

摘要

在过去的几年中，云计算一直备受关注，作为一种新的计算模式提供灵活的和按需的基础设施，平台和软件服务。在云计算领域，搜索云服务需要因续签了一些的云服务独有的特性而受到了挑战，比如动态，多样的服务在不同级别提供，以及缺乏标准化的描述语言。在本文中，通过搜索引擎的有效的云服务元数据解析超过五十万的可能联系之后，我们提出我们用来收集云服务的履带式引擎。

基于所收集的数据，我们进行了一系列的统计分析，并提出了本文的研究结果。这些统计结果，并提供对云服务的当前状态的总体视图。从我们调查的一些最有趣的网络 **ndings** 包括：i) 在云计算标准化的稀缺性，以及 ii) 从云计算的大力支持，很少有建立面向服务的计算（SOC）技术。

我们的网络 **ndings** 提供完善的云的一些进一步的见解服务发现和从这项研究收集数据集将是研究社会价值。

关键词：云服务，服务发现，履带式，WSDL，WADL

第一节

在过去的几年中，云计算正在获得相当大的动力，为移动服务提供了新的计算模式。随着云计算，用户享受的选项，他们的服务在一个大的资源池的网络之中，几乎没有资本投资和适度的运营成本部署[1]，[2]，[3]。尽管大量研究工作发生在寻址以及在云计算诸如数据处理和迁移[4]，[5]，[6]，知识管理[7]，无障碍各种挑战[8]，以及安全性和隐私中[9]，[10]，云服务被发现仍然在很大程度上仍然是一个不变区[2]。

事实上，随着云计算，服务发现的挑战，由于需要将被更新为多种原因。首先，云服务是在不同级别提供，不仅提供数据或业务逻辑，也有提供必要的基础设施的能力。其次，缺乏描述和发布云服务的标准。不同于使用标准语言，如 **Web** 服务描述语言（WSDL）和统一服务描述语言（USDL）揭露他们的接口和通用描述 **Web** 服务，发现和集成（UDDI）发布他们的服务，以服务“登记的发现时，多数的公开提供云服务不是基于描述标准[11]，[2] 这使得云服务发现一个具有挑战性的问题。例如，一些公开可用的云服务不提“云”在所有开发端（如 **Dropbox** 的）。而另一方面，一些商家什么都没有做与云计算有关的内容（例如，**cloud9carwash1**）也可以使用云描述其名称或服务。围绕云服务中心发现几个研究的问题如下所示：

1. Q1：如何识别网络上的服务是否是云服务？

2. Q2：许多云服务如何当前可用的网络上，是谁提供的这些服务？（即是云服务只能由主要厂商如微软，IBM，亚马逊，谷歌等提供的）？

3. Q3: 什么样的云服务提供商是否有在网络上显示?

4. Q4: 在何种程度上建立面向服务的计算 (SOC) 的标准有助于云计算?

5. Q5: 有没有公开可用的云服务数据, 研究人员可以在其云计算研究使用?

我们认为, 有必要确定, 收集和分析目前在 Web 上提供云服务, 这将有助于我们了解云服务的当前状态, 并获得在该地区未来的技术趋势有价值的见解。在本文中, 我们描述了我们一个履带式云服务引擎的设计。这种履带式引擎收集网络上真正的云服务解析而不是从搜索引擎的五十万可能存在的联系。基于所收集的信息, 我们对云服务的当前状态执行一组统计分析。据我们所知, 我们是最先进行这样一个全面的, 大规模的数据收集和分析云服务。从而使得我们的工作结果将为研究界带来显著的好处。我们的主要贡献有以下几点:

我们开发抓取搜索引擎收集网页上提供的云服务信息云服务履带式引擎。我们的抓取引擎有能力收集, 验证和分类云服务。通过在网络上连续地抓取资源, 也能够维持有效和高效的云服务发现向上最新云服务储存库。

从收集到的信息, 我们准备的现实世界的云服务几个大型数据集, 并会释放他们的研究 community²。这些数据集包括近 6000 个云服务。

我们目前收集到的数据, 包括许多云服务是如何访问和功能全面的统计分析, 以及 SOC 和云计算之间的关系。

本文的其余部分安排如下: 第二节概述了相关工作。第三节提出了云服务履带式引擎的设计。第四节讨论了履带式云服务的一些主要挑战。第五节提出了全面的分析和所收集的云服务的统计信息。最后, 第六部分讨论了云服务发现未来的研究方向, 并提供了一些结论性意见。

第二节

服务发现被认为是许多研究领域中的一个, 如无处不在的计算, 移动 Ad-hoc 网络的基本方法之一, 对等网络 (P2P) 和面向服务的计算[12], [2], [13][14]。虽然服务发现是一个非常活跃的研究领域, 特别是在过去十年中的 Web 服务, 云服务, 挑战, 需要重新考虑, 并进行有效的云服务发现解决方案是非常有限的[2], [15]。

一些研究人员建议使用本体技术, 云服务的发现。例如, 康和 Sim[15]提出了一种云服务发现系统 (CSDS), 它利用本体技术来找到那些更接近云服务消费者需求的云服务。特别是, 作者提出一个云本体哪里剂用于执行若干推理方法, 如相似的推理, 等效推理和数字推理。我们的工作是这样工作的补充, 因为他们的策略可以被修改为在更大的环境 (即, 万维网) 的使用。此外, 我们的工作在使用, 几个关系, 即定义了不同的云服务本体的贡献是, 一个是 - 不一, 有助于过滤掉噪音数据, 提高在搜索结果的准确度 (云服务的本体是在第三节中详细描述)。

其他研究人员建议使用更好的发现和云服务的负载均衡分布式哈希表（DHT 方法）。例如，Ranjan 等。 [16]提出云等体的延伸 DHT 叠加以支持多维范围查询（即，尺寸可包括服务类型，处理器速度，可用存储器等）用于服务发现的索引和匹配的概念。所提出的方法进行验证的公共云计算平台（Amazon EC2 上）。他们的工作主要集中在一个封闭的环境。相比之下，我们重点放在一个开放的环境（即万维网）发现云服务允许用户搜索云服务，适合他们的需求。

发现 Web 服务已经有一些很好的结果的活跃的研究领域。马斯里和马哈茂德·[13]通过抓取 UDDI 注册中心业务（UBRS），以及搜索引擎如谷歌，雅虎和百度收集 WSDL 文档。作者分析所收集的数据并呈现关于 Web 服务的一些详细的统计信息，例如活性与非活性的 Web 服务和对象的大小分布。 Li 等人。 [17]也通过谷歌 API 收集 Web 服务的数据，并提出相关的操作，大小，文字的 Web 服务的分布和功能多样性的一些有趣的统计信息。最近，客福有限公司等。 [18]对 RESTful Web 服务的现状报告了他们的发现。作者使用 17 种不同的 RESTful 服务的设计标准（例如，形式化描述，在陈述环节，资源类型的数量等情况而定）来分析 programmableWeb3 列出的前 20 RESTful 服务。不同于以往的作品，其中 Web 服务通过简单的收集接口文档（例如，WSDL 文件）和搜索 UDDI 企业注册，发现云服务发现拥有更多的挑战，如缺乏规范的描述语言的云 服务，这需要充分考虑的。

第三节

在本节中，我们将介绍云服务履带式引擎，基于 NIST [19]开发的云计算标准的云服务本体，并对收集的数据集进行架构设计。

欧安会架构

图 1 描绘了云服务检索器引擎，其包括六个层的主要成分：

云服务提供商层

这一层（如图 1 所示的右上部分）由不同的云服务提供商谁公开提供的云服务上的。这些云服务是通过网络门户访问和索引的搜索引擎如谷歌，雅虎和百度。

云服务本体层

这一层负责维护云服务本体。云服务本体包含一系列的概念和关系，使履带自动发现，验证和分类云服务。本体的维护是通过一个叫做本体更新模块来完成。在云服务本体的更多信息可以在第三节中。

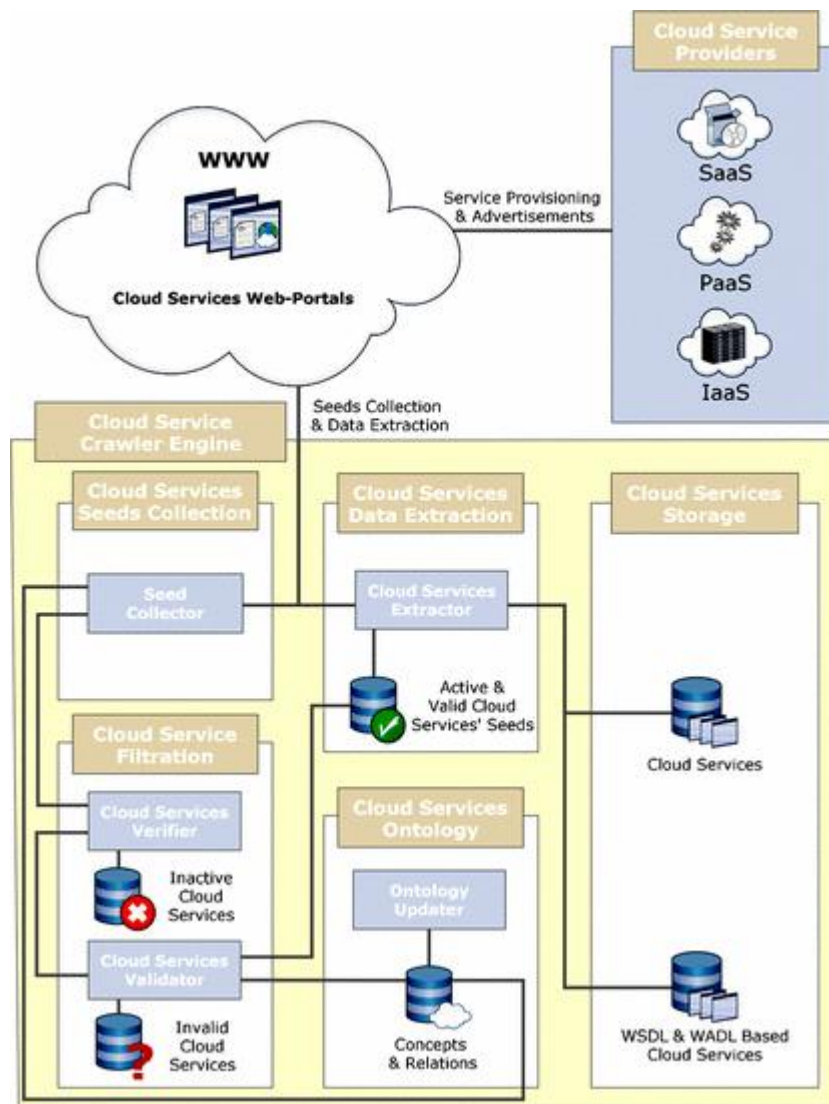


图 1 云服务履带式引擎的架构

云服务种子采集层

这一层负责收集可能的云服务的种子（即云服务的 URL）。我们开发了考虑在搜索引擎中的几个可能的资源如索引网页以及 WSDL 和 WADL（Web 应用程序描述语言）4 文档种子收集器模块。对于数据的采集，种子收集器使用的一些概念在云服务本体的前几个级别作为关键字来收集数据（例如，云服务的 IaaS，PaaS 的来说，SaaS 等），其中收集的种子，然后发送到云服务过滤层进行验证（例如，来自云服务本体知识用于推理）。

云服务过滤层

这一层负责从种子收集器收集过滤云服务“的种子。云服务验证首先确定云服务的种子是否是活动还是非活动之一。无效的种子保存在另一个检查（即，一些不活跃的种子可能是暂时不可用）和错误代码还捕获了无效云服务数据库。活性种子通过使用从云服务本体的概念传递给云服务验证进行验证。例如，如果种子的 Web 页包含了云服务，例如 IaaS 的，存储，基础设施等相关的概念，然后将种子被认为是一个有效的。然而，如果种子的 Web 页所包含的其他概念，如新闻，文章，论文，天气，等等，则该种子被认为是无效的（可以在

第 III-B 中找到有关概念的更多细节）。这意味着，收集种子可能是一个新闻网站发布关于云服务的文章。无效种子保持在无效云服务数据库和被传递到云服务数据提取用于进一步处理之前有效种子被分类（即，IaaS 的，PaaS 的或 SaaS）的。

算法 1 云服务发现算法

Step 1. Ontology:
The *Ontology Updater* updates the concepts and relations in the *Concepts & Relations* database.

Step 2. Discovery:
The *Seed Collector* (SC) gets the (*is-a*) relations from the *Concepts & Relations* database
SC collects possible seeds for cloud services
SC passes the collected seeds, S_c to the *Cloud Services Verifier*.

Step 3. Verification:
/*The *Cloud Services Verifier* determines whether a cloud service seed is active or inactive*/
foreach collected cloud service seed $s \in S_c$ **do**
 if s is inactive **then**
 - capture error code
 - store s in the Inactive Cloud Services database
 else
 - pass active seed s_a to the *Cloud Services Validator*
 end if
end foreach

Step 4. Validation:
The *Cloud Services Validator* gets the (*is-not-a/is-a*) relations from the *Concepts & Relations* database.
/*The *Cloud Services Validator* determines whether a cloud service seed is valid or invalid using *is-not-a* concepts*/
foreach active cloud service seed s_a **do**
 if cloud service seed s_a is invalid **then**
 - store s_a in the Invalid Cloud Services database
 else
 - categorize valid seed s_{av} using *is-a* concepts
 - pass s_{av} to the Active & Valid Cloud Services Seeds database
 end if
end foreach

Step 5. Data Extraction & Storage:
/*The *Cloud Services Extractor* extracts data from the valid seed s_{av} and store the data in the corresponding database*/
foreach s_{av} **do**
 Store data of s_{av} in the Cloud Services database.
 if s_{av} is WSDL or WADL **then**
 - store s_{av} 's data in the WSDL & WADL Cloud Services database.
 end if
end foreach
Go to Step 1

云服务的数据提取层

这一层负责提取主动和有效的云服务的信息。的云服务数据存储在云中的服务存储层用于进一步分析的相应数据库。通过云服务履带式引擎（欧安会）为云服务使用爬行整体算法总结了算法 1。

云服务本体

云服务本体（CSO）便于与元信息的检索器引擎和描述的云服务，这是在这个意义上至关重要的云服务可不必在使用标识词（例如，云，基础设施，平台和软件）数据语义其名称和描述。

当开发 CSO，我们认为由 NIST [19]开发的云计算标准。根据 NIST，云服务基础上建立五个基本特征，即：i) 按需自助服务，其中的云服务消费者能够自动提供计算资源，而不需要人的交互与每个云服务提供商，ii) 宽网络接入其中，云服务消费者可以通过网络访问可用的计算资源，iii) 资源池，其中的计算资源汇集到服务基于多租户模式的多个云服务的消费者，其中物理和虚拟计算资源是动态的按需重新分配，iv) 快速弹性，其中的计算资源弹性调配基于云服务消费者需要迅速扩大规模，和 v) 测量服务，在那里计算资源的使用情况进行监控，计量（即使用现收现付你去的机制），控制和据报道，为云服务提供商和消费者提供透明度。

云服务有三种不同的模式，包括基础设施作为基础云服务提供商和云服务之间不同的服务级别协议（SLA），服务（IaaS），平台即服务（PaaS）和软件即服务（SaaS）消费者[20]，[21]，[19]。

基础设施即服务（IaaS）。这种模式代表了云计算环境下，云服务消费者可以租储存，加工，并通过云提供商提供虚拟机通信的基础部分（例如，亚马逊的弹性计算云（EC2）5）。在此模型中，云服务提供商控制和管理底层的云环境，而云服务消费者拥有他/她的虚拟机，它包括存储，处理，甚至可以选择用于通信的一些网络部件的控制。

平台即服务（PaaS）。这种模式代表了云环境的集成部分，驻留在的 IaaS 层之上，支持的系统集成和虚拟化中间件。该允许的 PaaS 云服务消费者培养他/她自己的软件，其中云服务提供商规定的软件开发工具和编程语言（例如，谷歌 App6）。在 PaaS 的，云服务的消费者拥有底层的云基础设施没有控制（例如，存储网络，操作系统等），但拥有对部署的应用程序的控制。

软件作为服务（SaaS）。这种模式代表了云环境的应用程序的一部分，驻留在 PaaS 的层之上，支持远程访问，其中云服务的消费者可以远程访问存储在底层的云基础设施的数据，使用由云服务提供商提供的应用程序（例如，谷歌 Docs7）。同样地，在 SaaS 的，消费者拥有底层的云基础设施无控制（例如，存储网络，操作系统等），但有超过他/她的数据控制。

我们的 CSO 包含了一组概念和概念之间的关系，使云服务履带式引擎自动发现，确认和网上分类云服务。这些概念使云服务履带式引擎收录尽可能云服务'的种子，然后过滤掉无效的种子（见算法 1）。CSO 定义了两个不同的关系，即 IS-a 和是 - 不是一个。限于篇幅，我们只描绘了图 2 中的本体的一部分种子收集器使用的概念在云服务本体的前几个级别作为关键字来收集数据（例如，云服务的 IaaS，PaaS 的来说，SaaS 等）是与相关的是，一个关系（参见图 2（a）），从搜索引擎收集可能的云服务“的种子。在另一方面，云服务验证程序使用了与是 - 不是一个关系相关的概念（参见图 2（b））的云服务验证。例如，如果种子的 Web 页包含与相关的概念是，不-一个关系如新闻，文章，论文，天气，等等，这意味着，所收集的种子可以是一个新闻网站发布关于云服务的文章并应被视为无效种子。最后，云服务验证程序使用了与（IS-A）进行分类有效的云服务要么 IaaS 的，PaaS 的，SaaS 的，或者这些模型的组合关系相关的概念。

收集数据集

搜索引擎是主要来源，收集的云服务，因为我们不知道有任何业务注册（如 UDDI 注册中心为 Web 服务）的云服务。即使对于 Web 服务，这是众所周知的 UDDI 企业注册的想法是失败和 Web 服务的发现朝向搜索引擎[22]移动[13]。

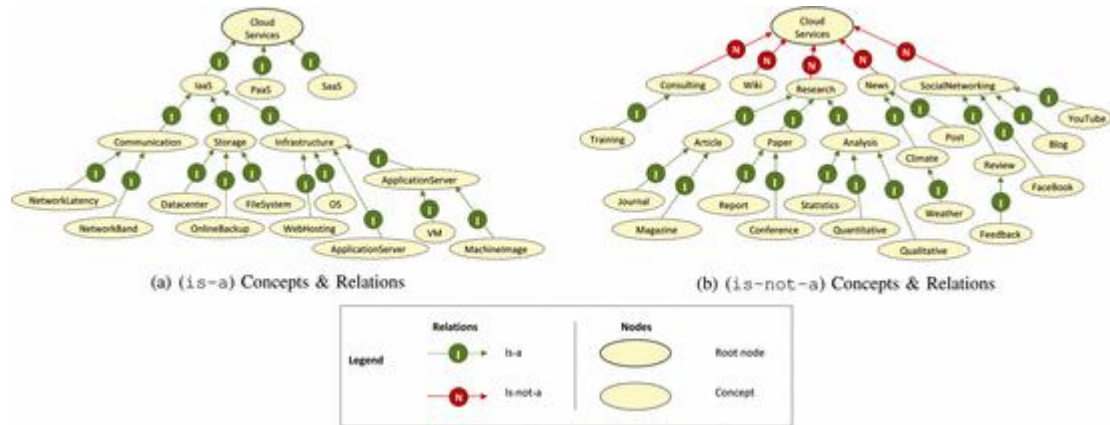


图 2 云服务本体的一部分

我们的云服务履带式引擎探讨在 Web 上使用现有的搜索引擎“的 API（例如，谷歌，雅虎和百度）的云服务。我们基本上选择云服务本体的前几个级别作为关键词，收集诸如云服务，IaaS 的，PaaS 的来说，SaaS，通信，存储，基础架构，在线备份，虚拟主机，虚拟桌面，虚拟机，软件，API 数据，等通过不断分析从搜索引擎返回的搜索索引结果可能收集云服务“种子履带。云服务履带式发动机成功地解析可能联系和发现了云服务可能的种子。应当指出的是，发现了云服务“的网页是充满噪声数据（例如，维基新闻，与相关的云服务的文章）被过滤这就需要的。过滤过程后，云服务履带式引擎识别独特的云服务，这是保存在云服务的数据库（见图 1）。

其中一个在我们的脑海重要的调查是要找出面向服务的计算(SOC)如何促成了云计算。为此，我们特别有兴趣了解多少云服务都采用 SOC 的标准，如 WSDL 得到落实。为了收集这些数据，履带引擎配置为搜索文件扩展名为诸如 WSDL 和 WADL。我们的云服务履带式引擎收集在 WSDL 和 WADL 云服务的可能联系。过滤过程后，我们的云服务履带式引擎识别使用 WSDL 和 WADL 实施有效的云服务。

第四节

自动发现，验证，确认，云服务的收集和分类并不是一个简单的任务。为了收集所需的数据集，这样的云服务履带式发动机的设计必须克服一些挑战，包括以下内容：

云服务的动态行为

云服务是在新的云服务出现在网络上，而老云服务可能全天候停止感动态。此外，云服务可能随时间发生变化。其结果，履带式发动机需要能够更新或为了保持软件库的最新定期重新云服务。它必须被重新（例如，新的云服务和那些经常更新）云服务和那些可以跳过（例

如，停产云服务或者那些不太更新版）区别开来是很重要的。由于网络上大量可能的云服务，能够将它们分成不同的组会大大提高性能爬行。

缺乏标准化

不像可以通过简单的收集 WSDL 文档或搜索 UDDI 注册中心，云服务缺乏规范的描述语言，这使得它们更难发现的被发现 Web 服务。这实际上是云服务发现的最显著的挑战。云服务履带式引擎可能会收集不必要的和嘈杂的数据（例如，维基，新闻，博客，评论，以及相关的研究论文云）。例如，我们的云服务履带式引擎最初收集网络上 36538 个可能的种子，只有 16.14% 的种子，实际上是云服务。

可扩展性

可用在网络上可大且日益增长的云服务，这可能严重影响云服务履带式发动机的性能的数量。因此，云服务履带式引擎的多个实例需要在以优化性能爬行从多台机器同时运行。显然，有必要协调这些多个实例，以避免多余的工作。

爬行堵塞

有高性能和资源消耗之间的明确的权衡。爬行搜索引擎和一些网站可能消耗的其他组织，这可能会导致访问这些服务被阻止的云服务检索器的资源。因此，云服务检索器引擎需要，以提高其性能，而不消耗服务提供商的资源并最小化网络上的负荷（例如，使用从不同位置的履带的几个实例）。

第五节

我们已经进行了云服务所收集的数据进行综合，统计分析，从许多不同的方面。这些结果也提供了答案，由于篇幅限制，我们在第一节中提出的问题，在本节中，我们将介绍其中的一些发现。

云服务标识

为了优化抓取性能，我们使用了云服务检索器引擎的三个不同的实例来从三个不同的机器上同时运行。表 1 描绘了云服务收集和验证结果的崩溃。

表 1 云服务解析结果

Error Code	Description	Percentage
101	The connection was reset.	13.66%
105	Unable to resolve the server's DNS address.	1.64%
107	SSL protocol error.	0.27%
118	The operation timed out.	0.27%
324	The server closed the connection without sending any data.	0.27%
330	Content decoding failed.	0.27%
400	Bad request.	0.82%
403	Access denied.	3.83%
404	The requested URL / was not found on this server.	10.11%
500	Server Error.	1.37%
503	The service is unavailable.	0.27%
504	Page not found.	0.27%
1005	URL does not exist.	66.95%
Total	-	100%

从表 I，我们可以看到，在收集过程中存在噪声数据的一个显著一部分。解析链接后可能获得 35,601 个种子，种子中 29189（超过 80%）是无效的。这在很大程度上贡献的事实，那就是缺乏描述和发布云服务的标准。因此，存在对云服务，例如接口和发现标准化的迫切需要。

还值得一提的是，不活动的云服务的总数量是显著低（仅 423，总的可能的种子为约 1.19%）。这并不奇怪，因为我们的抓取工具从搜索引擎，检查过期的链接，并从它们的索引中排除它们收集的数据。对于那些不活跃的云服务，我们的抓取工具也根据 W3C8 的 RFC 2616 状态代码定义，表二所示捕获的错误代码。从表中我们可以看到，比例最高（66.95%），进入错误代码“1005”（即网址不存在），这意味着大多数不活跃的云服务都停产了。

表 2 现有云服务的错误代码

Error Code	Description	Percentage
101	The connection was reset.	13.66%
105	Unable to resolve the server's DNS address.	1.64%
107	SSL protocol error.	0.27%
118	The operation timed out.	0.27%
324	The server closed the connection without sending any data.	0.27%
330	Content decoding failed.	0.27%
400	Bad request.	0.82%
403	Access denied.	3.83%
404	The requested URL / was not found on this server.	10.11%
500	Server Error.	1.37%
503	The service is unavailable.	0.27%
504	Page not found.	0.27%
1005	URL does not exist.	66.95%
Total	-	100%

云服务提供商分类

云服务广泛分为 IAAS, PaaS 和 SaaS 的, 由不同的云服务供应商提供。这将是有趣的, 看看不同类型的云服务提供商的百分比。如前所述, 经过我们的云服务履带完成云服务的种子的审定, 履带也归类到的 IaaS, PaaS 的或 SaaS 的这些云服务通过推理概念之间的云服务本体的更多详细信息在云服务本体的关系 (解释有关第三节)。

图 3 描述了分类的结果, 其中云服务提供商被分为六个不同的种类 (即, LAAS, PaaS 的来说, SaaS, IaaS 的+ PaaS, IaaS 的+的 SaaS, PaaS 的+ SaaS 的, 并且全部)。从图 3 中, 我们可以看到, 在云服务提供商相当程度的各种。特别是, 超过了云服务提供商的一半 (52.29%), 专注于提供的 IaaS 服务, 有近三分之一 (27.08%), 专注于提供 SaaS 服务, 以及 7.70% 专注于提供的 PaaS 服务。其余的 12.93% 提供一个以上的云服务模型 (即 IaaS 和 PaaS 的, IaaS 和 SaaS 的, PaaS 和 SaaS 的或所有云服务模型)。应当指出, 这些主要的球员, 如微软, 亚马逊, 谷歌等属于这一部分。

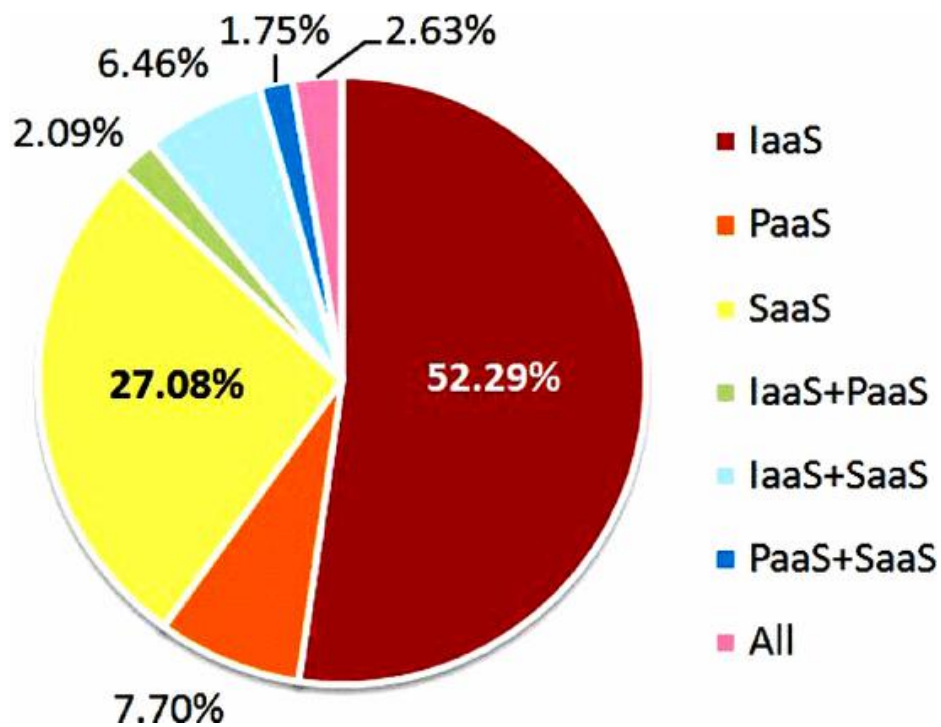


图 3 云服务提供商分类

云计算和 SOC

面向服务的计算 (SOC) 和 Web 服务被认为是最重要的使能技术云计算[23]中的一个, [2], [24]。这将是有趣的探讨在云计算利用 SOC 的当前状态。这是很难做到的关于这一主题的完整的研究。在本文中, 我们进行的基础上, 我们收集到的信息的一些初步的研究。结果仍然应该揭示了一些有用的信息。

我们首先调查多大从 SOC 的描述语言如 WSDL 已经用于云服务。要做到这一点, 我们比较有 WSDL (用于基于 SOAP 的 Web 服务) 或 Web 应用程序描述语言 (WADL, 对于

RESTful Web 服务) 的文档云服务的数量。图 4 描绘的结果。这是一个很大的惊喜地发现, 只有云服务 (仅 1.80%), 一个非常小的一部分使用 Web 服务接口语言实现的。我们想也强调, 实际上使用 SOC 一些云服务可能不会被我们的抓取工具进行检测。并非所有的 WSDL 文档是 Web [13] 在公开访问。此外, 大多数的 RESTful Web 服务提供没有正式的说明, 并依赖于非正式文件[18]。然而, 非常低的比例仍表示在应用 SOC 到云计算贫困状态。

从我们的分析, 它是清楚的是示出的 SOC 是打在实施云服务的积极作用没有强有力的证据。也有标准化的迫切需要, 以全面拥抱云计算。幸运的是, 这个问题已经得到了社会的实现和一些初步成果已经实现。例如, 分布式管理任务组 (DMTF) 刚刚发布, 于 2012 年 8 月 29 日, 在云基础设施管理接口 (CIMI) 规范, 标准化的云环境之间的相互作用来实现互操作的云计算基础架构的管理[25]。

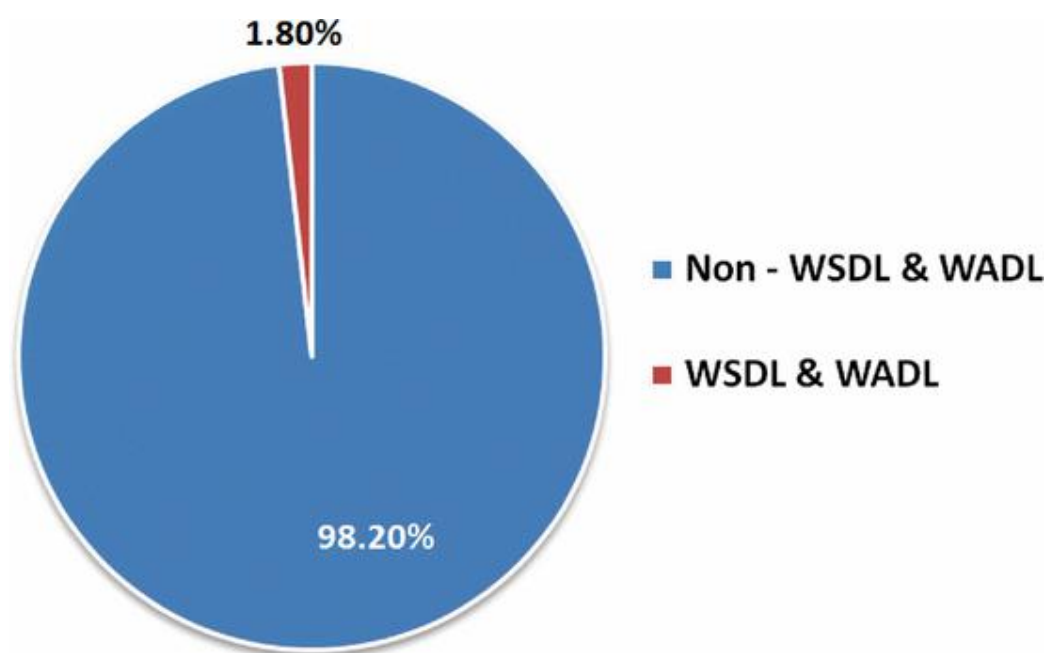


图 4 WSDL & WADL 云服务 VS 非 WSDL & WADL 云服务

第六节

随着越来越多地采用云计算, 有效地寻找客户相关的云服务正在成为一个重要的研究课题。在本文中, 我们已经进行了云计算服务目前在 Web 上提供了全面的调查。我们已经制定了收集, 验证和分类云服务的云服务履带式引擎, 并产生了许多存储所收集的云服务的信息数据集。根据收集到的信息, 我们已经进行了一些统计分析, 包括对云供应商分类的分布和云计算和 SOC 之间的关系。这些结果提供了对云服务的当前状态的总体视图。最有趣的发现是这样的事实, 没有强有力的证据表明 SOC 起着使云计算作为一种技术显著的作用。更多的调查应该按顺序进行, 以了解为什么是这种情况, 以及如何使 SOC, 以更好地促进云计算, 这样从研究和开发社区以前的努力不会白费。此外, 在目前的云产品和服务缺乏标准化, 这使得云服务发现一个非常艰巨的任务。

我们的云服务履带式解析云服务 619475 可能联系公众可在网站上，并最终确定和分类 5883 有效的云服务。据我们所知，这是在发现，收集，并在 WWW 规模分析云服务的第一次努力。所收集的数据集，这是被释放，将要研究界带来显著效益。

我们正在进行的研究包括通过发现使用 SOC 技术实现云服务更多的证据进一步调查 SOC 和云计算之间的关系。我们还计划扩大云服务履带式引擎收集到的云服务，这将有助于实现居云服务的任务执行更全面的 QoS 指标。

致谢

Talal H. Noor 和 Abdullah Alfazi 的工作由沙特阿拉伯王国的教育部高等教育事务处支持。

脚注

- 1.<http://www.cloud9carwash.com/>
- 2.有兴趣的读者都提到了项目网站，了解详情(<http://cs.adelaide.edu.au/~cloudarmor>)
- 3.<http://www.programmableweb.com/>
- 4.WADL 是具象状态传输 (REST 的)，可以用来描述 RESTful Web 服务的 WSDL 的等价物。统一服务描述语言 (USDL) 在我们的云服务发现最初尝试是排除由于结果不足。
- 5.<http://aws.amazon.com/ec2>
- 6.<http://www.google.com/apps>
- 7.<https://docs.google.com>
- 8.<http://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html>
- 9.这意味着，云服务供应商提供了两种类型的云服务：IaaS 的都和 PaaS 服务。

参考文献

1. M. Armbrust, A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A View of Cloud Computing," Communications of the ACM, vol. 53, no. 4, pp. 50-58, 2010.

2. Y. Wei and M. B. Blake, "Service-oriented Computing and Cloud Computing: Challenges and Opportunities," *Internet Computing*, IEEE, vol. 14, no. 6, pp. 72-75, 2010.
3. T. H. Noor, Q. Z. Sheng, S. Zeadally, and J. Yu, "Trust Management of Services in Cloud Environments: Obstacles and Solutions," *ACM Computing Surveys (CSUR)*, 2013, in press.
4. M. Hausenblas, R. Grossman, A. Harth, and C.-M. P., "Largescale Linked Data Processing - Cloud Computing to the Rescue?" in *Proc. of the 2nd Int. Conf. on Cloud Computing and Services Science (CLOSER'12)*, Porto, Portugal, Apr 2012.
5. M. Menzel and R. Ranjan, "CloudGenius: Decision Support for Web Server Cloud Migration," in *Proc. of the 21st Int. Conf. on World Wide Web (WWW'12)*, Lyon, France, Apr 2012.
6. H. L. Truong, S. Dustdar, and K. Bhattacharya, "Programming Hybrid Services in the Cloud," in *Proc. of the 10th Int. Conf. on Service-Oriented Computing (ICSOC'12)*, Shanghai, China, Nov 2012.
7. C. Fehling, T. Ewald, F. Leymann, M. Pauly, J. Rutschlin, and D. Schumm, "Capturing Cloud Computing Knowledge and Experience in Patterns," in *Proc. of IEEE 5th Int. Conf. on Cloud Computing (CLOUD'12)*, Honolulu, Hawaii, USA, Jun 2012.
8. B. White, "Accessibility Challenges of the Next Decade: Cloud and Mobile Computing and Beyond," in *Proc. of the 8th Int. Cross-Disciplinary Conf. on Web Accessibility*, Hyderabad, Andhra Pradesh, India, Mar 2011.
9. K. Ren, C. Wang, and Q. Wang, "Security Challenges for the Public Cloud," *IEEE Internet Computing*, vol. 16, no. 1, pp. 69-73, 2012.
10. S. Pearson and A. Benameur, "Privacy, Security and Trust Issues Arising From Cloud Computing," in *Proc. IEEE 2nd Int. Conf. on Cloud Computing Technology and Science (CloudCom'10)*, Indianapolis, Indiana, USA, Nov - Dec 2010.
11. P. Thibodeau, "Frustrations with Cloud Computing Mount," 2010, accessed: 25/11/2012, Available at: <http://www.computerworld.com/s/article/9175102/>.
12. E. Meshkova, J. Riihijärvi, M. Petrova, and P. Mähönen, "A Survey on Resource Discovery Mechanisms, Peer-to-Peer and Service Discovery Frameworks," *Computer networks*, vol. 52, no. 11, pp. 2097-2128, 2008.
13. E. Al-Masri and Q. Mahmoud, "Investigating Web Services on the World Wide Web," in *Proc. of the 17th Int. Conf. on World Wide Web (WWW'08)*, Beijing, China, Apr 2008.
14. A. Mian, R. Baldoni, and R. Beraldi, "A Survey of Service Discovery Protocols in Mobile Ad Hoc Networks," *Pervasive Computing*, IEEE, vol. 8, no. 1, pp. 66-74, 2009.
15. J. Kang and K. M. Sim, "Towards Agents and Ontology for Cloud Service Discovery," in *Proc.*

of the 2011 Int. Conf. on Cyber-Enabled Distributed Computing and Knowledge Discovery, Beijing, China, Oct 2011.

16. R. Ranjan, L. Zhao, X. Wu, A. Liu, A. Quiroz, and M. Parashar, Cloud Computing: Principles, Systems and Applications. Springer, 2010, ch. Peer-to-Peer Cloud Provisioning: Service Discovery and Load-Balancing, pp. 195-217.

17. Y. Li, Y. Liu, L. Zhang, G. Li, B. Xie, and J. Sun, "An Exploratory Study of Web Services on the Internet," in Proc. of IEEE Int. Conf. on Web Services (ICWS'07), Salt Lake City, Utah, USA, Jul 2007.

18. D. Renzel, P. Schlebusch, and R. Klamma, "Today's Top "RESTful" Services and Why They are not RESTful," in Proc. of the 13th Int. Conf. on Web Information Systems Engineering (WISE'12), Paphos, Cyprus, Nov 2012.

19. P. Mell and T. Grance, "The NIST Definition of Cloud Computing," Sep 2011, accessed: 05/06/2012, Available at: <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>.

20. I. Brandic, S. Dustdar, T. Anstett, D. Schumm, F. Leymann, and R. Konrad, "Compliant Cloud Computing (C3): Architecture and Language Support for User-Driven Compliance Management in Clouds," in Proc. of IEEE 3rd Int. Conf. on Cloud Computing (CLOUD'10), Miami, Florida, USA, Jul 2010.

21. K. Clark, M. Warnier, F. Brazier, and T. Quillinan, "Secure Monitoring of Service Level Agreements," in Proc. of the 5th Int. Conf. on Availability, Reliability, and Security (ARES'10), Krakow, Poland, Feb 2010.

22. J. Ma, Q. Z. Sheng, K. Liao, Y. Zhang, and A. H. Ngu, "WSFinder: A Framework for Similarity Search of Web Services," in Proc. of the 10th Int. Conf. on Service Oriented Computing (ICSOC'12), Shanghai, China, Nov 2012.

23. T. Dillon, C. Wu, and E. Chang, "Cloud Computing: Issues and Challenges," in Proc. of the 24th IEEE Int. Conf. on Advanced Information Networking and Applications (AINA'10), Perth, Australia, Apr 2010.

24. T. H. Noor and Q. Z. Sheng, "Trust as a Service: A Framework for Trust Management in Cloud Environments," in Proc. of the 12th Int. Conf. on Web and Information Systems (WISE'11), Sydney, Australia, Oct 2011.

25. DMTF.org, "DMTF Releases Specification for Simplifying Cloud Infrastructure Management," August 2012, accessed: 25/11/2012, Available at: <http://dmtof.org/news/pr/2012/8/dmtf-releases-specification-simplifying-cloud-infrastructure-management>.