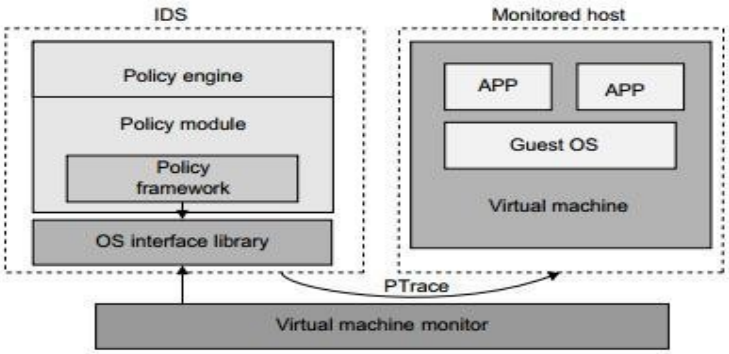
- **Steps 4 and 5: Commit and activate the new host.** After all the needed data is copied, onthe destination host, the VM reloads the states and recovers the execution of programs in it, and the service provided by this VM continues. Then the network connection is redirected to the new VM and the dependency to the source host isb cleared. The whole migration process finishes by removing the original VM from the source host.
- Before copying the VM with 512 KB files for100 clients, the data throughput was 870 MB/second. The first pre copy takes 63 seconds, during which the rate is reduced to 765 MB/second. Then the data rate reduces to 694MB/second in 9.8 seconds for more iterations of the copying process.
- The system experiences only 165 ms of downtime, before the VM is restored at the destination host.

3.9 VIRTUALIZATION FOR DATA-CENTER AUTOMATION

Part –B	1. What do you mean by Data Center Automation using Virtualization? (AU/April/May 2017)

- The dynamic nature of cloud computing has pushed data center workload, server, and even hardware automation to whole new levels.
- Now, any data center provider looking to get into cloud computing must look at some form of automation to help them be as agile as possible in the cloud world.
- New technologies are forcing data center providers to adopt new methods to increase efficiency, scalability and redundancy.
- Let's face facts; there are numerous big trends which have emphasized the increased use of data center facilities. These trends include:
    - More users
    - More devices
    - More cloud
    - More workloads
    - A lot more data
- A s infrastructure improves, more companies have looked towards the data center provider to offload a big part of their IT infrastructure. With better cost structures and even better incentives in moving towards a data center environment, organizations of all sizes are looking at colocation as an option for their IT environment.
- W ith that, data center administrators are teaming with networking, infrastructure and cloud architects to create an even more efficient environment. This means creating intelligent systems from the hardware to the software layer. This growth in data center dependency has resulted in direct growth around automation and orchestration technologies.
- N ow, organizations can granularly control resources, both internally and in the cloud. This type of automation can be seen at both the software layer as well as the hardware layer.
- V endors like BMC, ServiceNow, and Microsoft SCCM/SCOM are working towards unifying massive systems under one management engine to provide a single pain of glass into the data center workload environment

- Furthermore, technologies like the Cisco UCS platform allow administrators to virtualize the hardware layer and create completely automated hardware profiles for new blades and servers.
- This hardware automation can then be tied into software-based automation tools like SCCM. Already we're seeing direct integration between software management tools and the hardware layer



- Finally, from a cloud layer, platforms like CloudStack and OpenStack allow organizations to create orchestrated and automated fluid cloud environments capable of very dynamic scalability.
- Still, when a physical server or hardware component breaks – we still need a person to swap out that blade.

Server layer

- Server and hardware automation have come a long way. As mentioned earlier, there are systems now available which take almost all of the configuration pieces out of deploying a server.
- Administrators only need to deploy one server profile and allow new servers to pick up those settings. More data centers are trying to get into the cloud business.
- This means deploying high-density, fast-provisioned, servers and blades. With the on-demand nature of the cloud, being able to quickly deploy fully configured servers is a big plus for staying agile and very proactive.

Software layer

- Entire applications can be automated and provisioned based on usage and resource utilization.
- Using the latest load-balancing tools, administrators are able to set thresholds for key applications running within the environment.
- If a load-balancer, a NetScaler for example, sees that a certain type of application is receiving too many connections, it can set off a process that will allow the administrator to provision another instance of the application or a new server which will host the app.

Virtual layer.

- T he modern data center is now full of virtualization and virtual machines.
- I n using solutions like Citrix's Provisioning Server or Unidesk's layering software technologies, administrators are able to take workload provisioning to a whole new level.
- I magine being able to set a process that will kick-start the creation of a new virtual server when one starts to get over-utilized.
- N ow, administrators can create truly automated virtual machine environments where each workload is monitored, managed and controlled.

Cloud layer.

- T his is a new and still emerging field. Still, some very large organizations are already deploying technologies like CloudStack, OpenStack, and even Open Nebula. Furthermore, they're tying these platforms in with big data management solutions like MapReduce and Hadoop.
- O rganizations can deploy distributed data centers and have the entire cloud layer managed by a cloud-control software platform.
- E ngineers are able to monitor workloads, how data is being distributed, and the health of the cloud infrastructure.
- T he great part about these technologies is that organizations can deploy a true private cloud, with as much control and redundancy as a public cloud instance.

Data center layer.

- Although entire data center automation technologies aren't quite here yet, we are seeing more robotics appear within the data center environment.
- Robotic arms already control massive tape libraries for Google and robotics automation is a thoroughly discussed concept among other large data center providers. Many experts agree that eventually, data center automation and robotics will likely make its way into the data center of tomorrow.
- The need to deploy more advanced cloud solution is only going to grow.
- More organizations of all verticals and sizes are seeing benefits of moving towards a cloud platform. At the end of the day, all of these resources, workloads and applications have to reside somewhere. That somewhere is always the data center.
- In working with modern data center technologies administrators strive to be as efficient and agile as possible.
- This means deploying new types of automation solutions which span the entire technology stack. Over the upcoming couple of years, automation and orchestration technologies will continue to become popular as the data center becomes an even more core piece for any organization.

UNIT IV PROGRAMMING MODEL

Open source grid middleware packages – Globus Toolkit (GT4) Architecture , Configuration – Usage of Globus – Main components and Programming model - Introduction to Hadoop Framework - Mapreduce, Input splitting, map and reduce functions, specifying input and output parameters, configuring and running a job – Design of Hadoop file system, HDFS concepts, command line and java interface, dataflow of File read & File write.

4.1 OPEN SOURCE GRID MIDDLEWARE PACKAGES

- The Open Grid Forum and Object Management are two well- formed organizations behind the standards
- Middleware is the software layer that connects software components. It lies between operating system and the applications.
- Grid middleware is specially designed a layer between hardware and software, enable the sharing of heterogeneous resources and managing virtual organizations created around the grid.

The popular grid middleware are

1. BOINC -Berkeley Open Infrastructure for Network Computing.
2. UNICORE - Middleware developed by the German grid computing community.
3. Globus (GT4) - A middleware library jointly developed by Argonne National Lab., Univ. of Chicago, and USC Information Science Institute, funded by DARPA, NSF, and NIH.
4. CGSP in ChinaGrid - The CGSP (ChinaGrid Support Platform) is a middleware library developed by 20 top universities in China as part of the Open source grid middleware
5. Condor-G - Originally developed at the Univ. of Wisconsin for general distributed computing, and later extended to Condor-G for grid job management.
6. Sun Grid Engine (SGE) - Developed by Sun Microsystems for business grid applications. Applied to private grids and local clusters within enterprises or campuses.
7. gLight -Born from the collaborative efforts of more than 80 people in 12 different academic and industrial research centers as part of the EGEE Project, gLite provided a framework for building grid applications tapping into the power of distributed computing and storage resources across the Internet.

Grid Standards and APIs

- The Open Grid Forum and Object Management Group are two well-formed organizations behind those standards.
- OGSA - Open Grid Services Architecture
- GLUE - Grid Laboratory Uniform Environment
- SAGA - Simple API for Grid Applications
- GSI - Grid Security Infrastructure
- OGSI - Open Grid Service Infrastructure
- WSRE - Web Service Resource Framework
- The grid standards have guided the development of several middleware libraries and API tools for grid computing.
- They are applied in both research grids and production grids today.
- Research grids tested include the