

# **实际使用评测** 报告

# 大数据、高性能环境对存储需求

EMC Isilon在深圳华大基因使用评测

作者:杨恩博:中桥调研咨询分析师

王丛(Kim Wang):中桥调研咨询总经理兼首席分析师

彭雅芳:中桥调研咨询助理分析师 马艳:中桥调研咨询调研分析师

2012年12月



# 目录

简介	2
背景	2
高性能计算的应用	3
EMC Isilon在高性能应用	5
华大基因	
华大基因简介华大基因HPC环境华大基因IT挑战	7 
中桥实验室对华大基因IT中心的评测	10
高性能	14 16
结论	20

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

一直以来,高性能计算的主要目的就是通过提高运算速度来解决大规模科学计算和海量数据的处理问题。高性能计算每秒万亿次级的强大计算能力,使其成为石油、生物勘探、气象预测、生命科学研究等领域的重要技术选择。但是随着数据量以及数据价值的不断增长,金融、电信、互联网等领域对高性能计算的需求不断加大。随着技术的发展,高性能计算系统的处理能力越来越强,任务的计算时间越来越短,对业务的价值不断提高。但是,要想实现快速的任务计算处理,高性能计算系统的存储能力是关键。因为在计算开始,要从存储系统中读取数据;计算结束时,要向存储系统中写入计算后的结果。如果这之间的读取和写入速度不匹配,不仅会拖延高性能项目的完成周期,低延迟还会严重影响高性能创造价值的能力。通常,高性能计算要求存储系统能够满足性能、可扩展性要求,保护投资回报:吞吐量达到几个甚至几十个GB/s,容量能扩展至PB级;透明的访问和数据共享;集中式的智能化管理,高性价比;可按需独立扩展容量和性能等。中桥分析师在深圳华大基因研究院实地测试了EMC Isilon产品在其HPC环境下的运行情况,并记录下其结果。

## 背景

高性能计算(High Performance Computing—HPC)指通常使用很多处理器(作为单个机器的一部分)或者某一集群组织中几台计算机(作为单个计算资源操作)的计算系统和环境。长期以来,高性能计算应用的主要领域是科学与工程计算,诸如高能物理、核爆炸模拟、气象预报、石油勘探、地震预报、地球模拟、药品研制、CAD设计中的仿真与建模、流体力学的计算等。如今,像金融证券、政府信息化、电信行业、教育、企业、网络游戏等领域对HPC的需求也在迅猛增长。

#### 高性能计算的应用

高性能计算有着广泛的行业应用基础,下面列举几个行业对高性能计算的应用需求:

#### 1. 航空航天行业

在航空航天行业,随着中国航空航天事业的快速发展,尤其是载人航天技术的巨大成功,我国科技人员对空气动力学的数值模拟研究提出了越来越多的需求,常规的计算能力远远无法满足复杂的大型飞行器设计所带来的巨大需求。在航空航天企业的设计过程中,研究人员往往需要把飞机表面分成几百万甚至几千万个离散型的网格点,然后通过高性能计算平台求解方程,得出每个网格点的温度、速度、摩擦力等各种参数,并模拟出连续型的曲线,进而为飞机设计提供宝贵的参考资料。对这类计算来说,网格点分割得越细密,计算结果的精确度也就越好。但是这些大规模设计计算问题不但单个作业计算量庞大,且需不断调整、重复计算,因此高性能在航天航空行业中占据着举足轻重的地位。

#### 2. 能源行业

石油能源作为国家战略资源,对于国家经济、安全、军事等各方面都具有非常重要的战略意义。石油勘探承担着寻找储油构造、确定井位的重要任务。目前的主流做法就是人为的制造相应规模的地震(视勘探地区面积与深度不同),同时在相应的地层遍布若干震波收集点。由于不同材料的地质环境对地震波的影响是有规可循的,所以借助这一点,通过相关的算法,即可以通过对地震波的传递演算来"计算出"地质结构,从而找出我们所需要的能源位置。这种计算量无疑是异常庞大的,由于地震波法勘探收集的数据通常都以TB计,近年来海洋油气勘探所采集的数据甚至开始向PB规模发展。为此,只有借助高性能计算,才能在最短的时间内处理这些海量数据。

#### 3. 生命科学

在现代生命科学领域,以数据为驱动力的改变正引发着巨大的变革。海量生物数据的分析将会增强疾病的实时监控能力和对潜在流行病做出反应的能力,但海量数据的挖掘、处理、存储却面临着前所未有的挑战。特别是



随着新一代测序技术的迅猛发展,基因组学研究产生的海量数据正以每12-18个月10倍的速度增长,已远超越著名的摩尔定律,这使得众多生物企业和科研机构面临强大的数据分析和存储需求。

在国内,生物基因行业的发展势头也不可小觑。2011年1月30日,国家发改委已批复同意深圳依托华大基因研究院组建国家基因库,这是中国首次建立国家级基因库,首期投资为1500万元。深圳国家基因库是一个服务于国家战略需求的国家级公益性创新科研及产业基础设施建设项目,是目前我国唯一一个获批筹建的国家级基因库,是全球仅次美国、日本和欧洲三个国家级基因库之后的世界第四个国家级基因库。现在,该国家基因库已经收集了100万GB的生物数据,包含基因组、转录组、蛋白质组、代谢组及表型的数据,同时也积累了约四十万份生物样本。预计该基因库最终将达到10亿GB级别的数据容量。深圳国家基因库和国际上已有的基因库相比,它的特点是既有"湿库"也有"干库":前者把千万种实体的动植物、微生物和人类组织细胞等资源和样本纳入网络;后者汇集巨量的核酸、基因表达、蛋白、表型等多类数据信息,成为"大数据"生物学时代研究生物生长发育、疾病、衰老、死亡以及向产业化推广的利器。

#### 4. 金融行业

金融说到底就是数据。在金融市场中,拥有速度就意味着更高的生产力和更多的市场份额。金融计算模型相当复杂,数据收集越多,计算结果越精确。金融分析师都迫切地需要一个能模拟复杂现实环境,并进行精确处理的金融计算程序,以便对每个投资产品及时地评估投资收益,衡量投资风险,以期获得更好的投资回报。也正因此,高性能计算已经越来越多地应用到全球资本市场,以期在最短时间内实现对市场的动态响应与转换。

#### 5. 气象预报

世纪二十年代初,天气预报方程已基本建立。但只有在计算机出现以后,数值天气预报才成为可能。而在使用 并行计算机系统之前,由于受处理能力的限制,只能做到24小时天气预报。高性能计算是解决数值预报中大规 模科学计算必要手段。采用高性能计算技术,可以从提高分辨率来提高预报精度。

#### 6. 游戏动漫和影视产业

随着3D、4D电影的兴起和高清动漫趋热,由高性能计算(HPC)集群构成的"渲染农场"已经成为三维动画、影视特效公司不可或缺的生产工具。动漫渲染基于一套完整的程序进行计算,从而通过模型、光线、材质、阴影等元素的组合设定,将动漫设计转化为具体图像。以《玩具总动员》为例,如果仅使用单台工作站(单一处理器)进行动画渲染,这部长达77分钟的影片的渲染时间将会是43年,而采用集群渲染系统,只需约80天。

#### 大数据下高性能计算的挑战

尽管高性能计算的应用范围已经越来越广,但是其面临着大数据集带来的全新挑战。高性能计算如今要解决的计算难题极为复杂,其负载程度与十年前要解决问题的难度相比要高出多个数量级,并且复杂程度仍在不断增加,不断挑战着技术的极限。例如,当代石油物探高性能计算面临着地震勘探数据量海量增长的严峻形势。从上世纪80年代的2-36MB/km2的2D数据,增长至3D的30-300GB/km2。一个寻常勘探项目的原始数据通常都在十几TB左右,而要真正处理这些数据,至少要五倍于原始数据的存储空间。

高性能计算运行的应用程序一般使用并行算法,把一个大的普通问题根据一定的规则分为许多小的子问题,在集群内的不同节点上进行计算。之后,对这些小问题的结果进行处理并合并为原问题的最终结果。通常,这些小问题的计算是可以并行完成的,从而缩短问题的处理时间,提高系统的运算速度。

高性能计算最典型的三个数据流程包括: 创建输入数据、运行应用程序进行分析处理和结果归档管理。

1. 创建输入数据: 创建数据的数据安全可靠性和一致性非常重要。如果丢失了输入数据,通常可以通过重新运行应用程序,来重建丢失的数据。重新创建数据不仅费用高昂,很多高性能环境的数据往往是无法再次生成的。因此,高性能创建数据的价值往往不是能用金钱来衡量的。



- 2. 应用程序分析处理: 在分析处理阶段,保证执行应用程序的读/写性能是高性能分析效率和项目周期的关键。 这可能需要使用高性能可扩展性存储系统来满足吞吐量和存储容量的需求。
- 3. 数据归档: 高性能环境下,不同属性的数据在数据生命周期的不同阶段体现出来的价值是不同的。归档可以释放出主存储空间,使之用于主要的应用程序和项目。如何用不同存储介质存储不同数据,是高性能环境降低数据生命周期总成本的关键。

高性能计算的分析效率取决于计算能力、带宽和存储三方面。数据密集型计算如何保证存储为海量大数据并行处理 提供稳定的性能和可扩展的容量,在存储超大规模数据量的同时,满足多节点集群计算对存储I/O带宽的需求,是 保证高性能处理能力和效率的关键。

高性能计算集群系统中的节点,可分为计算节点和存储节点。其中,存储节点是指集群系统的数据存储器和数据服务器。如果需要存储TB级的数据,通常需要部署并行文件系统及多台IO服务器;计算节点功能则是执行计算。众多的计算节点带有IO流量瓶颈问题。当承载的计算任务被分布到众多的计算节点上实现,存储最终还是要汇总到一起。高性能计算中的计算节点可以是服务器、主机、工作站甚至PC笔记本等。计算节点对统一存储的必须性要求和各节点所汇集而来的IO流量对存储造成的冲击,是每个高性能计算中必须要考虑的因素。

#### 将横向扩展NAS用于大数据分析

#### Scale-out (横向扩展)

随着数据量的增加,数据的查询、存储、分析会越来越依赖于高性能、高可扩展存储和处理等设备。传统的scale-up(纵向扩展)架构存储,不仅扩展性、易部署管理性很难满足高性能需求;同时,容量的扩展还可能产生性能衰减,无法满足存储容量和性能线性升级的扩展需求。解决海量数据的处理、消除I/O瓶颈,实现数据存储安全和保护、降低升级运维复杂度,都成为了高性能计算用户保证高性能持续高性价比的最重要决定因素。

ESG调研结果曾表明,Scale-out存储已逐渐进入了商业IT,并且站稳脚跟。在商业和政府部门,scale-out网络存储从2010年的4,189 PB将增加到2015年的62,834 PB,达到了72%的复合年增长率(CAGR),远远超过了整体的外部网络存储54%的增长率。这个高速增长是基于供应商对于前景的预测,及早地从scale-up架构过渡到scale-out架构,并且他们的产品线中甚至是高端产品都提供scale-out架构。

横向扩展存储系统克服了物理机架和模块的限制,可作为单一系统,通过增加控制器或容量节点来实现性能和容量的独立升级,提高IT投入的回报率。同时,线性扩展能力为业务的长期高性价比提供了保障。解决了传统单一系统、模块化系统需要物理磁盘级别的管理、数据布局和性能调优的弊端。横向扩展平台不仅能够提升性能而且还可以降低操作成本,使单一系统在单一全局域名下,轻松扩展到若干PB容量范围,成为管理猛增数据的理想存储平台。

#### Hadoop

大规模数据的采集与处理已经让高性能计算用户备感压力,迫使企业需要通过一种新的方式去管理。一些新的分布式平台技术也不断出现,其中比较突出的是Hadoop。Hadoop是一个比较新的技术,它是一个建立在普通服务器上的高度可扩展、高可靠的分布式系统,专为大规模计算和海量数据存储而设计。Hadoop的数据来源可以是任何形式,在处理半结构化和非结构化数据上与关系型数据库相比有更好的性能,具有更灵活的处理能力。基于Hadoop,用户可编写处理海量数据的分布式并行程序,并将其运行于由成百上千个节点组成的大规模计算机集群上。

总之,在高性能计算系统中,计算节点的扩展、硬件升级能够进一步增强计算节点集群系统的处理能力,运算时间得到进一步缩短;但是存储设备的I/O能力也必须有所提高,提高I/O节点与存储设备之间的数据交换能力,从而提高整个高性能计算系统的计算效率。

在本报告中,中桥实地走访了华大基因数据中心并进行测评,还与华大基因研究院副院长方林先生进行了交流。下文记录了这次测评的结果。



# EMC Isilon在高性能应用

EMC Isilon做为横向扩展集群NAS存储解决方案,被全球150家大型生命科学和医药制造、石油天然气、航天航空以及科研研究等企业用于高性能分析环境。IsilonOneFS操作系统通过一个分布式管理器、统一缓存和集群内全局一致性的区域管理器,来保持节点的同步,以实现满足并行处理环境对存储性能的要求。同时还优化了系统可扩展性、简化了部署和管理。集群内所有节点具有同等地位,都能处理读写要求,正是这种整个集群节点间的全局一致性,避免了访问文件系统时的单点故障,保证高性能环境持续稳定的读取性能需求。OneFS可以在单独的磁盘上控制文件的布置,在卷、目录甚至是文件的层级上来控制存储系统的冗余水平,消除了其他存储系统通过RAID和卷管理层来传输数据,从而产生数据布局低效率的问题。

OneFS能够把单个文件系统从18TB扩展到20PB,可以直接在单独的磁盘上控制文件的布置,这种集群上文件的分配优化,极大地改善了磁盘子系统的性能。除此之外,采用的FlexProtect-AP™软件还能够主动监控和在故障发生前抢先将数据从有风险的组件迁移出去,利用整个集群内的空余存储空间来规避数据丢失的风险,在短时间内重建发生故障的磁盘。传统的企业存储系统,则是将等值重建操作局限在存储系统的子集内进行,这不仅导致了恢复的瓶颈,增加了重新保护数据的时间,还加大了单个组件发生故障时数据丢失的风险。相比之下,FlexProtect-AP能自动地将所有的数据和错误更正信息分布到整个集群上,其高效可靠的错误更正技术保证了所有数据的完整性,即使在同时出现多个组件故障的情况下数据也能方便地存取。

此外,EMC Isilon通过各种智能管理功能,让高性能用户实现按需采购、快速扩容、简化部署管理,提高了可靠性和安全性,实现了数据保护:

- SnapshotIQ<sup>™</sup>数据保护一提供了秒级快照的解决方案,和其他厂商的快照方案不同的是,snapshotIQ是基于目录或者子目录层次做快照,无需为快照预留空间,而且每个目录可以维护高达1024个快照。
- SmartConnect<sup>™</sup>数据访问一提供了前端网络的负载均衡模式,可以提供roundrobin、CPU忙闲、网络吞吐量、网络连接数等四种均衡策略。此外, SmartConnect提供了基于NFS协议的的IP failover方式,提供了对TCP连接的容错。
- SmartPools™资源管理一是一个数据分层管理的工具,数据通过使用策略,来决定何时存放在相应的节点上,以最大化存储管理利用率和使用效果。
- SmartQuotas™数据管理─SmartQuotas是空间访问限额模块,通过此模块可以对目录和用户实现空间限额。并且,该模块可以轻松加入客户环境中现有的AD、NIS、LDAP等用户控制域,更好地提供基于用户的访问空间限额。另外还可以提供thin provisioning的功能,实现跨用户空间的共享。
- SmartLock™数据存储保护—使用基于"一次写入多次读取"(WORM)的软件来防止关键数据因意外、过失和 恶意更改或删除而遭到破坏。
- InsightIQ™性能管理一是独立于集群存储之外的一个虚拟机,可以登录到集群存储后,对所有感兴趣的数据进行挖掘、整理,生成报表。
- SynclQ<sup>™</sup>数据复制—复制和分布大型任务关键性数据集到多个站点的多个共享存储系统之上,实现可靠的容灾。
- Aspera for EMC Isilon™—提供跨城域网WAN,实现快速数据同步。

此外,EMC Isilon的产品线提供了广泛的选项,以满足企业存储的不同需求。

下表是EMC Isilon各系列的产品技术规格。



# 表 1. EMC Isilon 各系列的产品技术规格

	S200	)		X200	X400	NL400	Accelera	tor	Backup Accelerator
目标访问模式		sactio		并发/连续性		近线	单数据》 / 10GbE	充/ IOPs	备份
前端网络	4铜缆1000 Base-T (GigE) 4xGigE(铜 缆)或2xGigE & 2x10GigE (SFP+或双绞 线) 2xGigE&2x10G igE(SFP+)		同 GigE E 双绞 <10G	基本配置: 4xGbE(双绞线) 可选配置: 4xGbE or 2XGbE & 2x10GE (SFP+或双绞线)	4xGbE或 2XGbE & 2x10GE (SFP+或 Twin-ax Copper)	4xGbE或 2xGbE & 2x10GE(SFP+ 或双绞线)	10GbE 型号	1GbE 型号	N/A
互连集群网 络			d连	2 InfiniBand连接,配置双数 据率(DDR)或四数据率 (QDR)链接		2 InfiniBand连接,配置双数据率(DDR)或四数据率 (QDR)链接	2 InfiniBand连接		
СРИ	Intel Xeon Processor			Intel Xeon Processor		Intel Xeon Processor	2xIntel Quad- Core 2.3 Ghz	1xIntel Quad- Core 2.0 Ghz	2xIntel Quad-Core 2.3 Ghz
内存	24 GB	48 GB	96 GB	基本配置: 6GB 可选配置: 6、12、24或 48Gb	24-192GB	12-48GB	8GB或 32GB	8GB	8GB
非易失性 RAM	失性 512MB			512MB		512MB	N/A	N/A	N/A



# 华大基因

#### 华大基因简介

华大基因自1999年成立以来,先后完成了国际人类基因组计"中国部分"(1%,承担其中绝大部分工作)、国际人类单体型图计划(10%)、水稻基因组计划、家蚕基因组计划、家鸡基因组计划、抗SARS研究、"炎黄一号"(100%)、大熊猫等多项具有国际先进水平的基因组科研工作;同时,建立了大规模测序、生物信息、克隆、健康、农业基因组等技术平台,其测序能力及基因组分析能力世界领先,测序仪全年产出数据量与全美国四大测序中心产出数据量总和大致相当。

华大基因拥有深圳、香港、北京、武汉、杭州等数个大型生物信息学超级计算中心,总峰值计算能力达到212T flops,总内存容量达到32.7TB,总存储能力达到17PB。其中位于深圳和香港的集群的峰值计算能力分列国内生物信息领域第一和第二位,有能力为海量生物信息学数据的存储、处理和分析提供稳定而高效的资源保障。下表所示为华大基因在中国各地所部属的超级计算集群及计算性能。

表 2. 华大基因在中国各地所部属的	团超级计算集群及计算性能

地点	核	内存	存储	T flops
深圳	11,000	22TB	8PB	117T flops
香港	7,776	9.7TB	6PB	83T flops
北京	300	500GB	500TB	1.5T flops
武汉	300	500GB	500TB	2T flops
CHOPS	300	800GB	500TB	2T flops
UC Davis	300	800GB	500TB	2T flops
Denmark	600	1.6TB	1PB	4T flops
Total	20,576	37.2TB	17PB	212T flops

面对基因组学研究中分析处理的数据量的飙升,华大基因与众多生物企业和科研机构面临着强大的存储和分析需求。华大基因在HPC软硬件系统上加大了各种资源的投入,建立了具有自己特色的云计算系统以及相应的运行、管理和研发队伍。主要目标是以更低的成本更高效地完成大量的数据处理计算,通过云计算平台,与全世界的科研人员共享拥有的基因数据和软件系统,帮助他们更快更好地进行研究工作,促进以基因信息指导临床诊断和治疗的贯穿应用。华大基因在强大的计算平台上运行着一系列复杂的软件,不断揭开基因编码生命的奥秘,解答遗传密码的起源进化和基因功能表达等挑战性的问题,为人类的健康和进步提供强有力的保障。

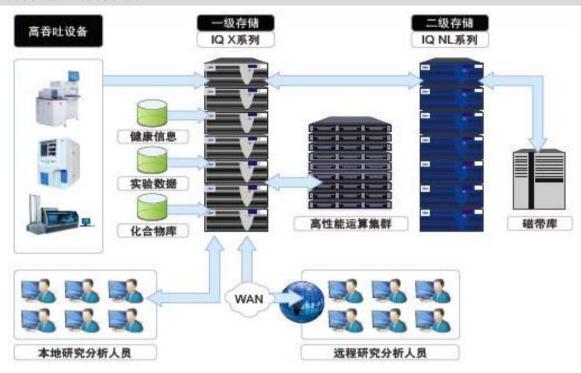
#### 华大基因HPC环境

在采访过程中,华大CIO表示: "我们的高性能计算主要应用于华大基因的科研项目并支持相关产业发展。HPC作为一个重要平台,支撑着我们的整体发展,并且在国家基因库等重大发展项目中担任核心支撑角色。我们曾参与了多项重大科研项目,如无创产前基因检测、单基因遗传病检测、地中海贫血检测、耳聋基因检测、新生儿遗传代谢疾病监测及人类白细胞抗原(HLA)高分辨基因分型等健康检测服务,上述项目均是在高性能计算的支撑下,进行大量基因组数据分析,并在相对较短的分析时间内提供准确的结果。不仅如此,我们还通过HPC技术为其合作伙伴提供在动植物基因组学、转录组学、微生物基因组学、宏基因组学、表观组学、复杂疾病基因组学、蛋白组学、药物基因组等领域的研究支持。在选择存储上,如何能降低闲置资源,根据业务发展快速扩容、简化大数据环境存储管



理复杂性和提供持续稳定的可靠性能,以及保证数据一致性和高可用性是我们评估的重要指标。同时,存储要能够支撑华大基因和全球基因协作的云计算平台,保证对Hadoop技术的支撑。经过反复评测和测试,我们采用了已经被全球多个生命科学企业采用的EMC Isilon方案。

图 1. 深圳华大基因IT架构示意图



#### 华大基因IT挑战

华大基因目前拥有4大平台,包括新一代测序平台、云计算平台、蛋白质谱平台及国家基因库。在新一代测序平台中包括罗氏旗下的454生命科学推出最新升级版Roche 454 GS FLX+测序仪、Illumina公司推出的HiSeq 2000测序仪、基于半导体芯片的新一代测序技术Ion Torrent等多种不同的测序技术平台。以Illumina Hiseq 2000测序仪为例,相对于传统测序的96道毛细管测序,高通量测序一次实验可以读取40万到400万条序列。读取长度根据平台不同从25bp到450bp不等,不同的测序平台在一次实验中,可以读取1G到14G不等的碱基数,这样庞大的测序能力是传统测序仪所不能比拟的。华大深圳数据中心138台高通量测序仪单次运行就能产生350GB数据,一天的数据产出量大概是10TB。

今年4月,华大基因推出了基于Hadoop架构的云计算平台,并在11月发布最新版的生物信息学软件,包括最新版SOAP系列软件、遗传变异检测软件、宏基因组测序数据分析软件Metacluster 4.0及两个基于云计算的软件Hecate2和Gaea2。这也是华大基因选择和持续采用Isilon集群存储的重要考虑之一。Isilon和Hadoop做了紧密技术整合,使存储能很好地支持Hadoop架构的云计算服务。如果使用传统存储,Hadoop仍有一些固有问题需要解决。例如,Apache™Hadoop™的NameNode存在单点故障问题。NameNode用于管理HDFS中存储的元数据文件,它是HDFS文件系统的核心部分,存放着文件系统中所有文件的目录树。客户端应用程序通过NameNode来定位、添加、修改、拷贝、移动和删除文件。如果NameNode发生故障,文件系统就会离线;Hadoop缺乏企业级的数据保护功能,必须手动设置HDFS的数据复制参数,生成的拷贝数有可能会导致存储利用率严重不足等问题。

仅仅将华大基因定义成高性能计算是不够精确的。我们自己定义成数据密集型的高性能计算。我们每年的业务量成倍增长,总的数据量,2010年4PB,2011年8PB,今年将10PB。我们的典型应用包含了结构化和非结构化的计算模式,同时对一个任务的数据量的规模非常大,对计算内存和存储带宽的需求也非常高。目前华大基因一天在计算集群上数据的吞吐量大概为320T,这相当于Google的十分之一。华大基因现在有接近2000个计算节点,总的计算核心



大概2万个。处理如此规模的数据量,对我们的IT提出了巨大的挑战。我们在解决存储瓶颈的问题时,需要寻找一个高性能、易扩展、部署简单、方便运维的存储系统。

华大基因存储需要解决的主要难题如下:

- 需要为华大基因的新一代测序平台、计算平台、蛋白质谱平台及国家基因库这四大计算平台提供服务
- 传统存储环境下,基于Hadoop架构的云计算平台中存储和CPU资源的利用率低、集成复杂性高、数据转移和加载进程效率低,缺少备份和灾难恢复功能
- 高性能、高并发需求强烈
- 随着业务发展,对容量和性能的需求不断增加



# 中桥实验室对华大基因IT中心的评测

#### 高性能

我们从2009年开始使用EMC Isilon的产品。EMC Isilon很适合华大基因这种数据计算需求,为华大基因在IT方面的发展提供了重要支撑。我们典型的应用模式是10套EMC Isilon的存储节点组成一个应用存储,性能能够达到2.5GB/s,这是相当出色的性能水准,能够满足我们的存储需求。

当一个计算任务被加载到集群系统时,各个计算节点首先从I/O节点获取数据,然后进行计算,最后再将计算结果写入I/O节点。在这个过程中,计算的开始阶段和结束阶段I/O节点的负载非常大。图2是文件系统吞吐量界面。从该界面中可以看到每个节点上统计出来的吞吐量性能数值。这里显示的统计结果反映了不同客户端上的应用对吞吐量的需求不同,最高的有2Gb/s左右,总吞吐量达到9Gb/s。传统的文件系统使用集中服务器来管理多个存储资源,导致存储系统单点故障,扩容到一定程度会产生性能瓶颈以及热点应用的性能问题。使用OneFS操作系统软件,EMC Isilon集群存储系统内的每个节点享有同等地位,因此任何节点都能处理请求,实现跨节点负载均衡,保持高性能和应用性能的稳定性。OneFS使用Infiniband®进行集群内通信和同步,使用拥有专利的B-tree技术,将所有数据(包括元数据)完全并行的分布到所有节点中,不需要专门的元数据服务器,从而杜绝了其他存储架构元数据服务器的性能瓶颈问题。

为了充分满足各种工作流的需求,EMC Isilon集群将文件分配到多个节点和磁盘中,优化I/O并行处理能力。在写操作过程中,OneFS跨所有节点提供了全局可访问的一致缓存。每个存储节点都包含标准DRAM(介于6GB至96GB之间),此内存主要用于存放经常被访问的缓存数据,以减少磁盘带来的延迟和处理资源消耗;在读取过程中,会先预取内存中数据以降低I/O压力。由于每个文件都被划分到集群中的不同节点和磁盘集合中,随机存取一系列文件时产生的总负载被分配到集群中的所有磁盘上,并且不会像基于RAID的系统一样受到有限的磁盘集合限制。由于所有节点都参与I/O处理,相比传统的单一磁头或RAID,可以支持使用更多的CPU和内存处理单元。

#### 图 2. 文件系统吞吐量



#### 高并发处理能力

高性能计算系统中的存储系统常常会是制约系统整体应用性能的关键因素。虽然存储系统容量提升和单位成本下降都很快,但传统存储技术架构对带宽和I/O能力的提高却非常困难且成本高昂。这造成了当原始数据量较大时,I/O



读写所占的整体时间就相当可观,成为高性能集群系统的性能瓶颈。华大基因研究院每天有数千个应用程序访问 EMC Isilon存储,EMC Isilon通过高速并行存取的方式来满足需求。

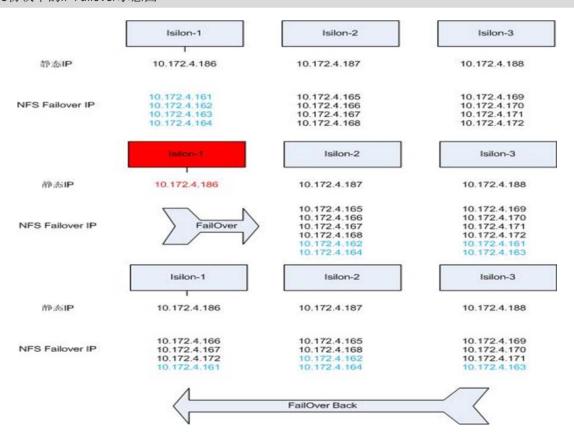
EMC Isilon在维持多用户并发性能方面有以下几处优势:

首先,EMC Isilon将文件数据和元数据分散保存在横向扩展的存储系统中,且系统集群中的多个节点具有同等地位,可以共同处理数据和元数据操作的I/O。在这种分布式架构中,I/O处理被均衡分布到集群中的各个节点上,任何单个节点都不会成为瓶颈或单一故障点。

其次,EMC Isilon还提供用户配额管理的功能一SmartQuotas™,通过此模块可以加入客户环境中现有的AD、NIS、LDAP等用户控制域,提供基于用户的访问空间限额,避免一个用户占用资源太多,影响其它并发用户的性能。此外还提供thin provisioning的功能,实现跨用户空间的共享。

第三,EMC Isilon专有的SmartConnect功能可以优化集群存储的性能和可用性。通过使用单一的主机名,SmartConnect支持客户端连接负载平衡和客户端连接在存储节点上的动态NFS故障切换和故障恢复,提供了最佳的集群资源利用率。无需安装客户端驱动程序,管理员就能轻松管理不断增加的众多客户端,确保即使一个系统发生故障,所有的读写操作也能成功完成。SmartConnect IQ提供了前端网络的负载均衡模式,包括Round Robin、CPU忙闲、网络吞吐量、网络连接数等四种均衡策略。此外,SmartConnect还提供了基于NFS协议的IP failover方式,提供了对TCP连接的容错。图3是NFS协议下的IP Failover示意图。当Isilon-1节点宕机时,该节点上的连接会自动被SmartConnect重新分配到集群中正常工作的节点上;当Isilon-1节点恢复工作后,SmartConnect再一次对集群中的所有连接进行重新分配,保证每个连接提供不中断的且理想的存储性能。

#### 图 3. NFS协议下的IP Failover示意图



下面来看下华大基因HPC环境下EMC Isilon SmartConnect功能工作情况。图4是集群中IP池信息。我们希望集群中的节点资源都能得到有效的利用,在这个实现过程中,EMC Isilon SmartConnect承担了这个任务,而且无须我们人工来优化节点连接。采用传统存储,当Client端发起连接时,会出现过多的用户被分配到一个节点上,该节点出现性



能瓶颈,而集群中其它的节点处于"空闲"状态,从而使得整个计算系统无法快速高效地完成用户作业任务。EMC Isilon SmartConnect通过4种策略设置使管理软件能够按照设定好的策略来设置新的Client连接,确保所有节点资源都得到有效利用,且无需人工介入。例如,我们在这里设定按照每个节点吞吐量的大小来分配连接,将新的连接分配到那些吞吐量不高的节点上,由此达到整个计算集群的平衡。从下图中可以看到,我们计算集群中的10个节点上都有一定数量的Client连接,并且尽量保证这些连接均衡地分配到各节点上。每个节点后面显示的IP地址即是当前的Client连接数量。

#### 图 4. 集群网络

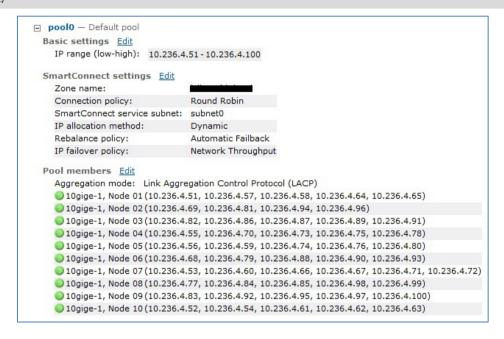


图5是Client Connection信息界面。从左边的列表中我们可以得知,当前该集群内有638个客户端连接。在右边的柱状图上,我们将鼠标放置在节点2(Node2)上,显示出共有80个客户连接。EMC Isilon的SmartConnect将该集群内活跃的客户端连接尽可能地平均分配到每个节点上,实现用户的负载均衡和动态迁移,保证资源的有效利用和存储故障时所有用户不掉线。

#### 图 5. Client Connection信息





#### 基于hadoop架构云计算平台下的存储性能

围绕生物云落地,并为了高效分析和妥善存储基因组数据,我们的弹性计算开发团队利用Apache™ Hadoop™ MapReduce框架,开发了SOAPhecate和SOAPgaea这两款基于分布式集群的新一代测序数据分析软件。如今DNA测序技术成本的下降幅度远远高于测序数据分析成本,云计算已成为基因组学研究领域中的一个日趋重要的工具或服务。

我们整个生物云计算核心分三大部分:一是基因的数据,把自己数据及从NCBI和EBI得到的国外数据提供给国内用户使用;二个是生物专有服务,比如以卖服务的方式测试,还有是生物应用软件接入;三是提供SDK,用户用自己的程序来调用数据、服务和软件。我们目前已经建立了新一代测序、生物信息分析、转基因、动物克隆、免疫标记、动植物基因组及分子育种、疾病研究、蛋白质组学、微生物、健康等技术与平台服务。基于云计算的软件服务给研究者提供了分析数据的平台,大大提升了运行速度并降低了成本,他们不再需要自己购买计算机集群即可完成海量数据的分析处理。现在,我们的客户遍布科研院所企业,国外的研究机构甚至政府机构有几万人同时要求做基因测序,国外客户的业务已经超过华大基因一半收入。华大还会在这一块继续投入大量的研发资源,提供更好的云计算平台。

我们的云计算平台的后端存储平台也是EMC Isilon,因为EMC Isilon是当时一款可与HDFS(HadoopFS)层本机集成的企业级横向扩展NAS存储平台。而且在选择云平台存储时,我们主要关注以下几点:

- 高可用性: EMC Isilon的OneFS操作系统从根本上消除了HDFS存储层中NameNode的单点故障,因为元数据分布在整个Isilon集群中,所以每个Isilon节点都相当于一个NameNode,这就扩大了利用Hadoop解决大数据分析的应用领域。
- 改进数据访问和加载: OneFS支持HDFS以及行业标准协议(如NFS和CIFS等),可以减少将数据移至Hadoop的需求,并很好地执行读写操作,以及将Hadoop操作附加到存储中。这改进了传统Apache Hadoop部署的HDFS繁复的数据转移和加载流程,消除了对过量文件拷贝的需求。Hadoop应用程序可以通过HDFS直接访问其他协议(如NFS等)加载到OneFS存储层的数据,而不必将原始数据拷贝或移动到其他文件系统。OneFS可消除HDFS中3倍的数据副本,实现80%的利用率,从而提高存储效率。
- 扩展和效率:与传统Hadoop的DAS部署相比,EMC的分布式Hadoop存储和计算解决方案可实现更好的存储和 CPU利用率,因为EMC Isilon可根据需求单独扩展性能或容量,并可以创建单独的存储池来满足单个系统中的不 同性能需求。相比于传统DAS模式中数据全部留在高昂的主存储上,SmartPools基于策略的自动分层还能够根据不同数据的价值和性能需求匹配相应的存储节点,在降低了成本的前提下进一步提升了存储效率。
- Isilon NAS存储解决方案提供高效的单个文件系统/单个卷,可扩展至20PB。
- Isilon SnapShotIQ和SyncIQ功能可以提供备份和灾难恢复以便将企业数据服务融入Hadoop。

下面看一下EMC Isilon在我们中心的使用情况。

图6是集群状态信息界面。从图中可以看到,该集群设置了10个节点,每个节点上都有活跃的连接。每个节点上的连接数都超过了19(通过Client Connection summary得知),并且每个节点的利用率保持一致,都达到93%。每个节点都得到了很好的利用,没有出现单个节点遭遇性能压力,而其它的节点资源处于空闲的状态。从界面左边可以看到每个节点上当前数据吞吐量的数值。作为对我们云平台的存储支持,我们将该集群开放给其他科研单位人员使用。从下面的统计图可以看到该集群当前吞吐量达到了6.81Gb/s,10个节点自动负载均衡支持大约330个客户端的应用。



#### 图 6. 集群状态信息

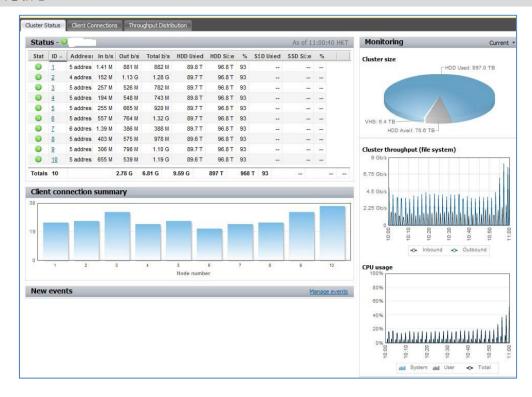
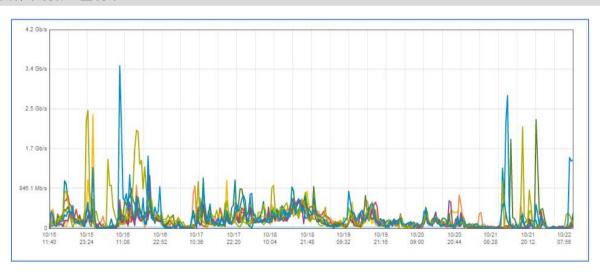


图7是文件系统吞吐量统计界面。通过该界面可以查看当前或历史时间下每个节点的总吞吐量走势。由图中可以看到,尽管不同时段下应用的性能增长随着压力的增加而提升,但仍然动态地满足了客户端对存储性能的需求。

#### 图7. 文件系统吞吐量统计



#### 高可扩展性

随着我们业务的飞速发展,数据量还在不断增长。2011年我们深圳中心的数据量达到了8PB,2012年将达到10P。相应的,EMC Isilon的节点也在增加,平均每年添加30-40个节点。添加节点非常简单,系统自动负载均衡能力强大,保证升级扩展不会对我们的业务造成影响。我们每套新的EMC Isilon存储,都是分配单独的应用空间,不会和其他的EMC Isilon的应用系统混合在一起。即使混在一起,升级过程也是非常平滑的,几十秒的时间就可完成。例如,我们最初购买的10个节点集群扩容成17个节点,就是由我们自己完成的。 EMC Isilon是由模块化节点构成,有两种节点: 一种是存储平台节点,一种是性能加速节点。系统的核心是存储平台节点,它由智能分布式文件系统提供支

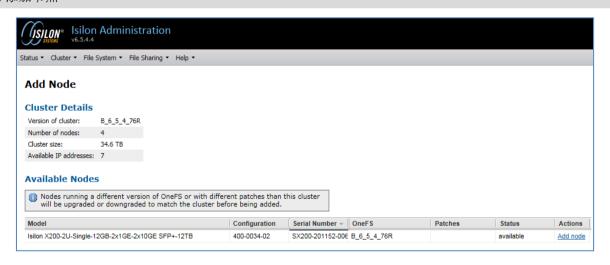


持。该文件系统将所有资源集中在一起,以处理几乎所有类型的应用程序工作负载的存储需要。如果工作流受到磁盘I/O限制,可以在系统运行时添加更多平台节点。同样,如果工作负载仅受CPU和内存限制,可以在系统运行时添加性能加速器节点。根据需求添加不同数量的存储平台节点或性能加速节点,使得EMC Isilon可以同时扩展两项性能,或独立扩展一项,以适应变化的应用程序环境。

相对于传统的NAS将所有数据从控制器出入造成性能和扩展性瓶颈不同,EMC Isilon支持大规模信息的集中访问,实现数据的共享,每台服务器都可以看到所有的存储空间。客户端能够随时按需安全地访问数据,单个卷可实现PB级文件系统支持,而且更易于安装、管理和扩展。

图8是添加节点的界面截屏。通过此界面,我们可以快速简便地添加新节点。由于EMC Isilon的OneFS(单一文件系统)操作系统实现了统一命名空间,并创建了一个共享存储池来集中管理集群中的存储资源。所以我们向集群中加入新的节点时,它能智能地将数据存放到集群里。此外,EMC Isilon存储节点采用infiniband高速互联,存储节点扩展时,性能可随着节点增加而线性扩展。只需要鼠标点击图8中Add node,不到60秒的时间就可以完成在线升级,且无需对原有设备及应用做任何改动,系统扩展的整个过程无需宕机。新加入的节点空间即刻就能在客户端看到并且立即可以投入使用。

#### 图8. 添加节点



EMC Isilon的SmartPools功能,还能够进一步扩展其横向扩展存储平台。图9是中桥分析师在POC过程中SmartPools的界面截屏,可以看到有S200和X200两个型号的存储节点组成一个集群。数据可以通过使用策略,决定何时存放在相应的节点上,达到利用率最大化的存储管理和使用效果。借助SmartPools,多个EMC Isilon IQ存储节点层可以共存于单个文件系统中,提供单一管理点。同时,策略驱动的SmartPools还可以定义工作流中数据的价值,自动将数据与适当性价比的存储层相匹配。存储管理员可以调整工作集到适当大小并加速工作集,然后以透明方式自动将不活跃的数据移动到性价比更高的存储设备内。此外,存储管理员还能够使用SmartPools来聚合和整合应用程序,实现工作流隔离、更高的利用率和独立的可扩展性。

SmartPools跨EMC Isilon横向扩展存储后端网络Infiniband移动数据,这消除了客户端网络上的流量并确保绝对无缝的数据移动,包括运行中读/写活动、锁定语义、备份应用程序交互和基础文件识别。SmartPools利用自动化策略、手动控制或API接口提供文件级粒度和控制,使IT管理员可以控制性能和布局、存储层协调以及文件级和目录级的保护设置并且对最终用户完全透明。



#### 图 9. SmartPools 界面

mart	Pools > Summa	ary							
	k Pools								
	Name	Resources In Pool	HDD Used	HDD Size	HDD % Used	SSD Used	SSD Size	SSD % Used	Default Protection
	James - Transport	Resources In Pool Nodes 2-4	HDD Used	HDD Size	HDD % Used 0.8%	SSD Used	SSD Size		Default Protection +2:1

#### 安全性

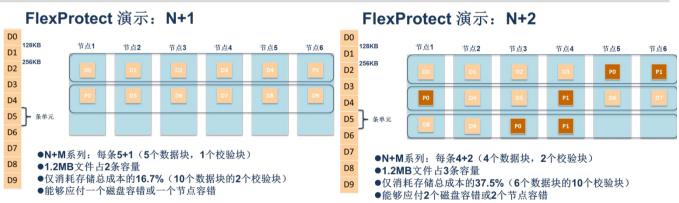
之前,华大曾比较了EMC Isilon和另一款产品,之所以选择EMC Isilon,就是因为其管理性、易用性和安全性很出色。 IsioIn可以达到N+2的安全级别,对于华大来说,这样的安全级别才能满足华大高性能计算环境的需求。

传统存储系统扩展时,在规模尚小时适用的技术在规模较大时变得不再适用,RAID就是最好的例子。只有在发生另一起故障前重新构建数据时,RAID才有效。但是,数据量增加时访问该数据的速度并未提高,发生其他故障的可能性仍会增加。EMC Isilon能够提供两个级别的容错: 1、节点级容错; 2、文件级容错,其FlexProtect功能无需基于硬件的RAID技术即可提供数据保护。

下图10以6节点为例。当一个1.2MB大小的文件写到EMC Isilon集群存储,文件被OneFS分成10个128KB的数据块,顺序存入到6个节点中,占用1条容量。如果采用了N+1的数据保护级别,每条5+1(5个数据块,1个校验块),1.2MB文件占2条容量,校验块容量占到占到16.7%,当某一个磁盘或者一个节点发生故障宕掉,FlexProtect能够判断并确定文件的哪些部分受到故障影响,并让多个节点参与受影响文件的重建。由于OneFS中的Autobalance功能将文件散布在整个群集中,可用来重建文件的磁盘轴和CPU数量远超过典型硬件RAID实施中可用的数量,数据仍然是安全完整的。如果N+1还不能满足要求,EMC Isilon可以最多支持4个节点,而且数据保护级别可以在线更换,数据仍然是完整的。这个容错机制跟业界其他产品容错机制完全不一样,这也是EMC Isilon集群存储非常有特点的一个产品特性。

对于N+1和N+2的数据保护类型,参见如下存放方式:

# 图 10. Isilon FlexProtect数据保护原理示意图(N+M)





#### 图 11. Isilon FlexProtect 数据保护原理示意图(N+M:B)

#### FlexProtect 演示: N+2:1



以将最重要的数据配置成N+4保护,将一般重要的数据配置成N+2保护。目前为了增加数据的使用效率,EMC Isilon集群存储又推出了N+M: B的数据保护模式。使得存储的使用效率和数据安全性提高了一倍以上。接下来看第二级容错即文件级的容错,EMC Isilon集群存储可以针对某一个目录或者某一个文件做多个链像系价。每一个链像都会分布在不同的节点上

EMC Isilon还支持分级保护。在一个文件系统中,可

群存储可以针对某一个目录或者某一个文件做多个镜像备份,每一个镜像都会分布在不同的节点上,为数据提供足够的数据保密度。即使节点或磁盘失效了,我们的数据仍然是安全的。此外,FlexProtect还可按照数据或者工作流的重要性进行安全级别划

#### 分,采取不同程度的保护策略。

图12是磁盘池设置信息界面。在该界面,我们看到当前10个节点的集群中磁盘池的数据保护策略为+2:1。这意味着我们的数据保护级别做到了N+1,即允许在损坏两块盘或者是1个节点的情况下,提供完好的数据保护。

#### 图12. 磁盘池设置信息

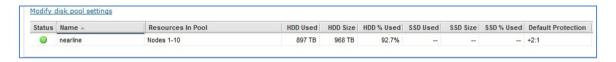
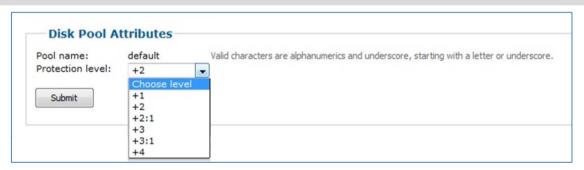


图13是磁盘池属性界面。通过该界面,我们可以对不同重要程度的磁盘设定不同等级的数据保护策略,只需要选中要设定保护方案的磁盘,然后选择数据保护策略即可。

#### 图 13. 磁盘池属性



#### 易管理性

在选择高性能存储系统时,还需要考虑的一点就是,要确保集群中的节点都能正常工作,保证管理人员在任何时刻 都能获取集群甚至是整个系统的运行状态。一旦出现错误及时定位出错点,并采取相应的措施。此外,系统的升级、 备份和打补丁等还要求操作上的简便性和时间上的快捷性。

对于EMC Isilon系统,我们认为它是非常可靠的,并且安装部署十分方便,运维工作也很简单轻松。平时只要看看控制面板有没有告警之类的信息即可。如果有告警,系统同时也会发邮件通知管理员。

EMC Isilon集群存储易管理性体现在以下方面:

#### 1. 安装和扩展简单



传统的SAN和NAS存储在配置过程中,需要提前做好规划工作,包括硬件上架、连线、存储网络配置、RAID group、LUN、卷的划分、创建文件系统等。任何一个环节出现问题,都将影响后续环节。而EMC Isilon集群存储在客户端看来没有任何RAID、LUN和卷的概念,因此,安装中就省除了那些配置操作,只需对硬件进行连线,配置好存储网络即可,存储系统就能交付使用。

而且在扩展中,借助OneFS的单一文件系统功能,管理员可以在不影响正常使用情况下,动态地添加节点到原有系统中,耗时仅为60秒,通过自动负载均衡,保证扩展不影响业务。

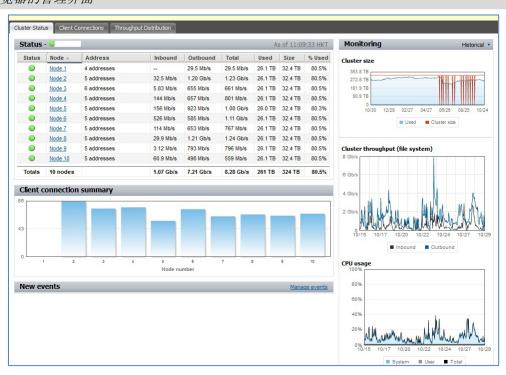
#### 2. 管理和配置简单

EMC Isilon集群存储的管理和配置很简单,存储容量的增加不会给存储管理带来任何压力。EMC Isilon集群存储的两种管理界面,一个是基于浏览器的管理界面,一个是全命令行的管理界面,所具备的功能完全一致。

看一下EMC Isilon在我们HPC环境内的工作情况。

图14是集群状态信息界面。该界面提供的信息非常丰富,包括集群中各个节点的状态信息,以及整个集群的状态监控,这便于我们及时掌控集群的整体运行状况。在这里,我们可以看到当前这个集群中搭建了10个节点,单个节点的裸容量为36T,换算后的容量为32.4T,整个集群的总容量为324T。每个节点当前均处于正常的运行状态。此外,还可以看到每个节点上Inbound和Outbound的数值变动详情,并且每个节点的容量使用率基本保持一致,都达到了80%以上。整个集群中每个节点都得到有效的利用,计算压力尽量均摊到每个节点上。Client connections summary呈现的是每个节点上的用户连接数量,表示当前有多少用户正在访问该节点的资源。由统计结果可以看到,除了节点1在目前的运行状态下没有活跃的连接,其它的节点上面都有一定数量的活跃连接。从显示的柱状图来看,用户连接尽量均摊到了每个节点上。在该界面的右侧则是对集群参数的监控,包括集群容量、集群吞吐量(文件系统)和CPU利用率。通过这些有规律的按时统计,我们可以随时掌握集群的容量、吞吐量和CPU利用率等状态。

#### 图14. 基于浏览器的管理界面



此外,除了通过图形化的界面来了解整个集群的运行状态之外,EMC Isilon还提供了命令行的查询方式(图15)。图形化界面的查询方式具有较强的直观性,对于新培训上岗的管理员来说图形化的界面减轻了管理员的负担,让他



们更加容易掌握整个系统,而命令行的形式也是一些经验丰富的管理员所喜爱的方式,他们对于命令行形式的管理 更加熟练。两种方式让我们的管理员能自由选择自己熟悉的管理方式。

#### 图 15. 全命令行的管理界面

	# isi statistic	s pstat	one Don Forond		
null	0.00/s		ons Per Second 1718.83/s	setattr	587.64/s
lookup	67.20/s		89.80/s	readlink	1.01/s
read	7275.10/s	write	1445.42/s		16.59/s
mkdir	0.00/s	symlink	0.00/s	mknod	0.00/s
remove	6.61/s	rmdir	0.00/s	rename	0.00/s
link	0.00/s	readdir	0.00/s	readdirplus	8.37/s
statfs	0.20/s	fsinfo	0.00/s	pathconf	0.00/s
commit	52.52/s	noop	0.00/s		
TOTAL	11269.28/s				
CPU Util	ization			one	FS Stats
user	0.2%			In	203.43 MB/s
system	51.4%			Out	1095.93 MB/s
idle	48.6%			Total	1299.36 MB/s
Network	Input	Networ	k Output		oisk I/0
	297.26	MB/s	943.45		45722.80 iops
Pkt/s	536312.20	Pkt/s	782152.80	Read	953.77 MB/s
Errors/s	0.00	Errors/s	0.00	Write	348.12 MB/s

#### 中桥观点

华大基因研究院通过EMC Isilon存储OneFS系统避免了单点故障、扩展性能瓶颈及热点问题;通过InfiniBand内通信技术降低了延迟;通过全局缓存提高了I/O响应;通过SmartQuotas访问空间限额、SmartConnect负载均衡和故障切换等功能提高了并发性能。在评测过程中,数千个应用程序在访问Isilon存储时,其高并发访问条件下的性能表现优异。

基于Hadoop架构的云计算平台得到了EMC Isilon存储很好的支持,并解决了传统Hadoop存储的一些固有问题。包括NameNode单点故障、缺乏企业级备份容灾功能等。通过OneFS改进了数据访问和加载流程,提高了存储效率。

EMC Isilon存储能在不停机或者不改变应用的情况下,在60秒的时间内线性或独立升级存储容量和性能。新加入到集群中的节点会自动继承现有的策略和设置,AutoBalance功能可以自动将数据分放到新扩展的集群里,因此克服了传统的存储系统升级容量时需要停机的弊端。

EMC Isilon存储使用FlexProtect-AP功能,能在集群、目录或文件层级动态设定数据保护策略,可以根据数据价值制定保护策略,最高提供N+4的数据保护,支持集群中4个节点发生故障的数据重建。

EMC Isilon存储利用一个基于Web的中心控制台提供对集群性能、容量利用、配额、监控、检测的实时单层管理,以及使用SynclQ软件管理复制作业。使用Web界面,只需点击鼠标,就能在集群中增加或者删除节点。



## 结论

对于生命科学和基因研究而言,并行处理能力、数据生命周期安全管理以及根据业务发展逐步升级扩展能力,决定 着业务的核心竞争力。

传统存储机构很难满足并行处理对IO性能的要求。不同分析技术对存储IO性能、吞吐量需求不同。如何为业务发展的不同阶段灵活提供存储IOPs、MBPs和容量需求,是高性能计算,尤其像华大基因这样大数据、高性能用户的最大挑战。此外,高性能环境对低延迟的要求,也在不断挑战着技术极限。

为了满足不同阶段对存储性能和容量的需求,用户常常需要持有相当水平的闲置存储资源,以保证业务增长扩容需求。这不仅导致存储资源利用率低下,也带来了存储管理、耗能的浪费。此外,传统存储随着容量扩展,不仅复杂也会影响业务正常运营,且消耗大量IT管理资源,阻碍了通过IT实现创新的能力。

横向扩展NAS系统易于安装和部署、成本合理、性能可靠,适合于大数据分析环境。ESG通常将大数据定义为使用超出正常处理能力范围和规模的数据集,需要用户采用非传统方法。从外部网络存储角度来看,横向扩展NAS正是这种"非传统"方法,它可以超出传统横向扩展系统的限制进行扩展,同时还可以随着数据的增长保持性能和可用性。

基于Hadoop架构的云计算平台与EMC Isilon横向扩展NAS之间的共性有利于聚合到一个解决方案中,确保华大基因能够利用企业级数据保护功能以及应用于Hadoop分析处理的共享存储模式;实现更高的存储和计算资源利用率;避免了NameNode单点故障;改进了数据访问和加载流程,消除HDFS中3倍的数据副本,实现80%的利用率,从而提高存储效率提高了存储效率。

应用EMC Isilon的OneFS操作系统软件,华大基因的研究人员可将每天研究产生的大量基因组数据统一存储到一个共享存储池中,不仅易于管理,还方便对这些数据进行随后的分析处理。避免了传统存储不得不跨越多个独立的信息孤岛来扩展数据,而造成的不必要的数据访问和分析障碍。

生命科学和基因研究都要产生大量的数据,并且这些数据还需要永远的保存起来。Isilon横向扩展NAS存储解决方案提供高效的单个文件系统/单个卷,可扩展至20PB。通过将存储整合成一个单一,可扩展卷,让任何人都可以来管理,Isilon存储增加了数据管理的简易性,减少了人员培训的时间,简化了大数据生命周期管理。

此外,计算节点(如果需要IOPs)和容量节点可以独立升级、线性扩展,通过这种灵活的模块化升级扩展方式,给用户提供灵活技术。同时,在不影响业务运营条件下60秒内即可轻松完成扩容(华大基因独立部署最新采购的10个节点),让用户可以最大限度地降低闲置资源的持有水平,提高存储资源利用率。

对于像华大基因这样的研究机构,数据是核心资产。如何保证输入(原始)数据的万无一失,如何确保存储系统可靠稳定,以及如何保证数据生命周期内的安全,是很多高性能用户最为关注的。EMC Isilon N+M架构让用户根据应用和数据生命周期价值的不同,选择存储安全配置策略,灵活满足不同业务、不同应用和不同数据对安全性的要求。Isilon的FlexProtect技术保证了数据的高可用和业务连续性。

目前,华大基因有8PB数据通过Isilon节点进行处理和保护,支持2000个计算节点,总计2万个计算核心。通过我们现场评测,结合与华大基因CIO的采访,华大基因对Isilon的并行处理支撑能力、性能、灵活快速的扩展性、管理的简易性都非常满意。Isilon让华大实现基因云计算平台的部署和实施,并将更多IT资源用于创新实现。

