

Numa Topology API

Paul Dorwin

IBM Linux Technology Center

Topology API

- There is a need for topology API to discover and manage system resources:
 - Performance tools can bind to specific CPUs and make measurements.
 - Diagnostics can bind resources into standard and non-standard configurations.
 - Applications can bind system resources for better performance.

Topology API

- The topology design is based on:
 - Kanoj Sarcar's topology patch.
 - Provides basic topology infrastructure / kernel API
 - A port of an IBM-NumaQ performance tool.
 - The tool utilizes various system calls from dynix/PTX which provides topology discovery and management.
 - The tool also utilizes dynix/PTX system calls to access i386 and platform specific performance counters.
 - The tool, patches and user library code will be released soon.

Topology API

- Additional patches of use
 - `cpus_allowed` -exports `cpus_allowed` to `/proc`
 - non-identical cpu patch
 - Nodes may contain non-identical cpus
 - I386 performance counters patch
 - Provides `/proc` access to i386 perf registers

Topology API

- Main points of the design
 - Provides capability to discover system topology.
 - Provides capability to online/offline resources.
 - Provides access to architecture/platform specific registers.
 - Adds a /proc interface for user/script access.
 - Adds a new system call for application access.
 - Adds a user library for standardized API.
 - Adds a set of user level tools.