BOE means back of envelop estimation. It drives our decision of system decision. It helps to make us decision do we really need that resource, how much capacity of the resource we need. Resources means load balancer, cdn, cache, application server, database etc. It is useful for making decision so we don not unutilized resources or over used resources unnecessarily.

We have to make some consideration to do BOE estimation.  
It is a rough (t-shirt size) estimation. Means its not accurate estimation. It is not the estimation of real world.  
Do not spend too much time in thinking. Because to make the system scalable you ultimately need all the resources. However you design you will conclude to the same thing.  
Keep the assumption value simple. Do not take users value like 275 million instead can take 100 million. Because it is easy to calculate.

Remember in mind, value changes after each three zeros.  
In case of traffic:  
3 zeros => thousand  
6 zeros => million   
9 zeros => billion   
12 zeros => trillion   
15 zeros => quadllion   
In case of storage:  
3 zeros => kb (Kilobyte)  
6 zeros => mb (megabyte)  
9 zeros => gb (gigabyte)  
12 zeros => tb (terabyte)  
15 zeros => pb (petabyte)

We have to deal with char, long/double, images. Lets say ascii char is 1 byte and unicode is 2 byte. So char is 2 byte. Long/double is 8 byte. Images are 300kb.

To do computation in BOE estimation, we will consider three things. They are: number of servers, ram and storage capacity.

Important calculation:  
X million users \* Y MB data = XY TB data. Because million is 6 zero and MB is 6 zeros so total 12 zeros denote to TB

We will do Facebook estimation. First we will do traffic estimation. Lets say there are 1 billion users. 25% of them (250 million) are DAU (Daily Active Users). Lets say each user do 5 read queries and 2 write queries daily. So there is 7 queries. So query per second is (250 million \* 7 query)/[100000](tel:100000) seconds = 18KB. [Need Later]

Now lets do storage estimation. Lets say every DAU do 2 posts every day. Lets assume each posts has 250 characters. Each character contains 2 byte. So each post contains 500 bytes [Need later]. As each DAU gives 2 posts so total 500 \* 2 = [1000](tel:1000) bytes = 1 KB. So total storage required 250 million \* 1 KB = 250 GB.  
Lets assume 10% of DAU (25 million) gives image post. Each image contains 300 KB. So total storage required 25 million \* 300 KB = [7500](tel:7500) GB = around 8 TB.  
To store this for 5 years, lets say 5 years has approximately [2000](tel:2000) days.  
So for posts,  
250 GB \* [2000](tel:2000) = 500 TB  
For images,  
8 TB \* [2000](tel:2000) = 16 PB  
Now we will do Ram estimation. Lets say last 5 posts of each DAU will be stored in cache. As each posts contain 500 bytes ( from previous calculation).  So 5 posts contain 5 \* 500 = [2500](tel:2500) bytes = 3KB. So total for 250 million users, we need 250 million \* 3KB = 750 GB.  
Lets say 1 machine can store 75 GB so we will need approximately 10 machines.  
Now lets calculate latency. Lets say approximately 95% request or query requires 500 ms to respond. So in each second we may get respond from 2 requests or query. 1 server has 50 threads means each thread can do 1 request in a moment. So total it can server 100 request per second at a time.  
1 server can do 100 query per seconds. But we need to provide 18 kb query per second support. So we need 180 servers.

Lets talk about trade off means CAP theorem. For Facebook we can ignore consistency. We need to do server and request always up. So we will pick AP.

===================  
Questions   
===================  
1. What is back of envelop estimation?  
2. Deeply understand different steps of back of envelope estimation.