

HyperCompression[®] 低延迟深度 压缩技术白皮书



轻网科技
— LightWAN —

目 录

1. 数据压缩及字节缓存基本原理	2
2. 低延迟深度压缩技术的独特优势	4
2.1 通过深度压缩达到最大压缩比	4
2.2 超低延迟引入.....	4
2.3 字节缓存双向性	4
2.4 设备自动检测.....	5
3. 用户案例：企业总部和分支机构之间专线加速	5

数据压缩和字节缓存是广域网优化的关键技术，其目的是减少广域网链路的数据量，从而提高链路的有效吞吐。两项技术的一个共同点是都需要双边部署，由数据发送方对数据进行相应处理以减少数据量，并由接受方将数据还原。两项技术的另一个共同点是数据无损性：即接受端还原后的数据保证跟发送方处理前的数据完全一致。所以对通信双方的应用层，这一过程完全是透明的。

1. 数据压缩及字节缓存基本原理

数据压缩也被称作基本压缩，或无损压缩，一般采用 LZ 系列压缩算法。数据压缩具有自包含性：即对端解压缩方只根据数据包本身即可进行解压还原，不需要其它任何信息。压缩比因数据类型而异：文本数据压缩比最大，各种网页、Windows office 文件(Excel、word 等等)、PDF 其次，对多媒体和已压缩数据基本无效。字节缓存技术又叫“字典缓存”或“超级压缩”等名称。它通过缓存的方式在内存和硬盘中记录下流经的数据流，并以一定的大小（例如 32 字节、64 字节或 128 字节等）为最小单位建立易于查询的索引。当以后流向广域网链路的数据流出现了大于最小单位的相同数据时，可以将该数据替换成某个更短的符号。远端的设备能够将相应符号还原成原始数据。

字节缓存与数据压缩的不同之处是，字节缓存不具有自包含性。要解开压缩数据，对端设备必须引用以前记录的历史缓存信息。为达到这一目的，字节缓存技术需要连接两端的缓存完全同步，于是需要在每条广域网链路两端的加速设备之间建立缓存数据同步机制。对比于数据压缩，字节缓存的另一优势是即使对本身不具可压性的数据类型，如多媒体或已压缩数据，只要曾经传输过

相同或类似数据，字节缓存即可以提取出全部或部分相同数据，从而大幅缩减需要传输的数据量。字节缓存技术通过使用大量的缓存信息（包括该数据流本身以前的数据历史和其他数据流的数据历史）来压缩当前数据，有时能够达到非常高的压缩比。

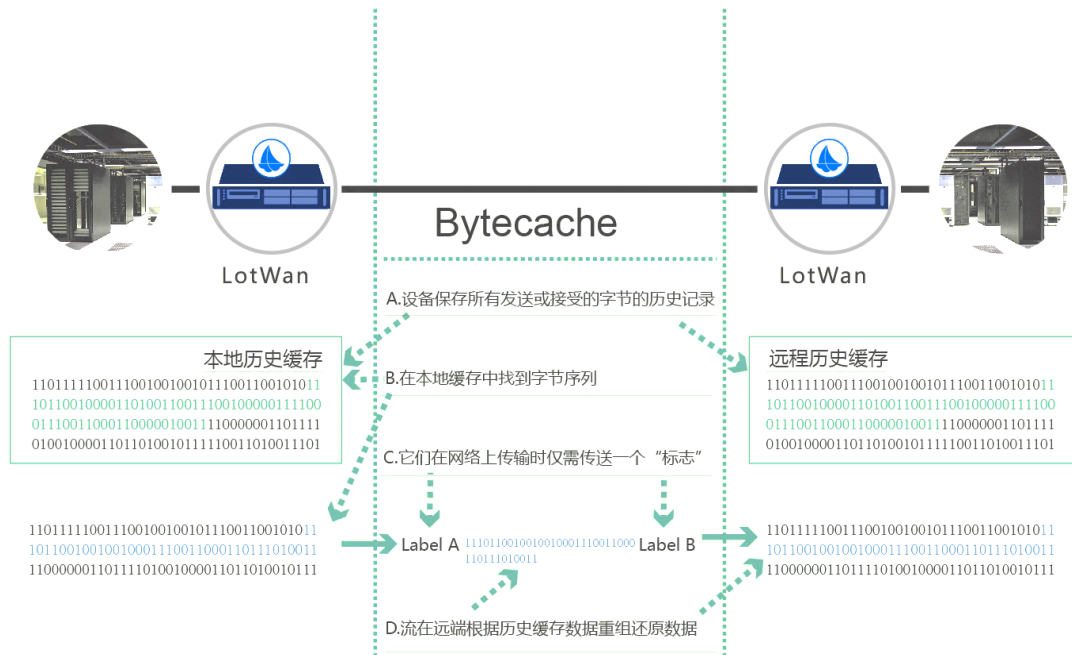


图 1.字节缓存工作原理示意图

2. 低延迟深度压缩技术的独特优势

尽管所有广域网厂商开发的数据压缩和字节缓存功能的基本工作原理类似，但不同的设计和实现收到的效果差别很大。低延迟深度压缩技术将数据压缩和字节缓存紧密融合，为用户带来最好的实际应用效果，相对于一般实现具有有以下优势：

2.1 通过深度压缩达到最大压缩比

首先使用全部缓存信息及字典索引对超过 64 字节的长数据模式进行快速智能匹配，再利用局部缓存信息匹配短数据串（超过 8 字节即可完成匹配），最后再对结果进行 LZ 压缩。通过三级压缩过程，数据的冗余成分被完全提取。实际的使用对比显示能够为用户带来最大的压缩比。

2.2 超低延迟引入

通常数据压缩和字节缓存功能由于对数据进行了复杂处理，并且处理中可能涉及到硬盘读写，往往引入可观的延迟。对某些实时性要求较高的应用，显著增加的延迟可能会带来用户体验的明显下降。为最大限度的降低这一副作用，设计了独特的内存硬盘二级缓存及字典结构，动态智能判断最可能被使用的数据，并将其置于内存中。同时根据当前连接匹配情况智能判断接下来需要的历史记录，并提前调入内存中。除此之外，还动态监控数据延迟，在必要时改变压缩方式（如只采用 LZ 压缩等方式）以控制延迟引入。这一独特的延迟的控制为用户带来整体应用体验的提升。

2.3 字节缓存双向性

有的字节缓存实现中缓存及字典信息只对同一方向的传输有效。比如，从分支机构 A 传输某文件到分支机构 B，当从分支机构 A 再次传输同样或经过修改的文件到分支机构 B 时，字节缓存能够起作用，但如果从分支机构 B 传输同样或修改过的文件到分支机构 A，前边积累的缓存及字典信息无法起作用。的优化实现能够在 B 到 A 的传输中仍然使用前次 A 到 B 传输积累的缓存及字典信息，获取同样的压缩比。

2.4 设备自动检测

一些广域网加速设备需要用户手工配置所有对端设备。这不但增加初次使用的配置复杂度，未来任何设备的增减都需要所有加速设备做配置改变，为使用带来极大的不便。为增加易用性，设计了设备自动检测功能。当一个连接新建时，都将自动探测该连接的对端是否有设备及是否对该连接打开了深度压缩功能。并根据检测结果决定是否打开深度压缩功能。这不但免去了用户部署设备时的配置工作，而且当用户在其它地点增加部署或改变部署时，系统都将自动适应并逐个连接的自动决定是否该启用深度压缩功能。这一全自动的设备及功能检测机制最大限度地提高了易用性，让用户不需劳神即可享受到低延迟深度压缩技术带来的带宽效率的巨大提升。

3. 用户案例：企业总部和分支机构之间专线加速

经过优化设计及实现的深度压缩功能不仅能为企业分支机构之间的各种应用带来性能的巨大提升，而且为企业节省了本来需要用于增加带宽的资金投

入。

以下通过一个实际用户使用案例说明深度压缩如何作用于具体应用及其使用效果。



图 2.深度压缩使用案例及效果

某省烟草公司在总部和分支机构之间部署。总部和分支机构之间使用 2M 的专线，没有部署前出现带宽紧张，文件传输及应用响应慢，有的应用跑不起

来。使用后，有效带宽得到成倍提升，文件传输及应用性能得到显著改善。

上图显示深度压缩带来的文件传输改善。首先从总部向一个分支机构上传一个有大量图片的 *PPT* 文件，由于该文件第一次传输并且内含的大量 *JPEG* 图片都不具可压缩性，深度压缩功能无法缩减传输数据量，只有 *TCP* 加速功能在起加速作用，并将 *2M* 带宽跑满。

整个上传持续了超过 *13* 分钟。过一段时间后，该 *PPT* 文件的作者对文件做了小的改动并更改了文件名，再次上传到同一分支机构。深度压缩功能在这次传输数据中发现了大量与前一次传输数据相同的数据串，字节缓存起到了显著的作用。几乎同样大小的文件传输仅用了 *23* 秒，*2M* 的专线有效带宽冲到了 *63M*，带来超过 *30* 倍的提高。深度压缩不但大大缩短了传送文件的时间，提高了工作效率，而且省下了宝贵的带宽给其它应用。总部从分支机构的下载类似应用案例也观察到了 *20* 倍的提升。