



# Next Generation POSIX Threading

## Moving Linux to the Enterprise

Bill Abt

Saurabh Desai

Dave McCracken

David Howell



# Agenda

- Project Overview, Design Points & Status
  - Bill Abt, IBM
- POSIX Testing, Performance, Compatibility
  - Saurabh Desai, IBM
- Linux Kernel Changes
  - Dave McCracken, IBM
- Case History
  - David Howell, Intel
- Futures
  - Bill Abt, IBM
- Q & A



# Project Overview

- Derived from the GNU Pth userlevel threading package.
- Licensed under the LGPL.
- Goals:
  - Provide a POSIX 1003.1 compliant threading package for Linux
  - Improve scalability by introducing an M:N threading model
  - Improve user application portability by making Linux more compatible with commercial UNIX offerings thru use of the POSIX API.
  - Make Linux threading ready for the Enterprise applications.



# Key Design Points

- Userlevel library, written in ANSI “C”, exporting the POSIX threading API.
- Uses a modified clone() interface to create Linux “kernel” threads.
- SMP support built in; locking, scheduling, etc.
- Key items configurable on a per process basis.
- Integrates into GLIBC as a LinuxThreads replacement.
- Provides application binary compatibility with LinuxThreads where POSIX compliance is not compromised.
- Runs on virtually all processors supported by Linux.



# Current Status

## ■ NGPT 1.0.0

- Released in June of 2001.
- Most initial project goals were met.

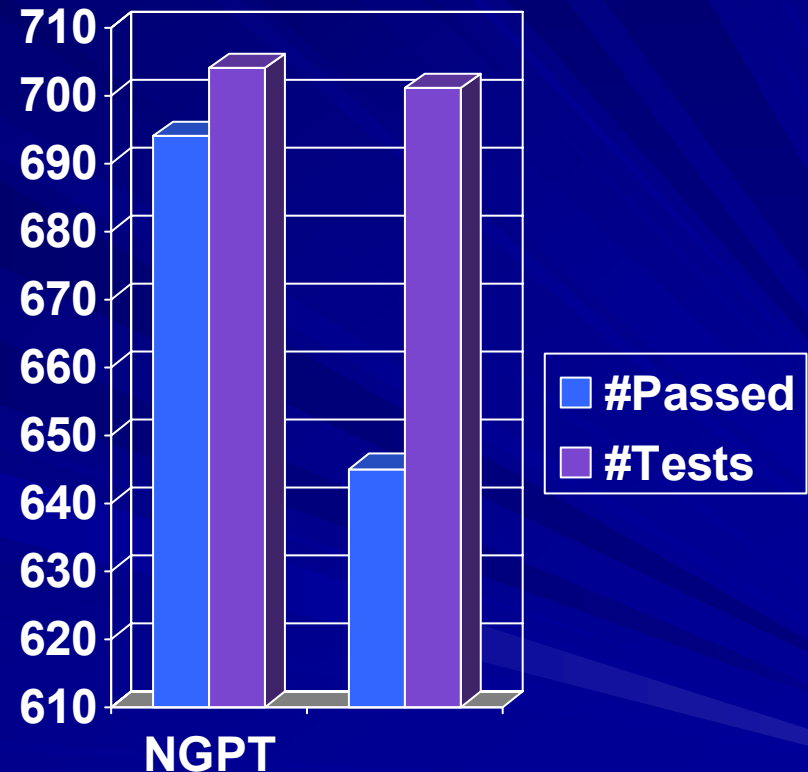
## ■ NGPT 1.1.2

- (To be) Released 25 January, 2002.
- 95% POSIX compliant.
- Performance on par with LinuxThreads on UP and superior on SMP.
- Near total binary compatibility with LinuxThreads.

# POSIX Compliance Testing using VSTH-lite



- More tests are attempted by NGPT, i.e. more POSIX functionality.
- More tests are passed by NGPT, i.e. more POSIX compliance.





# Performance

- Goals
  - Within 5% of LinuxThreads on a UP machine.
  - 10% or more increase on SMP.
- Statistics are still being gathered.
- While performance is always in the mix, we're concentrating on compliance and functionality.
- Goals are expected to be met in NGPT 2.0 release.



# Compatibility

- Designed a glibc plug-in replacement for existing libpthread shared library.
- Strives to be as binary compatible as possible with LinuxThreads.
- Enables most LinuxThreads applications that are coded to POSIX API to run without *recompile*.
- Where binary compatibility clashes with POSIX compliance, binary compatibility will be sacrificed.



# Thread Related Kernel Issues

- The Pthread process/thread model assumes state is process-wide.
- The Linux cooperating tasks model has much of its state per task.
- Linux has some hooks for sharing state via clone and thread groups.
- Obvious fix is to allow Linux to share more state between tasks.



# What State Are We Talking About?

- Linux is already capable of sharing address space, open files, signal handlers, and (limitedly) process id.
- The critical item missing is pending signals.
- Other missing items include system semaphore state and credentials.



# NGPT Kernel Patches

- Original signal fix funneled all signals to thread group leader. This patch is very small and isolated and therefore suitable for integration into the 2.4 stable tree.
- New signal patch implements fully shared pending signals. This patch is larger and not as isolated and is thusly, more suitable for the 2.5 development tree.
- Both patches support killing all tasks in a thread group on signal exit.



# Case History – Telecom Industry

- ALL have experience with Solaris.
- Has a mature threads implementation.
- POSIX threads conformance is key.
- Industry is very standards driven.
- For OS platforms POSIX API standards are expected.
  - Portability from RTOS (Blades/Server) or server OS requires this.
  - Real-time support a necessity.
- Thread performance and scalability are also key.
- Typical applications are multi-threaded.
- Several Telcom server apps are thread intensive
  - Softswitch, where at peak load 40K-50K concurrent threads are typical.

# NGPT is a better solution for threading on the Linux platform.



- Current POSIX threads conformance state is weak for LinuxThreads.
- NGPT improves on current LinuxThread conformance.
- Current LinuxThreads performance, scalability, and capacity are suspect.
- NGPT improves on current Linux thread support for all of these.
- Configurability of current LinuxThreads implementation is weak.
- NGPT improves this with increased configurability.
- Debugging of threads in the current LinuxThreads implementation is weak.
- NGPT has plans to improve the thread debugging in several areas.



# NGPT Integration Into Intel Telecom Linux Technology

- Initiative to improve Linux for the Telecom market.
- Drive feature set to Distributors interested in Telecom space.
- TLT 1.0 has incorporated NGPT 1.1.1 for evaluation (Available).
- TLT 1.5 will integrate NGPT 2.0 as replacement for existing threads.



# Success Criteria for NGPT

- Seamless binary replacement for existing LinuxThreads library.
- Provide API coverage equivalent to existing LinuxThreads implementation.
- Add key missing features for Telecom (real-time).
- Acceptability for embedded applications/true real time.
- Acceptance by Linux distributors, TLT is working to facilitate this.

# Multi-threaded Applications Being Qualified



- Apache Web Server
- Linux Enterprise Event Logging
- TAO CORBA ORB
- H323 Traffic Generator
- TLT Linux Resource Monitoring
- MySQL relational database



# NGPT 2.0

- POSIX semaphore support, including process shared semaphores.
- Process shared mutexes (support for the pshared thread attribute).
- More support for POSIX functions defined in the new draft.
- Native gdb (debugging) support.
- More platform support.
- Better performance.
- Better, easier configurable parameters.



# NGPT 2.0 GDB Support

- Provide `libthread_db` support to enable basic gdb live thread group debugging.
- Enables other tools to access thread state data using `libthread_db`.
- Enable NGPT to assist in improving kernel core dumping for thread groups.
- Extend gdb post-mortem analysis for thread groups.



# IBM Exploitation of NGPT 2.0

- IBM JVM for Linux
- WebSphere for Linux
- DB2 for Linux
- Carrier Space Linux



# Questions?

