AI在内核故障分析的应用实践

利用AI技术赋能运维场景

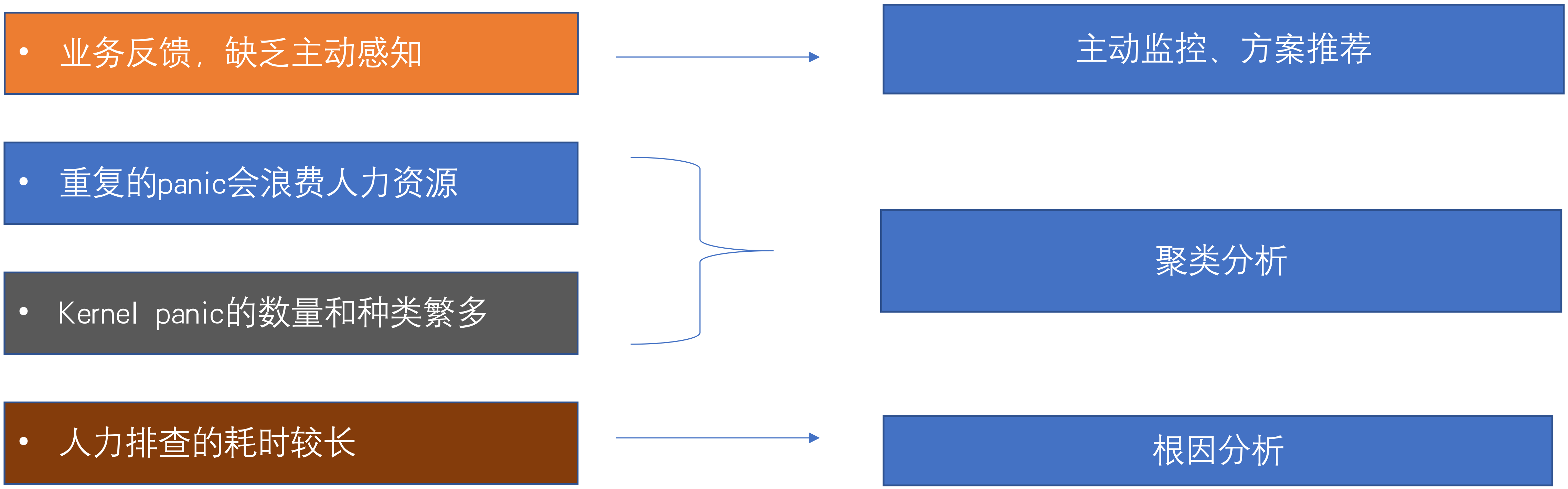
字节跳动系统部STE团队 － 王立新

ByteDance字节跳动

大纲

- 字节数据中心内核故障的概述
- 字节内核故障分析和监控
- 总结和展望

从手工运维到智能运维



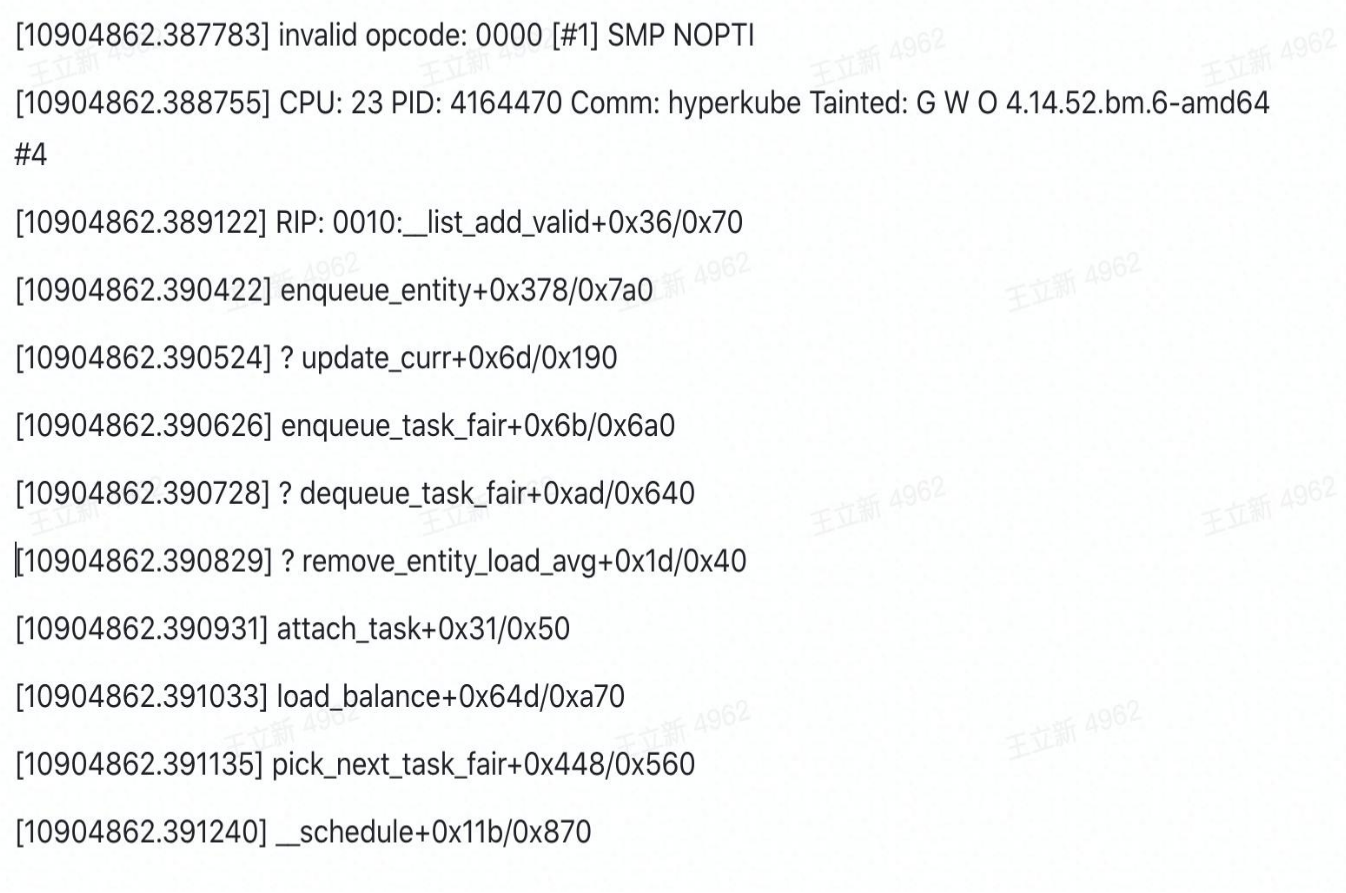
- 业务反馈，缺乏主动感知
- 重复的panic会浪费人力资源
- Kernel panic的数量和种类繁多
- 人力排查的耗时较长

主动监控、方案推荐

聚类分析

根因分析

内核故障分析之聚类分析



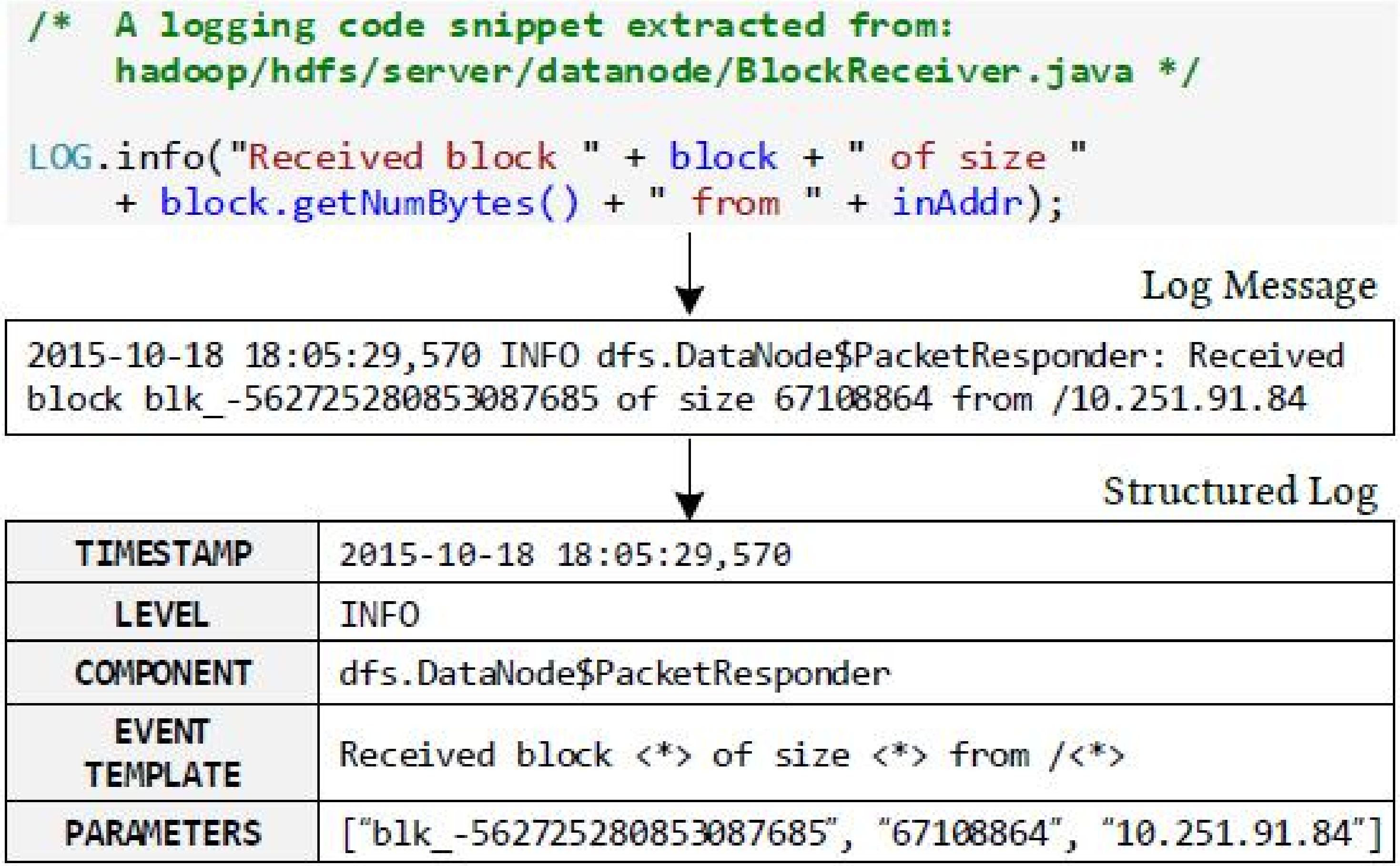
```
[10904862.387783] invalid opcode: 0000 [#1] SMP NOPTI
[10904862.388755] CPU: 23 PID: 4164470 Comm: hyperkube Tainted: G W O 4.14.52.bm.6-amd64
#4
[10904862.389122] RIP: 0010:__list_add_valid+0x36/0x70
[10904862.390422] enqueue_entity+0x378/0x7a0
[10904862.390524] ? update_curr+0x6d/0x190
[10904862.390626] enqueue_task_fair+0x6b/0x6a0
[10904862.390728] ? dequeue_task_fair+0xad/0x640
[10904862.390829] ? remove_entity_load_avg+0x1d/0x40
[10904862.390931] attach_task+0x31/0x50
[10904862.391033] load_balance+0x64d/0xa70
[10904862.391135] pick_next_task_fair+0x448/0x560
[10904862.391240] __schedule+0x11b/0x870
```

内核版本

故障类型

Call trace

聚类分析：日志解析



```
/*  A logging code snippet extracted from:
    hadoop/hdfs/server/datanode/BlockReceiver.java */

LOG.info("Received block " + block + " of size "
    + block.getNumBytes() + " from " + inAddr);
```

```
2015-10-18 18:05:29,570 INFO dfs.DataNode$PacketResponder: Received
block blk_-562725280853087685 of size 67108864 from /10.251.91.84
```

TIMESTAMP	2015-10-18 18:05:29,570
LEVEL	INFO
COMPONENT	dfs.DataNode$PacketResponder
EVENT TEMPLATE	Received block <*> of size <*> from /<*>
PARAMETERS	[“blk_-562725280853087685”, “67108864”, “10.251.91.84”]

基于聚类的算法

Logcluster等

基于共同模块

Spell，Darin等

基于语义

logEvent2vec

聚类分析：日志解析

1. 树深：h

2. Token的相似度

$$simSeq = \frac{\sum_{i=1}^{n} equ(seq_1(i), seq_2(i))}{n},$$

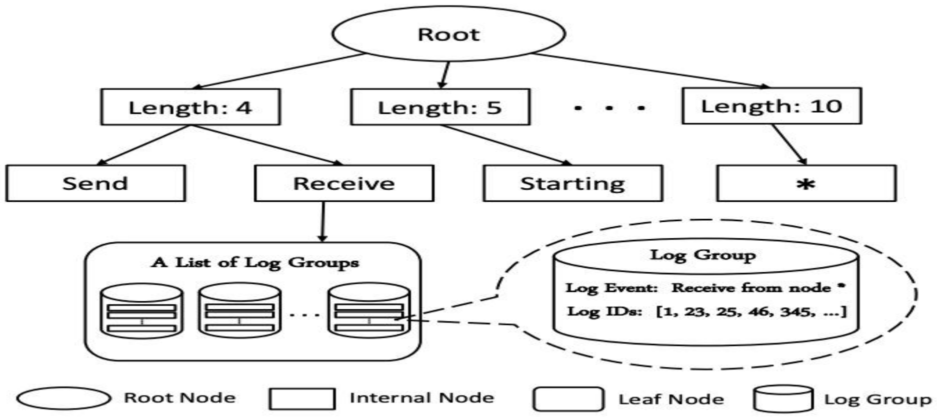
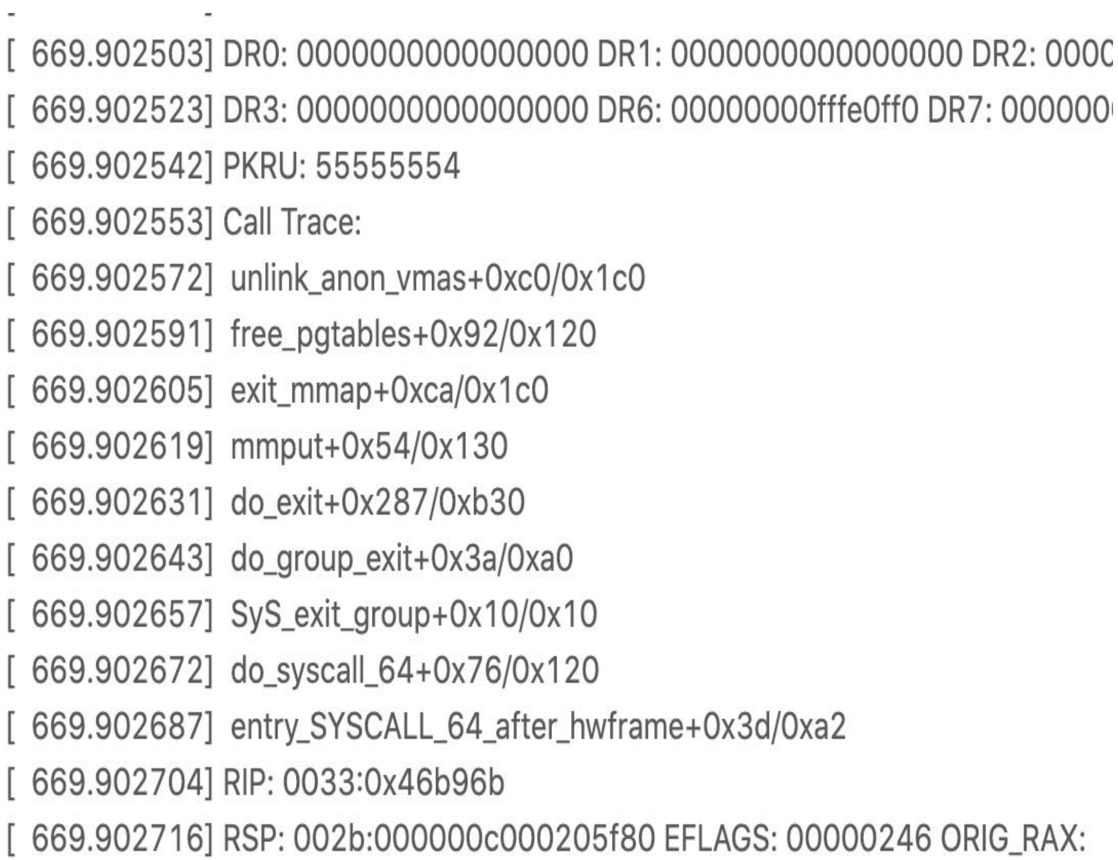


Fig. 2: Structure of Parse Tree in Drain ($depth = 3$)

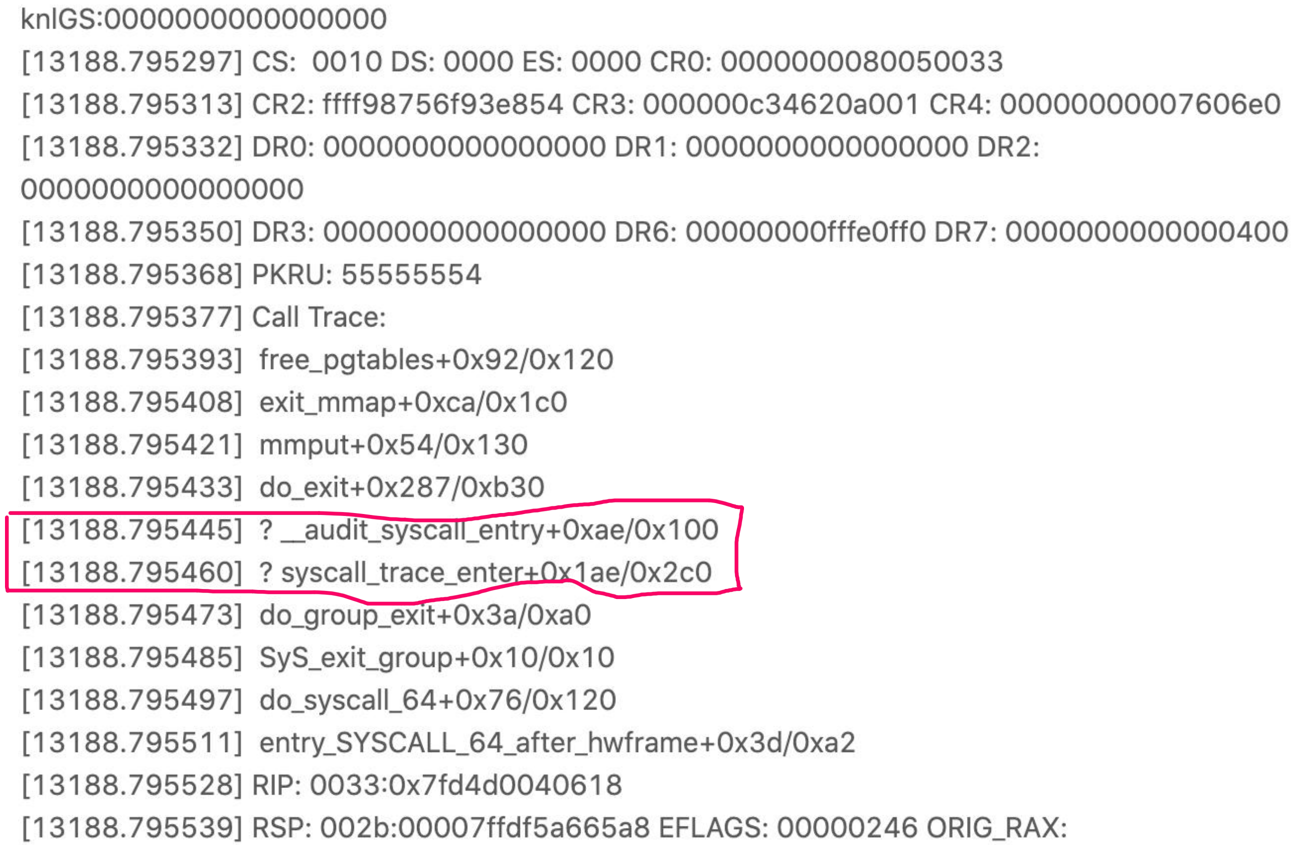
```
Kernel panic — not syncing：NMI：Not continuing
general protection fault：0000 [#1] SMP NOPTI
Oops：0000 [#1] SMP
BUG：unable to handle kernel paging request at ffff95a348f69ac8
Oops：0002 [#1] SMP NOPTI
Kernel panic — not syncing：Fatal hardware error！
```

```
Kernel panic — not syncing：NMI：Not continuing
general protection fault：* SMP NOPTI
Oops：* SMP
BUG：unable to handle kernel paging request at *
Oops：* SMP NOPTI
Kernel panic — not syncing：Fatal hardware error！
```

聚类分析：call trace



```
[  669.902503] DR0: 0000000000000000 DR1: 0000000000000000 DR2: 000
[  669.902523] DR3: 0000000000000000 DR6: 00000000fffe0ff0 DR7: 000000
[  669.902542] PKRU: 55555554
[  669.902553] Call Trace:
[  669.902572]  unlink_anon_vmas+0xc0/0x1c0
[  669.902591]  free_pgtables+0x92/0x120
[  669.902605]  exit_mmap+0xca/0x1c0
[  669.902619]  mmput+0x54/0x130
[  669.902631]  do_exit+0x287/0xb30
[  669.902643]  do_group_exit+0x3a/0xa0
[  669.902657]  SyS_exit_group+0x10/0x10
[  669.902672]  do_syscall_64+0x76/0x120
[  669.902687]  entry_SYSCALL_64_after_hwframe+0x3d/0xa2
[  669.902704] RIP: 0033:0x46b96b
[  669.902716] RSP: 002b:000000c000205f80 EFLAGS: 00000246 ORIG_RAX:
```



```
knlGS:0000000000000000
[13188.795297] CS:  0010 DS: 0000 ES: 0000 CR0: 0000000080050033
[13188.795313] CR2: ffff98756f93e854 CR3: 000000c34620a001 CR4: 00000000007606e0
[13188.795332] DR0: 0000000000000000 DR1: 0000000000000000 DR2:
0000000000000000
[13188.795350] DR3: 0000000000000000 DR6: 00000000fffe0ff0 DR7: 0000000000000400
[13188.795368] PKRU: 55555554
[13188.795377] Call Trace:
[13188.795393]  free_pgtables+0x92/0x120
[13188.795408]  exit_mmap+0xca/0x1c0
[13188.795421]  mmput+0x54/0x130
[13188.795433]  do_exit+0x287/0xb30
[13188.795445]  ? __audit_syscall_entry+0xae/0x100
[13188.795460]  ? syscall_trace_enter+0x1ae/0x2c0
[13188.795473]  do_group_exit+0x3a/0xa0
[13188.795485]  SyS_exit_group+0x10/0x10
[13188.795497]  do_syscall_64+0x76/0x120
[13188.795511]  entry_SYSCALL_64_after_hwframe+0x3d/0xa2
[13188.795528] RIP: 0033:0x7fd4d0040618
[13188.795539] RSP: 002b:00007ffdf5a665a8 EFLAGS: 00000246 ORIG_RAX:
```

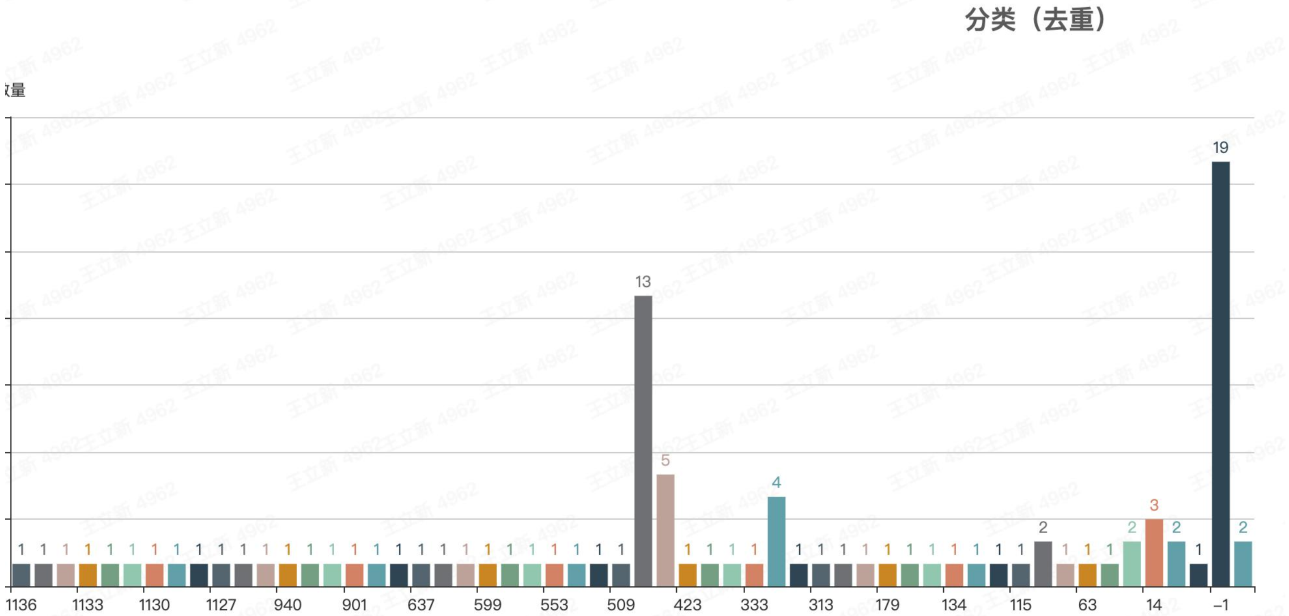
聚类分析：call trace匹配

函数调用一致顺序

可疑函数的权重处理

函数频次的权重处理

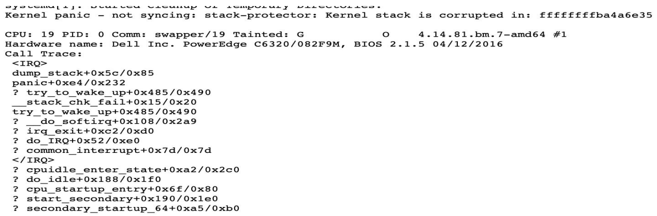
wi



$$sim = \frac{\max(\sum wi\ (seq\ \ 1, seq\ \ 2))}{l\ max}$$

内核分析之系统设计

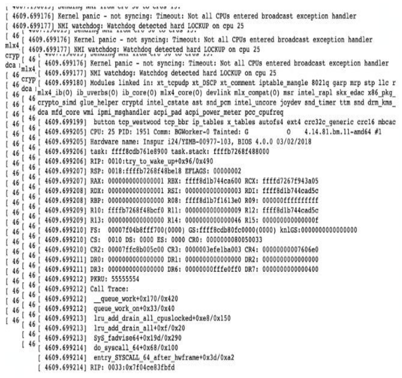
Online kernel panic



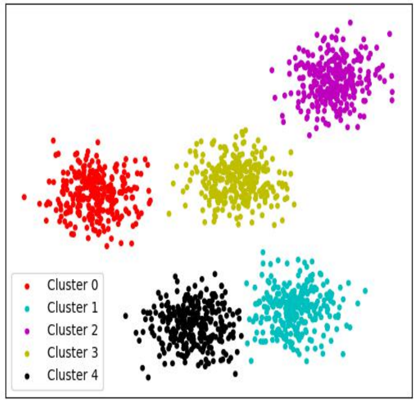
```
Kernel panic - not syncing: stack-protector: Kernel stack is corrupted in: ffffffffba4a6e35

CPU: 19 PID: 0 Comm: swapper/19 Tainted: G           O    4.14.81.bm.7-amd64 #1
Hardware name: Dell Inc. PowerEdge C6320/082F9M, BIOS 2.1.5 04/12/2016
Call Trace:
 <IRQ>
 dump_stack+0x5c/0x85
 panic+0xe4/0x232
 ? try_to_wake_up+0x485/0x490
 __stack_chk_fail+0x15/0x20
 try_to_wake_up+0x485/0x490
 ? __do_softirq+0x108/0x2a9
 ? irq_exit+0xc2/0xd0
 ? do_IRQ+0x52/0xe0
 ? common_interrupt+0x7d/0x7d
 </IRQ>
 ? cpuidle_enter_state+0xa2/0x2c0
 ? do_idle+0x188/0x1f0
 ? cpu_startup_entry+0x6f/0x80
 ? start_secondary+0x190/0x1e0
 ? secondary_startup_64+0xa5/0xb0
```

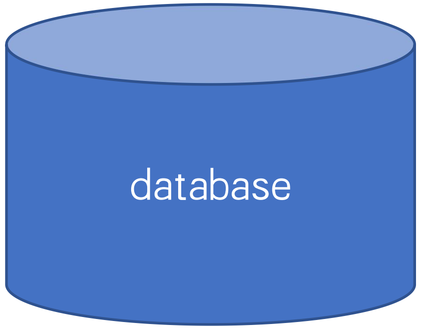
Historical kernel panic



```
[ 4609.699176] Kernel panic - not syncing: Timeout: Not all CPUs entered broadcast exception handler
[ 4609.699177] NMI watchdog: Watchdog detected hard LOCKUP on cpu 25
[ 4609.699176] Kernel panic - not syncing: Timeout: Not all CPUs entered broadcast exception handler
[ 4609.699177] NMI watchdog: Watchdog detected hard LOCKUP on cpu 25
[ 4609.699176] Kernel panic - not syncing: Timeout: Not all CPUs entered broadcast exception handler
[ 4609.699177] NMI watchdog: Watchdog detected hard LOCKUP on cpu 25
[ 4609.699180] Modules linked in: xt_tcpudp xt_DSCP xt_comment iptable_mangle 8021q garp mrp stp llc r
mlx4_ib(O) ib_uverbs(O) ib_core(O) mlx4_core(O) devlink mlx_compat(O) msr intel_rapl skx_edac x86_pkg_
crypto_simd glue_helper cryptd intel_cstate ast snd_pcm intel_uncore joydev snd_timer ttm snd drm_kms_
dca mfd_core wmi ipmi_msghandler acpi_pad acpi_power_meter pcc_cpufreq
[ 4609.699199]  button tcp_westwood tcp_bbr ip_tables x_tables autofs4 ext4 crc32c_generic crc16 mbcac
[ 4609.699205] CPU: 25 PID: 1951 Comm: BGWorker-0 Tainted: G           O    4.14.81.bm.11-amd64 #1
[ 4609.699205] Hardware name: Inspur i24/YZMB-00977-103, BIOS 4.0.0 03/02/2018
[ 4609.699206] task: ffff8cdb761e8900 task.stack: ffffb7268f488000
[ 4609.699206] RIP: 0010:try_to_wake_up+0x96/0x490
[ 4609.699207] RSP: 0018:ffffb7268f48be18 EFLAGS: 00000002
[ 4609.699207] RAX: 0000000000000001 RBX: ffff8d1b744ca600 RCX: ffffd7267f943a05
[ 4609.699208] RDX: 0000000000000001 RSI: 0000000000000003 RDI: ffff8d1b744cad5c
[ 4609.699208] RBP: 0000000000000000 R08: ffff8d1b7f1613e0 R09: 000000ffffffffff
[ 4609.699209] R10: ffffb7268f48bcf0 R11: 0000000000000009 R12: ffff8d1b744cad5c
[ 4609.699209] R13: 0000000000000000 R14: 0000000000000046 R15: 000000000000000f
[ 4609.699210] FS:  00007f04b8fff700(0000) GS:ffff8cdb80fc0000(0000) knlGS:0000000000000000
[ 4609.699210] CS:  0010 DS: 0000 ES: 0000 CR0: 0000000080050033
[ 4609.699210] CR2: 00007ffc8b005c00 CR3: 0000003efe1ba003 CR4: 00000000007606e0
[ 4609.699211] DR0: 0000000000000000 DR1: 0000000000000000 DR2: 0000000000000000
[ 4609.699211] DR3: 0000000000000000 DR6: 00000000fffe0ff0 DR7: 0000000000000400
[ 4609.699212] PKRU: 55555554
[ 4609.699212] Call Trace:
[ 4609.699212]  __queue_work+0x170/0x420
[ 4609.699212]  queue_work_on+0x33/0x40
[ 4609.699213]  lru_add_drain_all_cpuslocked+0xe8/0x150
[ 4609.699213]  lru_add_drain_all+0xf/0x20
[ 4609.699213]  SyS_fadvise64+0x19d/0x290
[ 4609.699214]  do_syscall_64+0x68/0x100
[ 4609.699214]  entry_SYSCALL_64_after_hwframe+0x3d/0xa2
[ 4609.699214] RIP: 0033:0x7f04ce83fbfd
```



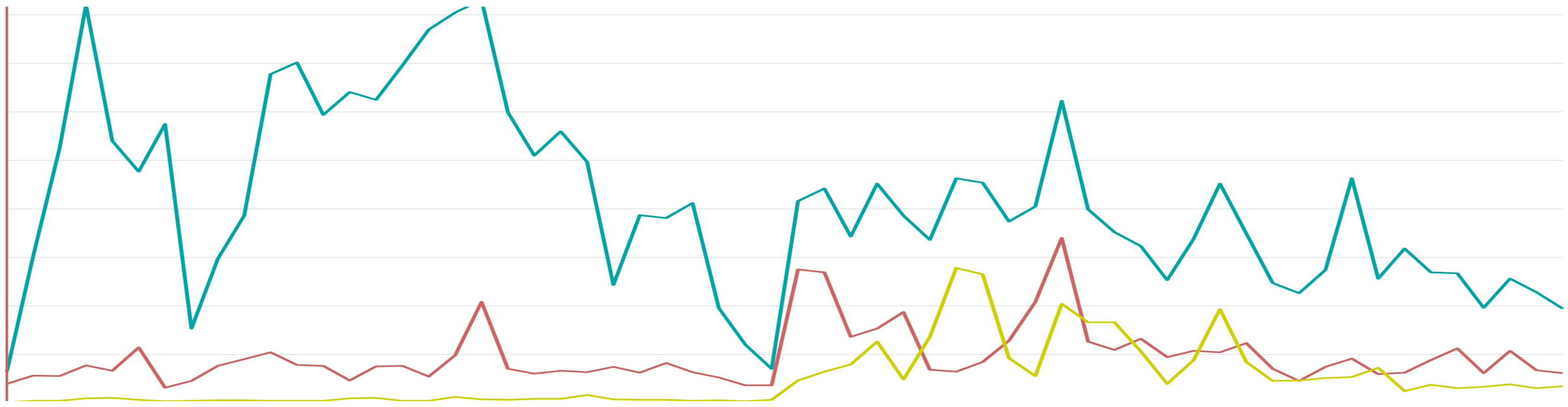
解决方案



Top 10/周

内核专家

内核分析之系统分析结果

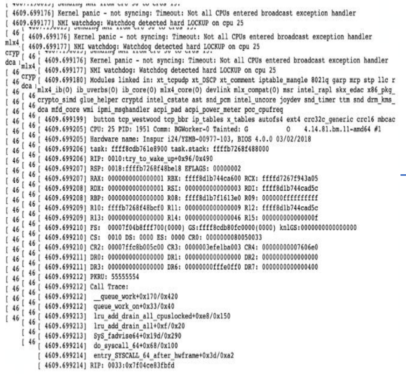


10%

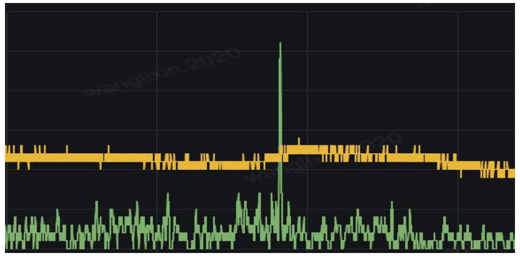
60～70%

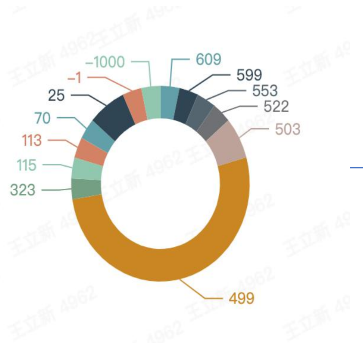
内核故障分析之主动监控

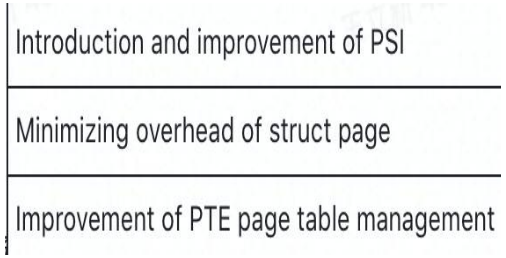
数据采集 → 时序异常检测 → 故障分析 → 解决方案



```
[ 4609.699176] Kernel panic - not syncing: Timeout: Not all CPUs entered broadcast exception handler
[ 4609.699177] NMI watchdog: Watchdog detected hard LOCKUP on cpu 25
[ 4609.699176] Kernel panic - not syncing: Timeout: Not all CPUs entered broadcast exception handler
[ 4609.699177] NMI watchdog: Watchdog detected hard LOCKUP on cpu 25
[ 4609.699176] Kernel panic - not syncing: Timeout: Not all CPUs entered broadcast exception handler
[ 4609.699177] NMI watchdog: Watchdog detected hard LOCKUP on cpu 25
[ 4609.699180] Modules linked in: xt_tcpudp xt_DSCP xt_comment iptable_mangle 8021q garp mrp stp llc r
mlx4_ib(O) ib_uverbs(O) ib_core(O) mlx4_core(O) devlink mlx_compat(O) msr intel_rapl skx_edac x86_pkg_
crypto_simd glue_helper cryptd intel_cstate ast snd_pcm intel_uncore joydev snd_timer ttm snd drm_kms_
dca mfd_core wmi ipmi_msghandler acpi_pad acpi_power_meter pcc_cpufreq
[ 4609.699199]  button tcp_westwood tcp_bbr ip_tables x_tables autofs4 ext4 crc32c_generic crc16 mbcac
[ 4609.699205] CPU: 25 PID: 1951 Comm: BGWorker-0 Tainted: G           O    4.14.81.bm.11-amd64 #1
[ 4609.699205] Hardware name: Inspur i24/YZMB-00977-103, BIOS 4.0.0 03/02/2018
[ 4609.699206] task: ffff8cdb761e8900 task.stack: ffffb7268f488000
[ 4609.699206] RIP: 0010:try_to_wake_up+0x96/0x490
[ 4609.699207] RSP: 0018:ffffb7268f48be18 EFLAGS: 00000002
[ 4609.699207] RAX: 0000000000000001 RBX: ffff8d1b744ca600 RCX: ffffd7267f943a05
[ 4609.699208] RDX: 0000000000000001 RSI: 0000000000000003 RDI: ffff8d1b744cad5c
[ 4609.699208] RBP: 0000000000000000 R08: ffff8d1b7f1613e0 R09: 000000ffffffffff
[ 4609.699209] R10: ffffb7268f48bcf0 R11: 0000000000000009 R12: ffff8d1b744cad5c
[ 4609.699209] R13: 0000000000000000 R14: 0000000000000046 R15: 000000000000000f
[ 4609.699210] FS:  00007f04b8fff700(0000) GS:ffff8cdb80fc0000(0000) knlGS:0000000000000000
[ 4609.699210] CS:  0010 DS: 0000 ES: 0000 CR0: 0000000080050033
[ 4609.699210] CR2: 00007ffc8b005c00 CR3: 0000003efe1ba003 CR4: 00000000007606e0
[ 4609.699211] DR0: 0000000000000000 DR1: 0000000000000000 DR2: 0000000000000000
[ 4609.699211] DR3: 0000000000000000 DR6: 00000000fffe0ff0 DR7: 0000000000000400
[ 4609.699212] PKRU: 55555554
[ 4609.699212] Call Trace:
[ 4609.699212]  __queue_work+0x170/0x420
[ 4609.699212]  queue_work_on+0x33/0x40
[ 4609.699213]  lru_add_drain_all_cpuslocked+0xe8/0x150
[ 4609.699213]  lru_add_drain_all+0xf/0x20
[ 4609.699213]  SyS_fadvise64+0x19d/0x290
[ 4609.699214]  do_syscall_64+0x68/0x100
[ 4609.699214]  entry_SYSCALL_64_after_hwframe+0x3d/0xa2
[ 4609.699214] RIP: 0033:0x7f04ce83fbfd
```







Introduction and improvement of PSI

Minimizing overhead of struct page

Improvement of PTE page table management

内核故障分析之未来规划

1. 扩展知识库

2. 增加故障分析能力

THANKS

ByteDance 字节跳动