**南通市第六人民医院分布式存储建设项目需求**

# 一、医院存储现状

目前医院私有云存储主要分为以下几块：1、超融合存储私有云：目前主要存储着医院现有部分主要核心业务系统的数据；2、全闪型存储：目前部分核心业务系统的数据，及超融合业务的数据备份；3、影像存储：现在医院目前的影像存储，目前可使用的空间已经达到了满载。4、全医无纸化业务系统上线，需要分布式存储的数据存储空间。

# 二、项目建设需求

目前医院对容量需求最大的主要是PACS系统，需要完成各种影像数据进行采集、存储、传输和处理，并进行全院共享，PACS影像具有单张数据量大、数量多的特点。

1、为保障良好的业务体验，从医生点击阅片到完成影像展示时间应在2~3秒以内，因此需要网络保障在1秒内完成数据传输；

2、PACS影像数据于其他医疗数据一样，需要长期可靠存储，根据国家法律法规要求，至少需要保证15年以上的数据存储。

另外众多的应用系统在运行中需要做一些配置和数据的备份，这部分数据对IOPS和延迟要求没有那么高，但容量需求比较大，对存取速度快、网络带宽要有一定性能保障，需要将这部分数据从全闪存存储中迁移出来，通过双活可以满足这部分数据对性能、带宽的要求，把稀缺的SSD存储资源用在关键性应用上。

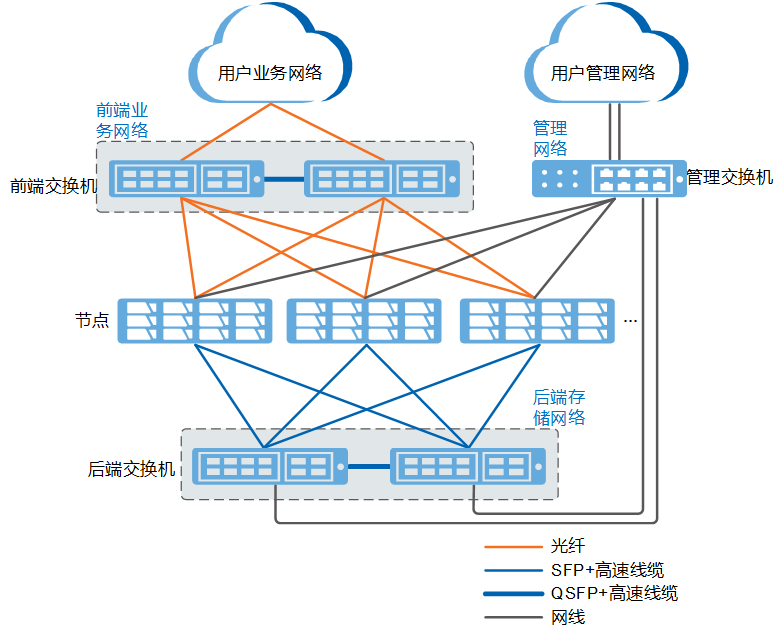
3、医院无纸化业务系统的建设，医院目前无纸化建设已经提上日程，无纸化业务中会产生大量的细小型文件，文件类型与现有PACS影像数据类似，我们统一定义为非结构化数据。考虑到医院现在的分布式存储的空间无法满足新的业务上线，需要对原有分布式存储进行数据扩容，以保证无纸化业务系统的上线.

4、医院目前的影像存储容量已使用达到了百分之九十以上的空间，后期为保证医院的影像类的数据的存储，调阅需要对影像存储的容量进行扩容，以满足未来的影像类数据的存放。

扩容后存储容量可以满足到原有PACS影像的数据存放，同时满足医院未来无纸化业务系统的部署，以及无纸业务系统使用后的数据容量增长存放。

# 三、分布式存储系统介绍

## 3.1架构



分布式存储系统具有大容量、高性能和易扩展性等方面的突出优势，同时提供丰富的增值特性，可广泛应用于广电媒资、高性能计算、数据中心、互联网运营和大型企业等各行业领域。

大数据是海量、高增长率和多样化的信息资产，大数据需要提供高效的分析处理模式才能提供更强的决策力、洞察发现力和流程优化能力。大数据的常见特点是：Volume、Velocity、Variety。大数据时代的到来，对存储系统的诉求也随之而来：

Volume特征包括容量和数量，需要存储具备灵活的扩展性。

Velocity特征需要采用分布式处理，采用Scale-out架构并使用通用硬件，在成本可控的前提下提供高性能。

Variety特征需要存储提供更丰富的接口，以应对不同业务访问存储需求。

分布式存储系统是针对大数据的特点，采用全对称架构，提供3～288节点的弹性无缝扩展

分布式存储采用全互联全冗余的组网机制，允许系统中所有节点并发访问整系统的任何数据；支持从多个节点并发访问相同数据的不同区域，实现高并发高性能读写，对外提供标准的NFS/CIFS/FTP接口。

## 3.2性能

分布式存储针对文件系统应用提升性能的关键技术包括：

并发访问

传统的NAS系统由引擎和存储单元组成。数据的全部并发访问需要先由NAS引擎处理，因此引擎容易成为性能瓶颈。

采用全对称的逻辑架构，每个节点均可提供外部业务访问。通过负载均衡设计，数据访问在集群内均匀分布，可大幅提升系统并发访问能力。

全局缓存

数据在缓存中的读写速度远高于硬盘，但单节点的缓存提升空间有限。同时如果各节点的缓存相互独立，则缓存命中率难以提升，数据的一致性保证机制也会趋于复杂。

通过全局缓存模式，整合所有节点缓存，使同一文件的数据在缓存池只缓存一份，任意节点均可命中，有效提高数据访问命中率，减少硬盘读写次数，降低访问时延，提升系统整体性能。

高速互联

分布式存储支持10GE高速以太网，单个节点可提供4个10GE端口。还同时支持IB（InfiniBand）网络，可满足更加苛刻的性能要求。

可适配用户的不同组网需求，通过节点间高速互联，降低系统内部时延，提供卓越性能。

SSD加速和自动分级存储

SSD硬盘可解决数据读写性能的瓶颈，提供低延迟和高吞吐量。分布式存储可自动识别小文件和热点数据，根据热度在系统内不同存储介质之间进行数据迁移，使用SSD硬盘存储热点文件，提升资源利用率。

SMB3 Multichannel

支持SMB3 Multichannel功能，充分利用客户端内的多核CPU和带宽资源，可大幅提升业务性能和可靠性，且在单个连接通道故障后使用其他通道继续传输数据，不影响业务。

内部文件快速复制

内部文件复制时，无需经过客户端，直接在系统内部高速完成。

NFS协议增强

传统NFS协议受到单网口带宽、块大小等限制，单客户端能够支持的极限带宽为1GB/s，无法满足更高的要求。分布式存储提供NFS协议增强特性，通过在NFS客户端上配置多个网口和安装NFS协议优化插件DFSClient，实现多网口并发连接和缓存优化，大幅度提升单客户端性能。

Mac OS X单客户端带宽达到1.5GB/s。

RHEL单客户端带宽达到2.5GB/s。

磁盘碎片整理

分布式存储提供磁盘碎片整理功能，在指定时间段，周期性的执行磁盘碎片整理任务，保证系统长时间运行的整体性能。

## 3.3安全性

安全维护具有保证系统安全、正常运行的作用，重视安全维护将保障业务的稳定与可靠。

目前，应用系统面临日趋严重的安全威胁，一旦出现问题，将面临业务中断、数据丢失，甚至系统崩溃的危险。因此，客户需要从多个层次构建、维护整个应用系统的安全屏障，提前发现并处理各种可能存在的安全问题。

另外，完全依赖技术很难彻底清除安全隐患、全面保证应用系统的安全。因此，客户需要根据安全维护建议和日常发现的问题，建立安全管理的制度，来保证系统安全、正常的运行。

账户安全

帐户的安全是系统安全的重要组成部分。涉及帐户的必要性、权限和管理等。帐户安全囊括了分布式存储系统中所有的默认系统帐户，给出系统帐户列表及其管理方法。

网络安全

网络安全的重点是保障分布式存储与周边网元间的组网规划及网络通信安全。网络安全涉及组网的安全规划、组网的安全变更和网络的日常维护。

安全证书与密钥

安全证书是保障系统日常使用安全及通信安全的重要因素，需要对安全证书和密钥进行管理。对分布式存储中包含的所有安全证书与密钥进行说明。

协议与加密算法

文件传输协议、网络协议、加密协议和算法在使用过程中均需要注意规避安全风险。对分布式存储使用的文件传输协议、网络通信协议、加密协议等协议层面的内容提供安全管理建议。

系统维护

在进行日常使用和系统维护的过程中，需要管理员注意安全的管理，防范于未然。系统维护着眼于管理员对分布式存储进行管理与维护时需要关注的安全事项，包括可能会对系统安全造成影响的端口、补丁、日志等。

## 3.4可扩展性

系统扩展性

采用Scale-out方式扩容：每个节点都具备CPU、内存、磁盘，且CPU/内存/磁盘比基本相同，因此在增加存储节点时，整个系统硬件的CPU/内存/磁盘比保持稳定，从硬件层面提供了Scale-out的能力；整个系统在软件层面通过统一的资源和IO调度，使得新接入的存储节点能够与已有的节点共同、并发的为上层业务提供存储服务。因此方案扩容能够保证容量与性能的线性增长。

方案扩容时，不但增加了磁盘空间，同时扩展了CPU、内存、网络端口等各种硬件资源。方案通过软件协调大量的存储节点并发对外提供服务，从而使得系统在扩展了存储容量的同时提升存储系统的性能。

负载均衡性

采用Scale-out存储架构，将海量数据压力分散到多个并发存储节点，避免资源争用，系统性能（吞吐量）按照比例扩展，并且各个存储节点之间负载均衡，有效避免单节点性能瓶颈。这种架构，不仅保持了对象存储系统高可靠、高性能的优点，而且使得云存储系统具有更好的扩展性。

存储能够根据业务增长进行平滑扩容，系统每增加一个存储节点，都能够自动识别，将该存储空间加入到整系统中，并根据负载均衡的原则，让更多的数据和负载迁移到新节点上，从而简化了系统扩容的管理，降低了操作成本。

容量、性能可扩展性：1分钟内完成单节点在线扩展，业务不中断，容量和性能同时提升，轻松扩展到PB级的可用容量和10GB/s以上的带宽，满足未来5年的业务需求；

接口可扩展性：随着用户业务的增长，更多服务器可以快速接入存储系统；

新技术可扩展性：方案采用的硬件和软件都是业界相关领域的开放型主流产品，采取开放性模块化设计，当新的存储技术出现后，可以通过升级相应的模块实现向新技术的扩展。

## 3.5可靠性

采用电信级可靠性理念设计，可靠性高达99.999%，具体可靠性设计技术如下：

存储网络冗余设计

内部交换模块具备电信级可靠性，且与用户IP环境组成双交换网络，通过IP协议确保任何一个网络端口、网络交换模块、网络链路发生故障，业务系统依然能够正常访问存储中的数据，从而保证系统可靠性。

硬件级可靠性

方案的核心组件采用电信级硬件，有效的提高系统的可靠性，使用的可靠性技术包括：

多BIOS设计，任何一个BIOS离线或故障，其余的BIOS将接管服务，对外业务不中断，从而使得硬件和Fireware支持在线升级；

风扇采用2+1冗余，支持热插拔、自动调速；电源采用1+1冗余。节点内热敏位置都布置了温度监控部件，通过这些部件实时监控系统中各个产生热敏位置的温度，根据温度的高低平滑的调控风扇转速；

可以有效防止系统过热，造成不稳定；

风扇无须时刻都工作在高转速状态，降低风扇的故障率和能耗；

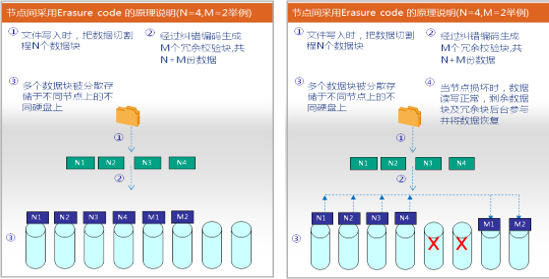
有效减少灰尘的进入、提高了系统的可靠性。

系统级可靠性

Erasure Code

分布式存储采用Erasure Code保障数据一致性，确保海量数据存储场景下的数据安全可靠。Erasure Code是RAID的超集，能够支持比传统RAID算法更高的可靠性和更灵活的冗余策略，其设计思路是对文件进行原始分片（N份），通过纠错编码生成M个冗余校验块文件。写入时，由客户端进行切片（N）和转码（M），一共生成N+M份数据，存储系统自动从各存储节点中选择一个硬盘组成一个N+M个磁盘的磁盘组，分别写入N+M份数据。任意一份数据损坏的情况下，可以通过其它数据恢复，最大可以支持M份数据的损坏。

分布式存储采用Erasure Code算法保障数据可靠性，将N+M份数据保存在N+M个不同的虚拟节点上，其中N支持2-16，M支持1/2/3/4。可以看出，它最大可以支持4份数据同时损坏，并且可以通过其它数据分片，恢复原始数据，而不影响系统业务的连续性，数据保护性强于传统的RAID5（只允许一份数据损坏）、RAID6（只允许两份数据损坏）。如下图所示：



专门针对小文件存储进行了优化。如果配置了N+M的保护方式，当一个文件的数据量不超过一个分片大小时，系统会自动将其转化为1+M份的存储方式，当数据量超过一个分片大小，系统才会对它进行校验计算。这种优化，可以提高小文件的读写性能和存储利用率。

元数据高可靠

分布式存储的元数据，采用存储多份副本的方式，并且对于配置了N+M保护方式的文件来说，其元数据的保存份数自动设定为M+2份，这使得分布式存储系统的元数据的可靠性高于其数据。

数据自愈

分布式存储采用的InfoProtector技术可以提供最多同时故障4个全节点的数据保护。并且硬盘故障时，InfoProtector确保将单个故障盘的重新构建时间缩至最短，达到2TB数据只需要1小时恢复时间。

数据存储时被分片打散到多个存储节点上，这些分片数据支持分布在不同的存储节点之间。在数据发生故障导致数据不一致时，分布式存储通过内部的自检机制，自动发现数据故障。发现故障后启动数据修复机制，在后台修复数据。由于数据被分散到多个不同的存储节点上保存，数据修复时，在不同的节点上同时启动修复，每个节点上只需修复一小部分数据，多个节点并行工作，有效避免单个节点修复大量数据所产生的性能瓶颈，对上层业务的影响做到最小化。

## 3.6可管理性

支持可视化自动部署，提供了专门的网管管理工具，它支持WEB UI形式，可以方便的显示所有设备的组网结构，所有设备的容量，CPU，内存等信息，以及所有进程的运行状态，业务的运行状态，心跳的状态信息，实时展示系统性能信息，保存和查询操作日志。管理维护高效集中，简洁易用。管理员可以通过集群内任意节点访问网管界面，完成对硬件、软件、集群和业务的统一管理。单个节点故障不影响网管的正常登录和使用。

免即时维护

采用自恢复全冗余空间技术，系统自行检查到系统故障时，启动自恢复流程，因采用了全局热备空间，只要还具备新的空余空间，那么系统将可以持续进行自恢复流程，无需手工干预。无需24小时内及时更换部件，可以采用固定的维护周期，如维护周期采用以半年或1年为单位，批量更换硬盘，简化系统维护难度，降低维护成本。

丰富、及时的告警信息

ISM提供丰富、及时的告警信息传递方式，保证系统出现故障时，能够通过多条途径、多种方式将告警信息传递到多个维护人员，从而协助系统故障的及时排除。

短信猫方式

可通过短信猫发送告警短信到指定手机进行告警管理，该告警传递方式直接有效，并可同时制定多个手机号码进行短信发送。

邮件告警

通过管理界面配置告警邮箱，存储设备产生的告警可通过管理网口发送到指定邮箱，管理员可通过查看邮件来管理阵列系统。

声音告警

系统发生告警后存储阵列可发出告警声音，也可以在管理界面所在终端发出告警提示声音。

指示灯告警

有紧急故障发生时，系统运行指示灯自动变为红色，以提示告警。

特有的自动管理

方案特有自动部署、故障自愈、自动升级等自动化管理特性，通过这些自动化管理特性，将用户从海量云节点的管理中解脱出来，帮助客户提高管理效率，降低管理和运维成本。

自动部署

支持自动部署，设备上电后管理软件通过远程远程管理自动对所有设备上的软件配置、安装，60分钟可完成系统的所有部署与配置，过程无需用户手动参与，真正实现即插即用。

故障自愈

通过自有的专项技术，系统中任何一个组件（包括磁盘、控制器、节点）发生故障，系统将自动检测、隔离该故障；如果出现多盘失效，系统自动通过其他节点在系统其余的可用空间中快速恢复受故障影响的数据。整个过程无需用户手动参与,整个过程无需中断业务，且对业务完全透明。

自动升级

可通过ISM一键完成海量存储节点的系统升级，升级过程无需人工干预和控制，系统自动判断各节点状态，并启动升级；升级过程无需用户手动参与；如果升级遇到问题，系统将自动回退；整个升级过程无需中断业务，对业务完全透明。

自动精简配置（Thin Provisioning）

自动支持Thin Provisioning，可以按需分配存储空间，并且这个空间是动态分配的，支持弹性扩容，实际需要使用多少空间，就为它分配多少空间，客户存储的数据容量增长时，分配的空间也弹性增长，为客户解决了在项目初期为媒资管理平台中不同应用精确存储空间规划的苦恼，也有效避免了资源浪费。

精细化配额管理

支持按照用户、用户组或目录的配额管理和配置，并且支持嵌套的配额管理，便于资源管理，匹配电视台组织结构。

基于目录的配额：管理员可以针对为空的共享目录设置配额，限制该目录的最大可用存储空间。

基于用户/用户组的配额：管理员可以对任意用户/用户组设置配额，限制该用户/用户组可以使用的最大可用存储空间。

# 四、预算及清单配置表

|  |  |  |  |
| --- | --- | --- | --- |
| 类别 | 品牌型号及配置参数 | 数量 | 预算价格 |
| 影像存储扩容 | 1. 配置4个分布式存储节点，每节点硬件配置：  配置两颗处理器,每颗主频≥2.6GHz,每颗物理核数≥32核；256G内存；2\*600G 10K SAS系统盘，1\*1.6T NVMe SSD缓存盘，12\*10T 7.2K SATA数据盘；4\*10GE光口(含光模块)；配置冗余电源和风扇；提供三年维保；  2. 软件配置：软件包(配额、分级存储、QoS、多租户、快照、异步复制、审计日志、回收站、WORM)； 3.存储交换机2套：48个10GE SFP+端口,6个100GE QSFP28端口,2个交流电源,端口侧进风；配置10个10G SFP+多模光模块，一根40G QSPF+高速线缆-3M；提供三年维保； | 1套 | 750000 |