



(12) 发明专利

(10) 授权公告号 CN 106850315 B

(45) 授权公告日 2020. 11. 20

(21) 申请号 201710156923.X

(56) 对比文件

(22) 申请日 2017.03.16

CN 103618627 A, 2014.03.05

(65) 同一申请的已公布的文献号

审查员 郭珩

申请公布号 CN 106850315 A

(43) 申请公布日 2017.06.13

(73) 专利权人 苏州浪潮智能科技有限公司

地址 215100 江苏省苏州市吴中区吴中经

济开发区郭巷街道官浦路1号9幢

(72) 发明人 李世平

(74) 专利代理机构 北京集佳知识产权代理有限

公司 11227

代理人 罗满

(51) Int. Cl.

H04L 12/24 (2006.01)

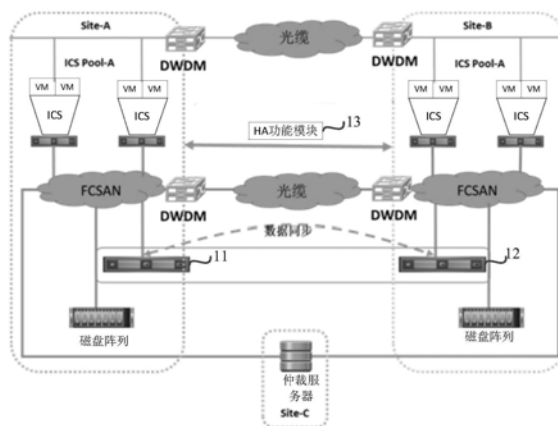
权利要求书1页 说明书6页 附图1页

(54) 发明名称

一种自动化容灾系统

(57) 摘要

本发明实施例公开了一种自动化容灾系统，通过将本地站点和备用站点设置于同一个池组中，当本地站点的服务器出现故障时，HA功能模块将服务器包含的虚拟机进行服务器的切换；通过在本地站点中设置第一存储虚拟化网关设备，在备用站点中设置第二存储虚拟化网关设备，这两个存储虚拟化网关设备可以形成虚拟化网关集群，用于实现两个站点之间数据的同步；当本地站点的存储出现故障，第一存储虚拟化网关设备将本地站点的IO指向备用站点；当本地站点出现故障，第一存储虚拟化网关设备和HA功能模块同时运行，将本地站点的业务系统切换至备用站点。该自动化容灾系统可以实现站点之间的数据同步，实现故障的自动切换以及自动进行数据恢复工作。



1. 一种自动化容灾系统,包括本地站点和备用站点,其特征在于,还包括第一存储虚拟化网关设备、第二存储虚拟化网关设备和HA功能模块;

所述本地站点和所述备用站点位于同一个池组,用于支持所述HA功能模块实现HA功能;

所述第一存储虚拟化网关设备设置于所述本地站点,用于获取所述本地站点存储的数据;

所述第二存储虚拟化网关设备设置于所述备用站点,与所述第一存储虚拟化网关设备形成虚拟化网关集群,用于实现所述数据的同步;

当所述本地站点的服务器出现故障,所述HA功能模块将所述服务器包含的虚拟机进行服务器的切换;

当所述本地站点的存储出现故障,所述第一存储虚拟化网关设备将所述本地站点的IO指向所述备用站点;

当所述本地站点出现故障,所述第一存储虚拟化网关设备和所述HA功能模块同时运行,将所述本地站点的业务系统切换至所述备用站点;

所述当所述本地站点的服务器出现故障,所述HA功能模块将所述服务器包含的虚拟机进行服务器的切换包括:

当所述本地站点包括有至少两个服务器时,若第一服务器出现故障,则所述HA功能模块将所述第一服务器包含的虚拟机切换至第二服务器,由所述第二服务器为所述虚拟机提供服务支持;所述第一服务器为所述至少两个服务器中的任意一个服务器;所述第二服务器为所述至少两个服务器中除所述第一服务器之外的任意一个服务器;

当所述本地站点包括的所有服务器出现故障时,则所述HA功能模块将所述本地站点包含的所有虚拟机切换至所述备用站点的服务器,由所述备用站点的服务器为所述虚拟机提供服务支持。

2. 根据权利要求1所述的自动化容灾系统,其特征在于,还包括:仲裁节点;

所述仲裁节点分别与所述本地站点和所述备用站点连接,用于当所述本地站点和所述备用站点出现数据差异时,对所述本地站点和所述备用站点的数据正确性进行判别。

3. 根据权利要求2所述的自动化容灾系统,其特征在于,所述仲裁节点为仲裁服务器。

4. 根据权利要求1-3任意一项所述的自动化容灾系统,其特征在于,还包括:报警模块;所述报警模块设置于所述本地站点,当所述本地站点出现故障时,进行报警提示。

5. 根据权利要求1-3任意一项所述的自动化容灾系统,其特征在于,所述本地站点还用于采用vmotion将所述本地站点的业务系统切换至所述备用站点。

一种自动化容灾系统

技术领域

[0001] 本发明涉及容灾建设技术领域,特别是涉及一种自动化容灾系统。

背景技术

[0002] 统计数据表明:93%的企业,一旦发生数据中心失效达10天的情况,通常在一年内就会破产。正常运行和数据保护对业务而言非常关键,灾难带来的员工工作效率下降、士气低落、收入下降、声誉受损和合规性下降等将对企业产生巨大的消极影响。

[0003] 灾难的发生往往是出乎人们意料的,当突然发生大的灾难,日常建立的控制措施已不再有效时,如何才能保护核心业务不被中断,使灾难造成的风险降到最低,这正是灾难恢复需要考虑的问题。

[0004] 数据中心的灾难会以不同形式发生,比如电力中断、硬件故障、人为操作失误、以及自然灾害。大量的事例表明,业务中断不是“假如”的问题,而是“何时”的问题。因此,只要是拥有IT信息系统的组织,都应为随时可能发生的中断做好准备。保证业务连续性也就是确保业务连续运作,不管发生什么情况,重要的系统和网络必须具有不间断的可用性。事实上,灾难恢复工作不仅仅局限于金融领域,政府、公安、医疗、教育等诸多行业都需要使用业务可用性解决方案来保证其业务的可用性,对于这些企业,机关,单位而言,当IT基础设施突然停运,不仅企业内部的业务流转、办公等面临瘫痪,更重要的是,对外与客户或合作伙伴的所有业务交流都不得不陷入停滞。

[0005] 随着虚拟化技术的普及,越来越多的企业将他们的核心业务放在虚拟化平台中,增加了业务的密集程度。在发生灾难时,为了保证业务连续运行,数据完整无丢失,并且在灾难恢复后尽快将整个环境恢复如初,需要进行容灾建设。

[0006] 容灾系统,对于IT而言,就是为计算机信息系统提供的一个能应付各种灾难的环境。当计算机系统在遭受如火灾、水灾、地震、战争等不可抗拒的自然灾难以及计算机犯罪、计算机病毒、掉电、网络/通信失败、硬件/软件错误和人为操作错误等人为灾难时,容灾系统将保证用户数据的安全性(数据容灾),甚至,一个更加完善的容灾系统,还能提供不间断的应用服务(应用容灾)。可以说,容灾系统是数据存储备份的最高层次。

[0007] 建立容灾系统的目的是为了在灾难发生后能够以最快速度恢复数据服务,所以容灾系统的设计指标主要与容灾系统的数据恢复能力有关,最常见的设计指标有恢复时间目标(Recovery Time Objective,RTO)和恢复点目标(Recovery Point Objective,RPO)。RTO是反映业务恢复及时性的指标,表示业务从中断到恢复正常所需的时间,RTO值越小,代表容灾系统的数据恢复能力越强。RPO是反映恢复数据完整性的指标,在同步数据复制方式下,RPO等于数据传输时延的时间;在异步数据复制方式下,RPO基本为异步传输数据排队的时间。

[0008] 现有技术中,浪潮服务器虚拟化平台(InCloudSphere,ICS)提供的灾难恢复功能(Disaster Recovery,DR),简称ICS DR,可以用于保护本地站点所有物理主机、存储遭到灾难性破坏,无法恢复时,可在备用站点进行重建,从而最大限度的减少对应用程序或用户造

成的停机时间。如图1所示,为ICS的基本架构的示意图,即容灾系统的结构示意图,Site-A表示本地站点,Site-B表示备用站点,ICS DR要求每个站点是一个池组(Pool),Site-A属于ICS Pool-A,Site-B属于ICS Pool-B,本地站点和备用站点之间通过光缆连接,每个站点的磁盘阵列可以用于存储数据,这两个站点之间的数据同步可以通过存储自身的远程复制功能完成,每个站点中可以包含有多台服务器,每台服务器可以为多台虚拟机(Virtual Machine,VM)提供服务支持。但是,现有技术提供的容灾系统的自动化程度低,在灾难发生时需要人工进行故障切换和灾难恢复。

[0009] 可见,如何提升容灾系统的自动化程度,是本领域技术人员亟待解决的问题。

发明内容

[0010] 本发明实施例的目的是提供一种自动化容灾系统,可以提升容灾系统的自动化程度,实现故障切换和灾难恢复的自动化。

[0011] 为解决上述技术问题,本发明实施例提供一种自动化容灾系统,包括本地站点和备用站点,还包括第一存储虚拟化网关设备、第二存储虚拟化网关设备和HA功能模块;

[0012] 所述本地站点和所述备用站点位于同一个池组,用于支持所述HA功能模块实现HA功能;

[0013] 所述第一存储虚拟化网关设备设置于所述本地站点,用于获取所述本地站点存储的数据;

[0014] 所述第二存储虚拟化网关设备设置于所述备用站点,与所述第一存储虚拟化网关设备形成虚拟化网关集群,用于实现所述数据的同步;

[0015] 当所述本地站点的服务器出现故障,所述HA功能模块将所述服务器包含的虚拟机进行服务器的切换;

[0016] 当所述本地站点的存储出现故障,所述第一存储虚拟化网关设备将所述本地站点的IO指向所述备用站点;

[0017] 当所述本地站点出现故障,所述第一存储虚拟化网关设备和所述HA功能模块同时运行,将所述本地站点的业务系统切换至所述备用站点。

[0018] 可选的,所述当所述本地站点的服务器出现故障,所述HA功能模块将所述服务器包含的虚拟机进行服务器的切换包括:

[0019] 当所述本地站点包括有至少两个服务器时,若第一服务器出现故障,则所述HA功能模块将所述第一服务器包含的虚拟机切换至第二服务器,由所述第二服务器为所述虚拟机提供服务支持;所述第一服务器为所述至少两个服务器中的任意一个服务器;所述第二服务器为所述至少两个服务器中除所述第一服务器之外的任意一个服务器;

[0020] 当所述本地站点包括的所有服务器出现故障时,则所述HA功能模块将所述本地站点包含的所有虚拟机切换至所述备用站点的服务器,由所述备用站点的服务器为所述虚拟机提供服务支持。

[0021] 可选的,还包括:仲裁节点;

[0022] 所述仲裁节点分别与所述本地站点和所述备用站点连接,用于当所述本地站点和所述备用站点出现数据差异时,对所述本地站点和所述备用站点的数据正确性进行判别。

[0023] 可选的,所述仲裁节点为仲裁服务器。

[0024] 可选的,还包括:报警模块;

[0025] 所述报警模块设置于所述本地站点,当所述本地站点出现故障时,进行报警提示。

[0026] 可选的,所述本地站点还用于采用vmotion将所述本地站点的业务系统切换至所述备用站点。

[0027] 由上述技术方案可以看出,在自动化容灾系统中,通过将本地站点和备用站点设置于同一个池组中,以便于使得本地站点和备用站点可以具有同一个存储位置,从而可以支持HA功能模块实现HA功能,当本地站点的服务器出现故障时,HA功能模块便可以自动将所述服务器包含的虚拟机进行服务器的切换;通过在本地站点中设置第一存储虚拟化网关设备,在备用站点中设置第二存储虚拟化网关设备,这两个存储虚拟化网关设备可以形成虚拟化网关集群,用于实现两个站点之间数据的同步;当所述本地站点的存储出现故障,所述第一存储虚拟化网关设备可以自动将所述本地站点的IO指向所述备用站点;当所述本地站点出现故障,所述第一存储虚拟化网关设备和所述HA功能模块同时运行,将所述本地站点的业务系统切换至所述备用站点。可见,依据于第一存储虚拟化网关设备、第二存储虚拟化网关设备以及HA功能模块,可以实现本地站点和备用站点之间的数据同步,在发生故障时实现故障的自动切换。并且,由于数据的同步,在故障恢复后,可以由第一存储虚拟化网关设备和第二存储虚拟化网关设备自动进行数据恢复工作。

附图说明

[0028] 为了更清楚地说明本发明实施例,下面将对实施例中所需要使用的附图做简单的介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0029] 图1为现有技术提供一种容灾系统的结构示意图;

[0030] 图2为本发明实施例提供一种容灾系统的结构示意图。

具体实施方式

[0031] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下,所获得的所有其他实施例,都属于本发明保护范围。

[0032] 为了使本技术领域的人员更好地理解本发明方案,下面结合附图和具体实施方式对本发明作进一步的详细说明。

[0033] 接下来,详细介绍本发明实施例所提供的一种自动化容灾系统。图2为本发明实施例提供一种容灾系统的结构示意图,相比于图1所示的容灾系统,本发明实施例提供的容灾系统,在本地站点中设置了第一存储虚拟化网关设备11,在备用站点中设置了第二存储虚拟化网关设备12,并且通过将本地站点和备用站点设置于同一个池组中,以便于使得本地站点和备用站点可以具有同一个存储位置,从而可以支持HA功能模块13实现HA功能。

[0034] 举例说明,HA功能是一种自动的故障切换机制,当某一主机发生故障时,服务器或虚拟机自动到另外的可用的主机上重启。

[0035] 实现HA功能的前提条件是,两个站点上的所有服务器共享一个外部存储,在本发

明实施例中,通过将本地站点和备用站点放置于同一个池组(pool)中,如图2所示,本地站点Site-A和备用站点Site-B均放置于ICS Pool-A这个池组中,从而可以支持HA功能模块13实现HA功能。

[0036] 在本发明实施例中,第一存储虚拟化网关设备11和第二存储虚拟化网关设备12可以形成一个虚拟化网关集群,从而实现本地站点和备用站点之间的数据同步。

[0037] 在具体实现中,第一存储虚拟化网关设备11可以通过LUN来获取本地站点存储的数据。其中,LUN指的是存储设备映射给存储虚拟化网关设备的磁盘空间。

[0038] 在本发明实施例中,每个站点的存储设备各自映射一个LUN给各自站点的存储虚拟化网关设备。两个站点的存储虚拟化网关设备将这两个LUN组成一个VDM卷,组成这个VDM卷的两个LUN之间通过存储虚拟化网关设备完成数据同步。虚拟化网关集群可以将这个VDM卷提供给上层ICS挂载。对于ICS中的所有虚拟机,其存储位置全在这个VDM卷中。

[0039] 由于第一存储虚拟化网关设备11和第二存储虚拟化网关设备12通过镜像映射的原理,可以实现数据的实时同步,当本地站点出现故障时,由于备用站点已经同步了本地站点的数据,从而可以实现RPO等于零,极大的提升了容灾系统的级别。由于数据的同步,在故障恢复后,可以由第一存储虚拟化网关设备和第二存储虚拟化网关设备自动进行数据恢复工作。

[0040] 设置第一存储虚拟化网关设备11,第二存储虚拟化网关设备12以及HA功能模块13的目的在于,当本地站点发生故障时,可以实现故障的自动切换。自动化切换相比于人工切换而言,将极大的缩短业务从中断到恢复正常所需的时间。本发明实施例提供容灾系统,可以使得RTO降低至2分钟之内。

[0041] 其中,本地站点的故障类型可以有多种,包括有服务器故障,存储故障和站点故障,其中,ICS可以对站点的运行情况进行检测,发现站点出现何种故障。

[0042] 对于服务器故障而言,当本地站点的服务器出现故障,所述HA功能模块13将所述服务器包含的虚拟机进行服务器的切换。

[0043] 考虑到本地站点中可以同时存在多台服务器,一台服务器可以为多台虚拟机提供服务支持,当其中一台服务器出现故障时,可以将该台服务器对应的虚拟机切换至其他未出现故障的服务器。具体的,当所述本地站点包括有至少两个服务器时,若第一服务器出现故障,则所述HA功能模块将所述第一服务器包含的虚拟机切换至第二服务器,由所述第二服务器为所述虚拟机提供服务支持。

[0044] 其中,第一服务器可以是所述至少两个服务器中的任意一个服务器;第二服务器可以是所述至少两个服务器中除所述第一服务器之外的任意一个服务器。

[0045] 例如,在本地节点中具有两台服务器。分别为服务器A和服务器B,服务器A为虚拟机1和虚拟机2提供服务支持,服务器B为虚拟机3和虚拟机4提供服务支持,当服务器A出现故障时,HA功能模块便可以将虚拟机1和虚拟机2切换至服务器B,由服务器B为虚拟机1和虚拟机2提供服务支持,从而保证虚拟机1和虚拟机2的正常工作。

[0046] 除上述类型的服务器故障外,也可能出现本地节点上的所有服务器都出现故障的情况,此时需要在本地节点和备用节点之间进行服务器的切换。具体的,当所述本地站点包括的所有服务器出现故障时,所述HA功能模块13将所述本地站点包含的所有虚拟机切换至所述备用站点的服务器,由所述备用站点的服务器为所述虚拟机提供服务支持。

[0047] 对于存储故障而言,在实际应用中,一般会按照默认的存储路径存储数据,当本地站点的存储发生故障时,若仍按照原来的存储路径,可能无法实现数据的存储,此时,本地站点的第一存储虚拟化网关设备11可以自动将本地站点的IO指向所述备用站点。

[0048] IO可以用于表示数据存储的路径信息,将本地站点的IO指向备用站点后,本地站点产生的数据会存储于备用站点,以保证数据存储的正常运行。

[0049] 对于站点故障而言,当本地站点出现故障时,此时本地站点已经无法工作,为保证业务的正常运行,需要进行站点的切换。具体的,当本地站点出现故障时,所述第一存储虚拟化网关设备11和所述HA功能模块13同时运行,将所述本地站点的业务系统切换至所述备用站点。

[0050] 需要说明的是,在本发明实施例中,虚拟化网关集群可以采用ESC (Enhance Stretch Cluster)的工作模式,上层应用在读写数据时可以实现优先本地读写,在本地站点未发生存储故障时,不会出现跨站点读写的情况。

[0051] 由上述技术方案可以看出,在自动化容灾系统中,通过将本地站点和备用站点设置于同一个池组中,以便于使得本地站点和备用站点可以具有同一个存储位置,从而可以支持HA功能模块实现HA功能,当本地站点的服务器出现故障时,HA功能模块便可以自动将所述服务器包含的虚拟机进行服务器的切换;通过在本地站点中设置第一存储虚拟化网关设备,在备用站点中设置第二存储虚拟化网关设备,这两个存储虚拟化网关设备可以形成虚拟化网关集群,用于实现两个站点之间数据的同步;当所述本地站点的存储出现故障,所述第一存储虚拟化网关设备可以自动将所述本地站点的IO指向所述备用站点;当所述本地站点出现故障,所述第一存储虚拟化网关设备和所述HA功能模块同时运行,将所述本地站点的业务系统切换至所述备用站点。可见,依据于第一存储虚拟化网关设备、第二存储虚拟化网关设备以及HA功能模块,可以实现本地站点和备用站点之间的数据同步,在发生故障时实现故障的自动切换。并且,由于数据的同步,在故障恢复后,可以由第一存储虚拟化网关设备和第二存储虚拟化网关设备自动进行数据恢复工作。

[0052] 考虑到在实际应用中,本地站点和备用站点可能会出现数据不统一的情况,造成数据不统一的原因可能是本地站点和备用站点之间的连接出现中断等。对于该种情况,可以在第三站点(如图2所示的Site-C)上设置仲裁节点,仲裁节点可以对本地站点和备用站点的数据进行判别,区分出哪个站点的数据是正确的。

[0053] 在具体实现中,仲裁节点分别与所述本地站点和所述备用站点连接,用于当所述本地站点和所述备用站点出现数据差异时,对所述本地站点和所述备用站点的数据正确性进行判别。

[0054] 其中,仲裁节点可以是一台服务器即仲裁服务器,也可以是一块磁盘,对于仲裁节点的具体形式,本发明实施例不做限定。

[0055] 为了便于工作人员可以及时获知本地站点出现故障,在本地站点上可以设置报警模块。当所述本地站点出现故障时,可以触发报警模块,进行报警提示。

[0056] 由上述介绍可知,当本地站点出现故障时,可以进行自动化的故障切换,该故障往往是一些突发的故障情况,也即是预先未知的故障。除此之外,还存在一些预先可知的故障,例如,需要对本地站点进行断电检修等,此时本地站点无法提供业务支持。对于该种情况,可以跨站点做计划迁移,具体的,本地站点可以采用在线迁移(vmotion)将本地站点的

业务系统切换至所述备用站点。

[0057] 其中,vmotion可以将正在运行的虚拟机从一台物理服务器移动至另一台物理服务器,而不影响业务的运行。

[0058] 以上对本发明所提供的一种自动化容灾系统进行了详细介绍。说明书中各个实施例采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似部分互相参见即可。对于实施例公开的装置而言,由于其与实施例公开的方法相对应,所以描述的比较简单,相关之处参见方法部分说明即可。应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以对本发明进行若干改进和修饰,这些改进和修饰也落入本发明权利要求的保护范围内。

[0059] 专业人员还可以进一步意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、计算机软件或者二者的结合来实现,为了清楚地说明硬件和软件的可互换性,在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本发明的范围。

[0060] 结合本文中所公开的实施例描述的方法或算法的步骤可以直接用硬件、处理器执行的软件模块,或者二者的结合来实施。软件模块可以置于随机存储器(RAM)、内存、只读存储器(ROM)、电可编程ROM、电可擦除可编程ROM、寄存器、硬盘、可移动磁盘、CD-ROM、或技术领域内所公知的任意其它形式的存储介质中。

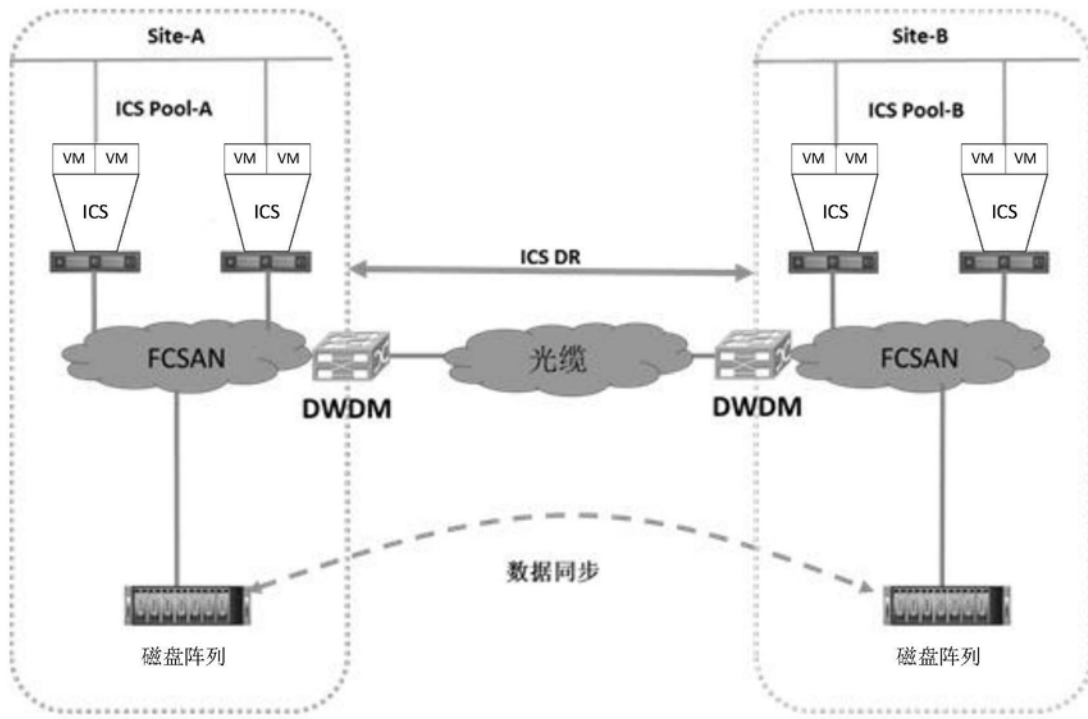


图1

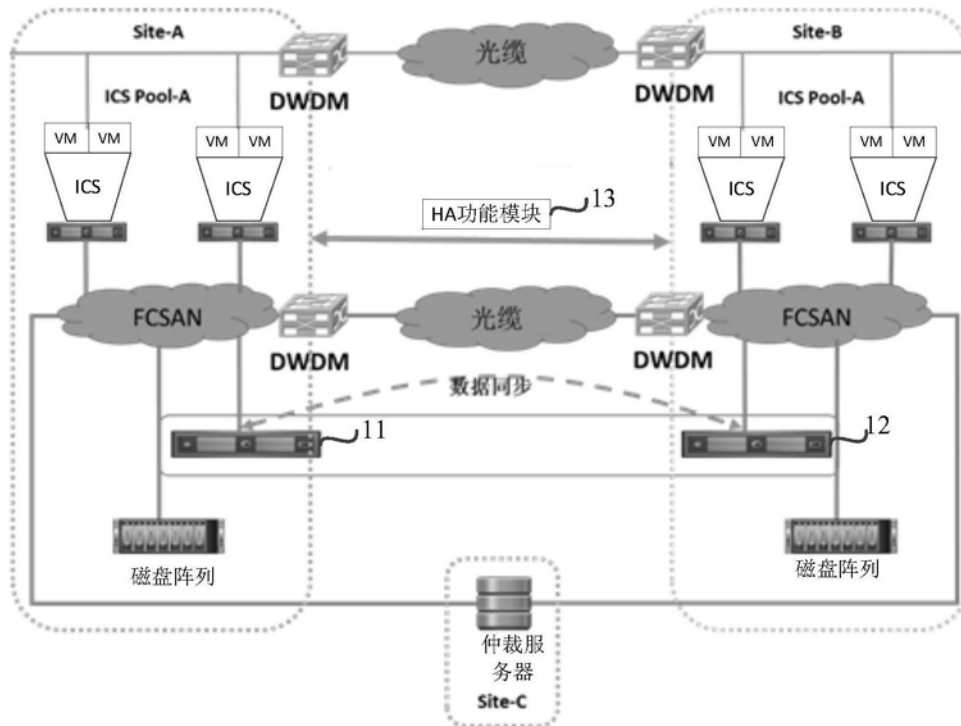


图2