How do programs operate? What is Virtualization? What is it good for? What is it not good for? Conclusion References

CS-495/595
Big Data
Virtualization

Dr. Chuck Cartledge

21 Jan. 2015
# Table of contents

1. How do programs operate?
2. What is Virtualization?
3. What is it good for?
4. What is it not good for?
5. Conclusion
6. References
A 50,000 foot view

What are the layers in this cake?

- User — the person (or thing) that wants something done
- Application — the program that does the work
- Hardware — the silicone, copper, other tangibles that generate heat
- Operating system — arbitrates between multiple programs and limited resources

**Figure:** Notional view of computer layers [3]
Focusing on the OS

What does it do?

- Provides a user interface (maybe a Command Line Interface)
- Schedules access to the hardware
- Schedules the functions of the CPU

Figure: A few OS details [2].

An OS is a program (albeit, a large program). What if we could write a program that would run an OS as an application??
Tricking the upper layer.

- Higher layers rely on lower layers for services
- Layers create interfaces
- Interfaces allow for hiding details

Virtualization software allows applications that previously ran on separate computers to run on one server machine.

What is real? And, what is virtual? [1]

As long as the lower layer supplies all the services, the upper layer won’t know where the services originated.
One hardware suite can run many OS in virtual machines.

- Ultimately the hardware determines how many virtual machines can be run
- Faster CPU(s), more RAM, more network connections, more disks, . . . , more is better
- Fewer actual machines usually means lower power, lower cooling, cheaper upgrade path

With clever software, almost anything can be virtualized. Hadoop is clever software.
Anything that has to be fast.

- Underlying hardware suite is shared across all “machines”
- Mission critical applications
In summary.

- To use virtual machines, or
- To not use virtual machines.

It depends on what is important.
What have we covered?

- There are lots of free virtual machine programs for you to try
- They are good for trying, testing, experimenting
- They are good for maximum use of available hardware
- They are not good for real-time applications
- HDFS is a virtual file system
References

