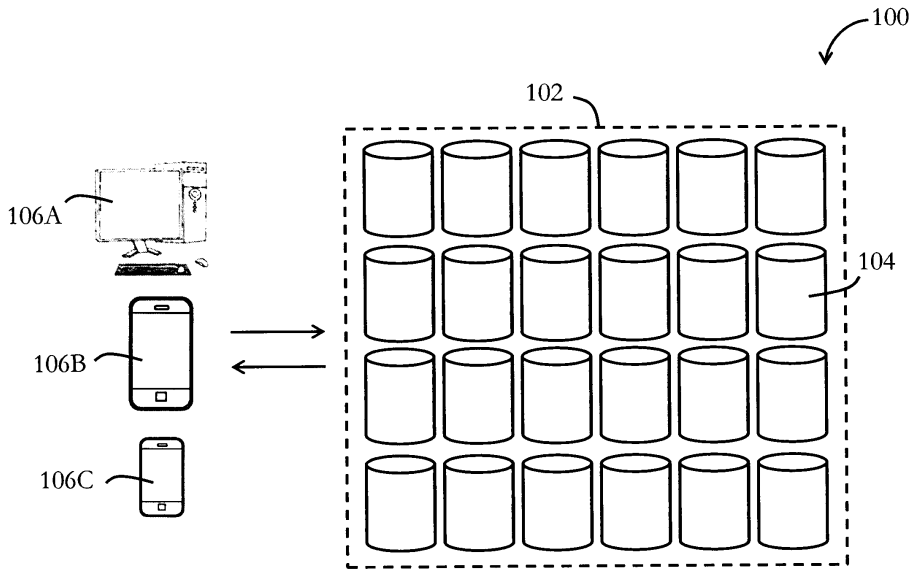
Knowledge Transfer Office
香港城市大學
City University of Hong Kong

Server Farm and Method for Operating the Same

Energy & Environment

Energy Conservation/Generation/Management/Storage (Battery)



IP Status
Patent granted

Technology Readiness Level (TRL) ? 2

Inventor(s)
Prof. ZUKERMAN Moshe
Prof. WONG Wing Ming Eric
Miss FU Jing
Prof. MORAN William
Dr. GUO Jun
Enquiry: kto@cityu.edu.hk

Opportunity

Data centers with server farms are essential to the functioning of computer systems in different applications and sectors in the modern economy. Generally, server farms in data centers include a large number of servers that consume power during operation to process and handle jobs or computational tasks. These servers account for the major portion of energy consumption of data centers.

Since excessive power consumption in server farm may increase operation cost and cause environmental concerns, various approaches have been proposed to optimize energy utilization in server farms. In one example, speed scaling is applied to control server speed. In another example, right-sizing of server farms is applied by powering servers on/off according to traffic load.

Rapid improvements in computer hardware have resulted in frequent upgrades of parts of the server farms, and this has led to server farms with different computer resources (heterogeneous servers) being deployed. The heterogeneity of servers in server farm significantly complicates the optimization of energy utilization. Therefore, there remains a need for server farm designers and/or operators to devise an optimal strategy in operating and managing server farms so as to conserve energy and maximize the effective energy efficiency of server farms.

Develop Concept
Proof Concept
Follow-on Funding
Build Value

Technology

The present invention relates to a system and method for operating a server farm, and particularly, although not exclusively, to an asymptotically optimal job assignment method for operating an energy-efficient processor sharing server farms.

A method for operating a server farm with a plurality of servers operably connected with each other includes: receiving a job request of a computational task to be handled by the server farm; determining, from the plurality of servers, one or more servers operable to accept the job request; determining a respective effective energy efficiency value associated with at least the one or more servers; and assigning the computational task to a server with the highest effective energy efficiency value. The effective energy efficiency value is defined by a service rate of the respective server divided by a difference between an energy consumption rate value when the respective server is busy and an energy consumption rate value when the respective server is idle. The present invention also relates to a server farm operated by the method.

Advantages

- The energy efficiency of a server farm

Applications

- A performance measure in ICT applications