

(Relatório Intermédio)

Linux Cluster with 3 or more Applications

GitHub:

<https://github.com/RubenABM/Sistemas-operativos>

Ana Bernardino (202000330), João Pires(20200459),
Marta Sousa (20200278) e Rúben Martins (20200453)

Licenciatura de Engenharia Informática

IADE – Faculdade de Design Tecnologias e Comunicação

Sistemas Operativos

Professor Pedro Rosa

19 de Março 2023

Introdução

Cluster (ou clustering) é o nome dado a um sistema que relaciona dois ou mais computadores para que estes trabalhem de maneira conjunta no intuito de processar uma tarefa. Estas máquinas dividem entre si as atividades de processamento e executam este trabalho de maneira simultânea. Caso um node falhe, os serviços presentes serão transportados para outro node.

Este relatório intermédio dá seguimento à proposta inicial, sendo analisados mais aprofundadamente 4 soluções de clustering, de modo a ser escolhida uma.

As soluções escolhidas foram as seguintes:

- SUSE Linux Enterprise High Availability Extension
- InfoScale Availability
- HPE ServiceGuard
- Corosync com Pacemaker

As quatro soluções de clustering foram avaliadas em relação ao seu enquadramento como produto, dificuldade de implementação, escalabilidade, suporte a cluster híbrido, os seus recursos disponíveis, user interface, facilidade de gestão, requisitos de software, e por fim a sua fiabilidade. Após uma análise das soluções tendo em conta estes aspetos, foi elaborada uma ordem da melhor para pior solução, e por fim feitas as conclusões finais sobre a solução escolhida para implementar no decorrer da parte prática deste trabalho.

SUSE Linux Enterprise High Availability Extension

- Enquadramento do produto
 - A instalação e utilização do software é gratuita. Os updates são feitos até 60 dias de utilização, após esse período de tempo, é possível continuar a utilizar gratuitamente mas sem serem realizados mais updates. Para manter o software em dia é preciso adquirir a versão paga.
- Dificuldade de implementação
 - A instalação da solução da SUSE para o master server é bastante fácil e rápida, havendo um wizard bastante semelhante a outras distro linux (ubuntu server, etc).
 - No sítio oficial da SUSE existe [documentação](#) extensiva em como implementar todos os diferentes aspetos do software.
- Escalabilidade
 - É bastante fácil integrar novos nodes no cluster e realizar a sua gestão devido à interface disponibilizada.
- Cluster híbrido
 - O software permite criar clusters linux entre servidores físicos e virtuais, aumentando a flexibilidade e melhor utilização de recursos.
- Recursos genéricos
 - Rules-based failover - transferência automática ou manual de processos para fora de áreas afetadas.
 - HAProxy - load balancing e proxying para aplicações TCP e HTTP, dividindo os pedidos entre múltiplos servidores.

- Cluster Join - possibilidade de adicionar novos nodes a um cluster sem necessidade de replicação manual das configurações (aumenta facilidade de gestão).
- Templates - templates e wizards disponíveis para setups básicos.
- Cluster Bootstrap - setup a partir de menu, aumentando a rapidez dos processos.
- Geo Clustering - proteção contra eventos disruptivos regionais, sendo que o não existe limite de distância entre serviços.
- Access control lists - role-based controls para garantir o acesso aos recursos pelas pessoas certas.
- Cluster-wide shell - É possível executar comandos shell em todos os nodes ao mesmo tempo, melhorando a gestão do cluster.
- History explorer - acesso aos logs do cluster.
- Cluster test drive - possibilidade de simular situações de failover nos clusters.
- Resource agents - integração com aplicativos open source terceiros, i.e., Apache, IPv6, DRBD, Postgres, etc.
- Clustered Samba (CTDB) - CTDB resources adicionados automaticamente e sincronizados com Active Directory.
- Node Recovery - possibilidade de criar snapshots completos de nodes de modo a restaurar serviços após failover.
- DRBD - DRBD cria uma imagem dos dados presentes no active node de um cluster para o seu standby node. Suporte a synchronous e asynchronous mirroring.
- Cluster-aware file system - MD RAID 1 é uma solução de storage para cluster.
- Suporta third-party hypervisors.

- **Fiabilidade**
 - Tendo em conta o historial da marca, o customer support disponibilizado, bem como toda a documentação presente sobre os diferentes aspetos do software, acreditamos ser uma solução muito viável tanto para o nosso caso-estudo como para implementações a mais alto nível, nomeadamente contexto empresarial. Acreditamos que todos os recursos disponibilizados tornam o SUSE Linux Enterprise High Availability Extension fiável.
- **User interface**
 - (High Availability Web Konsole) interface para gerir e monitorizar os clusters existentes ou adicionar e instalar novos. user-friendly
- **Facilidade de gestão**
 - Devido à interface disponibilizada, é fácil gerir os clusters sem grande complexidade técnica. É possível adicionar clusters, resources, constraints e dependências, gerir grupos de resources, aumentando a escalabilidade de clusters com maior dimensão. Aceder a listas de controlo, etc.
- **Requisitos de software**
 - Existem alguns requisitos ao clustering com o SUSE Linux Enterprise High Availability Extension, nomeadamente o uso da solução SUSE Server como OS do master node, que vem acompanhada dos seus requisitos próprios (ram mínima 1024MG, espaço em disco 2Gb(8.5GB all patterns, 32GB snapshot/rollback of OS). Os restantes nodes podem ser de qualquer outro OS linux, têm apenas de ter a mesma arquitetura, não sendo suportado clustering multi OS.

InfoScale Availability

- Enquadramento do produto:
 - É uma solução de infraestrutura definida por software que é integrado diretamente com aplicações para fornecer alta disponibilidade, e recuperação de desastres para serviços críticos de negócios.
 - Entrega uma plataforma de disponibilidade comum e de recuperação de desastre para plataformas físicas, virtuais e todas as plataformas principais de nuvem pública, fornecendo a flexibilidade para otimizar a disponibilidade para qualquer plataforma com base em requisitos de negócio.
- Dificuldade de implementação:
 - Para implementar este software a empresa disponibiliza um documento onde diz como instalar e que requisitos são necessários.
 - Para utilizar qualquer um dos softwares desta empresa é necessário ter uma conta veritas. É preciso também ter VEMS (veritas entitlement management system) Administrator ou Download privileges para conseguir fazer download.
- Escalabilidade:
 - Reduza as despesas migrando aplicações e dados para a nuvem. Move entre plataformas físicas, virtuais, de nuvem e de container.
- Cluster Híbrido:
 - Tem capacidade para lidar tanto com ambientes físicos e virtuais.
- Recursos genéricos:
 - Previsibilidade: Atender todos os objetivos de tempo e recuperação (RTOs) e objetivos de ponto de recuperação (RPOs).
 - Conformidade: Esteja em conformidade com os mandatos de continuidade de negócios com relatórios de auditoria e ensaios de recuperação sem interrupções.

- Automação: Livros de execução de recuperação que automatizam o failover e a orquestração de recuperação para aplicativos multicamadas, reduzindo assim intervenções manuais e tornando as ações de failover determinísticas.
- Mobilidade: Mova cargas de trabalho entre plataformas sem conversões de dados demoradas.
- Flexibilidade: Integra-se facilmente a ambientes locais existentes e tecnologias de nuvem com amplo suporte para ambientes físicos, virtuais e infraestruturas de nuvem.
- Fiabilidade:
 - Proporcionar resiliência com recuperação automática ou manual de desastres, e falha rápida para falhas nos nodes.
- User interface
 - Interface permite aos Admins ativar a proteção no nível de bloco ou arquivo para satisfazer SLAs e características de desempenho do aplicativo sem manipular a arquitetura do aplicativo e sem requisitos para tempo de inatividade.
- Facilidade de gestão
 - Ajuda a garantir que os serviços de IT físicos e virtualizados sejam protegidos contra paragens não planeadas por meio da monitorização de aplicações, visibilidade e discernimento.
- Requisitos de software
 - Reduz essa complexidade ao oferecer suporte a todos os principais sistemas operacionais e plataformas (Windows, Linux, Cisco, Red hat, etc...).
 - Principais fornecedores de nuvem pública Amazon Web Services (AWS), Google Cloud Platform e Microsoft Azure.

HPE ServiceGuard

- Enquadramento do produto
 - Solução de clustering para Linux que possui high availability (HA) e disaster recovery (DR). Aumenta o uptime de aplicações críticas, protegendo-as de uma variedade de falhas de infraestrutura e aplicações em ambientes físicos ou virtuais.
 - A instalação e utilização do software é gratuita, tendo um período experimental de 60 dias. Após esse período será necessário então comprar uma licença.
- Dificuldade de implementação
 - Suporta hypervisors específicos, sendo eles VMware, RHEV, Hyper-V, e KVM guests.
 - É necessário existir algum conhecimento prévio, sendo necessário seguir um conjunto de documentação.
 - A empresa disponibiliza [documentação](#) extensiva de como implantar os diferentes pontos do software.
- Escalabilidade
 - É possível adicionar vários nodes através da interface de gestão.
- Cluster híbrido
 - Tem capacidade para lidar tanto com ambientes físicos como virtuais.
- Recursos genéricos
 - Recuperação rápida de falhas
 - Downtime zero durante manutenções
 - Administração fácil

- Detecção eficaz de falhas, que automaticamente transporta aplicações críticas para nodes seguros.
 - Garante que a aplicação esteja sempre disponível, mesmo em caso de perda de um data center.
- Fiabilidade
 - Tendo por base os recursos descritos, acreditamos que o HPE ServiceGuard seja uma solução fiável, podendo ser uma possível escolha. Além disso, é uma empresa com experiência na área, aumentando ainda mais a fiabilidade.
- User interface
 - Interface simples, user-friendly, sendo acedida através do browser.
- Facilidade de gestão
 - Com a interface disponibilizada, acedida através do browser, é possível fazer uma gestão do cluster e das suas componentes. Com alguma prática, irá tornar-se cada vez mais fácil utilizar o software.
- Requisitos de software
 - Exclusivo para Linux
 - Java7 ou Java8 presentes no sistema para realizar a instalação do software

Corosync Cluster Engine com Pacemaker

- Enquadramento do Produto:
 - É um sistema de comunicação em grupo que permite que os membros de um cluster comuniquem entre si, de forma confiável, propaguem mudanças no estado partilhado, saibam quando os membros entram ou saem do cluster e formem quóruns.
 - Costuma ser utilizado com o Pacemaker que controla o comportamento do cluster.
 - É um programa de código aberto e pode ser obtido gratuitamente através do site oficial ou no GitHub.
- Dificuldade de implementação:
 - Varia dependendo do cenário e dos requisitos do cluster.
 - É sempre necessário instalar e configurar o Pacemaker para que tenhamos um cluster de alta disponibilidade.
- Escalabilidade:
 - É escalável na medida em que permite ajustar o tamanho e capacidade do cluster tendo em contas as necessidades do utilizador.
- Cluster híbrido:
 - O Corosync, por si só, não é um cluster híbrido, uma vez que apenas cria a comunicação entre clusters. Mas se estiver a trabalhar em conjunto com o Pacemaker, pode ser considerado um cluster híbrido se este estiver a gerir containers, serviços em nuvem ou máquinas que estejam alojadas em diferentes locais ou redes.
- Recursos:
 - Modelo de comunicação de grupo fechado com garantias de sincronização virtual para criar máquinas de estado replicadas.

- Serviço de membros de cluster que detecta e notifica sobre mudanças na composição do cluster, como a entrada ou saída de nós.
- Serviço de sincronização que permite que os processos do cluster sejam sincronizados em pontos específicos.
- Serviço de quorum que determina se o cluster tem um número suficiente de nós ativos para funcionar corretamente.
- Interface C API que permite aos programadores usar os recursos do Corosync nas suas aplicações.
- Utilização do Pacemaker ou de outro gestor de recursos de cluster.
- **Fiabilidade:**
 - A fiabilidade do Corosync depende de vários fatores, como a configuração do cluster, a rede, o hardware e o software. O Corosync oferece recursos para implementar alta disponibilidade em aplicações, mas não garante que o cluster seja sempre funcional ou resiliente a falhas.
- **User Interface:**
 - A User Interface do Corosync depende do gestor de recursos de cluster escolhido. Existem duas opções:
 - Interface Web do Pacemaker onde é possível criar e configurar clusters. É uma interface intuitiva e user friendly;
 - Linha de comandos
- **Facilidade de Gestão:**
 - Uma vez que o Corosync não faz a gestão do cluster, temos que avaliar o Pacemaker. Provavelmente ao início não será tão simples para quem não esteja familiar com clusters mas para quem já está familiarizado é fácil.
- **Requisitos de Software:**
 - O Corosync pode ser instalado em qualquer sistema operativo e, para além disso, são necessários os pacotes corosync e haveged. Para a gestão, é necessário instalar o pacemaker.

Análise final e solução escolhida

Com base nos diferentes pontos analisados acerca de cada software, criamos uma tabela (em anexo) em que classificamos cada um numa escala de 1 a 4, sendo 1 o melhor e 4 o pior, de modo a facilitar na escolha do software de clustering a implementar na parte prática do projeto.

Chegámos então à conclusão que de todos, o SUSE Linux Enterprise High Availability Extension (SLEHAE) é o que se adequa melhor aos nossos objetivos e arquitetura. Em praticamente todos os pontos, este software apresentou excelentes resultados, especialmente nos recursos disponíveis e facilidade de gestão de clusters.

Os principais recursos que o SLEHAE tem para nos oferecer é o anteriormente descrito Cluster test drive, que nos será útil para simular diretamente no software situações de failover e apresentar o comportamento da nossa implementação de cluster no contexto académico, os Resource Agents, que nos ajudaram a integrar mais facilmente software 3rd party com o cluster, Node Recovery, isto é, a possibilidade de criar snapshots dos nodes para recovery é de grande importância e uso, e por fim DRBD que facilitará a transição de serviços de nodes afetados para nodes com uso de mirroring.

Esta classificação é passível de ser alterada por outros, de modo a satisfazer diferentes requisitos.

Tabela de Classificações

	Enquadramento	Implementação	Escalabilidade	Híbrido	Recursos	Fiabilidade	User interface	Gestão	Requisitos
SUSE Linux Enterprise High Availability Extension	1	1	1	1	1	1	1	1	4
InfoScale Availability	3	3	4	3	4	2	1	3	3
HPE Serviceguard	2	2	2	1	3	2	1	2	1
Corosync Cluster Engine	4	4	3	4	2	4	1	4	2

Referências

https://documentation.suse.com/sles/15-SP4/pdf/article-installation_en.pdf

<https://www.suse.com/products/highavailability/>

<https://buy.hpe.com/pt/en/software/high-availability-disaster-recovery-software/serviceguard-software/serviceguard-for-linux/hpe-serviceguard-for-linux/p/3760>

<https://www.veritas.com/availability/infoscale>

<https://corosync.github.io/corosync/>