Sabancı@home

Acronym: S@h

Members:

- Ali Parlakçı
- Albert Deniz Levi
- Jeren Annagurbanova

Distributed computing systems are in our daily use, from telephone networks, aircraft control systems, the internet as a whole to banking systems. These systems are defined by computation units divided by distance sometimes spanning the globe communicating with each other to give a particular service. History of distributed computing systems can be traced way back to ARPANET, one of Internet's predecessors. Reason these systems are preferred over centralized systems is the reach, reach that is both physical but also the computation size that can be reached, despite the inefficiencies that come with a large distance between the computation units the possibility of public help makes the size limit for these networks infinite.

Sabancı@Home will be a template distributed computing system, looking to satisfy multiple computation buyers to sellers / donators. With a generalized view on computation heavy projects and their tendency to be parallelized, Sabancı@Home will connect donators and sellers of computation power with buyers while optionally storing the computed data. This project will require a DBMS that is capable of storing research / business data of infinitely diverse type; manage the purchase, selling, donation of power; as well as the fracturing and distribution of the project workloads.

The biggest distributed computing network is the Folding@home protein folding simulation project which was famously the first exaflop computing system. It used both supercomputers and donated public computers for scientific research. S@h is planned to have both donated and sold computational power as to facilitate both scientific, business and public endeavors that may benefit from a huge network of relatively cheap computation the public may want to donate to.

Currently existing distributed computing networks are mostly research based, this is the result of lack of security of the data that is produced by the public computers. Mathematics, cryptography, medicine, astrophysics are among the research fields that benefit from this technology.

The DBMS will consist of following entities:

- **Customers** are the buyers of computational power; they will provide projects to be fractured into chunks to be sent to the donors/sellers to be computed.
- **Donators** are the provider of the computation power; they may donate their power to specific projects or sell to customer's projects by their desire. They may have multiple computers.
- **Computers** are the computing units that the donators have introduced to the system, these will receive chunks of the projects the customers will provide.
- **Projects** are the subjects the customers want computation upon, this may be as simple as a mathematical formula, it may be a raytracing job from an animation studio, it may be an AI, it may be the processing of statistical data taken by many orbital and ground telescopes... Projects can also

be described formally as distributed programs. These programs are highly parallelized to effectively use the scattered computation power.

- **Chunks** are the parts of the projects the computers of donators will see, no full project will be given to a single computer, they will only receive a chunk of the project to compute.