



实际使用评测 报告

私有云和大数据分析创造业务价值

劲霸男装虚拟化和BI方案使用评测

作者：王丛 (Kim Wang)：中桥调研咨询总经理兼首席分析师

杨恩博：中桥调研咨询分析师

彭雅芳：中桥调研咨询助理分析师

2012年6月



目录

简介	3
背景	3
劲霸男装	4
劲霸简介	4
IT 挑战	4
劲霸虚拟化和云计算部署	5
中桥实验室对劲霸 IT 中心的评测	5
虚拟化提高应用部署效率	5
Greenplum 大数据分析提高商业智能效率	6
VNX 实现统一存储	13
Unisphere 实现统一集中管理	15
Networker 实现统一数据保护	16
结束语	19

所有商标和公司名称是其各自公司的财产。本出版物中包含的信息是由Sino-Bridges Research and Consulting Ltd.认为可靠的来源提供的，但Sino-Bridges不保证其可靠性。本出版物可能包含Sino-Bridges的观点，这些观点随时间可能会有所改变。本出版物的版权归Sino-Bridges所有。未经Sino-Bridges的明确许可，不得对本出版物的整体或部分以硬拷贝方式、电子方式或其他方式进行复制或将其分发给无权接收它的人，否则都将引起民事损害诉讼，乃至刑事诉讼。有任何问题请联系Sino-Bridges客户关系部：8610 85655510。

简介

自中国改革开放以来，中国零售业和服装制造业一直处于快速增长期。随着近年来国内原材料和生产成本力的持续攀升、跨国企业在中国市场投入的不断提高，以及除中国以外的零售市场的疲软，中国零售业和服装制造业的市场格局在发生着快速的变化。中国已经成为高端品牌零售的销售大国。从零售的角度来看，中国零售和服装业与跨国公司相比，信息化投资占总资产的比例平均不到2%，与国外大企业8%~10%的平均水平相比差距较大。信息化效率正逐步成为中国品牌消费品和中高端零售企业提高产业链价值的决定性因素。中国零售和服装行业普遍存在下列难题：应用无法互联互通、数据不能共享以及无法通过整合和分析各种数据来实现信息化价值；此外，传统以应用作为独立IT部署的模式，信息利用率低下、管理和新增采购成本上升快、业务应用部署效率低、信息化无法充分发挥其对供应链的价值。

在本报告中，中桥结合上述行业特征，对中国休闲男装制造商—劲霸男装的私有云数据中心做了实地评测和采访。了解劲霸是如何通过虚拟化和私有云，解决了数据中心面临的数据孤岛、IT资源使用率低、对业务响应慢等问题，提高了信息采集的实时性和准确性，加快了劲霸对市场的快速反应；同时优化了“信息链、资金链和物流链”，即提高采购、物流、分销、库存、客服信息化等信息链流程，改善了资金链回流和实现物流链的快速周转，显著提高了资源整合效率。劲霸的私有云数据中心将传统的数据“包袱”转换成“取之于民，用之于民”的共赢数据智能平台，给3000多门店提供了信息化支撑。此举不仅提高了劲霸的品牌含金量，也增强了对门店的价值和凝聚力。

背景

企业的数据中心正在成为新时代的“信息电厂”，成为知识经济的坚实基础。而数据则可以说是电厂的“发动机”所在，已经成为企业业务制胜的关键。企业通过有效地收集、分析和处理海量数据，及时获取有用信息，准确预测潜在的业务风险并快速响应新的市场需求，为用户提供优质的服务，从而赢得更多的市场份额。根据ESG 2011年对270家企业IT首要举措的调查数据显示，超过50%的各型企业用户将数据分析排在IT首要措施的前五位。

如何通过信息创造商业价值，是21世纪品牌服装和中高端零售业面临的巨大挑战。每天越来越多的信息被收集和存储起来，这些快速增加的数据量给传统的数据库架构和零售运营系统带来了难题。其中“大数据”（对零售和服装行业通常体现为数据库子集量大、数据分析效率低下）的分析效率是这些行业实现“信息创造价值”的关键。

对于服装行业而言，零售终端数据的分析是服装经营、市场拓展和库存及物流管理的根本。对终端数据进行快速、准确的收集以及整合和商业分析，有助于企业决策层挖掘出有用信息，从而提高供应链各环节的效率和优化企业的投资回报。比如，哪些是主要货品被广泛接受或不被接受的具体原因；哪一类客户是优质客户；如何提升优质客户互补产品的类别和服务；如何通过新品研发来改善产品和服务构成以及生产结构，保证产品在最短周期内为公司创造最大利润空间；同时提高优质客户资源的满意度等；如何根据不同地区的网点特色，以商业智能分析结果为导向，按销售预测准备库存和物流，满足节假日等品牌服装销售高峰期的促销活动；在对销售网点的高峰期提供优化支持的同时，不会带来不必要的库存资源浪费。

近年来，中国服装业集中化程度呈现加速趋势，行业资源配置进一步向优势企业倾斜和集中，市场竞争日益激烈。对于服装品牌企业来说，博弈的焦点集中在掌控有限的优质渠道资源和争夺有限的价值客户群资源。现在的竞争已经不再是企业之间的竞争，而是供应链之间的竞争。这里所说的供应链包含三个链路层，即物流链、信息链和资金链。而信息链是优化物流链和资金链效率的关键。

服装的季节特性十分明显，并且要兼具时尚性、地域性以及功能性。在服装行业，有一条“快、准、很”的准则。快：快速响应流行动态；准：准确把握时尚脉搏；很：很快速、很准确的覆盖所有区域。正是这一原因，使得竞争日益激烈的服装业在管理、订料、生产和销售等各供应链环节中无不突显出“信息”二字的重要和紧迫。

我国服装企业虽然已经采用了大量的信息化技术，但离跨国企业通过IT创造价值尚有很大差距。一是，随着近年来我国企业信息化建设的迅速发展，不少企业已经引进并实施了多个信息化管理系统，包括OA、ERP、DRP、CRM、

EPM、BI等，但这些系统软件互为割裂，容易形成信息孤岛，各业务模块的协同应用比较差，无法通过商业智能（BI）将数据转换为业务指南，无法优化供应链效率；二是，生产商、中间商与零售商之间信息的相对独立，以及分散的零售网点之间缺乏共有的信息交换平台，相关数据无法做到实时交换和共享，久而久之只能形成“越不交流，信息越不畅通；越不畅通，越做不到信息共享”的恶性循环。所以，如何打通信息化供应链，是所有服装企业必须面对的问题。

本报告中，中桥分析师对劲霸男装首席信息官林俊先生进行了采访，了解到这家企业如何采用EMC虚拟化和私有云方案，通过“统一商业智能、统一存储管理、统一保护”来实现IT创造价值。同时，通过EMC Greenplum大数据分析提高了企业内部供应链的效率，实现了“取之于民，用之于民”的门店业务智能服务。该文中还记录了劲霸首席信息官对于选择该方案时的考量和对该方案的评价。以下是他们的使用详情。

劲霸男装

劲霸简介

劲霸男装（上海）有限公司创立于1980年。劲霸男装专注茄克30余年。它用独特设计终结了茄克的单调，从而成为中国高级时尚茄克的领先者，同时引领茄克及配套服饰的研发设计，让休闲装更时尚。现在劲霸男装在全国拥有形象统一、规范管理的品牌专卖店3000多家。2011年，连续8年入选“中国500最具价值品牌”的劲霸男装，以175.65亿元的品牌价值排名50位，继续蝉联中国休闲男装第一价值品牌。

IT挑战

在对劲霸CIO林先生采访过程中，林先生提到：2010年，随着我们业务的高速发展，我们的IT建设（包括IT的系统化和规范化）目标是如何为公司的战略规划、战略实施提供更高效率的IT支撑。IT重点不仅仅在于系统本身的功能和技术，更多的关注如何去贴近我们的业务发展。也就是说，我们整个IT的规划和建设是站在一个新的高度和新的平台上，为业务发展提供一个弹性、可持续发展的IT平台，而不是重复性的建设。为此，我们采用了业内领先的技术，包括云计算平台、大数据分析、统一数据保护等。对劲霸来说，2011年是不寻常的一年，整个的组织变革是有史以来的一个开创性变化，而我们的信息化建设新战略也是在这样一个环境下进行的。

迄今为止，我们拥有的总数据量已经超过20TB。之前，传统以应用为核心，每增加一个应用就需部署一套设备的模式不仅成本高、配置资源使用效率低，同时分散的应用和系统所造成的数据孤岛使得我们难以通过商业智能分析来提高供应链的效率。劲霸通过部署虚拟化优化了IT资源利用率，保证了系统未来的扩展性；并且借助虚拟化将企业IT架构和应用进行了整合，初步建成私有云平台。公司的各种业务关键型应用在这次的项目中会逐步向虚拟化环境迁移，包括SAP。同时，会不断部署各种新应用。我们选择的Greenplum提供了各种业务数据的商业智能（Business Intelligent, BI）分析，保证了信息的价值化；VNX保证我们的各种结构化和非结构化应用对性能和可扩展性以及企业级功能的要求；同时Unisphere直观、统一的管理界面，不仅简化了整个系统的管理，还提高了系统资源的利用率。此外，Networker为我们在虚拟化和云计算演进过程提供了统一的数据保护，简化了对物理、虚拟和云计算多种环境数据保护的管理强度。

我们实施的解决方案有3个主要特征：

1. 各种业务关键型应用（SAP、ERP、CRM、DRP等系统）在虚拟化环境下运行
2. 云计算平台
3. 大数据商业分析平台

而在部署虚拟化和私有云前，劲霸数据中心面对的IT挑战包括：

1. 传统架构无法支撑快速部署新应用的需求
2. 不能通过高效商业智能提高业务运营效率

3. 数据孤岛不能满足品牌服装行业供应链需求
4. 管理资源无法满足业务需求
5. 无法提供统一的数据保护

劲霸虚拟化和云计算部署

2011年，EMC为我们提供了虚拟化数据中心部署计划，并提供了实施和服务。我们2011年基本完成了虚拟化和云计算的第一期部署，整个新数据中心业务构建在虚拟化平台上，上线后已经初步收到成效。同时，我们完成了Greenplum的一期部署，提高业务比对分析效率，初步实现通过信息创造价值。

我们部署虚拟化和私有云提高应用部署效率、资源利用率和业务部署能力。结合Greenplum，实现了四个“统一”：

1. 统一BI，对各种应用、物理和虚拟环境数据实现高效数据分析
2. 统一存储，满足各种应用对存储的需求
3. 统一管理，提高IT资源管理效率
4. 统一灾备，提高数据保护效率、降低数据保护成本

中桥实验室对劲霸IT中心的评测

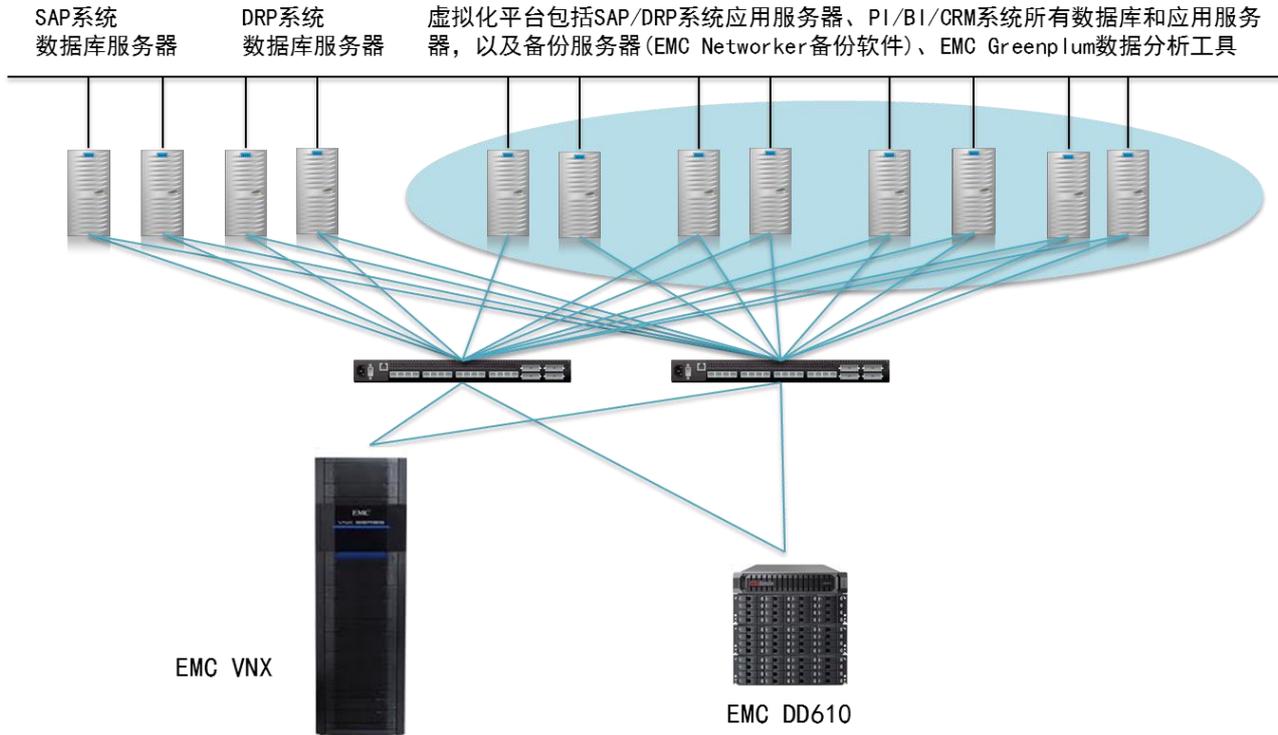
虚拟化提高应用部署效率

我们的IT环境比较复杂，拥有众多应用系统。各应用对硬件资源、操作系统版本、数据库版本的要求不一样；应用实施的厂家对操作系统、数据库的熟悉程度不一样；应用开发商对系统资源需求的规范也不一样，经常到应用实施时才发现硬件资源的不匹配。因此我们每部署一个新应用，都需要花费大量的时间来了解应用需求并反复测试。2011年开始，我们借助EMC专业的整体虚拟化数据中心的规划服务以及虚拟化实施阶段的评估和咨询服务，对内部IT面临的难题、未来数据中心的承载以及不同技术和方案作了全面分析，最终选择了EMC虚拟化技术来构建虚拟化数据中心以实现资源和业务整合，并逐步过渡到私有云平台。

2011年我们实施了一个“6纵6横”的ERP项目群。这意味着一群项目一起上线。这在服装行业也是首例。“6纵6横”包括OA、BI、CRM、DRP（分销资源管理）、EPM、ERP等几大应用，还加上配套的6个小项目同时上线。在具体实施时，我们本来打算分块进行后台的系统建设，但硬件采购预算远远高出我们的计划。采用EMC的方案，我们可以对上述众多系统的后台建设进行整合，这样不仅整个资金投入减少，以后每个系统上线的进度也会显著提高；同时还显著提升了资源利用率，让我们在未来不增加存储采购成本的前提下，可以上线更多应用。云计算平台也是在这样的背景下产生。在第一期项目执行时，除了一部分核心应用，例如SAP数据库和DRP的数据库是放在物理机上，其他的生产平台都移到了云计算平台上。

我们的IT架构图如下（图1）：我们的IT环境包括物理环境和虚拟环境。物理服务器包括IBM P550小型机和HP DL580 PC服务器。小型机用于运行SAP系统数据库和DRP系统数据库，而其它的各种应用如ERP、PI、BI、CRM、PLM、WMS、EHR、ETS等系统的数据库和应用服务器，以及备份服务器等，都是运行在VMware虚拟化平台上，虚拟机约为70多台。目前10多个应用如果采用传统方式需要40台服务器，但采用虚拟化方式，我们只需一个刀箱、16个刀片以及匹配相关软件即可，这不仅大大降低了总拥有成本，同时为未来应用部署提供了可用服务器资源。我们通过EMC Greenplum数据分析工具，对各种业务数据进行商业智能分析；同时，通过Greenplum多用户并行分析处理能力为门店提供BI服务。并且，采用EMC VNX作为统一存储平台，并通过Unisphere统一管理界面，实现所有应用数据的存储和管理。此外，还采用EMC NetWorker备份软件和Data Domain虚拟带库，对物理和虚拟环境下的各种应用和数据进行灾备，来实现数据的统一保护。

图1. 劲霸IT 架构



虚拟化的部署，提高了我们所有硬件资源的利用率，让我们能将更多可用资源用于部署各种应用。以前为某个应用进行资源配置，都是按服务器数量来计算，现在可以精细到CPU core和内存的分配，做到物尽其用。我们部署完云平台之后，不仅系统性能得到了保障，系统的可靠性、效率和可管理性都得到了优化。例如，现在我们上线一个系统并完成整个后台的部署，只需要两三天，简单应用可以在几个小时内完成部署，之前则需要半个月至一个月。并且可复用性、可扩容性以及整体成本（直接和间接成本）都有所改善。图2显示了现有数据中心已部署的各种应用，SAP运行在IBM P550（AIX OS）小型机上，可以提供高性能存储以保证性能需求。同时，一些主要应用和虚拟化运行在IBM P520（支持AIX和Linux）上，同时满足商业查询和分析的性能要求。

图2. 劲霸数据中心应用部署

Storage Group Name
blade13
DRP_d1580
P520
SAP-P550
vmwaregroup1
vmwaretest

Greenplum大数据分析提高商业智能效率

目前，我们拥有3000多家门店。原来的BI系统主要为总公司服务。数据从各店的终端抽取给总部进行分析，然后只有总部的分析人员和决策人员去看这些数据。这样所存在的一个最普遍的问题就是，抽取数据的有效性、准确性和实时性很难得到保证。今年，我们在建设BI的时候，决定双管齐下，在业务层面采取一种“取之于民，用之于民”的策略。数据从终端抽取之后，数据的使用者范围扩大到全国的门店。这样以来，数据的提供者同时也是使用者，以此提高终端数据有效性和准确性。在技术层面加大支持力度。原来只有总部的决策人员去分析数据，BI系统对大数据的并发性处理以及大数据的复杂性要求不高。但是现在，我们需要应对全国几百家店甚至数千家店同时进行数据的上传、快速查询、汇总及报表等，这就对大数据多用户的并行数据导入和查询性能提出了较高的要求。我们希望利用全国的门店和渠道一起来建设并维护这个数据。

我们已经部署了Oracle、DB2、SQL等数据库来满足业务对OLTP交易数据的处理。但是这些数据库在进行大数据分析时，往往需要很长的时间且消耗较多的资源才能完成一个业务指标的分析，甚至有时候即使耗费了大量的精力也无法进行数据抽取和比对。这不仅效率低，对业务的正常运行影响很大，还限制了我們实现“通过IT创造价值”的目的。

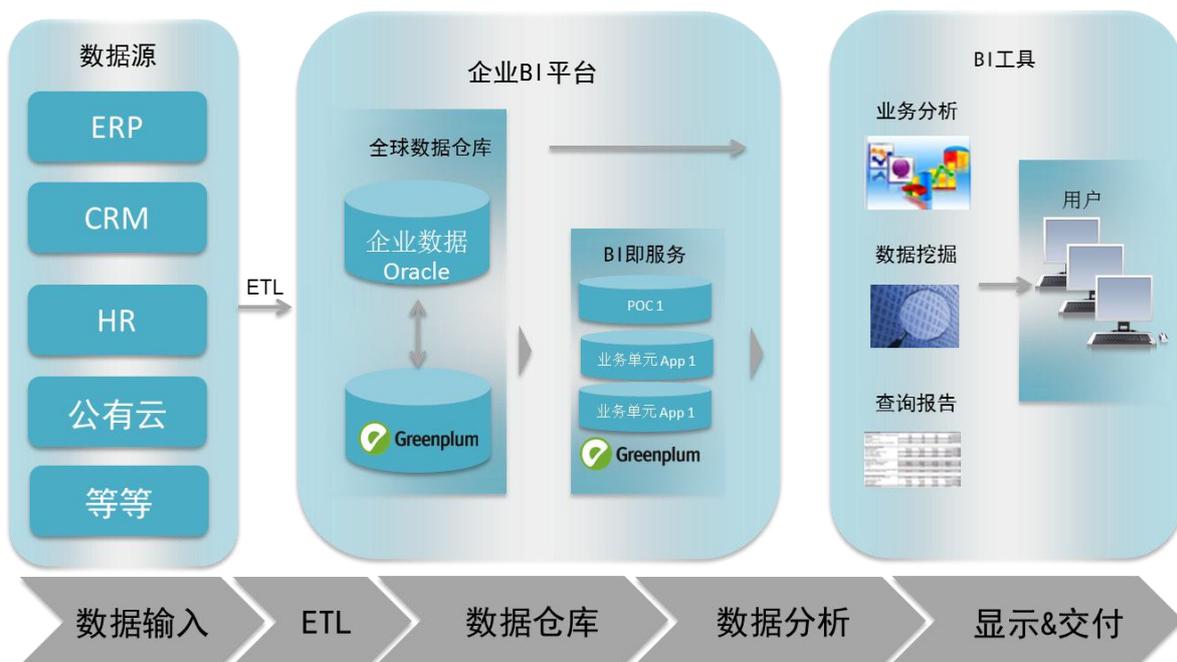
我们在选择大数据分析技术平台的主要指标包括：

1. 满足大数据快速智能分析，提高各种供应链的价值优化
2. 能支撑我们“取之于民，用之于民”的私有云BI战略
3. 高可靠性和高可管理性

在数据管理的过程中，数据最大的价值是完整真实地体现业务需求，通过整合和商业智能分析创造价值。我们部署私有云最主要的原因就是通过IT创造价值。之前，我们的商业分析效率，无论从终端数据的准确性、不同应用抽取数据的效率，还是查询和比对分析能力，都不能满足业务需求。一个简单的查询通常要几十分钟。而采用Greenplum，其ETL（Extract-Transform-Load的缩写，即数据抽取、转换、装载的过程）功能可以很好地解决从3000多家店中抽取数据的问题，可以从极度分散的数据中充分挖掘出所存在的商业价值。ETL作为BI/DW（Business Intelligence）的核心和灵魂，能够按照统一的规则集成并提高数据的价值，完成数据从数据源向目标数据仓库的转化，是实施数据仓库的重要步骤。图3显示了商业智能从数据采集到数据仓储、数据分析、数据报告呈现的整个流程。

无论对门店，还是对劲霸公司内部，部署Greenplum不仅提高了数据获取的准确性，也提高了从应用抽取数据（ETL）、清洗（Cleaning）、转换（Transform）、加载（Load）、分析（Analysis）到报告呈现整个实现商业智能的效率。

图3. Greenplum 提高整个商业智能效率



评估Greenplum技术主要从下面几个纬度进行：

Greenplum MPP大规模并行处理能力

Greenplum适合进行大数据、多用户的并行数据分析处理（OLAP，大范围全表扫描，大数据吞吐量）。这与交易系统OLTP数据事务处理（小数据量、快速响应）有着根本上的不同。我们对各种类似的技术做了对比测试。Greenplum的大数据处理能力（图4），尤其多用户处理能力，不仅提高了数据获取的准确和实时性，还提高了数据导入、分析过程以及报告呈现的效率，满足我们现在和未来对BI的需求。我们的常规数据库Oracle、DB2、SQL Server多是为满足OLTP而设计，在海量数据面前很难满足大数据分析的性能需求。现在，所有MPP集群分析应用都运行虚拟化主机平台，实现明细数据查询和深层数据挖掘。

而数据挖掘可以为公司的经营分析进行服务，经营分析体系在整个战略执行保障中起着非常重要的作用，起到支持战略执行过程中的业绩监控与高层决策的作用。具体而言，公司的经营分析体系就像一个诊断、照明系统，价值在于把握方向、支持决策、促进行动三个方面，最重要的是以此为切入点形成经营管理的机制：

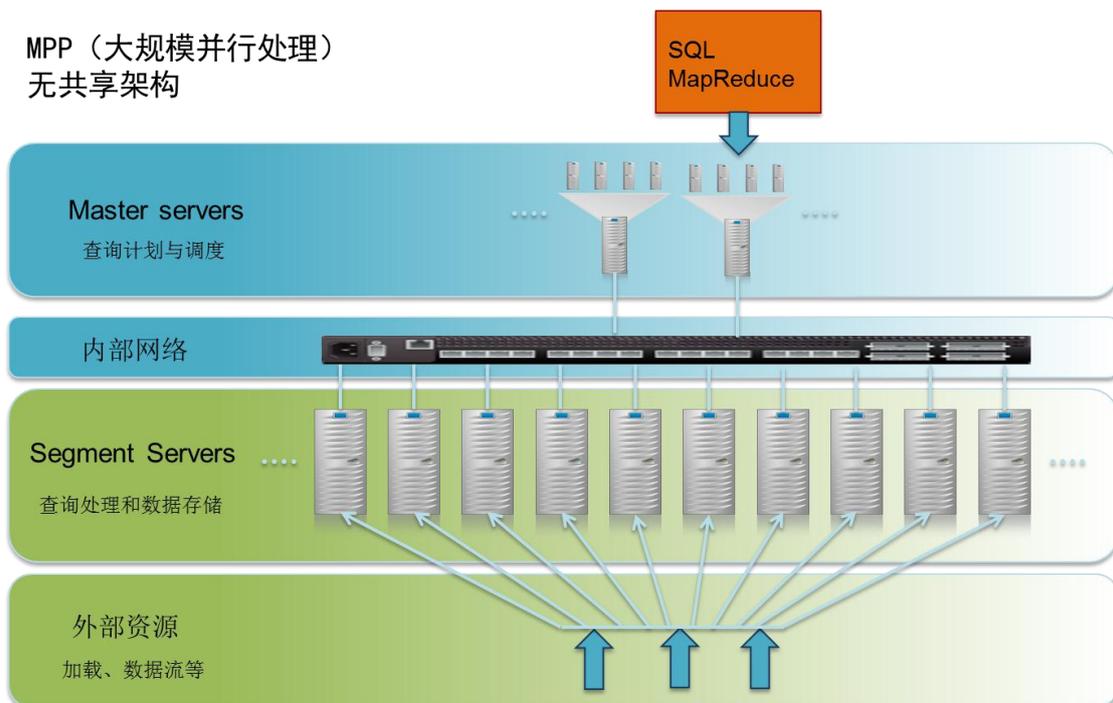
首先，帮助管理层把握战略方向。通过经营分析体系，清晰地传递管理要求，确保战略能够执行落地。同时，衡量公司与既定战略的距离。

其次，支持管理人员有效决策。对于公司决策来说，“时机”非常重要，战略框架下的经营分析体系保障了分析的客观完整和决策的更准确更透明。高层需要掌握每个职能领域的战略层和策略层指标，用数据来说话，深入分析，把握时机，动态决策。

最后，将战略和决策变为行动。在例行化的经营分析体系帮助下，及时发现公司经营的问题，追溯问题发生的根本原因。对未来的行动做出检讨和优化。面向未来思考，分析的结论以行动为导向，促进公司运营的改善。

此外，Greenplum数据库软件的无共享、大规模并行处理、按列存储数据库、数据库内压缩、MapReduce、不停机扩容、多级容错等特点，使得该数据库软件能很好地用于OLAP应用。

图4. Greenplum 的并行数据处理架构



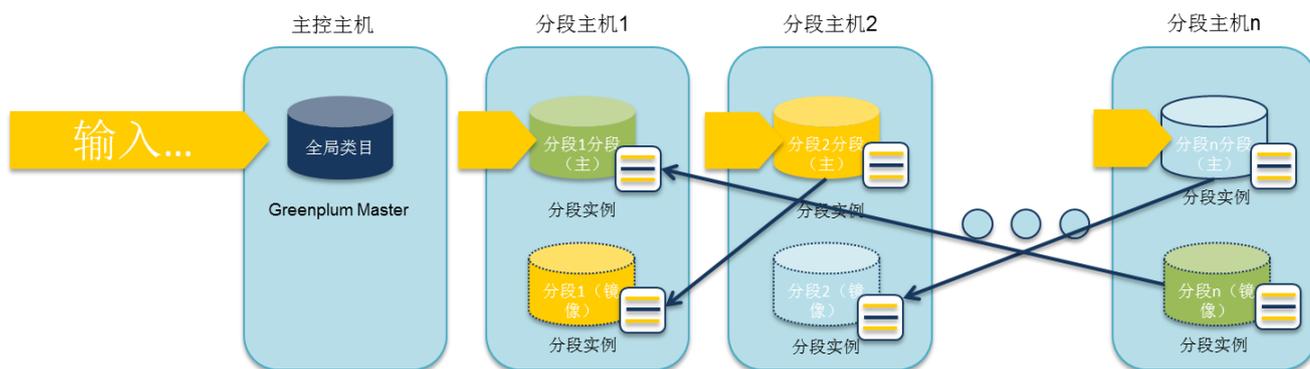
Greenplum的高可用性

Greenplum的高可用性具有以下特性，如图5所示：

- 系统自动侦测失败节点，并自动启动容错保护机制；
- 节点失败时和数据恢复时，数据库不宕机，可持续提供服务；
- 基于数据块（Block）的快速数据同步和恢复；

我们实施EMC Greenplum分布式数据库（见图5），以在ERP、DRP数据的基础上，对订单到货率、销售情况、生产及时率等生产经营数据做进一步的分析，从而改善了经营决策。

图5. Greenplum 分布式数据库分析生产经营数据



Greenplum统一智能管理

我们选择Greenplum的另一个原因是其统一的BI管理功能。我们业务是在跨物理、虚拟和云计算的平台上运行的。提高BI效率，就需要能够跨各种环境提供统一的BI。Greenplum图像化的统一管理界面，让我们能对物理、虚拟和云计算环境的各种应用以及数据进行快速的BI管理。同时，管理员可以轻松将任意选定查询项目或比对查询结果，以及相关内容由多种图表来呈现，包括饼状图、柱状图和趋势走向图，如图6和图7所示。

图6. Greenplum 统一管理界面



图7. 各种性能监控界面



Greenplum在劲霸环境下的数据导出和加载性能测试

中桥分别在物理环境（表1）和虚拟环境（表2）对Greenplum的数据加载性能进行了测试。劲霸的林先生表示，系统所呈现的加载速度在同类产品中处于领先状态。图8是从Oracle直接加载到Greenplum的性能，加载速度可达11万行/秒；图9是从Greenplum导出数据的性能，可达12万行/秒。

表1. 物理机加载性能

文件名	文件大小 (G)	记录条数	加载耗时	加载速度 (兆/秒)	加载速度 (行/秒)
dw_g_buhuo	243MB	1,975,358	1.3 秒	187M/秒	152 万行/秒
dw_g_endstore	2700MB	36,064,187	11 秒	245M/秒	327 万行/秒
dw_g_xiaos	4700MB	36,763,913	19 秒	247M/秒	193 万行/秒
dw_serverless	3600MB	26,718,268	15 秒	240M/秒	178 万行/秒

表2. 虚拟机加载性能

文件名	文件大小 (G)	记录条数	加载耗时	加载速度 (兆/秒)	加载速度 (行/秒)
dw_g_buhuo	243MB	1,975,358	2 秒	122M/秒	99 万行/秒
dw_g_endstore	2700MB	36,064,187	22 秒	123M/秒	164 万行/秒
dw_g_xiaos	4700MB	36,763,913	46 秒	102M/秒	80 万行/秒
dw_serverless	3600MB	26,718,268	26 秒	138M/秒	103 万行/秒

图8. 从 Oracle 加载到 Greenplum 的性能

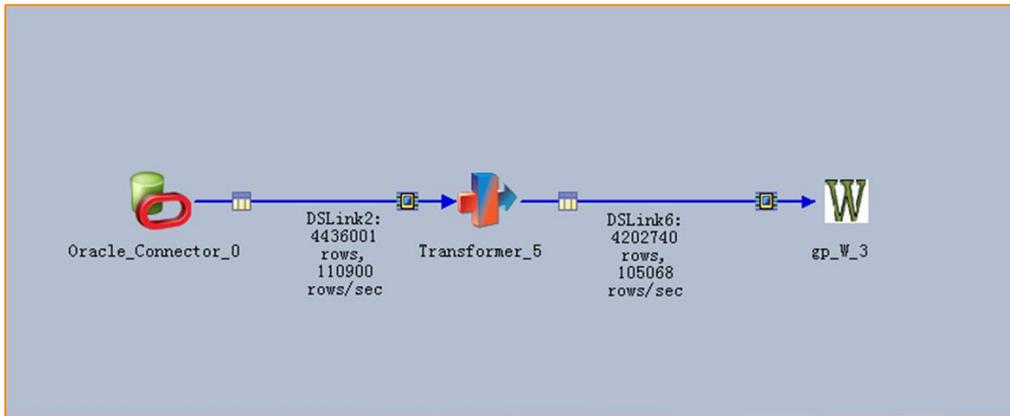


图9. Greenplum 的数据导出速度

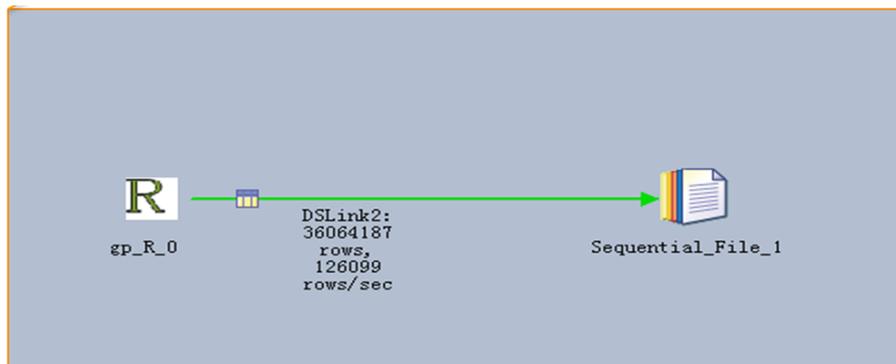
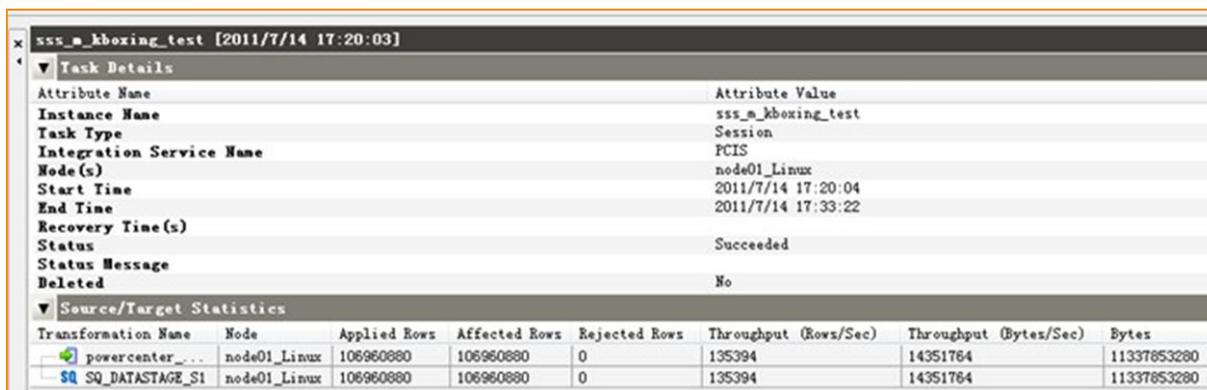


图10显示，Informatica能连接Greenplum，实现从Oracle读取数据，然后写入到Greenplum中。写入Greenplum的速度大概为135,394行/秒；加载速率约为14M/秒；1亿条数据约10分钟即可完成从Oracle到Greenplum的加载。

图10. Greenplum 的数据加载



The screenshot shows the task details for 'sss_m_kboxing_test' [2011/7/14 17:20:03]. The task is in a 'Succeeded' status. Below the task details, there is a table showing source and target statistics:

Transformation Name	Node	Applied Rows	Affected Rows	Rejected Rows	Throughput (Rows/Sec)	Throughput (Bytes/Sec)	Bytes
powercenter...	node01_Linux	106960880	106960880	0	135394	14351764	11337853280
SQ SQ_DATASTAGE_S1	node01_Linux	106960880	106960880	0	135394	14351764	11337853280

Greenplum的分析性能测试

下面记载了Greenplum在物理环境（表3）和虚拟环境（表4）下的查询性能。劲霸对各种同类的大数据分析软件做了性能对比测试。Greenplum呈现的分析性能，远远高出同类其它产品（几十倍甚至于上百倍），这是劲霸选择Greenplum最主要原因。

表3. 物理机测试结果统计

查询周期	当期	同期	性能耗时
1 个月	2009 年 9 月	2008 年 9 月	6 秒
6 个月	2009 年 6~12 月	2008 年 6~12 月	13 秒
12 个月	2009 年 1~12 月	2008 年 1~12 月	15 秒

表4. 虚拟机测试结果统计

查询周期	当期	同期	性能耗时
1 个月	2009 年 9 月	2008 年 9 月	9 秒
6 个月	2009 年 6~12 月	2008 年 6~12 月	15 秒
12 个月	2009 年 1~12 月	2008 年 1~12 月	17 秒

为了实现“供应云BI”，大数据分析性能的可扩展性对劲霸很关键。我们把同一组数据放大2倍和3倍，测试Greenplum在数据流快速翻倍前提下的分析能力。如下表（表5）显示，Greenplum大数据分析的可扩展性能可满足我们的未来需求。当数据量增加三倍，对“进销存情况查询”的时间从45秒上升到1分5秒。Greenplum在45秒完成的查询用Oracle则需要20分钟。如果数据量增长3倍，Greenplum在1分5秒完成的查询，通过Oracle则需要几十分钟。

表5. Greenplum 在数据量增长情况下的分析性能

	劲霸 oracle 测试结果（数据量×1）	Greenplum 测试结果（数据量×1）	Greenplum 测试结果（数据量×2）	Greenplum 测试结果（数据量×3）
进销存情况查询	20 分钟	45 秒	53 秒	1m5 秒
尺码分析	6 分钟	38 秒	52 秒	67 秒
销售趋势分析图	17 分钟	15 秒	20 秒	32 秒
新老货对比分析	3 分 45 秒	25 秒	27 秒	32 秒

下面是我们的多用户查询结果展示（图11）。分别显示了1个月（2008年8月与2009年8月的数据对比分析，查询耗时6秒）、6个月（2009年6~12月与2008年6~12月，6个月数据对比分析，查询耗时13秒）和12个月（2009年1~12月与2008年1~12月，12个月数据对比分析，查询耗时15秒）的全维度查询。我们使用之后发现多用户同时查询速度相当快，从传统的几十分钟缩短到数秒即可完成。

图11. 多用户查询结果对比分析

当期时间: 2009-1-1-2009-12-31 同期时间: 2008-1-1-2008-12-31									
商品名称	新老货	当期销售数量	同期销售数量	当期销占比	同期销占比	销占比升降幅	当期订货数量	同期订货数量	
CD盒	其他								
U盘	其他								
办公用品	老货								
	其他								
	新货								
抱枕	新货								
抱枕	其他								
	新货								
杯子	老货								
	其他								
长袖棉T恤	老货								
	新货								
长袖丝光棉T恤	老货								
	其他								
	新货								

当期时间: 2009-6-1-2009-12-31 同期时间: 2008-6-1-2008-12-31									
商品名称	新老货	当期销售数量	同期销售数量	当期销占比	同期销占比	销占比升降幅	当期订货数量	同期订货数量	
CD盒	其他								
U盘	其他								
办公用品	老货								
	其他								
	新货								
抱枕	新货								
抱枕	其他								
	新货								
杯子	老货								
	其他								
长袖棉T恤	老货								
	新货								
长袖丝光棉T恤	老货								
	其他								
	新货								

当期时间: 2009-8-1-2009-8-31 同期时间: 2008-8-1-2008-8-31									
商品名称	新老货	当期销售数量	同期销售数量	当期销占比	同期销占比	销占比升降幅	当期订货数量	同期订货数量	
CD盒	其他								X
U盘	其他								X
办公用品	老货	XX	XX	X%	X%	XX%	XXXX		
	新货	X	X	X%	X%	XX%	XXXX		
抱枕	新货	XX		X%			XXXX		
抱枕	新货	XXX		X%			XXXXX		
杯子	老货	XXXX	XXXX	X%	X%	XX%			X
长袖棉T恤	老货	XXXX	XXXXX	X%	X%	XX%	XXXXXX		
	新货	XXXXX	XXXXX	X%	X%	XX%	XXXXXX		
长袖丝光棉T恤	老货	XXXX	XXXXX	X%	X%	XX%	XXXXXX		
	新货	XXXX	XXXX	X%	X%	XX%	XXXXXX		
长袖丝光棉衬衫	新货	XXX		X%			XXXX		

在使用Greenplum技术之前，由于数据加载时间长，分析性能低下，我们很难满足明细查询要求。现在通过Greenplum我们可以进行明细数据的查询，而且能够快速进行高性能的加载和查询、并行的加载和导出。

VNX实现统一存储

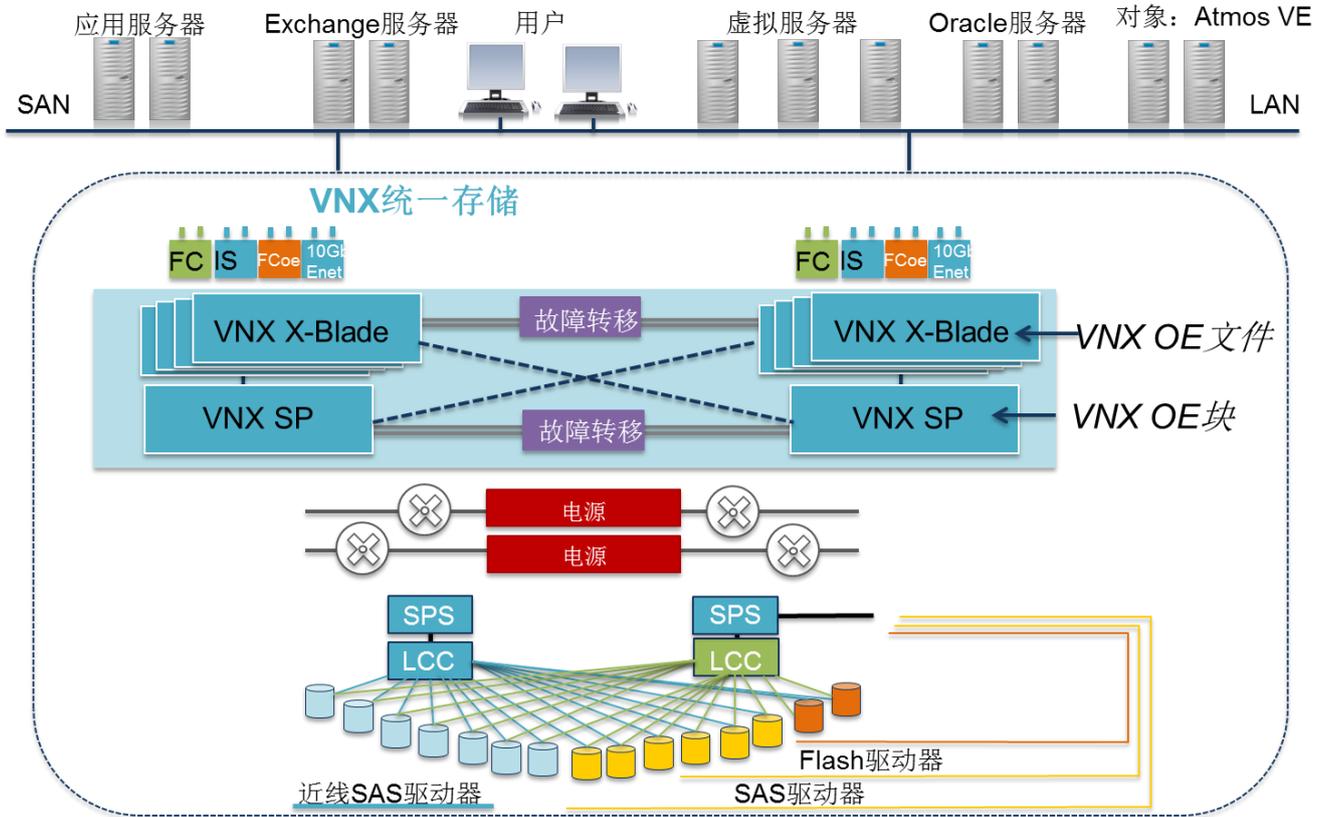
从去年的“6纵6横”（ERP集群）到今年的“9+1”，这整个的规划都对存储的易用性和扩展性提出了很高的要求。去年我们基本上2个月就有一个新系统上线。今年除了对去年系统的深化和优化，自然而然会考虑到系统整个扩展性，还有新系统的上线能力，以及各种应用在虚拟化环境的性能。我们选择VNX最主要的原因包括：

- VNX统一存储架构：支持结构化、非结构化数据以及目标存储技术，这对我们IT从“管理数据”到“通过信息创造价值”，逐步完善私有云数据中心十分重要。
- VNX结合FAST缓存功能大幅度优化性能并支持物理和虚拟环境下应用对性能的需求。
- VNX具有技术灵活性：支持NFS、CIFS、iSCSI、FCoE各种协议。
- VNX支持多种业务平台。我们现在的应用在各种异构平台运行，如Windows、Linux和Unix。

图12显示了VNX的统一存储架构。VNX作为新一代的统一存储平台，整合了原有的数据块存储、文件存储和直连存储，使企业可以动态增加、共享和经济高效地管理多协议文件系统以及多协议数据块存储访问，为企业数据中心从物理向私有云过渡提供了有力保证。

VNX操作环境支持Microsoft Windows和Linux/Unix客户端在多协议（NFS和CIFS）环境中共享文件。同时，它还支持高带宽和对延迟敏感的数据块应用程序的iSCSI、光纤通道和FCoE访问。

图12. VNX 统一存储架构图



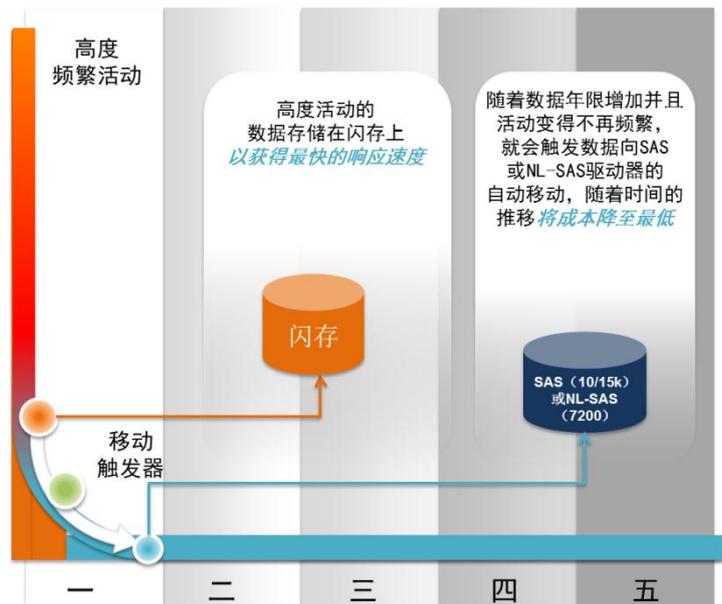
VNX提供FAST Cache，通过“自动优化”功能，获得最高的系统性能，同时将存储成本降至最低。除了整合到一个资源池这一基本价值外，“自动优化”描述了这样一种能力：有一个全自动的存储系统，它可以按数据使用情况、服务器、用户或应用程序，以最低成本提供最合适的性能。

FAST Cache（如图13所示）确保尽可能最高的层（通常是闪存固态硬盘）为“热”数据持续提供服务，以实现最快的响应速度。

如前面提到的，数据和应用程序的增长，以及因为预算无法随之提高导致的虚拟服务器增加，都放大了这一做法的必要性。解决此类难题和复杂性的唯一可行办法是优化—而实现这一优化的唯一可行办法是自动化。目前，我们采用了FAST Cache。在高峰期，应用特别繁忙的时候，FAST Cache可以起到一定缓冲作用，在高峰期，很多IO应用可以再在FAST Cache中进行“消化”。我们系统中几个数据库应用，包括虚拟平台上运行的数据库应用都使用了FAST Cache，作为缓存扩展。我们通过存储优化部署，为应用系统提供了更好的性能支撑和本地数据保护。

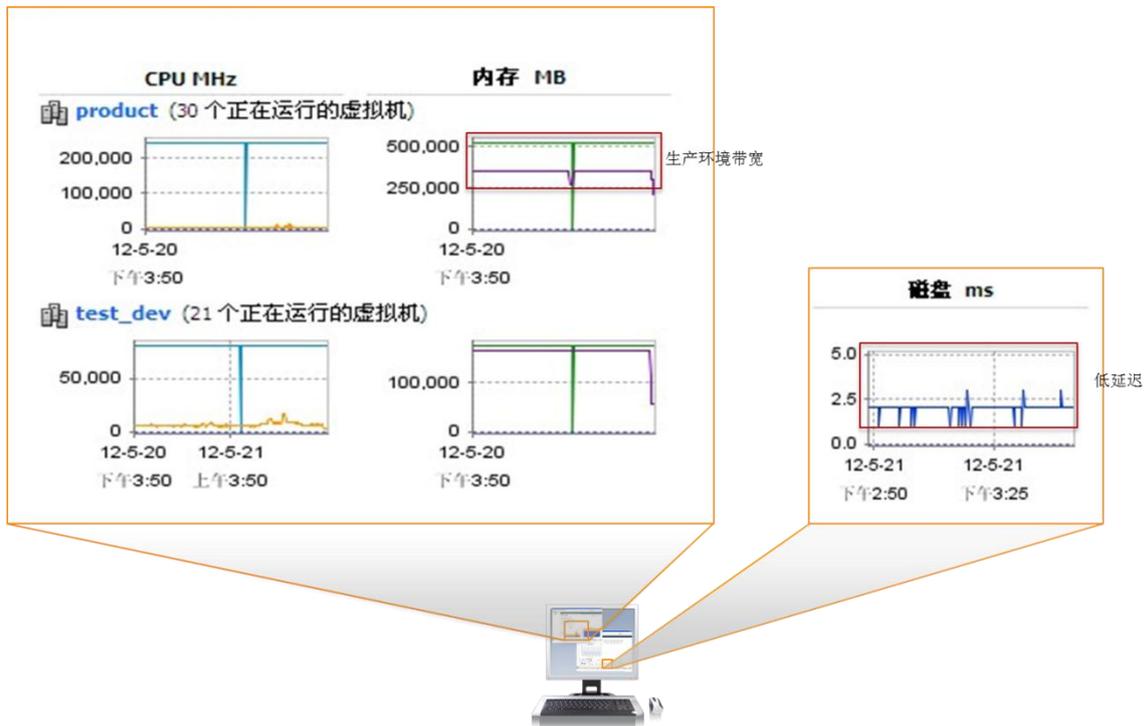
图14是我们虚拟化数据中心的整体情况。如果采用传统的数据中心方式，需要管理单台服务器，

图13. FASTCache 示意图



并且需要把机房分成PC区、小型机区、网络区进行归类管理。而对数据中心进行虚拟化之后，我们将数据中心资源分为开发测试群和生产群，只需管理这两个群即可，从而大大降低了我们的管理复杂度。在内存使用情况的小图中可看到，不管是生产环境下还是测试环境，内存的使用率都比较高（紫色水平线所示），主机这块内存比较大，在数据处理过程中（读或写）会有一些的压力。内存消耗比较大，说明I/O比较频繁。图14是中桥分析师在劲霸的现场截屏。从图14可以看到，由于各种业务关键型应用（包括运行在IBM AIX OS的应用，以及各种虚拟化运行在VNX上），生产环境内存大于350,000 MBPs，VNX的时间延迟一直在2.5ms，显示VNX满足业务关键型应用和虚拟化性能需求，保持了低延迟。

图14. 劲霸虚拟化数据中心整体概况



Unisphere实现统一集中管理

劲霸部署EMC Unisphere来简化存储管理和监控。Unisphere让存储管理员能够通过一个简单的基于web的管理工具来管理整个存储生态系统。Unisphere控制面板是一个可进行一览式管理、监控、管理权限设定、运行报告的界面，让管理员快速进行监控和管理。

Unisphere所具有的特性包括：

- 简单、直观的GUI界面，操作简单，并且提供报警、报告呈现等功能；
- 与VMware集成，提供端对端的虚拟机到物理机的映射，并且可以自动搜索虚拟机和ESX服务器；
- 统一的用户角色设定确保了系统的安全性。

我们可以通过EMC Unisphere（如图15）来管理复制本地和远程VNX阵列的存储容量的图形摘要信息和阵列的状态和细节。警报也显示在这里，左下角窗口是虚拟服务器的链接。要添加其他信息类型到面板中，管理员只需从面板选项拖动图标即可。通过顶部导航条就可在不同的系统视图之间进行切换。管理员可以通过任何平台，实时监控各种物理和虚拟存储资源运行情况，并采取必要措施。

图15. EMC Unisphere 管理界面

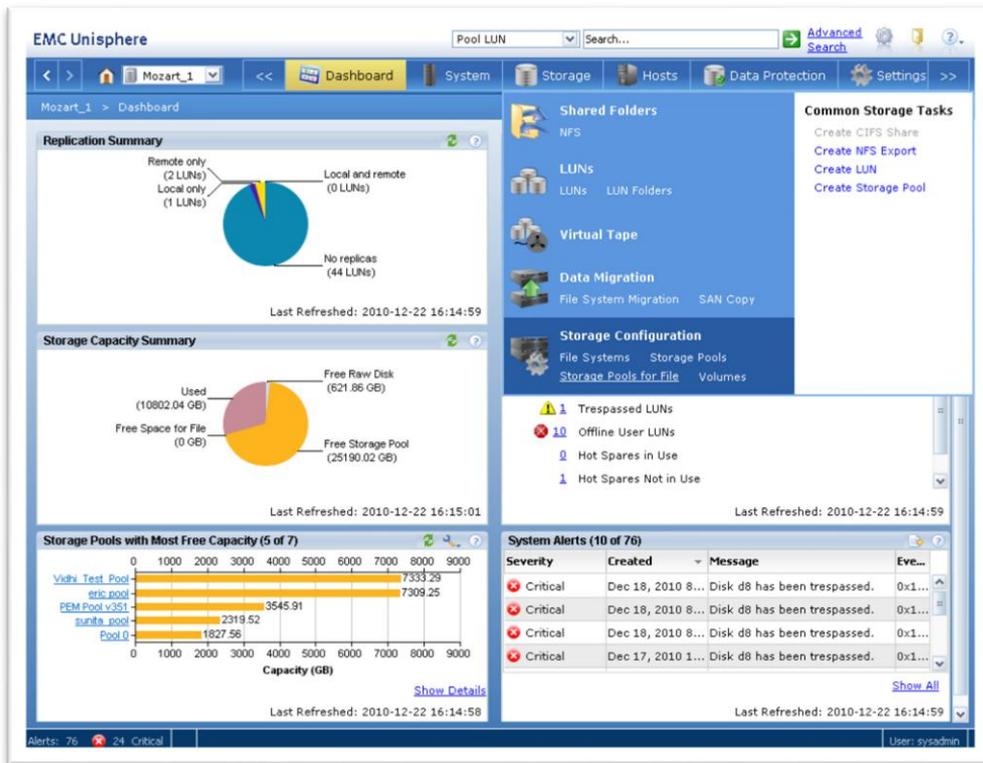


图15的截屏是根据客户需求进行定制的，其易于使用的下拉式菜单可进行常见任务的操作，如“创建存储池”或“创建LUN”。用户只需要轻点鼠标即可完成日常任务的操作，从而大大提供了管理效率和工作效率。

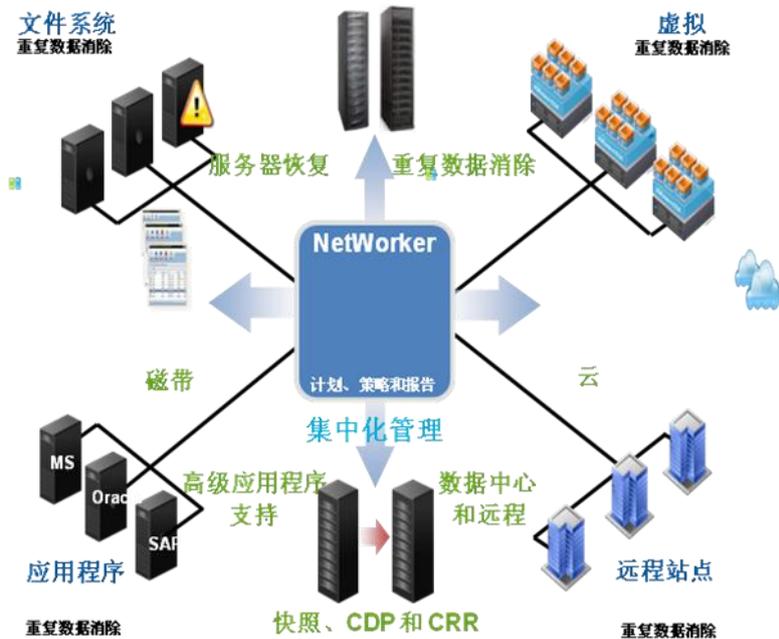
Networker实现统一数据保护

信息链最重要的组成部分就是数据。如何有效地对数据进行保护，降低数据保护成本和带宽需求，保证数据保护过程中系统的整体性，也是我们的重要考虑因素。虽然我们的数据中心超过80%的应用已经虚拟化，但还有20%在物理环境下运行。选择一个可以对物理和虚拟环境进行统一、高效的数据保护，并提供重复数据删除技术以提高备份效率和存储空间效率，是我们选择数据保护方案的最重要考虑因素。

在这方面，我们采用EMC Networker多流备份软件和数据域虚拟磁带库来作为数据库备份系统。Networker安全、简单、快速的备份为我们保护云计算数据，实现数据备份到云，提供了很好的技术选择。而与Data Domain重复数据删除的集成则显著降低了我们的存储容量需求，以及备份和复制的带宽需求，为我们实现统一备份、统一灾备提供了高自动化的数据保护，并降低了灾备成本。

据ESG 2012年《EMC Networker Backup and Recovery Solutions》用户实际使用测评报告显示，EMC Networker是一个值得信赖的备份和恢复解决方案。它使企业能够利用统一的平台来进行异构数据备份和恢复，并且不影响线上的应用。Networker可以在不同的计算环境下工作，包括多操作系统、SAN、NAS和DAS磁盘存储环境、磁带机和磁带库，还有云存储。Networker作为集中、统一的备份管理平台（如图16），可为本地和异地的所有数据提供全面的数据保护，包括数据库、文件系统、安全信息、用户profile、存取控制表、Windows注册表和事件日志。其集中式管理和监控界面，可提供定制的视图，并且可以通过客户端和设备设置向导，轻松启动恢复应用程序。其基于角色的访问安全性确保了系统整体的安全。Networker作为一款值得信赖的备份和恢复解决方案，支持跨物理和虚拟环境的数据和应用程序的各种数据保护选项，可轻松配置数据保护资源、设置时间表和策略，从而集中、自动且加快整个IT环境的数据保护。

图16. NetWorker 的集中管理提供了统一的数据保护



NetWorker 与 VMware vStorage API for Data Protection (VADP) 的集成消除了快照的临时性存储，并且支持 change block tracking (CBT)，从而提升了系统性能并改善了网络利用率和降低了管理负担。CBT 只追踪增量变化的虚拟机数据，而无需通过文件系统扫描来识别和备份变化的数据。VADP 还能够通过单一备份实现映像级和文件级的粒度恢复。物理或虚拟代理服务器都可用来实现更为高效、集中的虚拟机备份。此外，NetWorker 还能够与 VMware vCenter 通信在整个虚拟环境中自动搜索和显示虚拟映射，从而简化了整个管理强度（如图 17）。

NetWorker 内置的数据流并行技术能够为系统提供优越的性能，从而增强备份和恢复。如图 18 所示，其优化的多数据流设计，可以同时使用多个并发数据流进行数据恢复，并且备份、恢复克隆和转储可同步进行，从而显著提升了系统性能。

图17. NetWorker 和 vStorage API for Data Protection (VADP) 的集成

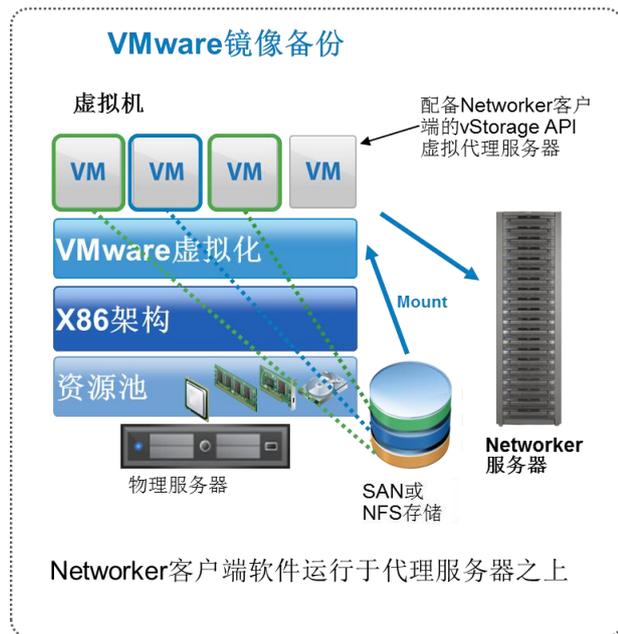
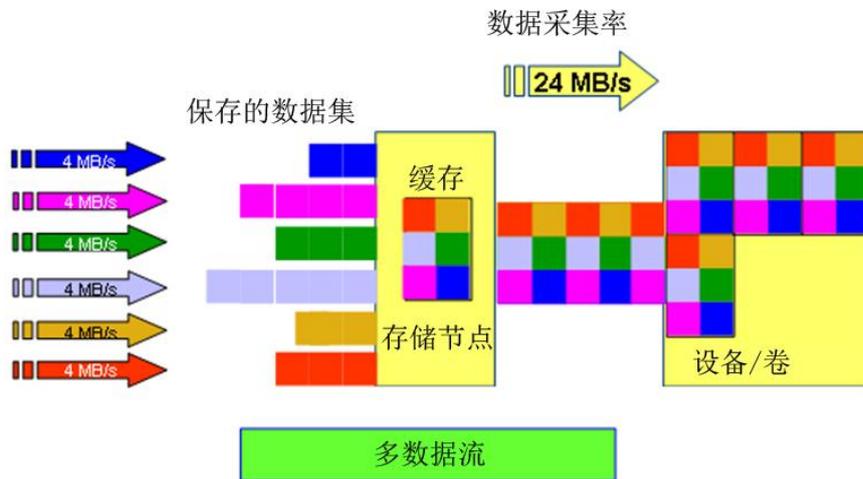


图18. 优化的多流设计提升系统性能

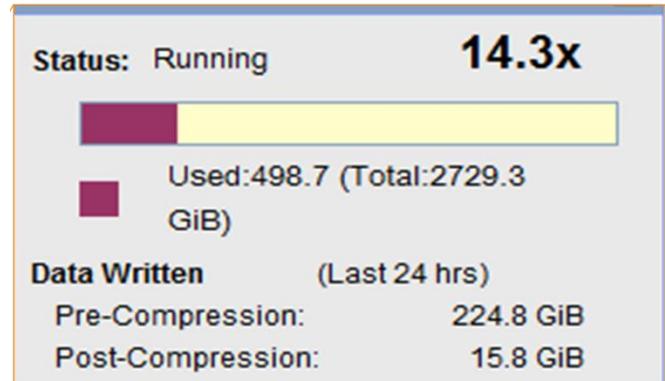


使用Networker之后，我们对于该方案提供的磁盘级的数据备份和恢复速度、通过重复数据删除技术减少了备份的数据量、提高备份效率等方面还是比较满意的。通过使用Data Domain，我们存储容量需求降低了80%，备份和复制带宽也大幅度下降。

Data Domain在备份的过程中就能进行重复数据删除，从而显著减少了磁盘存储量。Networker利用Data DomainBoost集成了EMC Data Domain重复数据删除存储系统。Data Domain系统可以作为Networker备份的目标磁盘，DD Boost还可额外地加速提高性能，提供对Data Domain系统信息的可视性，以及对将备份映像和复制到远程Data Domain系统的控制。我们能够通过网络器对Data Domain进行统一监控。从下图19可以看到，通过网络器可以看到Data Domain系统事件、信息及状态，从而简化了管理备份的流程。

图19是通过Data Domain进行数据压缩时的界面。从图中可以看到，压缩前的增量备份数据为224.8GB，压缩后，备份的数据不到16GB，重复数据删除比为14.3: 1。

图19. Data Domain 压缩界面



中桥观点

劲霸通过虚拟化和私有云提高了应用部署效率和业务对市场的反应速度，以及IT资源利用率和管理效率。

EMC Greenplum作为劲霸“同赢、高效”的商业智能平台的使能技术，提升了劲霸通过IT创造价值的能力，在提高产业链竞争优势的同时，还提升了优质渠道资源凝聚力和市場反应速度。

EMC VNX为劲霸提供高效的统一存储管理平台，保证了业务关键型应用在物理和虚拟环境的性能需求，并支持结构化、非结构化和半结构化应用对存储的需求。其FAST Cache为用户提供了多种提升性能、降低延迟的选择。

EMC Unisphere简单、直观的管理界面提供了全面的存储管理功能，增强了整合的虚拟化环境的简单性和存储效率。

EMC Networker结合Data Domain，为劲霸提供了高效、统一的数据灾备平台，降低了IT管理强度、提高灾备可靠性，并大大降低灾备存储容量和带宽资源需求。

结束语

中国是世界服装生产和消费的第一大国，同时也是全球高端零售商品第一大国。入世以来，中国零售和服装行业在迎接经济强增长和人均支出的快速上升带来的销售高增长的同时，也面对跨国企业带来的挑战。对于品牌服装生产和零售而言，信息链不仅决定着品牌价值，也同样决定着物流链和资金链的效率。提高信息化水平，是中国本土品牌服装和零售企业提升竞争力、优化产业链价值空间的关键。

在部署虚拟化和私有云计算中心前，劲霸IT的挑战是中国品牌服装和零售所面临的普遍问题：应用数据分离、IT资源利用率低下、IT对业务反应速度慢、数据保护机制不够完善。中桥对劲霸男装云数据中心部署及主要技术进行的评测，以及和劲霸数据中心CIO的采访结果显示，劲霸通过部署EMC的整体虚拟化和私有云方案解决了IT难题。虚拟化的实施使得劲霸男装每两个月部署一个新应用的部署时间从原来的2-4周缩短到几天甚至几个小时，实现了用更少IT资源支撑更多应用部署的目的。

如何通过IT创造价值，是劲霸男装部署虚拟数据中心和私有云的一个主要目标。劲霸男装遇到的终端数据获取准确性、数据快速导入导出、高效多用户业务数据分析和查询等挑战在中国服装行业应用很具有代表性。劲霸男装通过部署EMC Greenplum作为大数据分析技术平台，不仅提供了高效的多用户业务查询能力，还使得劲霸业务比对性能有了大幅提升，速度从几十分钟缩短到十几秒，提升了数百倍。同时，Greenplum的MMP大规模并行处理能力，提高了对3000多个门店终端数据的获取效率，为各店提供了优质的BI服务平台，让劲霸的“取之于民，用之于民”的“共赢”私有云BI部署成为可能。

劲霸男装部署的EMC VNX统一存储，不仅可以支持结构化、非结构化应用，同时可以支持跨各种平台，支持物理环境和虚拟环境的数据存储。此外，通过使用VNX的FAST Cache技术，保证各种业务关键型应用在虚拟化环境的应用性能。EMC Unisphere简单、直观的管理界面为整个系统提供了全面的存储管理功能，增强了整合的虚拟化环境的简单性和存储效率。通过部署EMC Networker，劲霸实现了对物理、虚拟和云计算数据的统一灾备。同时，通过多流备份和集成VMware的CBT (Change Block Tracking) 大大提高了备份效率。结合Data Domain压缩和重复数据删除技术，将存储容量需求降低了90%以上。

中桥实验室评测结果显示，劲霸通过Greenplum实现了统一的商业智能；通过EMC VNX实现了统一存储；通过Unisphere实现了直观、简化的统一管理；通过Networker和Data Domain实现了高效统一的数据保护。对于品牌服装制造和零售企业而言，这种解决方案能解决数据分离、资源低利用率、虚拟化业务关键型应用性能的不稳定，以及管理复杂和数据灾备等问题。让企业在集中管控整个IT环境的同时，实现通过IT创造价值的目的。



北京市朝阳区朝外大街26号朝外MEN财贸中心A座2306A室 | 电话: 8610 85655510 | www.Sino-Bridges.com