Cloud Computing

Course Faculty: Dr. Amit Praseed
(amit.p@iiits.in)

Dr. Neha Agrawal
(neha.a@iiits.in)

Course Objectives

- **Introduce the concept of cloud computing and associated technologies to students**
  - grid, fog, edge and cloud computing
  - enabling technologies of cloud computing such as Service Oriented Architecture and Virtualization
- **Cloud computing platforms**
  - AWS, Azure, GCP
- **Parallelization in the Cloud**
  - MapReduce and Hadoop.
- **Cloud Security**

Tentative List of Topics

- **Introduction to Cloud Computing and its Enabling Technologies :** Evolution of Computing, Grid and Utility Computing, The vision of Cloud Computing, Characteristics and Benefits, Distributed Computing, Virtualization, Web 2.0, Service Oriented Architecture

- **Cloud Computing Architecture** : Introduction, Cloud Service Models, Cloud Deployment Models, Open Challenges

- **Virtualization:** Introduction, Characteristics of virtualized environments, Taxonomy of virtualization techniques, Virtualization and cloud computing, Pros and cons of virtualization, Technology examples, Containers and Applications

- **Cloud Platforms:** Amazon Web Services, Google Cloud Platform, Microsoft Azure, Aneka, OpenStack, Cloud Automation using CHEF/Ansible

- **Hadoop and MapReduce:** Hadoop File System, GFS, Introduction to MapReduce, Information Retrieval using MapReduce, Page rank using MapReduce

- **Cloud Security:** Security Issues in Cloud Computing, Hypervisor and VM Security, Data Security in Cloud Environment, Identity and Access Management in Cloud

Tentative Evaluation Policy

- **Examinations: 50%**
  - Mid Semester Exam: 20%
  - End Semester Exam: 30%

- **Research Work / Assignments: 25%**

- **Surprise Quizzes: 10%**

- **Scheduled Quizzes: 15%**

So, let us begin…

Dr. Amit Praseed

Computing

- **Computing** consists of three things:
  - Managing,
  - Processing, and
  - Communicating information
- Over the years, computing has evolved – passing through a number of computational paradigms
  - More computational capacity
  - More efficiency
  - Less cost
  - Less size

Vacuum Tube **Transistors** **Integrated Circuit** **Microprocessor** **Quantum Computer**

1st Generation Computer **2nd Generation Computer** **3rd Generation Computer** **4th Generation Computer** **5th Generation Computer**

The Evolution of Commercial Computing

Centralized Era

first, large and expensive Computers
(since 1940s)

Microprocessors & Mass Production

Personal Computers
(since 1970s)

Miniaturization & Mobility needs

Notebook Computers
(since 1980s)

Timesharing Operating Systems & Local Area Networks

dedicated, on-premise Data Centers
(since 1980s)

Distributed Era

Distributed Computing (with flexible use cases)
(in development since 2018)

HiveNet

Artificial Intelligence

Distributed Computing (with limited use cases)
(in development since 2016)

Blockchain

Decentralized Era

Internet

Service-oriented Architecture

Public Cloud Computing
(since early 2000s)

Public Data Centers
(since late 1990s)

Dr. Amit Praseed

Networked Computing

- Need for networked computing:
  - Better communication between systems
  - Resource sharing
- Systems and associated resources, such as printers, were interconnected to form local networks (LANs)
- These small networks evolved to form ARPANET, which finally evolved to the Internet

Information Explosion

The Digital Universe: 50-fold Growth from the Beginning of 2010 to the End of 2020

(Exabytes)
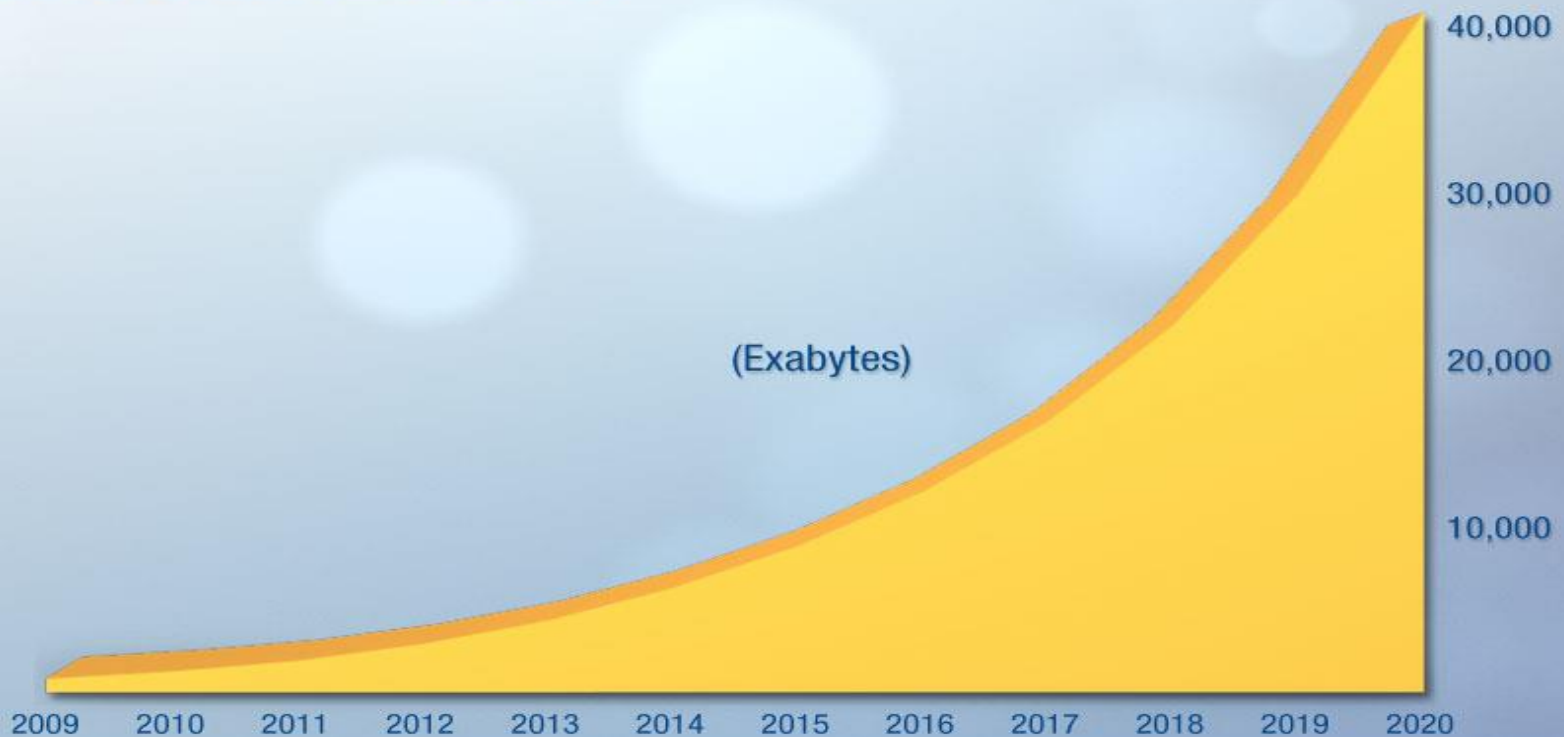
Source: IDC's Digital Universe Study, sponsored by EMC, December 2012

Dr. Amit Praseed

Identifying Mersenne Primes

- **Mersenne Primes :** Prime numbers of the form 2^n-1

- Exponent n needs to be prime

- Used in Elliptic Curve Cryptography (ECC)

- Require enormous amount of computation

The GIMPS Project

- The Great Internet Mersenne Prime Search (**GIMPS**) was launched in 1996

- Allows users to contribute a portion of their unused CPU/GPU for computation

- By the end of 1996, the project could identify the 35th Mersenne Prime ($2^{1,398,269}$-1)

- In 2018, the project identified the 51st known Mersenne Prime ($2^{82,589,933}$-1)

SETI

- SETI – Search for Extra Terrestrial Intelligence

  – Analyzes radio waves in outer space for patterns

- Requires enormous amount of computation

- SETI@home allowed users to dedicate a part of their unused CPU for this purpose

- SETI@home and GIMPS were some of the first instances of volunteer computing projects
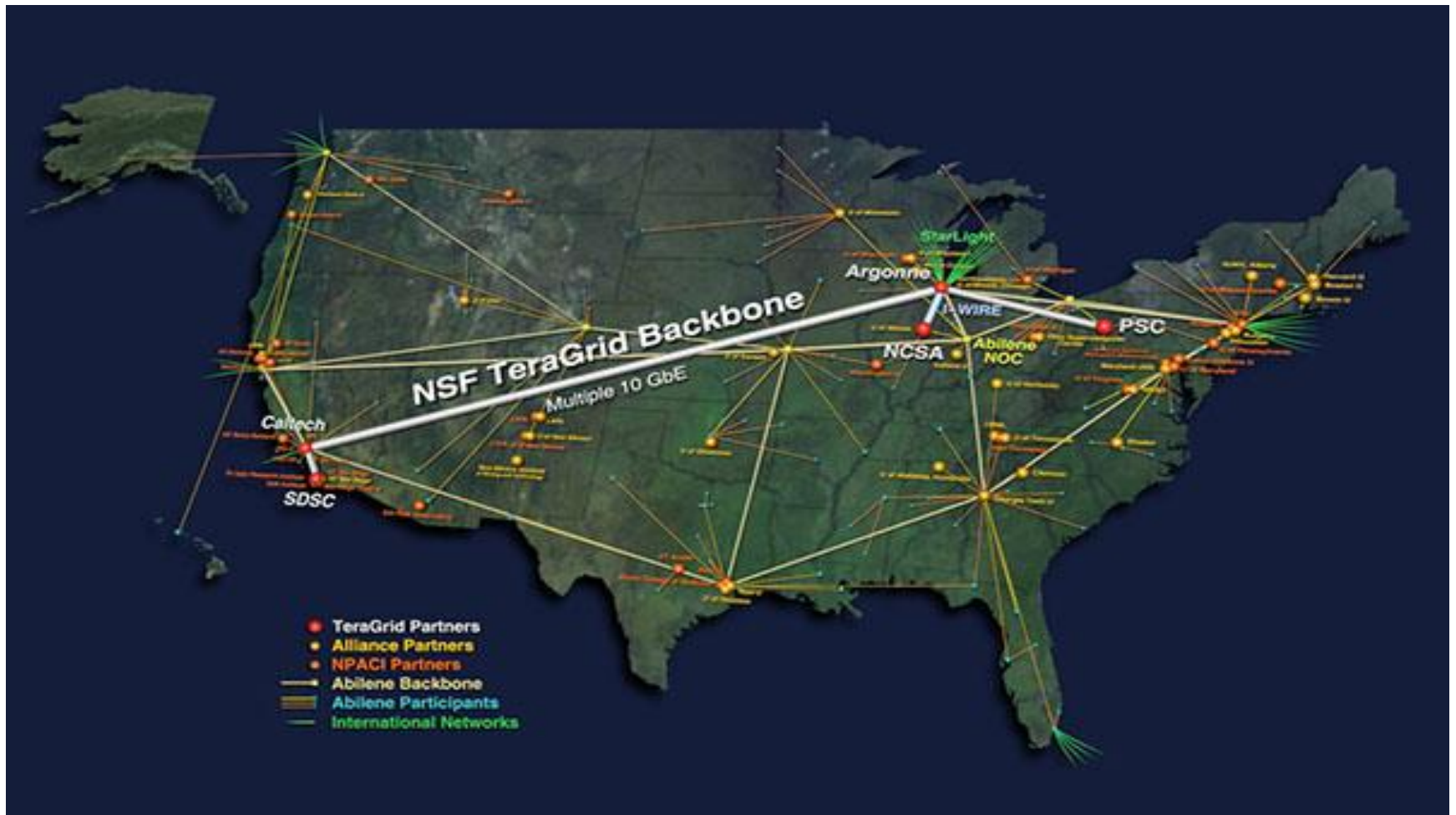
Grid Computing

- The volunteer computing projects are often compared to electric grids
  - Large number of connected nodes which act like a single entity
  - Users do not know which node serves their requests
- This comparison has led to the name "Grid Computing" being used to describe them

Grid Computing

- Multiple physically separated systems share data and resources for **performing a common task**

- Often, the original **task is split and distributed** among different systems or nodes

- Requires **special software** to be installed on the systems ("middleware")

TERAGRID

Dr. Amit Praseed

TERAGRID

- e-Science grid computing **spread over 11 sites** across USA
- **Petaflops of computing capability** and more than **30 petabytes of online and archival data storage** connected by **high speed optic fibre (10 Gbps)**
- Coordinated TeraGrid Software and Services (CTSS).
  - single-sign on
  - remote job submission
  - workflow support
  - distributed accounting and account management software
  - verification and validation software
  - set of compilers, programming tools etc

Advantages of Grid Computing

- Improved resource utilization
- General performance increase – parallel processing
- Easier collaboration
- Increased robustness