

媒体云转码的演进： MapReduce、DASH 与稳定婚姻

Alan Zhuang

cheedoong@acm.org

2014 年 3 月 10 日

Outline

1 背景

2 从 Cloud Transcoder 到 TranscX

- 前腾讯研究院 Cloud Transcoder
- 架平流媒体 TranscX

3 DASH 与稳定婚姻

- DASH
- 稳定匹配

Outline

1 背景

2 从 Cloud Transcoder 到 TranscX

- 前腾讯研究院 Cloud Transcoder
- 架平流媒体 TranscX

3 DASH 与稳定婚姻

- DASH
- 稳定匹配

多屏时代的挑战

■ 多种平台



多屏时代的挑战

■ 多种平台



多屏时代的挑战

- 多种平台





多屏时代的挑战

■ 多种平台



■ 多种屏幕大小



多屏时代的挑战

■ 多种平台



■ 多种屏幕大小



多屏时代的挑战

■ 多种平台



■ 多种屏幕大小



多屏时代的挑战

■ 多种平台



■ 多种屏幕大小



多屏时代的挑战

■ 多种平台



■ 多种屏幕大小



多屏时代的挑战

■ 多种码率



多屏时代的挑战

- 多种码率



多屏时代的挑战

■ 多种码率



多屏时代的挑战

■ 多种码率



多屏时代的挑战

■ 多种码率



多屏时代的挑战

■ 多种码率



■ 多种解码能力

联发科芯片	MT6572	MT6582	MT6588	MT6592
Display	960×540P	1280×720P	1920×1280P	1920×1280P
H.264 Decode	720P@30fps	1080P@30fps	1080P@30fps	1080P@30fps
HEVC Decode	N/A	N/A	720P@30fps	720P@30fps

多屏时代的挑战

■ 不同封装容器支持

mp4, mkv, avi, flv, wmv, rmvb, webm, mpeg-ts...

多屏时代的挑战

■ 不同封装容器支持

mp4, mkv, avi, flv, wmv, rmvb, webm, mpeg-ts...

多屏时代的挑战

- 不同封装容器支持

mp4, mkv, avi, flv, wmv, rmvb, webm, mpeg-ts...

- 不同编码标准支持

多屏时代的挑战

- 不同封装容器支持

mp4, mkv, avi, flv, wmv, rmvb, webm, mpeg-ts...

- 不同编码标准支持

多屏时代的挑战

■ 不同封装容器支持

mp4, mkv, avi, flv, wmv, rmvb, webm, mpeg-ts...

■ 不同编码标准支持

H.264(AVC), H.265(HEVC), VC-1, AVS, VP8/9, RealVideo...

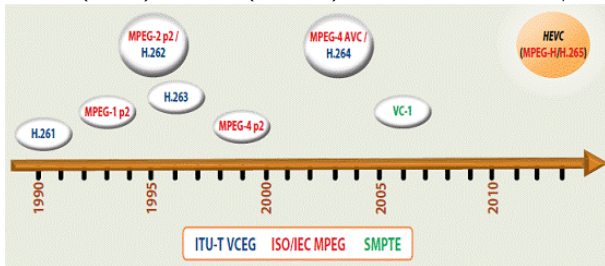
多屏时代的挑战

■ 不同封装容器支持

mp4, mkv, avi, flv, wmv, rmvb, webm, mpeg-ts...

■ 不同编码标准支持

H.264(AVC), H.265(HEVC), VC-1, AVS, VP8/9, RealVideo...



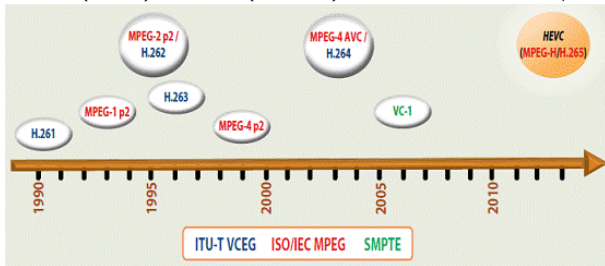
多屏时代的挑战

■ 不同封装容器支持

mp4, mkv, avi, flv, wmv, rmvb, webm, mpeg-ts...

■ 不同编码标准支持

H.264(AVC), H.265(HEVC), VC-1, AVS, VP8/9, RealVideo...



多屏时代的挑战

■ 巨头角力



Tencent 腾讯

多屏时代的挑战

■ 巨头角力



Tencent 腾讯

多屏时代的挑战

■ 巨头角力



Tencent 腾讯

幸运的是,

幸运的是,

决大多数设备都支持:

幸运的是,

决大多数设备都支持:

- 编码标准

幸运的是,

决大多数设备都支持:

- 编码标准

幸运的是,

决大多数设备都支持:

- 编码标准

H.264/AVC (ISO/IEC 14496-10; ITU-T H.264; MPEG-4 Part 10).

- 封装容器

幸运的是,

决大多数设备都支持:

- 编码标准

H.264/AVC (ISO/IEC 14496-10; ITU-T H.264; MPEG-4 Part 10).

- 封装容器

幸运的是,

决大多数设备都支持:

- 编码标准

H.264/AVC (ISO/IEC 14496-10; ITU-T H.264; MPEG-4 Part 10).

- 封装容器

MP4 (ISO/IEC 14496-14; MPEG-4 Part 14).

幸运的是,

决大多数设备都支持:

■ 编码标准

H.264/AVC (ISO/IEC 14496-10; ITU-T H.264; MPEG-4 Part 10).

■ 封装容器

MP4 (ISO/IEC 14496-14; MPEG-4 Part 14).

所以, 我们需要把用户/编辑上传的各种不同封装、不同编码的, 一般码率比较高的源视频转成若干种适合不同设备的不同码率的 H.264 编码、MP4 封装的视频。

但是,

但是，

- 媒体转码是件及其消耗计算资源的工作
尤其视频编码，对于目前常见的支持 SSE4 指令集的
x86/x64 CPU 的机器：

但是,

- 媒体转码是件及其消耗计算资源的工作
尤其视频编码, 对于目前常见的支持 SSE4 指令集的
x86/x64 CPU 的机器:
 - 编码 H.264 视频需要耗费播放时长的 1/3 到 2/3

但是,

- 媒体转码是件及其消耗计算资源的工作
尤其视频编码, 对于目前常见的支持 SSE4 指令集的 x86/x64 CPU 的机器:
 - 编码 H.264 视频需要耗费播放时长的 1/3 到 2/3
 - 编码 H.265 视频需要耗费播放时长的 30+ 倍

但是，

- 媒体转码是件及其消耗计算资源的工作
尤其视频编码，对于目前常见的支持 SSE4 指令集的 x86/x64 CPU 的机器：
 - 编码 H.264 视频需要耗费播放时长的 1/3 到 2/3
 - 编码 H.265 视频需要耗费播放时长的 30+ 倍
 - 单个 CPU 核通常最多可跑 1-2 个编码任务

但是，

- 媒体转码是件及其消耗计算资源的工作
尤其视频编码，对于目前常见的支持 SSE4 指令集的 x86/x64 CPU 的机器：
 - 编码 H.264 视频需要耗费播放时长的 1/3 到 2/3
 - 编码 H.265 视频需要耗费播放时长的 30+ 倍
 - 单个 CPU 核通常最多可跑 1-2 个编码任务
- 媒体文件大，再加上多码率副本，极其消耗存储

但是，

- 媒体转码是件及其消耗计算资源的工作
尤其视频编码，对于目前常见的支持 SSE4 指令集的 x86/x64 CPU 的机器：
 - 编码 H.264 视频需要耗费播放时长的 1/3 到 2/3
 - 编码 H.265 视频需要耗费播放时长的 30+ 倍
 - 单个 CPU 核通常最多可跑 1-2 个编码任务
- 媒体文件大，再加上多码率副本，极其消耗存储
- 潜在的带宽消耗

解决

解决

Criteria:

解决

Criteria:

- 单机内

解决

Criteria:

- 单机内

- 功能划分、数据局部性
宏块组粒度的并行

解决

Criteria:

■ 单机内

- 功能划分、数据局部性

宏块组粒度的并行

- 内存访问、CPU 核心/Cache 拓扑结构和转码格式

帧级或 GOP（图像组）级并行

对 NUMA 机器特别友好

解决

Criteria:

■ 单机内

- 功能划分、数据局部性

宏块组粒度的并行

- 内存访问、CPU 核心/Cache 拓扑结构和转码格式

帧级或 GOP（图像组）级并行

对 NUMA 机器特别友好

■ 分布式转码

解决

Criteria:

■ 单机内

- 功能划分、数据局部性

宏块组粒度的并行

- 内存访问、CPU 核心/Cache 拓扑结构和转码格式

帧级或 GOP（图像组）级并行

对 NUMA 机器特别友好

■ 分布式转码

- 存储

解决

Criteria:

■ 单机内

- 功能划分、数据局部性

宏块组粒度的并行

- 内存访问、CPU 核心/Cache 拓扑结构和转码格式

帧级或 GOP（图像组）级并行

对 NUMA 机器特别友好

■ 分布式转码

- 存储

- 路由

解决

Criteria:

■ 单机内

- 功能划分、数据局部性

宏块组粒度的并行

- 内存访问、CPU 核心/Cache 拓扑结构和转码格式

帧级或 GOP（图像组）级并行

对 NUMA 机器特别友好

■ 分布式转码

- 存储

- 路由

- 任务调度

Outline

1 背景

2 从 Cloud Transcoder 到 TranscX

- 前腾讯研究院 Cloud Transcoder
- 架平流媒体 TranscX

3 DASH 与稳定婚姻

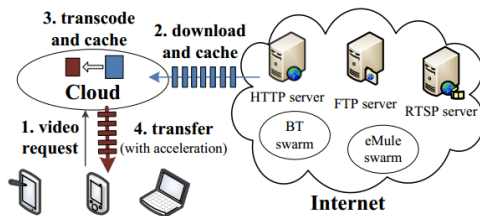
- DASH
- 稳定匹配

前腾讯研究院 Cloud Transcoder

**Done 2011-. Gale Huang et al. Cloud transcoder:
bridging the format and resolution gap between internet
videos and mobile devices. ACM NOSSDAV 2012.**

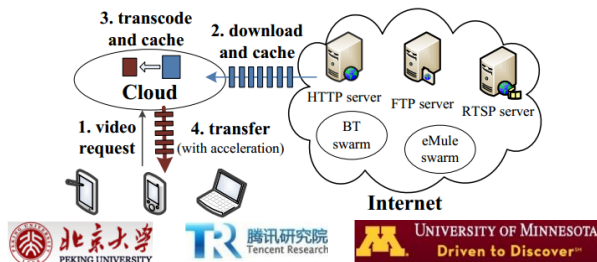
前腾讯研究院 Cloud Transcoder

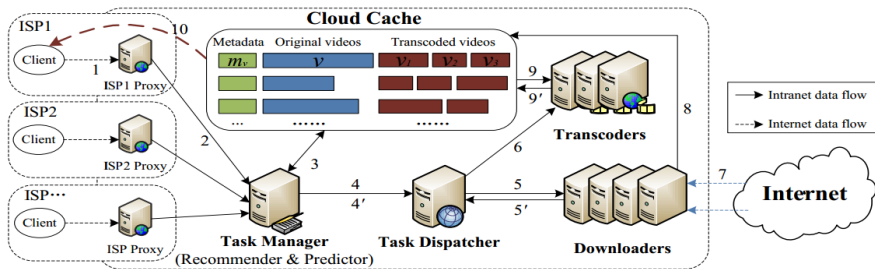
Done 2011-. Gale Huang et al. Cloud transcoder:
bridging the format and resolution gap between internet
videos and mobile devices. ACM NOSSDAV 2012.

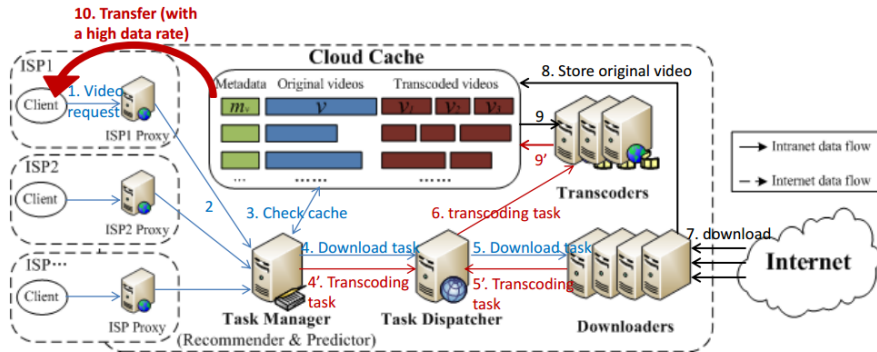


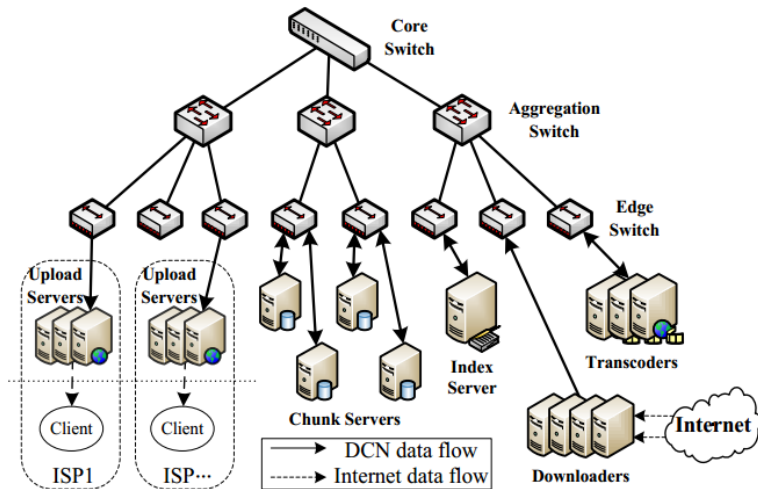
前腾讯研究院 Cloud Transcoder

Done 2011-. Gale Huang et al. Cloud transcoder:
bridging the format and resolution gap between internet
videos and mobile devices. ACM NOSSDAV 2012.





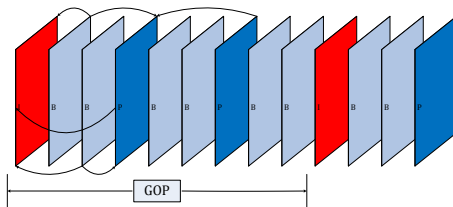




架平流媒体 TranscX

TranscX:

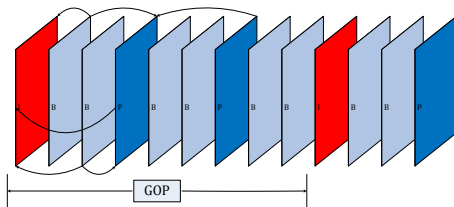
- Transcoding eXpress/eXperience/eXtreme...



架平流媒体 TranscX

TranscX:

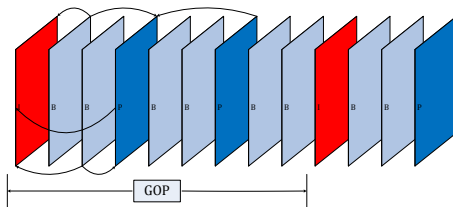
- Transcoding eXpress/eXperience/eXtreme...
- transc(x)



架平流媒体 TranscX

TranscX:

- Transcoding eXpress/eXperience/eXtreme...
- transc(x)



- GOP-level parallelism without REAL splitting

- Job/Map/Thread, accurate control and CPU & I/O limitation

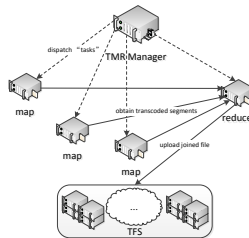
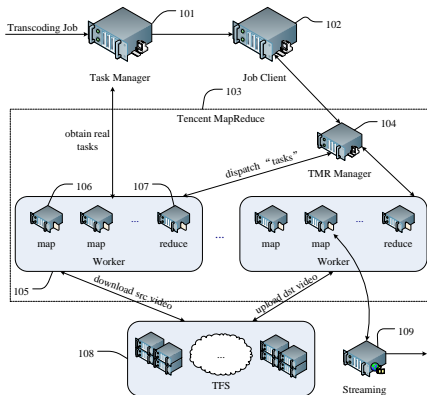
- Job/Map/Thread, accurate control and CPU & I/O limitation
- Real-time support for DASH and Live Broadcasting

- Job/Map/Thread, accurate control and CPU & I/O limitation
- Real-time support for DASH and Live Broadcasting
- Migration from Computing to Storage: soul of cloud tech

- Job/Map/Thread, accurate control and CPU & I/O limitation
- Real-time support for DASH and Live Broadcasting
- Migration from Computing to Storage: soul of cloud tech
- It later supported WeChat, Qzone and Weishi

- Job/Map/Thread, accurate control and CPU & I/O limitation
- Real-time support for DASH and Live Broadcasting
- Migration from Computing to Storage: soul of cloud tech
- It later supported WeChat, Qzone and Weishi
- 用 MapReduce 框架统一了起来, 减少了调度的复杂性, 增强了可扩展性

- Job/Map/Thread, accurate control and CPU & I/O limitation
- Real-time support for DASH and Live Broadcasting
- Migration from Computing to Storage: soul of cloud tech
- It later supported WeChat, Qzone and Weishi
- 用 MapReduce 框架统一了起来，减少了调度的复杂性，增强了可扩展性
- 复用了存储服务器的空间计算资源，节省了成本



Outline

1 背景

2 从 Cloud Transcoder 到 TranscX

- 前腾讯研究院 Cloud Transcoder
- 架平流媒体 TranscX

3 DASH 与稳定婚姻

- DASH
- 稳定匹配

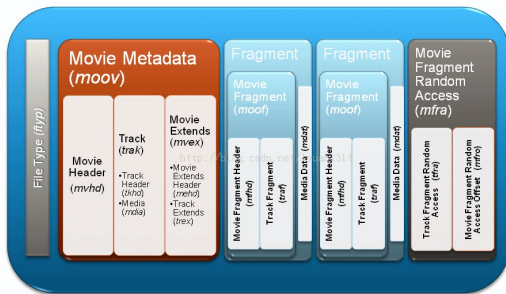
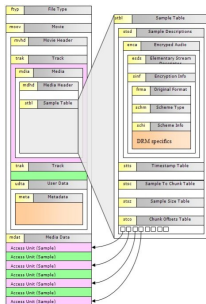
Why DASH?

DASH: Dynamic Adaptive Streaming over HTTP.

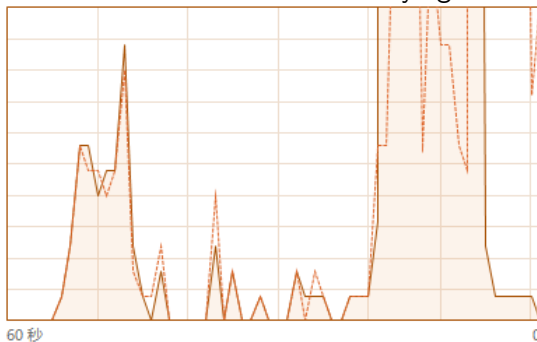
Why DASH?

DASH: Dynamic Adaptive Streaming over HTTP.

- 庞大的文件头，导致在线播放时较大的 initial/VCR delay



- 非平滑的码率间切换 due to varying download speed



DASH

DASH: Dynamic Adaptive Streaming over HTTP.

几种 DASH 标准:

DASH

DASH: Dynamic Adaptive Streaming over HTTP.

几种 DASH 标准:

- Apple HLS (HTTP Live Streaming) 2009

DASH

DASH: Dynamic Adaptive Streaming over HTTP.

几种 DASH 标准:

- Apple HLS (HTTP Live Streaming) 2009
- Microsoft HSS (HTTP Smooth Streaming) 2010

DASH

DASH: Dynamic Adaptive Streaming over HTTP.

几种 DASH 标准:

- Apple HLS (HTTP Live Streaming) 2009
- Microsoft HSS (HTTP Smooth Streaming) 2010
- Adobe HDS (HTTP Dynamic Streaming) 2010

DASH

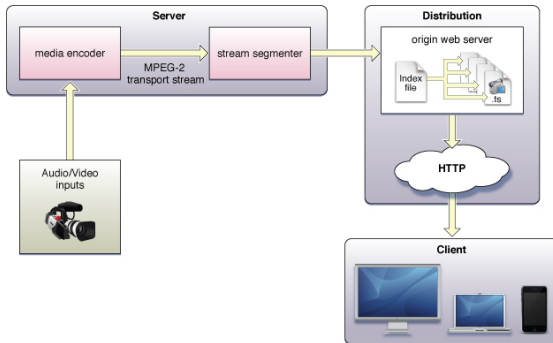
DASH: Dynamic Adaptive Streaming over HTTP.

几种 DASH 标准:

- Apple HLS (HTTP Live Streaming) 2009
- Microsoft HSS (HTTP Smooth Streaming) 2010
- Adobe HDS (HTTP Dynamic Streaming) 2010
- MPEG-DASH (ISO/IEC 23009-1) 2012

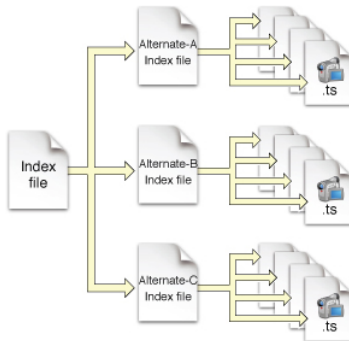
Apple HLS

■ Architecture



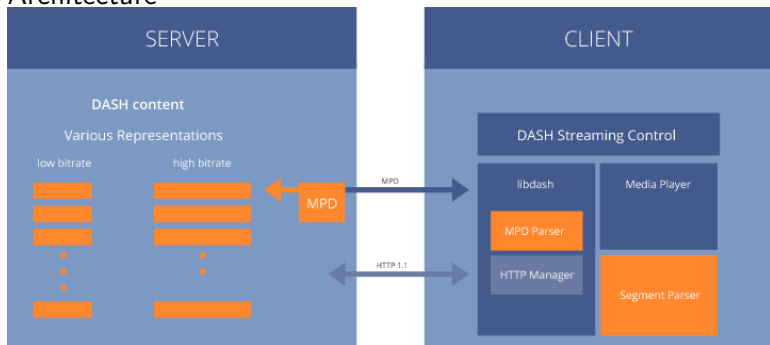
Apple HLS

■ Segment Indexing



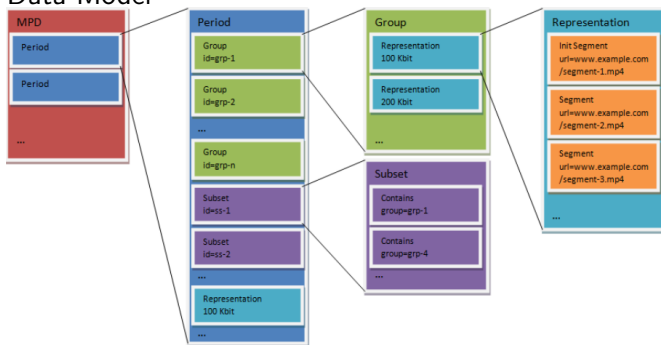
MPEG-DASH

■ Architecture



MPEG-DASH

■ Data Model



MPEG-DASH

■ Segment Indexing

Segment Index in MPD only

```
<MPD>
...
<URL sourceURL="seg1.mp4"/>
<URL sourceURL="seg2.mp4"/>
</MPD>
```

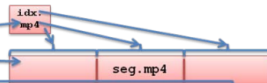


```
<MPD>
...
<URL sourceURL="seg.mp4" range="0-499"/>
<URL sourceURL="seg.mp4" range="500-999"/>
</MPD>
```



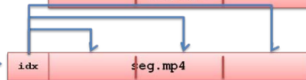
Segment Index in MPD + Segment

```
<MPD>
...
<Index sourceURL="idx.mp4"/>
<URL sourceURL="seg.mp4"/>
</MPD>
```



Segment Index in Segment only

```
<MPD>
...
<BaseURL>seg.mp4</BaseURL>
</MPD>
```



架平流媒体对 DASH 的使用

- 尽量只转封装，不转编码

架平流媒体对 DASH 的使用

- 尽量只转封装，不转编码
- 实时、按需转封装/转码码

Motivation

对大数据的分析，得出结论：

- 用户对 CDN 的 region 选择存在偏好

由于网络拓扑关系，用户去不同的 CDN 边缘节点下载的速率各不相同，而目前的调度算法，至少能得到一个次优解

Motivation

对大数据的分析，得出结论：

- 用户对 CDN 的 region 选择存在偏好
由于网络拓扑关系，用户去不同的 CDN 边缘节点下载的速率各不相同，而目前的调度算法，至少能得到一个次优解
- 不同不同地区的用户对码率存在偏好
在前面次优解的前提下，不同地区得到的平均服务质量有显著差异

优化问题

- 选择选择...
to be written...

优化问题

- 选择选择...
to be written...
- 稳定稳定匹配
to be written...