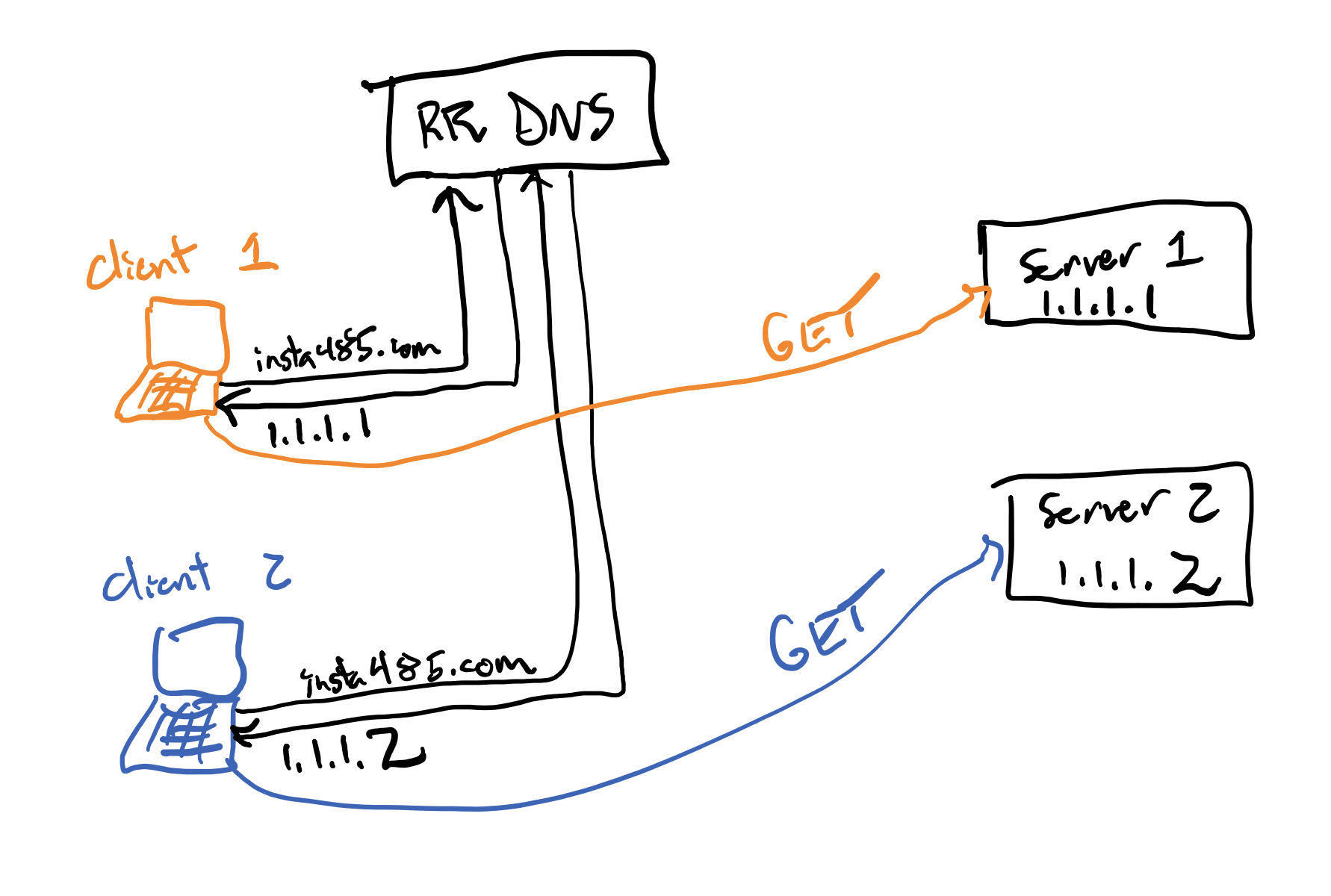
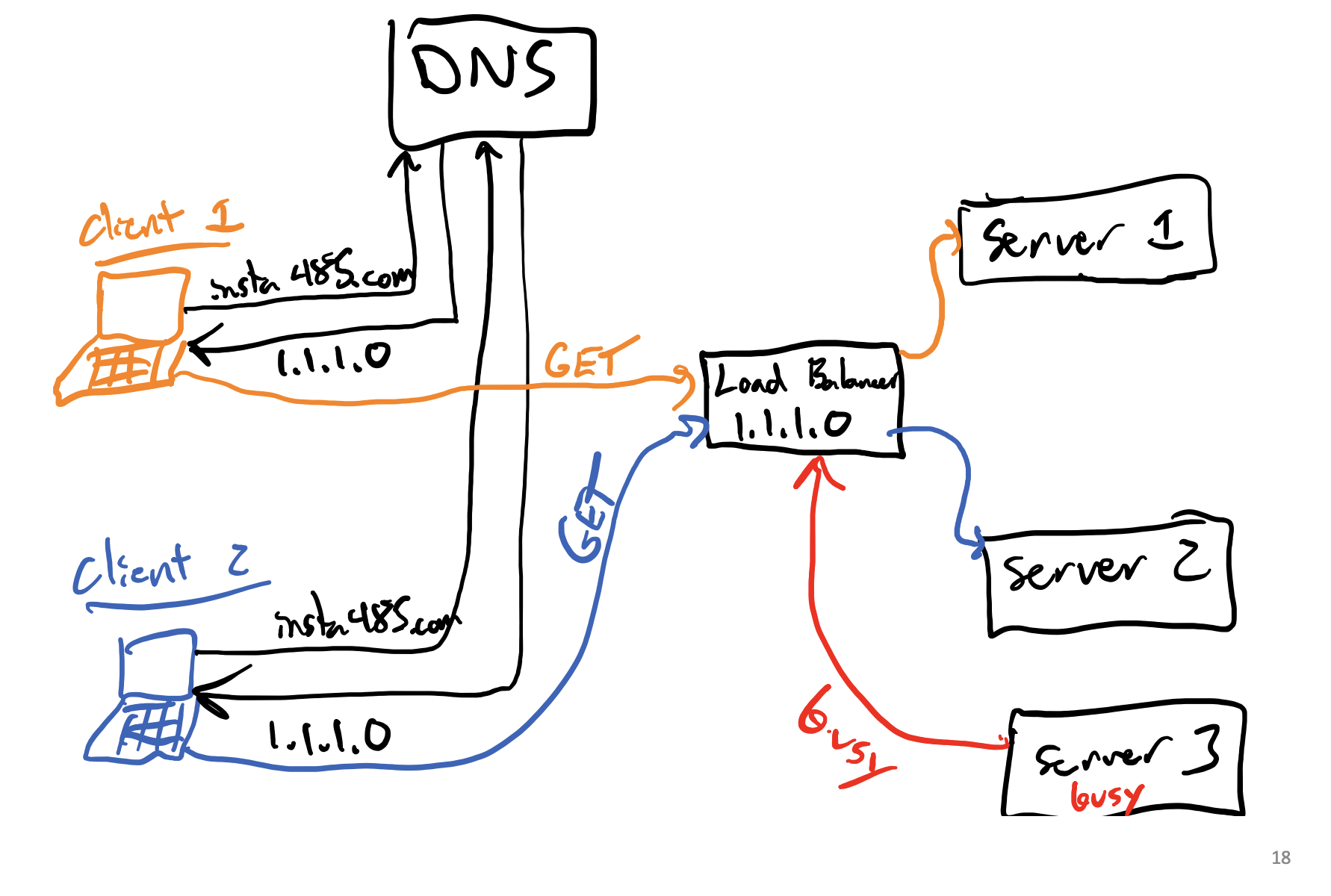
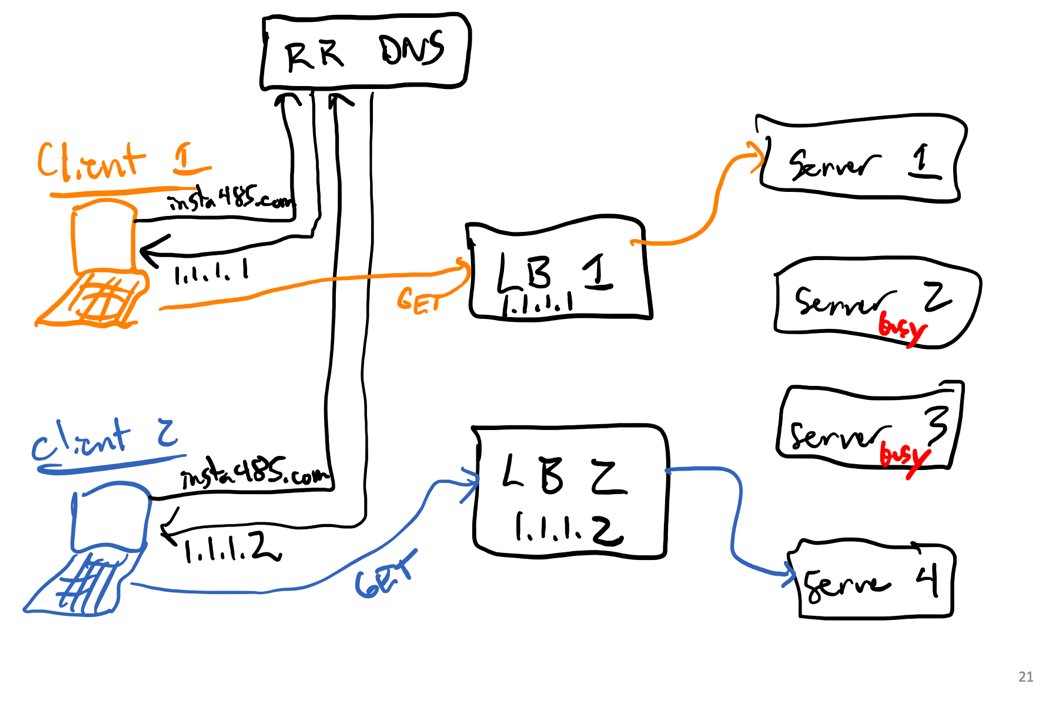
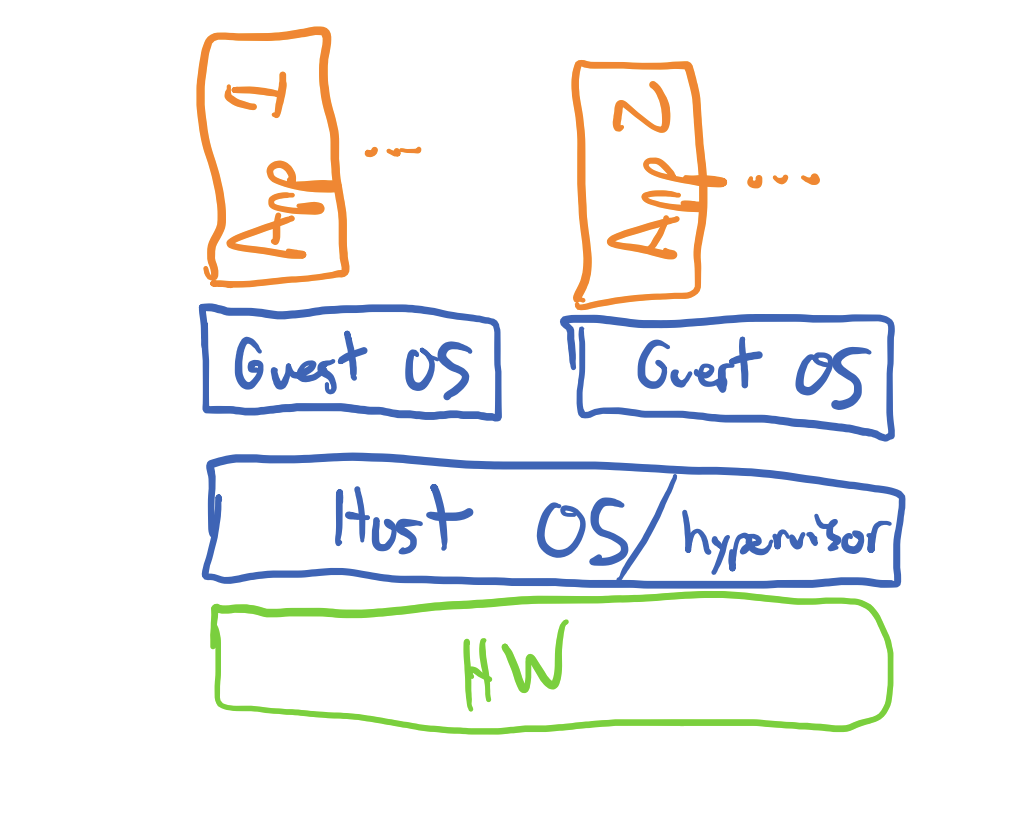
EECS 485 L19 Scaling Dynamic Pages

* Problem: Simultaenous requests go one at a time
  + Illusion of parallelism
    - Sol1: Async
      * Still runs on one CPU core
    - Sol2 Multithreading
      * Still one cpu threads take turns
    - Sol 3: Multiple processes on one CPU core
      * zStill runs on 1 scpu core
* Many servers
  + Prob: One multicore server not enough
  + Sol: buy more servers
    - EECS 281 AG uses 1 computer
    - Rest use more servers
  + • Problem: Two users make a request. Which server should respond? • Option 1: Round robin DNS
* DNS -> what’s ip adderss
* Sol: Round Robin DNS: Modified DNS server responds with different ips to different pages for web servers
  + Have 2 servers
  + 1st clent gets 1st server (ip)
  + 2nd gets second
  + 3rd gets first
    - How we do it in the project
  + Usually in same Datacenter



* + •Problem: round robin DNS doesn’t consider server load. One request might be more time consuming than another
  + Solution: Load Balancer
* Load Balancer
  + Instead different clients gert diffetnet ips
  + Differetn clients et same IP of load balancer
  + Sends to servers thata ren’t busy
    - 
  + Problem: Load balancercan get overwhelmed
    - Single point of failure
  + Solution Load balancer with Round Robin DNS
* Load balancer with Round Robin DNS
  + Multiple load balancers
  + Pools of servers managed by load balancers



* Hardware Virtualization
  + Datacenter Problems
    - IAAS providers use many different types of comps
      * P1: Wasteful to run same ssd page program on each server
      * Diifficlt to customize for each
      * Solution: Agreggate them and run on same computer ???
    - Isolation:
      * Provider has many customers
      * Different requirements for each
      * When working on same computer for 2 customers need isolation
      * Sol: Need to make virtual comp type stuff but still runon same computer
    - Migration
      * Many users visit site – sometimes few
      * Sometimes machines fail
        + Need to add more machines when load is high
        + ?? look more into
  + Problems
    - Wokrload Aggreagtion
    - Workload Isolation
    - Worklaod Migration
      * Solution to all – Hardware Virtualization
  + Hardware Virtualization: One computer runs multiple OS as programs
    - Green 1 physical comp
    - Multiple Blue os
      * Make physical computers essentially a program and run programs on same machine
      * 
    - Definitions
      * host: phyyiscal comp running os  
        guest os being run as programs
      * hypervisor: virtualization software runs guests on host
      * Hardware Emulation: How Host shares info like network/disk
  + **Hardware Virtualization Advantages – DIDN’T TELL US TO BUT PROB GOOD ON CHEAT SHEET**
    - Energy efficiency – Run each physical machine at max efficnecy
    - Diverse environments – Each service gets own OS, with libraries …
    - Security and Isolation
      * Compromised Guest OS won’t affect other guest VMs
    - Scaling
      * Add more VMs as more users hit your site
  + Hardware virtualization disadvantages
    - • Memory •
      * Guest includes a complete OS including drivers, binaries, and libraries •
    - Slow start •
      * 10s ~ 100s to boot a VM
  + AWS Elastic Compute Cloud (EC2) ex VM
* Containerization