Project Overview

- Companies, with sufficient demand and resources, are operating their own data centers using VMware's vSphere to guarantee autonomy and reliability.
- VMware uses a virtualization hierarchy shown to the right.
- Tasks or applications run on virtual machines (VMs), each of which run an OS abstracting the ESX host's physical hardware.
- VMs can migrate between ESX Hosts, within their cluster, to evenly distribute work.
- This mapping problem can be reduced to a vectorized bin-packing NP-Complete problem.
- Applications use a variable amount of resources (cpu, memory, network, disk), these resources can often become over- or under-utilized which means inefficient resource allocation.

Information Flow

Collector/Retriever (vSphere Java API)

Collects real time statistics every 20 seconds from vCenter.

SQL Set

Stores and organizes data from the collector.

SQL Database

Time Series Module (Java & R)

Retrieves, formats, and performs time-series analysis on statistics from the database.

RPC Call

Visualizes statistics and provides the user with functionality control for this application.

Virtualization Hierarchy

- vCenter
- Cluster 1
- ESX Host 1
- VM 1
- VM 2
- ESX Host 2
- VM 3
- VM 4
- VM 5
- ESX Host 3
- VM 6
- VM 7
- VM 8

Instructors and Mentors

UCSB
Chandra Krintz
Tim Sherwood
Stratos Dimopoulos

VMware
Ajay Gulati
Banit Agrawal

Solution

- Having an efficient allocation of resources means lower costs and energy usage.
- We provide a simple and intuitive interface to view stats on the resource utilization of a datacenter, which allows network administrators to make informed resource scheduling decisions.
- We provide the tools necessary to compare various statistics across multiple entities with the option of correlation coefficients.
- We also provide functionality to do in-depth investigations and forecasts of individual entities.
- Time-series analysis has not been applied to VMware's software and our application offers an insightful exploitation of the information available.