Application Consolidation Tutorial

Paper # 088
Wednesday, April 12th
8:00 - 10:00am

Hayden Brown
Senior Consultant
Consolidation
Advance Technology Center
Hewlett-Packard Co.
(613) 726-2191
hayden_brown@hp.com
Why Consolidation?
- The Business Issues

- It’s about controlling cost of management.
- While increasing the responsiveness of the lines of business.

Source: Gartner Group
Agenda

✓ Consolidation Process
  • Consolidation Platform
  • Consolidation Tools - ServiceControl
  • Application Consolidation
  • Case Studies
  • More Information/Contacts
Consolidation of computing systems reduces administration overhead, allows better utilization of computing resources, saves physical space, and reduces the burden of asset management. Consolidation requires careful planning, new strategies, and the right tools.
Agenda

- Consolidation Process
  - Consolidation Platform
    - HyperPlex
- Consolidation Tools - ServiceControl
- Application Consolidation
- Case Studies
- More Information/Contacts
Typical HyperPlex Solution Architecture

- **Virtual Backplane**
  - Primary
  - Secondary
- **Mass Storage Interconnect**
  - Primary
  - Secondary
- **Future SMP platforms**
- **Router(s)**
  - to clients
- **SMP**
  - V-Class
  - N-Class
  - D-Class
  - R-Class
- **SMP Management Station**
- **Model 30 disk array**
- **HP SureStore E Disk Array XP256**
- **FC/Model 12H disk array**
- **Future mass storage products**
- **SCSI MUX**
- **Tape library(ies) DLT 7000 or other mechanisms**

Backup - HyperPlex Router(s)
HP 9000 HyperPlex Solution Summary

**Central administration and management**
- Pre-integrated HP ServiceControl
  - ServiceControl Manager
  - Workload/Application control
  - Resource management
  - High Availability
  - Performance management tools

**Connectivity products**
- HP HyperFabric
- FDDI
- 10/100Base-T
- Gigabit Ethernet
- ATM

**Dense data center racking**
- Two V-Class stacked
- Four N-Class per rack
- Three K-Class per rack (field only)
- Five L-Class per rack
- Six D-Class per rack (field only)
- Six R-Class per rack

**Wide range of scalable multi-system mass-storage devices**
- EMC
- FC/Model 12H
- Model 30
- FC Arbitrated Loop hub
- FC SCSI multiplexer

**Support for multiple communications protocols**
- Standard TCP/IP
- Future, ultra-low-latency protocols

**Large performance capacity**
- Select from entire HP 9000 server family
- Mixed server model and HP-UX version
- Capacity to 64+ nodes (8,192 CPUs)
Agenda

- Consolidation Process
- Consolidation Platform
  ✔ Consolidation Tools - ServiceControl
    ‣ MC/ServiceGuard
    ‣ Process Resource Manager (PRM)
    ‣ HP-UX Workload Manager (HP-UX WLM)
    ‣ Memory Windows (MW)
- Application Consolidation
- Case Studies
- More Information/Contacts
HP 9000 ServiceControl Suite: Central Control and Capacity Planning

HP 9000 ServiceControl

Single Point, Multi-System Configuration Management
ServiceControl Manager (SCM)
SAM Ignite/UX SD/UX Online JFS
Secure Web Console Central Web Console
System Configuration Repository (SCR)

Fault Management
EMS EMS HA Monitors

Workload Management
HP-UX Workload Manager (HP-UX WLM) PRM
GlancePlus Pak MC/ServiceGuard
Web Workload Management
WebQoS

New: SCM in 2Q00
New: SCR in 1Q00
Enhanced PRM Dec 99
New: HP-UX WLM in 1Q00
Dynamic Application Rehosting within an HP HyperPlex with MC/ServiceGuard

- Minimize planned service interruptions
- Hardware/software upgrades or maintenance
- Re-Balance workloads and processing priorities with PRM
Benefits of MC/ServiceGuard

- Completely transparent to applications
- Intelligent cluster reconfiguration after node failure
  - Data Integrity: No 'split-brain' syndrome
  - Dynamic formation of new, viable cluster
- Flexible load balancing
- Mixed Series 800 class nodes
- Facilitates online hardware and software updates
- Highly available Enterprise Cluster
  - Fast switching of applications to alternate node (<60 seconds for basic system resources with JFS)
  - LAN failure protection (very fast local switch to standby LAN adapter inside same node)
- Application Packages
  - Easy application management
  - Flexible recovery options
- No idle resources
  - All systems run mission-critical applications
Workload Management

HP Process Resource Manager (PRM)

Allocate available processing resources according to business priority

- Provides greater control over the CPU by guaranteeing users, groups of users, or applications
  - Min/Max % of the CPU processing power
  - Min/Max % of Real Memory
  - Minimum % Disk I/O
- All entitlements are dynamically changeable
- Hard and soft CPU and memory limits
- GlancePlus integration allows automatic reconfiguration based on service level objectives
- Analysis tool, prm analyze, for resource accounting

- A must for every HP-UX server with 2 or more applications
- Enables a server consolidation strategy
- Compliments HP MC/ServiceGuard

Group 1
- 50% CPU
- 60% Real Memory

Group 2
- 25% CPU
- 20% Real Memory

Group 3
- 25% CPU
- 20% Real Memory
The PRM Advantage

- Each process gets equal priority - User 1 gets 66.6%
- UNIX Scheduler controls process priorities. Priority is lowered as processes consume more and more CPU
- CPU allocation in line with business priorities
- Provides a method to implement and manage service level objectives
- PRM entitlements determine % of 10ms execution timeslots
Using PRM with MC/ServiceGuard:
Consistent service delivery for critical applications

Strict response time requirement for Application C

Node 1

Applic C
100%

Applic A
80%

Applic B
20%

Node 2

If Node 1 fails, dynamically re-allocate processing resources of Node 2

- Dynamic allocation of CPU
- Load balancing for normal and post-failure operation
Application Consolidation

HyperPlex Solution Example: Initial State

(Multiple Apps. with High Availability)

Node 1
- App A
- App B
- App C
- App D
- App E

Node 2
- App F
- App G

Node 3
- App H
- App I
- App J
- App K
- App L

Node 4
- App M
- App N

Virtual Backplane

FC Switch

Disk Array

Small apps (5 per node) on Nodes 1, 3

Large apps (2 per node) on Nodes 2, 4
Application Consolidation

Hyperplex Solution Example: After Resource Balancing

(Multiple Apps with High Availability)

App D

Node 1

Virtual Backplane

Node 2

Node 3

Node 4

App A

App B

App C

App F

App G

App H

App I

App J

App K

App L

App M

App N

App E

App D Requires more Resources

FC Switch

FC Switch

Disk Array

Disk Array

Copyright Hewlett-Packard Co. 2000
Examples of Service Level Objectives (SLOs)

- Transactions will complete in less than 1 second.
- Batch job will finish in less than 1 hour.
- 50% of CPU allocation

Priority 1
Priority 2
Priority 3

*HP-UX WLM automatically reconfigures CPU entitlements depending on priority and set SLOs*
What is an SLO?

An SLO consists of a workload, a goal, and a priority for that goal.

- Transactions will complete in less than 1 second.
HP-UX Workload Manager (HP-UX WLM)

What else is in an SLO?

Application A

And...

Conditions:

- Mon-Fri, 9am-5pm

Constraints:

- No less than 10% CPU
- No more than 50%

Each goal has conditions under which it applies, and resource constraints.
HP-UX Workload Manager (HP-UX WLM)

How is goal achievement measured?

- Transactions will complete in less than 1 second.
- Collection of performance data via Application Response Measurement or non-invasive means.
- Batch job will finish in less than 1 hour.
- Collection of performance data from existing kernel instrumentation.
- 50% of CPU allocation
- No instrumentation required.

HP-UX WLM automatically reconfigures CPU entitlements depending on priority and set SLOs
HP-UX Workload Manager (HP-UX WLM)

Example uses of HP-UX WLM.

- OLTP Response-time goals
  - Ex: Online Retail price quote lookup in less than 2 seconds.

- Schedule PRM Entitlements.
  - Ex: End-of-month Financials.

- Event-based PRM entitlement change
  - Ex: Increase Oracle CPU entitlement by 5% for each additional user connection.

*Mix-and-match different goals on the same server. CPU allocation is driven by priority.*
What are Memory Windows?

- Memory Windows remove the 1.75 GB system-wide shared resources limit (for shared memory and memory-mapped files). Note: a per-process limit of 1.75 GB still exists for 32-bit apps.

- This was a major inhibitor to running multiple 32-bit applications on large memory systems such as V-class (16 GB real memory).

- With memory windows, there is one global shared resources window (max size 1.75 GB), and up to 128 private ones (max size 1 GB each).

- Provides isolation of shared resource regions across different applications.

- Very easy to implement using the "setmemorywindow" command.
Memory Usage with and without Memory Windows

Without Memory Windows

Overall system shared resources

With Memory Windows

Global Memory Region

Memory Window

Memory Window

Memory Window

Memory Window

Application

Application

Application
Agenda

- Consolidation Process
- Consolidation Platform
- Consolidation Tools - ServiceControl

✓ Application Consolidation
  - Hewlett-Packard’s Strategy
  - Types of Consolidation
  - Stacking Guidelines

- Case Studies
- More Information/Contacts
Select applications that will easily “stack” in order to...

- Reduce Risk
- Minimize Effort
Application Consolidation Strategy: Identifying Applications to Consolidate

1. Separate apps into groups by *Simple Factors*

2. Then split out groups by more *Complex Factors*

3. Deploy easy to stack applications
## Application Stacking Factors

<table>
<thead>
<tr>
<th>Technical Validation</th>
<th>Business Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Resource Consumption</td>
<td>• Organizational Boundaries</td>
</tr>
<tr>
<td>• System Impact</td>
<td>• Business Criticality</td>
</tr>
<tr>
<td>• Stability</td>
<td>• Security/Confidentiality</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Environment Commonality</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Operating System</td>
<td>• Vendor Support</td>
</tr>
<tr>
<td>• Production v. Development/Test</td>
<td>• Empirical Data</td>
</tr>
</tbody>
</table>
Types of Application Consolidation

- **Consolidation of a Distributed Client-Server or Multi-Tier Application**
  a) “Only change the hardware”
    - Run the application on fewer, larger systems
    - combine multiple instances onto one server
    - combine app servers and DB servers onto common servers
  b) Redundancy reduction
    - Reduce the number of instances of the application to process a given workload

- **Application Stacking**
  Consolidate multiple (diverse) applications on fewer, larger systems
Only Change the Hardware

Run the application on fewer, larger systems - without changing the application’s software architecture or number of application instances.

Before:
Four App Server Instances (ASI), 1 per box

ASI-1
ASI-2
ASI-3
ASI-4

DB Server

After:
Four App Server Instances, 2 per box.

ASI-1
ASI-2
ASI-3
ASI-4

DB Server
SAP Consolidation - example

Traditional:

- O/S Instance
- DB Instance
- Memory Window

- Shared Memory
  - App 1 / CI
  - O/S Instance
- Shared Memory
  - App 2
  - O/S Instance
- Shared Memory
  - App 3
  - O/S Instance

State of the art:

- DB Instance
- Memory Window

- Shared Memory
  - App 1 / CI
  - O/S Instance
- Memory Window
- Memory Window
- Memory Window

- Superior CPU Power
- RAM up to 32 GB

➔ Reduced complexity - reduced TCO

- Multiple OS to manage
- Up to 10% Network overhead
- Footprint
Redundancy Reduction

Reduce the number of instances of the application to process a given workload.

Before:
Four App Server Instances (ASI), 1 per box

After:
Two App Server Instances, 1 per box.
Application Stacking

Consolidate multiple (diverse) applications on fewer, larger systems

Before:

App A
App B
App C
App D

After:

Apps A, B, C, D

Four different applications on four servers (1 per box) to four different applications all on one server.
Guidelines for Running Multiple Applications per Server

- It is generally safe to run multiple instances of the same application on one OS image (e.g. multiple instances of Oracle, or multiple instances of SAP).

- It is often safe to combine application servers and database servers onto a common server - example: SAP and Oracle; Oracle Financials and Oracle Database.

- Applications that are intensive on different resources (e.g. CPU intensive vs I/O intensive) or at different times (example Batch and OLTP) are good consolidation candidates.
Guidelines - continued

- Firewall/security products require dedicated servers.

- Old home grown applications are generally not the first choice of application consolidation candidates. But ...

- It is generally not recommended to share servers across business units unless there is high level management commitment and sponsorship.

- There is NO substitute for testing.
Guidelines—continued

- Analyze Application Characteristics

- Making Them Work Together
  - ServiceControl - MC/ServiceGuard, PRM, Memory Windows

- Sizing System Hardware
  - HP Measureware Data for CPU, Memory, Disk I/O, Networking

- Kernel Tuning

- Security
Agenda

• Consolidation Process
• Consolidation Platform
• Consolidation Tools - ServiceControl
• Application Consolidation
✓ Case Studies
• More Information/Contacts
TCO for Systems Consolidation - Case Study

**EXISTING Environment**
- Entire environment: 56 Servers
- Chosen candidates: 43 Servers
- Entirely purchased, nothing financed
- High content of personnel cost
- Existing monthly cost of use = 382K$

**Phase I PROPOSED Environment**
- 17 existing Servers plus 6 new N-classes plus upgrades
- Finance proposal on 36 month lease
- 13.4% lower operational costs incl. start-up costs (excluding personnel savings)
- Future monthly cost of use = 331K$

**Phase II PROPOSAL**
- Decrease System Administration resources (26 to 15 people) expected savings 37.8%
- Future monthly cost of use = 248K$
**Situation**

- IT realized server proliferation was getting out of hand
- IT decided to consolidate servers and mass storage

**Solution: “the Compute Utility”**

- Two to three Oracle instances are stacked on a K-Class server where an EMC enclosure provides disk storage
- PRM guarantees CPU percentage for each Oracle instance
- IT sells CPU percentage and disk storage to the business units
Agenda

- Consolidation Process
- Consolidation Platform
- Consolidation Tools - ServiceControl
- Application Consolidation
- Case Studies

✔ More Information/Contacts
More Information

• White Papers on WEB Sites:
  – Memory Windows in HP-UX 11.0, White Paper
    • http://docs.hp.com/hpux/os/#papers
  – Using HP PRM with Databases
    • http://docs.hp.com/hpux/ha/#papers
  – MC/ServiceGuard Documentation
    • http://docs.hp.com/hpux/ha/#doc

• HP’s Consolidation WEB Site:
  – http://hp.com/go/consolidation

• Contact your local HP Sales Representative or Reseller

• Contact Us!
Dear HP Customer!

Would you like to influence new and existing HP products and solutions? Is your company implementing DataWarehousing or Server Consolidation?

If so, HP’s Business Critical Computing R&D labs might want to invite you to visit our site in Cupertino so that:

• Our R&D community can learn more about your company and how you use HP’s DataWarehousing and Consolidation solutions.
• You can provide us with input on our products.

If you are interested, please contact:

Carolyn Godfrey, 408-447-4074, carolyn_godfrey@hp.com