Distributed Systems PS [SS 2018]

Homework 01

Exercise 1: Distribution Transparency

You need to design a distributed system, which should achieve maximal scalability. Assume that only desktop machines will access to the system. Given the transparencies: access, location, migration, relocation, replication, concurrency and failure, which transparency(ies) will you implement? Explain your answers!

Solution:

Access
I do not select Access-Transparency, because it is not necessary to achieve high scalability.
Location
I do select Location-Transparency, because with dynamic IP's and Ports we can achieve higher scalability as with fixed ones.
Relocation
I do select Relocation-Transparency, because it should be possible to relocate objects to different servers if it is necessary for some (e.g.) computation, that is done on a different server.
Migration
I do select Migration-Transparency, because the user can continue to interact with the distributed system, even if a problem occurs on the server side.
Replication
I do select Replication-Transparency, because in a scalable systems there are many objects that need to be replicated to be accessible by multiple users.
Concurrency
I do select Concurrency-Transparency, because like that multiple threads can handle multiple requests, instead of just one main thread which is handling all the requests.
Failure
I do not select Failure-Transparency, because it is not necessary to achieve high scalability, but it can decreases it drastically.

Migration: User can Move from pc 1 to pc 2 and he should not see a difference.

Exercise 2: Scalability

Let T(R, L) denotes the execution time of an application processing a problem of size L when being distributed among R nodes in a distributed system. For example, a response time of a service assuming the presence of L users and R servers.

Part a)

Define efficiency E(pR, L) and scalability S(pR, NL) of an application through T. Assume that resources will be increased p times in order to handle the increased load of N times.

Solution:

$$E(pR, L) = \frac{T(R,L)}{T(pR,L) * p}$$

$$S(pR, NL) = \frac{N * T(R,L)}{p * T(pR,L)}$$

Part b)

Write possible values of E(pR, L) and S(pR, NL). Assume a well/badly-scaling system!
Expected output: $a < E(pR, L) < b; c < S(pR, NL) < d$

Solution:

a = 0; b = 1; c = 0; d = ∞

Examples:

Well Scaling System:

$E(pR, L) = \frac{T(10) = 10}{T(20) * p = 15} \rightarrow 66\%$	$S(pR, NL) = \frac{N * T(20) = 200}{p * T(30) = 40} \rightarrow 5$

Badly Scaling System:

$E(pR, L) = \frac{T(30) = 30}{T(50) * p = 90} \rightarrow 33\%$	$S(pR, NL) = \frac{N * T(10) = 30}{p * T(20) = 25} \rightarrow 1{,}2$