



Database as a Service (DaaS) or RDBMS?

JOÃO ALBERTO RODRIGUES DE SOUSA

Outubro de 2016

Database as a Service (DaaS) or RDBMS?

João Alberto Rodrigues de Sousa

**Dissertação para obtenção do Grau de Mestre em
Engenharia Informática, Área de Especialização em
Sistemas Computacionais**

Orientador: Ana Almeida - ISEP

Co-orientador: Jorge Bernardino - ISEC

Porto, Outubro de 2016

Dedicatória

Dedico este trabalho à minha mãe,
Adília Augusta

Resumo

A Cloud é um modelo de disponibilização e utilização de TIC (Tecnologias de Informação e Comunicação), que atualmente representa um grande desafio para a gestão das organizações. Uma das principais características da Cloud é a capacidade da disponibilização de recursos, estes podem ser retirados ou adicionados de uma forma elástica, outra das características é a capacidade de medir o serviço, tanto a nível de performance do mesmo, como ao nível da faturação. A Cloud introduz uma camada de abstração que elimina a complexidade tecnológica inerente aos serviços por ela prestados. Foi verificado durante a execução deste trabalho que os DaaS analisados cumprem as características essenciais da Cloud.

Este trabalho pretende fazer uma abordagem às DaaS (Database as a Service), que é um dos muitos serviços prestados na Cloud, tendo como objetivo a realização de uma análise de custos e de performance de bases de dados, comparando o uso de um SGBD (Sistema de Gestão de Base de Dados) tradicional com a utilização de um DaaS.

Foi realizado um teste de migração de uma base de dados de teste para o Azure SQL. Este tinha como principais objetivos a análise de performance obtida recorrendo ao uso deste tipo de serviço, bem como a análise de custos, para validar a hipótese de uma eventual migração. A base de dados encontrava-se num computador que já não dispunha de mais recursos. A migração decorreu com sucesso, pois foi possível aumentar a performance, bem como ficou saliente que a compra de novo equipamento fica, para este caso, mais caro que manter uma solução no Azure, foi tido em consideração para efeitos de análise os custos para um prazo de três anos.

Durante a realização deste trabalho, e mais concretamente, durante a realização do exemplo de migração efetuado, foi verificado que o utilizador não se apercebe da infraestrutura necessária para suportar este serviço, que pode modificar os níveis de serviço conforme a necessidade e sem quebra de serviço e ainda que pode verificar a performance do sistema, bem como pode a qualquer momento verificar o custo atual da sua fatura.

Este trabalho veio comprovar o aumento de performance quando se utiliza um serviço DaaS, bem como evidenciou que a compra de novo equipamento fica mais dispendiosa que manter uma solução no Azure SQL.

Palavras-chave: Cloud, Database, Database as a Service (DaaS), Relational DataBase Management System (RDBMS)

Abstract

The Cloud is a deployment and use model and use ICT (Information and Communication Technologies), which currently represents a major challenge for the management of organizations. One of the main Cloud characteristics is the ability of provision resources, these may be removed or added in an elastic manner, another characteristic is the ability to measure the service, measuring system performance and controlling costs. The Cloud introduces a layer of abstraction that eliminates the technological complexity of the services it provides. It was found during the execution of this work that DaaS analyzed meet the essential features of the Cloud.

This paper aims to make an approach to the subject DaaS (Database as a Service), this is one of many services available in the Cloud. The objective is performing a cost and performance analysis of databases, relating to the use of a traditional RDBMS, or the use of a DaaS and therefore choose to use one of them.

A migration test was performed to the SQL Azure, it was performed with a test database. This had as main objectives obtaining the performance analysis using this type of service, as well as a cost analysis, this to validate the hypothesis of a possible migration. The database was at a computer that had not more resources to supply. The migration was successful because it was possible to increase the performance and was projected that the purchase of new equipment was, in this case, more expensive than maintaining a solution on the Azure, it was taken into consideration, for the analysis of the costs a three years period.

During this work, and more specifically, during the realization of the performed migration example, it was found that the user is not aware of the necessary infrastructure to support this service, that you can modify the service levels as needed and during that operation there are not service breaks, that at any occasion the system performance can be verified, and may be at any time checked the current cost.

This work came to prove the increase in performance when using a DaaS service, and showed that the purchase of new equipment is more costly to maintain than a solution in SQL Azure.

Keywords: *Database, Database as a Service (DaaS), Relational DataBase Management System (RDBMS)*

Agradecimentos

Gostaria de deixar aqui expresso o meu agradecimento à Professora Doutora Ana Almeida, bem como ao Professor Doutor Jorge Bernardino, pelo apoio e disponibilidade prestados durante a realização deste trabalho.

Índice

1	Introdução	1
1.1	Problema	1
1.2	Objetivos	2
1.3	Critérios de Sucesso	2
1.4	Organização da Dissertação	2
2	Análise de Valor	5
2.1	Definições	5
2.1.1	Valor	5
2.1.2	Capex, Opex, Total Cost of Ownership	6
2.2	Fatores a considerar na aquisição de DaaS	7
2.2.1	<i>Service Level Agreement (SLA)</i>	7
2.2.2	Comunicações	8
2.2.3	Custos	8
2.3	Produto	8
3	Bases de Dados: Conceitos	11
3.1	Definição de Base de Dados	11
3.2	DataBase System (DBS) e DataBase Management System	11
3.3	Objetivos de um Sistema de Base de Dados	13
3.4	Vantagens de um Sistema de Base de Dados	14
3.5	Definição de DataBase Administrator (DBA)	14
3.6	Modelo Relacional	15
4	Relational DataBase Management System (RDBMS)	17
4.1	Definição	17
4.2	Principais RDBMS	17
4.2.1	Oracle	18
4.2.2	MySQL	21
4.2.3	Microsoft SQL Server	22
4.2.4	PostgreSQL	26
4.2.5	DB2	27
5	Database as a Service (DaaS)	31
5.1	Definição de Cloud Computing	31
5.1.1	Características Essenciais	32
5.1.2	Modelos de Serviço	32
5.1.3	Modelos de Implementação	33
5.1.4	Mitos	34

5.2	Database as a Service (DaaS)	34
5.3	Fornecedores de DaaS	35
5.3.1	Oracle Cloud	36
5.3.2	Azure SQL Database.....	39
5.3.3	Google Cloud SQL	42
5.3.4	Amazon RDS	44
6	Avaliação Experimental	49
6.1	Métodos	49
6.2	Metodologias a Aplicar.....	50
6.2.1	Testes a Aplicar.....	50
6.2.2	Exemplo.....	51
6.2.3	Conclusões.....	74
7	Conclusões e Trabalho Futuro	77

Lista de Figuras

Figura 1 – Modelo CANVAS	10
Figura 2 – Exemplo de um DBS, retirada de (Foster and Godbole, 2014, p. 4)	12
Figura 3 – Base de dados relacional, retirada de (Stephens, Plew, and Jones, 2011, p. 5)	15
Figura 4 – Quadrado Mágico da Gartner para SGBD Operacionais	36
Figura 5 – Modelo E-R pertencente ao <i>schema</i> Application.....	52
Figura 6 – Modelo E-R pertencente ao <i>schema</i> Purchasing	53
Figura 7 – Modelo E-R pertencente ao <i>schema</i> Sales.....	54
Figura 8 – Modelo E-R pertencente ao <i>schema</i> Warehouse	55
Figura 9 – Parametrização usada para testes	56
Figura 10 – Task Manager	57
Figura 11 – Performance Monitor.....	57
Figura 12 – Gráfico do nível de serviço / nível de desempenho	58
Figura 13 – Gráfico do nível de serviço / nível de desempenho da CPU	59
Figura 14 – Gráfico do nível de serviço / nível de desempenho para os lops.....	59
Figura 15 – Gráfico do nível de serviço / nível de desempenho para o transaction log.....	59
Figura 16 – Gráfico do nível de serviço / nível de desempenho para CPU, lops e Transaction Log.....	60
Figura 17 – Query Teste 1	61
Figura 18 – Plano de Execução da Query Teste 1	61
Figura 19 – Screenshot retirado do SSMS.....	62
Figura 20 – Query Teste 2	63
Figura 21 – Plano de Execução da Query Teste 2	63
Figura 22 – Query Teste 3	65
Figura 23 – Plano de Execução da Query Teste 3	65
Figura 24 – Multithreaded Order Insert Main – Nível de Serviço S0	67
Figura 25 – Monitorização de DTU – Nível de Serviço S0	68
Figura 26 – Multithreaded Order Insert Main – Nível de Serviço S1	68
Figura 27 – Monitorização de DTU – Nível de Serviço S1	69
Figura 28 – Multithreaded Order Insert Main – Nível de Serviço S2	69
Figura 29 – Monitorização de DTU – Nível de Serviço S2	70
Figura 30 – Multithreaded Order Insert Main – Nível de Serviço S3	70
Figura 31 – Monitorização de DTU – Nível de Serviço S3	71
Figura 32 – Multithreaded Order Insert Main – Nível de Serviço P1.....	71
Figura 33 – Monitorização de DTU – Nível de Serviço P1	72
Figura 34 – Script em Powershell.....	74
Figura A.1 – Orçamento HP Proliant ML150 Gen9.....	81
Figura A.2 – Orçamento HP Proliant ML30 Gen9.....	82
Figura B.1 – Custo de uma licença de Microsoft Windows 2012 Server Standard.....	83
Figura B.2 – Custo de uma licença de Microsoft SQL Server 2014 Standard.....	83

Figura C.1 – Retirado de (ERSE, 2016, P. 13)	84
Figura D.1 – Opções a modificar para restringir acesso à base de dados	85
Figura D.2 – SSMS Migração Ecrã 1	86
Figura D.3 – SSMS Migração Ecrã 2	87
Figura D.4 – SSMS Migração Ecrã 3	87

Lista de Tabelas

Tabela 1 – Comparação entre as versões do Oracle 12c	20
Tabela 2 – Datas de lançamento e fim de suporte de versões do SQL Server.....	22
Tabela 3 – Comparação entre as versões do SQL Server 2014	24
Tabela 4 – Comparação entre as versões do SQL Server 2016.....	24
Tabela 5 – Fatores de multiplicação para o licenciamento por core	25
Tabela 6 – Limites máximos PostgreSQL, transcrito de (Caldeira, 2015, p. 18).....	26
Tabela 7 – Datas de lançamento das versões do DB2	27
Tabela 8 – Opções de escolha rápida na criação de uma base de dados em Oracle Database Cloud Service.....	37
Tabela 9 – Opções existentes para o Oracle Database Exadata Cloud Service.....	38
Tabela 10 – Características básicas dos níveis de serviço, para uma base de dados	40
Tabela 11– Características básicas dos níveis de serviço do modelo <i>elastic</i>	40
Tabela 12 – Modelos existentes para o Google Cloud SQL Second Generation.....	43
Tabela 13 – Modelos existentes para o Google Cloud SQL First Generation	43
Tabela 14 – Tipos de instâncias passíveis de criação no Amazon RDS	45
Tabela 15 – Tempos de migração para Azure SQL, por nível de serviço	61
Tabela 16 – Tempos de execução da Query Teste 1.....	62
Tabela 17 – Tempos de execução da Query Teste 2.....	63
Tabela 18 – Resultados do T-Test para Servidor e Nível de Serviço S0	64
Tabela 19 – Resultados do T-Test para Servidor e Nível de Serviço S1	64
Tabela 20 – Tempos de execução da Query Teste 3.....	66
Tabela 21 – Resultados do T-Test para Servidor e Nível de Serviço S0	66
Tabela 22 – Resultados do T-Test para Servidor e Nível de Serviço S1	66
Tabela 23 – Resultados do T-Test para Servidor e Nível de Serviço S2	67
Tabela 24 – Quadro Resumo dos Testes Efetuados.....	72
Tabela 25 – Custo Total de Aquisição e Manutenção de hipóteses de servidor	73
Tabela 26 – Incrementos de Performance no Tempo de Execução.....	74
Tabela 27 – Percentagem de DTU's Usada no Azure	75
Tabela 28 – Custo a três anos do Nível S2 do Azure SQL.....	75

Acrónimos e Símbolos

Lista de Acrónimos

ACID	<i>Atomicity, Consistency, Isolation, Durability</i>
CAL	<i>Client Access Licence</i>
DaaS	<i>DataBase as a Service</i>
DBA	<i>DataBase Administrator</i>
DBMS	<i>DataBase Management System</i>
IaaS	<i>Infrastructure as a Service</i>
PaaS	<i>Plataform as a Service</i>
PME	Pequena e Média Empresa
RDBMS	<i>Relational DataBase Management System</i>
SGBD	Sistema de Gestão de Base de Dados
SaaS	<i>Software as a Service</i>
SSMS	<i>SQL Server Management Studio</i>
TCO	<i>Total Cost of Ownership</i>
TIC	Tecnologias de Informação e Comunicação
XML	<i>eXtensible Markup Language</i>

1 Introdução

Este trabalho descreve um dos problemas que existe atualmente na área das TIC, o problema em concreto é a implementação ou não das tecnologias Cloud na área das bases de dados. Também neste capítulo encontram-se os objetivos deste trabalho, bem como os seus critérios de sucesso, para além da descrição da organização dos capítulos seguintes.

1.1 Problema

A gestão eficiente de uma base de dados é um dos pontos fulcrais em Tecnologias de Informação e Comunicação e é uma tarefa altamente relevante para a ciência e para a indústria. As vantagens dos sistemas de base de dados incluem a possibilidade de armazenar dados persistentemente, garantindo a independência de dados física, a capacidade de processar consultas declarativas, e a sua usabilidade independentemente do uso de aplicações específicas.

Os primeiros sistemas a incluir as propriedades ACID (*Atomicity, Consistency, Isolation, Durability*) eram sistemas de bases de dados relacionais com a sua estrutura simples e de fácil operacionalidade. Depressa chegaram os conceitos de orientação a objetos, *triggers*, ou processamento de XML (*eXtensible Markup Language*) complementando a sua funcionalidade.

As aplicações web modernas necessitam de uma plataforma de gestão de dados mais flexível e de fácil uso que permita o desenvolvimento evolutivo das bases de dados e das aplicações. Atualmente estamos confrontados com a tendência de afastarmo-nos da perspectiva do mundo fechado dos sistemas de bases de dados clássicos, no sentido de um mundo aberto, como por exemplo, na recente tendência para SaaS (*Software as a Service*).

Um SaaS é um modelo de oferta de software, em que o utilizador não tem de se preocupar nem com a instalação do software nem com o hardware utilizado pelo fornecedor do serviço. O cliente não paga a aquisição do software, mas sim o direito à sua utilização. Exemplos de SaaS são o Gmail, o Hotmail e o Office 365.

Um dos principais objetivos da investigação em bases de dados é fechar a lacuna entre os sistemas de bases de dados clássicos com todas as suas restrições e as plataformas de gestão de dados simples, mas altamente escaláveis. Só então é que as aplicações SaaS centradas em dados da próxima geração, serão capazes de oferecer simultaneamente suporte a alta escalabilidade, extensibilidade e garantias não-funcionais.

Tal como acontece com o processo da escolha de um RDBMS (*Relational DataBase Management System*), a escolha de uma plataforma DaaS (*Database as a Service*) apropriado deve compreender algumas das características de um RDBMS, como por exemplo a capacidade de desempenho de carga de trabalho.

1.2 Objetivos

Esta dissertação tem como objetivos a:

- Análise de custos, que podem ir desde os custos de aquisição de equipamento, da instalação e da manutenção de uma base de dados;
- Análise da performance de uma base de dados;
- Avaliação e seleção da melhor ferramenta de DaaS, com vista a ser implementada numa PME (Pequena e Média Empresa), usando os critérios mencionados nos pontos anteriores.

1.3 Critérios de Sucesso

Como critérios de sucesso, pode-se ter:

- A satisfação do cliente;
- Os testes de benchmarking efetuados antes e depois de uma migração de bases de dados de um modelo tradicional de RDBMS para a Cloud;
- Redução de custos.

1.4 Organização da Dissertação

Esta dissertação encontra-se organizada do seguinte modo:

- Capítulo 2 – Análise do Valor, nele encontra-se o modelo de negócio passível de ser criado;
- Capítulo 3 – Neste capítulo é introduzido o conceito de base de dados, dos objetivos e vantagens de DBMS (*DataBase Management System*) e a definição de DBA (*DataBase Administrator*);

- Capítulo 4 – Aqui descreve-se sumariamente o que é um RDBMS (*Relational DataBase Management System*) e quais são os principais RDBMS existentes atualmente no mercado;
- Capítulo 5 – Neste capítulo é apresentada-se uma pequena introdução do que é o Cloud Computing, o que é um DaaS (*Database as a Service*) e quais são os principais fornecedores desse tipo de serviço.;
- Capítulo 6 – Descreve-se, neste capítulo a avaliação experimental de um DaaS, nele descrevem-se ainda os testes que irão ser realizados para a verificação da solução implementada.
- Capítulo 7 – Nele apresentam-se as principais conclusões deste trabalho, bem como uma possível continuação do mesmo.

2 Análise de Valor

Este capítulo é dedicado à análise de valor, tendo como objetivo a criação de um possível modelo de negócio sustentável, capaz de ser criado no final deste trabalho.

Dele consta a definição de valor, alguns fatores a ter em consideração antes da aquisição de um DaaS, a definição de produto e uma proposta de um modelo Canvas.

2.1 Definições

Neste subcapítulo são abordados a definição de valor, bem como a definição de Capex, Opex e TCO (*Total Cost of Ownership*).

2.1.1 Valor

A definição de uma proposta de valor para um produto ou para um serviço é, por vezes, uma tarefa mais complicada do que aquilo que parece. A proposta de valor é usada para descrever a forma como uma empresa se diferencia dos seus concorrentes, ou seja, tem como objetivo mostrar ao consumidor que comprar um produto ou serviço a ele é um melhor investimento do que compra a qualquer outro fornecedor.

A criação de valor é um conceito que é difícil de ser obtido, entendido, modelado e ou conceptualizado. A criação de valor é uma relação de compromisso entre os benefícios e sacrifícios percebidos pelos clientes durante a oferta de um fornecedor. A criação de valor é a chave para qualquer negócio, qualquer atividade de negócios é relativa à troca de um bem ou de um serviço tangível e / ou intangível e ter o seu valor aceite e recompensado por parte dos clientes.

Valor percebido é a avaliação sumária de um produto realizada pelo cliente, tendo em consideração os benefícios e o preço. O valor total percebido é igual aos benefícios percebidos mais o preço percebido. Quando o benefício percebido é maior que o preço percebido, o valor total percebido é positivo e os clientes ficam com a sensação de que realizaram um bom

negócio. De modo oposto, quando o benefício percebido é inferior ao preço percebido, os clientes concluem que o produto não tem um bom valor e sentem-se explorados.

A definição de valor é necessária para evidenciar aquilo que torna uma empresa ou entidade como sendo “únicos”, ou seja, aquilo que a diferencia das outras, também tem como finalidade mostrar aos seus clientes que o seu produto, embora possa ser semelhante ao dos seus concorrentes, é melhor que o deles.

2.1.2 Capex, Opex, Total Cost of Ownership

No mundo empresarial, particularmente nas multinacionais, Opex e Capex são termos muito frequentes. Capex significa *Capital expenditures* enquanto que Opex significa *Operational expenses*. Enquanto que Capex é o custo inicial de investimento necessário para começar uma atividade ou projeto, o Opex é o custo operacional recorrente que temos com uma determinada atividade ou projeto.

O Opex é o valor necessário para manter uma atividade ou projeto em funcionamento após o seu arranque, nele incluem-se os custos da energia elétrica que se gasta ao longo do tempo com o equipamento (servidores, ar-condicionado), o custo das reparações dos equipamentos, ou outro tipo de custos regulares, como por exemplo o licenciamento anual de antivírus.

Ao invés de comprar equipamento informático e o alugarmos estamos a evitar o Capex, mas tem a consequência de estarmos a aumentar o Opex, deixará de existir um investimento inicial e passará a existir uma renda. Esta opção pode fazer sentido tendo em consideração a disponibilidade financeira no momento do início da atividade ou projeto.

Segundo (Ferreira, 2015), nas TI, os custos recorrentes de um ativo, ou seja, o Opex, são complexos de ser calculados, de entre esses custos salientam-se os seguintes:

- Energia elétrica consumido pelos servidores;
- Energia elétrica do ar condicionado necessário para refrigerar os equipamentos informáticos existentes numa sala informática ou num *datacenter*;
- Licenças anuais de software;
- Custos com o pessoal técnico, sendo este da empresa ou de serviços contratados a entidades externas, pessoal este necessário a manter os sistemas em funcionamento, mantendo-os atualizados e resolvendo as avarias de hardware que ocorrem;
- Investimento nouro tipo de equipamentos complementares, tais como UPS, gerador, alarmes, switches, sistemas de monitorização, de entre outros;
- Tempo que a empresa desperdiça ao dedicar-se a uma atividade que poderá não fazer parte da sua missão, do seu *core business*, ou dos seus objetivos empresariais, sendo que este aspeto é difícil de ser calculado.

O conceito de *Total Cost of Ownership* (TCO), ou custo total de propriedade é o valor total que um determinado projeto ou atividade tem ao longo do seu tempo de vida. Normalmente o uso de serviço Cloud é mais baixo do que a aquisição e manutenção de uma solução tradicional.

Com o decorrer do tempo vão aparecendo novos projetos internos. Com estes novos projetos a dificuldade de fazer uma análise de “interno versus Cloud” torna-se mais difícil de ser realizada. Os custos operacionais da empresa acabam por se encontrar diluídos pelos vários projetos que esta tem.

Em (Ferreira, 2015) aparece mencionado que segundo a International Data Corporation (IDC), “o TCO da operação de uma plataforma informática deve considerar vários componentes de custo; entre estes os mais importantes são: custos de colaboradores (60%), custos de downtime (15%), custos de formação (8%), hardware (7%) e outros (3%)”. O software representa um custo de 7%.

A análise do custo da opção a tomar entre um RDBMS tradicional e um DaaS deverá ser feita não com base num custo mensal, mas sim com base no TCO, tendo como base os valores do Capex (custos iniciais) e o Opex (custos operacionais). Quanto mais curta for a duração de uma atividade, maior peso terá o custo inicial no custo total de propriedade, assim tornando-se mais evidente a contratação de serviços a um fornecedor externo. Quanto mais longa for a duração de uma atividade, fica mais evidente que o custo inicial fica mais diluído ao longo do tempo e que tem um peso menor no calculo do custo total de propriedade, podem acontecer casos em que não se justifique adquirir serviços a um fornecedor externo. Nos casos em que não se justifique a contratação de serviços externos, deve-se ter em conta outros fatores (nível de qualidade e fiabilidade do serviço, a tolerância a falhas, a rapidez de instalação, etc.).

2.2 Fatores a considerar na aquisição de DaaS

Neste subcapítulo descrevem-se os principais fatores a ter em consideração antes da escolha da aquisição de um serviço tipo DaaS, tais como o SLA, as comunicações e os custos associados.

2.2.1 Service Level Agreement (SLA)

Um SLA é o nível de qualidade estabelecido entre um fornecedor e um cliente, para a prestação de um ou mais serviços, fazendo parte do contrato de prestação de serviços. O SLA pode abranger vários aspetos de prestação de serviços, como por exemplo a disponibilidade do serviço, o tempo de resposta ou de intervenção no caso de um incidente ou mesmo o tempo máximo para a resolução de uma avaria. Na Cloud, o parâmetro mais frequentemente utilizado é a disponibilidade do serviço.

2.2.2 Comunicações

Ao adquirir um serviço na DaaS, é necessário ter em conta na fase de planeamento que não deverão existir falhas nas comunicações, falhas essas que poderão passar pela rede local ou pela ligação à Internet.

Aquando da decisão de se usar um serviço DaaS é necessário garantir que todos os equipamentos que pretendemos que tenham ligação à base de dados tenham acesso à Internet, e a tenham com qualidade e sem interrupções. Poderá ser necessário garantir parâmetros de qualidade de serviço na ligação à Internet para os dispositivos clientes poderem aceder às bases de dados.

Na ligação a bases de dados não deverá existir uma grande latência na rede nem falhas, pois as bases de dados podem ficar com dados incoerentes, com transações pendentes, ou com demasiadas sessões abertas. Demasiadas sessões abertas podem implicar problemas de performance na base de dados, bem como podem provocar um aumento de custos na fatura mensal do serviço.

2.2.3 Custos

Deve-se ter em atenção ao adquirir um serviço do tipo DaaS se o custo do mesmo durante o período de tempo que pretendemos que seja usado, não ultrapassa o TCO no caso de aquisição de servidores próprios ou reutilização de existentes.

Deve-se também ter em atenção a finalidade para a qual o pretendemos usar, pois para fins de desenvolvimento de produtos, existem soluções específicas e de baixo custo.

De referir que na grande maioria dos casos os serviços DaaS são tributados por base de dados, em termos de custos e para um cliente com múltiplas bases de dados que têm o mesmo RDBMS pode em termos de custos ficar mais barato a médio-longo prazo uma solução de RDBMS tradicional, que passa pela aquisição de hardware e software.

2.3 Produto

Esta dissertação tem como intuito ajudar na escolha entre um RDBMS ou um DaaS para a utilização numa PME. Não se refere propriamente à escolha de um produto, mas sim de um serviço.

Como referido este trabalho tem como público alvo o mercado das PME, e pretende fornecer esclarecimentos relativos aos serviços DaaS.

Nem todas as bases de dados podem ser colocadas em ambiente Cloud, existem normas em cada país onde se evidenciam, em Portugal deve-se ter em conta os pareceres emitidos pela Comissão Nacional de Proteção de Dados (CNPd), todas as bases de dados que contenham

dados pessoais de indivíduos, devem estar registados na CNPD. Em Portugal as bases de dados da área da Saúde são reguladas pela Lei nº 12/2005 de 26 de janeiro.

As principais vantagens no uso de um serviço DaaS relativamente a um RDBMS tradicional são :

- As atualizações ao software são realizadas de forma automática, sem necessidade de intervenção por parte do cliente;
- A base de dados pode ser acedida a partir de qualquer lugar, basta para isso que exista acesso à Internet;
- O cliente tem um melhor controlo dos custos, pois a grande maioria dos sistemas em Cloud fornece aplicações que permitem o controlo dos gastos de forma gratuita;
- Não é necessário pagar por uma licença integral de um RDBMS;
- Elimina a necessidade de manutenção da infraestrutura física do cliente, no que diz respeito ao servidor de base de dados;
- Deixa de ser necessário a intervenção / monitorização de forma habitual de um DBA;
- Poderá não existir o custo de Capex, o único custo de Capex poderá ser o caso de existir migração da base de dados de um RDBMS tradicional para o DaaS realizado por um profissional de TI.

Como principais desvantagens são assinaladas as seguintes :

- Necessidade de acesso a internet, caso exista perda de acesso, todos os sistemas dependentes da base de dados ficarão comprometidos, os sistemas também poderão ficar comprometidos se não existir uma boa largura de banda;
- Todos os serviços DaaS têm custos, normalmente taxados mensalmente e cobrados por cada base de dados;
- Dependendo da localização geográfica onde se situa o *datacenter* do fornecedor de DaaS, os dados / informação podem ser acedidos livremente por algumas entidades governamentais, um exemplo de um país desses é os Estados Unidos ao abrigo do PATRIOT Act (<https://www.congress.gov/bill/107th-congress/house-bill/3162>). Dependendo do ramo de negócio ou do nível de confidencialidade da informação é necessário ter em consideração a localização geográfica onde o fornecedor de serviços possui a sua infraestrutura física.

Neste momento existem diversos fornecedores de serviços DaaS, tendo estes diversas ofertas no mercado. Sendo dispendioso "montar" um negócio baseado em prestação direta de serviços DaaS, ou seja, possuir uma infraestrutura de hardware e software própria, bem como possuir uma equipa qualificada para a sua gestão.

Assim sendo o aconselhável seria criar uma empresa de consultadoria, cujo principal objetivo fosse o aconselhamento aos diversos clientes sobre o futuro das bases de dados existentes nas suas empresas, podendo este passar pela mudança da localização física das suas bases de

dados, para a Cloud, ou o inverso, do seu correto dimensionamento e da sua escalabilidade. Para isso teria de contar com uma equipa experiente de DBA's (administradores de base de dados) e de SysAdmin's (administradores de sistemas).

Na Figura 1 propõe-se um possível modelo CANVAS para o negócio.

<p>Principais Parceiros</p> <p>Fornecedores de hardware Fornecedores de software Fornecedores de DaaS</p>	<p>Principais Atividades</p> <p>Consultadoria Prestação de serviços de DBA Prestação de serviços de SysAdmin</p>	<p>Propostas de Valor</p> <p>Qualidade no atendimento Seriedade Rigor Cumprimento dos prazos acordados</p>	<p>Relacionamento com os clientes</p> <p>Atendimento personalizado Fidelização do cliente</p>	<p>Segmentos de Clientes</p> <p>PME's</p>
<p>Estrutura de Custos</p> <p>Despesas com os recursos humanos Despesas com o local de trabalho Publicidade Despesas de deslocação aos clientes Equipamento informático Formação</p>	<p>Principais Recursos</p> <p>Recursos humanos</p>		<p>Canais</p> <p>Social Media Email Distribuição de panfletos</p>	

Figura 1 – Modelo CANVAS

3 Bases de Dados: Conceitos

Descrevem-se, neste capítulo algumas definições que permitirão dar contexto a este trabalho, pode-se verificar uma definição de base de dados, a definição de DBMS (DataBase Management System). Também são enumerados os objetivos e vantagens de um DBMS, o que é um DBA e o que é o Modelo Relacional.

3.1 Definição de Base de Dados

Segundo (Conger 2011) uma base de dados, na sua forma mais simples, é um conjunto de dados relacionados. Os catálogos em cartão que as bibliotecas usavam eram considerados uma base de dados. O caderno espiral de um cientista onde ele mantém notas e observações pode ser considerado uma base de dados. Quando nos referimos a "base de dados", normalmente referimo-nos a bases de dados em suporte eletrónico.

3.2 DataBase System (DBS) e DataBase Management System

Tendo como referência (Foster and Godbole, 2014), um sistema de base de dados (DBS) é um sistema de registo informatizado com o principal propósito de manter informações e disponibilizá-las sempre que necessário. A base de dados normalmente armazena os dados relacionados num computador.

Um Sistema de Gestão de Base de Dados (SGBD), segundo (Foster and Godbole, 2014), é o software que gere o armazenamento, a manipulação e a pesquisa dos dados existentes na base de dados, funciona como uma interface entre as aplicações e os dados que são necessários para a execução dessas mesmas aplicações.

Tudo o que fazemos com a base de dados passa pelo SGBD, este ajuda a adicionar, atualizar, a realizar cálculos, a catalogar os dados, a controlar o acesso a informações, manter cópias de segurança e mais importante permite encontrar a informação que pretendemos. Sem um SGBD as bases de dados não funcionam.

Segundo (Foster and Godbole, 2014), as funções críticas de um sistema de gestão de base de dados são as seguintes:

- Definição de dados (relações, dependências, restrições de integridade, visualizações, etc.);
- Manipulação de dados (adicionar, atualizar, apagar, recuperar, reorganizar e agregar);
- Verificações de segurança e integridade dos dados;
- Apoio a linguagens de programação.

De acordo com (Foster and Godbole, 2014), os componentes de um sistema de base de dados incluem:

- O hardware;
- O sistema operativo;
- O sistema de gestão de base de dados;
- A base de dados;
- Aplicações ou sistemas adicionais;
- Os utilizadores finais.

É referido em (Foster and Godbole, 2014), que os utilizadores finais comunicam com os sistemas de software / aplicações, que por sua vez comunicam (através de interfaces de comunicação) com os DBMS. O DBMS comunica com o sistema operativo, o qual por sua vez comunica com o hardware para armazenar dados em e / ou extrair dados da base de dados, conforme ilustrado na Figura 2.

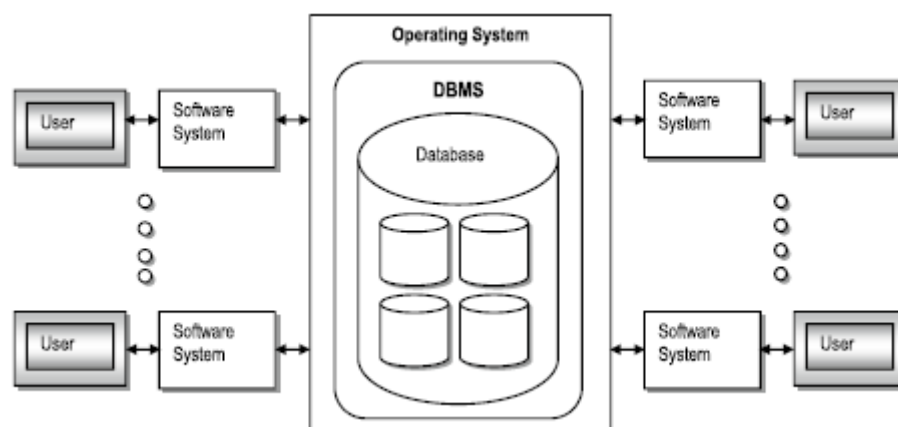


Figura 2 – Exemplo de um DBS, retirada de (Foster and Godbole, 2014, p. 4)

Os três níveis da arquitetura de um SGBD são:

- Nível físico – este nível corresponde à forma como os dados são armazenados e organizados internamente no sistema informático;
- Nível conceptual – corresponde à maneira como os dados são estruturados ou organizados ao nível da sua conceção lógica, ou seja, o número e o tipo de atributos em que a informação é estruturada, as relações entre os dados, etc.;

- Nível de visualização – corresponde à forma em quem os dados são apresentados aos utilizadores, através de interfaces gráficas proporcionadas por programas aplicativos ou pelos módulos do SGBD que trabalha com a base de dados. Este é o nível de abstração mais elevado da “arquitetura” da base de dados, pelo motivo que os utilizadores apenas têm de operar com as interfaces que permitem a visualização dos dados e não têm de conhecer ou se preocupar com a forma como os dados são tratados fisicamente nem como estes foram organizados conceptualmente.

3.3 Objetivos de um Sistema de Base de Dados

De acordo com (Foster and Godbole, 2014), existem vários objetivos primários e secundários de um sistema de base de dados que deveriam ser considerados pelos profissionais de informática.

Os objetivos principais são os seguintes:

- Segurança e proteção - prevenção de utilizadores não autorizados, proteção contra interferência inter-processo;
- Confiabilidade - garantia de estabilidade e desempenho previsível;
- Existência de múltiplos utilizadores;
- Flexibilidade - capacidade de obter dados via vários métodos;
- Facilidade de acesso a dados e alteração dos mesmos;
- Precisão e consistência;
- Clareza - padronização de dados para evitar ambiguidade;
- Capacidade de atender a solicitações imprevistas;
- Proteção do investimento intelectual;
- Minimização de proliferação de dados - novas necessidades das aplicações podem ser satisfeitas usando os dados já existentes em detrimento de criar novos ficheiros e programas;
- Disponibilidade - os dados estão disponíveis para os utilizadores sempre que necessário.

Os objetivos secundários são:

- Independência física dos dados - hardware de armazenamento e técnicas de armazenamento são isoladas a partir de programas de aplicação;
- Independência lógica dos dados - itens de dados podem ser adicionados ou subtraídos ou a estrutura lógica geral poder ser modificada sem que os programas sejam afetados;
- Controlo da redundância;

- Controlo de integridade - verificações e outros controlos devem impedir que dados inválidos entrem no sistema;
- Definição de dados - uso de dicionário de dados;
- Interface amigável para os utilizadores;
- Afinação - reorganização fácil da base de dados para melhorar o desempenho sem alterar a camada aplicacional;
- Reorganização automática de migração para melhorar o desempenho.

3.4 Vantagens de um Sistema de Base de Dados

Tendo novamente como referência (Foster and Godbole, 2014), um sistema de base de dados traz um elevado número de vantagens aos utilizadores finais, bem como às empresas, algumas dessas vantagens encontram-se abaixo mencionadas:

- A redundância pode ser reduzida;
- A inconsistência pode ser evitada;
- Os dados podem ser partilhados;
- Podem ser aplicadas restrições de segurança;
- A integridade dos dados pode ser mantida;
- Aumento do desempenho;
- Vistas lógicas dos dados podem ser facilmente criadas;
- Não é unicamente dependente de linguagens de programação de alto nível;
- Manutenção e recuperação de dados mais fácil;
- Melhor desempenho devido à velocidade de processamento, reduz o uso de papel.

3.5 Definição de DataBase Administrator (DBA)

Um administrador de base de dados, vulgo DBA, tem a responsabilidade do controlo do sistema de base dados a nível técnico. Algumas dessas funções são:

- Definição do esquema conceptual da base dados (nível lógico);
- Definição do esquema interno (nível físico);
- Definição das verificações de segurança e integridade dos dados;
- Definição de procedimentos para o backup e para a recuperação das bases de dados;
- Monitorização do desempenho dando resposta às possíveis mudanças de requisitos.

Em muitas organizações, a tendência é incluir a função de DBA nos descritivos das tarefas do engenheiro de software. Normalmente as grandes empresas em que as bases de dados necessitam de estar sempre disponíveis contratam os serviços de um ou mais DBAs.

De acordo com o relatório da Hays, disponível em: <http://guia-hays.pt/uploads/rte/Guia-do-Mercado-Laboral-2016-Hays-Portugal.pdf>, os profissionais desta área encontram-se entre os profissionais de Tecnologias de Informação (TI) mais bem pagos, tal pode ser verificado na página 82 do referido documento.

3.6 Modelo Relacional

Segundo (Foster and Godbole, 2014), o modelo relacional de bases de dados é de longe o modelo mais utilizado para o desenho de bases de dados. O modelo deve o seu sucesso ao facto de estar fortemente fundamentado em princípios matemáticos (teoria dos conjuntos e álgebra linear) que foram testados e comprovados. Como os princípios que o fundamentam, o próprio modelo foi testado e comprovado ao longo dos anos.

Na opinião de (Foster and Godbole, 2014), um Sistema Relacional de Base de Dados (RDBS) é uma coleção de relações normalizadas variáveis no tempo, geridas através de uma interface para o utilizador apropriado, com restrições desejáveis e com funcionalidades que permitam a gestão eficiente e eficaz das bases de dados.

É referido em (Stephens, Plew, and Jones, 2011), que uma base de dados relacional é uma base de dados dividida em unidades lógicas chamadas tabelas, essas tabelas estão relacionadas entre si. Este modelo de base de dados permite que os dados sejam divididos em unidades lógicas mais pequenas, permitindo assim uma manutenção mais fácil e também uma melhor performance de acordo com o nível da organização. Na Figura 3 podemos ver que as tabelas estão relacionadas entre si usando uma chave comum.

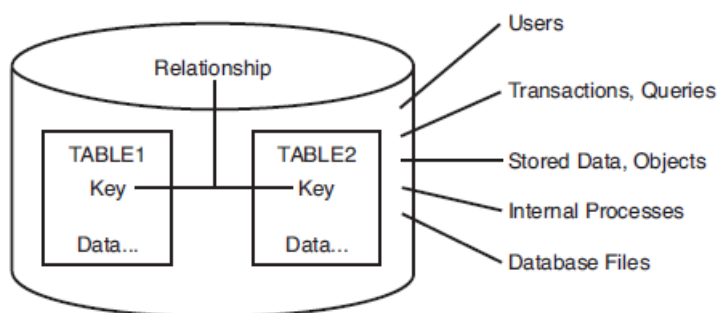


Figura 3 – Base de dados relacional, retirada de (Stephens, Plew, and Jones, 2011, p. 5)

A relação entre as tabelas, permite que os dados possam ser recuperados numa única consulta. Tendo chaves comuns ou campos entre tabelas relacionadas, os dados de múltiplas tabelas podem ser juntos para formar um grande conjunto de dados.

Concordando com a opinião de (Stephens, Plew, and Jones, 2011), os vendedores de bases de dados mais predominantes no mercado são a Oracle, a Microsoft, a Informix, a Sybase e a IBM. Estes vendedores distribuem várias versões de bases de dados relacionais pelo custo de uma

licença base e são consideradas de código fechado. Existem ainda fornecedores de soluções de código aberto de entre eles salienta-se o MySQL e o PostgreSQL.

4 Relational DataBase Management System (RDBMS)

A definição de um RDBMS (Relational DataBase Management System) e a descrição dos principais RDBMS existentes no mercado e uma breve introdução aos mesmos serão abordados neste capítulo.

4.1 Definição

Um Relational DataBase Management System, em português Sistema de Gestão de Base de Dados Relacional (SGBDR) é um SGBD desenhado especificamente para gerir a informação que está armazenada em uma ou mais tabelas, consoante o assunto a que dizem respeito.

O termo relacional refere-se ao modo como um SGBD espera que estejam organizados os dados que está a gerir. Uma relação, tal como é denominada na matemática, é simplesmente uma tabela de informação que está organizada num formato de linhas e colunas, ou seja, num formato tabular.

A obrigatoriedade de os dados estarem organizados num formato tabular pode aparentar ser uma restrição, mas não o é, quase toda a informação pode ser organizada de forma a ser armazenada em formato tabular, ou seja, numa ou mais tabelas.

4.2 Principais RDBMS

O site: <http://db-engines.com/en/ranking>, classifica os SGBD de acordo com a sua popularidade, este ranking é atualizado mensalmente. Em seguida apresentamos os seis primeiros classificados em fevereiro de 2016, que também o foram em janeiro de 2016. Os resultados encontram-se por ordem decrescente.

1. Oracle;
2. MySQL;
3. Microsoft SQL Server;
4. MongoDB;
5. PostgreSQL;
6. DB2.

Nos subcapítulos seguintes apresenta-se uma breve descrição dos cinco RDBMS pertencentes ao grupo acima mencionado pois segundo (Chodorow, 2013) o MongoDB é uma base de dados orientada a documentos e não uma base de dados relacional.

4.2.1 Oracle

A Oracle Corporation foi fundada em 1977 por dois programadores chamados de Larry Ellison e Bob Miner. Ambos tinham experiência prévia na construção de bases de dados em diferentes empresas. O primeiro projeto foi a construção de uma base de dados especial para a CIA (Central Intelligence Agency). Larry ficou com a parte de marketing e vendas do software enquanto que Bob cuidou do desenvolvimento do software.

Em 1978, eles criaram uma base de dados relacional usando SQL, segundo (Oracle Corporation story, 2016) a primeira do mundo, chamaram-na de Oracle RDBMS e permitia que os utilizadores usassem software normalizado, independentemente do hardware.

Foram a primeira empresa a vender software RDBMS e em 1982 tiveram uma receita de \$2.5 milhões, não estando satisfeitos com o sucesso instantâneo que tiveram, centraram-se novamente no desenvolvimento e investiram mais de 25% da sua receita em nova investigação. A empresa começou a expandir-se a um grande ritmo e em 1985 tiveram um volume de vendas de \$25 milhões, no ano seguinte chegou aos \$55 milhões. A Oracle centrou-se nos seus principais clientes que eram multinacionais do ramo automóvel, farmacêutico, aeroespacial e agencias governamentais.

A abordagem agressiva de marketing da Oracle permitiu-lhe estabelecer-se em mais de 50 países até o ano de 1987, nesse ano foi capaz de ultrapassar \$100 milhões em receitas e de se tornar a maior empresa de gestão de base de dados no mundo.

Até ao ano de 2001 a Oracle teve um volume de vendas de \$11 mil milhões sendo que \$2.6 mil milhões foram de lucro. O crescimento constante da Oracle foi possível por causa das constantes atualizações dos seus produtos e da incorporação da mais recente tecnologia nos seus produtos.

Em janeiro de 2010 anunciaram a compra da Sun Microsystems, avaliada em \$7 biliões, o que transformou a Oracle de uma empresa exclusivamente dedicada ao software para uma que produz software e hardware.

Em 2013 a Oracle tinha mais de 122 mil funcionários e um volume de vendas de \$37 biliões, sendo que \$11 biliões eram de lucro.

Passa-se a nomear abaixo as diferentes datas de lançamento do Oracle:

- 1979 – Oracle release 2 – o primeiro RDBMS comercial foi desenvolvido na linguagem assembler PDP-11, apesar de terem criado uma versão comercial em 1977, essa não estava disponível para venda até 1979 com o lançamento da versão 2. A empresa decidiu não começar pela versão 1 do produto, pois tiveram medo que o termo “versão 1” pudesse ser visto negativamente pelo mercado. Os primeiros clientes da Oracle foram a USAF (United States Air Force) e a CIA;
- 1983 – Oracle release 3 – esta versão foi desenvolvida em 1983, foi escrita em linguagem de programação C e podia ser executada em mainframes,

minicomputadores e computadores ou em qualquer hardware que tivesse um compilador de C. Permitia a execução de instruções SQL e transações;

- 1984 – Oracle release 4 –incluía suporte para leitura com consistência, o que a tornou mais rápida que qualquer versão anterior. Também nos trouxe a introdução de utilitários capazes de realizar exportação e importação, para além de um software para criar relatórios baseado em consultas;
- 1985 – Oracle release 5 – nesta versão foi introduzida a tecnologia de cluster, esta versão acrescentou alguns novos recursos de segurança, tal como a auditoria, o que permitiu ajudar a determinar quem e quando alguém acede à base de dados. Em 1986 foi lançada a versão 5.1 e permitiu apoiar nas consultas distribuídas, mais tarde, mas também naquele ano foi lançado o SQL Plus, que é uma ferramenta que oferece o acesso a dados ad hoc e na elaboração de relatórios, também nesse ano surgiu o SQL Forms;
- 1988 – Oracle release 6 – nesta versão foi introduzida a capacidade de backup a quente¹ e a bloqueio ao nível da linha. Antes desta versão os backups das bases de dados tinham de ser realizados pelos DBA's desligando as bases de dados, anteriormente também durante as operações de escrita a tabela inteira ficava bloqueada, passando apenas a bloquear a linha ou linhas que estão a ser usadas. Com a versão 6.2 foi introduzido o Oracle Parallel Server, esta nova funcionalidade fornecia alta disponibilidade, pois mais que um servidor pode aceder aos dados da base de dados, com o aumento da disponibilidade também existiu um aumento na performance do sistema;
- 1992 – Oracle release 7 – esta versão incluía uma vasta gama de novos recursos e capacidades em áreas como a segurança, a administração, o desempenho e o desenvolvimento. Fornecia total controlo de quem, quando e do que os utilizadores estavam a fazer na base de dados. Nesta versão passou a ser possível monitorizar cada comando. Com esta versão os utilizadores puderam passar a utilizar procedimentos e a usar gatilhos (triggers) para impor regras de negócio. Foi adicionada a funcionalidade de commit em duas fases para suportar transações distribuídas. Em 1996 foi lançada a versão 7.3 que oferecia aos clientes a possibilidade de gerir todos os tipos de dados, incluindo o vídeo, as imagens a cores, sons e dados espaciais, também trouxe o lançamento da autenticação biométrica;
- 1997 – Oracle release 8 –esta versão foi projetada para trabalhar com computadores de rede Oracle. Esta versão suportava Java, HTML e OLTP. Em 1998 foi lançado Oracle release 8i, que foi a primeira base de dados que suportava tecnologias Web, tal como o Java e o HTTP;
- 2001 – Oracle release 9i – O Oracle Real Application Cluster foi introduzido nesta versão, este recurso permite software de clustering e alta disponibilidade em ambientes Oracle. O suporte nativo a XML também foi introduzida nesta versão, e esta foi a primeira base de dados a fazê-lo. A versão 9i release 2 permitiu integrar bases de dados relacionais com multidimensionais. A versão 9i release 2 trouxe

¹ Um backup a quente, é um backup realizado com a base de dados em funcionamento.

consegue uma tecnologia especial que permitia a compressão das tabelas, permitindo assim a sua redução de tamanho entre 3 e 10 vezes, o que permitiu uma melhoria na performance ao aceder a essas mesmas tabelas;

- 2003 – Oracle release 10g – aqui apareceu a introdução de Grid Computing, esta também foi a primeira versão do Oracle com suporte para 64 bits em Linux. Com o Oracle Database 10g e o Real Application Cluster foi possível passar a usar uma infraestrutura mais barata, pois os mainframes eram muito dispendiosos, passou a ser possível usar servidores UNIX ou Windows que possuíam alta disponibilidade, escalabilidade e desempenho. Em 2005 a Oracle criou a versão 10g Express Edition, com esta introdução a Oracle deu a pequenas empresas uma opção viável para integrar o Oracle no mercado de trabalho, sem nenhum custo;
- 2007 – Oracle release 11g – esta versão introduziu mais funcionalidades do que as anteriores, incluía o Oracle Database Replay, uma ferramenta que permitia capturar consultas SQL e posteriormente realizá-las noutra base de dados, introduziu também uma funcionalidade que permitia backups em paralelo e passwords do tipo case sensitive de entre outras;
- 2013 – Oracle release 12c – esta é a última versão disponível no mercado, possibilita uma arquitetura que torna mais fácil a consolidação de bases de dados de uma forma mais rápida e geri-los como um serviço na cloud. O Oracle Database 12c também inclui recursos de processamento de dados na memória que oferecem um desempenho analítico revolucionário. Inovações adicionais oferecem novos níveis de eficiência, desempenho, segurança e disponibilidade. Existe agora em quatro edições para atender às necessidades do negócio e ao orçamento: Express Edition, Enterprise Edition, Standard Edition 2 e Standard Personal Edition.

Os dados acima encontram-se disponíveis e em maior detalhe no endereço: http://www.dba-oracle.com/t_history_oracle.htm.

Na Tabela 1 apresenta-se algumas das características da versão 12c.

Tabela 1 – Comparação entre as versões do Oracle 12c

	Express Edition	Standard Edition 2	Enterprise Edition
Número máximo de CPUs	1	2	Sem limite
Máximo de memória	1GB	Máx do SO	Máx do SO
Tamanho máximo da BD	11GB	Sem limite	Sem limite
Windows	Sim	Sim	Sim
Linux	Sim	Sim	Sim
Unix		Sim	Sim
Suporte a 64 bits		Sim	Sim

Mais informação relativa à comparação das versões existentes, pode ser encontrada em: <https://docs.oracle.com/database/121/DBLIC/editions.htm#DBLIC118>.

Em termos de licenciamento a Express Edition é gratuita. A Standard Edition 2 tem um custo que varia entre os €61 por utilizador da base de dados, para um ano, a licença perpétua tem o valor de €304, na Oracle Store o mínimo de encomenda é de 10 licenças, assim sendo o preço mínimo é de €610, se optarmos nesta versão pelo licenciamento por processador, o preço é de €3.039 por processador e para um ano, para ser perpétuo o custo é de €15.194 por processador. A Enterprise Edition tem um custo mínimo de €165 por utilizador de base de dados e na loja online existe uma quantidade mínima a adquirir que é de 25 licenças, se optarmos pelo licenciamento ao processador o custo é no mínimo de €8.248 por processador instalado, uma licença perpétua custa €41.240.

As licenças incluem a atualização do software, bem como o suporte ao mesmo. Em: https://shop.oracle.com/pls/ostore/product?p1=Database&p2=OracleDatabase&p3=&p4=&p5=&intcmp=ocom_database_oracledatabase encontram-se as diferentes versões para aquisição, existe a possibilidade de desconto na loja online.

4.2.2 MySQL

O MySQL apareceu em 1980, na Suécia, mais tarde, em 2008 a empresa que detinha os seus direitos, a MySQL AB foi adquirida pela Sun Microsystems, um ano mais tarde a Oracle adquiriu a Sun Microsystems. É um software do tipo open source.

Compete diretamente com outros RDBMS, nomeadamente com o SQL Server da Microsoft ou com o PostgreSQL. Utiliza linguagem do tipo SQL, para a criação da base de dados, seja para o seu controlo, acesso e manipulação dos dados.

Segundo (Tavares, 2015) é utilizado em milhões de páginas em todo o mundo, tendo-se tornado popular pela sua fiabilidade, robustez e facilidade de integração sobretudo com a linguagem PHP. Este é usado por grandes empresas, tais como a NASA, a Adobe, o LinkedIn, o Tumblr, mas também por grandes organismos, salientando-se a Casa Branca e a ONU. Em <https://www.mysql.com/industry/government/> pode-se encontrar uma listagem de instituições governamentais onde o MySQL é utilizado.

Existem 6 versões do MySQL, sendo que apenas a Community Version é gratuita, sob licença GPL, as versões disponíveis sob o licenciamento GPL podem ser encontradas em: <http://dev.mysql.com/downloads/>. O resumo das versões comerciais e as suas especificações estão disponíveis em: <http://www.mysql.com/products/>.

Os preços das versões comerciais variam entre os €1.740 e os €52.000, estes valores podem ser confirmados em: <https://shop.oracle.com/pls/ostore/product?p1=MySQL>.

O MySQL possui versões para as plataformas Windows, Linux, Unix e Mac OS X, tanto para 32 bits como para 64 bits. As versões para cada arquitetura podem ser constatadas no seguinte endereço: <http://www.mysql.com/support/supportedplatforms/database.html>.

Como opção pode-se instalar um conjunto de ferramentas existentes para ambientes Windows, Linux e Mac OS X disponíveis em: <http://dev.mysql.com/downloads/utilities/>. Para gerir e servir como ambiente gráfico ao MySQL e às suas diversas operações com as bases de dados, normalmente usa-se o MySQL Workbench, este é gratuito, encontrando-se disponível para download em: <http://dev.mysql.com/downloads/workbench/>.

Os limites máximos da última versão disponível do MySQL podem ser encontrados nas subpáginas do seguinte endereço: <https://dev.mysql.com/doc/refman/5.7/en/limits.html>.

Mais informações sobre o MySQL podem encontrar-se no livro MySQL de Frederico Tavares, editado pela FCA.

4.2.3 Microsoft SQL Server

Em 1988, a Microsoft juntou a Ashton-Tate e a Sybase para criar uma variante do Sybase SQL Server para o sistema IBM OS/2, que foi lançada em 1989.

Esta foi a primeira versão do Microsoft SQL Server, e serviu como entrada da Microsoft no mercado das bases de dados a nível empresarial, entrando assim em competição com a Oracle, a IBM e mais tarde com a Sybase.

Desde o lançamento do SQL Server 2000, têm sido realizados avanços no desempenho, em ferramentas IDE cliente e em vários sistemas complementares que vieram incluídos no pacote do SQL Server 2005. Estes são:

- SQL Server Integration Services (SSIS);
- Reporting Services;
- Analysis Services.

Segundo (Magalhães, 2015), “o SQL Server tem uma gestão relativamente simples e será talvez a opção ideal para as aplicações Web escritas em ASP/ASP.Net”.

Na Tabela 2 apresenta-se o ano de lançamento de algumas das versões do SQL Server, bem como a data da descontinuação do seu suporte pela Microsoft.

Tabela 2 – Datas de lançamento e fim de suporte de versões do SQL Server

Versão	Ano	Nome	Descontinuação do Suporte
7.0	1998	SQL Server 7.0	11/01/2011
8.0	2000	SQL Server 2000	09/04/2013
9.0	2005	SQL Server 2005	12/04/2016
10.0	2008	SQL Server 2008	09/07/2019
10.50	2010	SQL Server 2008 R2	09/07/2019
11.0	2012	SQL Server 2012	12/07/2022
12.0	2014	SQL Server 2014	09/07/2024
13.0	2016	SQL Server 2016	-

Mais informações relativas ao ciclo de vida do suporte para o Microsoft SQL Server podem ser verificadas no endereço: <https://support.microsoft.com/pt-pt/lifecycle?C2=1044>.

A versão SQL Server 2014, existe nas seguintes modalidades:

- Enterprise – segundo (Magalhães, 2015), esta é a versão mais completa e robusta sendo a mais indicada para empresas que necessitem de processamento de transações online e de aplicações de Data Warehousing. Tem características de performance e escalabilidade que permitem gerir grandes quantidades de dados com segurança e alta disponibilidade, bem como ferramentas que ajudam à redução da manutenção e minimizam o tempo de recuperação de desastre. Esta versão inclui um conjunto de componentes de Business Intelligence que podem transformar dados em importantes decisões empresariais. Mais detalhes em: <https://www.microsoft.com/pt-pt/server-cloud/products/sql-server-editions/sql-server-enterprise.aspx>;
- Standard – de acordo com (Magalhães, 2015), esta edição é pensada para um uso departamental em organizações médias e em infraestruturas que requeiram alta disponibilidade de dados e plataforma de análise. Tem suporte para características de alta disponibilidade, tais como Failover Clustering até 2 nós². Oferece também as funcionalidades de Business Intelligence, Integration Services, Analysis Services e Reporting Services. Mais detalhes em: <https://www.microsoft.com/pt-pt/server-cloud/products/sql-server-editions/sql-server-standard.aspx>;
- Business Intelligence – (Magalhães, 2015), diz-nos que esta é uma versão intermédia entre a versão Enterprise e a Standard, surgiu com a versão SQL Server 2012. Inclui todas as características da versão Standard e alguns extras, dos quais se salientam: DQS, MDS, PowerView, PowerPivot, Semantic Model e análises avançadas. Mais detalhes em: <https://www.microsoft.com/pt-pt/server-cloud/products/sql-server-editions/sql-server-business-intelligence.aspx>;
- Express – segundo (Magalhães, 2015) esta versão foi desenhada para permitir a construção de aplicações simples e dinâmicas, robustas e confiáveis. É de uso e distribuição livres, no caso de deixar de corresponder às necessidades do utilizador é facilmente atualizável para versões superiores. Mais detalhes em: <https://www.microsoft.com/pt-pt/server-cloud/products/sql-server-editions/sql-server-express.aspx>;
- Compact Edition – tendo em consideração (Magalhães, 2015), este é um sistema de base de dados compacto, de baixa manutenção e para aplicações clientes monoutilizador para as plataformas Windows, como os tablets, PC, PDA, smartphones, ou mesmo sistemas desktop. Esta versão é de uso e distribuição livres;

² Um failover cluster fornece alta disponibilidade a um sistema, neste caso de 2 nós, sendo um deles ativo e outro passivo. Neste caso o nó ativo tem as bases de dados em funcionamento, no caso de falha do nó ativo, os serviços são assegurados pelo nó passivo.

- Web Edition – foi concebido para suportar workloads com acesso à Internet, permitindo que as organizações implementem rapidamente páginas Web, aplicações, Web sites e serviços;
- Developer Edition – (Magalhães, 2015), escreve que esta versão inclui todas as características e potencialidades da versão Enterprise, tendo sido desenhada exclusivamente para o desenvolvimento de aplicações e bases de dados. Assim cada licença só permite o uso por um único utilizador e para os fins anteriormente descritos.

A versão SQL Server 2016, existe nas seguintes modalidades:

- Enterprise – tem um desempenho em memória abrangente e crítico para o negócio, com segurança sem rival, uma solução de business intelligence empresarial ponto a ponto com BI móvel incorporado e análise avançada;
- Standard – possui capacidades centrais de gestão de dados e de business intelligence com recursos de TI mínimos;
- Web - Plataforma de dados segura, rentável e altamente escalável para sites públicos, só está disponível para fornecedores de serviços de software externos;
- Developer – gratuita, e possui o mesmo conjunto de funcionalidades do SQL Server 2016 Enterprise Edition, esta versão permite que os programadores criem, testem e demonstrem aplicações num ambiente de não produção;
- Express – base de dados gratuita de nível de entrada, ideal para a implementação de bases de dados de pequenas dimensões em ambientes de produção.

Mais detalhes sobre as versões existentes encontra-se em <https://msdn.microsoft.com/en-us/library/ms144275.aspx>.

Na Tabela 3 e na Tabela 4 apresenta-se uma pequena comparação entre as versões disponíveis do SQL Server 2014 e do SQL Server 2016 respetivamente.

Tabela 3 – Comparação entre as versões do SQL Server 2014

Funcionalidades	Enterprise	Business Intelligence	Standard	Express
Numero máximo de cores	Máx do SO	16	16	4
Máximo de memória / instância	Máx do SO	128GB	128GB	1GB
Tamanho máximo da BD	524PB	524PB	524PB	10GB

Tabela 4 – Comparação entre as versões do SQL Server 2016

Funcionalidades	Enterprise	Standard	Express
Numero máximo de cores	Máx do SO	24	4
Máximo de memória / instância	Máx do SO	128GB	1GB
Tamanho máximo da BD	524PB	524PB	10GB

Em: <https://www.microsoft.com/pt-pt/server-cloud/products/sql-server-editions/>, podemos verificar uma lista mais detalhada das funcionalidades do SQL Server 2014. Esta comparação também pode ser verificada nas páginas 22-25 e 26-27 de (Magalhães, 2015).

O SQL Server 2014 e o SQL Server 2016 podem ser licenciados de acordo com os seguintes modelos:

- Licenciamento por core de CPU – este tipo de licenciamento, obriga a uma licença por cada 2 cores de processador físico. Este modelo destina-se às edições Enterprise e Standard. Exemplo: dois processadores com quatro cores cada, implica a compra de 4 licenças;
- Servidor mais CAL (Client Access License) – este licenciamento obriga a uma licença para o sistema que corre o motor de base de dados e a uma licença adicional para cada dispositivo que estabeleça uma ligação ao SQL Server. Este modelo está disponível para as versões Business Intelligence Standard. As licenças de CAL 2014 permitem aceder a base de dados de versões anteriores, o inverso não se verifica. Também cada CAL permite o acesso a múltiplos servidores de SQL Server;
- Licenciamento para virtualização – este modelo de licenciamento permite alguma maleabilidade, existindo duas opções, licenciar cada máquina virtual individualmente ou licenciar cada host físico em que cada máquina virtual é executada, este último é pensado para ambientes de cloud privada e para ambientes densamente virtualizados. De referir que para o primeiro caso a Microsoft só vende no mínimo quatro licenças de core para cada servidor.

Para o licenciamento por core deve-se ter em conta a Tabela 5, nela podemos constatar o fator de multiplicação a aplicar na fórmula do cálculo do licenciamento por core.

Tabela 5 – Fatores de multiplicação para o licenciamento por core

Tipo de Processador	Fator
Single Core	2
Dual Core	4
Processadores AMD das séries 31XX, 32XX, 33XX, 41XX, 42XX, 43XX, 61XX, 62XX, 63XX, com 6 ou mais cores	0,75
Todos os Outros	1

O licenciamento para ambientes de alta disponibilidade deve ter em consideração o número de servidores que compõem o cluster de base de dados e o número de instâncias instaladas. Se existir apenas uma instância, o cluster é ativo-passivo, ou seja, todas as bases de dados são executadas no mesmo nó do cluster, neste caso só é necessário adquirir licenciamento para um servidor. No caso de existirem mais instâncias, o cluster pode ser ativo-passivo, todas as instâncias são executadas no mesmo nó do cluster, sendo o licenciamento igual ao anteriormente descrito, mas distribuído as instancias pelos nós do cluster, passa a ser um

cluster ativo-ativo, assim sendo necessitamos de adquirir licenciamento para cada um dos nós que compõem o cluster.

Tanto no SQL Server 2014, como no SQL Server 2016, a versão Express é totalmente gratuita, a versão Developer tem um custo de cerca de €40 para a versão 2014, a versão 2016 é gratuita , para as restantes versões pode-se consultar o seguinte endereço: <http://mla.microsoft.com/default.aspx?language=Portuguese&country=PT>, no qual podemos encontrar um simulador para os diferentes produtos da Microsoft.

Mais detalhes sobre o SQL Server 2014 podem ser encontrados no livro SQL Server 2014 de Alberto Magalhães, editado pela FCA.

4.2.4 PostgreSQL

O PostgreSQL foi desenvolvido em 1986 pelo professor Michael Stonebraker da Universidade da Califórnia, mais precisamente da Universidade de Berkeley.

Michael Stonebraker e os seus alunos desenvolveram o Postgres entre 1986 e 1994. Andrew Yu e Jolly Chen, dois alunos de Stonebraker, substituíram em 1995 a linguagem de query do Postgres, que era denominada de Postquel pela linguagem SQL e batizaram o novo sistema como Postgres95.

A partir de 1996 o Postgres95 saiu do meio académico e foi adotado pela comunidade open source, sendo livre para descarregar e utilizar.

O PostgreSQL pode ser executado em diferentes sistemas operativos, tais como BSD, Linux, Mac OS X, Solaris e Windows. É fácil de instalar. Os limites máximos do sistema são impressionantes como se pode verificar na Tabela 6.

Tabela 6 – Limites máximos PostgreSQL, transcrito de (Caldeira, 2015, p. 18)

Objeto	Limite
Tamanho máximo da base de dados	Ilimitado
Tamanho máximo de uma tabela	32 Terabytes
Comprimento máximo de uma linha	1,6 Terabytes
Tamanho máximo de uma coluna	1 Gigabyte
Número máximo de linhas numa tabela	Ilimitado
Número máximo de colunas numa tabela	250-1600 (depende do tipo de dados)
Número máximo de índices numa tabela	Ilimitado

A página da Internet deste SGBDR é: <http://www.postgresql.org>

Em janeiro de 2016 foi lançado o PostgreSQL 9.5, as melhorias introduzidas nesta versão podem ser lidas em: <http://www.postgresql.org/docs/9.5/static/release-9-5.html>.

Mais detalhes sobre o PostgreSQL podem ser encontrados no livro PostgreSQL Guia Fundamental de Carlos Pampulim Caldeira, editado pelas Edições Sílabo.

4.2.5 DB2

O DB2 teve as suas origens no início dos anos 70, quando Edgar F. Codd, um investigador que trabalhava para a IBM escreveu a Teoria das Bases de Dados Relacionais e em junho de 1970 publicou o Modelo de Manipulação de Dados.

Em 1974, o centro de investigação da IBM em San Jose, desenvolveu um SGBD relacional intitulado de System R, isto para implementar os conceitos de Codd. Um desenvolvimento importante do projeto System R foi o SQL. Naquela época a IBM não acreditou nas ideias de Codd, deixando a implementação do projeto a um grupo de programadores que não esteve sob a sua supervisão, esta implementação violava vários dos fundamentos do modelo relacional de Codd, o resultado foi intitulado de Structured English QUery Language ou SEQUEL, na altura do lançamento comercial do produto mudaram aquele nome para Structered Query Language em abreviado SQL, pois a marca SEQUEL já se encontrava registada para o fabricante de aviação Hawker Siddeley.

O primeiro produto comercial de base de dados relacional da IBM, o SQL/DS foi lançado para DOS/VSE e VM/CMS em 1981.

O nome de DB2, ou IBM Database 2, apareceu apenas em 1983 quando a IBM o lançou para os seus mainframes Multiple Virtual Storage (MVS).

Durante alguns anos, apenas esteve disponível para ser usado em mainframes da IBM, mais tarde foi lançado para outras plataformas que incluíam o OS/2, o UNIX e o Microsoft Windows, depois foi lançado para o Linux, este processo decorreu na década de 1990.

Na Tabela 7 apresenta-se o ano de lançamento de algumas das versões do DB2.

Tabela 7 – Datas de lançamento das versões do DB2

Data	Versão	Nome de código
Meados dos anos 1990	DB2 Parallel Edition	
Meados de 2006	DB2 9	Viper
Outubro de 2007	DB2 9.5	Viper 2
Junho de 2009	DB2 9.7	Cobra
Inícios de 2012	DB2 10.1	Galileo
Junho de 2013	DB2 10.5	Kepler

De acordo com (Foster and Godbole, 2014), os principais componentes de uma suite DB2 são:

- DB2 Universal Database Core – contém os serviços de core do SGBD. No momento da aquisição deve ser especificado a edição a comprar, bem como o sistema operativo

onde vai ser instalado. Independentemente da versão e sistema operativo onde é executado, todos os aplicativos DB2 são acessíveis entre eles e mesmo de outro software de desenvolvimento da IBM, tais como o Rational e o WebSphere. Para além disso, usando drivers ODBC podem ser acedíveis outras bases de dados para além da DB2;

- IBM InfoSphere Information Server – é um conjunto de tecnologias que combina os SGBD, serviços Web, replicação e de data warehousing numa plataforma comum. É usado para integração de dados a partir de fontes heterógenas. Os dados de origem podem ser de diferentes SGBD (Oracle, Microsoft SQL Server, etc.). Pode-se transformar esses dados num formato que oferece fácil acesso à informação em toda a empresa, dando ao utilizador final a ilusão de que todos os dados residem na sua máquina;
- IBM Data Studio – este produto cobre todo o ciclo de vida (conceção, desenvolvimento, implementação e gestão) para todas as aplicações DB2;
- IBM InfoSphere Warehouse – com este recurso pode-se construir data warehouses (“armazéns de dados”) departamentais e simples, ou mesmo mais complexos para ser usado por uma empresa inteira ou mesmo por um grupo de empresas.

Podem existir outros componentes adicionais dependendo da versão.

Existe nas seguintes versões:

- DB2 Developer Edition;
- DB2 Express-C;
- DB2 Express Server Edition;
- DB2 Advanced Workgroup Server Edition;
- DB2 Enterprise Server Edition;
- DB2 Advanced Enterprise Server Edition.

Segundo o site da IBM, a primeira versão acima mencionada é para fins de desenvolvimento, as três seguintes são para o uso de pequenas e médias empresas, enquanto que as últimas duas são para grandes empresas.

A versão Express-C é gratuita para uso e distribuição, no entanto pode ser adquirido suporte à mesma mediante um pagamento anual. Esta versão pode ter até 15TB de tamanho por cada base de dados, no entanto tem as seguintes limitações:

- Máximo uso de 2 cores de CPU;
- Máximo de 16GB de memória RAM.

Mais informações sobre as diferentes versões podem ser encontradas no seguinte endereço: <http://www.ibm.com/developerworks/data/library/techarticle/dm-1311db2compare/>.

No que diz respeito aos custos, para Portugal os preços podem ser obtidos através de pedidos de cotação no site da IBM.

5 Database as a Service (DaaS)

Segundo (Bond, 2015), os DaaS são serviços que oferecem diversos tipos de bases de dados, como um recurso consumível da Cloud. Este serviço segue o modelo de Plataforma as a Service (PaaS).

Neste capítulo pode-se constatar uma definição de Cloud Computing, bem como algumas das suas características, os seus modelos de serviço e os seus modelos de implementação. Também, aqui, está pode ser verificada uma definição de DaaS, bem como alguns dos mais conceituados fornecedores de DaaS.

Em (Abadi, 2009) é referido que os SGBD são candidatos potenciais para serem implementados na Cloud, isto porque os sistemas de bases de dados corporativos normalmente têm custos iniciais proibitivos, tanto ao nível do hardware como do software. Também refere que para muitas empresas, especialmente para as startups e para as PME's o modelo de serviço da Cloud é muito atrativo, pois para além de ser pago conforme o uso, também deixa de existir a preocupação na manutenção do hardware. Este artigo foi publicado em 2009.

5.1 Definição de Cloud Computing

Em (Ferreira, 2015), aparece a definição de Cloud Computing do National Institute of Standards and Technology (NIST):

“Cloud Computing é um modelo que permite o acesso ubíquo, conveniente e a pedido, através da rede, a um conjunto de recursos de computação partilhados (redes, servidores, armazenamento, aplicações, serviços, etc.), que podem ser rapidamente provisionados ou libertados, com um mínimo de esforço e sem interação com o fornecedor.

O modelo Cloud é composto por cinco características essenciais, três modelos de serviço e quatro modelos de implementação.”

Como introdução ao mundo da Cloud, recomenda-se a leitura do livro *Introdução ao Cloud Computing*, de António Miguel Ferreira, editado pela FCA.

5.1.1 Características Essenciais

Segundo (Bond, 2015), as cinco características essenciais de Cloud Computing são:

- On-demand self-service (autosserviço a pedido): representa a capacidade de um utilizador poder unilateralmente aprovisionar recursos de computação, sem interagir com outras pessoas, mais concretamente com alguém do departamento comercial ou técnico do fornecedor de serviços que está a utilizar. Se for necessário a interação com outras pessoas e necessitar que alguém trate do pedido, então não está a ser usado um serviço Cloud;
- Broad network access (acesso generalizado à rede): as capacidades ou os recursos estão disponíveis na rede (Internet, por exemplo), estes podem ser acedidos através de mecanismos comuns, como por exemplo através de um browser, o que promove a utilização de plataformas clientes comuns. Isto significa que pode ser utilizado um computador normal, ou mesmo um smartphone para poder ter acesso aos serviços Cloud;
- Resource pooling (acesso partilhado a recursos): os recursos de computação do fornecedor são partilhados por vários clientes, sendo estes associados ou dissociados dinamicamente aos clientes, a pedido dos mesmos. Exemplos de recursos incluem o espaço em disco, a capacidade de processamento, a memória e a largura de banda. Os recursos contratados são para uma utilização exclusiva do cliente e são garantidos;
- Rapid elasticity (elasticidade rápida): os recursos são disponibilizados ou cancelados de uma forma elástica, nalguns casos até automaticamente dependendo da procura. É, por exemplo, ter a capacidade de num momento aumentar as características técnicas de um servidor, ou de aumentar o número de servidores, como noutra momento reduzir os recursos afetados;
- Measured service (serviço medido): a utilização de recursos pode ser monitorizada, controlada e reportada. Na Cloud os modelos de faturação são pay-per-use ou usage-based (paga apenas o que utiliza) e, na maioria dos casos, o que foi usado, apenas pode ser calculado no final de um determinado período, tipicamente um mês.

5.1.2 Modelos de Serviço

De acordo com (Ferreira, 2015), com (Bond, 2015) e com (Abourezq and Idrissi, 2016) existem os seguintes modelos de serviço da Cloud:

- Software as a Service (SaaS) – software como serviço, é a capacidade dada a um cliente ou utilizador de aceder a uma aplicação de software que é executada numa infraestrutura de computação. O utilizador não gera nem controla a infraestrutura de

computação, incluindo sistemas operativos, rede, memória, processamento, espaço em disco, entre outros. Um dos mais básicos exemplos de SaaS é o serviço de webmail como, por exemplo, o Gmail ou o Google Drive. Segundo (Abourezq and Idrissi, 2016) este é sem dúvida o mais conhecido e mais usado serviço disponibilizado pela Cloud;

- Platform as a Service (PaaS) – plataforma como um serviço, é a possibilidade de colocar na Cloud aplicações criadas ou adquiridas pelos utilizadores, que usam linguagens de programação, bibliotecas de funções, serviços e ferramentas suportadas pelo fornecedor. A infraestrutura da Cloud não é gerida nem controlada pelo utilizador, mas tem controlo sobre as aplicações por si instaladas e poderá também ter o controlo de algumas configurações no ambiente que as aloja. Um exemplo popular nos dias de hoje é o Heroku (<http://www.heroku.com>), um serviço PaaS para quem utiliza as linguagens Ruby, Node.js, Clojure, Java, Python ou Scala;
- Infrastructure as a Service (IaaS) – infraestrutura como um serviço, é a capacidade dada ao utilizador de instalar recursos fundamentais para a computação, como processamento, armazenamento, memória e rede, nesses recursos o utilizador pode instalar e correr software. O utilizador não gere nem controla a infraestrutura (não é ele que escolhe o hardware nem o instala), mas possui controlo sobre o sistema operativo e aplicações e possivelmente também sobre alguns aspetos da rede. Através do uso de um browser o utilizador pode escolher os recursos computacionais desejados (hardware, sistema operativo e o endereçamento IP desejado) para criar um servidor virtual numa questão de minutos ou mesmo de segundos. O exemplo mais conhecido a nível mundial de um fornecedor de IaaS é a Amazon Web Services (<http://aws.amazon.com>), que tem por exemplo como clientes a NetFlix, a rede social Reddit, o Financial Times, a NASA, a seguradora MAPFRE ou mesmo o Spotify.

5.1.3 Modelos de Implementação

(Ferreira, 2015), menciona que existem quatro modelos de implementação, estes definem a forma como o serviço nos é prestado. Os dois principais modelos são o *public* e o *private* e as duas variantes são a *hybrid* e a *community*. Abaixo encontra-se uma breve descrição dos modelos:

- Public Cloud ou cloud pública – a infraestrutura é disponibilizada para utilização do público em geral. Pode ser detida e gerida por uma empresa, uma instituição governamental ou uma instituição académica. Esta Cloud encontra-se nas instalações do fornecedor do serviço;
- Private Cloud ou cloud privada – a infraestrutura é para uso exclusivo de uma única organização. Pode ser detida ou gerida pela própria organização ou por uma entidade externa, podendo estar fisicamente localizada na própria organização ou num *datacenter* externo;
- Community Cloud ou cloud da comunidade – a infraestrutura de Cloud é disponibilizada para uso exclusivo de uma comunidade de utilizadores de empresas

que partilham determinadas preocupações. Esta pode ser detida ou gerida por uma ou mais das empresas da comunidade, ou por uma entidade externa, tal como na private cloud esta pode estar fisicamente localizada numa das empresas da comunidade ou num organismo externo;

- Hybrid Cloud, ou cloud híbrida – a infraestrutura desta Cloud é uma combinação de duas ou mais infraestruturas de Cloud, estas continuam separadas, mas partilham tecnologia que permite a portabilidade dos dados ou das aplicações.

Os mesmos modelos de serviço são mencionados em (Abourezq and Idrissi, 2016) e em (Bond, 2015), sendo que este último refere ainda um quinto modelo denominado de VPC (*Virtual Private Cloud*).

5.1.4 Mitos

É referido em (Ferreira, 2015), que existem mitos relativamente à Cloud, pelo que se descrevem três deles:

- A Cloud é imaterial:
A Cloud tem existência física, pois possui componentes de hardware, mesmo os dados alojados na Cloud possuem existência física, pois encontram-se gravados em suportes que têm forma física. Ela encontra-se espalhada nos diversos *datacenters* de fornecedores de Cloud, ou seja, dentro de edifícios que se encontram preparados para um funcionamento de 24 horas por dia, 365 dias por ano. Estes encontram-se localizados em determinadas zonas de países específicos.
Os utilizadores podem não saber onde a sua informação se encontra, mas o fornecedor do serviço sabe onde essa informação se encontra;
- A Cloud é insegura:
A Cloud não veio introduzir nem mais nem menos segurança no sistema, pois ter um servidor em qualquer lado ligado à Internet é tão potencialmente inseguro como ter um servidor virtual na Cloud. O fornecedor de serviço até poderá ter processos de gestão e operação mais rigorosos, ou mesmo recursos mais qualificados para proporcionar um maior nível de segurança;
- A Cloud é uma moda passageira:
A Cloud não é uma tecnologia, mas sim um conceito, as tecnologias tornam-se obsoletas ao final de algum tempo. A Cloud “é uma forma mais eficiente de se utilizarem as TI. Aquilo que a Cloud nos dá são essencialmente benefícios, na forma como lidamos com os recursos das TI” (Ferreira, 2015).

5.2 Database as a Service (DaaS)

As bases de dados são componentes essenciais para o arquivo organizado de informação, estando disponíveis em qualquer fornecedor de PaaS. No entanto e segundo (Ferreira, 2015)

existem outros fornecedores de serviços na Cloud, que, não sendo fornecedores de PaaS, disponibilizam também serviços de bases de dados. O exemplo mais conhecido é a Amazon, que sendo um fornecedor de IaaS, também fornece serviços de DaaS através do Amazon RDS (serviço de MySQL, Oracle e SQL Server).

De acordo com (Abourezq and Idrissi, 2016), DaaS pode ser simplesmente definida como um paradigma para a gestão de dados em que um fornecedor de serviço hospeda uma base de dados e fornece ao mesmo tempo todo suporte ao hardware e software inerente ao uso dessa mesma base de dados. As empresas que usam este tipo de serviço deixam de se preocupar com as operações de gestão da base de dados, desde a instalação do servidor e do SGBD até aos backups.

No caso deste serviço, e tendo em consideração as características da Cloud, o fornecedor irá cobrar o serviço mediante a sua utilização. Dependendo do fornecedor poderão ser utilizados um ou mais critérios para cobrar o serviço, tais como a carga de CPU, o espaço em disco ocupado, a replicação dos dados, os backups, o número de transações, de entre outros.

Tendo em consideração (Ferreira, 2015), a gestão de uma base de dados de forma a que esta seja eficiente na utilização dos recursos, seja rápida e escalável, assim como segura e resiliente é uma tarefa algo complexa, que normalmente é executada por DBA's.

Na utilização de um serviço DaaS evita-se a preocupação com os aspetos mais técnicos de administração de base de dados, que na grande maioria dos casos estão relacionados com a infraestrutura dos servidores, quem gere estes aspetos é o fornecedor do serviço.

Segundo (Ferreira, 2015) as bases de dados relacionais mais populares em ambiente Cloud são a MySQL e a PostgreSQL existindo também outras como a Oracle e o Microsoft SQL Server.

5.3 Fornecedores de DaaS

Neste subcapítulo descreve-se algumas das soluções de DaaS existentes no mercado, as quais seguem o modelo relacional, sendo elas:

- Oracle Cloud
- Azure SQL Database
- Google Cloud SQL
- Amazon RDS

Foi publicado pela Gartner um estudo denominado de “Magic Quadrant for Operational Database Management Systems”, publicado em 12 de outubro de 2015 e está disponível para consulta em <https://www.gartner.com/doc/reprints?id=1-2PO8Z2O&ct=151013&st=sb>.

A Oracle, a Microsoft e o Amazon Web Services aparecem no quadrado mágico da Gartner para Sistemas de Gestão de Bases de Dados Operacionais, como se pode verificar na Figura 4.



Figura 4 – Quadrado Mágico da Gartner para SGBD Operacionais

Para este estudo a Gartner definiu como Sistema de Gestão de Bases de Dados Operacional, um sistema que suporta múltiplas estruturas e tipos de dados, tais como XML, texto, JSON, áudio, imagem e vídeo, devendo estes incluir mecanismos para isolar os recursos e controlar os acessos dos utilizadores finais ao conteúdo das bases de dados.

Foram utilizados 15 critérios de avaliação neste estudo, a Oracle ficou como sendo forte na ampla gama de opções disponíveis, na funcionalidade e na inovação e na flexibilidade dos preços. A Microsoft é forte na visão de mercado, na performance e suporte. O Amazon AWS é forte pois fornece diversos tipos de serviço, pelo posicionamento geográfico e pela facilidade da criação de negócio com a Amazon.

Nos próximos subcapítulos descreve-se com algum detalhe os DaaS disponíveis pelas três empresas constantes no topo do ranking deste quadrado mágico, bem como o modelo de serviço disponível pela Google.

5.3.1 Oracle Cloud

A Oracle Cloud oferece várias opções de implementação, essas opções permitem que se comece a nível de custos e capacidade adequada para cada caso de uso e, em seguida, dá a flexibilidade para se adaptar conforme a mudança das necessidades. As opções incluem:

- Oracle Database Cloud Service;
- Oracle Database Exadata Cloud Service;
- Oracle Database Exadata Express Cloud Service - Managed;

- Oracle Database Schema Cloud Service - Managed.

O serviço Oracle Database Cloud Service e o Oracle Database Exadata Cloud Service, tem como unidade de medida o OCPU (Oracle CPU), esta medida equivale a um núcleo físico de um processador Intel Xeon com Hyper Threading habilitado. Cada OCPU corresponde a duas threads de execução, conhecidos com vCPUs.

5.3.1.1 Oracle Database Cloud Service

Oferece serviços de bases de dados elásticas para o desenvolvimento de aplicações, testes e implementação em produção. O serviço oferece uma interface web fácil de usar e uma API RESTful que permite criar e administrar bases de dados Oracle.

Permite escolher entre serviços pagos à hora ou por subscrição de bases de dados Oracle 11g ou 12c. Nos serviços pagos à hora, os preços podem variar entre os €0,88 e os €7 por cada hora de uso, para quem usar as subscrições, os preços variam entre os €260 e os €3.321 mensais.

Na criação de um serviço deste tipo, as escolhas que aparecem como rápidas encontram-se na Tabela 8.

Tabela 8 – Opções de escolha rápida na criação de uma base de dados em Oracle Database Cloud Service

Características	Escolhas rápidas			
	Pessoal	Equipa de Desenvolvimento	Arranque	Empresarial
OCPU	1	1	2	4
Tamanho máximo da BD	10GB	50GB	512GB	1TB
Cópia de segurança	Incluído	Incluído	Incluído	Incluído
Valor mensal	€ 211	€ 263	€ 690	€5.536

Em <https://shop.oracle.com/cloudstore/?product=database> podemos verificar as diversas opções disponibilizadas pela Oracle, para a criação deste modelo de implementação de DaaS da Oracle, os dados da Tabela 8 foram verificados em 2 de outubro de 2016.

5.3.1.2 Oracle Database Exadata Cloud Service

Esta implementação traz todo o poder do Exadata para a Cloud. O serviço inclui todos os benefícios de desempenho Exadata e 100% de compatibilidade com aplicações e bases de dados que sustentam modelos de negócio críticos.

O Oracle Exadata utiliza servidores de alta performance, bem como sistemas de armazenamento inteligente, que se encontram ligados por uma rede ultrarápida, denominada de InfiniBand e que possui uma baixa latência. Inclui armazenamento flash PCI, o qual oferece o mais alto rendimento e os melhores tempos de resposta. Usa a versão mais recente do

motor de base de dados que é o Oracle Database 12c Enterprise Edition. Incorporando ainda as seguintes funcionalidades:

- Arquitetura *multitenant*, suporta múltiplas bases de dados totalmente funcional conectáveis dentro da mesma base de dados física;
- Carregamento de dados em colunas diretamente na memória do hardware para obter maior velocidade analítica;
- Oracle Real Application Clusters (RAC), tecnologia de cluster da Oracle que suporta o uso de failover automático num cluster de alta velocidade.

Para esta implementação existem três tipos de oferta, que se encontram descritos na Tabela 9.

Tabela 9 – Opções existentes para o Oracle Database Exadata Cloud Service

Características	Configuração		
	Quarter Rack	Half Rack	Full Rack
Número de nós	2	4	8
Número mínimo de cores de CPU	16	56	112
Número máximo de cores de CPU	68	136	272
Máximo de memória	496GB	992GB	1964GB
Capacidade total de disco flash em Raw	19,2TB	38,4TB	79,8TB
Capacidade total de disco em Raw	144TB	288TB	576TB
Capacidade total de disco utilizável	42TB	84TB	168TB

Em <http://docs.oracle.com/cloud/latest/exadatacs/> podem ser encontrados mais detalhes sobre esta solução, estes dados foram verificados em 2 de outubro de 2016.

O preço mensal deste serviço varia entre os €34.728 e os €486.192, foi verificado em 2 de outubro de 2016, em https://cloud.oracle.com/en_US/opc/database/exadata/pricing.

5.3.1.3 Oracle Database Exadata Express Cloud Service – Managed

Este serviço, está assente no motor de base de dados Oracle 12c Release 2, e é executado sobre um Oracle Exadata.

Permite uma experiência completa com bases de dados Oracle entregue como um serviço Cloud, permite a criação de bases de dados em poucos minutos, isto a um preço de nível de entrada acessível. Este serviço é ideal para pequenas e médias bases de dados, bem como para a área da programação.

Este modelo tem associado a ele três produtos denominados de:

- Exadata Express – X20: permite o uso de uma base de dados em execução num Exadata, com um tamanho máximo de 20GB e com uma transferência máxima de 120GB de e para a base de dados;

- Exadata Express – X50: permite o uso de uma base de dados em execução num Exadata, com um tamanho máximo de 50GB e com uma transferência máxima de 300GB;
- Exadata Express – 50IM: permite também o uso de uma base de dados em execução num Exadata, com um tamanho máximo de 50GB, com uma transferência máxima de 300GB e com 5GB de memória para uso de InMemory.

O preço mensal deste serviço varia entre os €152 e os €825, os valores foram verificados em 2 de outubro de 2016 em https://cloud.oracle.com/en_US/opc/database/exadata-express/pricing.

5.3.1.4 Oracle Database Schema Cloud Service - Managed

Esta modelo é para quem não precisa de uma instância de bases de dados completa, mas para ainda quem quer o poder de uma base de dados Oracle para os seus projetos.

Este modelo tem associado a ele três produtos denominados de:

- Database S5, cujo tamanho máximo da base de dados é de 5GB e permite a transferência de e para a base de dados de 30GB por mês;
- Database S20, o tamanho máximo da base de dados é de 20GB e permite a transferência de e para a base de dados de 120GB por mês;
- Database S50, o tamanho máximo da base de dados é de 50GB e permite a transferência de e para a base de dados de 300GB por mês.

O preço mensal deste serviço varia entre os €152 e os €1.736, estes valores são à data de 2 de outubro de 2016.

Todos os preços mencionados acima para estes DaaS, podem ser verificados com maior detalhe em: https://cloud.oracle.com/en_US/database/enterprise/pricing.

5.3.2 Azure SQL Database

É uma base de dados relacional alojada na nuvem Azure, que pode ser considerada como Software as a Service (SaaS), mas também como Plataforma as a Service (PaaS). Este DaaS é propriedade da Microsoft, sendo alojada e mantida pela mesma. Tem compatibilidade com a maioria das características do Microsoft SQL Server.

Este DaaS é pago conforme o seu uso, existindo a possibilidade de mudar os recursos existentes. É otimizado para reduzir ao mínimo os custos com a aquisição e gestão das bases de dados, com ele é reduzida a administração que é feita de forma contínua aos sistemas, pois não é necessária a gestão de máquinas virtuais, de sistemas operativos ou de RDBMS. Está garantido a atualização dos sistemas, a alta disponibilidade do sistema e os backups das bases

de dados. Usando este DaaS torna-se possível reduzir de forma drástica o número de recursos de TI necessários para a gestão de bases de dados.

Atualmente está disponível em várias camadas de serviço, os quais são faturados por hora, usando uma taxa fixa, baseado no nível de serviço e no desempenho escolhido. Os níveis de serviço Básico, Standard e Premium foram estabelecidos para oferecer um desempenho previsível com vários níveis de desempenho, para atender às necessidades de pico das aplicações. Os níveis de serviço, bem como os níveis de desempenho podem ser alterados para atender às necessidades aplicacionais. Na Tabela 10 e na Tabela 11, podemos verificar alguma da informação mais importante relativa aos níveis de serviço.

Tabela 10 – Características básicas dos níveis de serviço, para uma base de dados

Características	Níveis de Serviço (1 base de dados)										
	Básico	Standard				Premium					
		S0	S1	S2	S3	P1	P2	P4	P6	P11	P15
Máximo de DTUs	5	10	20	50	100	125	250	500	1000	1500	4000
Tamanho máximo	2GB	250GB				500GB			1TB		

Tabela 11– Características básicas dos níveis de serviço do modelo *elastic*

Níveis de Serviço	Características			
	Armazenamento max. por conjunto	Num. max. de BDs	Num. max. de eDTUs por BD	
Básico	100	10GB	200	5
	200	20GB	400	5
	400	39GB	400	5
	800	73GB	400	5
	1200	117GB	400	5
Standard	100	100GB	200	10
	200	200GB	400	20
	400	400GB	400	50
	800	800GB	400	100
	1200	1,2TB	400	100
Premium	125	250GB	50	125
	250	500GB	50	250
	500	750GB	50	500
	100	750GB	50	1000
	1500	750GB	50	1000

Nos seguintes endereços podem encontrar-se informações adicionais:

<https://azure.microsoft.com/pt-pt/documentation/articles/sql-database-service-tiers/>

<https://azure.microsoft.com/pt-pt/documentation/articles/sql-database-elastic-pool/>

Em <https://azure.microsoft.com/pt-pt/pricing/details/sql-database/> pode-se verificar os diversos custos existentes para este DaaS. Para uma única base de dados o custo pode ir desde os €4,2 mensais do serviço Básico, até aos €13.496 mensais do serviço Premium denominado de P15. Para usufruir de múltiplas bases de dados, modelo de serviço *elastic*, os custos variam desde os €126 mensais do serviço básico com o nível mais baixo, até aos €7.059 mensais do serviço Premium mais elevado.

Não estão incluídos nos preços acima, os custos referentes a movimentação dos dados entre os clientes e o Azure. As operações de upload dos dados não são normalmente taxadas, mas o download é, sendo que os primeiros 5GB mensais são gratuitos o restante é taxado mediante o consumo e mediante a zona do globo, os preços podem variar entre os €0,0734 e os €0,1349 por GB, o tarifário detalhado pode ser encontrado em: <https://azure.microsoft.com/pt-pt/pricing/details/data-transfers/>.

Para reforçar a segurança da informação que circula entre o Azure e os clientes, pode recorrer-se ao uso de uma VPN do Azure, no entanto, este serviço também é taxado pelo seu uso, e o seu custo é taxado por cada hora de utilização.

As opções existentes dos níveis de serviço, bem como o seu preço e características podem ser alterados por parte da Microsoft a qualquer instante.

A Microsoft garante um SLA (Service Level Agreement) de 99,9% do tempo em que os clientes terão conectividade com a base de dados no Azure. Não está incluído neste tempo a indisponibilidade de serviços do operador de comunicações do cliente. Em caso de incumprimento do SLA a Microsoft fará um crédito ao cliente de 10% caso a indisponibilidade seja inferior a 99,99% do tempo de atividade mensal e de 25% caso esse valor seja inferior a 99%.

Algumas limitações deste DaaS são:

- não suporta autenticação Windows;
- não possui SQL agente, assim sendo não é possível realizar tarefas agendadas;
- não suporta o uso da instrução USE, ou seja, para alternar o contexto entre bases de dados deverá ser realizada uma nova conexão;
- não suporta o uso de FILESTREAM;
- não permite o envio de emails, pois não tem a funcionalidade de Database Mail;
- não permite o uso de tabelas temporárias globais;
- não suporta o uso de linked servers;
- não é possível realizar um shutdown.

Em <https://azure.microsoft.com/en-us/documentation/articles/sql-database-transact-sql-information/#features-not-supported-in-sql-database> pode ser encontrada uma listagem com todas as limitações existentes neste DaaS, relativamente a um sistema tradicional, ou a uma

máquina virtual em Azure contendo um SQL Server. Esta listagem pode ser atualizada, pois este tipo de serviço sofre um processo de constante evolução.

Pode-se usar REST API e PowerShell Cmdlets para automatizar e gerir operações administrativas para milhares de bases de dados.

Informação adicional relativa a este DaaS encontra-se em <https://azure.microsoft.com/pt-pt/documentation/articles/data-management-azure-sql-database-and-sql-server-iaas/>.

No caso de não ser viável uma implementação deste DaaS, o Azure disponibiliza máquinas virtuais com vários tipos de pré-instalações de SQL Server, o valor mais barato deste tipo de implementação é de €11,29 mensais, o custo com o armazenamento dos dados não está contabilizado no valor anterior. De salientar, que para o custo mencionado, o valor de ocupação médio da CPU ronda os 90%, esta máquina virtual só contém sistema operativo e motor de base dados SQL, não tendo nenhuma base de dados em uso.

5.3.3 Google Cloud SQL

Google Cloud SQL é um serviço DaaS onde é fácil configurar, manter e gerir as bases de dados MySQL. A Google Cloud SQL Second Generation oferece alto desempenho, escalabilidade e conveniência.

Alojado no Google Cloud Platform, o Cloud SQL fornece uma infraestruturas de base de dados para aplicações em execução em qualquer lugar. Até à data a Google tem três Datacenters nos Estados Unidos da América, um na Europa (St. Ghislain, na Bélgica) e outro na Ásia (Changua County, em Taiwan), está previsto até ao final de 2016 estar concluído o de Tokyo.

O Google Cloud SQL automatiza todos os backups, replicações e atualizações, garantindo ao mesmo tempo uma disponibilidade de serviço superior a 99,95%, em qualquer lugar do mundo.

Em <https://cloud.google.com/sql/docs/features> encontram-se as funcionalidades deste DaaS, de entre elas salientam-se as seguintes:

- As instâncias de MySQL podem estar em execução em datacenters localizados nos Estados Unidos, na Ásia ou na União Europeia;
- Criação e gestão das instâncias de MySQL usando a uma plataforma denominada de Google Cloud Platform Console;
- Permite a replicação de dados entre várias zonas com um failover automático;
- Automatização de backups;
- Importação e exportação de bases de dados usando o mysqldump , ou através do uso de ficheiros CSV.

Tal como os outros DaaS, existem ainda funcionalidades e instruções SQL que ainda não se encontram implementadas, estas são dadas não suportadas, elas podem também ser verificadas com maior detalhe em <https://cloud.google.com/sql/docs/features>, de entre elas, salientam-se as seguintes:

- Funções criadas pelos utilizadores;
- Ausência do Performance Schema que permite realizar uma monitorização da execução do MySQL a baixo nível.

As características básicas dos modelos de máquinas existentes para o Google Cloud SQL Second Generation encontram-se na Tabela 12.

Tabela 12 – Modelos existentes para o Google Cloud SQL Second Generation

Tipo	CPU's Virtuais	Memória
db-f1-micro	Partilhada	600MB
db-g1-small	Partilhada	1,7GB
db-n1-standard-1	1	3,75GB
db-n1-standard-2	2	7,5GB
db-n1-standard-4	4	15GB
db-n1-standard-8	8	30GB
db-n1-standard-16	16	60GB
db-n1-highmem-2	2	13GB
db-n1-highmem-4	4	26GB
db-n1-highmem-8	8	52GB
db-n1-highmem-16	16	104GB

Na Tabela 13, encontram-se os modelos de implementação e algumas das características básicas do Google Cloud SQL First Generation.

Tabela 13 – Modelos existentes para o Google Cloud SQL First Generation

Nível	Memória	Armazenamento incluído
D0	0,125GB	0.5GB
D1	0,5GB	1GB
D2	1GB	2GB
D4	2GB	5GB
D8	4GB	10GB
D16	8GB	10GB
D32	16GB	10GB

Os dados existentes na Tabela 12 e Tabela 13, foram retirados em 6 de outubro de 2016 de <https://cloud.google.com/sql/pricing>.

Tanto no Google Cloud SQL First Generation como no Second Generation o valor do serviço é dado pela soma de vários fatores, sendo estes o modelo a implementar, o armazenamento gasto (discos SSD têm valores mais elevados) e a rede usada (apenas é taxado o tráfego oriundo da Google Cloud, o de envio dos dados não é taxado).

Em <https://cloud.google.com/products/calculator/> encontra-se uma calculadora que permite determinar os custos deste DaaS, os valores apresentados encontram-se em US dólares. É necessário ter em atenção a advertência que se encontra no final da página, onde é referido que os valores são estimados, podendo estes no final ser maiores ou menores, e que estes devem apenas serem usados como referência. Recomenda-se o uso desta ferramenta para poder ser realizada uma estimativa dos custos, pois estes dependem de muitos fatores.

Em <https://cloud.google.com/sql/docs/concepts> pode-se encontrar toda a documentação necessária para um melhor entendimento deste DaaS.

5.3.4 Amazon RDS

O Amazon RDS (Relational Database Service), é um serviço de base de dados relacional, que permite usar bases de dados MySQL, Oracle, SQL Server, PostgreSQL, Amazon Aurora ou MariaDB.

O Amazon RDS executa as tarefas rotineiras que existem habitualmente nas bases de dados, como aplicação de atualizações (o utilizador pode escolher quando as atualizações serão realizadas), o backup, a recuperação e a deteção de falhas.

O backup automatizado do Amazon RDS está ativado por defeito e permite a recuperação para qualquer altura no tempo da instância de base de dados. O Amazon RDS faz o backup da base de dados e dos logs de transação, armazenando-os pelo período de retenção definido pelo utilizador. Isso permite o restauro da base de dados na posição de qualquer segundo durante o período de retenção, até os últimos cinco minutos. O período de retenção pode ser configurado até trinta e cinco dias.

Existe a possibilidade de alterar os recursos de computação e memória, para mais ou para menos, até ao máximo de 32 vCPUs e 244 GB de RAM. As operações de escalabilidade de computação normalmente são concluídas nalguns minutos.

O Amazon RDS pode enviar notificações de eventos, por meio de e-mail ou mensagem de texto SMS usando o Amazon SNS³.

³ O Amazon SNS (Simple Notification Service) é um serviço de notificações por push rápido, flexível e totalmente gerido, que permite enviar mensagens individuais ou distribuir mensagens para um grande número de destinatários.

É possível neste serviço o uso de snapshots das bases de dados, estes não passam de meros backups da base de dados, mas que foram criados explicitamente pelos utilizadores, sendo armazenados noutra serviço Amazon, neste caso só é taxado o armazenamento usado. Os snapshots são backups permanentes, só são eliminados através de ação explícita do utilizador. Podem ser usados para criar novas bases de dados.

Este serviço também permite a encriptação das bases de dados através do uso de chaves distribuídas por outro serviço da Amazon, todos os dados são encriptados, bem como os backups, os snapshots e as réplicas de leitura das bases de dados, caso essas existam.

Existem à data de outubro de 2016, doze zonas onde podem ser albergadas as bases de dados, sendo quatro delas nos Estados Unidos da América, duas na Europa, uma na América do Sul, uma na Austrália e quatro na Ásia.

Na Tabela 14, pode-se verificar os diversos tipos de instâncias passíveis de serem criadas no Amazon RDS, os dados foram revistos em 9 de outubro.

Tabela 14 – Tipos de instâncias passíveis de criação no Amazon RDS

Tipo de instância		vCPU	Memória (GB)	Desempenho de rede
Standard	db.m4.large	2	8	Moderado
	db.m4.xlarge	4	16	Alto
	db.m4.2xlarge	8	32	Alto
	db.m4.4xlarge	16	64	Alto
	db.m4.10xlarge	40	160	10 Gigabit
	db.m3.medium	1	3,75	Moderado
	db.m3.large	2	7,5	Moderado
	db.m3.xlarge	4	15	Moderado
	db.m3.2xlarge	8	30	Alto
Otimizadas para memória	db.r3.large	2	15	Moderado
	db.r3.xlarge	4	30,5	Moderado
	db.r3.2xlarge	8	61	Alto
	db.r3.4xlarge	16	122	Alto
	db.r3.8xlarge	32	244	10 Gigabit
Microinstâncias	db.t2.micro	1	1	Baixo
	db.t2.small	1	2	Baixo
	db.t2.medium	2	4	Moderado
	db.t2.large	2	8	Moderado

Mais detalhes podem ser encontrados em <https://aws.amazon.com/pt/rds/details/>.

Em <http://calculator.s3.amazonaws.com/index.html> encontra-se um simulador de custos para os diversos serviços existentes na Amazon AWS, para este caso deve ser escolhida a opção Amazon RDS.

Descrevem-se nos próximos subcapítulos os seis tipos de base de dados suportados no Amazon RDS:

- Amazon Aurora;
- Amazon RDS para MySQL;
- Amazon RDS para MariaDB;
- Amazon RDS para PostgreSQL;
- Amazon RDS para Oracle Database;
- Amazon RDS para SQL Server.

5.3.4.1 Amazon Aurora

Segundo a Amazon, o Amazon Aurora é uma base de dados relacional, compatível com o MySQL, que combina a velocidade e disponibilidade das bases de dados comerciais com a simplicidade e a economia das bases de dados de código aberto. Oferece um desempenho até cinco vezes maior que o MySQL com a segurança, disponibilidade e confiabilidade de um base de dados comercial a um décimo do custo.

O Amazon Aurora é totalmente compatível com o MySQL 5.6, isto significa que o código, as aplicações, os drivers e as ferramentas que normalmente se usam para as bases de dados em MySQL possam ser usadas com o Amazon Aurora com pouca ou nenhuma alteração. A migração das bases de dados MySQL existentes é facilitada pelo uso de ferramentas padrão de importação e exportação do MySQL ou pelo mecanismo de replicação do MySQL.

Este serviço foi criado para oferecer uma disponibilidade superior a 99,99%. A recuperação de falhas de armazenamento físico é transparente e o failover de instância normalmente é executado em menos de 30 segundos. O armazenamento do Amazon Aurora é tolerante a falhas e tem correção automática. Seis cópias de seus dados são replicadas em três zonas de disponibilidade com backup contínuo para o Amazon S3.

Este serviço pode ser usado e taxado de duas maneiras diferentes, por intermédio de uma instância on-demand, onde se é taxado à hora, ou por meio de uma instância reservada por períodos de um e três anos.

O uso de uma instância reservada permite três modelos de pagamento, o primeiro trata-se de um pagamento mensal, o segundo tem uma taxa de adiantamento e o restante em prestações mensais, enquanto que o terceiro é totalmente pago adiantadamente.

O uso de instância reservada normalmente é mais barato do que o uso de uma instância on-demand, quando o uso da base de dados é permanentemente em uso e com carga.

Deve-se ter em atenção que o tráfego de dados e o armazenamento de backups também é taxado. Mais detalhes, relativos à taxação deste serviço, podem ser encontrados em <https://aws.amazon.com/pt/rds/aurora/pricing/>.

5.3.4.2 Amazon RDS para MySQL

O Amazon RDS para MySQL foi desenhado para o uso de programadores ou de empresas que exigem todos os recursos e as capacidades de uma base de dados MySQL.

Tal como no Amazon Aurora, o modelo de taxa o   data de 9 de outubro   id ntico, ou seja, existe a hip tese do uso de uma inst ncia on-demand, ou de inst ncia reservada.

Mais detalhes podem ser obtidos em <https://aws.amazon.com/pt/rds/mysql/>, um maior detalhe relativo   taxa o encontra-se em <https://aws.amazon.com/pt/rds/mysql/pricing/>.

5.3.4.3 Amazon RDS para MariaDB

Este servi o permite o uso de bases de dados MariaDB. A partir de 1 de junho de 2016 suporta o uso da vers o 10.1 do MariaDB.

Tal como nos dois tipos de servi o anteriores, o Amazon Aurora e o Amazon RDS para MySQL, o modelo de taxa o   data de 9 de outubro   id ntico, ou seja, existe a hip tese do uso de uma inst ncia on-demand, ou de inst ncia reservada.

Mais detalhes podem ser obtidos em <https://aws.amazon.com/pt/rds/mariadb/>, um maior detalhe relativo   taxa o encontra-se em <https://aws.amazon.com/pt/rds/mariadb/pricing/>.

5.3.4.4 Amazon RDS para PostgreSQL

Este servi o permite o uso de bases de dados PostgreSQL, suportando   data de 9 de outubro as seguintes vers es:

- PostgreSQL 9.5.2
- PostgreSQL 9.4.7
- PostgreSQL 9.3.12

Mais detalhes est o dispon veis em <https://aws.amazon.com/pt/rds/postgresql/>, entra-se em <https://aws.amazon.com/pt/rds/postgresql/pricing/> mais detalhes referentes   taxa o deste servi o.

5.3.4.5 Amazon RDS para Oracle Database

Este servi o permite o uso de bases de dados Oracle, suporta as seguintes edi es:

- Standard Edition One (com licenciamento inclu do)
- Standard Edition One (com licenciamento do cliente)
- Standard Edition Two (com licenciamento inclu do)
- Standard Edition Two (com licenciamento do cliente)
- Standard Edition (com licenciamento do cliente)
- Enterprise Edition (com licenciamento do cliente)

Em <https://aws.amazon.com/pt/rds/oracle/> podem ser encontrados mais detalhes sobre este tipo de serviço, em <https://aws.amazon.com/pt/rds/oracle/pricing/> podem ser encontrados mais detalhes relativos à taxa  o deste servi o.

5.3.4.6 Amazon RDS para SQL Server

Este servi o permite o uso de bases de dados em SQL Server, as vers es suportadas s o:

- SQL Server 2008 R2
- SQL Server 2012
- SQL Server 2014

Este servi o, tal como o Amazon RDS para Oracle Database permite trazer o licenciamento do cliente para a Cloud, s  sendo suportado para as vers es Standard e Enterprise do SQL Server, no entanto existe a op  o de o licenciamento estar inclu do, estando dispon veis para as vers es Express, Web, Standard e Enterprise Edition do SQL Server.

Para ‘transportar’ o licenciamento existente de SQL Server para a Amazon Cloud SQL   necess rio cumprir os requisitos de licenciamento da Microsoft, est  dispon vel em <https://aws.amazon.com/pt/windows/resources/licensemobility/sql/> um breve resumo de como calcular o n mero de licen as necess rias.

Em <https://aws.amazon.com/pt/rds/sqlserver/> podem ser verificados mais detalhes sobre este servi o, em <https://aws.amazon.com/pt/rds/sqlserver/pricing/> podem ser encontrados mais detalhes sobre a taxa  o.

6 Avaliação Experimental

Neste capítulo são descritos os dois testes que se devem realizar para a concretização de uma migração de uma base de dados para um DaaS. Também nele se encontra um caso de uso de uma migração realizada para o Azure SQL, usando uma base de dados de teste disponibilizada pela Microsoft.

6.1 Métodos

Para realizar uma análise à performance dos sistemas, usam-se testes denominados de benchmarks, estes são índices ou indicadores de desempenho, baterias de testes que executam determinadas operações, com o objetivo de atribuir uma pontuação ao processador, servidor ou qualquer outro elemento em análise.

Em (Ferreira, 2015), é referido que na área dos processadores, a organização mais conhecida que publica benchmarks é a Standard Performance Evaluation Corporation (SPEC) (<http://www.spec.org>). Para a análise da performance de discos e de sistemas de armazenamento usa-se como principal atributo de comparação entre estes componentes o Input/Output Operations Per Second (IOPS) que mede o número de operações possíveis, por segundo, de realizar com um disco. Para além da fiabilidade dos discos, espaço de armazenamento, tipo de disco (SATA, SCSI, SAS, SSD), é necessário ter em atenção o seu custo. Na Internet em UserBenchmark (<http://hdd.userbenchmark.com/>) e em tom'sHARDARE (<http://www.tomshardware.com/charts/hard-drives,3.html>) podemos encontrar testes de benchmark efetuados a diversos modelos de discos. Existe também um utilitário bastante útil para medir a capacidade de leitura e escrita de um disco, podendo à posteriori realizar uma comparação dos dados num dos endereços acima mencionados, o seu nome é CrystalDiskMark e está disponível para download em <http://crystalmark.info/?lang=en>.

No caso de bases de dados é importante ter discos rápidos, para além da capacidade de processamento do sistema.

6.2 Metodologias a Aplicar

Com o intuito de analisar a qualidade da solução implementada iremos realizar um conjunto de testes.

Iremos utilizar como grandezas de medida o tempo e o TCO (*Total Cost of Ownership*) para um período de 3 anos, considerando também que nesse período nenhuma das soluções propostas sofrerá alterações, também não será contemplada a hipótese de migração entre DBMS, ou seja, se a base de dados for em MySQL, a migração para um DaaS seria para MySQL ou vice-versa.

6.2.1 Testes a Aplicar

Os testes a aplicar para a verificação de uma migração para a Cloud são:

- Teste de Performance;
- Análise de Custos.

6.2.1.1 Teste de Performance

Para este teste iremos usar como medida de comparação o tempo de execução de uma query pré-determinada, valor medido em milissegundos.

Antes deste teste vão ser retirados dados estatísticos relativos à performance das bases de dados, como por exemplo a memória utilizada, o consumo de CPU ou o número médio de IOPS. Estes valores vão servir para tentar equivaler os DBMS's a nível de características técnicas.

A query será escolhida de entre as que ocupam um maior tempo de processamento, será obtida através de ferramentas especializadas para o efeito e será obtida estando a base de dados em carga normal. A query a escolher deverá ser de preferência um select, pois esta instrução não tem influência nos dados existentes na base de dados.

Os testes serão realizados com a base de dados sem carga, ou seja, a única instrução SQL a ser executada seria a para obter a performance. O ideal será a realização de testes com a base de dados em carga, mas é altamente improvável que a carga seja a mesma de todas as vezes que seja efetuada a query.

Os testes têm de ser executados com as bases de dados contendo os mesmos dados, para o efeito o primeiro teste iria ser executado com uma base de dados instalada no cliente, sem carga, antes do início seria realizado um backup, para aquando do segundo teste os dados e a estrutura da base de dados serem os mesmos. Um dos testes irá ser executado num RDBMS tradicional e o outro num DaaS.

Para este efeito utilizaremos um método denominado de T-Test, visto os dados serem independentes, pois a plataforma onde se encontram será diferente. O que se irá comparar é o tempo de resposta, as queries serão executadas 20 vezes e posteriormente será calculada a sua média, quanto menor for o valor melhor, em caso de empate o critério de desempate será a variância. O T-Test é usado quando existe apenas um pequeno grupo de amostras a avaliar.

Para a realização destes testes poderemos utilizar ferramentas como: o Performance Monitor do Windows, o SQL Server Management Studio, o MySQL Workbench, o Toad DBA Suite para Oracle, o IBM Data Server Manager ou o SQL Monitor da Red-Gate.

6.2.1.2 Análise de Custos

Como existem vários fornecedores de DaaS, será de realizar um teste comparando os custos entre os principais *players* de mercado em que o RDBMS seja o mesmo (e com características similares), comparando o valor com o de uma solução de tradicional, para a comparação usar-se-á o TCO. Quanto menor for o custo para o cliente melhor.

6.2.2 Exemplo

Apresentamos neste subcapítulo um exemplo de como aplicar os métodos acima apresentados, para o efeito, usamos um PC antigo com as seguintes características:

Motherboard: Asus P5K-VM

Processador: Intel Pentium E2180 (2 cores, 2 GHz) a

Memória: 4 GB DDR2 a 800 MHz

Discos:

Sistema Operativo e Aplicações: SAMSUNG SSD 850 EVO 250GB

Bases de Dados: WDC WD10EACS-00D6B1 (SATA, 1 TB)

Windows Server 2012 R2

Microsoft SQL Server 2016

Para a realização desde teste usamos uma base de dados da Microsoft, disponibilizada para efeitos de teste, de seu nome WideWorldImporters-Standard, o nome do ficheiro do qual foi realizado download e foi utilizado para importar a base de dados é WideWorldImporters-Standard.bacpac, este encontra-se disponível em:

<https://github.com/Microsoft/sql-server-samples/releases/tag/wide-world-importers-v1.0>

Depois de importada a base de dados ocupa 645MB. Existem 4 *schemas* para as tabelas existentes na base de dados, os seus nomes são:

- Application
- Purchasing
- Sales
- Warehouse

Devido à quantidade de tabelas existentes, o modelo E-R (Entidade – Relação) encontra-se dividido, para melhor compreensão e visualização pelas Figura 5, Figura 6, Figura 7 e Figura 8 , conforme o *schema* a que estas pertencem.

O modelo E-R de cada um dos schemas foi importado diretamente da base de dados tendo sido usado para o efeito o Visual Paradigm, da Visual Paradigm International Ltd.

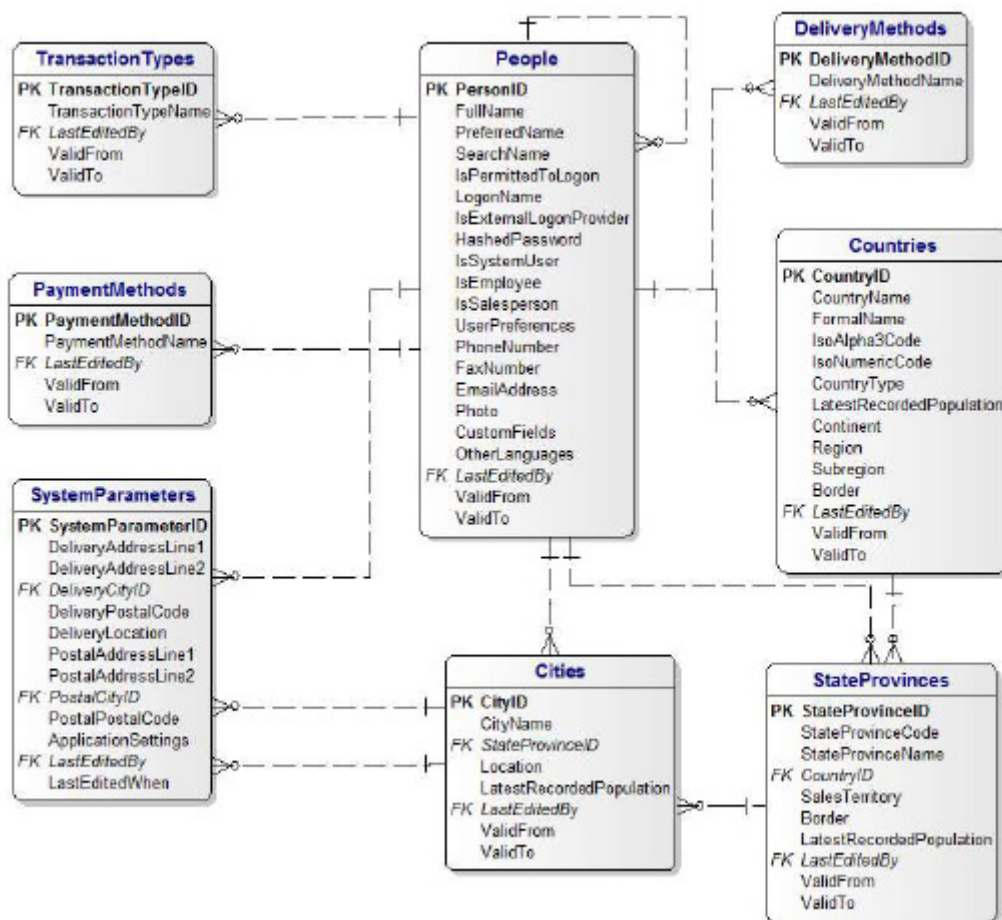


Figura 5 – Modelo E-R pertencente ao *schema* Application

Neste schema tem oito tabelas, de entre as quais se salienta a tabela People, que para além de relacionamentos com as outras tabelas existentes, tem um relacionamento com ela própria.

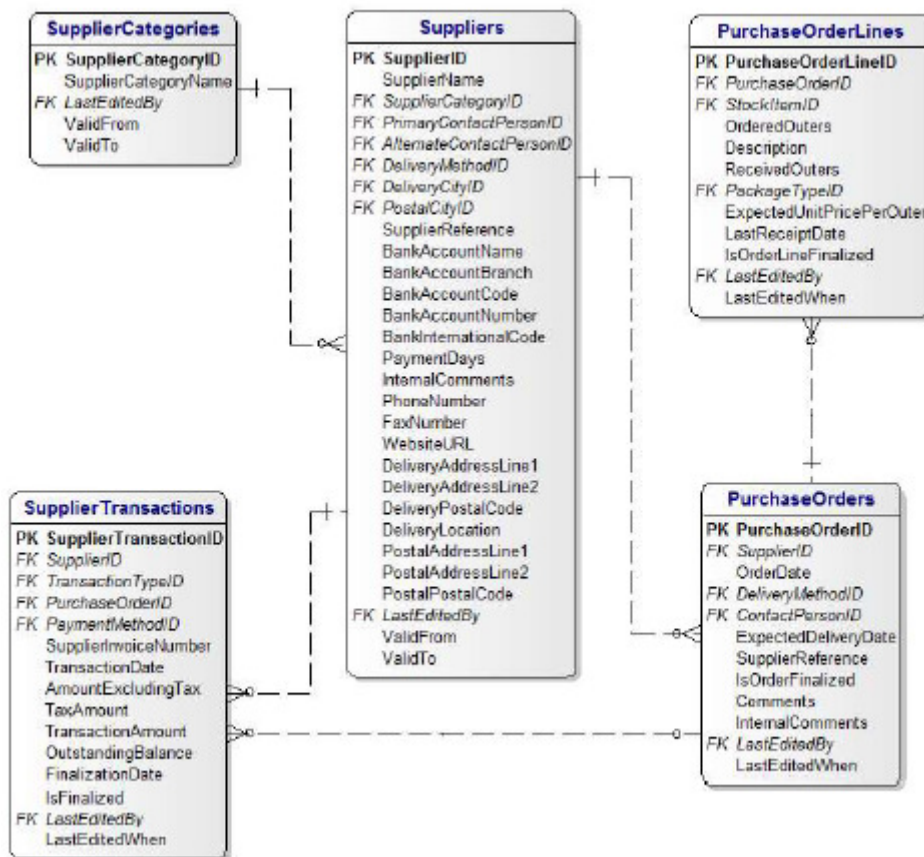


Figura 6 – Modelo E-R pertencente ao *schema* Purchasing

Este schema tem cinco tabelas, sendo de entre todos os existentes na base de dados o que possui um modelo E-R mais simples.

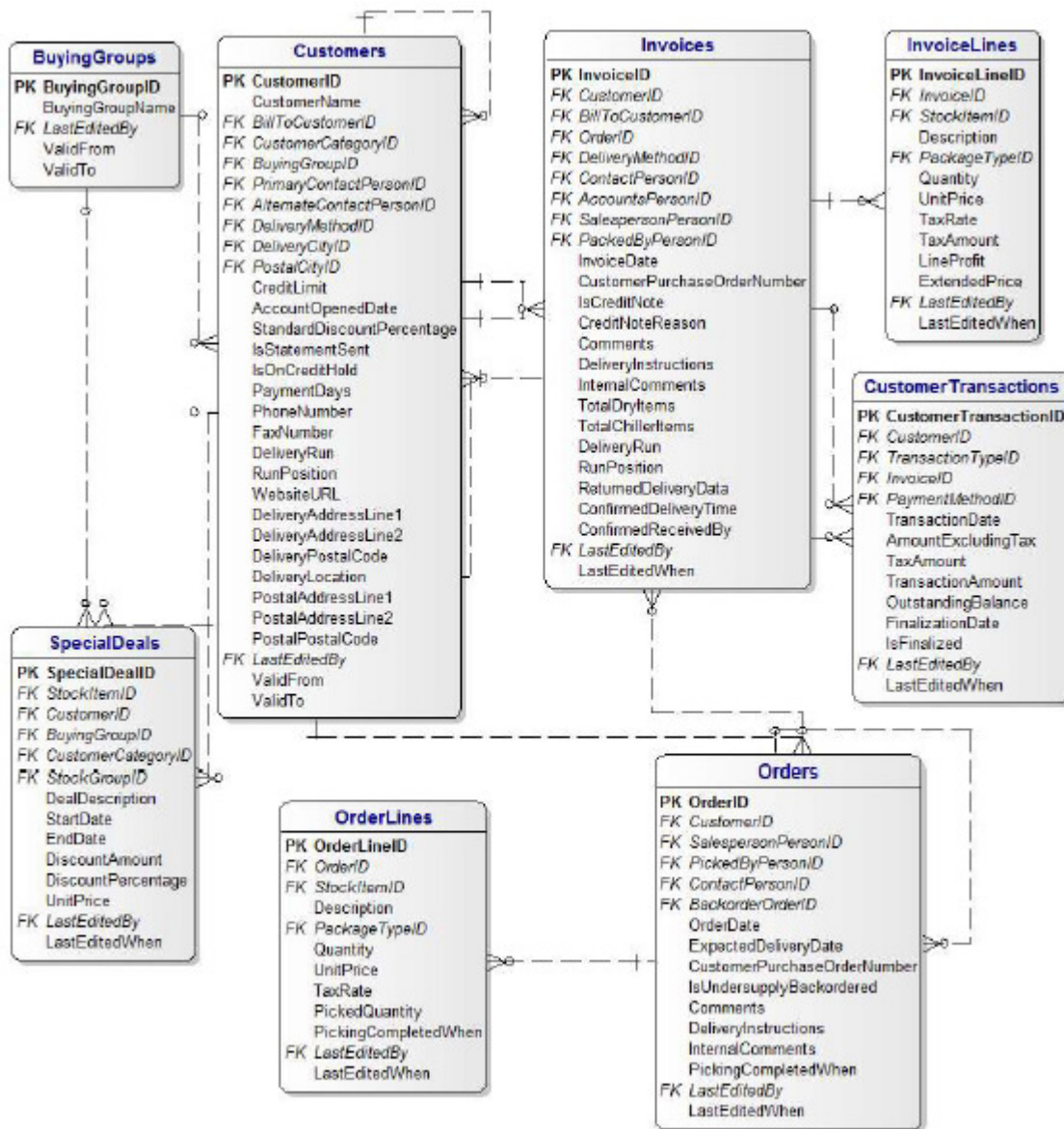


Figura 7 – Modelo E-R pertencente ao *schema* Sales

Este schema tem oito tabelas, sendo de entre os outros existentes na base de dados o mais complexo, a tabela Customers tem ou pode ter relacionamentos entre ela própria. Este schema é referente à área das vendas.

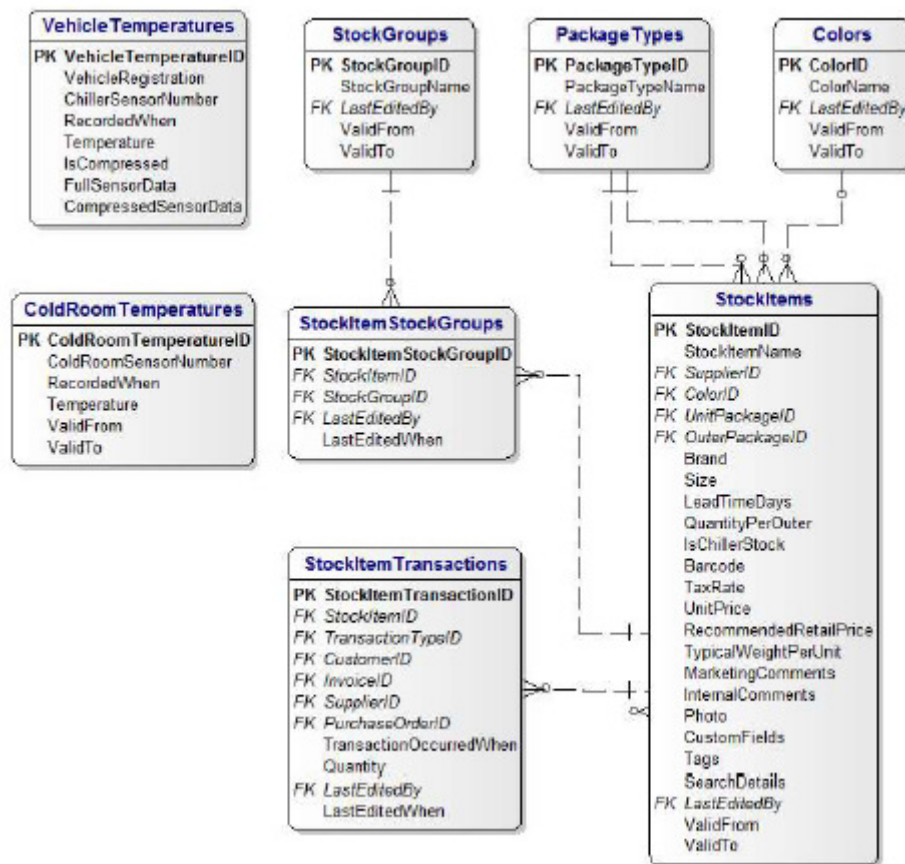


Figura 8 – Modelo E-R pertencente ao *schema* Warehouse

Este schema também tem oito tabelas e refere-se às existências em armazém, servindo assim para a gestão do inventário.

Para verificar a compatibilidade da base de dados com o Azure, foi executado o procedimento existente no site: <https://azure.microsoft.com/en-us/documentation/articles/sql-database-cloud-migrate-determine-compatibility-ssms/>.

O procedimento pode ser verificado no anexo 85D.

Se não existirem erros, após a execução do procedimento, pode concluir-se que a base de dados é compatível e podemos realizar a migração. Neste caso em concreto, não foram encontrados erros.

Se existirem erros e quisermos realizar na mesma a sua migração é necessário proceder à sua correção, ao lado do nome do teste aparece na coluna 'Result' a mensagem Error ao premir a mensagem aparecerá um ecrã com uma descrição do erro encontrado.

Para gerar carga no sistema foi usado um programa de seu nome MultithreadedOrderInsert, este foi obtido no site abaixo indicado. O ficheiro a realizar download é denominado de

workload-drivers.zip, depois de descompactado obtemos duas pastas, o ficheiro a usar encontra-se dentro da pasta order-insert.

<https://github.com/Microsoft/sql-server-samples/releases/tag/wide-world-importers-v1.0>

O teste foi executado usando os parâmetros que podem ser verificados na Figura 9.

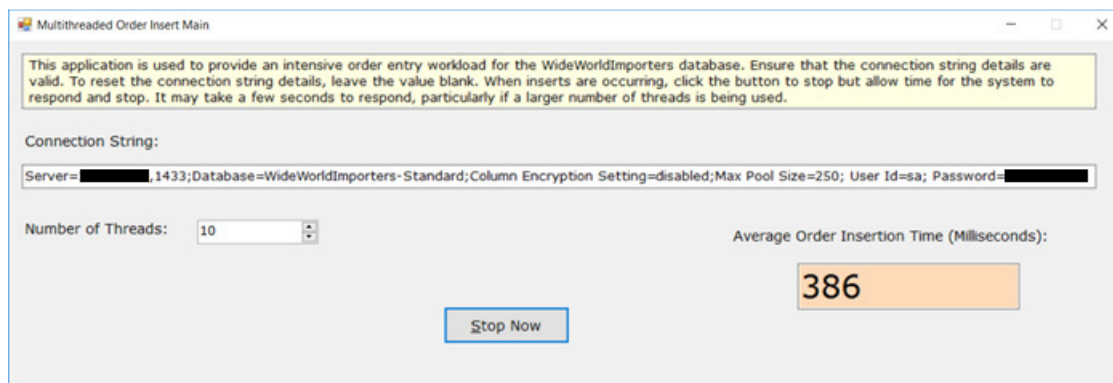


Figura 9 – Parametrização usada para testes

O parâmetro ‘Max Pool Size’ refere-se ao número máximo de conexões permitidas dentro de uma pool de conexão. Quando o número máximo de ligações é atingido, deixa de ser possível estabelecer ligações adicionais, ficando estas numa fila à espera que alguma das conexões existentes termine.

O parâmetro ‘Number of Threads’ refere-se ao número de threads que são executadas simultaneamente pelo programa, estas threads contêm as instruções SQL que são executadas na base de dados.

A caixa de texto abaixo do texto ‘Average Order Insertion Time (Milliseconds)’ indica-nos o tempo médio que cada thread (Order) demora a ser executada. No código da aplicação é possível verificar que a fórmula é dada pelo tempo de execução das threads em milissegundos a dividir pelo número de threads executadas.

Após o início do teste podemos verificar que o uso da CPU atinge os 93% e o da memória os 87%, sendo que o processo da base de dados ocupa 70,1% e cerca de 1,6 GB de memória, o mesmo pode ser comprovado na Figura 10.

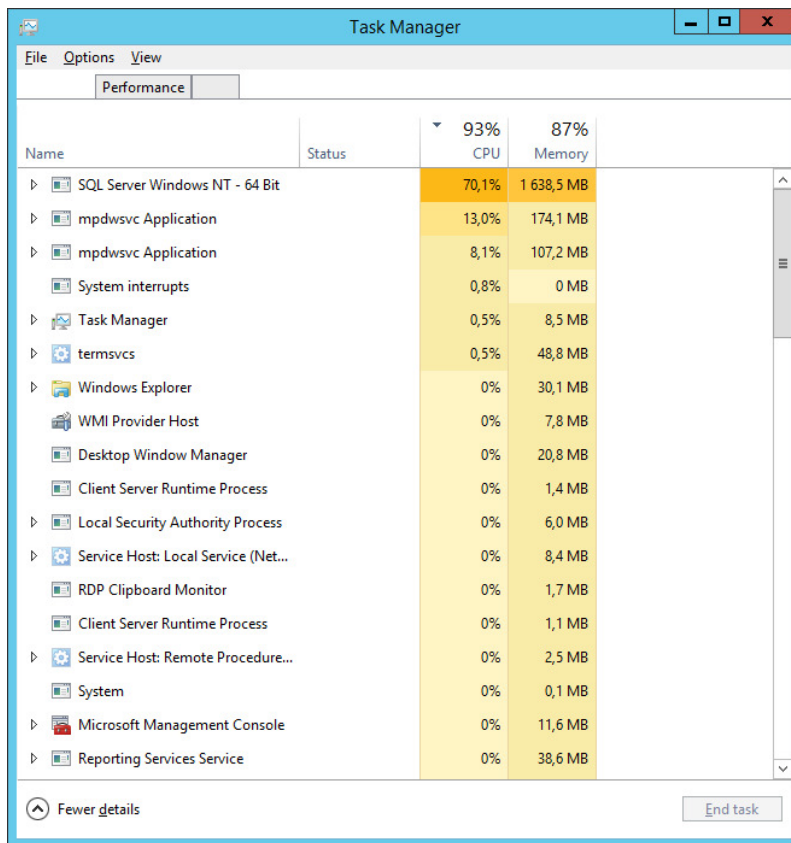


Figura 10 – Task Manager

Na Figura 11 é apresentado um printscreen retirado do sistema acima descrito, onde temos a verde a memória e a vermelho a CPU, esta análise decorreu por um período de 2 horas, num caso real, teria de ser realizada por um período mais alargado.

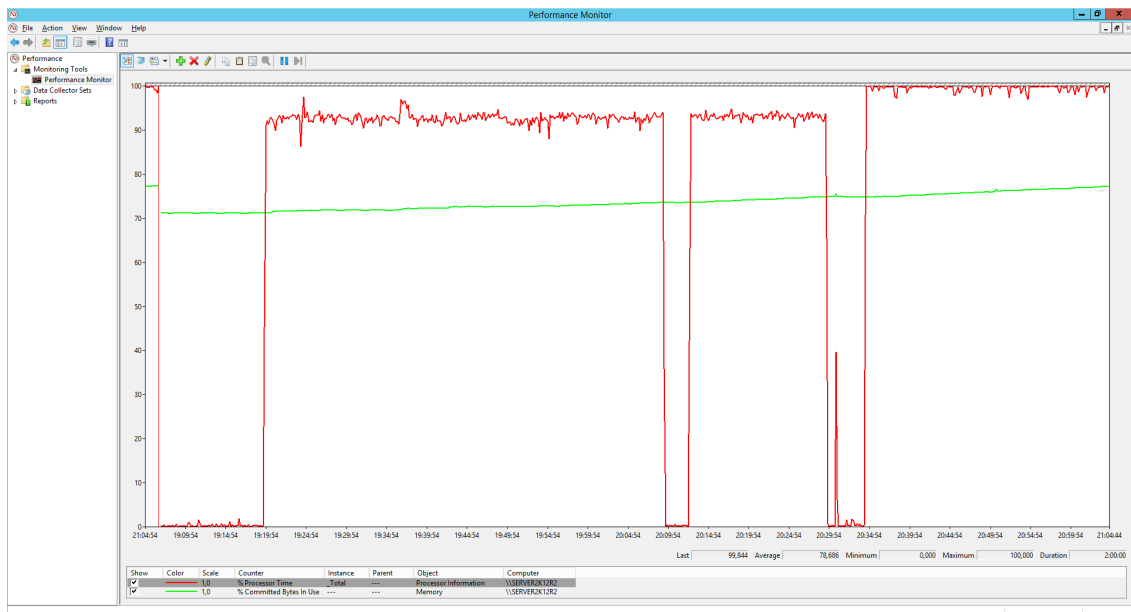


Figura 11 – Performance Monitor

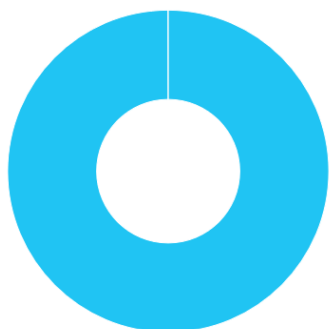
Durante os primeiros minutos, o simulador de carga está desligado, podemos verificar no gráfico que o gasto de CPU era quase nulo, mas nos momentos em que este se encontrava em execução o consumo de CPU ultrapassava os 90%.

Também foi retirado do sistema, enquanto o simulador de carga estava a ser executado, uma monitorização, esta foi retirada executando um script powershell de seu nome sql-perfmon.ps1, retirado do site: <http://dtucalculator.azurewebsites.net/>, também consta do site a advertência:

“Este site não é de modo algum afiliado com, nem foi autorizado, patrocinado, ou de outra forma aprovada pela Microsoft Corporation”

No entanto no próprio site da Microsoft é recomendado o uso desta ferramenta, para obter uma estimativa do nível de desempenho e do nível de serviço que a base de dados pode necessitar no Azure, esta afirmação pode ser comprovada em: <https://azure.microsoft.com/en-us/documentation/articles/sql-database-service-tiers/>. Esta monitorização tem a duração de uma hora, após a mesma é gerado um ficheiro Excel com a extensão CSV, o qual à posteriori foi submetido para o site do qual o script foi retirado, indicando também neste o número de cores da CPU. O resultado da análise dos dados submetida é uma recomendação do nível de serviço a adquirir, esta pode ser verificada nas figuras (Figura 12, Figura 13, Figura 14, Figura 15 e Figura 16) a seguir apresentadas.

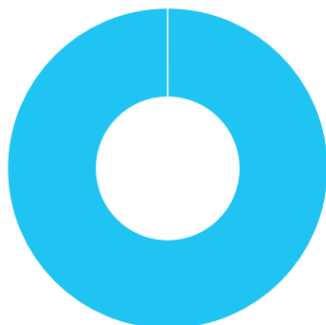
Service Tier/Performance Level



Based on your database utilization, we recommend you migrate your SQL Server workload to **Premium - P2**. This Service Tier/Performance Level should cover approximately **100.00 %** of your utilization.

Figura 12 – Gráfico do nível de serviço / nível de desempenho

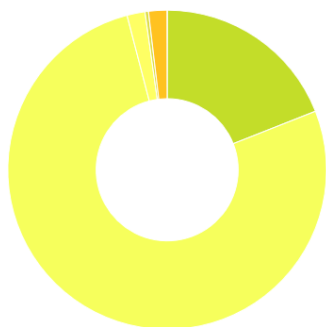
Service Tier/Performance Level for CPU



Based solely on **CPU** utilization, we recommend you migrate your SQL Server workload to **Premium - P2**. This Service Tier/Performance Level should cover approximately **100.00 %** of your CPU utilization.

Figura 13 – Gráfico do nível de serviço / nível de desempenho da CPU

Service Tier/Performance Level for Iops



Based solely on **Iops** utilization, we recommend you migrate your SQL Server workload to **Standard - S1**. This Service Tier/Performance Level should cover approximately **95.94 %** of your Iops utilization.

NOTE: There is approximately **4.06 %** of your workload that falls into a higher Service Tier/Performance Level. After migrating your database to Azure, you should evaluate your database's performance using the guidance mentioned in the [🔗](#) information section above.

Figura 14 – Gráfico do nível de serviço / nível de desempenho para os Iops

Service Tier/Performance Level for Log



Based solely on **Log** utilization, we recommend you migrate your SQL Server workload to **Standard - S0**. This Service Tier/Performance Level should cover approximately **100.00 %** of your Log utilization.

Figura 15 – Gráfico do nível de serviço / nível de desempenho para o transaction log

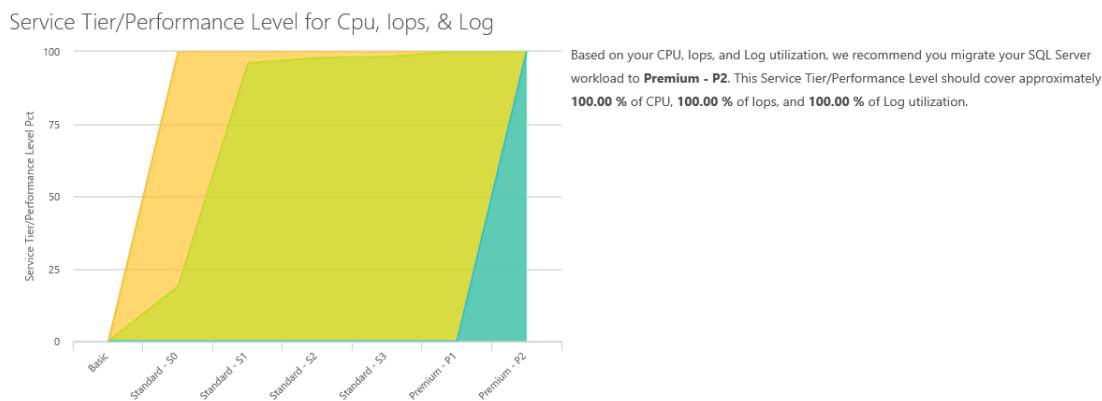


Figura 16 – Gráfico do nível de serviço / nível de desempenho para CPU, IOPS e Transaction Log

Pela descrição do gráfico da Figura 16, podemos constatar que nos é aconselhado a aquisição de um nível de serviço denominado de 'Premium – P2', no entanto no gráfico que nos apresenta o nível de serviço para os IOPS, é-nos indicado que o nível de serviço 'Standard – S1' é suficiente para satisfazer quase 100% da utilização de IOPS.

A base de dados tem neste momento cerca de 1870MB de tamanho, assim sendo uma mudança para um nível de serviço denominado de Básico não é viável, pois este tem uma capacidade máxima para a base de 2GB e esta já está no limiar do suportado.

Os primeiros passos a realizar, para a migração são, a criação de uma conta Windows Live, entrada no portal do Azure e proceder à verificação da conta, usando para o efeito um número de telemóvel para o qual irá ser enviada uma mensagem contendo esta um código de confirmação e por último proceder à inserção dos dados de um cartão de crédito, que será usado para a cobrança dos serviços usados.

Em seguida deverá ser criada uma nova subscrição, sugere-se a "Pago pelo Uso", após a criação da subscrição, deve-se proceder à criação de um grupo de recursos, aqui são pedimos a introdução do nome do grupo de recursos, a subscrição ao qual estará associado e a localização onde este grupo será criado. Após estes passos deve proceder-se à criação de um servidor SQL lógico (não é um servidor físico, nem virtual, trata-se de um apontador). Estes passos estão descritos em maior detalhe em: <https://azure.microsoft.com/pt-pt/documentation/articles/sql-database-get-started/>.

Foram efetuados alguns testes de migração desta base de dados para o Azure. As instruções para a migração da base de dados podem ser constatadas no anexo D.

Foram realizados 6 testes de migração com o intuito de determinar o tempo de migração desta base de dados, o resultado dos mesmos pode ser verificado na Tabela 15.

Tabela 15 – Tempos de migração para Azure SQL, por nível de serviço

Nível de Serviço		DTUs	Tempo de Migração (min)
Standard	S0	10	138
Standard	S1	20	93
Standard	S2	50	74
Standard	S3	100	61
Premium	P1	125	39
Premium	P2	250	31

Podemos assim afirmar que para além do desempenho das bases de dados, o nível de serviço também tem influência na migração das mesmas para o Azure.

Foram executados testes de performance aos tempos de processamento de uma das queries mais executadas pela aplicação de testes executada, esta pode ser verificada no código fonte da aplicação e consta da Figura 17.

```
SELECT TOP(7) 1 AS OrderReference, si.StockItemID, si.StockItemName AS [Description],
FLOOR(RAND() * 10) + 1 AS Quantity FROM Warehouse.StockItems AS si
WHERE IsChillerStock = 0
ORDER BY NEWID()
```

Figura 17 – Query Teste 1

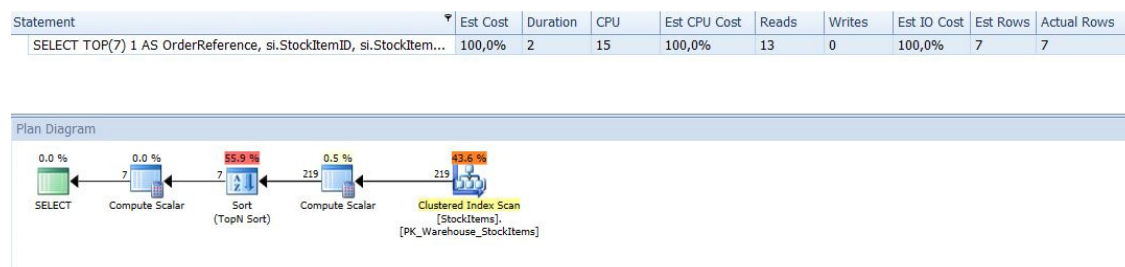


Figura 18 – Plano de Execução da Query Teste 1

Pode-se verificar pela Figura 18, que o tempo estimado para a execução desta query em servidor local, seria de 2ms, o plano de execução foi obtido através do uso do SQL Sentry Plan Explorer, da SQL Sentry, LLC.

A Tabela 16 seguinte apresenta os tempos de processamento dessa mesma query, tendo os tempos sido obtidos no servidor local e para os níveis de serviço S0 e S1 do Azure SQL.

Tabela 16 – Tempos de execução da Query Teste 1

Iteração	Tempo de execução (ms)		
	Servidor	S0	S1
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	0	0
10	0	0	0
11	0	0	0
12	0	0	0
13	0	0	0
14	0	0	0
15	0	0	0
16	0	0	0
17	0	0	0
18	0	0	0
19	0	0	0
20	0	0	0

Como podemos constatar, todos os tempos obtidos foram de 0 ms, assim sendo, não é possível determinar qual a melhor opção a tomar. Ao executar a query acima, todos os resultados são similares aos apresentados na Figura 19.

```
SQL Server Execution Times:
    CPU time = 0 ms,  elapsed time = 0 ms.

SQL Server Execution Times:
    CPU time = 0 ms,  elapsed time = 0 ms.
```

Figura 19 – Screenshot retirado do SSMS

Deste modo, optou-se por escrever uma query que permitisse obter tempos de processamento diferentes de 0. A nova query encontra-se na Figura 20, e permite listar por ordem alfabética todas as cidades existentes na tabela Cities que pertence ao schema Application, bem como a sua localização geográfica, população conhecida, identificador da última pessoa que alterou o registo e as datas de validade. A tabela encontra-se ainda ordenada pelo ID da província a que a cidade pertence.

```

SELECT CityID, CityName, StateProvinceID, Location, LatestRecordedPopulation,
LastEditedBy, ValidFrom, ValidTo
FROM Application.Cities
order by CityName, StateProvinceID

```

Figura 20 – Query Teste 2

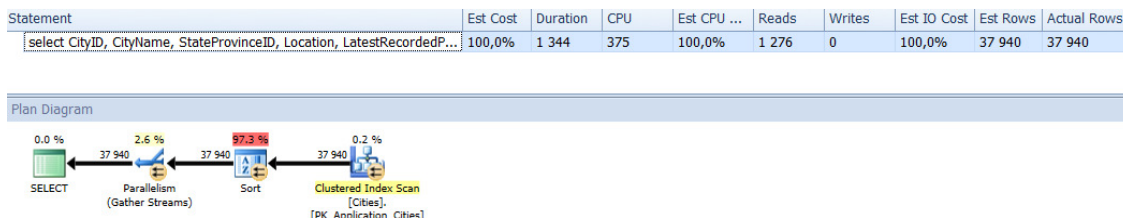


Figura 21 – Plano de Execução da Query Teste 2

Verifica-se pela Figura 21 , que o tempo estimado para a execução desta query seria de 1344ms no servidor local, este valor é superior ao valor anteriormente obtido com a query teste 1, valor esse que foi de 0ms.

A Tabela 17 apresenta os tempos de processamento desta query, tendo os estes sido obtidos, tal como no teste anterior, no servidor local e para os níveis de serviço S0 e S1 do Azure SQL.

Tabela 17 – Tempos de execução da Query Teste 2

Iteração	Tempo de execução (ms)		
	Servidor	S0	S1
1	1155	1010	858
2	1150	1127	391
3	1060	1254	787
4	1043	1107	487
5	949	1457	735
6	1032	1385	858
7	1114	1099	483
8	1078	1161	867
9	1014	1450	922
10	1029	1041	535
11	1054	1045	916
12	1150	1437	961
13	999	1106	379
14	1016	1354	424
15	1132	1019	692
16	993	1890	910
17	1065	1553	532
18	1067	1387	539
19	1106	1521	845
20	1085	1262	817

Para analisar os tempos obtidos utilizamos o método de T-Test com o intuito de obter as médias dos tempos apurados, o resultado encontra-se na Tabela 18 e Tabela 19.

Tabela 18 – Resultados do T-Test para Servidor e Nível de Serviço S0

SUMMARY		Hyp Mean Diff			0
<i>Groups</i>	<i>Count</i>	<i>Mean</i>	<i>Variance</i>	<i>Cohen d</i>	
Servidor	20	1064,55	3300,155		
S0	20	1283,25	52748,934		
Pooled			28024,54	1,30641	

Tabela 19 – Resultados do T-Test para Servidor e Nível de Serviço S1

SUMMARY		Hyp Mean Diff			0
<i>Groups</i>	<i>Count</i>	<i>Mean</i>	<i>Variance</i>	<i>Cohen d</i>	
Servidor	20	1064,55	3300,155		
S1	20	696,90	40941,252		
Pooled			22120,704	2,47192	

Como se pode verificar a média dos tempos obtida para a execução realizada no servidor local é de 1064,55 ms, a obtida para o nível de serviço S0 do Azure SQL é de 1283,25 ms e a do serviço S1 é de 696,90 ms.

Pode-se assim afirmar que a execução da query teste 2 é mais rápida no nível S1 do Azure do que no servidor local. Em termos percentuais a execução da query é 52,75% mais rápida no nível S1, pois a média dos tempos de resposta obtidos é inferior à obtida na execução no servidor local.

Optou-se ainda por escrever uma nova query, esta listará todas as faturas existentes na base de dados, ordenando os resultados pelo ID da fatura. Nela constam a identificação do cliente, o método de entrega, os intervenientes no processo de venda, a data da fatura, o número da nota de encomenda, a morada de entrega, o número de itens que necessitam de refrigeração, as coordenadas de entrega, os dados da confirmação da entrega e a data de modificação do registo, bem como a pessoa que o fez. Esta query é ideal para retirar mapas de vendas, ou ainda para um possível processo de BI (Business Intelligence). Essa nova query encontra-se na Figura 22.

```

SELECT invoices.InvoiceID as "Invoice ID", customers1.CustomerName as "Customer Name"
,customers1.CustomerName as "Bill to Customer", invoices.OrderID as "Order ID"
,deliverymethods.DeliveryMethodName as "Delivery Method", people1.FullName as "Contact Person"
,people2.FullName as "Accounts Person", people3.FullName as "Sales Person"
,people3.FullName as "Packed By", invoices.InvoiceDate as "Invoice Date"
,invoices.CustomerPurchaseOrderNumber as "Customer Purchase Order Number"
,invoices.IsCreditNote as "Is Credit Note"
,ISNULL(invoices.CreditNoteReason, '') as "Credit Note Reason"
,ISNULL(invoices.Comments, '') as "Comments"
,invoices.DeliveryInstructions as "Delivery Instructions"
,ISNULL(invoices.InternalComments, '') as "Internal Comments"
,invoices.TotalDryItems as "Total Dry Items"
,invoices.TotalChillerItems as "Total Chiller Items"
,JSON_Value(invoices.ReturnedDeliveryData, '$.Events[1].Longitude') as "Returned Delivery Data Longitude"
,JSON_Value(invoices.ReturnedDeliveryData, '$.Events[1].Latitude') as "Returned Delivery Data Latitude"
,JSON_Value(invoices.ReturnedDeliveryData, '$.Events[1].Status') as "Status"
,format(invoices.ConfirmedDeliveryTime, 'yyyy-MM-dd hh:mm') as "Confirmed Delivery Time"
,invoices.ConfirmedReceivedBy as "Confirmed Received By"
,people5.FullName as "Last Edited By"
,format(invoices.LastEditedWhen, 'yyyy-MM-dd hh:mm') as "Last Edited When"
FROM [WideWorldImporters-Standard].[Sales].[Invoices] as invoices
inner join [WideWorldImporters-Standard].[Sales].[Customers] as customers1
on invoices.CustomerID = customers1.CustomerID
inner join [WideWorldImporters-Standard].[Sales].[Customers] as costumers2
on invoices.BillToCustomerID = costumers2.CustomerID
inner join [WideWorldImporters-Standard].[Application].[DeliveryMethods] as deliverymethods
on invoices.DeliveryMethodID= deliverymethods.DeliveryMethodID
inner join [WideWorldImporters-Standard].[Application].[People] as people1
on invoices.ContactPersonID = people1.PersonID
inner join [WideWorldImporters-Standard].[Application].[People] as people2
on invoices.AccountsPersonID = people2.PersonID
inner join [WideWorldImporters-Standard].[Application].[People] as people3
on invoices.SalespersonPersonID = people3.PersonID
inner join [WideWorldImporters-Standard].[Application].[People] as people4
on invoices.PackedByPersonID = people4.PersonID
inner join [WideWorldImporters-Standard].[Application].[People] as people5
on invoices.LastEditedBy= people5.PersonID
order by invoices.InvoiceID

```

Figura 22 – Query Teste 3

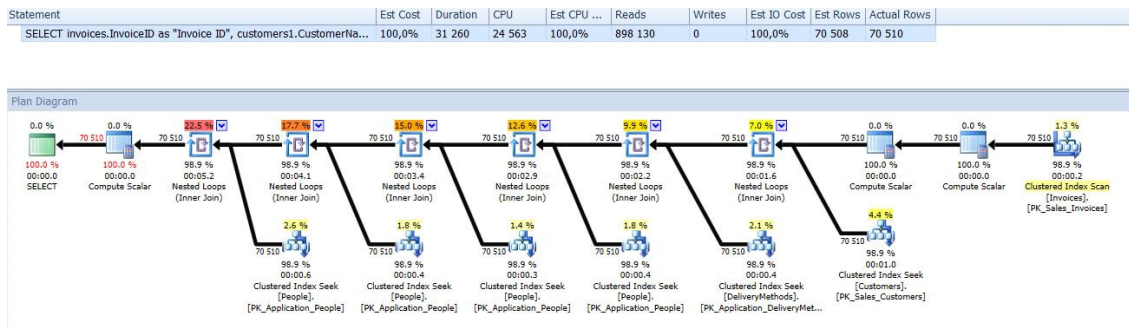


Figura 23 – Plano de Execução da Query Teste 3

No plano de execução, que consta da Figura 23, verifica-se que esta query demoraria 31260ms a ser executada no servidor local. A lentidão desta query deriva da execução de múltiplas instruções de inner join.

Foram realizados testes similares aos anteriores, com o intuito de obter os tempos de execução da query existente na Figura 22. Os testes foram realizados no servidor local, bem como para os níveis S0, S1 e S2 do Azure SQL.

Os resultados do teste encontram-se na Tabela 20.

Tabela 20 – Tempos de execução da Query Teste 3

Iteração	Tempo de execução (ms)			
	Servidor	S0	S1	S2
1	34 477	136 784	88 881	27 272
2	34 973	158 090	88 350	28 154
3	34 723	160 916	89 091	27 212
4	34 509	157 381	86 321	30 690
5	34 192	157 074	85 323	29 575
6	34 551	157 879	85 381	29 537
7	34 682	154 520	85 620	30 369
8	34 627	153 516	85 249	29 968
9	34 310	153 286	86 500	28 872
10	34 133	156 654	87 478	29 744
11	34 363	152 746	89 827	29 706
12	34 749	151 781	89 517	30 555
13	34 594	156 918	89 497	29 847
14	34 037	169 614	81 899	28 005
15	35 054	163 950	79 412	28 241
16	34 461	153 886	78 424	28 612
17	34 516	151 291	80 234	29 133
18	34 503	157 471	81 446	29 495
19	34 433	150 378	85 882	29 515
20	34 285	153 868	87 939	28 119

Novamente, para analisar a média dos tempos obtidos utilizamos o método de T-Test, os resultados encontram-se na Tabela 21 – Resultados do T-Test para Servidor e Nível de Serviço S0, Tabela 22 e Tabela 23.

Tabela 21 – Resultados do T-Test para Servidor e Nível de Serviço S0

SUMMARY			Hyp Mean Diff	0
<i>Groups</i>	<i>Count</i>	<i>Mean</i>	<i>Variance</i>	<i>Cohen d</i>
Servidor	20	34 509	65 973	
S0	20	155 400	39 793 336	
Pooled			19 929 654	27

Tabela 22 – Resultados do T-Test para Servidor e Nível de Serviço S1

SUMMARY			Hyp Mean Diff	0
<i>Groups</i>	<i>Count</i>	<i>Mean</i>	<i>Variance</i>	<i>Cohen d</i>
Servidor	20	34 509	65972,98947	
S1	20	85 614	12584039,52	

Pooled	6325006,257	20,32041
--------	-------------	----------

Tabela 23 – Resultados do T-Test para Servidor e Nível de Serviço S2

SUMMARY		Hyp Mean Diff			0
<i>Groups</i>	<i>Count</i>	<i>Mean</i>	<i>Variance</i>	<i>Cohen d</i>	
Servidor	20	34 509	65972,98947		
S2	20	29 131	1056104,261		
Pooled			561038,625	7,179399	

Como se pode verificar a média dos tempos obtida para a execução realizada no servidor local é de 34509 ms, a obtida para o nível de serviço S0 do Azure SQL é de 155400 ms , a do serviço S1 é de 85614ms e a do nível S2 é de 29131 ms.

Pode-se assim afirmar que a execução da query teste 3 é mais rápida no nível S2 do Azure do que no servidor local, pois a média dos tempos de resposta no nível S2 do Azure SQL são inferiores aos obtidos no servidor local. Em termos percentuais a execução da query é 18,46% mais rápida no nível S2.

Assim, pelos testes efetuados, pode-se afirmar que em termos de performance de execução, a migração pode tornar-se viável.

Em seguida terá de ser realizada uma análise à performance do sistema com carga, para o efeito iremos usar o mesmo simulador de carga usado anteriormente e obter os resultados para um período de 1 hora. Num caso real teria a base de dados ser submetida a uma carga de trabalho por um período mais alargado.

A seguir apresentam-se os 5 testes de performance realizados no Azure, estes foram obtidos utilizando níveis de serviço diferentes, começando pelo mais baixo, salienta-se de que na aplicação usada para simular a carga a caixa de texto ‘Average Order Insertion Time (Milliseconds)’ tinha o valor de 386ms.

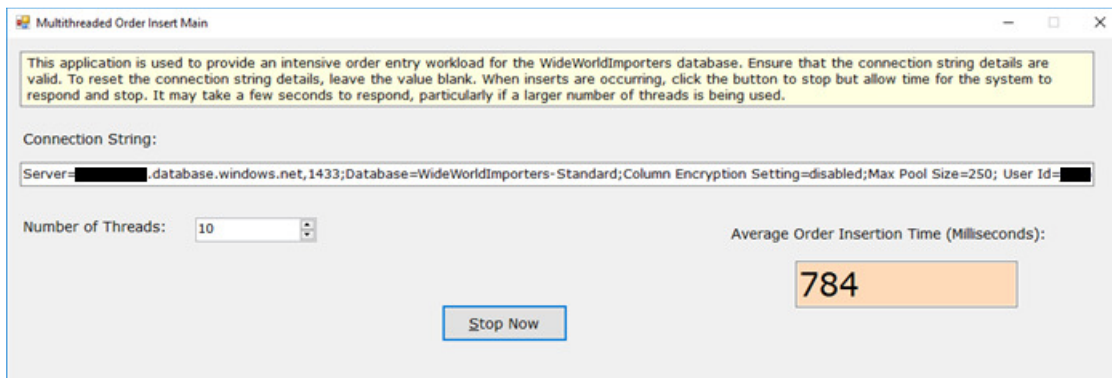


Figura 24 – Multithreaded Order Insert Main – Nível de Serviço S0

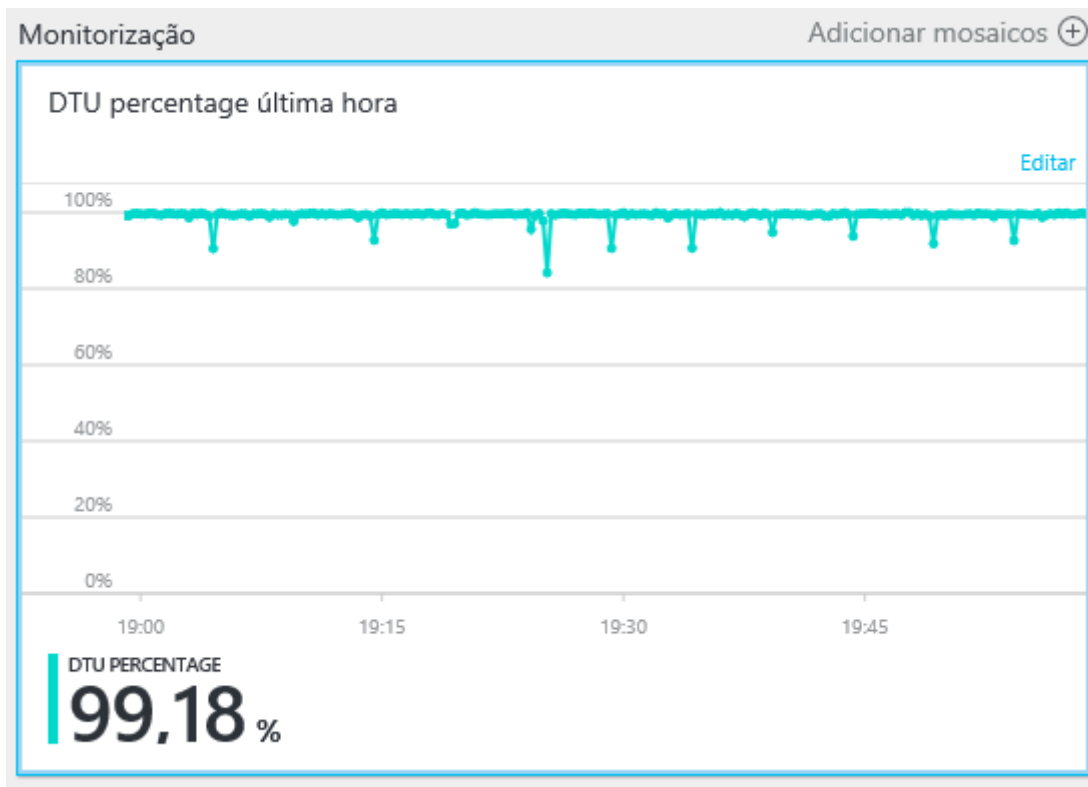


Figura 25 – Monitorização de DTU – Nível de Serviço S0

A Figura 24 e Figura 25 pertencem ao nível de serviço S0 do Azure SQL e demonstram que o tempo de execução médio é de 784ms e a ocupação média de DTUs é de 99,18%.

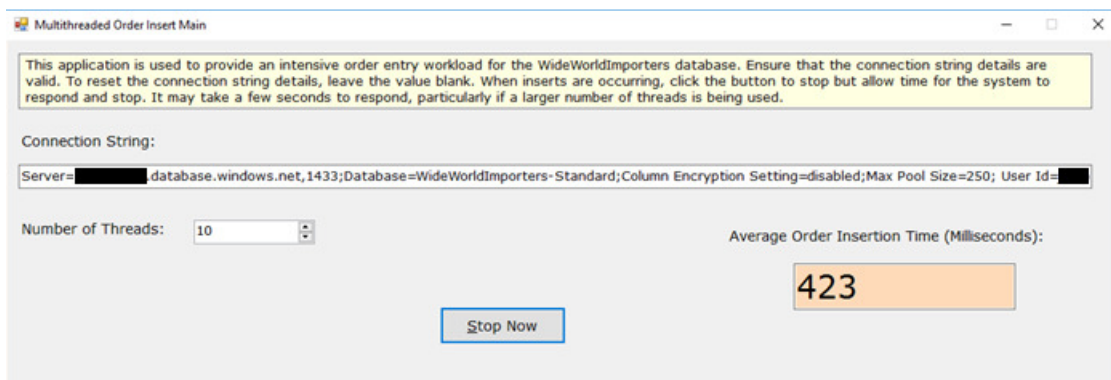


Figura 26 – Multithreaded Order Insert Main – Nível de Serviço S1

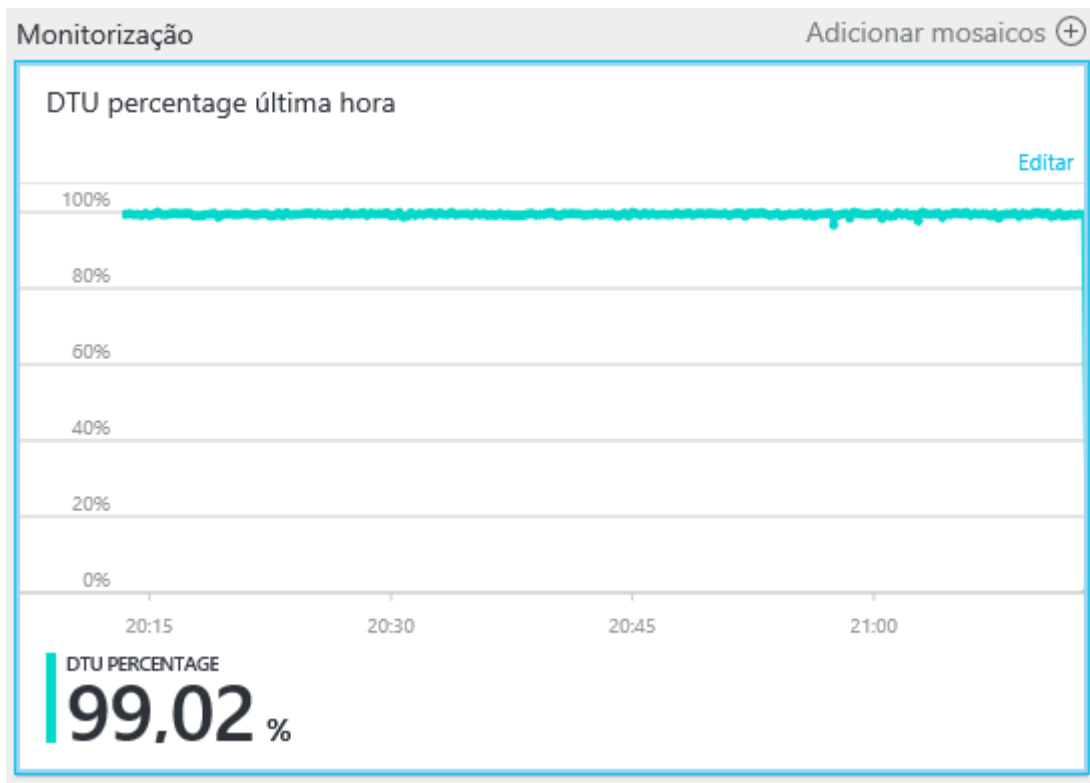


Figura 27 – Monitorização de DTU – Nível de Serviço S1

A Figura 26 e Figura 27 pertencem ao nível de serviço S1 do Azure SQL e demonstram que o tempo de execução médio é de 423ms e a ocupação média de DTUs é de 99,02%.

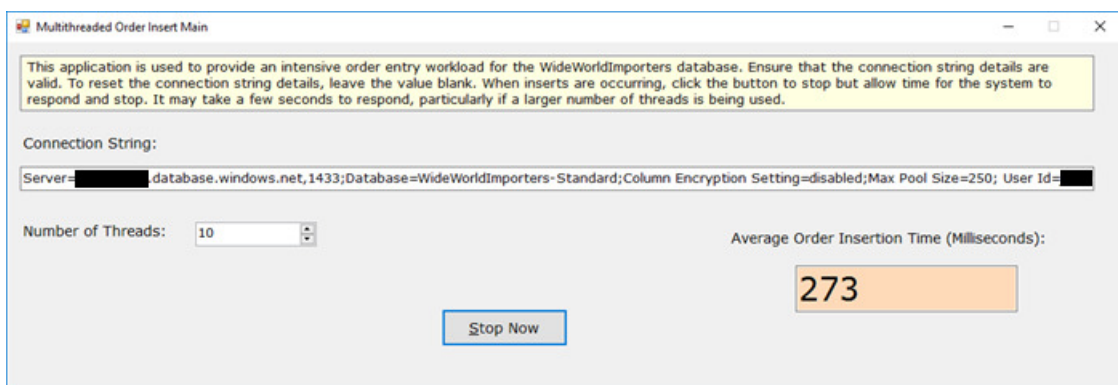


Figura 28 – Multithreaded Order Insert Main – Nível de Serviço S2

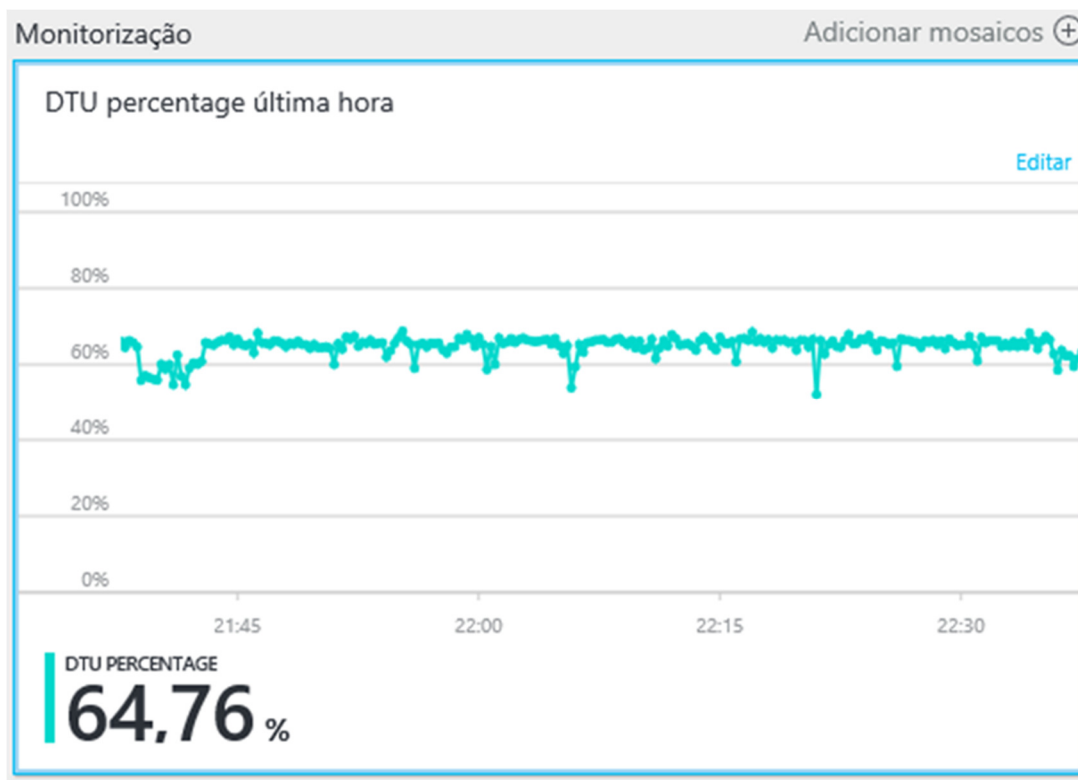


Figura 29 – Monitorização de DTU – Nível de Serviço S2

A Figura 28 e Figura 29 pertencem ao nível de serviço S2 do Azure SQL e demonstram que o tempo de execução médio é de 273ms e a ocupação média de DTUs é de 64,76%. Este tempo de execução já é inferior ao obtido no servidor local.

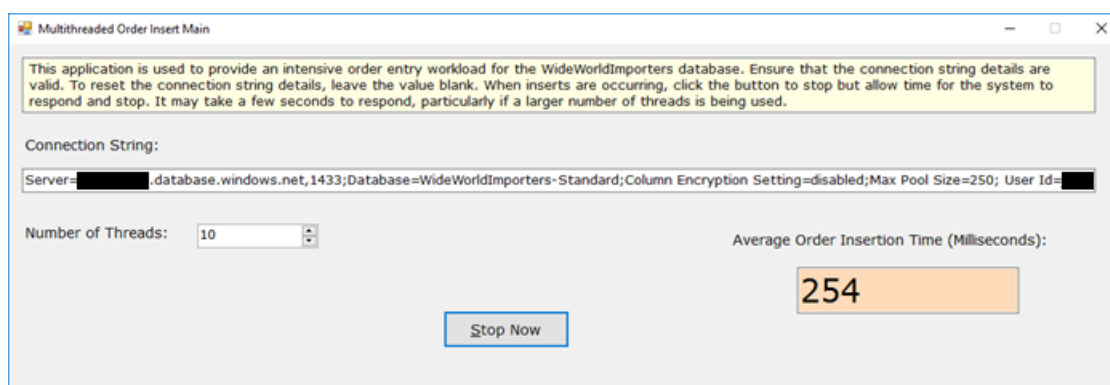


Figura 30 – Multithreaded Order Insert Main – Nível de Serviço S3

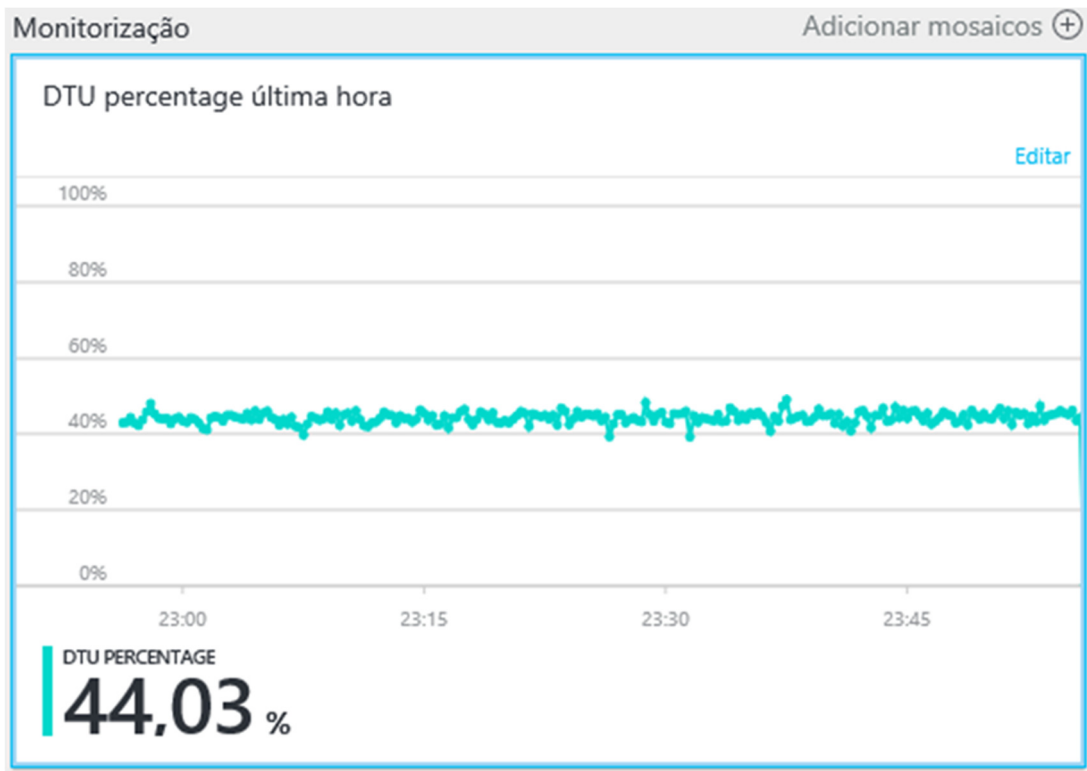


Figura 31 – Monitorização de DTU – Nível de Serviço S3

A Figura 30 e Figura 31 pertencem ao nível de serviço S3 do Azure SQL e demonstram que o tempo de execução médio é de 254ms e a ocupação média de DTUs é de 44,03%.

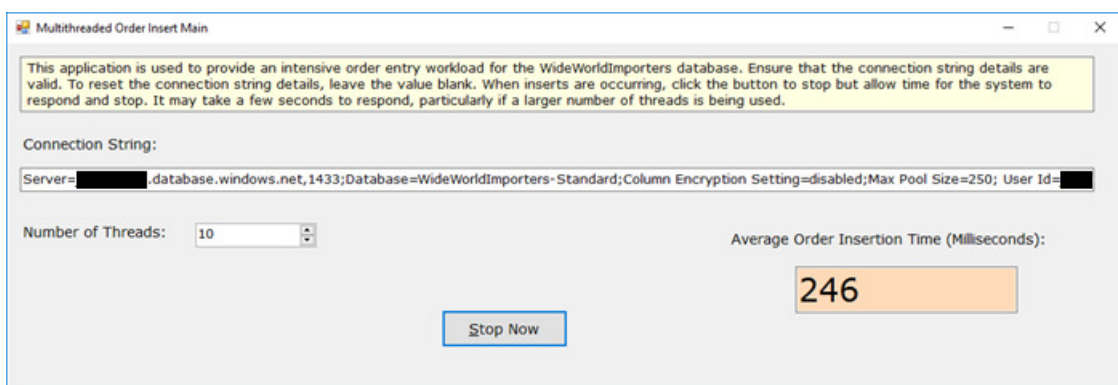


Figura 32 – Multithreaded Order Insert Main – Nível de Serviço P1

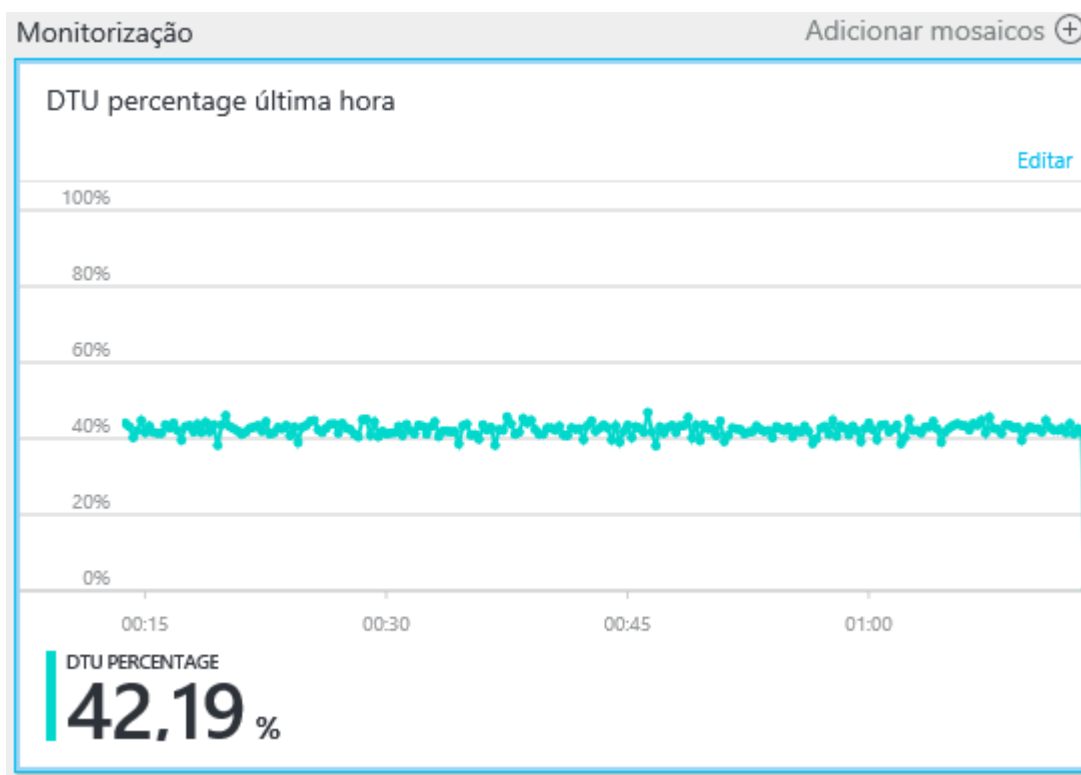


Figura 33 – Monitorização de DTU – Nível de Serviço P1

A Figura 32 e a Figura 33 pertencem ao nível de serviço S2 do Azure SQL e demonstram que o tempo de execução médio é de 246ms e a ocupação média de DTUs é de 42,19%.

Os valores dos cinco testes realizados, foram retirados após 1 hora do início da execução do simulador de carga. As figuras referentes à monitorização de DTUs, foram retiradas do portal do Azure.

Na Tabela 24 é contem o resumo dos testes efetuados.

Tabela 24 – Quadro Resumo dos Testes Efetuados

	Níveis de Serviço Azure SQL				
	S0	S1	S2	S3	P1
Número máximo de DTUs	10	20	50	100	125
Tempo médio de execução (ms)	784	423	273	254	246
Percentagem de DTUs usadas na última hora	99,18	99,02	64,76	44,03	42,19

Como podemos verificar pela Tabela 24, o tempo de execução médio do nível de serviço S2, já é inferior ao obtido pela execução em servidor local, também se pode confirmar que para o mesmo nível se serviço existe folga na percentagem de DTUs usadas na última hora, ou seja o sistema não está “estrangulado” a nível de recursos.

Como podemos constatar na Figura 11, os recursos de CPU já estavam esgotados, assim sendo, seria de equacionar a aquisição de um novo servidor para substituir o atual. Para estimar o custo de aquisição de um novo servidor, é necessário ter em conta tanto o custo do hardware, bem como todo o custo o software (neste caso sistema operativo e eventualmente o SGBD).

Do anexo A, constam a orçamentação de 2 servidores básicos da marca HPE, e do anexo B consta o custo de aquisição de uma licença de sistema operativo e de um SGBD, neste caso o SQL Server.

Também é necessário ter em atenção se o custo do licenciamento do SQL Server é necessário, pois pode ser possível a transferência do licenciamento do servidor em uso para o novo, isto depende de diversos fatores, sendo um deles a diferença do número de cores existente de um servidor para o outro. É de ter em atenção que só é possível adquirir o licenciamento do SQL mais atual, para as outras versões a Microsoft permite fazer o *downgrade*, migrar o SQL Server para versões mais atuais, só clientes com contrato de *Software Assurance*.

Não só os custos oriundos da aquisição dos produtos, mas também os tipos com a manutenção dos mesmos, nomeadamente os custos que dizem respeito aos tidos com a eletricidade, contam para o custo total da solução.

Na Tabela 25, todos os valores, relacionados com o hardware, contêm a manutenção do mesmo por um prazo de 3 anos. Também o custo da eletricidade é para o mesmo período, tendo sido considerado o valor de €0,1608 por kWh, dizendo este valor respeito ao uso da tarifa simples e para uma potência contratada entre os 10,35 e os 20,7 kVA, tendo este valor disso obtido no quadro existente no Anexo C. Foi considerado que o servidor consumia em média 300W/h e que este nunca seria desligado.

Tabela 25 – Custo Total de Aquisição e Manutenção de hipóteses de servidor

	HP Proliant ML150 Gen9 (6 cores)	HP Proliant ML30 Gen9 (4 cores)
Hardware	€4 601,00	€3 515,00
Sistema Operativo	€737,40	€737,40
Total Custo Básico	€5 338,40	€4 252,40
Licenciamento do SQL Server	€3 124,38	€2 082,92
Energia	€1 267,75	€1 267,75
Total	€9 730,53	€7 603,07

Como referido anteriormente o nível de serviço do Azure SQL S2 é suficiente em termos de DTUs para suportar a base de dados a migrar, à data de setembro de 2016, o custo mensal é de €63,26. A manter-se este preço, num período de 3 anos, o total é de €2.277,36. Este valor fica aquém de qualquer um dos valores estimados para a aquisição e manutenção de um servidor novo.

O custo pode ainda ser reduzido se existirem períodos determináveis em que o processamento na base de dados seja reduzido ou nulo, nesses períodos podemos usar scripts de powershell para modificar o nível de serviço do atual para, por exemplo o nível S0, abaixo apresenta-se um script para modificar o nível de serviço.

```

Set-ExecutionPolicy Unrestricted

$azureAccountName = "██████████@██████████.onmicrosoft.com"
$password = "██████████"
$azurePassword = ConvertTo-SecureString -AsPlainText -Force -String $password

$subscriptionId = "██████████"
$resourceGroupName = "██████████"
$location = "North Europe"
$serverName = "██████████"
$databaseName = "WideWorldImporters-Standard"
$newEdition = "Standard"
$newPricingTier = "S0"

Try
{
    $azureRMCredential = New-Object -TypeName System.Management.Automation.PSCredential($azureAccountName,$azurePassword)
}
Catch
{
    Write-Error ("Failed to create a PSCredential object.")
    Break
}
Try
{
    Login-AzureRmAccount -Credential $azureRMCredential
    #Add-AzureRmAccount
    Select-AzureRmSubscription -SubscriptionId $subscriptionId

    $scaleRequest = Set-AzureRmSqlDatabase -DatabaseName $databaseName -ServerName $serverName `
    -ResourceGroupName $resourceGroupName -Edition $newEdition -RequestedServiceObjectiveName $newPricingTier

    $scaleRequest
}
Catch
{
}

```

Figura 34 – Script em Powershell

Usando este script para baixar o nível de serviço de S2 para S0, por exemplo, aos fins de semana, admitindo 52 semanas por ano, pouparíamos cerca de €510 para um período de 3 anos.

6.2.3 Conclusões

Pelo exemplo acima, chega-se à conclusão que a migração realizada é viável em termos de custos e em termos de performance.

Na Tabela 26 **Error! Reference source not found.**, pode-se verificar os incrementos de performance do tempo de resposta da query 2 e da query 3, de salientar que esta informação foi obtida não tendo carga na base de dados.

Tabela 26 – Incrementos de Performance no Tempo de Execução

Query	Migração da BD	Incremento de Performance
2	Servidor - Nível S1 do Azure SQL	52,75%
3	Servidor - Nível S2 do Azure SQL	18,46%

A Tabela 27, apresenta a percentagem de DTU's consumida, no Azure SQL, durante uma hora de uso consecutivo do simulador de carga.

Tabela 27 – Percentagem de DTU's Usada no Azure

Nível de Serviço do Azure SQL	Percentagem usada /h
S0	99,18%
S1	99,02%
S2	64,76%
S3	44,03%
P1	42,19%

Pode-se verificar pela Tabela 27, que o gasto de DTU's para o nível S0 e S1 do Azure SQL, não são apropriados para a migração, mas que a partir do nível S2 a migração já se torna viável.

O custo mensal da solução proposta, que é a S2, é de €63,26. Este preço é relativamente baixo, visto que para além da disponibilização da base de dados, estão incluídos também os backups da mesma. A Tabela 28 apresenta os custos a três anos do uso do nível S2 do Azure SQL, admitindo a inexistência de alterações aos preços em vigor em outubro de 2016.

Tabela 28 – Custo a três anos do Nível S2 do Azure SQL

Descrição	Valor
Custo do Nível S2 do Azure SQL	€2 277,36
Poupança descendo do Nível S2 para S0 do Azure SQL aos fins de semana	€510,00
Total	€1 767,36

Comparando estes custos com os existentes na Tabela 25, verifica-se que os custos para o Azure SQL são de longe mais económicos do que os da aquisição de um novo servidor, do respetivo licenciamento e dos gastos energéticos, para um período de 3 anos, pois os valores para os possíveis novos servidores iniciam a partir dos €7600.

Chegou-se também à conclusão que o nível de serviço é importante para operações de migração de bases de dados para o Azure SQL, pois tem em atenção os DTUs disponibilizados pelo nível de serviço. Tal pode ser comprovado na Tabela 15.

No decorrer da experiência foi verificado a facilidade com que se pode criar e manter uma base de dados no Azure SQL, bem como a existência de diversa documentação útil sobre o assunto, também se pode encontrar com alguma facilidade scripts em powershell que ajudam a realizar operações de manutenção de base de dados.

O portal do Azure é de fácil usabilidade, o que permite que uma operação para mudar o nível de serviço demore apenas alguns segundos a ser realizada.

7 Conclusões e Trabalho Futuro

A gestão eficiente de uma base de dados é um dos pontos fulcrais das TIC, sendo que atualmente se coloca um paradigma relativamente à escolha entre o uso de SGBD ditos tradicionais ou o uso de serviços do tipo DaaS existentes na Cloud.

Este trabalho tinha como objetivos a realização de uma análise aos DaaS, tanto a nível de custos, como ao nível da performance prestada, bem como proceder à escolha de um DaaS para uso numa PME.

Foi descrito neste trabalho, os princípios básicos da Cloud e inerentemente os princípios associados a um serviço tipo DaaS, sendo que considero como mais importantes:

- a capacidade de atribuir ou retirar dinamicamente recursos, sem obrigar a interações com outras pessoas;
- a capacidade de monitorização dos sistemas e da faturação associado ao consumo dos recursos.

Verificou-se pela pesquisa efetuada que existem diversos modelos de implementação de DaaS, de diferentes fornecedores de serviço, alguns com modelos simples de implementar outros com modelos mais complexos.

Constatou-se que a informação disponibilizada pelos fornecedores de serviço é modificada com alguma frequência, podendo serem criados novos níveis de serviço, ou mesmo alteradas as especificações dos existentes. Os preços também variam conforme a localização geográfica.

O exemplo da migração para o Azure SQL apresentado, pretende demonstrar a escolha de uma ferramenta DaaS a ser implementada numa PME. Optou-se neste caso pelo uso de um DaaS fornecido pelo desenvolvedor do produto, neste caso a base de dados era em Microsoft SQL Server e foi migrada para o Azure SQL que é um dos serviços disponibilizados pelo Azure, sendo este propriedade da Microsoft.

Pode-se verificar pelo exemplo apresentado que a migração para Azure SQL, é fácil de ser realizada e foi conseguida com êxito. Neste momento o Azure SQL não suporta todas as funcionalidades de um SQL Server, para o caso de não ser possível a migração ou o uso do Azure SQL, a Microsoft disponibiliza outro tipo de serviço que não o DaaS, que passa por ter um SQL Server numa máquina virtual, sendo este serviço considerado um IaaS. Também pela experiência efetuado foi verificado que o utilizador mediante a utilização do portal, não se apercebe da infraestrutura que suporta todos os serviços prestados pelo Azure.

A avaliação experimental efetuada comprovou o aumento de performance quando se utiliza um serviço DaaS, bem como mostrou que a compra de novo equipamento fica mais dispendiosa que manter uma solução no Azure SQL. O exemplo utilizado neste trabalho pode também ser usado como um guia na migração de bases de dados existentes em SQL Server para o Azure SQL.

Do ponto de vista pessoal, este trabalho também contribuiu para a abertura ao novo mundo dos DaaS, percebendo minimamente os seus princípios de funcionamento, bem como das soluções existentes no mercado.

Como trabalho futuro propõe-se uma avaliação experimental de cada um dos DaaS apresentados, com a realização de um caso de uso para cada um deles, e também utilizando um benchmark, por exemplo do TPC – Transaction Processing Council.

Referências

- (Conger 2011) Steve Conger, 2011. Hands-On Database. 1 Edition. Prentice Hall
- (Foster and Godbole, 2014) Foster, E. and Godbole, S. (2014) Database systems: A pragmatic approach. Germany: APress
- (Stephens, Plew, and Jones, 2011) Stephens, R., Plew, R. D. and Jones, A. D. (2011) Sams teach yourself SQL in 24 hours. 5th edn. Indianapolis, IN: Sams Publishing.
- (Caldeira, 2015) Carlos Pampulim Caldeira, 2015. PostgreSQL Guia Fundamental. 1ª Edição. Lisboa: Edições Sílabo, Lda
- (Tavares, 2015) Frederico Tavares, 2015. MySQL. Lisboa: FCA
- (Magalhães, 2015) Alberto Magalhães, 2015. SQL Server 2014 Curso Completo. Lisboa: FCA
- (Chodorow, 2013) Chodorow, K. (2013) MongoDB: The definitive guide. 2nd edn. United States: O'Reilly Media, Inc, USA
- (Ferreira, 2015) António Miguel Ferreira, 2015. Introdução ao Cloud Computing. IaaS, PaaS, SaaS, Tecnologia, Conceito e Modelos de Negócio. Lisboa: FCA
- (Bond, 2015) Bond, J. (2015) The enterprise cloud: Lessons learned. United States: O'Reilly Media, Inc, USA.
- (ERSE, 2016) ERSE (2016) Preços de Referência no Mercado Liberalizado de Energia Elétrica e Gás Natural em Portugal Continental
- (Oracle corporation story, 2016) Oracle corporation story (2016). Disponível em: <https://successstory.com/companies/oracle-corporation> (Acedido em 5 October 2016).
- (Abadi, 2009) Abadi, D.J. (2009) 'Data management in the cloud: Limitations and opportunities', IEEE Data Eng. Bull., 32(1).
- (Abourezq and Idrissi, 2016) Abourezq, M. and Idrissi, A. (2016) 'Database-as-a-service for big data: An overview', International Journal of Advanced Computer Science and Applications, 7(1). doi: 10.14569/ijacsa.2016.070124.

Anexos

A. Orçamentos Servidores HPE

Os dois orçamentos abaixo apresentados foram realizados no simulador online existente no site: <https://sce-public.houston.hp.com/SimplifiedConfig/Welcome>.

QTD	NÚMERO DE PEÇA	DESCRIÇÃO
1	767063-B21	Servidor HP ProLiant ML150 Gen9 hot plug 4LFF, configuração segundo pedido
1	767063-B21 B19	HP ProLiant ML150 Gen9 Hot Plug 4LFF Europe-Multilingual Localization Configure-to-order Server
1	726660-L21	HP ML150 Gen9 Intel Xeon E5-2609v3 (1.9GHz/6-core/15MB/85W) FIO Processor Kit
4	726718-B21	Kit de memória CAS-15-15-15 HP de categoria única x4 DDR4-2133 de 8 GB (1 x 8 GB) registrado
4	726718-B21 OD1	Factory Integrated
2	657750-B21	Unidade de disco rígido SC Midline HP 1TB, 6G, SATA, 7,2K rpm, LFF (3,5 polegadas) com 1 ano de gara
2	657750-B21 OD1	Factory Integrated
2	658079-B21	Unidade de disco rígido SC Midline HP 2TB, 6G, SATA, 7,2K rpm, LFF (3,5 polegadas) com 1 ano de gara
2	658079-B21 OD1	Factory Integrated
1	726536-B21	Unidade óptica HP, 9,5 mm SATA DVD-ROM JackBlack Gen9
1	726536-B21 OD1	Factory Integrated
1	730941-B21	Kit de fonte de alimentação integrado de fábrica HP 550 W
1	784603-B21	Kit de ativação de unidade óptica HP ML150 Gen9 slim
1	784603-B21 OD1	Factory Integrated
1	H7J34A3	HPE 3Y Foundation Care 24x7 SVC
1	H7J34A3 YOL	HP ProLiant ML150 Gen9 Supp
1	HA113A1	HPE Installation SVC
1	HA113A1 5T6	HPE Installation DL160/DL360e Service

Total estimado da configuração EUR4,60100

Figura A.1 – Orçamento HP ProLiant ML150 Gen9

QTD	NÚMERO DE PEÇA	DESCRIÇÃO
1	823402-B21	Servidor com configuração sob pedido HPE ProLiant ML30 Gen9 hot-plug 4 LFF
1	823402-B21 B19	HPE ProLiant ML30 Gen9 Hot Plug 4LFF Europe-Multilingual Localization Configure-to-order Server
1	823373-L21	HPE ML30 Gen9 Intel Xeon E3-1240v5 (3.5GHz/4-core/8MB/80W) FIO Processor Kit
4	805669-B21	Kit de memória padrão HPE 8 GB (1x8 GB) de categoria dupla x8 DDR4-2133 CAS-15-15-15 sem buffer
4	805669-B21 OD1	Factory Integrated
2	657750-B21	Unidade de disco rígido SC Midline HP 1TB, 6G, SATA, 7,2K rpm, LFF (3,5 polegadas) com 1 ano de gara
2	657750-B21 OD1	Factory Integrated
2	658079-B21	Unidade de disco rígido SC Midline HP 2TB, 6G, SATA, 7,2K rpm, LFF (3,5 polegadas) com 1 ano de gara
2	658079-B21 OD1	Factory Integrated
1	726536-B21	Unidade óptica HP, 9,5 mm SATA DVD-ROM JackBlack Gen9
1	726536-B21 OD1	Factory Integrated
1	339778-B21	HP RAID 1 Drive 1 FIO Setting
1	503296-B21	Kit de fonte de alimentação HP 460W Common Slot Gold Hot Plug
1	503296-B21 OD1	Factory Integrated
1	822607-B21	Kit de ativação de fonte de alimentação redundante HPE ML30 Gen9 4U
1	822607-B21 OD1	Factory Integrated
1	820288-B21	Kit de ativação de unidade de disco óptico HPE ML30 Gen9 slim
1	820288-B21 OD1	Factory Integrated
1	H7J34A3	HPE 3Y Foundation Care 24x7 SVC
1	H7J34A3 XMP	HPE ProLiant ML30 Gen9 Supp
1	HA113A1	HPE Installation SVC
1	HA113A1 5T6	HPE Installation DL160/DL360e Service

Total estimado da configuração EUR3,515,00

Figura A.2 – Orçamento HP ProLiant ML30 Gen9

B. Custo de Aquisição de Software

As duas figuras abaixo apresentadas foram retiradas do site: <https://www.mbit.pt/>, e contêm IVA à taxa de 23%.



**MICROSOFT - WIN
SERVER STD 2012 R2
X64 ING 2CPU / 2VM O**

Marca: MICROSOFT
Código Produto:
P73-06165
UPC/EAN: 0885370659931

Disponibilidade:
✔ Online: envio em 1 a 2 dias úteis
✔ Nas lojas: entrega em 2 a 3 dias úteis

906,99 €

Comprar: 

Figura B.1 – Custo de uma licença de Microsoft Windows 2012 Server Standard



**MICROSOFT - MOLP
SQLSvrStd 2014/SGL
ONL**

Marca: MICROSOFT
Código Produto:
228-10344

Disponibilidade:
✔ Online: disponível
✔ Nas lojas: entrega em 2 a 3 dias úteis

1 280,99 €

Comprar: 

 Like 0

Figura B.2 – Custo de uma licença de Microsoft SQL Server 2014 Standard

Relativamente ao licenciamento, uma licença de Microsoft Windows 2012 Server Standard, licencia um servidor com no máximo dois processadores físicos, no entanto uma licença de Microsoft SQL Server 2014 Standard, apenas licencia 2 cores de CPU.

O custo de uma licença de Microsoft Windows 2012 Server Standard fica por €737,39 sem IVA, sendo que uma licença de Microsoft SQL Server 2014 Standard fica por €1.041,46.

C. Custo da Energia

A **Error! Reference source not found.** consta foi retirada de (ERSE, 2016), atualizado a 9 de Agosto de 2016 e refere-se a preços de energia praticados pela EDP Comercial, para o mercado empresarial e que utilizem débito direto como forma de pagamento, bem como tenham adesão à fatura eletrónica.

1.2.10 NEGÓCIOS (COM DD E FE)

Validade da proposta: **Válida até nova campanha**
 Duração do contrato: **12 meses**

Pot. Contratada (kVA)	Simples		Bi-Horário			Tri-Horário			
	Potência (€/Dia)	Energia (€/kWh)	Potência (€/Dia)	Energia (€/kWh)		Potência (€/Dia)	Energia (€/kWh)		
				Fora Vazio	Vazio		Ponta	Cheias	Vazio
1,15	0,0832	0,1408							
2,3	0,1461								
3,45	0,1552	0,1601	0,1918	0,1970	0,0928	0,1939	0,3179	0,1634	0,0934
4,6	0,2017		0,2360			0,2376	0,3174	0,1629	0,0929
5,75	0,2480		0,2801			0,2807	0,3173	0,1630	0,0924
6,9	0,2943		0,3242			0,3285	0,3182	0,1636	0,0936
10,35	0,4331	0,1608	0,4564	0,1974	0,0927	0,4590	0,3176	0,1632	0,0927
13,8	0,5718		0,5891			0,5922	0,3177	0,1633	0,0927
17,25	0,7106		0,7217			0,7249	0,3183	0,1641	0,0925
20,7	0,8495		0,8543			0,8643	0,3192	0,1651	0,0935
27,6						1,3379			
34,5						1,6674	0,2878	0,1438	0,0805
41,4						1,9969			

Notas:

- Aplicável a clientes que optem pelo Débito Direto como meio de pagamento.
- Obrigatoriedade de faturação eletrónica.

Figura C.1 – Retirado de (ERSE, 2016, P. 13)

D. Migração para o Azure SQL

Abaixo seguem as instruções necessárias para validar a possibilidade da migração de uma base de dados para o Azure SQL, as instruções encontram-se em inglês, é necessário selecionar primeiro a base de dados. As instruções devem ser realizadas no SSMS (SQL Server Management Studio), e são:

- Selecionar a opção 'Tasks' -> 'Export Data-Tier Application'
- Em 'Introduction' pressionar 'Next'
- Em "Export Settings" no separador 'Settings' escolher a localização onde gravar o ficheiro a ser gerado, depois em 'Advanced' retirar o visto na opção 'Select All', para não serem exportados os dados. Posteriormente pressionar novamente 'Next'
- Por fim pressionar 'Finish'.

Para efetuar a migração para o Azure SQL, as seguintes alterações deverão ser realizadas nas propriedades da base de dados:

- Colocar no campo 'Database Read-Only' a 'True', isto permite que não sejam efetuadas alterações aos dados.
- Colocar no campo 'Restrict Access' a 'RESTRICTED_USER', isto permite a ligação à base de dados a membros dos seguintes grupos: db_owner, dbcreator ou sysadmin.

Tal pode ser comprovado na Figura D.1

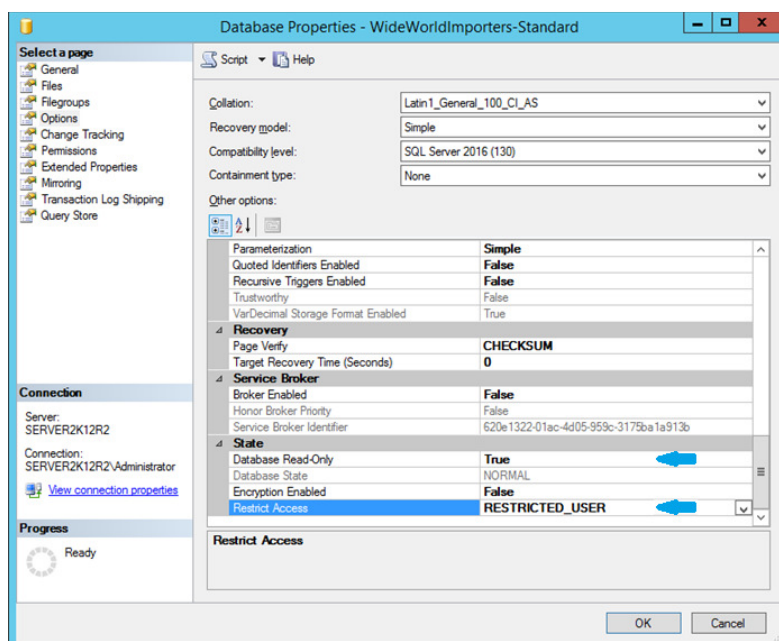


Figura D.1 – Opções a modificar para restringir acesso à base de dados

As migrações de base de dados para o Azure podem ser realizadas no SSMS, devendo para serem executados os seguintes passos:

- Selecionar a base de dados
- Pressionar o botão direito do rato, ir à opção 'Tasks' -> 'Deploy Database to Microsoft Azure SQL Database'
- Em 'Introduction' pressionar 'Next'
- Em 'Deployment Settings', pressionar o botão 'Connect' e introduzir os dados de acesso ao servidor de SQL criado anteriormente no Azure, atribuir o nome à base de dados, por defeito é igual à existente, escolher o nível de serviço que a base de dados irá ter no Azure. Posteriormente pressionar novamente 'Next'
- Em 'Summary', verificar as opções escolhidas e pressionar 'Finish'.

Na Figura D.2, Figura D.3 e Figura D.4 apresentam-se os ecrãs, executados no SSMS, nos quais se podem comprovar as operações acima mencionadas e realizadas.

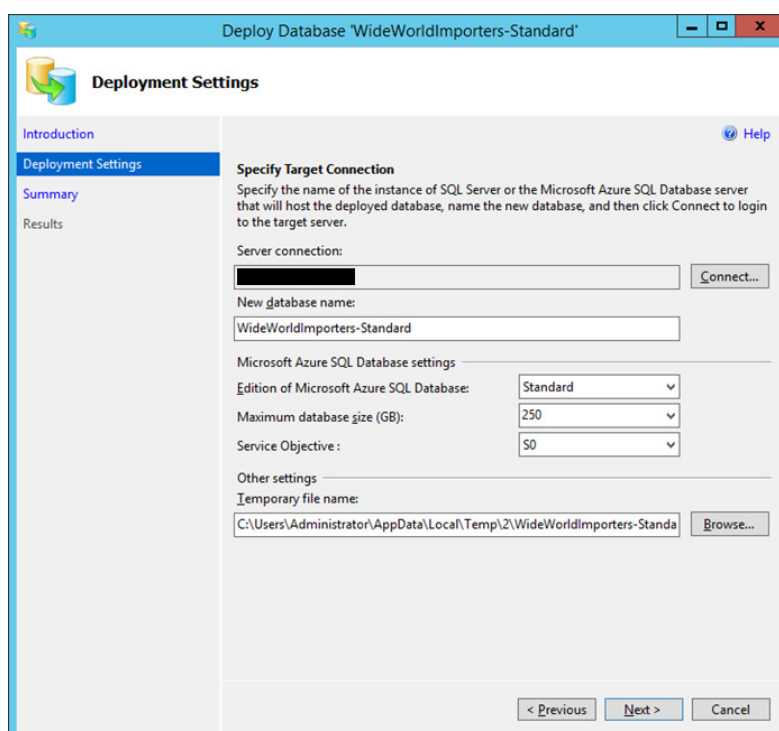


Figura D.2 – SSMS Migração Ecrã 1

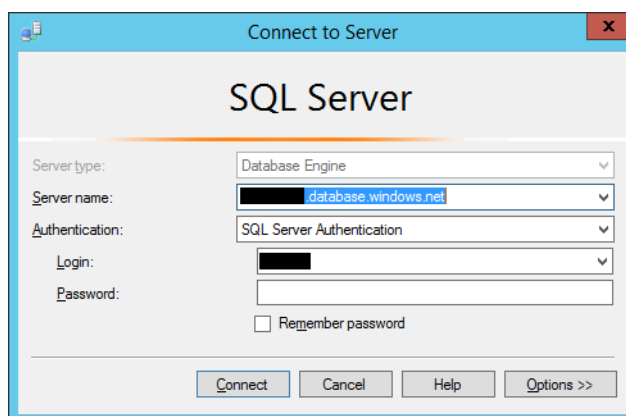


Figura D.3 – SSMS Migração Ecrã 2

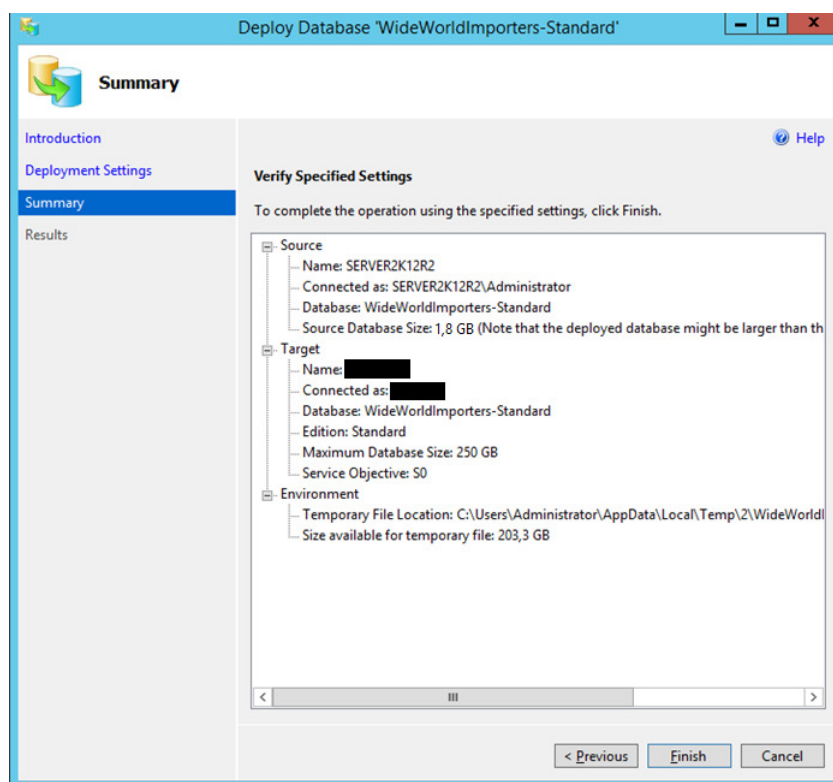


Figura D.4 – SSMS Migração Ecrã 3