System Architecture

- Topics:
  - Design Philosophy
    - how to go about thinking of architecture?
    - or, what drives the architecture design?
    - some examples
  - Internet architecture
    - history and evolution
    - current design: modular architecture
    - new challanges: security, ...
  - Interplay between theory & architecture design
    - digital communication and information theory
    - parallel programming and von Neumann bridge

Architecture

- Webster interpretation
  - Architecture = *the art or science of building (system)*
- A system design philosophy
  - List desired goals, properties, functional utility of the system in the order of importance
  - Derive design implications of the desired goals/properties/utility of system
  - Search for means/methods/technology for designing appropriate architecture
  - Repeat the above steps until one finds feasible, economical architecture that performs as desired

→ This philosophy leads to "an art of building system"

  - Next, we see some examples

An Example

- Consider the example of Telephone network
- Here is list of desired properties/utilities and their implications
  - Functionality: allow for real-time voice communication
    $\rightarrow$ need a very fast communication technology
  - Privacy: conversation should be private
    $\rightarrow$ communication should be encrypted or should happen over secure links
  - Accountability: resources utilized must be accountable
  - Pricing: cost must be proportional to usage
    $\rightarrow$ basic communication unit should be a *phone call*
  - Economical: system must be affordable
  - Scalability: should be able to accommodate growing demand
  - Robustness: must operate even when few components fail

...

Other Examples

- Engineered systems
  - Airline system
    - Primary purpose: safe, efficient and economical commuting
    - → Implications: regulations on checking of planes, design of optimized routes and schedules, inventory control, human resource, etc.
  - Energy distribution system
    - Primary purpose: efficient, ubiquitous and economical energy distribution
    - → Implications: design of optimized power-grid, set up of energy exchanges or markets, placement of energy plants, regulations on pricing energy resource etc.

Internet

- Next, we will study the design of Internet in detail

- We will talk about
  - A brief history of Internet
  - The utility or goals of Internet
    - initial goals
    - $\longrightarrow$ their implications and current design
  - Current challenges or new goals
    - network security, etc.

Brief History of Internet

- The original Internet project was started in the late 1970s
  - It was a DARPA funded project known as ARPANET
- Motivations behind the ARPANET project
  - To develop an effective architecture for multiplexed utilization of
    - existing interconnected, heterogeneous, independent and possible unreliable networks
  - Specifically, allow for access of ARPANET server
    - to field agent via ARPA radio network
- The Internet or ARPANET
  - Could have been *custom* designed for the above purpose
  - But, fortunately that was not the case
    - allowing for a general purpose network design
- Next, we see goals considered by the architects of ARPANET or Internet

Internet: Original Goals

- In addition to primary goal of ARPANET, the following were the goals considered by the architects to design first generation of Internet
  - Internet communication must continue despite loss of network at gateways
  - Internet must support multiple types of communications services
  - Internet architecture must accommodate a variety of networks operating independently
  - Internet arch. must permit distributed resource management
  - Internet architecture must be cost effective
  - Internet architecture must permit host attachment with low level of effort
  - The resources used in architecture must be accountable
- Reference
  - "The Design Philosophy of DARPA Internet Protocols," D. Clark.

Design Implications

- Next, we study the design implications of these goals
- The most important design goal is to allow
  - Multiplexing between hetergeneous independent networks
    - → packet switching architecture, where data is communicated into independent units called data-grams
  - Subsequently, networks must be interconnected via packet switches or gateways.
- In summary, structure of Internet should be such that
  - Packet-switched commication between heterogeneous networks
  - Different networks are connected via packet communication processors, called gateways
    - – which implement store and forward packet routing algorithms
- Now, the other goals

Design Implications

- Goal: survivability against failures
  - In the event of failure, the ongoing communications should resume from the point of failure after "re-configuration"

  $\rightarrow$ Maintain "state information" that should not be lost
- Implication:
  - Network is heterogeneous

  $\rightarrow$ Maintain "state information" at the end host that are communicating
    - that is, a state-free network
- An example of state maintained as end-hosts
  - Connection information and packet numbers

Design Implications

- Goal: support multiple types of services
  - ∘ Different services require different level of reliability, delay, jitter, etc.
    - – for example, voice data versus file transfer
- Implication:
  - – To support multiple services, need multi-levels of transport protocols
  - → architecture must support simultaneous multiple transport protocols
- Current example of transport protocols
  - ∘ TCP: reliable transport
  - ∘ UDP: low-delay transport

Design Implications

- Goal: design should work for heterogeneous networks
    - ◦ Allow for incorporation and utilization of variety of networks
    - ◦ For example, low-speed dialup versus high-speed cable network
- ◦ Implication:
    - ◦ Architecture must make minimal set of assumptions about underlying network
    - ◦ Architecture must allow for simple interface between underlying physical network and higher-level functionality
- ◦ Current example
    - ◦ All networks should provide interface for packet- or datagram-level communication
    - ◦ Network protocols should be utilizing only such interfaces
    - → Network protocol design is independent of underlying network

Design Implications

- Other goals
  - Distributed management
    - → network protocols must utilize only "local" information
  - Efficient utilization
    - → design overhead must be minimal
    - → protocols must provide high-performance
  - Economical
    - → network should be buildable with "cheap" technology
    - → network should be available to wide range of society
  - Accountability
    - → resource utilization should be accountable

Design Implications: Summary

- In summary, the architecture should be
  - ◦ Packet switched, i.e.
    - – packet or datagram is building block of communication
  - ◦ End-hosts should maintain connection information,
  - ◦ Different types of network should support packet level transmission and provide universal interface to other network protocols
  - ◦ Provision of multi-level of transports
  - ◦ Heterogeneous networks interconnected via packet processors or gateways
  - ◦ Protocols must utilize only local information, such as
    - – routing based on local topological and addressing information

→ This naturally leads to layered Internet architecture

Layered Architecture

- Basic principle of layered architecture
  - ◦ Overall function is divided into layers of "independent" modules
  - ◦ Different modules interact via pre-decided interface (inputs & outputs)
  - ◦ Details of one module is of no interest to the other module
    - – a module has effect on other only via interface
- Concept of layered architecture or modularity is as old as "engineering"
  - ◦ In the context of algorithm design
    - – known as *divide and conquer*
  - ◦ In the context of optimization
    - – known as *dual decomposition*

  $\longrightarrow$ We'll use these for architecture design
- Next, we describe layered Internet architecture
  - ◦ Reference: Chapter 1.3 of Data Networks, by Bertsekas-Gallager

Layered Architecture

Application layer	Applications are run here
Presentation layer	Encryption, compression, conversion
Session layer	Allows for establishing connections between end hosts
Transport layer	Transporting data between end-hosts and possibly converting unreliable network layer into reliable transport
Network layer	Determines routing of data packets, appends necessary information before giving to data layer
Data link layer	Convert unreliable physical layer into reliable higher level data link
Physical layer	Provide link for bit transmission

Layered Architecture: An Example

Application layer	Secure file transfer (SFTP)
Presentation layer	Encryption (SSL)
Session layer	Establish session between end hosts
Transport layer	Reliable transmission via flow contol (TCP)
Network layer	Appropriate header information is appended (IP)
Data link layer	Reliable data link (CDMA)
Physical layer	Transmission over physical link (Ethernet)

Layered Architecture

- General guidelines for modular design
  - More modules give independence of design
    - → better for supporting multiple types of services, robustness, interoperability, etc.
  - However, more modules may impose restriction on functionality of architecture
    - → a natural trade-off between functionality and design flexibility arises
- It is important to carefully select modules for desired flexibility and performance
  - We'll see how theory can help in choosing appropriate module
    - in the context of communication theory
    - in the context of parallel computation

Classification of Internet Protocols

- The essential Internet protocols are of two types
  1. Control protocols
     - related to what operation should be performed
     - e.g. how should data be routed
  2. Data protocols
     - related to actual data transfer
- The performance of Internet is affected mainly by control tasks
  - But, actual “work” is done by data tasks
- Treating them separately leads to better design
- This distinction is popularly called as
  - Control plane of Internet
  - Data plane of Internet
- We’ll see this distinction more closely through the course

New Architectural Challenges

- As we've seen, architecture is the result of desired goals

  → Newer goals will lead to changes

- Some new critical challenges
  - Quality-of-service (QoS)

  → Refined feedback from network to end hosts
    - we'll see relation between feedback and the task of interest
    - control theory helps !
  - Network security
    - protection of Internet resources from malicious users
    - e.g. viruses and worms, denial-of-service attack, etc.
    - see a recent article in MIT Tech Review, "The Internet is Broken," by D. Talbot, Dec 05/Jan 06 issue, page 62–69.

Homework I

- Choose your favorite (complex) engineering system, different from the one covered in the class. Some examples are
    - Network of health care system
    - Network of railway system for Europe
    - Network of Postal system
    - Communication system for poor rural country
- Given this chosen system, carry out the following steps in detail
    - List the desired performance goals of the system
    - Derive implications of the goals on the system architecture
    - Sketch the details of system architecture using layering, etc.

References

(1) Two papers

- "The Design Philosophy of The DARPA Internet Protocols," by D. Clark
- "Architectural Considerations for a New Generation of Protocols,, by D. Clark and D. L. Tennenhouse

(2) Data Networks, by Bertsekas–Gallager

- Chapter 1.3

(3) Expository article

- MIT Tech Review, Dec 05/Jan 06 issue, article, p. 62–69