Capstone Vision Statement

Null Terminators

January 21, 2013

1 DESCRIPTION

VMware is a company that provides virtualization and cloud computing infrastructure. The Distributed Resource Scheduler (DRS) is an autonomic computing system used by VMware that decides how resources will be allocated from a virtual pool. Currently the DRS uses predefined reservations and aggregated statistics to make efficient allocation decisions. However the DRS could function more efficiently if it could recognize underlying trends such as periodicity. More efficient utilization will translate into lower costs for computing power.

1.1 VIRTUALIZATION

Virtualization is the concept of adding a layer of software between hardware and applications, with the goal of abstracting the resources away from the application, providing standardization. This allows for much greater utilization, scalability, and autonomic computing.

1.2 TIME SERIES ANALYSIS

A time series is a sequence of data measured at regular intervals. Time series analysis consists of using mathematical techniques to extract meaningful statistics from time series data. Often times, there are trends hidden in what appears to be random data. These trends can in principle be identified through the use of sophisticated mathematics.

2 PROJECT GOALS

VMware would like to improve the DRS by using additional time series analysis techniques. Improving the DRS will result in higher virtual machine happiness (the percentage of the virtual machine's demand satisfied), which can translate into higher stability. Our project will involve building some of the components that go into performing time series analysis on data from virtualized pools of computing resources. These components comprise the following:

1. Collect performance statistics from hosts

2. Add visualization support to observe time series

3. Answer what-if queries based on time series analysis

4. Add support for other useful queries from administrators

3 IMPLEMENTATION SPECIFICS

Languages
 Java, C++, Python

Framework
 ESX Hosts, Virtual Center, Proprietary Statistics SDK

Tools
 VMware Time Series Analysis Libraries, Jenkins, GitHub