



Поиск для сайта:

зачем он нужен, как с ним работать,
как и зачем его улучшать

Дмитрий Вульбрун

Менеджер проектов

Я.Субботник, Минск, 2 июня 2012 года

Поиск для сайта

Возможности Яндекс.Поиска для сайта



Гибкие настройки

Настройте результаты поиска в соответствии с дизайном сайта



Поисковые подсказки

Используйте подсказки, учитывающие содержание вашего сайта



Статистика поисковых запросов

Узнайте, что искали и нашли посетители на вашем сайте



Уточнения

Уточнения помогают посетителям вашего сайта ограничить область поиска



Поисковые темы

Организуйте поиск по группе сайтов определенной тематики



Рекламная сеть Яндекса

Зарабатывайте на размещении контекстной рекламы в результатах поиска



Индексирование сайта

Настройте оперативную индексацию сайта поисковым роботом Яндекса



Плагины для популярных CMS

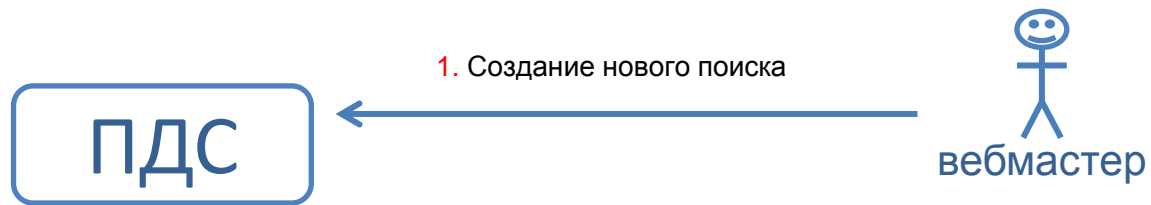
[Плагины-пингеры](#) автоматически уведомляют сервис "Поиск для сайта" о появлении новых документов для индексации

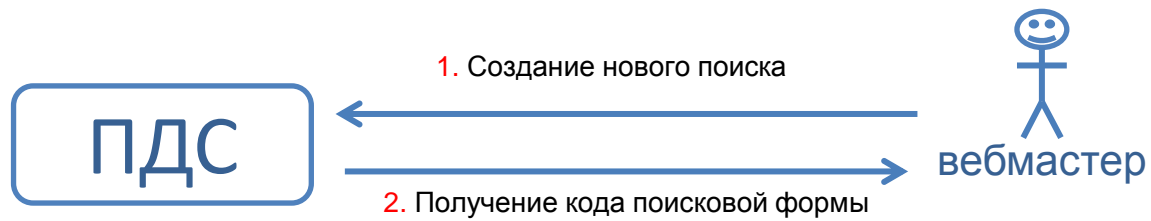


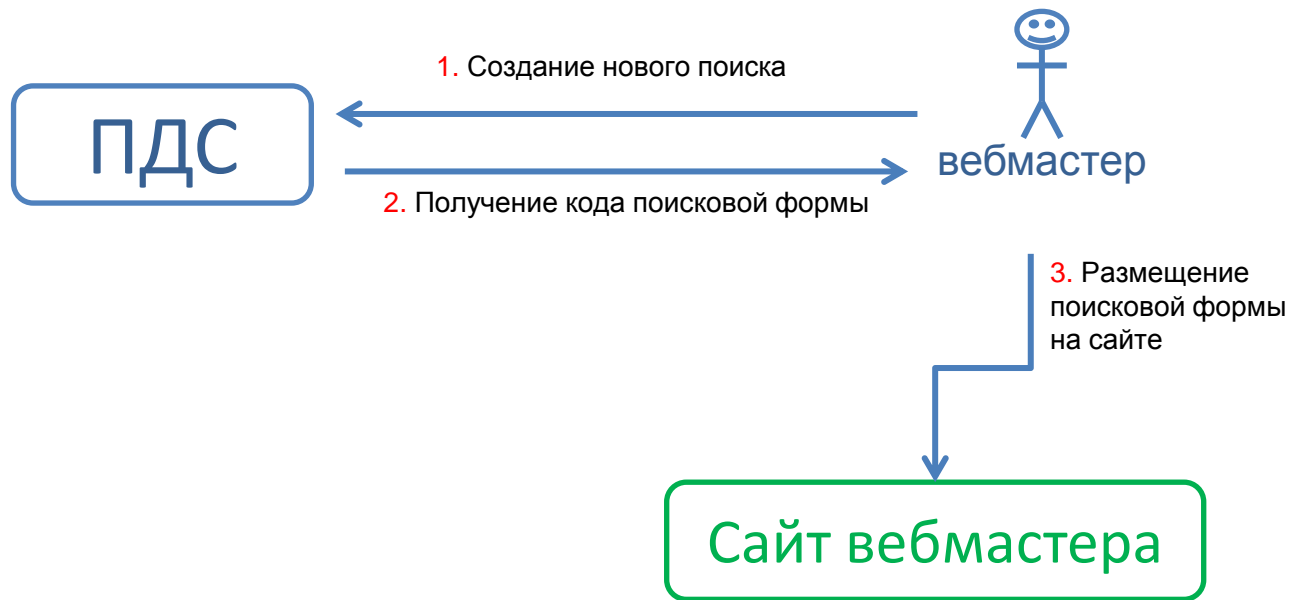
О чем мы сегодня поговорим

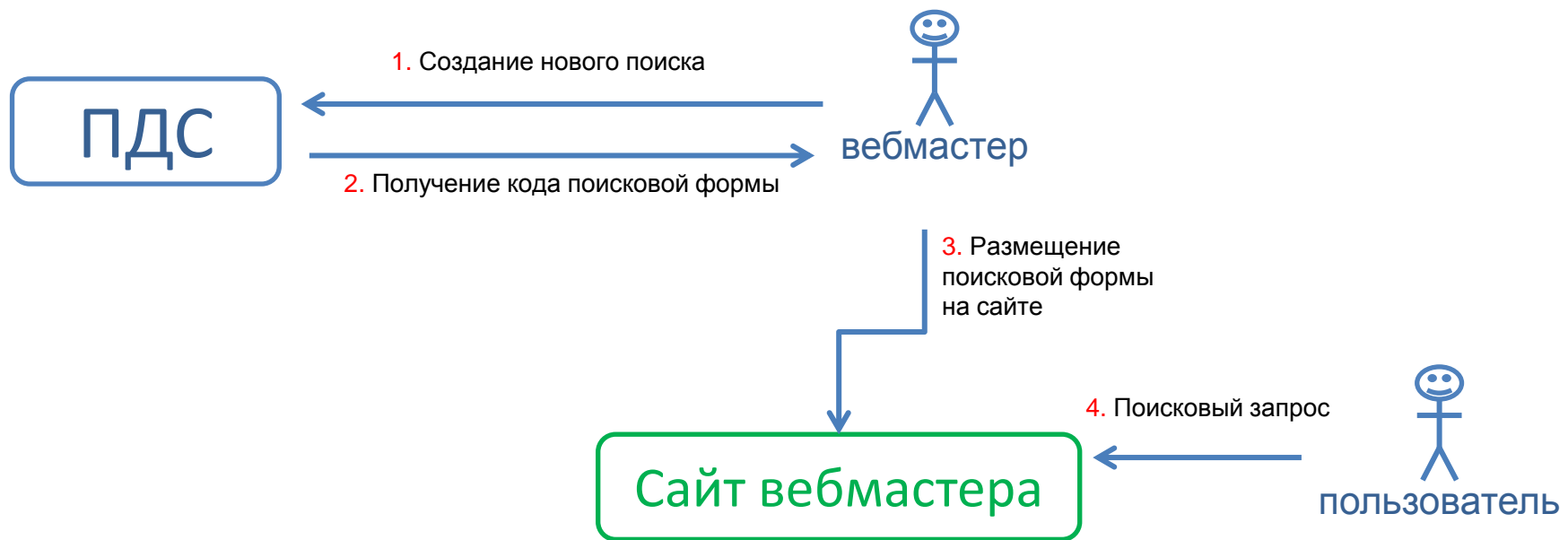
- Схема работы сервиса Яндекс.Поиск для сайта (ПДС)
- HTML-код поисковой формы и результатов поиска
- Уточнения поиска. Архитектура и принцип работы
- Индексирование по запросу. Пингеры

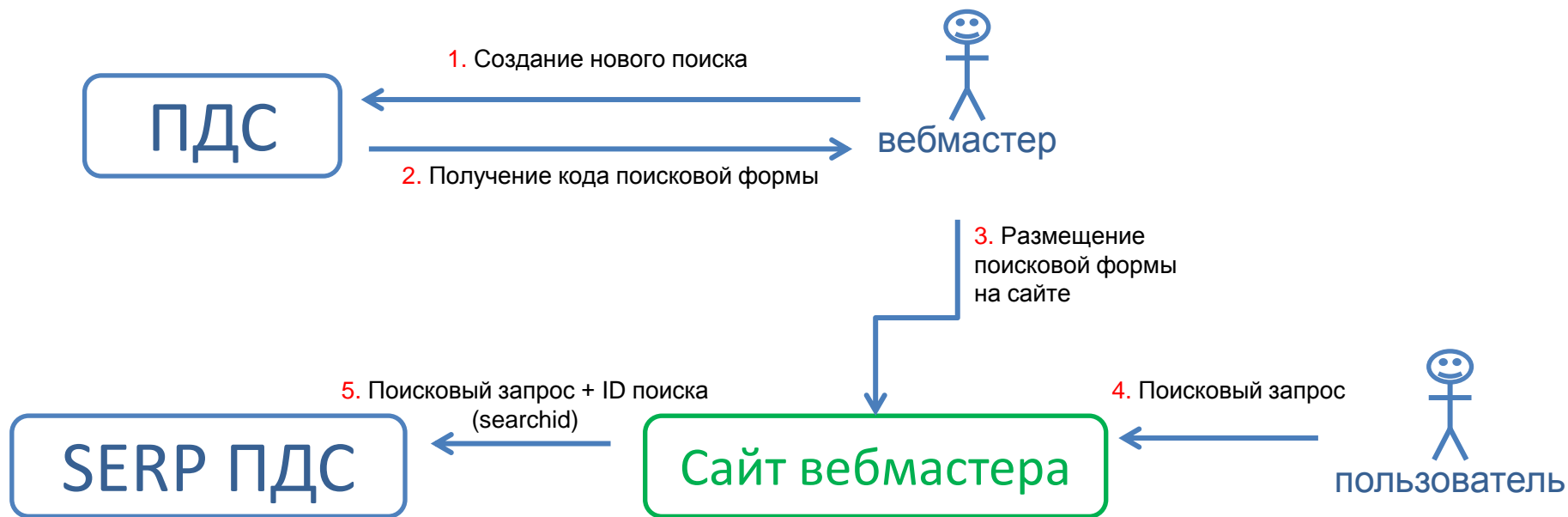
Общая схема работы сервиса



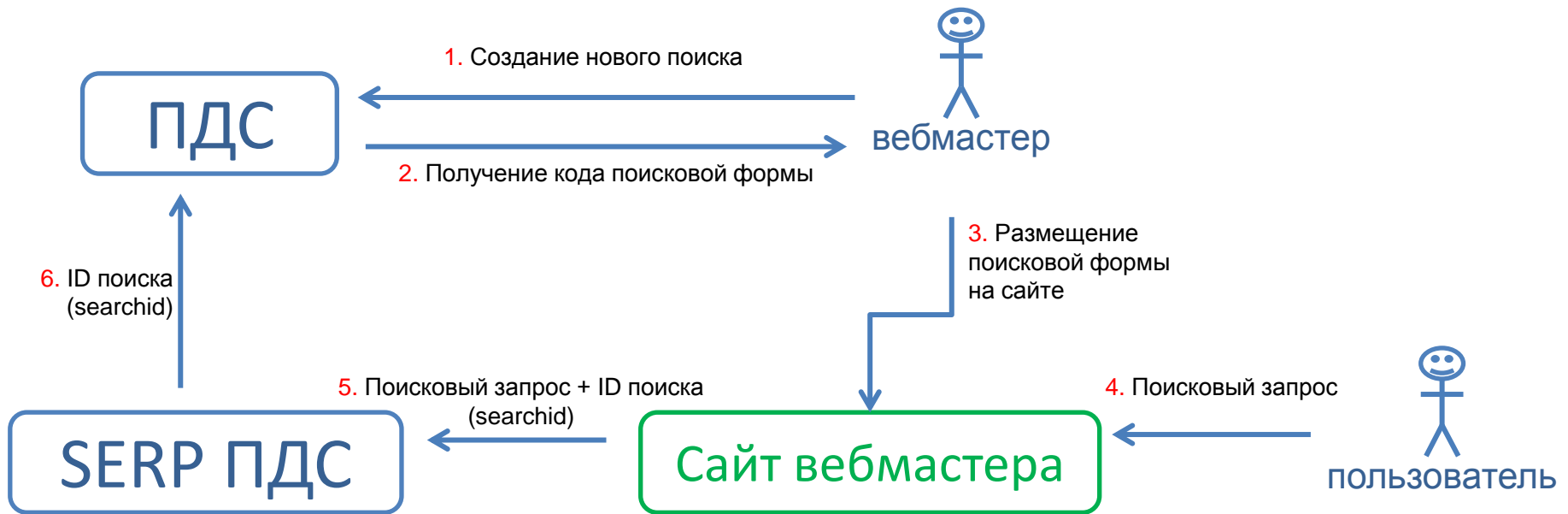




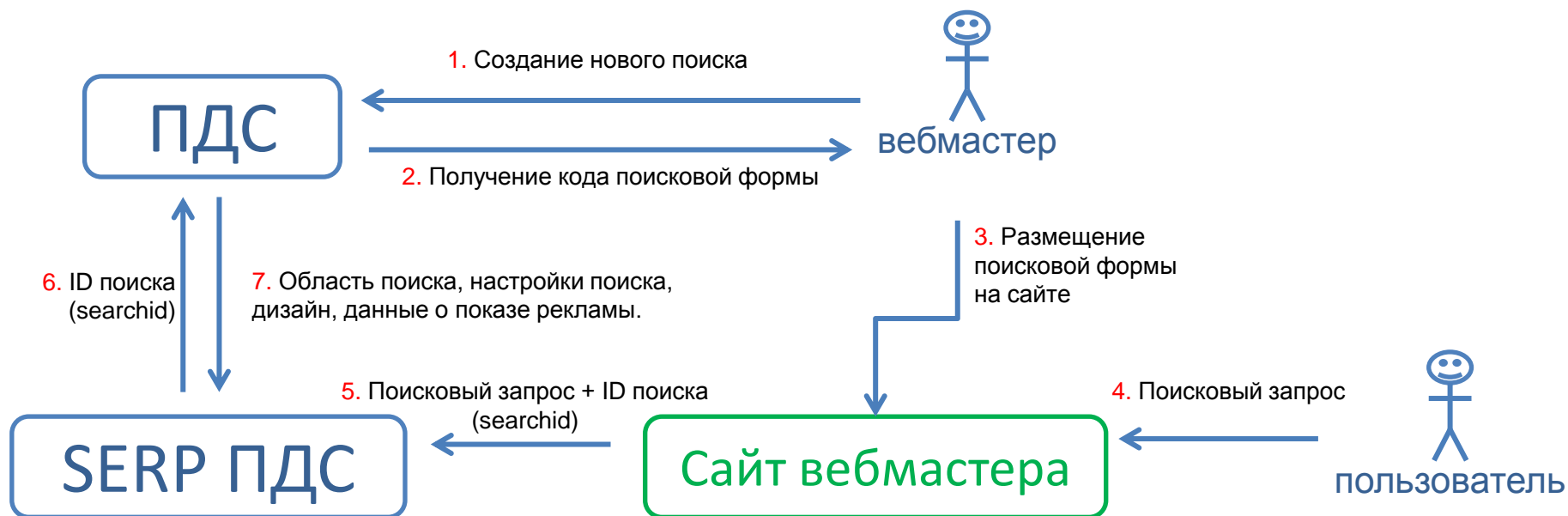




* SERP - Search engine results page



* SERP - Search engine results page



* SERP - Search engine results page





Поиск для сайта







HTML-код поисковой формы и результатов поиска

КОД ДЛЯ ВСТАВКИ НА САЙТ

Выберите кодировку вашего сайта

Выберите язык

Разместите код поисковой формы на вашем сайте

```
<div class="ya-site-form ya-site-form_initied_no" onclick="return {'bg': '#3399ff', 'target':  
'_self', 'language': 'ru', 'suggest': true, 'tld': 'ru', 'site_suggest': true, 'action':  
'http://mysite.ru/poisk.html', 'webopt': true, 'fontsize': 12, 'arrow': true, 'fg': '#ffffff',  
'searchid': '1902945', 'logo': 'rb', 'websearch': false, 'type': 0}"><form  
action="http://yandex.ru/sitesearch" method="get" target="_self"><input type="hidden"
```

Разместите этот код на странице <http://mysite.ru/poisk.html>

```
<div id="ya-site-results" onclick="return {'tld': 'ru', 'language': 'ru', 'encoding': ''}">  
</div><script type="text/javascript">(function(w,d,c){var  
s=d.createElement('script'),h=d.getElementsByTagName('script')
```

Код поисковой формы

```
<div class="ya-site-form ya-site-form_inited_no"
onclick="return {
    'searchid': '1902945',
    'action': 'http://mysite.ru/poisk.html',
    'type': 0,
    'arrow': true,
    'logo': 'rb',
    'webopt': true,
    'websearch': false,
    'bg': '#3399ff',
    'fg': '#ffffff',
    'fontsize': 12,
    'target': '_self',
    'suggest': true,
    'site_suggest': true,
    'language': 'ru',
    'tld': 'ru'
}">
<form action="http://yandex.ru/sitesearch" method="get" target="_self">
  <input type="hidden" name="searchid" value="1902945" />
  <input type="hidden" name="l10n" value="ru" />
  <input type="hidden" name="reqenc" value="windows-1251" />
  <input type="text" name="text" value="" />
  <input type="submit" value="Найти" />
</form>
</div>
<style type="text/css">.ya-page_js_yes .ya-site-form_inited_no { display: none; }</style>
<script type="text/javascript">(function(w,d,c){var s=d.createElement('script'),h=d.getElementsByTagName('script')[0]; s.type='text/javascript';s.async=true;s.charset='utf-8'; s.src=(d.location.protocol=='https:'?'https':'http:')+'//site.yandex.net/v2.0/js/all.js'; h.parentNode.insertBefore(s,h); (w[c]||(w[c]=[])).push(function(){Ya.Site.Results.init({})})(window,document,'yandex_site_callbacks'); </script>
```

```
'searchid': '1902945',
'action': 'http://mysite.ru/poisk.html',
'type': 0,
'arrow': true,
'logo': 'rb',
'webopt': true,
'websearch': false,
'bg': '#3399ff',
'fg': '#ffffff',
'fontsize': 12,
'target': '_self',
'suggest': true,
'site_suggest': true,
```

```
'language': 'ru',
'tld': 'ru'
```

```
<input type="hidden" name="searchid" value="1902945" />
<input type="hidden" name="l10n" value="ru" />
<input type="hidden" name="reqenc" value="windows-1251" />
```

JS переменные

Язык интерфейса поиска

```
'language' : 'ru'
```

Домен, в котором будет производиться поиск.

Различия в индексах и ранжировании.

```
'tld' : 'ru'
```

HTML теги формы

Домен, в котором будет производиться поиск.

Различия в индексах и ранжировании.

```
<form action="http://yandex.ru/ . . .
```

Язык интерфейса поиска:

```
<input type="hidden" name="l10n"  
value="ru" />
```

Кодировка сайта:

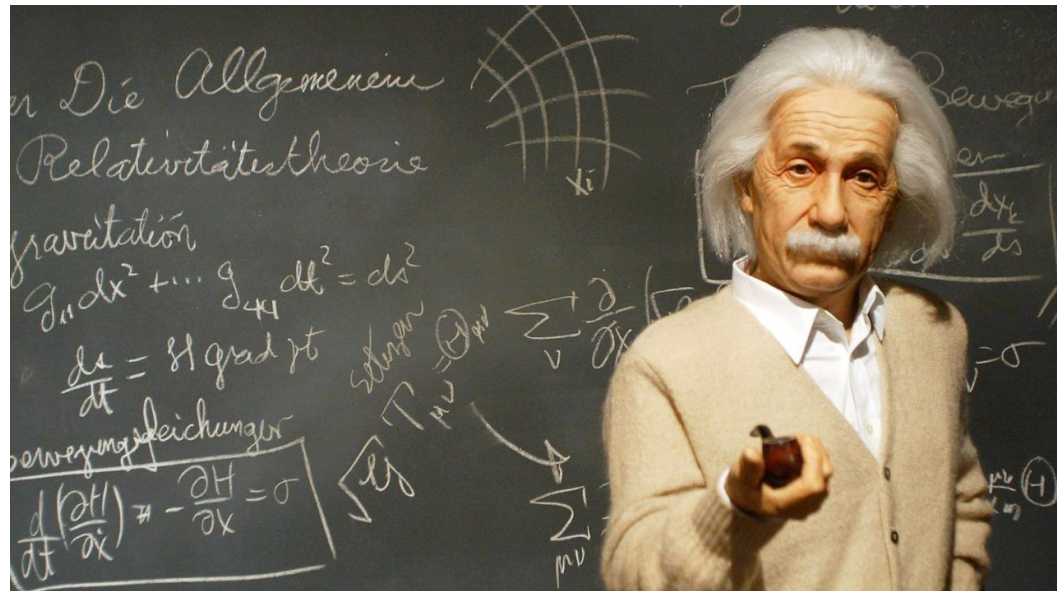
```
<input type="hidden" name="reqenc"  
value="windows-1251" />
```

Код результатов поиска

```
<div id="ya-site-results"
  onclick="return {
    'tld': 'ru',
    'language': 'ru',
    'encoding': 'windows-1251'
  }">

</div>
<script
  type="text/javascript">(function(w,d,c){var
  s=d.createElement('script'),
  h=d.getElementsByTagName('script')[0];
  s.type='text/javascript';
  s.async=true;s.charset='utf-8';
  s.src=(d.location.protocol==='https:'?'https
  ':'http:')+
  '//site.yandex.net/v2.0/js/all.js';h.parentNode.
  insertBefore(s,h);
  (w[c]||(w[c]=[])).push(function(){Ya.Site.Re
  sults.init()});
  (window,document,'yandex_site_callbacks');</
  script>
```

```
'tld': 'ru',
'language': 'ru',
'encoding': 'windows-1251'
```



JS переменные

Язык интерфейса поиска

```
'language' : 'ru '
```

Домен, в котором будет производиться поиск.
Различия в индексах и ранжировании.

```
'tld' : 'ru '
```

Кодировка сайта

```
'encoding' : 'windows-1251 '
```

Уточнения поиска **Архитектура и принцип работы**

Уточнения. Как выглядят.

The screenshot shows a Yandex search interface. The search bar contains the text "персонаж". Below the search bar, there are radio buttons for "на сайте" (selected) and "в интернете". There are also filters for "раздел любой", "ключевые слова нет", and "за любое число". A dropdown menu is open over the search results, listing categories: "Объявления", "Блог", "События", "Новости", "Флуд тоже Труд", "Об играх", and "Важно знать!". The search results list three items, all from "alltgame.ru".

Яндекс
Нашлось
11 тыс. ответов

персонаж Найти

на сайте в интернете

раздел любой ключевые слова нет за любое число

- 1. [Игровые объявления](#)
Продается отличный персонаж, синяя подзащита, 1000 IP, элитный воин, синяя подзащита.
[alltgame.ru/notice](#)
- 2. [Игровые объявления](#)
Доска игровых объявлений - продать персонажа, купить чара, продам перса...
[alltgame.ru/notice](#)
- 3. [Доска игровых объявлений - продать персонажа, купить чара, продам перса...](#)
Купить голда, купить адену, купить персонажа - Доска игровых объявлений. Вот захочется вам купить чара приличного. И голд купить, чтоб было где развернуться.
[alltgame.ru](#)

Объявления
Блог
События
Новости
Флуд тоже Труд
Об играх
Важно знать!

Уточнения. Как настраиваются.

Разделы сайта

Укажите разделы сайта (сайтов) для ограничения области поиска.

i URL должен быть указан в формате "http://имя_сайта/имя_раздела" и принадлежать области поиска.

<input type="checkbox"/> <input type="checkbox"/>	Объявления	http://alltosome.ru/notice	<input type="checkbox"/>
<input type="checkbox"/> <input type="checkbox"/>	Блог	http://alltosome.ru/blog	<input type="checkbox"/>
<input type="checkbox"/> <input type="checkbox"/>	События	http://alltosome.ru/blog/event	<input type="checkbox"/>
<input type="checkbox"/> <input type="checkbox"/>	Новости	http://alltosome.ru/blog/news	<input type="checkbox"/>
<input type="checkbox"/> <input type="checkbox"/>	Флуд тоже Труд	http://alltosome.ru/blog/flood	<input type="checkbox"/>
<input type="checkbox"/> <input type="checkbox"/>	Об играх	http://alltosome.ru/blog/ogames	<input type="checkbox"/>
<input type="checkbox"/> <input type="checkbox"/>	Важно знать!	http://alltosome.ru/blog/important	<input type="checkbox"/>

Название раздела URL раздела

Разрешить выбирать несколько разделов одновременно

разделы сайта все

- [Объявления](#)
- [Блог](#)
- [События](#)
- [Новости](#)
- [Флуд тоже Труд](#)
- [Об играх](#)
- [Важно знать!](#)

Уточнения. Как работают.

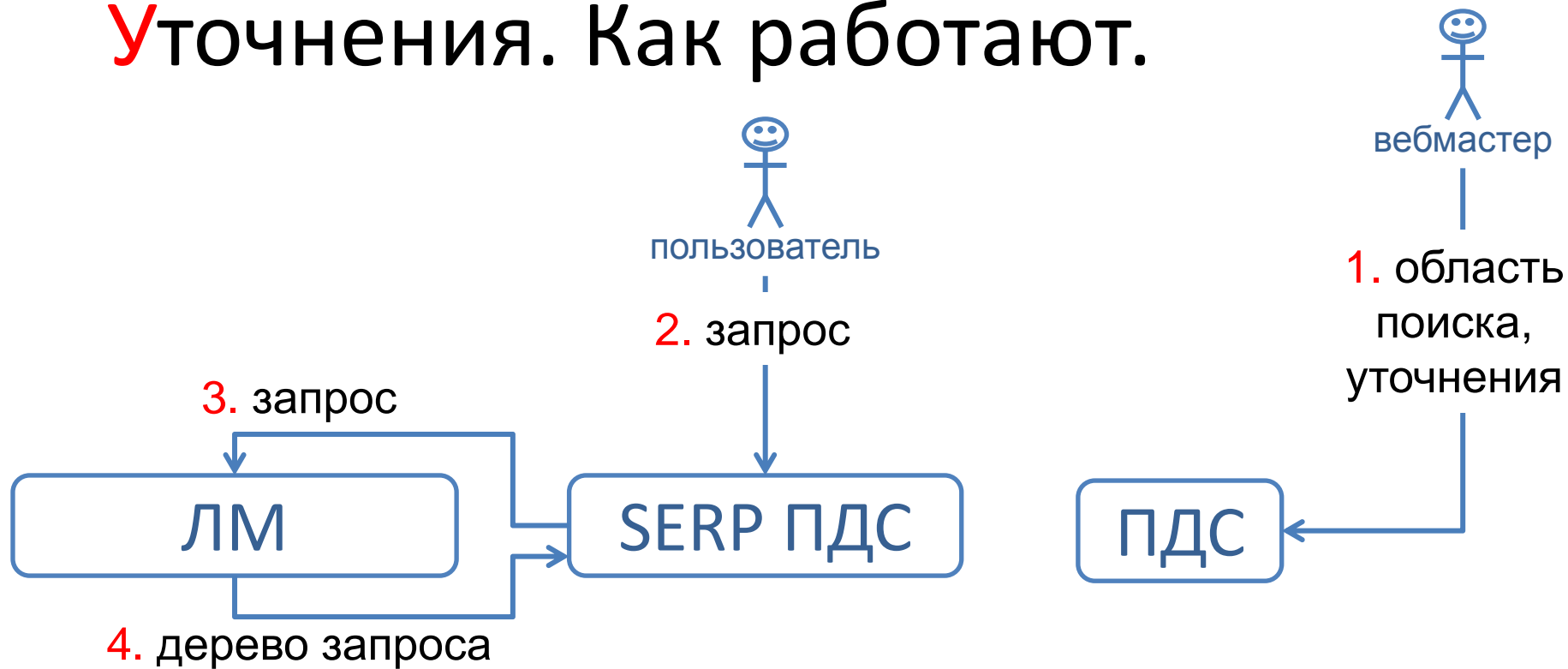
Уточнения. Как работают.



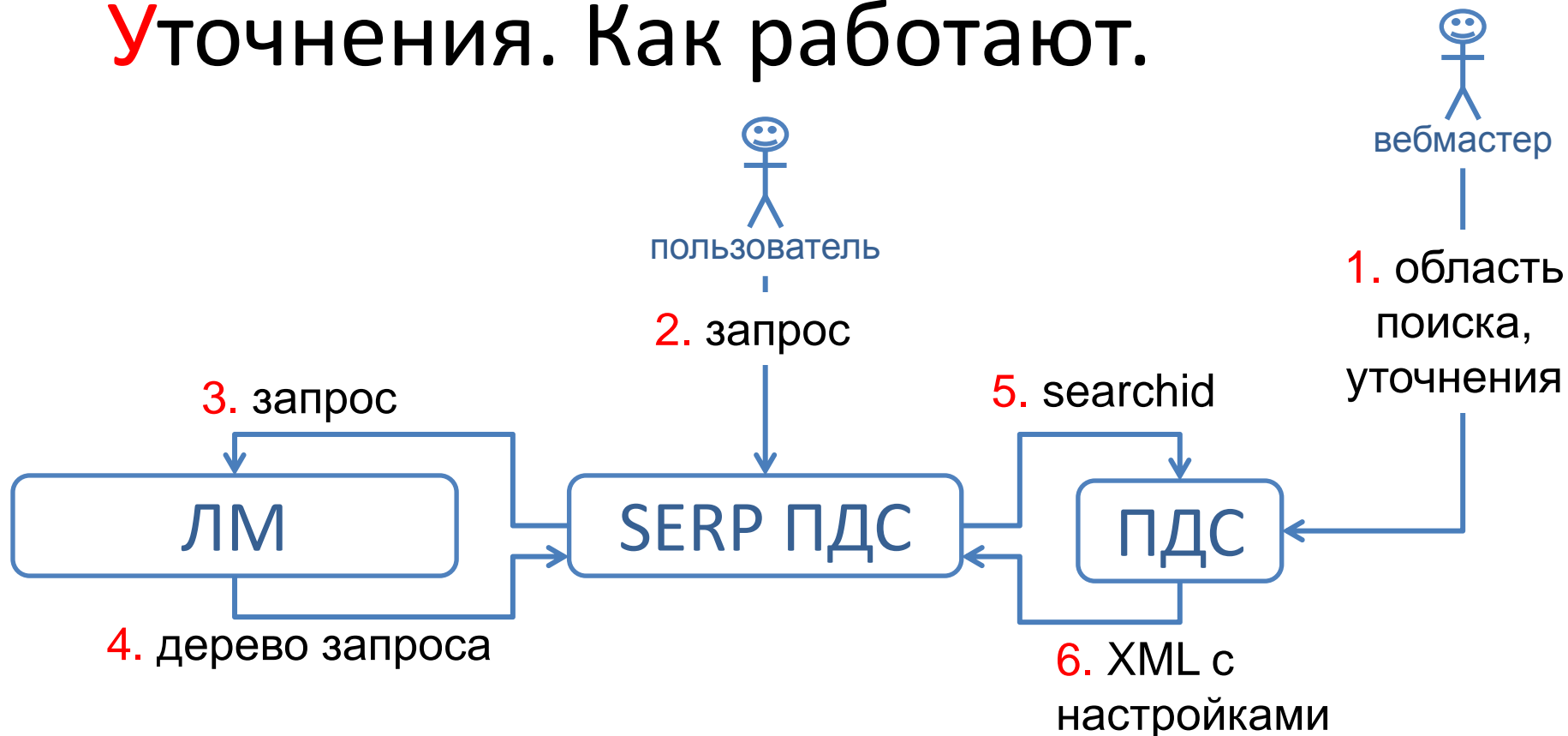
Уточнения. Как работают.



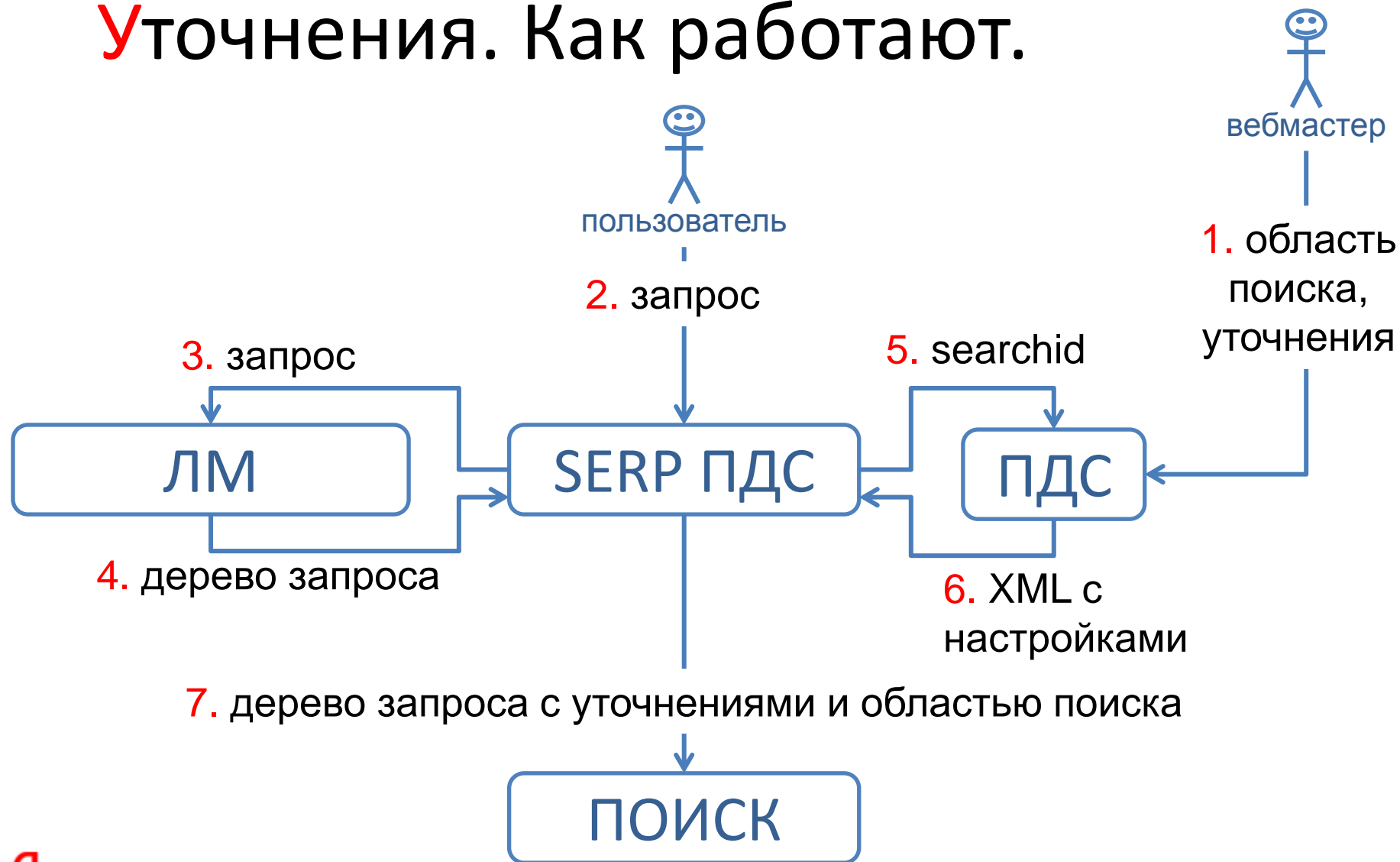
Уточнения. Как работают.



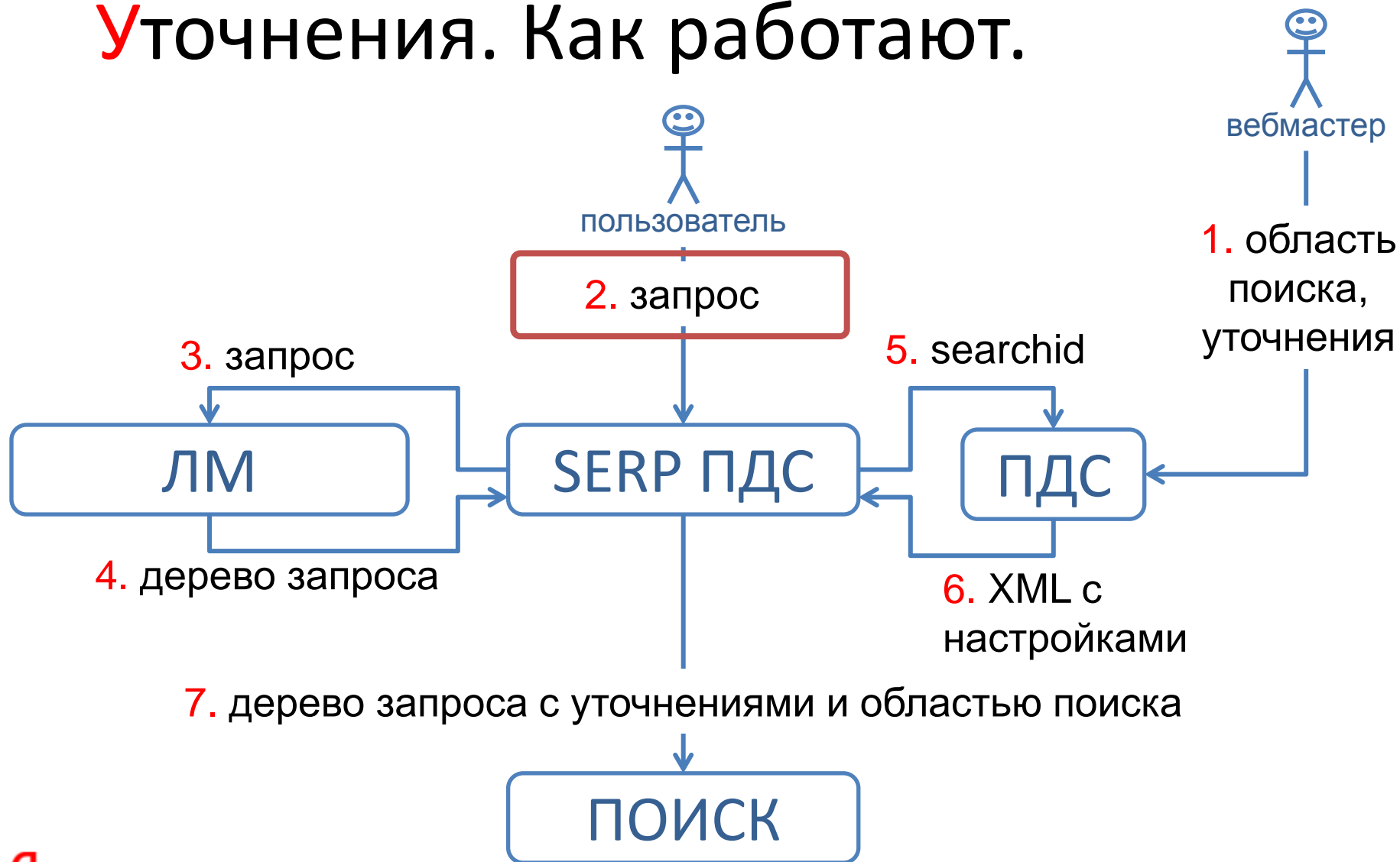
Уточнения. Как работают.



Уточнения. Как работают.



Уточнения. Как работают.



http запрос от пользователя в Яндекс

```
http://yandex.ru/sitesearch?
```

```
text=персонаж+lineage
```

```
&searchid=1775761
```

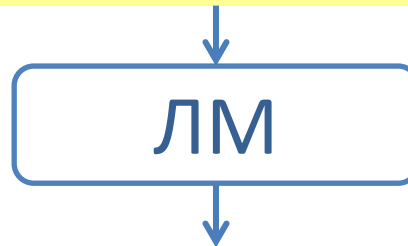
```
&constraintid=8
```

Уточнения. Как работают.



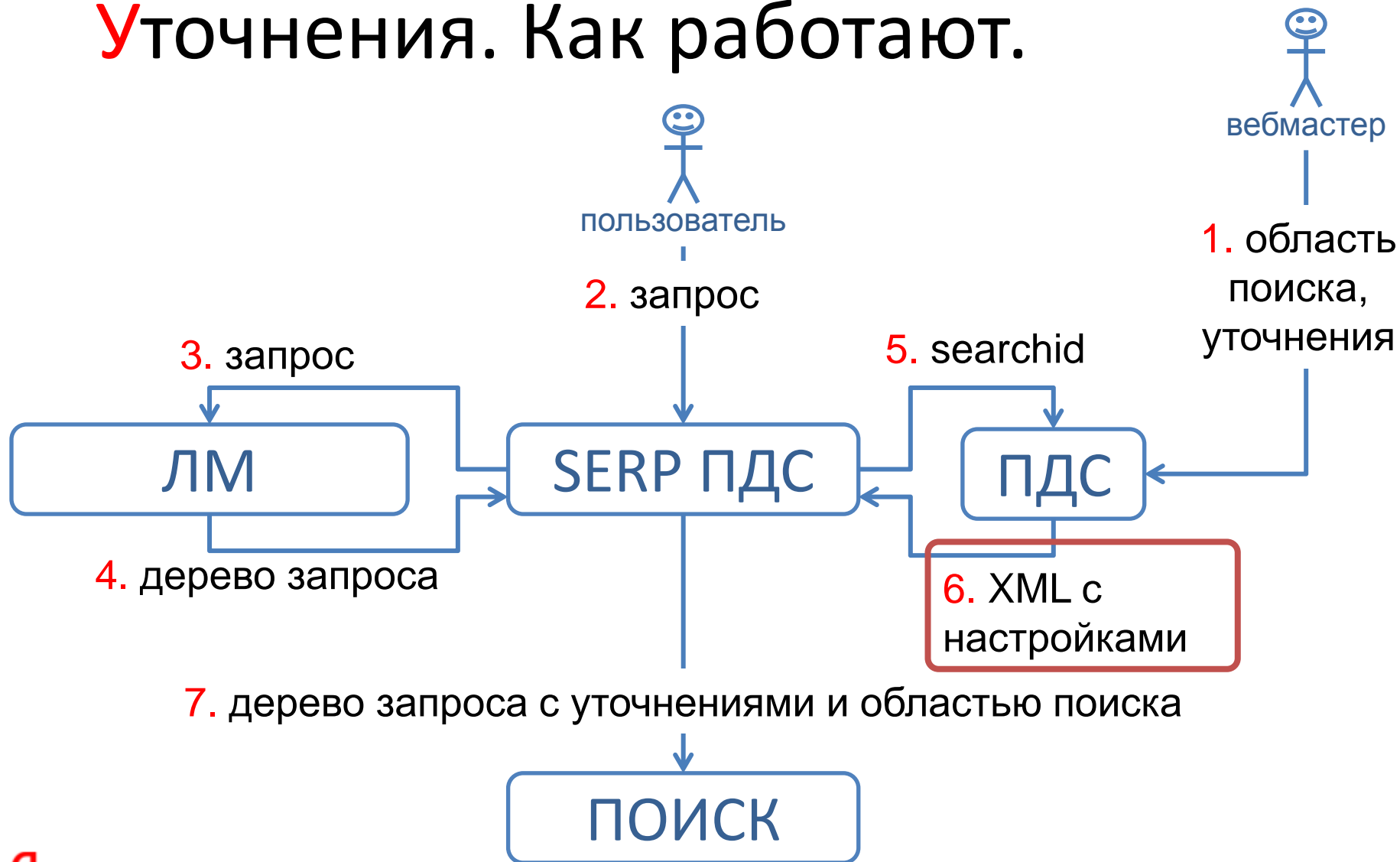
ЛИНГВИСТИЧЕСКИЙ МОДУЛЬ

персонаж lineage



```
(
    персонаж ^ personage
)
&&/ (-7 7)
(
    lineage ^ линейка ^ !!линейдж ^ !!линейка ^
    ((line &/ (1 1) age ) ^ (лайн &/ (1 1) age) ^
    лайнаге
)
```

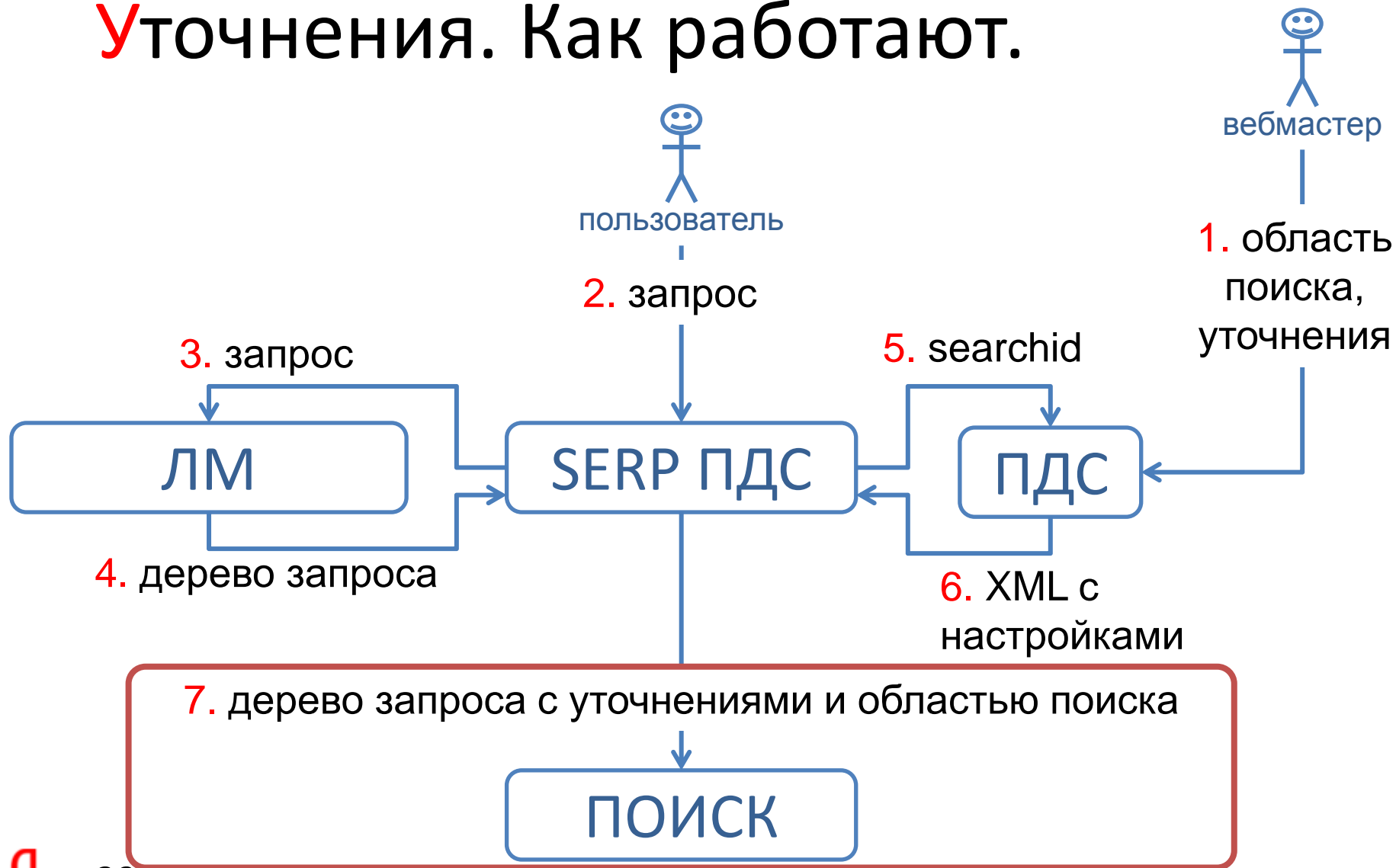
Уточнения. Как работают.



Данные в XML от ПДС в SERP

```
...
<urllist>
  <url loc="http://allto game.ru" ...></url>
  ...
</urllist>
...
<constraints>
  <constraint id="8">
    <multiselect value="false">
      <name value="Объявления">
        <type type-id="5">
          <value value="url:http://allto game.ru/notice |
            url:http://allto game.ru/notice*">
        </value>
      </type>
    </name>
  </multiselect>
</constraint>
...
</constraints>
...
```

Уточнения. Как работают.



Преобразование дерева запроса

```
(
  персонаж ^ personage
)
&&/(-7 7)
(
  lineage ^ линейка ^ !!линейдж ^ !!линейка ^
  ((line &/(1 1) age )) ^ (лайн &/(1 1) age) ^
  лайнаге
)
```


Преобразование дерева запроса

```
(
    персонаж ^ personage
)
&&/(-7 7)
(
    lineage ^ линейка ^ !!линейдж ^ !!линейка ^
    ((line &/(1 1) age )) ^ (лайн &/(1 1) age) ^
    лайнаге
)

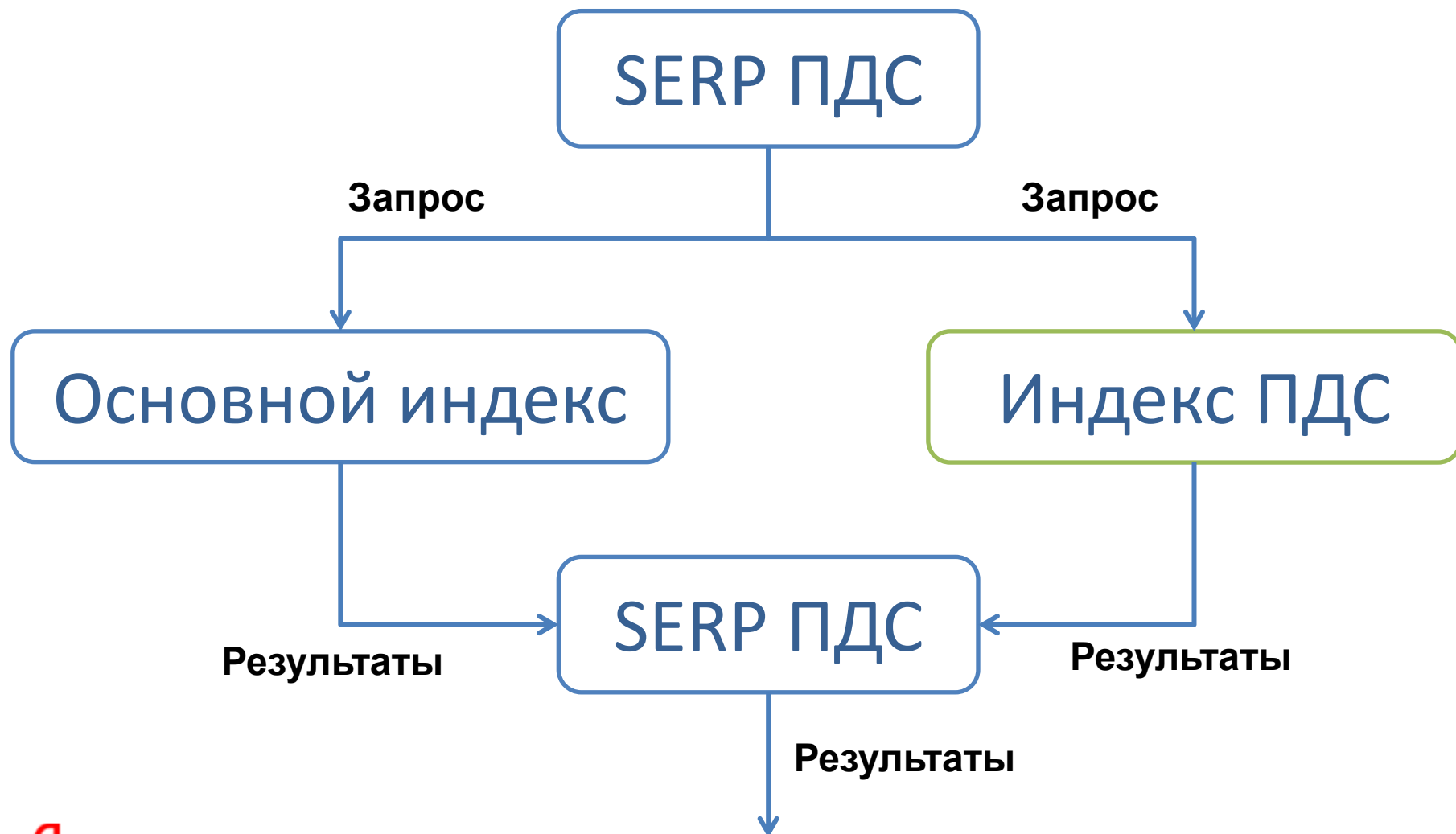
<< url:"allto game.ru/notice" | url:"allto game.ru/notice*"
```

Преобразование дерева запроса

```
(
  персонаж ^ personage
)
&&/(-7 7)
(
  lineage ^ линейка ^ !!линейдж ^ !!линейка ^
  ((line &/(1 1) age )) ^ (лайн &/(1 1) age) ^
  лайнаге
)
<< url:"alltosome.ru/notice" | url:"alltosome.ru/notice*"
    << (rhost:"ru.alltosome" | rhost:"ru.alltosome.*")
```

Индексирование по запросу Пинги и Пингеры

Дополнительный поисковый индекс



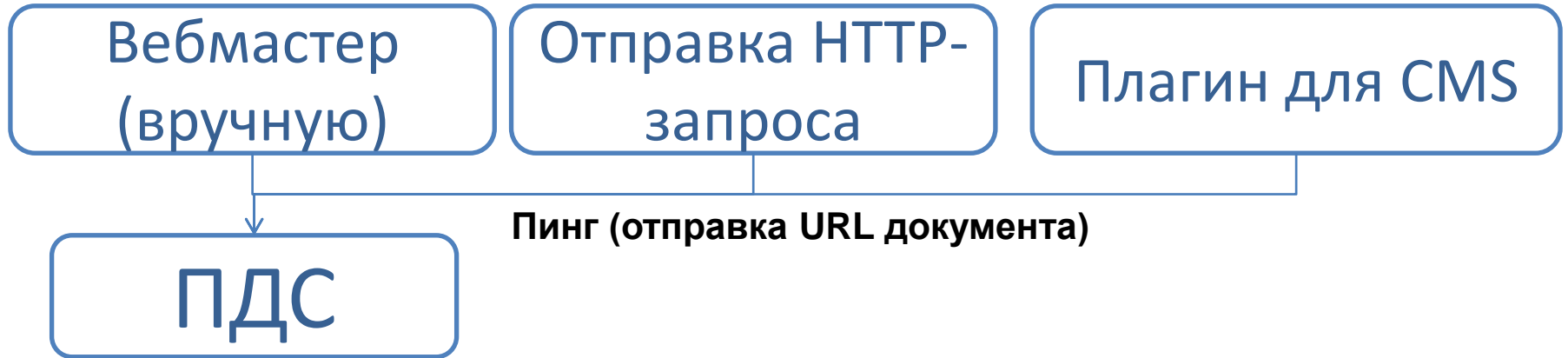
Индексирование по запросу

Вебмастер
(вручную)

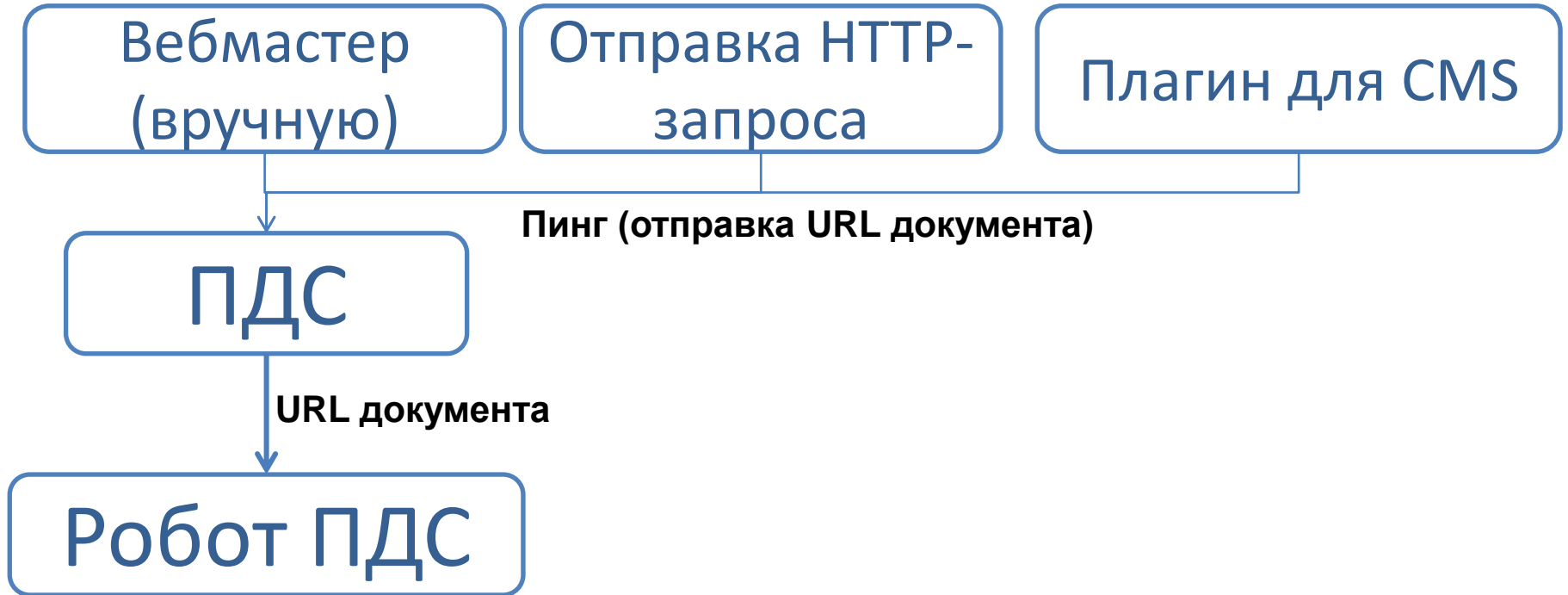
Отправка HTTP-
запроса

Плагин для CMS

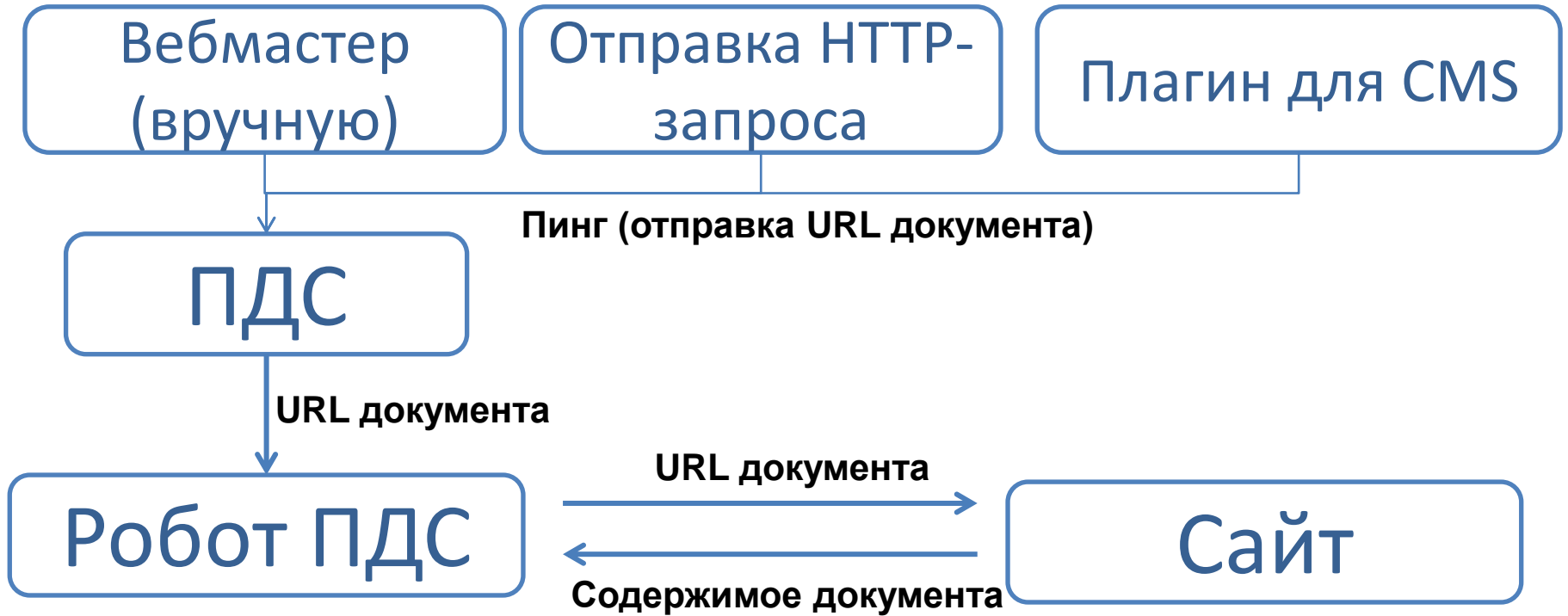
Индексирование по запросу



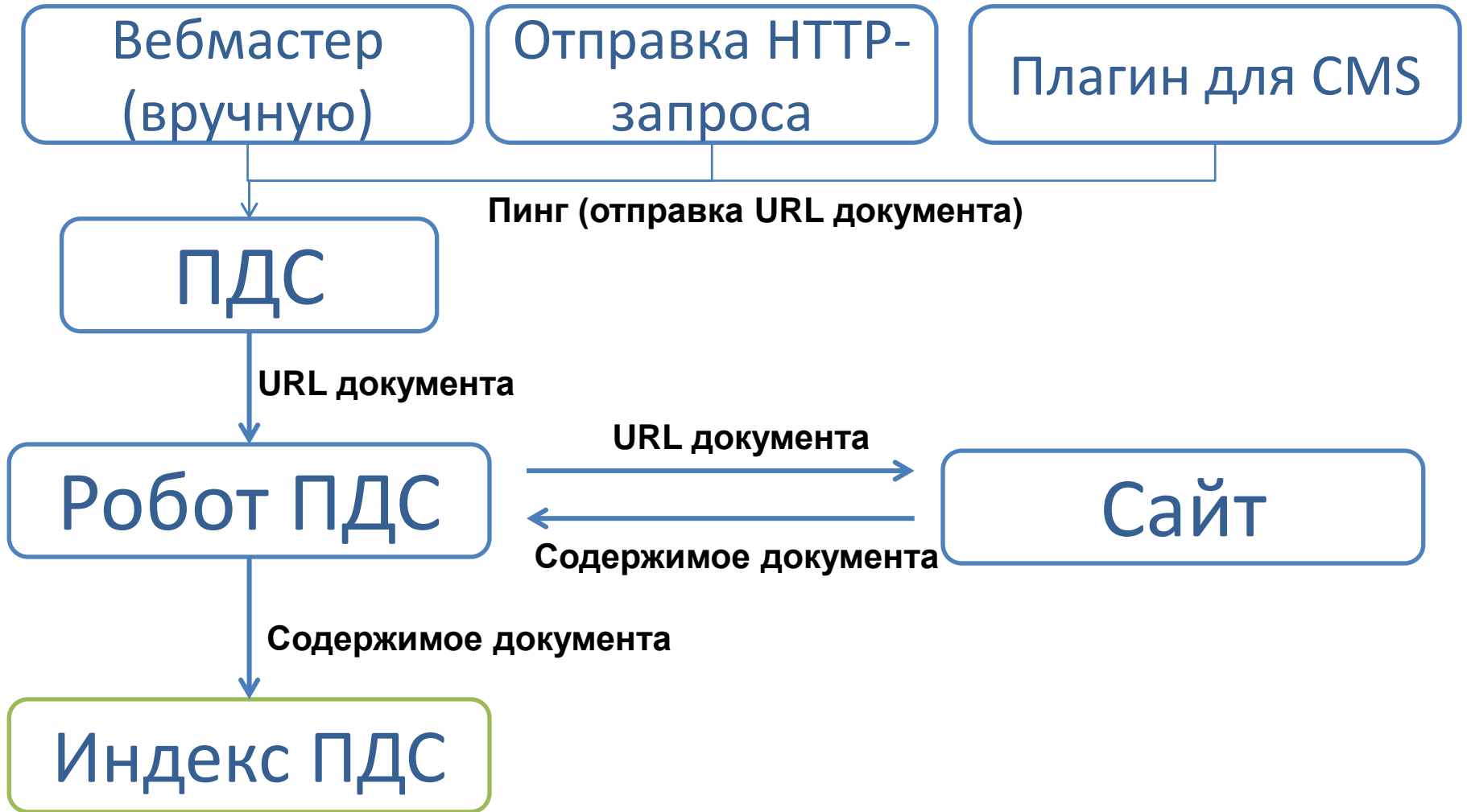
Индексирование по запросу



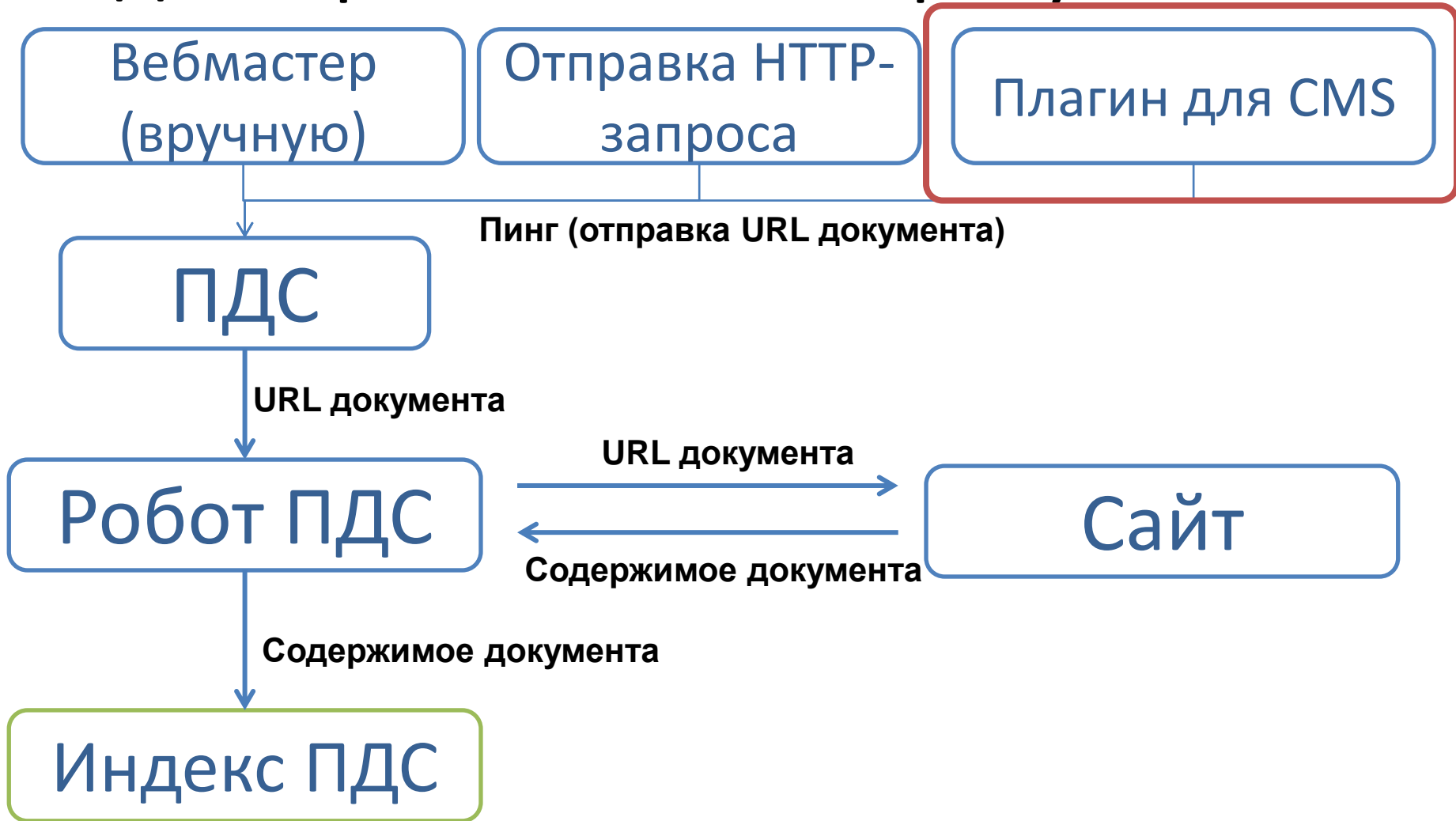
Индексирование по запросу



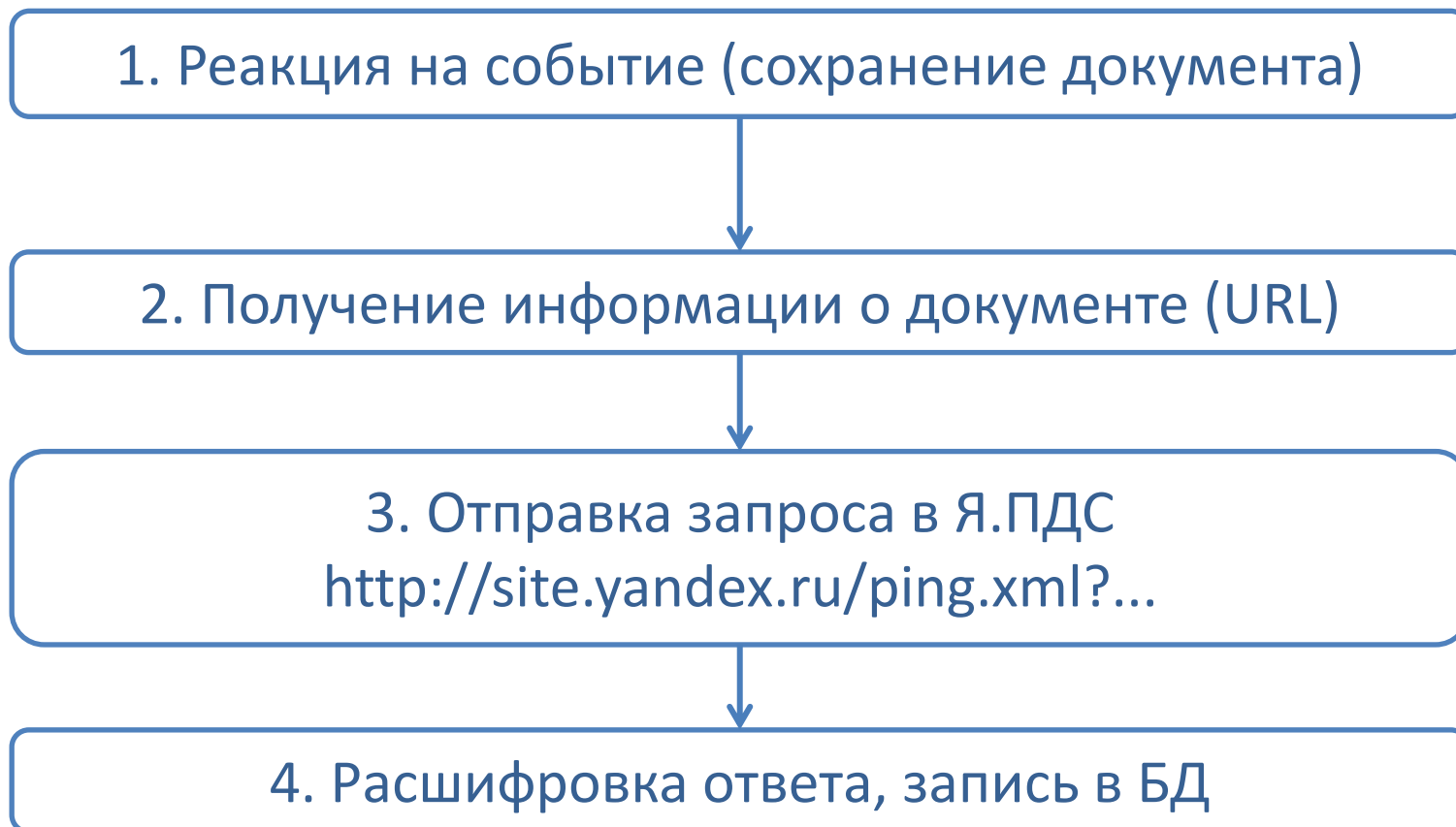
Индексирование по запросу



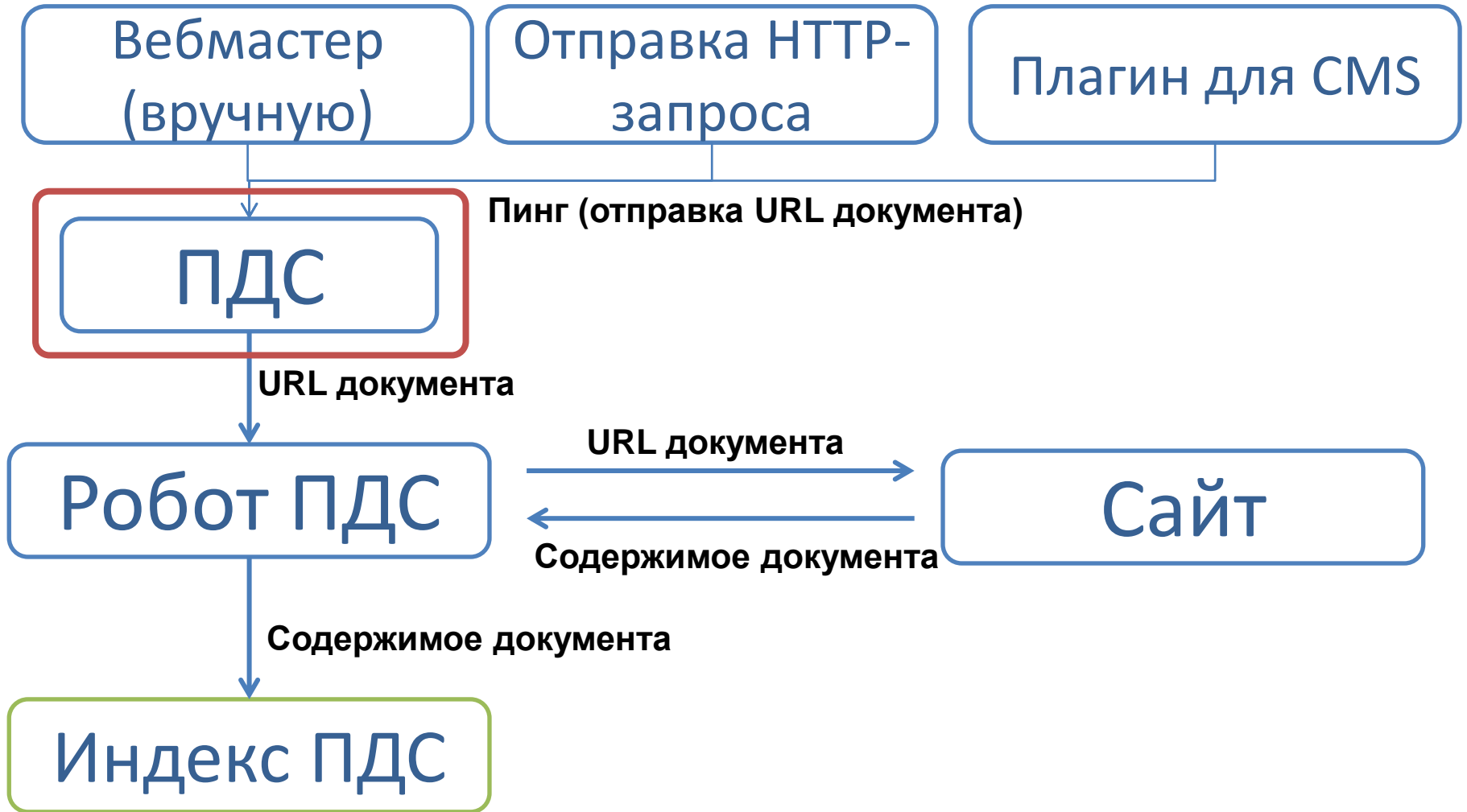
Индексирование по запросу



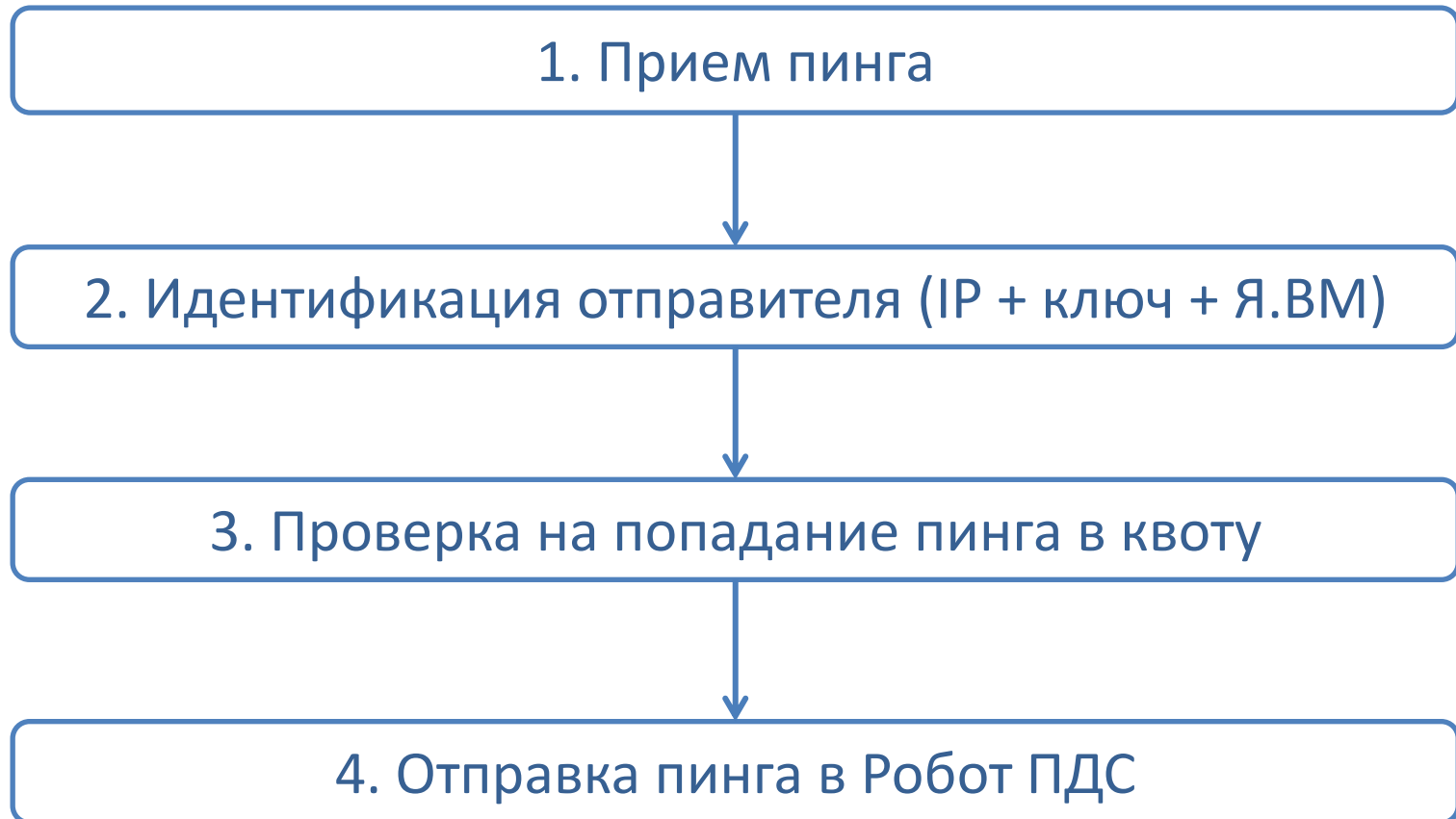
Плагины-пингеры для CMS



Индексирование по запросу



Прием и обработка пингов в ПДС



Формула для распределения квот скачивания для Робота ПДС

$$Q(\text{host}) = f(F1, F2, \dots, Fn)$$

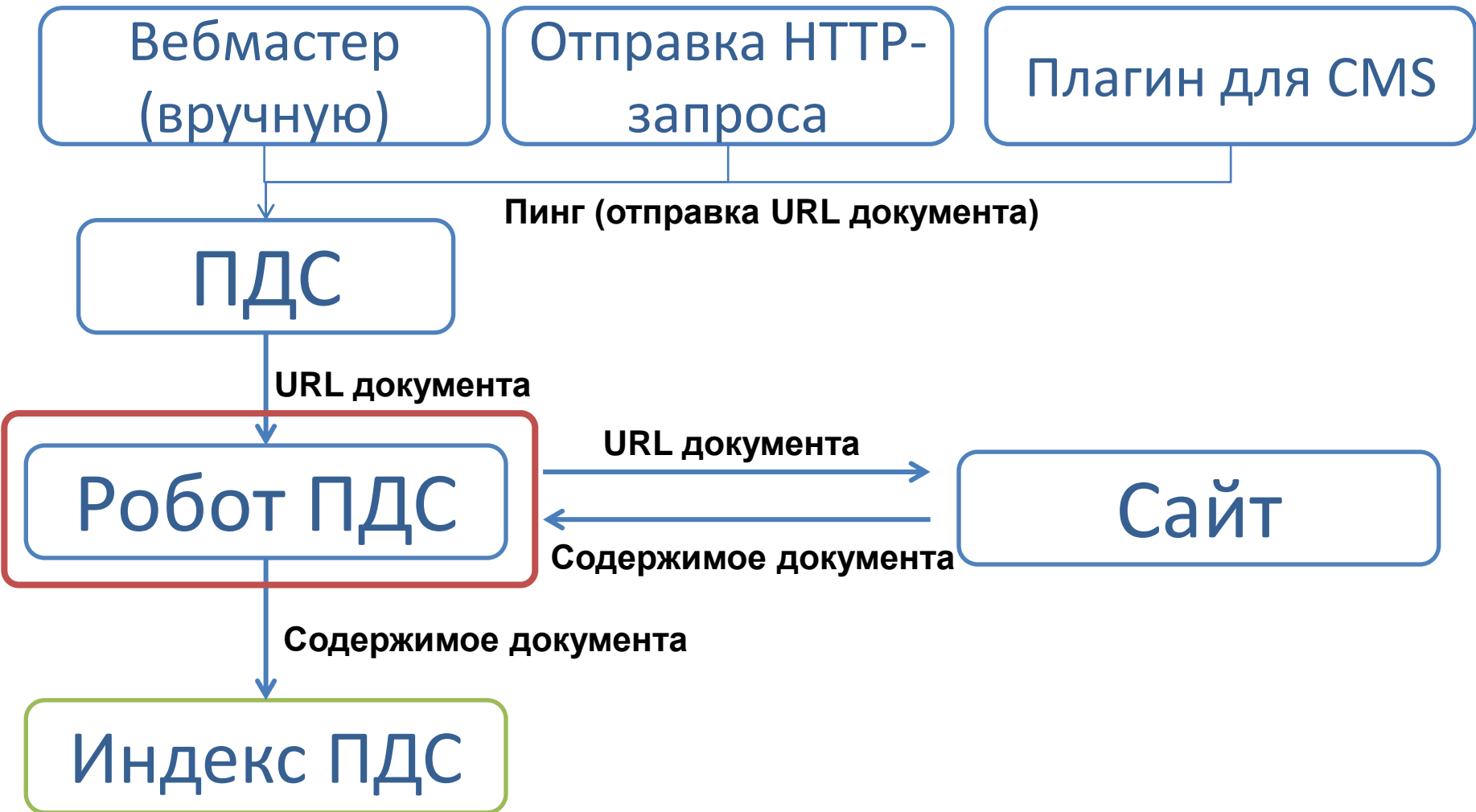
где:

F1 – Фактор кол-ва запросов к поисковой площадке

F2 – Признак того, что площадка недавно добавлена

...

Индексирование по запросу



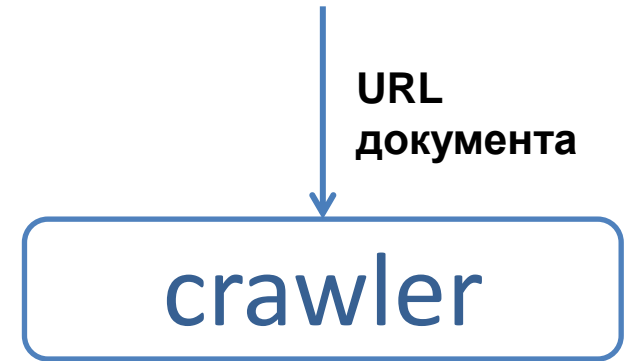
Робот ПДС



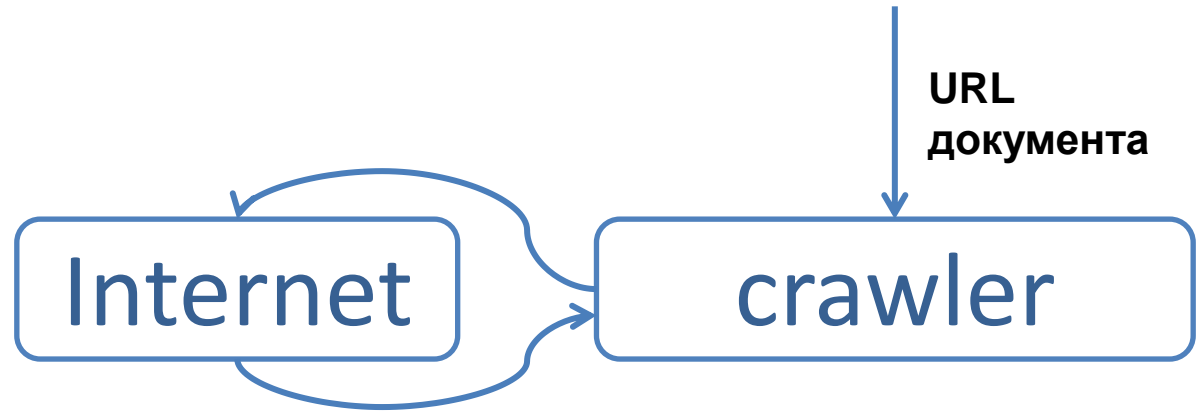
Робот ПДС



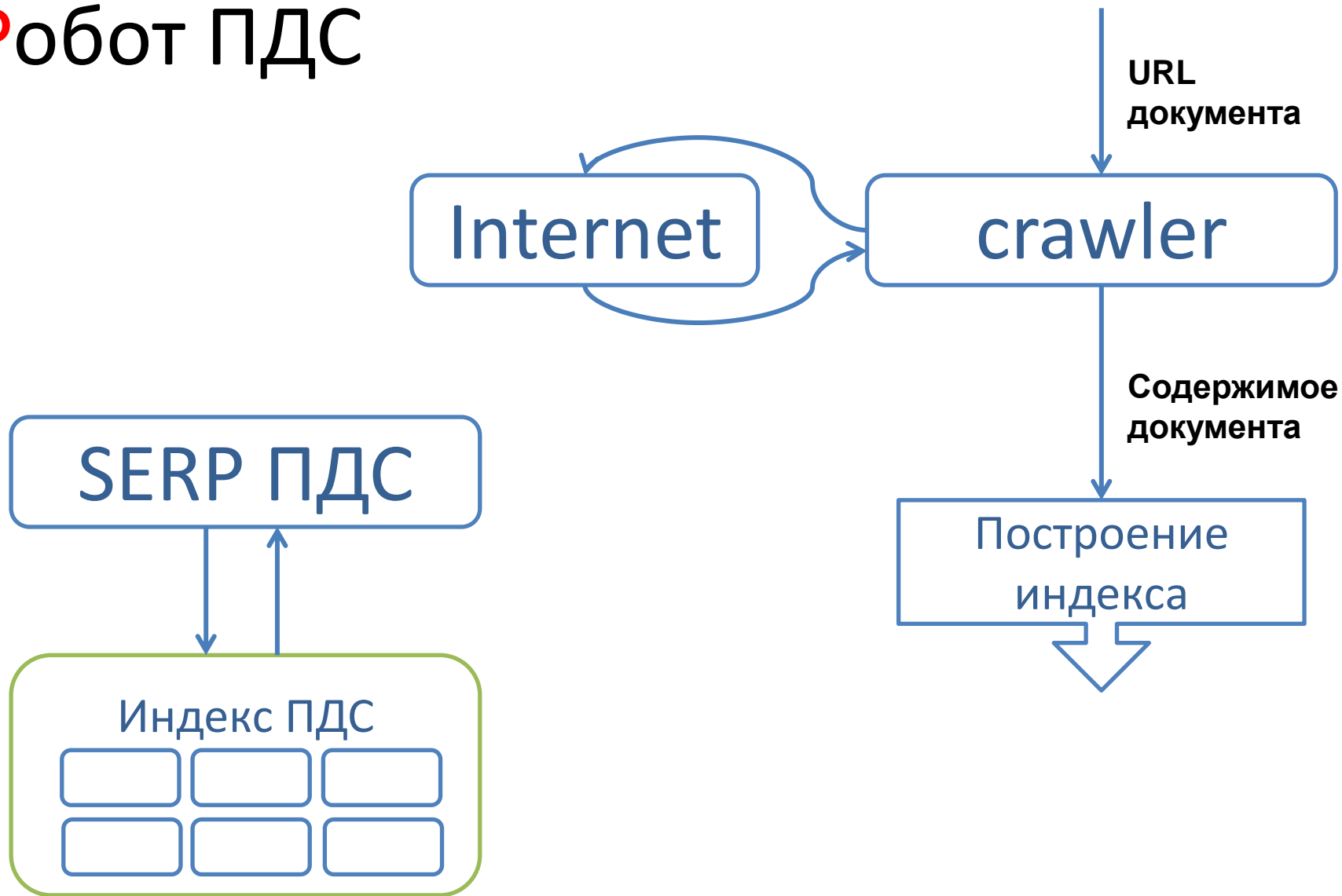
Робот ПДС



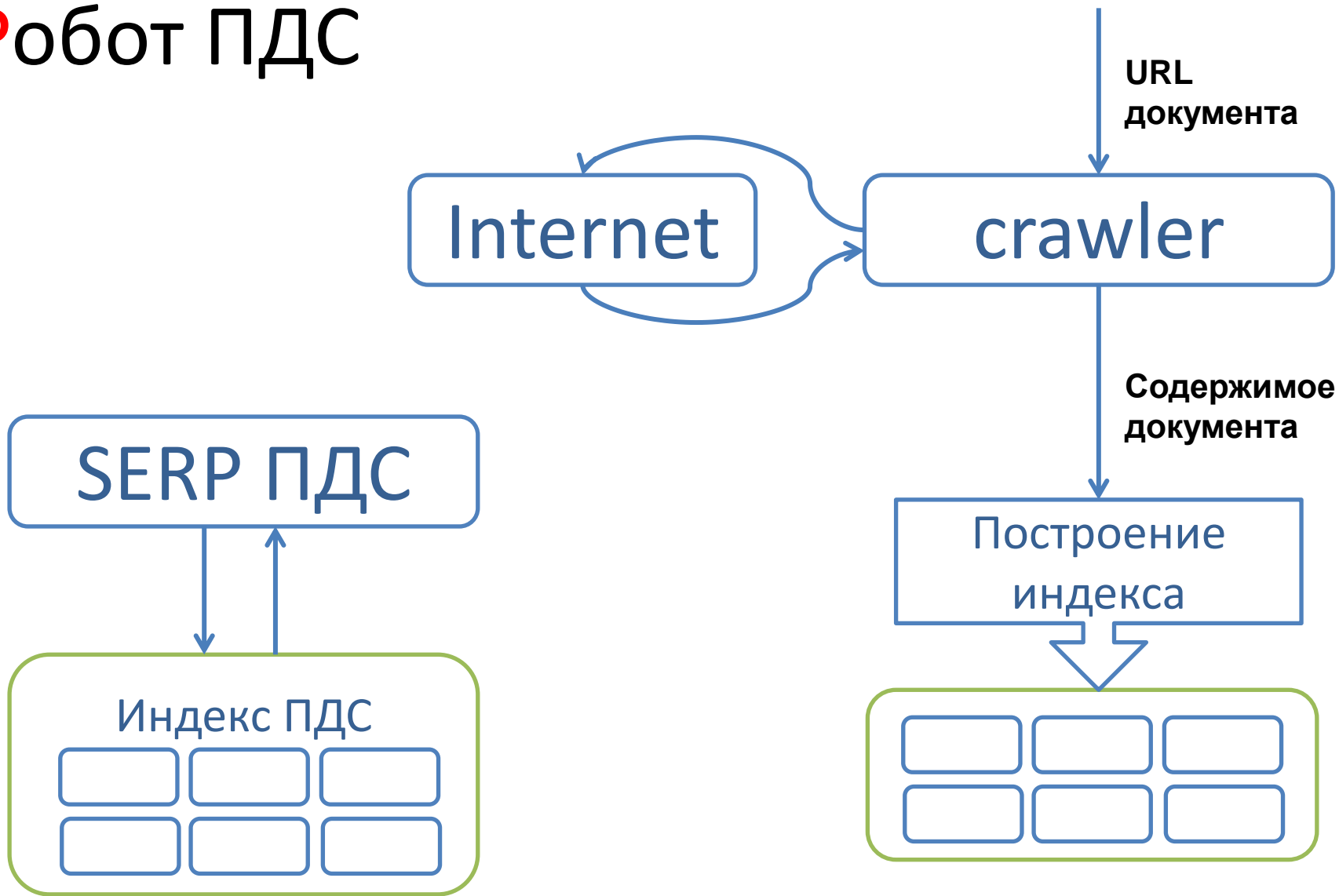
Робот ПДС



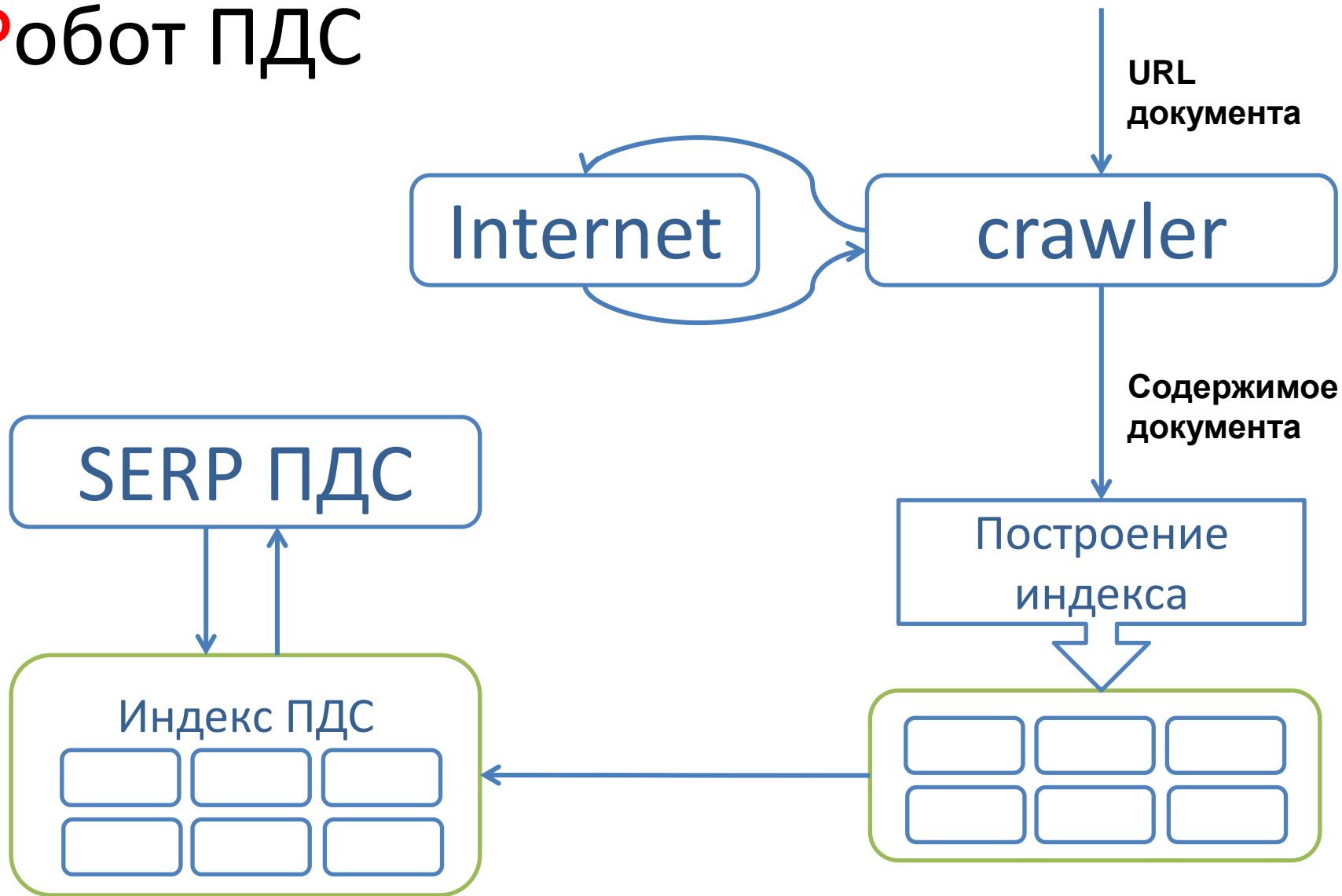
Робот ПДС



Робот ПДС



Робот ПДС



Ваши вопросы?





Дмитрий Вульбрун

Менеджер проектов

wulbrun@yandex-team.ru

@wulbrun