



vSMP Foundation
テクニカル・ワークショップ
スケーラブルシステムズ株式会社



vSMP Foundation

テクニカル・ワークショップ

vSMP Foundation では複数の物理サーバのリソースはアグリゲートされ、透過的に仮想マシンで使えるようになります。アグリゲートされた仮想マシンでは、構成する物理システムの数とそのリソースに応じた数多くのプロセッサ、I/Oデバイス、メモリフットプリントを、SMPシステムと同様の特性を持つVersatile SMPシステムとして利用可能です。

vSMP Foundationでは、NUMAと COMAを独自に組み合わせることで強力なキャッシュ技術を実現しています。このキャッシュラインに対して、異なるNUMAノードから書込みアクセスが行われた場合、または他のキャッシュラインのキャッシングによってキャッシュの追い出しが行われた場合は、そのキャッシュラインは最初のNUMAノードに対してローカルではなくなります。アプリケーションのスレッド、またはプロセス（以下、「タスク」）の割付けポリシーを慎重に選択することで、他の大規模マルチプロセッサ・システムと同様に、プロセッサとシステムのキャッシュメモリがアプリケーションのパフォーマンスを大幅に向上します。

スケーラブルシステムズでは、vSMP Foundationの利用技術に関するテクニカルワークショップで、効率的な利用方法をご紹介します。



目次/内容

目次	内容
①技術概要	vSMP Foundation による仮想化技術の概要とその仮想化によって実現するコンピューティングについての説明
②アプリケーション性能	vSMP Foundation によるスケーラブルSMPシステムでのアプリケーションの性能を紹介しています。
③導入・利用事例	導入事例とvSMP Foundation を導入しているユーザからの利用方法やアプリケーション性能に関する資料を紹介しています。
④アプリケーション実行ガイド	この資料ではオペレーティングシステムのタスク再配置ポリシーが、キャッシングのパフォーマンスに悪影響を及ぼす可能性があることに留意する必要があります。このため、タスクの再配置を最小限に抑えるためのタスク配置の方法などの事例をご紹介します。
⑤ サンプルプログラムと実行例	アプリケーション実行ガイドで紹介しましたプロセスの最適配置の実行例のプログラムと実行スクリプトのダウンロードが可能です。



ワークショップホームページ

[HOME](#) > [ソフトウェア](#) > [vSMP Foundation](#) > [テクニカル・ワークショップ](#)

テクニカル・ワークショップ - Mozilla Firefox

ファイル(E) 編集(E) 表示(V) 履歴(S) ブックマーク(B) ツール(I) ヘルプ(H)

テクニカル・ワークショップ

www.sstc.co.jp/biz/projects/vSMP_Foundation_Workshop.html

SSTC High Performance and Productivity スケーラブルシステムズ株式会社

お問い合わせ ▶ 検索

トップ 技術情報 プログラミング コンサルテーション 製品ホームページ サポート 会社案内 更新情報

HOME > ソフトウェア > vSMP Foundation > テクニカル・ワークショップ

Introducing Scalable Systems
FASTER
BETTER
AT LOWER COST

vSMP Foundation
テクニカル・ワークショップ
スケーラブルシステムズ株式会社

vSMP Foundation では複数の物理サーバのリソースはアグリゲートされ、透明的に仮想マシンで利用できるようになります。アグリゲートされた仮想マシンでは、構成する物理システムの数とそのリソースに応じた数多くのプロセス、I/Oデバイス、メモリー、プリント、SMPシステムと同様の特性を持つ Versatile SMPシステムとして利用可能です。

vSMP Foundationでは、NUMAとCOMAを独自に組み合わせることで強力なキャッシュ技術を実現しています。このキャッシュラインに付いて、異なるNUMAノードから読み込みアクセスが行われた場合、または他のキャッシュラインのキャッシュングによってキャッシュの読み出しが行われた場合は、そのキャッシュラインは最初のNUMAノードに付いてローカルではなくなります。アプリケーションのスレッド、またはプロセス（以下、タスク）の割付けポリシーを慎重に選択することで、他の大規模マルチプロセッサシステムと同様に、プロセッサとシステムのキャッシュメモリがアプリケーションのパフォーマンスを大幅に向上します。

スケーラブルシステムでは、vSMP Foundationの利用技術に関するテクニカルワークショップで、効率的な利用方法をご紹介します。ダウンロード【1.6MB】

このページに掲載してあります一部の資料のダウンロードにはパスワードが必要です。資料のダウンロードをご希望される場合には弊社にお問い合わせください。

vSMP Foundation アプリケーション・チューニングガイド ダウンロード【0.4MB】

vSMP Foundation Case Study (英文) (社外リンク)

vSMP Foundation を利用している事例を紹介しています。利用方法や利用している理由などを紹介しています。

vSMP Foundation Performance (社外リンク)

vSMP Foundation の性能については、vSMP Foundationの開発元である、ScaleMP社のホームページにも掲載されています。ここでは、商用アプリケーションや著名なオープンソースのアプリケーションのベンチマークを詳しく調査

資料	内容	ダウンロード
① 技術概要	vSMP Foundation による仮想化技術の概要とその仮想化によって実現するコンピュータ性能についての説明	
② アプリケーション性能	vSMP Foundation によるスケーラブルSMPシステムでのアプリケーションの性能を紹介しています。	
③ 導入・利用事例	導入事例とvSMP Foundation を導入しているユーザーからの利用方法やアプリケーション性能に関する資料を紹介しています。	
④ アプリケーション実行ガイド	この資料ではオペレーティングシステムのタスク再配置ポリシーが、キャッシングのパフォーマンスに悪影響を及ぼす可能性があることに留意する必要があります。このため、タスクの再配置を最小限に抑えるためのタスク配置の方法などの事例をご紹介します。	
⑤ サンプルプログラムと実行例	アプリケーション実行ガイドで紹介したプロセスの最適配置の実行例のプログラムと実行スクリプトのダウンロードが可能です。	

資料はホームページからダウンロード可能です。内容は適時更新しています。

Introducing Scalable Systems
FASTER
BETTER
AT LOWER COST

vSMP Foundation
技術概要
スケーラブルシステムズ株式会社

vSMP Foundation による仮想化環境
アプリケーション実行ガイド
スケーラブルシステムズ株式会社



アプリケーション実行ガイド

通常利用すべき手法

プリロードライブラリ : vsmpclib.so の利用

動的な配置をサポートする numabind コマンドの利用

プリロードライブラリ : hugetlb の利用

ジョブの
タイプ?

JAVA

スループット
(多重実行)

メッセージパッシング
MPI

マルチ
スレッド

シリアル
(逐次実行)

推奨オプショ

オプション

NUMA 向けのアプリケーション
適切なプログラミング方法が必要



アプリケーション向け 実行時注意点とサンプルスクリプト



vSMP Foundation でアグリゲートされた
仮想マシン用のアプリケーション実行ガイドライン

概要

vSMP Foundation は、複数の x86 ベース・システムを 1 つの論理サーバとして利用可能にする
トウェアです。これら物理サーバのリソースはアグリゲートされ、透過的に仮想マシンで使用でき
ります。アグリゲートされた仮想マシンでは、構成する物理システムの数とそのリソースに応じた
ロセッサ、I/O デバイス、メモリフットプリントを、SMP システムと同様の特性を持つ Versatile
テムとして利用可能です。本書では、vSMP Foundation が提供する仮想マシンを AVM (aggregated
machine: アグリゲートされた仮想マシン) として説明します。

vSMP Foundation の大きな特長は、高度なキャッシュ技術でシステムのパフォーマンスを強化
のローカルメモリの使用法にあります。アプリケーションのスレッド、またはプロセス（以下、「
割付けポリシーを慎重に選択することで、他の大規模マルチプロセッサ・システムと同様に、プロ
システムのキャッシュメモリがアプリケーションのパフォーマンスを大幅に向上します。vSMP F
は、マルチプロセッサ・システムに比べさらに強力なキャッシュ技術を実装しています。タスクを
体を通じて同じプロセッサ、システム、あるいは特定のシステム・セット上で継続的に実行する
Foundation のキャッシュ効果によってタスクのパフォーマンスが向上する可能性があります。

ScaleMP vSMP Foundation

MPIBLAST aggregation

Overview
MPI BLAST (Basic
BLAST program,
utilizing distribute
query segmentation
etc.).

This document describes
vSMP foundation,
integral part of the
<http://www.mpi-blast.org>

Please see the following
maximum performance

Building MPIBLAST
MPIBLAST 1.6.0
perform build and
As a part of building
command.

./configure --prefix

./configure --prefix

./configure --prefix

./configure --prefix

./configure --prefix

./configure --prefix

./configure --prefix

./configure --prefix

./configure --prefix

./configure --prefix

./configure --prefix

./configure --prefix

./configure --prefix

./configure --prefix

./configure --prefix

ScaleMP vSMP Foundation

Roche-454 / N aggregation

Overview

The Newbler is using
intensive phases.

Job Placement

As Newbler is using
while using the cpu
scheduler to place the

For example:
mnnactl --cpumodebin

Would place the job
the job would still be

Fast I/O

Newbler is using I/O
As such, it is recom

Use of Ramfs
In order to configure F

audio

audio

audio

audio

audio

audio

audio

audio

audio

audio

audio

audio

audio

audio

audio

audio

ScaleMP vSMP Foundation: Application Execution Guidelines

ScaleMP

Fluent 6.x and 12.x - Execution Guidelines for running applications in aggregated environment using ScaleMP's vSMP Foundation

Overview

Fluent is a multi-process application that uses MPI for inter-process communication. HP-MPI has been
set as the default MPI for the Fluent application. In addition Fluent supports MPICH2 as well.

While it is possible to run Fluent on the aggregation platform with any MPI implementation, using MPICH2
tuned for vSMP Foundation may yield a performance improvement of 5-15%.

Since there are some differences in the way versions 6.x and 12.x are launched, please see the relevant
section below.

Running Fluent 6.x with HP-MPI

HP-MPI has a built-in mechanism for assigning MPI processes to specific CPUs. Process placement is
controlled by environment variables named **MPI_BIND_MAP** and **MPIRUN_OPTIONS**. When these
variables are not set, process placement will not be performed.

Environment variables - HP-MPI

If you are running Fluent with HP-MPI, you should set the following environment variables prior to running
Fluent to yield the optimal performance:

```
export MPI_BIND_MAP=0,1,2,3,4,5,6,7 (For example)

export MPIRUN_OPTIONS="-cpu_bind=map_cpu,v"

export HMP_FRAGSIZE=131072

export MPI_SUBELEMENTS=16,24000000,4000000
```

MPI_BIND_MAP specifies a list of CPUs to which MPI ranks will be bound. You should replace the list
above with a list of integers, zero to 63. For more information on HP-MPI CPU affinity settings, refer to the HP-MPI user's guide available from
<http://docs.hp.com/en/B0000-95022/B0000-95022.pdf>.

Sample script for HP-MPI

```
export PATH=/opt/Fluent.lnx/bin:$PATH

export MPI_BIND_MAP=0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15

export MPIRUN_OPTIONS="-cpu_bind=map_cpu,v"

export HMP_FRAGSIZE=131072

export MPI_SUBELEMENTS=16,24000000,4000000

#Launches the sedan_4m benchmark with 16 processes

fluentbench.pl sedan_4m -t16 -mpi=hp
```

vSMP Foundation で利用されるアプリケーションの一部については、個別の利用ガイド
をご提供しています。掲載されていないアプリケーションの実行方法などについては、お
問い合わせください。



NUMA向けアプリケーション開発

- このワークショップの資料には、マルチスレッド・アプリケーション開発についての説明は含んでいません。
- アプリケーション開発などについては、「[マルチスレッド・アプリケーション開発のためのガイド](#)」などをご参照ください。vSMP Foundationでのアプリケーション開発については、NUMAメモリアーキテクチャ向けの最適化をご参照ください。
 - マルチスレッド開発ガイド：[3.5 NUMA 向けのアプリケーションの最適化](#)



アプリケーション開発

- NUMA 共有メモリー・アーキテクチャーにおいてパフォーマンスを管理する上で重要な点
 - プロセッサー・アフィニティー
 - 暗黙のメモリー割り当てポリシーによるデータの配置
 - 明示的なメモリー割り当て指示句によるデータの配置

NUMA Programming





ScaleMPホームページ

The screenshot shows the ScaleMP homepage with a navigation bar, a main banner, and several featured sections.

ScaleMP
CREATING THE POWER OF ONE

Products & Solutions | Appliances | Media-hub | Company | Support | Login

THE SCALE OF OUR SHARED-MEMORY SYSTEMS IS LIMITED TO 32,678 CPUS
Don't be hindered by lower limits
[Read More](#)

News & Events: Mar 18, 2013 - ScaleMP Expands Management Team [Click to beat SGI® performance](#)

- Large Memory Solutions**
Unlock Big Data
Reduce License Costs
- Shared Memory Solutions**
Accelerate Tech. Computing
Perform Complex Simulations
- On-Demand Solutions**
Increased Hardware Utilization
Cloud Based Flexible Datacenter
- Cluster Management**
Turbo-charge KVM Infrastructure
Simplify Cluster Management

ScaleMP is the leader in virtualization for high-end computing, providing higher performance and lower Total Cost of Ownership (TCO).
Tel: +1 (408) 342-0330
info@ScaleMP.com

Products & Solutions: Products, Solutions, Industries, Technology
Media-hub: News, Document Library, Events, Multimedia
Company: Company, Partners, Contact Us
Appliances
Support

© 2003-2013 ScaleMP, Inc. All rights reserved.

Document Library

[Data Sheets](#)

[White Papers](#)

[Partner Material](#)

[Customer Publications](#)

[Case Studies](#)

[Customer Quotes](#)

Logos: LBS, RWTH AACHEN UNIVERSITY, Institute for Technische Informatik

TraySearch: A ScaleMP Application Success Story

Dirk Schmidl, Christian Terboven, Dieter an Mey
{schmidl, terboven, anmey}@rz.rwth-aachen.de>
Center for Computing and Communication
RWTH Aachen University

14.10.2012, SC12, Salt Lake City, UT, USA

Foundation on SDSC's massive supercomputer

Lead Center

SDSC SAN DIEGO SUPERCOMPUTER CENTER
at the UNIVERSITY OF CALIFORNIA, SAN DIEGO UCSD

スケラブルシステムズ株式会社



vSMP Foundation活用方法・事例

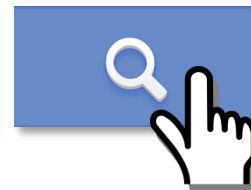
UFBS RWTH AACHEN UNIVERSITY Institut für technische Verbrennung

TraySearch: A ScaleMP Application Success Story

Dirk Schmidl, Christian Terboven, Dieter an Mey
{schmidl, terboven, anmey}@rz.rwth-aachen.de>
Center for Computing and Communication
RWTH Aachen University

14.10.2012, SC12, Salt Lake City, UT, USA

ScaleMP vSMP user guide



*Experiences with vSMP Foundation on SDSC's
Gordon data intensive supercomputer*

Robert Sinkovits
Gordon Applications Lead
San Diego Supercomputer Center



SDSC SAN DIEGO SUPERCOMPUTER CENTER
at the UNIVERSITY OF CALIFORNIA, SAN DIEGO **UCSD**

RWTH AACHEN UNIVERSITY

ScaleMP Applications

at RWTH Aachen University

Christian Terboven, Dirk Schmidl, Dieter an Mey
{terboven, schmidl, anmey}@rz.rwth-aachen.de



a HPC case
study with Dell

01.06.2010, ISC 2010
Hamburg, Germany



"LARGE MEMORY
SYSTEM WITH A
NEW LOOK"
by AMD, HP, and ScaleMP
White paper



AMD | **ScaleMP**
Creating the Power of One



お問い合わせ

0120-090715 

携帯電話・PHSからは（有料）

03-5875-4718

9:00-18:00 (土日・祝日を除く)

WEBでのお問い合わせ

www.sstc.co.jp/contact

この資料の無断での引用、転載を禁じます。

社名、製品名などは、一般に各社の商標または登録商標です。なお、本文中では、特に®、TMマークは明記していません。

In general, the name of the company and the product name, etc. are the trademarks or, registered trademarks of each company.

Copyright Scalable Systems Co., Ltd., 2005-2010. Unauthorized use is strictly forbidden.