

下一代互联网上的科学数据传输

文 / 吴桦 龚俭 周渔

摘要:

本文介绍了在CNGI上进行大型科学试验数据传输的AMS项目应用。描述了AMS数据传输系统的总体结构和需求,分析了东南大学数据处理中心的功能需求,并给出了目前的传输环境和传输测试结果。

关键字:

AMS 数据传输 CNGI IPv6

来自国际空间站的实验

AMS⁰¹(Alpha Magnetic Spectrometer 阿尔法磁谱仪)是由著名物理学家、诺贝尔物理学奖获得者丁肇中教授领导的一个大型空间物理实验,也是国际空间站上唯一的大型物理实验仪器。该实验的目的是研制一个能精密测量高能量带电原子核和粒子以及高能量光子的磁谱仪(阿尔法磁谱仪),放置在国际空间站上做为期3-5年的实验,以探测来自于宇宙深处的反物质,而寻找反物质、探索暗物质和测量宇宙线的来源是国际高能物理和天体物理学界极为关注的课题。因为这个实验的重要性,美国宇航局特别于1998年6月用“发现号”航天飞机将AMS(AMS-01)带入太空运行10天,获得了He核、质子和正反电子的空间和高能量分布规律,得到了新的结果,并发表在Phys. Lett.B等国际著名刊物上。AMS-02计划也得到了美国宇航局的优先支持,并继续允许用航天飞机将AMS02送到国际空间站进行实验。

由于AMS具有重大的科学意义,中国(包括台湾地区)、美国、瑞士、法国、德国、意大利、芬兰等15个国家和地区的诸多航天企业、科研院所、大学参与了AMS的研制工作,其中包括美国NASA、欧洲各国的宇航局及美国麻省理工学院、耶鲁大学、德国亚琛大学、俄罗斯莫斯科大学、瑞士日内瓦大学、苏黎世高工、意大利罗马大学等多所大学。东南大学(SEU)是中国大陆第一个被丁肇中教授邀请参加AMS研制的大学,其中的一项合作任务就是在东南大学建立一个地区地面数据中心(Regional Center, RC)。

在太空中由AMS-02采集到的大量观测数据将通过NASA的卫星网和地面的互联网传输到各个数据处理中心进行整理并提供给物理学家做后续的分析处理。AMS试验要求持续时间长的、稳定的网络传输支持,不能简单地使用常规的传输协议来支持,需要专门的构造,这是CNGI上的一个典型应用,因此我们将原来基于IPv4技术的AMS地面数据传输系统移植到了IPv6平台,并基于CERNET2在东南大学(南京)和日内瓦(CERN)之间进行了一系列Native IPv6的传输可行性测试实验,结果表明了方案的可行性。同时意大利国家核物理研究所(米兰)也基于欧洲教育网(GEANT)进行了一系列通过IPv6隧道和双栈网络的传输实验,获得了与我们的实验相似的结论。



AMS 数据传输的需要

图1给出了AMS实验数据传输系统的基本结构^{[1][3]}。安装在国际空间站ISS上的AMS-02磁谱仪不断地将检测到的宇宙带电粒子数据发送给位于NASA的Marshall宇航飞行中心(MSFC)的载体操作中心(POIC);地面支持计算机(GSC)接受从POIC传来的传输控制和科学数据,并将数据转发到位于瑞士日内瓦欧洲粒子物理研究中心(CERN)的负载控制中心(POCC)和数据操作中心(SOC)。POCC和NASA的飞行操作中心合作,在CERN接受实时的监测状态数据,发出控制命令;SOC接受全部的观测和分析数据,并分发给其他的地区地面数据中心。目前已经有包括美国麻省理工学院、耶鲁大学、英国剑桥大学、法国国家核物理研究院、德国卡尔斯路厄大学、意大利国家核物理研究所、中国东南大学等单位承担了地区地面数据中心的工作,其中位于东南大学的RC是亚洲唯一的地区地面数据中心。

从覆盖的网络范围看,AMS-02的观测数据将通过美国NASA的Ku波段卫星专用信道从国际空间站传送到POIC的存储系统,然后通过ESnet和美国的Abilene传送到位于CERN的SOC,然后再通过美国的Abilene、欧洲的GEANT和中国的

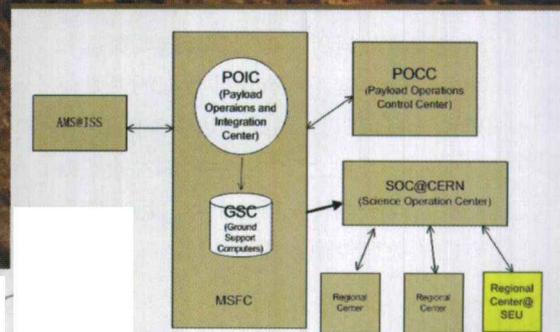
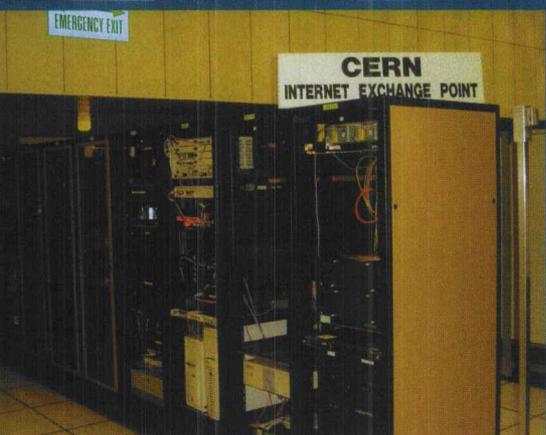
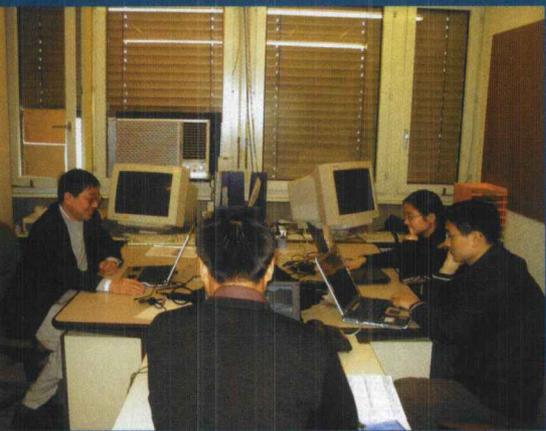


图1 AMS02 数据传输系统结构



更快、更稳、适应性更强的数据传输系统

· 高效率—满足需要高于 25M/s 以上的传输要求,而且必须根据 CERNET2 的使用管理政策将占用的带宽控制在允许的范围之内,不过度使用资源;

· 可靠性—要求能够不间断地运行 3—4 年的时间,对发生传输故障可以自行恢复,不发生数据丢失现象,保证接收数据的完整性,这要求系统能够有效地管理传输缓冲空间;

· 适应性—数据传输系统的配置在数据传输过程中应当能够根据网络的状态进行手工或自动的调整。

CERNET2 传送到世界各地的 RC。AMS-02 控制数据则通过 ESnet 由位于 MSFC 的 GSC 传送到 CERN 的 POCC。

AMS-02 需要传输的数据有多种类型,在设备升空之前是各个地区中心产生的蒙特卡洛仿真数据,设备升空之后采集到的原始数据,以及为访问和使用这些观测数据所派生的事件摘要数据和事件标签数据。所有这些数据的总量每年预计有 105TB,最终将达到 500TB。在 AMS-02 升空之后,这些观测实验数据需要不间断地转发至各个 RC,因此需要专门的数据传输系统的支持。

选择 IPv6

RC 最主要的功能是存储数据,如果提供所有数据的拷贝,数据量预计将达到 500TB 左右,是个超大规模数据存储系统。但是不同类型的数据的重要性是不一样的,这样导致可以采用不同的策略存储原始数据、摘要数据和蒙特卡洛数据。系统要具备合理的调度功能,使数据在高速硬盘存储系统和低速磁带存储系统之间的分配与用户访问时间之间有良好的平衡。此外系统需要具有目录,以便对数据的有效性和可用性进行管理 (BOOKKEEPING)。

东南大学 RC 直接从位于 CERN 的 SOC 获取 AMS 实验数据,基本带宽要求达到 25MB/s。AMS-02 实验过程中各课题组成员之间协同工作也需要一定的带宽(例如视频会议),但与数据传输相比,这个问题几乎可以忽略。由于目前基于 IPv4 的 CERNET 一直处于流量较为饱和的状态,以它作为 AMS-02 传输平台并不合适,而电信运营商的商业互联网更不可能为这个科学研究项目单独做任何带宽保障工作,因此使用基于 IPv6 技术的 CERNET2 是一个自然而合理的选择。随着 CERNET2 主干网(东南大学为其中的南京节点)在 2004 年 12 月率先开通,CERN 也与欧洲和美国的多个 IPv6 试验网实现了连接,从而为将 AMS-02 地面数据传输系统向 IPv6 网络移植创造了条件。CERNET2 设计的国际出口带宽超过 1G,可以满足 AMS02 的传输需求。

数据传输需求的满足不仅依赖于物理信道的带宽保证,还需要通过软件保证传输的稳定性。由于 AMS 数据传输是数年不能间断的特殊应用,需要通过预约带宽的方式保证传输服务的稳定性。但是在实际环境中,可用带宽会受其它流量的影响,被传输的数据可能由于信

道的故障而被积压在发送端,因此需要对传输 AMS 数据的互联网信道的状态进行监测,并根据实际情况不断调整参数,加快或放缓发送速率,保证传输任务的正常进行。另外需要开发相应的管理系统来监控系统的平稳运行。

图 2 给出了东南大学 RC 的总体结构。

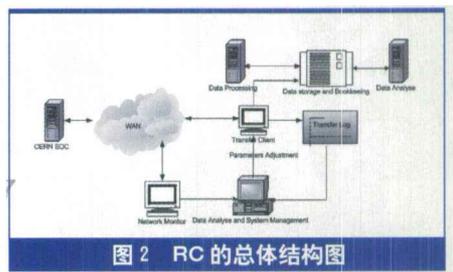


图 2 RC 的总体结构图

成功的可行性测试实验

CERNET2 主干网的开通为数据在 IPv6 上的传输创造了条件,我们需要一个良好的传输环境,CERNET2 毫无疑问满足了这种需求。我们采取了合理的方案对整个传输环境进行了测试。

我们使用如图 3 所示的 Native IPv6 连接方式进行数据传输实验,避免了一般情况下隧道造成的不良影响。

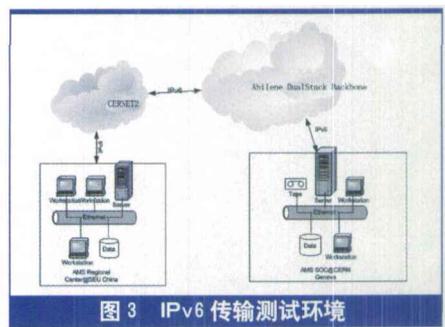


图 3 IPv6 传输测试环境

首先进行的是线路可用性分析。使用 traceroute6,我们可以发现从 CERN 到 SEU 和从 SEU 到 CERN 经过的网络连接是对称的,路由是稳定的,这避免了由于不确定路由给传输带来的影响。通过 ping6,我们测试得到的平均往返时间 (RTT),从 CERN 到 SEU 平均是

基于 IPv6 技术的
CERNET2 是 AMS-02
地面数据传输一个
自然而合理的选择

412ms, 从SEU到CERN平均是420ms。BDP (Bandwidth Delay Product) [4] 给出了为了最优使用可用带宽必须配置的最小发送窗口数目, 是提高数据传输速度的关键参数。我们使用 Iperf 对这两点之间的 IPv6 信道进行测量, 将测得的可用带宽与 RTT 相乘, 得到了如图 4 所示的 BDP 测试结果。

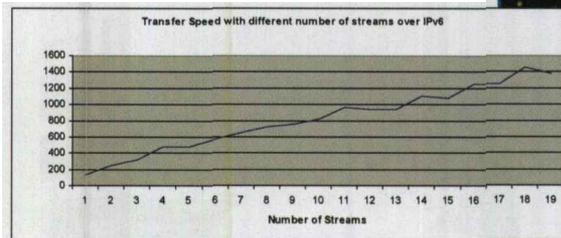


图5 并发流数目对传输速度

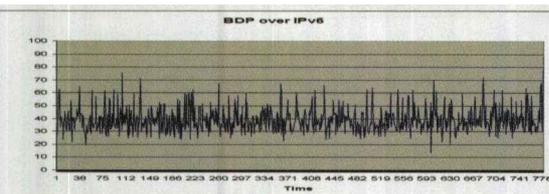


图4 南京与日内瓦之间IPv6 通路上的BDP

从图上可以看出来, BDP 的值基本上在 40Kbits 左右, 随着时间的变化, 与流量的周期规律没有很明显的差别, 这反映了目前默认的 64K 窗口长度是符合实际情况的。

随后在此环境下进行了文件的传输试验。AMS 实验一直使用的传输软件是 BBFTP [5], 它是 IPv4 下的高性能的文件传输软件, 支持并发传输和 TCP 连接窗口的可调节, 由法国里昂 IN2P3 计算机中心的 Gilles Farrache 开发, 特别对大文件的传输比普通 FTP 具有更高的效率。AMS 的试验数据每个文件大小都在百兆以上, 选择这个工具充分利用了一些传输参数的可调性。为了能够在 IPv6 协议的网络上传输, 我们将原有基于 IPv4 的 BBFTP 移植到了 IPv6 环境。图 5 为在目前的情况下传输 10M 文件的传输速度, 试验使用的发送和接收窗口是 64K, 并发流的个数是 1 到 20。从图中可以看出, 随着并发流的增加, 传输速度几乎线性增加, 但是当数目达到 18 以后, 由于系统资源的限制, 变化趋势有所改变。这证明可以利用增加并发流个数的方法提高传输速度。

试验对网络的新挑战

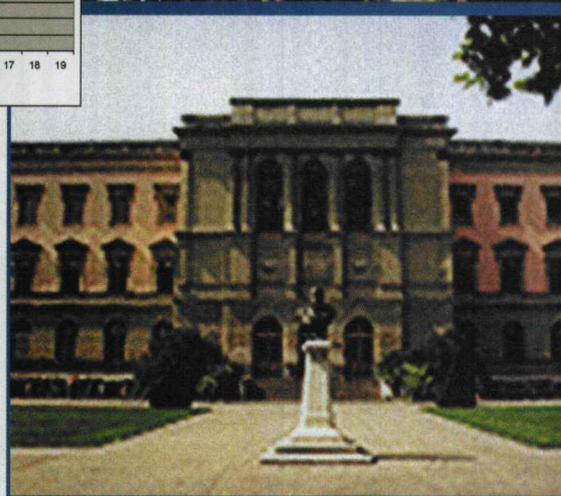
AMS 是一项符合国家目标、原创性的基础研究工作, 参与 AMS 项目将有助于我国空间科学与技术的水平的提高, 并有可能取得划时代的原创性科研成果。

我们依托 CERNET 和 CERNET2 所进行的一系列 AMS 数据传输测试实验的结果表明: 非对称路由在 IPv4 互联网中是常见现象, 它对于传输带宽的可用性有影响, 而 IPv6 网络由于目前结构尚比较简单, 因此路由的对称性很好; 另外, 现有的带宽测量工具是不精确的, 其测量方法需要改进。

CERNET2 是支撑基于海量数据处理的新型应用的一个十分有效的基础平台, 它为我国科学工作者依托互联网与国外同行进行大规模科学合作研究活动提供了可能, 因此是我国开展高水平科研的重要基础设施。CERNET2 隔绝了常规的互联网访问传输活动, 排除了这些活动对 AMS 实验数据传输的干扰, 这是教育网独特作用的一个范例。

虽然 AMS 是一个物理实验, 但它仍然给我们带来了网络研究方面的新的挑战。例如对系统带宽占用的自我约束与网络中继系统带宽管理的相互协作, 就是 AMS 地面数据传输系统需要进一步解决的问题, 而这个问题的解决对于改进 E2E 传输系统的行为特性具有积极的意义。

(作者单位: 东南大学计算机科学与工程系)



龚俭, 1957年8月出生, 工学博士, 现任东南大学计算机科学与工程系教授, 博士生导师。中国教育和科研计算机网 CERNET 专家委员会委员; 教育部现代远程教育专家委员会委员; CERNET 华东(北)地区网络中心主任, 江苏省计算机网络技术重点实验室主任; IEEE 会员, 中国电子学会高级会员, 江苏省互联网协会副理事长。20世纪80年代中期开始从事计算机网络方面的教学与科研工作, 涉及的研究领域包括开放系统互连理论及其应用、开放分布式处理理论及其应用、计算机互联网络工程、网络管理和网络安全等方面。



参考文献:

- [1] The AMS Collaboration, AMS on ISS Construction of a particle physics detector on the International Space Station, 2004
- [2] V. Choutko, A. Klimentov and M. Pohl, Computing Facilities for the AMS-02 ISS Mission, AMS Note AMSnote-2002_01_02
- [3] A. Elin, A. Klimentov, Data transmission programs for the AMS-02 ISS Mission, AMS Note 2001-11-02
- [4] Matt Mathis, Raghu Reddy, Enabling High Performance Data Transfer
<http://www.psc.edu/networking/projects/tcptune/>
- [5] <http://doc.in2p3.fr/bbftp>