

NBP

NETWORK

NBP-NETWORK

浅探百度网盘下载逻辑

01

基本准备

02

网页端 VS 客户端

03

SVIP VS PoorMan

04

买一送一 -> RESTful框架 双栈下载

优雅使用wireshark:

<https://zhuatlan.zhihu.com/p/36669377>

对PC端抓包:

<https://www.cnblogs.com/gancuimian/p/14010628.html>



Wireshark

Fiddler+proxifier

01 从网页端下载 21386 字节 tupian.png 和 26014284 字节 testppt.pptx

02 从清华镜像网站下载 4167029 字节 ubuntu-22.04.3-live-server-amd64.iso.zsync

03 从百度客户端登录并下载 21386 字节 tupian.png 和 26014284 字节 testppt.pptx

12对比, 观察百度网盘 (PanBD) 网页端和一般下载网站的下载区别

13对比, 观察PanBD网页端和客户端的下载区别

01 从网页端下载 21386 字节 tupian.png 和 26014284 字节 testppt.pptx

02 从清华镜像网站下载 4167029 字节 ubuntu-22.04.3-live-server-amd64.iso.zsync

02

网页端 VS 客户端

TCP下载证明：
以包02为例
找到下载的包的报文，
通过IP反向观察wireshark数据流

The screenshot displays a web browser window at the top, showing a list of files for download. The selected file is 'ubuntu-22.04.3-live-server-amd64.iso.zsync' with a size of 4,167,029 bytes. Below the browser, the Wireshark network traffic analysis tool is shown. The packet list pane highlights packet 481, which is a TCP segment. The packet details pane shows the following information:

```

Frame 481: 1514 bytes on wire (12112 bits), 1514 bytes captured (12112 bits) on interface \Device\NPF_{C8C6AA08-F938-4801-99F5-A7FE55676AF2}, id 0
Ethernet II, Src: 82:0d:9e:86:db:09 (82:0d:9e:86:db:09), Dst: e0:0a:f6:a9:e3:39 (e0:0a:f6:a9:e3:39)
Internet Protocol Version 4, Src: 101.6.15.130, Dst: 10.20.243.61
Transmission Control Protocol, Src Port: 443, Dst Port: 8683, Seq: 2921, Ack: 851, Len: 1460
  Source Port: 443
  Destination Port: 8683
  <Source or Destination Port: 443>
  <Source or Destination Port: 8683>
  [Stream index: 6]
  [TCP Segment Len: 1460]
  Sequence number: 2921 (relative sequence number)
  Sequence number (raw): 3949671960
  [Next sequence number: 4381 (relative sequence number)]
  
```

01 从网页端下载 21386 字节 tupian.png 和 26014284 字节 testppt.pptx

02 从清华镜像网站下载 4167029 字节 ubuntu-22.04.3-live-server-amd64.iso.zsync

02

网页端 VS 客户端

TCP下载证明：
对01包和03包做相同处理，
结果一致
证毕

The image displays two screenshots of Wireshark network traffic analysis. The top screenshot shows a successful TLS handshake and data transfer for a file named 'tupian.png'. The bottom screenshot shows a similar process for a file named 'testppt.pptx', including a FIN segment indicating the end of the transfer.

No.	Time	Source	Destination	Protocol	Length	Info
4429	24.646388	10.20.243.61	119.167.143.56	TLSv1.2	180	Client Key Exchange, Change Cipher Spec, Encrypted Handshake Message
4432	24.684081	119.167.143.56	10.20.243.61	TLSv1.2	328	New Session Ticket, Change Cipher Spec, Encrypted Handshake Message
4433	24.684081	119.167.143.56	10.20.243.61	TCP	60	443 → 5809 [ACK] Seq=3461 Ack=308 Min=30464 Len=0
4436	24.695511	10.20.243.61	119.167.143.56	TCP	458	[TCP Previous segment not captured] 5809 → 443 [PSH, ACK] Seq=170
4448	24.785799	119.167.143.56	10.20.243.61	TCP	60	[TCP ACKed unseen segment] 443 → 5809 [ACK] Seq=1735 Ack=2164 Win=
4458	25.148397	119.167.143.56	10.20.243.61	TCP	1514	443 → 5809 [ACK] Seq=3735 Ack=2164 Min=34048 Len=1460 [TCP segment
4459	25.148631	119.167.143.56	10.20.243.61	TCP	1514	[TCP Previous segment not captured] 443 → 5809 [ACK] Seq=16875 Ad
4460	25.148631	119.167.143.56	10.20.243.61	TCP	1514	[TCP Out-Of-Order] 443 → 5809 [ACK] Seq=15415 Ack=2164 Min=34048
4461	25.148631	119.167.143.56	10.20.243.61	TCP	1514	[TCP Out-Of-Order] 443 → 5809 [ACK] Seq=13965 Ack=2164 Min=34048

No.	Time	Source	Destination	Protocol	Length	Info
51644	98.053615	10.20.243.61	113.137.57.249	TCP	66	13525 → 443 [ACK] Seq=152 Ack=3457 Min=132096 Len=0 SLE=2921 SRE=
51645	98.055482	10.20.243.61	113.137.57.249	TLSv1.2	180	Client Key Exchange, Change Cipher Spec, Encrypted Handshake Message
51646	98.086883	113.137.57.249	10.20.243.61	TLSv1.2	296	New Session Ticket, Change Cipher Spec, Encrypted Handshake Message
51647	98.086883	113.137.57.249	10.20.243.61	TCP	60	443 → 13525 [ACK] Seq=3457 Ack=278 Min=30464 Len=0
51648	98.094535	10.20.243.61	113.137.57.249	TCP	54	13525 → 443 [FIN, ACK] Seq=278 Ack=3699 Min=131840 Len=0
51668	98.125707	113.137.57.249	10.20.243.61	TCP	60	443 → 13525 [ACK] Seq=3699 Ack=279 Min=30464 Len=0
51661	98.126270	113.137.57.249	10.20.243.61	TCP	60	443 → 13525 [FIN, ACK] Seq=3699 Ack=279 Min=30464 Len=0
51682	98.126359	10.20.243.61	113.137.57.249	TCP	54	13525 → 443 [ACK] Seq=279 Ack=3700 Min=131840 Len=0
52946	106.303571	10.20.243.61	113.137.57.249	TCP	66	13542 → 80 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 WS=256 SACK_PERM=
52947	106.334935	113.137.57.249	10.20.243.61	TCP	66	80 → 13542 [SYN, ACK] Seq=0 Ack=1 Min=8192 Len=0 MSS=1452 WS=32 S
52948	106.335101	10.20.243.61	113.137.57.249	TCP	54	13542 → 80 [ACK] Seq=1 Ack=1 Min=132096 Len=0
52949	106.338562	10.20.243.61	113.137.57.249	TCP	66	13543 → 80 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 WS=256 SACK_PERM=
52950	106.369707	113.137.57.249	10.20.243.61	TCP	66	80 → 13543 [SYN, ACK] Seq=0 Ack=1 Min=8192 Len=0 MSS=1452 WS=32 S
52951	106.369874	10.20.243.61	113.137.57.249	TCP	54	13543 → 80 [ACK] Seq=1 Ack=1 Min=132096 Len=0

01 从网页端下载 21386 字节 tupian.png 和 26014284 字节 testppt.pptx

02 从清华镜像网站下载 4167029 字节 ubuntu-22.04.3-live-server-amd64.iso.zsync

02

网页端 VS 客户端

一般下载网站 VS PanBD网页端:

先看一般网站的下载，可以发现来往报文 http 首部相对简洁，自定义字段并不多，也没有 cookies 或者 request URL 自定义字段在图中给出了说明当然，这只是个例网站，仍有以偏概全之嫌

The screenshot shows a network traffic analysis tool (WinConfig) displaying a list of HTTP requests. The selected request (ID 7) is for the file 'ubuntu-22.04.3-live-server-amd64.iso.zsync' from the URL 'mirrors.tuna.tsinghua.edu.cn/ubuntu-releases/22.04/ubuntu-22.04.3-live-server-amd64.iso.zsync'. The content length is 4,167,029 bytes, and the server IP is 101.6.15.130. A red box highlights the IP address, with the text '通过IP反向查询TCP流' (Reverse query TCP flow by IP) next to it.

The detailed view of the request shows the following headers:

```

Host: mirrors.tuna.tsinghua.edu.cn
Connection: keep-alive
Sec-CH-UA: "Not_A_Brand";u="a", "Chromium";v="120", "Google Chrome";v="120"
Sec-CH-UA-Mobile: ?0
Sec-CH-UA-Platform: "Windows"
Upgrade-Insecure-Requests: 1
User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/120.0.0.0 Safari/537.36
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/avif,image/webp,image/apng,*/*;q=0.8,application/signed-exchange;v=b3;q=0.0
Sec-Fetch-Site: same-origin
Sec-Fetch-Mode: navigate
Sec-Fetch-User: ?1
Sec-Fetch-Dest: document
Referer: https://mirrors.tuna.tsinghua.edu.cn/ubuntu-releases/22.04/
Accept-Encoding: gzip, deflate, br
Accept-Language: zh-CN,zh;q=0.9,en;q=0.8
  
```

The response headers show:

```

HTTP/1.1 200 OK
Date: Wed, 20 Dec 2023 06:23:12 GMT
Content-Length: 4167029
Content-Type: application/octet-stream
Etag: "605c0d1-39579"
Last-Modified: Thu, 10 Aug 2023 18:33:05 GMT
  
```

Red boxes highlight the 'Content-Length: 4167029' and 'Content-Type: application/octet-stream' fields, with the text '包的大小' (Package size) next to them.

01 从网页端下载 21386 字节 tupian.png 和 26014284 字节 testppt.pptx

02 从清华镜像网站下载 4167029 字节 ubuntu-22.04.3-live-server-amd64.iso.zsync

02

网页端 VS 客户端

一般下载网站 VS PanBD网页端:

先看一般网站的下载，可以发现来往报文 http 首部相对简洁，自定义字段并不多，也没有 cookies 或者 request URL 自定义字段在图中给出了说明当然，这只是个例网站，仍有以偏概全之嫌

Progress Telerik Fiddler Classic

Get Started | 请求统计 | 会话详细 | 自动转发 | 请求编辑 | Fiddler Orchestra Beta | FiddlerScript | Log | Filters | Timeline

协议头 | 文本查看 | 语法查看 | 网络视图 | 十六进制查看 | 认证 | Cookies | 数据 | JSON | XML

Name: | Value:

This request did not send any cookie data.

Find ... (press Ctrl+Enter to highlight all) | 在记事本中打开

Transformer | 协议头 | 文本查看 | 语法查看 | 网络视图 | 十六进制查看 | 认证 | 缓存 | Cookies | 数据 | JSON | XML

```

1 GET https://mirrors.tuna.tsinghua.edu.cn/ubuntu-releases/22.04/ubuntu-22.04.3-live-server-amd64.iso.zsync HTTP/1.1
2 Host: mirrors.tuna.tsinghua.edu.cn
3 Connection: keep-alive
4
5 sec-ch-ua: "Not A Brand";v="8", "Chromium";v="120", "Google Chrome";v="120" 表示浏览器的品牌和版本。Not A Brand是一个用于防止服务器端指纹识别的标记
6 sec-ch-ua-mobile: ?0 浏览器是否在移动端
7 sec-ch-ua-platform: "windows"
8 Upgrade-Insecure-Requests: 1
9 User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/120.0.0.0 Safari/537.36
10 Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/avif,image/webp,image/apng,*/*;q=0.8,application/signed-exchange;v=b3;q=0.7
11 Sec-Fetch-Site: same-origin 表示请求的来源和资源的来源之间的关系
12 Sec-Fetch-Mode: navigate
13 Sec-Fetch-User: ?1 是否由用户激活发起
14 Sec-Fetch-Dest: document
15 Referer: https://mirrors.tuna.tsinghua.edu.cn/ubuntu-releases/22.04/
16 Accept-Encoding: gzip, deflate, br
17 Accept-Language: zh-CN,zh;q=0.9,en;q=0.8
18
19
20
21 HTTP/1.1 200 OK
22 Server: nginx/1.18.0
23 Date: wed, 20 Dec 2023 06:23:12 GMT
24 Content-Type: application/octet-stream
25 Content-Length: 4167029
26 Last-Modified: Thu, 10 Aug 2023 18:33:05 GMT
27 Connection: keep-alive
28 ETag: "64d52d61-3f9575"
29 Strict-Transport-Security: max-age=31536000
30 X-TUNA-MIRROR-ID: neomirrors 标识了请求发送到的镜像站点，用于负载均衡和缓存目的
31 Accept-Ranges: bytes
32
33 zsync: 0.6.2
34 Filename: ubuntu-22.04.3-live-server-amd64.iso
35 MTime: Thu, 10 Aug 2023 05:06:27 +0000
36 Blocksize: 4096
37 Length: 21333
38
39 *** FIDDLER: RawDisplay truncated at 128 characters. Right-click to disable truncation. ***

```

绿框内为常见字段

01 从网页端下载 21386 字节 tupian.png 和 26014284 字节 testppt.pptx

02 从清华镜像网站下载 4167029 字节 ubuntu-22.04.3-live-server-amd64.iso.zsync

一般下载网站 VS PanBD网页端：

再看PanBD，我们发现复杂度激增，连接请求逻辑和自定义首部数量与一般下载网站截然不同

我们将从下载服务器开始，尽可能的梳理出网页端下载时的服务器请求链与特殊字段

从右图可见，下载服务器并不是pan.baidu.com，他们似乎都是baidupcs.com的子域名服务器，命名也都有迹可循

同时，不像一般网站的下载，网盘并不是在点击下载时直接跳转到该服务器，我们也看见了如d.pcs.baidu.com/mbd.baidu.com等等服务器，他们的作用也是我们感兴趣的地方

下面，我们以d.pcs.baidu.com这一服务器为例，带大家从http首部剖析其作用

259	200	HTTPS	pan.baidu.com	/api/analysis?clienttyp...	43	182.61.200.162	no-c...	image/jpeg; cha...	thro...	[#257]
260	200	HTTPS	pan.baidu.com	/api/list?clienttype=0&...	4,883	182.61.200.162	no-c...	application/son...	thro...	[#258]
261	200	HTTPS	thumbal0.baidupcs.com	/thumbal/34e0571ccj...	3,538	182.61.200.152	max...	image/jpeg	thro...	[#259]
262	200	HTTPS	mbd.baidu.com	/abd?data=%7B%22d...	2	180.97.107.3		application/son...	thro...	[#260]
263	200	HTTP	Tunnel to	103.30.235.171:443	729	103.30.235.171	pro...			[#261]
264	200	HTTPS	pan.baidu.com	/api/analysis?clienttyp...	43	182.61.200.162	no-c...	image/jpeg; cha...	thro...	[#262]
265	200	HTTPS	pan.baidu.com	/api/analysis?clienttyp...	43	182.61.200.162	no-c...	image/jpeg; cha...	thro...	[#263]
266	200	HTTPS	pan.baidu.com	/api/analysis?clienttyp...	43	182.61.200.162	no-c...	image/jpeg; cha...	thro...	[#264]
267	200	HTTPS	pan.baidu.com	/api/analysis?clienttyp...	43	182.61.200.162	no-c...	image/jpeg; cha...	thro...	[#265]
268	200	HTTP	Tunnel to	mbd.baidu.com:443	1,569	153.3.237.224			thro...	[#266]
269	200	HTTPS	pan.baidu.com	/dsk/cmsdata?clienttyp...	206	182.61.200.162		application/son...	thro...	[#267]
270	200	HTTPS	pan.baidu.com	/dsk/cmsdata?clienttyp...	3,228	182.61.200.162		application/son...	thro...	[#268]
271	304	HTTPS	staticws.cdn.bcebos.com	/ms-program%2Fmag...	0	60.188.66.35	Expir...	image/png	thro...	[#269]
272	200	HTTPS	pan.baidu.com	/api/gettemplatevariabl...	545	182.61.200.162	no-c...	application/son...	thro...	[#270]
273	200	HTTPS	mbd.baidu.com	/zbox?action=zpblog&...	44	153.3.237.224		application/son...	thro...	[#271]
274	200	HTTPS	pan.baidu.com	/api/download?clienttyp...	393	182.61.200.162	no-c...	application/son...	thro...	[#272]
275	200	HTTP	Tunnel to	d.pcs.baidu.com:443	815	182.61.200.15			thro...	[#273]
276	302	HTTPS	d.pcs.baidu.com	/file/34e0571ccj2db543...	50	182.61.200.15		text/plain; char...	thro...	[#274]
277	200	HTTP	Tunnel to	xafj-c111.baidupcs.com...	815	113.137.57.111			thro...	[#275]
278	200	HTTPS	xafj-c111.baidupcs.com	/file/34e0571ccj2db543...	21,386	113.137.57.111	max...	image/png	thro...	[#276]
279	200	HTTPS	pan.baidu.com	/api/analysis?clienttyp...	43	182.61.200.162	no-c...	image/jpeg; cha...	thro...	[#277]
280	200	HTTPS	pan.baidu.com	/api/analysis?clienttyp...	43	182.61.200.162	no-c...	image/jpeg; cha...	thro...	[#278]
281	200	HTTPS	pan.baidu.com	/api/analysis?clienttyp...	43	182.61.200.162	no-c...	image/jpeg; cha...	thro...	[#279]
282	200	HTTPS	pan.baidu.com	/api/analysis?clienttyp...	43	182.61.200.162	no-c...	image/jpeg; cha...	thro...	[#280]
283	200	HTTP	Tunnel to	mbd.baidu.com:443	1,569	153.3.237.224			thro...	[#281]
284	200	HTTPS	pan.baidu.com	/dsk/cmsdata?clienttyp...	3,228	182.61.200.162		application/son...	thro...	[#282]
285	200	HTTPS	pan.baidu.com	/api/download?clienttyp...	391	182.61.200.162	no-c...	application/son...	thro...	[#283]
286	200	HTTPS	mbd.baidu.com	/zbox?action=zpblog&...	44	153.3.237.224		application/son...	thro...	[#284]
287	200	HTTP	Tunnel to	d.pcs.baidu.com:443	815	182.61.200.15			thro...	[#285]
288	302	HTTPS	d.pcs.baidu.com	/file/32eb81f1cp13cb8f...	50	182.61.200.15		text/plain; char...	thro...	[#286]
289	200	HTTP	Tunnel to	alal02.baidupcs.com:443	815	119.167.143.56			thro...	[#287]
290	200	HTTPS	alal02.baidupcs.com	/file/15eb81f1cp13cb8f...	26,014,284	119.167.143.56	max...	application/vnd...	thro...	[#288]
291	302	HTTP	Tunnel to	ab-edf.google.com:443	582		no-c...	text/html; char...	thro...	[#289]
292	200	HTTP	Tunnel to	www.googleapis.com:443	0	172.217.160.74			thro...	[#290]
293	200	HTTP	Tunnel to	103.30.235.171:443	729	103.30.235.171	pro...			[#291]
294	200	HTTP	Tunnel to	asfr.lenovoem.com:443	729	103.30.235.171	leno...			[#292]
295	406	HTTPS	osfr.lenovoem.com	/report2	512	103.30.235.171	no-c...	text/html; char...	thro...	[#293]
296	200	HTTP	Tunnel to	103.30.235.171:443	729	103.30.235.171	pro...			[#294]
297	200	HTTP	Tunnel to	103.30.235.171:443	729	103.30.235.171	pro...			[#295]
298	502	HTTP	Tunnel to	172.217.163.42:443	582		no-c...	text/html; char...	thro...	[#296]
299	502	HTTP	Tunnel to	172.217.163.42:443	582		no-c...	text/html; char...	thro...	[#297]
300	200	HTTP	Tunnel to	61.178.98.116:80	0	61.178.98.116			thro...	[#298]
301	403	HTTP	cd.huoying666.com	/static/profiles.pro	330	175.6.29.181		application/xml	thro...	[#299]
302	200	HTTP	Tunnel to	180.163.249.205:80	0	180.163.249.205			thro...	[#300]
303	200	HTTP	profile.se.360.cn	/proxyen.php	0	180.163.249.205		application/octe...	thro...	[#301]

文件下载

01 从网页端下载 21386 字节 tupian.png 和 26014284 字节 testppt.pptx

02 从清华镜像网站下载 4167029 字节 ubuntu-22.04.3-live-server-amd64.iso.zsync

一般下载网站 VS PanBD网页端:

让我们先回到wireshark的抓包流，将IP锁定为下载服务器的IP（这里指xafj-ct11服务器）可以看见下方 113段IP为下载服务器。而在此之前紧跟着的是IP182.61.200.15的tcp握手与数据传输。图中没有显示的是在于182交流完毕后主机马上就利用DNS寻找113服务器，说明d.pcs.baidu.com可能是一个响应传输请求的服务器

让我们回到fiddler去实锤

No.	Time	Source	Destination	Protocol	Length	Info
3740	15.765046	182.61.200.162	10.20.243.61	TCP	66	[TCP Dup ACK 3738#1] 443 → 5693 [ACK] Seq=22370 Ack=64402 Win=1500 Len=0
3744	15.789673	182.61.200.162	10.20.243.61	TCP	60	[TCP ACKed unseen segment] 443 → 5693 [ACK] Seq=22370 Ack=65854 Len=0
3745	15.792298	182.61.200.162	10.20.243.61	TCP	60	443 → 5693 [ACK] Seq=22370 Ack=68589 Win=167936 Len=0
3746	15.877428	182.61.200.162	10.20.243.61	TLSv1.2	792	Application Data
3747	15.894367	10.20.243.61	182.61.200.15	TCP	66	5797 → 443 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 WS=256 SACK_PERM=1
3748	15.919047	10.20.243.61	182.61.200.15	TCP	54	5693 → 443 [ACK] Seq=68500 Ack=23188 Win=131328 Len=0
3749	15.935900	182.61.200.15	10.20.243.61	TCP	66	443 → 5797 [SYN, ACK] Seq=0 Ack=1 Win=8192 Len=0 MSS=1452 WS=32 SACK_PERM=1
3750	15.936130	10.20.243.61	182.61.200.15	TCP	54	5797 → 443 [ACK] Seq=1 Ack=1 Win=132096 Len=0
3751	15.939343	10.20.243.61	182.61.200.15	TLSv1.2	229	Client Hello
3757	16.079630	182.61.200.15	10.20.243.61	TCP	60	443 → 5797 [ACK] Seq=1 Ack=176 Win=30336 Len=0
3758	16.082606	182.61.200.15	10.20.243.61	TCP	1514	[TCP Previous segment not captured] 443 → 5797 [ACK] Seq=1461 Ack=176 Win=30336 Len=0
3759	16.082606	182.61.200.15	10.20.243.61	TCP	1514	[TCP Out-Of-Order] 443 → 5797 [ACK] Seq=1 Ack=176 Win=30336 Len=0
3760	16.082700	10.20.243.61	182.61.200.15	TCP	54	5797 → 443 [ACK] Seq=176 Ack=2021 Win=132096 Len=0
3761	16.082892	182.61.200.15	10.20.243.61	TLSv1.2	555	Server Hello, Certificate, Server Key Exchange, Server Hello Done
3762	16.082892	182.61.200.15	10.20.243.61	TCP	555	[TCP Retransmission] 443 → 5797 [PSH, ACK] Seq=2021 Ack=176 Win=30336 Len=0
3763	16.082950	10.20.243.61	182.61.200.15	TCP	66	5797 → 443 [ACK] Seq=176 Ack=3422 Win=131584 Len=0 SLE=2921 SRE=1
3764	16.085677	10.20.243.61	182.61.200.15	TLSv1.2	147	Client Key Exchange, Change Cipher Spec, Encrypted Handshake Message
3765	16.185542	182.61.200.15	10.20.243.61	TLSv1.2	224	New Session Ticket, Change Cipher Spec, Encrypted Handshake Message
3766	16.185542	182.61.200.15	10.20.243.61	TCP	60	443 → 5797 [ACK] Seq=3422 Ack=269 Win=30336 Len=0
3767	16.189363	10.20.243.61	182.61.200.15	TCP	832	[TCP Previous segment not captured] 5797 → 443 [PSH, ACK] Seq=317 Ack=176 Win=30336 Len=0
3768	16.190320	10.20.243.61	13.107.21.230	TCP	66	[TCP Retransmission] 5780 → 443 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 WS=256 SACK_PERM=1
3770	16.208258	182.61.200.15	10.20.243.61	TCP	60	[TCP ACKed unseen segment] 443 → 5797 [ACK] Seq=3592 Ack=1721 Win=30336 Len=0
3771	16.208898	182.61.200.15	10.20.243.61	TCP	60	443 → 5797 [ACK] Seq=3592 Ack=3951 Win=37760 Len=0
3772	16.309201	182.61.200.15	10.20.243.61	TLSv1.2	1262	Application Data
3773	16.309392	182.61.200.15	10.20.243.61	TCP	139	[TCP Previous segment not captured] 443 → 5797 [PSH, ACK] Seq=547 Ack=176 Win=30336 Len=0
3774	16.309392	182.61.200.15	10.20.243.61	TLSv1.2	724	[TCP Out-Of-Order] , Application Data
3775	16.309470	10.20.243.61	182.61.200.15	TCP	54	5797 → 443 [ACK] Seq=3951 Ack=5555 Win=132096 Len=0
3776	16.309008	182.61.200.15	10.20.243.61	TLSv1.2	88	Application Data
3777	16.309008	10.20.243.61	182.61.200.15	TCP	54	5797 → 443 [ACK] Seq=3951 Ack=5589 Win=132096 Len=0
3780	16.423979	10.20.243.61	113.137.57.111	TCP	66	5799 → 443 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 WS=256 SACK_PERM=1
3781	16.424794	182.61.200.15	10.20.243.61	TLSv1.2	88	[TCP Spurious Retransmission] , Application Data
3782	16.424843	10.20.243.61	182.61.200.15	TCP	66	[TCP Dup ACK 3777#1] 5797 → 443 [ACK] Seq=3951 Ack=5589 Win=132096 Len=0
3783	16.490492	113.137.57.111	10.20.243.61	TCP	66	443 → 5799 [SYN, ACK] Seq=0 Ack=1 Win=8192 Len=0 MSS=1452 WS=32 SACK_PERM=1
3784	16.490732	10.20.243.61	113.137.57.111	TCP	54	5799 → 443 [ACK] Seq=1 Ack=1 Win=132096 Len=0
3785	16.493602	10.20.243.61	113.137.57.111	TLSv1.2	236	Client Hello

01 从网页端下载 21386 字节 tupian.png 和 26014284 字节 testppt.pptx

02 从清华镜像网站下载 4167029 字节 ubuntu-22.04.3-live-server-amd64.iso.zsync

02

网页端 VS 客户端

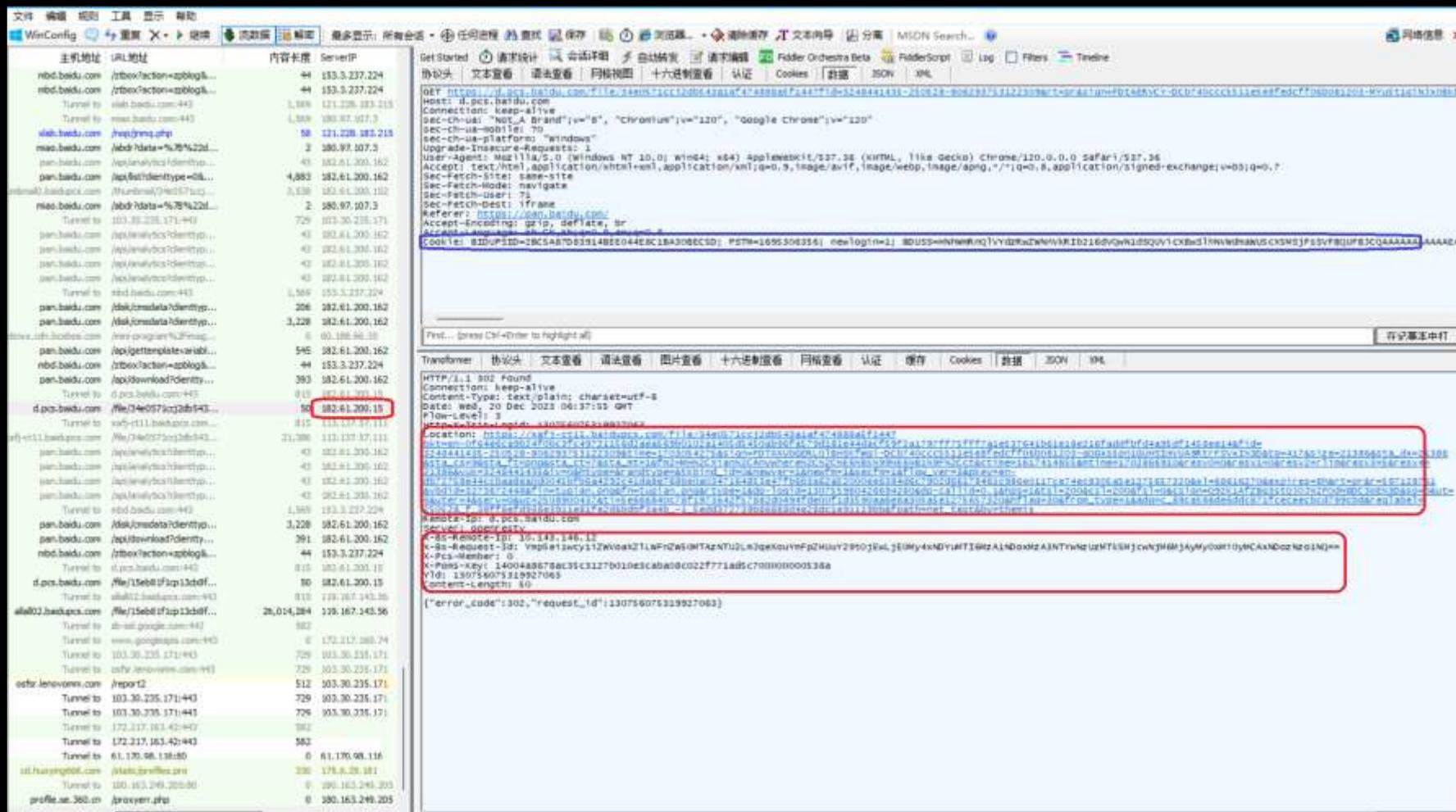
一般下载网站 VS PanBD网页端:

182段IP证实刚刚看见的tcp握手确实是 d.pcs.baidu.com

再看应答与请求

应答中的Location字段纯纯显眼包，仔细一看开头正是xaji下载服务器，这说明就是这个服务器给host提供了下载的地址。同时下面也有一些自定义的X字段，推测是认证信息，此处不予证明。

而请求报文看起来则是风平浪静，与一般网站相差无几。但仔细看它超出屏幕的cookie就知道事情不简单



PanBD网页端：

故接下去我们按照分析d.pcs.baidu.com的逻辑
将网页端主要（出现频率高）的子域名服务器作了分析

因时长有限，我们直接给出结论

其中

请求服务器用于处理所有的动作请求并给出下一指令应答

推送服务器用于建立全双工通信以实时获取数据交互

略缩图服务器用于预览图片，子域名中的X代表数字

静态资源服务器负责存储框架类web前端文件

埋点服务器在特定动作埋伏，获取用户特征画像以优化运营策略

文件传输服务器分为专线和通用，每次连接的服务器为host发送请求时的相对最优解。专线服务器的前四位‘aabb’代表地区和用途，‘cu/cm/ct’代表三大运营商，XX代表number

通用服务器一般以allall或者地区+all为域名，所在地服务器规模一般较大

具体的结构框图会在本节最后给出

具体的结构框图会在本节最后给出

01 从网页端下载 21386 字节 tupian.png 和 26014284 字节 testppt.pptx

02 从清华镜像网站下载 4167029 字节 ubuntu-22.04.3-live-server-amd64.iso.zsync

‘根’服务器:pan.baidu.com/pcs.baidu.com/baidupcs.com

百度账号登录服务器:passport.baidu.com

账号数据云服务器:pcsdata.Baidu.com

请求服务器:d.pcs.baidu.com

推送服务器:webpush.pan.baidu.com //websocket 全双工通信

略缩图服务器:thumbnailX.baidupcs.com

静态资源服务器:bdstatic.com

埋点服务器 mbd.baidu.com Mobile Big Data / hm.baidu.com hao123 monitor

文件传输服务器

专线服务器 aabb-cu/ct/cmXX.baidupcs.com

通用服务器: allall/地点all01.baidupcs.com

PanBD客户端:

同理（省略一万字），我们整理出客户端从登录到下载所接触的服务器，并与网页端对比

右图中黑色为增添服务器，白色为保留服务器，灰色为不可见服务器

其中，host会规律性的向更新服务器发送用户状态，包括签名，传输沟道，VIP信息等
与P2P服务器的连接出现在登录初始化时，可能用于所谓“高速下载”？

插件服务器提供内核、基础库、升级程序等插件，封装了web端的一些功能服务器

我们会发现相比于网页端，客户端在下载服务器上并无太大区别，而功能服务器似乎更加集中密闭，从某一程度上确实提高了安全性。

找到了交互服务器的区别，接着再聚焦于下载时的字段有无明显区别

01 从网页端下载 21386 字节 tupian.png 和 26014284 字节 testppt.pptx

03 从百度客户端登录并下载 21386 字节 tupian.png 和 26014284 字节 testppt.pptx

‘根’服务器

更新服务器:update.pan.baidu.com

P2P连接服务器:aa.t.bcsp2p.baidu.com

百度账号登录服务器

账号数据云服务器

请求服务器

推送服务器

略缩图服务器

静态资源服务器

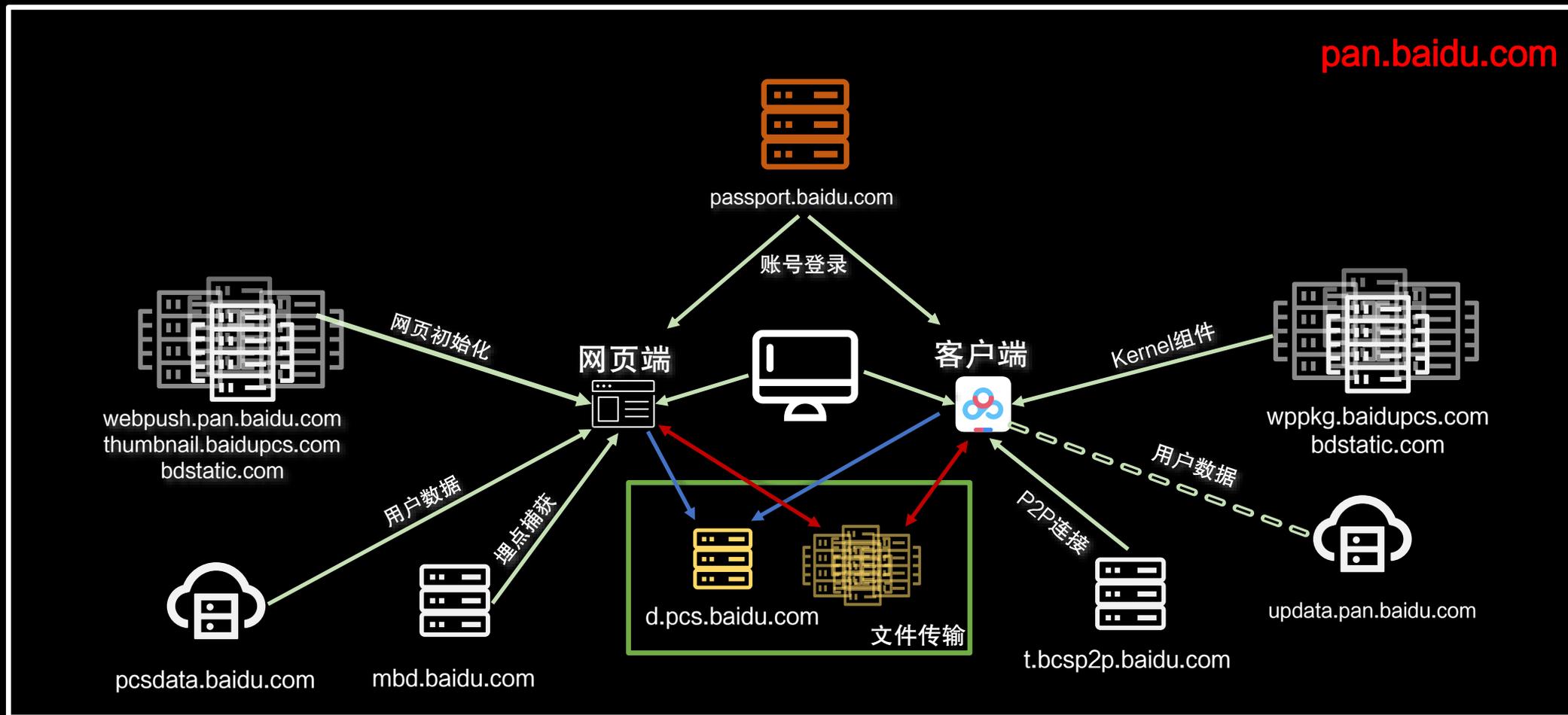
埋点服务器

插件服务器:wppkg.baidupcs.com

文件传输服务器



总结框图



SVIP VS PoorMan

这里我们采用拿他人API接口获取SVIP直链的方法
可以看见原本的d.pcs.baidu.com变成了三个api的请求
分别为数据列表提交、身份认证和token验证

普通用户和SVIP用户所用的服务器是一样的
从requestURL的角度，除了iv字段从0变成2，还有一些随机数的不同几乎没有差别。

而后，我们开始怀疑tcp的窗口限制

04 用第三方工具下载 21386 字节 tupian.png 和 26014284 字节 testppt.pptx

The screenshot displays the Fiddler Classic interface. The left pane shows a list of network requests. Three requests to www.94speed.com are highlighted in red, with their respective URLs: `/api.php/`, `/api.php`, and `/api.php/token`. The right pane shows the details of a request to www.94speed.com. The 'iv' field in the request body is highlighted in blue and has a value of '2'. The response content is truncated at 126 characters, showing a truncated image file named 'tupian.png'.

序号	状态	请求	主机地址	URL地址
537	200	HTTP	Tunnel to 124.237.224.33:443	
538	206	HTTPS	bddx<td>1.baidu.com	/file/15eb02f3p13cbff...
539	200	HTTP	Tunnel to 124.237.224.33:443	
540	206	HTTPS	bddx<td>1.baidu.com	/file/15eb02f3p13cbff...
541	206	HTTPS	bddx<td>1.baidu.com	/file/15eb02f3p13cbff...
542	206	HTTPS	bddx<td>1.baidu.com	/file/15eb02f3p13cbff...
543	206	HTTPS	bddx<td>1.baidu.com	/file/15eb02f3p13cbff...
544	206	HTTPS	bddx<td>1.baidu.com	/file/15eb02f3p13cbff...
545	206	HTTPS	bddx<td>1.baidu.com	/file/15eb02f3p13cbff...
546	0	HTTP	Tunnel to content-audit8.google...	
547	200	HTTPS	o2ywhn1de.feishu.cn	/apace/api/obj_status/...
548	200	HTTP	Tunnel to www.google-analytics.com	
549	200	HTTP	Tunnel to internal-api-tark-ap.fei...	
550	200	HTTPS	o2ywhn1de.feishu.cn	/apace/api/obj_status/...
551	204	HTTPS	www.google-analytics.com	/mp/collect?newuame...
552	200	HTTPS	internal-api-tark-ap.fei...	/accounts/status/heart...
553	200	HTTPS	mca-bd.feishu.cn	/v1/lat
554	200	HTTP	Tunnel to internal-api-tark-ap.fei...	
555	200	HTTPS	internal-api-tark-ap.fei...	/accounts/status/heart...
556	200	HTTPS	o2ywhn1de.feishu.cn	/apace/api/explorer/hec...
557	200	HTTPS	o2ywhn1de.feishu.cn	/like/v2/api/obj_high...
558	200	HTTPS	o2ywhn1de.feishu.cn	/apace/api/extract/...
559	200	HTTPS	mca-bd.feishu.cn	/v1/lat
560	200	HTTPS	mca-bd.feishu.cn	/v1/lat
561	200	HTTPS	mca-bd.feishu.cn	/v1/lat
562	200	HTTPS	mca-bd.feishu.cn	/v1/lat
563	200	HTTP	Tunnel to 124.237.224.33:443	
564	200	HTTPS	mca-bd.feishu.cn	/v1/lat
565	200	HTTPS	o2ywhn1de.feishu.cn	/apace/api/extract/...
566	200	HTTP	Tunnel to d.pcs.baidu.com:443	
567	0	HTTP	Tunnel to www.google.com:443	
568	200	HTTPS	o2ywhn1de.feishu.cn	/apace/api/explorer/hec...
569	200	HTTPS	o2ywhn1de.feishu.cn	/apace/api/obj_status/...
570	0	HTTP	Tunnel to www.google.com:443	
571	200	HTTP	Tunnel to content-audit8.google...	
572	200	HTTPS	www.94speed.com	/api.php/
573	200	HTTPS	www.94speed.com	/api.php
574	200	HTTPS	www.94speed.com	/api.php/token
575	200	HTTP	Tunnel to 124.237.224.33:443	
576	200	HTTP	pan.baidu.com	/percent#?identity=...
577	0	HTTP	Tunnel to content-audit8.google...	
578	200	HTTP	Tunnel to 153.120.57.111:443	
579	302	HTTP	Tunnel to 20.198.183.76:443	
580	200	HTTPS	bddx<td>1.baidu.com	/file/15eb02f3p13cbff...
581	206	HTTP	Tunnel to 113.121.57.111:443	
582	HTTP	Tunnel to alq@live.com:443:8080...		
583	200	HTTP	Tunnel to 124.237.224.33:443	
584	200	HTTPS	update.pan.baidu.com	/statistics?identity=9...
585	0	HTTP	Tunnel to www.google.com:443	
586	200	HTTP	Tunnel to 124.237.224.33:443	
587	200	HTTPS	update.pan.baidu.com	/statistics?identity=9...

03

SVIP VS PoorMan

但事实上，两者的实际窗口大小相差无几，怀疑不成立

接着我们想到了线程的控制，并对比了网页端、客户端和SVIP的区别，有了一些发现：

先看网页端，网页端用浏览器下载只有单线程，难怪恹慢

Window size value: 1040 SVIP [Calculated window size: 33280] [Window size scaling factor: 32] Checksum: 0xd846 [unverified]	Window size value: 952 [Calculated window size: 30464] [Window size scaling factor: 32] Checksum: 0xb2aa [unverified]
---------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------

序号	状态...	请求...	主机地址	URL地址	内容长度	ServerIP
290	200	HTTPS	pan.baidu.com	/api/download?clientty...	391	182.61.200.162
291	200	HTTPS	mbd.baidu.com	/ztbox?action=zblog&...	44	153.3.237.224
292	200	HTTP	Tunnel to	d.pcs.baidu.com:443	815	182.61.200.15
293	302	HTTPS	d.pcs.baidu.com	/file/15eb81f1cp13cb8f...	50	182.61.200.15
294	200	HTTP	Tunnel to	allall02.baidupcs.com:443	815	119.167.143.56
295	200	HTTP	Tunnel to	110.242.69.174:80	0	110.242.69.174
296	200	HTTP	pan.baidu.com	/api/filediff?block_list=...	467	110.242.69.174
297	200	HTTPS	allall02.baidupcs.com	/file/15eb81f1cp13cb8f...	26,014,284	119.167.143.56

03

SVIP VS PoorMan

左图是小文件客户端下载，速度在5M/s左右，已经是飞速了

可以看见网盘对线程数量并没有作过多的限制，而是对每个线程的下载上限作出约束，像左图都是2097152封顶

小文件封顶影响不显著，但一旦下载较大文件就要命了，上限掉到原来的1/4，意味着与服务器连接所花费的时间激增，自然速度就下来了

再看SVIP，在截图所在时间的网速下并没有触碰到上限，这样能让网络自己调整到最高效的下载节奏上即不限速

综上，我们得出结论，网盘是通过对其一线程的上限作出限制而达成限速的目的~

进程ID	进程名	协议	本地地址	远程地址	内容长度	速度
56	206	HTTP	app01.baidu.com	/file/15eb81f1cp13cb8f...	2,097,152	2,097,152
57	206	HTTP	app01.baidu.com	/file/15eb81f1cp13cb8f...	2,097,152	2,097,152
58	206	HTTP	app01.baidu.com	/file/15eb81f1cp13cb8f...	2,097,152	2,097,152
59	206	HTTP	app01.baidu.com	/file/15eb81f1cp13cb8f...	2,097,152	2,097,152
60	200	HTTP	Tunnel to	113.137.57.75:80	0	0
61	200	HTTP	Tunnel to	113.137.57.75:80	0	0
62	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	26,014,284	2,097,152
			Tunnel to	124.237.224.10:443	815	0
63	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	25,248,344	2,097,152
			Tunnel to	124.237.224.10:443	815	0
64	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	12,534,402	2,097,152
			Tunnel to	124.237.224.10:443	815	0
65	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	6,267,202	2,097,152
			Tunnel to	124.237.224.10:443	815	0
66	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	5,988,396	2,097,152
			Tunnel to	124.237.224.10:443	815	0
67	200	HTTP	Tunnel to	124.237.224.10:443	815	848,460
68	200	HTTP	Tunnel to	124.237.224.10:443	815	5,520
69	200	HTTP	Tunnel to	124.237.224.10:443	815	0
70	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	2,920,844	2,097,152
			Tunnel to	124.237.224.10:443	815	0
71	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	2,658,529	2,097,152
			Tunnel to	124.237.224.10:443	815	0
72	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	2,322,486	2,097,152
			Tunnel to	124.237.224.10:443	815	0
73	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	2,128,167	2,097,152
			Tunnel to	124.237.224.10:443	815	0
74	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	1,280,230	2,097,152
			Tunnel to	124.237.224.10:443	815	0
75	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	915,856	2,097,152
			Tunnel to	124.237.224.10:443	815	0
76	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	714,983	2,097,152
			Tunnel to	124.237.224.10:443	815	0
77	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	640,148	2,097,152
			Tunnel to	124.237.224.10:443	815	0
78	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	1,911,780	2,097,152
			Tunnel to	124.237.224.10:443	815	0
79	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	559,248	2,097,152
			Tunnel to	124.237.224.10:443	815	0
80	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	510,097	2,097,152
			Tunnel to	124.237.224.10:443	815	0
81	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	457,961	2,097,152
			Tunnel to	124.237.224.10:443	815	0
82	206	HTTP	bddx-ct01.baidu.com	/file/15eb81f1cp13cb8f...	1,291,492	2,097,152
			Tunnel to	124.237.224.10:443	815	0

进程ID	进程名	协议	本地地址	远程地址	内容长度	速度
57	200	HTTP	Tunnel to	113.167.143.118:80	0	0
58	206	HTTP	qdbn-cu00.baidu.com	/file/89f160ed5n246b1...	524,288	524,288
59	206	HTTP	qdbn-cu00.baidu.com	/file/89f160ed5n246b1...	524,288	524,288
60	206	HTTP	qdbn-cu00.baidu.com	/file/89f160ed5n246b1...	524,288	524,288
61	206	HTTP	qdbn-cu00.baidu.com	/file/89f160ed5n246b1...	524,288	524,288
62	206	HTTP	qdbn-cu00.baidu.com	/file/89f160ed5n246b1...	524,288	524,288
63	200	HTTP	pan.baidu.com	/rest/2.0/membership/...	14	14
64	200	HTTP	pan.baidu.com	/rest/2.0/membership/...	14	14
65	206	HTTP	qdbn-cu00.baidu.com	/file/89f160ed5n246b1...	524,288	524,288
66	206	HTTP	qdbn-cu00.baidu.com	/file/89f160ed5n246b1...	524,288	524,288
67	403	HTTP	pan.baidu.com	/rest/2.0/membership/...	80	80
68	200	HTTP	Tunnel to	update.pan.baidu.com:...	0	0
69	200	HTTPS	update.pan.baidu.com	/statistics?dentype=8...	11	11
70	206	HTTP	qdbn-cu00.baidu.com	/file/89f160ed5n246b1...	524,288	524,288
71	206	HTTP	qdbn-cu00.baidu.com	/file/89f160ed5n246b1...	524,288	524,288
72	206	HTTP	qdbn-cu00.baidu.com	/file/89f160ed5n246b1...	524,288	524,288

我们在研究的过程中了解了RESTful框架统一接口，并在抓包中发现了IPv6、IPv4的双栈传输，本欲详细说明，怎奈时间有限，只能贴出网址了T.T

优雅提速直链获取网站: 94speed.com

什么是RESTful API:

<https://zhuatlan.zhihu.com/p/334809573>

什么是双栈结构:

<https://zhuatlan.zhihu.com/p/569641486>

基于wireshark和fiddler+proxifier两个平台，我们抓取了网页端和客户端、普通用户和SVIP用户的下载流程链，作出对比分析，最终得到了下载逻辑框图和限速原因，并给出了科学合法有效的提速方法，浅显地解决了自己的问题。当然，我们仍未清晰的弄明一些请求字段的含义与作用，这是我们此次研究的遗憾。最后，这里是NBP-宁波小学队，希望我们的调查能给你一些收获~

