Linear Scalability

- Ideally, adding 2x more servers should support 2s more users
- However:
  - more servers introduce **overhead** and add the need of **synchronisation** which also adds delays
  - both the data and the compute power needs to be partitioned across multiple servers