Cluster Extension on Windows

- Windows 2000
  - Advanced Server
  - Datacenter Edition
- Windows 2003 (32/64bit)
  - Enterprise Edition
  - Datacenter Edition

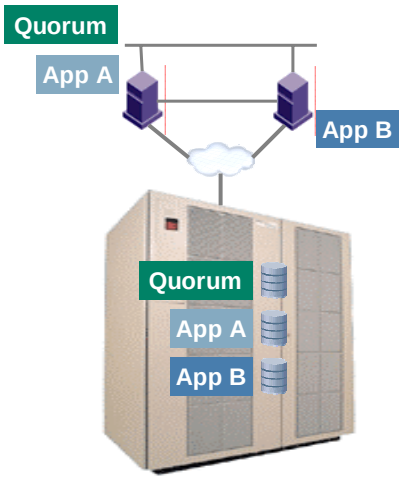
Anton.Vogel@hp.com
Storage Consultant

© 2004 Hewlett-Packard Development Company, L.P.
The information contained herein is subject to change without notice

Agenda

- Identifying the building blocks of CLX on Windows
  - How does a local cluster with a single Quorum work?
  - A new cluster model: How do Majority Node Sets work?
  - What is needed for a distributed environment?
  - Parameter Settings for CLX Resources

Local Microsoft Cluster - using a single Quorum

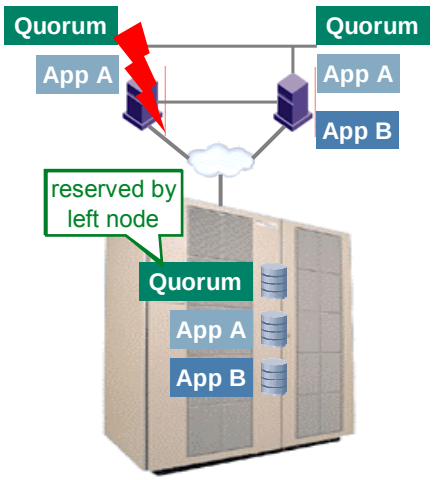


Shared disks are used
- to keep the application data

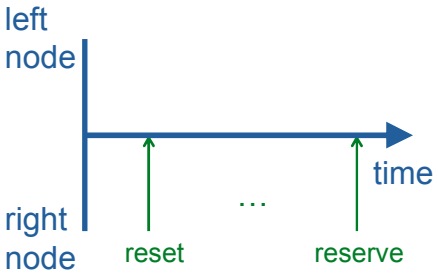
The Quorum disk is used
- to keep the Quorum-Log
- to keep a copy of the Cluster configuration
- to propagate Registry Checkpoints
- for arbitration, if LAN-connectivity is lost

4/22/2004 3

Local Microsoft Cluster - using a single Quorum

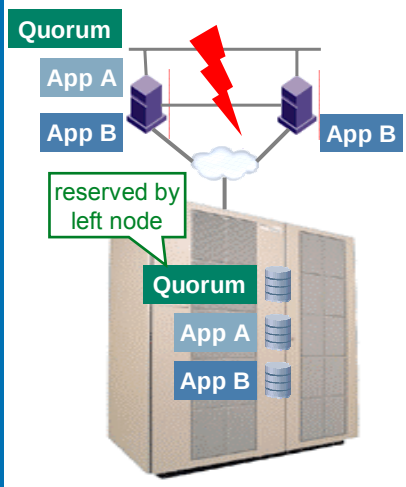


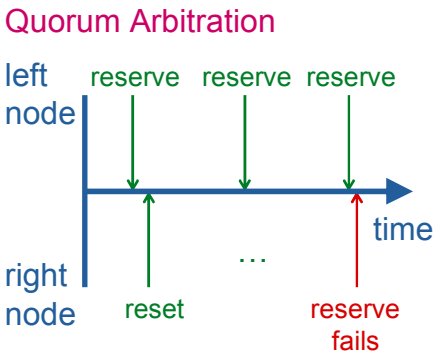
Quorum Arbitration



4/22/2004 4

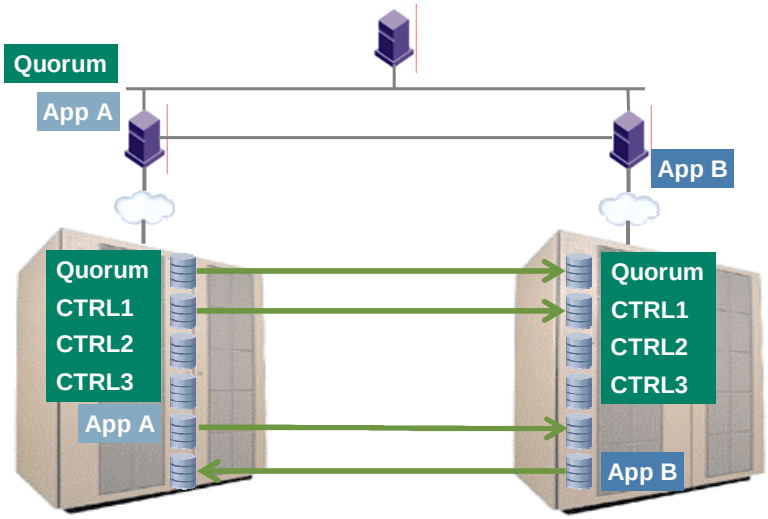
Local Microsoft Cluster - using a single Quorum





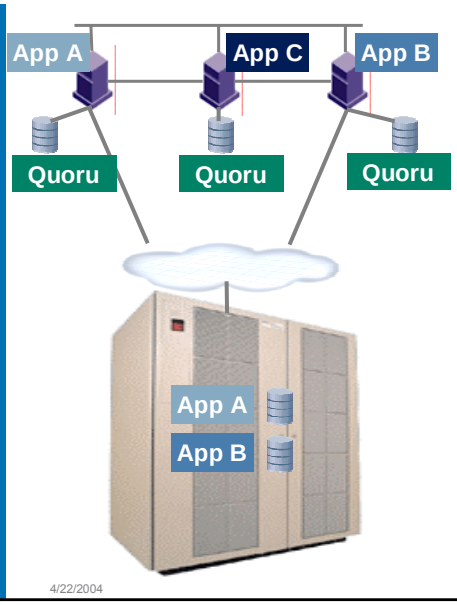
4/22/2004 5

Cluster Extension on Windows - using a distributed quorum



4/22/2004 6

Microsoft Cluster with Majority Node Sets



Shared disks can be used

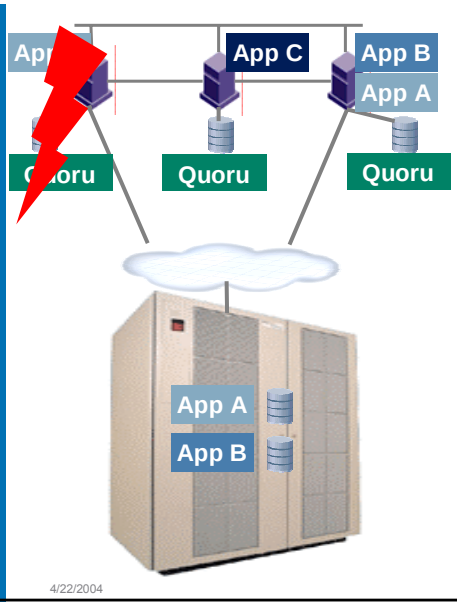
- to keep the application data, but they might not be necessary for all applications

The Quorum disks are

- used to keep a copy of the Cluster configuration
- sychronized by the Cluster Service

4/22/2004 7

Microsoft Cluster with Majority Node Sets



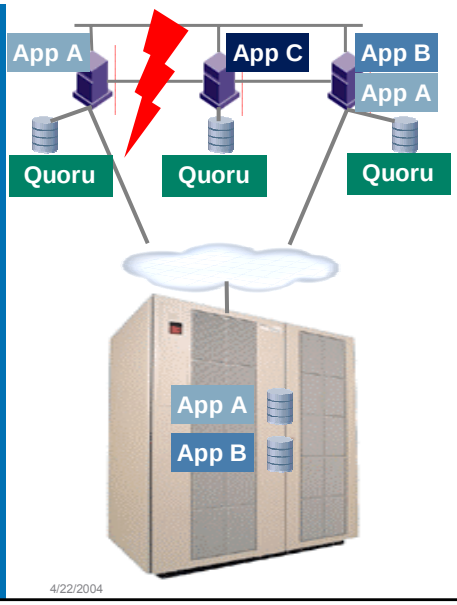
- In case of a failure, the cluster will survive, if more than 50% of the nodes are still available
- In case of a split site situation, the site with more than 50% of the nodes will survive

- No common quorum log and no common cluster configuration available => changes in the cluster configuration are only allowed, when all nodes are online AND LAN-connectivity is available for all surviving nodes!

- Microsoft suggests to use this model for geographically dispersed clusters

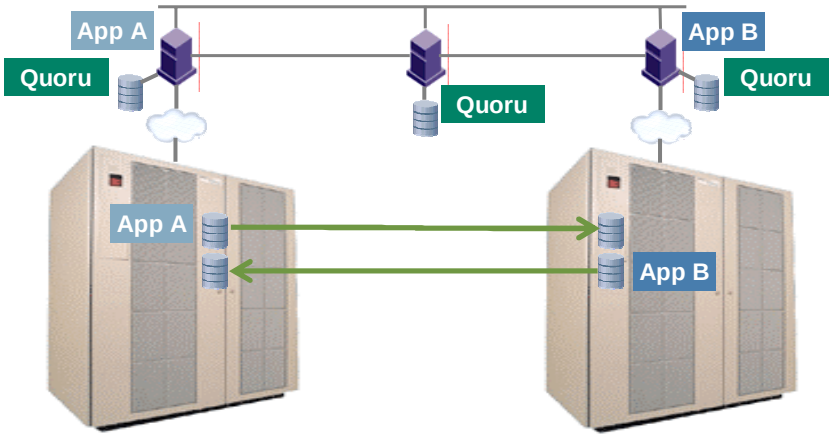
4/22/2004

Microsoft Cluster with Majority Node Sets



4/22/2004 9

Cluster Extension on Windows - using Majority Node Sets



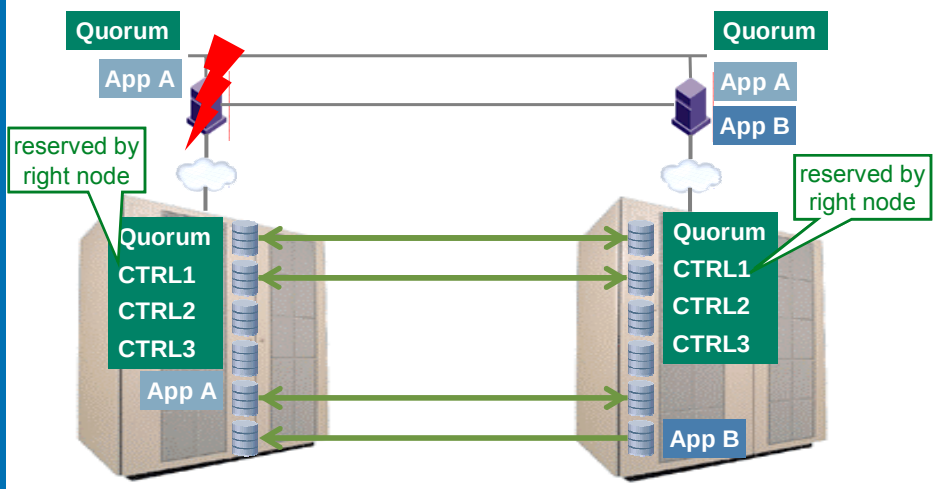
4/22/2004 10

Agenda

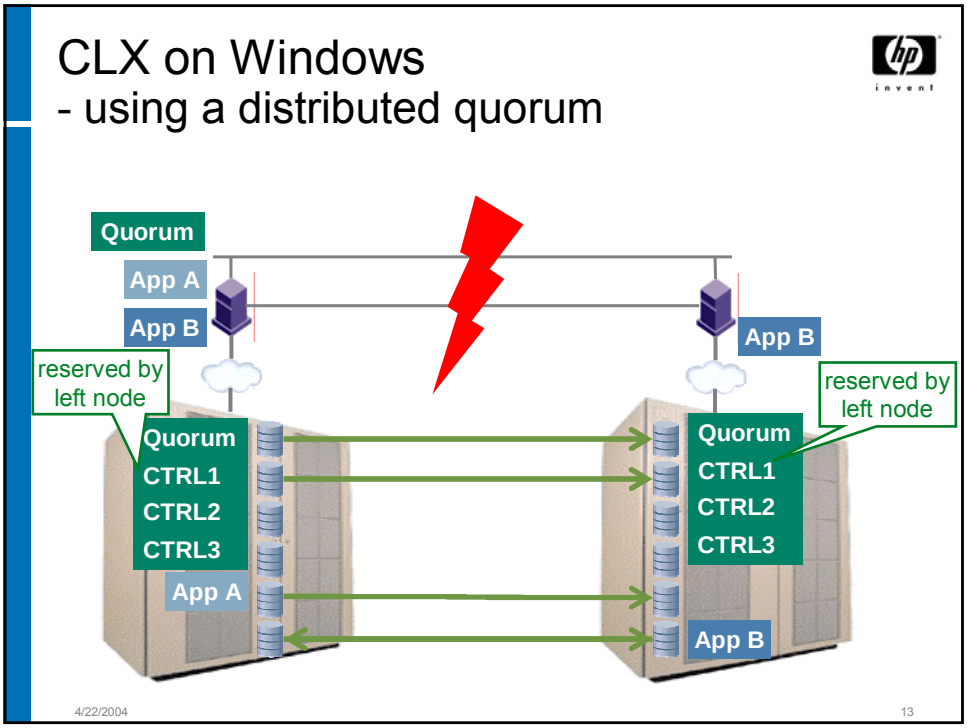
- Integration of CLX into MSCS
  - Using a Distributed Quorum and External Arbitration
  - Using Majority Node Sets with CLX
  - Handling Shared Disks: CLX Resource Type, Cluster Administrator Extension, Resource Dependencies,

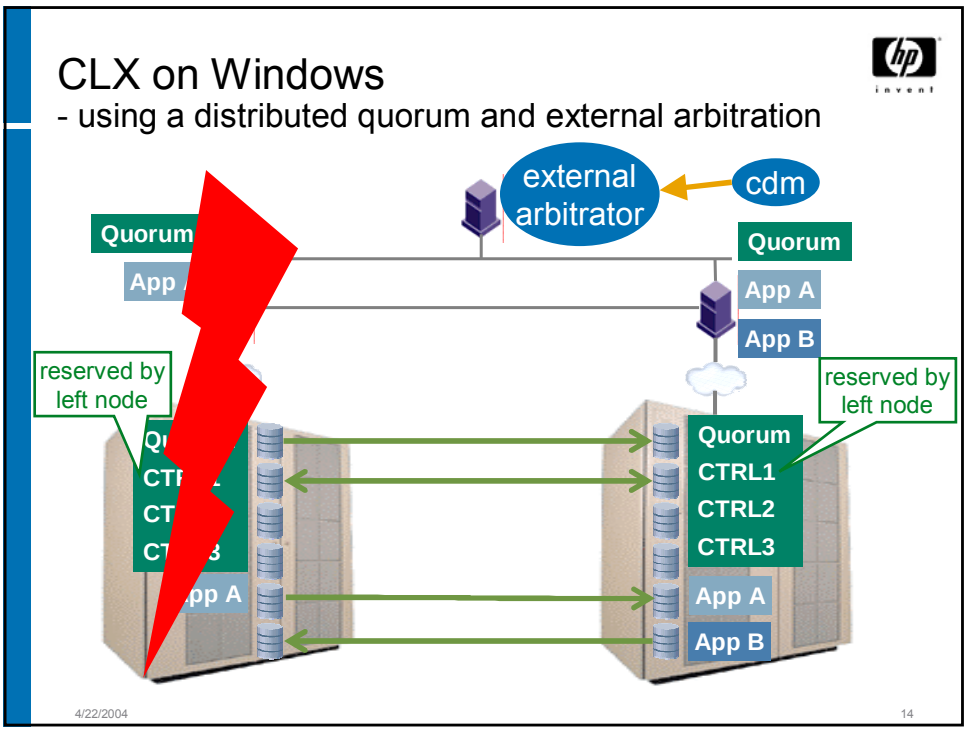
4/22/2004 11

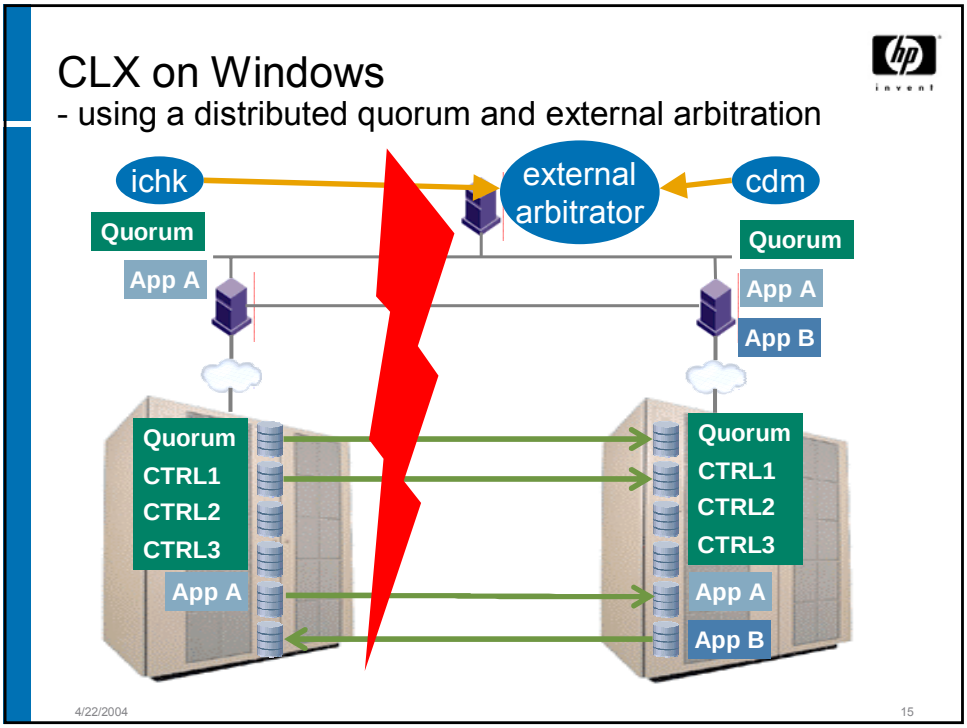
CLX on Windows
- using a distributed quorum

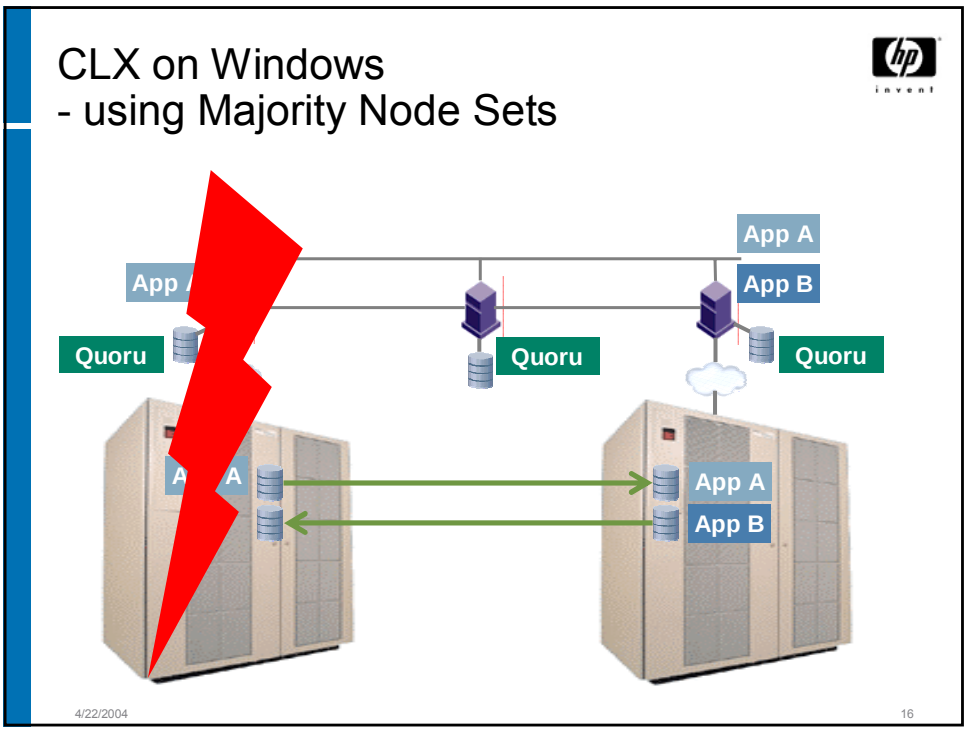


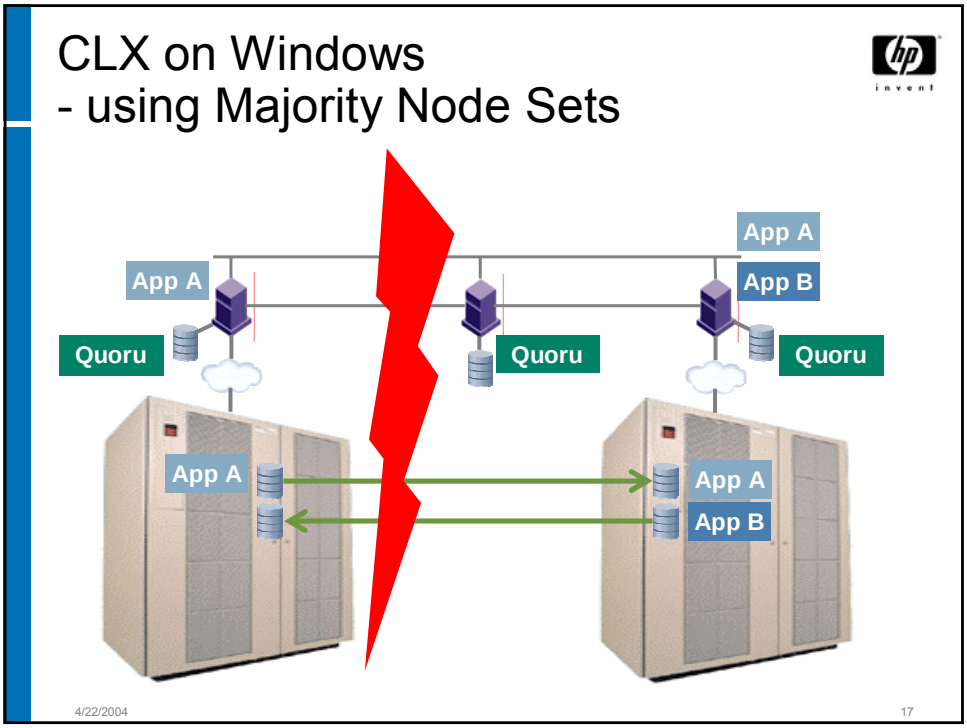
4/22/2004 12

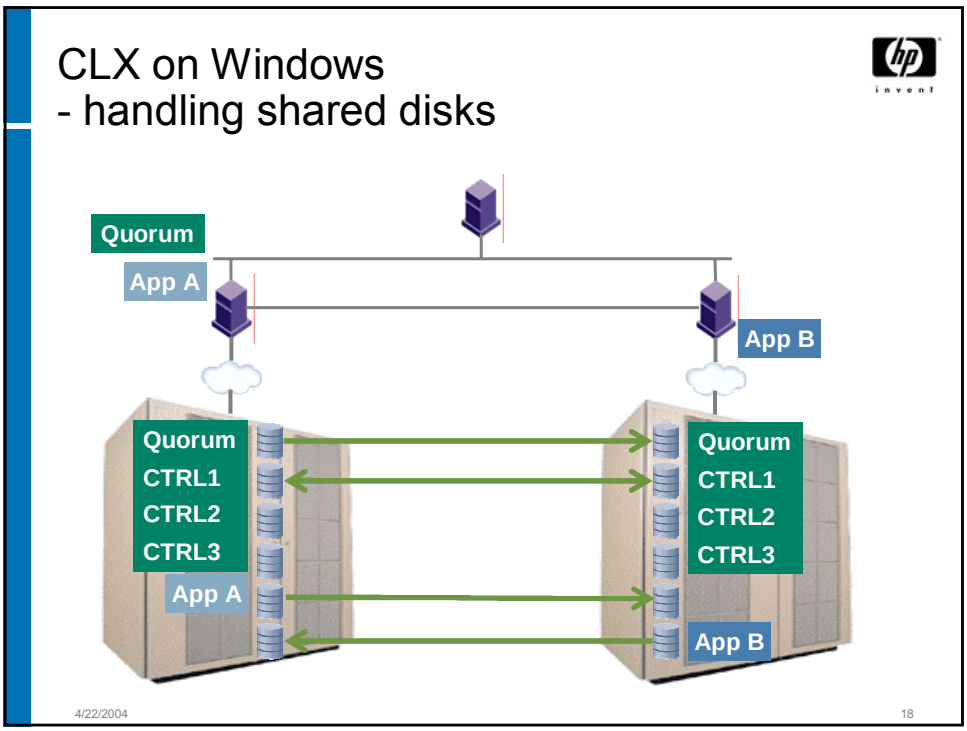

CLX on Windows
- using a distributed quorum
Quorum
App A
App B
reserved by left node
Quorum
CTRL1
CTRL2
CTRL3
App A
App B
reserved by left node
Quorum
CTRL1
CTRL2
CTRL3
App B
4/22/2004
13


CLX on Windows
- using a distributed quorum and external arbitration
external arbitrator
cdm
Quorum
App A
App B
reserved by left node
Quorum
CTRL1
CTRL2
CTRL3
App A
App B
reserved by left node
4/22/2004
14

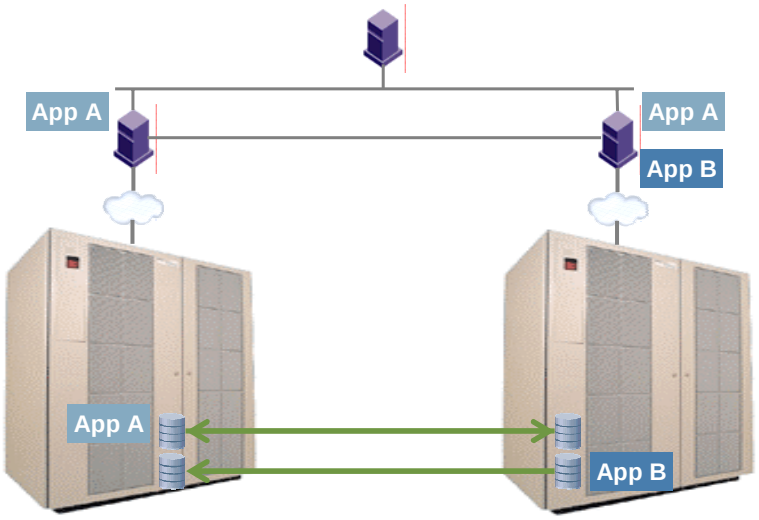

CLX on Windows
- using a distributed quorum and external arbitration
ichk
external arbitrator
cdm
Quorum
App A
Quorum
App A
App B
Quorum
CTRL1
CTRL2
CTRL3
App A
Quorum
CTRL1
CTRL2
CTRL3
App A
App B
4/22/2004
15


CLX on Windows
- using Majority Node Sets
App A
App B
Quoru
Quoru
Quoru
App A
App B
4/22/2004
16


CLX on Windows
- using Majority Node Sets
App A
App B
App A
Quoru
Quoru
Quoru
App A
App A
App B
4/22/2004
17

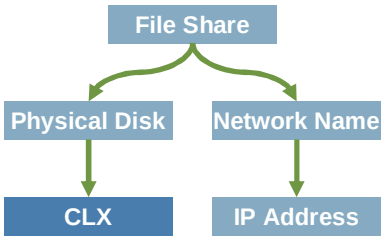

CLX on Windows
- handling shared disks
Quorum
App A
App B
Quorum
CTRL1
CTRL2
CTRL3
App A
Quorum
CTRL1
CTRL2
CTRL3
App B
4/22/2004
18

CLX on Windows
- handling shared disks



4/22/2004 19

Resource Dependency Trees
Example Virtual File Share



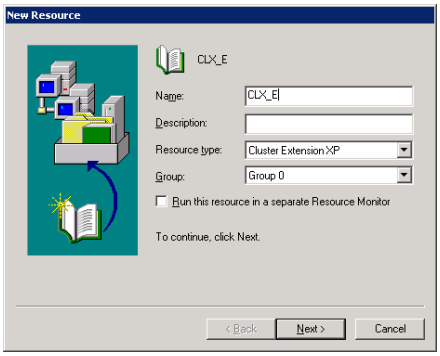
- All Physical Disk resources of one Resource Group depend on a CLX resource
- Very smooth integration

4/22/2004 20

Resource Type „Cluster Extension XP“

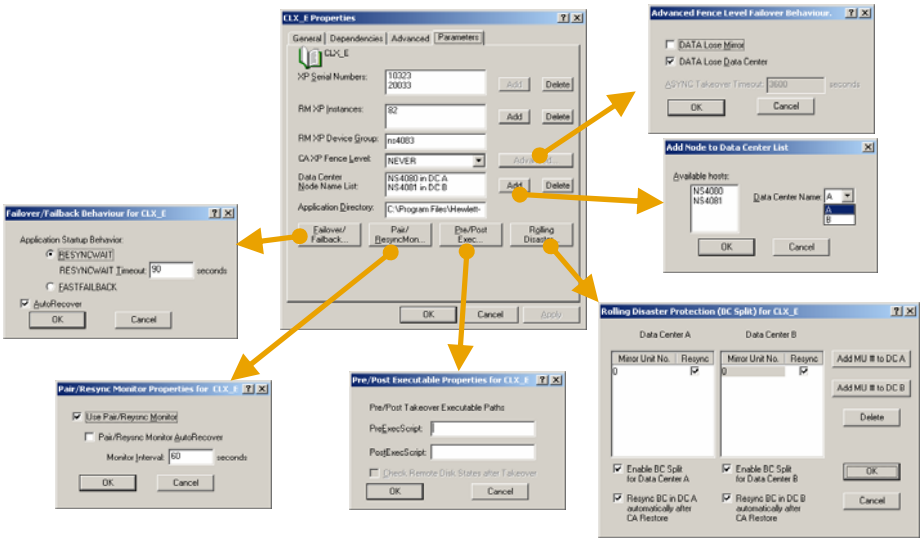
What does it do?

- responsible for takeover actions
- triggers the Pair-/Resync-Monitor, if configured
- double-checks fence-level, XP serial number, etc.
- uses RaidManager instance



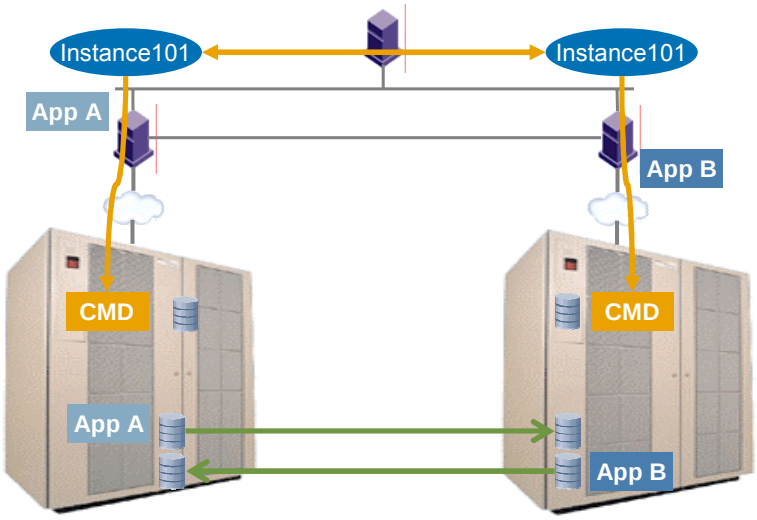
4/22/2004 21

Cluster Administrator Extension DLL for the CLX Resource Type



4/22/2004 22

RaidManager XP



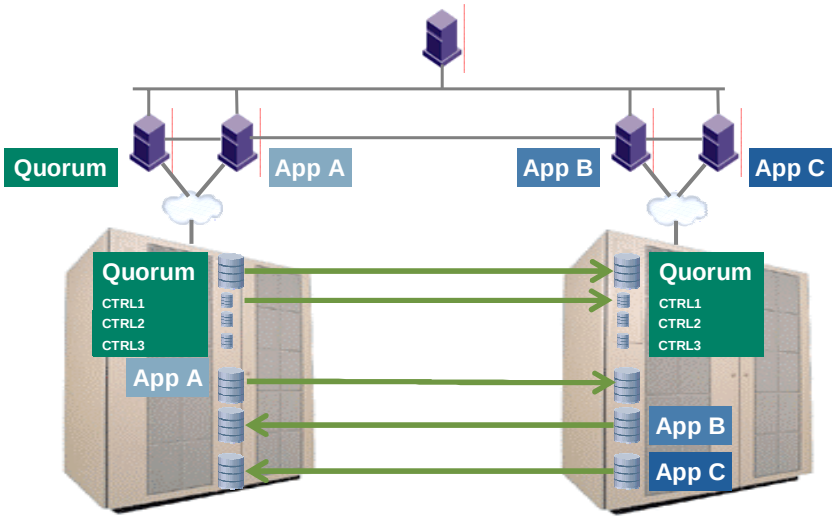
4/22/2004 23

Agenda

- Best practices
  - Choose a Cluster Model:
    Distributed Quorum or Majority Node Sets?
  - LUN-Mapping
  - Parameter Settings for CLX Resources

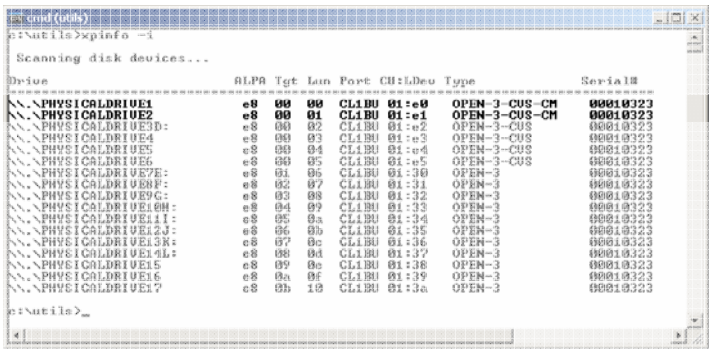
4/22/2004 24

CLX on Windows
- using a distributed quorum



4/22/2004 25

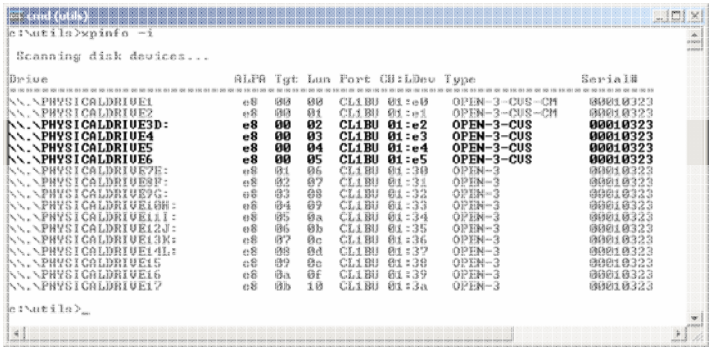
LUN-Mapping



```
c:\utils>xpinfo -i

 Scanning disk devices...

Drive                          ALPA  Tgt  Lun  Port  CH:LDev  Type          Serial#
\\.\PHYSICALDRIVE1             e8    00   00   CL1BU 01:e0    OPEN-3-CUS-CM 00010323
\\.\PHYSICALDRIVE2             e8    00   01   CL1BU 01:e1    OPEN-3-CUS-CM 00010323
\\.\PHYSICALDRIVE3D:           e8    00   02   CL1BU 01:e2    OPEN-3-CUS    00010323
\\.\PHYSICALDRIVE4             e8    00   03   CL1BU 01:e3    OPEN-3-CUS    00010323
\\.\PHYSICALDRIVE5             e8    00   04   CL1BU 01:e4    OPEN-3-CUS    00010323
\\.\PHYSICALDRIVE6             e8    00   05   CL1BU 01:e5    OPEN-3-CUS    00010323
\\.\PHYSICALDRIVE7E:           e8    01   06   CL1BU 01:30    OPEN-3        00010323
\\.\PHYSICALDRIVE8F:           e8    02   07   CL1BU 01:31    OPEN-3        00010323
\\.\PHYSICALDRIVE9G:           e8    03   08   CL1BU 01:32    OPEN-3        00010323
\\.\PHYSICALDRIVE10H:          e8    04   09   CL1BU 01:33    OPEN-3        00010323
\\.\PHYSICALDRIVE11I:          e8    05   0a   CL1BU 01:34    OPEN-3        00010323
\\.\PHYSICALDRIVE12J:          e8    06   0b   CL1BU 01:35    OPEN-3        00010323
\\.\PHYSICALDRIVE13K:          e8    07   0c   CL1BU 01:36    OPEN-3        00010323
\\.\PHYSICALDRIVE14L:          e8    08   0d   CL1BU 01:37    OPEN-3        00010323
\\.\PHYSICALDRIVE15            e8    09   0e   CL1BU 01:38    OPEN-3        00010323
\\.\PHYSICALDRIVE16            e8    0a   0f   CL1BU 01:39    OPEN-3        00010323
\\.\PHYSICALDRIVE17            e8    0b   10   CL1BU 01:3a    OPEN-3        00010323

c:\utils>_
```

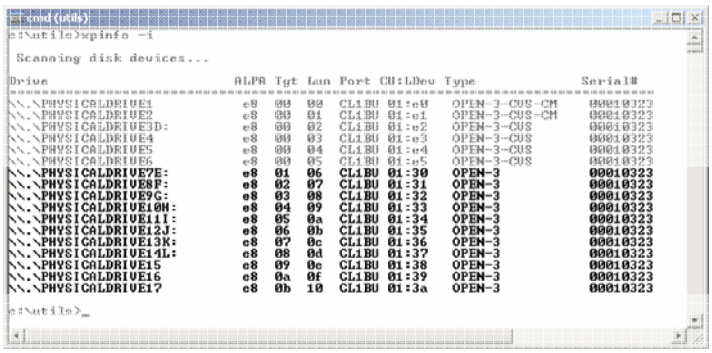
4/22/2004 26

LUN-Mapping
Quorum and CTRL-Disks



4/22/2004 27

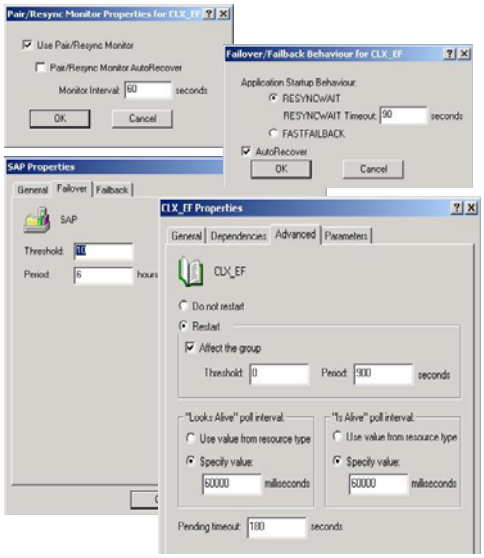
LUN-Mapping
Quorum and CTRL-Disks



4/22/2004 28

Parameter Settings for CLX-Resources

- Fence level
  - for higher availability: never
  - for guaranteed data currency and consistency: data
  - To avoid latency over long distances: async
- Failover-/Failback-Behavior
  - Two nodes: resyncwait
  - more nodes: fastfailback
- Use Autorecover with care
- Timeout dependencies: failover period (group), pending timeout (clx resource), RaidManager timeout (horcmX.conf)



Agenda

- Identifying the building blocks of CLX on Windows
  - How does a local cluster with a single Quorum work?
  - A new cluster model: How do Majority Node Sets work?
  - What is needed for a distributed environment?
- Integration of CLX into MSCS
  - Using a Distributed Quorum and External Arbitration
  - Using Majority Node Sets with CLX
  - Handling Shared Disks: CLX Resource Type, Cluster Administrator Extension, Resource Dependencies, RaidManager
- Best practices
  - Choose a Cluster Model: Distributed Quorum or Majority Node Sets?
  - LUN-Mapping
  - Parameter Settings for CLX Resources

invent