

DEPARTAMENTO DE COMUNICAÇÃO, ARTES  
E TECNOLOGIAS DA INFORMAÇÃO

*INFORMÁTICA DE GESTÃO*

## Trabalho Final de Curso

*Ano lectivo 2009/2010 – 2ºSemestre*



### Framework para backup de dados de negócio, aplicado a um exemplo simplificado da CSJD

Alunos:

Nº: a53067    Pedro Santos Bessa    Email: [pbessa@norvico.pt](mailto:pbessa@norvico.pt)

---

Nº: a20071174    José Saraiva Sousa    Email: [saraiva.sousa@edp.pt](mailto:saraiva.sousa@edp.pt)

---

Professor Orientador:

Engº Mestre Sérgio Guerreiro

---

## Índice

Abstract.....	6
Sumário.....	7
1. Introdução .....	7
1.1 Os custos resultante da perda de informação.....	8
2. Enquadramento teórico.....	8
2.1 Objectivo .....	8
2.2 Âmbito.....	9
2.3 Enquadramento .....	10
2.3.1 ETL - Extract, transform and load.....	10
2.3.2 Opções técnicas .....	10
3. Diagrama Global de Actividades.....	14
3.1 Pressupostos .....	15
3.2 Ferramentas .....	15
3.2.1 Robocopy.....	15
3.2.2 NTBackup.....	16
3.2.3 Blat .....	17
3.2.4 Composição das máquinas virtuais.....	17
3.2.5 MySQL e MSSQL .....	19
3.2.6 7Zip.....	19
3.2.7 Plataforma de Monitorização.....	20
3.3 Metodologias e Casos de Uso.....	21
3.3.1 Plataforma de Monitorização.....	22
3.3.2 Plataforma de Monitorização - Internet Information Service.....	23
3.3.3 Normalização dos artefactos produzidos pelo TFC.....	23
3.3.4 Programador de tarefas .....	25
3.3.5 Tarefa de Restauro.....	26
3.3.6 Identificação de espaço em disco por servidor .....	27
3.3.7 Transformação dos dados .....	28
3.3.8 Off-line Storage Backup.....	29
4. Resultados.....	31
4.1 O Sistema de Extracção, Transformação e Carregamento (ETL) .....	32
5. Conclusões e Trabalho Futuro.....	38
5.1 Conclusões.....	38
5.2 Trabalho futuro .....	39
6 - BIBLIOGRAFIA.....	40

## Índice de Figuras

Figura 1– Diagrama Global de Actividades .....	14
Figura 2– Ferramenta de Backup “NTBackup” .....	16
Figura 3– Configurações e papéis dos Servidores virtualizados .....	18
Figura 4– Representação esquemática do ambiente virtualizado e em rede .....	18
Figura 5– VMWare Player, utilizado para executar as máquinas .....	19
Figura 6 – Caso de Uso para o Actor “Administrador de SI”[15].....	20
Figura 7– Caso de Uso para o Actor “Operador de SI”.....	21
Figura 8– Plataforma de Monitorização – duas representações .....	22
Figura 9– Estrutura de directórios para as Bases de Dados.....	23
Figura 10 – Estrutura de directórios para o “Estado do Sistema” .....	23
Figura 11– Estrutura de directórios para os Shares de Rede .....	24
Figura 12– Estrutura de directórios para os Scripts a executar .....	24
Figura 13– Estrutura de directórios para o armazenamento de dados .....	24
Figura 14– Estrutura de directórios para o armazenamento de logs.....	25
Figura 15– Caso de Uso para o Actor “Operador de SI” .....	25
Figura 16– Diagrama de Actividades Operador de SI IV .....	26
Figura 17– Caso de Uso para o Actor “Operador de SI” .....	27
Figura 18– Caso de Uso para o Actor “Backup System” .....	27
Figura 19 - Representação do esquema a implementar para identificação dos espaços em disco	28
Figura 20 – Macro em Visual Basic - Fluxo de execução.....	29
Figura 21– Caso de Uso para o Actor “Backup System” .....	30
Figura 22– Estratégia para Off-line Storage.....	30
Figura 23– Diagrama de Instalação, representando cada Máquina Virtual.....	32
Figura 24– Representação global dos resultados obtidos [14] .....	33
Figura 25– Diagrama de Actividades Backup System Server XPTO .....	34
Figura 26– Execução de um processo de backup a um “Estado de Sistema” .....	34
Figura 27– As Bases de Dados são depositadas nos respectivos directórios .....	35
Figura 28– Caixa de correio do Operador, contendo os logs diários de todas as acções .....	35
Figura 29– Os logs são colocados conforme o esquema apresentado .....	36
Figura 30–PDF, contendo informação rica, depositado na Plataforma de Monitorização.....	36
Figura 31– Representação do processo de recolha de informação sobre os espaços em disco .....	37
Figura 32– Ficheiro gerado pela acção anterior .....	37

## Índice de Tabelas

Tabela 1 - Robocopy.....	11
Tabela 2 - Estado do Sistema .....	11
Tabela 3 - NT Backup .....	11
Tabela 4 - Blat .....	12
Tabela 5 - Máquinas Virtuais .....	12
Tabela 6 - VMWare Player.....	12
Tabela 7 - MySQL.....	13
Tabela 8 - MSSQL.....	13

## Tabela de abreviaturas e acrónimos

ADO	-	ActiveX Data Objects
ASP	-	Active Server Pages
BD	-	Base de Dados
CJSD	-	Clínica de São João de Deus
Cron	-	Command Run On
CSS	-	Cascading Style Sheets
DW	-	Data Warehouse
ETL	-	Extract, Transform and Load
GUI	-	Graphical user interface
HTTP	-	HyperText Transport Protocol
IIS	-	Internet Information Service
ISACA	-	Information Systems Audit and Control Association
LINUX	-	Linus Torvald's UNIX
MSSQL	-	Microsoft Structured Query Language
RFC	-	Request for Comments
RH	-	Recursos Humanos
ROI	-	Return On Investment
SGBD	-	Sistema Gestor de Base de Dados
SI	-	Sistemas de Informação
SLA	-	Service License Agreement
SO	-	Sistema Operativo
SQL	-	Structured Query Language
SRS	-	Software Requirements Specification
TCO	-	Total Cost of Ownership
TFC	-	Trabalho Final de Curso
UNIX	-	Uniplexed Information and Computing System
VB	-	Visual Basic
VBA	-	Visual Basic for Applications
VBScript	-	Visual Basic Scripting Edition
VM	-	Virtual Machine
Web (ou WWW)	-	World Wide Web
WS2K3	-	Windows Server 2003

## Abstract

*(The Computer ) ICT world is vulnerable to mistakes, but it is also enabled to correct them. Failures being less and less frequent create a fake notion of security, causing that a vital task/procedure (the backup) for the organization is regarded secondarily. The backup has little visibility, therefore is not chosen among other tasks perceived as having a higher added value.*

*In a globalized world facing an economical crisis, all IT systems are subject to availability and efficiency requirements, being observed from the economical point of view on ROI (return on investment).*

*This peculiar view on IS makes every time more difficult to justify expenditures with a systems perceived as unnecessary. Hardly a backup is easily faced as an insurance: it is seen simply as a cost, till misfortune happens.*

*Through this economical perspective, future is shaped as: «cloud computing». This new approach is expected to decrease considerable TCO (total costs of ownership), to increase ROI and efficiency reducing costs with the service «pay-as-you-go». But that reality implies several risks not to disregard: technical (net connections failures) and security (who does data control and how to retrieve data).*

*With this work we try to demonstrate, in the actual context, that it is possible to create a backup system with progressive capacity, easy to control and maintain, silent for all elements SI included, at low maintaining costs, without neglecting the necessary reliability.*

**Keywords:** backup, restauro, eficiência, salvaguarda, escalabilidade, alarmística, controlo, inovação, racionalização de custos

## Sumário

O mundo da informática está sujeito a erros, mas também está habilitado a combatê-los. As falhas, sendo cada vez menos frequentes, originam uma falsa sensação de segurança que faz com que uma tarefa fundamental para qualquer organização (o backup) seja encarada como algo secundário. Porque pouco visível, o backup acaba por ser preterido por outras tarefas mais facilmente entendíveis como de maior valor acrescentado.

Num mundo globalizado e em crise económica, todos os sistemas de TI estão sob exigência quanto à disponibilidade e eficiência, sendo observados do ponto de vista económico através do retorno do investimento (ROI). Esta visão peculiar dos SI torna cada vez mais difícil a existência de elevados gastos com um sistema que se espera não necessário. Um backup é facilmente encarado como um seguro: é simplesmente um custo, enquanto não temos um azar.

Através dessa perspectiva económica, o futuro começa a ganhar forma: “cloud computing”. Espera-se que esta nova abordagem venha reduzir consideravelmente os custos de propriedade (TCO), aumentar consideravelmente os ROI, para além do aumento da eficiência e uma redução de custos através do serviço “pay-as-you-go”. Mas essa realidade ainda comporta vários riscos não negligenciáveis, tanto sob o ponto de vista técnico (falha na ligação à rede) como de segurança (quem controla os dados).

Com este trabalho pretendemos demonstrar que, ainda no actual contexto, é possível ter um sistema de backup com capacidade evolutiva, fácil de controlar e de manter, silencioso mesmo para os elementos do SI, com baixos custos de manutenção, sem nunca descurar a indispensável fiabilidade.

## 1. Introdução

Um sistema integrado de backup [32] e de restauro [32] é essencial para qualquer empresa, independentemente da sua dimensão. Infelizmente, os responsáveis pelas infra-estruturas de tecnologias de informação (TI), nem sempre têm o orçamento que precisam e os Sistemas de Backup raramente estão adequados às necessidades.

Todos estão de acordo de que as empresas devem ter um Sistema de Backup eficiente mas, o esforço financeiro necessário para que o consigam atingir, nem sempre espelha as reais necessidades [30] - “*You need champagne backup on a beer budget*”.

## 1.1 Os custos resultante da perda de informação

A primeira consideração a fazer é a de saber o tipo de dados que estão a ser salvaguardados. Conhecer o valor intangível da informação é importante, bem como é fundamental saber o custo que representará a indisponibilidade de serviço, por menor que esta possa ser.

Na realidade, quando se fala de informação, de tecnologias e da infra-estrutura de suporte físico, numa abordagem de baixo nível, é considerado o pormenor do bit e do byte, dos 0 e 1, onde as falhas e erros são constantes, suportados por mecanismos de controlo [31] que, fornecem (numa visão de alto nível) uma sensação de falsa estabilidade que leva a comportamentos negligentes e, consequentemente, à perda de informação.

Para evitar essas situações, há que adoptar metodologias de backup, adequadas aos ambientes onde serão implementadas.

Abordagens do tipo Disk-to-disk [30], off-line storage [30] ou tape de backup [30], cada uma representa uma forma de salvaguarda dos dados produzidos diariamente.

O Backup é, em primeiro lugar, o principal mecanismo para a garantia da informação mas, para ter a certeza de que esses mecanismos estão a funcionar como se espera que estejam, há que prever acções de restauro que garantam a efectiva fiabilidade da informação armazenada pelo processo de backup.

Novas abordagens de salvaguarda de dados estão a ser consideradas. Existe uma aposta no conceito remoto de backups, personificado pela abstracção representativa da “*nuvem*” - “The promise of cloud computing is arguably revolutionizing the IT services world by transforming computing into an ubiquitous utility” [33].

## 2. Enquadramento teórico

### 2.1 Objectivo

O propósito deste trabalho é o de propor uma solução detalhada sobre um Sistema Integrado de Backup e sua interacção com uma Plataforma de Monitorização. Esta última componente terá como função a apresentação rica dos conteúdos produzidos pelos processos de backup do sistema integrado.



## 2.2 Âmbito

O Sistema a implementar, contempla um conjunto significativo de variáveis, que integram os sistemas operacionais [29], ferramentas de monitorização, programação de macros [18], envio de emails [8], execução de rotinas de backup [7] e geração de logs [5] e alertas.

Este trabalho pretende ser um contributo importante para a implementação de um conjunto de valências que têm por objectivo a consolidação de processos de backup para um ambiente heterogéneo, permitindo a concentração dos dados produzidos pela infra-estrutura, num repositório final, que garanta a correcta salvaguarda da informação.

Todos estes dados são posteriormente disponibilizados num front-end (denominado de Plataforma de Monitorização [3]) e que permite a todos aqueles que lhe estejam autorizados a aceder, a possibilidade de análise, verificação e validação de todos os processos de backup que ocorrem na infra-estrutura.

Estão habilitados a aceder à plataforma, o Operador dos Sistemas de Informação (SI), o responsável pelos SI e também os próprios sistemas que alimentam a solução, sendo eles (neste momento) protótipos [9, 10], servindo como base de teste para validação de todos os workflows produzidos pelas acções dos backups diários.

Este sistema integrado pretende resolver uma lacuna identificada na Clínica de São João de Deus (CSJD), permitindo aos seus responsáveis de SI, o pleno conhecimento dos processos de geração de conteúdos e a salvaguarda desses bens produzidos. A identificação de eventuais falhas nos processos de backup serão mais fáceis de compreender e a resolução mais eficiente de executar.

Todos os intervenientes terão o seu papel definido de forma clara, havendo uma objectiva atribuição de tarefas para cada stakeholder (Anexo I – 1.3).

É importante referir que o projecto a que nos propomos (**Definição e implementação de processos de backup – TFC nº44**), que inicialmente contemplava apenas uma componente conceptual de backups, foi redesenhado, no sentido de providenciar um verdadeira ferramenta, embora conceptual (apesar de havermos construído todo o ambiente virtualizado [9, 10] para a inequívoca demonstração da sua viabilidade em ambientes de produção), que seja mais objectiva em relação ao problema apresentado.

## 2.3 Enquadramento

### 2.3.1 ETL - Extract, transform and load.

Este processo ajuda na obtenção dos dados de sistemas-fonte [29] (bases de dados relacionais [28], folhas de excel [18], texto e outros sistemas operacionais, geradores de dados) de qualquer organização, transforma e carrega esses dados num data warehouse que é um repositório de dados centralmente gerido e integrado, que acumula dados dos sistemas de origem e fontes operacionais.

O principal objectivo do ETL [26] é a transformação dos dados de origem para a obtenção de informação valiosa. As ferramentas de ETL são feitas para simplificar o trabalho com a integração de dados, migração e transformação, de acordo com o modelo ETL considerado.

A proposta de TFC apresentada, pretende adoptar este conceito de ETL, numa perspectiva de obter informação valiosa. Para tal, estão presentes as 3 fases do processo (Extracção – Transformação - Carregamento):

- 1 – Em primeiro lugar são processados os dados de todas as origens para um destino definido.
- 2 – O processo anterior é monitorizado, gerando logs por cada evento existente. Esses eventos são rastreados por uma Macro (descrição detalhada no Anexo V) que, mediante a utilização de linguagem de programação, identifica e transforma os dados recolhidos em informação útil.
- 3 – Por fim, executa-se um processo de carregamento desses resultados para um repositório de dados que poderá ser consultado através de uma Plataforma (Anexo III – Source Code - I) elaborada para esse efeito.

### 2.3.2 Opções técnicas

Nas tabelas seguintes apresentam-se as diferentes aplicações utilizados neste trabalho, listando as suas referências, descrição, comandos associados e principalmente a opção tomada face aos concorrentes.

Para detalhes relativos à descrição e concorrentes, consultar "Anexo IV - Documentação Técnica - I".

Para consulta de todos os comandos, para cada um dos utilitários existe um documento guardado no respectivo directório do CD anexo ao TFC – ver índice de Anexos.

### 1 - Robocopy

Referências	Observações	Comandos batch
Windows Server 2003 Resource Kit Tools <a href="http://www.microsoft.com/downloads/details.aspx?familyid=9d467a69-57ff-4ae7-96ee-b18c4790cffd&amp;displaylang=en">http://www.microsoft.com/downloads/details.aspx?familyid=9d467a69-57ff-4ae7-96ee-b18c4790cffd&amp;displaylang=en</a>	Este utilitário foi desenvolvido pela Microsoft Corporation (MSFT) e introduzido pela primeira vez no Windows Server 2000 Resource Kit Tools.	Manual: Robocopy.doc (documento no CD)
<b>Concorrentes:</b> XXCopy: <a href="http://www.xxcopy.com/xxcopy30.htm">http://www.xxcopy.com/xxcopy30.htm</a> . RobocopyPlus: <a href="http://www.robocopyplus.com/">http://www.robocopyplus.com/</a> . XCopy: <a href="http://technet.microsoft.com/en-us/library/bb491035.aspx">http://technet.microsoft.com/en-us/library/bb491035.aspx</a> . Comparação entre Xcopy e Robocopy em <a href="http://www.xxcopy.com/xxcopy30.htm">http://www.xxcopy.com/xxcopy30.htm</a>		
<b>Escolha:</b> <u>Robocopy</u> porque embutido no SO e porque mais fácil de manusear, executando todas as tarefas indispensáveis à operação pretendida.		

Tabela 1 - Robocopy

### 2 - Estado do Sistema

Referências	Observações	Comandos batch
Dados de estado de sistema <a href="http://technet.microsoft.com/pt-pt/library/cc785306(WS.10).aspx">http://technet.microsoft.com/pt-pt/library/cc785306(WS.10).aspx</a> Efectuar cópia de segurança de dados de estado do sistema <a href="http://technet.microsoft.com/pt-pt/library/cc781353(WS.10).aspx">http://technet.microsoft.com/pt-pt/library/cc781353(WS.10).aspx</a>	Ver nos links technet	Ver nos links technet
<b>Concorrentes:</b> Não Aplicável		

Tabela 2 – Estado do Sistema

### 3 - NTBackup

Referências	Observações	Comandos batch
Artigo sobre NTBACKUP: Preserve os dados com uma aplicação que o sistema já possui: <a href="http://technet.microsoft.com/pt-pt/magazine/2005.05.ntbackup(en-us).aspx">http://technet.microsoft.com/pt-pt/magazine/2005.05.ntbackup(en-us).aspx</a>	Ferramenta de backup introduzida nos Sistemas Operativos Windows desde a sua versão XP e manteve-se disponível até à versão Server 2003 R2.	Manual: NTBackup.doc (documento no CD)
<b>Concorrente:</b> Symantec Backup Exec. - Solução comercial, não embutida no sistema, condicionada pela necessidade de existência de licenciamento e pagamento anual.		
<b>Escolha:</b> <u>NTBackup</u> , porque: A escolha deste componente do SO Windows Server 2003 acontece pelo facto de: <ol style="list-style-type: none"> <li>1 – Ser nativo do SO.</li> <li>2 – Satisfazer as necessidade de cópia do estado de sistema, exchange server (se existente).</li> <li>3 – Utilização de linha de comandos, prescindindo do interface GUI, para efeitos de tarefas diferidas.</li> <li>4 – Necessidade em garantir cada estado do sistema dos servidores existentes na infra-estrutura.</li> </ol>		

Tabela 3 – NT Backup

## 4 - BLAT

Referências	Observações	Comandos batch
<a href="http://www.blat.net">http://www.blat.net</a>	O Blat é um utilitário de linha de comando, utilizado para o envio de email através de SMTP Filosofia: • Blat deve ser: pequeno, rápido, fácil de utilizar e RFC Compliant. <a href="http://www.blat.net/?RFC.html">http://www.blat.net/?RFC.html</a>	Manual: Blat.doc (documento no CD)
<b>Concorrente:</b> <b>Sendmail</b> - <a href="http://www.sendmail.org/">http://www.sendmail.org/</a> História: <a href="http://www.icsi.berkeley.edu/cgi-bin/events/event.pl?ID=000155">http://www.icsi.berkeley.edu/cgi-bin/events/event.pl?ID=000155</a> Adaptação para Windows: O sendmail é nativo de unix, apesar de poder ser emulado e executado em ambientes win32.		
<b>Escolha:</b> <b>Blat</b> porque integra uma sintaxe de alto nível, sendo a sua codificação simples de executar em ambientes win32, podendo ser executado através da linha de comandos		

Tabela 4 - Blat

## 5 - Máquinas virtuais - composição

Referências	Observações	Comandos batch
VMware vSphere HypervisorTM (ESXi) <a href="http://www.vmware.com/products/vsphere-hypervisor/">http://www.vmware.com/products/vsphere-hypervisor/</a>	VMware vSphere 4, uma solução que aumenta os limites de expansão de servidores e máquinas virtuais. Com o melhor suporte da indústria para as novas funções de auxílio à virtualização de hardware e um sistema de I/O altamente otimizado, o VMware vSphere 4 é o primeiro sistema operativo do mercado para desenvolvimento da nuvem interna. Essas novas melhorias de arquitectura possibilitam correr as aplicações mais críticas e com maior quantidade de transacções, como SAP, Exchange e suas bases de dados SQL e Oracle, numa nuvem interna 100% virtualizada gerada pelo VMware vSphere 4.	Não aplicável
<b>Concorrente:</b> Citrix – XEN 5.5 <a href="http://www.citrix.com/English/ps2/products/feature.asp?contentID=2300351">http://www.citrix.com/English/ps2/products/feature.asp?contentID=2300351</a> O Citrix XenServer™ é a única plataforma de virtualização de classe empresarial cloud-proven que oferece os recursos essenciais de migração a quente e gestão centralizada de múltiplos servidores sem nenhum custo.		
<b>Escolha:</b> <b>VMware vSphere HypervisorTM</b> porque há total compatibilidade com a ferramenta utilizada para à posteriori, o VMWare Player. Havia a possibilidade de escolher a CITRIX – XEN 5.5 como software concorrente para a criação das máquinas virtuais mas, para a posterior execução das mesmas, seria necessária a conversão das vm's para outros clientes, como sejam o Virtual Box, da Sun, o que tornaria todo o processo mais longo, sem garantir a total compatibilidade fornecida pela VMWare.		

Tabela 5 – Máquinas Virtuais

## 6 - VMware Player

Referências	Observações	Comandos batch
<a href="http://www.vmware.com/products/player/">http://www.vmware.com/products/player/</a>	VMware Player é a maneira mais fácil de executar vários sistemas operativos ao mesmo tempo no seu PC. Com sua interface amigável, o VMware Player torna fácil para qualquer um experimentar o Windows 7, Chrome ou versões mais recentes do Linux, ou criar máquinas virtuais isoladas com segurança e testar o novo software de navegação na web. VMware Player também pode ser usado para executar uma cópia virtual de um computador antigo para que possa reciclar máquinas antigas.	Não aplicável
<b>Concorrente:</b> Sun VirtualBox - <a href="http://www.virtualbox.org/">http://www.virtualbox.org/</a> O VirtualBox é um excelente software de virtualização distribuído em GPL, para uso doméstico e/ou avaliação, pela Sun Microsystems, embora nascido alemão, criado pela empresa Innotek posteriormente adquirida por aquela empresa norte-americana.		
<b>Escolha:</b> <b>VMware Player</b> Utilizamos o VM Player por causa da compatibilidade deste com as virtual machines produzidas pelo ESXi. Estas não necessitam de conversão e correm de forma transparente no VMware Player. Outras tecnologias carecem de conversão antes de poderem ser utilizadas.		

Tabela 6 – VMWare Player

## 7 - My SQL

Referências	Observações	Comandos batch
<a href="http://www.mysql.com/about/">http://www.mysql.com/about/</a>	Com o seu rápido desempenho, alta fiabilidade, facilidade de utilização e enormes reduções de custos, o MySQL é a base de dados open source mais famosa do mundo e uma alternativa segura para a propriedade de software de bases de dados. O MySQL Enterprise oferece um conjunto de aplicações - desde sistemas de transacção empresarial vitais ao comércio electrónico online ou a sistemas de armazenamento, registo e análise de dados.	Não aplicável
<b>Concorrente:</b> Neste caso não se trata de concorrência, mas sim utilização paralela. A escolha pela utilização paralela é fruto da necessidade em retratar o mais possível os ambientes de produção, sejam eles este cliente em particular ou qualquer outro onde a solução possa ser enquadrável.		

Tabela 7 – MySQL

## 8 - MS SQL

Referências	Observações	Comandos batch
<a href="http://www.microsoft.com/portugal/sql/prodinfo/overview/default.aspx">http://www.microsoft.com/portugal/sql/prodinfo/overview/default.aspx</a>	Plataforma de Dados SQL Server O SQL Server é uma solução de dados abrangente, integrada e “end-to-end” que reforça o poder dos utilizadores em toda a organização através da oferta de uma plataforma mais segura, fiável e produtiva para as aplicações empresariais de dados e BI. O SQL Server 2005 proporciona ferramentas poderosas e familiares aos profissionais de TI, bem como aos trabalhadores da informação, reduzindo a complexidade de criar, implementar, gerir e utilizar aplicações empresariais de dados e de análise em plataformas que vão desde dispositivos móveis a sistemas de dados empresariais. Através de um abrangente conjunto de funcionalidades, interoperabilidade com sistemas existentes e automatização de tarefas de rotina, o SQL Server 2005 fornece uma solução de dados completa para empresas de qualquer dimensão	Não aplicável
<b>Concorrente:</b> Não aplicável		

Tabela 8 - MSSQL

### 3. Diagrama Global de Actividades

O diagrama ao lado representa de forma globalizada o sistema a implementar (*Anexo I, secção 3.2.2.1 Visão geral*).

Para a execução de todas as tarefas que compõem o projecto, foi necessário estabelecer metodologias que permitissem satisfazer os requisitos de desenvolvimento e consequente implementação. Cada pentágono a azul representa o capítulo da respectiva metodologia ou caso de uso.

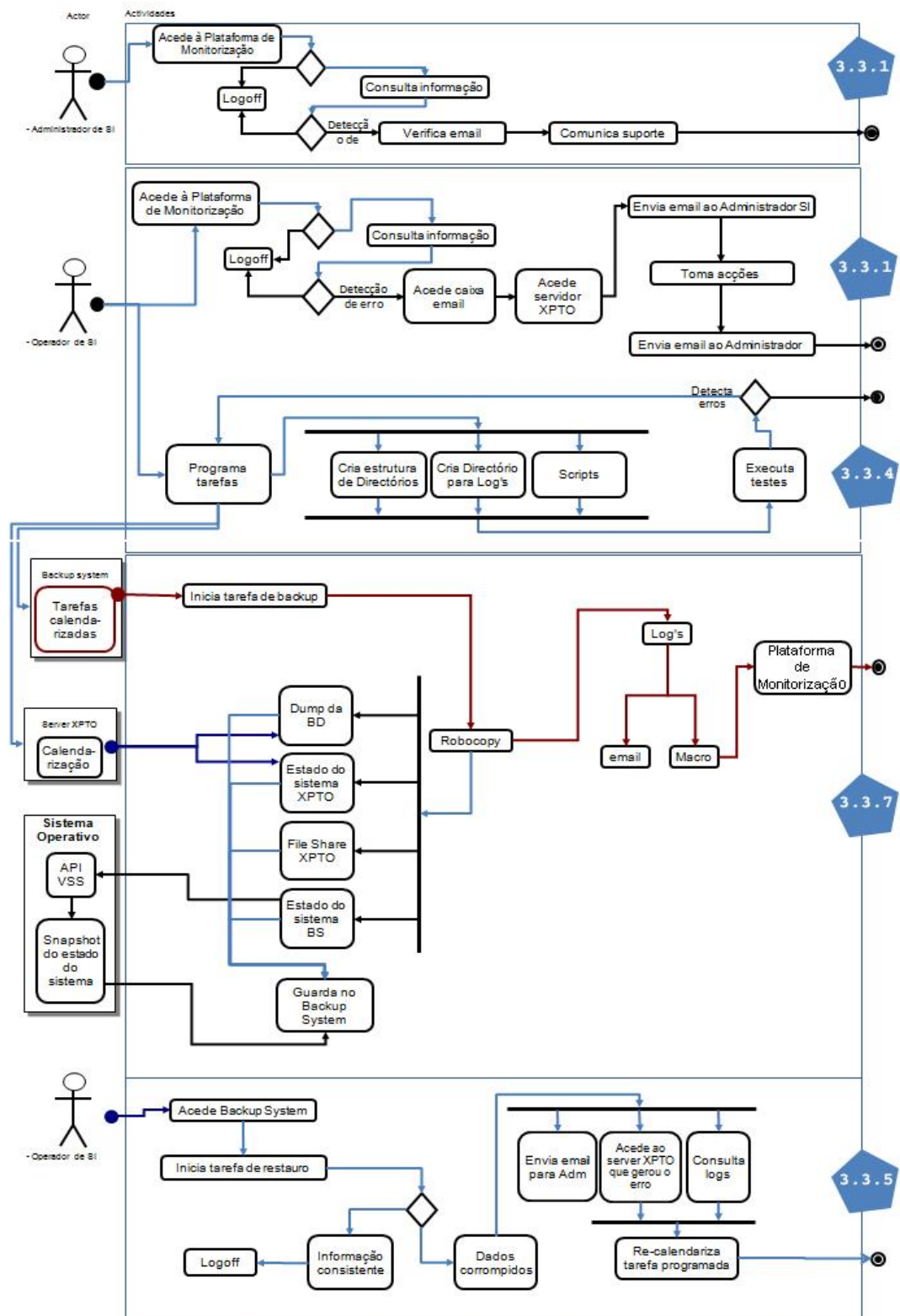


Figura 1– Diagrama Global de Actividades



### 3.1 Pressupostos

Foram considerados para a demonstração dos pressupostos, um conjunto de “ferramentas” indispensáveis para a implementação dessas metodologias (3.3 - Metodologias e Casos de Uso).

### 3.2 Ferramentas

Neste capítulo trataremos de identificar os utilitários contemplados, sua atribuição, contexto e explicação pela opção tomada em relação a outros utilitários existentes e que poderiam desempenhar as mesmas tarefas.

**3.2.1 Robocopy** [5]: Este utilitário foi desenvolvido pela Microsoft Corporation (MSFT) e introduzido pela primeira vez no Windows Server 2000 Resource Kit Tools. Posteriormente, pelo seu comportamento robusto, (Robocopy = Robust file copy) veio substituir um outro utilitário (xcopy - na cópia de dados através da linha de comandos) a partir do Windows Vista (Anexo IV – Documentação técnica – II).

Na actualidade, está presente nos Sistemas Operativos (SO) Windows 7 e Windows Server 2008 R2 [19], de forma nativa. A escolha deste programa em relação a outros (ver 2.3.2 – Opções Técnicas – Tabela 1) aconteceu pela integração com os SO's escolhidos e pela documentação existente sobre os comandos a implementar (Anexo IV – Documentação Técnica - II).

No projecto, o Robocopy será invocado pelo programador de tarefas (Anexo I – 1.3), que contém os scripts de backup (Anexo III – Source Code - III). Quando estes são executados, o Robocopy inicia a cópia dos arquivos (sejam eles bases de dados, estado do sistema (Anexo I – 3.2.1.6) ou shares de rede) das diversas origens para o repositório central no Backup System (Anexo I – 1.3, 2.2.4, 3.2.2.1 – Backup System + Server System). Neste processo, o utilitário regista todos os fluxos de dados em logs que são depositados num directório predefinido.

A estrutura sintáctica para a execução do Robocopy num dos scripts será a seguinte (exemplo):

```
REM @ECHO OFF
REM --- EXECUTA ROBOCOPY E REGISTA LOGS - POR SERVIDOR ---
REM --- BACKUP_SYSTEM ---
robocopy "\\Backupsystem\C$\Backups\SYSTEM_STATE"
"E:\Backups\dados\2feira\Backupsystem\SYSTEM_STATE" /E /Z
/LOG:"E:\Backups\logs\diarios\2feira\BackupSystem_System_State.log" /NP /R:0 /W:0 /MIR
robocopy "\\Backupsystem\C$\share1" "E:\Backups\dados\2feira\Backupsystem\Share1" /E /Z
/LOG:"E:\Backups\logs\diarios\2feira\BackupSystem_Share1.log" /NP /R:0 /W:0 /MIR
robocopy "\\Backupsystem\C$\Inetpub\wwwroot" "E:\Backups\dados\2feira\Backupsystem\wwwroot" /E /Z
/LOG:"E:\Backups\logs\diarios\2feira\BackupSystem_Wwwroot.log" /NP /R:0 /W:0 /MIR
REM --- CSJD001 ---
```

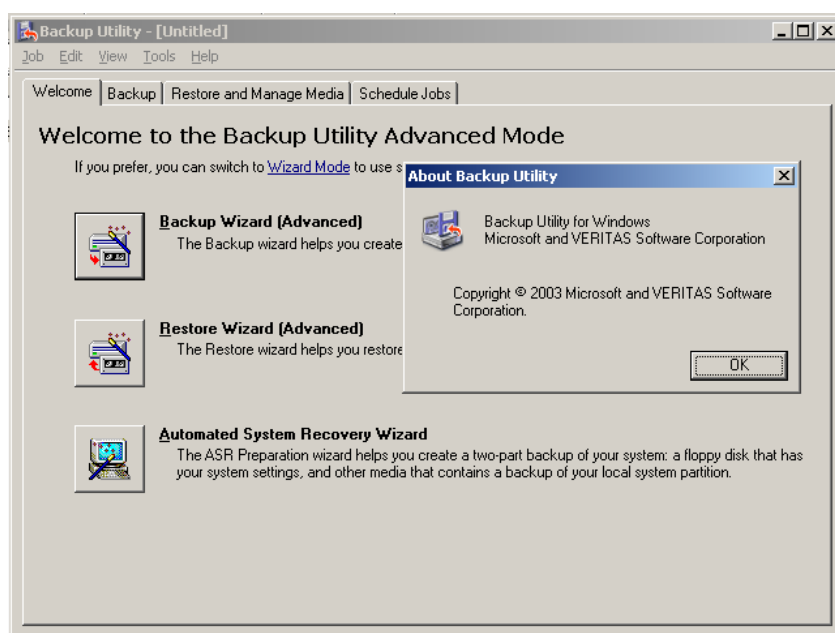
```

robocopy "\\CSJD001\C$\Backups\SYSTEM_STATE" "E:\Backups\dados\2feira\CSJD001\SYSTEM_STATE" /E /Z
/LOG:"E:\Backups\logs\diarios\2feira\CSJD001_System_State.log" /NP /R:0 /W:0 /MIR
robocopy "\\CSJD001\C$\share1" "E:\Backups\dados\2feira\CSJD001\Share1" /E /Z
/LOG:"E:\Backups\logs\diarios\2feira\CSJD001_Share1.log" /NP /R:0 /W:0 /MIR
REM --- CSJD003 ---
robocopy "\\CSJD003\C$\Backups\SYSTEM_STATE" "E:\Backups\dados\2feira\CSJD003\SYSTEM_STATE" /E /Z
/LOG:"E:\Backups\logs\diarios\2feira\CSJD003_System_State.log" /NP /R:0 /W:0 /MIR
robocopy "\\CSJD003\C$\Backups\SQL" "E:\Backups\dados\2feira\CSJD003\SQL" /E /Z
/LOG:"E:\Backups\logs\diarios\2feira\CSJD003_Sql.log" /NP /R:0 /W:0 /MIR
REM --- PREPARACAO PARA ENVIO DE LOGS VIA EMAIL ---
"C:\Program Files\7-Zip\7z.exe" a -tzip "C:\Temp\2feira.zip" "E:\Backups\logs\diarios\2feira"
REM --- ENVIO EMAIL BLAT ---
blat "C:\Temp\2feira.zip" -f "logs@norvico.pt" -to logs@norvico.pt -subject "[TFC] Log 2FEIRA -
BACKUP_SYSTEM" -server smtpa.netcabo.pt -port 25 -u norvico -pw bessal234 -uencode
REM --- APAGAR LOG ZIPADO ---
del /q "C:\Temp\2feira.zip"
call macro.cmd

```

Depois de executado, os logs resultantes da instrução são analisados por um outro componente – Macro (Anexo V) do qual é gerado um documento em pdf, depositado na pasta correspondente ao dia a que diz respeito, podendo ser consultado pelos intervenientes do sistema, através da Plataforma de Monitorização, mediante autenticação prévia. Este processo é acompanhado pelo envio de todos os registos de eventos gerados pela acção, para os emails do Operador e Administrador de SI (Anexo I – 1.3).

**3.2.2 NTBackup [7]:** Este programa, desenvolvido pela MSFT e Veritas, foi introduzido nos Sistemas Operativos Windows desde a sua versão XP e manteve-se disponível até à versão Server 2003 R2. Nos SO's mais recentes, como sejam o Windows 7 ou Windows Server 2008 R2, a ferramenta de backup integra conceitos de cópia de volumes e não somente a componente de ficheiros.



**Figura 2– Ferramenta de Backup “NTBackup”**



A escolha deste componente do SO Windows Server 2003 acontece pelo facto de:

- 1 – Ser nativo do SO.
- 2 – Satisfazer as necessidade de cópia do estado de sistema, exchange server (se existente), garantindo cada um dos servidores existentes na infra-estrutura.
- 3 – Utilização de linha de comandos, prescindindo do interface GUI (Graphical user interface) [18], para efeitos de tarefas diferidas.

Há no mercado um conjunto de aplicações que poderiam ser utilizadas para desempenhar esta tarefa (ver 2.3.2 – Opções Técnicas – Tabela 3) mas, sendo este componente nativo, integrado e gratuito, não se demonstrou necessária outra solução.

A estrutura sintáctica para a execução do NTBackup num dos scripts será a seguinte (exemplo):

```
REM --- BACKUP SYSTEM STATE PARA PASTA LOCAL C:\BACKUPS\SYSTEM_STATE ---  
C:\WINDOWS\system32\ntbackup.exe backup "@C:\Scripts\System_State\SYS.BKS" /n "Backup dia %date%"  
/d "Sys_State do dia %date%" /n "Backup Estado do Sistema dia %date%" /v:no /r:no /rs:no /hc:on  
/m normal /j "CSJD %date%" /l:s /f "C:\Backups\SYSTEM_STATE\Sys_CSJD003.bkf"
```

**3.2.3 Blat** [8]: Este utilitário de distribuição livre é um win32 SMTP mailer, que permite o envio de correio electrónico através de linha de comandos. É uma ferramenta RFC compliant (Anexo IV – Documentação Técnica – I – 4.2.2), compacta e eficiente. Há alternativas a considerar, pelo que estas estão descritas (ver 2.3.2 – Opções Técnicas – Tabela 4).

O blat foi utilizado no projecto para permitir o envio dos relatórios por email (Anexo I - 3.2.1.7 e 3.2.1.8), quer estes sejam resultantes dos backups diários ou da verificação do espaço em disco (Anexo III – Source Code II) nos diversos Servidores da infra-estrutura.

A estrutura sintáctica para a execução do Blat num dos scripts será a seguinte (exemplo):

```
REM --- ENVIO EMAIL BLAT ---  
blat "C:\Temp\Sys_CSJD003.zip" -f "logs@norvico.pt" -to logs@norvico.pt -subject "[TFC] Log  
System State - CSJD003" -server smtpa.netcabo.pt -port 25 -u norvico -pw bessal234 -uuencode
```

**3.2.4 Composição das máquinas virtuais** [9, 10]: Para a implementação do TFC, houve a necessidade de criar um ambiente virtualizado para que fosse possível testar os processos a desenvolver. Para o desempenho desta tarefa, foi necessária a criação de 3 máquinas virtuais para a simulação dos pressupostos.

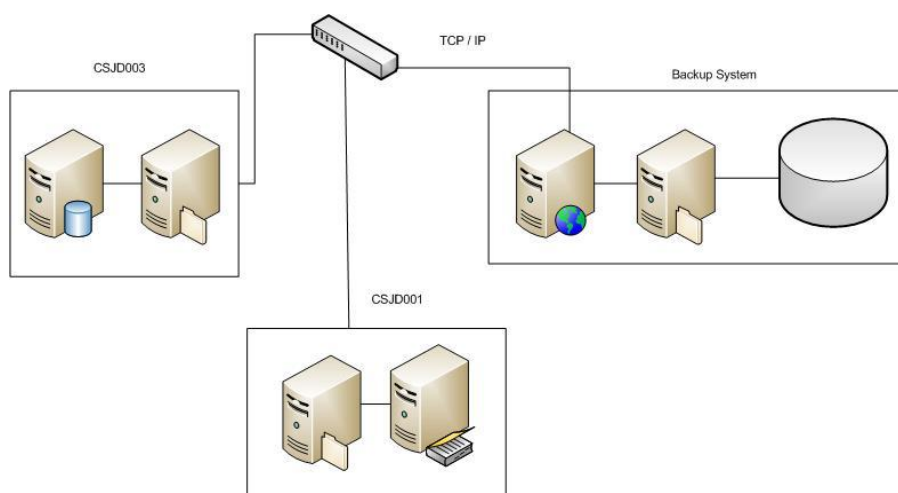
Servidor	IP	Sub-net	Gateway	DNS	Roles
CSJD001	10.0.0.100	255.255.255.0	10.0.0.254	10.0.0.100	DNS Server
				10.0.0.254	DHCP Server
					NTP Server
					File Share
					Print Server
					Active directory

Servidor	IP	Sub-net	Gateway	DNS	Roles
CSJD003	10.0.0.101	255.255.255.0	10.0.0.254	10.0.0.100	Database Server
				10.0.0.254	File Share

Servidor	IP	Sub-net	Gateway	DNS	Roles
BackupServer	10.0.0.102	255.255.255.0	10.0.0.254	10.0.0.100	Backup Server
				10.0.0.254	Plataforma intranet
					IIS6
					File Share

**Figura 3– Configurações e papéis dos Servidores virtualizados**

Foi considerada a solução VMWare vSphere Hypervisor TM (ESXi) [9] para a criação das máquinas virtuais pois posteriormente elas seriam executadas pelo VMWare Player [10].



**Figura 4– Representação esquemática do ambiente virtualizado e em rede**

Havia a possibilidade de escolher a CITRIX – XEN 5.5 (ver 2.3.2 – Opções Técnicas – Tabela 5) como software concorrente para a criação das máquinas virtuais mas, para a posterior execução das mesmas, seria necessária a conversão das vm's para outros clientes, como sejam o Virtual Box, da Sun (ver 2.3.2 – Opções Técnicas – Tabela 6), o que tornaria todo o processo mais longo, sem garantir a total compatibilidade fornecida pela VMWare.

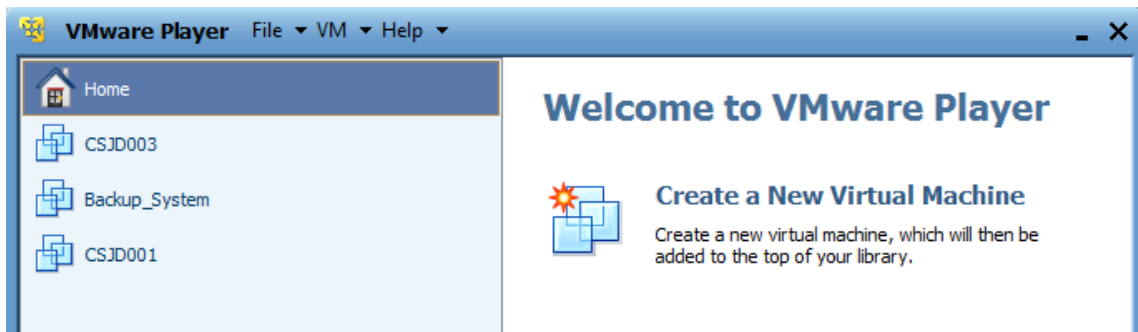


Figura 5– VMWare Player, utilizado para executar as máquinas

As máquinas virtuais ajudaram à construção de todo o modelo e permitiram a correcção de erros, bem como serviram para verificar os pressupostos constantes no documento em anexo (Anexo I – Relatório de Especificação de Requisitos).

**3.2.5 MySQL e MSSQL:** Para conferir consistência ao projecto, tornando-o mais ajustado à realidade das organizações, foi necessário instalar dois conhecidos Sistemas de Gestão de Bases de Dados (SGBD's) para a demonstração de procedimentos de backup relativos às bases de dados de produção. As principais características e referências estão contidas em “2.3.2 – Opções Técnicas – Tabelas 7 e 8”. Foi necessário considerar procedimentos que garantissem a salvaguarda das bases de dados e, através de scripts contendo comandos específicos para cada uma das duas tecnologias, executar os respectivos dumps (Anexo I – 3.2.1.5) das bases de dados simuladas.

A estrutura sintáctica utilizada para a execução destas tarefas será a seguinte (exemplo):

#### MySQL:

```
REM --- DUMP À BASE DE DADOS E REGISTO DA OPERAÇÃO EM FICHEIRO LOG ---
cd "C:\Program Files\MySQL\MySQL Server 5.1\bin\"
mysqldump.exe -u root -p"bessal234" --log-error="C:\Backups\SQL\MYSQL\BD1\classicmodels.txt"
classicmodels >C:\Backups\SQL\MYSQL\BD1\classicmodels.sql
```

#### MSSQL:

```
REM --- DUMP À BASE DE DADOS E REGISTO DA OPERAÇÃO EM FICHEIRO LOG ---
"C:\Program files\Microsoft SQL Server\90\Tools\Binn\osql" -E -S CSJD003\SQLEXPRESS -Q "BACKUP
DATABASE AdventureWorks TO DISK = 'C:\Backups\SQL\MSSQL\BD1\AdventureWorks.bak' WITH INIT" >
C:\Backups\SQL\MSSQL\BD1\AdventureWorks.txt
```

**3.2.6 7Zip[13]:** Este utilitário permite a compressão de arquivos. A escolha deste em detrimento de outros aplicativos do género, como sejam o Winzip ou Winrar (Anexo IV – Documentação Técnica – I - 9) deveu-se à possibilidade de instalação gratuita, quer seja em ambientes académicos ou de produção.

Revelou-se necessária a utilização deste programa, na medida em que era importante a compactação dos resultados obtidos em cada acção e enviar essa informação de forma uniforme. O 7zip, à semelhança de

todos os outros, permite a execução através de linha de comandos, tornando o processo de automatização simplificado.

Toda a sintaxe deste utilitário está contida no “Anexo IV – Documentação Técnica – V”.

A estrutura sintáctica utilizada para a execução destas tarefas será a seguinte (exemplo):

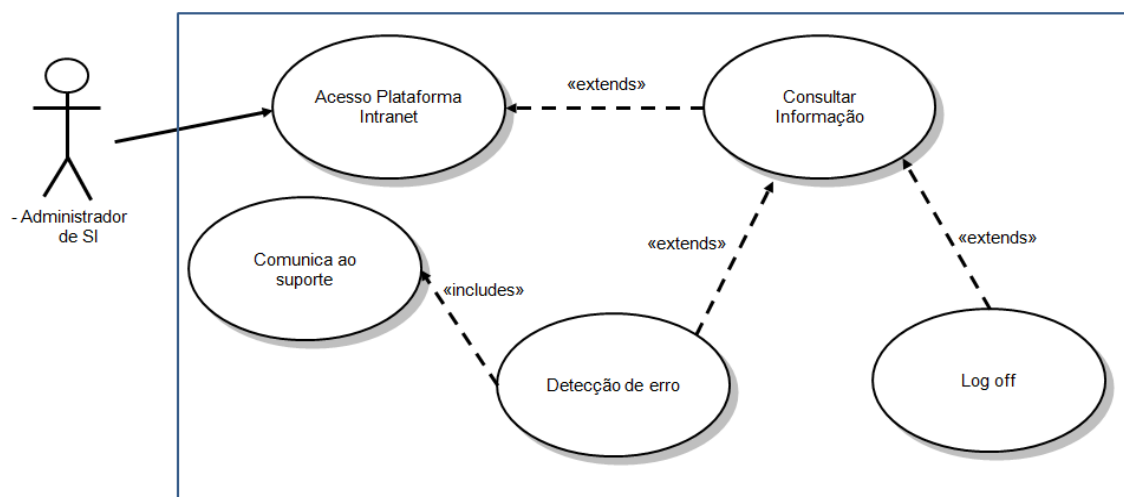
REM --- PREPARACAO PARA ENVIO DE LOGS VIA EMAIL ---

"C:\Program Files\7-Zip\7z.exe" a -tzip "C:\Temp\classicmodels.zip"

"C:\Backups\SQL\MYSQL\BD1\classicmodels.txt"

**3.2.7 Plataforma de Monitorização:** Os dados resultantes dos processos de backup são disponibilizados num front-end, que foi denominado de Plataforma de Monitorização e que permite, a todos aqueles que lhe estejam autorizados a aceder, a possibilidade de análise, verificação e validação de todos os processos de backup que ocorrem na infra-estrutura. Esta plataforma foi desenvolvida em tecnologia ASP (active server pages) [1] e vba (visual basic for applications) [2]. Nativamente, a aplicação criada por M.Blokdijk ([maarten@blokdijk.com](mailto:maarten@blokdijk.com)) [3], não estava habilitada a permitir níveis de acesso. Este protótipo foi adaptado de forma a criar diferentes ambientes em função do login apresentado (Anexo III – Source Code - I). Existem dois tipos de acesso:

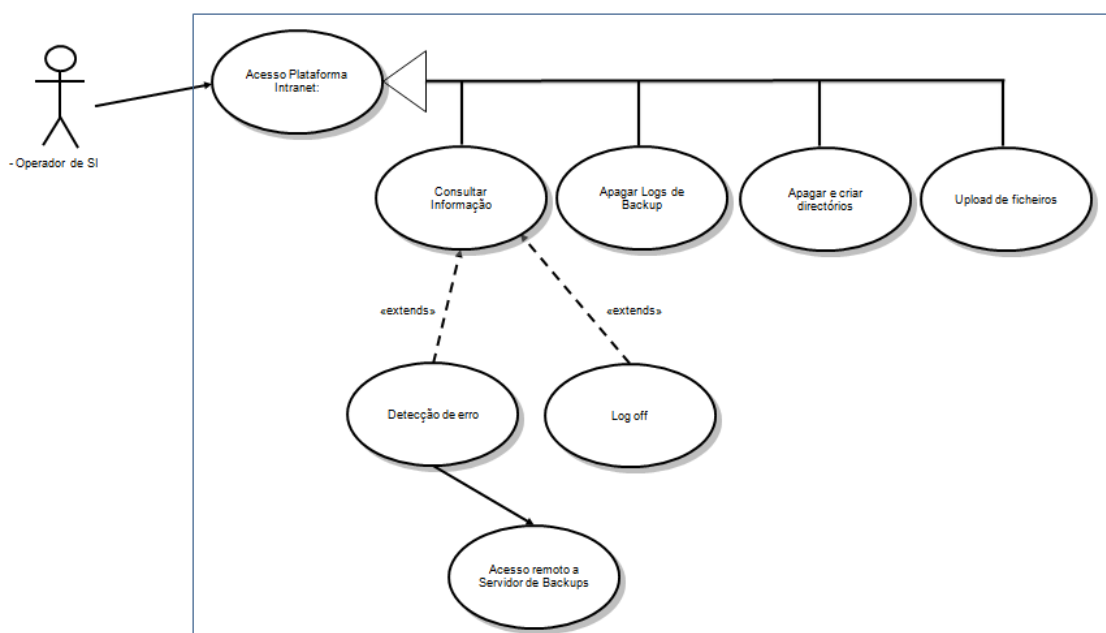
**1 – Acesso condicionado:** O Utilizador (este será o Administrador de SI) executa o login e só poderá consultar a informação decorrente dos processos de backup. Este conjunto de acções está descrito no “Anexo I, 3.2.1.1 e 3.2.1.2”



**Figura 6 – Caso de Uso para o Actor “Administrador de SI”[15]**

Todo o detalhe de interacção entre o Administrador de SI e a Plataforma de Monitorização pode ser consultado no “Anexo I, 2.2.1”

**2 – Acesso Total:** O Utilizador (este será o operacional de SI) não estará condicionado, pelo que poderá apagar ou criar directórios/ficheiros. Poderá, também, consultar os relatórios diários gerados automaticamente pelo sistema interno gerador de relatórios. Pode fazer upload de ficheiros e apagar logs de backup. É responsável por, em caso de erro, aceder ao Backup System, (informação detalhada no Anexo I - 1.3 Definições, acrónimos e abreviaturas – Backup System) sendo também da sua responsabilidade a programação de tarefas e teste de recuperação de ficheiros. Este conjunto de acções está descrito no “Anexo I, 3.2.1.1 e 3.2.1.2”.



**Figura 7– Caso de Uso para o Actor “Operador de SI”**

Todo o detalhe de interacção entre o Operador de SI e a plataforma de Intranet pode ser consultado no “Anexo I, 2.2.2”.

Estão habilitados a aceder à Plataforma, o Operador dos Sistemas de Informação (SI), o responsável pelos SI e também os próprios sistemas que alimentam a solução, sendo estes protótipos que servem de base para teste e validação de todos os workflows produzidos pelas acções de backup diários.

### 3.3 Metodologias e Casos de Uso

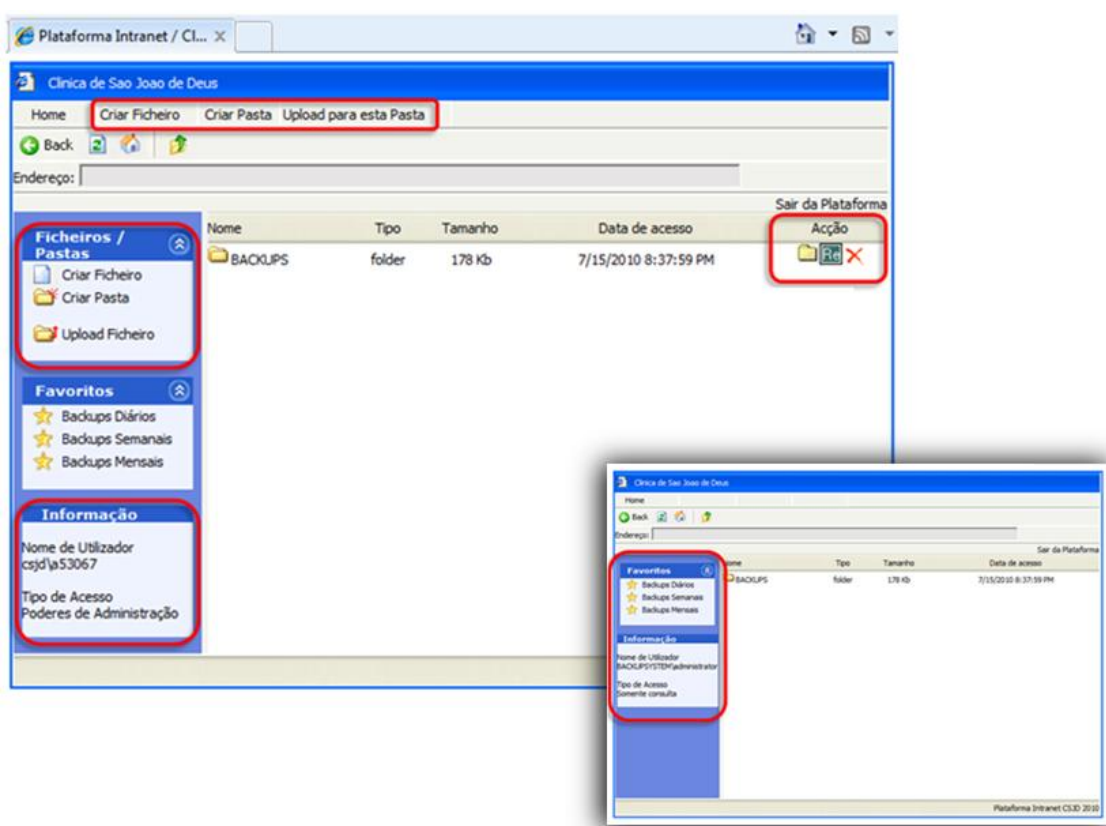
Todos os passos referentes a cada um dos actores mencionados acima, estão referidos de forma detalhada no “Anexo I, 2.2 Especificação dos Requisitos Funcionais”.

### 3.3.1 Plataforma de Monitorização

Na Plataforma de Monitorização, em função do login, diversas funcionalidades estão disponíveis ou desabilitadas. A solução encontrada para a implementação desta condição, foi a criação de código (Anexo III – Source Code - I) que prevê, mediante um conjunto de nomes contidos num ficheiro de texto, carregar essa informação e verificar se o utilizador autenticado está habilitado a ter poderes de administração.

Se este for identificado como Operador de SI, todas as funcionalidades estão disponíveis, caso contrário, apenas serão apresentadas as opções referentes à consulta da informação gerada pelos processos de backup.

A imagem seguinte representa o front-end da plataforma que, em todos os aspectos se assemelha a uma pasta do windows, apesar de ser uma aplicação baseada na web.



**Figura 8– Plataforma de Monitorização – duas representações**

Todas as opções a limitadas pelo vermelho serão dinâmicas, sendo o campo “informações” apresentado a todos os utilizadores, diferindo apenas o conteúdo que, em função do login apresentará o utilizador registado e suas permissões no sistema. As restantes opções a vermelho estão disponíveis apenas aos utilizadores com poderes de administração. Todos os pormenores e acções referentes à plataforma estão descritos no “Anexo I, 3.2.1.1, 3.2.1.2 e 3.2.2.1”.

### 3.3.2 Plataforma de Monitorização – Internet Information Service

Para o suporte à Plataforma de Monitorização, foi necessário contemplar um web service que fizesse o hosting do site. Foi escolhido o Internet Information Service 6 (IIS) [4] por ser nativo do Sistema Operativo Windows Server 2003. Este serviço terá que ficar instalado no Backup System, onde os relatórios serão registados.

### 3.3.3 Normalização dos artefactos produzidos pelo TFC

Para a correcta implementação da solução proposta, foram definidas normas para a criação de estrutura de directórios, nomes de ficheiros, ficheiros batch, locais de colocação dos dumps das bases de dados e estados do sistema de cada servidor.

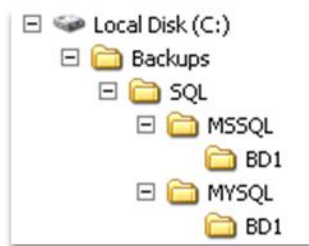
#### *Para Bases de dados:*

Unidade:\Backups\SQL\tipo de base de dados

\MSSQL\BDn\nomebd.sql + ficheiro de log

\MYSQL\BDn\nomebd.sql + ficheiro de log

Exemplo:

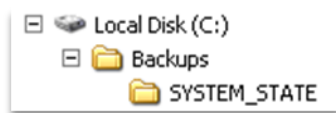


**Figura 9– Estrutura de directórios para as Bases de Dados**

#### *Para SYSTEM\_STATE*

Unidade:\Backups\SYSTEM\_STATE

Exemplo:



**Figura 10 – Estrutura de directórios para o “Estado do Sistema”**

**Para FILE\_SHARES**

Unidade:\share\estrutura organizacional

Exemplo:



**Figura 11– Estrutura de directórios para os Shares de Rede**

**Para SCRIPTS**

Unidade:\scripts\Tipo\SYSTEM\_STATE\_SYS\_CSJD003.CMD (exemplo)

Unidade:\scripts\diario.cmd (chama todos os scripts a executar)

Exemplo:



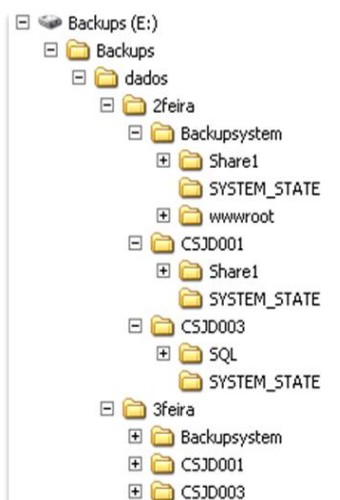
**Figura 12– Estrutura de directórios para os Scripts a executar**

**Estrutura de armazenamento de dados no Backup System:**

Unidade 2:\Backups\dados\2feira - 6feira\servidororigem\SYSTEM\_STATE - SQL (MSSQL - MYSQL) - File\_Shares

Unidade 2:\Backups\dados\1sabado - 5sabado\servidororigem\SYSTEM\_STATE - SQL (MSSQL - MYSQL) - File\_Shares

Exemplo:



**Figura 13– Estrutura de directórios para o armazenamento de dados**

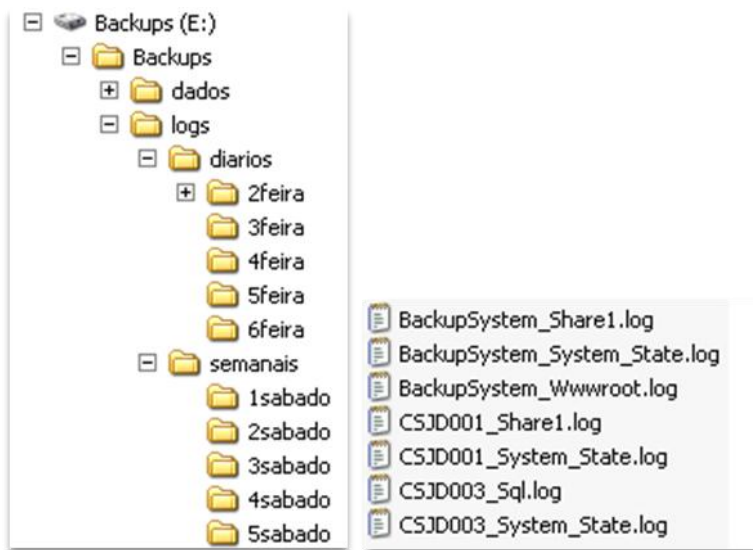


### *Estrutura de armazenamento de logs no Backup System:*

Unidade 2:\Backups\logs\diarios\2feira - 6feira\servidororigem\_tipo\_backup (SQL - SYSTEM\_STATE - File\_Shares)

Unidade 2:\Backups\logs\1sabado - 5sabado\servidororigem\_tipo\_backup (SQL - SYSTEM\_STATE - File\_Shares)

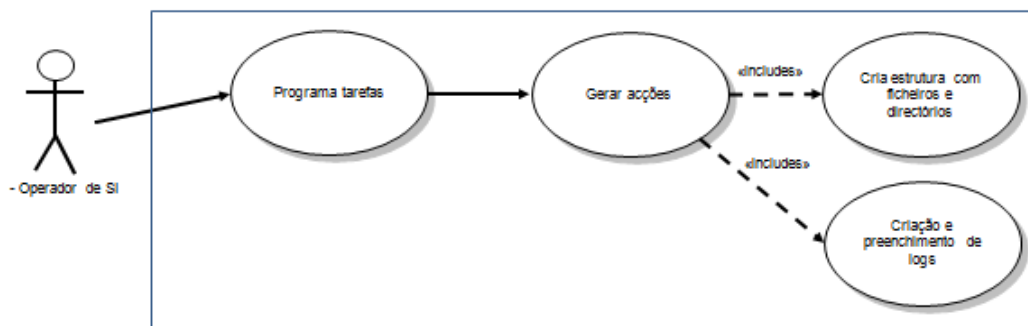
Exemplo:



**Figura 14– Estrutura de directórios para o armazenamento de logs**

### **3.3.4 Programador de tarefas**

Para a execução das rotinas de backup, foi considerada uma ferramenta existente no sistema operativo Windows Server. Trata-se de uma funcionalidade nativa, denominada de “Task Scheduller”, semelhante ao cron para linux e que permite a programação diferida dos processos.

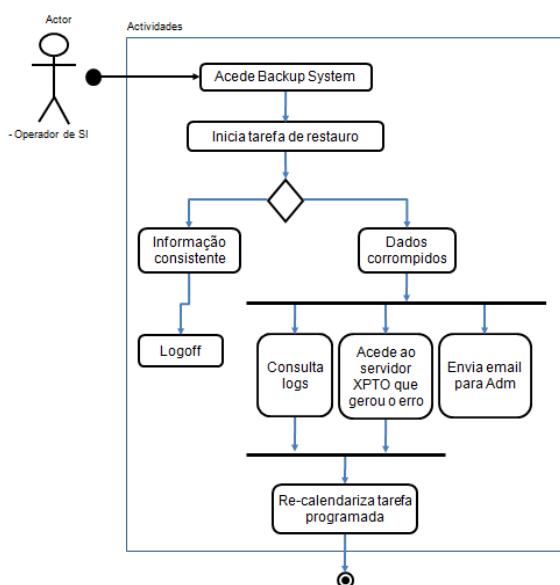


**Figura 15– Caso de Uso para o Actor “Operador de SI”**

A imagem representa um caso de uso para o Operador de SI, no momento em que este programa uma tarefa agendada. Todo este processo pode ser verificado no “Anexo I, 3.2.1.3”.

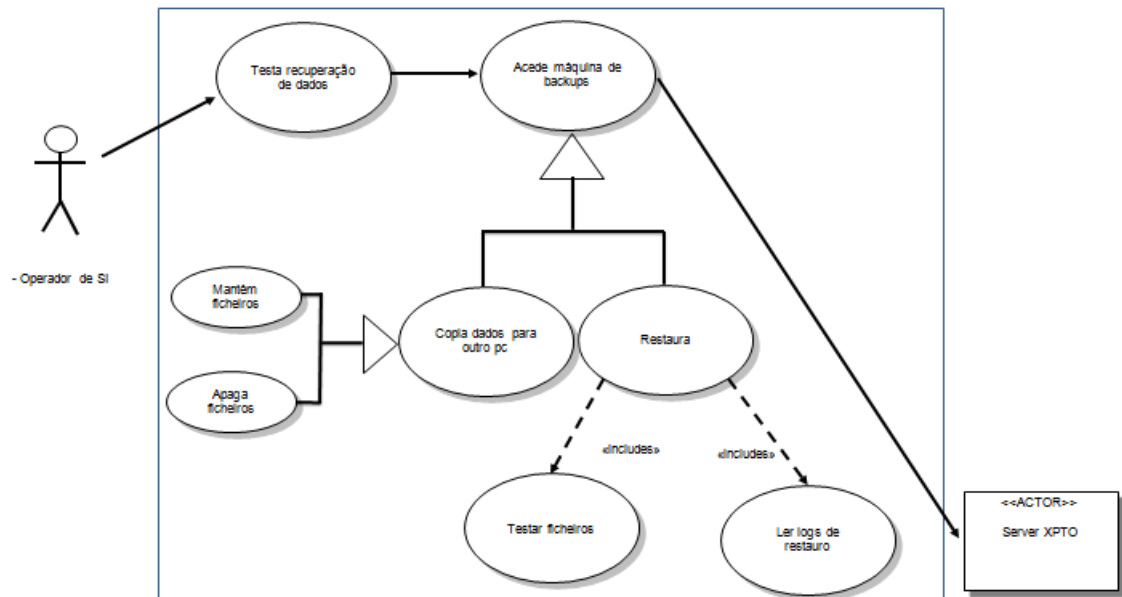
### 3.3.5 Tarefa de Restauro

Para que um sistema integrado de backups seja eficiente, não basta a salvaguarda da informação, sendo necessário garantir que esses dados estão íntegros e possíveis de ser restaurados a qualquer momento. Tendo esse aspecto como prioritário, foi estabelecido como metodologia a execução periódica (a periodicidade dependerá do tipo de negócio) de restauros que validassem de forma inequívoca todo o processo de backups implementados.



**Figura 16– Diagrama de Atividades Operador de SI IV**

Estabeleceu-se que o restauro dos backups estará a cargo do Operador de SI e que este, depois de terminar uma operação desta natureza, obrigatoriamente elaborará um relatório da ocorrência ao Administrador de SI e que o envia por email.

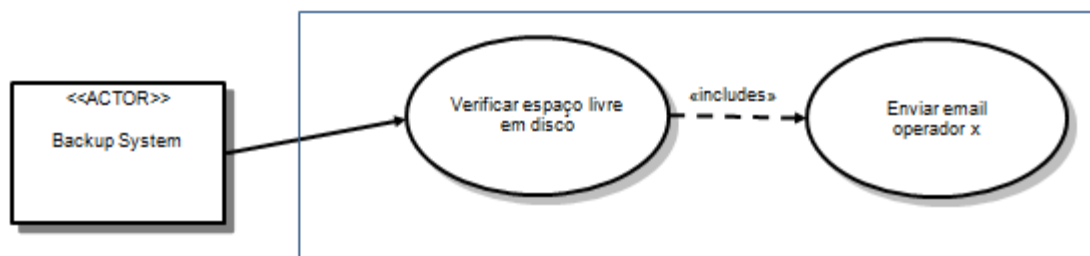


**Figura 17– Caso de Uso para o Actor “Operador de SI”**

Os detalhes desta acção estão contidos no “Anexo I, 3.2.1.4”.

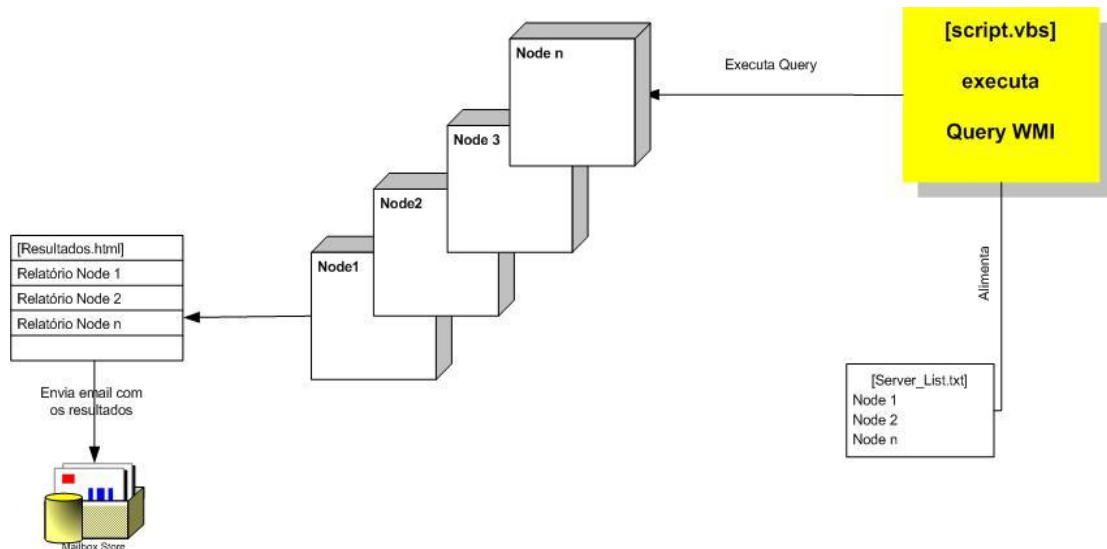
### 3.3.6 Identificação de espaço em disco por servidor

A tarefa de gestão da informação deve ser acompanhada por mecanismos de alarmística que auxiliem os sistemas a identificar potenciais problemas antes de eles se tornarem uma realidade.



**Figura 18– Caso de Uso para o Actor “Backup System”**

A solução pretende implementar uma ferramenta desenvolvida em VBA, e que executa uma query WMI [1] a todos os servidores, que devolvem os respectivos espaços em disco.



**Figura 19 - Representação do esquema a implementar para identificação dos espaços em disco**

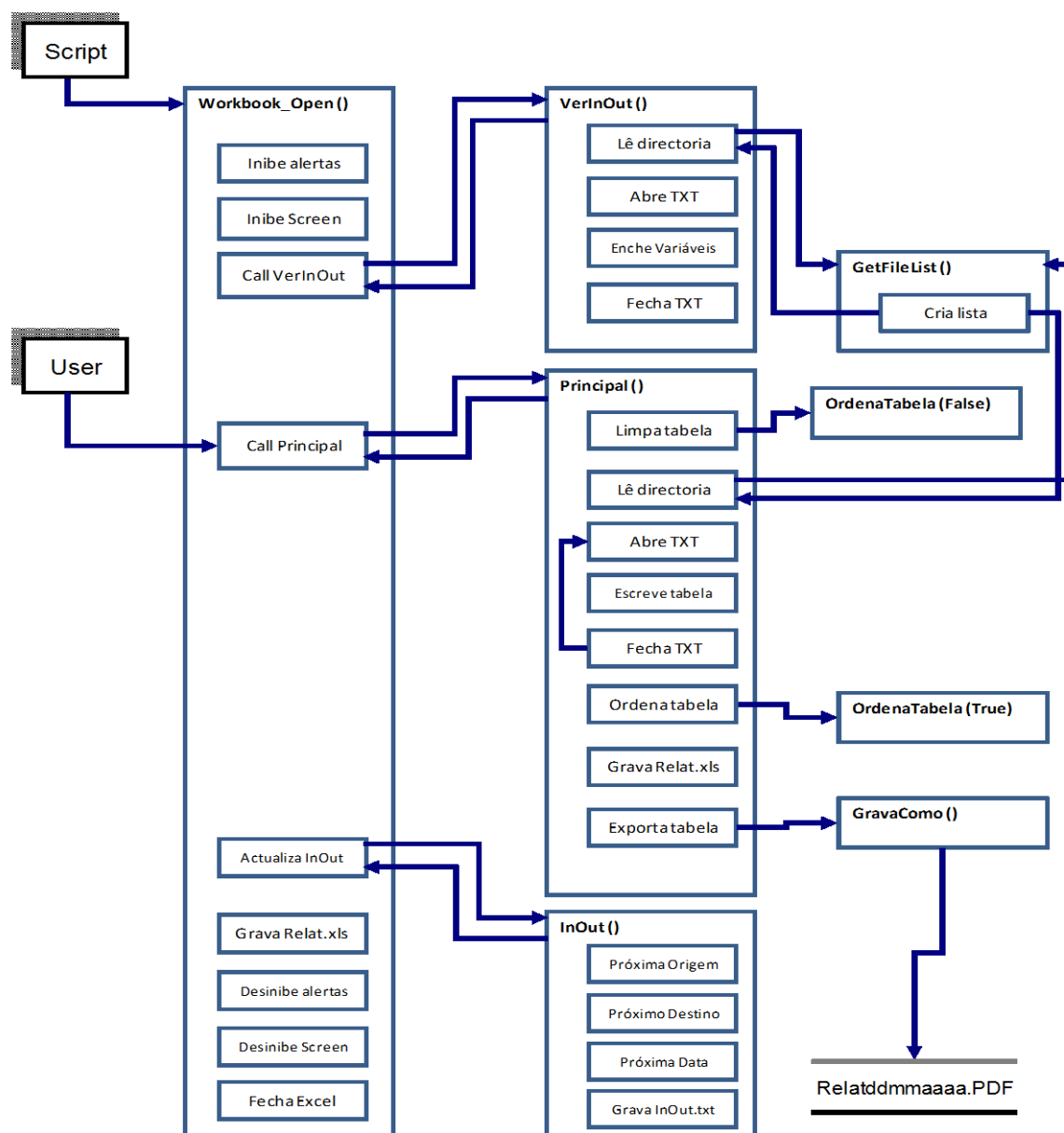
Esta acção irá decorrer mediante uma tarefa agendada e que será programada pelo Operador de SI. A implementação deste script está contida no “Anexo III - Source Code – II”.

### 3.3.7 Transformação dos dados

A transformação dos dados em informação que se pretende dar ao Operador de SI e Administrador de SI, é processada por um mecanismo automático, denominado de “Macro”. Este processo, desenvolvido em linguagem VBA, tem como suporte o software Excel.

A execução deste processo acontecerá após o processo de backup, “Anexo I – 3.2.1.8)” e, despoletado pelo script, fará a abertura do ficheiro (Macro) que por sua vez abrirá o Excel.

Nessa altura, a macro abrirá os ficheiros resultantes dos processos de cópia, executando um processo interno de transformação dos dados, conforme representado na imagem seguinte.

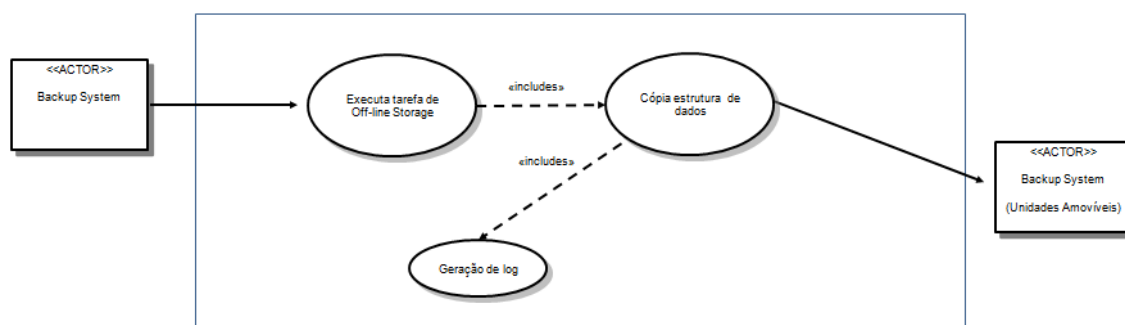


**Figura 20 – Macro em Visual Basic - Fluxo de execução**

Para análise mais detalhada sobre todo o processo, consultar “Anexo V”.

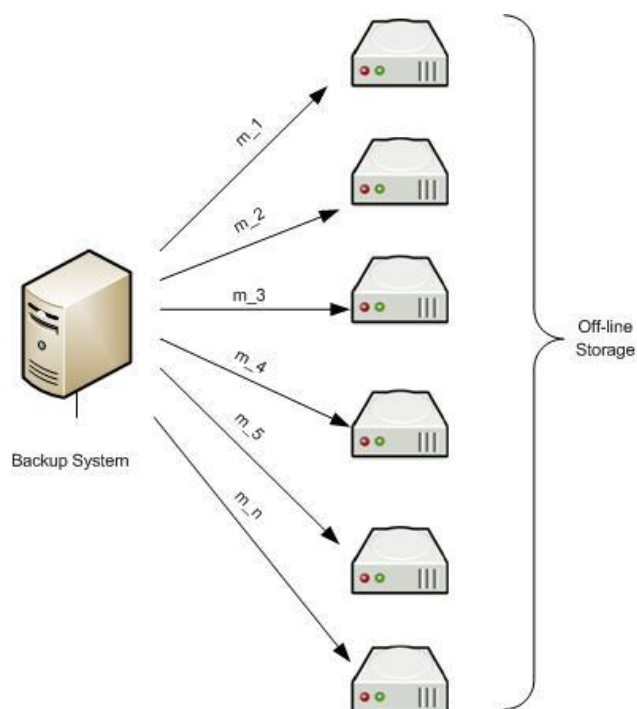
### 3.3.8 Off-line Storage Backup

Está prevista a existência de uma tarefa agendada, que deverá correr mensalmente, e que acontecerá no Backup System. Esta tarefa difere de todas as outras, no sentido em que faz executar uma cópia de dados ao Backup System (através do Robocopy), mais objectivamente à unidade onde está toda a estrutura e dados das cópias de segurança.



**Figura 21– Caso de Uso para o Actor “Backup System”**

Estes dados são copiados para “n” discos amovíveis [30], de grande capacidade e, quando terminada a tarefa, este deverá ser entregue ao Administrador de SI, que colocará o disco num local fisicamente distante das instalações da CSJD.



**Figura 22– Estratégia para Off-line Storage**

## 4. Resultados

A abordagem prática adoptada para a realização da tarefa nº44 – “*Definição e implementação de processos de backup*”, permitiu-nos executar um processo de ETL, simulando em ambientes virtualizados uma realidade adaptável à Clínica de São João de Deus.

Essa virtualização produziu resultados (embora sendo estes académicos) que demonstram a aplicabilidade de uma “implementação de processos de backup” para a CSJD, ou outra organização que careça de um referencial [3] para a salvaguarda da informação.

Para este projecto, foram utilizados diversos utilitários (3.2 - Ferramentas) que auxiliaram à extracção da informação dos diversos Servers XPTO “Anexo I – 2.2.3 e 3.2.2.1 –Backup System e Server System”, à transformação da informação gerada pelos processos de backup (logs) e por fim, ao carregamento dessa informação devidamente tratada (Anexo V) numa Plataforma de Monitorização, permitindo a todos os actores deste projecto, a identificação formal dos processos diários de fluxos de dados entre cada um dos sistemas “ilhas” e um repositório central de dados denominado de Backup System.

O processo de ETL construído permitiu obter os seguintes resultados:

- **Extracção:** criação de scripts de backup locais, considerando todas as bases de dados [26, 29, 31] existentes, bem como os “estados de sistema”. Esta lógica permite a construção de um referencial (3.3 – Metodologias e Casos de Uso) que não só identifica dados críticos de negócio, como permite o backup às bases de dados locais e aos “estados de sistema”. Depois de terminada a extracção destes dados, segue-se uma outra execução (remota – Backup System) que salvaguarda esta informação num repositório de dados estruturado (3.3.3 - Normalização).
- **Transformação:** Todos os processos que geram fluxos de dados, são acompanhados por logs que registam todas as acções, possibilitando no final de cada processo, a transformação desses dados em informação útil para quem interage com a implementação. Foi construída uma Macro para o tratamento desses dados, filtrando apenas a informação que permite identificar de forma resumida, o sucesso ou insucesso de todas as acções anteriores, criando uma semântica específica, adaptada à CSJD (esta Macro pode ser adaptada a todo o tipo de input textual sob a forma de log).
- **Carregamento:** Toda a informação tratada é guardada de forma automatizada (Anexo V) na Plataforma de Monitorização, residente no Backup System. Os dados estão disponíveis a todos os que se autenticarem nessa Plataforma e permite a identificação, rastreio e consequente acção, em função dos dados apurados durante os processos de backup.

Todas as acções referidas nos processos de ETL, quer seja na fase de extracção, transformação ou carregamento, geram emails com os respectivos resultados, devidamente identificados, conforme apresentado na Figura 24. Esses registos ficam depositados nas respectivas caixas de correio, disponíveis para consulta em caso da ocorrência de erros [22] ou falhas [22] nos processos de backup. Estes logs contêm todos os pormenores que a macro não considera, aprofundando na identificação do erro quer este seja no servidor de origem, dados dessa origem ou falha na cópia para o Backup System.

#### 4.1 O Sistema de Extracção, Transformação e Carregamento (ETL)

A representação da imagem seguinte diz respeito ao diagrama de instalação que foi considerado para a obtenção dos resultados. Estão caracterizados os nós “tipo” de uma implementação de rede, onde existe a dimensão “controlador de domínio”, “servidor de Bases de Dados” e um sistema de “Backup” com um front-end para apresentação dos resultados. Estes componentes estão referenciados no anexo de requisitos e a sua descrição mais pormenorizada pode ser consultada nas secções 2.2.3, 2.2.4, 2.4, 3.2.2.1, 3.2.2.2.

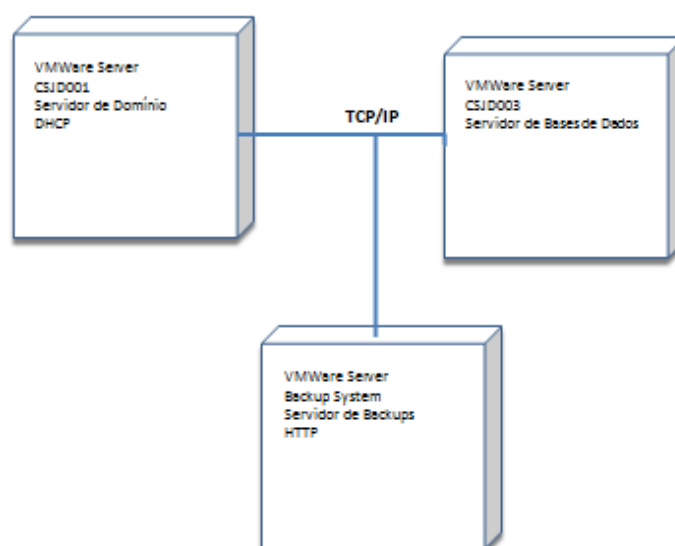


Figura 23– Diagrama de Instalação, representando cada Máquina Virtual



O resultado da nossa implementação, visto de forma global, é representado pelo seguinte diagrama:

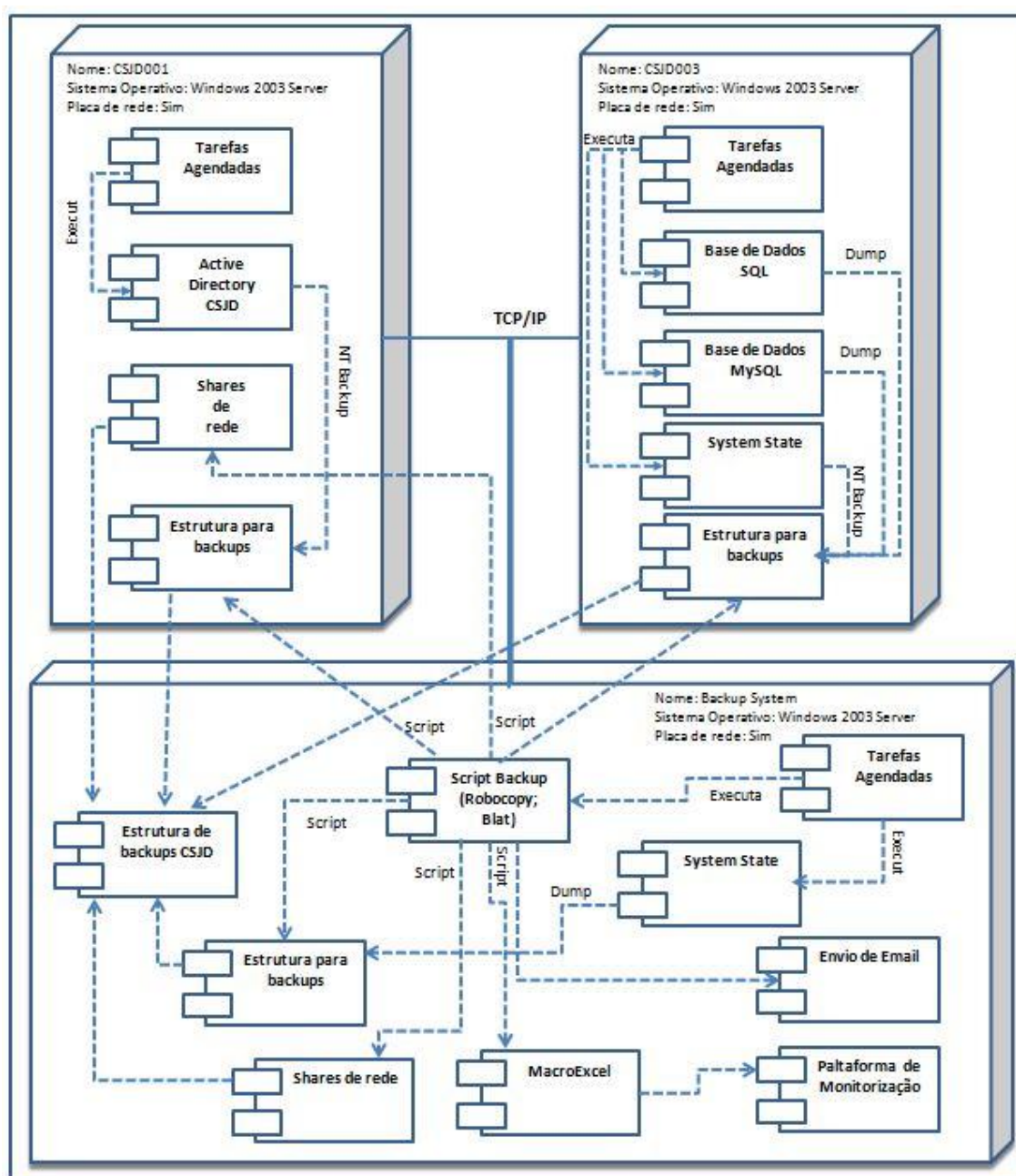
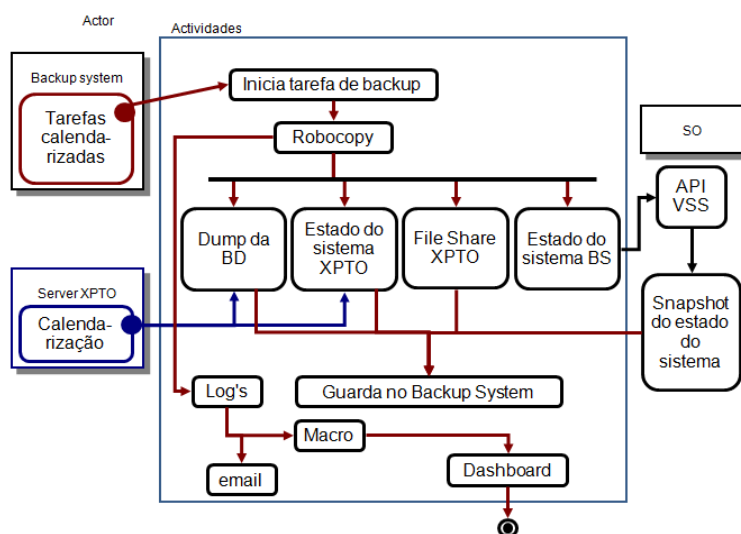


Figura 24– Representação global dos resultados obtidos [14]

Estes são os fluxos de dados produzidos entre as origens de dados e o sistema de backup. A implementação de máquinas virtuais, em articulação como os pressupostos enunciados no Anexo I, permitiu a verificação da totalidade das acções inscritas na proposta deste trabalho. Cada um dos processos foi testado, todos os caminhos foram rastreados, desde o pressuposto até à implementação, permitindo a verificação dos requisitos numa abordagem bottom-up[35].

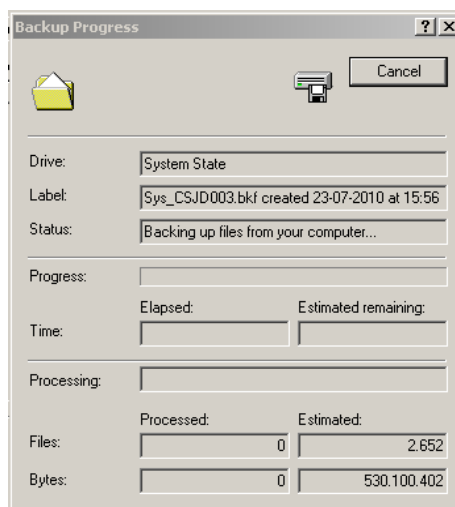
O diagrama é composto por equipamentos virtualizados e construídos para a representação de uma realidade em produção. Nestes estão contidos os componentes necessários para a demonstração de um sistema de backups eficiente e capaz de ser implementado num cenário real.

O processo de ETL inicia-se em cada uma das máquinas consideradas e, mediante a execução de tarefas agendadas (Anexo I - 3.2.1.3), são iniciadas as correspondentes acções:



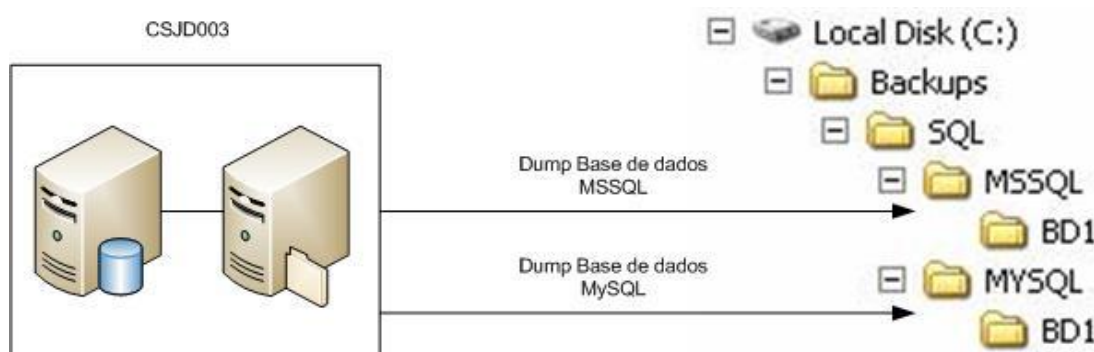
**Figura 25– Diagrama de Actividades Backup System Server XPTO**

- Invocam o NTBackup (3.2.2 – NTBackup) - processo ocorre em todos os servidores - para a salvaguarda do System State na estrutura para backups (Anexo I - 3.2.1.6).



**Figura 26– Execução de um processo de backup a um “Estado de Sistema”**

- Dump das bases de dados (3.2.5 – MySQL e MSSQL) em todos os servidores onde elas estejam presentes. Estes são depositados na estrutura para backups (Anexo I, 3.2.1.5).



**Figura 27– As Bases de Dados são depositadas nos respectivos directórios**

Estas tarefas são programadas para se executar fora da janela diária (Anexo I - 3.2.1.8) e acontecem antes das tarefas de backup (**Extracção**) a ocorrer no Backup System (Anexo I - 3.2.2.2).

Terminadas as execuções anteriores, inicia-se a recolha dos dados pelo Backup System. Neste servidor é despoletada uma tarefa agendada que processa toda a programação incluída nos scripts de backup (Anexo III). Esta acção de backup começa (Anexo I, 3.2.1.8) e é invocado o utilitário robocopy (3.2.1 - Robocopy) que recolhe os dados existentes nas estruturas remotas de dados (de todos os servidores) e os coloca numa estrutura de backups previamente definida, conforme explicado na implementação dos processos de normalização estabelecidos (3.3.3 – Normalização).

Em simultâneo, todos os fluxos de dados são registados em logs para que seja possível rastrear o sucesso ou insucesso de todas as operações - esses logs são depois enviados por email (para o Operador e Administrador), através do Blat (3.2.3 – Blat).

[TFC] Log 2FEIRA - BACKUP_SYSTEM	dom 18-07-2010 01:48	7 KB
[TFC] Log ESPAJO DISCO POR SERVIDOR	dom 18-07-2010 01:47	2 KB
[TFC] Log 2FEIRA - BACKUP_SYSTEM	dom 18-07-2010 01:24	7 KB
[TFC] Log 2FEIRA - BACKUP_SYSTEM	sáb 17-07-2010 23:39	7 KB
[TFC] Log 2FEIRA - BACKUP_SYSTEM	sáb 17-07-2010 23:29	7 KB
[TFC] Log 2FEIRA - BACKUP_SYSTEM	sáb 17-07-2010 23:29	8 KB
[TFC] Log System State - CSJD001	sáb 17-07-2010 05:59	2 KB
[TFC] Log System State - CSJD001	sáb 17-07-2010 05:42	2 KB

**Figura 28– Caixa de correio do Operador, contendo os logs diários de todas as acções**

Depois de todas as execuções terem terminado, é chamada de forma automática uma Macro em Excel (descrição e detalhes técnicos em Anexo V) que analisa os logs produzidos e transforma esses dados em informação visualmente imediata para os intervenientes dos SI.

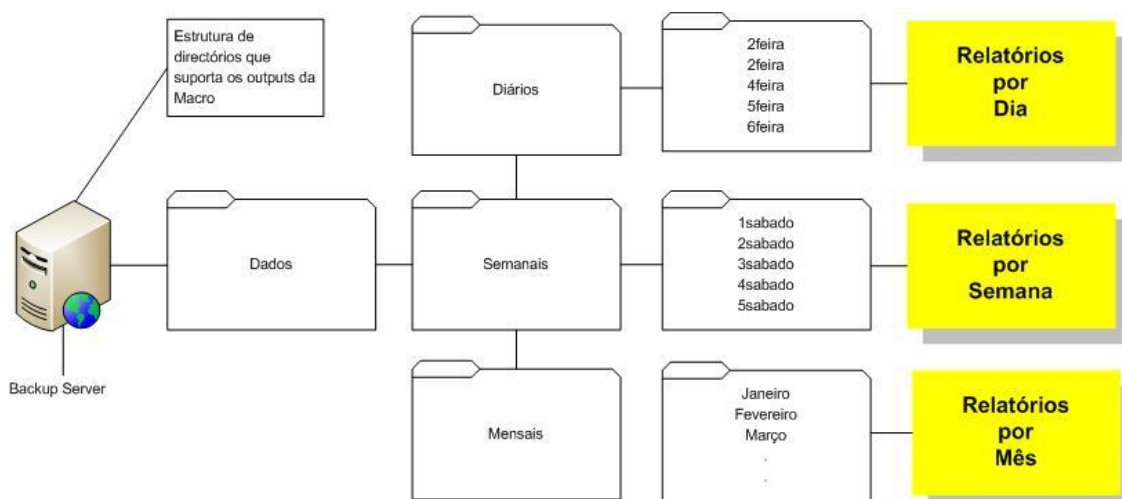


Figura 29– Os logs são colocados conforme o esquema apresentado

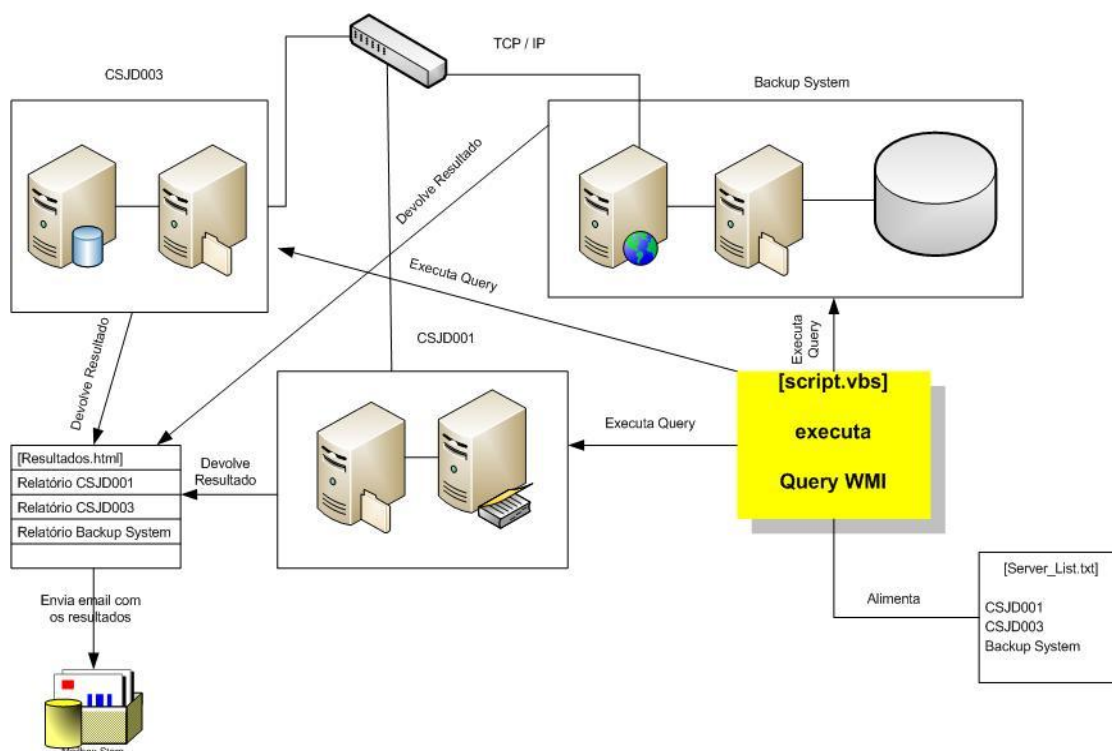
Depois de transformados, os ficheiros resultantes são carregados para a Plataforma de Monitorização (3.3.1 – Plataforma de Monitorização e Anexo III – Source Code - I), ficando disponíveis para consulta.

Data	Source	Dics				Files				Bytes				TEMPO Total	Nome Log
		Total	Copied	Skipped	FAILED	Total	Copied	Skipped	FAILED	Total	Copied	Skipped	FAILED		
1 19-07-2010 15:30	\\Backupsystem\CS\Back	1	0	1	0	2	0	2	0	523,49 m	0,00-	523,49 m	0-	0:00:00	BackupSystem_System_State.log
2 19-07-2010 15:30	\\Backupsystem\CS\inet	63	29	34	0	138	70	68	0	471,70 k	305,50 k	166,20 k	0-	0:00:01	BackupSystem_Wwwroot.log
3 19-07-2010 15:30	\\Backupsystem\CS\share	6	0	6	0	0	0	0	0	0,00-	0,00-	0,00-	0-	0:00:00	BackupSystem_Share1.log
4 19-07-2010 15:30	\\CS\0001\CS\Backups\S	-1	-1	34	-1	-1	-1	-1	-1	-1,00-	-1,00-	-1,00-	-1-		CS\0001_System_State.log
5 19-07-2010 15:30	\\CS\0001\CS\share1\	-1	-1	34	-1	-1	-1	-1	-1	-1,00-	-1,00-	-1,00-	-1-		CS\0001_Share1.log
6 19-07-2010 15:31	\\CS\0003\CS\Backups\S	-1	-1	34	-1	-1	-1	-1	-1	-1,00-	-1,00-	-1,00-	-1-		CS\0003_Sql.log
7 19-07-2010 15:30	\\CS\0003\CS\Backups\S	-1	-1	34	-1	-1	-1	-1	-1	-1,00-	-1,00-	-1,00-	-1-		CS\0003_System_State.log

Figura 30–PDF, contendo informação rica, depositado na Plataforma de Monitorização

Todo este processo permite a identificação rápida e eficiente de eventuais falhas ocorridas nos diversos Servidores, resultando daí uma resposta adequada às situações que se demonstrem críticas para o normal funcionamento da organização.

A alarmística foi um dos aspectos considerados fundamentais. Conforme identificado no capítulo 3 (3.3.6 – Identificação de espaço em disco por servidor) e no “Anexo I, 3.2.1.7, foi implementada uma rotina que permitiu a identificação do espaço em disco em todos os sistemas críticos da infra-estrutura.



**Figura 31– Representação do processo de recolha de informação sobre os espaços em disco**

A Figura 31 representa o resultado da implementação.

A query WMI é feita a todos os servidores constantes num ficheiro de configuração [Server\_List.txt]. O resultado desta é colocado num ficheiro em html [Resultados.html] e enviado para as caixas de correio do Administrador e Operador de SI.

O resultado atingido está representado na figura seguinte:

Relatório Diário - Espaço em disco por Servidor				
CSJD001				
Drive / Mount	Total Capacity (in GB)	Used Capacity (in GB)	Free Space (in GB)	Freespace %
C:	49,99	6,21	43,78	88%
CSJD003				
Drive / Mount	Total Capacity (in GB)	Used Capacity (in GB)	Free Space (in GB)	Freespace %
C:	49,99	7,27	42,72	85%
Backup_System				
Drive / Mount	Total Capacity (in GB)	Used Capacity (in GB)	Free Space (in GB)	Freespace %
C:	49,99	7,27	42,72	85%

**Figura 32– Ficheiro gerado pela acção anterior**

## 5. Conclusões e Trabalho Futuro

### 5.1 Conclusões

Os dados suportados pelas tecnologias de informação, representam um dos maiores bens que as Empresas possuem. Essa realidade implica o pleno conhecimento da informação gerada pelos processos diários e, conseqüentemente, a garantia de que esses activos estarão sempre disponíveis. Pretende-se atingir o triplo compromisso: *Fiabilidade, Integridade e Disponibilidade*.

Num momento de intensa competitividade, onde as palavras de ordem são: *eficiência, eficácia, empreendedorismo e racionalização de custos*, há que criar as melhores condições para que as empresas possam implementar estratégias de gestão, suportadas nos seus Sistemas de Informação.

O projecto que desenvolvemos permitiu-nos compreender a importância que o planeamento da infra-estrutura representa para o funcionamento das organizações. É fundamental a existência de planos estratégicos para os SI, tal como é importante dimensionar os projectos, assumindo compromissos de estabilidade, sem nunca esquecer a necessidade de salvaguardar todos os activos que compõem a plataforma operacional.

Os SI devem ser exaustivamente conhecidos, toda a infra-estrutura deve estar mapeada, referenciada, documentada e adequada ao papel que desempenha na organização. Todos os processos devem ser testados e sujeitos a mecanismos de controlo interno que garantam a correcta utilização da informação, permitindo às organizações atingir níveis de maturidade superiores (CMMI) [24], alavancando para melhores resultados.

A proposta de TFC deste relatório, pretendeu demonstrar, de forma inovadora, a possibilidade de implementação de uma “**Framework de Backup**”, adaptada a uma realidade em particular. Para tal, fez-se valer da criatividade, da imaginação e da criação de (programação) código orientado(a) aos objectivos propostos.

Havia a necessidade de criar níveis de eficiência na gestão dos SI de um determinado “cliente” e de atingir o conhecimento dos fluxos de dados, fazendo destacar o factor económico. Para tal, foram utilizados programas gratuitos que, em conjunto com todos os atributos anteriores, permitiram representar um sistema *integrado, fiável e consistente*.

A salvaguarda da informação, através de mecanismos de backup é e será sempre, uma das mais silenciosas tarefas que uma empresa tem de garantir para que o negócio esteja suportado de forma adequada e segura.

Os pressupostos de ETL foram demonstrados no Capítulo 4, e aí todos os caminhos foram validados, passo a passo, suportados pelo *Anexo I - Relatório de Especificação de Requisitos*, conferindo à solução os níveis de eficiência que esperávamos atingir.

Conseguimos deste modo, atingir o objectivo a que nos propusemos, que era o de criar um Sistema de Backup que, embora académico e virtualizado, demonstrasse a possibilidade de implementação de uma “*Framework*” orientada a um ambiente concreto, podendo até considerar um âmbito mais generalizado, como sejam outros sistemas heterogéneos, independentemente da dimensão a considerar.

O processo de aprendizagem daqui em diante irá continuar, não sendo este momento, o último passo na solução apresentada. Gostaríamos de poder ver a proposta de TFC implementada num ambiente em produção, como forma de reconhecimento do valor teórico que representa, tendo em conta as valências que integra e o baixo custo de concretização. Para finalizar, as palavras-chave que destacamos: *eficiência, salvaguarda, escalabilidade, alarmística, controlo e racionalização de custos*.

## 5.2 Trabalho futuro

Há aspectos que gostaríamos de implementar no futuro, destacando como principais os seguintes:

- 1 - Criação de Macro, independente de qualquer software proprietário, que integre um interface GUI para suporte de configurações, adaptadas a ambientes diversos.
- 2 – Melhorar a interface da Plataforma de Monitorização, integrando módulo de configuração de utilizadores, permitindo a parametrização de acessos na própria ferramenta. A este módulo teria acesso único, o Administrador de SI, apesar de este não ter a função de Operacional dos SI. Esta medida funcionaria como mecanismo de controlo nas atribuições de permissão de acessos à Plataforma.
- 3 – Integração nos mecanismos de carregamento de informação tratada pela Macro, de um conjunto de logs de sistema que neste momento não são considerados. Destes, fazem parte os alertas gerados pelos “event viewer”, na sua dimensão de erros e alertas, não considerando os eventos informativos.
- 4 – Criação de escala efectiva na implementação da solução de backups em ambientes com outra dimensão, independentemente da infra-estrutura existente, contemplando mecanismos de adaptabilidade e capacidade de resistência às constantes alterações dos paradigmas de computação. Este último aspecto integra conceitos de “cloud computing”, infra-estruturas como serviço (infrastructure as a service), software como serviços (software as a service) e plataformas como serviço (platform as a service).

## BIBLIOGRAFIA



1. **MSDN.** Windows Management Instrumentation. *MSDN Library*. [Online] Microsoft Corporation. [http://msdn.microsoft.com/en-us/library/aa394582\(VS.85\).aspx](http://msdn.microsoft.com/en-us/library/aa394582(VS.85).aspx).
2. —. Visual Basic Developer Center. *MSDN Library*. [Online] Microsoft Corporation. <http://msdn2.microsoft.com/pt-br/vbasic/default.aspx>.
3. **Blokdijk, Maarten.** Easy-content Products. *Easy-content File Manager (ECFM)*. [Online] <http://www.xs4all.nl/~blokdyk/>.
4. **Technet.** IIS 6.0 Operations Guide. *Microsoft Technet*. [Online] Microsoft Corporation. [http://technet.microsoft.com/en-us/library/cc785089\(WS.10\).aspx](http://technet.microsoft.com/en-us/library/cc785089(WS.10).aspx).
5. **Microsoft.** Windows Server 2003 Resource Kit Tools. [Online] Microsoft Corporation. <http://www.microsoft.com/downloads/details.aspx?familyid=9d467a69-57ff-4ae7-96ee-b18c4790cffd&displaylang=en>.
6. **Technet.** Dados de estado do sistema. *Windows Server TechCenter*. [Online] Microsoft Corporation. [http://technet.microsoft.com/pt-pt/library/cc785306\(WS.10\).aspx](http://technet.microsoft.com/pt-pt/library/cc785306(WS.10).aspx).
7. **Shaw, Jay.** Preserve Your Data With The Backup Tool You Already Have. *TechNet Magazine*. [Online] Microsoft Corporation. [http://technet.microsoft.com/pt-pt/magazine/2005.05.ntbackup\(en-us\).aspx](http://technet.microsoft.com/pt-pt/magazine/2005.05.ntbackup(en-us).aspx).
8. **Pedro Mendes, Mark Neal, Gilles Volland.** Easily mail any file from the command line. *BLAT for Windows*. [Online] SourceForge.net. <http://www.blat.net>.
9. **VMware, Inc.** Begin the Journey to a Private Cloud with Datacenter Virtualization. *VMware vSphere 4: Private Cloud Computing, Server and Data Center Virtualization*. [Online] VMware, Inc. <http://www.vmware.com/products/vsphere/>.
10. **VMware, Inc.** VMware Player - The Easiest Way to Run a Virtual Machine. *VMware Player: Run Windows 7, Chrome OS - Free Download for a Virtual PC*. [Online] VMware, Inc. <http://www.vmware.com/products/player/>.
11. **Oracle Corporation.** The world's most popular open source database. *MySQL :: The world's most popular open source database*. [Online] Oracle Corporation. <http://www.mysql.com/>.
12. **Microsoft.** Microsoft SQL Server. *Descrição Geral do SQL Server 2005*. [Online] Microsoft Corporation. <http://www.microsoft.com/portugal/sql/prodinfo/overview/default.mspx>.
13. **Pavlov, Igor.** Welcome to the 7-Zip Home! *7-Zip*. [Online] <http://www.7-zip.org/>.
14. **Mauro Nunes, Henrique O' Neill.** Fundamental de UML. s.l. : FCA - EDITORA DE INFORMÁTICA.
15. **Dean Leffingwell, Don Widrig.** *Managing Software Requirements*. s.l. : Addison Wesley, 1999. 0-201-61593-2.
16. **Postel, Jonathan B.** SIMPLE MAIL TRANSFER PROTOCOL. *RFC 821 - Simple Mail Transfer Protocol*. [Online] <http://tools.ietf.org/html/rfc821>.

17. **IEEE Standard Association.** IEEE Std 830-1998 IEEE Recommended Practice for Software Requirements Specifications. *IEEE Standards Description: 830-1998*. [Online] IEEE. [http://standards.ieee.org/reading/ieee/std\\_public/description/se/830-1998\\_desc.html](http://standards.ieee.org/reading/ieee/std_public/description/se/830-1998_desc.html).
18. **Sousa, Manuel de.** *Dicionário de termos informáticos*. s.l. : Milénio – Inglês / Português.
19. **Wikipedia.** Windows Server. [Online] [http://en.wikipedia.org/wiki/Windows\\_Server](http://en.wikipedia.org/wiki/Windows_Server).
20. —. Web service. [Online] [http://pt.wikipedia.org/wiki/Web\\_service](http://pt.wikipedia.org/wiki/Web_service).
21. —. Active Directory. [Online] [http://pt.wikipedia.org/wiki/Active\\_Directory](http://pt.wikipedia.org/wiki/Active_Directory).
22. **Pressman, Roger S.** *Software engineering: a practitioner's approach / Roger S. Pressman*. —5th ed. s.l. : Thomas Casson, 2001. 0073655783.
23. **Correia, José.** *Análise de Sistemas Informáticos*. [Online] <http://paginas.ispgaya.pt/~jcorreia/asi/>.
24. **Ubuntu.** CronHowto. *Ubuntu - How To*. [Online] Creative Commons. <https://help.ubuntu.com/community/CronHowto>.
25. **Linux Online Inc.** The History of Linux. [Online] Linux Online Inc. [http://www.linux.org/info/linux\\_timeline.html](http://www.linux.org/info/linux_timeline.html).
26. **Ralph Kimball, Joe Caserta.** *The Data Warehouse ETL Toolkit*. s.l. : Wiley Publishing, Inc., 2004. 0-764-56757-8.
27. **Squidoo, LLC.** Data warehousing, ETL and data integration applications and tools in business intelligence. *Data Warehousing and ETL Tools*. [Online] Squidoo, LLC. <http://www.squidoo.com/etltools>.
28. **Gonçalves, A. Moreira.** *Apontamentos da cadeira de Base de Dados*. Lisboa : Apontamentos do Professor A. Moreira Gonçalves, 2008. Vol. 1 a 7.
29. **Ribeiro, Rui.** *Apontamentos da cadeira de Sistemas de Suporte à Decisão*. Lisboa : Apontamentos do Professor Rui Ribeiro, 2009. Vol. 0 a 7.
30. **W. Preston.** *Backup & Recovery*. s.l. : O'Reilly, December 01, 2006. 0-596-10246-1, cap. 1.1.
31. **Abraham Silberschatz, Peter Baer Galvin, Greg Gagne.** *Operating System Concepts - Seventh Edition*. s.l. : Wiley, 2005. 0-471-69466-5, cap. 6.2 pp.193.
32. **Preston, W.** *Backup & Recovery*. s.l. : O'Reilly, December 01, 2006. 0-596-10246-1.
33. **ISACA.** Isaca.org. *Cloud Computing: Business Benefits with security, governance and assurance perspectives*. [Online] 2009. <https://www.isaca.org/Pages/default.aspx>.
34. **Pressman, Roger S.** *Software engineering: a practitioner's approach / Roger S. Pressman*. —5th ed. s.l. : Thomas Casson, 2001. 0073655783, pp 24.
35. —. *Software engineering: a practitioner's approach / Roger S. Pressman*. —5th ed. s.l. : Thomas Casson, 2001. 0073655783, pp 490.