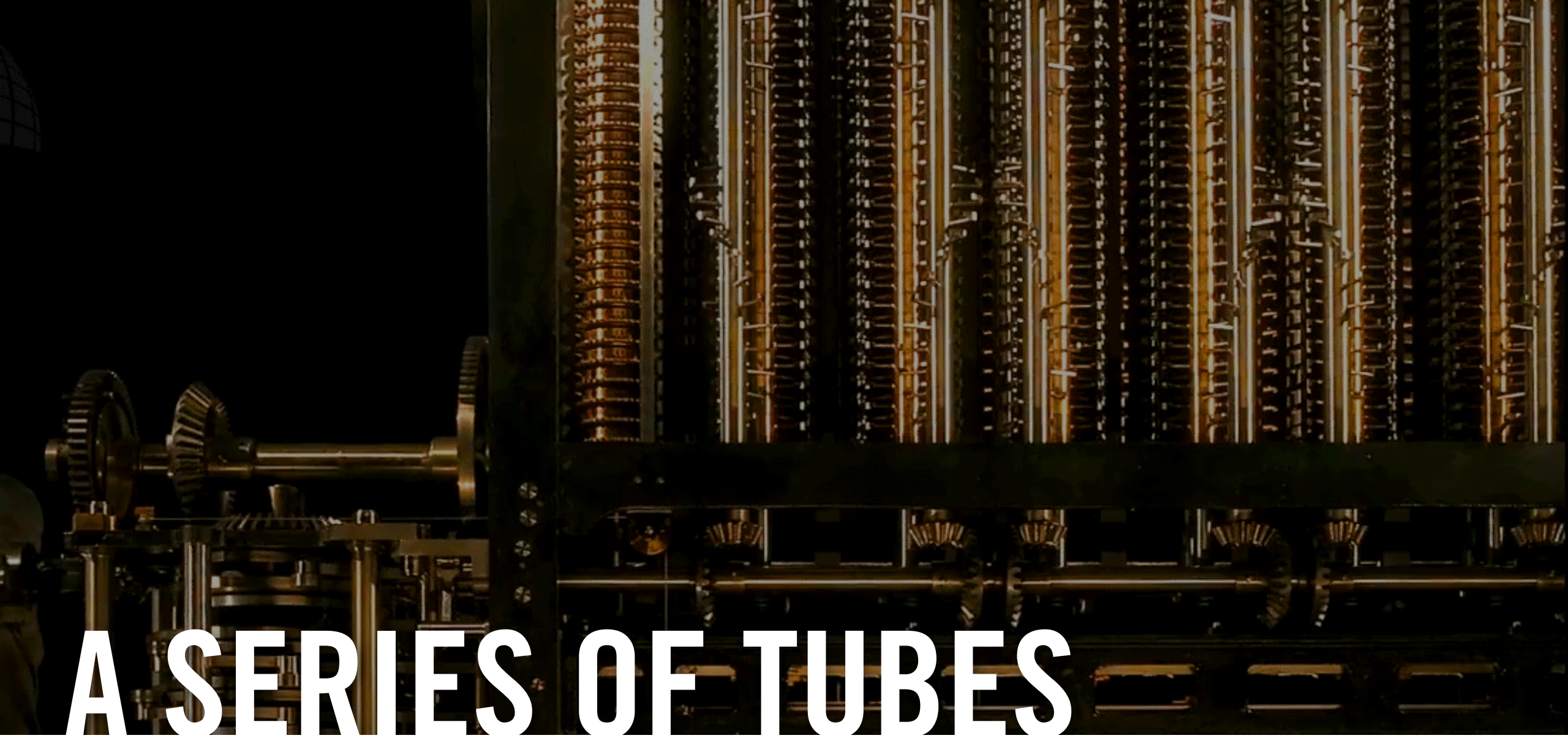


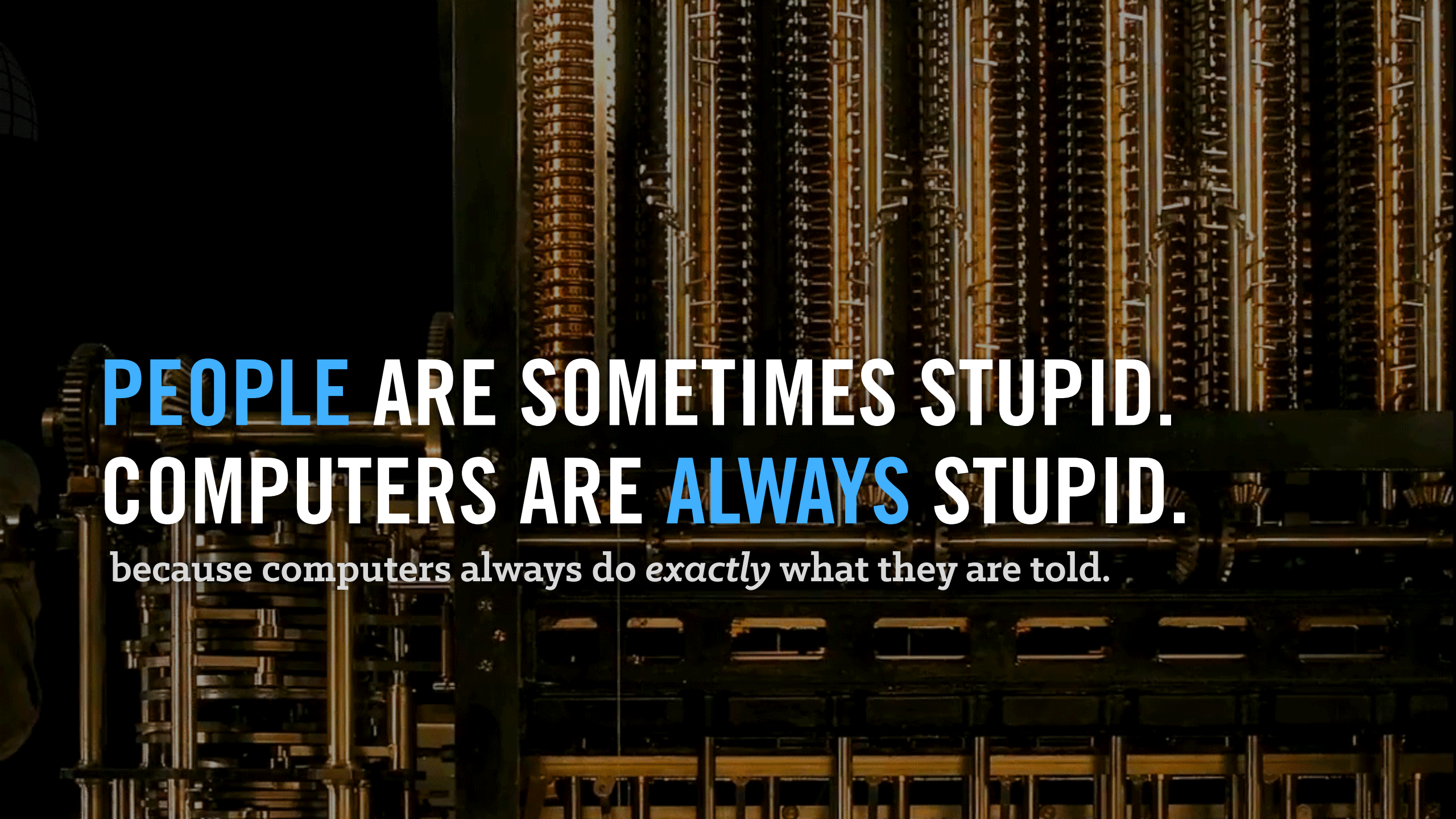
A SERIES OF TUBES

Prof. Wright | COMD2451

A SERIES OF TUBES?

...the Internet is not something that you just dump something on. It's not a big truck. **It's a series of tubes.**

~Ted Stevens, US Senator (really)


PEOPLE ARE SOMETIMES STUPID.
COMPUTERS ARE ALWAYS STUPID.
because computers always do *exactly* what they are told.

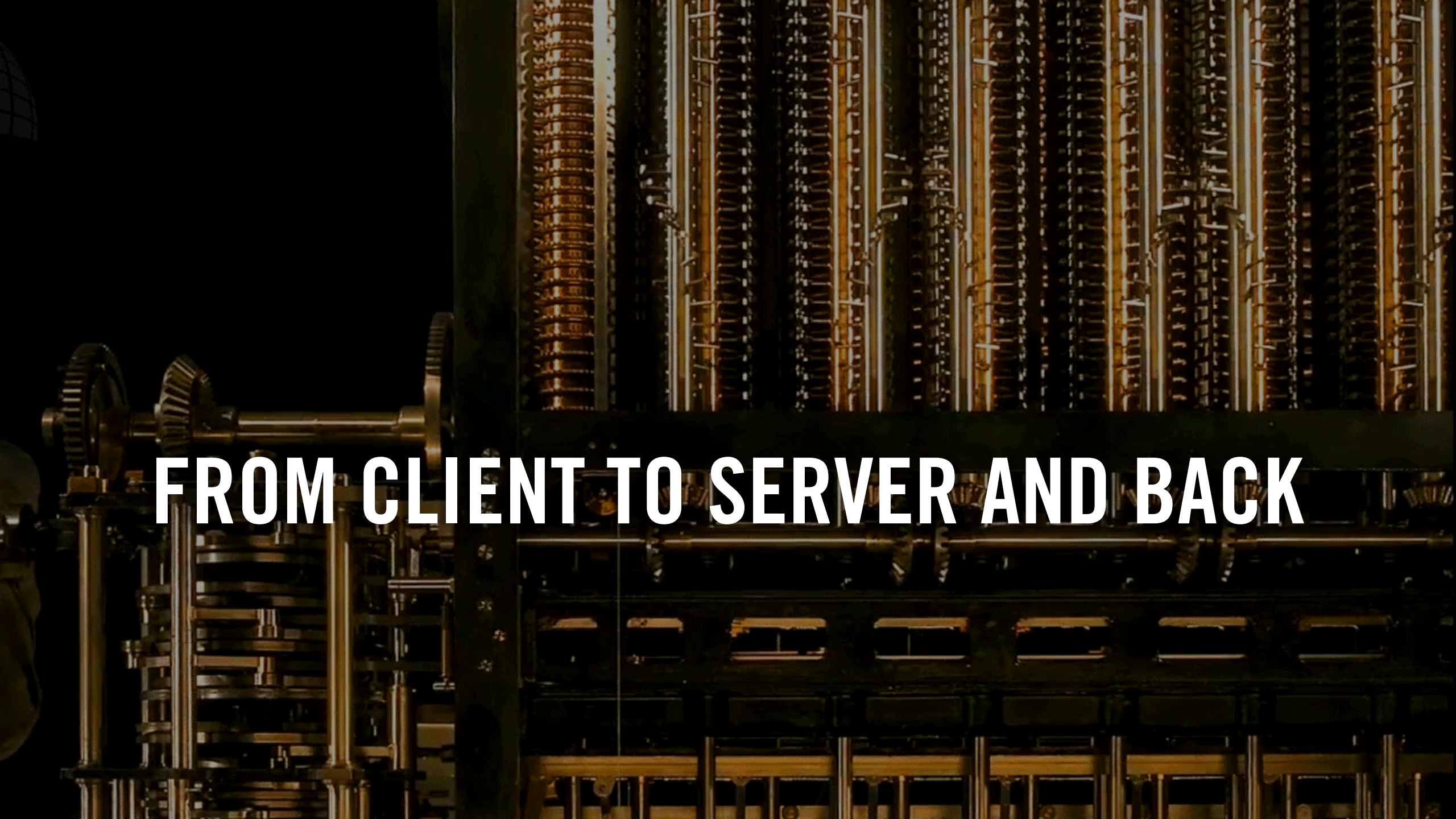
HOW COMPUTERS THINK

- computers process information using Boolean (binary) logic
  - in binary logic, the smallest pieces of data have two possible values: true or false
  - this is represented in binary code, which handles data as strings of ones and zeros
- binary code is *much* older than computers
  - 1679 – Gottfried Leibniz develops the modern binary number system, inspired by the Chinese I Ching
  - other binary number systems existed in ancient Egypt, India, and China
  - 1847 – mathematician George Boole invents Boolean algebra
    - yes-no, on-off approach with three basic operations: AND, OR, and NOT
  - 1930s – Claude Shannon, a grad student at MIT, noticed similarities between Boolean algebra and electric circuits
  - Shannon's thesis led to the use of binary in computing and electronics

DATA UNITS

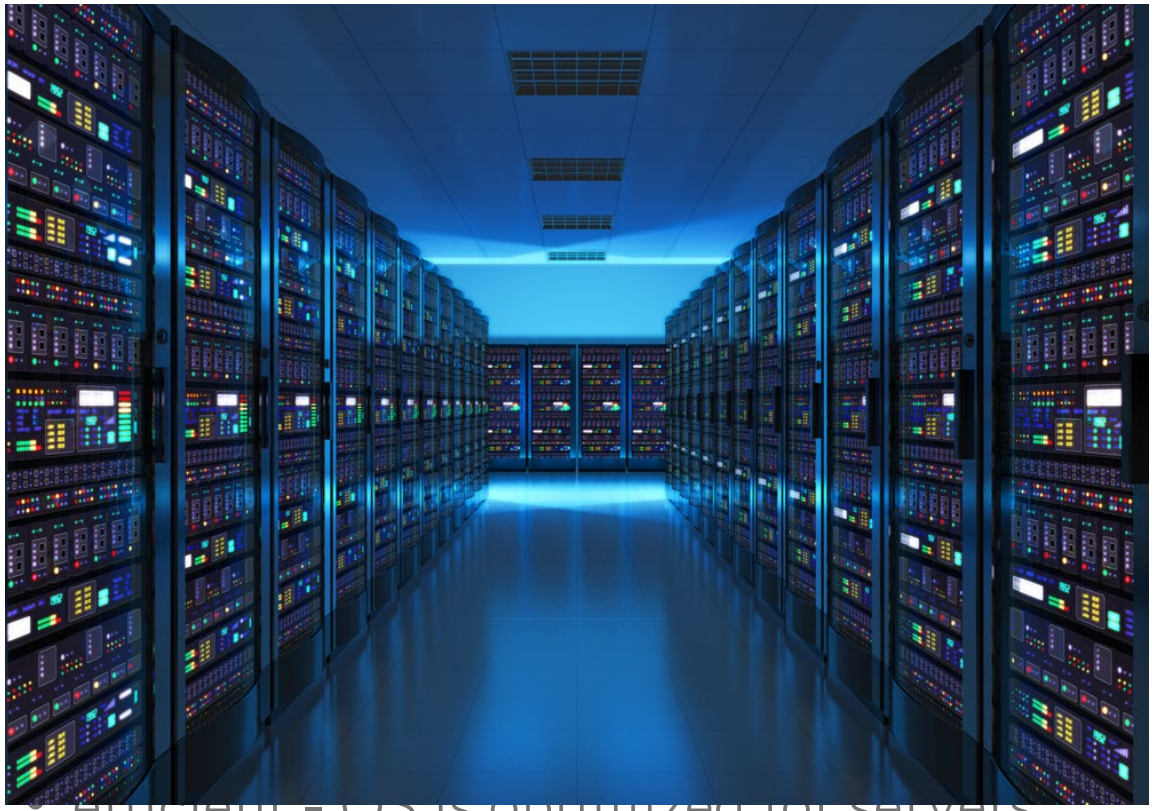
- bit - smallest unit of data possible
  - a bit is a single "0" or "1"
- kilobit - 1000 bits
  - or 125 bytes
- megabit - 1000 kilobits
  - or 125 kilobytes
- gigabit - 1000 megabits
  - or 125 megabytes
- terabit - 1000 gigabits
  - or 125 gigabytes

- byte - a string of 8 bits
  - can express 256 different values
- kilobyte - 1000 bytes
  - or 8000 bits
- megabyte - 1000 kilobytes
  - or 8000 kilobits
- gigabyte - 1000 megabytes
  - or 8000 megabits
- terabyte - 1000 gigabytes
  - or 8000 gigabits


FROM CLIENT TO SERVER AND BACK

SERVERS

- servers are computers configured to *serve* requested data to client
- ALL web data is hosted on servers
  - an individual HTML document might load or link to data across multiple servers
- any computer *can* be a server
  - even a $40 Raspberry Pi
  - but if a server has to handle more than a handful of users, it's usually built for the job



- efficient - OS is optimized for servers with consumer/desktop functionality switched back (usually Linux or Windows Server)

IP ADDRESSES

- IP Address - a unique string of numbers, separated by periods, that identifies each computer using the Internet Protocol to communicate over a network
  - all devices connected to the Internet are assigned an IP address
  - usually automatic through DHCP (Dynamic Host Communications Protocol)
- you can access a site with only an IP Address; URLs are not necessary!
  - URL - Uniform Resource Locator, aka web address
  - http://104.162.103.77 - copy of classwright.net hosted on my home server
  - no URL assigned; content on this server is *only* accessible via IP Address
- ...but numeric IP addresses aren't practical for navigating the web
  - plus popular sites maintain duplicate IP addresses to balance traffic

THE DOMAIN NAME SYSTEM

- Domain Name System (DNS) translates URLs to assigned IP addresses
  - when you enter a URL into a browser, a DNS server resolves the URL into the IP address for the assigned web server
  - meaning: you only need to remember "amazon.com," not 54.239.25.208
- IP Address : satellite coordinates :: URL : street address
  - coordinates are easier for computers to process, while humans have an easier time understanding and navigating street addresses
  - but both put you in exactly the same place

DOMAIN NAMES

- a domain name consists of one or more labels, separated by periods:

developer.mozilla.org

- each label is a subdomain of the label that comes after it
- the last (right-most) label is the top-level domain (TLD)
  - TLDs are often (inaccurately) referred to as domain extensions
  - in this case the top-level domain is "org"
  - "mozilla" is a subdomain of "org"
  - "developer" is a subdomain of "mozilla"

CHOOSING A DOMAIN NAME

- **there are no technical advantages to any top-level domain**
  - e.g. "com" is not faster than "net" or "org"
  - but unusual TLDs can confuse users...
  - ...and many inexpensive TLDs are used extensively by spammers
- **com is the best option, but not always available/affordable**
  - best alternatives: net, org, io (for programmers and developers)
  - also consider country-code TLDs (ccTLDs): us, uk, de, etc
- **generally avoid:**
  - biz, info (very popular with spammers)
  - anything long and/or unusual (club, guru, photography) - most people are used to a small handful of short extensions, and might have trouble remembering these