ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.19, No.6, June 2008, pp.1481−1490
DOI: 10.3724/SP.J.1001.2008.01481
© 2008 by *Journal of Software*. All rights reserved.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

支持网格应用的光网络控制和管理*

任勇毛[1,2+], 唐海娜[1], 李 俊[1], 钱华林[1]

[1](中国科学院 计算机网络信息中心,北京 100190)
[2](中国科学院 研究生院,北京 100049)

Optical Network Control and Management for Grid Applications

REN Yong-Mao[1,2+], TANG Hai-Na[1], LI Jun[1], QIAN Hua-Lin[1]

[1](Computer Network Information Center, The Chinese Academy of Sciences, Beijing 100190, China)
[2](Graduate School, The Chinese Academy of Sciences, Beijing 100049, China)
+ Corresponding author: E-mail: renyongmao@cstnet.cn

Ren YM, Tang HN, Li J, Qian HL. Optical network control and management for grid applications. *Journal of Software*, 2008,19(6):1481−1490. http://www.jos.org.cn/1000-9825/19/1481.htm

Abstract: There are many research projects based on grid in many research fields like high energy physics (HEP). Grid applications including massive data transfer have very high requirement for network on QoS parameters such as bandwidth and delay. The traditional IP routing network can not satisfy these requirements. However, the connection-oriented optical network has this capability, but it still faces many challenges. The authors firstly analyzed the features of grid applications and their special requirements for optical network control and management, and emphatically analyze several existing methods of optical network control and management. Subsequently, current issues which are open are summarized. At last, some novel research issues such as Lambda Grid and OVPN are put forward.
Key words: grid; optical network; control and management; Lambda Grid; UCLP; GMPLS; ASON

摘 要: 在高能物理等科研领域有许多基于网格的科研项目,包括海量数据传输在内的网格应用对网络带宽、延迟等 QoS 指标等有很高的要求.传统的 IP 路由网络无法满足这些要求.面向连接的光网络能够较好地满足这些要求,但也面临着许多挑战.分析了科研网格应用的特点及其对光网络控制和管理的特殊要求,重点分析比较了几种现有的光网络控制和管理方式,归纳和总结了此领域尚待研究的问题,最后提出了 Lambda Grid 和 OVPN 等新的研究方向.
关键词: 网格;光网络;控制和管理;Lambda Grid;UCLP;GMPLS;ASON
中图法分类号: TP393 **文献标识码**: A

随着 Internet 的飞速发展,科研活动的网络化趋势越来越明显,以网格应用为代表的 E-science 受到了全世

* Supported by International Cooperation and Communications Project of National Ministry of Science and Technology of China under Grant No.2006DFA1990 (国家科技部国际合作与交流项目); the National High-Tech Research and Development Plan of China under Grant No.2007AA01Z214 (国家高技术研究发展计划(863))
Received 2007-01-24; Accepted 2007-09-13

© 中国科学院软件研究所 http://www.jos.org.cn

界的广泛关注.在高能物理、天文、地理等领域都有许多基于网格应用的研究项目.例如,在粒子物理研究领域,大型的国际合作和实验产生了巨大的数据量(PB 级/年),需要采用基于非常先进的高速网络的高能物理网格来支持如此海量数据的传输和处理.中国科学院高能物理研究所和欧洲粒子物理研究中心(CERN)之间有粒子物理方面的合作研究项目,位于 CERN 的大型粒子对撞机每年产生几十 PB 的数据,这些数据要采用网格传送到分布在全球的各个研究中心,以便处理和分析,这对网络有着非常高的迫切要求.在天文学领域,天文学家采用 VLBI(very long baseline interferometry,甚长基线干涉测量法)来获得详细的图像,相关的实验由全球分布式的仪器采集数据,通过网络传送到一个中心点.高性能计算和可视化研究系统中,需要远程传送 TB 级数据,需要有高带宽的链路(约 Gbps 或几百 Mbps).其他还有大气研究,e-health 以及 HDTV 等应用.

网格应用产生的网络流量通常具有带宽要求高,QoS 要求高,以及突发性等特点.图 1 为中国科学院高能物理研究所某段时间的 MRTG 网络流量图(数据来源:2006 年 11 月 11 日,11:05:06,中国科技网(China Science & Technology Network,CSTNET)网络中心).

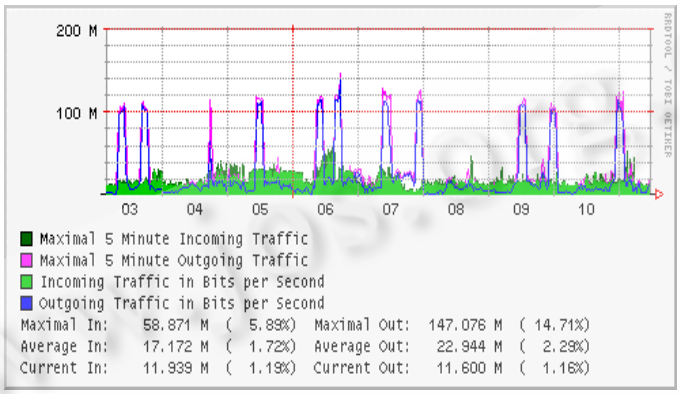


Fig.1 Network traffic graph

图 1 网络流量图

从图 1 可以看出,流量的突发性非常明显,突发流量对带宽要求很高.据预测,仅高能物理研究领域,高能物理实验对网络带宽的要求在 2013 年将达到 Tbps 级[1].

针对网格应用对网络的要求,传统的 IP 网络显得力不从心.首先,路由器的接口速率在做到 10Gbps 后,难以提高,Cisco 公司最新发布的 CRS-1 路由器,其接口速率是 40Gbps,这是目前业界的最高水平,再往上,技术难度大,而且代价昂贵.在设计方面,现有的传统 IP 网络,是基于三层路由的“尽力而为”的传输方式.一方面,采用分组交换,数据需要路由器中转.所有数据被分成用地址标记的包中.每个遇到包的路由器再决定数据包应该采用什么路径到达目的地.对于小而短时间的数据包(比如远程登录、E-mail 和 Web 应用)来说,这是有效的.但是,对固定点间的大数据流来说,这是费时且昂贵的方法.转发 1 500B 的每个包对于 E-mail 或 10KB~100KB 的网页来说很有效,但对于转发 1.5TB 的数据来说,意味着沿路径上的每个路由器都需要转发大约 10 亿次;另一方面,IP 网络的“尽力而为”的无连接方式本身无法提供严格的高质量 QoS 保证.

相对于 IP 网络,光网络具有很大的优势.一方面,光网络能提供巨大的带宽.采用 WDM 技术,实验室里已经做到了 10.9Tbps(273×40Gbps)带宽速率,国内的烽火科技公司开发了 3.2Tbps(80×40Gbps)带宽的 WDM 系统,国际上,1.6Tbps WDM 技术已经开始大量商用;另一方面,面向连接的光电路交换方式能够提供端到端的 Lightpath.在 Lightpath 上传输数据,无须路由,用户独享带宽,相当于真正的信息高速公路,具有低延迟、高可靠性等 QoS 和安全保障.分组交换方式可以共享带宽,节约带宽资源,但这是以 QoS 和安全性差为代价的,采用光网络,带宽资源极其丰富,电路交换方式能够提供高 QoS 和安全保障.由此可见,相对于传统 IP 网络,光网络能够更好地满足网格应用的需求.

本文第 1 节讨论国内外在支持网格应用的光网络控制和管理方面的有关研究工作和重大研究项目.第 2 节

© 中国科学院软件研究所 http://www.jos.org.cn

提出光网络支持网格应用所面临的各种挑战.第 3 节分析、比较基于管理系统和基于信令两种连接指配方式,具体分析 UCLP(user controlled lightpath,用户可控光通路)和 ITU-T G.7713 以及 GMPLS(generalized multi-protocol label switching,通用多协议标记交换)几种现有的控制和管理方式.第 4 节提出本领域尚待研究的问题.第 5 节总结全文并提出新的问题和方向.

1 相关工作

高速、高性能的网络是网格等 E-Science 应用的基础设施之一.光网络作为一种新的网络技术,以其巨大的带宽和优良的服务质量(QoS)以及端到端(end-to-end,简称 e2e)优势而受到普遍关注.但是,光网络的控制和管理面临诸多挑战,国际上许多先进科研和教育网络(R&E 网络)正在开展光网络控制和管理技术方面的研发、测试和部署.这方面有非常多的重大研究项目,比较典型的有欧洲的 PHOSPHORUS 项目[2].PHOSPHORUS 项目主要研究面向网格应用的光网络控制和管理问题,主要目标是在欧洲和国际范围跨多域多厂商设备的科研网络测试床上提供按需(on-demand)、端到端的网格网络服务.美国有几个大的研究项目,如:CHEETAH[3],DRAGON[4],UltraScience Net[5],HOPI[6]等,其中,UltraScience Net 是美国国防部(DOE)资助的一个实验性大规模科研测试床,用于研发混合光联网(hybrid optical packet networking)和相关技术,以满足大规模科学应用的极大需求.其他研究项目有本文将详细分析的加拿大 CANARIE 主导的 UCLP(user controlled lightpath)[7]技术研发项目,以及荷兰 SURFnet 正在进行的 OPN(optical private network)[8]项目.

中国科技网参与建设了中美俄环球科教网络(global ring network for advanced applications development, GLORIAD[9])项目.GLORIAD 是由中美俄三国发起,以及韩国、荷兰、加拿大等国参与的环绕北半球的高速光网络.GLORIAD 主要是为了支撑全球的科研合作,服务于各个成员国的 E-Science 等科研应用,并提供光网络、IPv6 技术等下一代互联网络技术的实验床.GLORIAD 骨干网络带宽为 OC-192(10Gbps),能够提供端到端的光通路(Lightpath,以下用 Lightpath 专指端到端的光通路).将光网络技术引入 R&E 网络,支持网格等 E-Science 科研应用,这是网络的一大飞跃,但是,网格应用对网络有何具体的要求和特点,如何控制和管理这种光网络,以满足各种网格应用的需要,有非常多的问题需要研究.基于 GLORIAD 高速光网络实验床,同 GLORIAD 的部分成员国家的技术人员一起,我们做了相关的研究和测试.

2 光网络支持网格应用面临的挑战

随着 WDM,OADMs,OXCs 等硬件系统的快速发展,光网络具备了提供高带宽的能力.但是,如何充分地利用这种带宽资源,提供各种服务,满足各种应用的需求,在光网络的体系结构、控制和管理等方面还有诸多问题亟待研究.光网络的控制和管理介于光网络的硬件和应用之间,是连接光网络与应用的桥梁.针对各种不同的网络应用,需要首先分析这些应用对光网络的要求.

不同的网格应用对网络的要求也不同,比如对带宽的要求,可接受的延迟以及丢包率.我们重点考虑有高带宽需求的应用.这些应用,有的对丢包率很敏感,要求可靠的数据传输.有些网格应用要求实时(低延迟)、持续较长时间的相对小的带宽,但是有很多同时运行的用户.尽管对网络的要求不同,但许多的网格应用有两个主要的共同的要求:大量便宜的按需预订的带宽和用户或应用对网络资源的控制和管理.

光网络是随着电信应用而发展起来的,最初只是考虑各种电信方面的应用,对于网格应用的支持,光网络面临许多新的挑战,总结起来主要有:

可扩展、灵活和可配置的网络基础设施;

能够提供超长距离的高带宽——支持 TB 级甚至 PB 级海量科学数据的传输;

具有按需带宽预订能力——波长级或子波长级,短期或长期(时间可以是几分钟,几小时,甚至数天);

提供端到端(e2e)连接——具有低抖动、低延迟等 QoS 保证;

网络资源与其他重要的网格资源协调——CPU、存储、以及可视化显示;

广播/多播能力;

© 中国科学院软件研究所 http://www.jos.org.cn

一些网格应用要求用户对光网络具有一定的控制能力[10],能够直接由用户控制网络带宽的提供,跨多个管理域,建立和删除端到端的 Lightpath.对用户来说,要求光网络能够根据应用需求自主、动态地建立跨越不同管理域的端到端的网络连接.

3 光网络控制和管理方式

光网络的基本用途是为不同的区域之间提供通信带宽,比如,在分散于两个不同地方的 IP 路由器之间提供一条链路.光网络控制和管理的核心问题是如何提供连接,包括连接的建立、修改、保护和恢复以及删除等.连接指配就是在光网络的两端点(这里,端点包括路由器、交换机或主机等)之间建立一条连接.传统上,它是光网络运营商接受到用户的服务请求以后作出的反应,由网络管理人员手动配置来完成,但这是费时、费力的办法,可能需要几天的时间,而且容易出错,这种方法不能满足频繁的、不同的连接建立请求,基本上是一种静态配置方法,无法满足网格应用对光网络控制和管理的要求.最近的发展趋势是把这个过程自动化、智能化,这就是智能光网络的思想.动态提供带宽是智能光网络的一个主要特征.

从连接指配方式来分,有两种方法可以建立连接[11].第 1 种方法是集中式的,通过网络管理系统向网元发送命令来建立连接;第 2 种方法是分布式的,通过在网元间传递控制信令的方式来建立连接.下面将详细分析这两种方法及具体的实现方案.

3.1 基于管理系统的连接指配

基于管理系统的连接指配方式,由管理系统首先计算连接的路由,然后使用 TL-1(transaction language 1)、简单网络管理协议(simple network management protocol,简称 SNMP)、通用管理信息协议(common management information protocol,简称 CMIP)以及通用对象请求代理结构(common object request broker architecture,简称 CORBA)等协议与网元通信,向各个网元分别发送连接建立命令,连接路径中的每一个网元独立地接收来自管理系统的指令,建立正确的交叉连接.这种方式,单独的网元只需完成自己的操作,而无须了解整个网络的拓扑状况.

UCLP(user controlled lightpath,用户可控光通路)是采用这种方式的一种具体实现技术.UCLP 软件是一种控制和管理连接的管理系统软件.UCLP 是 在 CISCO Canada 赞助及加拿大 CANARIE 主导下,由加拿大的通讯研究中心(CRC)、Ottawa 大学、Waterloo 大学和 Quebec 大学共同研发出控制光通路的软件.UCLP 作为一种新出现的技术,现已扩展为多国的合作研究计划,许多国家的 R&E 网络机构正在研究 UCLP 技术.

UCLP 软件系统架构[12]如图 2 所示.UCLP 系统采用 Jini 和 JavaSpaces 等技术开发,提供网格应用接口,以便网格应用可以访问系统来动态确定高带宽 Lightpath.交换通信服务(SCS)是分布式的,发现网络的物理资源并管理系统所有资源.UCLP 系统在底层与网元间通过 TL-1 进行通信,在顶层向用户提供网格服务访问接口(grid service access point,简称 GSAP).用户通过网格服务访问接口或从客户端 GUI 发送连接建立请求,UCLP 系统接受到请求后,自动计算路由(也可以由用户来指定路由),然后采用 TL-1 命令方式向各个网元发送连接建立命令,来完成连接的建立.UCLP 设计给用户自主权,用户可以根据应用需求自己设定连接的一系列参数,如带宽、连接持续时间、路由路径等.UCLP 系统提供连接请求预订服务,用户可以提前预订连接.采用 UCLP 系统管理方式,网络管理员不需要登录每个网元来进行连接指配.由于 TL-1 是一种标准化协议,绝大部分厂商产品都支持 TL-1,提供 TL-1 管理接口,因此,对于多厂商设备混合组网来说,采用 UCLP 系统来管理可以很好地解决跨不同厂商设备的连接建立问题.

UCLP 对网格应用有很好的支持.UCLP 系统将各种光网络资源(比如:带宽、接口、Lightpath 等)视为向用户提供的各种服务.UCLP 系统提供 3 种服务:Jini Service,Web Service 和 Grid Service.网格应用可以通过 UCLP 系统提供的网格服务访问点(GSAP)调用各种网络资源服务,包括端到端的 lightpath 连接服务.因此,UCLP 这种控制和管理方式对网格应用提供了很好的支持.

© 中国科学院软件研究所 http://www.jos.org.cn

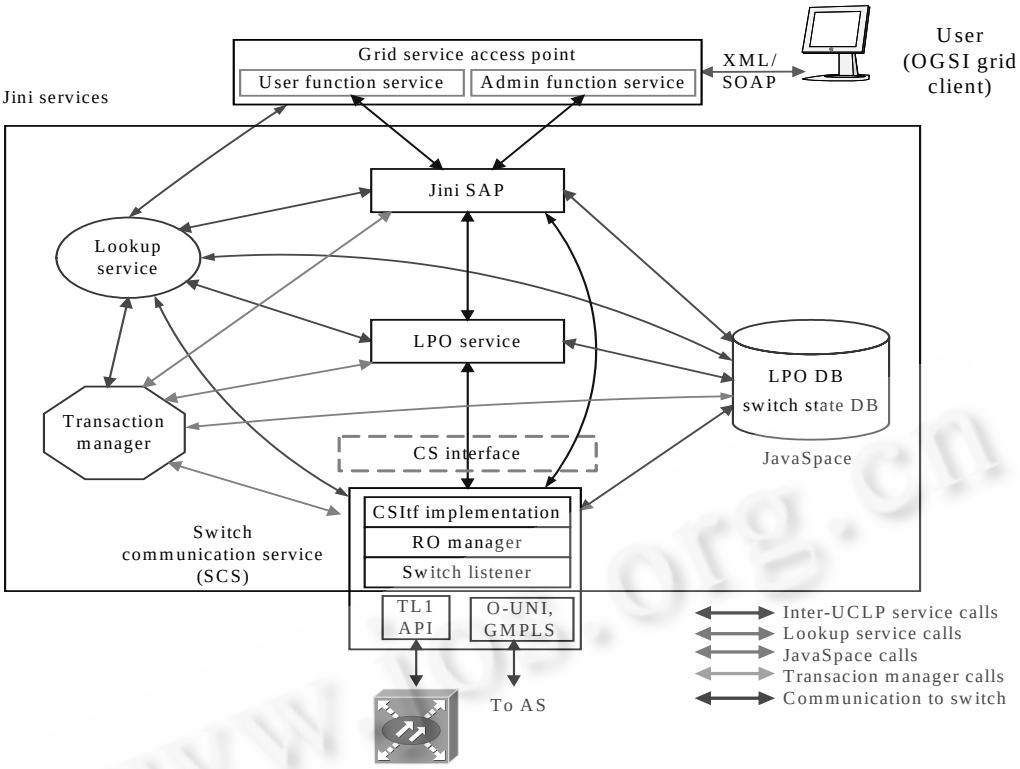


Fig.2 UCLP system architecture

图 2 UCLP 系统架构

中国科技网参与了国际上 UCLP 相关的国际交流和合作.在加拿大 CA*net 4 和西班牙 i2CAT 的技术人员的帮助下,我们对 UCLP 技术进行了研究,在 Cisco 中国公司的帮助下,采用 Cisco ONS 设备搭建了硬件实验环境,在本地以及加拿大CA*Net4 的实验环境中部署UCLP系统,进行了多次测试,并和外方技术人员一起解决了软件中存在的部分问题和缺陷.测试结果表明,采用 UCLP 系统可以跨不同厂商设备(比如 Cisco 的 ONS 和 Nortel 的 OME 等)较好地实现 Lightpath 的指配和管理.测试中也发现 UCLP 系统存在一些问题,比如支持的设备类型有限,连接保护和恢复功能不完善,路由计算过于简单,以及安全问题等等.当然,UCLP 技术仍然在发展中,UCLP 系统软件也在不断完善中.

3.2 基于信令的连接指配

基于信令的连接指配方式,使用连接路径中网元之间的控制通信,这种通信是一种信令,控制消息流从源网元开始,经过中间网元,到达宿网元.信令协议定义了控制消息的语法、语义和决定网元如何发送、接收与处理这些控制消息的规则.在许多信令协议中,指配请求消息在源网元和宿网元之间要往返多次.连接指配也是从管理系统开始执行的,由管理系统触发源网元去建立连接,这个网元将计算路由,或者该路由由管理系统来计算,并随指配命令一起提供给网元.

与连接指配相关的信令已经被 ITU-T、OIF 和 IETF 标准化了.这些标准化组织开发了具体的体系结构模型.ITU-T 提出了自动交换光网络(auto-switched optical network,简称 ASON)的概念,在 ITU-T G.8080 建议中定义了 ASON 的体系结构.ITU-T 的建议 G.7713 描述了光网络中的分布式呼叫和连接管理模型.尽管有些抽象,此建议规范了许多信令相关的概念.GMPLS(generalized multi-protocol label switching,通用多协议标记交换)是 IETF 的成果,它也处理光网络中信令相关的内容.G.7713 定义了信令接口和功能需求,GMPLS 定义了具体的信令消息、对象和过程.GMPLS 信令因此可以在 G.7713 中定义的任何信令接口中使用.

© 中国科学院软件研究所 http://www.jos.org.cn

3.2.1 ITU-T G.7713

ITU-T G.7713 用于处理符合 ITU-T G.8080 体系结构的光网络的分布式呼叫和连接管理.这篇建议描述了与信令相关的参数和流程,但是它独立于任何具体的信令协议.它定义了一些信令的功能需求和体系模型.G.7713 的主要作用是明确了信令的接口,发展了光网络中信令控制平面和数据传送平面相分离的体系结构,这也包含了信令消息流的描述、呼叫和连接控制有限状态机的描述.

在 G.7713 模型中,逻辑控制平面功能和它的物理实现是分离的.这样的好处是便于处理不同类型的子网,在子网间采用标准的 NNI(网络-网络 接口)信令,来实现端到端的连接指配.连接指配过程如图 3 所示.图 3 中,端到端连接被指配为一系列段,其中一些在单条链路上,另一些则在子网中.在连接指配的起始阶段,每一个处理信令消息的控制器都要决定转发信令消息的下一个控制器以及连接在本地的资源分配.决定下一个控制器的路由信息可能在也可能不在信令消息中.如果不在,这个控制器就必须计算一条局部路由.本地资源的分配将取决于连接的参数.此外,如果是子网控制器,那么,在分配本地资源时,也许还需要为连接段在子网内分配完整的路径.G.7713 建议只是描述了通用规范,遵循这个模型的具体协议各有独立的定义.到目前为止,支持 G.7713 功能的协议有:P-NNI、RSVP-TE 和 CR-LDP.

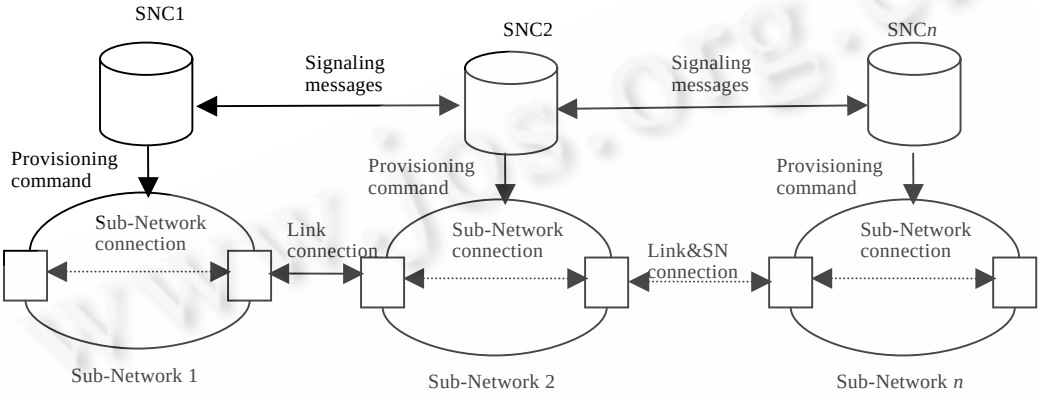


Fig.3 End-to-End connection provision cross sub-network of G.7713

图 3 G.7713 跨子网的端到端连接指配

3.2.2 GMPLS

GMPLS 对 MPLS 进行了扩展,标签不再被包含在数据中进行传送,被用来标识一个时隙、波长或端口,通过它们,连接数据可以被交换(按照接口类型).GMPLS 支持的接口类型包括:支持分组交换的接口(如路由器上的 MPLS 接口),支持时分复用(TDM)的接口(如 SONET 和 SDH 接口),支持波长交换的接口以及支持光纤交换的接口.GMPLS 信令的目的,本质上是沿着连接的每一个 LSR 可以分配标签,并与邻居交互,使得邻居可以建立正确的交叉连接.

GMPLS 的连接指配,既可以由管理系统发起(软永久连接),又可以由客户设备通过信令发起(交换连接).两种方式,都是在源网元上接受命令.不论是由源网元计算的还是由指配发起者给出的,如果连接路由可用,源网元都向连接路由中的下一个网元发起请求,请求分配一个标签(如标识端口或时隙).请求消息中包含连接的端点标识和其他参数.如图 4 所示,在路径中,每一个网元处理来自前一个网元的请求,同时再向下一个网元发送请求,直到请求到达目的网元.然后,响应从目的网元逐跳返回到源网元.响应消息包含下游网元为上游网元分配的标签.具体的连接建立由 GMPLS RSVP-TE[13]在指配连接的请求和响应阶段分别使用 Path 消息和 Resv 消息来通知网元配置交叉连接,最终建立起一条端到端的连接.

在实际网络部署中,连接指配主要通过 GMPLS 分布式控制平面来实现.GMPLS 分布式控制平面具有 3 个主要的功能使其能够较好地满足网格应用的要求[14].它们是:分布式路由,近乎实时的采用信令方式的连接建立,网络资源和拓扑自动发现机制.由 GMPLS 软件控制的每个网元使用 OSPF-TE(open shortest path first with

© 中国科学院软件研究所 http://www.jos.org.cn

traffic engineering)路由协议,RSVP-TE(resource reservation protocol with TE)信令协议和 LMP(link management protocol)链路管理协议.GMPLS 控制平面对于需要紧急建立连接的应用很有用.对于可以提前预订的大带宽连接,近乎实时的提供和分布式路由能力在提高网络的带宽利用率方面作用不大.但是,无论是什么类型的应用,控制平面引入的自动发现功能都很有价值.

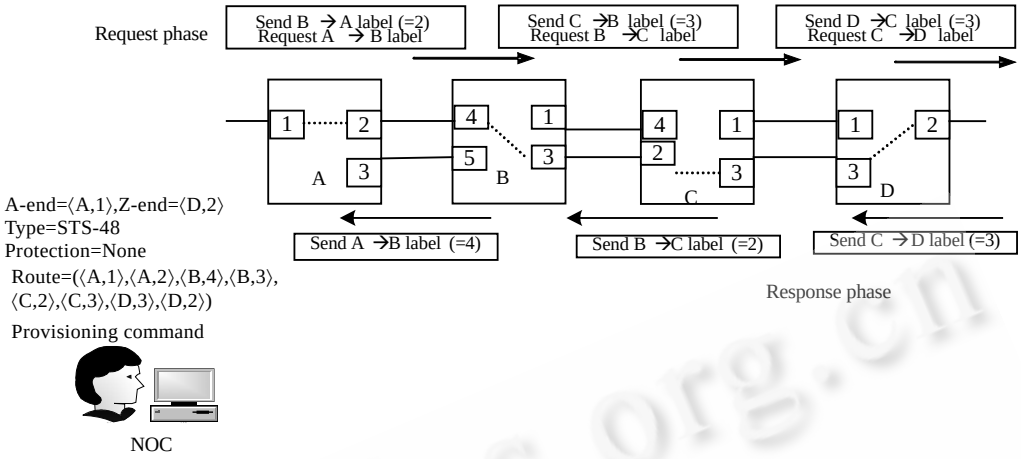


Fig.4 Label request and response process in connection provision of GMPLS

图 4 GMPLS 连接指配中的标签请求和响应过程

3.3 比较分析

基于管理系统的连接指配和基于信令的连接指配两种方式有其各自的特点.

从技术实现的角度看,采用基于管理系统的连接指配的好处是容易实现.所有网络都需要一个高度可靠的管理系统,无论是否使用信令建立连接,这个系统必须与所有网元连通,而且,系统维护整个网络的状态信息,因此,使用它去指配连接显得理所当然.采用信令方式的好处是,减轻当前的网络管理系统的负荷,并将使得网元对某些需要重新指配连接的失效(如节点失效)有更加快速的反应.

从用户应用的角度看,GMPLS/ASON 由用户应用请求连接,但是不能按需控制 lightpath 的指配(如 lightpath 路由选择),也就是说,网络实现了自动连接指配,用户却对此毫无控制.UCLP 与此不同,UCLP 设计为用户可控光网络(user controlled optical network),允许让用户来控制连接的建立和释放,包括路由选择、带宽分配、使用时间等,而这些参数在网格应用所需网络连接中常常有特殊的要求.UCLP 允许用户修改和拆分 lightpath,与其他用户共享 lightpath,以此可以分担成本.

从设备厂商角度看,由于现有设备大都提供有 TL-1 等控制和管理协议接口,UCLP 在技术上相对容易实现.但是,采用信令方式,需要有标准协议,需要各厂商设备来支持,这将是一个比较漫长的过程.

从技术进展的角度看,UCLP 技术相对比较成熟,在全球 R&E 网络中进行了多次测试,并且有结合 E-Science 具体应用的测试,比如国际间高能物理网格海量数据传输、HDTV 传输演示、国际视频会议等.以 ITU-T G.8080 和 G.7713 为代表的 ASON 技术,涉及整个网络体系结构,比较繁杂,到目前为止,只是定义了一套框架、组件功能和各种接口需求,有待进一步细化和通过具体的协议来实现.ASON 概念目前在国内很热门,但具体的实现还需要做大量的研究和测试工作.关于 GMPLS,国际上有不少的研究项目,如美国的 DRAGON 等正在对 GMPLS 进行相关的研究和部署测试.

网格应用经常涉及到国际间的合作,建立端到端 lightpath 常常需要跨越多个管理域,每一个管理域内都可能包含不同厂商的设备,而且,每一个管理域内的连接指配方式都可能不同,比如,一个管理域中使用基于管理系统的指配,另一个使用基于信令的指配.如何跨多域异构网络提供端到端的 lightpath 还有诸多问题需要研究和解决.

© 中国科学院软件研究所 http://www.jos.org.cn

4 尚待研究的问题

光网络控制和管理技术本身还在发展,针对网格应用的特定需求,光网络的控制和管理还有许多问题需要研究.以下将总结已知的尚待研究的一些问题.

1) 控制平面相关的问题.控制平面是光网络控制和管理的关键.对控制平面的研究已经取得了较大的进展,但是还有不少问题仍待研究.比如,物理损伤相关问题,现有的路由算法只是以跳数(hops)或连接长度作为代价参数,对物理层进行了理想的假定,如假定物理层特征位错率(BER)等是固定的[15,16],然而实际情况却并非如此,因此,基于损伤约束的路由(impairment-constraint-based routing,简称 ICBR)问题需要研究.文献[17]提出了基于约束的动态路由和波长分配算法(CDRWA),但此算法也没有考虑跳数和位错率之间的关系.另一个问题是失效定位问题,失效包括由故障和攻击引起的,现有的方法适用于不透明光网络,到现在为止,尚没有透明光网络的失效定位方法[18].

2) 面向连接的光网络的异构性问题.由于光网络目前不能完全代替 IP 网络,当前引入光网络主要是与 IP 网络共存的混合网络(hybrid network),既提供无连接的 IP 服务,也提供面向连接的电路服务.在引言部分,我们提到的好几个研究项目,如 CHEETAH,DRAGON,UltraScience Net,以及 HOPI 等等,都是为了探索创建和使用这种混合网络的方法.当前的研发主要集中在具体的面向连接网络(如 WDM,SONET/SDH)的技术和指配算法.缺乏研究的一个关键问题是控制平面的互操作问题以及相关子问题.比如,域间路由问题,跨异构网络的分布式路径计算算法和分布式带宽管理算法.我们用“Multi-X”来描述异构性(heterogeneity),包括有 Multi-Service(多服务),Multi-Layer 和 Multi-Technology(多层和多技术),Multi-Domain(多域),Multi-Provi-der(多提供商)以及 Multi-Vendor(多厂商)等.许多国际间的网格应用需要跨越多个管理域的网络,如何跨多域异构网络提供端到端 lightpath,国际上这些大的研究项目都还在对此进行研究之中.比如,针对光网络域间路由问题,加拿大的 CANARIE 提出了扩展 BGP 协议针对光网络的 OBGP(optical border gateway protocol)协议[19].

3) 其他问题.比如协议的性能问题,连接建立的时间不能太长,以及连接的快速恢复等.针对网格等具体应用的需求,也有许多问题亟待研究.具体的实现方,比如 UCLP 和 GMPLS 也存在不少问题.光网络与传统 IP 网络共存的混合网络的组网、控制和管理以及对各种应用的支持,目前还有许多问题.第 1 节提到的许多国际研究和测试项目正在尝试对各种方案和技术进行实验测试,在此过程中也将不断发现新问题,提出新的解决方案.光网络控制和管理的标准与协议也在发展之中,不同厂家的光网络设备的互通问题也比较突出.

5 结束语

不断出现的网格应用需求对当前的网络技术提出了很大的挑战.为了满足这些应用的要求,网络技术,尤其是网络的控制和管理技术正在进行重大的变革.将相对于传统 IP 路由网络具有很大优势的光网络技术引入 R&E 网络来支持科研网格应用,这是一个巨大的进步.由于光网络技术本身正处在不断发展之中,网格应用对光网络有着特殊的要求,因此,支持网格应用的光网络在控制和管理方面面临着诸多挑战.基于管理系统的连接指配和基于信令的连接指配两种光网络控制和管理方式,具有其各自的特点,它们具体的实现方案以及相关路由、信令协议均在发展中.在光网络的控制平面、跨异构网络的连接指配等许多方面,还有诸多尚待研究的问题.尽管如此,许多国家都在积极开展此领域的研究,对此问题的研究,推动了光网络技术和科研网格应用的发展,呈现出了一些新的趋势.

传统的网格是基于 IP 路由网络,由于光网络技术的引入,现在还出现了 Optical Grid、Photonic Grid(光子网格)[10]、Lambda Grid(波长网格)[19,20]等概念.这些概念尚无统一定义,但都是采用光网络来构建网格.Lambda Grid 的一个定义[19]是采用光交换机的波长路由网络,提供动态按需的具有数个 Gbps 带宽的 lightpath 以互连共享的用户域的计算资源.Lambda Grid 更多地关注光联网的问题,而不是网格的问题.Lambda Grid 刚刚提出,还有许多问题需要研究.除了光联网技术之外,在 OSI 模型的其他层也产生了相应问题需要研究,比如传输层协议的性能问题[21−23].我们在 GLORIAD 网络的实际测试中发现了此问题.传统的 TCP 协议在低带宽、低延迟的网络

© 中国科学院软件研究所 http://www.jos.org.cn

环境中,工作很好,但是对于高带宽、稳定可靠、长途传输(高延迟)的光网络来说,TCP 的慢启动、拥塞机制等机制不再适用,反而阻碍了吞吐率的提高,以至于远远不能有效利用高带宽.因此,需要改进 TCP 协议或设计新的协议,或采用其他方法来改进传输层的性能.另外,在 OSI 的 1 层,荷兰 SURFnet 提出 OPN(optical private network)[8]技术,采用租用波长链路等方式,为网格应用中各个不同的学科分别建立它们各自的 1 层的 VPN 网络,在逻辑上形成各个学科的网格社区,OPN 技术在过去几年取得了许多进展,但是,具体的实施、算法、性能等仍有待进一步展开研究.

致谢 感谢加拿大 CANARIE 提供 UCLP 实验环境和技术帮助,感谢西班牙 i2CAT 技术人员的帮助,感谢思科公司和中国科技网的李宏高级工程师等人帮助建立 ONS 光网络实验环境.

References:

[1] Simeonidou D, Nejabati R. Optical networking for grid services. In: Proc. of the 47th Annual IEEE Global Telecommunications Conf. (Globecom 2004). 2004. 385−389. http://ieeexplore.ieee.org/iel5/9692/30597/01417609.pdf?arnumber=1417609

[2] PHOSPHORUS [Online]. 2008. http://www.phosphorus.pl/

[3] Zheng X, Veeraraghavan M, Rao NSV, Wu QS, Zhu MX. CHEETAH: Circuit-switched high-speed end-to-end transport architecture testbed. IEEE Communication Magazine, 2005,43(8):11−17.

[4] Lehman T, Sobieski J, Jabbari B. DRAGON: A framework for service provisioning in heterogeneous grid networks. IEEE Communications Magazine, 2006, 84−90.

[5] Rao NSV, Wing WR, Carter SM, Wu Q. Ultrascience net: Network testbed for large-scale science applications. IEEE Communications Magazine, 2005,43(11):12−17.

[6] Summerhill R. Next generation networking and the HOPI testbed. In: Proc. of the CANS 2005. 2005. http://networks.internet2.edu/hopi/presentations/hopi-cans-summerhill-2-nov-2005.pdf

[7] Communications Research Centre. User controlled light paths. 2005. http://www.uclp.ca/

[8] SURFnet. GigaPort Annual Report 2005. Technical Report, 2005. http://www.surfnet.nl/info/en/innovation/gigaport/results/home.jsp

[9] Tang HN, Ren YM. GLORIAD: Supporting global research cooperation. China Computer World, 2006,13(44):B13 (in Chinese)

[10] Wu J, Savoie M, Campbell S, Zhang HX, Bochmann GV, St Arnaud B. Customer-Managed end-to-end lightpath provisioning. Int'l Journal of Network Management, 2005,15:349−362.

[11] Bernstein G, Rajagopalan B, Saha D. Optical Network Control: Architecture, Protocols and Standards. Boston: Addison Wesley, 2003.

[12] CANARIE. UCLP System Overview. Technical Report, 2005. http://www.canarie.ca/canet4/uclp/index.html

[13] Berger L. Generalized multi-protocol label switching (GMPLS) signaling resource ReserVation protocol-traffic engineering (RSVP-TE) extensions. IETF RFC 3473, 2003.

[14] Habib IW, Song Q, Li ZM. Deployment of the GMPLS control plane for grid applications in experimental high-performance networks. IEEE Communications Magazine, 2006.

[15] Gerstel O, Kutten S. Dynamic wavelength allocation in all-optical ring networks. In: Proc. of the IEEE ICC'97. 1997. 432−436. http://ieeexplore.ieee.org/Xplore/login.jsp?url=/iel3/4646/13277/00605331.pdf?arnumber=605331

[16] Ramaswami R, Sivarajan KN. Routing and wavelength assignment in all-optical network. IEEE/ACM Trans. on Networking, 1995,3(5):489−500.

[17] Wu WJ. Modeling and analysis of GMPLS-based automatically switched optical network [Ph.D. Thesis]. Arizona: University of Arizona, 2003.

[18] Report of "US/EU Workshop on Key Issue and Grand Challenges in Optical Networking". Technical Report, Brussels Workshop, 2005. http://networks.cs.ucdavis.edu/~mukherje/US-EU-wksp-June05/US-EU-wksp-June05-Final-Report.pdf

[19] Canarie. Optical BGP (OBGP): InterAS lightpath provisioning. Internet draft, 2001. http://obgp.canet4.net/

[20] Yu OT, DeFanti TA. Collaborative user-centric Lambda-grid over wavelength-routed network. In: Proc. of the ACM/IEEE Supercomputing (SC 2004). 2004. 32. http://ieeexplore.ieee.org/iel5/9595/30315/01392962.pdf

© 中国科学院软件研究所 http://www.jos.org.cn

[21] Wu XR, Chien AA. Evaluation of rate-based transport protocols for Lambda grids. In: Proc. of the IEEE Conf. on High-Performance Distributed Computing (HPDC-13), 2004. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1323499

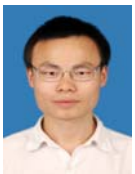
[22] Optical network infrastructure for grid. draft-ggf-ghpn-opticalnets-1. In: Proc. of the Global Grid Forum. 2002. http://anf.ne.kr/document/draft-ggf-ghpn-opticalnets-1.doc, 2005

[23] Huang XM, Lin C, Ren FY. Recent development of high speed transport protocols. Chinese Journal of Computers, 2006,29(11):1901−1908 (in Chinese with English abstract).

附中文参考文献:

[9] 唐海娜,任勇毛.GLORIAD:支撑全球科研合作.计算机世界,2006,(44):B13.

[23] 黄小猛,林闯,任丰源.高速传输协议研究进展.计算机学报,2006,29(11):1901−1908.



任勇毛(1981－),男,湖南邵阳人,博士生,CCF 学生会员,主要研究领域为高速网络,传输协议.

李俊(1968－),男,博士,研究员,主要研究领域为下一代互联网,网络监测.

唐海娜(1977－),女,工程师,主要研究领域为网络管理,网络监测.

钱华林(1940－),男,研究员,博士生导师,主要研究领域为下一代网络体系结构.

www.jos.org.cn

www.jos.org.cn

© 中国科学院软件研究所 http://www.jos.org.cn