

Роман

Вопрос 1

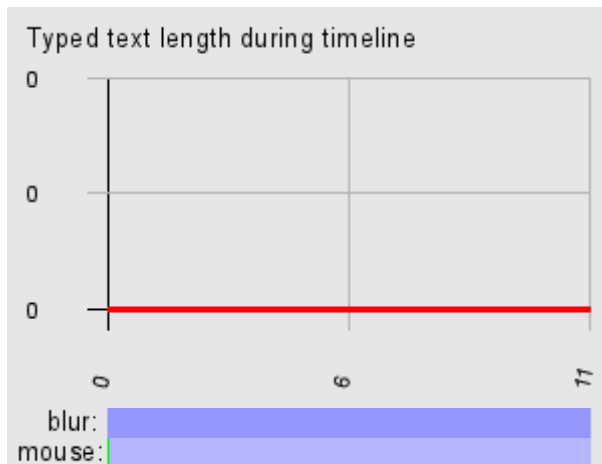
Знакомы ли Вы и работали ли с CoffeeScript?

Если да, расскажите, пожалуйста, подробно, в каких проектах Вы его использовали и что реализовали.

Правильный ответ

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36



Ответ за 11 секунд(ы), открыт в 11:34:06 21.06.2016

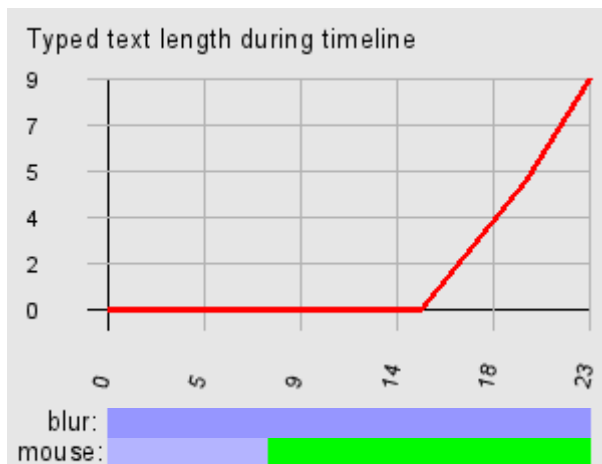
Вопрос 2

Какой уровень дохода для Вас является минимально допустимым, а какой - комфортным?

Правильный ответ

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36



Ответ за 23 секунд(ы), открыт в 10:26:16 24.06.2016

75к
100к

Вопрос 3

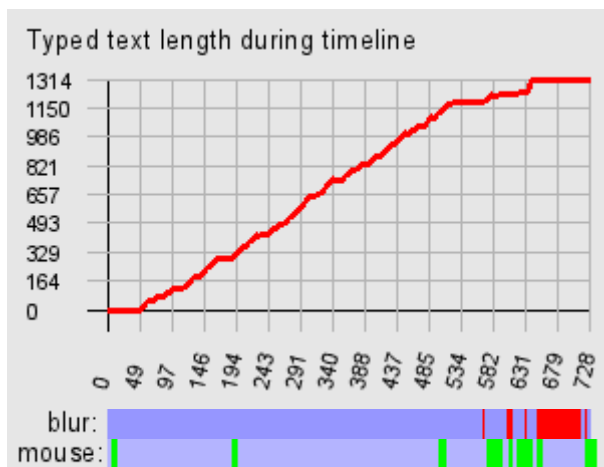
Есть ли у Вас опыт создания или участия в разработке высоконагруженных систем?

Расскажите, пожалуйста, о нём подробнее.

Правильный ответ

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36



Ответ за 12 минут(ы), открыт в 10:26:39 24.06.2016

Был проект с несколькими сотнями тысяч запросов в час. Использовали кэш, переписали некоторую ЦПзатратную логику на Cython. Вот результат <https://pp.vk.me/c630119/v630119967/22134/nSDFU3oiNbQ.jpg>

Был проект с постоянным инсертом в БД, БД висела на HDD и не справлялась. Было принято решение добавить прослойку в виду Redis'а, т.к. потеря данных не была критичной. Сначала все данные шли в Redis, редис был немного переконфигурирован на более редкие сохранения (данных было несколько тысяч записей в секунду). Там они "отстаивались", т.к. были дубликаты данных, в результате, в редисе просто инкрементировался счётчик. Раз в полчаса воркер обходил редис (да, я знаю, что это дорогая $O(N)$ операция, но другого выхода я не видел) и собирал те записи, у которых счётчик достиг определённого уровня и складывал их в БД. Чтобы в редисе не залеживались записи со счётчиком, который не достиг "уровня перехода в БД", было добавлено поле "last_updated" и вчерашние записи шли в БД независимо от счётчика. Да, можно было бы продумать более плавную нагрузку на БД, но этого хватило.

Был проект, который не выдерживал постоянный поток запросов, а на ab тесте выдавал всего лишь 5 RPS. Проект был не мой, я не стал оптимизировать SQL, а просто поставил кэш с большим TTL и инвалидацией на сохранение. RPS поднялся до 170.

Вопрос 4

Напишите, пожалуйста, что будет напечатано в результате выполнения следующего кода:

```

1 def func(type_='s'):
2     if type_ == 's':
3         return 'Mark'
4     elif type_ == 'i':
5         return 20
6
7 def dec(func, type_):
8     x = 8
9     def wrapper():
10        value = func(type_)
11        if isinstance(value, int):
12            return value * x
13        elif isinstance(value, basestring):
14            return 'Hi ' + value
15    return wrapper
16
17 print dec(func, 'i')()

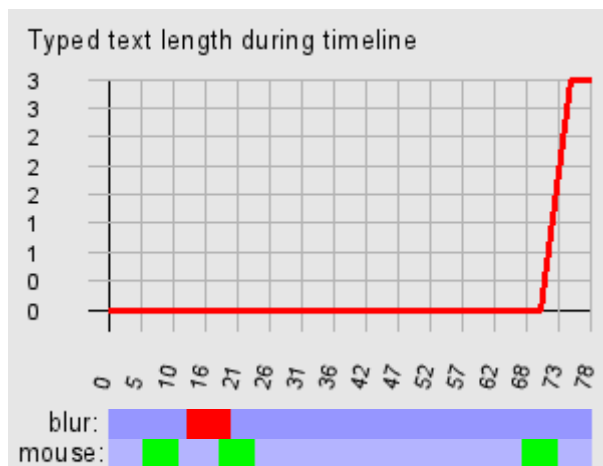
```

Правильный ответ

160

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36



Ответ за 1 мин 18 с, открыт в 10:38:47 24.06.2016

160

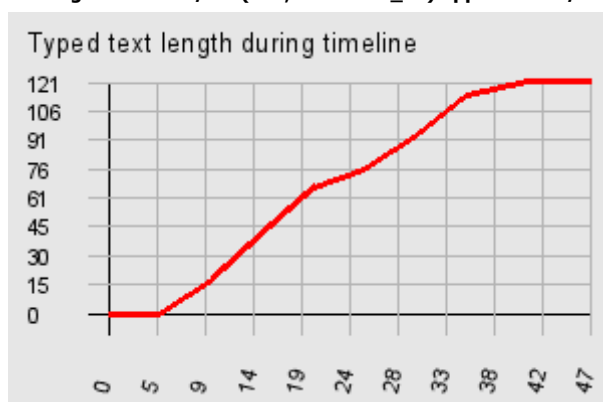
Вопрос 5

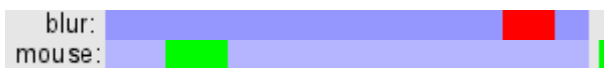
Уточните, пожалуйста, в связи с чем Вы решили сменить место работы?

Правильный ответ

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36





Ответ за 47 секунд(ы), открыт в 10:40:05 24.06.2016

В связи с тем, что работа была проектная и проект поставили на паузу из-за отсутствия вёрстки. Она очень сложная, правда.

Вопрос 6

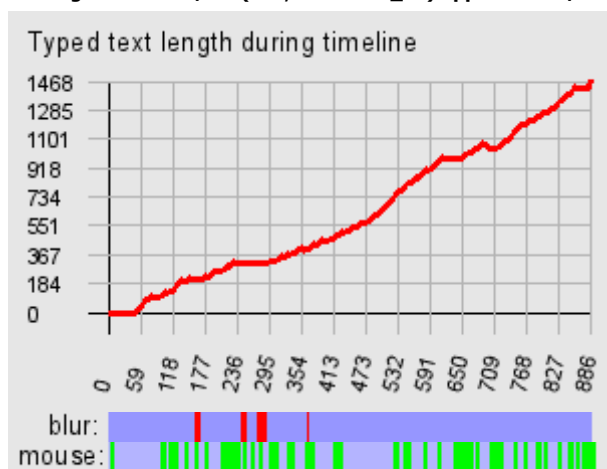
Представьте, что у Вас есть сайт, состоящий примерно из 300-400 страниц. Напишите на Python скрипт, который будет анализировать сайт и выводить перечень ссылок, находящихся на страницах сайта и ведущих на несуществующие страницы. На входе программы - адрес сайта.

Поскольку в данной задаче может быть множество деталей и "подводных камней", нужно написать первую версию скрипта, которая, на Ваш взгляд, должна работать, а также перечень возможных улучшений, которые можно было бы внести в будущем.

Правильный ответ

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36



Ответ за 14 минут(ы), открыт в 10:40:52 24.06.2016

Что-то вроде такого:

```
from bs4 import BeautifulSoup as BS
import requests
```

```
def invalid_links(url, session=None, content=None):
    if session is None:
        session = requests.session()
    if content is None:
        content = session.get(url).content
    soup = BS(content)

    result = []

    for link in soup.find_all('a', href=True):
        link_url = link['href']
        if link_url in result:
            continue
        r = session.get(link_url)
        if r.status_code != 200:
            result.append(link_url)
        else:
            content = r.content
            result.extend(invalid_links(link_url, session, content))
```

```
return result
```

Скрипт не тестированный, писался прямо в этот недо-блокнот, так что, всё, что отсюда можно увидеть - это идею.

Очевидно, что все запросы нужно обернуть в `try..except` и добавить им таймаут.

Вторым улучшением будет переход на асинхрон (`eventlet` или `aiohttp`).

Не знаю, какой именно фронтэнд, может быть, это SPA, тогда использовать Selenium будет просто необходимо вместо BS и requests.

Избавиться от циклических ссылок, сохранением в какое-то множество урлов уже посещённых страниц (мы же предполагаем, что GET запрос на одну и ту же строку адреса выдаёт одинаковый ответ?). Но это нужно переписывать в класс, либо перекидывать расшаренную переменную (при условии, что у нас неплюский тип данных), но это, я считаю, является дурным тоном.

Я имею ввиду это:

```
def something(lst=[]):
```

Вопрос 7

Даны две SQL-таблицы:

employee - сотрудник

- id
- first_name - имя
- last_name - фамилия
- salary - размер зарплаты
- department_id - отдел

department - отдел

- id
- name - название отдела

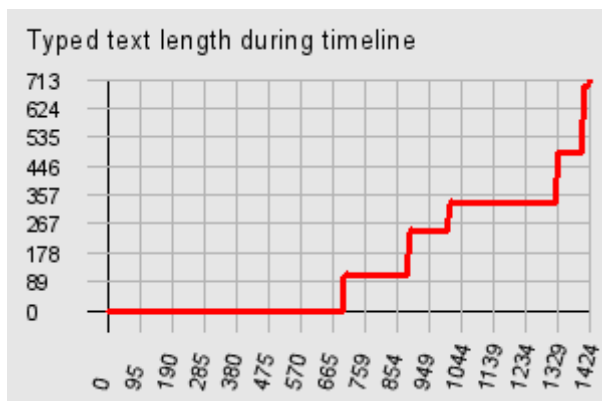
Напишите запросы к ним, которые будут выводить:

1. Список имён и фамилий всех сотрудников с названиями отделов, к которым каждый из них привязан
2. Список отделов с количеством сотрудников в каждом из них
3. Список отделов, в которых нет сотрудников
4. Список отделов, к которым привязано более двух сотрудников
5. Список отделов с указанием средней и максимальной зарплаты в каждом из них, отсортированный по убыванию средней зарплаты

Правильный ответ

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36





Ответ за 23 минут(ы), открыт в 10:55:38 24.06.2016

```
select first_name, last_name, d.name from
  employee inner join department as d on department_id = d.id;
```

```
select name, count(1) as cnt from
  department left join employee
  on department.id = employee.department_id
  group by department.id;
```

```
select name from department where id not in (select department_id from
  employee);
```

```
select name, count(1) as cnt
  from department
  left join employee
  on department.id = employee.department_id
  group by department.id
  having cnt > 2;
```

```
select name, Avg(salary) as avg_salary, Max(salary) as max_salary
  from department
  left join employee
  on department.id = employee.department_id
  group by department.id
  order by avg_salary desc;
```

Вопрос 8

Что будет напечатано в результате выполнения следующего кода?

```
1 class Parent(object):
2     x = 1
3
4 class Child1(Parent):
5     pass
6
7 class Child2(Parent):
8     pass
9
10 print Parent.x, Child1.x, Child2.x
11 Child1.x = 2
12 print Parent.x, Child1.x, Child2.x
13 Parent.x = 3
14 print Parent.x, Child1.x, Child2.x
```

Правильный ответ

The output of the above code will be:

```
1 1 1
1 2 1
3 2 3
```

What confuses or surprises many about this is that the last line of output is 3 2 3 rather than 3 2 1. Why does changing the value of Parent.x also change the value of Child2.x, but at the same time not change the value of Child1.x?

The key to the answer is that, in Python, class variables are internally handled as dictionaries. If a

variable name is not found in the dictionary of the current class, the class hierarchy (i.e., its parent classes) are searched until the referenced variable name is found (if the referenced variable name is not found in the class itself or anywhere in its hierarchy, an `AttributeError` occurs).

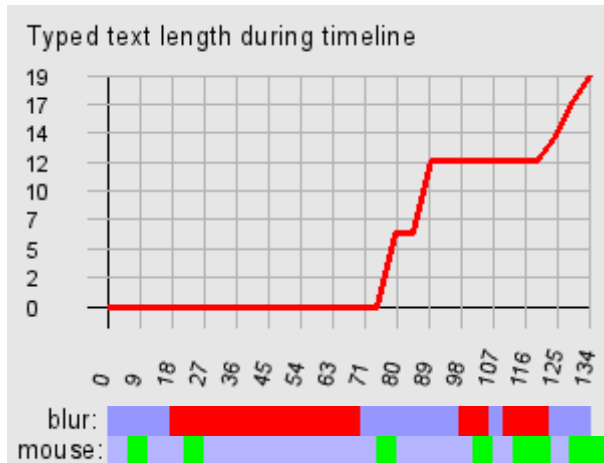
Therefore, setting `x = 1` in the Parent class makes the class variable `x` (with a value of 1) referenceable in that class and any of its children. That's why the first print statement outputs `1 1 1`.

Subsequently, if any of its child classes overrides that value (for example, when we execute the statement `Child1.x = 2`), then the value is changed in that child only. That's why the second print statement outputs `1 2 1`.

Finally, if the value is then changed in the Parent (for example, when we execute the statement `Parent.x = 3`), that change is reflected also by any children that have not yet overridden the value (which in this case would be `Child2`). That's why the third print statement outputs `3 2 3`.

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36



Ответ за 2 мин 14 с, открыт в 11:19:22 24.06.2016

1 1 1
1 2 1
3 2 3

Вопрос 9

Насколько хорошо Вы знакомы с фреймворком Flask?

Много ли проектов Вы на нём реализовали?

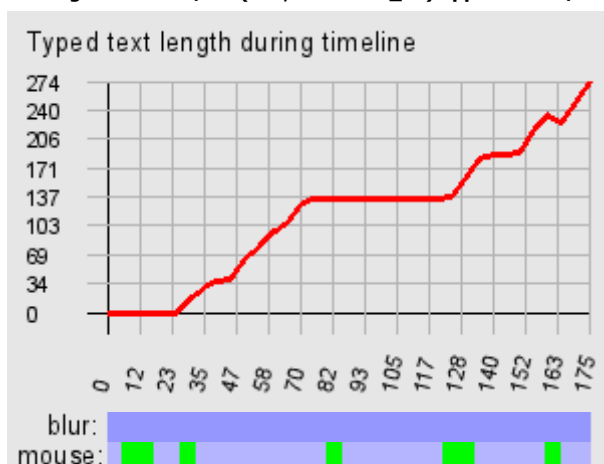
С какими основными особенностями и модулями фреймворка Вы познакомились?

Расскажите, пожалуйста, кратко о самых сложных проектах на Flask, с которыми Вам приходилось работать или которые Вы реализовали.

Правильный ответ

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36



Ответ за 2 мин 55 с, открыт в 11:21:36 24.06.2016

Знаком немного.

Пока хватало Django, хотя первым фреймом на питоне я изучал именно Flask.

Flask нужен в том случае, если в Django не устраивает всё: ORM, сессии, авторизация, формы и т.д.

Flask нужно очень сильно допиливать самому, чтобы комфортно им пользоваться.

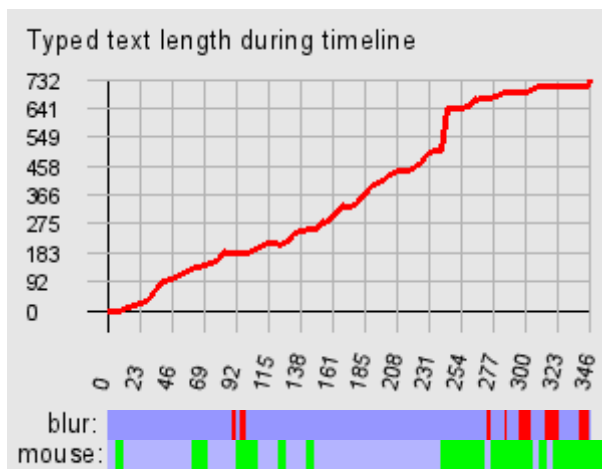
Вопрос 10

Спроектируйте структуру SQL-таблицы для хранения дерева папок.

Правильный ответ

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36



Ответ за 5 мин 46 с, открыт в 11:24:31 24.06.2016

Буду использовать самый простой SQLite синтаксис.

```
create table folder (  
    id integer primary key,  
    name text,  
    parent_id integer,  
    foreign key(parent_id) references folder(id)  
);
```

Вообще, этого хватит и это будет являться нормальной формой. Но запросы по древовидным структурам по такой таблице будут очень неоптимальные, поэтому для хранения древовидных структур я обычно использую MPTT. К тому же, в Django есть удобный модуль для этого.

Да, MPTT долгий на инсерты, но зато самый быстрый на селекты.

```
create table folder (  
    id integer primary key,  
    name text,  
    parent_id integer,  
    tree_id integer,  
    lft integer,  
    rght integer,  
    level integer,
```



```
foreign key(parent_id) references folder(id)
);
```

Вопрос 11

Взгляните на SQL-запрос:

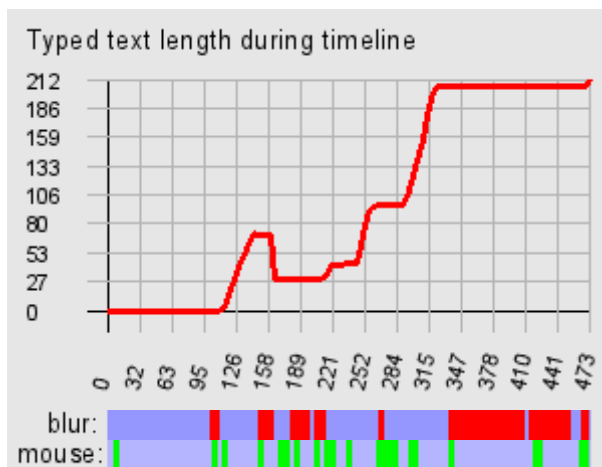
```
1 SELECT DISTINCT name, id
2 FROM authors
3 WHERE room-id IN (
4     SELECT id
5     FROM rooms
6     WHERE levelId=7
7 )
```

Есть ли в нём какие-то проблемы? Если да, то какие и как их исправить? Напишите новый запрос, в котором не будет данных проблем.

Правильный ответ

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36



Ответ за 7 мин 53 с, открыт в 11:30:17 24.06.2016

- 1) room-id написано неверно.
- 2) Стилистика. Обычно используют snake_case для именования столбцов

```
SELECT DISTINCT name, id
FROM authors
WHERE room_id IN (
    SELECT id FROM rooms WHERE level_id=7
);
```

Вопрос 12

Напишите, пожалуйста, что будет напечатано в результате выполнения следующего кода:

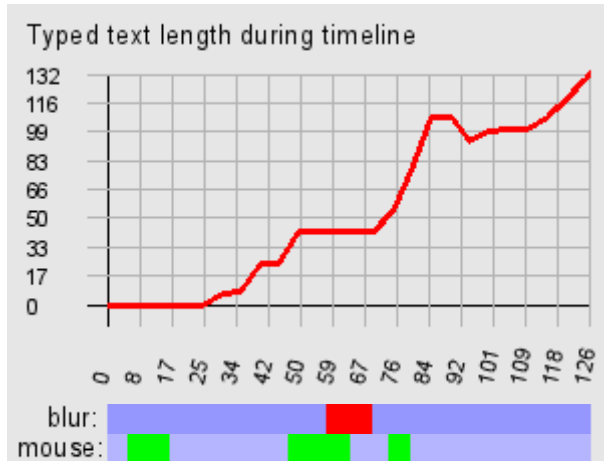
```
1 some_guy = 'Fred'
2
3 first_names = []
4 first_names.append(some_guy)
5
6 another_list_of_names = first_names
7 another_list_of_names.append('George')
8 some_guy = 'Bill'
9
10 print (some_guy, first_names, another_list_of_names)
```

Правильный ответ

Bill ['Fred', 'George'] ['Fred', 'George']

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36



Ответ за 2 мин 6 с, открыт в 11:38:10 24.06.2016

Bill ['Fred', 'George'] ['Fred', 'George']

Строки - плоские структуры и не передаются, как ссылки.
Списки передаются как ссылки.

Вопрос 13

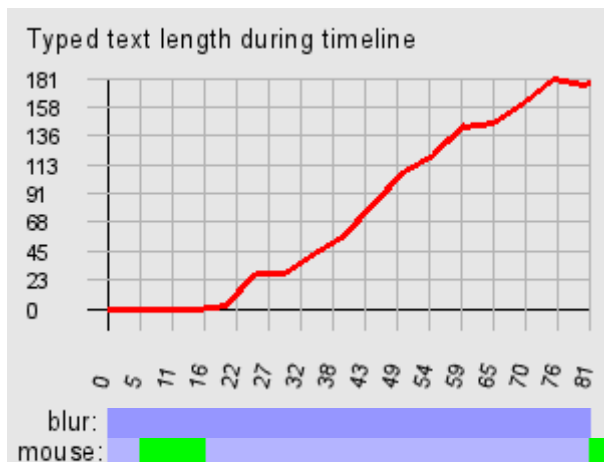
Если Вы в настоящий момент проживаете не в Москве, планируете ли Вы переезд и готовы ли будете переехать в Москву в случае получения оффера?

Если готовы или планируете, то в какие сроки?

Правильный ответ

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36



Ответ за 1 мин 21 с, открыт в 11:40:16 24.06.2016

Возможно, после длительного сотрудничества. К примеру, год.
Иначе нет никакого смысла бросать всё и ехать, если мы, в итоге, не сработаемся и это выяснится через пару месяцев.

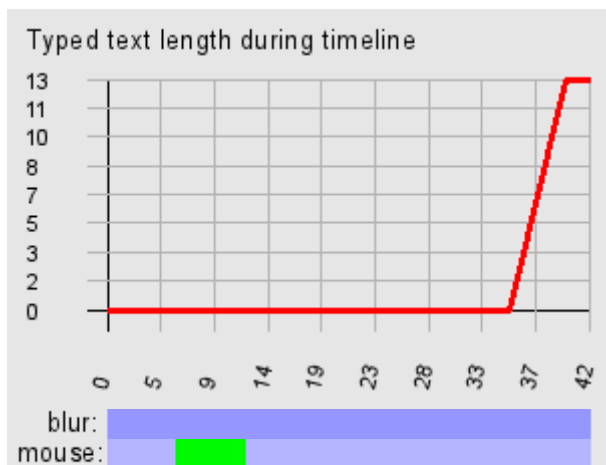
Вопрос 14

Как быстро Вы сможете приступить к работе в случае нашего обоюдного положительного решения о сотрудничестве?

Правильный ответ

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36



Ответ за 42 секунд(ы), открыт в 11:41:37 24.06.2016

Прямо сейчас.

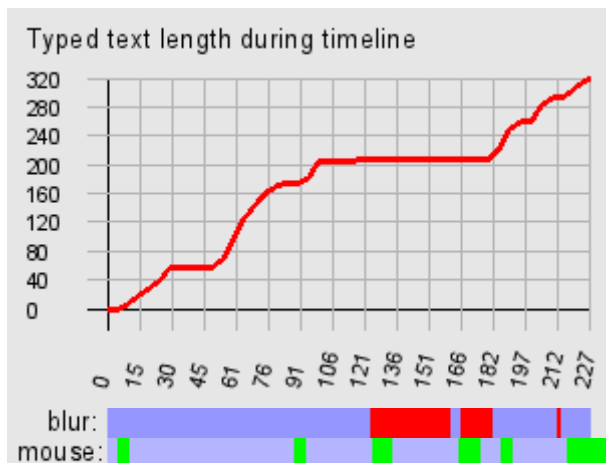
Вопрос 15

В какой мере и с какими NoSQL-хранилищами Вы знакомы? Расскажите, пожалуйста, подробно.

Правильный ответ

Statistics

User agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/51.0.2704.103 Safari/537.36



Ответ за 3 мин 47 с, открыт в 11:42:19 24.06.2016

Redis. Пользуюсь почти везде и часто. Очень люблю редис.
memcached. Использовал раньше в качестве кэша, но его быстро вытеснил редис со своими типами данных и дампами на диск.
MongoDB. Пользовался немного.
Планирую в ближайшем времени познакомиться с Cassandra. Очень привлекает кластеризация.
PostgreSQL + hstore :)