



Five Nines with OpenMail and MC/ServiceGuard

Richi Jennings, Technical Product Manager
Hewlett-Packard
Nine Mile Ride
Wokingham
Berkshire RG40 3LL
United Kingdom

T +44 (0)1344-365870
F +44 (0)1344-763429
richi@hp.com

OpenMail®



Agenda

2

OpenMail[®]

How To Implement Highly-Available Messaging and Collaboration

- ◆ **OpenMail backgrounder**
- ◆ **Planning for reliability and scalability**
- ◆ **HA**
- ◆ **Backup and disaster recovery**

Lots to cover in 50 minutes, so don't expect huge depths of knowledge. However, I'll give you pointers for further study.



1. OpenMail Backgrounder

OpenMail®



OpenMail is...

- OpenMail
- ◆ **Scalable**
 - ◆ **Flexible**
 - ◆ **Reliable**
 - ◆ **Highly-available**
 - ◆ **Low-cost**
 - ◆ **Feature-rich**
 - ◆ **Future-proof**



2. Planning for Reliability and Scalability

OpenMail[®]



"Reliability" vs. "High Availability"

OpenMail

- **My definitions...**
 - *HA*: using redundant machines to improve uptime in the event of failure;
 - *Reliability*: everything else!
- **Before thinking about HA, think about *design for reliability***
- **The cost of going from 99% to 99.9% is "small", compared with going from 99.9% to 99.99%**
- **Many OpenMail customers report high uptimes *without* HA**
- **Think hard about your requirements—perhaps only *some* of your users need that extra percentage point of availability?**
 - Design your network and write your SLAs accordingly



Scalability

7

OpenMail

◆ **Key is I/O performance**

- Key to I/O performance is spreading the load over several spindles and interfaces

◆ **Sometimes a scalability vs. reliability tradeoff**

- Or, strictly-speaking, cost more to provide similar scalability with improved reliability

◆ **Fewer, larger servers are less complicated to administer**

- e.g. routing and directory synchronization

◆ **More, smaller servers...**

- may affect fewer users on failure
 - common fallacy: “smaller servers reduce impact of downtime”
- less expensive hot spares
- have more potential for expansion

◆ **The above is true for almost any messaging system (not just for OpenMail)**



RAID

OpenMail

A realistic business messaging system uses a lot of disks

- 50MB per user is not unusual
 - 250GB for 5000 users!
 - 10s of spindles, decreasing reliability by an order of magnitude
- Plan for expansion!

Mirroring

- You should plan for disks to fail, so you should probably use mirroring (and hot spares are nice)
- Performance tradeoff
- Great for fast disaster recovery

Striping

- Good for improving I/O performance in a few “hot spots”, but use *thin* striping (not *byte* or *extent*)
- Generally, not necessary with the bulk of OpenMail data storage, because the application automatically does this (`omdisksprd`)

If you're serious, you'll probably use an independent disk array

(e.g. Nike, AutoRAID, SureStore E Disk Array, EMC), rather than “JBOD”



OpenMail Filesystems

- OpenMail
- /opt/openmail/
 - Static data (mostly binaries) so don't worry about RAID
 - /var/opt/openmail/
 - Hot spots live here, so stripe and mirror
 - Special treatment of temporary data helps (symlink to fstemp, JFS nolog, re-create fstemp on each boot)
 - /var/opt/openmail/data/
 - OpenMail spreads data across mounted filesystems under here

/opt	/opt	/opt	/opt	/opt
/opt/fstemp	/opt/fstemp	/opt/fstemp	/opt/fstemp	/opt/fstemp
/opt/condat1	/opt/condat1	/opt/condat1	/opt/condat1	/opt/condat1
...
/opt/condat1	/opt/condat2	/opt/condat3	/opt/condat4	/opt/condat5
OS	Mail Queue	spare	spare	spare



3. High Availability

10

OpenMail[®]



MC/ServiceGuard

11

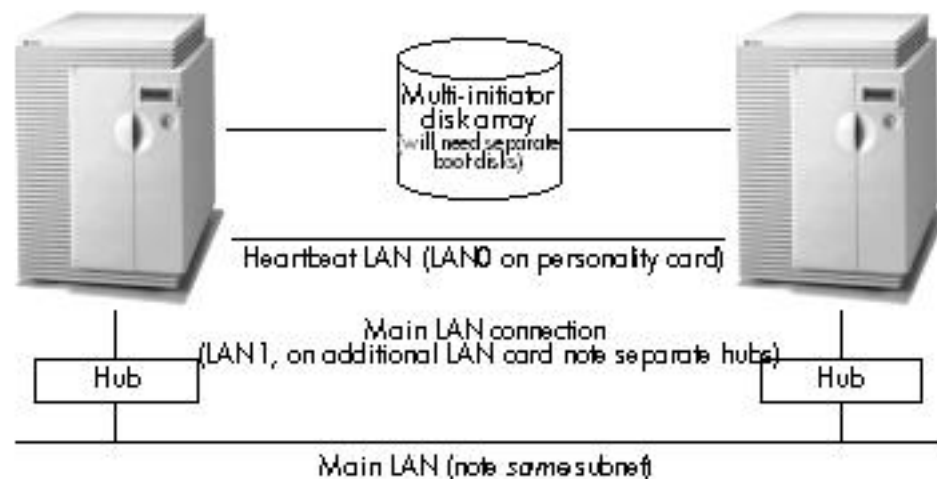
OpenMail

- **Up to this point, we have talked about resilience against disk subsystem failures**
- **MC/ServiceGuard protects against processor failures**
- **OpenMail supports both active/passive and active/active MC/ServiceGuard cluster configurations**
- **OpenMail supports multiple running installations on a single machine, so can load-share on failure**
- **Also RSF1 for Linux and Solaris (High Availability.com)**



Simple Two-node Example

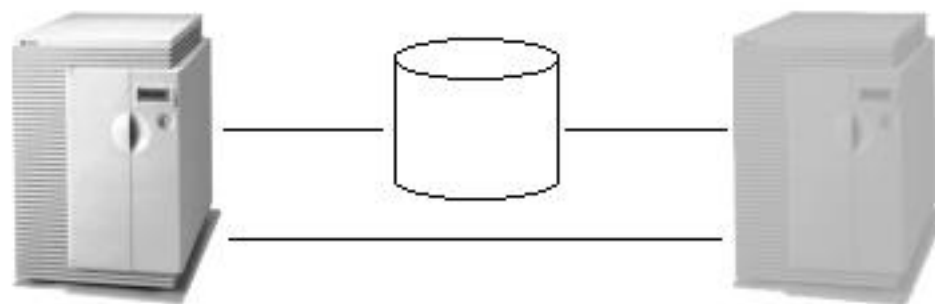
OpenMail[®]





Active/Passive Example

OpenMail[®]

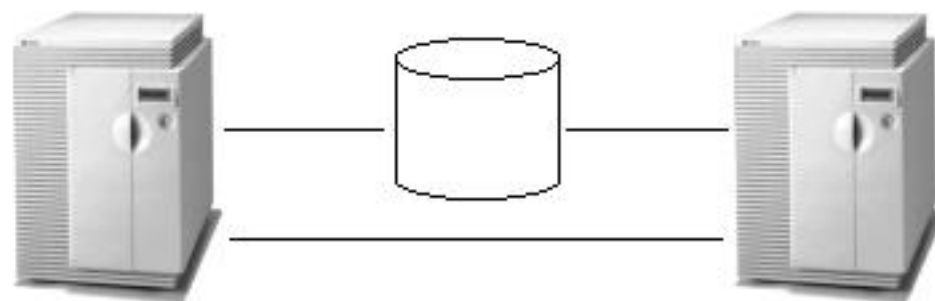


- ◆ OpenMail runs on *one* machine
- ◆ Other machine unused: a hot spare
- ◆ Expensive
- ◆ To save your sanity, configure both machines identically!



Active/Active Example

OpenMail[®]

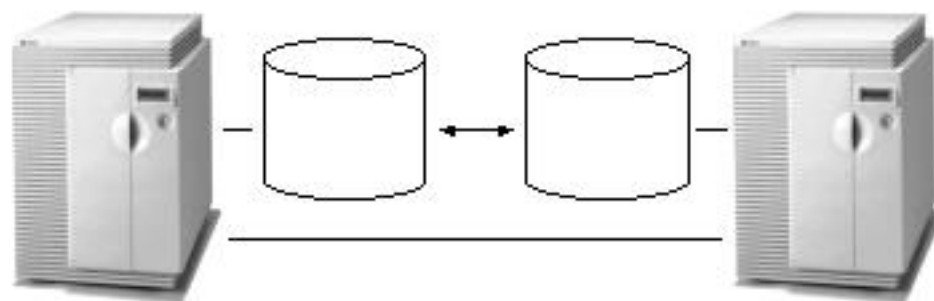


- ◆ OpenMail runs on *both* machines
- ◆ On failure, remaining machine runs *both* OpenMail services, simultaneously
- ◆ More taxing on CPU, but disk I/O is usually the bottleneck, so not as big a problem as it might appear



What if the Disk Array Fails?

OpenMail[®]



- ◆ Use MirrorDisk/UX to provide physical replication of data
- ◆ See later for an additional use of this configuration



4. Backup and Disaster Recovery

OpenMail[®]



Requirements?

17

OpenMail

- ◆ **Fast backup**
- ◆ **Fast recovery**
- ◆ **What about changes since last backup?**
- ◆ **Bombs, etc.**



Backup/restore Performance; Availability

- OpenMail
- **With a DLT array, OpenMail customers have reported backup speeds of 150MB/minute**
 - **Recovery tends to be significantly slower**
 - **To provide "365x7" availability, consider splitting the disk mirror...**
 - omsuspend, split the mirror, omsuspend -r
 - Backup the offline mirror to tape
 - Re-merge the mirror
 - **To provide lightning-fast recovery, consider adding another layer of mirroring**
 - Use the offline mirror as a hot backup, obviating the need to recover from tape!



What About Changes Since Last Backup?

OpenMail

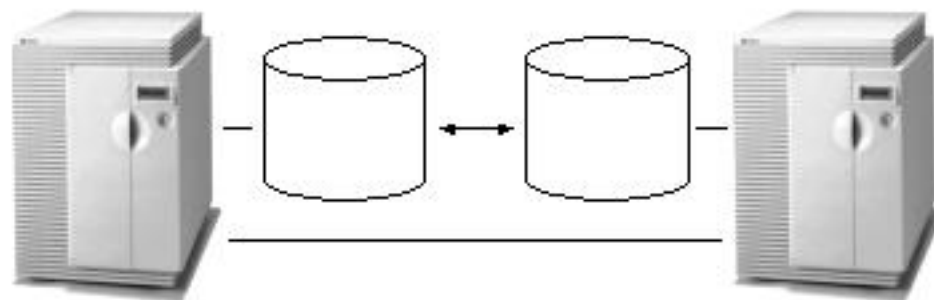
- ◆ **OpenMail includes the “Archive Server”**
 - Keeps a copy of all messages transported by this system
- ◆ **Keep the archive logs on a separate filesystem, immune to common-mode failures of the main message store**
- ◆ **Once you’ve recovered the store, re-play the archive logs**



Bombs, etc.

OpenMail

- **What if your data center is destroyed?**
- **Campus-wide architectures are possible, using MirrorDisk/UX or SAN hardware products**
- **Can permit physical separation of several miles between servers**





OpenMail is...

OpenMail

- ◆ **Scalable**
- ◆ **Flexible**
- ◆ **Reliable**
- ◆ **Highly-available**
- ◆ **Low-cost**
- ◆ **Feature-rich**
- ◆ **Future-proof**



OpenMail

**Business Messaging for the next E.
E-services.**

www.hp.com/go/openmail