ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.20, No.8, August 2009, pp.2255−2268
doi: 10.3724/SP.J.1001.2009.03386

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

并行下载最优机制*

王 正[1,2+], 罗万明[1], 阎保平[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)

Optimal Mechanism of Parallel Downloading

WANG Zheng[1,2+], LUO Wan-Ming[1], YAN Bao-Ping[1]

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

Wang Z, Luo WM, Yan BP. Optimal mechanism of parallel downloading. *Journal of Software*, 2009,20(8): 2255−2268. http://www.jos.org.cn/1000-9825/3386.htm

Abstract: Parallel downloading is modeled. And based on the model, the optimal mechanism is analyzed and proposed in terms of the performance of downloading nodes. The optimal mechanism includes the optimal source nodes set selected and file blocking scheme. Theoretical analysis proves that this mechanism can minimize the cost function of downloading nodes. Simulation results justify the validity of this mechanism.
Key words: parallel downloading; optimal mechanism; bandwidth allocation; file block; peer-to-peer system

摘 要: 建立并行下载模型,并基于这一模型,从下载节点效益的角度出发,分析并提出最优机制.这种最优机制包括最优的选择源节点集合和文件分块方案.理论分析表明,这种机制能够使下载节点代价函数最小.仿真结果说明了这种机制的有效性.
关键词: 并行下载;最优机制;带宽分配;文件块;对等系统
中图法分类号: TP393 **文献标识码**: A

下载节点能够同时向不同的拥有目标资源的源节点发出下载请求,下载一个目标资源的不同部分,这种对下载节点有利的下载方式通常可称为并行下载.并行下载对下载节点的直接好处是下载的总带宽增加了,而且当源节点突然离开网络或者发生故障时,并行下载也能提供更好的稳健性.并行下载对下载节点也有好处,因为它们不必单独为一个文件提供下载服务,而是可以和其他拥有同一文件的节点分担责任.

现今大量的对等(peer-to-peer,简称 P2P)系统利用了并行下载的技术,如 eMule[1],Kazaa[2],Limewire[3], Overnet[4].但是这些系统有的是把一个文件分为相同大小的块,从不同的节点请求下载这些块(如 eMule 和 Overnet),有的是向源节点反复发出下载小块的请求(如 Kazaa 和 Limewire).针对对等网络并行下载进行分析的文献并不多见,如文献[5]提出利用机器学习的方法来帮助节点从存有需要文件的源节点中选取一个,而不是利用所有的源节点的带宽之和,着重从源节点提供的带宽和网络下载时间的角度出发,找到最可靠的源节点,而没

* Supported by the National Basic Research Program of China under Grant No.2003CB314807 (国家重点基础研究发展计划(973))
Received 2007-11-01; Revised 2008-03-11; Accepted 2008-05-05

有利用并行下载的思想.文献[6,7]研究利用节点的协作来增加整个系统的存储容量,但只是针对节点能够相互协作的应用.文献[8–10]讨论了一种共享和提供资源的架构:OceanStore.尽管这些文献深入分析了这种架构,并以并行下载的普遍使用为前提来提供性能和冗余的保证,但没有涉及到下载的分块和相关的策略.其他研究[11–13]探讨了对等系统的性能,但没有从节点的角度分析并行下载的情形.文献[14]针对高丢包率和高延迟的动态网络,提出以 UDP 为传输层协议的并行下载方法,但没有分析最优的节点选择和分块策略.

并行下载的实现主要需要解决 3 个方面的问题:(1) 如何获取所有源节点的列表;(2) 如何得到文件的长度;(3) 如何在源节点之间分块.其中最重要的是第(3)点,因为如果分块不合理,下载的效率则可能很低,有时甚至要比单个源节点的下载更低.现有的方法大致可以分为两种:(1) 平均分块:把整个文件平均分为 n 块,分别从每一个源节点下载其中的一块[1–4];(2) 按比例分块:根据从源节点的历史记录获取的下载性能(带宽、吞吐量等等),按照与每一个源节点性能成比例的方式分块.从充分利用各个源节点性能的角度,按比例分块要比平均分块更优.高性能的源节点承担了更多的下载任务,会使下载的时间更小.如文献[15]提出了用于代理服务器的可变块下载方法,这种方法在不增加下载时间的前提下可以减小冗余的通信量和代理服务器的缓存大小;文献[16]从提高分块的稳定性和减小下载时间出发,提出基于预测的动态分块方法;文献[17]提出根据传输带宽的变化,自适应地动态分配下载任务的方法.此外,并行下载机制的研究还包括拓扑感知的源节点选择[18]、并行下载的博弈论模型[19]以及源节点编码[20–22]等.

现有的并行下载实现方案大都是从最小化下载时间的目标出发,而没有考虑下载节点的系统资源消耗.现今的对等网络对一些文件往往可以提供几百个源节点,对于一些热门资源,甚至可能达到上千[23].同时,利用大量的源节点来完成下载,从理论上来说固然可以加速下载,但同时也给系统带来沉重的负担(如下载线程的资源消耗、分块的完整性检查等).考虑到相当数量的低性能源节点可能对下载的时间贡献很小但却消耗了大量的系统资源,而简单地限制最大源节点数的方法[24]则有可能对下载时间造成过多的限制,本文研究了同时优化下载时间和系统消耗的并行下载机制,从理论上得到了综合考虑下载时间和系统消耗的分块方案,即在下载时间最小和系统消耗最小意义上的令节点效益最大的并行下载策略,并分别针对是否考虑源节点策略的情况提出并证明了最优的源节点选择策略和分块策略.仿真结果表明,对比传统的两种分块方法,这种并行下载策略是最优的,相对于其他策略对节点效益有较为明显的改善,能够在降低系统消耗的同时令下载时间保持一个较小的值.

1 并行下载策略

1.1 并行下载模型

系统中有一个节点的集合 A,A 中的每一个节点可以发起下载请求,如果它有某一目标文件,那么它同时也可以成为源节点.因此,整个系统可以表示为 $A=I\cup N\cup \Upsilon$,这里 I,N 和 Υ 分别是发起下载请求的下载节点、提供目标文件的源节点和空闲节点的节点子集.可以有 $I\cap N\neq\varnothing$,也就是说,节点可以在从其他节点下载文件的同时,也为其他节点提供下载文件.

为了简化问题,我们假定在节点下载期间,源节点的上传带宽基本保持稳定或者变化很小.从另一角度来看,在源节点的上传带宽变化较大时,我们可以把整个下载过程分为若干阶段,在每一阶段,源节点的上传带宽近似不变.当然,节点可以根据源节点上传带宽的变化随时调整它们的策略,但这种调整必然引入附加的代价(如发起新的下载请求造成的时延、增加同时下载的节点数引起的线程数的增加等).当引入的代价超过调整策略带来的额外代价时,节点宁愿选择保持策略不变.总之,我们仅讨论在源节点的上传带宽不变的情况下,下载节点/源节点的策略.

每一个节点用于对等的上传带宽和下载带宽分别为$[B_u,B_d]$,我们假设节点连接过程中没有发生网络拥塞,这可以使我们从端到端的角度建立模型,而不考虑底层网络的实际拓扑.

系统中的节点 i 发送下载大小为 O_i 的请求,它接收来自于节点集合 N_i 的反馈.这里,$N_i\subset N$.然后,节点 I 将形成自己的最优分块策略,把下载文件分为从节点集合 N_i 并行下载的小块.最初,节点 i 会下载一些小的分块

$O_{i,j}[0]$,这里,在时刻 $t[0]$,$j\in N_i$.因为节点不知道 $B_{u,j}[0]$的大小,$\forall j\in N_i$,它将从所有各个节点下载同样大小的分块.因此,在时刻 $t[0]$,我们有 $O_{i,j}[0]=O_{i,k}[0],\forall j,k\in N_i$.

节点 i 可以测量从每一个源节点的下载时间,计算对每一个源节点可用带宽的粗略估计 $B^*[0]$,即在时刻 $t[0]$的观测带宽.节点 i 的策略就是如何在各个源节点之间分配分块的大小并下载这些块.我们令 $O_{i,j}[n]$为在时刻 $t[n]$节点 i 从节点 j 下载的块.系统的约束为

$$\sum_{j\in N_i}\sum_{n=0}^{T} O_{i,j}[n]=O_i,\ \forall i\in I \tag{1}$$

$$\sum_{j\in N_i} B_{u,ji}[n]\le B_{d,i},\ \forall i\in I \tag{2}$$

$$B_{u,ji}[n]\le B_{u,j},\ \forall j\in N_i \tag{3}$$

这里,T 是下载文件的总时间.这些约束的含义是,总下载块之和应等于文件的大小,源节点上传带宽之和不应超过下载文件节点的下载带宽,任何一个节点被请求上传的带宽不应超过它的上传带宽.

1.2 下载节点最优策略

由于我们假定节点的上传带宽不变,因此上述约束可以简化为

$$\sum_{j\in N_i} O_{i,j}=O_i,\ \forall i\in I \tag{4}$$

$$\sum_{j\in N_i} B_{u,ji}\le B_{d,i},\ \forall i\in I \tag{5}$$

$$B_{u,ji}\le B_{u,j},\forall j\in N_i \tag{6}$$

我们定义下载节点的代价函数,分析它是如何影响节点的分块行为的.当节点 i 得到 N_i 时,它会发送下载分块的请求.此时,节点 i 必须决定如何分块,是否利用 N_i 中所有节点或者部分节点.在整个策略空间中,一个极端是节点 i 利用 N_i 中所有节点,这样可以获得最大的总带宽资源从而使下载时间最小,但代价是节点 i 需要耗费更多的系统资源来管理线程(假定一个线程对应从一个源节点的下载);另一个极端是节点 i 只利用 N_i 中带宽最大的源节点,这样只需管理单个线程即可,但获得的带宽资源将只有单个节点的带宽,这无疑将增大下载时间.

为了定义每一个节点的代价函数,我们知道每一个节点都尽量使自己的下载越快越好,同时使管理线程所耗费的系统资源最小.因此,我们定义节点 i 的代价函数 u_i 为

$$u_i=\alpha\max_j t_{ij}+\beta\sum_j c\cdot t_{ij}=\alpha\max_j\left\{\frac{O_{i,j}}{B_{u,ij}}\right\}+\beta\sum_j c\cdot\frac{O_{i,j}}{B_{u,ij}},\ j\in N_i \tag{7}$$

其中,α,β是归一化因子,表示节点 i 如何评价下载快慢和管理线程所耗费系统资源的大小.其中,管理线程所耗费系统资源与线程数成正比,与线程的持续时间成正比;c 为单线程持续单位时间所需的系统资源.节点 i 的目标是使公式(7)最小.

命题 1. 对任意一个下载节点,当它从所有节点的下载时间相等时,其公式(7)的第 1 项最小.

证明:我们有:

$$t_{i,j}=\frac{O_{i,j}}{B_{u,ij}} \tag{8}$$

令所有节点的下载时间相等为 t_i,即

$$t_i=t_{i,j_1}=t_{i,j_2},\forall j_1,j_2\in N_i \tag{9}$$

我们要证明 t_i 为最优解.假定有一个更优的解 t_i^*,即 $t_i^*<t_i$.由定义和公式(9)可知,t_i^* 是下载时间最大的.由于所有节点不在同一时间完成下载,则至少有一个节点 m 在 t_i^* 之前完成下载.因此,我们有:

$$\frac{O_{i,m}}{B_{u,im}}<t_i^*\Rightarrow\frac{O_{i,m}}{B_{u,im}}<\frac{O_i^*}{B_{u,i}^*} \tag{10}$$

如果我们从 O_i^* 中取走足够小的一部分ε,利用节点 m 下载这一部分,并且令 $\frac{O_{i,m}+\varepsilon}{B_{u,im}}<\frac{O_i^*-\varepsilon}{B_{u,i}^*}$, 因为 $\frac{O_i^*-\varepsilon}{B_{u,i}^*}<$

$\frac{O_i^*}{B_{u,i}^*}$,故新的下载时间小于原下载时间.也就是说,我们找到了一个比 t_i^* 更优的解,而这与我们最初的假设相矛盾. □

由命题 1,对于令公式(7)第 1 项最小的最优解应有 $t_i = t_{i,j_1} = t_{i,j_2}$,故有:

$$\frac{O_{i,j_1}}{O_{i,j_2}} = \frac{B_{u,ij_1}}{B_{u,ij_2}^*}, \forall j_1, j_2 \in N_i \tag{11}$$

也就是说,节点 i 应按源节点上传带宽的比例分配从各个源节点下载文件分块的大小.即

$$O_{i,j_k} = \frac{B_{u,ij_k}}{\sum_{j_l \in N_i} B_{u,ij_l}} \cdot O_i, \forall j_k \in N_i \tag{12}$$

故下载时间为

$$t_i = t_{i,j_k} = \frac{O_{i,j_k}}{B_{u,ij_k}} = \frac{O_i}{\sum_{j_l \in N_i} B_{u,ij_l}}, \forall j_k \in N_i \tag{13}$$

由公式(12)可见,最优解等价于从上传带宽等于所有源节点上传带宽之和的单个节点下载文件.因此,从令公式(7)的第 1 项最小的角度来看,下载节点的集合应为 N_i 的全集.

命题 2. 对任一下载节点,当它的下载节点集合为下载带宽最大的源节点时,其公式(7)的第 2 项最小.

证明:公式(7)的第 2 项的大小等价于 $\sum_j \frac{O_{i,j}}{B_{u,ij}}$ $(j \in N_i)$的大小,当节点 i 的下载节点集合为下载带宽最大的源节点时,其公式(7)的第 2 项为

$$\frac{O_i}{\max_{j \in N_i}(B_{u,ij})} \tag{14}$$

假设节点 i 的下载节点集合为 $N_i' = \{j_{k_i} \in N_i, L \ge 2 \mid j_{k_1}, j_{k_2}, ..., j_{k_L}\}$,其公式(7)第 2 项为

$$\sum_{j_{k_l} \in N_i'} \frac{O_{i,j_{k_l}}}{B_{u,ij_{k_l}}} \tag{15}$$

我们把整个文件 O_i 分为与 $O_{i,j_{k_l}}$ 相等的 L 块,即 $O_i = \sum_{j_{k_l} \in N_i'} O_{i,j_{k_l}}$,则公式(15)等价于

$$\sum_{j_{k_l} \in N_i'} \frac{O_{i,j_{k_l}}}{\max_{j \in N_i}(B_{u,ij})} \tag{16}$$

由于集合 N_i' 的元素数 $L \ge 2$,则 N_i' 中至少有 1 个元素 j_{k^*} 使 $B_{u,ij_{k^*}} < \max_{j \in N_i}(B_{u,ij})$,则有:

$$\frac{O_{i,j_{k^*}}}{B_{u,ij_{k^*}}} > \frac{O_{i,j_{k^*}}}{\max_{j \in N_i}(B_{u,ij})} \tag{17}$$

故

$$\sum_{j_{k_l} \in N_i'} \frac{O_{i,j_{k_l}}}{B_{u,ij_{k_l}}} > \sum_{j_{k_l} \in N_i'} \frac{O_{i,j_{k_l}}}{\max_{j \in N_i}(B_{u,ij})} \tag{18}$$

即

$$\sum_{j_{k_l} \in N_i'} \frac{O_{i,j_{k_l}}}{B_{u,ij_{k_l}}} > \frac{O_i}{\max_{j \in N_i}(B_{u,ij})} \tag{19}$$

因此,从令公式(7)的第 2 项最小的角度来看,应只从下载带宽最大的源节点下载. □

由命题 1 可知,从令公式(7)的第 1 项最小的角度来看,下载节点的集合中源节点个数应为 N_i 中源节点的个

数;而由命题 2 可知,从令公式(7)的第 2 项最小的角度来看,下载节点的集合中源节点个数应为 1.因此,公式(7)的最优解应在下载源节点个数上取一定的折衷.

命题 3. 对任意一个下载节点,当它的下载节点集合中源节点的个数一定且为 n 时,令其公式(7)最小的最优下载节点集合是下载带宽最大的 n 个源节点.

证明:我们假设当节点 i 的下载节点集合中源节点的个数一定且为 n 时,最优的下载节点集合为 $N_{i,n}^*=\{j_1^*,j_2^*,...,j_n^*\}$ 且 $B_{u,ij_1^*}>B_{u,ij_2^*}>...>B_{u,ij_n^*}$,而下载带宽最大的 n 个源节点的集合为 $N_{i,n}=\{j_1,j_2,…,j_n\}$ 且 $B_{u,ij_1}>B_{u,ij_2}>...>B_{u,ij_n}$, $N_{i,n}^*\neq N_{i,n}$. 令 $N_{i,n}^*$ 对应的最优分块为 $O_{i,j,n}^*$, $\forall j\in N_{i,n}^*$,下面我们证明当下载节点集合是 $N_{i,n}$ 时的最优分块一定比 $O_{i,j,n}^*$ 要更优,即其对应的公式(7)的值比 $O_{i,j,n}^*$ 对应的公式(7)的值要小.

因为 $N_{i,n}^*\neq N_{i,n}$,则必有 $B_{u,ij_1}\geq B_{u,ij_1^*},B_{u,ij_2}\geq B_{u,ij_2^*},...,B_{u,ij_n}\geq B_{u,ij_n^*}$ 且不能同时取等号,则有:

$$\frac{O_{i,j_1^*,n}^*}{B_{u,ij_1}}\leq\frac{O_{i,j_1^*,n}^*}{B_{u,ij_1^*}},\frac{O_{i,j_2^*,n}^*}{B_{u,ij_2}}\leq\frac{O_{i,j_2^*,n}^*}{B_{u,ij_2^*}},...,\frac{O_{i,j_n^*,n}^*}{B_{u,ij_n}}\leq\frac{O_{i,j_n^*,n}^*}{B_{u,ij_n^*}} \tag{20}$$

且不能同时取等号.

故对于公式(7)的第 1 项,有:

$$\max_k\left\{\frac{O_{i,j_k^*,n}^*}{B_{u,ij_k}}\right\}\leq\max_k\left\{\frac{O_{i,j_k^*,n}^*}{B_{u,ij_k^*}}\right\} \tag{21}$$

对于公式(7)的第 2 项,有:

$$\sum_k\frac{O_{i,j_k^*,n}^*}{B_{u,ij_k}}<\sum_k\frac{O_{i,j_k^*,n}^*}{B_{u,ij_k^*}} \tag{22}$$

因此,对两项之和,有:

$$\max_k\left\{\frac{O_{i,j_k^*,n}^*}{B_{u,ij_k}}\right\}+\sum_k\frac{O_{i,j_k^*,n}^*}{B_{u,ij_k}}<\max_k\left\{\frac{O_{i,j_k^*,n}^*}{B_{u,ij_k^*}}\right\}+\sum_k\frac{O_{i,j_k^*,n}^*}{B_{u,ij_k^*}} \tag{23}$$

这意味着当下载节点集合是 $N_{i,n}$ 时,若其分块与 $O_{i,j,n}^*$ 相同,则其对应的公式(7)的值比 $O_{i,j,n}^*$ 对应的公式(7)的值都要小.显然,其最优分块必定也比 $O_{i,j,n}^*$ 对应的公式(7)的值要小. □

命题 4. 对任意一个下载节点,令其公式(7)最小的最优下载节点集合是 $N_{i,n}=\{j_1,j_2,…,j_n\}$(下载节点集合的节点数为 n)且 $B_{u,ij_1}>B_{u,ij_2}>...>B_{u,ij_n}$,则对应的最优分块 $O_{i,j,n}$,$\forall j\in N_{i,n}$,有 $\frac{O_{i,j_1,n}}{B_{u,ij_1}}\geq\frac{O_{i,j_2,n}}{B_{u,ij_2}}\geq...\geq\frac{O_{i,j_n,n}}{B_{u,ij_n}}$.

证明:我们假设最优下载节点集合是 $N_{i,n}$ 对应的最优分块 $O_{i,j,n},\forall j\in N_{i,n},\exists j',j''\in N_{i,n}$,有 $B_{u,ij'}>B_{u,ij''}$ 且 $\frac{O_{i,j',n}}{B_{u,ij'}}<\frac{O_{i,j'',n}}{B_{u,ij''}}$.我们要证明这必然不是最优分块,只需证明存在另一种分块比它更优即可.

我们从 $O_{i,j'',n}$ 中取出一部分 ε 加到 $O_{i,j',n}$,即此时 j' 的分块为 $O_{i,j',n}+\varepsilon$,j'' 的分块为 $O_{i,j'',n}-\varepsilon$,并且令 $\frac{O_{i,j',n}+\varepsilon}{B_{u,ij'}}\leq\frac{O_{i,j'',n}-\varepsilon}{B_{u,ij''}}$ (只需令 ε 足够小即可),而其他分块保持不变.我们考察公式(7)值的变化,此时有:

$$\max\left(\frac{O_{i,j',n}+\varepsilon}{B_{u,ij'}},\frac{O_{i,j'',n}-\varepsilon}{B_{u,ij''}}\right)=\frac{O_{i,j'',n}-\varepsilon}{B_{u,ij''}} \tag{24}$$

而对于原分块,有:

$$\max\left(\frac{O_{i,j',n}}{B_{u,ij'}},\frac{O_{i,j'',n}}{B_{u,ij''}}\right)=\frac{O_{i,j'',n}}{B_{u,ij''}} \tag{25}$$

故

$$\max\left(\frac{O_{i,j',n}+\varepsilon}{B_{u,ij'}},\frac{O_{i,j'',n}-\varepsilon}{B_{u,ij''}}\right)<\max\left(\frac{O_{i,j',n}}{B_{u,ij'}},\frac{O_{i,j'',n}}{B_{u,ij''}}\right) \tag{26}$$

我们先考察公式(7)的第 1 项,原分块第 1 项为 $\max\limits_{j\in N_{i,n}}\left\{\frac{O_{i,j,n}}{B_{u,ij}}\right\}=\max\left\{\max\limits_{j\in N_{i,n},j\neq j',j''}\left\{\frac{O_{i,j,n}}{B_{u,ij}}\right\}+\max\left\{\frac{O_{i,j',n}}{B_{u,ij'}},\frac{O_{i,j'',n}}{B_{u,ij''}}\right\}\right\}$,从 $O_{i,j'',n}$ 中取出一部分ε加到 $O_{i,j',n}$ 后,公式(7)的第 1 项为

$$\max_{j\in N_{i,n}}\left\{\frac{O_{i,j,n}}{B_{u,ij}}\right\}=\max\left\{\max_{j\in N_{i,n},j\neq j',j''}\left\{\frac{O_{i,j,n}}{B_{u,ij}}\right\}+\max\left\{\frac{O_{i,j',n}+\varepsilon}{B_{u,ij'}},\frac{O_{i,j'',n}-\varepsilon}{B_{u,ij''}}\right\}\right\} \tag{27}$$

由不等式(26)得:

$$\max\left\{\max_{j\in N_{i,n},j\neq j',j''}\left\{\frac{O_{i,j,n}}{B_{u,ij}}\right\}+\max\left\{\frac{O_{i,j',n}+\varepsilon}{B_{u,ij'}},\frac{O_{i,j'',n}-\varepsilon}{B_{u,ij''}}\right\}\right\}\le\max\left\{\max_{j\in N_{i,n},j\neq j',j''}\left\{\frac{O_{i,j,n}}{B_{u,ij}}\right\}+\max\left\{\frac{O_{i,j',n}}{B_{u,ij'}},\frac{O_{i,j'',n}}{B_{u,ij''}}\right\}\right\} \tag{28}$$

可知,从 $O_{i,j'',n}$ 中取出一部分ε加到 $O_{i,j',n}$ 后,公式(7)的第 1 项至少不会比原来大.

下面我们考察公式(7)的第 2 项.从 $O_{i,j'',n}$ 中取出一部分ε加到 $O_{i,j',n}$ 后,由于 $B_{u,ij'}>B_{u,ij''}$,故

$$\frac{O_{i,j',n}+\varepsilon}{B_{u,ij'}}+\frac{O_{i,j'',n}-\varepsilon}{B_{u,ij''}}<\frac{O_{i,j',n}}{B_{u,ij'}}+\frac{O_{i,j'',n}}{B_{u,ij''}} \tag{29}$$

改变分块后,公式(7)的第 2 项为

$$\sum_{j\in N_{i,n}}\frac{O_{i,j,n}}{B_{u,ij}}=\sum_{j\in N_{i,n},j\neq j',j''}\frac{O_{i,j,n}}{B_{u,ij}}+\frac{O_{i,j',n}+\varepsilon}{B_{u,ij'}}+\frac{O_{i,j'',n}-\varepsilon}{B_{u,ij''}} \tag{30}$$

原分块公式(7)的第 2 项为

$$\sum_{j\in N_{i,n}}\frac{O_{i,j,n}}{B_{u,ij}}=\sum_{j\in N_{i,n},j\neq j',j''}\frac{O_{i,j,n}}{B_{u,ij}}+\frac{O_{i,j',n}}{B_{u,ij'}}+\frac{O_{i,j'',n}}{B_{u,ij''}} \tag{31}$$

由不等式(29)可得:

$$\sum_{j\in N_{i,n},j\neq j',j''}\frac{O_{i,j,n}}{B_{u,ij}}+\frac{O_{i,j',n}+\varepsilon}{B_{u,ij'}}+\frac{O_{i,j'',n}-\varepsilon}{B_{u,ij''}}<\sum_{j\in N_{i,n},j\neq j',j''}\frac{O_{i,j,n}}{B_{u,ij}}+\frac{O_{i,j',n}}{B_{u,ij'}}+\frac{O_{i,j'',n}}{B_{u,ij''}} \tag{32}$$

可知,从 $O_{i,j'',n}$ 中取出一部分ε加到 $O_{i,j',n}$ 后,公式(7)的第 2 项比原来要小.

综上所述,从 $O_{i,j'',n}$ 中取出一部分ε加到 $O_{i,j',n}$ 后,两项之和即公式(7)的值将比原来小,即存在比原分块更优的分块,所以原分块不是最优分块. □

命题 5. 对任意一个下载节点,令其公式(7)最小的最优下载节点集合是 $N_{i,n}=\{j_1,j_2,\ldots,j_n\}$(下载节点集合的节点数为 n)且 $B_{u,ij_1}>B_{u,ij_2}>\ldots>B_{u,ij_n}$,则对应的最优分块 $O_{i,j,n}$,$\forall j\in N_{i,n}$,有 $\frac{O_{i,j_1,n}}{B_{u,ij_1}}=\frac{O_{i,j_2,n}}{B_{u,ij_2}}=\ldots=\frac{O_{i,j_{n-1},n}}{B_{u,ij_{n-1}}}\ge\frac{O_{i,j_n,n}}{B_{u,ij_n}}$.

命题 5 是命题 4 的推论,我们可以根据命题 4 证明命题 5.

证明:我们假设对于最优分块 $O_{i,j,n}$,$\forall j\in N_{i,n}$,$\exists j',j''\in N_{i,n}$,$j''\neq j_n$,使得 $B_{u,ij'}>B_{u,ij''}$,$O_{i,j',n}>O_{i,j'',n}$.我们只要证明存在比 $O_{i,j,n}$ 更优的分块即可.

由于 $j''\neq j_n$,则$\exists k\ge 0$,$B_{u,ij''}>B_{u,ij_{n-k}}>\ldots>B_{u,ij_{n-1}}>B_{u,ij_n}$,我们从 $O_{i,j_{n-k},n},\ldots,O_{i,j_{n-1},n},O_{i,j_n,n}$ 中任意一个或几个分块取一部分ε加到 $O_{i,j'',n}$,且保证 $\frac{O_{i,j'',n}+\varepsilon}{B_{u,ij''}}\le\frac{O_{i,j',n}}{B_{u,ij'}}$.实际上,由于 $B_{u,ij'}>B_{u,ij''}$,$O_{i,j',n}>O_{i,j'',n}$,则必有 $\frac{O_{i,j'',n}}{B_{u,ij''}}<\frac{O_{i,j',n}}{B_{u,ij'}}$,我们只要令$\varepsilon$足够小即可.

我们先考察公式(7)的第 1 项.由命题 4,原分块公式(7)的第 1 项为 $\max\limits_{j\in N_{i,n}}\left\{\frac{O_{i,j,n}}{B_{u,ij}}\right\}=\max\limits_{j\in\{j_1,j_2,\ldots,j'\}}\left\{\frac{O_{i,j,n}}{B_{u,ij}}\right\}$,改变分块后,$O_{i,j_{n-k},n},\ldots,O_{i,j_{n-1},n},O_{i,j_n,n}$ 中一个或几个分块减小,令此时的分块为 $O'_{i,j_{n-k},n},\ldots,O'_{i,j_{n-1},n},O'_{i,j_n,n}$,则 $O'_{i,j_l,n}\le O_{i,j_l,n}$,$O'_{i,j_l,n}-O_{i,j_l,n}=\varepsilon_{j_l}\ge 0$,$\sum\limits_{l}\varepsilon_l=\varepsilon$,$l=n-k,\ldots,n-1,n$.由命题 4,改变分块前有:

$$\max_{j\in\{j_{n-k},\dots,j_{n-1},j_n\}}\left\{\frac{O_{i,j,n}}{B_{u,ij}}\right\}\le\frac{O_{i,j'',n}}{B_{u,ij''}} \tag{33}$$

故改变分块后有:

$$\max_{j\in\{j_{n-k},\dots,j_{n-1},j_n\}}\left\{\frac{O'_{i,j,n}}{B_{u,ij}}\right\}<\frac{O_{i,j'',n}+\varepsilon}{B_{u,ij''}}\le\frac{O_{i,j',n}}{B_{u,ij'}}\le\max_{j\in\{j_1,j_2,\dots,j'\}}\left\{\frac{O_{i,j,n}}{B_{u,ij}}\right\} \tag{34}$$

改变分块后公式(7)的第 1 项为

$$\max_{j\in N_{i,n}}\left\{\frac{O_{i,j,n}}{B_{u,ij}}\right\}=\max\left\{\max_{j\in\{j_{n-k},\dots,j_{n-1},j_n\}}\left\{\frac{O'_{i,j,n}}{B_{u,ij}}\right\},\frac{O_{i,j'',n}+\varepsilon}{B_{u,ij''}},\max_{j\in\{j_1,j_2,\dots,j'\}}\left\{\frac{O_{i,j,n}}{B_{u,ij}}\right\}\right\}=\max_{j\in\{j_1,j_2,\dots,j'\}}\left\{\frac{O_{i,j,n}}{B_{u,ij}}\right\} \tag{35}$$

因此,改变分块前后,公式(7)的第 1 项不变.

下面我们考察公式(7)的第 2 项.改变分块前,

$$\sum_{j\in N_{i,n}}\frac{O_{i,j,n}}{B_{u,ij}}=\sum_{j\in\{j_1,j_2,\dots,j'\}}\frac{O_{i,j,n}}{B_{u,ij}}+\frac{O_{i,j'',n}}{B_{u,ij''}}+\sum_{j\in\{j_{n-k},\dots,j_{n-1},j_n\}}\frac{O_{i,j,n}}{B_{u,ij}} \tag{36}$$

改变分块后公式(7)的第 2 项为

$$\begin{aligned}\sum_{j\in N_{i,n}}\frac{O_{i,j,n}}{B_{u,ij}}&=\sum_{j\in\{j_1,j_2,\dots,j'\}}\frac{O_{i,j,n}}{B_{u,ij}}+\frac{O_{i,j'',n}+\varepsilon}{B_{u,ij''}}+\sum_{j\in\{j_{n-k},\dots,j_{n-1},j_n\}}\frac{O_{i,j,n}-\varepsilon_j}{B_{u,ij}}\\&=\sum_{j\in\{j_1,j_2,\dots,j'\}}\frac{O_{i,j,n}}{B_{u,ij}}+\frac{O_{i,j'',n}}{B_{u,ij''}}+\sum_{j\in\{j_{n-k},\dots,j_{n-1},j_n\}}\frac{O_{i,j,n}}{B_{u,ij}}+\frac{\varepsilon}{B_{u,ij''}}-\sum_{j\in\{j_{n-k},\dots,j_{n-1},j_n\}}\frac{\varepsilon_j}{B_{u,ij}}\end{aligned} \tag{37}$$

而由于 $B_{u,ij_{n-k}}>\dots>B_{u,ij_{n-1}}>B_{u,ij_n}$, 故

$$\frac{\varepsilon}{B_{u,ij''}}-\sum_{j\in\{j_{n-k},\dots,j_{n-1},j_n\}}\frac{\varepsilon_j}{B_{u,ij}}<\frac{\varepsilon}{B_{u,ij''}}-\sum_{j\in\{j_{n-k},\dots,j_{n-1},j_n\}}\frac{\varepsilon_j}{B_{u,ij_{n-k}}}=\frac{\varepsilon}{B_{u,ij''}}-\frac{\sum_{j\in\{j_{n-k},\dots,j_{n-1},j_n\}}\varepsilon_j}{B_{u,ij_{n-k}}}=\frac{\varepsilon}{B_{u,ij''}}-\frac{\varepsilon}{B_{u,ij_{n-k}}} \tag{38}$$

而由于 $B_{u,ij''}>B_{u,ij_{n-k}}$,故

$$\frac{\varepsilon}{B_{u,ij''}}-\sum_{j\in\{j_{n-k},\dots,j_{n-1},j_n\}}\frac{\varepsilon_j}{B_{u,ij}}<0 \tag{39}$$

故改变分块后公式(7)的第 2 项减小.综上所述,由两项之和可知,改变分块后,公式(7)比原分块变小,即存在比 $O_{i,j,n}$ 更优的分块. □

命题 6. 对任意一个下载节点,令其公式(7)最小的最优下载节点集合是 $N_{i,n}=\{j_1,j_2,\dots,j_n\}$(下载节点集合的节点数为 n)且 $B_{u,ij_1}>B_{u,ij_2}>\dots>B_{u,ij_n}$, 则对应的最优分块 $O_{i,j,n},\forall j\in N_{i,n}$,有 $\frac{O_{i,j_1,n}}{B_{u,ij_1}}=\frac{O_{i,j_2,n}}{B_{u,ij_2}}=\dots=\frac{O_{i,j_n,n}}{B_{u,ij_n}}$.

命题 6 是命题 5 的推论,我们可根据命题 5 证明命题 6.

证明 1:由命题 5,对最优分块 $O_{i,j,n},\forall j\in N_{i,n}$,有 $\frac{O_{i,j_1,n}}{B_{u,ij_1}}=\frac{O_{i,j_2,n}}{B_{u,ij_2}}=\dots=\frac{O_{i,j_{n-1},n}}{B_{u,ij_{n-1}}}\ge\frac{O_{i,j_n,n}}{B_{u,ij_n}}$.当最后的大于等于号取等号时,即为命题 6.因此,我们只需证明 $\frac{O_{i,j_1,n}}{B_{u,ij_1}}=\frac{O_{i,j_2,n}}{B_{u,ij_2}}=\dots=\frac{O_{i,j_{n-1},n}}{B_{u,ij_{n-1}}}>\frac{O_{i,j_n,n}}{B_{u,ij_n}}$ 不是最优解即可.

当 $\frac{O_{i,j_1,n}}{B_{u,ij_1}}=\frac{O_{i,j_2,n}}{B_{u,ij_2}}=\dots=\frac{O_{i,j_{n-1},n}}{B_{u,ij_{n-1}}}>\frac{O_{i,j_n,n}}{B_{u,ij_n}}$ 时,我们可以把 O_i 分为两部分: O'_i 和 O''_i. O'_i 为

$$O'_i=\sum_{j\in\{j_1,j_2,\dots,j_{n-1}\}}O'_{i,j,n} \tag{40}$$

其中,对 $O'_{i,j,n}$ 有:

$$\frac{O'_{i,j_1,n}}{B_{u,ij_1}}=\frac{O'_{i,j_2,n}}{B_{u,ij_2}}=\dots=\frac{O'_{i,j_{n-1},n}}{B_{u,ij_{n-1}}}=\frac{O_{i,j_n,n}}{B_{u,ij_n}} \tag{41}$$

O_i'' 为

$$O_i'' = \sum_{j \in \{j_1, j_2, \ldots, j_{n-1}, j_n\}} O_{i,j,n}'' \tag{42}$$

其中,对 $O_{i,j,n}''$ 有:

$$\frac{O_{i,j_1,n}''}{B_{u,ij_1}} = \frac{O_{i,j_2,n}''}{B_{u,ij_2}} = \ldots = \frac{O_{i,j_{n-1},n}''}{B_{u,ij_{n-1}}} \tag{43}$$

且对 $O_{i,j,n}'$ 与 $O_{i,j,n}''$ 之和,有:

$$O_{i,j,n}' + O_{i,j,n}'' = O_{i,j,n}, j \in \{j_1, j_2, \ldots, j_{n-1}\} \tag{44}$$

我们把原优化问题分解为两部分,分别考虑对于公式(41)和公式(43)的最优解.实际上,这可以看作是分别对文件 $O_i^1 = O_i' + O_{i,j_n,n}$ 和文件 O_i'' 的最优下载方案.不妨先考虑对于公式(41)的最优解.由命题 3 可知,最优下载节点的集合可以简化为确定下载节点的个数.已知下载节点间分块也已确定,因此,问题可以简化为确定下载节点的个数.

设下载带宽最大的 n 个源节点的集合为 $N_{i,n}=\{j_1,j_2,\ldots,j_n\}$,则有:

$$\frac{O_{i,j_1,n}'}{B_{u,ij_1}} = \frac{O_{i,j_2,n}'}{B_{u,ij_2}} = \ldots = \frac{O_{i,j_{n-1},n}'}{B_{u,ij_{n-1}}} = \frac{O_{i,j_n,n}}{B_{u,ij_n}} = \frac{\sum_{j \in \{j_1, j_2, \ldots, j_{n-1}\}} O_{i,j,n}' + O_{i,j_n,n}}{\sum_{j \in N_{i,n}} B_{u,ij}} = \frac{O_i' + O_{i,j_n,n}}{\sum_{j \in N_{i,n}} B_{u,ij}} = \frac{O_i^1}{B_{u,i}^n} \tag{45}$$

其中, $B_{u,i}^n = \sum_{j \in N_{i,n}} B_{u,ij}$, 即为下载带宽最大的 n 个源节点带宽之和.

公式(45)的含义为,每个源节点的下载时间等价于原文件从一个带宽等于 n 个源节点带宽之和的源节点下载的时间.节点 i 的代价函数 u_i 为

$$u_i = \alpha \frac{O_i^1}{B_{u,i}^n} + \beta \cdot c \cdot n \cdot \frac{O_i^1}{B_{u,i}^n} = \frac{\alpha + \beta cn}{B_{u,i}^n} \cdot O_i^1 \tag{46}$$

由公式(46)可知,在下载文件给定的情况下,u_i 只与下载节点的个数有关.而下载节点的个数取决于 $f(n)$的大小. $f(n)$为

$$f(n) = \frac{\alpha + \beta cn}{B_{u,i}^n} \tag{47}$$

$f(n)$最小的节点个数 n 即为最优解.我们注意到,$f(n)$的值与下载文件的大小无关,即对文件 O_i^1 最优的节点个数 n,对于文件 O_i'' 同样为最优.而对于原假设,若对文件 O_i^1 最优的节点个数为 n,则对于文件 O_i'' 最优的节点个数为 $n-1$,故原假设不成立. □

证明 2:对最优分块 $O_{i,j,n}, \forall j \in N_{i,n}$,同证明 1,我们只需证明 $\frac{O_{i,j_1,n}}{B_{u,ij_1}} = \frac{O_{i,j_2,n}}{B_{u,ij_2}} = \ldots = \frac{O_{i,j_{n-1},n}}{B_{u,ij_{n-1}}} > \frac{O_{i,j_n,n}}{B_{u,ij_n}}$ 不是最优解即可.

我们考察当 $\frac{O_{i,j_1,n}}{B_{u,ij_1}} = \frac{O_{i,j_2,n}}{B_{u,ij_2}} = \ldots = \frac{O_{i,j_{n-1},n}}{B_{u,ij_{n-1}}} > \frac{O_{i,j_n,n}}{B_{u,ij_n}}$ 时, $O_{i,j_n,n}$ 的变化对节点 i 的代价函数 u_i的影响.令 $O_{i,j_n,n}$ 变化一个足够小的量ε,相应地, $O_{i,j_1,n}, O_{i,j_2,n}, \ldots, O_{i,j_{n-1},n}$ 也分别变化一个量 $\varepsilon_{i,j_1,n}, \varepsilon_{i,j_2,n}, \ldots, \varepsilon_{i,j_{n-1},n}$,且 $\sum_{k=1}^{n-1} \varepsilon_{i,j_k,n} = \varepsilon$.由命题 3 可知,此时有:

$$\frac{O_{i,j_1,n} + \varepsilon_{i,j_1,n}}{B_{u,ij_1}} = \frac{O_{i,j_2,n} + \varepsilon_{i,j_2,n}}{B_{u,ij_2}} = \ldots = \frac{O_{i,j_{n-1},n} + \varepsilon_{i,j_{n-1},n}}{B_{u,ij_{n-1}}} > \frac{O_{i,j_n,n} - \varepsilon}{B_{u,ij_n}} \tag{48}$$

即对于 $\varepsilon_{i,j_1,n}, \varepsilon_{i,j_2,n}, \ldots, \varepsilon_{i,j_{n-1},n}$,有:

$$\frac{\varepsilon_{i,j_1,n}}{B_{u,ij_1}}=\frac{\varepsilon_{i,j_2,n}}{B_{u,ij_2}}=\ldots=\frac{\varepsilon_{i,j_{n-1},n}}{B_{u,ij_{n-1}}}=\frac{\sum_{k=1}^{n-1}\varepsilon_{i,j_k,n}}{\sum_{k=1}^{n-1}B_{u,ij_k}}=\frac{\varepsilon}{B_{u,i}^{n-1}} \tag{49}$$

我们考察代价函数的变化,第 1 项的增量为 $\alpha\cdot\frac{\varepsilon}{B_{u,i}^{n-1}}$, 第 2 项的增量为 $\beta\cdot c\cdot\left[\frac{(n-1)\varepsilon}{B_{u,i}^{n-1}}-\frac{\varepsilon}{B_{u,ij_n}}\right]$. 因此,代价函数的增量$\Delta u_i$为

$$\Delta u_i=\alpha\cdot\frac{\varepsilon}{B_{u,i}^{n-1}}+\beta\cdot c\cdot\left[\frac{(n-1)\varepsilon}{B_{u,i}^{n-1}}-\frac{\varepsilon}{B_{u,ij_n}}\right]=\left[\frac{\alpha+\beta c(n-1)}{B_{u,i}^{n-1}}-\frac{\beta c}{B_{u,ij_n}}\right]\varepsilon \tag{50}$$

令 $g(n)$为下载节点个数 n 的函数,有:

$$g(n)=\frac{\alpha+\beta c(n-1)}{B_{u,i}^{n-1}}-\frac{\beta c}{B_{u,ij_n}} \tag{51}$$

$g(n)$的正负决定了ε对代价函数的影响.当 $g(n)>0,\varepsilon<0$ 时,$\Delta u_i<0$,即代价函数变小.这说明当 $O_{i,j_n,n}$ 减小,相应地 $O_{i,j_1,n},O_{i,j_2,n},\ldots,O_{i,j_{n-1},n}$ 增加时,得到的解更优,原方案 $\frac{O_{i,j_1,n}}{B_{u,ij_1}}=\frac{O_{i,j_2,n}}{B_{u,ij_2}}=\ldots=\frac{O_{i,j_{n-1},n}}{B_{u,ij_{n-1}}}>\frac{O_{i,j_n,n}}{B_{u,ij_n}}$ 不是最优解.实际上,只要 $O_{i,j_n,n}>0$,这种减小的过程可以一直持续下去,不断得到更优的解,直到 $O_{i,j_n,n}=0$ 为止.此时就是命题 5 全取等号的情形,只是下载节点个数变为 $n-1$.

类似地,当 $g(n)<0,\varepsilon>0$ 时,$\Delta u_i<0$,即代价函数变小.这说明 $O_{i,j_n,n}$ 增加,相应地,在 $O_{i,j_1,n},O_{i,j_2,n},\ldots,O_{i,j_{n-1},n}$ 减小时,得到的解更优,原方案 $\frac{O_{i,j_1,n}}{B_{u,ij_1}}=\frac{O_{i,j_2,n}}{B_{u,ij_2}}=\ldots=\frac{O_{i,j_{n-1},n}}{B_{u,ij_{n-1}}}>\frac{O_{i,j_n,n}}{B_{u,ij_n}}$ 不是最优解.

实际上,只要 $\frac{O_{i,j_1,n}}{B_{u,ij_1}}=\frac{O_{i,j_2,n}}{B_{u,ij_2}}=\ldots=\frac{O_{i,j_{n-1},n}}{B_{u,ij_{n-1}}}>\frac{O_{i,j_n,n}}{B_{u,ij_n}}$,这种减小的过程可以一直持续,不断得到更优的解,直到 $\frac{O_{i,j_1,n}}{B_{u,ij_1}}=\frac{O_{i,j_2,n}}{B_{u,ij_2}}=\ldots=\frac{O_{i,j_{n-1},n}}{B_{u,ij_{n-1}}}=\frac{O_{i,j_n,n}}{B_{u,ij_n}}$ 为止.此时就是命题 5 全取等号的情形.

上面没有考虑的一个特例是 $g(n)=0$ 的情况.此时,无论ε的取值正负,都有$\Delta u_i=0$,即代价函数不变.ε变化的下限是令 $O_{i,j_n,n}=0$, 此时就是命题 5 全取等号的情形,只是下载节点个数变为 $n-1$.ε变化的上限是令 $\frac{O_{i,j_1,n}}{B_{u,ij_1}}=\frac{O_{i,j_2,n}}{B_{u,ij_2}}=\ldots=\frac{O_{i,j_{n-1},n}}{B_{u,ij_{n-1}}}=\frac{O_{i,j_n,n}}{B_{u,ij_n}}$,此时也是命题 5 全取等号的情形,下载节点个数变为 n.ε在此范围内变化,代价函数始终保持不变.

综上所述,若 $g(n)=0$,则ε分别取下限和上限,下载节点个数分别为 $n-1$ 和 n,代价函数相同.按证明 1,下载节点个数分别为 $n-1$ 和 n,代价函数相同,这意味着 $f(n-1)=f(n)$.因此,下面我们看 $g(n)=0$ 与 $f(n-1)=f(n)$是否等价.由 $f(n-1)=f(n)$有:

$$\frac{\alpha+\beta c(n-1)}{B_{u,i}^{n-1}}=\frac{\alpha+\beta cn}{B_{u,i}^{n}} \tag{52}$$

又有

$$B_{u,i}^{n}=B_{u,i}^{n-1}+B_{u,ij_n} \tag{53}$$

故有

$$\frac{\alpha+\beta c(n-1)}{B_{u,i}^{n-1}}=\frac{\alpha+\beta cn}{B_{u,i}^{n-1}+B_{u,ij_n}}=\frac{\beta c}{B_{u,ij_n}} \tag{54}$$

因此有 $g(n)=0$.同理,我们可由 $g(n)=0$ 推得 $f(n-1)=f(n)$.因此,$g(n)=0$ 和 $f(n-1)=f(n)$等价. □

我们自然会想到下面是否存在 3 个下载节点个数等价的情形.下面我们来证明 $f(n-k)=\ldots=f(n-1)=f(n),k>1$ 的情况并不存在.

命题 7. 若 $f(n-k)=\ldots=f(n-1)=f(n)$,则 $k=1$.

证明:我们假设 $k\geq 2$,则 $f(n-k)=\ldots=f(n-2)=f(n-1)=f(n)$,我们只要证明后 3 项之间的连等关系不存在即可.若 $f(n-2)=f(n-1)=f(n)$,则对于 $f(n-1)=f(n)$有公式(54),对于 $f(n-2)=f(n-1)$有:

$$\frac{\alpha+\beta c(n-2)}{B_{u,i}^{n-2}}=\frac{\alpha+\beta c(n-1)}{B_{u,i}^{n-1}}=\frac{\beta c}{B_{u,ij_{n-1}}} \tag{55}$$

由公式(54)和公式(55),有:

$$B_{u,ij_n}=B_{u,ij_{n-1}} \tag{56}$$

这与 $B_{u,ij_n}>B_{u,ij_{n-1}}$ 矛盾. □

综上所述,下载节点令其代价函数最小的最优策略是首先计算 $f(n)$,得到令 $f(n)$最小的节点个数 n^*,下载节点集合为 n^*个下载带宽最大的节点.分块策略是按下载节点带宽大小的比例在下载节点之间分块,即对下载节点 i,其最优下载策略可分为 3 部分:(1) 节点 i 得到大小为 N 的可用源节点集合,并向所有源节点发送带宽测试文件,由下载时间估算各源节点的可用下载带宽;(2) 计算源节点个数 n 对应的 $f(n)$,得到令 $f(n)$最小的节点个数 n^*;(3) 按照 n^*个带宽最大的下载节点的带宽比例在下载节点之间分块.

1.3 考虑源节点策略的下载节点最优策略

设有 n 个节点 $j_1,j_2,\ldots,j_n$ 向源节点 i 发出下载请求,请求下载的分块大小分别为 $O_{j_1,i},O_{j_2,i},\ldots,O_{j_n,i}$.源节点 i 的上传带宽为 $B_{u,i}$,其策略空间是分配给 n 个下载节点 $j_1,j_2,\ldots,j_n$ 的上传带宽: $B_{u,j_1i},B_{u,j_2i},\ldots,B_{u,j_ni}$,且 $\sum_{k=1}^{n}B_{u,j_ki}=B_{u,i}$.为了保证对不同下载节点的公平性,我们假定源节点的策略是使各个下载节点的下载时间相等,即令

$$\frac{O_{j_1,i}}{B_{u,j_1i}}=\frac{O_{j_2,i}}{B_{u,j_2i}}=\ldots=\frac{O_{j_n,i}}{B_{u,j_ni}} \tag{57}$$

我们令 $O_{-j_k,i}$ 为除了 j_k 节点以外,其他节点向源节点 i 发出下载请求块之和,即

$$O_{-j_k,i}=\sum_{l\neq k}O_{j_l,i} \tag{58}$$

则 j_k 节点从源节点 i 得到的下载带宽 B_{u,j_ki} 为

$$B_{u,j_ki}=\frac{O_{j_k,i}}{O_{j_k,i}+O_{j_{-k},i}}\cdot B_{u,i} \tag{59}$$

前述的下载节点最优策略假定节点的下载带宽与对应的分块大小无关,而下载带宽实际上是分块大小的函数,下面我们讨论此时下载节点的最优策略.

由公式(57)可知,下载节点 j_k 从源节点 i 的下载时间为

$$t_{j_k,i}=\frac{O_{j_k,i}}{B_{u,j_ki}}=\frac{O_{j_k,i}+O_{-j_k,i}}{B_{u,i}} \tag{60}$$

对比公式(13),公式(60)中用源节点的总上传带宽取代了下载节点得到的下载带宽,分块大小增加了与下载节点竞争同一个源节点上传带宽的其他下载节点的请求块大小之和.

与命题 6 的证明类似,我们可以证明最优下载节点分块策略如命题 8 所述.

命题 8. 若源节点的策略为使各个下载节点的下载时间相等,对任意一个下载节点,令其公式(7)最小的最优下载节点集合是 $N_{i,n}=\{j_1,j_2,\ldots,j_n\}$(下载节点集合的节点数为 n)且 $B_{u,j_1}>B_{u,j_2}>\ldots>B_{u,j_n}$,则对应的最优分块 $O_{i,j,n},\forall j\in N_{i,n}$,有 $\frac{O_{i,j_1,n}+O_{-i,j_1}}{B_{u,j_1}}=\frac{O_{i,j_2,n}+O_{-i,j_2}}{B_{u,j_2}}=\ldots=\frac{O_{i,j_n,n}+O_{-i,j_n}}{B_{u,j_n}}$.

由命题 8,当下载节点集合的节点数为 n 时,节点 i 的下载时间 t_i' 为

$$t_i' = \frac{O_i + \sum_{k=1}^{n} O_{-i,j_k}}{\sum_{k=1}^{n} B_{u,j_k}} \tag{61}$$

节点 i 的代价函数 u_i' 为

$$u_i' = (\alpha + \beta cn)t_i' = \frac{(\alpha + \beta cn)\left(O_i + \sum_{k=1}^{n} O_{-i,j_k}\right)}{\sum_{k=1}^{n} B_{u,j_k}} \tag{62}$$

因此,为确定下载节点集合的个数 n,比较 n 对应的 u_i',最优的 n'^* 应使 u_i' 最小.

综上所述,在源节点公平分配带宽的前提下,下载节点令其代价函数最小的最优策略是首先计算 $u_i'(n)$,得到令 $u_i'(n)$ 最小的节点个数 n'^*,下载节点集合为 n'^* 个总下载带宽最大的源节点.分块策略应使源节点下载总分块大小与源节点的总下载带宽成比例,即对下载节点 i,其最优下载策略可分为 3 部分:(1) 节点 i 得到大小为 N 的可用源节点集合,并向所有源节点发送带宽测试文件,由下载时间估算各源节点的可用总下载带宽;(2) 计算源节点个数 n 对应的 $u_i'(n)$,得到令 $u_i'(n)$ 最小的节点个数 n'^*;(3) 按照使 n'^* 个总下载带宽最大的源节点的下载总分块大小与其总下载带宽成比例的方法在下载节点之间分块.

2 仿真结果

为了验证并行下载策略的有效性,我们在仿真中令 $\alpha=1,\beta c=0.2$,源节点总数为 $n=100$,源节点带宽为在区间[0,100]内按均匀分布随机生成的 100 个带宽,文件的总大小为 $O=1000$.首先,我们比较最优并行下载策略和利用所有源节点的按比例分块策略,考察目标函数随源节点个数的变化.为了比较目标函数中下载时间和系统消耗随源节点个数的变化,我们同时也考察下载时间和系统消耗分别按 α 和 βc 加权后的值,如图 1 所示.由图 1 可知,系统消耗随源节点个数的增加单调递增,而下载时间随源节点个数的增加单调递减,综合考虑二者最优的节点个数为 27.而按比例分块策略的源节点个数为 100,与之相比,最优并行下载策略的下载时间增加很小,而系统消耗则大为降低.我们再比较最优并行下载策略与平均分块策略,考察目标函数随源节点个数的变化,如图 2 所示.由图 2 可知,受少数低带宽节点的制约,增加源节点的个数并不能显著减少平均分块的下载时间.无论就系统消耗还是下载时间而言,平均分块都远远逊于最优并行下载策略,同时也不如按比例分块策略.

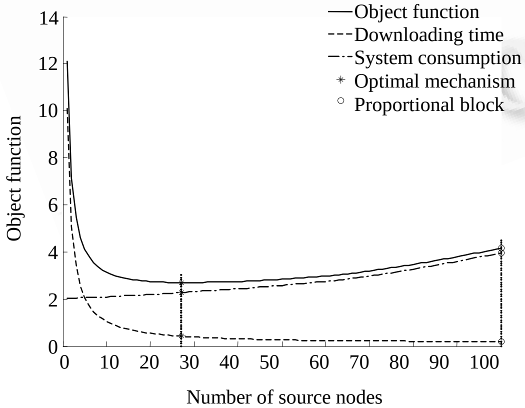


Fig.1 Optimal mechanism and proportional block

图 1 最优机制和按比例分块

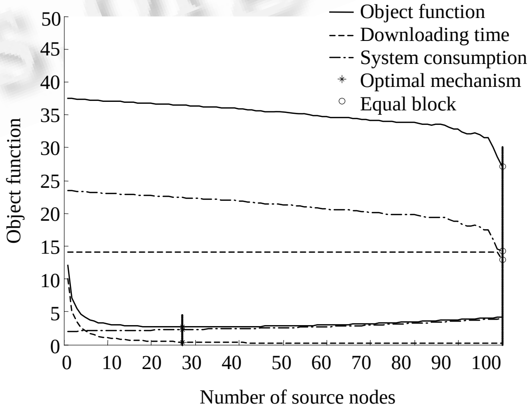


Fig.2 Optimal mechanism and equal block

图 2 最优机制和平均分块

我们令从各源节点处下载的其他下载节点请求总分块大小为在区间$\left[0,2\cdot\frac{O}{n}\right]$均匀分布的 100 个随机值,即其均值为$\frac{O}{n}$(均匀分块大小),其他参数同上.在源节点采用公平策略的情况下,首先我们比较最优并行下载策略和利用所有源节点的按比例分块策略,如图3所示.此时,最优的节点个数为18.我们再比较最优并行下载策略与平均分块策略,如图4所示.

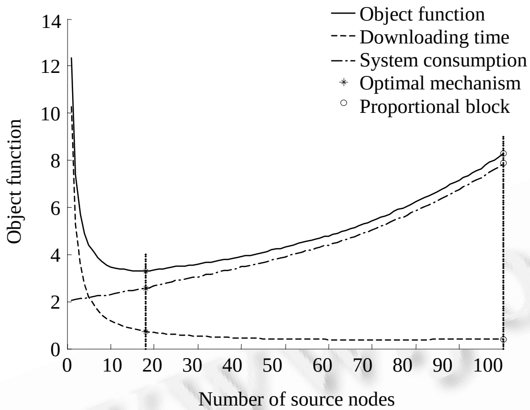


Fig.3 Optimal mechanism and proportional block (considering source node)

图3 最优机制和按比例分块(考虑源节点)

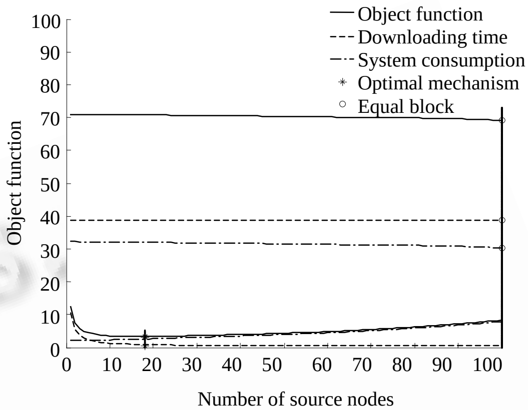


Fig.4 Optimal mechanism and equal block (considering source node)

图4 最优机制和平均分块(考虑源节点)

我们令βc=0.5,其他参数同上,图5和图6分别为此时不考虑源节点策略和考虑源节点策略的最优并行下载策略和按比例分块策略的比较.此时,最优节点个数分别为 20 和 13.由于βc 相对于α变大,系统消耗的权重增大,最优节点个数应减小.

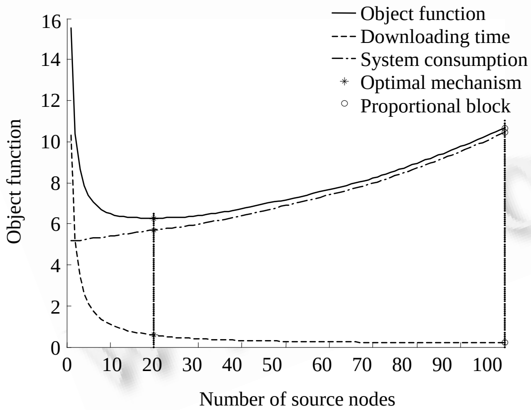


Fig.5 Optimal mechanism and proportional block (βc=0.5)

图5 最优机制和按比例分块(βc=0.5)

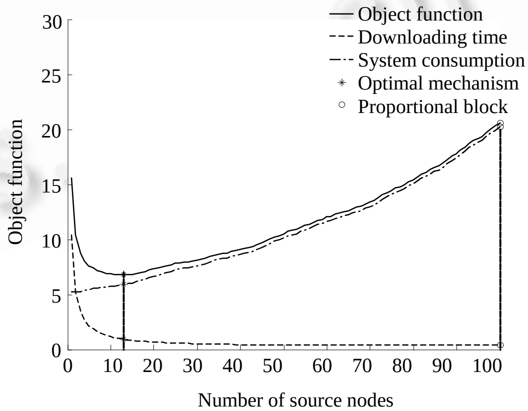


Fig.6 Optimal mechanism and proportional block (considering source node, βc=0.5)

图6 最优机制和按比例分块(考虑源节点,βc=0.5)

3 结束语

并行下载可以大幅度提高对等系统的下载速率和下载的稳健性,但现有对等系统的并行下载机制都没有在理论上按最优策略进行优化.本文从下载节点效益的角度出发,分析了同时优化下载时间和系统消耗的并行

下载策略,为优化并行下载机制提供了理论基础.由于实际网络中的上传、下载带宽是时变的,并行下载机制应在一定的代价效率下适应带宽的变化,这将是进一步研究的内容.

致谢 在此,我们向对本文的工作给予支持和建议的同行,尤其是中国科学院计算机网络信息中心 CNNIC 实验室的老师和同学表示感谢.

References:

[1] eMule. http://www.emule.org/

[2] KaZaA. http://www.kazaa.com/

[3] Lime wire. http://www.limewire.com/

[4] Overnet. http://www.overnet.com/

[5] Bernsteing DS, Feng ZZ, Levine BN, Zilberstein S. Adaptive peer selection. In: Kaashoek F, Stoica I, eds. Proc. of the 2nd Int’l Workshop on Peer-to-Peer Systems. LNCS 2735, Heidelberg: Springer-Verlag, 2003. 237−246.

[6] Byers J, Considine J, Mitzenmacher M, Rost S. Informed content delivery across adaptive overlay networks. IEEE/ACM Trans. on Networking, 2004,12(5):767−780.

[7] Fuqua AC, Ngan TWJ, Wallach DS. Economic behavior of peer-to-peer storage networks. In: Proc. of the Workshop on Economics of Peer-to-Peer Systems. 2003. http://www2.sims.berkeley.edu/research/conferences/p2pecon/index.html

[8] Bolosky WJ, Douceur JR, Ely D, Theimer M. Feasibility of a serverless distributed file system deployed on an existing set of desktop PCs. In: Alexandre B, Jim K, Phillippe N, eds. Proc. of the ACM SIGMETRICS 2000: Conf. on Measurement and Modeling of Computer Systems. New York: ACM Press, 2000. 34−43.

[9] Kubiatowicz J, Bindel D, Chen Y, Czerwinski S, Eaton P, Geels D, Gummadi R, Rhea S, Eaton P, Geels D, Gummadi R, Rhea S, Weatherspoon H, Weimer W, Wells C, Zhao B. OceanStore: An architecture for global-scale persistent storage. In: Larry R, Anoop G, eds. Proc. of the 9th Int’l Conf. on Architectural Support for Programming Languages and Operating Systems. New York: ACM Press, 2000. 190−201.

[10] Rhea S, Wells C, Eaton P, Geels D, Zhao B, Weatherspoon H, Kubiatowicz J. Maintenance-Free global data storage. IEEE Internet Computing, 2001,5(5):40−49.

[11] Qiu DY, Srikant R. Modeling and performance analysis of BitTorrent-like peer-to-peer networks. In: Raj Y, Ellen Z, Jennifer R, eds. Proc. of the ACM SIGCOMM 2004: Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communications. New York: ACM Press, 2004. 367−378.

[12] Yang XY, De Veciana G. Service capacity of peer to peer networks. In: Victor OKL, Bo L, Marwan K, eds. Proc. of the INFOCOM 2004: The 23rd Annual Joint Conf. of the IEEE Computer and Communications Societies. Washington: IEEE Computer Society, 2004. 2242−2252.

[13] Adler M, Kumar R, Ross K, Rubenstein D, Turner D, Yao DD. Optimal peer selection in a free-market peer-resource economy. In: Proc. of the 2nd Workshop on Economics of Peer-to-Peer Systems. 2004. http://www.eecs.harvard.edu/p2pecon/program.html

[14] Funasaka J, Takemoto Y, Ishida K. Parallel downloading method using HTTP over UDP for high loss rate and delay networks. In: Kane K, Yoshiaki K, Eckhard M, Tsai WT, eds. Proc. of the 8th Int’l Symp. on Autonomous Decentralized Systems. Washington: IEEE Computer Society, 2007. 555−561.

[15] Kawano A, Funasaka J, Ishida K. Parallel downloading using variable length blocks for proxy servers. In: Sajal KD, Jiannong C, Cheng-Zhong X, Mohan K, eds. Proc. of the 27th Int’l Conf. on Distributed Computing Systems Workshops. Washington: IEEE Computer Society, 2007. 59.

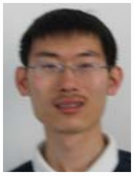
[16] Fu JM, Fan H. Assigning block size based on speculation for parallel downloading. In: Sunshin A, Daming W, eds. Proc. of the 6th IEEE Int’l Conf. on Computer and Information Technology. Washington: IEEE Computer Society, 2006. 119.

[17] Zhou X, Lu XL, Hou MS, Zhan C. A speed-based adaptive dynamic parallel downloading technique. Computer Science, 2005, 32(4):168−170 (in Chinese with English abstract).

[18] Higashi Y, Ata S, Oka I, Fujiwara C. Topology-Aware server selection method for dynamic parallel downloading. In: Naohisa O, Raouf B, eds. Proc. of the IEEE Consumer Communications and Networking Conf. Washington: IEEE Computer Society, 2005. 325−330.
[19] Song JT, Sha CF, Zhu H. Nash equilibria in parallel downloading with multiple clients. In: Ming TL, Yutaka M, eds. Proc. of the 24th Int'l Conf. on Distributed Computing Systems. Washington: IEEE Computer Society, 2004. 94−101.
[20] Dairaine L, Lancérica L, Lacan J. Enhancing peer to peer parallel data access with PeerFecT. In: Stiller B, Carle G, Karsten M, Reichl P, eds. Proc. of the 5th Int'l Workshop on Networked Group Communications and the 3rd Int'l Workshop on Internet Charging and QoS Technologies. LNCS 2816, Heidelberg: Springer-Verlag, 2003. 254−261.
[21] Lacan J, Lancérica L, Dairaine L. Speedup of data access using error correcting codes in peer-to-peer networks. In: Proc. of the 6th IEEE Int'l Conf. on Computer and Information Technology. Washington: IEEE Computer Society, 2003. 471−481.
[22] Lacan J, Lancérica L, Dairaine L. When FEC speed up data access in P2P networks. In: Boavida F, Monteiro E, Orvalho J, eds. Proc. of the Joint Int'l Workshops on Interactive Distributed Multimedia Systems and Protocols for Multimedia Systems: Protocols and Systems for Interactive Distributed Multimedia. LNCS 2515, Heidelberg: Springer-Verlag, 2002. 26−36.
[23] Pouwelse JA, Garbacki P, Epema DHJ, Sips HJ. A measurement study of the BitTorrent peer-to-peer file-sharing system. Technical Report, PDS-2004-007, Delft: Delft University of Technology, 2004.
[24] Emule. http://www.emule.org.cn/faq/doc/pref_connection.htm

附中文参考文献:

[17] 周旭,卢显良,侯孟书,詹川.一种速度自适应的动态并行下载技术.计算机科学,2005,32(4):168−170.



王正(1979－),男,山东胶南人,博士生,主要研究领域为下一代互联网.

阎保平(1950－),女,博士,研究员,博士生导师,CCF 高级会员,主要研究领域为下一代互联网,大规模科学数据共享技术.

罗万明(1973－),男,博士,副研究员,CCF 高级会员,主要研究领域为下一代互联网.