

年度深入报道：suprnova 终结：是 BT 的末日还是未来？

引子：Napster 和 kazza 的故事

前几天看到报道，Napster 现在已经在 Nasdaq 上市，由盗版 Mp3 的代言人成功转型为有偿下载收听音乐的一个公司，不由得慨叹世事无常，担架卖得可真快啊^_^。想当年偶在 Napster 下载了无数在国内很难见得到的 Mp3 歌曲。大学时我在电台听到的却羞于阮囊羞涩无法入手的《Maxwell》的《Whatever, whenever, wherever》独自沉醉于缭绕仙乐中，那种感觉。。。让我对 P2P 情有独钟。Napster 被关后，偶就迅速加入到了 kazza 一族中去。BT 刚在 03 年初在国内提及的时候，让偶觉得很不解，也因为在内网的缘故，浅尝 BT 后，因为速度和种子的问题而抛弃了它。之后一直混迹于 kazza 每天 2 到 3 百万的用户群中。

得益于 Bittorrent 的彻底的共享精神，国内编程高手也为 BT 做出了符合中国国情的改进，代表人物中有开发 bitcomet 的灿烂微笑等人，内网现在不用这个程序的还是少。而在 BT@china 联盟发布页上，文件的搜索变得越来越方便，这一切都令 BT 的吸引力越来越大。

第一章 Suprnova 的故事

一项研究表明，在 Kazza 上 70% 以上的歌曲是不完整或者虚假的，而 kazza 为防止“文件污染”而让用户对文件的质量进行投票的设计在后期失去了作用。据说 P2P 唯一屹立不倒的另一巨人 Edonkey 电骡也正受到这一威胁。花了几天才下载的电影打不开，气愤可想而知。

而在 suprnova.org 上，文件受污染的情况很少，发布在 suprnova.org 上的种子文件极少是无效的。需要多少人力和时间才能够做到这一点，本文将为你揭示 suprnova 这一切背后的秘密。

开启 BT 客户端程序，并不能从中找到需要下载内容。打个比喻来讲，用户硬盘就象是仓库，BT 客户端程序就象是车队，但免费的货物从哪里来？答案就在 suprnova.org。

BT 最为人称道的是其下载的速度，使用 BT 的每一个用户既上传又下载还都是服务器，在下载上传的同时提供分流的服务。而 torrent 文件或者说种子文件则仍然需要依赖传统互联网的 Http 服务。Suprnova 就是一个这样的提供种子文件的网站，在去年底由 Johan Pouwelse 博士所发表的一篇深入报道中，作者对 BT 及与其唇齿相依的 Suprnova.org 在网上的活动进行了深入的剖析。（原文见 <http://www.st.ewi.tudelft.nl/~pouwelse/>）

先来咬文嚼字一番吧。Peer 在英语里大意指的具有相似年龄、教育程度及家庭背景的阶层。任何 P2P 如果脱离了草根阶层也就不能叫做 Peer to Peer 程序了。一个好的 P2P 程序最直观的指标无非就是在线用户数。想当年偶上 Kazza 时，同时在线数经常稳定在 2.5 至 3.5 百万之间，什么东西都找得到。P2P 得以流行的几个主要条件：一是文件容易获取，这里免费是一个基本条件。二是用户能够找到他们想要的文件。三是系统能够承载突发的下载用户群。最后一点是下载速度。

该文采集数据量大及时间跨度长。不但专门设计了收集数据的软件，并且动用了 DAS 也即分布式超级计算机系统（类似于 SETI@home 的系统），可以说是迄今为止对 suprnova 最详尽及客观的深度报道。文章发表时恰逢 suprnova 宣布 suprnova.org 黯然退出的时间，使得此文更有纪念意义。下面请大家跟我一起来一窥 suprnova 的运作内幕吧。

第二章 技术背景

从 03 年 6 月至 04 年 3 月为期 8 个月的跟踪观测，对 suprnova 在互联网上有关的二千多个部份进行了观测。除此之外，还对下载其中一个 torrent 文件的 90,155 个下载者和其注入到消亡期间进行了跟踪。在两周期间内对于下载上百个新加入的文件的 54,845 个下载者的带宽进行了测量。这些工作使得此次评测为史上最大的。

采集数据的测量软件包括两个部分，两个部分各包括三个脚本。软件的第一部分用于监测互联网上的 BT\Suprnova 的各个组成部分的活动，其中包括了一个 Mirror 脚本，这个脚本负责测量 suprnova 的镜像站点的上线时间和反应时间；一个 Html 脚本，这个脚本负责收集和分析 suprnova 的镜像的超文本网页和下载所有新的 torrent 文件；第三个 tracker 脚本负责分析 torrent 文件和检查 tracker 的状态。

软件的第二大部分是用来监视下载用户的。为了在以分钟为间隔单位的情况下跟踪成千上万的用户，作者使用了分布式超级计算机系统里的 100 个节点。

脚本的工作原理如下：猎手脚本(Hunt Script)选择一个 torrent 文件进行跟踪并且开始测量所有下载此文件的用户；用户获取脚本(Getpeer script)与 tracker 连接并且获得下载用户的 IP 地址；第三个脚本 Peering 脚本与并行下载的用户连接，并且使用 BT 协议来测量他们的下载进程和上传时间。Hunt 脚本每隔一分钟就会扫描所有活动的 suprnova 镜像，寻找新发布的文件，一旦某个 torrent 文件被选中，Getpeer 和 Peering 脚本会同时启动，以同样的时间间隔（每分钟）。这种方法的话我们可以在第一时间得到注入元文件的用户的 IP 地址和估计每个下载者的平均下载时间。

在收集数据阶段主要有两个问题。首先是软件无法穿透众多用户的防火墙。当用户处于防火墙后时，Getpeer 脚本可以获得这个用户的 IP 地址，但 Peering 脚本则无法向其发送信息。因此，所测得的下载速度只是针对无防火墙用户而言的。第二个问题，我们无法从 tracker 直接得到所有用户的 IP 地址。BT 协议规定 tracker 只提供随意选取的一定数量的用户 IP 地址（默认数量是 20 个用户）。因此我们获得的用户 IP 地址只是实际用户的一部分。

第三章 专家解读图表

Johan Pouwelse 博士现于达夫特科技大学攻读博士后，自 1999 年起开始发布有关 P2P 的文章。在 BT 之前，也对 kazza 做出了深入研究。而有关 BT/Suprnova 的这篇文章是其最新的研究成果。文章中分别用收集的数据形成了七个图表，包括：种子文件的可获得性(availability)；BT/Suprnova 的三种网站的平均在线时间排名；X 轴为用户做种时间排行；Suprnova 上的内容提交者提交文件的排名表；个案报告：指环王第三部的突发下载效应；用户的平均下载速度；种子数量与文件寿命的统计。

这个部分的数据是在大约为一个月获得的（2003 年圣诞期间）下载或做种的用户的数量。然后是 BT/Suprnova 的整体测量结果，包括文件的获得性、完整度、突发下载效应和下载速度。

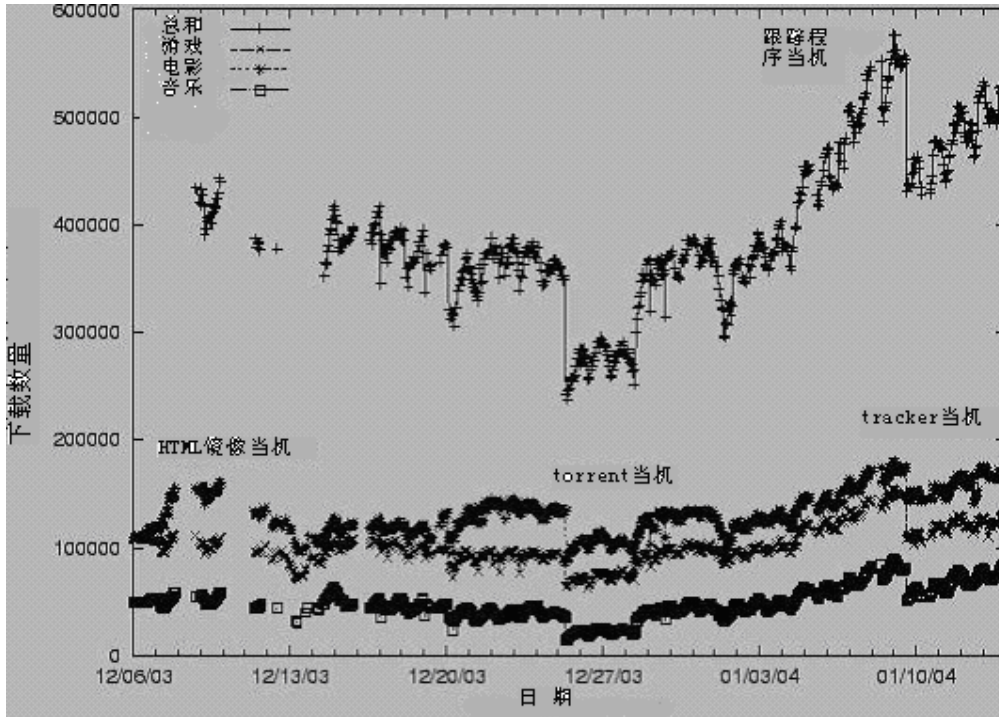


图 1: 下载及做种用户数. X 轴为时间 Y 轴为下载用户数。(2003 年 12 月到 2004 年 1 月)

图表里对三种类型文件（游戏、电影和音乐）的下载进行了统计，图表的最上方是所有文件类型下载的统一。选择这个时间段是因为在这段时间内，由于几次 BT/Suprnova 镜像服务器的当机而显示出来较大的区别。下载用户数由最低的 237,500(圣诞当天)到最高的 576,500 (1 月 9 日)。

图 1 中有两件事情必须注意：首先是重复性。每天的最大和最小用户数约会在每天的同一时间出现。第二点如在 12 月 8 号和 10 号，由于 Suprnova 的 8 个镜像服务器中有 6 个当机，在图表中出现了很大的缺口。尤其是在圣诞节当天，一个主要镜像服务器当机达 98 个小时，这导致了可下载电影数从 1675 个下降到 1017 个，从而使得用户数大幅下降。从一月 5 号到 10 号，镜像系统也离线若干次，使得 suprnova.org 无法登陆，镜像也无法更新，这导致的缺口在图表中清晰可见。但一旦 suprnova 服务正常后，迅速回升的用户数表明下载者并不会因此而不访问 suprnova.org。

结论：BT/Suprnova 系统所提供的种子文件与用户数成正相关关系。

BT/Suprnova 的结构脆弱性是因为其互联网四个部分的潜在危险而造成的。Suprnova 的主服务器有几次更换了 IP 地址和当机。各种不同的镜像服务器很少能在每天超过 1,200,000 的访问量下撑过几天的。(2004 年 10 月份数据)，而且有几次只有少于五个镜像服务器在线，甚至偶尔会发生没有任何服务器在线的情况，使得任何新的下载都无法进行。总体上来说，tracker 是经常遭受 DOS (denial-of service) 攻击的对象，而且由于每天数以 G 计的带宽流量，使得维持其运行成本很高。

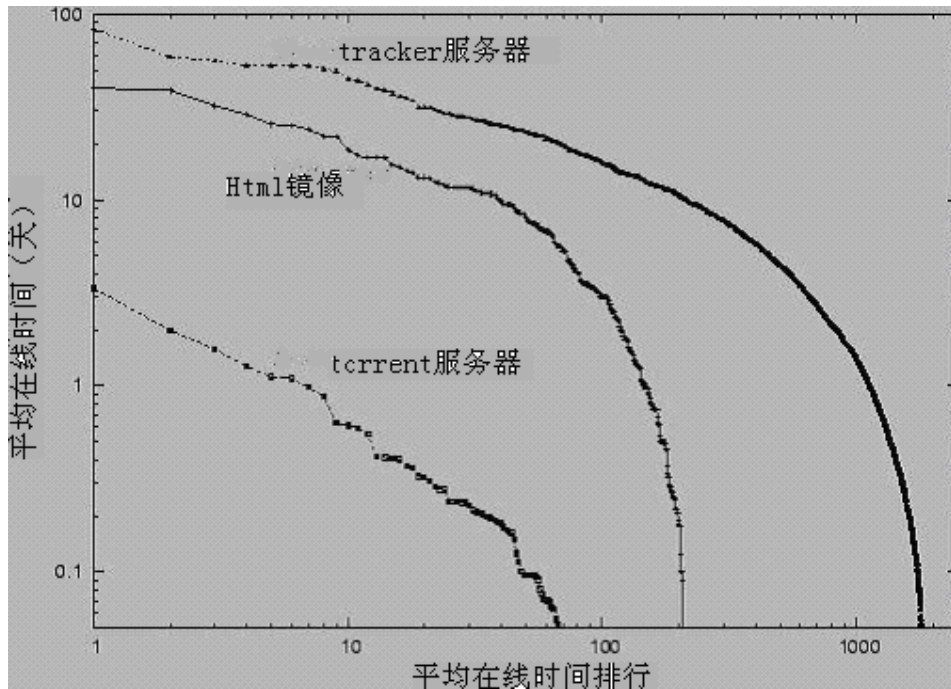


图 2: BT/Suprnova 的三种网站的平均在线时间排名 (X 轴为平均在线排行.Y 轴为平均在线时间。)

图 2 显示了 suprnova 的 234 个镜像服务器, 95 个 torrent 文件服务器和 1941 个种子追踪器的数据 (suprnova.org 本身并没有包括在内)。从图表中可以看出, 只有一半的镜像的平均在线时间超过了 2.1 天, 这也是文件可获得性的一个很好的指标。除此之外, 只有 39 个镜像的连续在线时间是长于 2 周的。可靠的虚拟主机的服务是一个问题。而据图表显示, torrent 文件服务器的服务则更加不可靠。有些 tracker 的可靠性是挺高的, 其中一个 tracker 显示连续在线超过 100 天。一半以上的 tracker 的平均在线时间超过了 1.5 天, 排名前一百名的 tracker 的平均在线时间高于 15.7 天。

结合图 1 和图 2 来看, 如果要使得 BT/Suprnova 的服务更易获得, 更加稳定, 有必要使其互联网的服务分散化。但所有使得 BT/Suprnova 成功的原因 (单击下载、高效文件、高速下载) 都高度依赖传统的网络部分。使得这一切象是个悖论

2003 年 12 月 10 号 Ubisoft 的著名游戏《撕裂的天堂》Beyond Good and Evil 被加入到 Suprnova, 到 2004 年 3 月 11 号最后一个种子消失。在此期间, 我们使用 Getpeer 脚本并且获得了下载此文件的 90, 155 的用户的 IP 地址。其中只有 53, 883 个 IP 地址是不在防火墙之后的而且是可被 Peerping 脚本所跟踪的。每隔一分钟我们就测算一次不在防火墙之后的用户的做种时间。

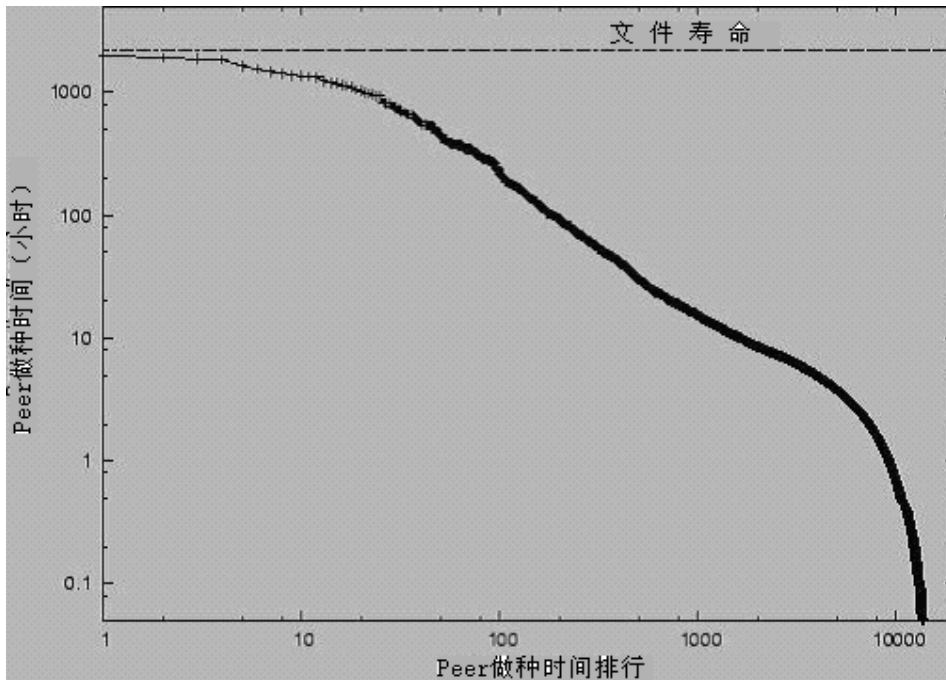


图 3: 用户做种时间排行和做种时间

图 3 是上传时间的统计。最长的做种时间为 83.5 天。排名 10 到 5000 的上传时间较为平稳，排名 5000 后则呈直线下降，这表明大部分下载者在完成下载后只会做种几十分钟到几个小时。但考虑到此游戏的平均下载时间需要几天的时间。另外，即使元文件很易获得，但做种者仍然较少。在这个例子内，53,883 个用户中只有 9,219 个用户（17%）在下载完成后做种时间超过一个小时。超过 10 小时的下降到只有 1,649 个用户即 3.1%，做种超过 100 小时的只有 183 人，只有 0.34%。

此两个图表展示了将来 P2P 程序在设计架构上仍然需要改进的地方。为了使整个系统更易获取，BT 的互联网部分的功能（即依赖 http 协议发布种子等）必须被分散（分散文件的发布工作是 kazza 的长处，在 exeem 中也做到了），例如由普通下载者发布。但是，由于现在长期做种用户很少，应有一种激励机制来使用户延长上传时间。

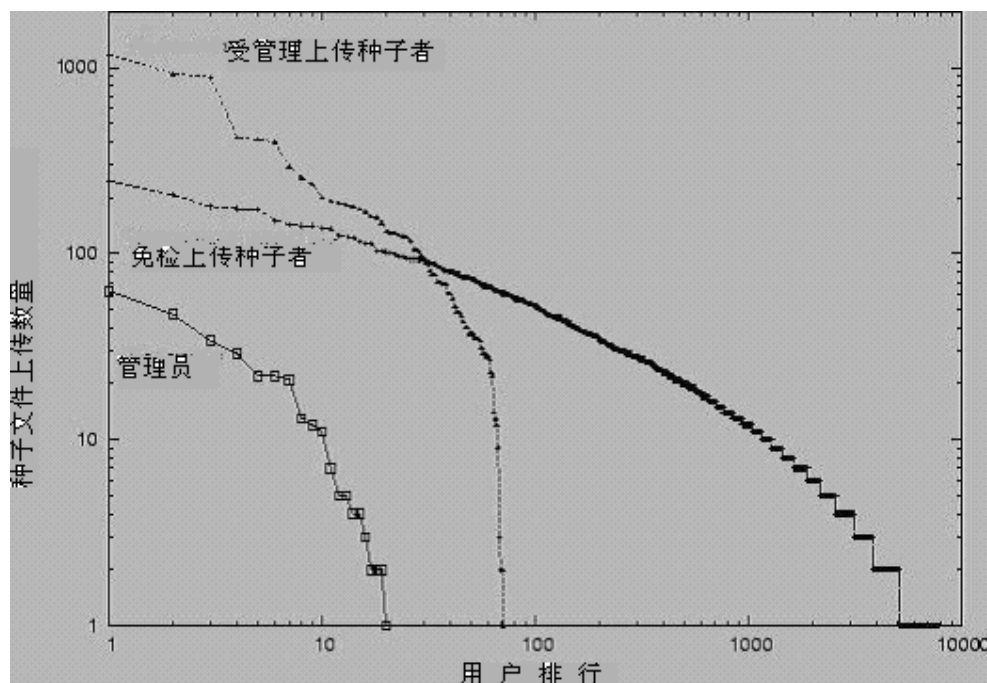


图 4. Suprnova 上的内容提交者提交文件的排名表。

这个部分研究了 BT/Suprnova 上的种子文件和元文件的完整性，这在 P2P 软件中一直为人所诟病。为了测试元文件的完整性，我们向 Suprnova 免费提供了一个镜像服务器的帐号，通过在 Html 代码中加入间谍软件，我们登记了每个种子文件的下载并且能够轻易地破坏元文件。

结论：使用捐献的服务器来存储元文件存在完整性被破坏和个人隐私遭侵权的危险。

至于内容的完整性，在 P2P 论坛及其他渠道都强调 BT/Suprnova 基本上是无虚假文件的。但在技术上直接测量虚假文件是很难的，而手工去检测海量的数据也行不通。实际上，我们积极地想发布虚假文件，通过在几台计算机上登记了不同的帐号，然后利用这些帐号上传了虚假的种子文件，但未能成功，管理员过滤掉了这些虚假的文件。

管理员的体制看上去在移除虚假和遭破坏的文件上非常有效，而且效率高。图 4 中可以看出 20 个管理员，71 个免检上传者，和 7,933 个受管理的上传者（活动期间从 2003 年 6 月至 2004 年 3 月）。排名前十的受管理上传者共注入了 5,191 个文件，免检上传者贡献了 1,693 个文件，而管理员注入了 274 个文件。只有 20 人组成的管理员队伍能够如此有效地管理不计其数的数据，令人吃惊。但不幸的是，这种管理机制严重地依赖于高度集中的互联网网站。

图 5: 突发下载效应

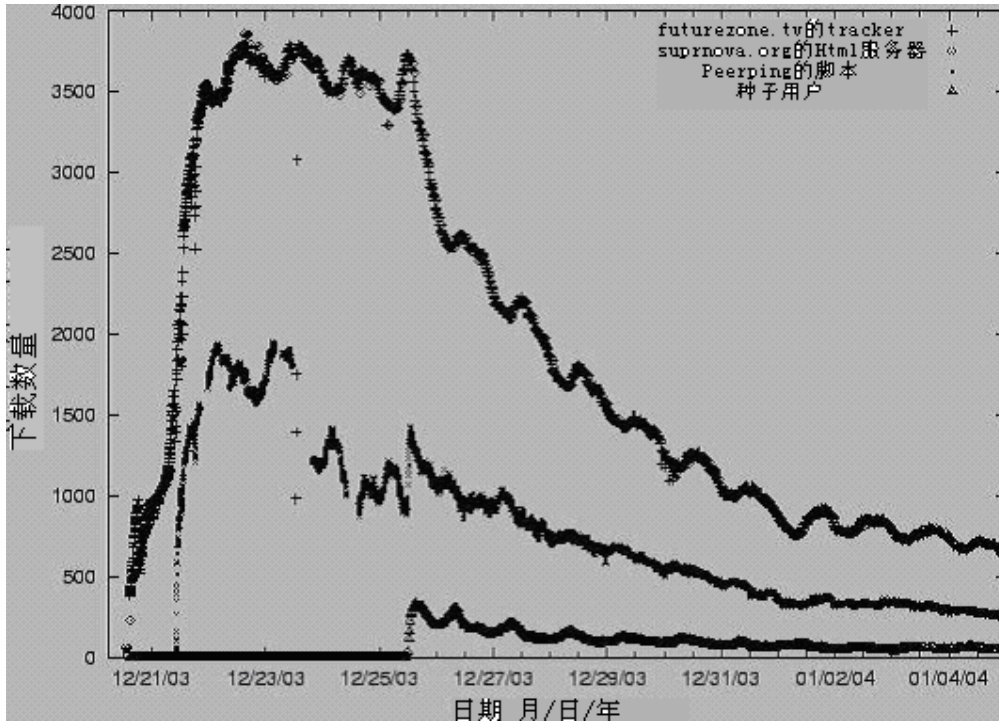


图 5: 指环王第三部的突发下载效应

选中 futurezone.TV 这个 tracker 是因为它提供了对详细统计数据的访问。最上层的曲线显示了正在下载的用户数总和，最底层的曲线显示了种子的数量。在文件提供下载的前五天，没有用户完成下载而元文件提供者一直在线做种。这么长的时间足以确认侵犯版权者的身份了。

HTML 脚本每隔一个小时从 suprnova 读取数据，从 suprnova 返回的数据和 futurezone.tv 这个 tracker 所提供数据非常一致，在图表上的曲线几乎是重叠的。只是在 2003 年 12 月 23 日 tracker 出现了几分钟的故障，但在 suprnova 没有返回这样的数据。Peerping 脚本返回的数据显示（即三条曲线中位于中间的曲线）则远比 tracker 显示的要少，这是由于防火墙造成的。40%的用户位于防火墙后，这部分用户在 peerping 中无法显示出来。而这条曲线中的缺口是因为运行此测算程序的 DAS 系统的存储限额而造成的。从这张图表我们认为 BT/Suprnova 的系统能够有效地处理非常高的突发下载群。

图 6: 下载速度

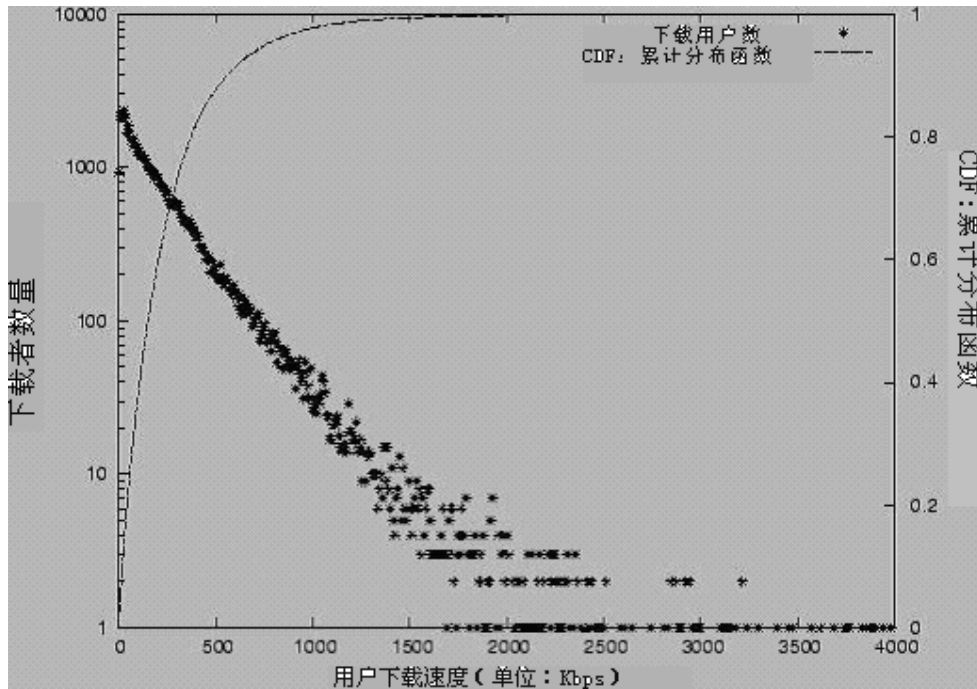


图6:用户的平均下载速度

在为期两周的试验期内，我们测算了 54, 845 名用户的平均下载带宽。为了获得这些数据，我们的 Hunt 脚本跟踪了在 2004 年 3 月 10 日被注入 *suprnova* 的前 108 个文件的下载情况。图表还显示了具有一定下载速度的用户的 CDF 即累计分布函数的统计数据（用短横线连接者）。结果显示 90% 的用户下载速度低于 520k; 平均 240K 的下载速度使用户能够在一天之内把大文件拉到自己的硬盘。

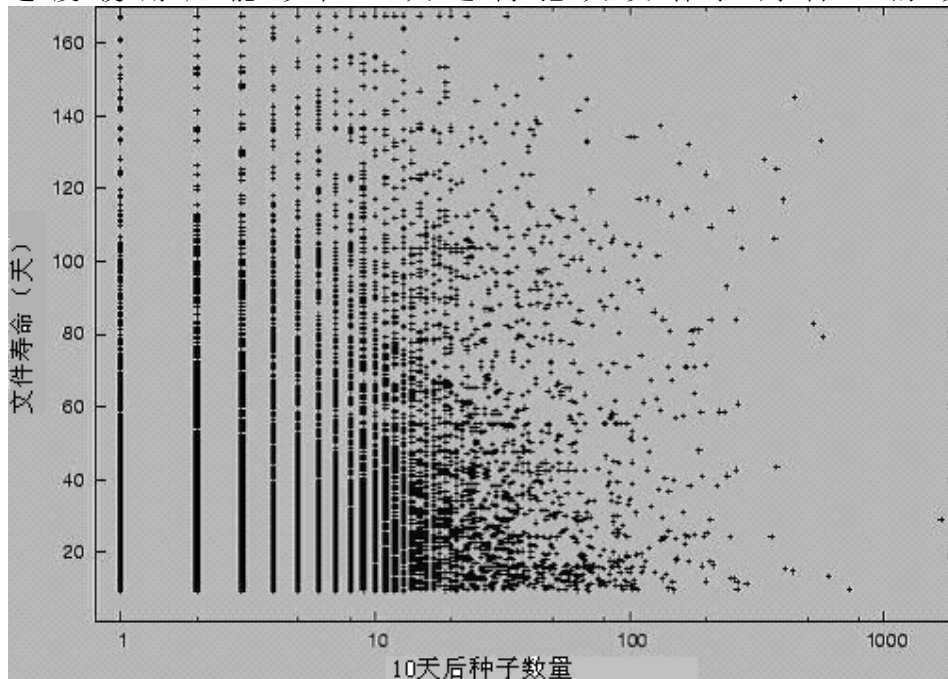


图7:超过 14, 000 个文件在加入 *suprnova* 十天后种子数量的统计.

图 7 显示的是我们跟踪的至少 500M 的所有文件的寿命。横轴显示的是文件被注入 BT/Suprnova 系统 10 天后的种子的数量。竖轴是文件寿命。很重要的一个发现是种子数量少并不意味着文件寿命短，只有一个种子的文件的寿命也可以很长。

结论：BT 上的文件的可获得程度是无法预计的。当最后一个种子下线时，文件也就寿终正寝了目前做种者在做种时其上传带宽被用到极致以利于其它下载者，这更象是一种惩罚措施。应该有一种激励机制来鼓励做种的行为，比如说做种者自己选择上传的带宽

我们相信现在所有的 BT 协议都有一个根本的设计上的缺陷。使用 Kazza 无需担心文件的可获得性，但其提供的文件的完整性很差，充斥了虚假文件。而需 20 名管理员就能够高效运转的 suprnova 的人员及机制建构又高度依赖于其高度集中的互联网部分。两者如果互相取长补短，则有望开创 P2P 系统的新纪元。

第四章：末日还是未来

正如文章开头所说，BT 与种子发布页的关系是唇齿相依的，suprnova 的关闭希望不会导致“唇亡齿寒”的情况。最好能够《功夫》中露臀沟的“酱爆”在面对星爷威胁时，还能抛出“还有千千万万个我”这样颇具烈士就义意味的警句那样。倒下一个 suprnova，将有更多的 suprnova 站起来。

Suprnova 在 2004 年 12 月 19 日宣布终结这颗超新星的光芒后，颇具深意地留下了一个不是句号的句号，宣布推出了新一代的 P2P 系统 exeem，文章截稿前其最新版本是 exeem0.21Beta。exeem 最大的特性是兼具 BT 与 kazza 的特点，也就是不使用 suprnova 提供的种子文件，在 exeem 内即可使用 BT 的高速下载网友共享的数据。

我们拭目以待。