**Conduta do Administrador de Sistemas**

• A integridade do AS está acima de suspeita

• O AS não deve interferir nas liberdades dos utilizadores

• O AS deve dialogar com todos os interlocutores com o máximo respeito e profissionalismo

• O processo de aprendizagem do AS é contínuo

• A ética profissional do AS deve ser exemplar

• O profissionalismo do AS deve estar presente em todas as atividades

**Como Construir uma Infraestrutura Computacional**

**1 - Controlo de Versões (git, CVS, SVN...)**

• Registar a evolução dos processos/programas.

• Permitir retroceder para versões anteriores - Change Management.

• Gestão de diferentes sistemas.

**2 - Gestão de incidentes/alterações (GNU, keystone...)**

• Registar problemas/pedidos dos utilizadores.

• Documentar a resolução dos mesmos.

• Ferramenta de apoio na resolução de problemas semelhantes.

• Útil na resolução de conflitos.

**3 - Gold Server - Centralização de Configurações**

• Servidor de configuração de todas as máquinas da infraestrutura.

• As acções são reproduzíveis, recuperáveis, traceable, portáveis e integráveis noutras infraestruturas.

• Migração do “administrador de Sistemas” para “engenheiro de infraestrutura”

• Passivo. Os clientes fazem push da informação.

**4 - Instalação automática**

• Métodos para instalação automática/ sem intervenção humana.

• O mesmo método instalação/ administração independente de servidor ou desktop.

• Imagens de instalação, patches, scripts de gestão, ficheiros de configuração sempre obtidos do gold server.

• Instalação por omissão do sistema + rc.local para atualizações.

**5 - Resolução de Problemas Graves / Falhas na Instalação**

• Deve evitar-se o uso de atualizações por upload e ferramentas r/s-\* como ferramentas usuais de administração.

• Tornam cada máquina um caso particular e por isso mais difícil de substituir.

• Aumentam o TCO e dificulta a recuperação em caso de falha grave

• Ferramentas ad-hoc só para forçar a utilização do gold-server/ correcção problemas graves da instalação

**6 - Como Construir uma Infraestrutura Computacional**

• Serviços de Diretoria -- LDAP, DNS , DHCP

• Servidores de autenticação -- LDAP, Kerberos

• Sincronização de tempo -- NTP

• Servidores de Ficheiros de Rede -- NFS, AFS, SMB

• Servidores de Replicação de Ficheiros -- SUP

• Acesso a filesystems remotos -- automount, AMD, autolink

• Actualizações do SO -- rc.config, configure, make, cfengine

• Gestão da Configuração -- CFEngine, Bcfg2, Chef, puppet, Quattor, Ansible, Juju, OpenLMI

• Gestão de aplicações de Clientes -- autosup, autolink

• Mail -- SMTP

• Servidores de Impressão -- Linux/SMB to serve both NT and UNIX

• Monitorização -- syslogd, paging

**Armazenamento de dados**

**Requisitos**

• Redundância de nível de host

• Tolerância ao erro e desastres

• Acesso de alta velocidade

• Baixa latência

• Interoperabilidade

• Alta escalabilidade

• Facilidade de gestão

**Modos**

• DAS - armazenamento de conexão direta (Escalabilidade complexo e cara)

• NAS - armazenamento conectado à rede (TPC/IP)

• SAN - rede de área de armazenamento

• iSCSI - Internet SCSI

• FCIP - Canal de fibra sobre IP

**Benefícios do SAN**

• Consolidação de recursos de armazenamento

• Acesso simultâneo por múltiplos hosts

• TCO reduzido

• Dados sem LAN e sem servidores transferências

• Distância máxima entre nós e usar em DR

• Alta performance

• Escalabilidade e flexibilidade

• Servidor cluster

**Uma imagem com mesa

Descrição gerada automaticamente**

**Topologias switch fabric:**

Star

Cascada

Ting

Mesh

tree

**RAID**

• Padrão para organizar grupos de discos

• Endereços: Capacidade, Redundância e Performance

• Diferentes níveis de RAID equilibram esses fatores e preço

**LINEAR**

• Concatenação linear de discos:

• N discos com blocos M

• Obtemos um disco com blocos N \* M

• setores 0 a M-1 estão no disco 0

• setores M a 2 \* M-1 estão no disco 1

• Aborda apenas a capacidade do disco

• Vulnerável à falha de um único disco

**RAID 0 Striping**

• N discos com blocos M

• Obtemos um disco com blocos N \* M

• Cada faixa com blocos F

• Setores 0 a F-1 no disco 0

• Setores F a 2 \* F-1 no disco 1

• Aborda desempenho e capacidade

• Extremamente vulnerável à falha de um único disco

**RAID 1 Mirroring**

• 2 \* N discos com blocos M

• Obtemos um disco com blocos N \* M

• O setor i é lido de qualquer disco

• Cada escrita no setor i deve ser duplicada em cada disco

• Endereça apenas redundância

**RAID 4**

• Reserva um disco para paridade (XOR):

• N discos com blocos M

• Obtemos um disco com (N-1) \* blocos M

• Endereçamento semelhante ao stripping

• Cada gravação força o recálculo do setor em o disco de paridade

• Aborda a redundância com pouco impacto na capacidade

• Tolera a falha de um disco

• O disco de paridade limita o desempenho

**RAID 5**

• Paridade em todos os discos:

• N discos com blocos M

• Obtemos um disco com (N-1) \* blocos M

• Endereçamento semelhante ao stripping

• Cada faixa tem um disco de paridade diferente

• Melhor redundância com um pequeno impacto sobre capacidade

• Tolera a falha de um disco

**RAID 0+1**

• Striping antes mirror

• Stripe os disco em 2 grupos

• A falha de um disco desliga uma stripe

• A falha de um disco na outra stripe perde dados

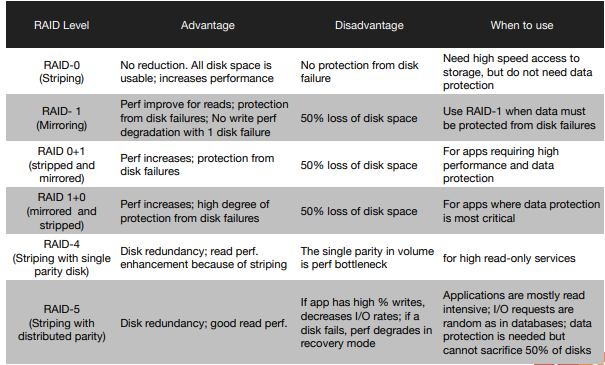
**RAID 1+0 (10)**

• Mirror antes Striping

• Stripe em discos

• A falha de um disco não fecha descendo a faixa

• Duas falhas dos discos espelhados perdidos dados



**QUAL RAID?**

• Linear inadequado para a maioria dos formulários

• RAID 0 extremamente vulnerável a falhas

• RAID 1 não tira vantagem de paralelismo de discos

• RAID 4 e 5 têm desempenho limitado:

• Com aplicações de grandes quantidades de dados gravados durante a recuperação

• Aplicativos com alto compartilhamento de leitura (> 3/4)

- RAID 5:

• Servidores de arquivos

• Servidores web

• Aplicativos com alta participação de gravações RAID 1 + 0:

• BD OLTP

**Partições**

• Divisão (estática) de um dispositivo físico em diversos dispositivos lógicos

• Uma partição ocupa sectores contíguos

• Utilidade:

• diferentes sistemas operativos

• alocação de espaço para diferentes finalidades dentro do mesmo sistema operativo (quotas, users, log, sistema, swap, ...)

• facilidade de manutenção (backups, atualizações, ...)

**Causas de indisponibilidade**

**• Modificar o tamanho de partições:**

• paragem das aplicações

• umount das partições em causa

• na melhor das hipóteses, modificação directa do tamanho da partição (se existir espaço contíguo)

• caso contrário, cópia de segurança seguida de reinicialização das partições e reposição dos dados

• mount das partições

• reinício das aplicações

**• Mover uma partição para um novo disco:**

• paragem das aplicações

• umount das partições em causa

• cópia dos dados para a nova localização

• mount das partições

• reconfiguração do sistema para localizar os dados

• reinício das aplicações

**• Cópias de segurança coerentes:**

• paragem das aplicações

• cópias de segurança

• reinício das aplicações

**Gestão de volumes lógicos**

• Constrói dispositivos lógicos com funcionalidade acrescida:

• isolamento

• redundância

• virtualização

• Funcionalidade com semelhanças à gestão de memória virtual!

**Arquitectura**

• Volumes físicos (PV) são agregados num grupo de volumes (VG):

• o espaço disponível num VG pode ser usado independentemente do PV originário

• Camada que aloca espaço de um VG para volumes lógicos (LV):

• a alocação pode obedecer a diferentes critérios

• Podemos fazer a correspondência entre: • VG e discos • LV e partição

**Capacidade variável**

• O espaço alocado a um LV não precisa de ser fisicamente contíguo

• Um LV pode estar espalhado por vários discos físicos

• Modificar o tamanho de uma partição é sempre possível, desde que haja espaço no VG • É possível acrescentar novos discos a um VG aumentando o espaço disponível para LVs

**Migração**

• Havendo necessidade de substituir um PV:

• acrescenta-se ao VG um novo PV com tamanho suficiente

• mudam-se os sectores do PV original, sendo ocupado o novo PV

• A troca é invisível para os utilizadores

• Não é necessário parar as aplicações nem fazer umount

**Snapshot**

• É possível congelar um LV num determinado instante no tempo, tornando-o num novo LV

• Qualquer modificação ao LV original aloca espaço no VG deixando intacto o LV cópia

• O snapshot pode ser usado para efectuar cópias de segurança

**Coordenação com FS**

• A gestão de volumes lógicos pode ser coordenada com sistemas de ficheiros:

• modificar o tamanho do LV sem fazer umount

• fazer um snapshot coerente sem fazer umount

• A coordenação tem que ser feita modificando o código do próprio sistema de ficheiros

**Cuidado!**

• A falha de um PV pode inutilizar os dados contidos em diversos LVs!

• A utilização de VGs com vários discos só é aconselhada se combinada com RAID

• Com RAID em hardware, usa-se a gestão lógica por cima de RAID

• Com RAID em software, são possíveis ambas as combinações

**Sistema de Ficheiros**

1. **Comandos essenciais que devem estar sempre disponíveis** PASTA BIN
2. **Users de administração essenciais para o arranque, recuperação, restauro e reparação do sistema** PASTA SBIN
3. **Todos os ficheiros necessários para o arranque da máquina** PASTA BOOT
4. **Devices do sistema (discos, placas rede, impressoras, teclados, ratos)**
5. **Deve existir o comando MAKEDEV, para criar os devices caso não existam** PASTA DEV
6. **Ficheiros de configuração e directorias específicas do sistema** PASTA ETC
7. **Configurações do sistema** PASTA SYSCONFIG
8. **Scripts de arranque do sistema e controlo de serviços** PASTAS INIT.D, RC0.D RC5.D
9. **Arranque de serviços do sistema** systemd
10. **Bibliotecas partilhadas necessárias para o arranque e programas localizados em** /bin e /sbin.
11. **Ficheiros temporários das aplicações** MNT, MEDIA, TMP, ROOT
12. **Ficheiros gerados em tempo de execução** var, lib, lock, yp, mail, cache

**Domain Name System (DNS)**

• Tradução hostname em IP e vice-versa

• Hierárquico

• Delegação de autoridade

**Servidores**

• Zonas

• resolução de nomes num domínio

• reverse (zona: c.b.a.in-addr.arpa)

• Servidores recursivos e não recursivos

**Resource Records**

• Owner

• type • A, CNAME, HINFO, MX, NS, PTR, SOA

• class • IN, CH

• TTL

• RDATA • depende do tipo de RR

• SOA start of authority

• MNAME name server responsável

• RNAME email do responsável pela zona

• SERIAL • REFRESH intervalo max para refresh pelos secundários

• RETRY espera antes do próx. refresh qdo falhou

• EXPIRE intervalo a partir do qual a zona passa a inválida

• MINIMUM TTL mínimo

• NS auth. name server

• A hostname-to-ip map

• CNAME host alias • HINFO

• MX Servidor de mail

• PTR ip-to-hostname map

• SRV service advertise

• AAAA ipv6 host-to-ip map

**Dynamic Host Configuration Protocol DHCP**

• Automatização da configuração TCP/IP

• DHCP lease

• Atribuição de IP estática ou dinâmica

• Atribui IP a partir do MAC address (12 digitos hexadecimais)

• RFC 2131, 2132, 2939, 1497, 2610, 3004, 3679, 3046,2241, 3118, 2485, 3442, 3594, 3825

**Mensagens**

• DHCPDISCOVER - Transmissão do cliente para localizar servidores disponíveis.

• DHCPOFFER - Servidor para cliente em resposta a DHCPDISCOVER com oferta de configuração parâmetros.

• DHCPREQUEST - Mensagem do cliente para os servidores

• (a) solicitação de parâmetros oferecidos de um servidor e implicitamente recusando ofertas de todos os outros

• (b) confirmar a exatidão de alocados anteriormente endereço após, por exemplo, reinicialização do sistema, ou

• (c) estender o aluguel em uma rede particular endereço.

**Opções**

• [128, 254] reservadas para opções locais

• 1 - subnetmask

• 2 - time offset

• 3 - router

• 4 - time server

• 6 - domain name server

• 12- host name

• 15 - domain name

• 28 - broadcast address

• 33 - static routing

• 42 - NTP Server

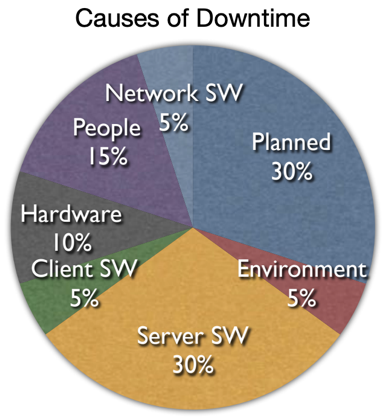
• 44 - Netbios over TCP name server

• 46 - Netbios node type

• 69 - SMTP server

• 70 - POP3 server

**Data Center Basics**



**Sem tempo para Tempo de inatividade**

**Requisitos**

• Disponibilidade 24 horas por dia

• Rápido crescimento de dados

• Dependência de dados

• O tempo de inatividade custa e a sobrevivência do negócio

• Confiabilidade

• Disponibilidade

**O custo da indisponibilidade**

• Desliga o negócio

• Diga aos clientes para irem para outro lugar

**Causas de tempo de inatividade planeado**

• Backup

• Substitua ou atualize o hardware

• Atualização de software ou aplicativo

• Manutenção ou reconfiguração de software

• Atualização do sistema operacional

• Instalação de patch

**Causas de tempo de inatividade NÃO planeado**

• Tempo de inatividade planejado estendido

• Erro humano

• Falha do aplicativo

• Falha do sistema operacional

• Falha de hardware, como disco, CPU, memória

• Incompatibilidade / conflito entre a aplicação parâmetros

**The HA Continuum**

High Availability Clusters (HAC)

• Clusters de dois ou mais nós com uma série de interconexões, como discos compartilhados e pulsação privada rede.

High Performance Clusters (HPC)

• Também chamados de clusters de computação paralela. Cada cluster tem um

grupo de computadores, ligados para trabalhar ao mesmo tempo num problema, não como backups um do outro.

Disaster Recovery (DR)

No contexto das aplicações online, é um período prolongado de interrupção de serviço ou dados de missão crítica, causada por eventos como incêndio e ataques terroristas que danificam a instalação inteira.

**HA Metrics**

Mean Time Between Failures (MTBF)

• O intervalo de tempo médio (geralmente em horas) entre duas falhas consecutivas de um certo componente ou sistema

Mean Time To Repair (MTTR)

• O tempo médio necessário para completar uma ação de reparo

**Confiabilidade de Redes de Dados**

**Entraves à Confiabilidade da Rede**

• Falhas são transientes

• A definição de rede é difusa

• Aumentos de latência podem confundir-se com falhas

• Ataques por negação de serviço (DoS)

• Balanceamento aumenta a escalabilidade não a confiabilidade

• Falha de segurança no acesso

**Tipos de Falha da rede**

Interface local

Cabos

Falha de serviços fundamentais

Latência

**Falhas Físicas**

Complicadas de diagnosticar

Cabo estragado

Falha num switch

Falha numa placa de rede

**Falhas na camada IP**

Problemas de configuração

Problemas de routing

**Orientações para o Desenho e Manutenção da Rede**

Conhecer a rede, os equipamentos e os protocolos utilizados

Utilizar routers, filtros e firewalls para protecção

Utilização de Caches

**Endereços IP virtuais**

Distinguir a máquina que fornece um serviço do serviço fornecido

Cada máquina tem um IP/nome (privado/para adm)

**IP Network route**

Clientes seleccionam qual a ligação

Só funciona com nº reduzido de clientes/ servidores

**Orientações para evitar problemas de routing**

• Evitar um único ponto de acesso

• Rotas por omissão desde que exista + que 1

• Redirecção dinâmica de pacotes (ICMP redirect)

• Monitorização e intervenção pro-activa

**Load Balancing & Network Redirection**

• Round-Robin DNS

• Network Redirector

**Monitorização**

Permite identificar quais os servidores/serviços operacionais

Antecipar falhas e agir pro-activamente

**Algoritmo genérico**

Periodicamente, cada monitor obtém uma resposta do monitorizado: o monitor faz periodicamente pedidos

**ICMP Ping/Echo**

O pedido é um pacote ICMP (ping)

A resposta é enviada pelo kernel (echo)

Verifica que:

a máquina está ligada

a rede está fisicamente ligada

**ATENÇÃO**

A existência de resposta ao pedido de monitorização apenas verifica a disponibilidade dos componentes utilizados na transmissão e processamento do pedido de monitorização!

É frequente um computador responder a ICMP ping mesmo quando não está em condições de efectuar trabalho útil!

**UDP Echo**

**Cuidado!**

Se a máquina ou a rede estão sobrecarregadas mas não indisponíveis, a monitorização pode errar

O pedido é um pacote UDP para a porta 7

A resposta é enviada pelo xinetd

Verifica que: o xinetd está a correr

**UDP Heartbeat**

Usa um servidor próprio, o pedido pode ser implícito

Funcionamento: periodicamente é enviado um pacote para o monitor

**HTTP HEAD**

A resposta são os cabeçalhos, incluindo a data de modificação

Verifica que: o httpd está correctamente configurado e a correr

**HTTP GET e POST**

A resposta são os cabeçalhos e os dados

**Estratégias para a monitorização**

Uma máquina monitoriza-se a si própria

Uma máquina monitoriza outras máquinas

Várias máquinas monitorizam-se a si próprias

**Watchdog**

Hardware especializado que automaticamente reinicia a máquina em caso de indisponibilidade

Funcionamento: escrever algo periodicamente para /dev/ watchdog

**fping**

Para monitorização de outra máquina

Monitorização usando ICMP ping/echo de um conjunto de máquinas

Funcionamento: envia um conjunto de pings

**Linux-HA mon**

Monitorização de aplicações com módulos para:

detectar indisponibilidade ao nível da aplicação

**ldirectord**

Monitorização ao nível da aplicação de serviços http, https e ftp

**Linux-HA Heartbeat**

Para monitorização mútua

Produz periodicamente pacotes em UDP multicast ou broadcast

**Disco partilhado**

Um sector de disco por cada máquina onde é guardado um inteiro

Cada máquina periodicamente: lê todos os sectores

**Split-Brain**

Se ao fazer monitorizaçao mútua o sistema de monitorização falha, cada uma das máquinas considera a outra indisponível

**Coerência**

Monitorização para escolha de um líder

Exemplo: Escolher qual o servidor num cluster que pode escrever num disco partilhado

**I/O Fencing**

Permitir que uma máquina impeça outra de aceder a um recurso comum (e.g. um disco)

**STOMITH / STONITH**

Cada máquina controla a fonte de alimentação de outra

Após uma suspeita de indisponibilidade, a máquina é desligada

**Data Center Arquitetura**

**Requisitos**

• Fornece um local físico seguro para servidores, armazenamento e equipamento de rede.

• Fornece conectividade de rede 24x7 para equipamentos dentro do data center para dispositivos fora dos dados Centro

• Fornece a energia necessária para operar todos os equipamentos

• Fornece um ambiente onde a temperatura e a umidade são controladas dentro de uma faixa estreita e ar trocado em uma taxa adequada.

**Área Física Requerida para Equipamentos e espaço desocupado**

• servidores

• armazenamento

• dispositivos de rede

• HVAC (aquecimento, ventilação e ar condicionado)

• painéis de energia

• disjuntores

• piso para suportar a altura do equipamento

**Energia necessária para Execute todos os dispositivos**

• A UPS deve estar presente para proteção contra falhas de energia.

• Geradores podem ser necessários se as interrupções forem frequentes e demorem

grandes.

**Resfriamento necessário e HVAC**

• Mantenha os dispositivos frios

• Manter baixa umidade no data center

• Plano para o crescimento

**Required Weight**

• Determine o peso de todas as prateleiras vazias

• Determine o peso de todos os equipamentos

• Determinar se o data center existente chão é forte o suficiente para o peso

**Guias de planeamento a Centro de dados**

• Planear de antemão

• Plane para o pior

• Plano de crescimento

• Simplifique seu design

**Centro de dados Estruturas**

• Piso não elevado ou elevado

• Corredores

• Rampa

**Best Practices**

• O design deve ser modular

• Rotule tudo

• Documente tudo

• Isole os cabos

• Use telhas de alumínio fundido

**Points of Distribution (PODs)**

• Rack com dispositivos

• switches de rede

• servidores de terminal

• painéis de remendo

**Centro de dados Manutenção**

**Operações de Rede Centro (NOC)**

Instalação dedicada com pessoal que monitoram a disponibilidade de todos dispositivos e serviços dentro dos dados e responder a qualquer data center problemas.

**Monitoramento de rede**

• Coleta de dados em tempo real e sua classificação em desempenho e interrupções

• Usado para prever a necessidade de escalonamento futuro

• Monitoramento in-band e out-of-band

**Data Center Físico Segurança**

Todos os pontos de entrada e portas devem ser controlados por leitores de cartão ou pessoas

que monitoram e restringem o acesso a aqueles que entram no data center

**Limpeza de data center**

• Não requer limpeza regular ou aspirador.

• Materiais de limpeza aprovados.

• Limpeza da superfície do piso. evite cabos perturbadores

**Best Practices**

• Um alto nível de segurança física, de rede e de host deve ser aplicada em todos os momentos.

• A limpeza regular do data center é importante para remover toda poeira e sujeira, o que pode danificar os dispositivos.

• Inspeção física do local todos os meses.

**Power Distribution**

**Estimar Necessidades de energia**

• Alimentação para racks e autônomo equipamento

• Energia para o sistema HVAC

• Geradores

• UPS

**HVAC**

• Deve ser redundante para lidar com falhas e manutenção

• Deve ser eficiente. Representam 40 a 60% de a energia usada no DataCenter

Optimal range 21ºC to 23ºC 45-50%

Acceptable Range 10ºC to 32ºC 25-75%

**Faixa estreita em Temperatura e RH**

• Servidores e armazenamento geram quantidades substanciais de calor

• Certas áreas exigem mais refrigeração

• O tempo é imprevisível

• Mudanças de equipamentos dentro do data center

• Numerosas trocas de ar

• Alta UR causa corrosão e curtos-circuitos

• Baixa UR causa problemas de ESD

• Atividades de data center que podem afetar a temperatura perfil

**Sistemas de ar condicionado**

• Ar condicionado frio-líquido para regiões quentes e desérticas

• Ar-condicionado seco para regiões úmidas

• Circulação de ar

**Best Practices**

• Deve haver corredores frios e quentes. Os racks devem ser colocados costas com costas apontando para o corredor quente e frente a frente voltada para o corredor frio

• Os dispositivos são projetados para expelir seu próprio calor em um ótimo maneira. O padrão de fluxo de ar deve coletar ar quente e transportá-lo para o HVAC

• Diferentes níveis de perfuração para azulejos no equilíbrio do corredor frio

**Key points**

• As cargas de calor dentro de cada data center continuam a aumentar em um ritmo rápido

• O objetivo do HVAC é fornecer o suficiente fluxo de ar frio para manter uma temperatura que

atende aos requisitos de equipamento

• Sistemas de ar condicionado líquido frio e seco são comumente usados para resfriamento DC

• Os data centers são organizados em frio e quente corredores