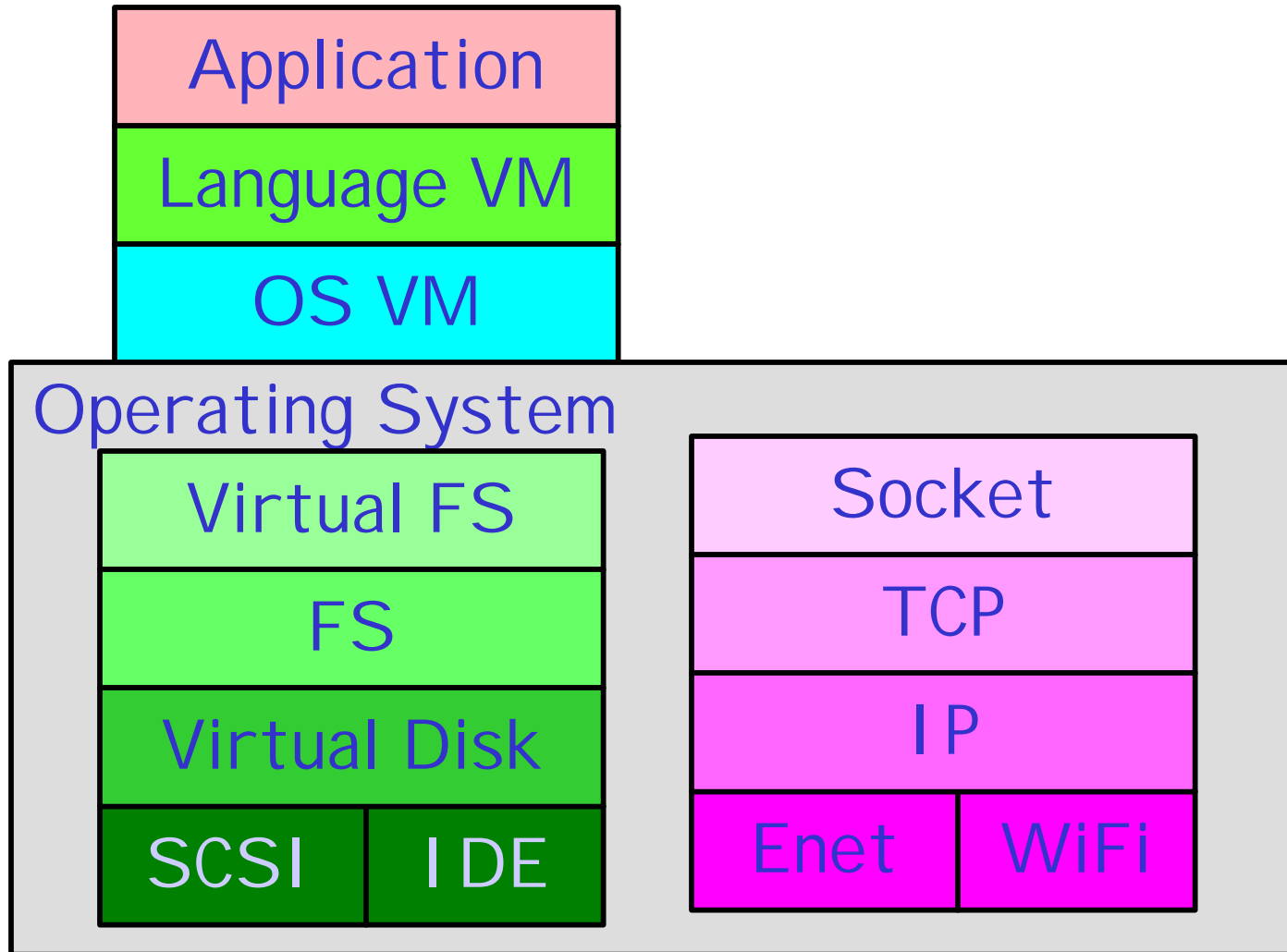

*Impact of Layering and Faults on Availability
and its End-to-End Implications*

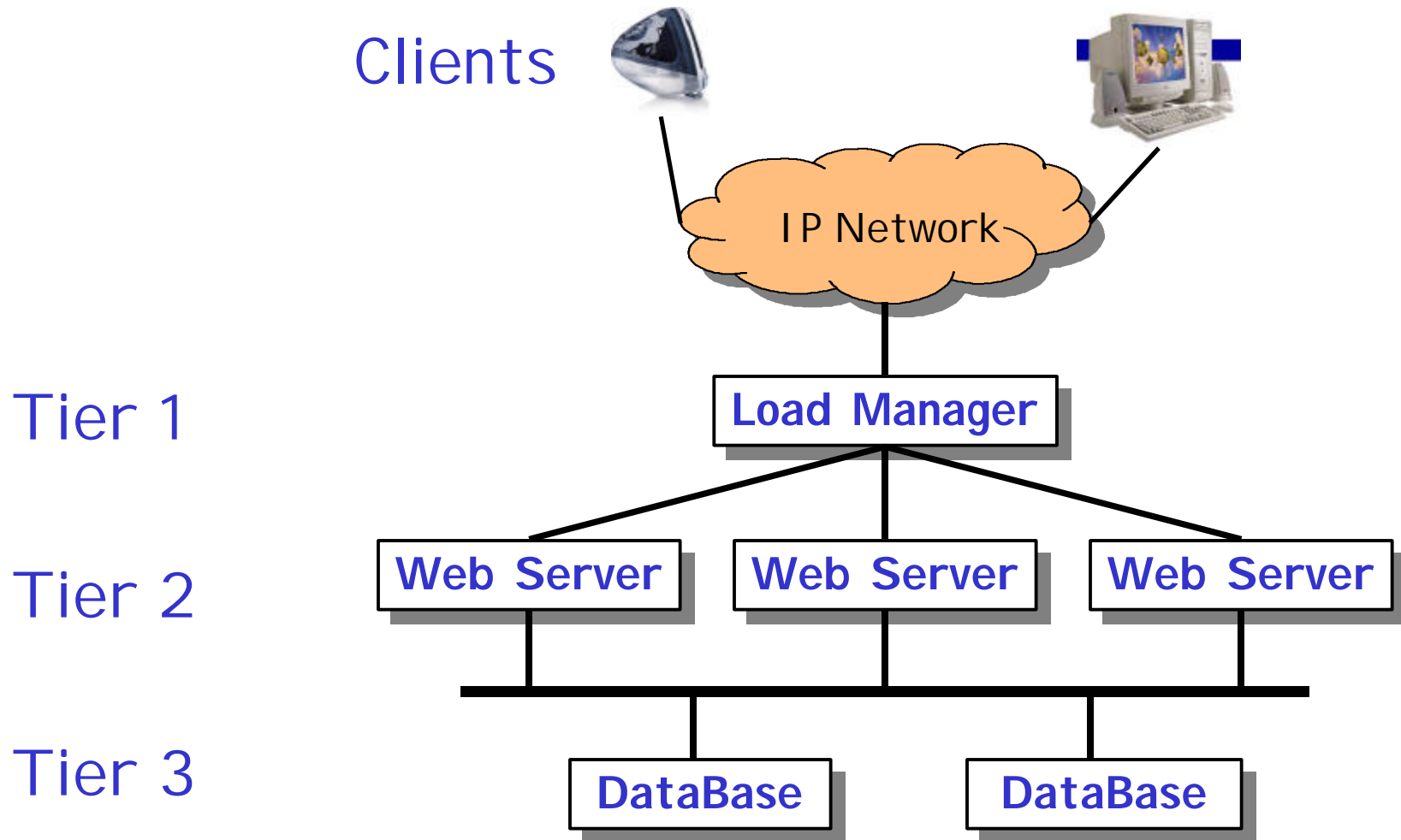
Ricardo Bianchini, Richard Martin, Thu Nguyen

Department of Computer Science
Rutgers University

Layers on Layers on Layers



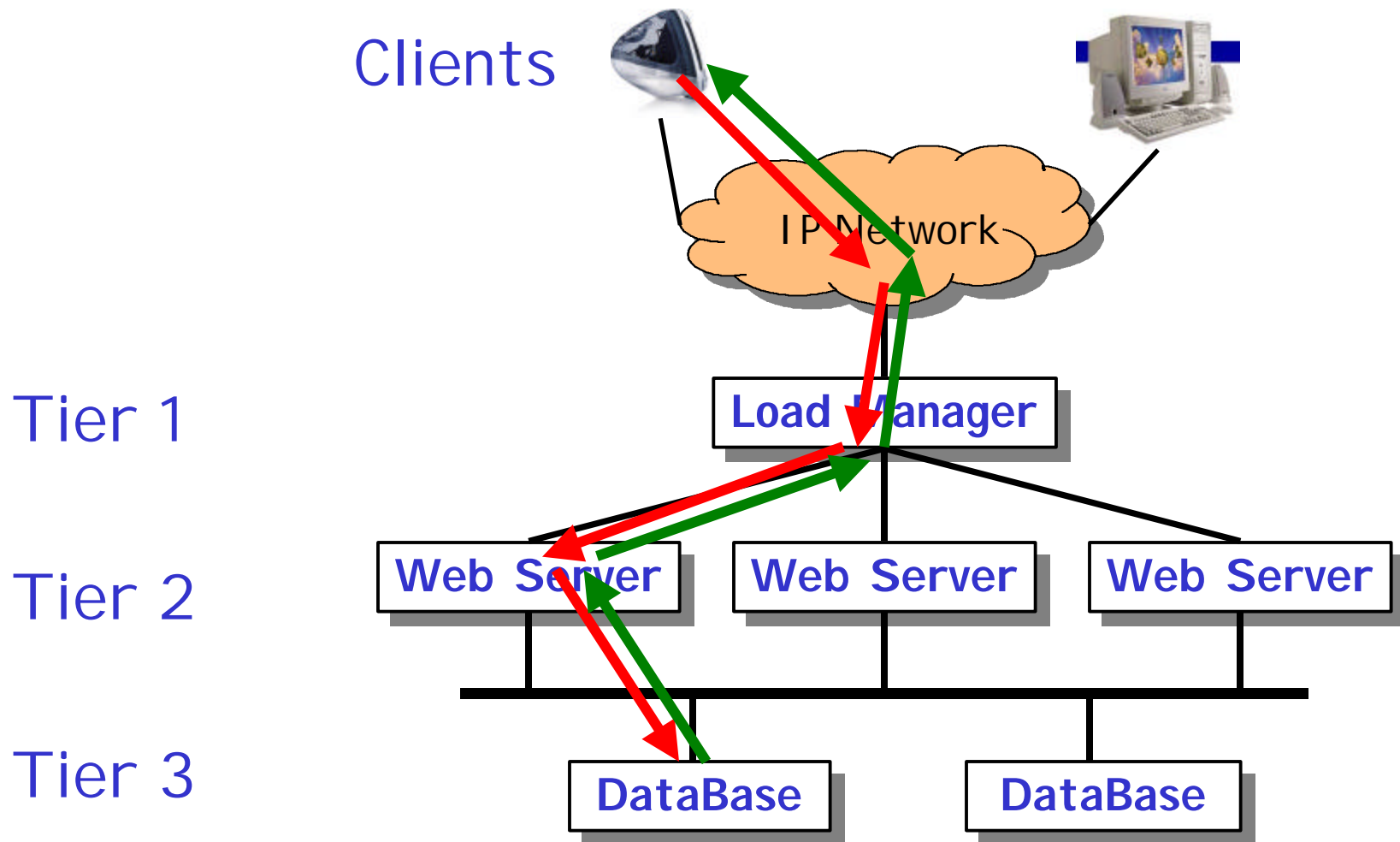
Global Layering



Availability

- ❖ Correct response within accepted time bound
 - ❖ Fraction correct/90th percentile response time
- ❖ Two components:
 - ❖ Correctness
 - ❖ Time bound
- ❖ To get to 99.999 nines (5 min unavailability/year)
 - ❖ Not a lot of time to mess around

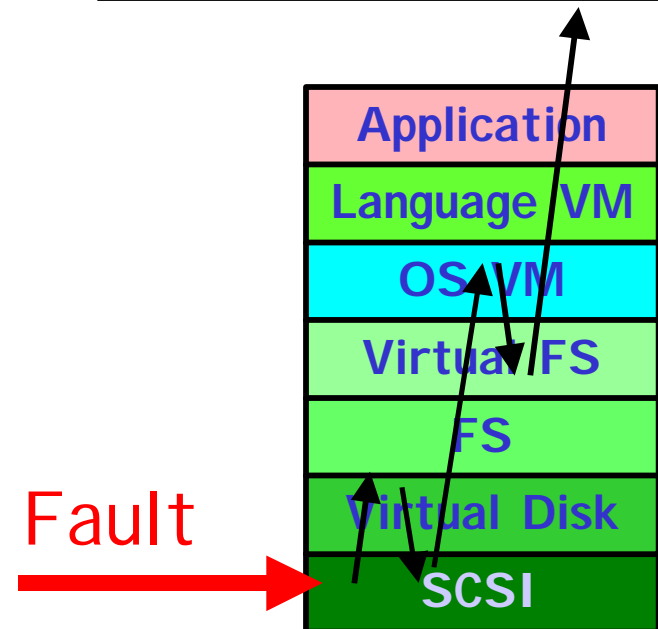
2 Second Reponse time



Impact of Faults and Layers

- ❖ Each layer built independently
- ❖ Intermediate layers hide exceptional conditions
 - ❖ Buffer and continue
 - ❖ Retry N
 - ❖ Crashes
 - ❖ Punt
- ❖ Hard to build available systems
 - ❖ React quickly enough?
 - ❖ Diagnosis difficult
 - ❖ Prediction

```
try {  
    do_stuff();  
} catch(e) {  
    freak_out();  
}
```



End-to-End Implications

- ❖ **Traditional Philosophy**
 - ❖ End-to-end checks necessary for correctness,
 - ❖ Intermediate checks only a performance enhancement
- ❖ **Availability oriented:**
 - ❖ End-to-end checks necessary for correctness
 - ❖ Intermediate checks and timely cross-layer propagation of information necessary
- ❖ **How to achieve these properties while maintaining layering?**