CSE 60641: Operating Systems

- **The duality of memory and communication in the implementation of a multiprocessor operating system. Young, M., Tevanian, A., Rashid, R., Golub, D., and Eppinger, J. SOSP '87**
  - **Implementing user level memory system using the communication primitives in Mach**

Mach

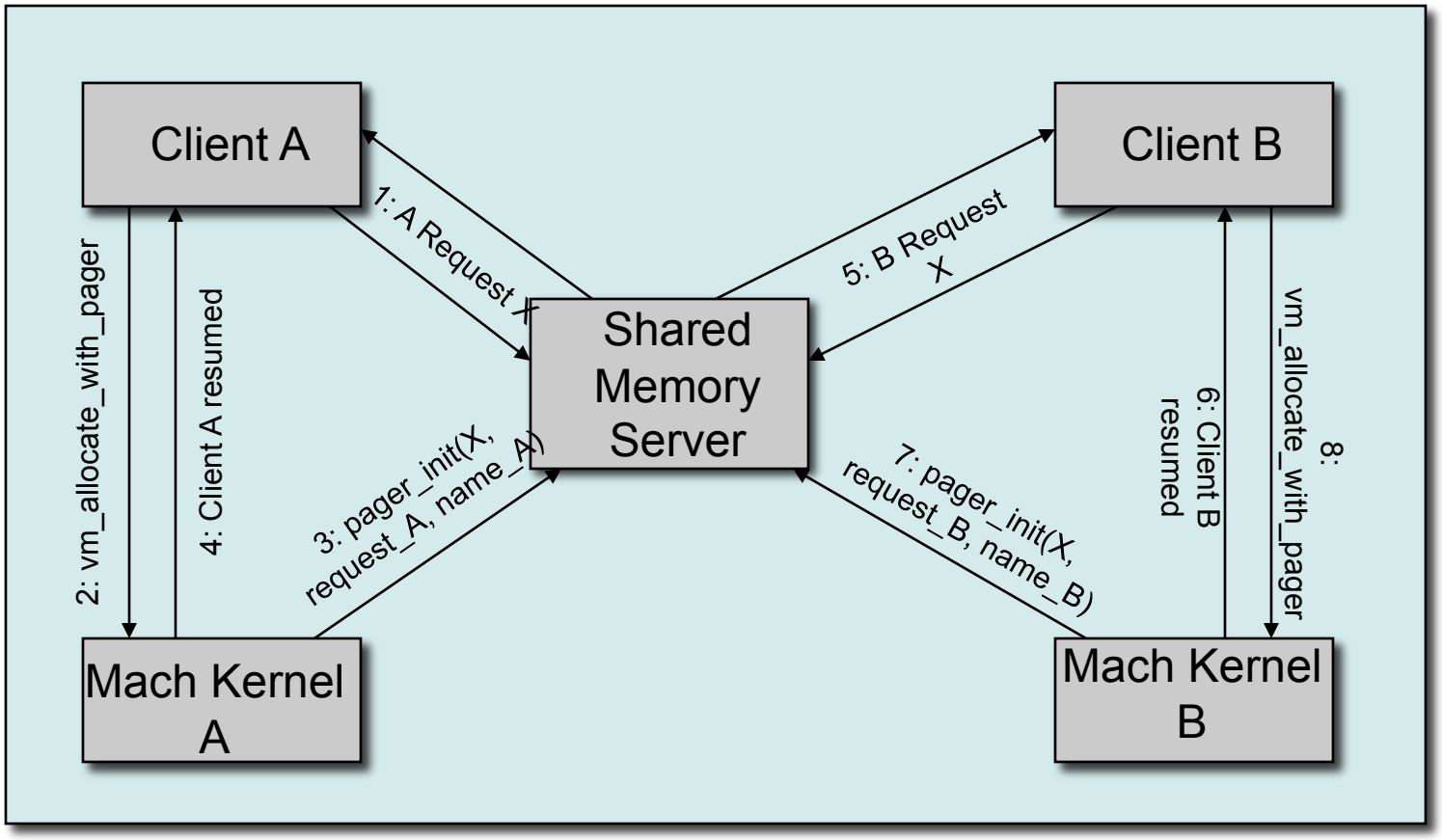
- Task, thread, port, messages, memory objects
- Ports
  - Inter-process communication
  - Protected bounded queue within the kernel
  - Access to port is granted by receiving a message containing a port capability (to send or receive messages)
  - Any number of senders, only one receiver
- Message
  - Fixed length header and variable size collection of typed data objects. Messages may contain port capabilities or imbedded pointers as long as they are properly typed.

External Memory management

- Unified way for user level management of backing store, file system, networked file system etc.
- Calls made by an application program to cause a memory object to be mapped into its address space
- Calls made by the kernel on the data manger
- Calls made by the data manager on the Mach kernel to control use of its memory objects
  - Calls are made via ports and are asynchronous
  - All the components can be distributed
  - Can support multiple clients for the same objects (share)

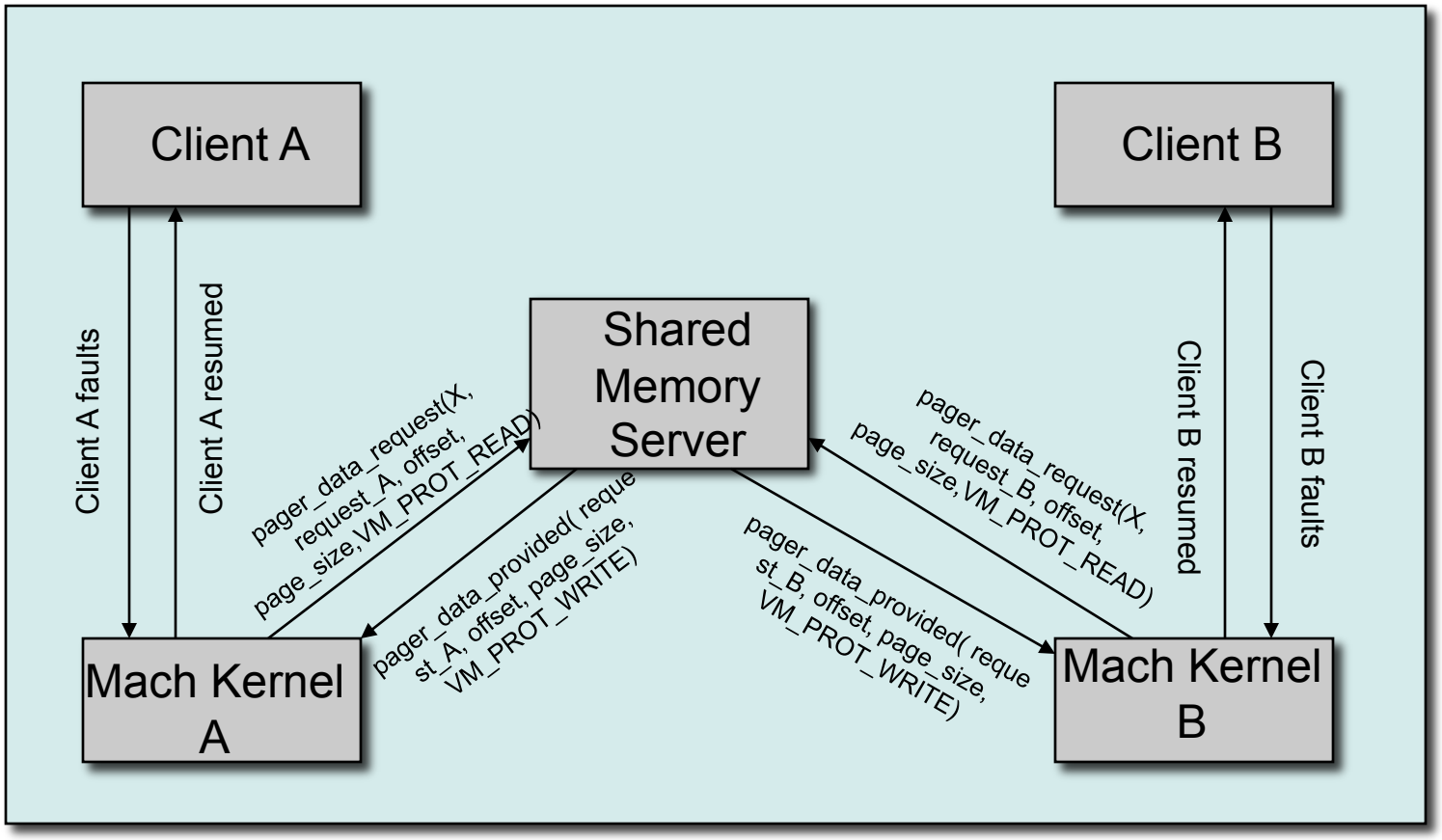
- Application to kernel interface
  - vm_allocate_with_pager: library can call this to implement a file system (by mapping secondary storage manager)
- Kernel to data manager interface
  - pager_init(), pager_data_request(), pager_data_write(), pager_data_unlock(), pager_create()
- Data manager to kernel interface
  - Pager_data_provided(), pager_data_lock(), pager_flush_request(), pager_clean_request(), pager_cache(), pager_data_unavailable()
- All interactions via messages to ports and asynchronous. Need extra buffers because of delay

Consistent Network Shared Memory (Initialization)



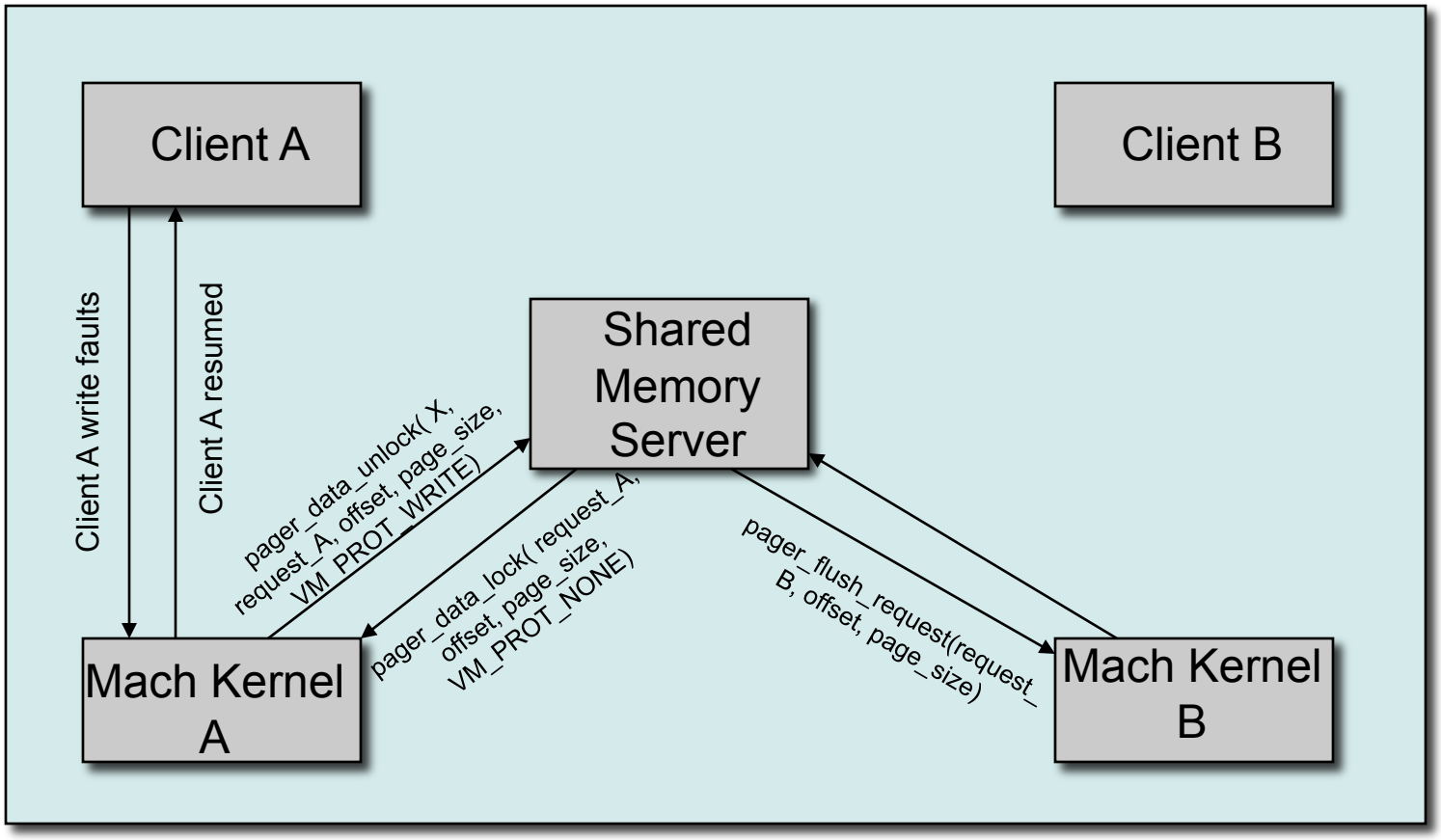
Courtesy: Rajesh Sudarsan @ VirginaTech

Consistent Network Shared Memory (Read)



Courtesy: Rajesh Sudarsan @ VirginaTech

Consistent Network Shared Memory (Write)



Courtesy: Rajesh Sudarsan @ VirginaTech

What are the key ideas?

- Implement memory as a communications mechanism
  - Hardware systems use similar ideas:
    - UMA – Uniform memory access
    - NUMA – Non-UMA
    - NORMA – No Remote memory access

- Applications:
  - Copy on reference process migration
  - Database management: Camelot
  - AI Knowledge base: Agora

Performance evaluation

- Sketchy
  - Avoid deadlocks using extra threads
  - Memory mapped file system can be fast – at the expense of storage safety

- The data manager is trusted to not be malicious and respond within reasonable amounts of time