

陆颖华, 马廷淮, 曹浩, 等. 大量气象数据小文件自适应优化传输. 应用气象学报, 2014, 25(5): 629-637.

# 大量气象数据小文件自适应优化传输

陆颖华<sup>1)</sup> 马廷淮<sup>1)\*</sup> 曹浩<sup>1)</sup> 李德泉<sup>2)</sup>

<sup>1)</sup>(南京信息工程大学计算机与软件学院, 南京 210044) <sup>2)</sup>(国家气象信息中心, 北京 100081)

## 摘 要

为满足气象实时资料传输中大量气象数据小文件的高时效传输需求, 对其中的数据运输服务进行优化, 提出一种基于实时网络状况的自适应数据传输优化方法。该方法采用优化网络传输协议并使用文件压缩技术, 通过获取网络传输线路上的实时参数, 实时调整压缩参数和网络传输参数以优化传输性能。在自适应压缩时, 通过试验分析和归纳, 确立了气象数据小文件标准为文件小于 50 KB; 根据网络实时状况, 设计了基于实时网络状况自适应调整压缩等级的算法。在自适应传输参数调优中, 研究了 TCP 缓冲区大小和 TCP 并发连接数在 GridFTP 协议中的重要性, 针对实时网络状况, 分别设计了自适应调整 TCP 缓冲区大小和 TCP 并发连接数的算法, 算法提升传输性能 65%。对以上提出各自适应参数调整算法进行试验验证表明, 融合压缩和网络传输的自适应调优方法能显著提升气象小数据文件的传输性能近 500 倍。

**关键词:** 气象数据; 小文件; 压缩; 传输优化

## 引 言

气象观测资料作为各类观测的有效数据记录, 是对天气过程进行预测及对大气活动规律进行研究的重要依据, 其有效交换是气象业务和科研的基础。目前, 通过国内通信系统、国际通信系统及同城服务系统组成的通信业务系统, 建立了台站、省级、国家级 3 级资料传输与服务体系<sup>[1-2]</sup>, 该体系承担绝大部分气象观测资料的传输任务, 实现了稳定可靠、完整、及时地收集、存储和共享实时气象观测资料。

近几年, 随着气象现代化业务的发展及精细化预报、短时临近预报的不断发展, 对资料传输的信息量、准确性和时效性不断提出新要求。如何进一步提高各类数据传输的时效性也一直是提升气象业务服务能力的热点问题。

根据目前的业务现状, 数据传输与服务分为两种典型情况: 一种是小数据文件传输与服务, 如区域自动气象站和雷达产品数据。对于自

动气象站资料, 区域自动气象站(全国约 4 万)正常情况下为每小时上传, 国家级自动气象站未来将加密至分钟级观测并上传, 该类型文件大小多为千字节量级。多普勒天气雷达每次体扫实时生成约 40 种产品, 每种雷达产品的数据量不等, 按平均 50 KB 计算, 以上数据每 6 min 生成 1 次。另一种情况是单个数据容量大, 数据传输及服务往往需分割, 如数值预报产品和卫星数据等。数值预报产品中, 欧洲中期天气预报中心(ECMWF)1 个时次的集合预报产品大约 50 GB, 其大气集合预报模式(中国周边区域)单个文件为 330 MB。

数据传输服务需要解决两个问题: 一是大量异构文件的传输问题。由于气象部门采用的多种探测仪器、软件系统和通信系统, 产生了不同类型的文件, 比如文本文件、二进制文件; 以及不同数据内容的文件, 比如气温数据或是降水数据。为了实现气象数据在气象各部门之间的共享, 需要考虑异构数据在网络环境下的传输问题。二是数据传输的实时性问题。由于本文重点关注大量气象数据小文

2013-11-07 收到, 2014-05-26 收到再改稿。

资助项目: 江苏省教育厅青蓝工程中青年学术带头人项目“基于云计算的气象数据共享平台”, 江苏省“六大人才高峰”高层次人才项目(2012-WLW-022)

\* 通信作者, email: thma@nuist.edu.cn

件,因此该问题就转变为如何从海量异构文件中过滤出大量小文件并高效传输,以提高传输时效。通过对以上两个问题的解决方法分析,本文旨在给出一个整体的解决方案,充分利用整个气象信息网络上各个节点的计算机资源、存储资源,发挥网络资源的整体效益,进而切实提高气象部门的信息共享水平。

解决以上问题首先要过滤大量异构文件并进行高效压缩,其次要优化网络传输协议的参数设置,实现压缩数据的快速传输。

由于气象数据中存在大量异构小文件,将小文件压缩成一个大文件传输将有效减少输入/输出(I/O)访问次数,将有效提高数据传输效能。数据压缩技术一般分为两类:无损压缩和有损压缩。无损压缩要求经压缩还原后能产生精确的数据副本,有损压缩则是以一定的精度损失换取较高的数据压缩率。气象数据传输过程中不可失真,拟采用基于 LZO(Lempel Ziv Oberhumer)<sup>[3]</sup>算法的无损压缩技术。

在气象领域,提高数据传输时效的一个解决方案是建立高效的数据网格,通过 GridFTP 部件实现网格环境中的数据传输任务。文献[4-6]采用网络技术设计建立气象数据网格环境,并将传输服务作为其核心部分,尝试高效管理各类气象文件,从而使管理繁杂的气象数据变得更加高效和稳定。其中,文献[6]给出了国家气象计算网格的设计说明。数据网格同样可以应用于气象学领域<sup>[7-8]</sup>,使用 Globus Toolkit 4.0 构建网格服务,通过 GridFTP 协议传输大规模气象数据,并大量使用了并行传输技术。

采用 GridFTP 并对其传输优化是常用方法。针对 GridFTP 的研究主要分为两个方面:一是修改 GridFTP 协议本身的实现方式,GridFTP 为开发者提供了可以扩展的接口<sup>[9]</sup>,Tian 等<sup>[10]</sup>描述了一个兼容远程内存直接存取(RDMA)的远程扩展输入输出接口(RXIO),它基于 GridFTP 中扩展输入输出接口 XIO 系统设计实现,兼容 RDMA 特性。试验结果显示它能够显著提升 GridFTP 性能,降低 32% 的网络时延。Alberto 等<sup>[11]</sup>通过优化数据存储接口(DSI)来体现如何读写存储系统,并结合并发文件系统(MAPFS)设计了一个新的 DSI。二是动态调整 GridFTP 协议的网络参数,文献[12-13]提出了 GridFTP 自动参数配置机制,该机制利用网络状态试验结果来优化 TCP 并发连接数,如吞吐量和 GridFTP 数据通道的往返时间。文献[14-17]描述

了如何将自适应调整缓冲区大小机制整合到 GridFTP 中,以此来提升 TCP 性能。另外,OTN/WDM 网络并行传输优化框架<sup>[18]</sup>可支持高速因特网。文献[19-20]呈现了一组新模型——使用最低信息和最低开销来预测并发数和缓冲区大小的最优值。

## 1 自适应数据压缩算法

考虑到气象数据中存在大量异构小文件且要求传输过程中不失真,将小文件采用无损压缩技术生成一个大文件传输将有效减少 I/O 访问次数,并提高数据传输效能。

### 1.1 大量小文件的定义

#### 1.1.1 文件大小阈值的设定

气象数据文件大小多为 2 KB~60 MB,可通过试验方法确定需要压缩的文件大小阈值。根据气象数据文件大小,分别考察了 1 KB,10 KB,50 KB,100 KB,250 KB,500 KB 和 1000 KB 共 7 种数据,并分别为各种数据准备了 1000 个和 10000 个文件组成的两组数据。所有数据采用相同传输方法,计算性能提升百分比(图 1)。

图 1 表明,两组数据呈相同的变化趋势,可以得出将大量小文件压缩为几个有限的体积较大的文件,其传输效率会有很大提升。试验结果显示,当文件超过 50 KB 时,其性能提升度明显下降。因此,将数据文件低于 50 KB 的气象数据文件定义为气象数据小文件,传输前宜先进行压缩。

#### 1.1.2 压缩包文件数设定

压缩技术对网络传输的影响主要包含两个方面:压缩/解压时间和压缩后的文件大小。压缩技术对网络传输性能提升在不同的网络环境下不同。一般来说,网络带宽越小,压缩技术对传输性能的提升越明显。带宽越大,网络传输本身性能尚可,但仍需压缩和解压时间,总体性能是否提升有待考察。所以采用压缩技术应满足以下条件:

$$t_T > t_{zip} + t_{ZT} + t_{uzip} \quad (1)$$

式(1)中, $t_T$  为某组文件不压缩时的传输时间, $t_{zip}$  为该组文件的压缩时间, $t_{ZT}$  为该组文件压缩后的传输时间, $t_{uzip}$  为该组压缩文件传输后的解压时间。

小数据文件数量很多,在传输前将其压缩为一个文件,然后进行传输,最后接收端进行解压,期待提升整个文件的传输效率。压缩时需要确定每个压缩文件中所包含的文件数量。

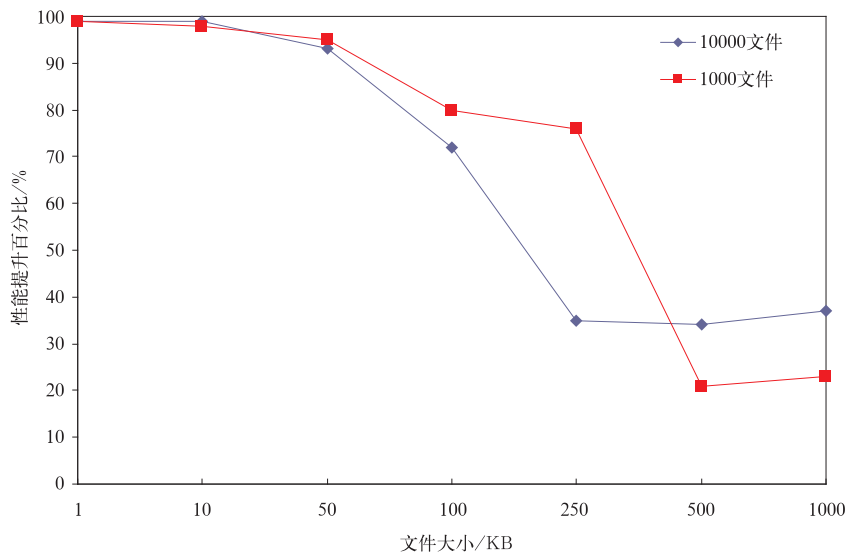


图 1 不同大小文件压缩后传输性能提升情况

Fig. 1 Different transmission performances with different file sizes

基于式(1),全部文件的传输时间表示为

$$T(N, n, s) = T_c ns + t(\beta + \frac{ns}{r} T_i) + T_d ns + R$$

$$= (T_c + T_d) ns + \frac{N\beta}{n} + \frac{Ns}{r} T_i + R. \quad (2)$$

式(2)中,  $N$  表示所有的文件数量,  $n$  表示每组文件数量,  $t = \frac{N}{n}$  表示组数,  $s$  为文件大小,  $T_c$  表示每个字节所需的压缩时间(则每组压缩时间为  $T_c ns$ ),  $T_d$  表示每个字节所需的解压时间(每组解压缩时间为  $T_d ns$ ),  $r$  为压缩倍数(压缩后的文件大小为  $\frac{ns}{r}$ ),  $T_i$  表示 1 个字节的传输时间,  $\beta$  为网络时延,  $R$  表示其他开销。

最佳传输方案要求  $T(N, n, s)$  取最小值, 因此基于上述公式对  $n$  求偏导得到其最小值,

$$n = \sqrt{\frac{\beta N}{(T_c + T_d)s}}. \quad (3)$$

由式(3)可知, 需要确定  $\beta$ ,  $T_c$ ,  $T_d$  和  $s$  才可确定每组的文件数  $n$ 。其中,  $s$  为 50 KB。使用 ping 命令确定  $\beta$ , 参考 LZ0 相关数据确定  $T_c$  和  $T_d$ :  $25 \times 10^{-3} \text{ s} \leq \beta \leq 500 \times 10^{-3} \text{ s}$ ,  $0.241 \times 10^{-7} \text{ s/B} \leq T_c \leq 1.35 \times 10^{-7} \text{ s/B}$ ,  $0.9 \times 10^{-8} \text{ s/B} \leq T_d \leq 1.23 \times 10^{-7} \text{ s/B}$ 。代入式(3), 得

$$\sqrt{1.9084N} \leq n \leq \sqrt{75.5290N}. \quad (4)$$

结果表明:  $n$  与文件总量密切相关, 气象文件总量大, 则  $n$  的取值区间也比较大。由图 1 可知, 当批

量传输文件时, 文件数少, 传输性能高。因此,  $n$  的取值应尽可能大, 即减少压缩包数量。

### 1.2 基于 LZ0 的改进压缩算法

本文基于 LZ0 算法<sup>[3]</sup>, 使用现有的 lzop 算法库, 通过设置 lzop 命令的各个参数控制整个压缩流程。lzop 命令共有 9 个压缩级别: 级别 1 表示压缩速度最快, 但压缩率最低; 级别 9 表示压缩速度最慢, 但压缩率最高。

自适应的压缩算法要根据网络实时状况调整压缩级别。本算法使用往返时延(RTT)来判断当前网络的拥塞状况, 比较当前 RTT 和上一次测量的 RTT, 最终判断是否需要改变压缩等级。其中, 当需要增加压缩等级时, 需要进行多次传输进行确认, 以免出现瞬时网络波动现象, 本算法设定该传输次数初始值为 0, 传输次数阈值为 2。

算法分为两个部分, 其流程如图 2 所示。流程图左边是主线程的压缩数据文件, 右边是获取实时优化压缩等级参数进程。算法执行中, 实时测量 RTT 的值, 将计算得出的压缩等级  $L$  和当前 RTT 值存入数据库, 同时将保留一份到内存中, 以供主线程和下次计算过程使用。在主线程中, 如果内存中当前压缩等级为空, 则使用 lzop 默认的压缩等级 3 来压缩该组文件。其中, 为了减少压缩时间, 采用了 lzop 最新支持的命令 parallel 进行并行压缩, 其命令格式为 \$ parallel lzop ::: \*.lzo。

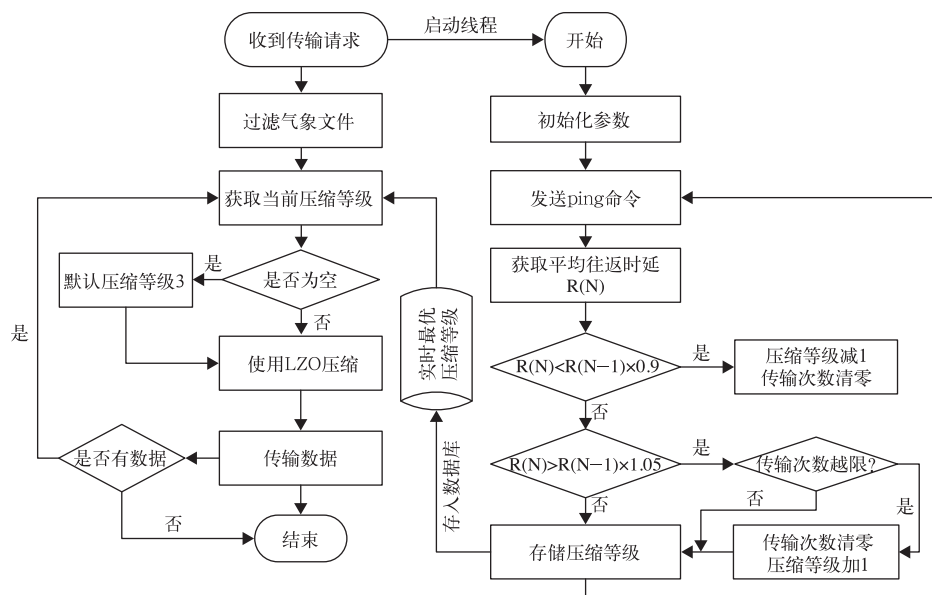


图 2 自适应压缩算法流程图

Fig. 2 Flow chart of adaptive compression algorithm

### 1.3 自适应压缩传输试验

试验的硬件环境设置为 3 台服务器,其中服务器含有两个网络接口卡,即将服务器用作路由器,由此可以实现类似路由器的配置,另外两台服务器通过该服务器互连。每个服务器包含 dual 2.66-GHz Intel(R) Core(TM) 处理器、2 GB 内存和 1000 Mbps 的网络接口卡(NIC)。数据传输采用 GridFTP 服务,使用默认参数设定的 GridFTP 传输协议,在保证传输协议和传输参数等相同的前提下评估自适应数据压缩算法。

试验数据为南京信息工程大学气象台气温、空气湿度、土壤湿度等观测数据,大小为 230 MB,其中小文件共 194245 个。共进行 3 组试验:试验 A1 传输过程中不使用压缩技术,试验 A2 传输数据前压缩所有的小文件,试验 A3 传输过程中应用自适应压缩算法。在 3 次试验中,保持了网络传输参数的一致。每一组试验进行 10 次操作,取平均值作为参考依据。其试验结果:试验 A1 传输时间为 3425 s,试验 A2 为 10 s,试验 A3 为 7 s,试验 A1 和 A2 体现了应用压缩技术的优势,试验 A2 和 A3,其差距并不明显,试验中仅相差 3 s,但试验结果也说明自适应压缩算法具有一定优势。

## 2 自适应传输参数调优算法

数据网格 Globus 平台采用 GridFTP 作为文件

传输协议,是一个基于 ftp 的高性能、安全、可靠的数据传输协议。宽带网 CMAnet 具有分布式传输、网络带宽低、网络时延高等特点。本文采用 NETEM 工具模拟宽带网 CMAnet,设置网络时延为 30 ms,丢包率为 1%,包重复率为 1%,包损坏率为 0.2%。该协议针对高带宽广域网进行了优化,提供了一些非常适合数据网格环境的服务<sup>[9,21-22]</sup>。其中,优化并行传输和 TCP 缓冲区大小是有效提高传输效率的两个重要方面。

### 2.1 并行传输和 TCP 缓冲区

为检验并行传输和 TCP 缓冲区大小对传输性能的影响,本文设计了两组试验来分别验证并行传输条数和缓冲区大小的优化设置的可能性。两组试验均在搭建的 Globus 平台上,采用 GridFTP 传输大小为 10 MB 的文件从一台服务器传输到另一台服务器。

试验结果表明,随着并行流条数的增加,文件传输更加高效,但当并行流条数超过 5 时,传输时间趋于稳定。

随着 TCP 缓冲区不断扩大,文件传输也更加高效。但当 TCP 缓冲区超过 100 KB 后,传输时间趋于平缓,且有增大的趋势。这表明对于每次传输,TCP 缓冲区大小都有一个最优值,即 BDP(带宽-时延积)值,因此,传输前将 TCP 缓冲区大小设置为 BDP 变得非常必要。

TCP 缓冲区和并行传输均消耗 CPU 和内存资

源。同时,设置过多的并行流条数和过大的 TCP 缓冲区,可能导致网络拥塞。因此在改进 GridFTP 的时候,需综合考虑,合理分配各种资源,达到最优化。

### 2.2 基于 GridFTP 的自适应参数调整算法

气象网格中网络环境经常变动,带宽和时延并不稳定,而二者都对文件传输有明显影响。因此,为了获得理想性能,可以针对带宽和时延进行优化,即根据当前网络状况实时地改变网络参数值。

#### 2.2.1 自适应 TCP 缓冲区调整试验

文献[13]指出,当 TCP 缓冲区大小设置为 BDP (带宽-时延积)的值时,可达到最优值。对于给定的

链路,本文将 TCP 缓冲区大小设置为当前 BDP 的值。

自适应 TCP 缓冲区调整算法流程如图 3 所示。该算法分为两个部分:流程图主线程采用 ABUF (Auto-Negotiate Buffer Size)调整缓冲区大小,并实时计算 BDP 值。在保持其他网络传输条件的一致性的前提下,本文通过两组试验来验证自适应 TCP 缓冲区调整的有效性,试验 B1 使用固定的 TCP 缓冲区传输数据,即 200 KB;试验 B2 是在传输过程中采用自适应 BDP 调整算法。试验数据同 1.3 节,试验结果如图 4 所示。

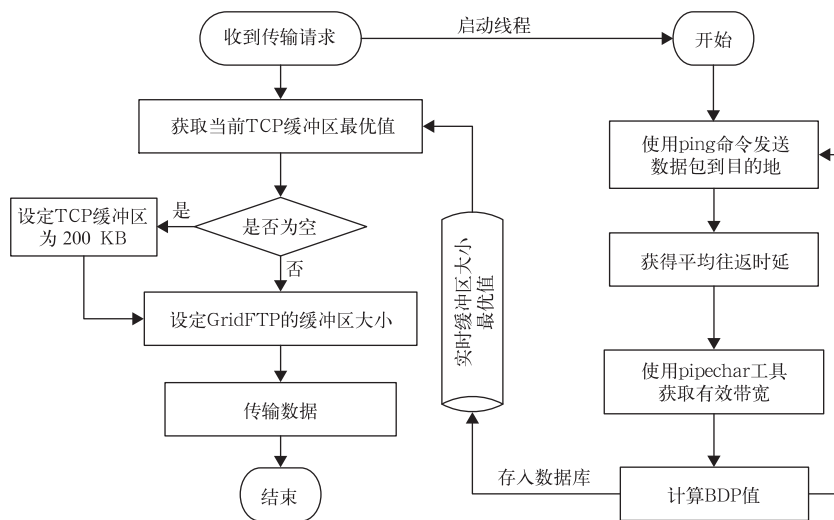


图 3 自适应 TCP 缓冲区调整算法流程图

Fig. 3 Flow chart of adaptive TCP buffer size adjustment

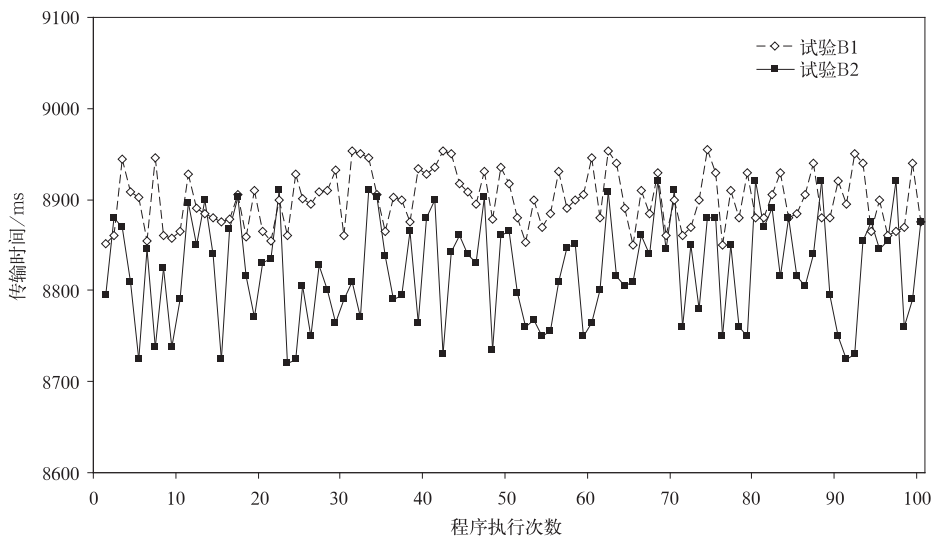


图 4 自适应 TCP 缓冲区调整试验结果

Fig. 4 Result of adaptive buffer size adjustment

由图 4 可知,对于这两组试验,自适应 TCP 缓冲区调整比固定缓冲区下传输性能有明显提高。

### 2.2.2 自适应 TCP 并发数设置试验

采用 TCP 并发传输技术可以进一步提高网络传输性能。如何得到最优 TCP 连接数,是 TCP 并发传输中的一个关键问题。目前,已经有了这方面的研究成果,如 GridFTP-APT 算法<sup>[18]</sup>。该算法基于式(5)和式(6)<sup>[17]</sup>:

$$G \approx \min \left[ \frac{NW}{R}, \frac{N(1-p^*)}{2R} \left( -3 + \frac{\sqrt{6+21p^*}}{\sqrt{p^*}} \right) \right], \quad (5)$$

$$p^* \approx \left[ -2 + \frac{2BR}{N} + \frac{2}{3} \left( \frac{BR}{N} \right)^2 \right]^{-1}. \quad (6)$$

其中, $N$ 为 TCP 并发连接数, $W$ 为每个 TCP 连接的缓冲区大小, $B$ 为链路瓶颈带宽, $R$ 为 TCP 连接的往返时延, $p^*$ 为链路丢包率。上式表明带宽  $G$  会随着  $N$  的数目增加而不断增加,而  $N$  过大时将降低所获取的带宽  $G$ 。

Chen 等<sup>[18]</sup>指出,当 TCP 连接数不断增长时,传输性能将有极大值出现,其最优 TCP 连接数为 10~30。当 TCP 并发连接数增加时,数据的整体传输速率随之也提升,但到后期,数据传输时间趋于平稳。另外,网络测量为即时测量,网络背景流量在时刻变化,传输时间有着极大的瞬时性,波动较大。所以,并不能仅通过一次测量来判断网络状况。针对这个问题,GridFTP-APT 作出以下假设:使用 1 s 的时间传输数据,所获得的传输数据的值能够准确反映网络状况。

使用 GridFTP-APT 得到最优 TCP 连接数后,将在后面的所有传输过程中一直使用,假设网络状况稳定,不会改变。然而,网络实时流量变化使最优 TCP 连接数不断发生改变,因此,需要一种动态算法,可以随着网络状况的变化自适应改变 TCP 并发连接数,始终能够保证得到较好的传输性能。

但如果 TCP 连接数过大,将导致网络吞吐量下降,这是因为在总窗口大小不变的前提下,每条 TCP 链接的窗口容量将会很小,容易造成某些链路发生超时、丢包、重传等网络状况;同时 GridFTP 服务器和客户端用于处理 TCP 协议栈的开销增加。目前并没有比较充分的研究成果,如何在多变的网络环境中决定 TCP 并发连接数仍是一个有待解决的难题。

根据网络实时状况自适应调整 TCP 并发连接

数。其中,网络实时状况的量化方法是基于链路往返时延实现的,即实时获取链路的 RTT 值。通过比较当前网络时延和上一次获取的网络时延,确定当前网络的状况,进而判断当前使用的 TCP 连接数是否需要更改。判断准则:若实时网络时延降低 10%,则将 TCP 并发连接数增长 20%;若实时网络时延增长 5%,且满足限定条件,则将 TCP 并发连接数减 1<sup>[17]</sup>;若实时网络时延在这两者之间,则维持当前 TCP 并发连接数不变。

通过 3 组试验来验证自适应 TCP 缓冲区调整的有效性:试验 C1 串行传输数据,试验 C2 传输数据时使用固定的 TCP 并发连接数 4,试验 C3 传输过程中采用自适应 TCP 并发连接数自适应调整算法。在 3 组试验中,保持其他网络传输条件一致,即 3 组试验设置了相同的 TCP 缓冲区大小,三者使用了 globus-url-copy 命令将文件从客户端传输到服务器。试验数据仍然采用 1.3 节中的试验数据进行传输试验。采用串行传输需 10093 s,采用固定并行和自适应并行传输分别为 5134 s 和 3827 s,从试验结果可看出采用并行传输的优势。而对于固定并行传输和自适应并行传输,其差距不明显,但该试验结果仍可以说明自适应调整 TCP 并发连接数算法的优越性。

## 3 自适应综合传输试验

本研究将整合基于 LZ0 的改进压缩算法和自适应传输参数调优算法,设计一套整体的大量气象小文件传输算法框架,并用试验加以验证,算法流程如图 5 所示。在收到数据请求后,立即启动新线程,用于采集实时网络状况,并计算实时最优缓冲区值、实时最优 TCP 连接数和实时最优压缩等级,并将其分成两份,一份放入数据库供后期统计使用,一份放入内存中供主线程使用。需要说明的是,压缩过程和传输过程分属于不同的线程,即存在压缩线程和传输过程,读写文件的路径都为数据请求中指定的路径,压缩线程将压缩后的文件放在相同路径下,传输线程传输该路径中的文件。

基于网络试验平台,将自适应算法应用于数据传输中。通过 3 组试验来验证自适应优化传输的有效性:试验 D1 使用 GridFTP 默认参数传输未经压缩的文件,试验 D2 使用 GridFTP 固定传输参数传输已压缩的文件,其中 TCP 缓冲大小设为 100 KB,

TCP 并发连接数设为 4, LZO 压缩等级设为 3, 试验 D3 在传输过程中采用本文自适应算法。在 3 组试验中, 保持了其他网络传输条件的一致, 三者使用了 globus-url-copy 命令将文件从客户端传输到服务器。试验数据与 1.3 节中的试验数据一致。

试验 D2 相对于试验 D1 在性能上有明显提升, 由试验 D1 的 3400 s 降低到 D2 的 10 s。但本文在此基础上, 针对本文特定应用进行深入研究, 试验 D3 性能(7 s)相对于试验 D2 又有了进一步的提升。

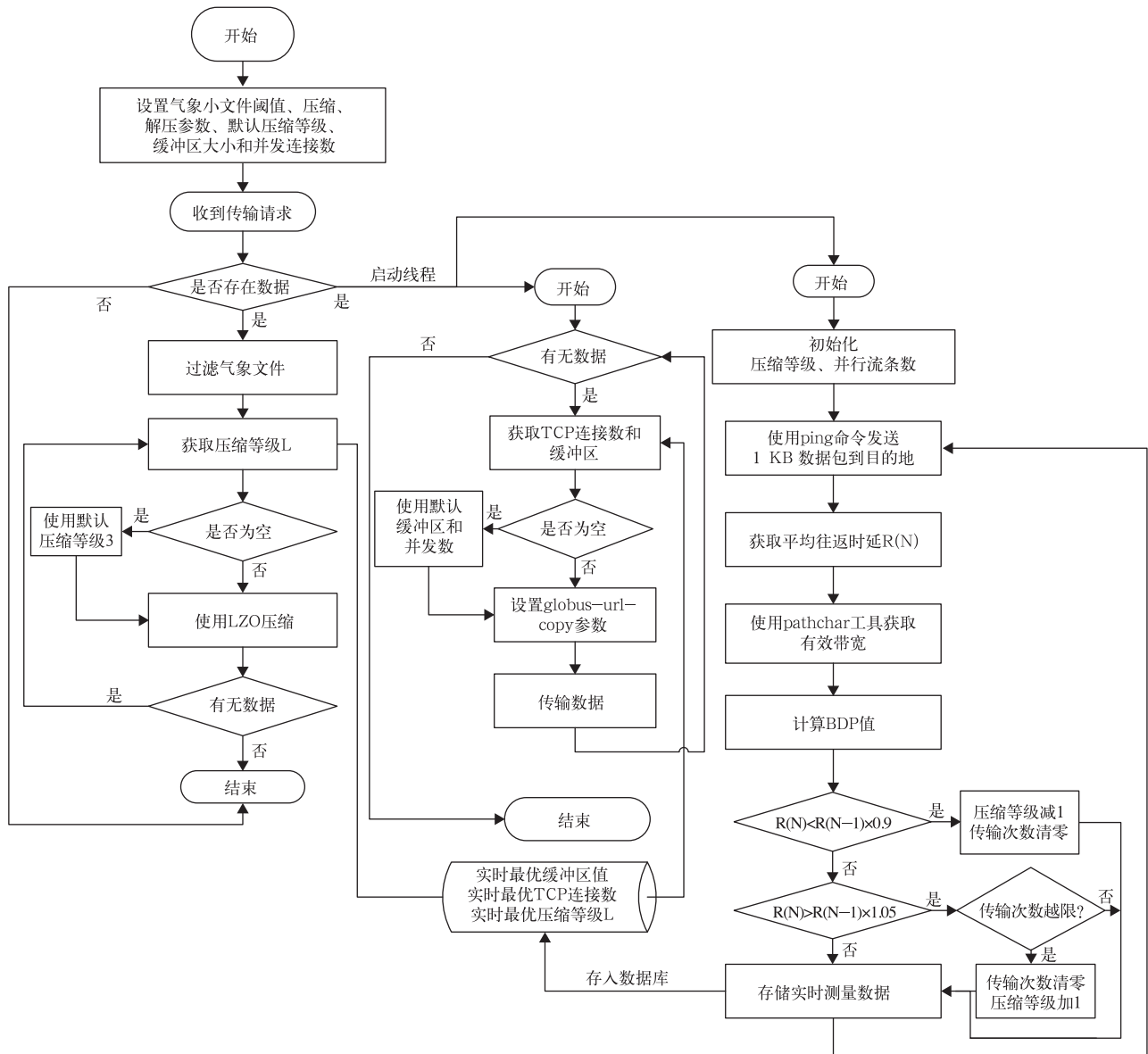


图 5 自适应综合传输算法流程

Fig. 5 Flow chart of combined adaptive transmission algorithm

#### 4 小 结

本文以国家气象数据网格为应用背景, 研究了如何高效传输大量气象数据小文件, 给出了一套整

体优化方案。具体表现在以下几点:

- 1) 在分析各类气象文件种类及其特性的基础上, 得到大量气象数据小文件的限定范围为小于 50 KB。
- 2) 根据网络带宽和往返时延(RTT)经常变动

特性,设计基于实时计算带宽-时延积(BDP)的自适应 TCP 缓冲区大小和并发连接数的设置方法,并自适应调整文件压缩等级。

3) 基于相同的试验环境和数据,本文提出的气象数据小文件传输方法与传统方法相比,传输速度提升近 500 倍,特别是小文件压缩传输带来了显著的传输性能提高。

### 参考文献

- [1] 李集明,熊安元. 气象科学数据共享系统研究综述. 应用气象学报,2004,15(增刊 I):1-9.
- [2] 高梅,接连淑,张文华. 气象科研数据共享系统建设. 应用气象学报,2004,15(增刊 I):17-25.
- [3] [2012-05-12]. <http://zh.wikipedia.org/wiki/LZO>.
- [4] 邓莉,王国复,孙超,等. 基本气象资料共享系统建设. 应用气象学报,2004,15(增刊 I):33-38.
- [5] 祝婷,李湘. WMO 信息系统中气象元数据的设计与实现. 应用气象学报,2012,23(2):238-244.
- [6] Ma Tinghuai, Ge Jian, Cao Hao, et al. Design and Implementation of Virtual Resources Management in Meteorology Grid. 9th International Conference on Grid and Cooperative Computing (GCC), 2010:58-63.
- [7] 高峰,王国复,喻雯,等. 气象数据文件快速下载服务系统的设计与实现. 应用气象学报,2010,21(2):243-249.
- [8] 周铮嵘,王铮,何文春. 分布式气象元数据同步系统的探索研究. 应用气象学报,2010,21(1):121-128.
- [9] Allcock W. GridFTP: Protocol Extensions to FTP for the Grid// Global Grid Forum. 2003.
- [10] Tian Y, Yu W K, Vetter J S. RXIO: Design and implementation of high performance RDMA-capable GridFTP. *Computers & Electrical Engineering*, 2012, 38(3):772-784.
- [11] Alberto S, Maria S P, Pierre G, et al. A Parallel Data Storage Interface to Gridftp// Robert M, Zahir T. OTM Conferences (2), Lecture Notes in Computer Science. 2006:1203-1212.
- [12] Takeshi I, Hiroyuki O, Makoto I. Automatic Parameter Configuration Mechanism for Data Transfer Protocol GridFTP. 2006 International Symposium on Applications and the Internet (SAINT'06), 2006:32-38.
- [13] Thulasidasan S, Feng W, Gardner M K. Optimizing GridFTP Through Dynamic Right-sizing. Proceedings of IEEE International Symposium on High Performance Distributed Computing, 2003.
- [14] Ma Teng, Luo Junzhou. Optimizing Large File Transfer on Data Grid. Lecture Notes in Computer Science, 2005:455-460.
- [15] Ito T, Ohsaki H, Imase M. On Parameter Tuning of Data Transfer Protocol GridFTP in Wide-area Grid Computing. Proceedings of Second International Workshop on Networks for Grid Applications (GridNets 2005), 2005:415-421.
- [16] Ito T, Ohsaki H. GridFTP-APT: Automatic Parallelism Tuning Mechanism for Data Transfer Protocol GridFTP. Proceedings of the Sixth IEEE International Symposium on Cluster Computing and the Grid, 2006:454-461.
- [17] Patrick M, Ezra K, Martin S, et al. MNEMONIC: A Network Environment for Automatic Optimization and Tuning of Data Movement over Advanced Networks. Proceedings of 18th International Conference on Computer Communications and Networks, 2009:1-7.
- [18] Chen X, Jukan A. Optimized Parallel Transmission in OTN/WDM Networks to Support High-Speed Ethernet with Multiple Lane Distribution (MLD). *IEEE/OSA Journal of Optical Communications and Networking (JOCN)*, 2012:248-258.
- [19] Yildirim E, Yin D P, Kosar T. Prediction of optimal parallelism level in wide area Data transfers. *IEEE Trans Parallel Distrib Syst*, 2011, 22(12):2033-2045.
- [20] Yin D P, Yildirim E, Kulasekaran S, et al. A data throughput prediction and optimization service for widely distributed many-task computing. *IEEE Trans Parallel Distrib Syst*, 2011, 22(6):899-909.
- [21] [2012-02-15]. <http://www.globus.org/datagrid/gridftp.html>.
- [22] Bresnahan J, Link M, Kettimuthu R, et al. GridFTP Pipelining. Teragrid Conference, 2007:1-6.



## Adaptive Optimization in Small Size File Transmission of Massive Meteorological Data

Lu Yinghua<sup>1)</sup> Ma Tinghui<sup>1)</sup> Cao Hao<sup>1)</sup> Li Dequan<sup>2)</sup>

<sup>1)</sup> (School of Computer & Software, Nanjing University of Information Science & Technology, Nanjing 210044)

<sup>2)</sup> (National Meteorological Information Center, Beijing 100081)

### Abstract

The data transfer and service architecture constructed by National Meteorological Information Center is the fundament for most meteorological data transmission. How to improve the timeliness of transmission of various data is a hot topic to enhance capabilities of meteorological services.

According to requirements of transmission performance of massive small files, transmission parameters are optimized. And a self-adapting data transmission method is proposed based on real-time network status, which emphasizes network transmission protocol and file compression. Compression parameters and network transmission parameters are adjusted in real-time operation.

Meteorological data include a great amount of heterogeneous small files, therefore compressing small files into a big file when being transformed will effectively reduce I/O accesses. First, 50 KB is defined as the threshold for small meteorological data files through experiments. Then, by analyzing the file transfer time, the appropriate file amount in compressed packages is calculated to achieve the best transmission efficiency. Finally, considering the variability of network conditions and real-time network conditions, a self-adapting compression methods based on real-network is designed by means of real-time adjusting the compression level. This entire compression process is controlled by setting various parameters of lzop commands on the basis of the lzop algorithm library and the LZO algorithm. To achieve the goal of adjusting compression levels according to real-time network conditions, RTT (round trip time) is taken advantage of judging the current state of the network congestion. By comparing current RTT and previous RTT, changing the compression level or not is decided.

In network transmission optimization, conclusions are made that TCP buffer and parallel transmission will consume memory resources according to experiments in Globus platform. At the same time, more parallel streams and larger size of TCP buffers will result in network congestion. Then, the self-adapting adjustment algorithm of TCP buffer size and the concurrent connection number algorithm of TCP based on real-network parameters are designed. Finally, the entire transmission framework of massive small files is designed by combining self-adapting compression method and transmission parameters optimization. Complete experiments are carried out based on the integration of self-adapting algorithm, showing that proposed optimization methods can improve the transmission performance sharply.

**Key words:** meteorological data; small files; compression; transmission optimization