



ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ  
UNIVERSITY OF CRETE

# HY590.45

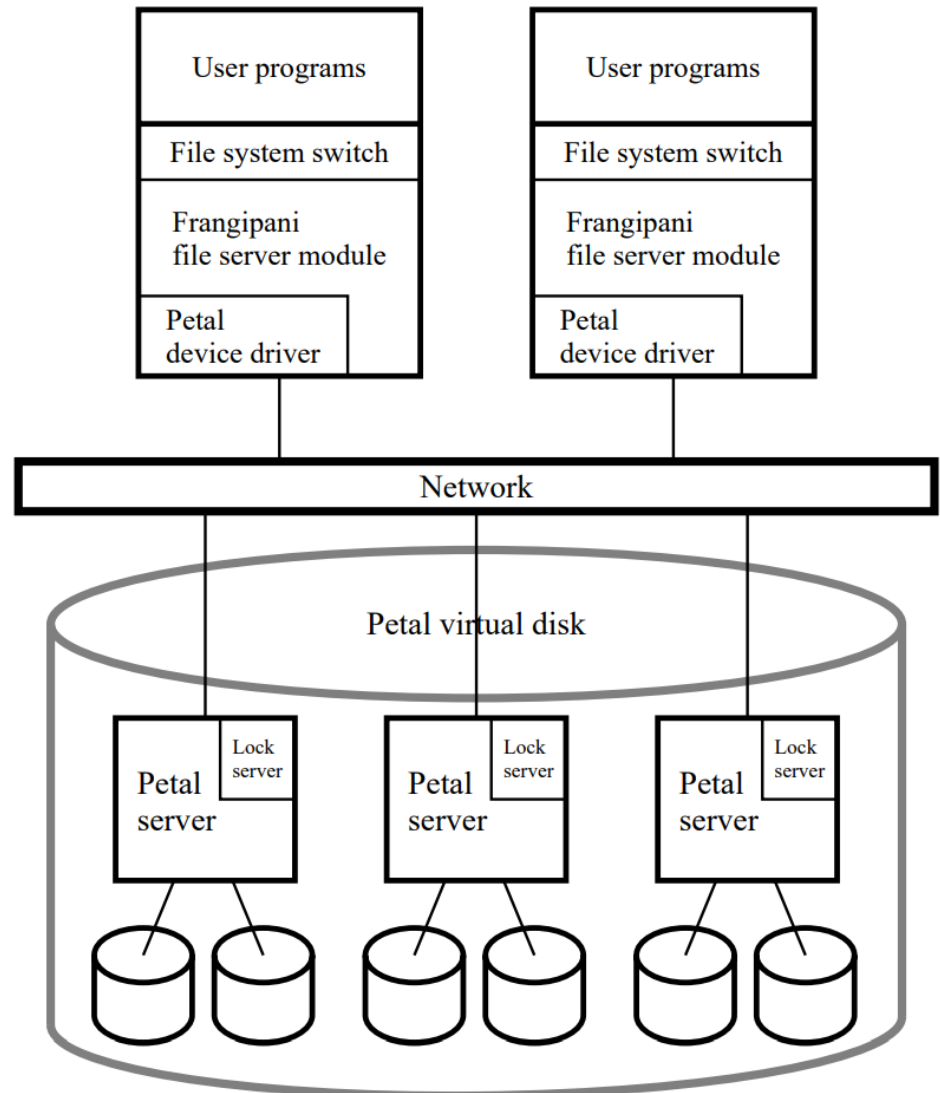
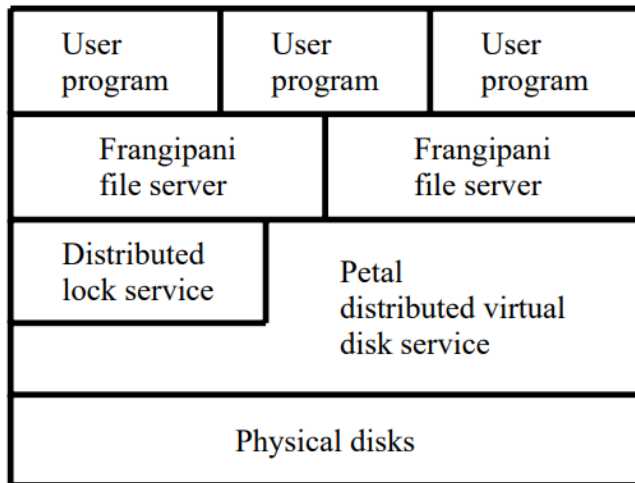
## Modern Topics in Scalable Storage Systems

Kostas Magoutis

magoutis@csd.uoc.gr

<http://www.csd.uoc.gr/~hy590-45>

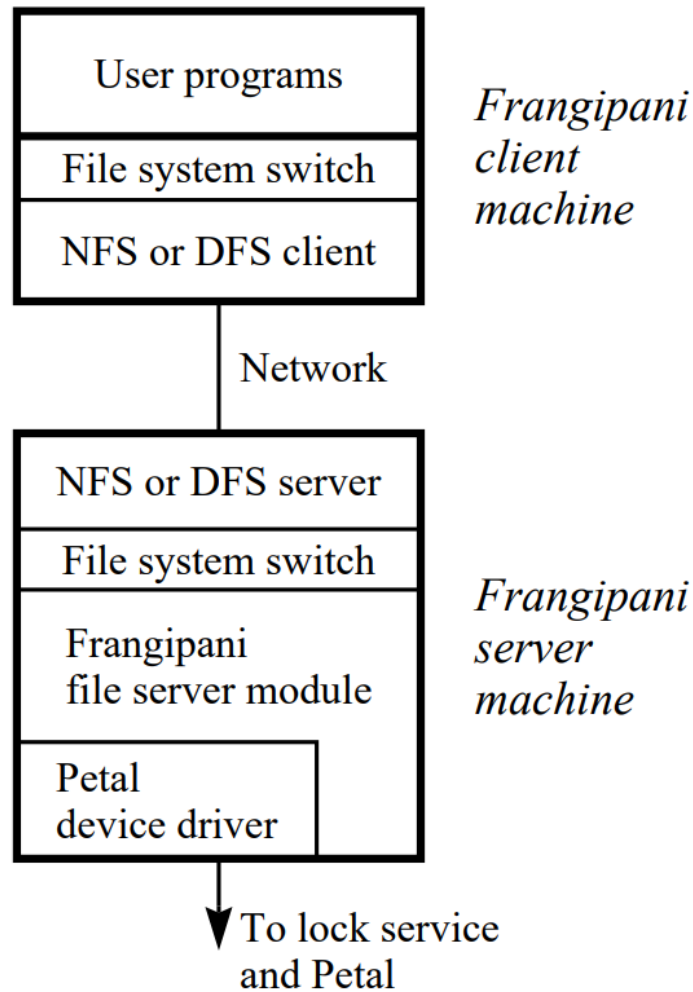
# Frangipani: Layering and structure



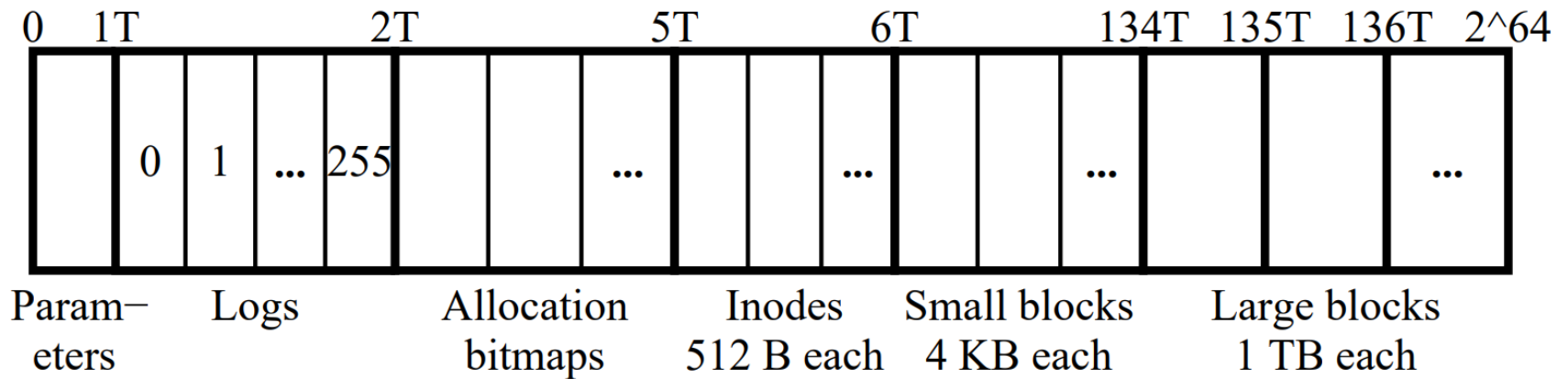
# Goals

- Coherent, shared access to files across many users
- Ability to scalably add more servers, disk capacity
- Add new users without worrying about management
- Full, online, consistent backup of entire file systems
- Ability to tolerate machine, network, disk failures

# Client/server configuration



# Disk layout



- Example: Lookup file `/dir/file1`
- Example: Create file `/dir/file2`

# Metadata logging

- Each metadata update is recorded in the log
- Logs are periodically written out to Petal
- In-place update performed by updated, every 30sec
- Logs are 128KB
- Recovery is run on a log after failure is detected

# Locking

- Multiple readers/single writer
  - Sticky locks
  - Using leases (30")
- Four basic operations
- Three implementations
  - Single lock server with volatile state
  - Primary/backup lock server, state stored on Petal
  - Distributed lock server with volatile state
- Global state replicated across all lock servers