

云泥殊路

云上环境数据泄露的探讨

by : 姚 威 ID: p0tt1





全球技术领导力峰会

Geekbang | TGO 鲲鹏会
极客邦科技

500+ 高端科技领导者与你一起探讨 技术、管理与商业那些事儿



🕒 2019年6月14-15日 | 📍 上海圣诺亚皇冠假日酒店



扫码了解更多信息

姚威

ID: p0tt1



广州凌晨网络科技有限公司 (LCS)
CEO



腾讯Wi-Fi安全实验室
联合发起人

坐标
广州



[!]"凌晨网络科技是一家以安全大数据为核心能力的科技公司，致力于用大数据为网络安全赋能，用安全为大数据保驾护航"

[!]"腾讯Wi-Fi安全实验室是由凌晨网络LCS及腾讯手机管家/腾讯Wi-Fi管家联建实验室，致力于在Wi-Fi网络空间中将安全大数据落地实践"

目录

CONTENTS

01

云上数据泄露事件回顾

Historical Events

02

云上数据安全的“现实问题”

Realistic Problems

03

深入分析数据泄露手段

The Analysis

04

云上数据安全的建议

The Suggest

00

前言
Intro

先聊一下云上环境的不同理解



入住与入驻之争

住宅论

我在X云地产购置了一套房，所有的东西都属于我，包括门牌号和整个房屋的空间，当然我也有70年产权，我按揭付费，除此之外我还要交水电费和物业费。



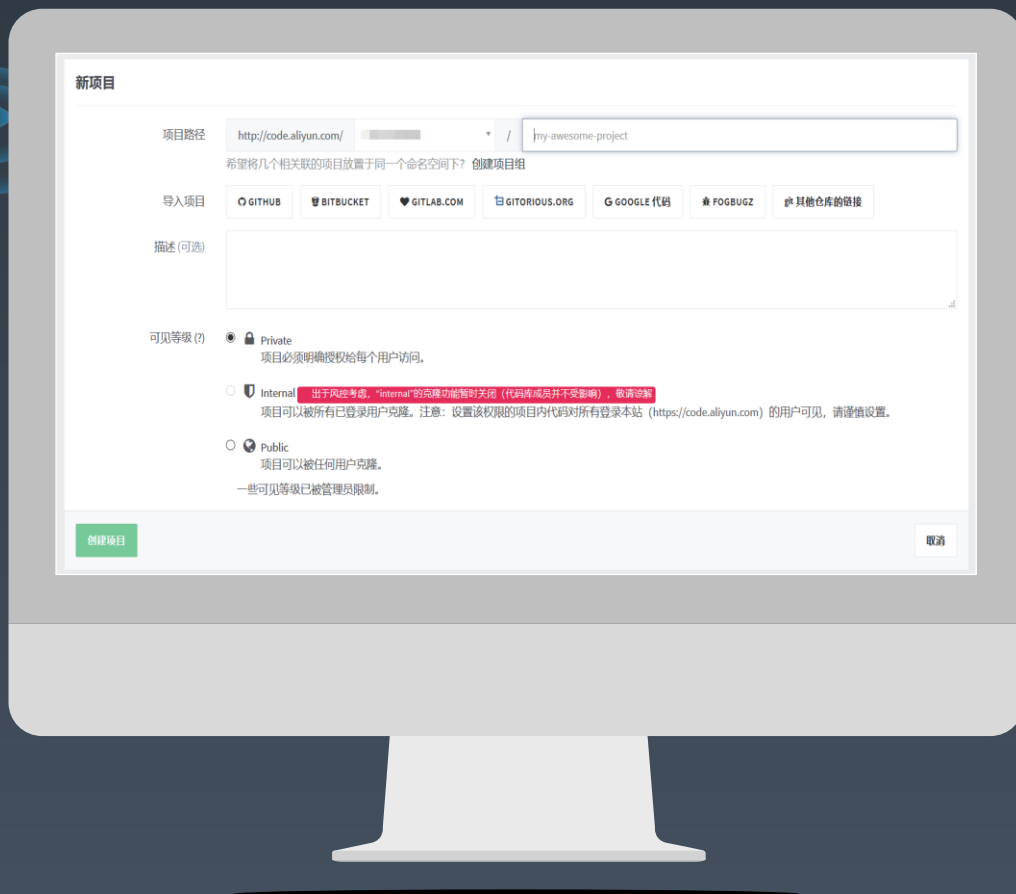
商铺论

我在X云地产租赁了一间商铺，所有的东西都是我带来的当然属于我，我的租金也为我换来了门牌号和整个商铺的空间，我按月交租金，除此之外我还要交水电费和物业费。

01 云上数据泄露 事件回顾

Historical Events

云上数据泄露事件回顾



云用户出现的源代码泄露

由于某云厂商代码托管平台的项目权限设置存在歧义，导致开发者操作失误，造成至少40家以上企业的200多个项目代码泄露，其中涉及到万科集团、咪咕音乐、51信用卡旗下51足迹、百度无人车合作伙伴ecarx等知名企业。

本质原因

- 阿里云代码托管平台code.***yun.com访问权限设置中的“Internal”选项存在理解歧义，Internal怎么理解因人而异：
- Internal：使用某云厂商代码托管平台的人公开
- Internal：企业内部用户公开

产品设计和用户体验使得用户对产品功能理解上存在歧义，导致了不正确的配置，这确实一部分是某云厂商产品的问题，同时更重要的是使用者安全意识的问题。

云上数据泄露事件回顾

某云厂商澄清网络传闻“可重置任意某云服务器root密码”为虚假信息

01.真实情况

某客户自身的管理员AK（access key）泄露所导致的独立事件。

02.典型案例

CodeSpaces的破产的就是因为上云之后AccessKey泄露了，黑客勒索未遂结果彻底删除CodeSpaces的所有数据以及数据备份。

注：AccessKey包括AccessKeyId和AccessKeySecret
AccessKeyId用于标识用户；AccessKeySecret是用来验证用户的密钥



AccessKey公开在Github等代码托管平台上，相当于把取款密码写在银行卡上，然后再把银行卡扔在大街上。

云上数据泄露事件回顾

Amazon S3云存储对象提供公开访问



医疗数据数据泄露

医疗设备公司Patient Home Monitoring的医疗数据存储纪录包含47GB医疗数据文件的Amazon S3 云存储对象提供公开访问，包含多达315,363份PDF档案，涉及近15万患者的姓名、地址、医生和病例纪录以及周常血液检查结果等隐私信息。



美国选民数据泄露

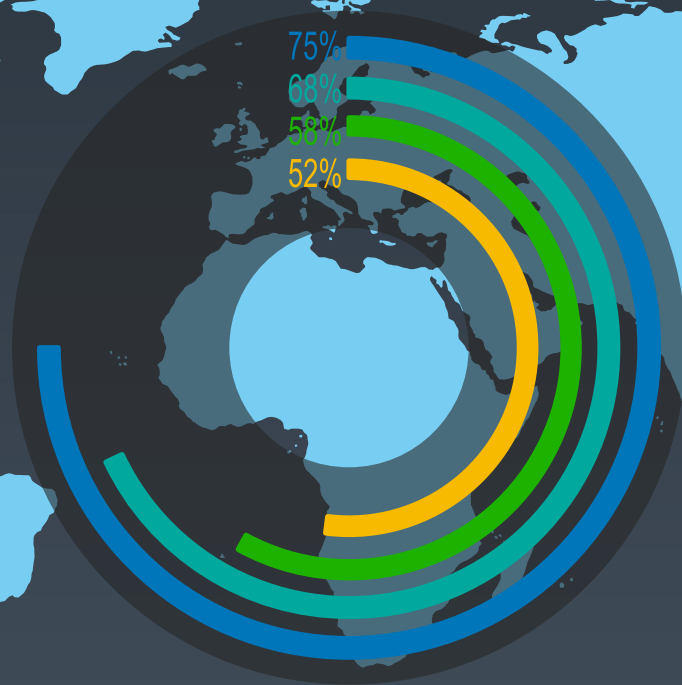
美国共和党全国委员会（RNC）合作的数据分析商Deep Root Analytics、TargetPoint以及Data Trust放在AWS S3的1.1 TB数据发生泄露。其中包含超过1.98亿名美国选民的敏感个人资料，例如姓名、出生日期、住址、电话号码以及选民注册细节信息。



GoDaddy数据泄露

一个不安全的GoDaddy的亚马逊一个不安全的S3 bucket其中包含与31000多个GoDaddy系统相关的敏感信息，被泄露的信息包括主机名、操作系统、工作负载、AWS区域、内存和CPU规格等字段。

云上数据泄露事件回顾



安全问题

信息泄露 (GitHub、Gitee、Blog)
脆弱安全设置 (弱密码、安全组问题)、应用缺陷 (应用服务缺陷, 中间件缺陷)

信息泄露 安全设置 应用缺陷 其他

02 云上数据安全“现实问题”

Realistic Problems

云上数据是否安全



数据作为企业的核心资产，云端数据安全的关注程度越来越高，把数据放到云上是否足够安全成为各个企业思考最多的问题。

任何的安全都是相对的，没有绝对的安全

相对安全的云厂商

顶尖的安全专家
专业的安全基础设施
完善的安全保障机制
同样是遭遇攻击，云厂商的安全专家能够更快更好地处理。

不安全的内在因素

数据安全的内在因素是用户自身，暴露在互联网中用户，用户需要为自身的安全做出响应。回顾前面的数据泄露事件，主要问题基本都出现在用户侧。



就像再安全的汽车也需要安全驾驶一样，在云计算行业也是类似。

用户关注云平台的安全问题

+ 01.云厂商的安全认证

- CSA STAR
- 可信云服务认证
- 等级保护
- ISO 27001

+ 02.云厂商的基础安全服务

- 基础安全（DDoS防护/Web应用防火墙/主机安全等）
- 数据安全（数据加密/敏感数据保护/密钥管理等）
- 业务安全（内容安全/风险识别等）
- 访问控制（云资源访问控制的精细程度）



03 深入分析数据泄露手段

The Analysis

现状与问题

主流云厂商通过行业内的权威安全认证
自己的安全研发团队建立了完善的基础安全服务

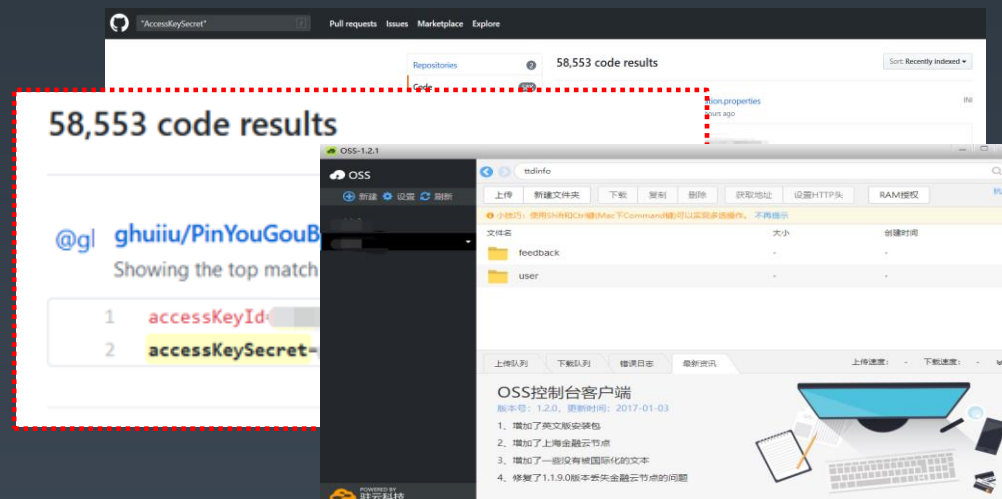
- 问题1：用户的安全意识及习惯
- 问题2：传统安全继续困扰云平台用户

深入分析数据泄露手段

用户的安全意识及习惯

某云厂商AccessKey泄露

GitHub搜索accessKeySecret有几万条记录，都是程序员将AccessKey公开在Github等代码托管平台上。



资源

- Github、Gitee等代码托管平台
- 各大云厂商提供的控制权限体系，如AccessKey等



获取&调用

- 通过搜索获取在互联网的上泄露的AK
- 获取到AK以后攻击者可以针对阿里云的各种服务调用相应API执行任意操作



常见使用

- OSS上传
- 下载
- 删除任意文件

深入分析数据泄露手段

ECS修改虚拟机实例

- ECS修改虚拟机实例的部分信息，包括实例密码、名称、描述、主机名和自定义数据等。
- 通过AK调用阿里云相关api，修改虚拟机实例密码。

```
https://ecs.aliyuncs.com/?Action=ModifyInstanceAttribute  
/?InstanceId=i-instance1  
&Action=ModifyInstanceAttribute  
&CreditSpecification=Standard  
&DeletionProtection=false  
&Description=InstanceAttribute  
&HostName=LocalHost  
&InstanceName=EcsInstance  
&Password=EcsV587!  
&Recyclable=  
&UserData=echo hello ecs!  
&<公共请求参数>
```

这就是上文提到的“可重置任意阿里云服务器root密码”

深入分析数据泄露手段

用户的安全意识及习惯

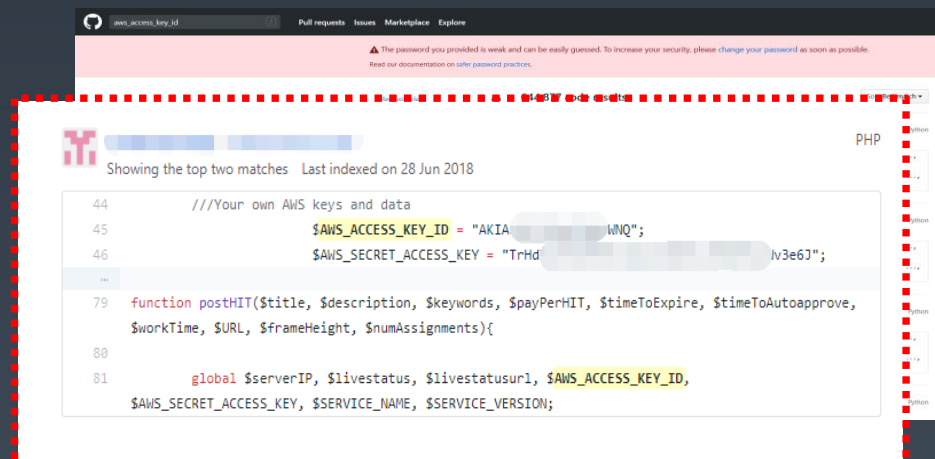
AWS AccessKey泄露

GitHub搜索aws_access_key_id部分程序员会将aws_access_key公开在Github代码托管平台上。

虽然AWS提供了强大的安全控制保障，但许多公司S3设置不合理或者泄露AK，允许未经授权的访问他们的数据，这势必会给企业带来巨大的安全隐患。

攻击者可以做哪些事情

- 直接下载bucket数据，下载敏感数据、系统备份，源代码等。
- 甚至可以下载到日志文件，其中可能还包含了用户名，密码，数据库查询等敏感信息；



The screenshot shows a GitHub search result for the query 'aws_access_key_id'. The search interface includes a header with navigation links like 'Pull requests', 'Issues', 'Marketplace', and 'Explore'. A warning message at the top states: 'The password you provided is weak and can be easily guessed. To increase your security, please change your password as soon as possible. Read our documentation on safer password practices.' Below this, the search results show 'Showing the top two matches' and 'Last indexed on 28 Jun 2018'. The code snippet is in PHP and contains the following lines:

```
44 //Your own AWS keys and data
45 $AWS_ACCESS_KEY_ID = "AKIA[REDACTED]";
46 $AWS_SECRET_ACCESS_KEY = "TrHd[REDACTED]";
...
79 function postHIT($title, $description, $keywords, $payPerHIT, $timeToExpire, $timeToAutoapprove,
$workTime, $URL, $frameHeight, $numAssignments){
80
81     global $serverIP, $livelstatus, $livelstatusurl, $AWS_ACCESS_KEY_ID,
$AWS_SECRET_ACCESS_KEY, $SERVICE_NAME, $SERVICE_VERSION;
```

- 实际生产环境中，S3 bucket常被用于存放一些静态资产，如图像和Javascript库。
- 但可以上传文件，例如某些恶意的Javascript；

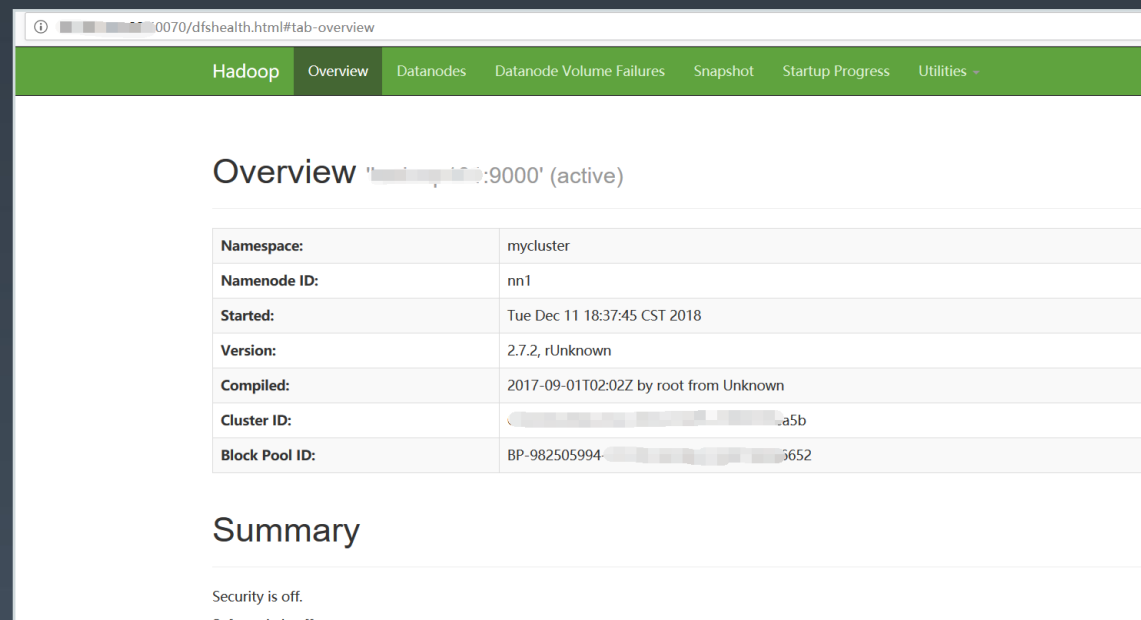
深入分析数据泄露手段

传统安全问题

云服务未授权访问

随着云计算和大数据的发展，云服务提供商基本都提供了云存储和大数据处理平台的融合，也推出了以存储为中心的轻计算处理框架，对数据处理，挖掘数据的价值和管理数据等带来前所未有的便利。

但有些企业将业务数据迁移到云服务以后缺乏必要的安全意识，导致大量敏感数据的未授权访问。这里我们仅以部分迁移到云服务商的hadoop平台为例简单了解下数据泄露的普遍现象。



通过搜索引擎可以搜到大量部署在云服务器上并且对外开放了访问端口服务的hadoop平台。

深入分析数据泄露手段

云服务未授权访问

50070/explorer.html#/user/hive/warehouse/online/ods_user/yfq_user							
Hadoop Overview Datanodes Snapshot Startup Progress Utilities							
Browse Directory							
Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	deploy	supergroup	609.65 MB	2018/9/10 上午11:09:50	1	256 MB	1
-rw-r--r--	deploy	supergroup	5.92 MB	2018/9/12 上午10:41:11	1	256 MB	1
-rw-r--r--	deploy	supergroup	2.99 MB	2018/9/13 上午10:41:13	1	256 MB	2
-rw-r--r--	deploy	supergroup	5.51 MB	2018/9/14 上午10:41:12	1	256 MB	3
-rw-r--r--	deploy	supergroup	5.02 MB	2018/9/15 上午10:41:14	1	256 MB	4
-rw-r--r--	deploy	supergroup	3.85 MB	2018/9/16 上午10:41:15	1	256 MB	4
-rw-r--r--	deploy	supergroup	3.61 MB	2018/9/17 上午10:41:03	1	256 MB	5
-rw-r--r--	deploy	supergroup	3.82 MB	2018/9/18 上午10:41:13	1	256 MB	7
-rw-r--r--	deploy	supergroup	3.58 MB	2018/9/19 上午10:41:12	1	256 MB	1
-rw-r--r--	deploy	supergroup	3.46 MB	2018/9/20 上午10:41:02	1	256 MB	3
null	0	1	null	null	135	115	83
null	0	1	null	null	130	103	24
null	0	1	null	null	177	78	86
null	0	1	null	null	136	262	51
null	0	1	null	null	131	29	75
null	0	1	null	null	13	12	84
null	0	1	null	null	13	47	26
null	0	1	null	null	18	22	69
null	0	1	null	null	17	1	43
null	0	1	null	null	15	4	71
null	0	1	null	null	13	35	75
null	0	1	null	null	18	2	62
null	0	1	null	null	13	5	75
null	0	1	null	null	13	50	99
null	0	1	null	null	15	3	50
null	0	1	null	null	15	3	38
null	0	1	null	null	18	2	88
null	0	1	null	null	13	33	16
null	0	1	null	null	13	12	15
null	0	1	null	null	13	05	86
null	0	1	null	null	13	36	51
null	0	1	null	null	13	38	018
null	0	1	null	null	13	12	599
null	0	1	null	null	13	12	168
null	1	1	null	null	15	1	41
null	0	1	null	null	137	1	41
null	0	1	null	null	131	0	51
1001	1	1	null	null	177	7	40

由于权限设置不合理，任意地址可以访问服务端口，可以直接下载数据文件，我们可以看到该企业在云服务器上泄露了大量隐私数据，包括手机号码等敏感信息。



深入分析数据泄露手段

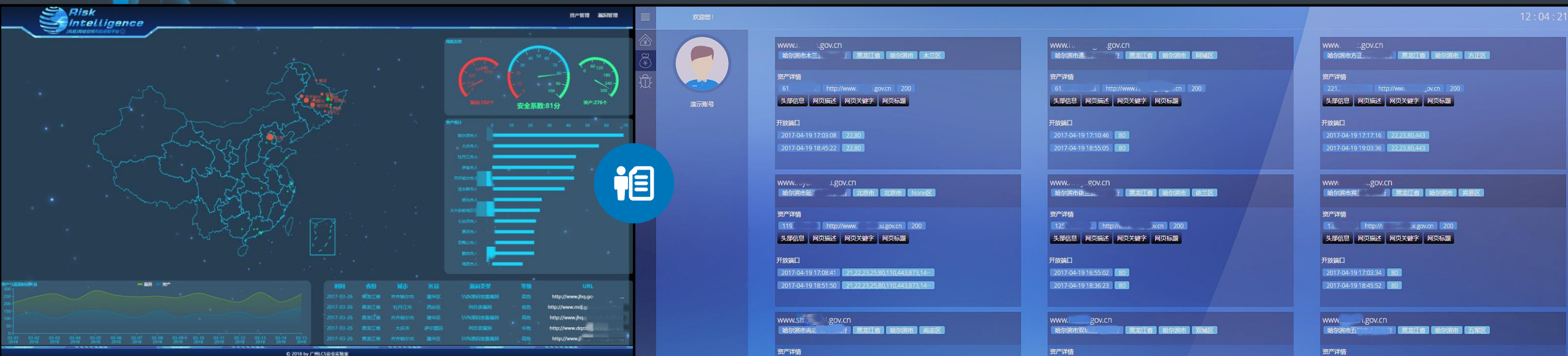
类似的问题还有很多Mongodb、Elasticsearch，这些问题是因为用了云才出现的问题吗？

04

云上数据安全的
建议

The Suggest

云上数据安全的建议



云商安全 + 用户安全

- 云用户的研发、运维流程的定期review;
- 云基础安全, 使用云服务器控制台, 合理配置安全组策略, 服务器出入口流量等进行监控预警;
- 对自身网络资产、漏洞进行发现和管理;
- 对应用缺陷威胁进行实时预警;

云上数据安全于使用者的建议



1. 选择通过权威安全认证的主流云厂商
2. 合理使用云厂商提供的基础安全服务
3. 不要泄露AK等敏感信息
4. 建议遵循云安全最佳实践，定期重置所有的证书、密码和API访问密钥
5. 合理设置云服务访问权限，使用子用户来进行API调用
6. 云上敏感数据加密存储

云上数据安全的风险发现手段



1. 云厂商所提供的监测工具
2. 体系化的安全运维标准
3. 搭建自己的资产管理平台（成熟开源项目）
4. 搭建自己的通用漏洞测试平台（成熟开源项目）
5. 使用更完善的云上资源安全方案（商业安全产品）
6. 建立自身数据泄露情报通道（成熟开源项目）

云上数据安全与否，
是因为谁用了云服务，而不是用了谁的云服务。

TGO 鲲鹏会

汇聚全球科技领导者的高端社群

 全球12大城市

 850+高端科技领导者

使命
Mission

为社会输送更多优秀的
科技领导者

愿景
Vision

构建全球领先的有技术背景
优秀人才的学习成长平台



扫描二维码，了解更多内容

想做团队的领跑者 需要迈过这些“槛”

成长型企业，易忽视人才体系化培养
企业转型加快，团队能力又跟不上

VS

从基础到进阶，超100+一线实战
技术专家带你系统化学习成长

团队成员技能水平不一，
难以一“敌”百人需求

VS

解决从小白到资深技术人所遇到
80%的问题

寻求外部培训，奈何价更高且
集中式学习

VS

多样、灵活的学习方式，包括
音频、图文 和视频

学习效果难以统计，产生不良循环

VS

获取员工学习报告，查看学习
进度，形成闭环



课程顾问「橘子」

回复「QCon」
免费获取
学习解决方案

极客时间企业账号 # 解决技术人成长路上的学习问题

THANKS! | QCon IOth