IIT KHARAGPUR

NPTEL ONLINE
CERTIFICATION COURSES

CLOUD COMPUTING

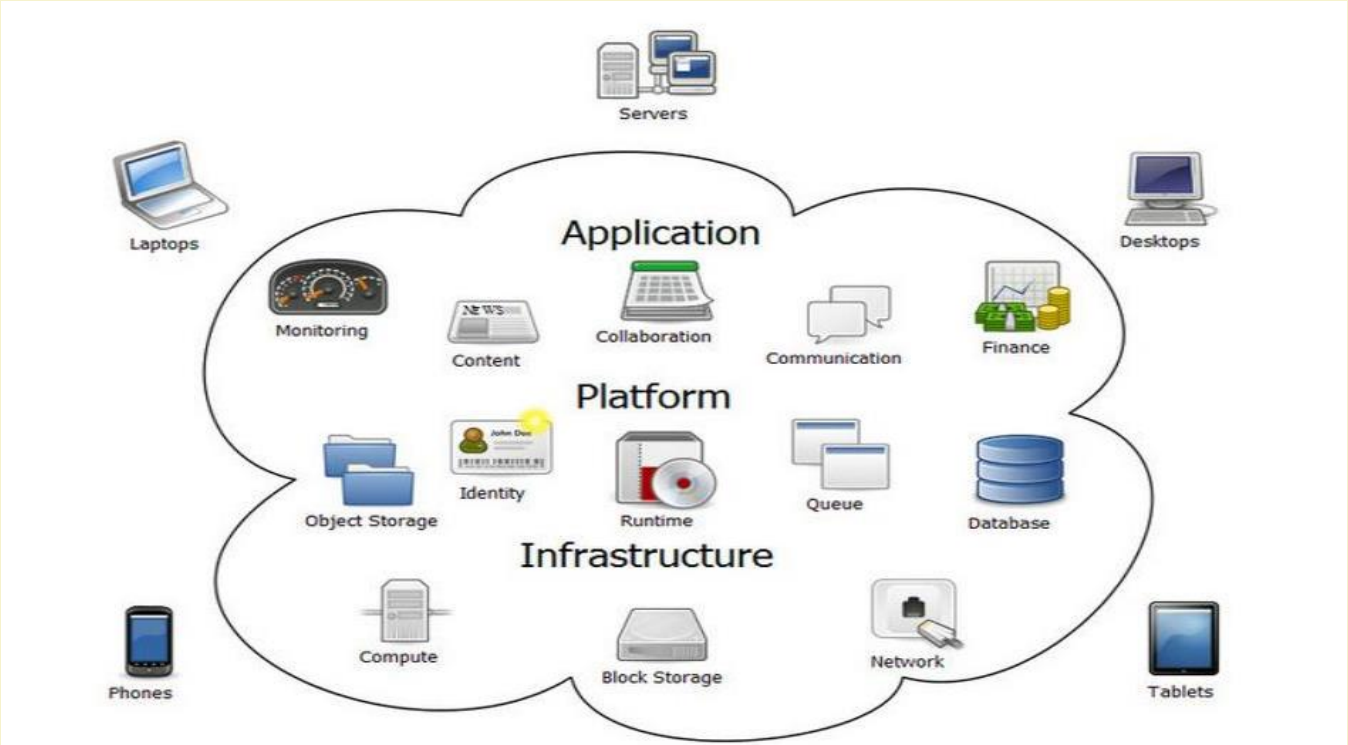
Resource Management - I

PROF. SOUMYA K. GHOSH

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

IIT KHARAGPUR

Different Resources in Computing



Source: http://www.cse.hcmut.edu.vn/~ptvu/gc/2012/GC-pp.pdf

Resources types

- **Physical resource**
  - ❑ Computer, disk, database, network, scientific instruments.
- **Logical resource**
  - ❑ Execution, monitoring, communicate application .

Source: http://www.cse.hcmut.edu.vn/~ptvu/gc/2012/GC-pp.pdf

Resources Management

- The term ***resource management*** refers to the operations used to control how capabilities provided by Cloud resources and services cane be made available to other entities, whether users, applications, services in an *efficient* manner.

Source: http://www.cse.hcmut.edu.vn/~ptvu/gc/2012/GC-pp.pdf

Data Center Power Consumption

- Currently it is estimated that servers consume 0.5% of the world's total electricity usage.
- Server energy demand doubles every 5-6 years.
- This results in large amounts of CO_2 produced by burning fossil fuels.
- Need to reduce the energy used with minimal performance impact.

Ref: Efficient Resource Management for Cloud Computing Environments, by Andrew J. Younge, Gregor von Laszewski, Lizhe Wang, Sonia Lopez-Alarcon, Warren Carithers,

Motivation for Green Data Centers

Economic

- New data centers run on the Megawatt scale, requiring millions of dollars to operate.
- Recently institutions are looking for new ways to reduce costs
- Many facilities are at their peak operating stage, and cannot expand without a new power source.

Environmental

- Majority of energy sources are fossil fuels.
- Huge volume of CO_2 emitted each year from power plants.
- Sustainable energy sources are not ready.
- Need to reduce energy dependence

Green Computing ?

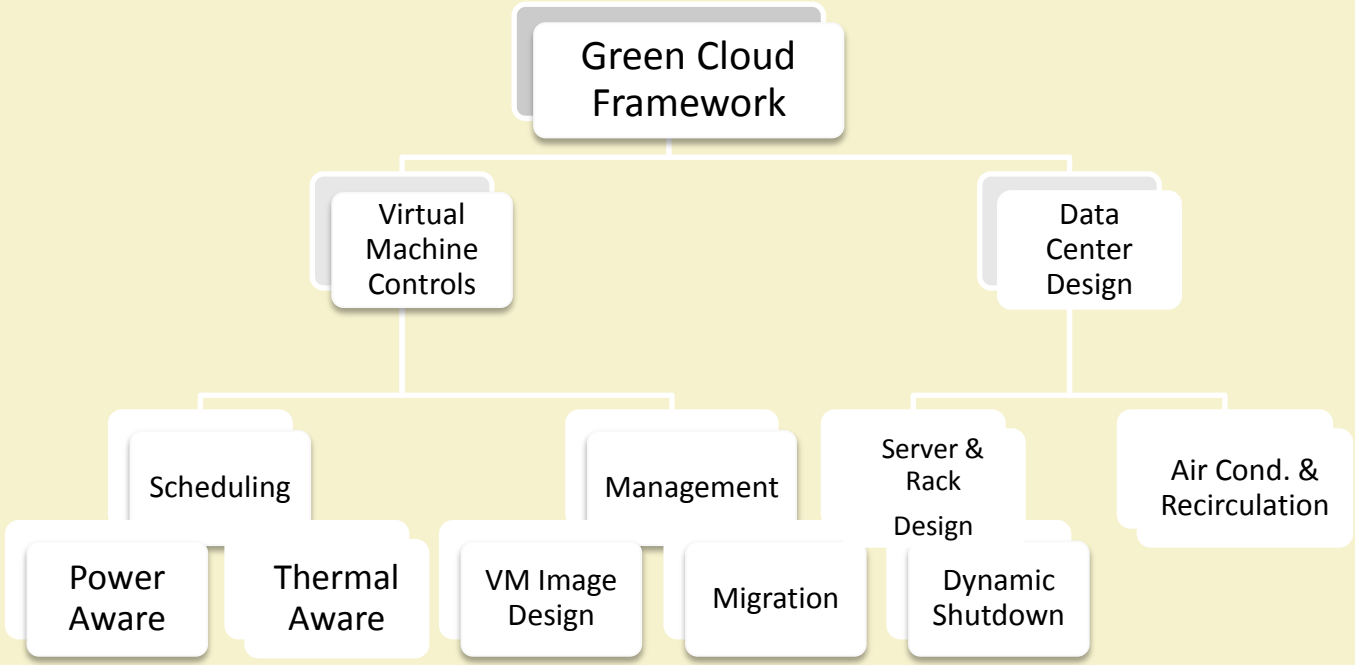
- Advanced scheduling schemas to reduce energy consumption.
  - Power aware
  - Thermal aware
- Performance/Watt is not following Moore's law.
- Data center designs to reduce Power Usage Effectiveness.
  - Cooling systems
  - Rack design

Research Directions

How to conserve energy within a Cloud environment.

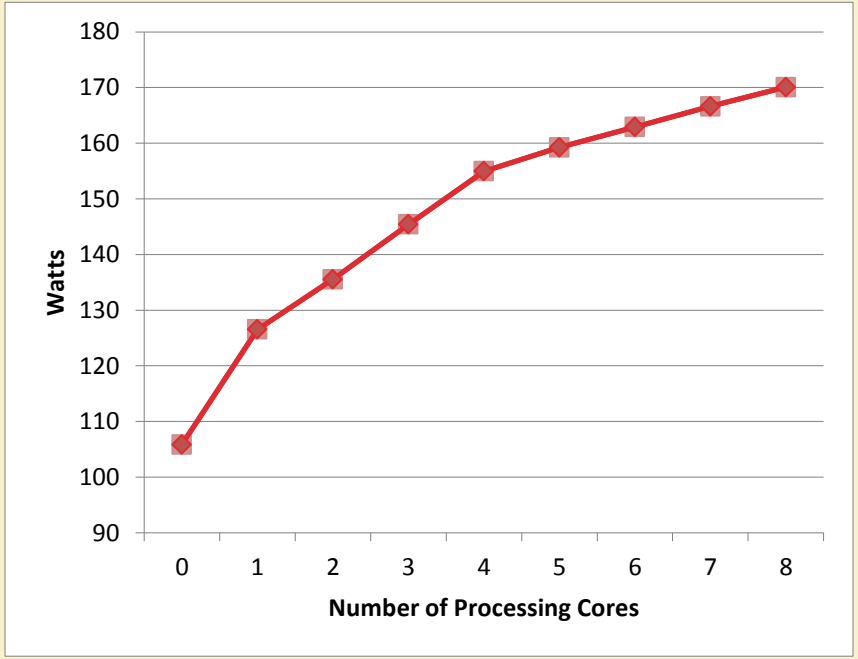
- Schedule VMs to conserve energy.
- Management of both VMs and underlying infrastructure.
- Minimize operating inefficiencies for non-essential tasks.
- Optimize data center design.

Steps towards Energy Efficiency



VM scheduling on Multi-core Systems

- There is a nonlinear relationship between the number of processes used and power consumption
- We can schedule VMs to take advantage of this relationship in order to conserve power



Power consumption curve on an Intel Core i7 920 Server (4 cores, 8 virtual cores with Hyperthreading)

Scheduling

Power-aware Scheduling

- Schedule as many VMs at once on a multi-core node.
  - Greedy scheduling algorithm
  - Keep track of cores on a given node
  - Match VM requirements with node capacity

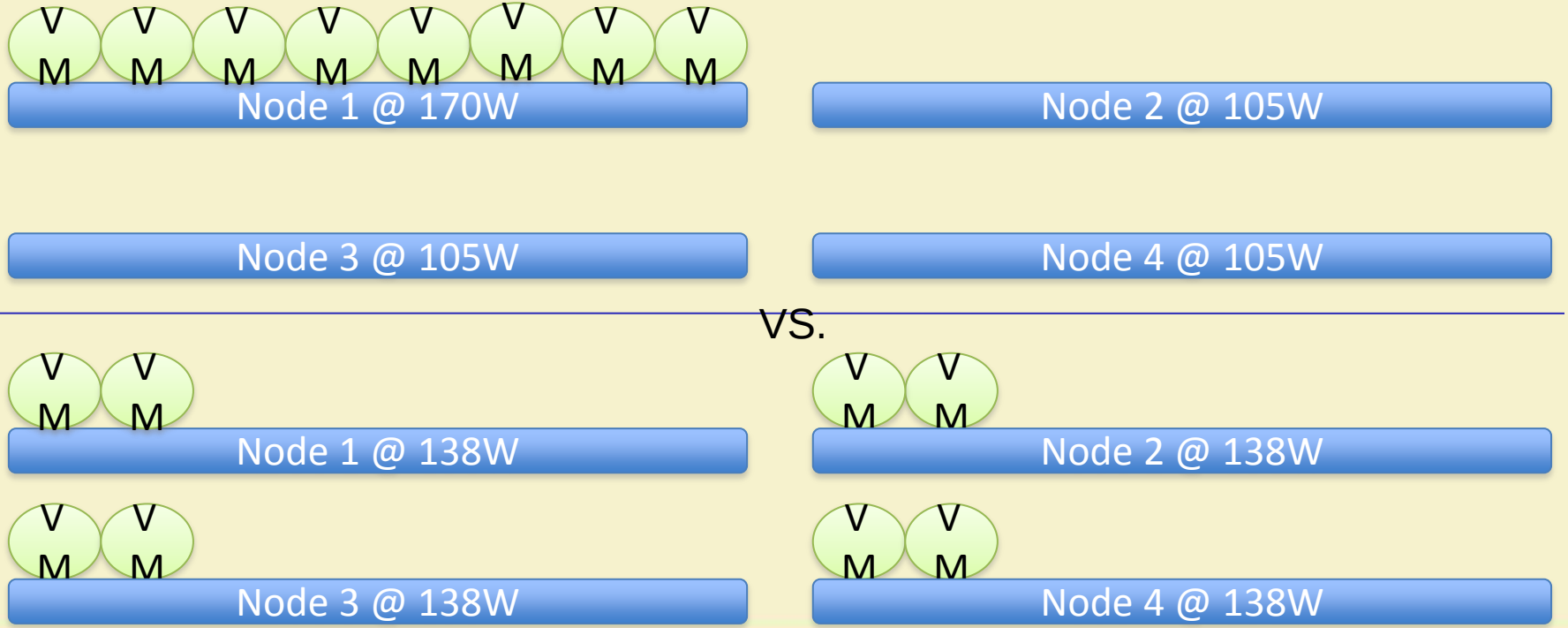
Algorithm 1 Power based scheduling of VMs

```
FOR i = 1 TO i ≤ |pool| DO
  pe_i = num cores in pool_i
END FOR

WHILE (true)
  FOR i = 1 TO i ≤ |queue| DO
    vm = queue_i
      FOR j = 1 TO j ≤ |pool| DO
        IF pe_j ≥ 1 THEN
        IF check capacity vm on pe_j THEN
          schedule vm on pe_j
          pe_j − 1
        END IF
      END IF
    END FOR
  END FOR
  wait for interval t
END WHILE
```

Scheduling

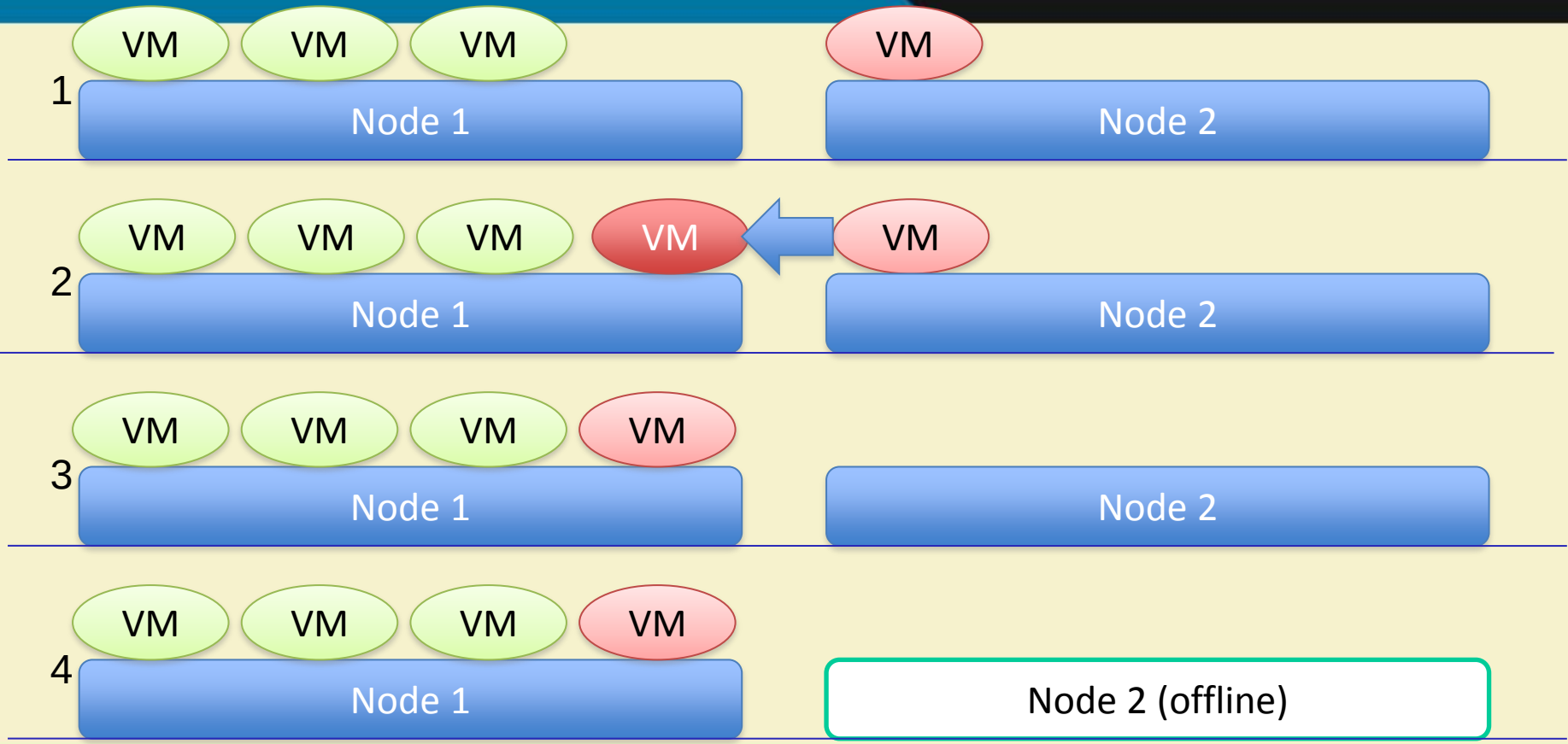
485 Watts vs. 552 Watts !



VM Management

- Monitor Cloud usage and load.
- When load decreases:
  - Live migrate VMs to more utilized nodes.
  - Shutdown unused nodes.
- When load increases:
  - Use WOL to start up waiting nodes.
  - Schedule new VMs to new nodes.

Management


1
VM
VM
VM
Node 1
VM
Node 2
2
VM
VM
VM
VM
Node 1
VM
Node 2
3
VM
VM
VM
VM
Node 1
Node 2
4
VM
VM
VM
VM
Node 1
Node 2 (offline)

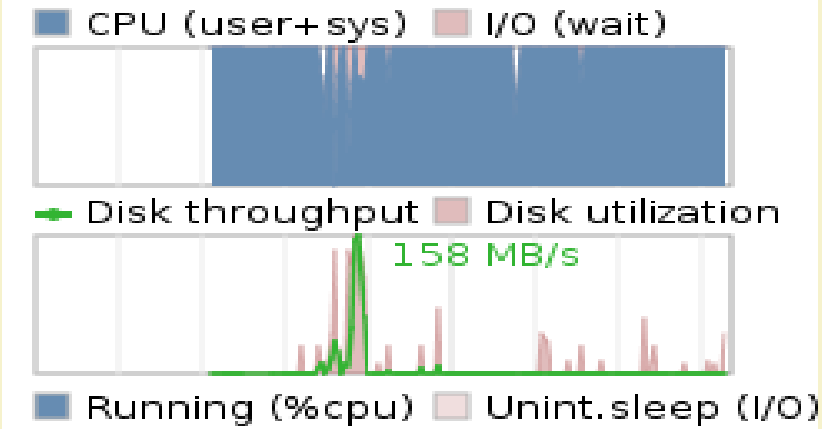
Minimizing VM Instances

- Virtual machines are loaded!
  - Lots of unwanted packages.
  - Unneeded services.
- Are multi-application oriented, not service oriented.
  - Clouds are based off of a Service Oriented Architecture.
- Need a custom lightweight Linux VM for service oriented science.
- Need to keep VM image as small as possible to reduce network latency.

Management

Typical Cloud Linux Image

- Start with Ubuntu 9.04.
- Remove all packages not
- required for base image.
  - No X11
  - No Window Manager
  - Minimalistic server install
  - Can load language support on demand (via package manager)
- Readahead profiling utility.
  - Reorder boot sequence
  - Pre-fetch boot files on disk
  - Minimize CPU idle time due to I/O delay
- Optimize Linux kernel.
  - Built for Xen DomU
  - No 3d graphics, no sound, minimalistic kernel
  - Build modules within kernel directly



VM Image Design

Energy Savings

- Reduced boot times from 38 seconds to just **8** seconds.
  - 30 seconds @ 250Watts is 2.08wh or .002kwh.
- In a small Cloud where 100 images are created every hour.
  - Saves .2kwh of operation @ 15.2c per kwh.
  - At 15.2c per kwh this saves $262.65 every year.
- In a production Cloud where 1000 images are created every minute.
  - Saves 120kwh less every hour.
  - At 15.2c per kwh this saves over 1 million dollars every year.
- Image size from 4GB to 635MB.
  - Reduces time to perform live-migration.
  - Can do better.

Summary - 1

- Cloud computing is an emerging topic in Distributed Systems.
- Need to conserve energy wherever possible!
- Green Cloud Framework:
  - Power-aware scheduling of VMs.
  - Advanced VM & infrastructure management.
  - Specialized VM Image.
- Small energy savings result in a large impact.
- Combining a number of different methods together can have a larger impact then when implemented separately.

Summary - 2

- Combine concepts of both Power-aware and Thermal-aware scheduling to minimize both energy and temperature.
- Integrated server, rack, and cooling strategies.
- Further improve VM Image minimization.
- Designing the next generation of Cloud computing systems to be more efficient.

Thank you!