Better TCO & Higher Availability than Oracle RAC with HP NonStop

Azucena Ubierna
HP NonStop PreSales
GTUG 2015
azucena.ubierna@hp.com
How should we build Mission Critical Infrastructure?

Business is extremely cost sensitive

Open Source SW is cheap .. even under EE distributions

Commercial SW focuses on Always On features

Growing demand for Always ON

Standard servers grow on RAS features

Storage Subsys is the Corp Data “Safe Box”
- Mission Critical

Maximize Data Protection

Minimize Downtime
- Mission Critical
Impact of Failures - IDC Availability Spectrum

HP NonStop systems are AL4

NonStop delivers

**Impact of component failure**

- **Availability Level 4**
  - Switch to alternate resources is not perceptible to end users

- **Availability Level 3**
  - Short outage is needed for failover to take place

- **Availability Level 2**
  - Balancing may not be perceptible to end-users because of retry

- **Availability Level 1**
  - Need to switch to redundant resources before processing resumes.

### IDC Availability Level 4

**SMP Enterprise Servers**
- **Superdome**
  - Server built for resiliency.
  - One big SMP box but can have isolated failure areas.
  - IDC Level 4
  - Server built for resiliency.
  - Two big SMP boxes and SW is resiliency aware.

- **Sysplex**
  - Server built for scalability.
  - Several SMP boxes to run the DB SW & serveral SMP boxes to drive the access to storage.
  - IDC Level 4
  - IF used in PAIRs IDC Level 4

- **Exadata**
  - Full SW stack is HW & SW Fault Tolerance aware.
  - IF used in PAIRs IDC Level 4

**Fault Tolerant Servers**
- **MPP HP NonStop**
  - Server built for Fault Tolerance.
  - Full SW stack is HW & SW Fault Tolerance aware.
  - IDC Level 4

- **Stratus, NEC..**
  - Server built for Fault Tolerance.
  - OS SW is HW Fault Tolerance aware.
  - Guest SW is not Fault Tolerance aware.
  - IDC Level 4

**Standard Server Clusters**
- **Virtual Servers**
  - Most vendors provide High Availability & Fault Tolerant options.
  - Same names but not necessarily same meanings.

- **Stand by Clustering**
  - Servers built for standard resiliency.
  - Two SMP boxes, one active and one standby.

- **Farm Clustering**
  - All active boxes.
  - When failure happens activity needs to be redirected.
## Ensuring Data Protection & Data Recovery Capability

### Access Granted to All Committed Business Operations

<table>
<thead>
<tr>
<th>DB Coherence &amp; Logs</th>
<th>Storage System</th>
<th>Back Up Solution</th>
<th>DR Site - RTO/RPO requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACID DBMS</td>
<td>ESS</td>
<td>RAID</td>
<td>Cold</td>
</tr>
<tr>
<td>BASE DBMS</td>
<td>RAID</td>
<td></td>
<td>Warm</td>
</tr>
<tr>
<td></td>
<td>ESS</td>
<td>Simple Tape Drives sized for the required duty along with Catalog Management SW</td>
<td>Hot</td>
</tr>
<tr>
<td></td>
<td>RAID</td>
<td>Compulsory for some clustered HA options: - Stand By - Multi node RDBMS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Standalone ESS Or with Sync Replicated ESS</td>
<td>Directly Attached to Standalone Fault Tolerant Servers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Directly Attached to Standard servers in Clustered Scale-Out &quot;RAID&quot; deployments</td>
<td>Directed Attached to Standalone Fault Tolerant Servers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RAID</td>
<td>Directly Attached to Standalone Fault Tolerant Servers</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Full integrated Enterprise Backup Systems that may include encryption, de-duplication, replica...</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RAID</td>
<td></td>
<td>RTO = 0 &amp; RPO = 0</td>
</tr>
<tr>
<td></td>
<td>RAID</td>
<td>Simple Tape Drives sized for the required duty along with Catalog Management SW</td>
<td>RTO = 0 &amp; RPO &gt; 0</td>
</tr>
<tr>
<td></td>
<td>RAID</td>
<td>Compulsory for some clustered HA options: - Stand By - Multi node RDBMS</td>
<td>RTO = 0 &amp; RPO &gt; 0</td>
</tr>
<tr>
<td></td>
<td>RAID</td>
<td>Directly Attached to Standalone Fault Tolerant Servers</td>
<td>Minutes &lt; 10</td>
</tr>
<tr>
<td></td>
<td>RAID</td>
<td>Full integrated Enterprise Backup Systems that may include encryption, de-duplication, replica...</td>
<td>Minutes &lt; 10</td>
</tr>
<tr>
<td></td>
<td>RAID</td>
<td></td>
<td>RTO = 0 &amp; RPO &gt; 0</td>
</tr>
<tr>
<td></td>
<td>RAID</td>
<td>Compulsory for some clustered HA options: - Stand By - Multi node RDBMS</td>
<td>RTO = 0 &amp; RPO &gt; 0</td>
</tr>
<tr>
<td></td>
<td>RAID</td>
<td>Directly Attached to Standalone Fault Tolerant Servers</td>
<td>Minutes &lt; 10</td>
</tr>
<tr>
<td></td>
<td>RAID</td>
<td>Full integrated Enterprise Backup Systems that may include encryption, de-duplication, replica...</td>
<td>Minutes &lt; 10</td>
</tr>
<tr>
<td></td>
<td>RAID</td>
<td></td>
<td>RTO = 0 &amp; RPO &gt; 0</td>
</tr>
</tbody>
</table>

### DB Coherence & Logs
- Atomicity
- Consistency
- Isolation
- Durability
- Extensive Sets of Tools for Dumps, Replica and Recovery

### Storage System
- Standalone ESS
  - Or with Sync Replicated ESS
  - Typically left to whatever the underlying FS provides.
  - Cold snapshots.
- Directly Attached to Standalone Fault Tolerant Servers
- Compulsory for some clustered HA options:
  - Stand By
  - Multi node RDBMS

### Back Up Solution
- Simple Tape Drives sized for the required duty along with Catalog Management SW
- Fully integrated Enterprise Backup Systems that may include encryption, de-duplication, replica...

### DR Site - RTO/RPO requirements
- Cold: RTO >> 0 & RPO >> 0 Minutes >> 10
- Warm: RTO > 0 & RPO > 0 Minutes < 10
- Hot: RTO = 0 & RPO = 0

*Based on:
- Stand By resources
- DB Restore
- Restart of Services

© Copyright 2014 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.
Processing Power Scalability & DATA Scalability

Scalability

SMP

Scale UP
Mainframe Applications
"Standard" Applications
Mainframe

Scale OUT
Stateless Applications
Superdome

MPP

External Interconnect
Cluster aware ACID DBMS
Oracle RAC
Cluster aware BASE DBMS
Teradata Netezza

Internal Interconnect
Big Data Oracle App
Oracle App
MPP aware applications
HP NonStop

Big Data

Oracle App

Exadata

© Copyright 2014 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.
Scalability & Availability

Scalability

SMP
- Scale UP
  - Mainframe Applications
  - Parallel Sysplex
  - Fault Tolerance Managed by OS

- "Standard" Applications

- Stateless Applications

- Cluster aware Applications

MPP
- Scale OUT
  - Cluster aware ACID DBMS
  - Cluster aware BASE DBMS

- External Interconnect

- Internal Interconnect
  - Big Data Apps
    - Oracle RAC
  - NonStop Applications & DB
    - Teradata
    - Netezza
    - Exadata * 2
      - Oracle RAC Managed HA
  - HP NonStop
    - Fault Tolerance Managed by OS

Stateless
- Applications

Parallel
- Sysplex

Fault Tolerance
- Managed by OS

Cluster aware
- Applications

ACID
- DBMS

BASE
- DBMS

Oracle RAC

Oracle RAC
Managed HA

Managed HA
Availability, Data Integrity & Scalability
Design Principles and a few “Not-So-Easy-Tasks”

**Eliminate Single Points of Failure**
- Usage of redundant resource
- Smart provisioning

**Minimize the Impact of Failures**
- Smaller areas of failure

**Build Fault Tolerance for each and every layer**
- “All Active” Clustering
- “Stand-by” Clustering

**Data Preservation and Recovery**
- ACID properties preservation
- Database Backup
- Online Replica

**Data Scalability**
- Multi-node Database Pros & Cons for OLTP

© Copyright 2014 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.
### Mission Critical Infrastructure & Features Summary

<table>
<thead>
<tr>
<th>Scalability</th>
<th>High Availability</th>
<th>Data Integrity</th>
<th>Recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMP HW</td>
<td>Fork Lift &amp; Node Addition</td>
<td>Multiplicity</td>
<td>Platform Dependent</td>
</tr>
<tr>
<td>NonStop MPP HW</td>
<td>Up Lift &amp; Incremental</td>
<td>Built-in</td>
<td>At every component</td>
</tr>
<tr>
<td>Linux OS - Scale UP</td>
<td>Passive – NUMA constrains</td>
<td>Stand By Clustering</td>
<td>Passive</td>
</tr>
<tr>
<td>Linux OS – Scale OUT</td>
<td>Passive – App/DB dependent</td>
<td>SL Retry or Clustering</td>
<td>Passive</td>
</tr>
<tr>
<td>NonStop OS</td>
<td>Active</td>
<td>Active</td>
<td>End-To-End</td>
</tr>
<tr>
<td>Oracle S1, SE, EE</td>
<td>One Node (2, 4, N sockets)</td>
<td>Stand By Clustering</td>
<td>Redo logs</td>
</tr>
<tr>
<td>Oracle EE RAC</td>
<td>Multi Node (3 nodes max ?)</td>
<td>RAC</td>
<td>Multi - Redo logs</td>
</tr>
<tr>
<td>NonStop SQL/MX</td>
<td>16 Blade units per Node</td>
<td>Built-In</td>
<td>TMF – Audit Trail</td>
</tr>
<tr>
<td>Application</td>
<td>HW Aware, State Less, ??</td>
<td>HW Aware, State Less, ??</td>
<td>DB Dependent</td>
</tr>
</tbody>
</table>
Provisioning IT Resources for Mission Critical OLTP Business

Some Relevant Costs Added

• Business Apps:
  • Resilient & Scalable SW
• Operations:
  • MC Staff,
  • Change Management …
• HW & SW Infrastructure
  • Processing Capacity
  • Storage Resources
  • Basic SW Stack
  • Data Base Services

Some Extra Payback Expected

• HW Infrastructure & Basic SW stack:
  • High Availability
  • Capability to Scale & Perform
• Data Base Services:
  • Data Integrity
  • Capability to Scale & Perform
  • Data Protection & Recovery
What are we comparing?
Scale-out SMP servers & MPP NonStop servers

Two approaches to Availability, Scalability & Data integrity

Load Balancer

Application Nodes: from 2 to N

RAC Database Nodes: from 2 to ..?

External Storage Subsystem

External Storage Subsystem

MPP Architecture
NonStop OS
NonStop SQL/MX

Integrated Storage Subsystem
Integrated Storage Subsystem

DB App DB App DB

Stateless or Cluster Aware

Interconnect & Coherence
Real Application – Entry Systems
NonStop MPP vs
SMP Linux + Oracle
Pure StandBy Clustering at Each Site
ISV suggested configuration includes DB Replica to DR Site

- **Primary Site**
  - IP Comms
  - Clustering SW
    - App Server 2
  - Clustering SW
    - App Server 1
  - Clustering SW
    - RDBMS
      - DB Server 1
  - FC or SAS
    - HA Storage SubSystem
      - App FS
      - LUN
      - Data Base
  - Active DB Node
    - Active DB Node @ Primary Site
    - Replicates to
    - Active DB Node @ DR Site

- **DR Site**
  - IP Comms
  - Clustering SW
    - App Server 2
  - Clustering SW
    - App Server 1
  - Clustering SW
    - RDBMS
      - DB Server 2
  - FC or SAS
    - HA Storage SubSystem
      - App FS
      - LUN
      - Data Base
  - Active DB Node
    - Active DB Node @ DR Site
HP c7000 Proliant BL460c 1 Proc q Core & 3PAR 7200

Specific ISVs 100 to 200 tps rate - Generic

- **OS, App, DB**
  - RHEL Lic E-LTU
  - (No Compilers included)
  - ServiceGuard for Linux Advanced Edition
  - Oracle EE, Adv Sec & RAC

- **Operations & Management**
  - HP Insight Control E-LTU
  - HP 3PAR 7200 Reporting Suite E-LTU

- **Processing Power DB & App**
  - 2 HP BL460 Gen8 1P 4 core 32GB RAM
  - 2 HP BL460 1P 4 core 32GB RAM

- **Storage Subsystem**
  - 6 HP 82B PCIe 8Gb FC Dual Port HBA
  - 2 HP Virtual Connect FlexFabric, 10Gbps/24p
  - 1 HP 3PAR StoreServ 7200 2-N Storage Base
  - 24 HP M6710 300GB 6G SAS 10K HDD

- **Communication Subsystem**
  - 2 HP Virtual Connect FlexFabric, 10Gbps/24p

- **Rack Elements**
  - HP BLc7000 CTO 3 IN LCD Plat Enclosure
  - HP 642 1075mm Shock Intelligent Rack
  - HP R5KVA UPS 3U IEC309-32A HV Intl Kit
HP Integrity NonStop NS2302

Extrapolation for a 100 tps rate from an ISV/Hp Benchmark

- **OS, App, DB**
  - NONSTOP OS, MISSION CRITICAL EDITION
  - NS OS SECURITY ENHANCEMENT J SER
  - NONSTOP TS/MP (ACS)
  - C/C++ COMPILER, TNS/E NATIVE
  - NONSTOP SERVER FOR JAVA 6.0
  - NONSTOP JDBC TYPE 4 DRIVER-100+USERS
  - NONSTOP SQL/MX ENGINE & TABLE

- **DR Synchronization**
  - REMOTE DATABASE FACILITY/IMPX
  - NONSTOP AUTOSYNC SOFTWARE
  - EXPAND

- **Operations, Management & Performance**
  - WEB VIEWPOINT
  - MEASURE
  - I/O ESSENTIALS - WINDOWS CONSOLE & HOST
  - NONSTOP CLUSTER ESSENTIALS-WINDOWS & HOST
  - NS CLUSTER ESSENTIALS - BLADES HOST &
  - NS PERF ESSENTIALS-HP SIM CLIENT & HOST
  - NS SOFTWARE ESSENTIALS R2 - WIN CONS & HOST

- **Processing Power**
  - 2 HP NONSTOP CPUs (Pulson based)
  - 1 Core CAPABLE
  - 32GB RAM per CPU

- **Storage Subsystem**
  - 26 146 GB 15K SAS 6G SFF HDD
  - 2 HP NONSTOP SAS HBA 6G R3
  - 2 HP NONSTOP STORAGE CLIM R4
  - 2 SAS 6G SFF ENCLOSURE

- **Communication Subsystem**
  - VIO Pair with VIO extension
  - 2x6 Ethernet ports

- **Rack Elements**
  - 1 HP NONSTOP SINGLE PHASE UPS INTL R2
  - 1 NONSTOP EXTENDED RUN TIME MODULE R2
  - Rack Mounted Console
From Pure StandBy Clustering to Cluster AWARENESS

Building Resiliency for DB nodes failures
NonStop to Linux 3y TCO comparison
Real Application – High End
NonStop MPP vs
SMP Linux + Oracle
From NO HA to Pure StandBy Clustering

Building Minimal HA Features on top of ISV suggestion
HP Proliant DL385p clusters & 3PAR 7200

Specific ISVs 300 tps rate for Specific Customer

- **OS, App, DB**
  - RHEL Lic E-LTU
  - (No Compilers included)
  - Oracle Min, Oracle EE and Oracle RAC

- **Operations & Management**
  - HP Insight Control E-LTU
  - HP 3PAR 7200 Reporting Suite E-LTU

- **Processing Power DB & App**
  - 2 HP DL385p Gen8 8-SFF CTO Server, 128 GB RAM
  - 1 HP DL385p Gen8 8-SFF CTO Server, 64 GB RAM(*)

- **Storage Subsystem**
  - 6 HP 82B PCIe 8Gb FC Dual Port HBA
  - 2 HP 8/24 Base 16-ports Enabled SAN Switch
  - 1 HP 3PAR StoreServ 7200 2-N Storage Base
  - 8 HP M6710 300GB 6G SAS 15K 2.5in HDD
  - 8 HP M6710 100GB 6G SAS 2.5in SLC SSD

- **Communication Subsystem**
  - 3 HP Ethernet 10GbE S30FLR-SFP+ FIO Adptr
  - 3 HP Ethernet 10Gb 2P 530SFP+ Adptr

- **Rack Elements**
  - HP 642 1075mm Shock Intelligent Rack
  - HP RSKVA UPS 3U IEC309-32A HV Intl Kit
HP Integrity NonStop Blade System NB54004

Extrapolation for a 300 tps rate from Specific Customer Production Environment

• OS, App, DB
  • NONSTOP OS, MISSION CRITICAL EDITION
  • NS OS SECURITY ENHANCEMENT J SER
  • C/C++ COMPILER, TNS/E NATIVE
  • NONSTOP SQL/MX ENGINE & TABLE

• Operations, Management & Performance
  • MEASURE
  • I/O ESSENTIALS - WINDOWS CONSOLE
  • I/O ESSENTIALS - BLADESYSTEM HOST
  • NONSTOP CLUSTER ESSENTIALS-WINDOWS
  • NS CLUSTER ESSENTIALS - BLADES HOST
  • NS PERF ESSENTIALS-HP SIM CLIENT
  • NS PERFORMANCE ESSENTIALS-NSK HOST
  • NS SOFTWARE ESSENTIALS R2 -WIN CONS
  • NS SOFTWARE ESSENTIALS R2 NB HST

• Processing Power
  • 4 HP NONSTOP CPUS (Tukwila based)
  • 4 Core CAPABLE BLADE
  • 48GB RAM per BLADE

• Storage Subsystem
  • 44 146 GB 15K SAS 6G SFF HDD
  • 2 HP NONSTOP SAS HBA 6G R3
  • 2 HP NONSTOP STORAGE CLIM R4
  • 2 SAS 6G SFF ENCLOSURE

• Communication Subsystem
  • 2 HP NONSTOP IP CLIM 5C R4
  • 2 HP NONSTOP IP CLIM 3C/2F R4

• Rack Elements
  • 2 HP NONSTOP SINGLE PHASE UPS INTL R2
  • 1 NONSTOP EXTENDED RUN TIME MODULE R2
  • Rack Mounted Console
# Sizing for the Different tps rates on both Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Configuration</th>
<th>tps Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>NB54014</td>
<td>32GB/cpu, 100 HDD (14G)</td>
<td>1200</td>
</tr>
<tr>
<td>NB54008</td>
<td>48GB/cpu, 44 HDD (300)</td>
<td>600</td>
</tr>
<tr>
<td>NB54004</td>
<td>48GB/cpu, 44 HDD (14G)</td>
<td>300</td>
</tr>
<tr>
<td>1*DL580</td>
<td>2*DL980, HP 3PAR 7000</td>
<td>1200</td>
</tr>
<tr>
<td>1*DL385p</td>
<td>2*DL580, HP 3PAR 7000</td>
<td>600</td>
</tr>
<tr>
<td>1*DL385p</td>
<td>2*DL385p, HP 3PAR 7000</td>
<td>300</td>
</tr>
</tbody>
</table>

(*) Just one App server in the picture and quotes
NonStop to StandBy Linux - 3 year TCO comparison -

Stand-by Failover TIME

DB recovery UNCERTAINTY

DB Replica – DataGuard
Not possible without EE

© Copyright 2014 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.
From Pure StandBy Clustering to Cluster AWARENESS

Building Resiliency for DB nodes failures
NonStop to Linux RAC DB - 3 year TCO comparison -

300 tps rate

600 tps rate

1200 tps rates

© Copyright 2014 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.
Real Application – High End NonStop MPP vs SMP Linux + Oracle

Update NonStop X & Proliant Gen 8/9 servers
NonStop X to Linux RAC DB - 3 year TCO comparison -

300 tps rate  600 tps rate  1200 tps rate

© Copyright 2014 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.
Comments please?
&
Thank you

In case you need further details, pls do not hesitate and contact the NonStop PreSales team