



Compaq OpenVMS

e

Digital Networks

**Relatório de teste em
Ambiente de Cluster
OpenVMS**

14 de agosto de 2001



Resumo executivo

Testes foram realizados com equipamentos Digital Networks (DNmultilayer 1200 e DNswitch 800) para determinar sua viabilidade em um ambiente de Cluster OpenVMS.

O teste de hardware utilizou um 4 node AlphaServer GS com OpenVMS Cluster com 4 CPU's EV68.

O software usado foi o OpenVMS Versão 7.3 com OpenVMS Cluster e Host Based Volume Shadowing (HBVS) para OpenVMS.

O cenário do teste incluiu a configuração do cluster 4-node, 2 nodes com acesso direto para armazenamento e 2 nodes com MSCP (Mass Storage Control Protocol). Cada switch foi utilizada como Cluster Interconnect.

Os testes provaram que cada switch trabalharia em um ambiente OpenVMS Cluster.

Os testes provaram uma melhoria significativa do desempenho quando utilizando Gigabit Ethernet como Cluster Interconnect.

Estudo de viabilidade adicional foram feitos com o sistema Digital Networks GIGAswitch/FDDI para a Digital Network VNswitch com capacidade de Gigabit Ethernet.

Este estudo provou a viabilidade do upgrade para o sistema GIGAswitch/FDDI que utiliza o módulo VNswitch 900 GV.

Objetivos do Teste

O teste dos equipamentos Digital Networks DNswitch 800 e DNmultilayer 1200 aconteceram durante a semana de 22 de julho de 2001 no laboratório de testes da Compaq Computer Corporation VMS, em Nashua, NH. Estes testes foram administrados para testar a viabilidade e desempenho de 2 switches em um ambiente de cluster OpenVMS. Em um teste secundário, a Digital Networks testou que considera ser o legado de ambiente padrão, utilizando o sistema Digital Networks GIGAswitch/FDDI e um MultiSwitch 900 com opções VNswitch para testar o FDDI para Gigabit Ethernet OpenVMS Cluster. Uma explicação mais detalhada está no final deste relatório.

Cenário do Teste

O hardware utilizado no teste foi um AlphaServer GS160 com 16 CPU's EV68 (1000 Mhz), 64 gigabytes de memória principal, 4 slots PCI, 2 controladoras de disco KGPSA, 4 controladoras HSG60, unidades de disco de 18, 24 e 36 gigabytes; 4 controladoras de Ethernet DE602-AA 10/100 e 4 controladoras de Gigabit Ethernet DEGPA-SA.

O hardware do teste foi configurado em 4 partições que consistem em 1 Quad Building Block (QBB) com 4 CPU's EV68, 16 gigabytes de memória, 1 slot PCI e 1 controladora de Ethernet DE602-AA. Duas das partições tinham controladoras de disco KGPSA e 2 não. Esta configuração foi projetada para causar I/O's no MSCP de uma das partições, aumentando assim a carga na switch testada.

O ambiente de software usado foi o Compaq OpenVMS Versão 7.3 incluindo os softwares OpenVMS Cluster e OpenVMS Host Based Volume Shadowing.

O hardware de teste foi configurado com o OpenVMS rodando em um único disco do sistema, ligado às controladoras HSG80. A partir daí, só 2 exemplos do hardware em teste tiveram acesso direto às controladoras HSG e ao disco do sistema, os outros 2 exemplos foram conectados ao disco do sistema usando a conexão de rede local (LAN), que eram as switches da Digital Networks em teste naquele momento.

Teste de equipamentos Digital Networks

Foram administrados testes em 2 switches Digital Networks, o DNmultilayer 1200 configurado com 16 conexões TX 10/100 Ethernet e 3 conexões Gigabit (1000 BASE SX), e o DNswitch 800 com 8 conexões Gigabit (1000 BASE SX) Ethernet. Depois da conclusão, alguns equipamentos de legado adicional foram testados. Um sistema GIGAswitch/FDDI e um MultiSwitch 900 foram configurados para testar o upgrade de sistemas FDDI baseados em clusters OpenVMS para Gigabit Ethernet. O MultiSwitch 900 consistiu em 5 módulos: um 916TXG, um VNswitch 900EF, um VNswitch 900FX, um DECConcentrator 900MX e um VNswitch 900GV.

As famílias DNmultilayer e DNswitch utilizam uma base de hardware e software comum em todos os seus produtos. Então, estes testes servem de exemplo para todos os produtos das famílias DNmultilayer e DNswitch.

Objetivo dos testes

O objetivo dos procedimentos era testar as switches Digital Networks para determinar se elas trabalhariam com sucesso em um ambiente de clusters OpenVMS, sem qualquer queda de desempenho sério para tráfego de LAN em clusters. O DNmultilayer 1200 pode trabalhar em ambas as velocidades de 10/100 Mbs e a velocidades de Gigabit, assim o teste inicial feito nesta switch foi para determinar a viabilidade e o diferencial de desempenho. O teste da 1200 consistiu em criar um 2 nodes OpenVMS Cluster rodando em uma LAN de 100 Mbs, cada node com acesso direto a subsistemas do disco, então foram adicionados os 2 nodes restantes em sucessão com uma conexão em LAN 100 Mbs, tendo o Mass Storage Control Protocol (MSCP) acesso ao subsistema de armazenamento. A conexão a disco MSCP implica que os nodes de serviço podem executar tarefas de intervenção em nome dos nodes servidos. Isto utiliza a CPU e largura de banda da rede. Os testes executados no OpenVMS cluster consistiram em atividades normais de cluster como configuração de cluster e setup, como também espelhamento de disco e cópias de fusão. Em complemento, numerosas transferências de disco para disco usando comandos VMS COPY e BACKUP foram feitas para utilizar o backbone da rede. Após a conclusão com sucesso do teste de 100 Mbs, o mesmo cenário foi usado utilizando conexões Gigabit Ethernet no 1200. O segundo set de testes foi realizado então no DNswitch 800.

Figuras 1 e 2 mostram as configurações usadas:

Test Configuration #1

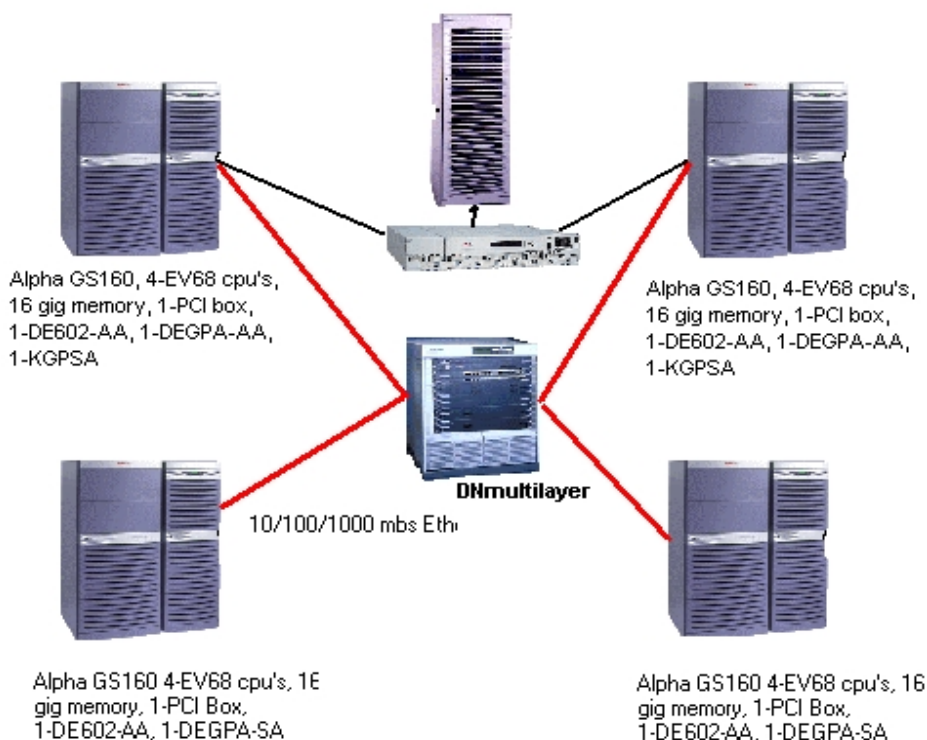


Figure 1



Test Configuration #2

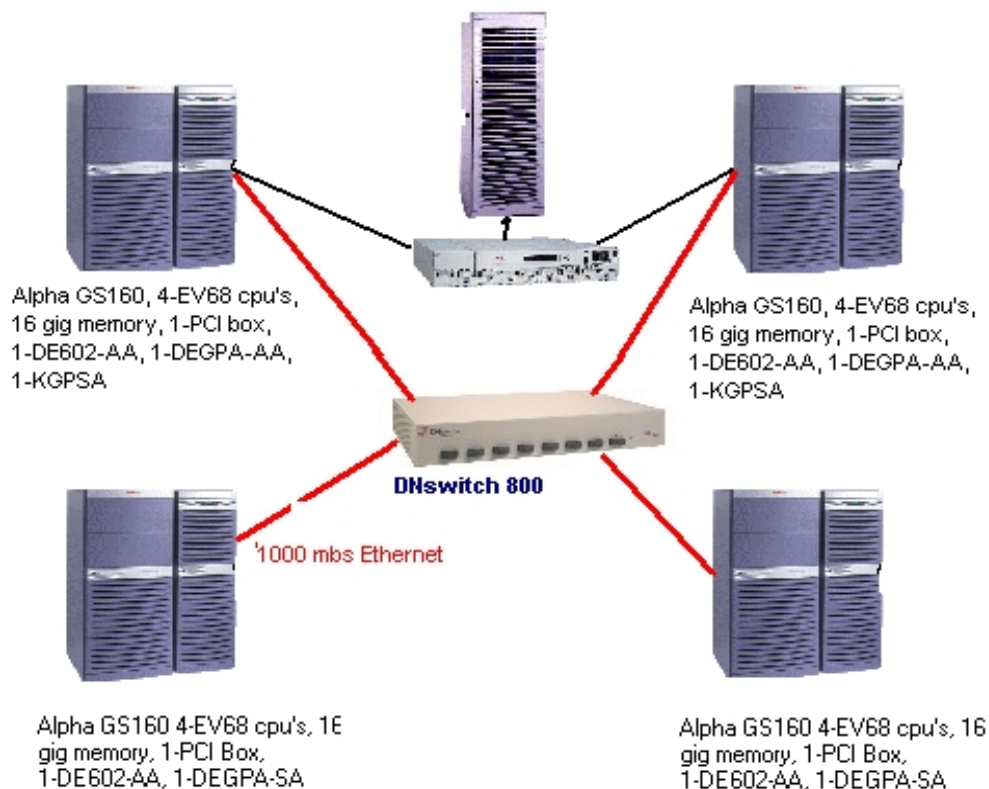


Figure 2

Critérios de sucesso

O sucesso desta verificação de sistema foi baseado em dois marcos principais. Primeiro, as switches tiveram que bootar e configurar um OpenVMS cluster composto de 4 nodes sem perda mensurável de pacotes de LAN, e segundo, cada switch teve que passar MSCP I/O's sem impacto sério de desempenho em I/O ou na atividade da CPU. As medições foram tomadas usando ferramentas de desempenho padrão OpenVMS como Monitor and System Dump Analyzer (SDA). A média do tempo de transferência de pacotes na LAN (RTT) e o número de colisões foram consideradas importantes e efetivas na medição de tráfego do cluster. Geralmente a utilização da CPU e o tempo Kernel mode CPU também são importantes para o desempenho do cluster. Se uma grande quantidade de tempo Kernel mode CPU é gasto na atividade de cluster, poderia rapidamente degradar todo o sistema e a performance do cluster.

Resultados dos testes

Os resultados dos testes conduzidos mostraram que ambas as switches passaram com sucesso em todos os critérios analisados. O DNmultilayer 1200 estava habilitado para configurar e rodar os 4 nodes OpenVMS Cluster a velocidades de 100 Mbs e Gigabit sem problemas de desempenho perceptíveis.

A média de cada pacote RTT na LAN para cada um dos nodes de cluster em modo de 100 Mbs era de 1400-1500 microsegundos para cada um dos 4 nodes, com a maior média comum registrada de 4400 microsegundos. O tráfego de RTT na LAN para o DNmultilayer usando Gigabit Ethernet era significativamente menor, na ordem de 280-400 microsegundos para as mesmas atividades, indicando uma melhoria de desempenho significativa quando utilizando o Gigabit Ethernet como Cluster Interconnect. De modo geral, a utilização da CPU foi extremamente baixa para operações de cluster normais, menos de 5% da atividade total do sistema, e a atividade de modo Kernel foi de menos de 3%, até mesmo durante cópias de volume pesadas e operações de backup entre nodes usando acesso MSCP. O pacote RTT para a DNswitch 800 foi similar ao DNmultilayer Gigabit Ethernet RTT, o que já era esperado.



Uma exigência de configuração importante que foi notada nestes testes é que os usuários têm que verificar que adaptadores Gigabit Ethernet no Alpha sejam configurados para Full Duplex. Isto pode ser feito no prompt usando o seguinte comando:

```
>>> SET EWA0_MODE FD
```

Se o dispositivo Ethernet não estiver configurado para Full Duplex, a seguinte mensagem de erro será exibida:

```
PEA0 - Numerous excessive packet losses on LAN Path...
```

Conclusões

As famílias DNmultilayer e DNswitch 800 provaram ser uma solução de LAN Backbone extremamente satisfatória para clusters OpenVMS, especialmente com Gigabit Ethernet. A família DNmultilayer de switches fornecem um eficiente ambiente de multi-tecnologia, com múltiplas portas Gigabit Ethernet e Fast Ethernet para permitir uma fácil migração de clusters Ethernet 10/100 para o ambiente Gigabit mais rápido. O DNswitch 800 é um excelente cluster backbone Gigabit Ethernet para 4 até 8 nodes devido a seu alto desempenho e baixo custo.

Legado de Ambiente de Redes OpenVMS

A divisão de equipamentos de rede da Digital Equipment Corporation desenvolveu o sistema de GIGAswitch/FDDI. Um grande número destes dispositivos se tornaram backbones para redes como também o subsistema de I/O para OpenVMS Clusters e tem trabalhado de forma confiável durante anos. Hoje no entanto, com o advento do padrão Gigabit Ethernet, o FDDI está caindo lentamente e os usuários estão buscando ampliar as capacidades de seus ambientes FDDI ou estão procurando modos para integrar FDDI com Fast Ethernet e Gigabit Ethernet. Pensando nisto, a Digital Networks escolheu definir e validar uma rota para estes clientes seguirem.

O ambiente típico que usuários Digital normalmente possuem, é um backbone FDDI que usa um ou mais sistemas de GIGAswitch/FDDI junto com um ou mais MultiSwitch 900 (antigamente conhecido como o DECHub 900). O acesso ao ambiente FDDI é normalmente realizado por DECconcentrators como o DECconcentrator 900MX ou mais recentemente com um dispositivo de VNswitch como o VNswitch 900FX. Estes dispositivos podem fornecer acesso único ou dual. As máquinas Alpha podem ser ligadas diretamente ao backbone FDDI utilizando um adaptador DEFPA-xx através do sistema GIGAswitch/FDDI por um DEFGL-xx line card. Uma vez que o OpenVMS suporta Clusters usando FDDI como interconexão, assim como Fast Ethernet e Gigabit Ethernet, a meta da Digital Networks é permitir os clientes uma perfeita integração FDDI em ambientes Fast Ethernet e Gigabit Ethernet.

Ambiente proposto

A atualização de um ambiente FDDI é atualmente bastante simples e pode ser feita de diversas maneiras. Os usuários com configurações específicas devem conferir com o suporte técnico da **DECmaster** se sua configuração específica é suportada. Para introduzir um Gigabit Ethernet cluster a um FDDI cluster existente, um dos clusters precisa ser equipado com um adaptador DEGPA-SÁ Gigabit Ethernet PCI que é ligado a uma VNswitch 900GV ou VNswitch 900CG encontradas em uma MultiSwitch 900. Esta switch multilayer consta de uma única porta modular Gigabit Ethernet (GBIC). Os produtos da família VNswitch fornecem acesso ao VNbus, com uma tecnologia de 400 Mbs independente do barramento do MultiSwitch 900. A conectividade para Gigabit Ethernet é fornecida pelo VNbus através de um módulo VNswitch FDDI, tal como a VNswitch 900EF. As 900EF são conectadas ao GIGAswitch/FDDI line card ou DEFGL, que automaticamente padroniza a conexão full duplex FDDI quando conectado em uma configuração ponto a ponto. O mesmo DEFGL (ou outro cartão) é conectado então aos sistemas de Alpha/VAX no ambiente legado. O adaptador de FDDI neste cenário foi um Single Attached Station (SAS). O sistema GIGAswitch/FDDI automaticamente configura o dispositivo de entrada, porém, no caso de configurar um anel FDDI, é importante configurar o VN900EF como uma "Entrance into HUB from outside Ring" na seção de configuração FDDI. O node FDDI nesta configuração não teve acesso direto à ordem de armazenamento, utilizou o MSCP para ganhar acesso aos discos. Este cenário seria igual a um node satélite ou o que é geralmente conhecido como Local Area VMS Cluster ou LAN cluster. Também era pretendido colocar uma carga maior nas switches envolvidas.

A figura 3 mostra um exemplo desta configuração.

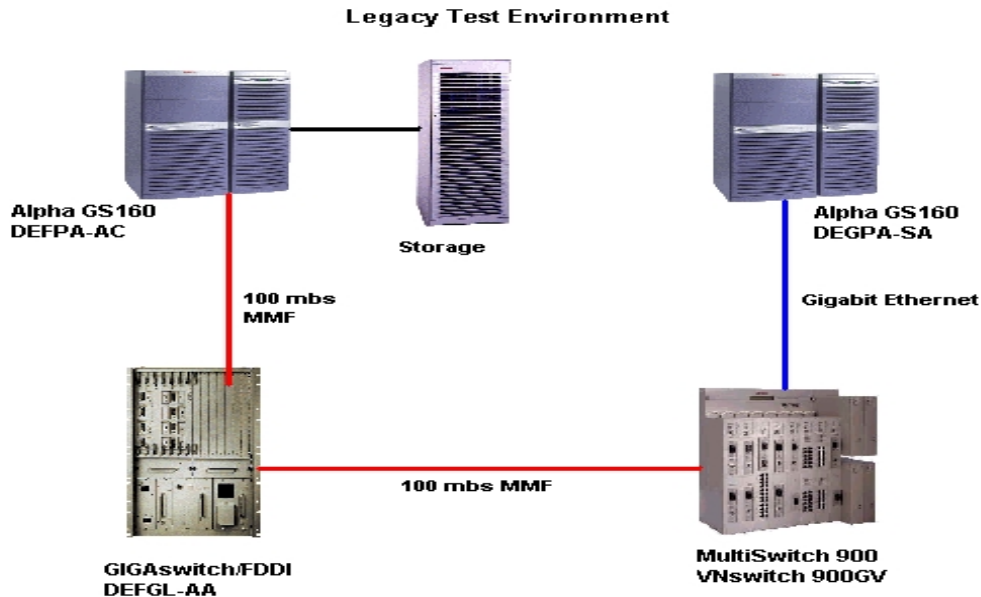


Figure 3

Plano do teste

A configuração anterior foi montada dentro do laboratório OpenVMS e submetida à mesma bateria de testes dos DNmultilayer e DNswitch 800.

Resultados do teste

Várias conclusões surgiram desta configuração de teste. Primeiro, provou ser uma configuração válida sem perda de conexão e demonstrou desempenho sustentável. Pacotes RTT do node Gigabit para o node FDDI tiveram média de 3632 microsegundos e pacotes RTT do node FDDI para o node Gigabit, média de 5228 microsegundos. Isto é mais lento que os tempos médios para Fast Ethernet devido ao número de saltos de pacotes enviados. Os erros de congestionamento nesta configuração foram mínimos e foram manifestados como mensagens retransmitidas. Em todos os casos observados o número de erros nunca excedeu 2% dos pacotes totais enviados, o qual nós sentimos está certamente dentro de limites aceitáveis. Não houve nenhum fechamento de circuito virtual informado e as estatísticas de PEdriver não mostraram nenhum erro, isto determinou que estes eram provavelmente erros de I/O relacionados ao tráfico de MSCP, que é controlado pela rede. De modo geral, a utilização da rede foi baixa, nunca ultrapassando mais que 5% da largura de banda total disponível, mesmo com 10 grandes acessos de cópia de disco para disco simultâneos.

Conclusões

O teste do OpenVMS FDDI para Gigabit Ethernet provou ser uma solução de hardware estável, solução esta que pode ser adotada por usuários que desejam integrar o seu sistema FDDI baseado em OpenVMS Cluster com o padrão Gigabit Ethernet.