ISSN: 0975-766X
CODEN: IJPTFI
Research Article

Available Online through
www.ijptonline.com

VIRTUALIZATION IN GRID COMPUTING

M.Sriram[1], Dr.R.M.Suresh[2]
Research Scholar, Department of Computer Science and Engineering, Bharath University, Chennai[1]
Professor, Department of Computer Science and Engineering, SLACE, Chennai[2]
Email: msr1sriram@gmail.com

Received on: 15.10.2016 *Accepted on: 22.11.2016*

Abstract

Cluster and grid computing have become extremely popular, significantly for scientific and engineering applications. Higher information measure of laptop Networks has any boosted development of cluster computing. Native and geographically distributed readying of computing resources is that the most outstanding differentiating issue between a cluster and grid computing primarily based setups. Cluster essentially needs readying of multiple nodes at one location connected via a high speed switch, whereas grid computing relies on access to remote computing resources, which can be a cluster. desegregation software system with multi-layer integration has been the most thrust of analysis and development within the field of cluster and grid computing. we have a tendency to introduced the construct of virtual workspaces which might be organized, managed and deployed in an exceedingly Grid surroundings. Since clusters underlie most vital Grid deployments these days, during this paper. we have a tendency to extended the notion of virtual workspaces to incorporate virtual clusters. we have a tendency to describe changes to Grid design and judge virtual cluster creation and management, the impact of death penalty in virtual clusters on applications furthermore because the chance of running many virtual clusters on one physical cluster. The study focuses on the impact within the space of virtual cluster and virtual workspaces in virtual machines by supporting virtualization.

Keywords: Virtualization – Virtual Cluster – Grid computing - virtual machine – virtual space.

1. Introduction

Cluster Computing/Grid Computing technology is changing into standard for scientific and engineering applications, thanks to improved high performance/ low price magnitude relation. With the big increase in computing power on desktop PCs in recent years, it's become attainable to integrate distributed computing Resources into one image, therefore providing increased computing power. Cluster refers to 2 or a lot of nodes connected along either at intervals a cupboard or over a local area network, giving the users one system image [1] and usually it's options like high performance, expandability, measurability, high turnout, high handiness and everyone these at low price. Presently, table high PCs ar integrated as a cluster resource and their power is devoted for the applications running on the cluster [2]. Cluster primarily based computing is quick maturing as a technology and it's poised to require an enormous leap by Integrating the distributed cluster computing power into a grid primarily based infrastructure.

Grid Computing refers to a gaggle of clusters operative at geographically distributed locations, connected over high-speed networks.

2. Virtualization and Grid Computing

Virtualization may be a framework or methodology of dividing the resources of a laptop into multiple execution environments, by applying one or a lot of ideas or technologies like hardware and software system partitioning, sharing, partial or complete machine simulation, emulation, quality of service, and lots of others. Virtualization of resources will be achieved employing a variety of technologies. A virtual machine [5] is AN emulation of lower layers of a laptop abstraction on behalf of upper layers. A VM illustration contains a full image of RAM, disk, and different devices. A virtual machine monitor (VMM) may be a software system method that manages the hardware resources of the important machine among instances of VMs, therefore permitting multiple instances of VMs to run at the same time on a similar hardware. With superior isolation properties, fine-grained resource management, and also the ability to instantiate severally organized guest environments on a number resource, virtual machines give a decent platform for Grid computing [8, 9]. The In-Vigo project [13, 14] and also the associated Virtuoso project [15] explored a number of the problems concerned in combining Grid and virtual machine technology particularly as relates to networking and readying. Our approach differs in this it focuses on virtual workspaces, superior entities that require to be managed severally of their readying, treating virtual machines collectively of their implementations Driven by community requirements; we have a tendency to additionally specialise in clusters as a primary Grid platform. Such a spotlight has been recognized by different teams. The Cluster on Demand infrastructure [4] initial introduced the notion of a virtual cluster (albeit in its initial iteration not mistreatment virtual machines).

2.1 Virtual Clusters

Distinguish 2 varieties of nodes in an exceedingly virtual cluster: a head-node and employee nodes. the aim and configuration of a head-node arE usually completely different from those of employee nodes, particularly in software system and operational setup. Although employee node configurations arE similar, they will be allotted totally completely different names or their standing could also be different (for example, some nodes might not be operational). For these reasons, I represent every node of a cluster by a separate atomic space, every with its own data and image handle, as delineated in [10]. a collection of atomic workspaces representing the nodes of the cluster is then wrapped by AN XML section containing the data regarding the cluster as an entire like its kind (cluster/atomic), name, variety of nodes, or time it absolutely was instantiated. All different info a few space comes from the data of the atomic workspaces describing those nodes.

Virtual clusters work by suggests that of:

• Collections of atomic workspaces.

• a technique of describing collections of atomic workspaces.

• a technique of managing collections of atomic workspaces.

The space data schema is extended to incorporate aggregations of atomic workspaces. Workspaces currently become combination workspaces. AN combination space contains sets of spaces; every set contains solid workspace configurations. it's through the mixture of sets at intervals AN combination space, that a heterogeneous cluster could also be delineated employing a single space data schema. significantly the flexibility of combination workspaces to outline heterogeneous clusters permits each service and employee nodes to be delineated. Virtual cluster workspaces for grid computing Virtual machines give a promising platform for machine Grids. Virtualization of underlying hardware they permit mental representation of a replacement, severally organized guest surroundings on a number resource. additionally, they provide the advantages of isolation and fine-grain social control and, given the flexibility to set up their state and migrate, supply multiplied flexibility to environments within the Grid. To take advantage of this new technology in Grid computing, introduced the construct of virtual workspaces which might be organized, managed and deployed in an exceedingly Grid surroundings. Since clusters underlie most vital Grid deployments these days, during this paper I extended the notion of virtual workspaces to incorporate virtual clusters. I describe changes to Grid design and judge virtual cluster creation and management, the impact of death penalty in virtual clusters on applications furthermore because the chance of running many virtual clusters on one physical cluster.

3. Virtual Workspaces

Grid computing involves the sharing of heterogeneous Grid resources between VOs, every with probably conflicting resource needs. Such shared use typically leads to under-utilization of Grid resources caused by a match between resources offered and also the application needs [4] of the shared users'. but sharing isn't the sole reason behind resource below utilization. [10] Draw attention to the potential incompatibility of a cluster's put in libraries across VOs as another issue with current Grid use. Virtual Workspaces (VW) are an endeavor to handle the subsequent 3 issues with

Grid clusters:

• Lack of performance isolation.

• very little management over resource sharing.

• Fine grained usage troublesome to enforce

3.1 Describing Virtual Workspaces

A space description ought to contain adequate info for a readying service to make the surroundings painted by this space. This info is of 2 kinds:

(1) Description of packages or different knowledge that require to be obtained from probably external sources and place along (such as a software system installation package or a VM image), and

(2) Readying supplying info, that must be taken and organized at readying time (such as network affiliation configuration for a VM).

The quantity of data that has to transfer and also the quantity of readying-time configuration can rely upon each space implementation (installation-based readying versus deploying a VM image) and also the readying service implementation (deploying a VM supported a pre-configured image [16] or purification configuration at deployment time [17]). Thus, a good vary of approaches to space description ar attainable, from the easy however inflexible (e.g., a pointer to a VM image and a default readying configuration) to the advanced however powerful (e.g., capricious on-the-fly configuration of AN image).

3.2 Workspaces as Virtual Machines

A virtual machine (VM) [3] provides a virtualization of a physical host machine.

Software running on the host, usually known as a virtual machine monitor (VMM) or hypervisor is chargeable for supporting this abstraction by intercepting and emulating directions issued by the guest machines. A hypervisor additionally provides AN interface permitting a consumer to begin, pause, serialize, and stop working multiple guests. A VM illustration (VM image) consists of a full image of a VM RAM, disks (or partition) pictures, and configuration files. Recent exploration of para virtualization techniques [9] has crystal rectifier to substantial performance enhancements in Virtualization technologies, creating virtual machines a pretty choice for superior applications. Virtual machines enable a consumer to make a custom execution surroundings organized with a needed package, software system stack and access policies and so deploy it on any resource running a hypervisor. Further, VM state could also be serialized into a VM image, permitting the consumer to pause or stop working VM operation, and resume it at {a completely different|a special|a unique|a distinct} time and in an exceedingly different location, decoupling image preparation from its readying and enabling migration. additionally, virtual machines supply wonderful social control of resource usage: usually, a virtual machine is organized with a particular memory and disk size and a few, like [9], enable those qualities to be managed throughout readying. mistreatment schedulers, such as , a consumer will assign a proportion of hardware to a given virtual machine effectively control the hardware usage of the cluster of processes encapsulated in it. For these reasons, VMs give a wonderful implementation choice for workspaces: the configuration of a VM image will mirror a workspace's software system needs whereas the hypervisor will make sure the social control of hardware properties. Overall, VMs have the advantage of each flexibility and speed of readying. the flexibleness stems from the VM construct, that provides AN abstract illustration of state which will be deployed anyplace a hypervisor is gift. In fashionable hypervisors such readying is quick: show that deploying a VM will take but a second [4], that is resembling the overhead induced by the Grid tools. additionally to the present, hypervisor's ability to supply fine-grain social control makes virtual machines a perfect resolution for short readying of unambiguously organized workspaces requiring controlled resource usage. software [7], a cluster of virtual machines organized with the software system configuration needed by Open Science Grid (OSG) [8], and a collection of physical machines organized with Xen hypervisor [20] all represent a space.

3.3 Virtual Cluster Workspaces

Workspaces enforced via any of the ways simply delineated will be classified to make virtual clusters of varied topologies. A virtual cluster space will be made via, as an example, the Cluster- on- demand (COD) infrastructure [7], AN existing cluster with tools for dynamically enabling access, or a cluster of virtual machines. When employing a virtual machine implementation, a virtual cluster space is enforced in terms of multiple VM pictures that will represent specialised nodes like employee or head nodes.

4. Performance Problems and Analysis

Performance comparison of VCG as AN integrated system of pooled-in resources against networked resources, that ar accessible to users, relies on multiple and varied parameters, that ar mentioned during this section.

4.1 Job Size Assessment Criteria

Size of a task or employment will be evaluated in several ways that; following ar the foremost ways during which job size assessment will be made: Job size definition: usually, quality of Computing task is evaluated on the idea of hardware time required; total I/O certain dependency, total computer file size, total swap/ temporary file area size etc. once multiple tasks are handled by a hardware, the hardware time gets divided among the tasks, whereas different higher than mentioned parameters stay guaranteed to the task itself. Small job vs. giant job: it's necessary to know the associated problems once a tiny low task runs on a machine as compared to the massive tasks, that runs on a machine. Even before we have a tendency to assess the problems, the definition of tiny and enormous jobs ar basically to be processed. tiny jobs gets completed in an exceedingly affordable timeframe that may well be starting from jiffy to few hours, whereas the massive jobs might take abundant larger time, from few tens of hours for rader. Statistically speaking giant jobs additionally want giant swap and file storage, however this can not be a generalized statement. CPU certain vs. Input/output certain: Some tasks are extremely hardware bound, which means they consume a lot of hardware power than the memory area, either the most memory or the swap/ temporary area. Whereas some tasks may well be extremely Input/output certain which means that they consume giant memory area and use very little of hardware time. Of course, it's price mentioning here that some tasks may well be a mixture of the 2 options, resulting in moderate use of hardware power and affordable memory usage.

4.2 Network Resources VS. Pooled Resources

Networked resources: once the resources are solely connected on a network and ar accessible to the user as ‘stand alone’ resource, task submission and execution of the task is ruled by the subsequent factors:

Access right to submit the task on the resource: The user is needed to form a ‘login’ to the system, that the user wants a ‘login id’ and ‘password’. The user additionally wants disc space on the node for supply files, input and output files etc. Job preparation for execution is additionally to be done on every machine by the user, by login to the system. Management of user’s knowledge (program and knowledge files – input and output) for backup etc. is additionally to be taken care on every machine separately; this can be a fancy job for the computing resource supplier. Job monitoring: The user must monitor standing of the task specifically from the node on that task has been submitted and will cause plenty of inconvenience in terms of user’s own time utilization. Object code compatibility: The user must prepare the task as per the resource to satisfy the thing code compatibility needs.

Parameters on the performance of the pooled resources:

Network latency: The time needed to repeat, Executable and input file files guide the initiation of the task on remote resource. Network latency plays a important role for this parameter. Network latency additionally plays a task for task observation, that is mechanically updated within the job standing table by task daemon. Operating System over head: The VCG will load the package of the pooled resource, however doesn't have any further load as compared to a task which can be submitted to the node directly. Memory management: Since the memory of the Pooled resource goes to be consumed, even if Temporarily, by the tasks coming back from remote users, it's not extremely within the management of the owner of the resource and projected VCG design doesn't take under consideration memory management of the pooled resources.

5. Connected Work

Past comes on configuration and customization of computing environments for clusters and Grids embody the work of Krsul at al. [11] and Papadopoulos et al. [14]. Krsul et al. projected a VM creation and customization framework known as VMPlant [11] for virtual Grid environments like In-VIGO [1]. It provides a graph-based configuration interface to the user, that encodes every configuration as a node and dependency between every configuration as directed edges. The package-based system installation in our framework is analogous to their graph-based configuration; In essence, the package-dependency structure will be painted as a DAG. Their approach be 2 points. First, exploit AN existing infrastructure for cluster configuration management and extend it for virtual clusters, whereas they have confidence their own custom configuration schema. Therefore, whereas our framework needs the user to arrange solely declarative configuration files in most eventualities, they need the user and also the resource supplier to supply actual implementation, usually written in scripts, of such configurations furthermore.

Second, they are doing not enable the user to make VMs on the fly; rather they assume offline-created VM pictures on the market for each user request. Such assumption is unlikely to be command in multi-organization, multi-site, heterogeneous Grid environments. Papadopoulos et al. conferred their cluster installation and management tool known as Rocks in [14]. potency and measurability of installation has been another necessary topic in providing user-specific environments since the size of underlying physical clusters has continuing to extend. VMPlants support recycle of part organized VMs through partial graph matching [11].

6. Conclusion

Cluster primarily based computing has provided a replacement thrust to computing, by approach of mixing the computing power of enormous variety of resources into one system image. A example model of Virtual Computing

Grid has been deployed and it's giving promising results. System will be increased in future to supply improved error handling and system flow. We tried to supply an outline on this and future use (and benefit) of virtualization techniques for Grids. The selection of usage eventualities might not be intensive however most likely offers a quite thorough list of unreal changes on however Grids ar used or deployed. In our read this trend isn't actively steered or influenced by the Grid communities. Instead, we have a tendency to see that the common trend of moving towards virtualization in knowledge centres is going on straight away and can stick with United States of America for an extended time. This affects however we have a tendency to contend with resources normally, together with Grids The paper is meant to supply a outline on the impact of this trend and supply food of thought for current developments. We conferred many eventualities that we have a tendency to see as relevant for adopting virtualization within the light-weight of Grids environments. a number of these eventualities ar presently below development by completely different teams, corresponding results have become on the market within the close to future. another eventualities want a lot of technological advancements in numerous areas. Thus, it's unlikely that single analysis teams are able to overcome those challenges in an exceedingly short timeframe. Here, we have a tendency to expect AN biological process approach towards these eventualities once partial options become unremarkably on the market. The current interest and success in cloud computing was created attainable through the adoption of virtualization. Cloud computing limits a number of the challenges and difficulties that Grid computing is facing. Similarly, Grids and their current users can adopt a lot of virtualization models and become a lot of freelance from the specifics of resources and applications. Thus, we have a tendency to believe that the conferred models can become key use cases within the close to and mid-term future.

7. References

1. Rajkumar, Buyya: design Alternatives for Single System Image Clusters. Conference on High Performance Computing on Hewlett-Packard Systems (HiPer'99), Tromse, Norge (1999).
2. Rajkumar, Buyya: High Performance Cluster Computing: design and Systems Volume one. learner Hall PTR, NJ, USA(1999).
3. Rajkumar, Buyya: High Performance Cluster Computing: Programming and Applications Volume two. learner Hall PTR, NJ, USA (1999).
4.

4. Meyer, R.A. and L.H. Seawright, A Virtual Machine sharing System. IBM System Journal, 1970. 9(3): p. 199-218.
5. Rajkumar, Buyya: Steve, Chapin, David, DiNucci, ‘Architectural Models for Resource Management within the Grid. the primary IEEE/ACM International Workshop on Grid Computing (GRID 2000), Springer Verlag LNCS Series, Bangalore, India (Dec. 2000).
6. A.Matsunaga et al. 2005. On the employment of Virtualisation and repair Technologies to Enable Grid-Computing.
7. Raman, R., M. Livny, and M. Solomon, Matchmaking: AN protractible Framework for Distributed Resource Management. Cluster Computing: The Journal of Networks, software system Tools and Applications, 1999. 2: p. 129-138.
8. Space Management Service: http://www.mcs.anl.gov/workspace/.
9. X.Zhang et al. 2005, Virtual Cluster Workspaces for Grid Applications.
10. I. Krsul, A. Ganguly, J. Zhang, J. A. B. Fortes, and R. J. Figueiredo. Vmplants: Providing and managing virtual machine execution environments for grid computing. In Proceedings of the 2004 ACM/IEEE conference on Supercomputing,pages 7–18, Pittsburgh, PA, November 2004.
11. Figueiredo, R., P. Dinda, and J. Fortes. A Case for Grid Computing on Virtual Machines.
12. S. Adabala, V. Chadha, P. Chawla, R. Figueiredo, J. Fortes, I. Krsul, A. Matsunaga, M. Tsugawa, J. Zhang, M. Zhao, L. Zhu, and X. Zhu. From virtualized resources to virtual computing grids: the in-vigo system. Journal of Future Generation Computing Systems,
13. P. M. Papadopoulos, M. J. Katz, and G. Bruno. Npaci rocks: Tools and techniques for simply deploying manageable UNIX system clusters. In Proceedings of the International Conference on Cluster Computing, 2001.
14. Chervenak, A., E. Deelman, I. Foster, L. Guy, W. Hoschek, A. Iamnitchi, C. Kesselman, P. Kunst, M. Ripenu, B. Schwartzkopf, H. Stockinger, K. Stockinger, and B. Tierney. Giggle: A Framework for Constructing scalable duplicate Location Services. in SC'02: High Performance Networking and Computing. 2002.
15. Keahey, K., I. Foster, T. Freeman, X. Zhang, and D. Galron. Virtual Workspaces within the Grid. in Europar. 2005. Lisbon, Portugal.
16. Krsul, I., A. Ganguly, J. Zhang, J. Fortes, and R. Figueiredo. VMPlants: Providing and Managing Virtual Machine Execution Environments for Grid Computing. in SC04. 2004. Pittsburgh, PA.