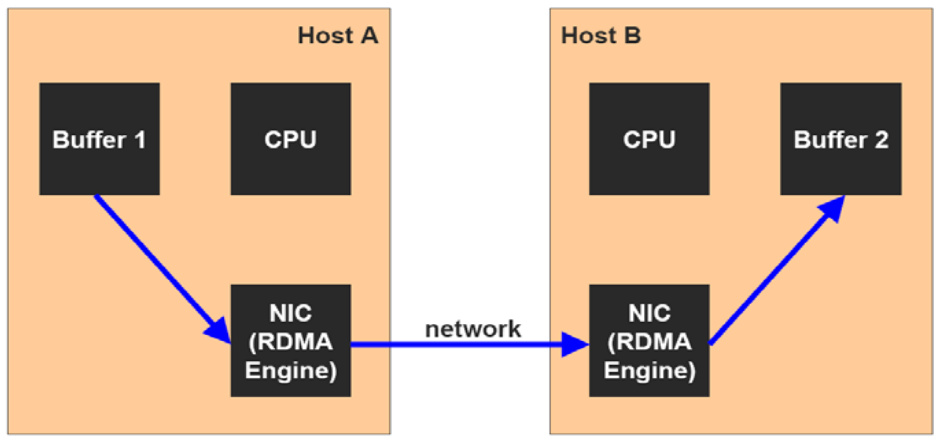
DAPL testing

CERN openlab II quarterly review
20 September 2006

Ryszard Jurga

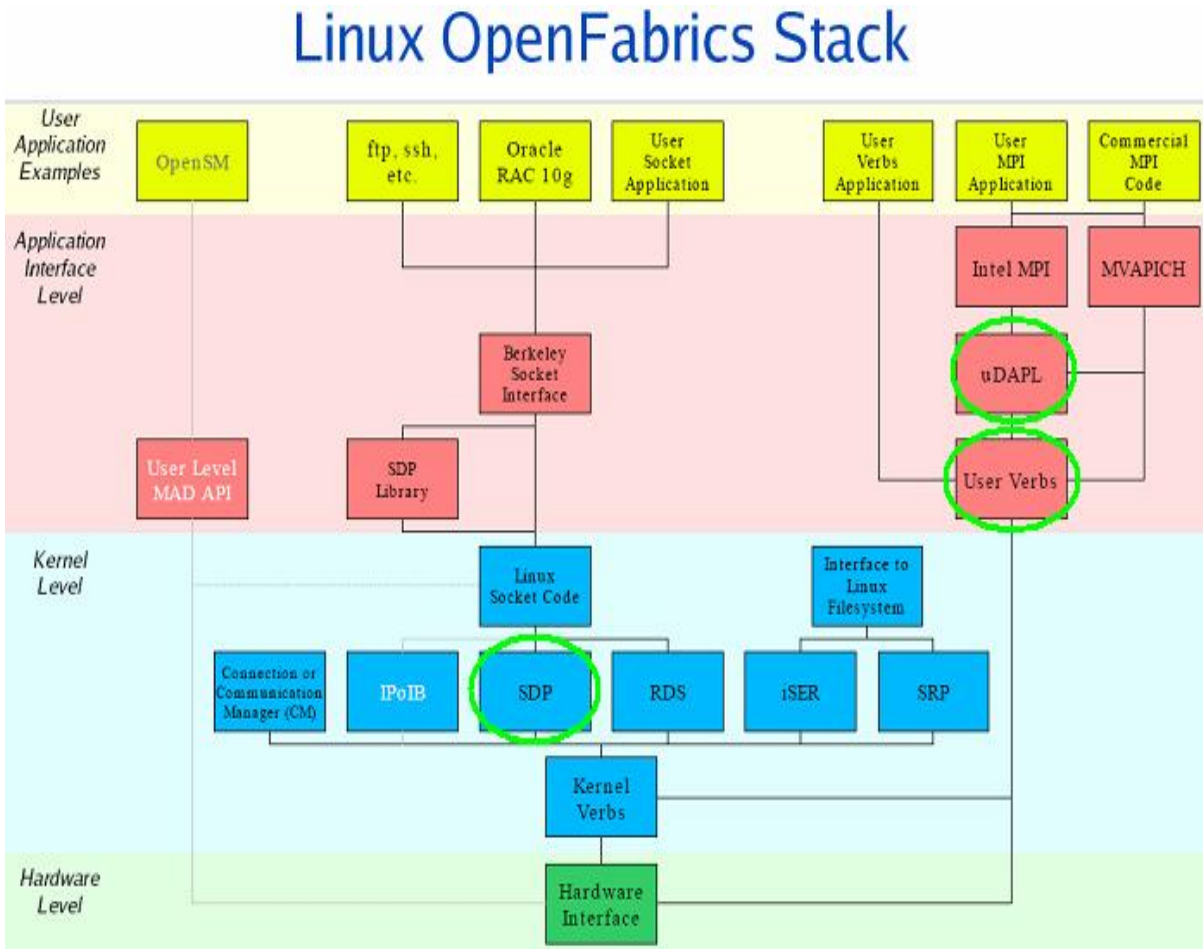
Introduction

- DAPL
  - Direct Access Programming Library – API that expoilts the capabilities of Remote Direct Memory Access (RDMA)
- Origin
  - An increase of a network speed
  - CPU bottleneck
    - memory to memory copying (3x,4x memory bandwidth requirements)
    - kernel context switching
- TCP/IP Offload Engine (TOE)
  - data is still copied
- RDMA
  - no copying of data
  - kernel bypass
  - WR/RD mode
- OpenFabrics Alliance
  - OpenFabrics/Infiniband Workshop at CERN, 26th June 2006



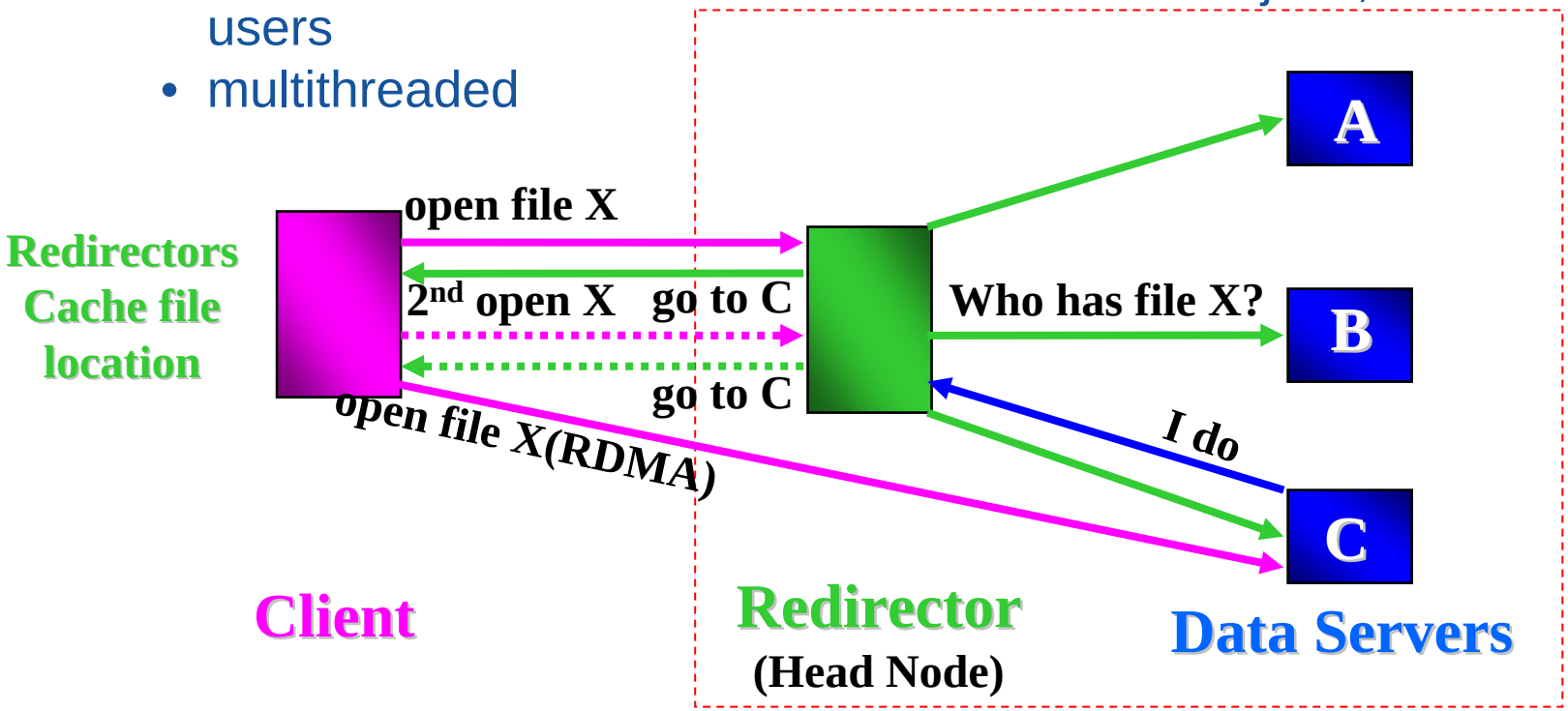
Openib stack

- Sockets Direct Protocol (SDP) and SDP Library
  - compatible sockets interface with Berkeley Socket
  - ~50% bandwidth gain (SDP vs. IPoIB)
- User verbs
  - Direct access to Hardware Interface, used directly by user applications
- uDAPL
  - Interface between user applications and user verbs
  - dapltest



Region of application of RDMA - xrootd

- xrootd introduction
  - toolkit for file base data access
  - thousands of tousands files scattered among multiple servers
  - thousands of concurrent accesses from batch jobs, end users
  - multithreaded



So far...

- Installation of the openib stack on 4 machines
- Initial tests with libsdp, dapltest
- **point-to-point transport prototype for sending files, RDMA Read and Write, Send/Recv type of transmission, DAPL**
- brainstorming with Andy Hanushevsky
- xrootd installation and study (Martin Swany, University of Delaware)

Questions & Future plans

- DAPL/RDMA
  - one port/many connections?
  - user verbs
  - measurements by means of the prototype
- xrootd
  - exchanging parameters about a new connection between a client and a server
  - Server side: maintaining a list of available network resources (free ports, remote buffers, etc.), life cycle
  - implementation of RDMA connection on the basis of our prototype
  - measurements