

Sketch Engine — Infrastructure, Support and Customer Services

Jan Bušta



jan.busta@sketchengine.co.uk

4th Sketch Engine Workshop
Tallinn, October 16, 2013

Challenge

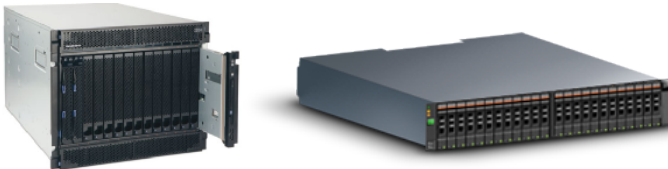
Meet users' requirements.

Simple facts:

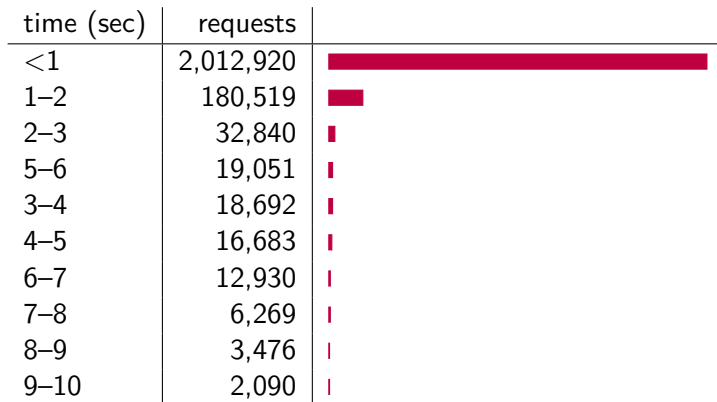
- **Thousands** of users.
- About **1.5 TB** are transferred per month.
- **150,000 requests** a day.
- **435** preloaded corpora.
- **8,222** users' corpora.
- Handling more than **50 TB** of data.
- Availability at **99,83%**.
- **All requirements are fast growing...**

Servers

- IBM Blade servers for High-Availability Clusters.
 - Remote administration.
 - Preventing data corruption.
 - Automatic recovery.
- Storage system with hard disk replacement without stopping services.
- Online monitoring system with real-time notification.

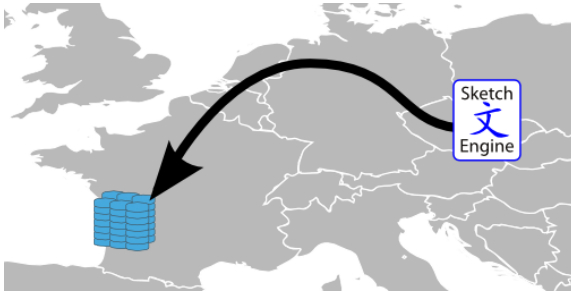


Request processing



Backup

- Backup twice a week.
- Geographically remote location of backup server.
- Providing services during primary servers maintenance.

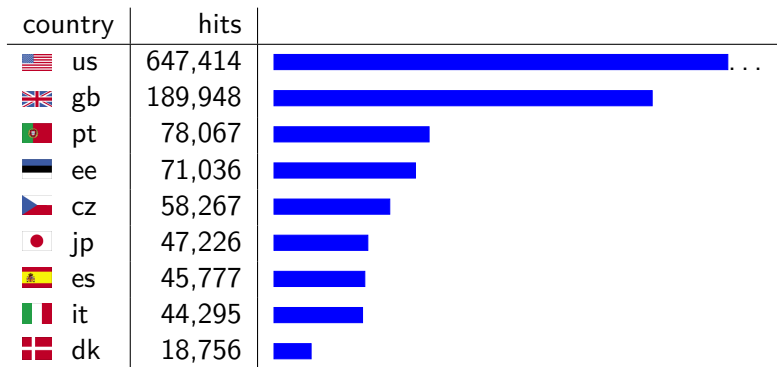


Local Installations

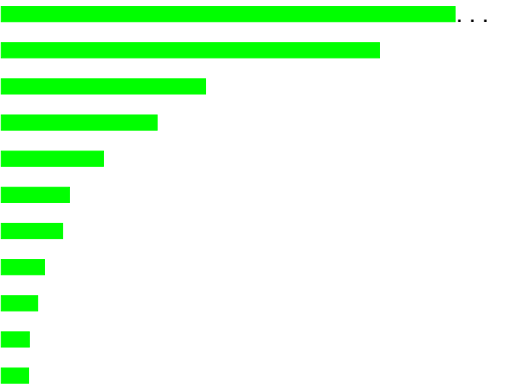
- Solution for ambitious customers and big companies.
 - Servers & data are placed in area of customer's wish on rent or customer's servers.
 - Meets legal limitation of data access.
 - Sketch Engine can be customized for specific purposes.



Usage ordered by country (Sep 2013)



Corpora querying statistics (Sep 2013)

corpus	queries	
ententen12_1	273,416	
pttenten11	31,718	
bnc2	17,105	
estonianRC	13,050	
eseutenten11	8,575	
jpwac	5,716	
frtenten	5,173	
eflux	3,632	
ittenten	3,071	
rutenten	2,332	
ukwac3	2,331	

Support Service

- Bug tracking system. (Preferred.)
 - ⇒ **62%** of bugs are fixed during the same day.
 - ⇒ **94%** solved in the first week after reporting.
- User support via e-mail. (Less reliable.)
- Documentation system (readings, examples, tutorials)
<http://www.sketchengine.co.uk/documentation>
- One-off data processing and development requests.

Sketch Engine infrastructure
is always growing
to provide a fine service.

Thank you!

Thank you for your attention!