

# PIOsimHD

Julian M. Kunkel

Institute for Computer Science  
Parallel and Distributed Systems  
Ruprecht-Karls-Universität Heidelberg

07.08.2008

# Ziele

- MPI & MPI-IO Simulierbar
- Einsatz/Nutzbarmachung von Standard-Tools zur Analyse der Ausgabe
- Validierung ist wichtig (Glaubwürdigkeit des Simulators)
- Skalierbarkeit (bedingt realisierbar, sequentieller Simulator)
- Neue Algorithmen/Verhalten schnell und reproduzierbar testen
- (Varianten der einzelnen Funktionsimplementierungen auswählbar)
- Was wäre wenn der Supercomputer (oder das Cluster) diese und jene Hardware hätte?
- Wie ist die Last-balancierung? Wieviel Sequenziellen Anteil hat die Anwendung?

# Komponenten

- GUI um die Komponenten zu spezifizieren
- Simulator-kern (Java), nimmt Komponenten-Spec und mapped eine oder mehrere Applikation-Specs hinzu.
- PMPI-Wrapper erzeugt Eingabedatei für Simulator-anwendungen
- Benchmark-tool erzeugt Konfiguration für Komponenten Spezifikation (erste überlegungen)
- Tau-Trace-generator - konvertierung für jumpshot möglich, profiling mit paraprof möglich
- Integration in Scalasca, (Cube3) wäre möglich um Expertensystem zur Analyse zu nutzen (Late-Sender)

# Status

- Simulator-Kern simuliert Client Operationen mit wählbaren Varianten:
  - Rechnen(Zyklen)
  - Send/Receive (Tags, Wildcards) mit Eager + Rendezvous Protokoll
  - Barrier (Virtual)
  - Allreduce (Virtual, One-Gather-Scatter, Binary-tree)
  - File-Write/Read (PVFS-like)
  - Weitere sehr schnell implementierbar (Werden in Schritte zerlegt, vgl. simpler Endlicher-Automat)...
- Switchtopologie möglich (im Moment nur eine Route).
- Server simuliert I/O-Read und Write
  - Mit simplem Write-Behind, verwaltet Cachespeicher und gibt diesen sofort wieder frei.
  - (Hier gibts noch viel zu tun)

## Status (2)

- GUI erlaubt Komponenten zu platzieren, Eigenschaften auszuwählen (Dulip Praktikum)
- Simulator erstellt für Client-Operationen TAU-Trace
- Experimenteller PMPI Wrapper erstellt für genannte Operationen (außer I/O) XML-Trace Datei
- Benchmark bestimmt Latency für Netzwerk und gibt Zeiten für Allreduce, Barrier und Max-Größe der Nachricht für Eager-Protokoll

- Jumpshot Demo :-)