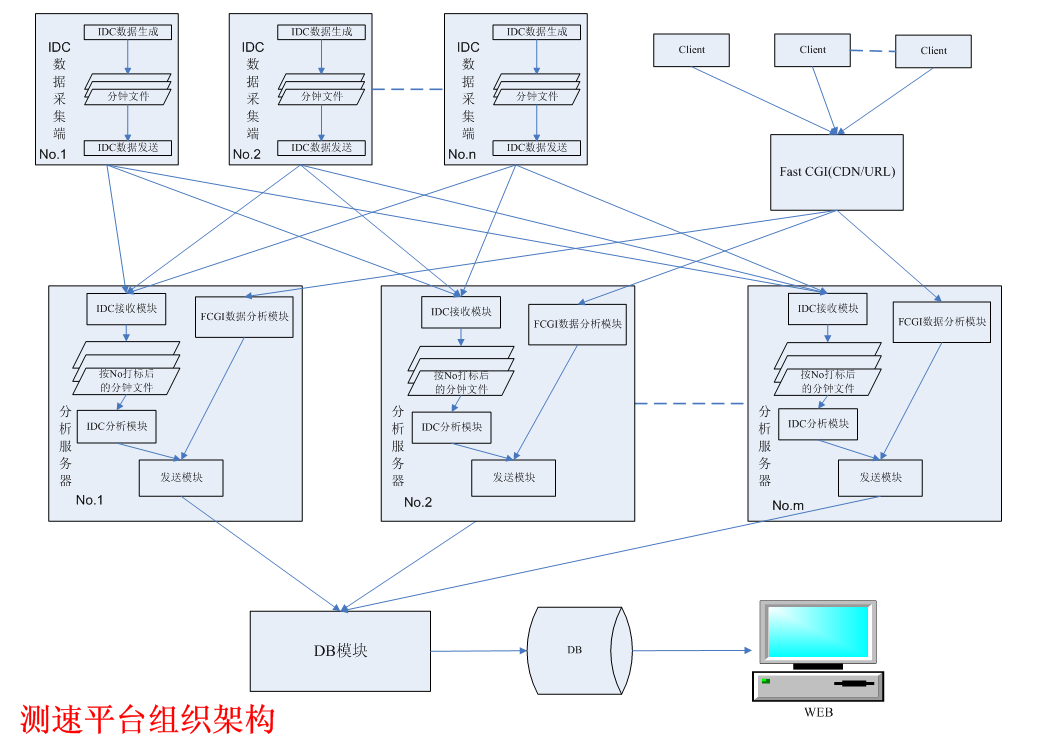
测速平台概要设计文档

## 测速平台架构



**CDN测速网络拓扑图**

IDC采集端通过TCP协议将采集到的IDC分钟文件发送到分析端；用户访问页面时以Http协议，通过DNS，将CDN测速数据或URL测速数据上传到FCGI。FCGI与分析端位于同一台服务器上，分析端计算分钟平均值后，将结果发送到DB模块，DB模块根据不同的协议和字段更新DB的不同数据库和表。

IDC采集端的功能：

每分钟探测指定的IP地址集合，将探测结果保存在分钟文件里，将分钟文件发送到分析端的接收模块(可选功能：根据域名解析服务器IP，并从IP组中选择负载最低的分析端服务器)。

FCGI的功能：

将来自Client的数据，以tcp协议传送给分析端，由于FCGI接收到的数据很多，所以FCGI只管发数据，分析端只管收数据，发完不断开连接(可选功能：根据域名解析服务器IP，并从IP组中选择负载最低的分析端服务器)。

分析端的功能：

IDC接收模块：将数据采集端发送的分钟数据按照idc机房编号存在分钟文件里，响应来自外部的负载查询请求；IDC分析模块：轮询分钟文件，计算分钟平均值，存入IDC MMAP；FCGI数据分析模块：接收来自FCGI的数据，并计算分钟平均值，存入CDN MMAP和URL MMAP；发送模块：对于IDC的MMAP，处理完一个分钟文件向DB模块发送一次，对于CDN 以及URL的MMAP，每分钟发送一次。

DB模块的功能：

接收分析端发过来的数据，并按接收数据的协议更新对应的数据库和数据表。

# CDN测速

## 数据收集

用户在登录时，上报下载速率到服务器。

需要上报的字段：

Type: 上报数据类型

UserID： 用户米米号、

ProjNo： 项目编号、

CDNIP： 客户端首先使用\*.61.com/ip获取cdn的ip、

Rate： 下载速度。

FCGI获取上报数据的Client IP(客户端IP)，连同上报的数据，发送到服务器。

服务器平均每秒钟大概需要处理2000条消息。

## DB设计

### 根据CDN IP查看

DB中分为4类表。

分钟表cdn\_rate\_ min：

记录每分钟的平均值，当收到分析端的一条消息时，根据cdn\_ip、citycode、isp和当前时间得到的分钟seq(timestamp/60)插入新记录或者更新已有记录的平均值。由于分钟表的数据量比较大，每个月自动生成一张新表。

10分钟表cdn\_rate\_10min：

记录每十分钟的平均值，当收到分析端一跳消息时，根据cdn\_ip、citycode、isp和当前时间得到的10分钟seq(timestamp/600)插入新记录或者更新已有记录的平均值。

小时表cdn\_rate\_hour：

记录每小时的平均值，当收到分析端一跳消息时，根据cdn\_ip、citycode、isp和当前时间得到的60分钟seq(timestamp/3600)插入新记录或者更新已有记录的平均值。

天表cdn\_rate\_day：

记录每天的平均值，当收到分析端一跳消息时，根据cdn\_ip、citycode、isp和当前时间得到的1440分钟seq(timestamp/86400)插入新记录或者更新已有记录的平均值。

### 根据CDN 节点查看

注：CDN 节点是同一CDN运营商下的所有CDN IP的集合。

根据CDN IP 除以256取整得到的结果将CDN IP分类，即可得到CDN节点的所有信息，查看方法同上。

### 查看某CDN 节点在不同地区运营商的结果

将CDN IP按除以256的值归类到不同的数据表中，即可在单张表中提取同一CDN节点的所有数据，然后按地区运营商分类，即可查看在不同的地区，不同的运营商对该CDN节点的平均访问速度随时间的变化曲线图。

### 根据地域查看

根据客户端的ip地址所属地域来划分，**不同省/市，一张表**，与CDN节点类似，下面的每一种划分方法，也需要4张表。

按市/区来查看，需要如下4张表：

Cdn\_rate\_city\_net\_min 用于查看一小时/天数据，数据量100\*10\*1440\*365 = 500M

Cdn\_rate\_city\_net\_l2 用于查看一周的数据

Cdn\_rate\_city\_net\_l3 用于查看一月的数据

Cdn\_rate\_city\_net\_l4 用于查看一年的数据

### 查看某地域下所有CDN IP的信息

上述表中每条记录都有一个CDN字段，标记每个客户端IP获取素材时用到的CDN IP，对于指定的地区运营商，可以按CDN IP聚类，即可查看该地区不同CDN IP的平均访问速度随时间的变化曲线图。

## 画图

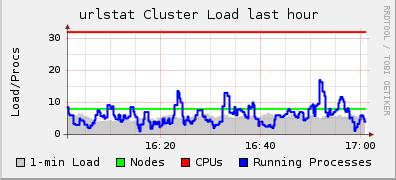
下面的曲线图，都需要查看到最近一个小时，最近一天，最近一周，最近一个月，最近一年的数据

* 各CDN节点的用户平均访问速度随时间变化的曲线图

算法：在一分钟内，对相同的CDN IP，计算 **速率和/人数**

DB中，每分钟存储1条记录，包含“速率和”和“人数”

**查看最近一个小时**的曲线时，从cdn\_rate\_min中读取最近60条数据。分12格展示，如：



**查看最近一天**时，从cdn\_rate\_min中读取最近1440条记录。

**查看最近一周**时，从cdn\_rate\_10min中读取最近144\*7条记录。

**查看最近一月**时，从cdn\_rate\_hour中读取最近24\*30条记录。

**查看最近一年**时，从cdn\_rate\_day中读取最近365条记录。

* 以省/市为单位计算CDN网络的用户平均访问速度随时间变化的曲线图

算法：根据client IP，从IP库中查到用户所属地区，精确到市级，对于北上广深，精确

到区。在一分钟内，对相同·的市/区，计算 **速率和/人数**

* 按地域运营商查看用户平均速度随时间变化的曲线图
* CDN网络总的用户平均速度随时间变化的曲线图

算法：在一分钟内，对所有上报的速率，计算 **速率和/人数**

* 查看从起始日起始时间到终止日终止时间的历史数据

# WEB页面监控

## 数据收集

客户端在J S中设置多个段点，作为一个页面阶段性加载完成的标点。使用POST向测速平台上报两类数据：

1. 连续两点之间消耗时长。
2. 整个页面加载完成消耗时长。

上报数据格式：

Userid： 用户米米号

Prj NO：项目编号

Report\_type：1，上报两点耗时；2，上报页面耗时

ID：PointID / PageID

time：耗时(ms)

## 断点图

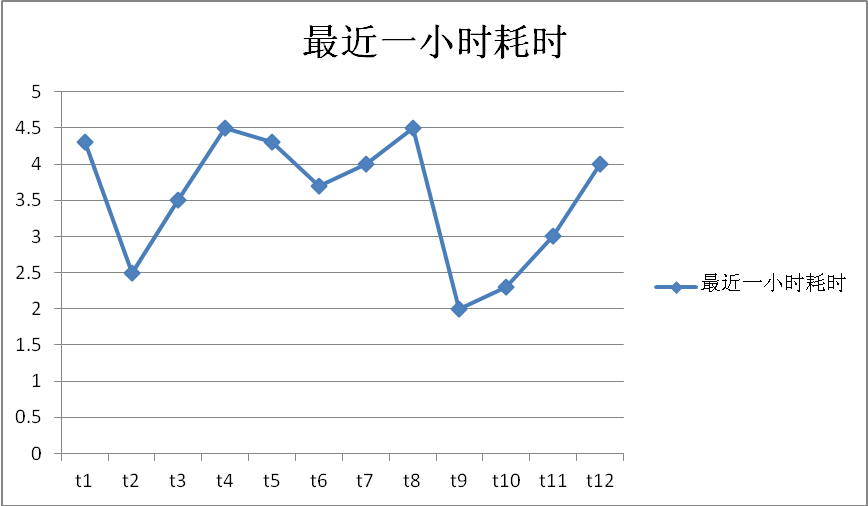
* 以（PointID，市/区，ISP）为关键字，画出具有相同关键字的平均耗时随时间变化的曲线图（耗时 / 时间）。

算法：根据client IP，从IP库中查到用户所属地区，精确到市级，对于北上广深，精确

到区。在一分钟内，对相同的（PointID，市/区，ISP）计算 **耗时和/人数。**

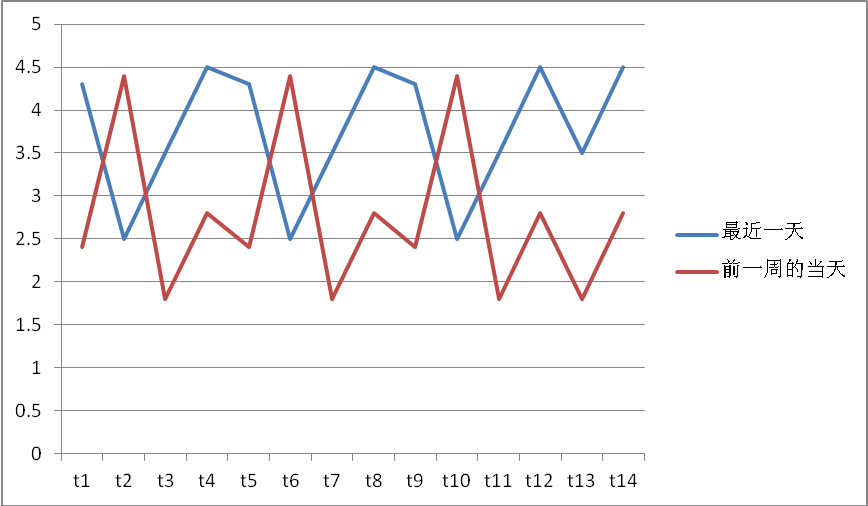
画图：以（PointID，市/区，ISP）

1，小时图：



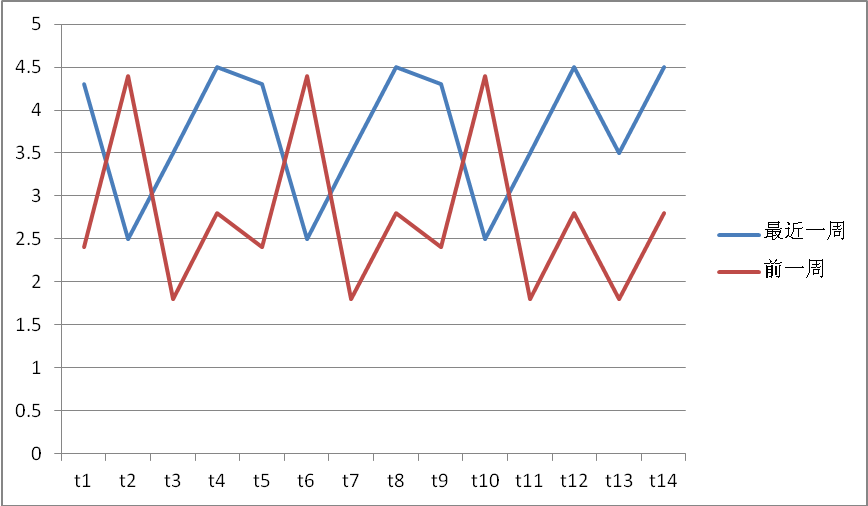
**最近一小时随时间变化的耗时图**

2，一天图/前一周的对比图：



**最近一天随时间变化的耗时图**

3，两周对比图：



**最近两周随时间变化的耗时对比图**

## 页面耗时图

* 以（PageID，市/区，ISP）为关键字，画出具有相同关键字的平均耗时随时间变化的曲线图（耗时 / 时间）。

算法：根据client IP，从IP库中查到用户所属地区，精确到市级，对于北上广深，精确

到区。在一分钟内，对相同的（PageID，市/区，ISP）计算 **耗时和/人数。**

画图：同PointID的画图。

## 断点耗时全记录

以userid为关键字，记录用户一个月内的所有断点耗时记录。方便校吧定位问题。

|  |  |  |
| --- | --- | --- |
| PointID | time | Time\_used |
| 1 | 23394904890 | 300ms |
| …… | …… | …… |

**某时间段用户50015的断点耗时记录**

# 全网探测

## 数据收集

监控数据：丢包率，延迟，跳数

方法：在每个IDC机房中，放置一台自动探测机，原则电信机房探测电信IP集合，网通机房探测网通IP集合。每半分钟探测一次，将探测结果上报到数据处理服务器。

上报数据格式：

Probe\_type：1，丢包；2，延迟；3，跳数

Net： 网通或电信

IP： 探测IP

TIME：服务器探测时间

result：探测结果

## DB设计

与CDN测速类似，有如下数据表：

main\_ip\_lost\_pkg\_min

main\_ip\_lost\_pkg\_l2

main\_ip\_lost\_pkg\_l3

main\_ip\_lost\_pkg\_l4

main\_ip\_delay\_min

main\_ip\_delay\_l2

main\_ip\_delay\_l3

main\_ip\_delay\_l4

main\_ip\_ttl\_min

main\_ip\_ttl\_l2

main\_ip\_ttl\_l3

main\_ip\_ttl\_l4

## 画图

* 全网探测全国各主要IP，对每个IP提供一个丢包率，延迟，跳数随时间变化的统计图

由运维部门配置主要IP集合，可以按照省/市来展现

* 按地域统计丢包率，延迟，跳数随时间变化的统计图

数据处理服务器在收到探测结果时，根据探测IP，从IP库中找到相应的地区，计算一分钟内，该地区平均丢包率，延迟，跳数。

地域应精确到市级，北上广深精确到区级。

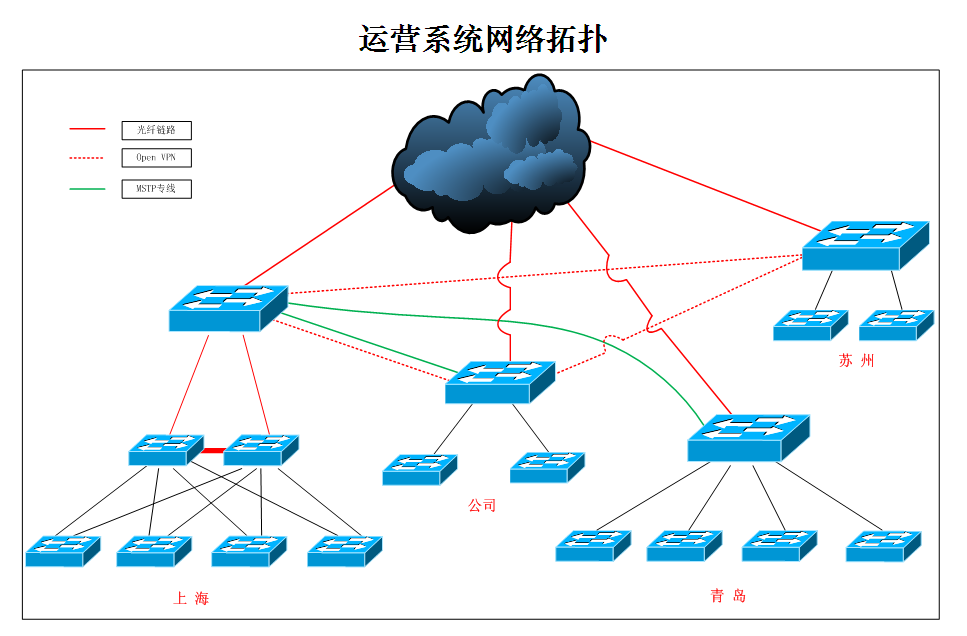
* 按城市运营商，统计丢包率，延迟，跳数随时间变化的统计图
* 以表格的形式统计出在哪一跳丢包，实时数据

根据丢包所在跳，找出所属地区，按照时间从后向前排列出丢包表格。

# IDC之间的网络状况

IDC主要有，青岛，南汇，苏州，加上公司，分三个层面来展现-----全局平均，IDC平均，单条链路。

链路分布见下图：



## 数据收集

监控数据：丢包率，延迟，跳数

方法：在每个IDC机房中，放置一台自动探测机，针对每条链路探测相应的交换机接口。每分钟探测一次，将探测结果上报到数据处理服务器。

上报数据格式：

Probe\_type：1，丢包率；2，延迟；3，跳数

link： 探测链路

IP： 探测IP

TIME：服务器探测时间

result：探测结果

## DB设计

与CDN测速类似，有如下数据表：

根据链路的不同：

Idc\_link\_xx\_lost\_pkg\_min

Idc\_link\_xx\_lost\_pkg\_l2

Idc\_link\_xx\_lost\_pkg\_l3

Idc\_link\_xx\_lost\_pkg\_l4

根据机房的不同

Idc\_ xx\_lost\_pkg\_min

Idc\_ xx\_lost\_pkg\_l2

Idc\_ xx\_lost\_pkg\_l3

Idc\_ xx\_lost\_pkg\_l4

总的网络状况：

Idc\_ lost\_pkg\_min

Idc\_ lost\_pkg\_l2

Idc\_ lost\_pkg\_l3

Idc\_lost\_pkg\_l4

## 画图

* 对于单条链路

根据每分钟的探测数据，画出丢包率，延迟，跳数随时间变化的统计图

* 对整个IDC机房做平均值统计

根据该机房的所有链路做丢包率，延迟，跳数的平均值计算，画出随时间变化的统计图

* 对公司所有IDC机房做平均值统计

根据所有链路做丢包率，延迟，跳数的平均值计算，画出随时间变化的统计图