NXLib on Workstation Clusters[a]

Stefan Lamberts

Institut für Informatik
Lehrstuhl für Rechnertechnik und Rechnerorganisation
Technische Universität München
D-80290 München

Tel.: +49–89–2105 2382
email: lamberts@informatik.tu-muenchen.de

[a]This project was partially funded by a research grant from the Intel Foundation

Agenda

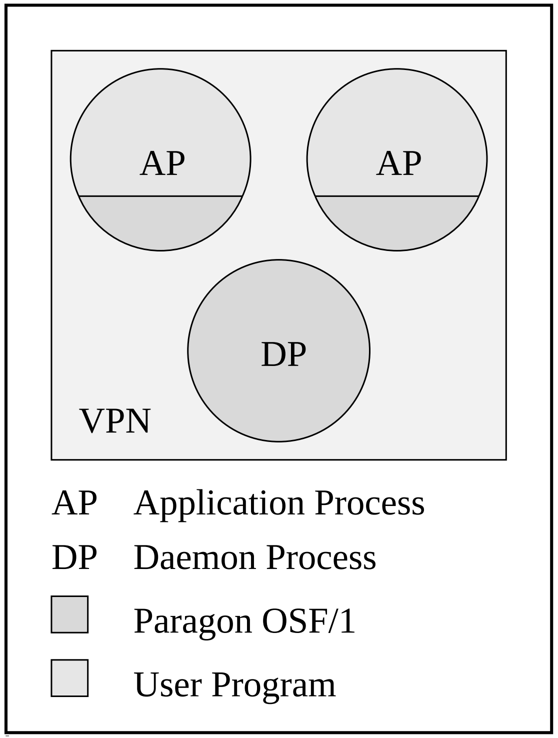
- ☐ Purposes of NXLib
- ☐ Emulating the Paragon System
- ☐ Implementation of NXLib
- ☐ Performance Results
- ☐ Obtaining NXLib

Purposes of NXLib

- education: workstations are anywhere accessible
  - first contact to parallelism in a familiar environment
  - withdraws teaching classes from the multicomputer
- development platform for parallel applications
  - only final performance tuning on the multicomputer
  - frees computational resources for production runs
- additional execution platform for suitable applications (coarse grain, little communication)
- partly substituting multiprocessor systems

Emulating the Paragon System

- software environment to provide a *virtual* Paragon on workstations
  - library with NX message passing calls
  - user commands, e.g. compiler scripts providing `-nx, -lnx` switches
  - system programs to emulate Paragon hard– and software properties
- Virtual Paragon Node as basic concept
- partitions established by mapping tables
- automatic application start–up on a network of workstations with corresponding command line switches



Implementation of NXLib

- heavyweight Unix processes
- TCP/IP socket connections for communication
- SIGIO signal on message arrival at socket

  ⇝ no busy waiting

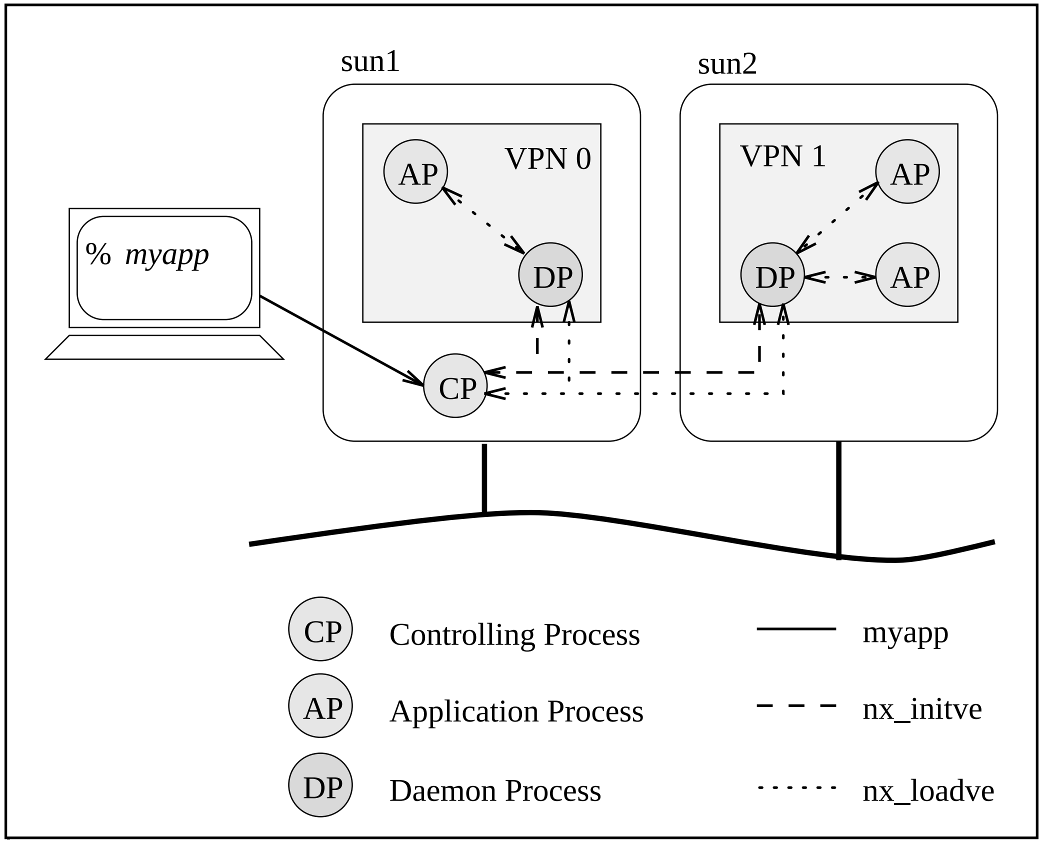
  ⇝ interrupt driven messages (`hrecv()`)
- remote machine access with `rsh` command

  ⇝ security

  ⇝ `stdout, stderr` output on local screen

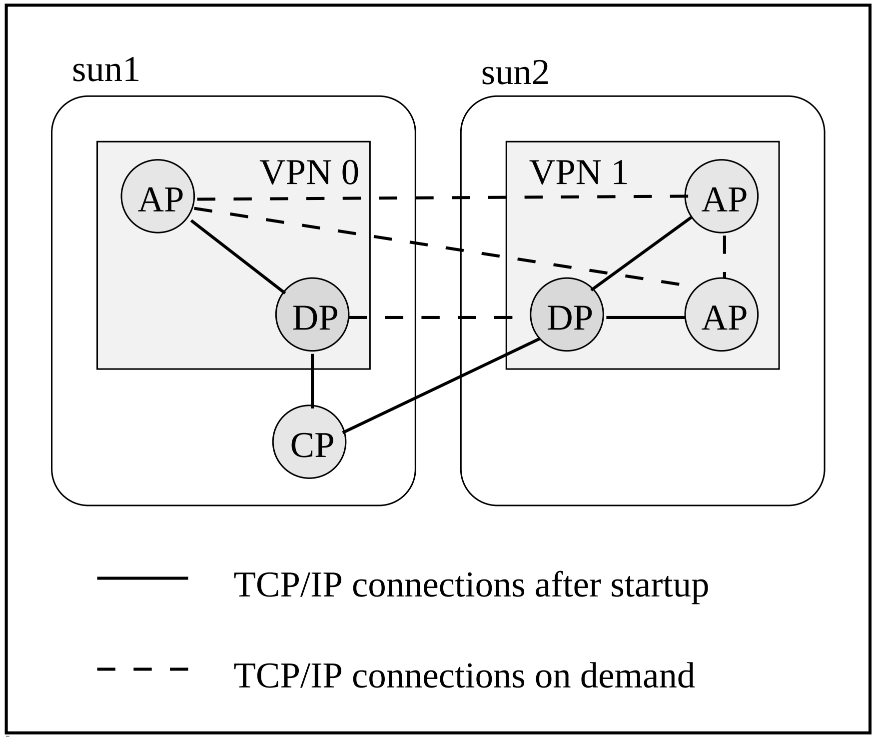
Start–Up of a NXLib Application

- controlling process on local machine
- daemon processes started with `rsh` on remote machines
- application processes forked by daemon processes

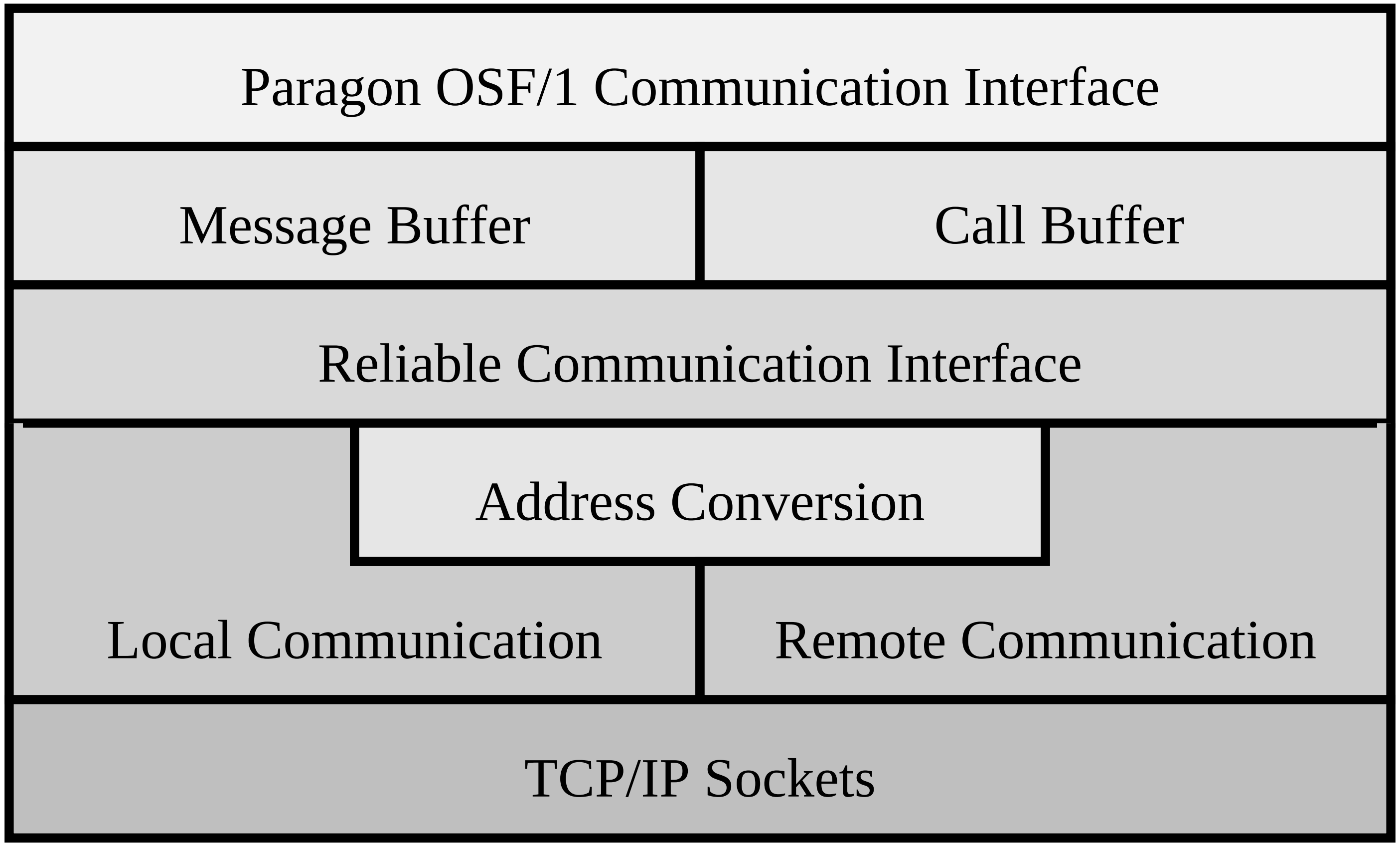


TCP/IP Connections

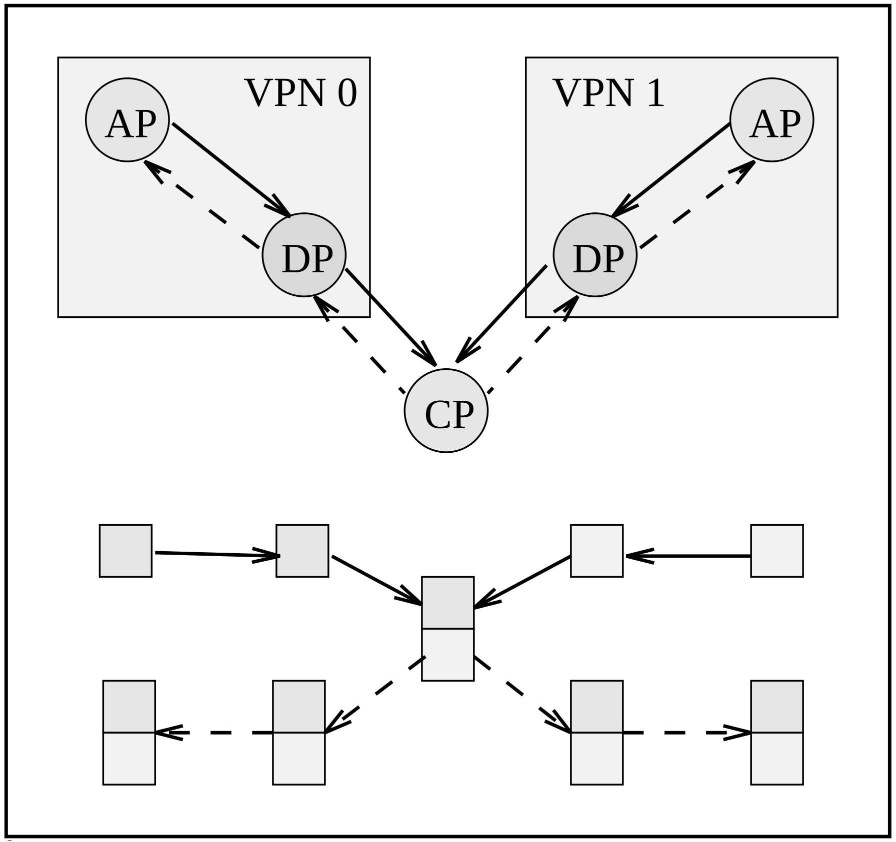
- TCP/IP connection set up on demand (during first communication)
- **one** TCP/IP connection between communicating processes
- no connection setup latencies after the first communication
- full interconnection of all communicating processes
- Address resolution
  - daemons know addresses of other daemons and of their application processes
  - address request via daemons



Communication Layers of NXLib



Global Operations



- controlling process gathers data via daemon processes
- distribution of data via daemon processes
- calculation within every process
  - $\leadsto$ user defined global operation
  - $\leadsto$ easy implementation of different distribution algorithm

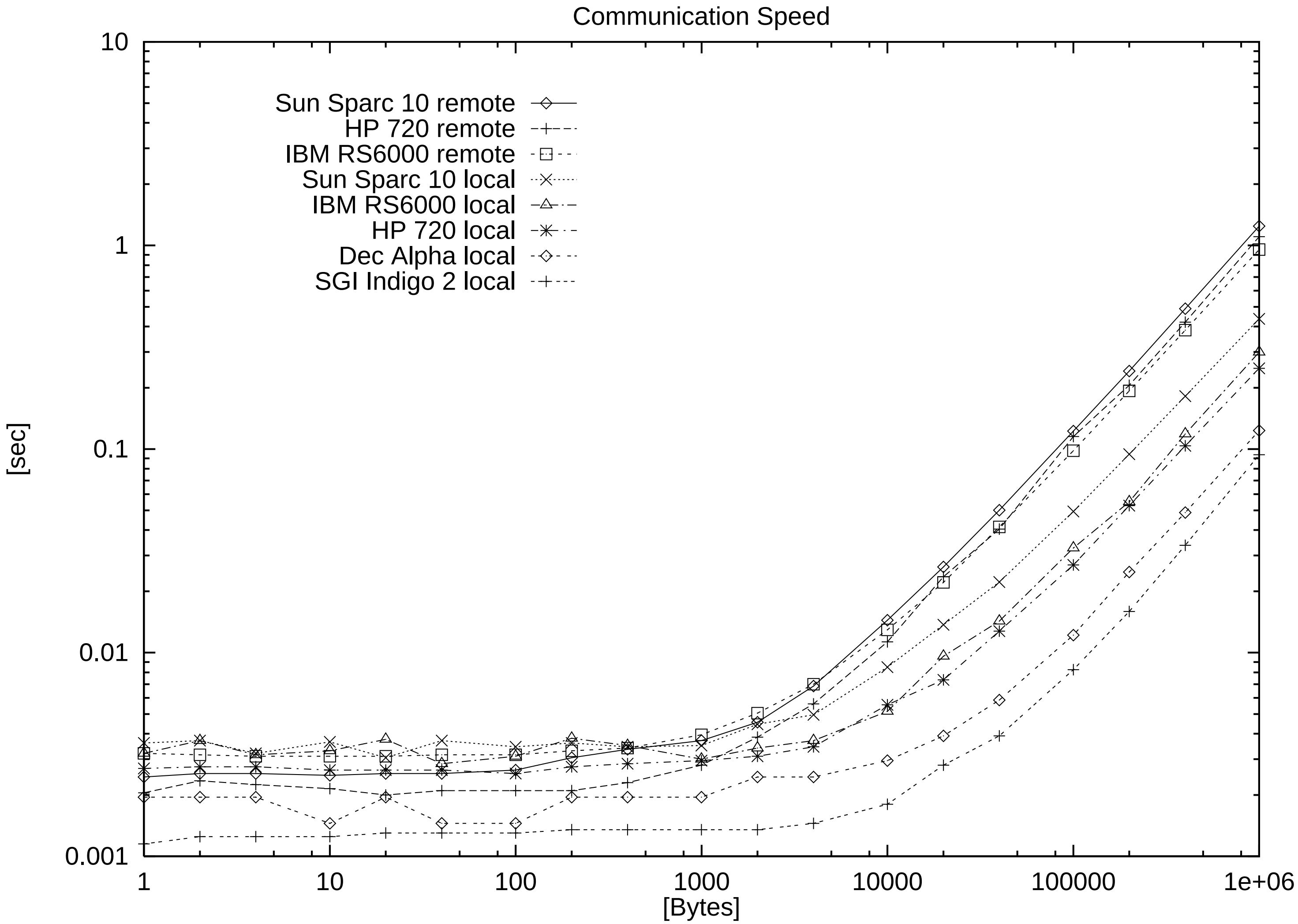
Summary

- special shell scripts to provide `-nx` and `-lnx` switches
- library with almost all Paragon NX calls
- automatic start of the application
- application command line switches (`-pn`, `-sz`, `-pt`, `-on`, `\;`)
- Ctrl-C/Ctrl-Z kills/stops all processes
- differences due to hardware and operating system properties
  - no single system image
  - buffer configuration and management
  - partition management
- limitations
  - iPSC compatibility calls
  - `nx_nfork()`
  - reloading of processes with previously existing ptype

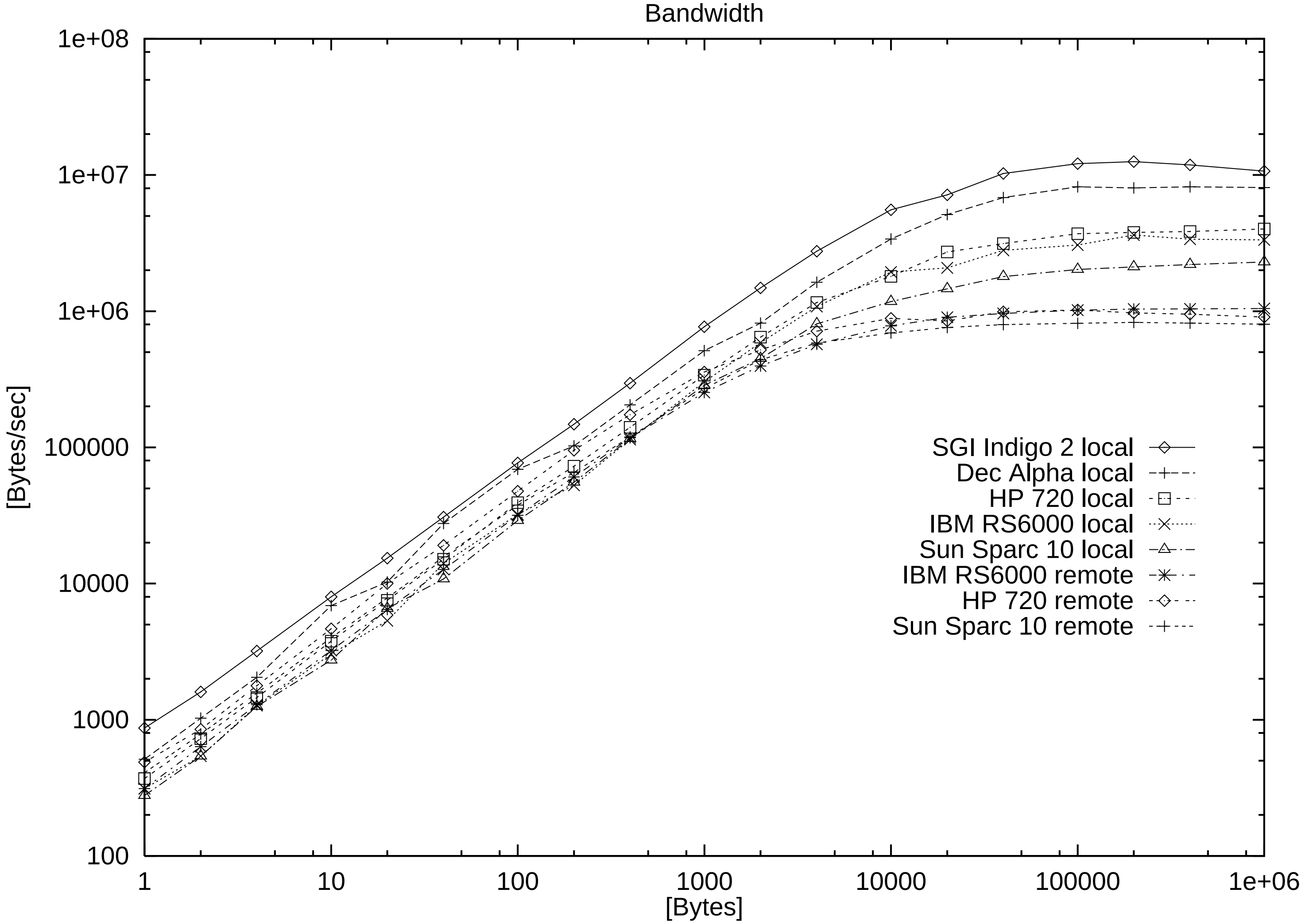
Performance Results

- Communication Speed
- Bandwidth
- Latency

Communication Speed



Bandwidth


Bandwidth
1e+08
1e+07
1e+06
100000
10000
1000
100
[Bytes/sec]
1
10
100
1000
10000
100000
1e+06
[Bytes]
SGI Indigo 2 local
Dec Alpha local
HP 720 local
IBM RS6000 local
Sun Sparc 10 local
IBM RS6000 remote
HP 720 remote
Sun Sparc 10 remote

Performance Summary

- Ethernet bandwidth: 1.25 MByte/s
- NXLib bandwidth

  IBM 1.04 MBytes/s (83 % of bandwidth)

  HP 1.02 MBytes/s (81 % of bandwidth)

  Sun 0.81 MBytes/s (64 % of bandwidth)
- Latency

  HP 1.7 ms

  Sun 2.2 ms

  IBM 2.6 ms

Obtaining NXLib

- Gnu library and Gnu general public license
- anonymous ftp: `ftpbode.informatik.tu-muenchen.de [131.159.0.156]`
- WWW page: `http://wwwbode.informatik.tu-muenchen.de/~lamberts/NXLib/NXLib.html`
- email: `nxlib@informatik.tu-muenchen.de`
- About 200 different ftp accesses on our ftp server