ICSC Socket API Extensions Working Group

David Edmondson

dme@sun.com

Introduction

The Interconnect Software Consortium

- Created under the auspices of the Open Group
- ICSC was formed to develop new APIs and extensions to the Sockets API suite in order to improve application offerings on Unix-based OS offerings
- Includes major Unix vendors and others
  - ▷ *Compaq, Fujitsu, HP, IBM, Sun*
  - ▷ *NetApp, VIEO, Intel, Infiniswitch*
  - ▷ *Lane15, Oracle, IB Source*
- A steering committee oversees the activities of the consortium

Working Groups

- Sockets API Extensions
- Interconnect Transport API
- Fabric Management API

Socket API Extensions

The socket API has seen no significant changes in over ten years

These changes are driven by:

- new application requirements,

- customer-drive performance expectations (bandwidth and latency),

- new application paradigms:
  - ▷ *a shift to async operations,*
  - ▷ *desire for OS bypass / protocol off-load,*
  - ▷ *copy avoidance,*
  - ▷ *increased use of RDMA-based technology,*
  - ▷ *unified completion model,*
  - ▷ *...*

- some perception in the ISV community that the current API does not provide the performance and functionality to meet current requirements

Of the three working groups, one is specific to Infiniband (Fabric Management).

Working Group Approach I

Charter

- The Sockets API Extensions Working Group will specify additions to the standard UNIX (Berkeley) Sockets API that improve the efficiency of network programming. The interfaces will be defined for user-space subsystems and will facilitate protocols such as SDP that bypass the TCP/IP software stack. The interfaces may additionally be adapted for kernel-space subsystems. The primary goal is to preserve the large existing base of sockets applications, enabling portability where reasonable, while providing a natural evolution to more efficient transport protocols in an industry-standard manner. It is intended that these extensions will be submitted to the Austin Group for consideration in a future revision of the Single UNIX Specification.

The group has two co-chairs:

- Satya Sharma, IBM Corporation
- David Edmondson, Sun Microsystems

Requirements specification

- A product of the working group, this refines the charter to identify specific areas of interest and effort.
- Subject to approval by the steering committee

Working Group Approach II

Definition of the new and extended APIs

- The principal output of the group
- In the form of the Single Unix Specification

Per-transport protocol or protocol family details

- Detailing the mapping of the API to relevant protocol features, as appropriate

Implementors Guide

Users Guide

Test Suite Specification

Current Areas of Discussion

Support for asynchronous IO and control operations

Completion notification for asynchronous operations

Improved memory management techniques (buffer pinning, etc.)

Support for user-level socket library and protocol implementation

New protocol features (SCTP multiple addresses per socket, etc.)

Codify common existing API extensions (e.g. sendfile())

ISV Participation is Key

ISV adoption is **the biggest** hurdle for any new API

Key applications and relevant ISV partners need to be identified to participate in the working group:

- iPlanet, Java, Oracle, DB2, Notes, ...

The majority of the current members are OS developers

The API changes are not the whole story - to get anywhere near 10Gb/s throughput will require effort on every front:

- application, API, VM, network stack, network hardware, ...
- developer support, technical marketing, ...

These areas are opportunities for companies to innovate within the standard

How to participate

Various levels of involvement are possible:

- Badger your ICSC representative to ensure that your opinions are known,
- Review and comment on the requirements and specifications as they are drafted,
- Participate directly in the development process.

Twice weekly 1 hour teleconference with representatives from member companies, with a set agenda

Working Group Timescales

May 2002: Requirements and Work Plan documents ready for SC approval

August 2002: Phase 1 specification

September 2002: Phase 2 requirements

November 2002: Phase 2 specification

???: Available for consideration for incorporation into the Single Unix Specification

More Information

Interconnect Software Consortium

- Web site: `http://www.opengroup.org/icsc/`
- Contact `dme@sun.com` or `ssharma@us.ibm.com` with concerns, questions, etc.

Infiniband Trade Association

- Web site: `http://www.infinibandta.org`