

饿了么 MySQL运维平台介绍

自我介绍

1 蔡鹏

2 MySQL DBA@饿了么(2015~)

3 目前负责运维平台建设、日常运维

4 MySQL、Go、Python

大纲

- 发展阶段
- MySQL架构的演变
- DBA面临哪些挑战
- 平台的演进过程
- 主要功能介绍

发展阶段



实例30+

2015年初

人肉+脚本阶段
python,shell

实例800+

2016初

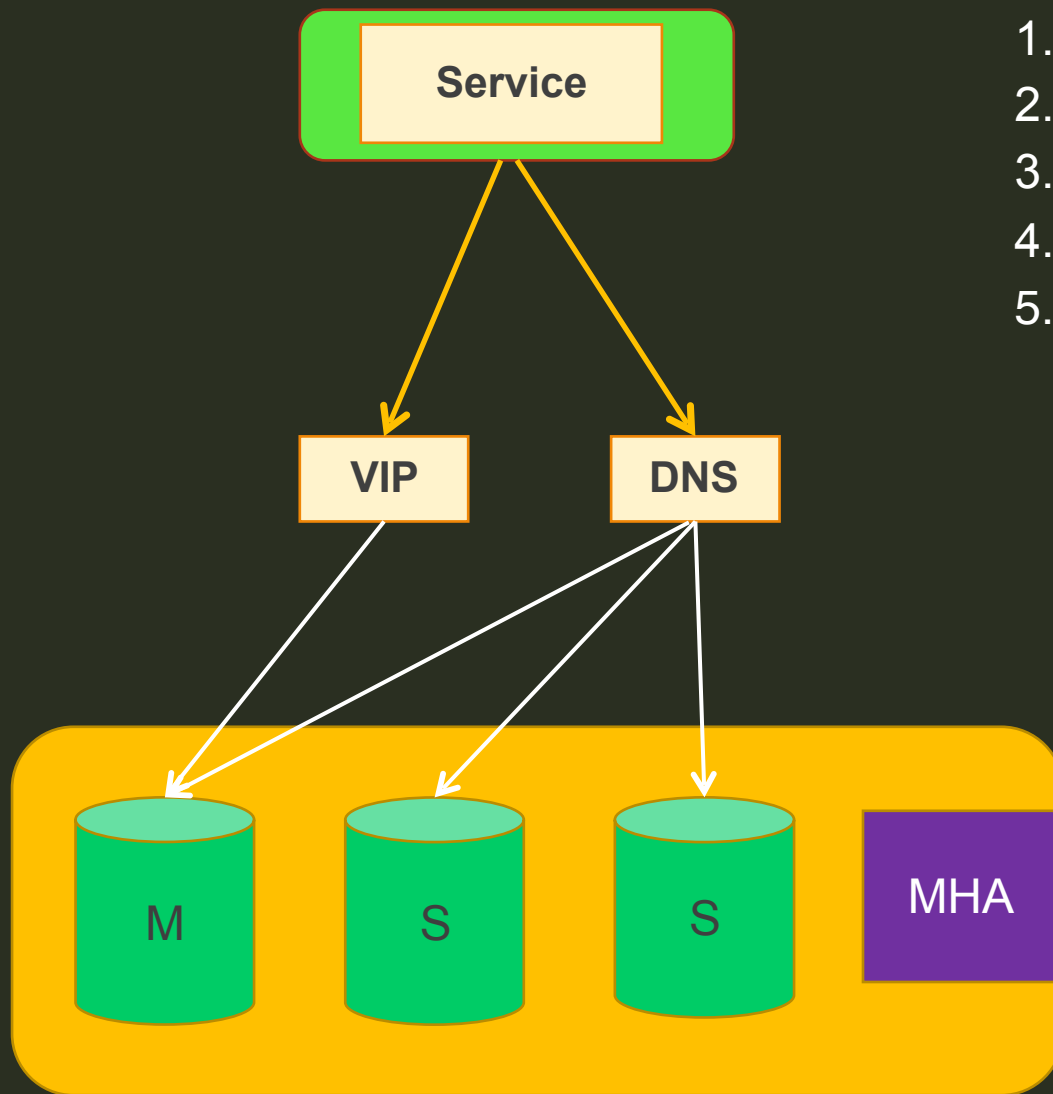
工具半自动化阶段
Django
Bootstrap

实例2000+

2017至今

平台化
Django
GO
React

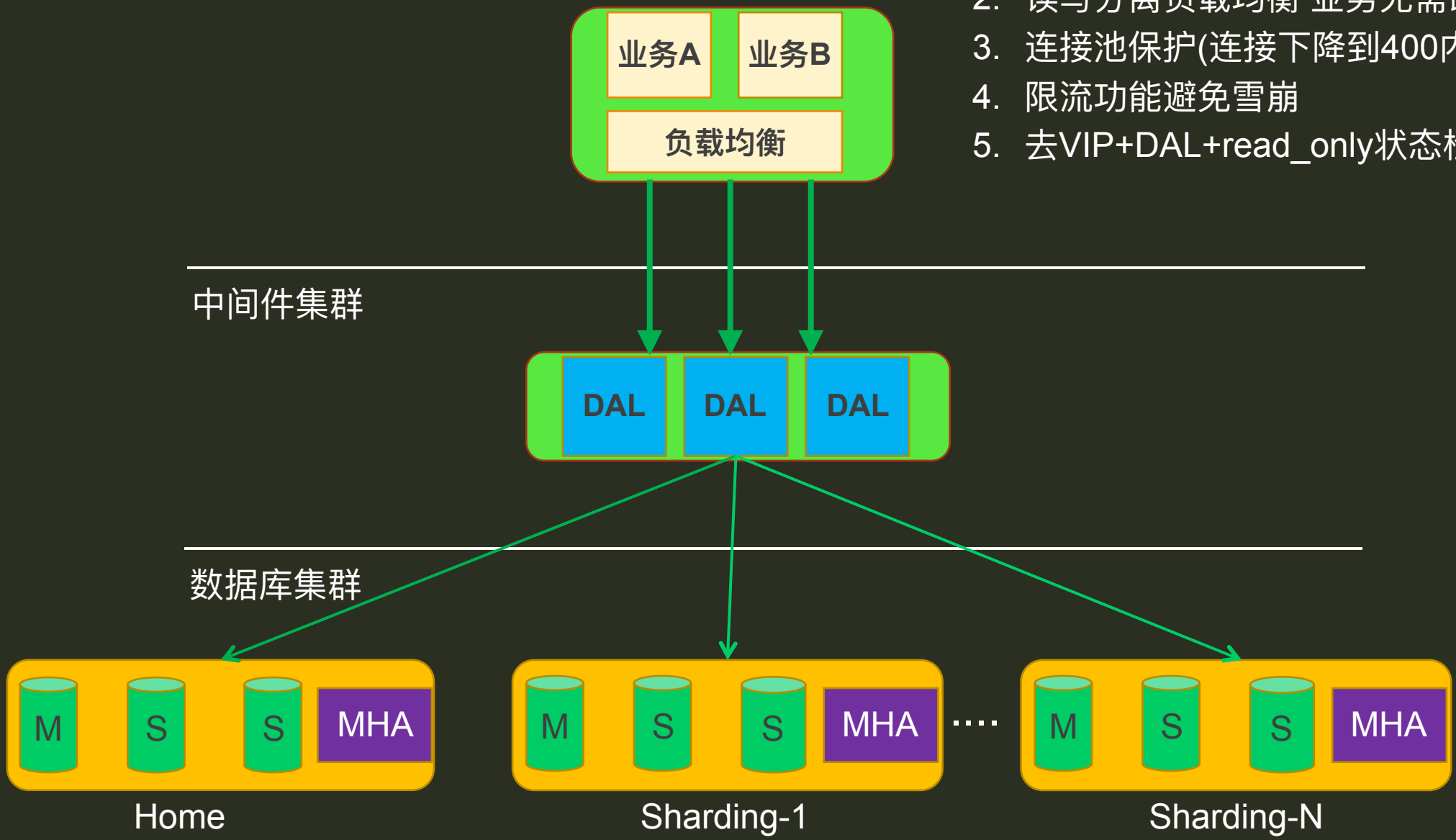
架构演变1.0



1. 难扩展(出现单集群1主11从的超大集群)
2. 读写分离、负载均衡业务改造成本高
3. DB缺乏限流机保护制高并发下易雪崩
4. 无连接池,连接数维持在10000+高风险
5. MHA切换容易脑裂

架构演变2.0

- 1. 易扩展(sharding)
- 2. 读写分离负载均衡 业务无感知
- 3. 连接池保护(连接下降到400内)
- 4. 限流功能避免雪崩
- 5. 去VIP+DAL+read_only状态检查防止脑裂



中间件集群

数据库集群

Home

Sharding-1

Sharding-N

架构演变3.0

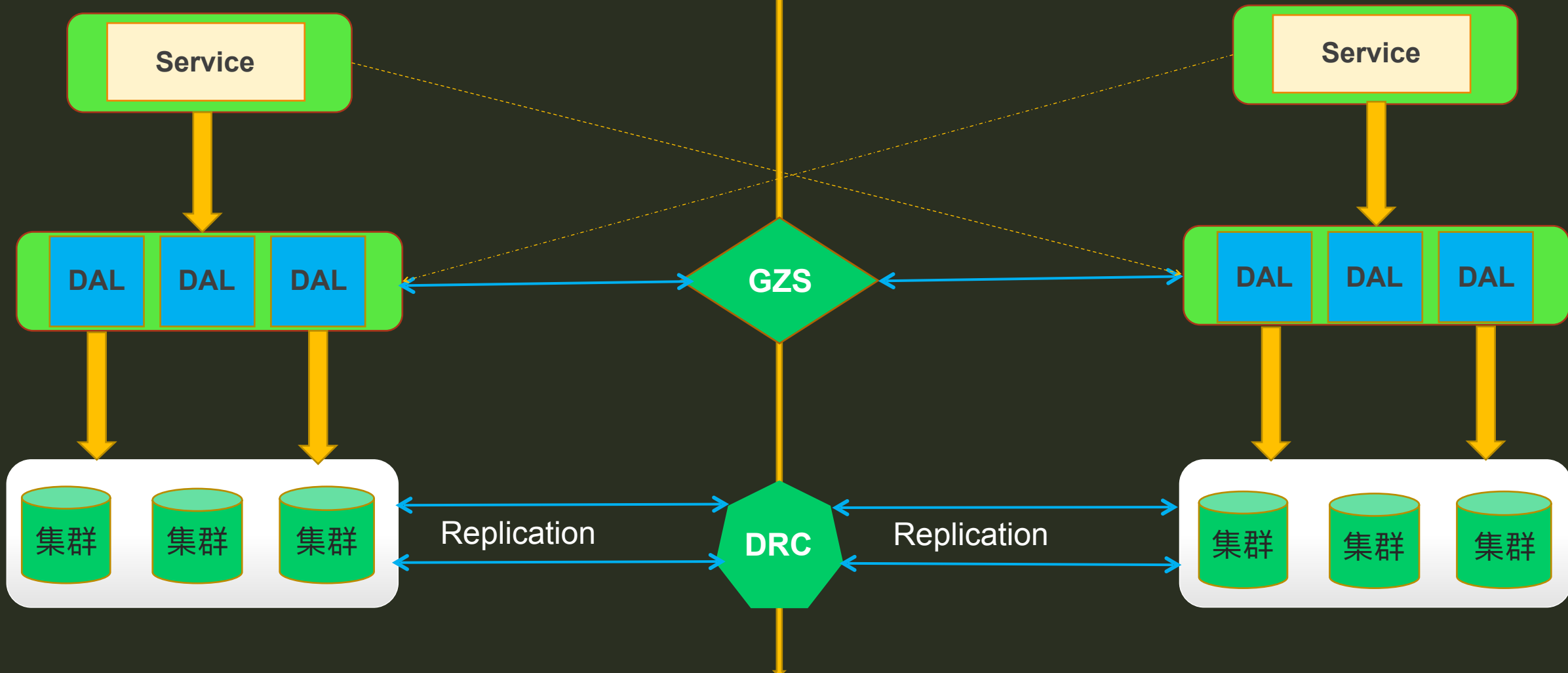
北京机房

上海机房



IT大咖说
知识分享平台

1. 如何DDL(最大挑战)
2. 如何自发现(集群拓扑结构)
3. 监控



DBA面临的挑战

1

- ◆ 如何快速定位问题快速解决问题?
- ◆ 如何从急于上午救火、下午忙于打杂、夜里疲于发布中解脱?
- ◆ 人为操作失误太多如何解决? 军规? SOP?(不能写进代码里的标准化, 军规, sop最终都是徒劳的)
- ◆ 我们需要自动化工具, 谁来做?

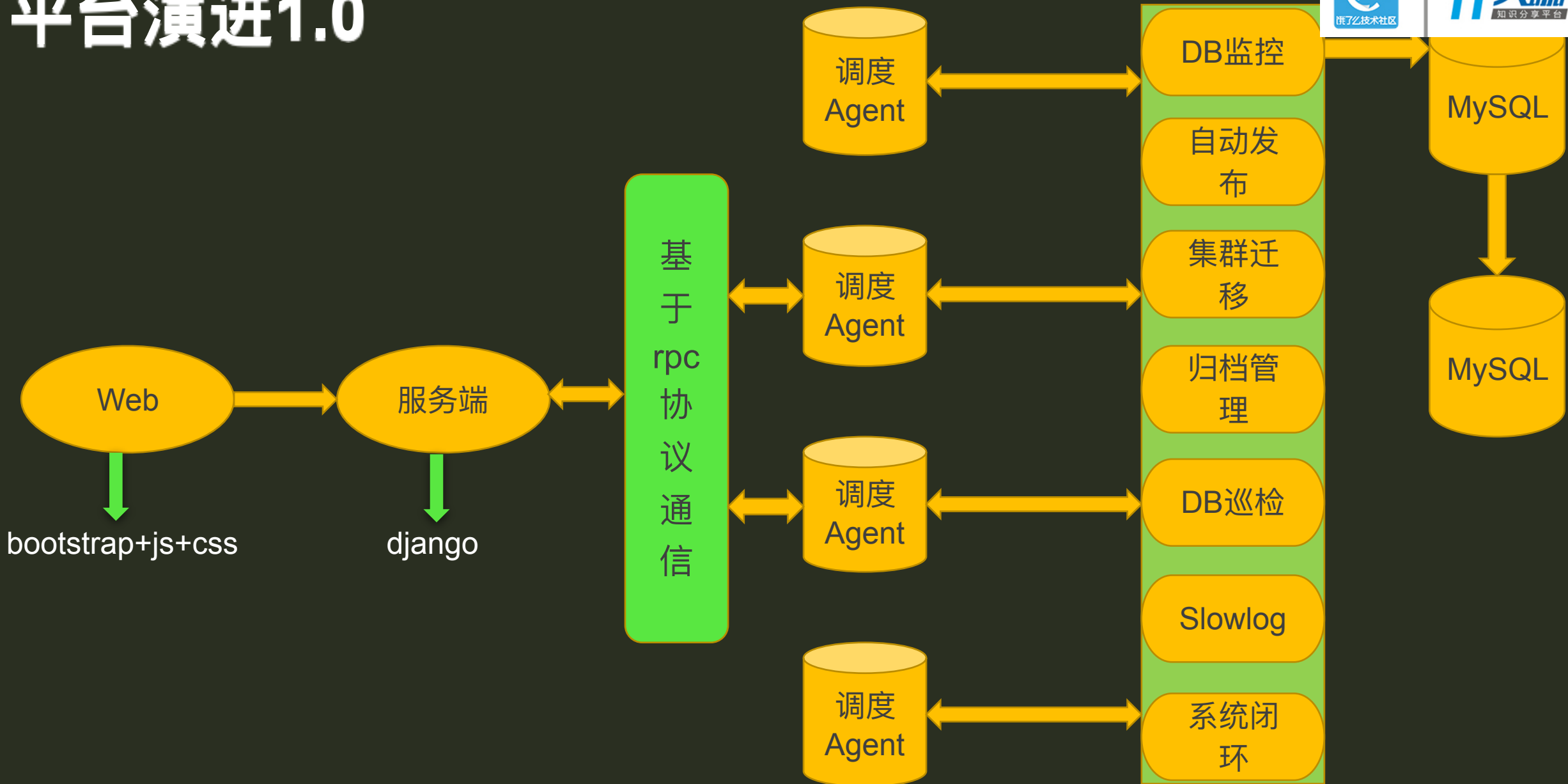
2

- ◆ 如何做到快速扩容?
- ◆ 如何应对DB体积的快速膨胀带来的一系列问题?
- ◆ 如何满足每天数量繁多的发布(sharding后又该如何应对)?

3

- ◆ 多活后实例翻倍原有的集中式监控采集无法满足实时性需求如何解决?
- ◆ 如何满足每天数量繁多的发布(sharding+drc(多活)后又该如何应对)?
- ◆ 数量庞大的归档需求该如何满足, 如果做到维护成本最低
- ◆ 多活后集群数量翻倍如何管理(每天都有实例挂掉、上下线MHA配置真的准确吗?)

平台演进1.0

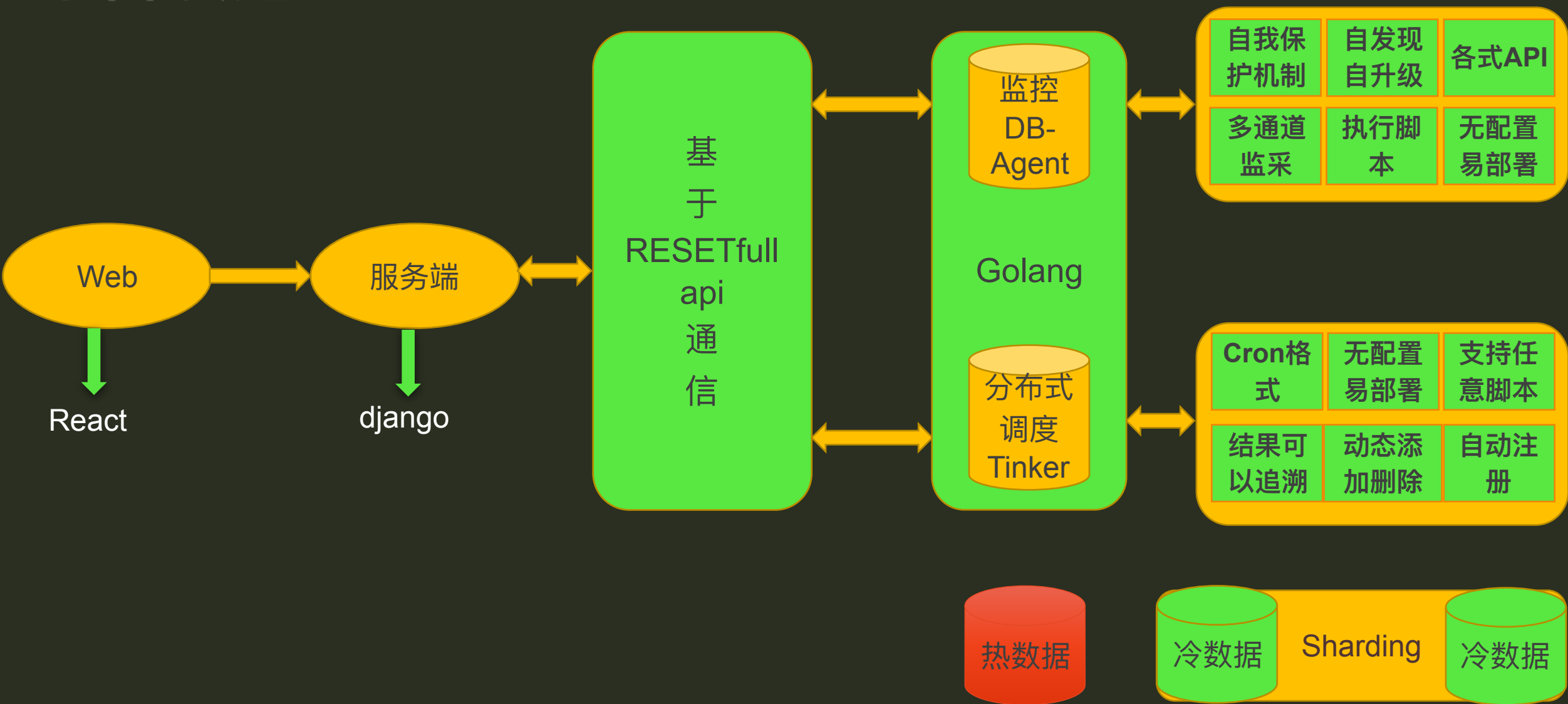


平台1.0不足



- ◆基于python实现的任务调度不够轻盈、部署繁琐、扩展性不佳
- ◆集中式DB监控性能差、不精细、横向扩展能力有限、无法采集到OS指标、无法对服务器执行操作(不能消灭手工、不能完成更多自动化功能赋予的需要)

平台演进2.0



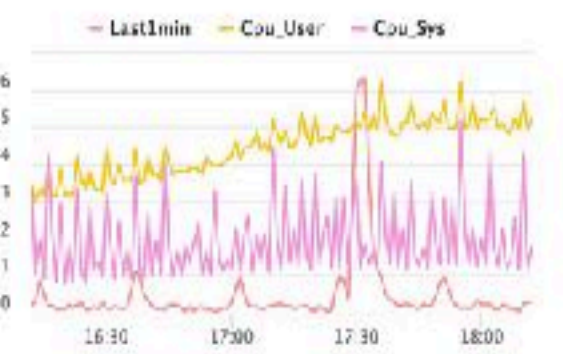
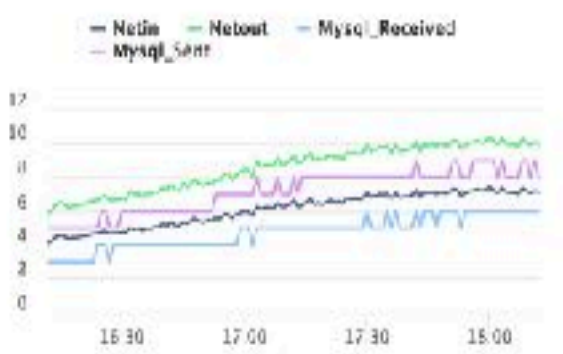
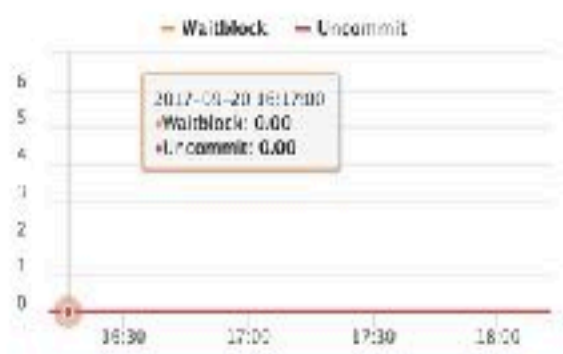
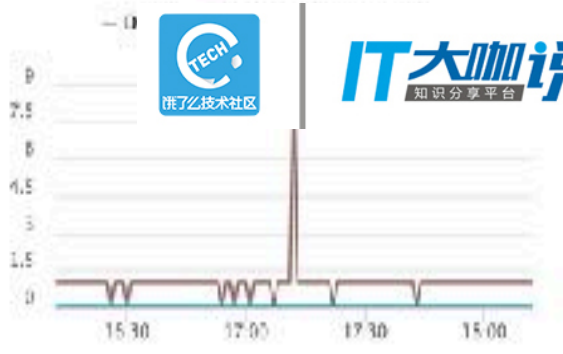
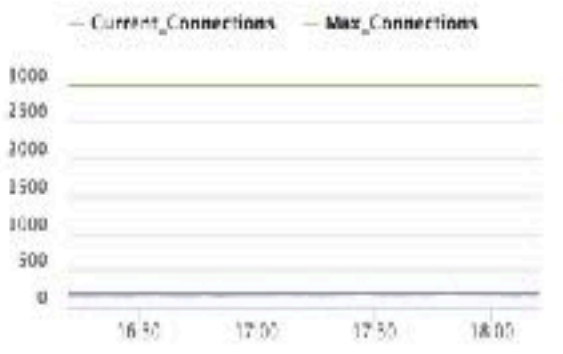
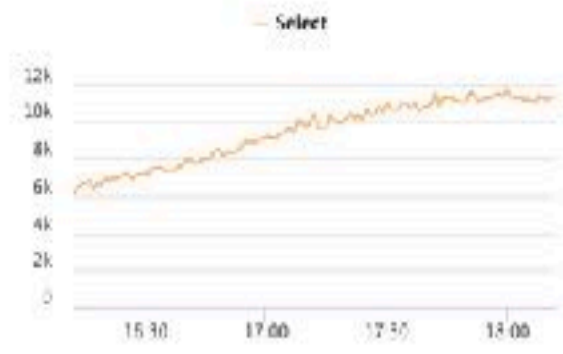
- ◆ 监控、调度扩展能力强
- ◆ 分布式DB监控采集高性能、可扩展、采集指标全面、可对服务器做操作为众多自动化功能提供支撑、不在依赖SSH的各种玩法

主要功能介绍

- ◆实时监控、快速排障
- ◆数据库发布
- ◆任务调度
- ◆MHA分布式管理
- ◆数据迁移
- ◆资源池，一键安装
- ◆备份系统的实现
- ◆授权

实时监控、快速排障

- ✓ 全局监控仪表盘(秒级监控), 问题快速呈现
- ✓ AWR报表趋势快速呈现, 囊括各种细节
- ✓ Processlist展示, 执行计划呈现, 锁分析
- ✓ 异常快速处理(kill sql, kill block)
- ✓ 常见复制异常快速修复



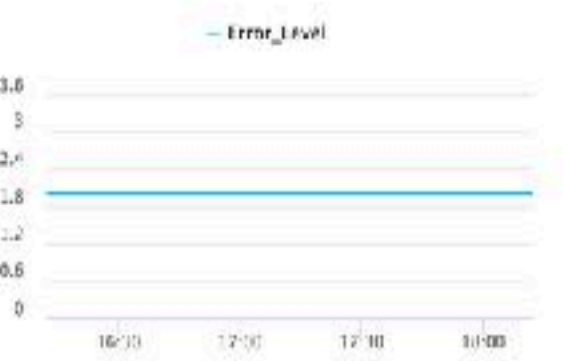
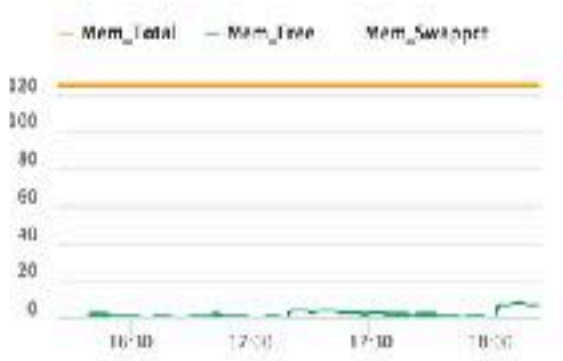
查看processlist快照



查看表的ops分布



查看sql相应时长分布



- ProcessSnapshots
- TableOps
- Sqlrt
- FilelogicOps
- UserConns

数据库文件逻辑IO次数



kill:46446261 2017-09-20 10:28:05 d... report 10.0.2.15 50987 ... /# EMAVCB31E224-682B-46A1-B7C8-B0CD8CB07082/ynn@ele.me *) select substr(date_sub(a.data_dt,interval 7 day),6,5) as chart_x, round(a.gmv/a.order_num,2) as `客单价`, concat(round(a.month_gmv*100/b.target*a.data_day/a.data_last_day,2),'%') as `GMV达成率`, concat(round(a.subsidy_amt*100/a.no_group_gmv,2),'%') as `机动补贴力度_已使用`, concat(round(a.qiandian_qianmian_subsidy*100/a.no_group_gmv,2),'%') as `千店千面补贴力度` from (select data_dt,day(data_dt) as data_day, day(last_day(data_dt)) as data_last_day,regiment_id,sum(gmv+group_gmv) as gmv,sum(month_gmv+month_group_gmv) as month_gmv, sum(gmv) no_group_gmv,sum(subsidy_amt) as subsidy_amt,sum(qiandian_qianmian_subsidy) qiandian_qianmian_subsidy, sum(order_num) order_num from important_shop_data where data_dt=>DATE_sub('2017-09-19',INTERVAL 7 DAY) and bu_flag='SIG' and regiment_id = '3208' group by data_dt,regiment_id) a left join (select month,org_id,sum(target1) target from st_report_warreport where war_id in (1,2) and org_type='regiment' and org_attr in ('CKA','SIG') and month=>substr(DATE_sub('2017-09-19',INTERVAL 30 DAY),1,7) group by month,org_id) b on substr(a.data_dt,1,7)=b.month and b.org_id=a.regiment_id order by a.data_dt

SQL state-->Sending data query time-->28 LockByID:

id	select_type	tables	type	key	key_len	rows	Extra
1	PRIMARY	<derived2>	ALL	NULL	NULL	57977	Using filesort
1	PRIMARY	<derived3>	ref	<auto_key0>	66	10	Using where
3	DERIVED	st_report_warreport	ALL	NULL	NULL	113853	Using where; Using temporary; Using filesort
2	DERIVED	st_report_prd_important_shop_data	index	ix_data_dt_regiment_id	13	11595421	Using where

kill:46446264 2017-09-20 10:28:05 d... report 10.0.2.15 50909 ... /# EMAV095E90CF-5199-44D3-A432-204A08151D31/ynn@ele.me *) select substr(date_sub(a.data_dt,interval -7 day),6,5) as chart_x, round(a.gmv/a.order_num,2) as `客单价`, concat(round(a.month_gmv*100/b.target*a.data_day/a.data_last_day,2),'%') as `GMV达成率`, concat(round(a.subsidy_amt*100/a.no_group_gmv,2),'%') as `机动补贴力度_已使用`, concat(round(a.qiandian_qianmian_subsidy*100/a.no_group_gmv,2),'%') as `千店千面补贴力度` from (select data_dt,day(data_dt) as data_day, day(last_day(data_dt)) as data_last_day,regiment_id,sum(gmv+group_gmv) as gmv,sum(month_gmv+month_group_gmv) as month_gmv, sum(gmv) no_group_gmv,sum(subsidy_amt) as subsidy_amt,sum(qiandian_qianmian_subsidy) qiandian_qianmian_subsidy, sum(order_num) order_num from important_shop_data where data_dt=>DATE_sub('2017-09-19',INTERVAL 37 DAY) and data_dt=>DATE_sub('2017-09-19',INTERVAL 7 DAY) and bu_flag='BL' and regiment_id = '3208' group by data_dt,regiment_id) a left join (select month,org_id,sum(target1) target from st_report_warreport where war_id in (1,2) and org_type='regiment' and org_attr='BL' and month=>substr(DATE_sub('2017-09-19',INTERVAL 30 DAY),1,7) group by month,org_id) b on substr(a.data_dt,1,7)=b.month and b.org_id=a.regiment_id order by a.data_dt

kill 该sql SQL state-->Sending data query time-->25 LockByID:

id	select_type	tables	type	key	key_len	rows	Extra
1	PRIMARY	<derived2>	ALL	NULL	NULL	57977	Using filesort
1	PRIMARY	<derived3>	ref	<auto_key0>	66	10	Using where
3	DERIVED	st_report_warreport	ALL	NULL	NULL	113853	Using where; Using temporary; Using filesort
2	DERIVED	st_report_prd_important_shop_data	index	ix_data_dt_regiment_id	13	11595421	Using where

kill:46446269 2017-09-20 10:28:05 d... report 10.0.2.15 2146 ... /# EMAV04FA0085-179A-4281-B9A6-81FE4C2CFB4E/ynn@ele.me *) select substr(a.data_dt,6,5) as chart_x, round(a.gmv/a.order_num,2) as `客单价`, concat(round(a.month_gmv*100/b.target*a.data_day/a.data_last_day,2),'%') as `GMV达成率`, concat(round(a.subsidy_amt*100/a.no_group_gmv,2),'%') as `机动补贴力度_已使用`, concat(round(a.qiandian_qianmian_subsidy*100/a.no_group_gmv,2),'%') as `千店千面补贴力度` from (select data_dt,day(data_dt) as data_day, day(last_day(data_dt)) as data_last_day,regiment_id,sum(gmv+group_gmv) as gmv,sum(month_gmv+month_group_gmv) as month_gmv, sum(gmv) no_group_gmv,sum(subsidy_amt) as subsidy_amt,sum(qiandian_qianmian_subsidy) qiandian_qianmian_subsidy, sum(order_num) order_num from important_shop_data where data_dt=>DATE_sub('2017-09-19',INTERVAL 30 DAY) and regiment_id = '3208' group by data_dt,regiment_id) a left join (select month,org_id,sum(target1) target from st_report_warreport where war_id in (1,2) and org_type='regiment' and month=>substr(DATE_sub('2017-09-19',INTERVAL 30 DAY),1,7) group by month,org_id) b on substr(a.data_dt,1,7)=b.month and b.org_id=a.regiment_id order by a.data_dt

SQL state-->Sending data query time-->24 LockByID:

id	select_type	tables	type	key	key_len	rows	Extra
1	PRIMARY	<derived2>	ALL	NULL	NULL	579771	Using filesort
1	PRIMARY	<derived3>	ref	<auto_key0>	66	10	Using where
3	DERIVED	st_report_warreport	ALL	NULL	NULL	113853	Using where; Using temporary; Using filesort

Master_Host	10.200.20.11.3307
Master_User	root
Master_Log_File	mysql-bin.000050
Relay_Master_Log_File	mysql-bin.000050
Slave_IO_Running	Yes
Slave_SQL_Running	No
Exec_Master_Log_Pos	47296130
Seconds_Behind_Master	None
Last_IO_Error	
Last_SQL_Error	Could not execute Write_rows event on table ... Duplicate entry '123490' for key 'PRIMARY', Error_code: 1062; handler error HA_ERR_FOUND_DUPP_KEY; the event's master log mysql-bin.000050, end_log_pos 47296698
error	

跳错

Stop Slave

Start Slave

累计杀死Block线程数:0

kill关键字

time:4 v

KillProc

KillBlock

KillSleep

Stop

数据库发布

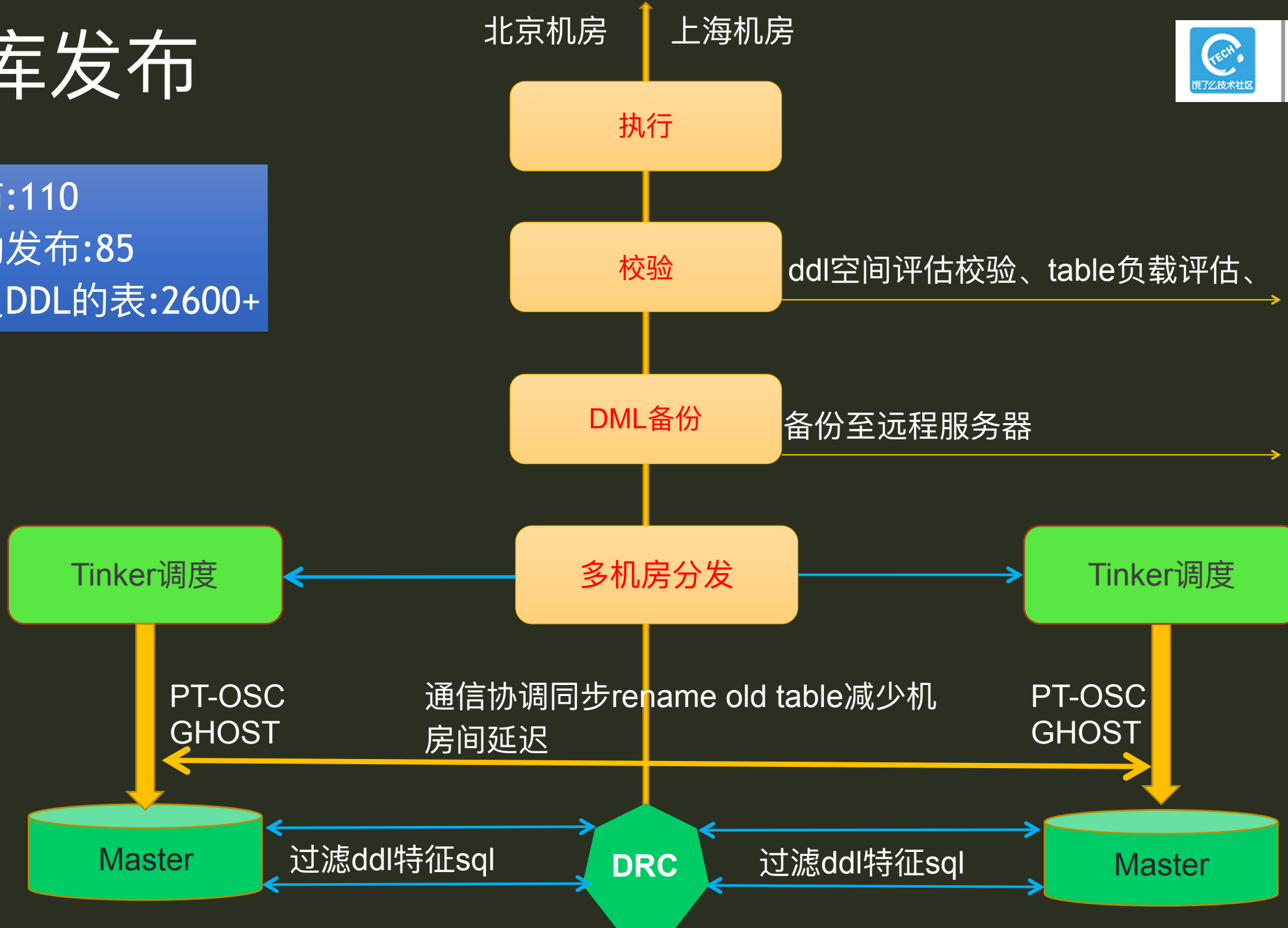
北京机房

上海机房



IT大咖说
知识分享平台

日均发布:110
日均自动发布:85
日均涉及DDL的表:2600+




序号	定时任务	IP	Port	DB	TableName	Rows	
1	时:分		3306			0	DE
2	时:分	10.18.124.11	3306	oss	im_ag	-36	咖啡_咖啡DB
3	时:分				咖啡_ag	0	咖啡_CoffeeDB
4	时:分					0	咖啡_CoffeeDB
5	时:分					0	咖啡_CoffeeDB
6	时:分					0	咖啡_CoffeeDB
7						0	
8	时:分					0	咖啡_CoffeeDB
9	时:分					0	咖啡_CoffeeDB
10	时:分	10.18.124.24	3306	atapi	tb_message_diffinfo	0	Archiver
11	时:分	10.18.124.64	3306		risk_alliance	0	Risk
12	时:分	10.18.124.50.30	3306		tb_message_diffinfo	48,027	Lpd
13	时:分	10.18.124.35	3306	lpd	tb_message_diffinfo	48,027	Lpd
14	时:分	10.18.124.30.12	3306	commodity	tb_message_diffinfo	0	Google
15	时:分	10.18.124.95	3306	commodity	tb_message_diffinfo	0	Google
16	时:分	10.18.124.31	3306	tbl	tbl_weather	0	tbl

自由定义发布时间

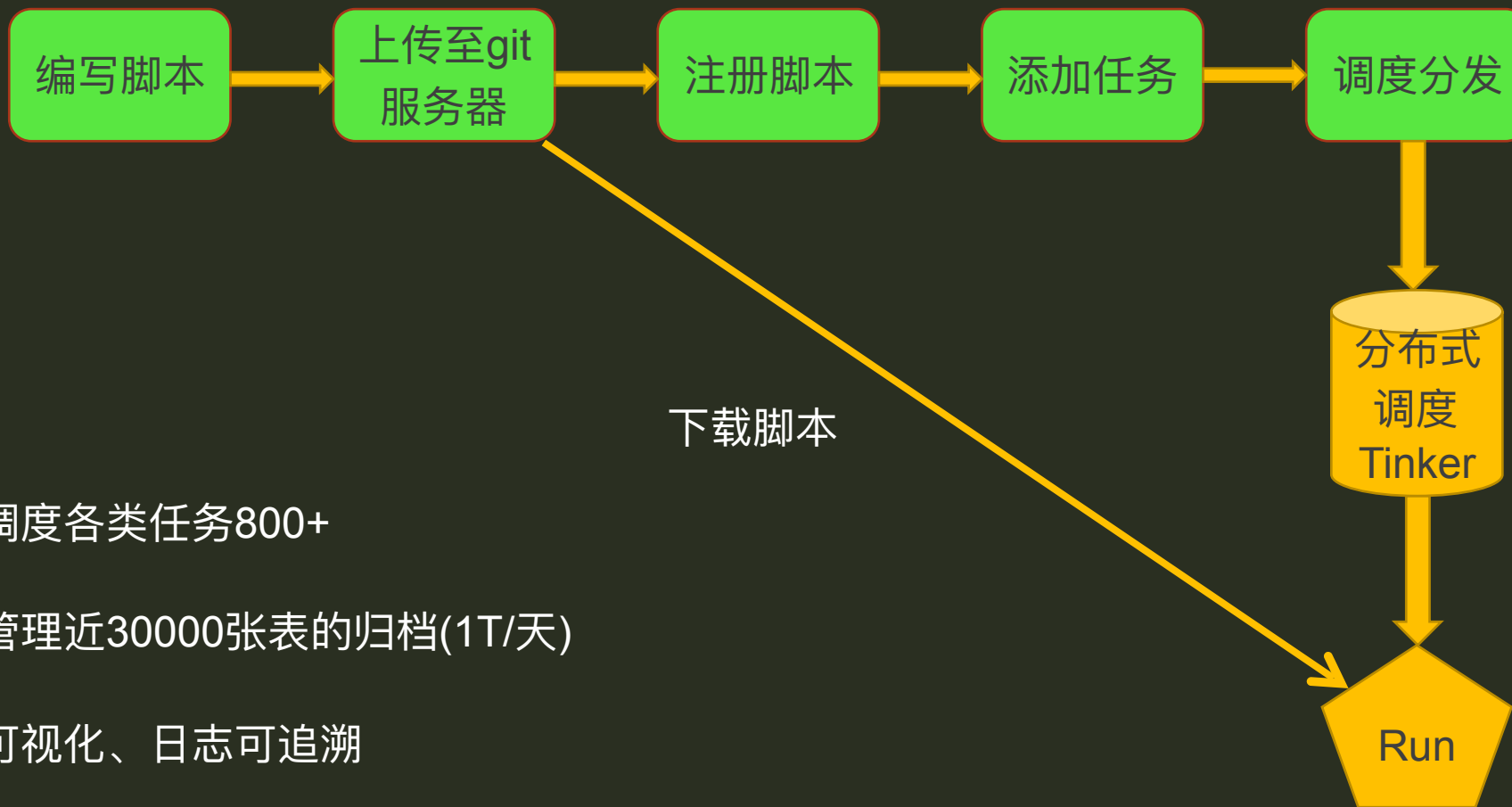
**查看该表的读写情况
用来选择合理的执行时间**

— 逻辑读
— 逻辑写

JobID	SQL  <input type="checkbox"/> 过滤[已执行]	备注	执行	状态		
13194	CREATE TABLE `dal_se` 	2017-09-18 21	执行 	执行成功		
13194	CREATE TABLE `dal_se` 	2017-09-18 21	原生Alter执行	执行成功	100%	日志
15594	CREATE TABLE `tb_ret` 	2017-09-18 13	PT执行	执行成功	100%	日志
15594	CREATE TABLE `tb_ret` 	2017-09-18 13	mm-ost执行	执行成功	100%	日志
15650	create table retail_e 	2017-09-18 12	取消执行	执行成功	100%	日志
15650	create table retail_e 	2017-09-18 12	执行 	执行成功	100%	日志
16130	delete from taoist_c	2017-09-18 14	执行 	执行成功	100%	日志
16338	delete from taoist_c	2017-09-18 14	执行 	执行成功	100%	日志
16450	ALTER TABLE payment_	发布依赖 2017-09-1	执行 	执行成功	100%	日志
17506	alter table page_set 	2017-09-18 17	执行 	执行成功	100%	日志
17506	alter table page_set 	2017-09-18 17	执行 	执行成功	100%	日志
17706	INSERT INTO account(2017-09-18 14	执行 	执行成功	100%	日志
17706	INSERT INTO `bu_user	2017-09-18 14	执行 	执行成功	100%	日志
17898	CREATE TABLE `handle` 	发布依赖 2017-09-1	执行 	执行成功	100%	日志
17898	CREATE TABLE `handle` 	发布依赖 2017-09-1	执行 	执行成功	100%	日志
17938	ALTER TABLE com modit 	2017-09-18 11	执行 	执行成功	100%	日志



任务调度



每天调度各类任务800+

有效管理近30000张表的归档(1T/天)

结果可视化、日志可追溯



for Owner

任务名称	任务服务器	任务脚本	任务计划	开始时间	结束时间	执行时间	Owner	开关	操作	状态
自增id溢出巡检	10.0.38.47	table_overflow_monitor.py	1 21 8 * * *	2017-09-20 08:21:01	2017-09-20 08:26:31	15分	曹飞	<input checked="" type="checkbox"/>	Delete 编辑	SUCCESS
MWIA每日巡检							程博雨	<input checked="" type="checkbox"/>	Delete 编辑	SUCCESS
报警汇总任务							曹飞	<input checked="" type="checkbox"/>	Delete 编辑	SUCCESS
基础数据收集							曹飞	<input checked="" type="checkbox"/>	Delete 编辑	SUCCESS
数据库自动发布							曹飞	<input checked="" type="checkbox"/>	Delete 编辑	SUCCESS
DT分区管理							曹飞	<input checked="" type="checkbox"/>	Delete 编辑	SUCCESS
DB_S3_W3备份							高博峰	<input type="checkbox"/>	Delete 编辑	未执行
收集Table Cps							曹飞	<input checked="" type="checkbox"/>	Delete 编辑	SUCCESS
数据库自动发布(周三)							程博雨	<input type="checkbox"/>	Delete 编辑	未执行
DRC表字段巡检							曹飞	<input checked="" type="checkbox"/>	Delete 编辑	SUCCESS
数据库性能统计-每天							曹飞	<input checked="" type="checkbox"/>	Delete 编辑	SUCCESS
数据库性能统计-每小时							曹飞	<input checked="" type="checkbox"/>	Delete 编辑	SUCCESS
Project参数一致性检测							曹飞	<input checked="" type="checkbox"/>	Delete 编辑	SUCCESS
MySQL配置巡检	10.0.11.27	check_mysql_param.py	0 20 0 * * *	2017-09-20 09:20:00	2017-09-20 09:24:05	4分	曹飞	<input checked="" type="checkbox"/>	Delete 编辑	SUCCESS
DAI配置巡检	10.0.11.76	daiconfig_check.py	0 45 0 * * *	2017-09-20 09:45:00	2017-09-20 09:58:02	13分	曹佳文	<input checked="" type="checkbox"/>	Delete 编辑	SUCCESS
analyze table bdi_ba.st_trc_heatmaps_restaurant	10.0.11.127	analyze_bdi_ba.py	0 0 0 * * *				曹飞	<input checked="" type="checkbox"/>	Delete 编辑	未执行

任务编辑 注册脚本

* 任务名称: * 任务服务器: * 任务计划:

* 任务脚本:

* 脚本参数:

参数信息:

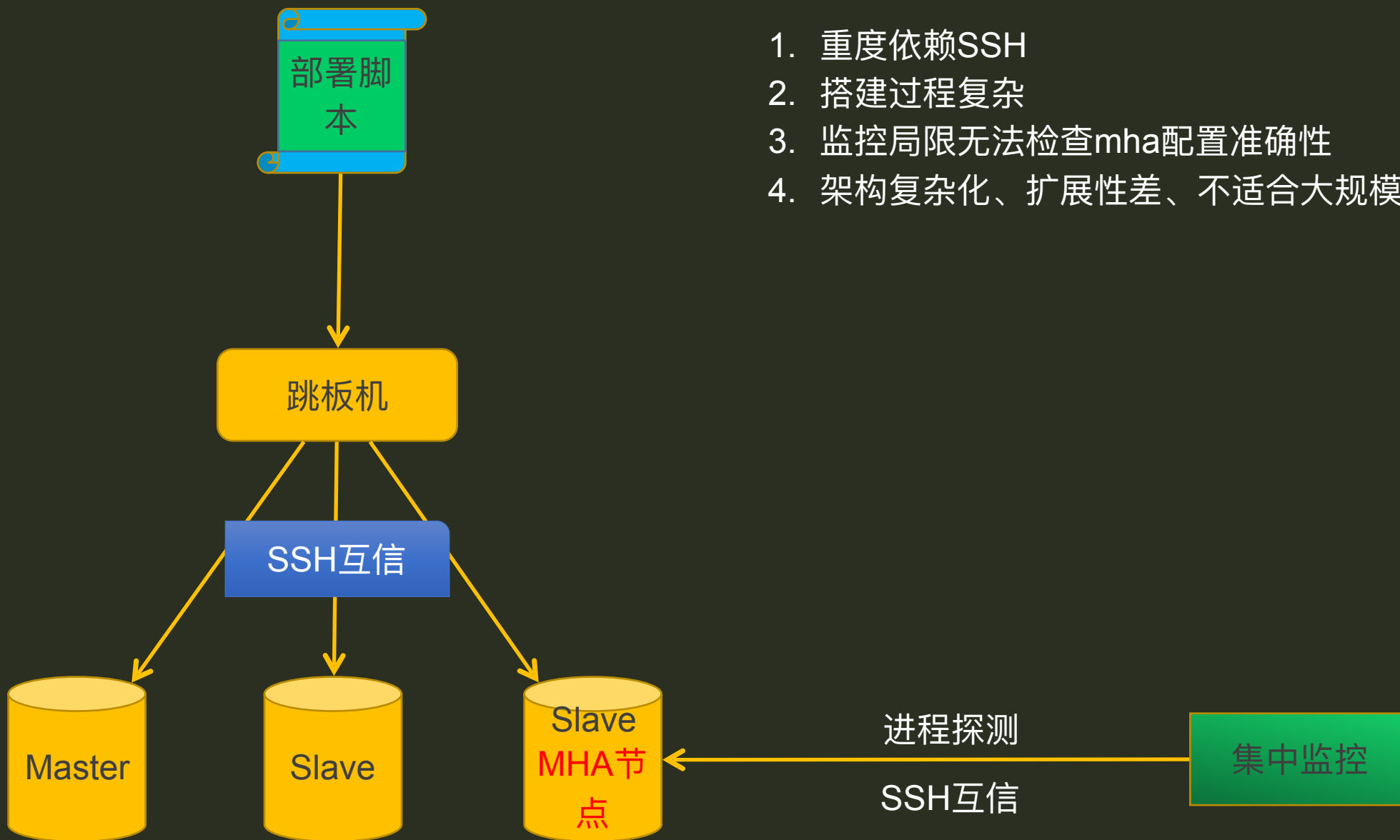
```

--project 要检查的库表
--threads 线程

```

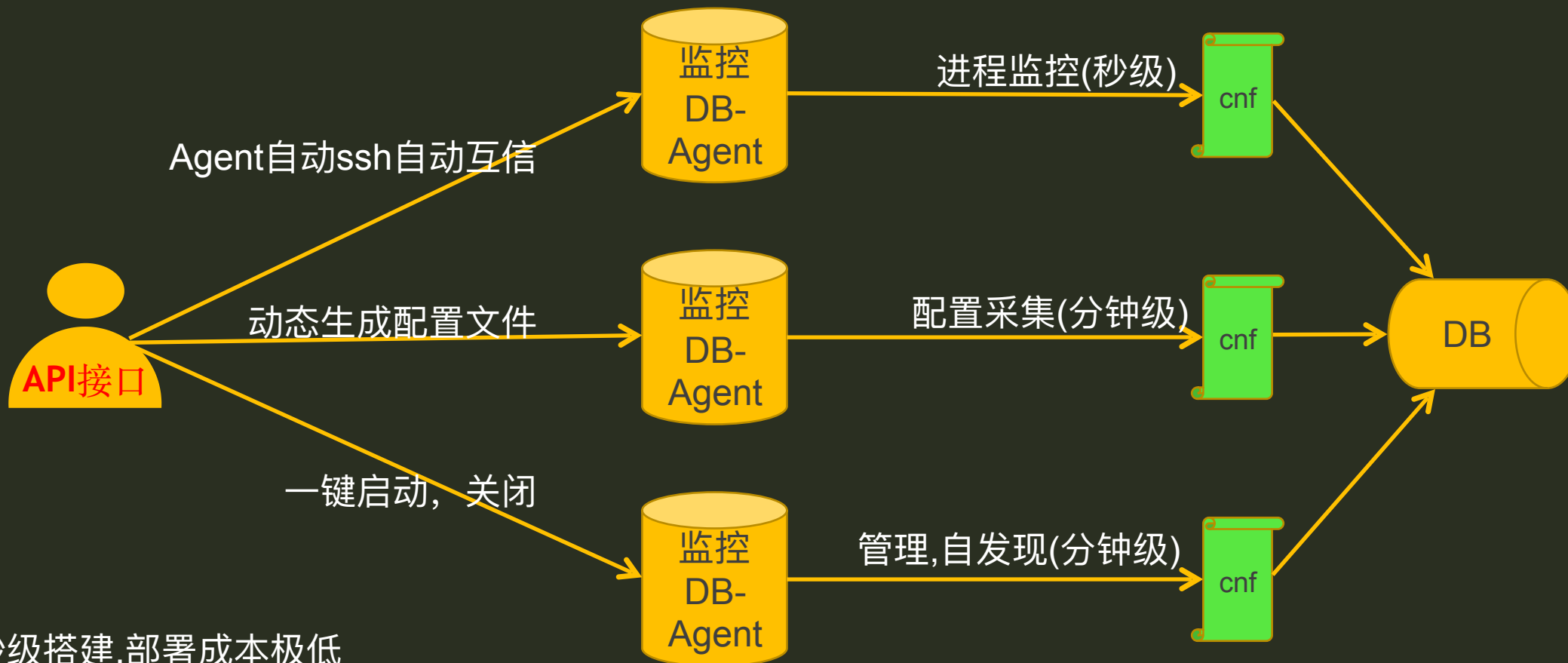
只保留最新一条日志

MHA管理1.0



1. 重度依赖SSH
2. 搭建过程复杂
3. 监控局限无法检查mha配置准确性
4. 架构复杂化、扩展性差、不适合大规模管理

MHA管理2.0



1. MHA秒级搭建,部署成本极低
2. 相比集中式管理更加轻量级、隐患少(跨网络心跳检测问题, manager单节点问题)
3. 通过任务调度完成mha自动化管理(启动、配置更新、切换等)
4. 自动化部署彻底摆脱基于SSH方案实现自动化的各种烦恼

MHA:EU

自动识别集群拓扑结构，生成配置文件
自动判断是否是多活集群并生成多机房配置信息



xg

wg

10.10.10.1:3306=>master

10.10.10.1:3306=>master

10.10.10.2:3306=>mha

10.10.10.2:3306=>mha

10.10.10.3:3306=>slave

10.10.10.3:3306=>slave

10.10.10.4:3306=>slave

启动MHA

关闭MHA

10.10.10.5:3311=>slave

启动MHA

关闭MHA

关闭

* 数据源: * DB: * 目标源:

* 备份线程: * 还原线程: * 复制类型: M->S S->S Delay

开始迁移

- 实例详细性检查 检查通过
- 版本检查 检查通过
- 数据源实例Binlog开启检查 检查通过
- 目标实例角色检查 检查通过
- 数据源实例Binlog_Format 检查通过
- 字符集检查 检查通过
- 复制用户检查 检查通过
- 额外的DB检查 检查通过

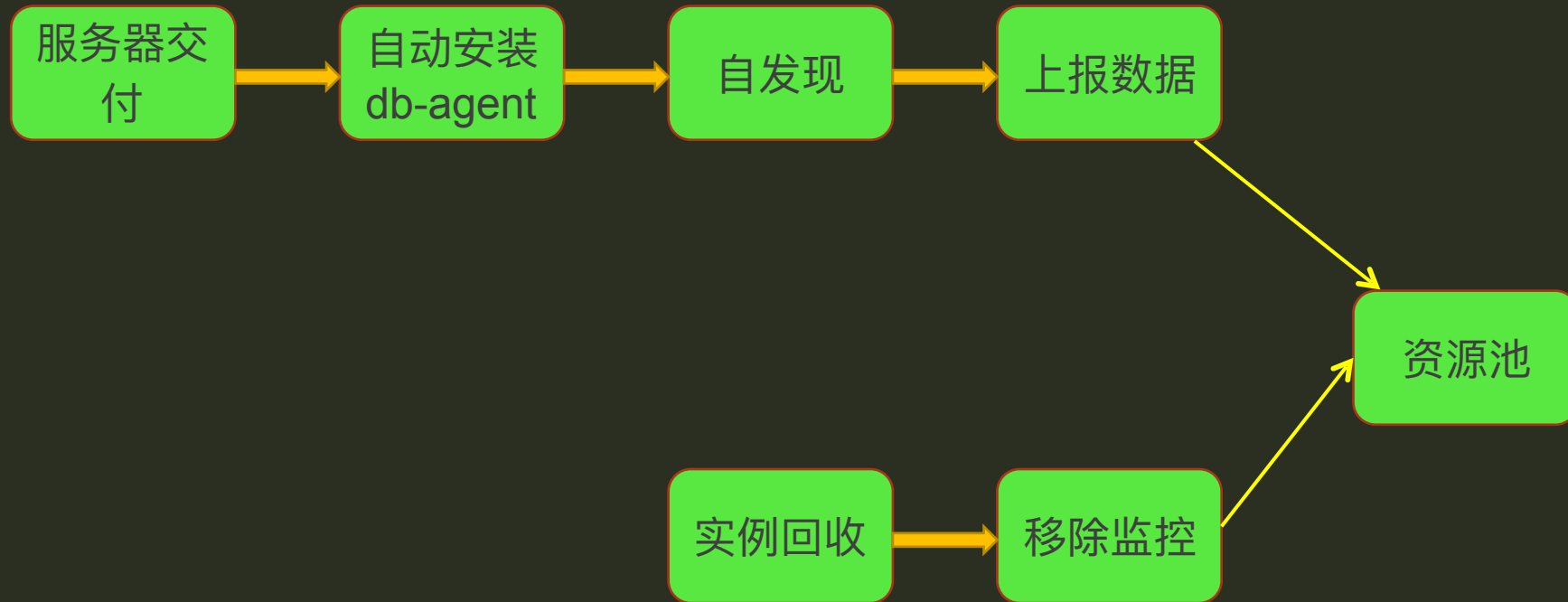
1 系统会自动将迁移任务分发到距离目标源最近的机房并从中找到空间合适、系统负载低的任务服务器(Tinker)

2 通过与资源池整合后该功能可以由调度系统自动调用完成搭建slave, 集群迁移等工作

3 接口化之后该功能有力的支撑了我厂DB灾备、DB上云、异地多活等重大项目

4 高效可靠的大规模实例迁移, 以上云为例:24小时内完成300套集群的迁移任务

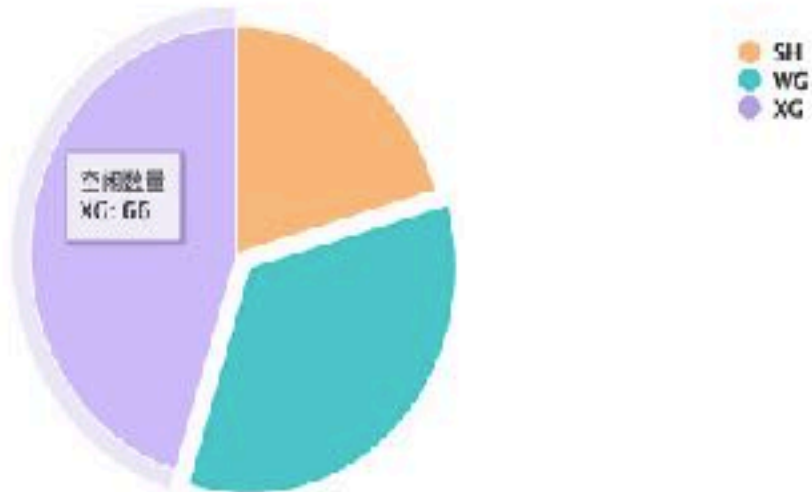
资源池管理



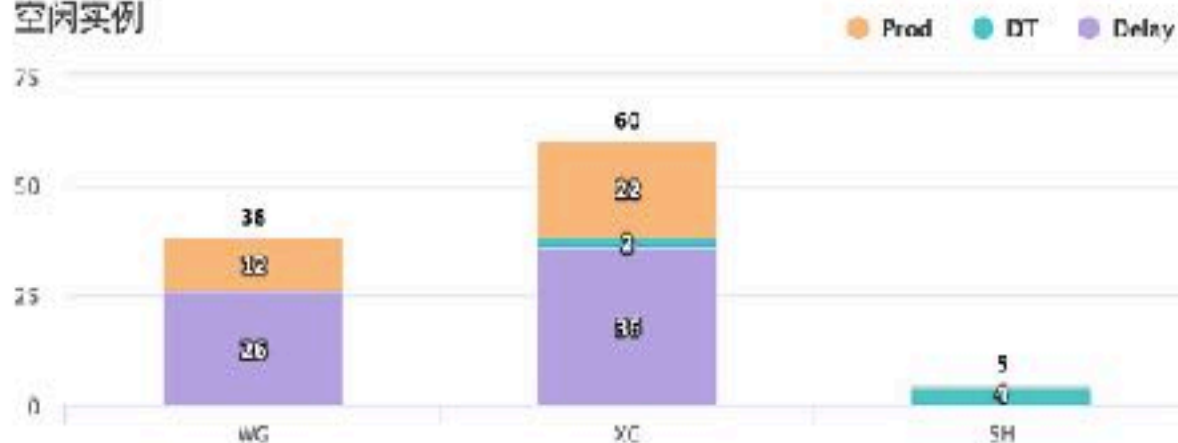
资源池管理



空闲机器

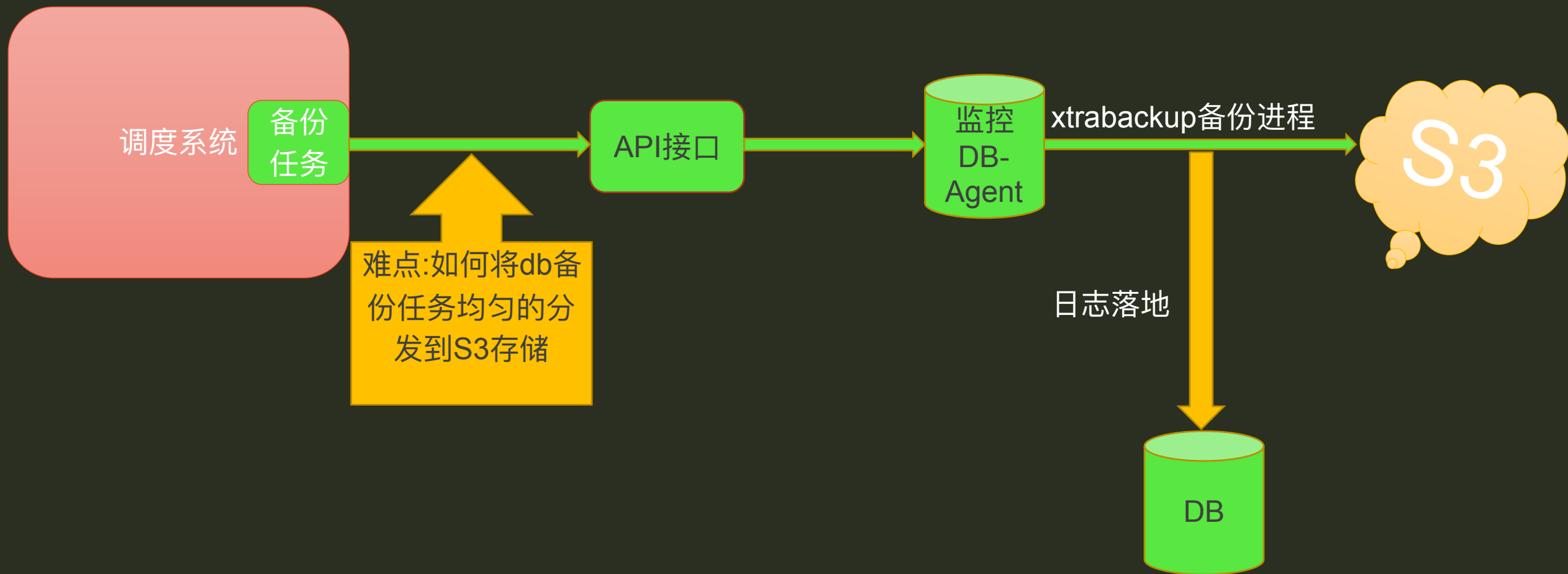


空闲实例



host	ServerName	agent_version	idc	服务型类型	操作
10.0.0.114	wg tidb-1	1.3.0	WG	5.6	3306 安装
10.0.0.135	wg rfo-master-1	1.3.0	WG	5.6	3306 安装
10.0.0.147	wg tom-java-1	1.3.0	WG	5.6	3306 安装
10.0.0.148	wg ind-active-1	1.3.0	WG	5.6	3306 安装
10.0.0.111	wg-dm-ny-354	1.3.0	WG	5.6	3306 安装
10.0.0.116	wg-dm-mysql-356	1.3.0	WG	5.6	3306 安装
10.0.0.119	wg-dm-mysql-352	1.3.0	WG	5.6	3306 安装
10.0.0.121	wg-dm-mysql-355	1.3.0	WG	5.6	3306 安装

备份系统的实现



授权



用户预览

General

Server级权限

DB级权限

Table级权限

SQL Preview

授权用户

超级口令

创建新用户

Save

MySQL 5.7.20 4.0.10.1 - MySQL 5.7.20

Host	DB	User	Select	Insert	Update	Delete	Create	Drop	Grant	References	Index	Alter	Create_tmp_table	Lock_tables	Create_view	Show_view	Create_routine	Alter_routine
10.0%	mysql	z_mysql_cal	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.0%	mysql	z_mysql_cal	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.0%	mysql	z_mysql_cal	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.0%	mysql	z_mysql_cal	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.0%	mysql	z_mysql_cal	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



过去

2015-2016(高压下快成长)

人肉

标准化

自动化

2018精细运营、数据驱动

API化

VDBA

当下

平台化

自助化

2017技术转型、赋能产研

智能化

数据驱动





IT大咖说
知识分享平台

Thanks