THE HIGH-END VIRTUALIZATION COMPANY
SERVER AGGREGATION – CREATING THE POWER OF ONE

Corporate Presentation

Michael Schulman
mschulman@scalemp.com
Introduction to ScaleMP
Doing more

Big Data
Technical Computing
Hardware Utilization

Paying less

Hardware Cost
Software Licensing
Cluster Management
ScaleMP Fact-sheet

Founded in 2003
Proven and mature technology

Scalable Virtual Machine (VM)
SMP replacement driving application capability and system manageability

World’s largest SMP
32,768 CPUs & 256 TB RAM

Processor & interconnect agnostic
Intel/AMD x86, Intel Xeon Phi (1Q13)
Mellanox IB, Intel TrueScale IB (1Q13)

Product shipping since 2006
Over 4000 systems in hundreds of deployment across 30 countries

Channel-only business model

Creating the Power of One

N x Servers  N x OS

1 VM  1 OS

Over 4,000 systems in hundreds of deployment across 30 countries
Selected Customers and Applications

RDBMS & Analytics
- Actian VectorWise
- Oracle
- SAP HANA
- Sybase

Manufacturing
- Structural Mechanics
  - ABAQUS/Explicit
  - ABAQUS/Standard
  - ALTAIR Radioss
  - ANSYS Mechanical
  - LSTC LS-DYNA
  - NASTRAN
  - TNO Diana
- Fluid Dynamics
  - ANSYS CFX
  - ANSYS Fluent
  - ANSYS TGrid
  - AVL FIRE
  - EXA PowerFlow
  - EZNSS
  - GeoDict
  - MHD3D
  - NASA Cart3D
  - STAR-CCM+
  - STAR-CD
- Other
  - Comsol
  - inTrace OpenRT

Energy
- IMEX
- Norsar 3D
- Paradigm GeoDepth
- Schlumberger ECLIPSE

Finance
- KX
- Wombat

Bio-informatics
- 454/Newbler
- Abyss
- Bowtie
- CLC Bio
- FASTA
- HMMER
- Illumina
- mpiBLAST
- SOAPDenovo
- Velvet

EDA
- Cadence
- HSPICE
- Mentor
- Quartus
- Silvaco SmartSpice
- Synopsys

Weather Forecasting
- MITgcm
- MM5 (MPI & OpenMP)
- MOM4
- WRF

Numerical Simulations
- Octave
- R
- MathWorks MATLAB
- Wolfram Mathematica

…and many more homegrown applications

Life Sciences
- Comp. Chemistry
  - AMBER
  - CFOUR
  - DOCK
  - GAMESS
  - Gaussian
  - GOLD
  - NWChem
  - Octopus
  - OpenEye FRED
  - OpenEye OMEGA
  - Schrödinger Jaguar
  - Schrödinger Glide
  - SCM ADF
  - VASP

Molecular Dynamics
- GROMACS
- MOLPRO
- NAMD
- OpenEye ROCS
- Schrödinger Desmond
- Turbomole

Financial
- KX
- Wombat

…and many more homegrown applications

ScaleMP
Server Aggregation Virtualization
Server Virtualization

**PARTITIONING**
Subset of the physical resources

**AGGREGATION**
Concatenation of physical resources

Virtual Machines

- App
- OS

Hypervisor or VMM

Virtual Machine

- App
- OS

Hypervisor or VMM

- Hypervisor or VMM
- Hypervisor or VMM
- Hypervisor or VMM
- Hypervisor or VMM

vmware®  red hat®  CITRIX®  ORACLE®  Microsoft®  ScaleMP™
More Hardware Sales

Hardware Consolidation

- Lowered OPEX
  - Manage a single entity
- Reduced CAPEX
  - Scale up from industry standard servers
- Increased Performance
  - Fastest processors and memory management
- Greater Flexibility
  - Application-driven, on-demand SMP for mixed workloads

Application Consolidation

- Manage a single entity

Processing Consolidation

- Scale up from industry standard servers

Infrastructure Consolidation

- Application-driven, on-demand SMP for mixed workloads

Partioning - Subset of the physical resources

- Virtual Machines
  - App
    - OS
  - App
    - OS
  - App
    - OS

Aggregation - Concatenation of physical resources

- Virtual Machine
  - App
    - OS

ScaleMP
Solutions
vSMP Foundation - Solutions

**Single Operating System**
Cluster management and server consolidation

**Single (large) System**
Compute and memory demanding applications

**Single Infrastructure**
Cloud enabler – on-demand infrastructure

- **vSMP Foundation for Cluster**
  - Aggregated by: vSMP Foundation
  - OPEX savings

- **vSMP Foundation for SMP**
  - Aggregated by: vSMP Foundation
  - CAPEX savings

- **vSMP Foundation for Cloud**
  - Aggregated by: vSMP Foundation
  - Flexibility (CAPEX and OPEX savings)
vSMP Foundation - Solutions

**Single Operating System**
Cluster management and server consolidation
vSMP Foundation for Cluster

**Single (large) System**
Compute and memory demanding applications
vSMP Foundation for SMP

**Single Infrastructure**
Cloud enabler – on-demand infrastructure
vSMP Foundation for Cloud

Consolidation
IT Driven

Compute Intensive
Large Memory
I/O Intensive
Application Driven

vSMP Foundation - Solutions

Single System
Single OS
Single Infrastructure
Simplified Clustering:
- Single Operating System
- Storage:
  - Fast internal storage – excellent for scratch space
  - Eliminate the need for cluster file system
- Ease of programming: OpenMP, POSIX threads, MPI, PVM, Global Arrays, and any other common parallelism paradigms
- Versatile: Run jobs that require more memory than single physical server
- From unboxing to production in less than 3 hours

Before
- Difficult to manage
- Complex
- Clustered storage required

After
- Single OS instance
- Seamless integration
- Built-in storage

Aggregated by:
vSMP Foundation

OPEX savings
## Clusters vs. vSMP Foundation

<table>
<thead>
<tr>
<th></th>
<th>Clusters</th>
<th>vSMP Foundation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of Operating Systems</strong></td>
<td>Requires an OS per node</td>
<td>One OS only</td>
</tr>
<tr>
<td><strong>Scheduling</strong></td>
<td>Requires installation of job schedulers</td>
<td>OS scheduler is free and effective</td>
</tr>
<tr>
<td><strong>Work modes</strong></td>
<td>♦ Interactive is close to impossible</td>
<td>Interactive, batch, multi-user: all are ready out of the box</td>
</tr>
<tr>
<td></td>
<td>♦ Batch requires installation of batch schedulers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>♦ Multi-user requires installation of resource managers or job scheduler</td>
<td></td>
</tr>
<tr>
<td><strong>Installation</strong></td>
<td>Days, sometimes weeks</td>
<td>Within minutes from un-boxing to production, even for a novice</td>
</tr>
<tr>
<td><strong>I/O device consolidation</strong></td>
<td>♦ Scale I/O for parallel jobs only</td>
<td>♦ Provides scalable performance even for serial job</td>
</tr>
<tr>
<td></td>
<td>♦ Requires cluster file systems</td>
<td>♦ Readily available by the OS</td>
</tr>
<tr>
<td></td>
<td>♦ Requires additional hardware</td>
<td>♦ Requires no additional hardware</td>
</tr>
<tr>
<td><strong>SAN connectivity</strong></td>
<td>Requires an HBA per node, or front-end IO node serving the SAN data via network</td>
<td>Easily via low-number of HBA connections</td>
</tr>
<tr>
<td><strong>Parallel Programming Models</strong></td>
<td>Only distributed solutions: MPI, PVM</td>
<td>OpenMP, POSIX threads, MPI, PVM, Global Arrays, and any other common parallelism paradigms</td>
</tr>
<tr>
<td><strong>Shared memory</strong></td>
<td>Limited to RAM of a single node</td>
<td>Aggregate memory of all nodes</td>
</tr>
</tbody>
</table>
Highly Scalable Shared-memory System:
- Up to 256TB RAM
- Up to 32,768 CPUs
- Support for latest CPU architecture
- Ease of programming
- Best price-performance
Cost-effective Scaling

**Compute**

- Compute scaling with best-of-breed processor
- 35% to 60% cost savings compared to the alternatives

**Memory**

- Memory scaling without limits
- 35% to 50% cost savings compared to the alternatives
- Additional savings on socket-based application licenses

### List-price for 8 socket system
(processor and software cost)

- Westmere Systems (2.4, 10cores)
- Sandy-bridge systems (2.6, 8cores)
- vSMP Foundation - Sandy-bridge systems (2.6, 8cores)
- vSMP Foundation - AMD Opteron systems (2.6, 16cores)

### List-price for 4TB RAM
(processor, memory and software cost)

- 8 socket - Westmere Systems (2.4, 10cores)
- 16 socket Sandy-bridge systems (2.6, 8cores)
- 4 socket - vSMP Foundation - Sandy-bridge systems (2.6, 8cores)
- 4 socket - vSMP Foundation - AMD Opteron systems (2.6, 16cores)
Analytics and Big-data Solution

vSMP Foundation – Memory and I/O Extension Solutions

vSMP Foundation – Memory and I/O Extension
- Extension nodes could be connected directly (DC2) or thru a switch (up to 128 nodes total).
- Full mesh connectivity required for DC2 topology.
- Extension nodes could have different architecture than compute nodes.
- It is recommended to use low-cost processors in Extension nodes.
- I/O of extension nodes available for OS – for example can be used for fast or large storage pool.
- Target applications: analytics, in-memory databases, big-data, bio-informatics.

<table>
<thead>
<tr>
<th>4 socket</th>
<th>4 socket</th>
<th>4 socket</th>
<th>4 socket</th>
</tr>
</thead>
<tbody>
<tr>
<td>48 DIMMs</td>
<td>96 DIMMs</td>
<td>144 DIMMs</td>
<td>240 DIMMs</td>
</tr>
</tbody>
</table>

- Target analytics with integrated solution.
- Very large memory: - up to 7.5 TB RAM without InfiniBand switch
- up 48 TB RAM with InfiniBand switch
Unified Infrastructure for Mixed Workloads:
• On demand SMP
• No need for workload-dedicated systems:
  — Distributed applications use native nodes
  — Shared-memory use virtualized nodes
• Match workload requirements with system capabilities
• Provision and re-provision as needed:
• Integrated with world’s best provisioning systems: Bright Cluster Manager, HP CMU, IBM XCAT, IBM Platform HPC, ROCKS and others
Legacy vs. On-demand Datacenter

- **Mixed infrastructure:**
  - The best machine for each workload
  - One size doesn’t fit it all

- **High maintenance**

- **Complicated provisioning**

- **Can’t accommodate for changes in demand over time**

- **Budget bet!**
Legacy vs. On-demand Datacenter

- 25% more compute resources
- Smaller rack-space footprint
- Lower power consumption
- Pay as you grow
The Virtualized Datacenter

Legacy Datacenter
- 128 nodes cluster:
  - 128 x 16 (2,048) cores
  - 4GB RAM / core
- Shared memory system: 256 cores
- Fat node cluster: 4 x 512 (2,048) GB RAM

On-demand Datacenter
- 192 nodes cluster:
  - 192 x 16 (3,072) cores
  - 75%: 4GB RAM / core
  - 25%: 8GB RAM / core
- 756GB RAM VM
- 512GB RAM VM

- 256 cores VM
- Big Data VMs
- Compute VM

• Single infrastructure, many workloads. Allows for big-data problems in cloud environment
• Integrated with leading provisioning software.
On-demand Infrastructure

- **More capacity per $**
  - Single, cost-effective platform across the board

- **Increased utilization**
  - Shape the resource to the application usage model
  - Larger, unified and less-fragmented run-queues

- **Greater customer satisfaction**
  - On-demand resource allocation

- **Choice of provisioning system**
  - Adaptive Computing’s MOAB
  - Bright Cluster Manager Suite
  - HP’s Insight CMU
  - IBM’s xCAT
  - IBM Platform HPC
  - ROCKS
vSMP Foundation: Memory Expansion

Power Metric (QphH)

Capacity Scaling

Out of memory

- Raw dataset: 100 GB
  Query mem.: 43 GB
  Power Metric: 1.0
- Raw dataset: 200 GB
  Query mem.: 92 GB
  Power Metric: 2.0
- Raw dataset: 300 GB
  Query mem.: 138 GB
  Power Metric: 2.5
- Raw dataset: 400 GB
  Query mem.: 183 GB
  Power Metric: 3.0
- Raw dataset: 500 GB
  Query mem.: 226 GB
  Power Metric: 3.3
- Raw dataset: 600 GB
  Query mem.: 270 GB
  Power Metric: 3.5

Notes:
- Power Metric (QphH) indicates the performance metric.
- Capacity Scaling (Raw) and (Query memory) show the scaling for raw and query memory.
- Total OS memory: 4 * 96 GB * 90% = 345 GB RAM
- Scaling dataset up to 500 GB in memory (nominal query memory size: 45% of dataset)
Products Portfolio
Products Comparison

Flexibility

vSMP Foundation Advanced Platform

On-demand SMP

Up to 128 nodes

vSMP Foundation

Static VM

Up to 32 nodes

Virtual SMP

Single-rail InfiniBand

Node-locked license

Performance

Multi-rail InfiniBand

Floating license

Licensing

Detailed comparison: http://www.ScaleMP.com/compare
## Product Comparison

<table>
<thead>
<tr>
<th>Feature</th>
<th>vSMP Foundation™</th>
<th>vSMP Foundation™ Advanced Platform</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Highly scalable SMP</strong></td>
<td>Up to 32 nodes</td>
<td>Up to 128 nodes</td>
</tr>
<tr>
<td>with up to 32,768 cores and 256 TB RAM by virtualizing the aggregate resources of multiple nodes to a single VM</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tailor-made SMP</strong></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>allowing to select processor type (AMD or Intel), node-size (2-socket, 4-socket and 8-socket) and I/O capabilities (1GigE, 10GigE, Fibre-Channel, SAS, etc.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Built-in high-performance storage</strong></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>providing parallel, non-blocking scratch volume with linear scalability</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Unmatched RAS characteristics</strong></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>with node level fault isolation and automatic system recovery – providing maximum system uptime without user intervention</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Seamless InfiniBand failover</strong></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>for increased availability using active-passive dual-rail InfiniBand support</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pay as you grow</strong></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>allows for expansion with no price or performance penalties when more resources are needed</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>On-demand SMP</strong></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>allows repurposing cluster nodes in an ad-hoc basis. Readymade integrations with multiple provisioning systems: Bright Cluster Manager, Insight CMU, ROCKS and XCAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Active-active multi-rail InfiniBand</strong></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>for increased performance with up to four HCAs providing maximum of up to 224Gbps (FDR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SMP partitioning</strong></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>allows creation of multiple, isolated and independent smaller SMPs</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Shared InfiniBand fabric for multiple VMs</strong>, providing infrastructure cost savings</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td><strong>Licensing model</strong></td>
<td>Node-locked</td>
<td>Floating</td>
</tr>
<tr>
<td><strong>Per-socket pricing</strong></td>
<td>Starting at $400/socket</td>
<td>Starting at $800/socket</td>
</tr>
</tbody>
</table>

2/14/2013
VM on VM: Partition and Aggregate

Virtual Machine

App
OS

Virtualization for Aggregation

Virtual Machines

App
OS

App
OS

Hypervisor or VMM

Virtual Machines

App
OS

App
OS

App
OS

Hypervisor or VMM

Virtual Machines

App
OS

App
OS

App
OS

Hypervisor or VMM

Virtual Machines

App
OS

App
OS

App
OS

Hypervisor or VMM
COPROCESSOR ADD-ON

Traditional offload programming model

AGGREGATION

Unified processing and memory
THE HIGH-END VIRTUALIZATION COMPANY
SERVER AGGREGATION – CREATING THE POWER OF ONE

sales@scalemp.com