

Active Storage Networks

John A. Chandy

Department of Electrical and Computer Engineering

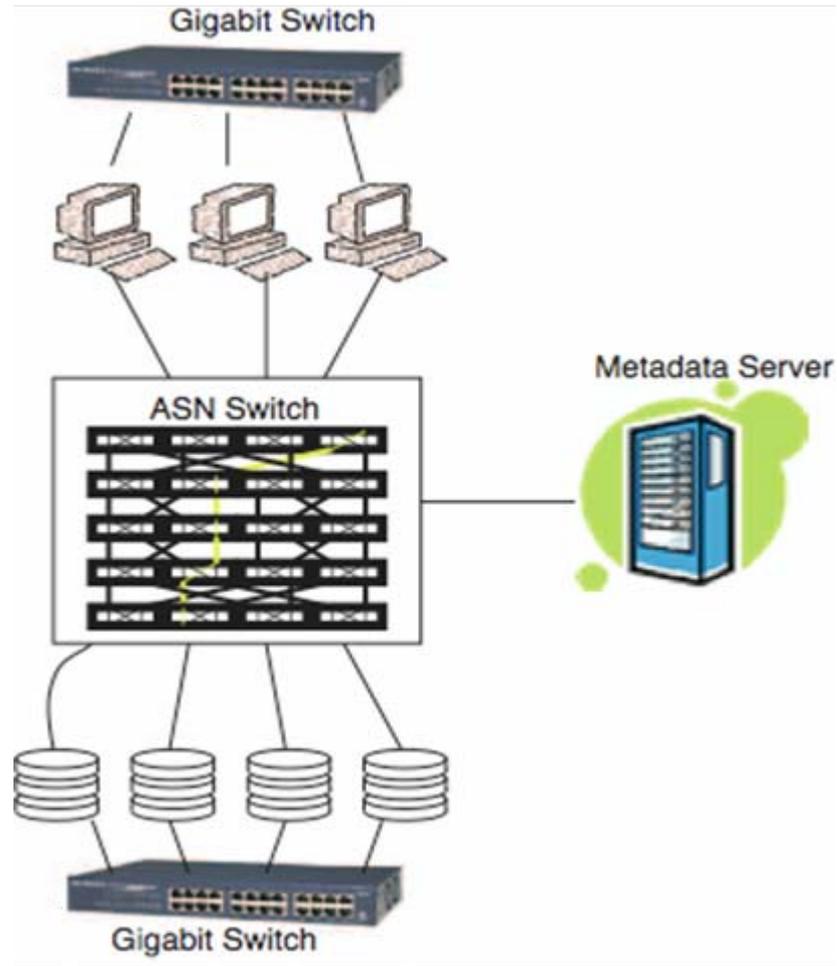


University of
Connecticut

Active Storage Networks

- Active Disks
 - Intelligence at the disk can distribute computation to parallel disks
 - Process data in streams
 - Disks only have local view of data
- Active Storage Network
 - Network has a global view of data
 - Distributed caching of file system metadata and data
 - Redundancy optimizations

Active Storage Network



Active Storage Networks

- Application operations
 - Reduction operations - database, scientific
 - Transformational operations - streaming
- File System Caching
- Redundancy optimizations

ASN Switch Implementation

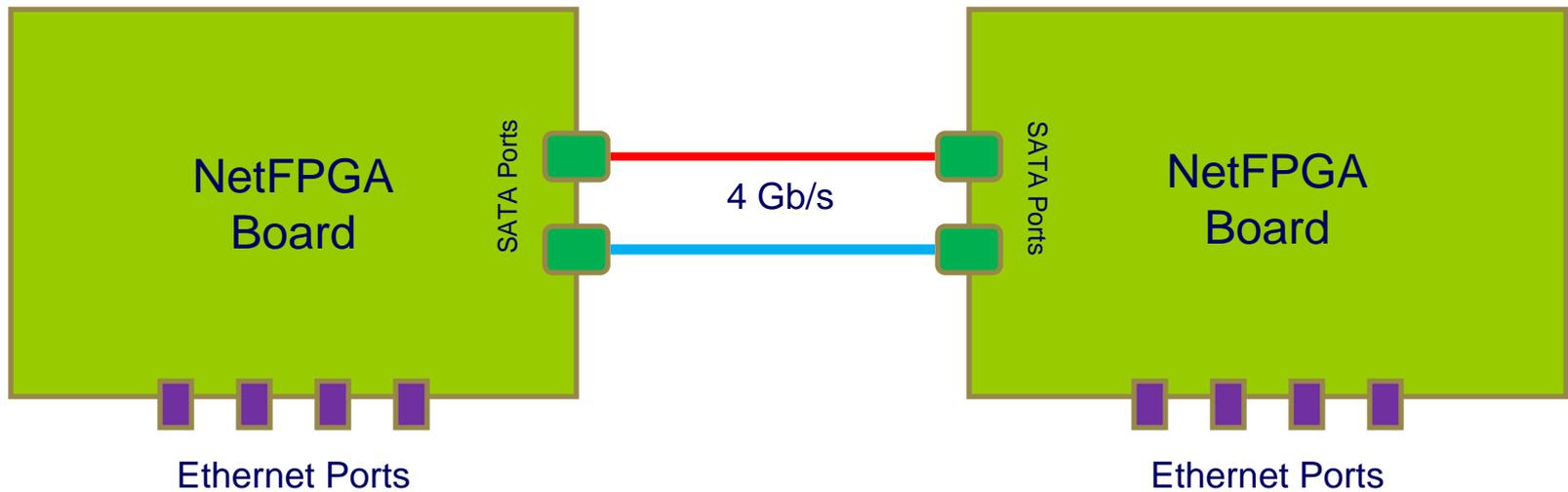
- Reconfigurable elements (FPGA)
 - Faster design and lower cost than ASIC
 - Fast enough to handle gigabit speeds
- Downloadable functions (netlets)
 - Software functions on embedded processors
 - Hardware functions on FPGAs

Hardware Implementation

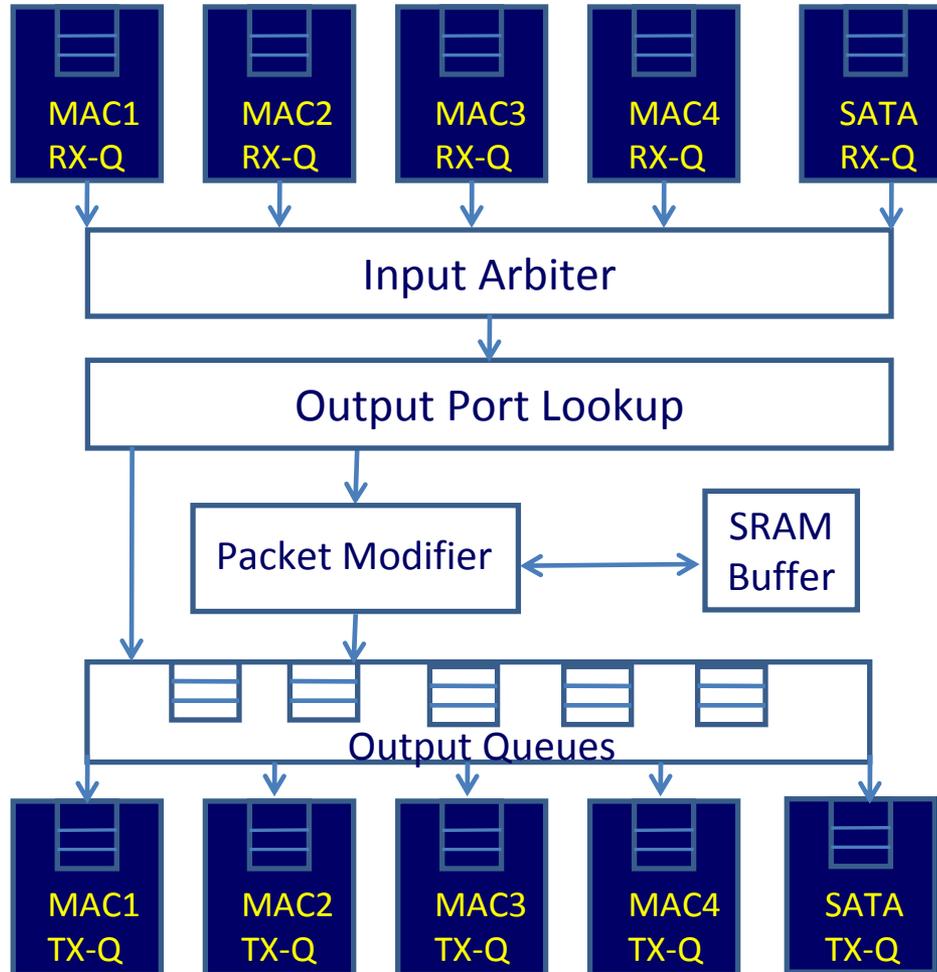
QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

- NetFPGA board from Stanford
 - Nick McKeown and John Lockwood
- 4 GigE connects
- 2 SATA connectors for node to node communication
- PCI bus for node to node communication

Hardware implementation

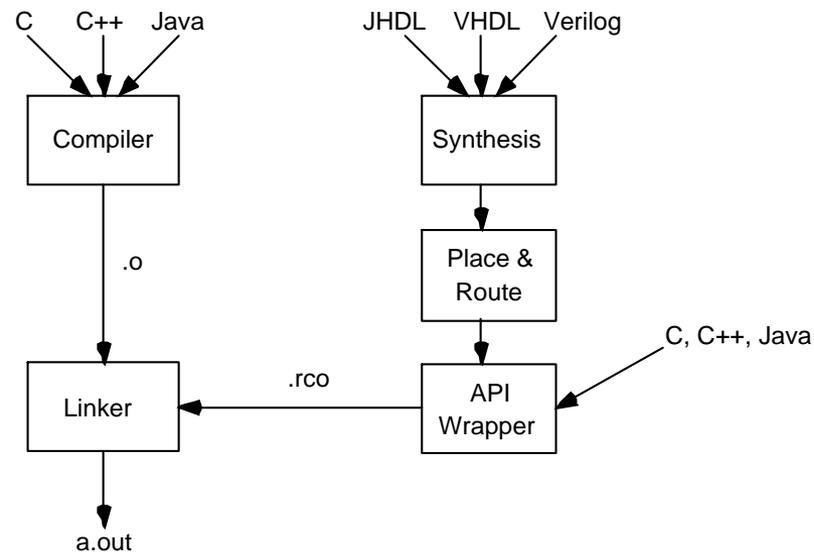
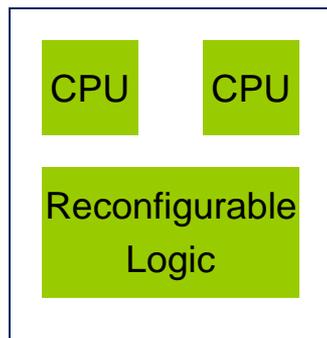


Node architecture



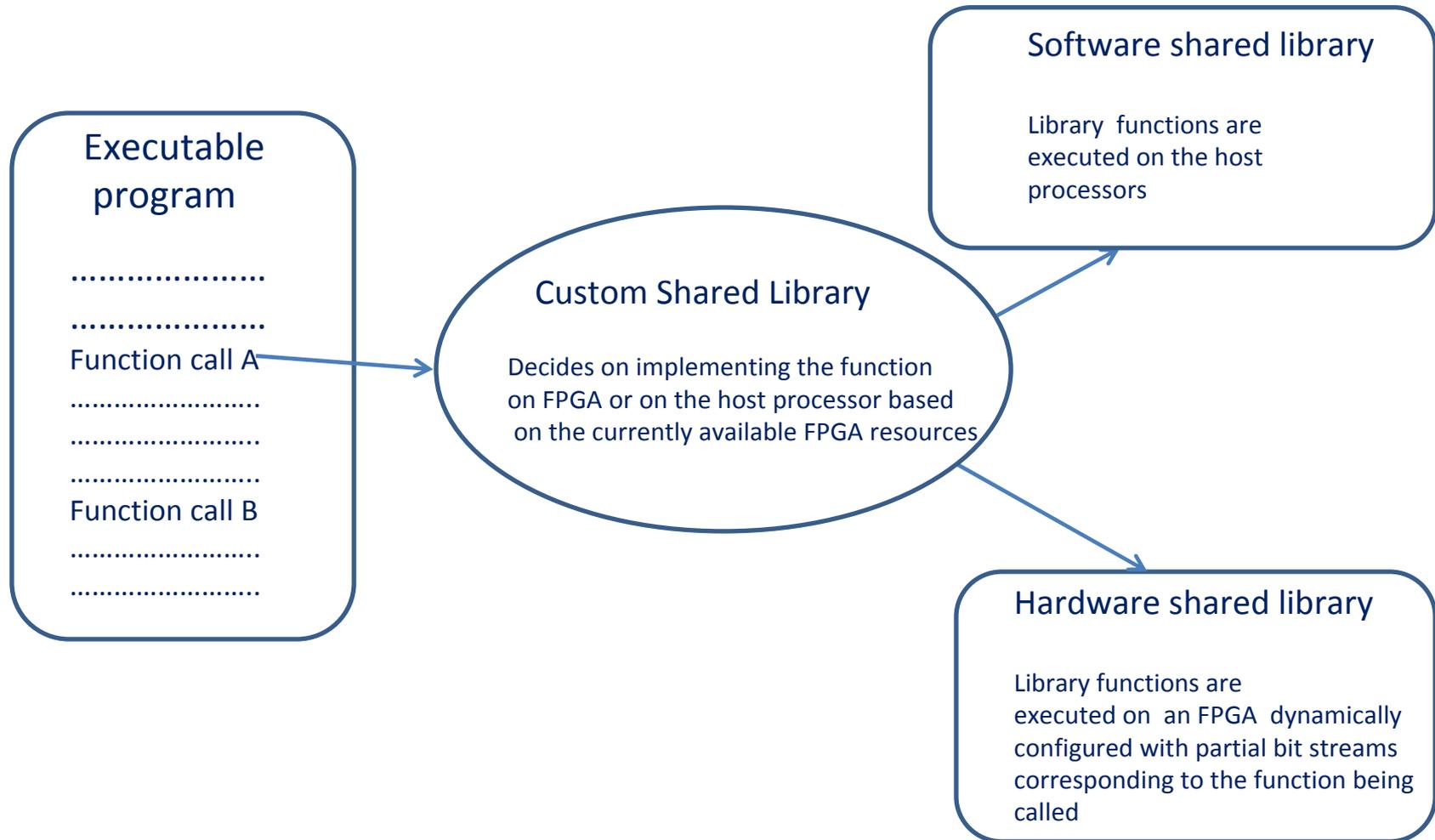
Adaptive Computing

- Hybrid Computing Tools



- Applications: Embedded Systems, High-Performance Computing, Compute-intensive algorithms
- Algorithms: String matching, Numerical Methods

Adaptive Computing



Active Disks using OSD

- Challenges
 - What is the programming model?
 - How do you download code to the OSD target?
 - How do you execute code on the OSD?

Active Disks using OSD

- Programming Model
 - Object-oriented
 - Attach object types to storage objects
 - Define methods for object types
 - RPC
 - Call methods on OSD remotely

Active Disks using OSD

- Example:
 - Record list object

```
List {  
    addRecord();  
    sortRecords();  
    searchRecord(string);  
};
```

Active Disks using OSD

- Distribute List across multiple objects - one per OSD target
- Client has a single unified List view
 - Proxy class coordinates methods on client's List with method calls on OSD Lists
 - Client proxy manages objects in distributed List
- Code written in Java

Active Disks using OSD

- How do you move Java code from client to target within OSD framework?
 - OSD objects allow users to set attributes on an object
 - We set an attribute on each object where the attribute is the .class file or .jar file associated with the object methods
 - One code attribute page with multiple code attributes
 - Allows different users to define their own code

Active Disks using OSD

- How do you execute the method remotely within the OSD framework?
 - New EXECUTE OSD command so that we can invoke a method
 - Recent change to OSD spec allows for CDB continuations
 - i.e. CDBs can accommodate longer
 - Introduced for support of scatter/gather and query commands
 - We use the CDB continuation to specify the method and parameters
 - Results (if any) returned directly or written to a new object

From T10/08-185r5 changes to OSD-2

Active Disks using OSD

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

From T10/08-185r5 changes to OSD-2

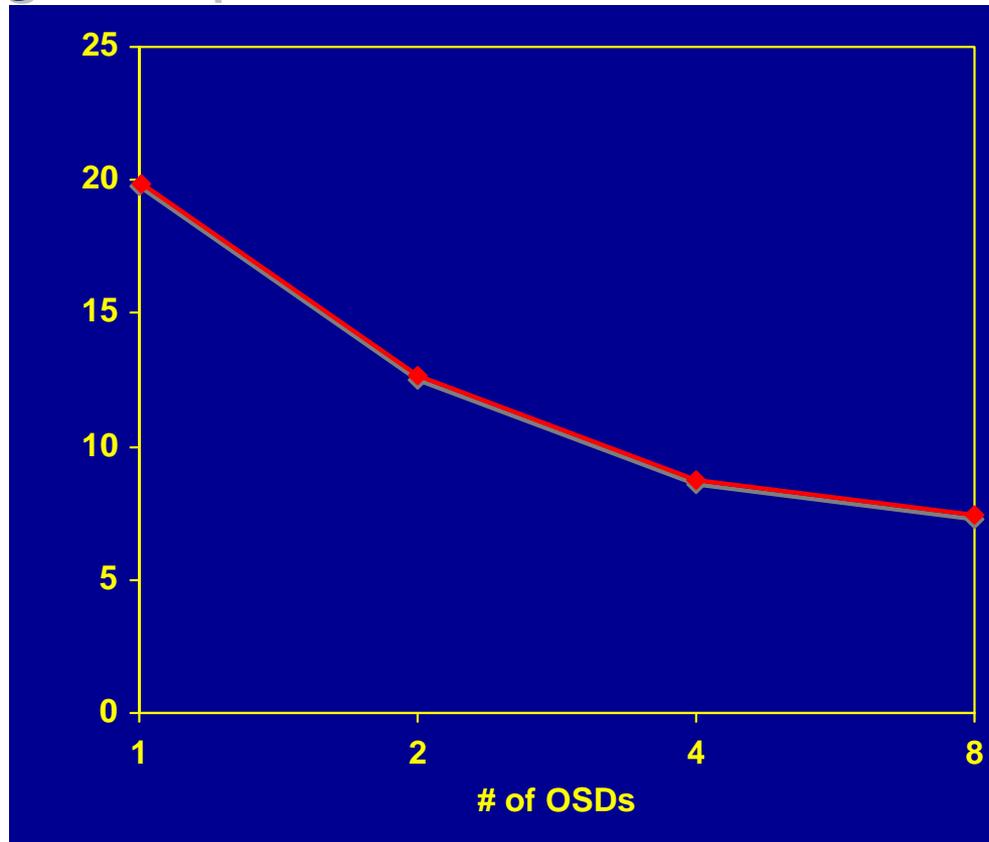
Active Disks using OSD

- CDB continuation descriptor

CODE ATTRIBUTE NUMBER	
METHOD IDENTIFIER	NUMPARAMS
PARAM_1_LEN	
PARAM_1	
...	
PARAM_N_LEN	
PARAM_N	

Simple List example

- Sorting on up to 8 OSDs

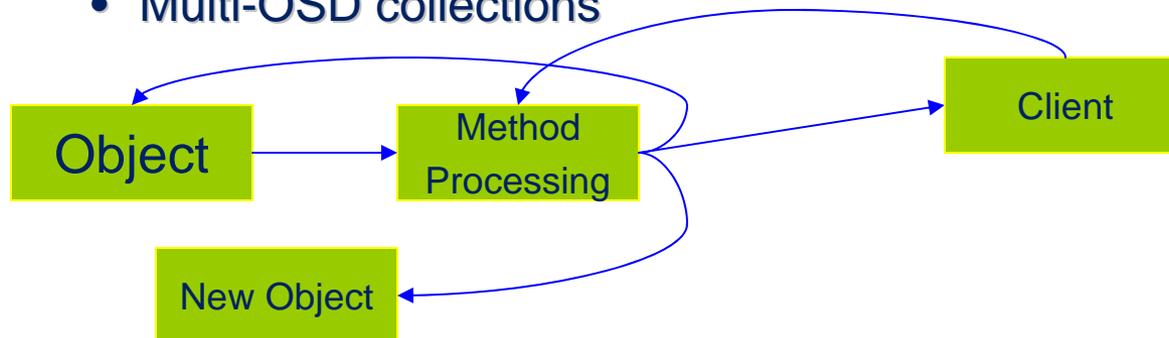


Active Disks using OSD

- Issues
 - Atomicity
 - Can we guarantee atomicity of operations?
 - Safety
 - Should we place limits on active disk computation?
 - CPU, memory, disk
 - Programming model
 - Java is portable - but still version dependent
 - Download a VM?

Active Disks using OSD

- Issues
 - What about other command data patterns?
 - Read only
 - Write only
 - Read/write
 - Read object - write to another object
 - Read collection - write to another collection
 - Multi-OSD collections



Active Disks using OSD

- Other issues
 - Is OO RPC the right model?
 - We are looking into functional models like MapReduce/Hadoop
 - What system services should the OSD target provide?
 - Local file system access?
 - Process/Thread management?
 - Application adoption is needed
 - Database is the most promising
 - Reductions
 - Update operations

Summary

- Investigation of ASN topologies and architectures
- Creation of ASN switch from reconfigurable components
- Programmable functions for ASNs
- OSD implementation of Active Storage
- Students
 - Tina John, Sumit Narayan, Janardhan Singaraju, Ajithkumar Thamarakuzhi, Anu Thiruvenkata Ramani