

# 大型券商的ELK日志运维实践









































李凯平 11-10,2018 议题

01 背景和目标

03 应用效果

05 后续工作

02 技术方案

04 实践中的问题及解决方案

#### 1、背景和目标 | 背景

01

指标类监控工具多,但 缺乏对日志的利用,完 整的运维数据包括日志、 指标和配置 02

日志数据分散,保存时间短,无法满足运维即 席检索及监管审计要求 03

日志中的业务价值有待 挖掘,如用户行为、应 用性能、安全攻击等

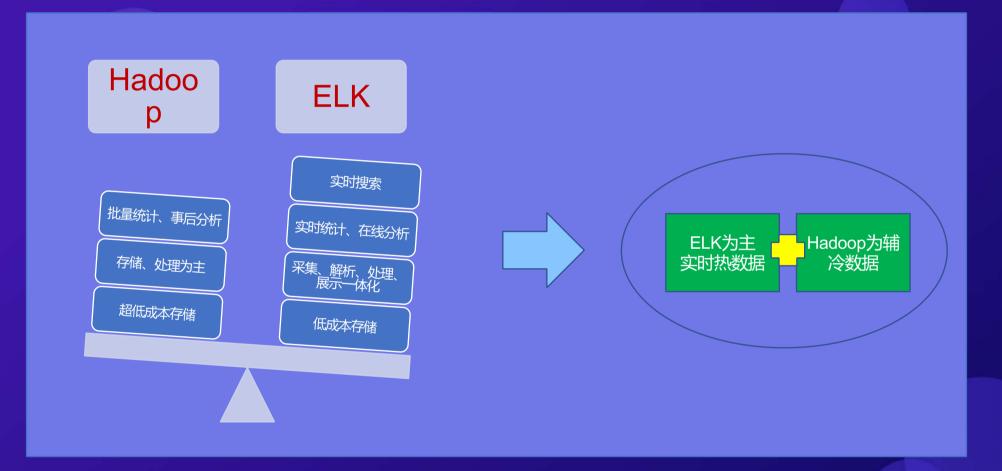
#### 1、背景和目标 | 目标

- 增加日志处理能力,结合配置数据,增强现有的运维能力
- 业务监控,应用监控,智能告警,容量辅助决策
- 建立可持续扩展的运维数据仓库,对接更多的数据源,提供数据消费能力

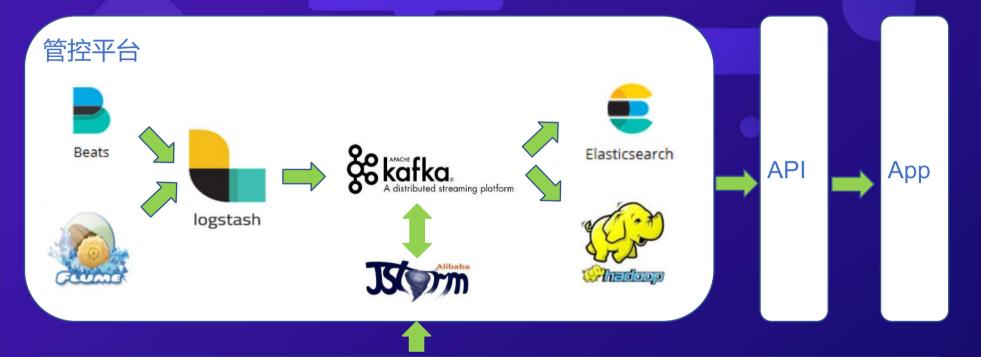
调用链 指标 强调明细 强调聚合 日志 强调事件



# 2、技术方案 | 选型



# 2、技术方案 | 系统架构



#### 管控平台







#### 2、技术方案 | 其他关注点

企业级应用还需要关注以下几点:

1 2

数据规范:建 数据标签: 与CMDB结合, 数据经过治 理才能有效 应用

容量管理:
了解重要组
件的容量性
能基线
才能高效运

管控与自助 化:具有集 中化的管控 和自助化能 力,才能有 效管理

4

存储规划: 区分冷热数 据,做好数 据的存储和 使用规划

# 2、技术方案 | 更多思考 日志的分类 -分类:业务、中间件、操作系统、网络、存储、安全 -要素5w:when、where、what... —级别:trace、debug、info、warn、error、critical 日志的管理 -模版化 日志的存储 一与CMDB结合 -冷热数据分离 4 日志的展示 -业务视角

# 3、应用效果 | 被采集系统



50+个系统 4500+主机



采集了全部网络设备日志、重要系统应用日志、 Linux及windows系统日志



集中化管理

#### 3、应用效果 | 集群情况

- 100+解析节点、20个ES服务器节点,日采集3.5T日 志
- 秒级查询, 秒级告警





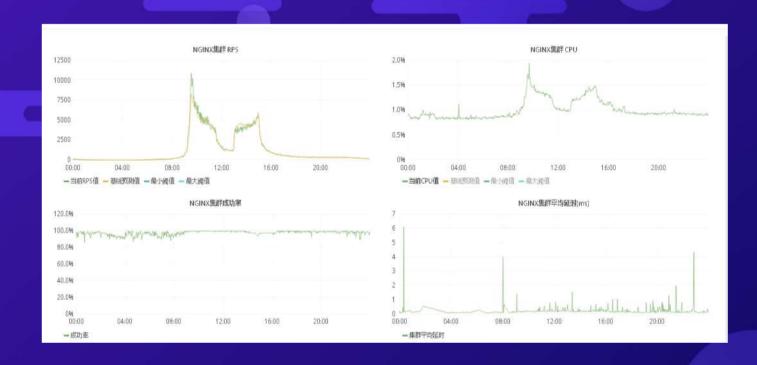
#### 3、应用效果 | 应用情况

- 日志检索
- 业务指标分析
- 调用链



# 3、应用效果 | 应用情况

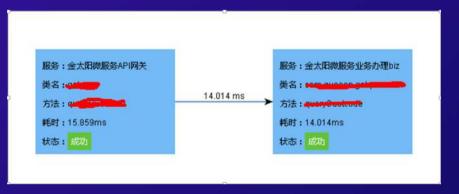
- 日志检索
- 业务指标分析
- 调用链



#### 3、应用效果 | 应用情况

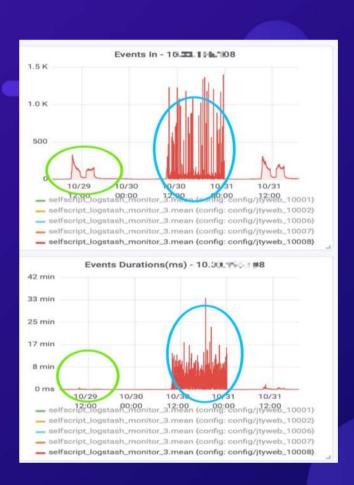
- 日志检索
- 业务指标分析
- 调用链





#### 4、实践中的问题及解决方案 | 问题

- 1.时延
- 2.数据的漏采和非标



#### 4、实践中的问题及解决方案 | 方案

做好自监控 做好自监控 做好自监控 建立好基线

建立好基线

建立好基线

重要的事情说三遍!!!

#### 4、实践中的问题及解决方案 | 方案

FileBeat Logstash 数据校验 Logstash 数据校验 logstash kafka logstash jStorm ES

- 1. 时间标记,时延监控
- 2. 日志自监控
- 3. 漏文件和文件解析失败
- 4. 文件级别采集监控

name \$	lang 💠	script ≑
delay_log_to_filebeat	expression	doc['normalFields.collecttime']-doc['timestamp']
delay_filebeat_to_prekafka	expression	doc['normalFields.logchecktime']-doc['normalFields.collecttime']
delay_istorm_to_index	expression	doc['indexTime']-doc['deserializerTime']
delay_prekafka_to_jstorm	expression	doc['deserializerTime']-doc['normalFields.logchecktime']
delay_log_to_index	expression	doc['indexTime']-doc['timestamp']

#### 找出Filebeat、Logstash能力基线 – 示例

实例数	pipeline workers	output workers	old heap(%)	CPU(%)	CPU(总%)	TPS	备注	肘延
1-	4.	1.	-0	54%	54%	3200	ÿ.	j.
1/	8.	1.	~	86%	86%	4900	ē	2
1.	12 <i>-</i>	1-		90%	90%	5400	ę.	٥
1-	16-	1.	->	90%	90%	5500₽	ě.	٥
		<u>1</u>		22	100 M		e	- 4
2.	4.	14	<b>-</b> a	43%	86%	2700₽	p.	ø
2	4.	1.0	~	43%	86%	2700	内存为 2G	1.5ms
2	4.	1.	18%	35%	70%	3700	grok + geoip	1ms
2	8.	1.	25%	45%	90%	4800	grok + geoip	2ms
2-	8-	1,-	70%	47%	94%	2900		
2	12.	1.	70%	48%	96%	2900	Ē	ý
ē		10		0	W. W.		Ø.	ě
4.	4.	1.	70%	24%	96%	1530-	ø	d.
4	8-	1.	70%	24%	96%	1580	e .	6ms

Logstash it	上滤器头	9空。							
Logstash		512	2s	8294	5266944	18%	0.7%	1.	44
Logstash	-,,	1024	2s.	8806	5592064	17.5%	0.8%	1.	
Logstash		2048	25-	8601	5462016	16%	0.9%	1	+1
Logstash	-,-	1024	4s	8601	5462016	16%	0.8%	1.	20
Logstash	-,,	1024	8s	8396	5331968	16%	0.8%	1.	- 41
Logstash	1-	1024	2s.	8601	5462016	17.4%	0.8%	1	
Logstash	2	1024	2s.	8601	5462016	17%	0.8%	1.	41
Logstash	4	1024	2s.	8601	5462016	16.8%	0.8%	1	- 2
Logstash	8	1024	2s.	8192-	5201920	16.6%	0.8%	1.	45
Logstash it	上滤器头	过空,file	beat f	立于不同机子	40	1/2		a)	at .
Logstash	-,-	1024	2s.	3072	1950720	6.2%	0.6%	4	11763
Logstash	74	1024	2s.	4096-	2600960	8.1%	0.5%	3	11353
Logstash		1024	2s.	6144	3901440	11.9%	0.6%	2	11112
Logstash	-,-	1024	2s.	9420	5982208	19.9%	0.8%	1	9117
Logstash 过件,2个文 Logstash			F-17 - 115	at 进程,多个 总:- 9011。	个文件源,一 文件源:4。	19%	原 <b>6G,</b> 0.4%。	一个文 1-	件源是空 8700
Logstash 过	滤器:	grok, g	geoip,	file					
Logstash		1024	2s.	4300-	2731008	8.8%	0.5%	1.	e.
Logstash 현	滤器:	grok, g	geoip,	file; filebea	t 有 4 个实	例位于不	同机子		

- > 综上, filebeat -> logstash
  - 1. logstash 的处理能力: 1.2w-
  - 2. filebeat 单个进程处理能力: 9000
- ▶ 基于以上2个限制。
  - 1. n 个 filebeat 对接到一个 logstash,则每个 filebeat 的处理能力为 min{1.2w / n, 9000}
  - 2. 1个 filebeat 接 m 个文件源,则每个文件的读取速度为 min{1.2w/n,9000} / m-

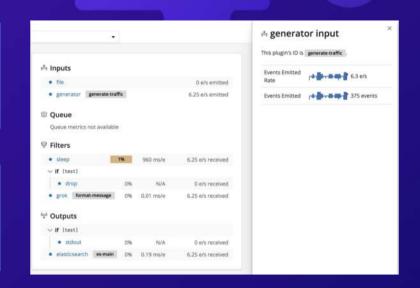
# Logstash性能优化



问:Logstash性能受到很多因素的影响,除了做好参数调优,通过实验得到基线,还需要做什么工作?



答:利用pipeline工具发现性能不足的原因,修改代码实现方式,持续的观察性能瓶颈点,经过持续改进,直到达到要求。优化的代码可以提升性能好多倍。



Pipeline工具可以看到代码级耗时情况,是 性能调优利器。

#### ES性能优化



问:服务器异构,如何让shard更多的分布在SAS盘服务器?如何提高磁盘的利用率?

问:同一个index的shard数远小于es实例数,并且统一零点创建index,导致index实例只会分布在部分磁盘



答:配置zone,进行shard个数优化



问:在提升了磁盘利用率之后,所有磁盘iostat util都非常高,写入频次很高,吞吐量很低,如何提高磁盘吞吐量?



答:调整压缩方式、refresh时间间隔、merge线程数、translog sync模式、translog flush阈值

# ES优化后效果对比

_		>	,
4	-	'n	
ď	J	U	
X	J	ے	
1	1	Ī	

叉	
<del></del>	
八文	
11	

_	7
7	
ď	ä
ï	ī
ı	н

avg-cpu:	Suser.	inice	'aystem	%iowait	Asteal	*idle									
	2.59	0.00	0.42	15.89	0.00	61.11									
Device:		EEqm/s	wrqm/s	E/8	w/=	rkB/s	wkB/s	avgrq-sz	avgqu-sz	await	r_await	w_await	svetm	*util	
5 dm		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
adl.		0.00	288.00	0.00	146.00	0.00	1856.00	25.42	1.84	12.64	0.00	12.64	6.82	99.60	
sdh.		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
adi.		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
sdd		0.00	182.00	0.00	116.00	0.00	1316.00	22.69	1.18	10.31	0.00	10.31	7.97	92.50	
adg		0.00	395.00	0.00	136.00	0.00	2312.00	34.00	2.19	16.15	0.00	16.15	7.33	99.70	
sda		0.00	278.00	0.00	114.00	0.00	1572.00	27.58	1.26	11.04	0.00	11.04	8.69	99.10	
nde		0.00	393.00	0.00	293.00	0.00	25872.00	176.60	15.07	51.38	0.00	51.38	3.25	95.20	
sdj		0.00	226.00	0.00	120.00	0.00	1492.00	24.87	1.40	11.73	0.00	11.73	6.01	96 10	
adb		0.00	67.00	0.00	51.00	0.00	2764.00	108.39	0.38	7.49	0.00	7.49	1.35	6.90	
sdc		0.00	24.00	0.00	5.00	0.00	112.00	44.60	0.04	9.00	0.00	9.00	9.00	4.50	
a dis		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
sdf		0.00	56.00	0.00	53.00	0.00	648.00	24.45	0.39	7.26	0.00	7.26	2.13	11.30	
dia-O		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
dm-1		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
dm-Z		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
dm-3		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

Device:	rrqm/s	WEGM/S	E/8	w/s	rkB/s	wkB/s	avgrq-sz	avgqu-sz	await	r_await	w_await	syctm	suti1
sdw	0.00	0.00	0.00	1.00	0.00	4.00	8.00	0.00	0.00	0.00	0.00	0.00	0.00
adb	0.00	155.00	0.00	126.00	0.00	1412.00	22.41	0.89	7.09	0.00	7.09	6.38	80.40
adc	0.00	121.00	0.00	125.00	0.00	1048.00	16.77	0.90	7.22	0.00	7.22	6.44	80.50
øde	0.00	156.00	0.00	133.00	0.00	1236.00	18.59	0.84	6.32	0.00	6.32	5.47	72.70
sdg	0.00	137.00	0.00	139.00	0.00	1172.00	16.86	0.98	7.07	0.00	7.07	5.74	79.80
odi.	0.00	90.00	0.00	113.00	0.00	868.00	15.36	0.67	5.95	0.00	5.95	5.94	67.10
tbs	0.00	597.00	0.00	222.00	0.00	3844.00	34.63	2.40	10.18	0.00	10.18	4.39	97.50
sdn	0.00	76.00	0.00	115.00	0.00	848.00	14.37	0.70	5.96	0.00	5.96	5.75	67.80
sdo	0.00	402.00	0.00	228.00	0.00	3088.00	27.09	2.64	11.58	0.00	11.58	4.39	100.00
er diba	0.00	151.00	0.00	156.00	0.00	3700.00	47.44	1.85	12.90	0.00	12.90	4.59	71.60
phe	0.00	116.00	0.00	109.00	0.00	968.00	17.76	0.66	6.05	0.00	6.05	5.55	60.50
sda	0.00	256.00	0.00	180.00	0.00	2332.00	25,91	2.96	15.88	0.00	15.88	5.47	98.50
qbp	0.00	107.00	0.00	115.00	0.00	960.00	16.70	0.80	6.94	0.00	6.94	6.51	74.90
adf	0.00	104.00	0.00	118.00	0.00	944.00	16.00	0.73	6.19	0.00	6.19	5.75	67.80
sdv	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
bbc	0.00	126.00	0.00	134.00	0.00	1140.00	17.01	0.87	6.51	0.00	6.51	6.15	82.40
20 CRIM	0.00	93.00	0.00	129.00	0.00	972.00	15.07	0.96	7-41	0.00	7.41	7.01	90.40
sdl	0.00	73.00	0.00	125.00	0.00	548.00	13.57	0.77	6.15	0.00	6.15	5.75	71.90
sdr	0.00	99.00	0.00	99.00	0.00	1068,00			8.66	0.00	8.66	7.92	78.40
a cile	0.00	80.00	0.00	112.00	0.00	828.00		0.82	7.25	0.00	7.25	6.95	77.80
ada	0.00	127.00	0.00	121.00	0.00	1060.00	17.52	0.77	6.37	0.00	6.37	5.97	72.20
sdt	0.00	329.00	0.00	234.00	0.00	2632.00	24.21	2.42	10.24	0.00	10.24	4.26	99.70

avg-cpu:	Super		*system		*steal	sidle		(A) (P) **	CD + 100 - 1	1				
	36.67	0.00	2.79	7.76	0.00	52.79								
Devices		EEqm/s	wrqm/s	E/8	10/0	EkB/s	wkB/s	avgrq-sz	avgqu-sz	await	E_await	w_await	SUCTE	suti1
ar chiar		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
or ct.ca		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
db		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
sde		0.00	83.00	0.00	247.00	0.00	27828.00	225.33	25.72	104.14	0.00	104.14	1.42	35.10
sche		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
# clf		0.00	27.00	0.00	23.00	0.00	500.00	50.43	0.23	10.30	0.00	10.30	5.96	13.70
mcl.i.		0.00	240.00	1.00	377.00	4.00	42976.00	227.41	42.51	112.46	23.00	112.70	1.02	60.00
ired ed.		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
a cher		0.00	145.00	0.00	22.00	0.00	692.00	62.91	0.19	8.95	0.00	8.95	6.09	13.40
CDS		0.00	20.00	0.00	12.00	0.00	148.00	24.67	0.08	6.67	0.00	6.67	5.58	6.70
r cDh.		0.00	102.00	0.00	22.00	0.00	528.00	48.00	0.19	0.02	0.00	0.02	6.73	14.80
r clm.		0.00	14.00	0.00	15.00	0.00	2092.00	305.60	0.25	16.47	0.00	16.47	2.93	4.40
releo:		0.00	83.00	0.00	36.00	0.00	5700.00	300.00	0.51	13.45	0.00	13.45	3.45	13.10
ed.s		0.00	41.00	0.00	16.00	0.00	268.00	33.50	0.14	8.44	0.00	8.44	6.50	10.40
rdq:		0.00	46.00	0.00	12.00	0.00	256.00	42.67	0.07	5.50	0.00	5.50	5.50	6.60
s dp		0.00	186.00	0.00	228.00	0.00	35276.00	309.44	20.40	163.86	0.00	163.86	2.16	49.20
sdt.		0.00	37.00	0.00	16.00	0.00	256.00	32.00	0.10	6.56	0.00	6.56	6.30	10.20
m char		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
e chie		0.00	9.00	0.00	20.00	0.00	4108.00	410.80	0.46	22.75	0.00	22.75	3.55	7.10
o chu		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
m cl. k		0.00	21.00	1.00	15.00	4.00	2808.00	351.50	0.26	16.44	6.00	17.13	3.19	5.10
er clara.		0.00	23.00	0.00	6.00	0.00	532.00	177.33	0.04	6.50	0.00	6.50	5.00	3.00
rdx		0.00	136.00	0.00	134.00	0.00	3620.00	54.03	5.91	44.11	0.00	44.11	1.04	24.60
Im - O		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3m-1		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3m-2		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
dm-3		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

充分均衡IO负载,彻底解决ES写入 石吐量不够的问题,性能显著提高5+倍

# 5、后续工作

Zabbix 4.0的原生 Elasticsearch解决方

案



全业务链监控



平台的集中管控和 自助化能力



数据服务API的梳理

#### 5、公司招聘





诚聘英才!

# 谢谢聆听!





专业、垂直、纯粹的 Elastic 开源技术交流社区 <a href="https://elasticsearch.cn/">https://elasticsearch.cn/</a>

