效利用的数据只占全部数据的一小部分也并非言过其实。之所以形成这种状态, 是因为企业的信息利用结构不能应对在现场被累积的数据的多样“变化”。

应对数据的种类及形式的变化是一件至关重要的事情。营业日报表或有关维护信息等报表一般都是由普遍使用的办公室软件工具 (微软® Word, Excel, 文本编辑器等) 编辑而成。这些数据以独自的二进制数据为主, 另外还包扩一些手写形式的数据。此外, 个人简历, 医院的电子病历卡等结构极为松散的数据群的管理也是一件繁琐而复杂的事情。

在企业使用的数据中, 本身就为数值型数据的营业额或成本价格等有关收支的数据可以原封不动地使用。数值型数据作为各商品, 各客户的销售业绩, 销售成本, 赢利情况, 零售店的销售业绩等数值使用来进行分析和比较已经足够。但是, 仅依据数值型数据, 很难决定企业的运行方针和战略决策。每个企业都需要运用所掌握的信息制定新企画, 开发新客户, 所以, 在业务和经营方面对现有数据进行充分而有效的利用也变得越来越重要。

此外, 如何利用被积累的数据也成为一大问题。我们在和客户接触的过程中, 或是在各种业务现场收集到大量的数据, 这些数据不仅来自于营业现场, 工厂, 客户等商业最前线, 也来自于连锁超市, 连锁便利店的 POS (Point-Of-Sale) 系统。这些销售管理数据和客户管理数据日益积累; 另一方面, 来自于信息咨询服务中心等部门的客户管理数据和支援服务信息也日益增多。从这些被积累的庞大数据中, 根据利用者的需求, 用任意的关键词提取信息, 并以此作为解决问题的线索进行有效利用并非是一件容易的事情。

今后, 在对信息进行有效利用时, 无论从现

场收集的数据是什么种类、形式、及规模,都必须对其进行一元化管理,以迅速实现数据的共享和有效利用,建立起“有效利用来自现场的第一手信息”的体系,使信息的利用与企业的快速经营相连接。

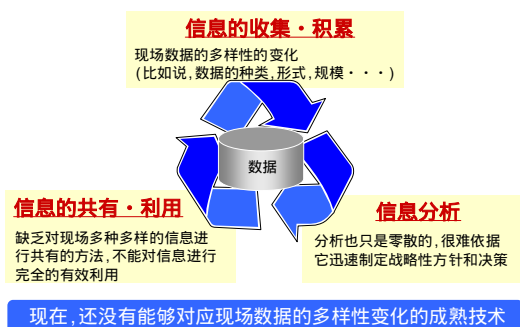


图1 在信息的有效利用方面存在的问题

1.3 对变化应对自如的「XML型」数据结构

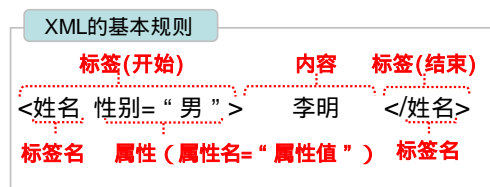
如上所述,在对信息进行有效利用方面存在很多问题。作为解决这些问题的方法,被称为「XML(eXtensible Markup Language)型数据」的数据形式倍受瞩目。XML型数据具有擅长应对数据结构变化的特长,因而成为解决问题的关键。

首先,XML型数据是以文本形式进行记述,适合于在多个平台的环境中进行数据交换。例如,企业使用文档作成软件和表格等各种办公室软件工具作成各种各样的数据,然后,各自以独自的二进制形式进行储存。所以,为了确保用其它工具对储存的数据进行利用,必须将这些数据转换为文本形式。数据被转换为文本形式之后,即使在不同的平台上也能进行读解。XML型数据正是采用文本形式进行记述,从而使跨平台的数据利用成为可能。

此外,XML型数据利用标签(tag)来记述数据,既能够保证在不改变数据的意义和结构的情况下进行数据交换,又能灵活地对应规格的变更或不同系统之间的数据交换。

XML型数据具有卓越的表现力。它使用标签来记述数据,标签的名称,阶层结构等等可以由使用者自由定义。因而,增设新的数据项时,只要在所要记述的数据项内容前后加上作为数据项的标签,就可以自由地进行数据项目追加。如上所述,XML型数据是一种擅长应对结构的变更和扩展的数据形式,因此,人们对其在信息的有效利用领域所起的作用寄予很大的希望。但是,XML也有缺点。XML型数据为文本形式,并且使用标签的记述方式,比二进制数据冗长。在数据量庞大的情况下,数据的利用性能将出现很大的问题。这个缺点使XML型数据在信息的有效利用领域的运用受到限制。

使用标签或属性对数据进行定义的文本信息



擅长应对构造变更或扩张等“变化”的XML型数据

图2 XML型的数据结构

1.4 仅靠RDB已不能适应市场的需求

在目前的信息处理系统中,使用得最多的数据管理系统是关系型数据(RDB: Relational Database)。RDB使用被称为“表格”的二元表格对所有的数据进行储存,结构简单易懂;字段(field)与字段之间用“关系”进行设定,使复杂的数据结构也能进行表现。由于具备这些优点,RDB被普遍使用。另外,RDB通过设定“索引键”,实现了高速查询。

但是,在业务处理多样化的今天,出现很多RDB只能勉强应对或仅靠RDB无法应对的局面。其技术方面的原因在于这种以“索引”及“关系”为必要组成的“表格结构”的本身。

RDB是以严密的结构化为前提的数据库。

也就是说，在设计初期，表格的内容及数据之间的关系不确定的话，就无法进行编程。即使强制进行了编程，设计方面发生的变更也让人难以应对。而且，即使运用中的绝大部分信息为已经构造好的定型信息，随着业务情况的变化，也会产生大量经常发生变化的流动性信息。为了对应这些动态信息而进行的数据库结构变更，应用程序变更等给管理者带来巨大的负担。

索引确实是实现高速查询的有效手段，它在对于定型数据的查询方面具有相当优异的成果。但是，对于流动数据的查询而言，使用这种方法，数据库的设计，开发和维护非常繁琐和复杂，所需费用也很高。

为了解决这些问题，把 XML 型数据直接向数据库登录，并进行有效利用的数据库技术引起人们的关注。确实，利用这个技术，到目前为止难以处理的数据可以简单地得到利用，另外，RDB 的表格及关系的设计和开发，运用中结构的变更等麻烦也可以得到避免。但是，这种方法只是解决了建表问题，而没有解决检索问题。运用这种技术，我们仍然必须用索引键来进行查询。这样，问题并没有得到根本性的解决。能解决这个问题的是不需要使用索引的新型 XML 型数据库技术。

富士通推出把来自于实际应用中的信息进行 XML 化，并使其顺利得到充分利用的信息基础「Interstage Shunsaku Data Manager(以下称为 Shunsaku “瞬索”)」。

瞬索是由院校和企业经过 10 年共同努力而研究成功的突破了传统思维的 XML 型数据库引擎。它采用了文本形式的 XML 型数据技术和在应用方面具有突破性意义的高性能刀锋服务器技术，解决了以 RDB 为核心的现存技术所难以解决的“设计阶段的性能保障、削减开发成本、运行简单”等问题。

瞬索把到目前为止不能进行利用的不同种类或形式的数据库统一为 XML 型数据，使信息得到快速的有效利用。它也适用于在现存系统中要求对查询性能进行改善和实现高速查询的领域。并且，使用瞬索的话，不需要对在不同的事业单位或部门进行单独管理的数据进行再构筑，而是将其 XML 化后，快速地进行统一和联合，实现跨越多个系统的信息系统构筑。瞬索是从小规模到大规模的范围内，对现存 RDB 技术所难以应对的业务领域进行弥补的综合性技术。

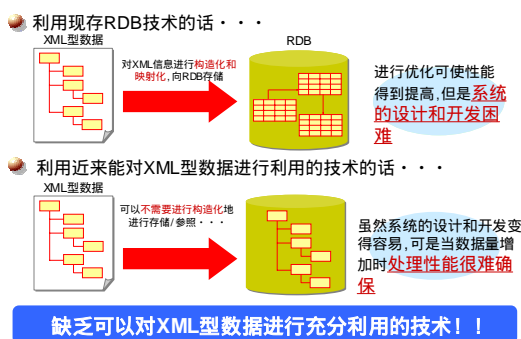


图3 现存 RDB 技术和利用 XML 型数据的技术

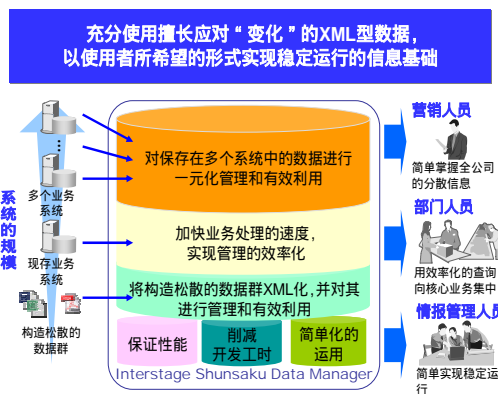


图4 瞬索所提供的服务

1.5 富士通的研究成果

针对当前信息有效利用方面所存在的问题，

1.6 对 IT 技术进行最大限度的有效利用

随着处理器技术的发展,CPU 的性能年年得到提高,存储器的容量也不断地扩大。网络等通信基础设施的传送速度也在持续提高,使宽带网技术不断发展。

但是,IT (Information Technology) 技术并没有因为这些技术革新而得到充分有效的利用,发挥其应有的作用。今天,因特网的普及率已经很高。ERP (Enterprise Resource Planning), CRM (Customer Relationship Management), SCM (Supply Chain Management) 等 IT 技术对企业提供了强有力的支援。但是,真正导入了这些系统的企业并不多。B to B (Business to Business), B to C (Business to consumer) 等 EC (Electronic Commerce) 的普及率也并不高。

导致这种情况的主要原因是高速发展的技术革新和实际运用之间所存在的差距。相对于 IT 的运用方面的水准而言,技术革新的发展速度要快得多,而且,二者之间的差距仍在不断扩大。

为了打破这种局面,必须开拓新思路,使技术革新成果得到充分利用和发挥。例如,为了提高应用软件的生产性,一直以来就有把软件进行零部件化的想法,但是由于存在运行速度不足或不能确保必要的性能等问题,从前没有得到实现。在 CPU 性能大幅度提高的今天,由于彻底地实现了软件的零部件化,生产性得到大幅度提高,大大地降低了成本。在可靠性方面,不仅是在软件上下工夫,在硬件方面也应采取有效措施,比如说对系统进行双机热备等。

也就是说,在今后的信息系统中,为了对飞速发展的 IT 技术进行有效的利用,从而发现新的价值和可能性,必须在硬件和软件两方面双管齐下。

除了使用 XML 型数据技术的研究成果之外,瞬索还采用了高密度型的服务器---“刀锋

服务器”,以保证性能和降低成本。刀锋服务器是指在同一个机箱中,装入多个具有服务器功能的板状物(刀锋服务器),数据可以被分散到这些刀锋服务器上进行并列处理。这样,能够根据数据规模简单地进行扩展,实现了以低成本获取高可靠性的服务器系统。

瞬索把实现了高速查询的软件技术与有效地利用了刀锋服务器的硬件技术结合起来,给信息系统带来具有革新意义的突破。

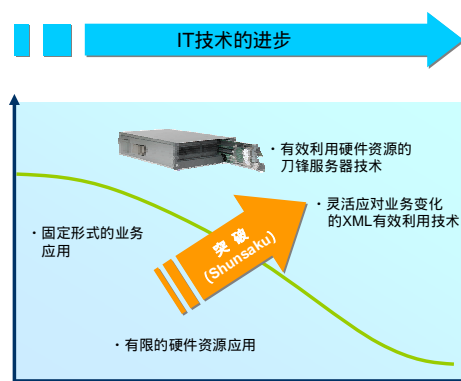


图5 有效利用了 IT 技术的 Shunsaku 技术

第 2 章 瞬索的特长

2.1 给信息系统带来革新的技术

使用以往的 RDB 技术, 为了实现必要的响应速度和性能, 通常需要对表格和关系反复进行设计。并且, 优化处理及实际测试所花费的开发成本也非常庞大。即使完成了编程及系统安装, 在数据库的运行中, 数据的增加和转移, 故障的应对处理, 以及对不断增加的新数据的对应也让人头疼不已。

Shunsaku(瞬索)为您解决使用以往的信息系统所不能避免的这些难题: Shunsaku(瞬索)可以在设计阶段保证性能, 从而迅速向开发阶段进行转移; Shunsaku(瞬索)不使用索引, 所以也不需要反复进行优化处理和性能测试; 由于使用了刀锋服务器技术, 运行开始后, 能够随着数据量的增加, 灵活地应对数据的增加、转移、以及故障。对于新增项目, 也不再需要一次又一次地进行性能测试, 而只要把必要的项目内容作为 XML 型数据进行追加, 就能迅速完成对应。如上所述, Shunsaku(瞬索)给现存信息系统带来了革新。

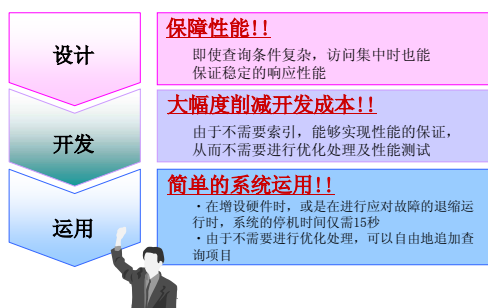


图 6 信息系统的生命周期(Shunsaku)

2.2 Shunsaku(瞬索)的全貌

Shunsaku(瞬索)是由两种服务器所构成的 XML 型数据库。这两种服务器是: 把查询对象

的数据作为 XML 型数据进行储存, 受理查询要求, 并进行访问文件管理和结果应答的“主导服务器”, 和接受主导服务器的要求, 实行高速查询处理的引擎部分的“搜索服务器”。搜索服务器有效利用了刀锋服务器的功能, 把数据分散到多个刀锋服务器上, 实行并列查询处理。构成刀锋服务器的最小查询单位是一片一片的刀锋。Linux/Windows 版的 Shunsaku(瞬索)可以把主导服务器和搜索服务器配置在同一个服务器中。

因为 Shunsaku(瞬索)是由这些不同形式的, 高性能的系统所构成, 所以它可以灵活对应从低端(MB)到高端(TB)的数据规模, 快速构筑起“对来自现场的第一手信息进行有效利用”的体系。

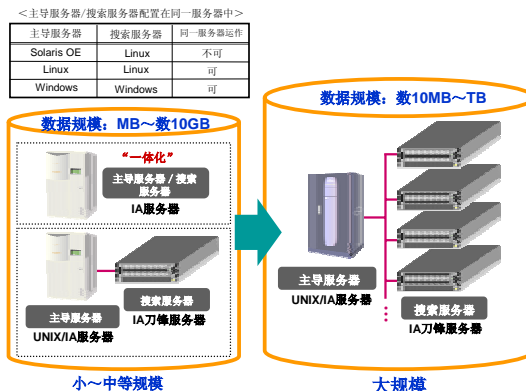


图 7 Shunsaku 的全貌

2.3 设计

不需要进行结构化的设计方式

一般来说, RDB 是以严密的数据结构设计为基础, 所以设计之前必须事先掌握全部数据。技术人员必须首先设定最初所管理的全部的数据类型, 并把数据项作为字段进行设定, 对所储存的数据的尺寸(字节数)也必须进行严格设定。

在 RDB 表格间的关系设计方面, 为了避免在同一表格内出现重复信息, 各个表格之间通过唯一的区分符(主键)直接关联各数据。这就是所谓的 RDB 的“正规化”的构成设计。正规化确保 RDB 中经常使用的处理或影响性能的处理。技术人员一边测定键值所对应的数据量和访问性能, 一边继续对数据反复进行正规化, 一直到第 n 次。此外, 为了实现处理性能的高速化, 对于性能上的重要数据项, 必须设定索引。但是, 如果只是一味地增加索引, 会给更新等处理带来弊端。即使对于经验丰富的技术人员来说, 找到索引设计的最佳方案也是一件相当繁琐而艰巨的任务。这样, RDB 的严密结构性本身导致了数据库设计的复杂化。设计阶段的工作进行得不完善的话, 在以后的运行中, 进行项目追加时, 或是想要追加比事先设定的规模量大的数据时, RDB 的复杂结构将使其难以灵活应对这些局面, 而必须进行数据库的重新设计。

Shunsaku(瞬索)所管理的数据均为 XML 型数据结构, 所以不再需要对可变长的数据和数据项, 以及重复的数据项进行设计处理, 而是灵活地吸收数据。因而, 在设计数据库时, 不再需要进行表格和字段的设计, 以及正规化处理等, 大大地减轻了负担。

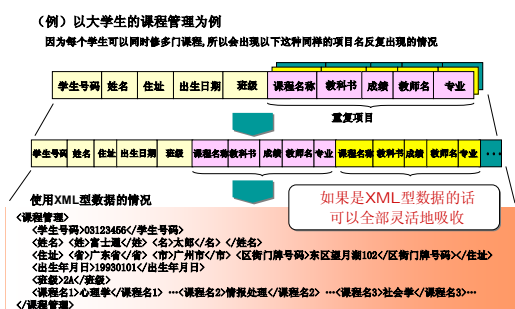


图8 不需进行结构化的 XML 型数据

设计阶段的性能保证

RDB 在严密设计的基础上, 经过反复的优

化处理及实际测试, 而达到一定的使用性能。也就是说, 不进行优化处理和实际测试的话, 就不能确保系统的响应性能。

但是, 由于 Shunsaku(瞬索)的系统构筑采用了把所有的数据从头开始以一定的速度按顺序进行读入的简单的“全文查询方式”, 所以不再需要严密的设计过程。一般来说, 不设定索引的话, 全文查询的查询速度会比较低, 但是, Shunsaku(瞬索)由于采用了「SIGMA 技术」, 「高速多载运行技术」, 以及「刀锋服务器技术」, 所以, 即使是在查询条件复杂或访问集中, 查询量大的情况下, 也能保证稳定的查询性能。

对多项目·多方面的查询进行对应 (SIGMA 技术)

Shunsaku(瞬索)使用的 SIGMA 查询技术, 不论是在只有一个查询条件时, 还是在使用了“AND”, “OR”, 或“NOT”等一大堆连接词的复杂查询条件时, 都能实现稳定的查询性能。无论查询关键词有多少个, 都能通过 SIGMA 技术把这些关键词合成起来生成 automaton, 这个 automaton 对查询对象数据依次进行匹配, 并读入符合查询条件的数据。所以, 无论查询条件多么的复杂, 也与只有一个查询条件的时候一样, 只要对所有查询对象数据进行一次查询处理, 就能实现无遗漏的, 稳定的查询。这个查询技术是企业 and 院校经过十年共同研究和开发的成果, 并已在运用方面取得了相当的成绩。在使用这种技术的查询系统上, 对含有 100 多个数据项的合成查询条件进行多方面的查询时, 所需要的时间是使人惊叹不已的“1 秒钟之内”!

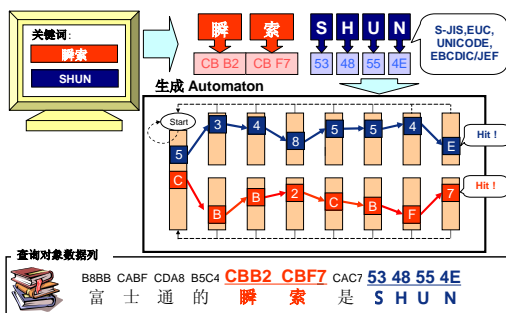


图9 SIGMA 技术模式图

访问集中时的对应
(高速多载运行技术)

高速多载运行技术是在使用者多，查询要求量大的情况时，也能保证稳定的查询性能的技术。Shunsaku(瞬索)通过把大量的查询要求进行一体化处理，可以一次性地完成查询处理，并把查询结果分散反馈给应用程序。因此，即使是在使用人数多，使用者不确定的因特网的环境中，也能确保稳定的查询性能。这与以往使用 RDB 的查询系统相比，格外显得成果显著。

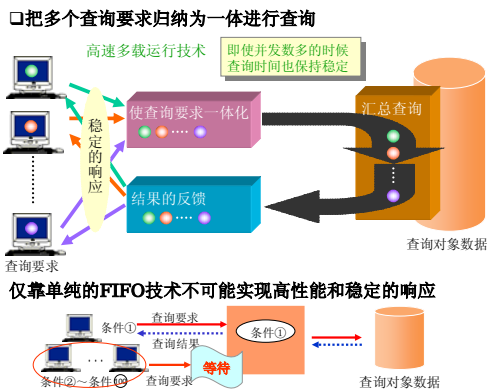


图10 高速多载运行技术模式图

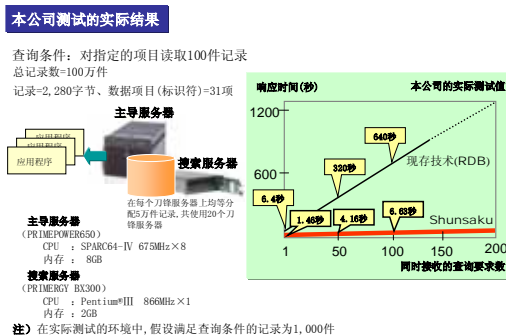


图11 并发访问性能的测定结果 (本公司的测定结果)

对数据量的增加进行的对应
(刀锋服务器的有效利用技术)

刀锋服务器的有效利用技术是指在查询处理时，数据自动地分散到多个刀锋服务器上同时进行并列处理，所以，即使是在处理大量的数据时也能保证稳定的查询性能。在以往的系统里，把 CPU 的计算能力全部有效利用到数据的处理上本来就是不可能的事情，而且，即使把 CPU 数增加到两倍，处理性能也不一定增长为原来的两倍。Shunsaku(瞬索)通过使用 SIGMA 技术和高速多载运行技术，可以在各种情况下保证稳定的查询性能，只要将数据分散到各个 CPU(刀锋服务器)上,通过这些 CPU(刀锋服务器)的同时运作，就可以实现全 CPU 的同时查询处理。也就是说，当 CPU(刀锋服务器)数增加为两倍时，查询性能也增长为原来的两倍。这样，利用刀锋服务器技术，只要根据查询数据量和查询时间的要求，简单地增设 CPU(刀锋服务器)，就能使性能得到保证。

> 查询时间只需要以往查询方式的1/刀锋服务器数

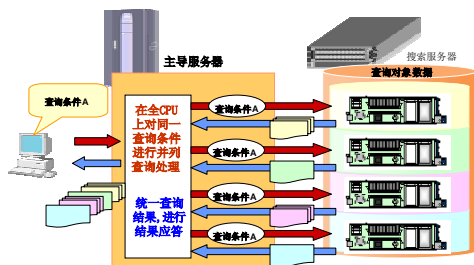


图 12 刀锋服务器的有效利用技术

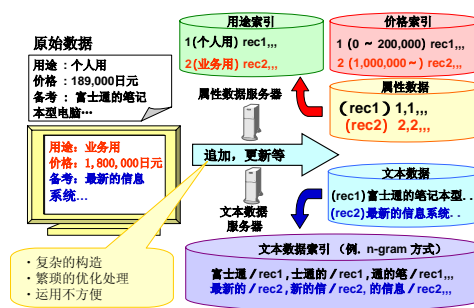


图 13 数据的登录 (使用 RDB)

2.4 开发

不需要索引

在开发商品信息查询系统时，以往我们常常使用索引键来建立一个查询系统。在这个系统中，用 RDB 对商品的用途或价格等信息进行管理，用文本 DB 实现对各种文字列的全文查询功能。这里使用索引键的目的是为了确保一定的查询性能，但是为了对商品用途及价格的属性信息建立索引键，并确定索引关系，还必须对这些属性信息建立文本索引。对文本数据也必须抽出关键词，建立索引。

Shunsaku(瞬索)由于使用 XML 型数据，从数据的登录，到运用中的追加，删除和更新都可以用简单的文本操作方式进行。在性能方面，由于使用了不需要索引的全文查询方式，使性能得到了保证，所以不再需要进行 RDB 所必不可少的优化处理。Shunsaku(瞬索)在性能方面的优越性是:只要根据查询数据的规模，设置必要的 CPU，在硬件方面采取措施就能解决问题。刀锋服务器的有效利用使得这种“在硬件方面解决问题”的作业相当容易。

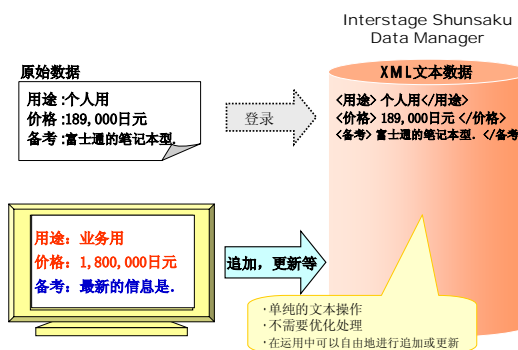


图 14 数据的登录 (使用 Shunsaku)

用导入简单的 GUI (Graphical User Interface) 操作方式进行设置

Shunsaku(瞬索)使用简单的 GUI 操作方式就可以完成安装和配置(标准配置)。配置方法可以选择以下两种：一种是使用一台服务器的系统构筑方式；另一种是把搜索服务器与主导服务器分开配置的系统构筑方式。利用 GUI 进行配置后，立刻就可以开始运行，使以往繁琐的配置工作大大得到简化。对其它的系统构筑方式，可通过 GUI 操作，先进行导入，然后再以手动方式按照自己的需要对配置进行详细设定。所以说瞬索可以对应各种不同的系统构筑方式。



图 15 配置画面

2.5 运用

现有资源的 XML 化

Shunsaku(瞬索)提供把 CSV 形式的数据库数据，或是储存在 RDB 中的数据转换成 XML 型数据输入文件的标准工具。这个工具能够定义变换对象的数据项目和 XML 数据的标识之间的关系，并方便地将变换对象的数据转换成 XML 数据。RDB(Symfoware, Oracle 等)中所储存的数据的 XML 变换可通过 JDBC 来进行。

数据变换应用程序用 FTP, JDBC 等调出需变换的数据，再调出提供 XML 型数据变换工具的 API，然后利用变换工具把数据输出为 XML 数据(文件)。这些 XML 数据可以立刻被 Shunsaku(瞬索)所利用。

这样一来，企业现有资产的 CSV 形式数据和储存在 RDB 中的数据就可以通过 Shunsaku 得到有效的利用。

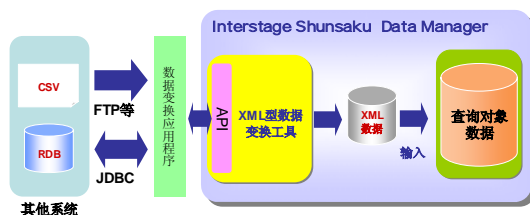


图 16 XML 型数据变换功能

简单的运用方式

Shunsaku(瞬索)通过使用刀锋服务器，实现了简单的系统运用。刀峰服务器的易扩张性保证了数据库在数据量增加时的响应性能。另外，Shunsaku(瞬索)还具备应对故障的自动退缩运行功能，可以保证故障发生时的持续运行。

Shunsaku(瞬索)在 CPU 增设或故障发生时，停机时间仅仅为 15 秒钟。而且即使是在这 15 秒钟之中，应用服务器仍然在接受查询要求，所以从实际意义来说，整个服务并没有完全停止。综上所述，Shunsaku(瞬索)不需要预备应对数据增加或故障的备用机。Shunsaku(瞬索)能使您以最小的投资，开始系统运行，并使系统逐步地进行扩展。

自动增设

Shunsaku(瞬索)为了维持数据增加时的性能，采用了增设 CPU 的方法。只需要向机箱中插入新的刀锋服务器，就可以使增加的数据向新增设的 CPU(刀锋服务器)上自动配置，而不需要停止数据库的运行。

- 只要增设刀锋服务器就能对应数据量的增加
- ▶ 即使是在增设过程中数据的查询服务也不会停止

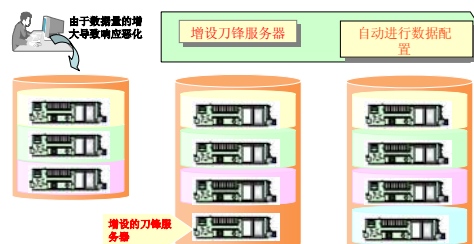


图 17 自动增设

自动退缩运行功能

当某个刀锋服务器的 CPU 发生故障时，它会自动地从整个系统中分离出去。并且，数据也会自动地从故障 CPU 中退缩出来，向正常的 CPU 上进行再配置。

□刀锋服务器发生故障时的自动退缩运行
>不需要应对故障的预备机

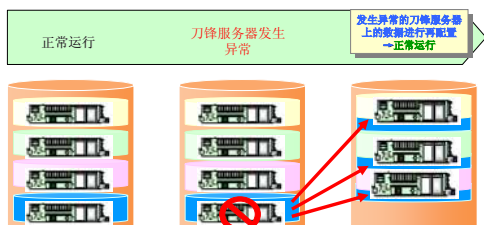


图 18 退缩运行

开发·运用成本大幅度降低

Shunsaku(瞬索)根据 XML 型数据和刀锋服务器的有效利用技术, 实现了 RDB 技术所不能实现的简单的开发和运用方式, 大大降低了开发和运用成本。

开发成本

- 不需要索引的开发方式使 Shunsaku(瞬索)不再需要反复进行优化处理和性能测试, 大大缩短了工时。
- 通过使用 GUI 操作方式实现了快速安装和配置。

运用成本

- 不用制作应用程序, 而是把现有资产进行 XML 化。
- 在运用方面, 不再需要进行到目前为止必不可少的定期性的索引再制作等需要停止系统的维护处理。
- 因为 Shunsaku(瞬索)使用 XML 型数据, 所以容易进行数据追加, 也不需要随着数据的增加进行优化处理和性能测试。
- 由于采用了自动增设和自动退缩运行功能, 不再需要准备应对转移和故障的备用机。
以数据规模为 20GB 的信息查询系统的构

筑为例, 使用 RDB 的话, 开发周期为 4 个月, 可是, 使用 Shunsaku(瞬索)的话, 开发周期只需要 1 个月, 也就是说使用 Shunsaku(瞬索)时所需开发周期为使用 RDB 时的四分之一。

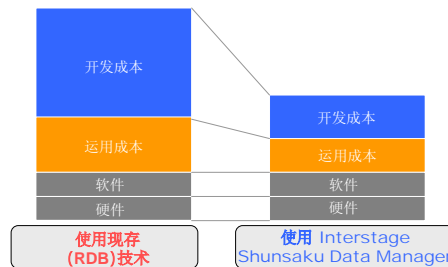


图 19 削减开发·运用成本

2.6 应用程序开发

通过应用程序实现远距离操作

Shunsaku(瞬索)可以开发利用 Java 语言和 C 语言的业务应用程序。

开发应用程序时, Shunsaku(瞬索)通过 Java 语言或 C 语言作成的应用程序, 提供了在远程操作下的高速查询、合计以及更新功能的 API (Application Programming Interface)。更新功能包括两种方式, 一种是要素单位的更新, 另一种是批量更新。Shunsaku(瞬索)提供了高效率的 API 接口, 所以应用范围非常广泛。

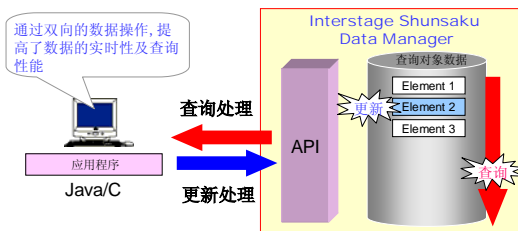


图 20 应用程序利用模式图

有利于快速实现信息的有效利用的统计功能

Shunsaku(瞬索)提供有利于快速实现信息的有效利用的排序和统计功能。例如, 它提供把 Shunsaku(瞬索)的查询结果中的任意的项

Technical Whitepaper
「Interstage Shunsaku Data Manager V6」

目(标识符)进行重新排列的排序功能。排序处理功能把搜索服务器的查询结果在主导服务器中进行重新排列。

此外, Shunsaku(瞬索)还可以通过应用程序取得任何项目(标识符)的合计值, 平均值, 最大值或最小值等统计结果。Shunsaku(瞬索)通过利用刀锋服务器的并列处理技术, 可以在刀锋服务器中的多台机器(BLADE)上并列进行统计处理, 然后再把得到的多个统计结果在主导服务器中进行再统计, 从而实现了高速的统计处理。

Shunsaku(瞬索)通过支持排序和统计功能, 提高了业务应用的开发效率, 有利于快速实现信息的有效利用。

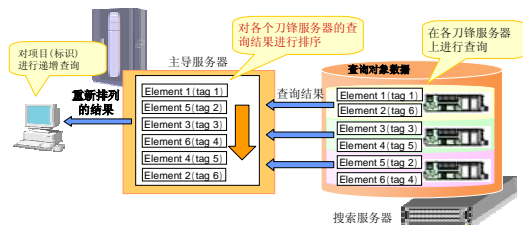


图 21 排序处理模式图

第 3 章 业务应用模式

3.1 Shunsaku 的应用模式 (概要一览)

Shunsaku (瞬索) 打破了以 RDB 技术为基础的以往的信息系统的常规, 在保证性能的同时大大削减了开发工时。它通过简单的运用方式, 实现了从小规模到大规模的业务领域内, 到目前为止难以实现的业务运用。

以下为具有代表性的 Shunsaku (瞬索) 应用模式:

应用模式	应用领域
改善查询性能	品质管理, 顾客支持, 商品的主要数据查询
自由查询 (横向查询多个 DB)	专利信息查询, 品质管理, 顾客记录查询
并发数多的查询	就业信息查询, 各种公众活动信息查询, 招生信息查询, 各种入场券的贩卖信息查询
多项目·多方面的 条件查询	院校招生信息(学部, 学科, 科目等信息的组合)查询, 人材交流中心或工作介绍所的信息管理
顾客信息 管理·分析	客户关系管理(CRM), 销售支援(SFA)
多种商品的 主要数据查询	商品查询数据库
改善查询品质	检查管理系统
在 CRM 方面的应用	电话呼叫中心/联络中心, 信息管理

图 22 具有代表性的 Shunsaku 应用模式

3.2 Shunsaku 的应用模式 (详细介绍)

改善查询性能

使用现存系统时, 由于预料之外的数据增

加或查询条件复杂等情况而导致系统响应恶化, 并给业务带来障碍时, 瞬索可以助您一臂之力。

不需改变现存系统的运用方式, 只要把查询系统更换为 Shunsaku (瞬索), 就可以使查询性能得到提高。Shunsaku (瞬索) 所采用的刀锋服务器技术保证了数据量增加时的稳定查询性能, 所以, 不需要因为数据的增加而进行系统的重新构筑。通过把瞬索与现存数据库联合使用, 使之各司其职的运用方式, 实现了瞬索和现存系统的共存, 并且, 因为不需要建立索引以及进行优化处理, 大幅度地缩短了开发周期。

这种应用模式, 通过把瞬索与现存系统结合使用的方式, 提高了响应性能, 实现了查询的高速化, 从而避免了由于响应的恶化给业务处理带来的不良影响。

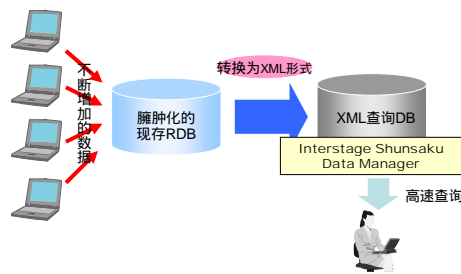


图 23 改善查询性能的模式

自由查询 (横向查询多个数据库)

在多个系统同时运作的情况下, 运用瞬索可以把分散储存在各个系统中的相同类型的信息进行统一处理。

瞬索可以从多数个系统中抽出所需数据, 并将其转换为 XML 数据, 进行一元化管理, 从

而实现横向查询。由于现有系统照旧运作，所以也不会产生因为系统的构筑所引起的业务现场混乱。

使用这个自由查询模式，可以实现以往系统所难以实现的“对储存在多个系统中的数据群进行横向查询”，有利于从多个角度实现快速的信息有效利用。

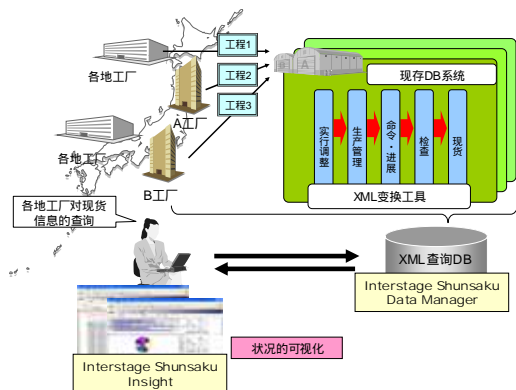


图 24 自由查询 (横向查询多个数据库)模式

并发数多的查询

当新的促销活动出台，或是换季时节到来时，使用者对因特网的访问在短时间内大量集中。为了避免访问集中时数据库查询性能的下降，可以借 Shunsaku (瞬索) 来助您一臂之力。

Shunsaku (瞬索) 所具有的“将多个查询要求汇总后统一处理”的特长，使其可以保证稳定的响应性能。此外，它还具有对不同形式的数据库进行汇总查询和不需要建立索引等优势，所以可以轻松自如地应对短时间内的大量数据更新。

这个模式最适用于在因特网等使用者数不确定的环境中，保证高通量情况下的响应性能；也适用于必须定期进行大量数据更新的信息系统。

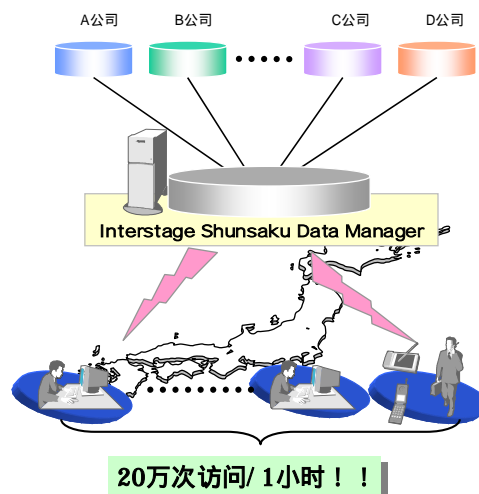


图 25 并发数多的查询模式

多项目·多方面的信息查询

使用现存查询系统时，当查询条件为 AND、OR、NOT 等条件组合的复杂查询时，会影响系统的响应性能。Shunsaku (瞬索) 可以帮助您解决这个烦恼。

使用 Shunsaku (瞬索) 时，无论查询条件多么复杂，都只要用一次全文查询就能完成查询处理，保证稳定的响应性能。另外，在 RDB 的查询中，对 SQL 语句不能表述的查询条件，在 Shunsaku (瞬索) 中也能进行高速查询。所以，Shunsaku (瞬索) 对于到目前为止在 RDB 中受到限制的条件，也可以用任意的关键词进行查询。

使用这个模式，能够轻松实现高速的查询性能，有利于提高信息系统的可用性和减轻业务负担。



图 26 多项目·多方面的信息查询模式

顾客信息的管理和分析

在现存系统，由于输入数据的长度或个数等受到限制而导致形式或格式不确定的数据信息的输入或利用发生困难时，Shunsaku 可以助您一臂之力。

Shunsaku (瞬索) 采用的是 XML 型数据形式，可以不受数据的长度或个数等限制地把数据进行直接输入。所以，它能够高效，灵活地处理不同形式的数据组合、可变化的项目名、以及循环项目等冗长的数据组合。而且，维护也变得更加容易。

这个应用模式能把 RDB 不能处理的不同种类和形式的现有文件转换为 XML 型数据形式，并在数据库对其进行一元化管理，从而使数据的查询和分析更加准确和丰富。

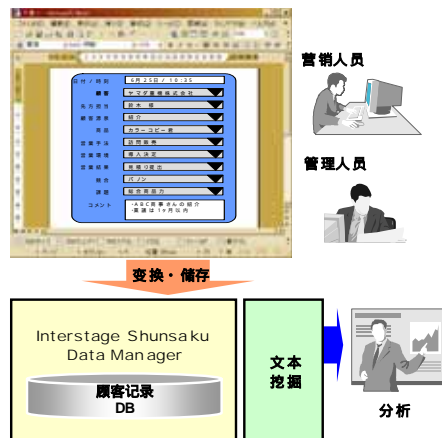


图 27 顾客信息的管理·分析模式

多种商品的主要数据查询

使用 RDB 的话，根据业务的变化进行查询项目的追加时，必须进行索引的再建等优化处理和性能测试。Shunsaku (瞬索) 可以为您免除这些麻烦，减轻您的作业负担。

Shunsaku (瞬索) 采用的是 XML 型数据形式，可以简单地进行数据项目的变更。而且，由于 Shunsaku 使用了不需要索引的全文查询方式，所以也不需要进行优化处理。在整个系统中，充分发挥 Shunsaku 的这些特长，用它来应对流动性的数据变更，可以实现系统的灵活运用。

对于频繁发生数据追加等变更的系统来说，使用这个应用模式可以大大减轻运用负担。使用这个应用模式，任何人都可以轻松自如地对那些以往由于查询项目复杂而导致运用困难的系统进行操作。

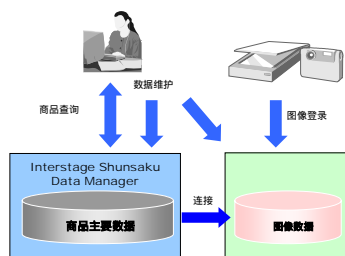


图 28 多种商品的主要数据查询模式

改善查询品质

在品质监督或有关安全事项的业务查询方面，疏忽和遗漏绝对不可容许。进行这些要求高严密性，高品质的信息查询时，Shunsaku 可以充分地发挥其特长，确保高度的查询性能。

Shunsaku 能够对 XML 型数据毫无遗漏地进行全文查询，所以可以保证可靠的查询品质。此外，刀锋服务器的有效利用也使查询性能得到进一步的保障。

在要求高严密性，高品质的信息查询领域，这个应用模式能使业务效率得到飞跃性的提高。

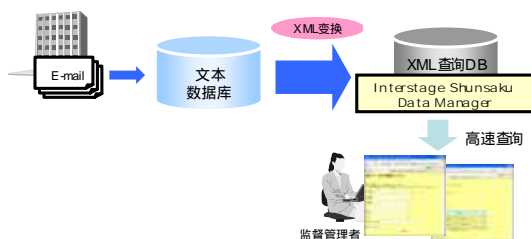


图 29 改善查询品质模式

在 CRM 方面的应用

CRM 系统是把从顾客信息，商品信息以及顾客投诉信息等各方面收集到的信息进行有效利用的系统。Shunsaku 可以帮助您实现 CRM

系统的构筑。

通过利用瞬息，顾客信息，商品信息，以及顾客投诉信息等在不同系统中进行管理的数据可以容易地被统合起来，从而实现跨越多个系统的横向查询。另外，向各操作端等进行项目追加时，利用 Shunsaku，可以灵活地根据现场的需求进行规格的变更。

使用这个应用模式，可以把 Shunsaku 作为综合 CRM 数据库来使用，轻松建成诸如 FAQ (Frequently Asked Questions)，专门技术诀窍以及电子目录的建立和维护之类的业务信息系统，从而实现快速的 CRM 的基础构筑。

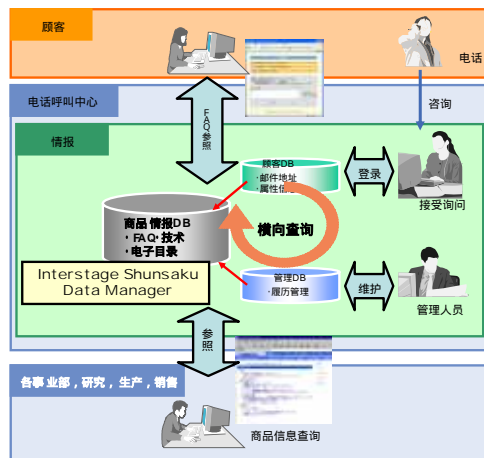


图 30 CRM 的应用模式

第 4 章 今后的研究课题

4.1 关于 Shunsaku (瞬索)V7 的研究

Shunsaku V6 作为信息有效利用的基础，在 RDB 难以应对的业务领域中所起的作用越来越大。Shunsaku V6 的推出强化了 API，推进了信息的有效利用。

在今后的工作中，我们将使 Shunsaku(瞬索)作为 XML 型数据库引擎得到进一步的发展和完善。Shunsaku V 7 将采用查询范围局部化的技术和实际文件虚拟化的技术。它通过采用多重文件的运行方式实现运用的效率化。并且，和商业用 RDB 一样，对 XML 型数据导入事务处理控制功能以确保数据的可靠性。另外，使用 GUI 不仅可以进行安装和配置，还可以进行其他整个操作，从而提供了简单便捷的操作环境。

Shunsaku V7 不仅仅是高度的信息利用基础，而且也是具有高可靠性的数据综合管理基础。

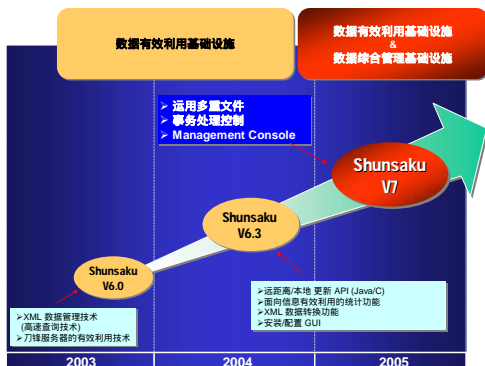


图 31 Shunsaku 的研发进程图

4.2 Shunsaku(瞬索) V7 的主要构成技术 多重文件的应用

为了实现根据不同的目的对 XML 型数据进行管理，Shunsaku V7 支持多重文件的应用。采用在同一个系统中使多重文件并行的运行方

式，可以排除与查询无关的数据，提高查询性能，也有助于实现数据的载入，备份以及恢复的效率化。并且，可以对分割文件进行统一，实现横向查询。

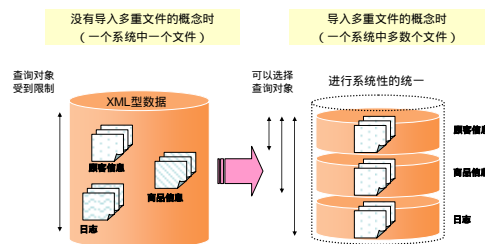


图 32 多重文件的运用

事务处理控制

Shunsaku V7 为保证应用程序处理的完整性提供了事务处理控制 API，以每一个元素（相当于一条记录）为单位进行加锁保护和日志管理。并通过支持回滚处理（恢复到备份时状态）、故障停机恢复（防止在处理过程中出现数据的不匹配）、介质恢复（恢复到最近点）等功能实现高信赖性的数据运用。

综合操作环境 (Management Console)

Shunsaku V7 提供 GUI 综合操作环境 (Management Console)，在整个操作过程中，均提供 GUI 的操作支持。

例如，在构筑与其他的 Interstage 产品关联的系统时，使用一个 GUI 视图就可以进行多个产品的环境设定、服务器运用、文件运用和运行监视等操作。而且，用一个 GUI 视图还可以对在多数个服务器中进行运用的 Shunsaku 进行操作。