CouchDB 迁移与查询

# 本地安装

## 安装

从<http://couchdb.apache.org/>下载最新的安装包, 按步骤安装即可.

注意: CouchDB要求安装目录不能有空格, 可以安装在c:\ProgramData文件夹, 或根目录下.

本前最新版本为CouchDB 2.3.1

## 创建管理员

打开<http://localhost:5984/_utils/index.html>, 点击左侧导航栏的Admin Party!, 创建管理员即可

# 创建用户

普通用户的信息在系统数据库\_users中保存, 如果没有这个库, 直接新建即可, CouchDB能够识别, 并自动添加一个design文件; 这个库里除design document之外的, 每一条都是单独的用户信息.

## 手动新建

在\_users库中新建一个document, 文件的ID必须是org.couchdb.user:<username>, 如下 (注意, 用户名一旦建立, 不可修改):

CouchDB 1.2.0或更高版本使用这个:

{

"\_id": "org.couchdb.user:dbreader",

"name": "dbreader",

"type": "user",

"roles": [],

"password": "plaintext\_password"

}

使用CouchDB <1.2.0时, 用户文档需要如下所示:

{

"\_id": "org.couchdb.user:dbreader",

"name": "dbreader",

"type": "user",

"roles": [],

"salt": "54935938852dd34f92c672ab31e397cedaf0946d",

"password\_sha": "42253ea4461a604f967813aaff90b139d7018806"

}

## 通过请求新建

按下例发送PUT请求 (只有管理员可以设计roles):

def main():

username = '<username>'

url = 'http://localhost:5984/\_users/org.couchdb.user:{}'

r = requests.Session()

r.headers['Accept'] = 'application/json'

r.headers['Content-Type'] = 'application/json'

body = {

'name': username,

'password': 'a2data',

'roles': [],

'type': 'user'

}

html = r.put(url.format(username), data=json.dumps(body))

print(html)

print(json.dumps(html.json(), indent=4, ensure\_ascii=False))

## 更新密码

与Section 1.2完全一样, 另外添加一个revision版本号即可. 即:

r.headers['If-Match'] = '1-3409a4c39c0954ce8e3d5492e2e58196'

可能需要admin权限, 直接在url里加入就行, 如:

url = 'http://admin:password@localhost:5984/\_users/org.couchdb.user:{}'

## 参考文档

<http://docs.couchdb.org/en/stable/intro/security.html>

# 权限管理

## 字段说明

通过GET /<database>/\_security 获取, 返回两个字段admins和members, 显示具备对应权限的用户或角色列表:

* **members**: 读写doc, 但不可以design doc
* **admins**: 包括members的权限, 另外加上design doc, 添加/删除databases / members, 设置数据库revision上限

members和admins对象都包含两个数组类型的数据:

* **names**: 用户名列表
* **roles**: 用户角色列表

## 权限说明

如果names和roles两个数组均为空, 代表这个数据库没有admins或members, 即:

* 没有admins, 即只有server admins更新design doc或其他管理员级别的修改
* 没有members, 代表任意用户均可以读写doc (除了design doc)

## 参考网址

<http://docs.couchdb.org/en/stable/api/database/security.html>

# 数据迁移

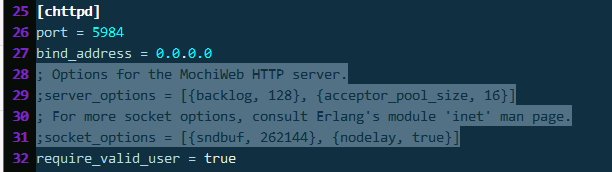
本机的IP地址, 可以通过在CMD中运行ipconfig获取

## 服务器远程访问设置

### 直接修改配置文件

在CouchDB的安装目录, %COUCHDB\_PATH%/etc/couchdb/local.ini, 找到[chttpd], 修改bind\_address (去掉开头的分号), 同时设置require\_valid\_user=true,

如下:



### 在Fauxton界面中修改

点击左侧导航栏Configuration, 找到对应的key, 单击即可以修改, 完成后点击绿色确定键即可

## 重启服务器

除了杀掉CouchDB的进程重启, 也可以使用如下命:

* net.exe stop “apache couchdb”
* net.exe start “apache couchdb”

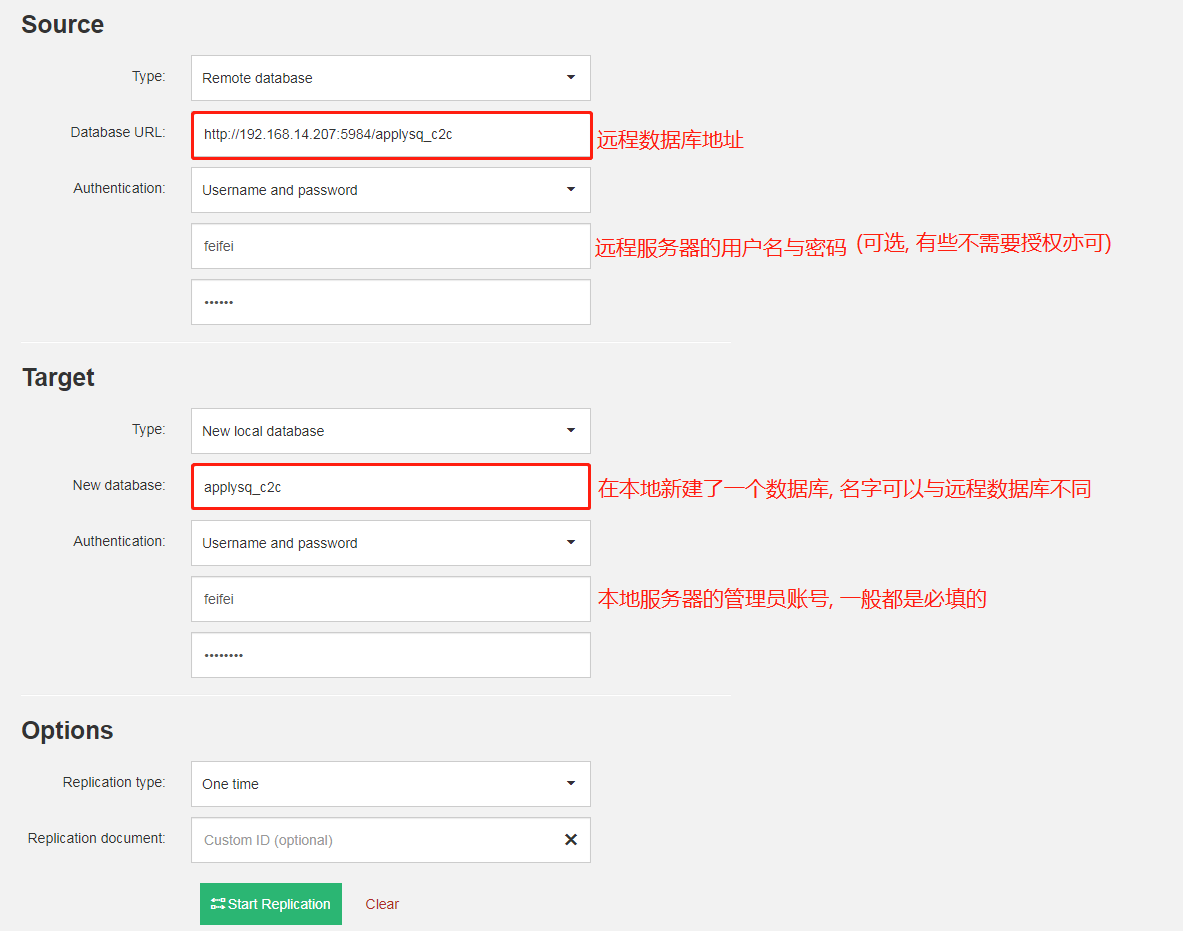
重启CouchDB之后, 即可在内网任何一台PC上访问.

可以通过IP地址或是机器名称访问, 以下两种方法均可以 (需要登录):

* http://192.168.14.207:5984/\_utils/
* http://win-svr-2016-dc:5984/\_utils/

## 数据迁移

打开<http://localhost:5984/_utils/index.html>, 点击左侧导航栏的Replication, 点击右上角New Replication:



所需时间取决于要迁移的数据库大小

# 数据查询

以下查询要求目标CouchDB至少2.0.0以上版本, 否则会报400错误:

Referer header required.

## URL

http://127.0.0.1:5984/<*dbname*> /\_find

## 请求

采用POST提交, POST的参数是一个JSON, 要求详见 <http://localhost:5984/_utils/docs/api/database/find.html#find-combination-operators>

如下例, 是查询\_id = TOWER: 27116的记录:

* 行20-21设置请求头
* 行23-30是POST提交的数据, 其中:
  + selector设置筛选的参数  
    如果是嵌套字段, 可以用点连接, 如name.age这样  
    详细可参考 <http://localhost:5984/_utils/docs/api/database/find.html#find-selectors>
  + fields设置返回的字段
  + sort返回数据的排序
  + execution\_stats执行的简单统计
* 行31是执行POST请求, 注意传的是一个JSON的字符串, 而非字典
* 行33-34打印返回的结果



# Excel连接CouchDB

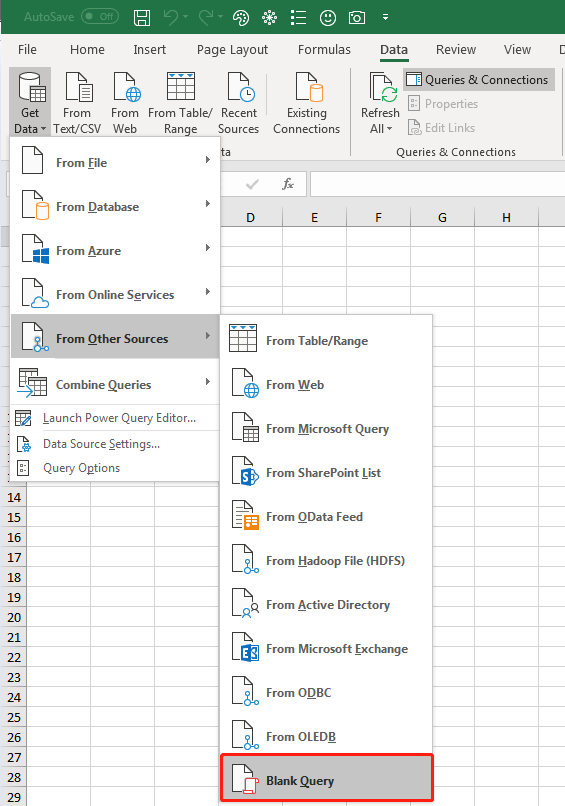
使用Power Query向CouchDB发送POST请求, 进而获取数据

## 版本要求

至少Excel 2013及以上版本

## 创建查询

点击Data 🡪 Get Data 🡪 From Other Sources 🡪 Blank Query, 新建一个空白查询



## 修改查询

在弹出的Power Query窗口, 点击View 🡪 Advanced Editor. 然后, 我们以Section 3.2为例, 将以下代码全部粘至编辑窗口:

let

url = "http://127.0.0.1:5984/data\_log/\_find",

body = "{

""selector"": {

""\_id"": ""TOWERTASK:27116""

},

""fields"": [""\_id"",""\_rev"",""score""],

""sort"": [{""\_id"": ""asc""}],

""execution\_stats"": **true**

}",

Source = Json.Document(Web.Contents(url,

[

Headers = [#"Accept"="application/json", #"Content-Type"="application/json"],

Content = Text.ToBinary(body)

]

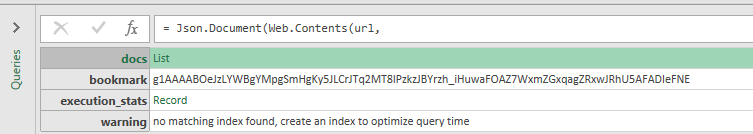
))

in

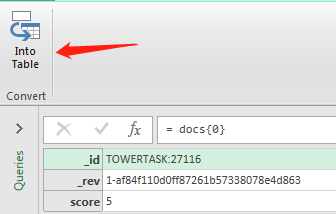
Source

|  |
| --- |
| 特别提示: JSON里面布尔值是小写, 跟Python一样写成首字母大写会报错: |

点击确定, 即可得到JSON结果, 如下图:



其中, docs字段为查询得到的数据内容, 点击List, 会展开第一个条目:



如果需要把它们转化为普通的表单, 点击左上角的Into Table即可

## 导出数据

在Power Query中可以对数据作各种清洗, 这里不详细展开.

数据清洗完成后, 可以导出至sheet或是Power Pivot数据模型, 作进一步的分析

使用这种方法创建的工作簿可以作为模版使用, 每次只需要更新查询, 即可同步更新数据