

High End Computing with K42

Paul H. Hargrove and
Katherine Yelick

Lawrence Berkeley National Lab

Angela Demke Brown and
Michael Stumm

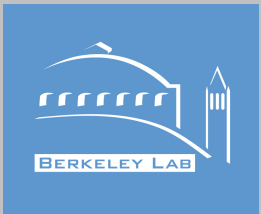
University of Toronto

Patrick Bridges

University of New Mexico

Orran Krieger and
Dilma Da Silva

IBM



Project Motivation



The Answer To ...

HIGH END COMPUTING WITH K42

- The HECRTF and FastOS reports enumerate unmet needs in the area of Operating Systems for HEC, including
 - Availability of Research Frameworks
 - Performance Visibility
 - Adaptability to Application Requirements
 - Support for Architectural Innovation
- This project uses the K42 Operating System to address these needs

What Is K42?



The Answer To ...

HIGH END COMPUTING WITH K42

- **K42 is a GPLed research O/S from IBM**
- **Framework for PERCS (DARPA HPCS) work**
- **API and ABI compatible w/ Linux on PPC64**
- **Runs most Linux kernel modules (fs, etc)**
- **Many O/S services implemented in user-space**
- **Object oriented**
 - Every virtual or physical instance is an object
 - Every class may have multiple implementations
- **Has extensive performance/tracing capabilities**
- **Design/implementation is very SMP-scalable**

First Year Activities



The Answer To ...

HIGH END COMPUTING WITH K42

- **Activities in our three main areas**
 - Making K42 more “available” as a research framework for HEC
 - Lowering the barrier(s) to entry
 - Porting to additional platforms
 - Providing an HEC environment
 - Adaptability to Application Requirements
 - Tailor the runtime to the application’s needs
 - Performance Visibility
 - Expose performance data

Availability of Research Frameworks 1



The Answer To ...

HIGH END COMPUTING WITH K42

- **k42 . sh** tool from LBNL, automates installation and use of K42
 - Configure/prep/install console machine
 - Download, patch and compile of K42
 - Configure OpenFirmware on victim
 - Works on both real H/W and Mambo*
 - Boot/run K42

*Mambo is IBM's full system PPC simulator

Availability of Research Frameworks 2



The Answer To ...

HIGH END COMPUTING WITH K42

- **UNM porting K42 to Opteron**
 - PPC port of K42's Protected Procedure Calls from gcc-3.x to gcc-4.x
 - Reimplemented Xen MiniOS demo on Xen 3.0/Opteron
 - Will make available to FastOS community
 - Restructuring scheduling code
 - Will allow replacement of PPC-specific bits
 - Will ease construction of HEC-specific bits
 - Revising w/ IBM documentation on the key internals of K42

- **IBM work in several areas**
 - **K42-on-Mambo improvements**
 - **Rewrite of boot code for ease of porting**
 - **Stabilization of K42's TCP/IP networking**
 - **Enables ssh & MPI**
 - **Added MPI to default root fs**
 - **Transition to `git` for Linux sources**

Hardware Performance Counters



The Answer To ...

HIGH END COMPUTING WITH K42

- **UofT infrastructure for hardware performance counters**
 - **Precise event counts**
 - Cache misses & invalidations, branch misprediction, etc
 - **Statistical events**
 - Number and reason for processor stalls

The “Arbiter Infrastructure”



The Answer To ...

HIGH END COMPUTING WITH K42

- **Work at UofT and IBM**
 - Masters thesis of Raymond Fingas
- **Key Features**
 - Dynamic dispatch of method calls for all runtime objects
 - “Hot swap” or “prelude/postlude”
 - Transparent to both caller and callee
 - Can be traded for performance
 - Encapsulates the ugly details

Performance Monitoring Via Arbiters



The Answer To ...

HIGH END COMPUTING WITH K42

- **Provides gprof-like profiling data**
 - Execution time
 - Invocation count
 - Exact, not statistical in nature
- **Will soon incorporate the H/W performance counters work**
- **Ultimately can be used to drive automatic adaptation**

Performance Visualization



The Answer To ...

HIGH END COMPUTING WITH K42

- **IBM and UofT working on a GUI**
 - Display the K42 object structure
 - Display associated performance data
 - In real time
 - On a per-object basis
- **Currently a basic version is working**
 - Displays object structure
 - Percent of CPU spent in each object

Next Steps



The Answer To ...

HIGH END COMPUTING WITH K42

- **Continue porting HEC environment**
 - Scalable Systems Software Suite
 - Global Address Space Languages
- **Complete the Opteron port**
- **Tie together arbiters and hardware performance counters**
- **Extended GUI to direct performance gathering interposition**