



河南工业大学

《云计算》课程 考核报告

姓名：_____郭治洪_____

班级：_____物联网 1603_____

学号：_____201616070320_____

学期：_____2017-2018-2_____

指导教师：_____魏蔚_____

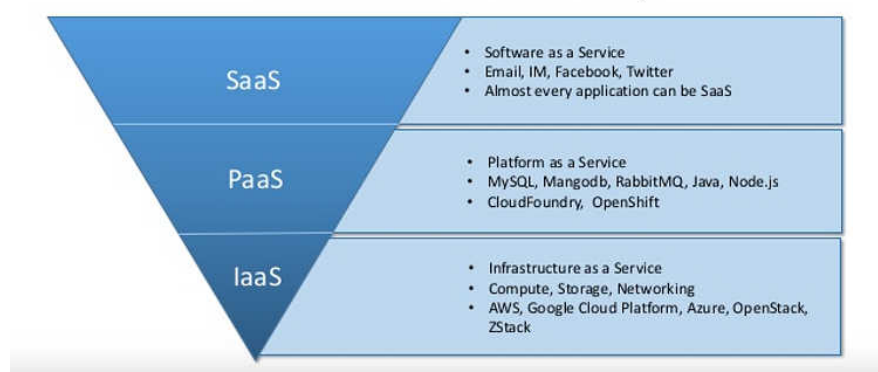
完成时间：_____2018-4-30_____

一、云计算有别于一般分布式计算的六大特点是什么？三种云计算类别是什么，之间有什么区别和联系？（10分）

云计算有以下六大特点：

- (1) 超大规模：使用多台计算机，借助自动化的管理平台和体系架构作为集群处理大量的数据满足需求。
- (2) 虚拟化：机器采用的封装好的系统镜像，用以快速通过各种终端设备控制部署到云服务器上，方便启动关闭，控制操作，添加删除。
- (3) 通用性：云计算适用性强，不针对某些特定条件环境使用，无需用户了解架构的部署和调度，只要花钱即可使用。
- (4) 高可靠性：云计算使用大公司的数据中心的集群机器，拥有完备的冗余环境（电源、网络环境），同时采用多种方式进行冷热备份，多机备份，异地备份等等减少异常错误。
- (5) 高可伸缩性：云计算借助虚拟化以及自动化调度工具，可以根据不同的需求选择调节硬件配置，满足所需。
- (6) 廉价且按需求服务：借助硬件的发展，用户可以根据需求花费极低的费用满足计算需求，同时根据需要选择服务以及使用期限，减少不必要开支。

三种云计算分别是 IAAS（基础措施作为服务），PAAS(平台作为服务)，SAAS(软件作为服务)。



每种云计算提供的服务等级和专有程度不同，使用费用也不同，偏重的使用对象也不同和可以配置的空间也不同。

一般来说上级服务针对的需求更专一，费用也越高，服务越细致，适合直接需要成果的中小型公司。在中层和下层的服务将会提供大的可操作空间，但是不会提供更专一的服务，

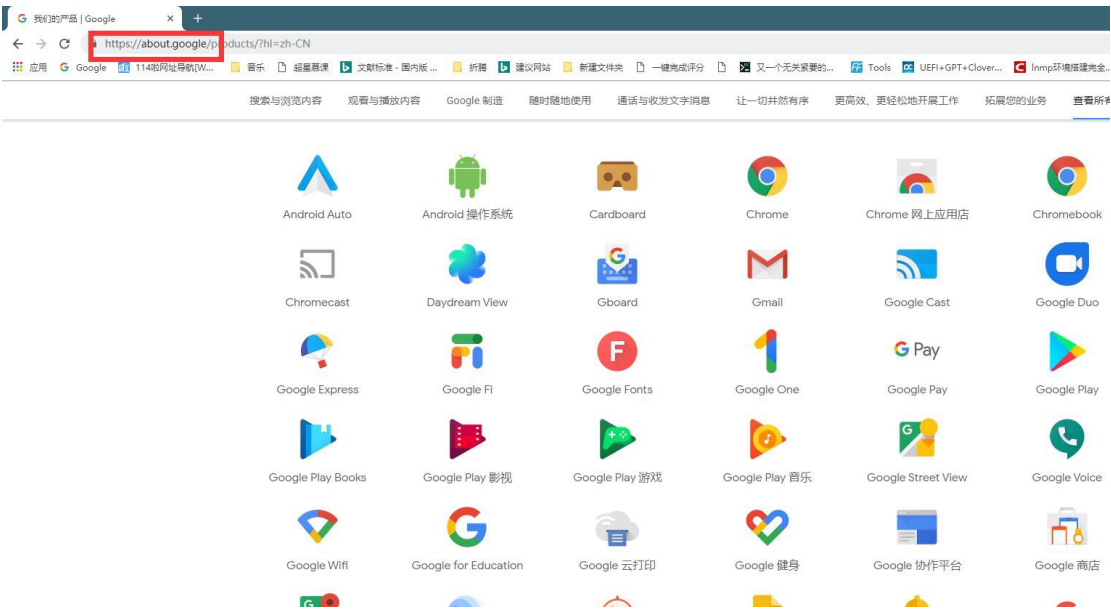
费用也会降低，开发需要的一定的成本，因此只有大型和巨型互联网公司才拥有。

个人认为上层更偏重应用，下层更加偏重基础，一般只有有能力大公司和国企才能提供下层的服务，新的中小型公司都是租借下层而开发应用使用。

不过我想说一点是，上层应用层随时有被新应用取代趋势，下层平台和基础才是目前云计算重心，个人觉得越往深入越接近底层才有实力立足做大。其实我个人觉得上层应该叫做云服务，下层才能叫做云平台和云计算。

云计算虽然是个范围很广的词，但是我们应该通过现象看到本质。

二、请列举 Google 最好的三个 SaaS 应用，并分别阐述相对其他同类产品的优点。（10 分）



谷歌现在产品线拉的很长，竞争对手很多，但是最重要的核心业务几乎达到了无法匹敌境界：搜索，邮箱和 YouTube 视频。

个人认为谷歌最大优点还是技术。在 03 年时候谷歌就提出了分布式云计算的概念，而且谷歌很有远见收购很多创新科技公司培养（例如 Deep Wind）。谷歌提出的这些框架或问题解决方法，收购这些科技创新公司的产品能够吸引人们的目光，引起新的潮流方向。

以国内来说，现在国内没有一家互联网企业能够在全球占有优势的，国内现在目前环境还是偏重快速应用型，说白了就是好赚钱。相比国外来说差距很多，虽然表面上看起来很火，但是实际都是因为吸引到大量使用用户，感觉并没有很强的竞争优势，所以可能仅仅用在国内环境。而谷歌包括国外其他科技公司不同，提出一个个又一个框架和思想方法，并且引领学习潮流，我觉得还有很长的路要走。如果国内也有像谷歌，微软，脸书，推特这样公司崛

起，我们国家或许真的强大。

总结下是因为有核心竞争力，因为强大到没有对手，先进到引领世界，所以成功。

这是来自一个知乎的回答，可以论证我的想法：

下面列幾個重要的而且公開資料較多的基礎技術：

- **大規模機群建設與管理。** Google 的情況可以參見 [2] The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines, Second Edition. Google 擁有世界上最大的計算機集群，論機器數量的話能在量級上超過所有其他公司。同時，它有一整套自動化管理軟件，以便工程師申請和使用這些硬件資源（大致可以理解成一套 Amazon EC2）。就我的了解，現在在普通工程師使用機群硬件資源的方便程度和可以使用的量上，百度還是遠遠不及。
- **大規模計算與存儲。** Google 論文老三篇 GFS, MapReduce, BigTable 不再贅述，近年 Google 在這些方面的研發和進步沒有停滯甚至在加快。當然百度也在努力追趕，百度不僅使用 Hadoop，而且基於 Hadoop 做了大量改進和擴展，並貢獻回 Hadoop 開源社區。百度在 SSD 存儲技術等方面也很有心得，比如 flash 存儲方面最近中了一篇 ASPLOS '14 SDF: Software-Defined Flash for Web-Scale Internet Storage System.
- **機器學習和人工智能。** 被吹得神乎其神的 deep learning 和 Google Brain 等等。在 deep learning 這個相對較新的領域，百度追趕的更快，水平也更接近。

三、Google 的 GFS 文件系统相对现存文件系统的优点，主要为解决哪两个核心问题？。（12 分）

1. 在不可靠的硬件系统上建设可靠的文件服务。廉价和易损的硬件容易损坏，数据可能会丢失，因此要借助软件和自动系统完成数据的存储备份和异常处理。

2. 大文件以及流式文件的存储处理。因为数据量大，需要能够分布式储存；这些数据需要批量处理就像生产线一样不间断，并不需往复修改。

四、分布式计算核心要解决的两个问题是什么？MapReduce 架构如何做到尽可能通用化？（12 分）

1. 任务的分配，程序的拆分，数据分发计算。因为要进行分布式计算，就需要将之前写的程序进行改写和任务计算数据调度分配，这些数据计算完还要进行收集统计汇总。

2. 计算中可能产生的错误。某些节点产生的异常不能够影响整体的计算结果。

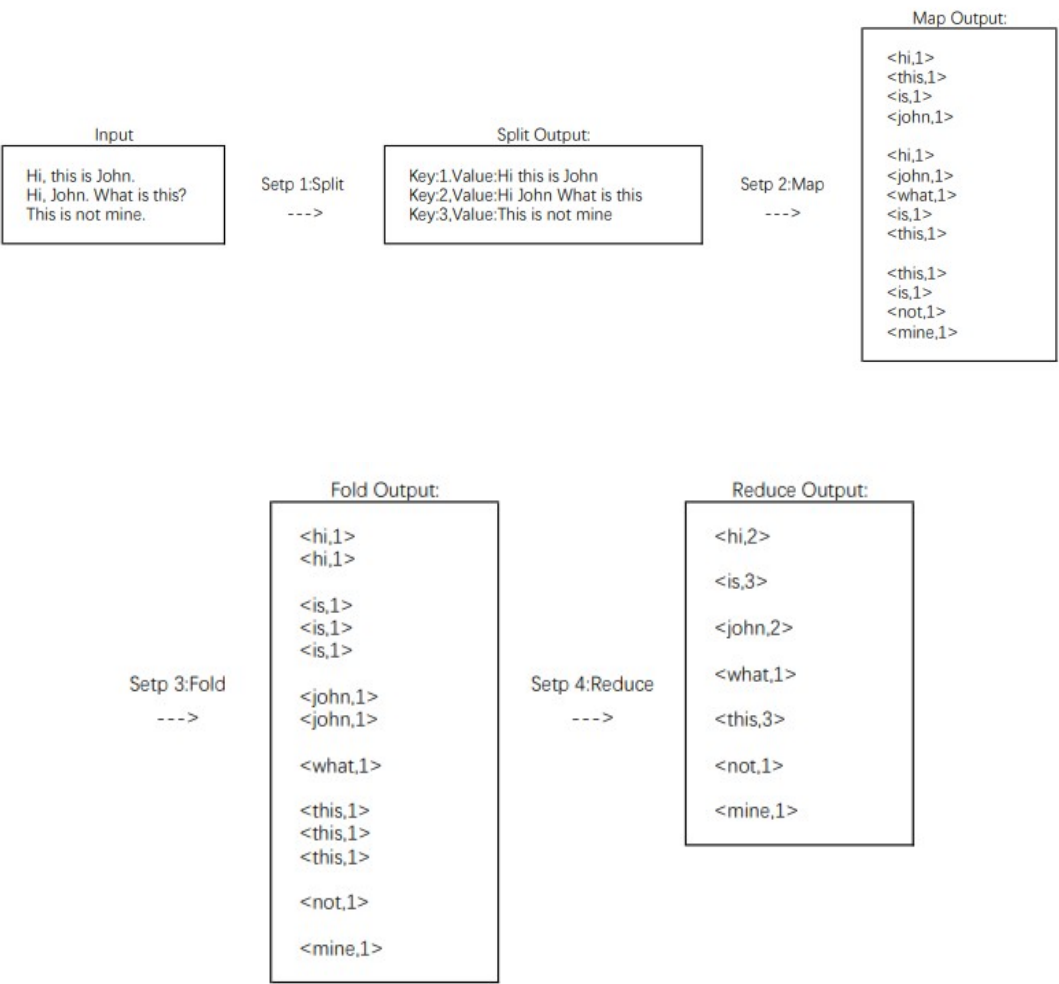
MapReduce 框架使用了框架结构，将分布式计算简化成任务分配和结果的整理收集，只需要按照格式填写 Map 和 Reduce 函数相关内容即可，无需关心如何调度分配，负载均衡以及异常处理等操作。

五、设有三台 **Mapper** 机器和三台 **Reducer** 机器，给定一个文件有如下三行，画简单图表描述 **MapReduce** 架构如何为每个单词计数,不计大小写不包括标点符号。参考课程资料中"**MapReduce** 过程图"，给出每个环节的输入和输出内容 (下方空白如不够可写在试卷背面)（12 分）

Hi, this is John.

Hi, John. What is this?

This is not mine.



六、亚马逊云计算 AWS 提供哪些服务？并简述每种服务的特点。（14 分）

1. 弹性计算云 EC2：虚拟主机服务，可以根据用户的需要调节区域，配置，带宽。可以从映像新建实例，也可以将实例保存成映像。挂载数据盘 EBS 实现存储。
2. 简单存储服务 S3：以对象保存文件，可以通过网络提供用户对数据的操作，文件储存在桶中，使用键值对的 nosql 数据库储存文件元素信息。
3. 简单队列服务 SQS：解决低耦合系统间的通信问题，支持分布式计算机系统之间的工作流。将任务以队列形式进行调度分配，充分调动计算资源，提升效率，系统间消息沟通灵活，资源分配合理。
4. 简单数据库服务 SimpleDB：结构化数据库，使用树状结构储存数据，无需预定义模式，一个属性有多个值，支持自动索引，但是不支持复杂的联合查询，没有事务操作，存储类型过于单一。

七、微软的 Azure 云计算平台提供哪些服务？简述每种服务的特点(15 分)

1. 虚拟机。提供 IAAS 服务，用户可以新建 Linux 和 Windows 镜像的虚拟机。在上面部署开源或者微软自己的企业应用程序；选择公有云，私有云或者混合云满足企业不同所需。同时可以根据需求进行弹性云进行扩展，满足不同情况下所需。
2. Windows 虚拟桌面。在 Azure 上部署含有 Windows 和 Office 软件的 Windows 虚拟桌面环境，通过 Azure 简化管理和部署体验，提供多会话支持。
3. APP 部署服务（Web 应用程序部署服务）。提供 SAAS 服务，直接使用开发语言构建 Web，移动端，API 接口的应用程序，结合 Visual Studio 的优势以及 Github 开源社区的支持，使得部署更简单。
4. SQL 数据库。Azure 上的关系云数据库服务，提供广泛的 SQL Server 兼容性支持，提供高容量，高性能，最大化收益，提供安全保障，结合 APP 部署服务构建应用。

八、试阐述云计算理论研究有哪些前沿课题，分别解决云计算环境中的什么问题？ (15 分)

1. 云计算的网络安全。隐私泄露，网络安全一直算热门话题，未来在大数据，物联网时代更是如此，我们不希望个人信息随着网络爆炸一样传播，被人用来贩卖赚钱，因此我们必须重视这个问题，如何进行信息的安全存储传输以及处理，如何保证资料不外泄，不被过度滥用一直也是研究的重点。

2. 云计算的设计，资源调度分配，异常的处理，数据的存储，性能的优化，成本的分析。任何投资之前，都需要进行风险评估，收益的计算，云计算算是科技发展的产物之一，同样也离不开这些研究，只有充分利用计算资源，存储更多的信息，更少的避免异常，价格更低才会更有竞争力。

3. 云计算与其他方向结合。将云计算的优点运用于其他学科的研究，如何结合现有平台和框架，能够发挥更好的应用。

九、请提供一个设置 **VirtualBox** 虚拟机名字的实验截图，虚拟机名字是你自己的名字。

