☰ Sections

Option C - Web Science (/topic/C-web-science.html)

C.3 Distributed approaches to the web

Different distributed system types and its role for future developments.

> ## C.3.1-2 Define the terms: mobile computing, ubiquitous computing, peer-to-peer network, grid computing. Compare the major features.

Mobile computing

Mobile Computing is a technology that allows transmission of data, voice and video via a computer or any other wireless enabled device without having to be connected to a fixed physical link. The main concept involves,

- Mobile communication
- Mobile hardware: portable laptops, smartphones, tablet Pc's, Personal Digital Assistants
- Mobile software

Characteristics

Portability: The Ability to move a device within a learning environment or to different environments with ease.

Social Interactivity: The ability to share data and collaboration between users.

Context Sensitivity: The ability to gather and respond to real or simulated data unique to a current location, environment, or time.

Connectivity: The ability to be digitally connected for the purpose of communication of data in any environment.

Individual: The ability to use the technology to provide scaffolding on difficult activities and lesson customization for individual learners.

Advantages

- Increase in productivity- as they would be used out in the field of various companies, as it would reduce the time and cost for the client.
- Entertainment- Mobile devices can be used for the entertainment purposes, for personal and even presentations to people and clients.
- Cloud computing- Saving documents on online server and being able to access them anytime and anywhere when you have a connection to the internet.

- Portability- not restricted to one location in order for you to get jobs done or even access email on the go.

Disadvantages

- Quality connectivity- mobile devices will need to either WIFI connectivity or mobile network. such as GPRS, 3G
- Security concerns- Mobile VPNs are unsafe to connect to, and also syncing devices might also lead to security concerns. accessing a WiFi network can also be risky because WPA and WEP security can be bypassed easily.
- Power consumption, due to the use of the batteries

Ubiquitous computing (pervasive computing)

Definition

Ubiquitous computing is the idea of computing being available everywhere and anytime.

- Idea of invisible computing
- Embedded computing (microprocessors)
    - Need for low cost, low power computing with connectivity
- Usually includes a variety of sensors
- Smart designs: different architectures
    - Need for standards and protocols

Peer-to-peer computing

Definition

PCs handling data locally instead of servers(becomes client and server); individual computers connect directly and communicating with each other as equals.

"A peer-to-peer (P2P) network is created when two or more PCs are connected and share resources without going through a separate server computer"

Characteristics:

- Decentralized
    - If one peer falls out not the whole network affected
    - But data recovery of one peer that is shutdown is not possible
        - Requires independent backup
- Each peer acts as client and server
- Resources and contents shared amongst all peers and shared faster than client <-> server
- Has to be done by some software to enable this
- Malware can be faster distributed

Grid computing

Definition

Grid computing is a computer network where each computer shares its resources with all other computers in the system.

Characteristics/Features

- All computers are spread out but connected to each other
- Grid computing develops a 'virtual supercomputer' in a system

Advantages

- Solves larger more complex problems in less time
- Easier collaboration and interaction with other organizations
- Makes efficient use of existing hardware
- Less chances of failure

Disadvantages

- Software and standards still developing
- Non-interactive job submission–> unreliable

C.3.3 Distinguish between interoperability and open standards

Interoperability can be defined as "the ability of two or more systems or components to exchange information and to use the information that has been exchanged". In order for systems to be able to communicate they need to agree on how to proceed and for this reason standards are necessary. A single company could work on different systems that are interoperable through private standards only known to the company itself. However, for real interoperability between different systems open standards become necessary.

Open standards are standards that follow certain open principles. Definitions vary, but the most common principles are:

- public availability
- collaborative development, usually through some organization such as the World Wide Web Consortium (W3C) or the IEEE
- royalty-free
- voluntary adoption

The need for open standards is described well by W3C director and WWW inventor Tim Berners-Lee who said that "the decision to make the Web an open system was necessary for it to be universal. You can't propose that something be a universal space and at the same time keep control of it."

Some examples of open standards include:

- file formats, e.g. HTML, PNG, SVG
- protocols, e.g. IP, TCP
- programming languages, e.g. JavaScript(ECMAScript)

C.3.4 Describe the range of hardware used by distributed systems

This of course depends on the different types of distributed systems, but most generally speaking on a low level multiple CPUs need to be interconnected through some network, while at a higher level processes need to be able to communicate and coordinate. For each approach to distributed system, more specific types of hardware could be used:

- Mobile computing: wearables (e.g. Fitbit ⬀ (https://www.amazon.com/gp/product/B01K9S260E/ref=as_li_tl? ie=UTF8&camp=1789&creative=9325&creativeASIN=B01K9S260E&linkCode=as2&tag=csibrev- 20&linkId=aeffbe18bad4c0b4300d5454f9e5bdf3)), smartphones, tablets, laptops, but also transmitters and other hardware involved in cellular networks
- Ubiquitous computing: embedded devices, IoT devices, mobile computing devices, networking devices
- Peer-to-peer computing: usually PCs, but can include dedicated servers for coordination
- Grid computing: PCs and servers
- Content delivery networks (CDNs) is a system of distributed servers. They can cache content and speed up the delivery of content on a global scale
- Blockchain technology(e.g. Bitcoin (https://bitcoin.org/en/), Ethereum (https://www.ethereum.org/)) are decentralized and based on multiple peers, which can be PCs but also server farms
- Botnets can probably be considered a form of distributed computing as well, consisting of hacked devices, such as routers or PCs

This list is probably not very complete, if you have any further suggestions, please let me know in the comments!

C.3.5 Explain why distributed systems may act as a catalyst to a greater decentralization of the web

Distributed systems consist of many different nodes that interact with each other. For this reason they are decentralized by design, which you can see in this comparison.
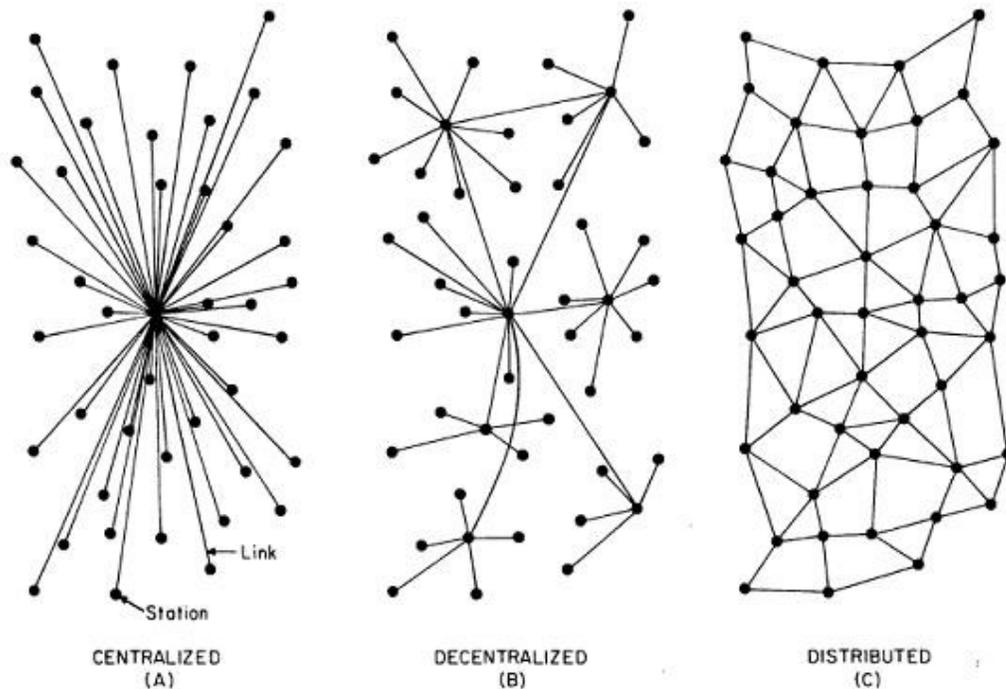
Figure 1: Comparison of centralized, decentralized and distributed networks

Therefore, the importance of distributed systems for a decentralized web lies in their benefits and disadvantages compared to classic centralized client-server models.

Benefits

- higher fault tolerance
- stability
- scalability
- privacy
- data portability is more likely
- independence from large corporations such as Facebook, Google, Apple or Microsoft
- potential for high performance systems

Disadvantages

- more difficult to maintain
- harder to develop and implement
- increased need for security

Personal conclusion

While some decentralized systems such as Bitcoins are gaining traction and some other systems like Git or Bittorrent have been around for a good time already, most part of the internet is still centralized, as most web applications follow the client-server model, which is

further encouraged by corporations wanting to make profit. I found this post from Brewster Kahle's Blog (http://brewster.kahle.org/2015/08/11/locking-the-web-open-a-call-for-a-distributed-web-2/) on the topic very interesting.

C.3.6 Distinguish between lossless and lossy compression

Lossy compression algorithms

Definition: Lossy compression or irreversible compression is the class of data encoding methods that uses inexact approximations and partial data discarding to represent the content. These techniques are used to reduce data size for storage, handling, and transmitting content.

General features:

- Looks for common patterns in data to compress a file –> usually used for multimedia files (images, audio, video)
- Part of original data are lost
- Compresses to really low file sizes
- Usually include settings for the compression quality –> allows for balance between quality and file size
- As data become more compressed, the quality deteriorates –> to certain degree not noticeable by humans

Examples:

- JPEG, GIF
- MP3, MP4, OGG
- H.264, WMV

Lossless compression algorithms

Lossless data compression algorithms usually exploit statistical redundancy to represent data without losing any information, so that the process is reversible.

You need lossless compression when compressing installation files and programs and it can only compress files by 50% of their original file size. But important is that the information/data that is compressed does not affect a loss in information.

Examples:

- Images: BMP(Bitmap), PNG, RAW
- Audio: WAV (Waveform Audio File Format), FLAC (Free Lossless Audio Codec), ALAC (Apple Lossless Audio Codec)
- Graphics: PNG – Portable Network Graphics, TIFF – Tagged Image File Format, WebP – (high-density lossless or lossy compression of RGB and RGBA images)
- 7zip, WinRAR

C.3.7 Evaluate the use of decompression software in the transfer of information

Evaluation of lossy compression

- Significant reduction of file size –> important for file storage, transfer of data over the internet
  - E.g. image files can be reduced to be around 90% smaller before quality degradation is noticeable
- Most important use is streaming multimedia files and VoIP –> bandwidth is usually limited
- However, doesn't work with all file types –> text files or binary data cannot be compressed in a lossy way, as the meaning of the data are lost
- Different things to consider:
  - Compression speed
  - Decompression speed
  - Compression ratio
    - Think about streaming and reducing file size

Evaluation of lossless Compression

- When compressing a file if decompressed will have the same data/Information as the initial file
- Important when compressing an installation file and programs
  - It is required that the installation files and programs' information is the same in the compressing and decompressing phase
  - No loss in quality in lossless compression in Images and Audio files
    - Larger file sizes than lossy compressions

Sources:

- Tech-Faq (http://www.tech-faq.com/lossy-compression.html)