

CANN
6.3.RC2

故障处理 (推理)

文档版本	01
发布日期	2023-08-07



版权所有 © 华为技术有限公司 2023。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

目 录

1 故障处理流程	1
1.1 内容简介	1
1.2 处理流程	1
2 Error Code 故障	4
2.1 使用说明	5
2.2 GE Errors	5
2.2.1 E10001 Invalid Argument	5
2.2.2 E10002 Invalid --input_shape Argument	6
2.2.3 E10003 Invalid Argument	6
2.2.4 E10004 Invalid Argument	6
2.2.5 E10005 Invalid Argument	6
2.2.6 E10006 Invalid Argument	7
2.2.7 E10007 Invalid Argument	7
2.2.8 E10008 Invalid Argument	7
2.2.9 E10009 Invalid Dynamic Shape Argument	8
2.2.10 E10010 Invalid --log Argument	8
2.2.11 E10011 Invalid --input_shape Argument	8
2.2.12 E10012 Invalid --input_shape Argument	9
2.2.13 E10013 Invalid Argument	9
2.2.14 E10014 Invalid Argument	9
2.2.15 E10015 Invalid Argument	9
2.2.16 E10016 Invalid Node Name	10
2.2.17 E10017 Invalid Argument	10
2.2.18 E10018 Invalid Dynamic Shape Argument	10
2.2.19 E10019 Invalid --input_shape Argument	11
2.2.20 E10020 Invalid --dynamic_image_size Argument	11
2.2.21 E10021 Invalid Argument	11
2.2.22 E10022 Invalid Argument	12
2.2.23 E10023 Invalid --singleop Argument	12
2.2.24 E10024 Invalid --singleop Argument	12
2.2.25 E10025 Invalid --singleop Argument	12
2.2.26 E10026 Invalid --singleop Argument	13
2.2.27 E10027 Invalid --singleop Argument	13

2.2.28 E10029 Invalid --singleop Argument.....	13
2.2.29 E10030 Invalid --singleop Argument.....	14
2.2.30 E10031 Invalid --input_shape Argument.....	14
2.2.31 E10032 Invalid Json File.....	14
2.2.32 E10034 Invalid --input_fp16_nodes Argument.....	15
2.2.33 E10035 Invalid Dynamic Shape Argument.....	15
2.2.34 E10036 Invalid Dynamic Shape Argument.....	15
2.2.35 E10037 Invalid Dynamic Shape Argument.....	16
2.2.36 E10038 Invalid Dynamic Shape Argument.....	16
2.2.37 E10039 Invalid Dynamic Shape Argument.....	16
2.2.38 E10040 Invalid --input_shape Argument.....	17
2.2.39 E10041 Invalid --framework or --model Argument.....	17
2.2.40 E10044 Insufficient Memory.....	17
2.2.41 E10045 Invalid Dynamic Shape Argument.....	18
2.2.42 E10046 Invalid Dynamic Shape Argument.....	18
2.2.43 E10047 Invalid Argument.....	18
2.2.44 E10048 Invalid --input_shape_range Argument.....	19
2.2.45 E10049 Invalid --input_shape_range Argument.....	19
2.2.46 E10050 Invalid --input_shape_range Argument.....	19
2.2.47 E10051 Invalid --job_id Argument.....	20
2.2.48 E10052 Invalid AIPP Configuration.....	20
2.2.49 E10053 Invalid Argument.....	20
2.2.50 E10054 Invalid Argument.....	20
2.2.51 E10055 Unsupported Feature.....	21
2.2.52 E10056 Invalid Argument.....	21
2.2.53 E10401 Invalid Operator Input Count.....	21
2.2.54 E10402 Invalid Input Buffer Allocation for Operator Execution.....	21
2.2.55 E10403 Invalid Operator Output Count.....	22
2.2.56 E10404 Invalid Output Buffer Allocation for Operator Execution.....	22
2.2.57 E10405 Inconsistent Input Buffer Count and Input Tensor Count for Operator Execution.....	22
2.2.58 E10406 Inconsistent Output Buffer Count and Output Tensor Count for Operator Execution.....	23
2.2.59 E10410 Invalid Argument.....	23
2.2.60 E10501 Unsupported Operator.....	23
2.2.61 E11001 Caffe Model Data Error.....	24
2.2.62 E11003 Caffe Model Data Error.....	24
2.2.63 E11004 Caffe Model Data Error.....	24
2.2.64 E11005 Invalid --input_shape Argument.....	25
2.2.65 E11008 Caffe Model Data Error.....	25
2.2.66 E11009 Unsupported Caffe Operator.....	25
2.2.67 E11012 Caffe Model Data Error.....	25
2.2.68 E11014 Caffe Model Data Error.....	26
2.2.69 E11015 Caffe Model Data Error.....	26

2.2.70 E11016 Invalid --out_nodes Argument.....	26
2.2.71 E11017 Invalid --out_nodes Argument.....	27
2.2.72 E11018 Caffe Model Data Error.....	27
2.2.73 E11021 Caffe Model Data Error.....	27
2.2.74 E11022 Caffe Model Data Error.....	27
2.2.75 E11023 Caffe Model Data Error.....	28
2.2.76 E11024 Caffe Model Data Error.....	28
2.2.77 E11027 Caffe Model Data Error.....	28
2.2.78 E11029 Caffe Model Data Error.....	29
2.2.79 E11032 Caffe File Error.....	29
2.2.80 E11033 The Caffe weight file is invalid.....	29
2.2.81 E11035 Caffe Model Data Error.....	29
2.2.82 E11036 Caffe Model Data Error.....	30
2.2.83 E11037 Caffe Model Data Error.....	30
2.2.84 E12004 Operator Prototype Registration Error.....	30
2.2.85 E12009 TensorFlow Model Data Error.....	31
2.2.86 E12013 TensorFlow Model Data Error.....	31
2.2.87 E12029 TensorFlow Model Data Error.....	31
2.2.88 E14001 Invalid Argument for Operator Compilation.....	32
2.2.89 E14002 Failed to Obtain Tensor Attribute.....	32
2.2.90 E16001 ONNX Model Data Error.....	32
2.2.91 E16002 ONNX Model Data Error.....	32
2.2.92 E16004 ONNX Model Data Error.....	33
2.2.93 E16005 ONNX Model Data Error.....	33
2.2.94 E19000 Invalid Directory.....	33
2.2.95 E19001 Failure to Open File.....	33
2.2.96 E19002 Too Long File Directory.....	34
2.2.97 E19003 Failure to Read File.....	34
2.2.98 E19004 Failure to Write File.....	34
2.2.99 E19005 Failure to Parse File.....	35
2.2.100 E19009 Operator Name Conflict.....	35
2.2.101 E19010 Unsupported Operator.....	35
2.2.102 E19014 Operator Data Verification Failure.....	35
2.2.103 E19015 File size invalid.....	36
2.2.104 E19018 Failure to Parse File.....	36
2.2.105 E19022 Insufficient Memory.....	36
2.2.106 E19023 Too Large OM Model.....	37
2.2.107 E19024 Invalid Environment Variable.....	37
2.2.108 E19025 Input Tensor Error.....	37
2.2.109 E19026 Input Path Name Error.....	37
2.2.110 E19027 Input Config Item Error.....	38
2.2.111 E19999 System Terminated.....	38

2.2.112 W11001 Operator Missing High-Priority Performance.....	38
2.2.113 W11002 Invalid Weight Configuration.....	39
2.3 FE Errors.....	39
2.3.1 E20001 Op Compilation Failure.....	39
2.3.2 E20002 Invalid Environment Variable.....	39
2.3.3 E20003 Invalid Content.....	40
2.3.4 E20007 Fusion Pass Execution Failure.....	40
2.3.5 E20101 Invalid CANN Initialization Argument.....	40
2.3.6 E20103 Invalid Platform Initialization Argument.....	41
2.3.7 E21001 File Opening Failure.....	41
2.3.8 E21002 File Read Failure.....	41
2.3.9 E22001 Compile result is not found.....	42
2.4 AI CPU Errors.....	42
2.4.1 E30003 NN Process Count Limit Exceeded.....	42
2.4.2 E30004 NN Process Bin Error.....	42
2.4.3 E30005 Device Connection Failure.....	42
2.4.4 E30006 OPP Verification Failure.....	43
2.4.5 E30007 Cgroup Add Failure.....	43
2.4.6 E30008 AI CPU Timeout.....	43
2.4.7 E39001 Invalid Argument.....	44
2.4.8 E39002 Driver Function Call Failure.....	44
2.4.9 E39004 AI CPU Scheduler Initialization Failure.....	44
2.4.10 E39005 TSDaemon Process Abnormal.....	45
2.4.11 E39006 Failed to Send OPP to device.....	45
2.4.12 E39007 Device Subprocess Startup Timeout.....	45
2.5 TEFusion Errors.....	45
2.5.1 E40001 Failure to Import Python Module.....	46
2.5.2 E40002 Failure to Call Python Function.....	46
2.5.3 E40003 Failed to Open JSON File.....	46
2.5.4 E40020 Failure to Import Python Module.....	47
2.5.5 E40021 Failure to Compile Op.....	47
2.5.6 E40022 Invalid Argument.....	47
2.5.7 E40023 Invalid Path.....	47
2.5.8 W40010 Invalid Path.....	48
2.5.9 W40011 Failure to Create Disk Cache Directory.....	48
2.5.10 W40012 Invalid Argument.....	48
2.6 Vector Operator Plugin Errors.....	49
2.6.1 E76002 ONNX Model Data Error.....	49
2.7 Vector Operator Errors.....	49
2.7.1 E80001 Invalid Config Value.....	49
2.8 TBE Pass Compilation Tool Errors (Back-End).....	49
2.8.1 EB0000 Invalid IR.....	49

2.8.2 EB1000 Primitives Error.....	50
2.8.3 EB3000 Memory Error.....	50
2.8.4 EB4000 Emit Insn Error.....	50
2.8.5 EB9000 Common Compilation Optimization Error.....	51
2.9 Auto Tune Errors.....	51
2.9.1 EC0000 Remote Device Connection Failure.....	51
2.9.2 EC0001 Remote Device Login Failure.....	51
2.9.3 EC0002 Invalid NFS Mount Path.....	52
2.9.4 EC0003 File Not Found.....	52
2.9.5 EC0004 Permission Denied.....	52
2.9.6 EC0005 Invalid Argument.....	52
2.9.7 EC0006 Directory Operation Failure.....	53
2.9.8 EC0007 Invalid Environment Variable.....	53
2.9.9 EC0008 Invalid Environment Variable.....	53
2.9.10 EC0009 Resources Insufficient.....	54
2.9.11 EC0010 Failure to Import Python Module.....	54
2.10 RTS Errors.....	54
2.10.1 EE1001 Invalid Argument.....	54
2.10.2 EE4001 Model Binding Errors.....	55
2.10.3 EE4002 Model Unbinding Errors.....	55
2.10.4 EE4004 Profiling Enable Errors.....	55
2.11 LxFusion & AutoDeploy Errors.....	55
2.11.1 EF0000 Invalid Directory.....	56
2.11.2 EF0001 Invalid SoC Version.....	56
2.11.3 EF0003 Invalid Environment Variable.....	56
2.11.4 WF0000 File Read Failure.....	56
2.11.5 WF0001 Invalid Config Value.....	57
2.11.6 WF0002 Model Performance Not Optimal.....	57
2.11.7 EF1001 Invalid Hierarchical Machine Configuration.....	57
2.11.8 EF1002 Repository File Not Exist.....	58
2.11.9 EF1003 Repository File Format Incorrect.....	58
2.11.10 EF1004 Fused Optimizer Operator Partition Failure.....	58
2.11.11 EF1005 pulp Module Import Failure.....	58
2.11.12 EF1006 coin-or-cbc Not Installed or Correctly Configured.....	59
2.11.13 EF1007 Python Environment Failure.....	59
2.11.14 EF1008 Dynamic Shape Not Supported.....	59
2.11.15 EF1009 Invalid Slice Number for a Node.....	60
2.11.16 EF1010 Multiple Slice Numbers Set for a Node Scope.....	60
2.12 AOE Errors.....	60
2.12.1 EG0000 Invalid Command Line Argument.....	60
2.12.2 EG0001 Invalid Configuration Argument.....	61
2.12.3 EG0002 Invalid Environment Variable Argument.....	61

2.12.4 EG0003 Invalid Option Argument.....	61
2.12.5 EG0010 Directory Operation Failure.....	61
2.12.6 EG0011 File Operation Failure.....	62
2.12.7 EG0012 Invalid Component.....	62
2.12.8 EG0013 Device Operation Failure.....	62
2.12.9 EG10000 Network Communication Failure.....	62
2.12.10 EG10001 Private Key/Certificate Verification Failure.....	63
2.12.11 WG0000 Memory Insufficient.....	63
2.13 ACL Errors.....	63
2.13.1 EH0001 Invalid Argument.....	63
2.13.2 EH0002 Null Pointer.....	64
2.13.3 EH0003 Invalid Path.....	64
2.13.4 EH0004 Invalid File.....	64
2.13.5 EH0005 Invalid AIPP Argument.....	64
2.13.6 EH0006 Unsupported Feature.....	65
2.14 HCCL Errors.....	65
2.14.1 EI0001 Invalid Environment Variable Configuration.....	65
2.14.2 EI0002 Communication Operation Timeout.....	65
2.14.3 EI0003 Invalid Collective Communication Op Argument.....	66
2.14.4 EI0004 Invalid Ranktable Configuration.....	66
2.14.5 EI0005 Inconsistent Collective Communication Arguments Between Ranks.....	67
2.14.6 EI0006 Get Socket Timeout.....	67
2.14.7 EI0007 Allocation Failure.....	67
2.14.8 EI0008 Inconsistent CANN Versions.....	68
2.15 HCCP Errors.....	68
2.15.1 EJ0001 HCCP Process Initialization Failure.....	68
2.15.2 EJ0002 Environment Error.....	68
2.15.3 EJ0003 Bind Failed.....	69
2.15.4 EJ0004 Check Ip Failed.....	69
2.16 Profiling Errors.....	69
2.16.1 EK0001 Invalid Argument.....	69
2.16.2 EK0002 Interface Call Sequence Error.....	69
2.16.3 EK0003 Profiling Config Setting Error.....	70
2.16.4 EK0004 API Support Error.....	70
2.16.5 EK0201 Memory Allocation Error.....	70
2.16.6 EK9999 System Terminated.....	71
2.17 Driver Errors.....	71
2.17.1 EL0001 Device Absent/Abnormal.....	71
2.17.2 EL0002 Invalid Device ID.....	71
2.17.3 EL0003 Invalid Argument.....	71
2.17.4 EL0004 Memory Allocation Failure.....	72
2.17.5 EL0005 Resource Busy.....	72

2.17.6 EL0006 Insufficient Resources.....	72
2.17.7 EL0007 No Permission.....	73
2.17.8 EL0008 Insufficient Event Resources.....	73
2.17.9 EL0009 Insufficient Stream Resources.....	73
2.17.10 EL0010 Insufficient Notify Resources.....	74
2.17.11 EL0011 Insufficient Model Resources.....	74
2.17.12 EL0012 Full Service Queue.....	74
2.17.13 EL0013 Insufficient CDQM Resources.....	74
2.17.14 EL0014 Operation Unsupported.....	75
2.17.15 EL0015 Invalid Device Access.....	75
2.18 Queue Schedule Errors.....	75
2.18.1 EM9001 Group Attaching Failure.....	75
2.18.2 EM9002 Queue_Schedule Initialization Failure.....	76
2.19 DVPP Errors.....	76
2.19.1 EN0001 Invalid Argument.....	76
2.19.2 EN0002 Some Argument Invalid.....	76
2.19.3 EN0003 Null Pointer.....	77
2.19.4 EN0004 Unsupported Feature.....	77
2.19.5 EN0005 Channel Resource Error.....	77
2.19.6 EN0006 Memory Address Error.....	77
2.19.7 EN0007 Operation Timeout.....	78
2.19.8 EN0008 Unsupported JPEG File.....	78
2.19.9 EN0009 Unsupported PNG File.....	78
2.20 AMCT Errors.....	78
2.20.1 EO0001 Invalid compression_optimize_conf Value.....	79
2.20.2 EO0002 Invalid compression_optimize_conf Path.....	79
2.20.3 EO0003 Invalid Model.....	79
2.20.4 EO0004 Command Error.....	79
2.20.5 EO9999 System Terminated.....	80
2.21 Public Operator Errors.....	80
2.21.1 EZ0501 Unsupported Operator.....	80
2.21.2 EZ3002 Unsupported Operator.....	80
2.21.3 EZ3003 Unsupported Operator.....	81
2.21.4 EZ9010 Unsupported Operator.....	81
2.22 E*9999 Inner Error.....	81
3 故障案例集.....	82
3.1 AscendCL 常见故障.....	82
3.1.1 背景知识.....	82
3.1.2 常见故障分析与处理.....	82
3.1.2.1 内存未释放.....	82
3.1.2.2 APP 使用 dvpp 接口编译失败.....	83
3.1.2.3 使用 dump 功能未获取 dump 结果.....	84

3.1.2.4 Event 数量超过上限导致 aclrtRecordEvent 接口返回失败.....	85
3.1.2.5 执行单算子产生 coredump 的定位处理.....	85
3.1.2.6 进程异常退出后重新执行任务失败.....	86
3.1.2.7 进程异常时资源清理的处理建议.....	86
3.1.2.8 单算子匹配失败.....	87
3.1.2.9 ACL 执行无输出无报错.....	88
3.1.2.10 动态 shape 推理申请内存失败.....	88
3.1.2.11 异步拷贝调用查询接口报错.....	89
3.1.3 数据预处理常见问题（昇腾 310P）.....	89
3.1.3.1 视频解码丢帧/丢包.....	89
3.1.3.2 视频解码花屏.....	91
3.1.3.3 视频解码性能问题.....	92
3.1.3.4 视频解码失败.....	92
3.1.3.5 JPEG 图片解码失败.....	93
3.1.3.6 VPC 参数校验失败.....	94
3.1.3.7 VENC 创建通道失败.....	95
3.1.3.8 VENC 送入待编码图像帧失败.....	96
3.1.3.9 VENC 编码无输出.....	97
3.1.3.10 模型推理精度下降.....	98
3.1.4 数据预处理常见问题（昇腾 310&昇腾 910）.....	99
3.1.4.1 视频解码丢帧/丢包.....	99
3.1.4.2 视频解码花屏.....	101
3.1.4.3 视频解码性能问题.....	102
3.1.4.4 视频解码失败.....	103
3.1.4.5 视频帧解码失败不触发回调函数.....	103
3.1.4.6 JPEG 图片解码失败.....	104
3.1.4.7 VPC 调用失败.....	105
3.1.4.8 retCode 返回值设置错误，导致 VDEC 视频解码异常.....	107
3.1.4.9 模型推理精度下降.....	107
3.2 GE 常见故障.....	109
3.2.1 背景知识.....	109
3.2.2 常见故障分析与处理.....	109
3.2.2.1 算子插件未注册报错.....	109
3.2.2.2 算子原型未注册报错.....	110
3.2.2.3 动态 shape 模型用户输入和模型推导结果不匹配.....	111
3.2.2.4 动态 shape 模型输入大小校验失败.....	112
3.3 Driver 常见故障.....	113
3.3.1 背景知识.....	113
3.3.2 常见故障分析与处理.....	113
3.3.2.1 内核版本差异导致驱动子包安装失败.....	113
3.3.2.2 PCI 总线上没检测到昇腾 AI 处理器.....	114
3.3.2.3 调用 Device 失败.....	115

3.3.2.4 ssh 登录 Device 侧失败.....	116
3.3.2.5 查询到的 Device 数量与实际设备数量不匹配.....	118
3.3.2.6 I2C 通信失败.....	120
3.3.2.7 带外通信查询异常.....	121
3.3.2.8 HDC 建链异常.....	122
3.3.2.9 内存申请失败, 出现 OOM 情况.....	122
3.3.2.10 虚拟地址抢占导致 mmap 失败.....	123
3.3.2.11 容器内 HDC 字符设备打开失败问题.....	124
3.3.2.12 DVPP 驱动引擎异常返回码.....	124
3.4 Runtime 常见故障.....	127
3.4.1 背景知识.....	127
3.4.2 问题定位流程&方法.....	129
3.4.3 常见故障分析与处理.....	129
3.4.3.1 用户进程异常退出后重启进程失败.....	129
3.4.3.2 注册算子数超过最大规格.....	130
3.4.3.3 AI Core 算子执行报错.....	131
3.4.3.4 AI CPU 算子执行报错.....	133
3.4.3.5 Memcpy 异步拷贝算子执行报错.....	134
3.4.3.6 Runtime 报 Notify wait 错误.....	137
3.5 HCCL&HCCP 常见故障.....	138
3.5.1 背景知识.....	138
3.5.2 常见故障分析与处理.....	139
3.5.2.1 通信操作超时 (EI0002)	140
3.5.2.2 socket 建链超时 (EI0006)	141
3.5.2.3 Get P2P status 超时.....	143
3.5.2.4 HCCP ra 初始化失败, 返回-17.....	144
3.5.2.5 CANN 版本不一致.....	145
3.6 getnext 故障处理.....	146
3.6.1 getnext 算子超时.....	146
3.6.2 mbuf 入队场景出现驱动报错.....	147
3.6.3 mbuf 入队场景出现 OOM 报错.....	149
4 故障定位工具.....	150
4.1 故障诊断工具 (一键收集)	150
4.1.1 工具介绍.....	150
4.1.2 工具使用方法.....	151
4.1.3 工具使用 FAQ.....	152
4.2 集群故障日志分析.....	153
4.2.1 工具介绍.....	153
4.2.2 工具使用方法.....	154
4.3 集群故障日志分析 (云环境)	155
4.3.1 工具介绍.....	155

4.3.2 工具使用方法..... 156

1 故障处理流程

1.1 内容简介

1.2 处理流程

1.1 内容简介

本文主要以开发者在执行推理、训练过程中可能遇到的各类异常故障现象为入口，提供自助式问题定位、问题处理方法，方便开发者快速定位并解决故障。

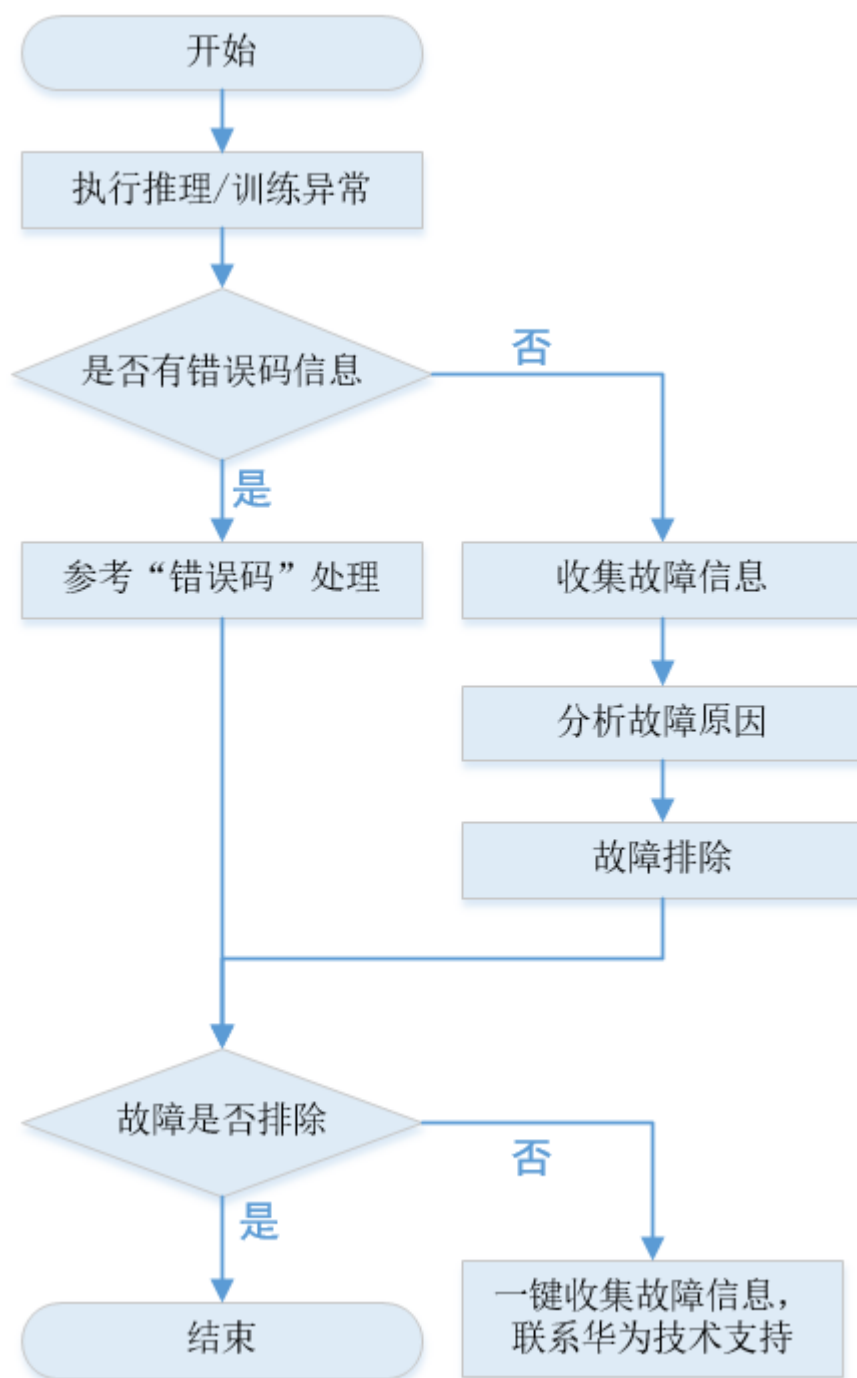
内容包括：

- 屏幕打印的错误码信息及处理方法。
- 一键式日志收集。
- 常见典型故障案例。
- 以及高阶版各类问题定位工具使用。

1.2 处理流程

故障处理总体流程主要包括以下过程：收集故障信息、分析故障原因、故障排除。具体实施过程如[图1-1](#)所示。

图 1-1 故障处理流程



- 收集故障信息
故障信息是故障处理的重要依据，故障处理人员应尽可能多的收集故障信息，包括但不限于日志、环境信息等。
- 分析故障原因
分析故障原因是指从众多可能原因中找出故障原因的过程。通过一定的方法或手段分析、比较各种可能的故障成因，不断排除可能因素，最终确定故障发生的具体原因。

- 故障排除
故障排除是指根据不同的故障原因清除故障的过程。
- 记录故障处理过程
故障排除后应记录故障处理要点，给出针对此类故障的防范和改进措施，避免同类故障再次发生。

说明

您也可以将故障处理案例分享到[华为开发者社区论坛](#)，分享您的经验、提供给其他开发者参考，形成良性循环，丰富社区内容，大家共同受益。

本文提供的故障处理步骤中涉及的第三方工具（如eseye u、Netron），均为举例，非必须工具，请根据您自己实际情况参考使用或替换成其他类似工具。

2 Error Code 故障

- 2.1 使用说明
- 2.2 GE Errors
- 2.3 FE Errors
- 2.4 AI CPU Errors
- 2.5 TEFusion Errors
- 2.6 Vector Operator Plugin Errors
- 2.7 Vector Operator Errors
- 2.8 TBE Pass Compilation Tool Errors (Back-End)
- 2.9 Auto Tune Errors
- 2.10 RTS Errors
- 2.11 LxFusion & AutoDeploy Errors
- 2.12 AOE Errors
- 2.13 ACL Errors
- 2.14 HCCL Errors
- 2.15 HCCP Errors
- 2.16 Profiling Errors
- 2.17 Driver Errors
- 2.18 Queue Schedule Errors
- 2.19 DVPP Errors
- 2.20 AMCT Errors
- 2.21 Public Operator Errors
- 2.22 E*9999 Inner Error

2.1 使用说明

产生机制

用户执行模型构建、推理运行、训练脚本等场景时发生异常错误，例如检查到输入错误（命令行输入参数错误、API输入参数错误、输入文件错误、算子不支持、Shape/Format不支持……）、环境错误，此时通过调用REPORT_INPUT_ERROR()、REPORT_EVN_ERROR()等接口函数，在用户界面展示报错的错误码信息。

错误码展示方式说明

由于场景不同、用例不同、发生故障的原因不同，造成打印的错误码信息有区别，因此，本文中以[%s]变量形式替代实际的打印日志，[%s]替代的实际日志以屏幕打印为准。例如E10035错误码在手册中的表示形式为：

E10035: The [--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims] argument has [%s] profiles, which is less than the minimum [%s].

用户界面实际报错示例：

E10035: The [--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims] argument has [1] profiles, which is less than the minimum [2].

📖 说明

错误码以6位字符形式体现，各段含义如下：

字段1	字段2	字段3
级别（1位）	模块（1位）	错误码（4位）

其中，级别分为E、W、I，分别表示错误、告警、提示类；错误码0000~8999为用户类错误，9000~9999为内部错误码。一般情况下，内部错误码需要联系华为进行故障排查，您可以到昇腾社区（<https://gitee.com/ascend>）提交issue获取帮助。

本文介绍的错误码相关信息已在屏显报错时全部展示，包括可能的原因、解决方案等，因此，本文仅将相关内容罗列供参考。

2.2 GE Errors

2.2.1 E10001 Invalid Argument

Symptom

Value [%s] for parameter [%s] is invalid. Reason: %s.

Possible Cause

N/A

Solution

Try again with a valid argument.

2.2.2 E10002 Invalid --input_shape Argument

Symptom

Value [%s] for parameter [--input_shape] is invalid. Reason: %s. The value must be formatted as [%s].

Possible Cause

N/A

Solution

The valid format is [input_name1:n1,c1,h1,w1;input_name2:n2,c2,h2,w2]. Replace [input_name#] with node names. Ensure that the shape values are integers.

2.2.3 E10003 Invalid Argument

Symptom

Value [%s] for parameter [--%s] is invalid. Reason: %s.

Possible Cause

N/A

Solution

Run the "**atc -h**" command to view the usage. For details, see ATC Instructions.

2.2.4 E10004 Invalid Argument

Symptom

Value for [--%s] is empty.

Possible Cause

N/A

Solution

N/A

2.2.5 E10005 Invalid Argument

Symptom

Value [%s] for parameter [--%s] is invalid. The value must be either [true] or [false].

Possible Cause

N/A

Solution

N/A

2.2.6 E10006 Invalid Argument

Symptom

Value [%s] for parameter [--%s] is invalid. The value must be either 1 or 0.

Possible Cause

N/A

Solution

N/A

2.2.7 E10007 Invalid Argument

Symptom

[--%s] is required. The value must be [%s].

Possible Cause

N/A

Solution

N/A

2.2.8 E10008 Invalid Argument

Symptom

[--weight] must not be empty when [--framework] is set to 0 (Caffe).

Possible Cause

N/A

Solution

1. If the source model framework is Caffe, try again with a valid [--weight] argument.
2. If the source model framework is not Caffe, try again with a valid [--framework] argument.

2.2.9 E10009 Invalid Dynamic Shape Argument

Symptom

[--dynamic_batch_size], [--dynamic_image_size], [--input_shape_range], and [--dynamic_dims] are mutually exclusive.

Possible Cause

N/A

Solution

1. In dynamic shape scenarios, include only one of these options in your command line.
2. In static shape scenarios, remove these options from your command line.

2.2.10 E10010 Invalid --log Argument

Symptom

Value [%s] for parameter [--log] is invalid.

Possible Cause

N/A

Solution

Select the value from [debug], [info], [warning], [error], and [null].

2.2.11 E10011 Invalid --input_shape Argument

Symptom

Value [%s] for parameter [--input_shape] is invalid. Shape values must be positive integers. The error value in the shape is %s.

Possible Cause

N/A

Solution

1. In static shape scenarios, set the shape values in [--input_shape] to positive integers.
2. In dynamic shape scenarios, add the related dynamic-input option in your command line, such as [--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims].

2.2.12 E10012 Invalid --input_shape Argument

Symptom

[--dynamic_batch_size] is included, but the dimension count of the dynamic-shape input configured in [--input_shape] is less than 1.

Possible Cause

N/A

Solution

1. In static shape scenarios, remove the [--dynamic_batch_size] option from your command line.
2. In dynamic shape scenarios, set the corresponding axis of the dynamic-shape input in [--input_shape] to -1.

2.2.13 E10013 Invalid Argument

Symptom

Value [%s] for [--%s] is out of range [%s].

Possible Cause

N/A

Solution

Run the "**atc -h**" command to view the usage. For details, see ATC Instructions.

2.2.14 E10014 Invalid Argument

Symptom

Value [%s] for parameter [--%s] is invalid.

Possible Cause

N/A

Solution

Run the "**atc -h**" command to view the usage. For details, see ATC Instructions.

2.2.15 E10015 Invalid Argument

Symptom

Value [%s] for [--%s] is invalid.

Possible Cause

Try again with a valid argument.

Solution

Run the "**atc -h**" command to view the usage. For details, see ATC Instructions.

2.2.16 E10016 Invalid Node Name

Symptom

Opname [%s] specified in [--%s] is not found in the model, confirm whether this node name exists, or node is not split with the specified delimiter ';'.

Possible Cause

N/A

Solution

N/A

2.2.17 E10017 Invalid Argument

Symptom

Input Op [%s] specified in [--%s] is invalid. The Op type must be Data.

Possible Cause

N/A

Solution

Run the "**atc -h**" command to view the usage. For details, see ATC Instructions.

2.2.18 E10018 Invalid Dynamic Shape Argument

Symptom

Value [%s] for shape [%s] is invalid. When [--dynamic_batch_size] is included, only batch size N can be -1 in [--input_shape].

Possible Cause

When [--dynamic_batch_size] is included, only batch size N can be -1 in the shape.

Solution

Try again with a valid [--input_shape] argument. Make sure that non-batch size axes are not -1.

2.2.19 E10019 Invalid --input_shape Argument

Symptom

When [--dynamic_image_size] is included, only the height and width axes can be -1 in [--input_shape].

Possible Cause

When [--dynamic_image_size] is included, only the height and width axes can be -1 in the shape.

Solution

Try again with a valid [--input_shape] argument. Make sure that axes other than height and width are not -1.

2.2.20 E10020 Invalid --dynamic_image_size Argument

Symptom

Value [%s] for parameter [--dynamic_image_size] is invalid.

Possible Cause

N/A

Solution

The value must be formatted as [imagesize1_height,imagesize1_width;imagesize2_height,imagesize2_width]. Make sure that each profile has two dimensions. For details, see ATC Instructions.

2.2.21 E10021 Invalid Argument

Symptom

Path for parameter [--%s] is too long. Keep the length within %s.

Possible Cause

N/A

Solution

N/A

2.2.22 E10022 Invalid Argument

Symptom

Path [%s] for parameter [--%s] does not include the file name.

Possible Cause

N/A

Solution

Add the file name to the path.

2.2.23 E10023 Invalid --singleop Argument

Symptom

Value [%s] for parameter [--singleop] is invalid.

Possible Cause

The path does not exist or the file name is incorrect.

Solution

N/A

2.2.24 E10024 Invalid --singleop Argument

Symptom

Failed to open file [%s] specified by [--singleop].

Possible Cause

N/A

Solution

Check the owner group and permission settings and ensure that the user who runs the ATC command has enough permission to open the file.

2.2.25 E10025 Invalid --singleop Argument

Symptom

File [%s] specified by [--singleop] is not a valid JSON file. Reason: %s.

Possible Cause

N/A

Solution

Check that the file is in valid JSON format.

2.2.26 E10026 Invalid --singleop Argument

Symptom

Empty Op name in the file specified by [--singleop].

Possible Cause

N/A

Solution

Check that no Op name is empty in the file.

2.2.27 E10027 Invalid --singleop Argument

Symptom

Attribute [%s] of [%s] tensor [%s] for Op [%s] is invalid when [--singleop] is specified.

Possible Cause

N/A

Solution

Try again with a valid tensor dtype and format.

2.2.28 E10029 Invalid --singleop Argument

Symptom

Attribute name of Op [%s] is empty in the file specified by [--singleop].

Possible Cause

N/A

Solution

Check that no Op attribute name is empty in the file.

2.2.29 E10030 Invalid --singleop Argument

Symptom

There is an invalid value for attribute name [%s] of Op [%s] in the file specified by [--singleop].

Possible Cause

N/A

Solution

Check that no Op attribute value is invalid in the file.

2.2.30 E10031 Invalid --input_shape Argument

Symptom

[--dynamic_batch_size] is included, but none of the nodes specified in [--input_shape] have a batch size equaling -1.

Possible Cause

As [--dynamic_batch_size] is included, ensure that at least one of the nodes specified in [--input_shape] has a batch size equaling -1.

Solution

1. In static shape scenarios, remove the [--dynamic_batch_size] option from your command line.
2. In dynamic shape scenarios, set the corresponding axis of the dynamic-shape input in [--input_shape] to -1.

2.2.31 E10032 Invalid Json File

Symptom

Parse json file [%s] failed, reason: %s. %s

Possible Cause

N/A

Solution

N/A

2.2.32 E10034 Invalid --input_fp16_nodes Argument

Symptom

Nodes (for example, [%s]) connected to AIPP must not be of type fp16.

Possible Cause

N/A

Solution

1. To enable AIPP, remove the nodes connected to AIPP from the [--input_fp16_nodes] argument.
2. If AIPP is not required, remove the [--insert_op_conf] option from your ATC command line.

2.2.33 E10035 Invalid Dynamic Shape Argument

Symptom

[--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims] has [%s] profiles, which is less than the minimum ([%s]).

Possible Cause

N/A

Solution

Ensure that the number of profiles configured in [--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims] is at least the minimum.

2.2.34 E10036 Invalid Dynamic Shape Argument

Symptom

[--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims] has [%s] profiles, which is greater than the maximum ([%s]).

Possible Cause

N/A

Solution

Ensure that the number of profiles configured in [--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims] is at most the maximum.

2.2.35 E10037 Invalid Dynamic Shape Argument

Symptom

The profiles configured in [--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims] have inconsistent dimension counts. A profile has %s dimensions while another has %s dimensions.

Possible Cause

N/A

Solution

Ensure that the profiles configured in [--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims] have the same dimension count.

2.2.36 E10038 Invalid Dynamic Shape Argument

Symptom

Dimension size [%s] is invalid. The value must be greater than 0.

Possible Cause

N/A

Solution

Set the shape values of each profile to positive in [--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims].

2.2.37 E10039 Invalid Dynamic Shape Argument

Symptom

The [--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims] argument have duplicate profiles.

Possible Cause

N/A

Solution

Check that the profiles configured in [--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims] are unique.

2.2.38 E10040 Invalid --input_shape Argument

Symptom

As the [--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims] argument is included, the corresponding nodes specified in [--input_shape] must have -1 axes and can not have '~'.

Possible Cause

In dynamic shape scenarios, when [--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims] argument is included, the corresponding shape value in the input node must be set to -1 and can not have '~'.

Solution

1. In static shape scenarios, remove the [--dynamic_batch_size], [--dynamic_image_size] or [--dynamic_dims] option from your command line.
2. In dynamic multi-batch scenarios, set the corresponding axis of the dynamic-shape input in [--input_shape] to -1.
3. In dynamic shape scenarios, remove the [--dynamic_batch_size], [--dynamic_image_size] or [--dynamic_dims] option from your command line and set [--input_shape] to -1 or n1~n2.

2.2.39 E10041 Invalid --framework or --model Argument

Symptom

Failed to load the model from [%s].

Possible Cause

N/A

Solution

1. Check that the model file is valid.
2. Check that the weight file or path is valid when model is more than 2G.
3. Check that the [--framework] argument matches the actual framework of the model file.

2.2.40 E10044 Insufficient Memory

Symptom

The available memory is [%s KB], which is less than the required minimum [%s KB].

Possible Cause

N/A

Solution

Stop unnecessary processes and ensure that the required minimum system memory is available.

2.2.41 E10045 Invalid Dynamic Shape Argument

Symptom

The number of -1 axes in the [--input_shape] argument exceeds the dimension count per profile in [--dynamic_dims].

Possible Cause

N/A

Solution

Ensure that the number of -1 axes in the [--input_shape] argument matches the dimension count per profile in [--dynamic_dims].

2.2.42 E10046 Invalid Dynamic Shape Argument

Symptom

The total number of -1 axes in the [--input_shape] argument is greater than the dimension count per profile in [--dynamic_dims].

Possible Cause

N/A

Solution

Ensure that the total number of -1 axes in the [--input_shape] argument is less than the dimension count per profile in [--dynamic_dims].

2.2.43 E10047 Invalid Argument

Symptom

[--%s] and [--%s] are mutually exclusive.

Possible Cause

N/A

Solution

Remove either of them and try again.

2.2.44 E10048 Invalid --input_shape_range Argument

Symptom

Value [%s] for parameter [--input_shape_range] is invalid. Reason: %s. The value must be formatted as [%s].

Possible Cause

N/A

Solution

Try again with a valid argument.

2.2.45 E10049 Invalid --input_shape_range Argument

Symptom

Dimension count [%s] configured in [--input_shape_range] does not match dimension count [%s] of the node.

Possible Cause

N/A

Solution

Set the dimension count in [--input_shape_range] according to the dimension count of the node.

2.2.46 E10050 Invalid --input_shape_range Argument

Symptom

Current dimension size [%s] is not in the range of [%s-%s] specified by [--input_shape_range].

Possible Cause

N/A

Solution

Set the dimension size according to [--input_shape_range].

2.2.47 E10051 Invalid --job_id Argument

Symptom

Value [%s] for parameter [--job_id] exceeds the allowed maximum ([%s]).

Possible Cause

N/A

Solution

Try again with a valid argument.

2.2.48 E10052 Invalid AIPP Configuration

Symptom

AIPP configuration is invalid. Reason: %s.

Possible Cause

N/A

Solution

Run the "atc -h" command to view the usage. For details, see ATC Instructions.

2.2.49 E10053 Invalid Argument

Symptom

There is an error in the command line. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.2.50 E10054 Invalid Argument

Symptom

The required parameter [%s] for ATC is empty. Another possible reason is that the value of some parameter is not enclosed by quotation marks ("").

Possible Cause

N/A

Solution

N/A

2.2.51 E10055 Unsupported Feature

Symptom

The operation is not supported. Reason: %s

Possible Cause

N/A

Solution

N/A

2.2.52 E10056 Invalid Argument

Symptom

Cannot config both parameters [%s] and [%s] simultaneously.

Possible Cause

N/A

Solution

N/A

2.2.53 E10401 Invalid Operator Input Count

Symptom

The number of operator inputs ([%s]) exceeds the allowed maximum ([%s]).

Possible Cause

N/A

Solution

Try again with a valid number of inputs.

2.2.54 E10402 Invalid Input Buffer Allocation for Operator Execution

Symptom

Input indexed [%s] requires a %s buffer, but %s (aligned) are allocated.

Possible Cause

N/A

Solution

Check whether the data type, dimensions, and shape are correctly set. For details, see the `aclGetTensorDescSize` API description in AscendCL API Reference.

2.2.55 E10403 Invalid Operator Output Count

Symptom

The number of operator outputs ([%s]) exceeds the allowed maximum ([%s]).

Possible Cause

The number of outputs configured for operator execution does not match that described in the operator specifications.

Solution

Check whether the number of elements in `numoutputs` is correctly set. The `aclopCompile`, `aclopExecuteV2`, and `aclopCompileAndExecute` APIs may be involved. For details, see AscendCL API Reference.

2.2.56 E10404 Invalid Output Buffer Allocation for Operator Execution

Symptom

Output indexed [%s] requires a %s buffer, but %s (aligned) are allocated.

Possible Cause

N/A

Solution

Check whether the data type, dimensions, and shape are correctly set. For details, see the `aclGetTensorDescSize` API description in AscendCL API Reference.

2.2.57 E10405 Inconsistent Input Buffer Count and Input Tensor Count for Operator Execution

Symptom

The number of input buffers is [%s], which does not match the number of input tensors ([%s]).

Possible Cause

N/A

Solution

Check whether the number of elements in inputDesc and inputs of the operator is correctly set. The aclopExecuteV2 and aclopCompileAndExecute APIs may be involved. For details, see AscendCL API Reference.

2.2.58 E10406 Inconsistent Output Buffer Count and Output Tensor Count for Operator Execution

Symptom

The number of output buffers is [%s], which does not match the number of output tensors ([%s]).

Possible Cause

N/A

Solution

Check whether the number of elements in outputDesc and outputs of the operator is correctly set. The aclopExecuteV2 and aclopCompileAndExecute APIs may be involved. For details, see AscendCL API Reference.

2.2.59 E10410 Invalid Argument

Symptom

File [%s] does not exist.

Possible Cause

The file specified by the [--keep_dtype] or [--compress_weight_conf] argument does not exist.

Solution

Try again with a valid file directory.

2.2.60 E10501 Unsupported Operator

Symptom

IR for Op [%s, otype [%s]], is not registered.

Possible Cause

1. The environment variable ASCEND_OPP_PATH is not configured.

2. IR is not registered.

Solution

N/A

2.2.61 E11001 Caffe Model Data Error

Symptom

[input_dim] and [input_shape] are mutually exclusive in [NetParameter] for Caffe model conversion.

Possible Cause

N/A

Solution

Remove either of [--input_dim] and [--input_shape] from your atc command line.

2.2.62 E11003 Caffe Model Data Error

Symptom

The number of [input_dim] fields in the model is [%s], which is not 4x the input count [%s].

Possible Cause

N/A

Solution

Modify your Caffe model and try again.

2.2.63 E11004 Caffe Model Data Error

Symptom

The number of input shapes is [%s], which does not match the number of inputs [%s].

Possible Cause

N/A

Solution

Modify your Caffe model and try again.

2.2.64 E11005 Invalid --input_shape Argument

Symptom

Shape is not defined by using [--input_shape] for input [%s].

Possible Cause

N/A

Solution

Modify your Caffe model, or add the shape of the input to the [--input_shape] argument in your atc command line.

2.2.65 E11008 Caffe Model Data Error

Symptom

Optype DetectionOutput is unsupported.

Possible Cause

N/A

Solution

Modify your Caffe model and replace DetectionOutput operators with FSRDetectionOutput or SSDDetectionOutput.

2.2.66 E11009 Unsupported Caffe Operator

Symptom

No Caffe parser is registered for Op [%s, optype [%s]].

Possible Cause

N/A

Solution

N/A

2.2.67 E11012 Caffe Model Data Error

Symptom

Unknown bottom blob [%s] at layer [%s]. The bottom blob is indexed [%s].

Possible Cause

N/A

Solution

Modify your Caffe model and try again.

2.2.68 E11014 Caffe Model Data Error

Symptom

Failed to find the top blob for layer [%s].

Possible Cause

The top blob has no corresponding node in the source Caffe model.

Solution

Modify your Caffe model and try again.

2.2.69 E11015 Caffe Model Data Error

Symptom

Failed to find the bottom blob for layer [%s].

Possible Cause

The bottom blob has no corresponding node in the source Caffe model.

Solution

Modify your Caffe model and try again.

2.2.70 E11016 Invalid --out_nodes Argument

Symptom

Failed to add Op [%s] to NetOutput. Op output index [%s] is not less than [%s].
NetOutput input_index [%s] is not less than [%s].

Possible Cause

N/A

Solution

Try again with a valid [--out_nodes] argument.

2.2.71 E11017 Invalid --out_nodes Argument

Symptom

Failed to find node [%s] specified by [--out_nodes].

Possible Cause

N/A

Solution

Try again with a valid [--out_nodes] argument.

2.2.72 E11018 Caffe Model Data Error

Symptom

Op name [%s] contains invalid characters.

Possible Cause

N/A

Solution

Allowed characters include: letters, digits, hyphens (-), periods (.), underscores (_), and slashes (/). Modify the Op name and try again.

2.2.73 E11021 Caffe Model Data Error

Symptom

Model file [%s] contains [layers] structures, which have been deprecated in Caffe and unsupported by ATC.

Possible Cause

N/A

Solution

Replace [layers] with [layer].

2.2.74 E11022 Caffe Model Data Error

Symptom

Invalid prototxt file.

Possible Cause

No [layer] structures are found in the Caffe model.

Solution

N/A

2.2.75 E11023 Caffe Model Data Error

Symptom

Weight file contains [layers] structures, which have been deprecated in Caffe and unsupported by ATC.

Possible Cause

N/A

Solution

Replace [layers] with [layer].

2.2.76 E11024 Caffe Model Data Error

Symptom

Invalid Caffe weight file.

Possible Cause

No [layer] structures are found in the Caffe weight file.

Solution

N/A.

2.2.77 E11027 Caffe Model Data Error

Symptom

Op [%s, optype [%s]] in the Caffe model has an input node with shape size 0.

Possible Cause

N/A

Solution

N/A

2.2.78 E11029 Caffe Model Data Error

Symptom

Op [%s] exists in model file but not found in weight file.

Possible Cause

N/A

Solution

Try again with a valid Caffe model or weight file. Ensure that the two files match with each other.

2.2.79 E11032 Caffe File Error

Symptom

Failed to parse message [%s]. The error field is %s. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.2.80 E11033 The Caffe weight file is invalid.

Symptom

Failed to convert the weight file. Blob [%s] of size [%s] is invalid. Reason: %s.

Possible Cause

The blob size of the node in the Caffe weight file does not match the number of elements calculated based on its shape.

Solution

Try again with a valid Caffe model or weight file. Ensure that the two files match with each other.

2.2.81 E11035 Caffe Model Data Error

Symptom

The top size of data node [%s] is not 1 but [%s].

Possible Cause

N/A

Solution

N/A

2.2.82 E11036 Caffe Model Data Error

Symptom

Data nodes have duplicate top blobs [%s].

Possible Cause

N/A

Solution

N/A

2.2.83 E11037 Caffe Model Data Error

Symptom

Op [%s] has zero outputs.

Possible Cause

N/A

Solution

Nodes in the Caffe model must have at least one output.

2.2.84 E12004 Operator Prototype Registration Error

Symptom

Failed to register the prototype of Op [%s]. If input index is less than 0, then input index [-%s] (absolute value) must be less than the input count [%s]

Possible Cause

N/A

Solution

N/A

2.2.85 E12009 TensorFlow Model Data Error

Symptom

Op [%s]'s input [%s] is not found in graph_def.

Possible Cause

The input name of the node is not found in the graph.

Solution

Try again with a valid TensorFlow model.

2.2.86 E12013 TensorFlow Model Data Error

Symptom

Failed to find a subgraph by the name [%s].

Possible Cause

N/A

Solution

1. To use function subgraphs to convert a TensorFlow model, place the subgraph .proto description file in the same directory as the model file and name it [graph_def_library.pbtxt].
2. Run the [func2graph.py] script in the ATC installation directory to save the subgraphs to [graph_def_library.pbtxt].

2.2.87 E12029 TensorFlow Model Data Error

Symptom

Failed to find the subgraph library.

Possible Cause

The model to convert contains function subgraphs, but the [graph_def_library.pbtxt] file is not found.

Solution

1. To use function subgraphs to convert a TensorFlow model, place the subgraph .proto description file in the same directory as the model file and name it [graph_def_library.pbtxt].
2. Run the [func2graph.py] script in the ATC installation directory to save the subgraphs to [graph_def_library.pbtxt].

2.2.88 E14001 Invalid Argument for Operator Compilation

Symptom

Argument [%s] for Op [%s, optype [%s]], is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.2.89 E14002 Failed to Obtain Tensor Attribute

Symptom

In the current process, obtaining the attribute of %s must be successful. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.2.90 E16001 ONNX Model Data Error

Symptom

Model has no [%s] node.

Possible Cause

N/A

Solution

N/A

2.2.91 E16002 ONNX Model Data Error

Symptom

No ONNX parser is registered for optype [%s].

Possible Cause

N/A

Solution

Check the version of the installation package and reinstall the package. For details, see the operator specifications.

2.2.92 E16004 ONNX Model Data Error

Symptom

ONNX model has no graph.

Possible Cause

N/A

Solution

N/A

2.2.93 E16005 ONNX Model Data Error

Symptom

The model has [%s] [--domain_version] fields, but only one is allowed.

Possible Cause

N/A

Solution

N/A

2.2.94 E19000 Invalid Directory

Symptom

Path[%s] is empty. Reason: %s.

Possible Cause

The file does not exist.

Solution

Try again with a valid directory.

2.2.95 E19001 Failure to Open File

Symptom

Failed to open file[%s]. Reason: %s.

Possible Cause

N/A

Solution

Fix the error according to the error message.

2.2.96 E19002 Too Long File Directory

Symptom

Directory [%s] is too long. Keep the length within [%s] characters.

Possible Cause

N/A

Solution

Try again with a valid file directory.

2.2.97 E19003 Failure to Read File

Symptom

Failed to read file [%s]. Reason: %s.

Possible Cause

Failed to read the file.

Solution

Fix the error according to the error message.

2.2.98 E19004 Failure to Write File

Symptom

Failed to write file [%s]. Reason: %s.

Possible Cause

N/A

Solution

Fix the error according to the error message.

2.2.99 E19005 Failure to Parse File

Symptom

Failed to parse file [%s].

Possible Cause

N/A

Solution

Check that a matched Protobuf version is installed and try again with a valid file.
For details, see section "--framework" in ATC Instructions.

2.2.100 E19009 Operator Name Conflict

Symptom

Op [%s] has a name conflict in the graph.

Possible Cause

N/A

Solution

Ensure that the operators in the graph have unique names.

2.2.101 E19010 Unsupported Operator

Symptom

No parser is registered for Op [%s, optype [%s]].

Possible Cause

N/A

Solution

Check the version of the installation package and reinstall the package. For details, see the operator specifications.

2.2.102 E19014 Operator Data Verification Failure

Symptom

Value [%s] for Op [%s] is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.2.103 E19015 File size invalid

Symptom

File[%s] size %s is out of valid range[0, %s].

Possible Cause

N/A

Solution

N/A

2.2.104 E19018 Failure to Parse File

Symptom

Failed to parse file [%s] through [google::protobuf::TextFormat::Parse].

Possible Cause

The file may not be in valid protobuf format.

Solution

N/A

2.2.105 E19022 Insufficient Memory

Symptom

Model %s requires [%s] memory, which exceeds system limit [%s].

Possible Cause

N/A.

Solution

1. Reduce the batch size.
2. If the error persists when the batch size is reduced to 1, modify the model structure to reduce the model size.

2.2.106 E19023 Too Large OM Model

Symptom

Model %s has size [%s], which exceeds system limit [%s].

Possible Cause

The generated OM model is too large and therefore cannot be dumped to the disk.

Solution

Reduce the model size.

2.2.107 E19024 Invalid Environment Variable

Symptom

Value [%s] for environment variable [%s] is invalid when %s.

Possible Cause

N/A

Solution

Reset the environment variable by referring to the setup guide.

2.2.108 E19025 Input Tensor Error

Symptom

Input tensor is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.2.109 E19026 Input Path Name Error

Symptom

Input path [%s] is invalid. Reason: %s

Possible Cause

N/A

Solution

N/A

2.2.110 E19027 Input Config Item Error

Symptom

Failed to connect to the peer address[%s].

Possible Cause

ipaddr, port or token is invalid.

Solution

Check ipaddr, port or token in the configuration file, ensure that the configuration is correct.

2.2.111 E19999 System Terminated

Symptom

Unknown error occurred. Please check the log.

Possible Cause

System terminated abnormally without valid error messages.

Solution

In this scenario, use SdkInfoCollector to collect the logs generated when the fault occurs and locate the fault based on the logs.

2.2.112 W11001 Operator Missing High-Priority Performance

Symptom

Op [%s] does not hit the high-priority operator information library, which might result in compromised performance.

Possible Cause

The operator does not hit the high-priority operator information library, which might result in compromised performance.

Solution

Submit an issue to request for support at <https://gitee.com/ascend>.

2.2.113 W11002 Invalid Weight Configuration

Symptom

In the compression weight configuration file [%s], some nodes do not exist in graph: %s.

Possible Cause

N/A

Solution

N/A

2.3 FE Errors

2.3.1 E20001 Op Compilation Failure

Symptom

Failed to compile Op [%s]. (optype: [%s])

Possible Cause

The operator has an invalid argument.

Solution

For a custom operator, check the operator implementation and arguments based on the error log. For a Huawei built-in operator, submit an issue to request for support at <https://gitee.com/ascend> or contact Huawei technical support.

2.3.2 E20002 Invalid Environment Variable

Symptom

Value [%s] for environment variable [%s] is invalid when %s.

Possible Cause

N/A

Solution

Reset the environment variable by referring to the Installation Guide.

2.3.3 E20003 Invalid Content

Symptom

Configuration file [%s] for parameter [%s] has invalid content. Reason: [%s]

Possible Cause

The content or format of the configuration file does not comply with the specifications provided in the documentation.

Solution

Modify the configuration file by referring to the specifications in the user guide.

2.3.4 E20007 Fusion Pass Execution Failure

Symptom

Failed to run graph fusion pass [%s]. The pass type is [%s]

Possible Cause

N/A

Solution

1. If the pass code is custom, check the error log and the verification logic.
2. If the pass code is not custom, perform a complete or partial dump by using npucollect.sh and then send the dump to Huawei technical support for fault locating.

2.3.5 E20101 Invalid CANN Initialization Argument

Symptom

Value [%s] for parameter [%s] is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.3.6 E20103 Invalid Platform Initialization Argument

Symptom

Value [%s] for parameter [%s] is invalid. The value must be in the range of (0, %s].

Possible Cause

N/A

Solution

N/A

2.3.7 E21001 File Opening Failure

Symptom

Failed to open file [%s].

Possible Cause

1. The configured file path does not exist.
2. The file permissions are insufficient.
3. An error occurred when installing the system environment.

Solution

1. Configure the file path correctly.
2. Configure user permissions correctly.
3. Reinstall the system environment.

2.3.8 E21002 File Read Failure

Symptom

Failed to read file [%s]. Reason: %s.

Possible Cause

N/A

Solution

Configure the file path correctly.

2.3.9 E22001 Compile result is not found

Symptom

Compile result is not found for Op [%s, %s].

Possible Cause

N/A.

Solution

N/A.

2.4 AI CPU Errors

2.4.1 E30003 NN Process Count Limit Exceeded

Symptom

The number of started NN processes exceeds the limit.

Possible Cause

Something wrong with queue_schedule.

Solution

Wait for 10s after the last training process exited and try again.

2.4.2 E30004 NN Process Bin Error

Symptom

The binary file of the NN process is damaged or the file permissions are incorrect.

Possible Cause

The driver package is incorrectly installed.

Solution

Reinstall the driver package.

2.4.3 E30005 Device Connection Failure

Symptom

The device failed to be connected.

Possible Cause

1. The device is broken.
2. The device failed to be started.

Solution

1. Restart the device.
2. Submit an issue on Gitee at <https://gitee.com/ascend>.

2.4.4 E30006 OPP Verification Failure

Symptom

Failed to verify the OPP.

Possible Cause

The OPP is incorrect.

Solution

Use a correct OPP.

2.4.5 E30007 Cgroup Add Failure

Symptom

Failed to add the AI CPU scheduler to the Cgroup.

Possible Cause

1. The User have no sudo Permission.
2. The Cgroup configuration is incorrect.

Solution

1. Check whether the user have sudo Permission or not.
2. Submit an issue on Gitee at <https://gitee.com/ascend>

2.4.6 E30008 AI CPU Timeout

Symptom

AI CPU operator execution time out.

Possible Cause

1. For the GetNext operator, its preprocessing duration may be too long.

2. For a custom operator, its logic may be improper.

Solution

1. For the GetNext operator, check its preprocessing or set OpExecuteTimeOut to a larger value.
2. For a custom operator, make sure its logic is proper.

2.4.7 E39001 Invalid Argument

Symptom

The argument is invalid.

Possible Cause

N/A

Solution

Check the argument sent by the training service daemon (TSD).

2.4.8 E39002 Driver Function Call Failure

Symptom

Failed to obtain the number of devices and the device list.

Possible Cause

Failed to obtain the device information from the driver.

Solution

Check the device logs to see whether the device driver service is normal.

2.4.9 E39004 AI CPU Scheduler Initialization Failure

Symptom

The AI CPU scheduler is abnormal.

Possible Cause

Something wrong with aicpu_scheduler

Solution

Check the initialization information from device logs.

2.4.10 E39005 TSDaemon Process Abnormal

Symptom

The TSDaemon process is abnormal.

Possible Cause

The TSDaemon process failed to be started.

Solution

Check the device logs to see if TSDaemon is normal.

2.4.11 E39006 Failed to Send OPP to device

Symptom

Failed to send the OPP to the device.

Possible Cause

- 1.The HDCD process is abnormal.
2. The device failed to be connected.

Solution

Check the host and device logs to see if HDC is normal.

2.4.12 E39007 Device Subprocess Startup Timeout

Symptom

Starting a subprocess on the device timed out.

Possible Cause

- 1.The subprocess failed to be started.
2. HDC is abnormal.

Solution

1. Check the device logs to see if the subprocess is started successfully.
2. Check the host and device logs to see if HDC is normal.

2.5 TEFusion Errors

2.5.1 E40001 Failure to Import Python Module

Symptom

Failed to import the Python module: [%s]

Possible Cause

A component installation is missing or its installed version does not match.
Alternatively, the specified Python path is incorrect.

Solution

Check that all required components are properly installed and the specified Python path matches the Python installation directory.

2.5.2 E40002 Failure to Call Python Function

Symptom

Failed to call function [%s] with arguments [%s].

Possible Cause

An invalid argument is passed or the Python function has a bug.

Solution

This error is caused by an internal problem. Check the code of the Python function call.

2.5.3 E40003 Failed to Open JSON File

Symptom

Failed to open the JSON file: %s

Possible Cause

1. The file has been deleted.
2. You do not have the permission to open the file.
3. The file is locked.

Solution

Run your program again in a single-user environment (to ensure that no other user accesses the file).

2.5.4 E40020 Failure to Import Python Module

Symptom

Failed to import Python module [%s].

Possible Cause

N/A

Solution

Check that all required components are properly installed and the specified Python path matches the Python installation directory. (If the path does not match the directory, run set_env.sh in the installation package.)

2.5.5 E40021 Failure to Compile Op

Symptom

Failed to compile Op [%s]. (oppath: [%s], optype: [%s])

Possible Cause

N/A

Solution

See the host log for details, and then check the Python stack where the error log is reported.

2.5.6 E40022 Invalid Argument

Symptom

Value [%s] for parameter [%s] is invalid. The value must be in the range of [%s].

Possible Cause

N/A

Solution

N/A

2.5.7 E40023 Invalid Path

Symptom

Path [%s] for [%s] is invalid. Result: %s. Reason: %s.

Possible Cause

The path does not exist.

Solution

N/A

2.5.8 W40010 Invalid Path

Symptom

Path [%s] for [%s] is invalid. Result: %s. Reason: %s.

Possible Cause

The path is empty or does not exist.

Solution

Try again with a valid path.

2.5.9 W40011 Failure to Create Disk Cache Directory

Symptom

Failed to create disk cache directory [%s]. Result: %s. Reason: %s.

Possible Cause

You do not have the permission for the directory or the directory name is invalid.

Solution

Modify the permission or directory name and try again.

2.5.10 W40012 Invalid Argument

Symptom

Value [%s] for parameter [%s] is invalid. The value must be in the range of [%s] and defaults to [%s].

Possible Cause

N/A

Solution

N/A

2.6 Vector Operator Plugin Errors

2.6.1 E76002 ONNX Model Data Error

Symptom

No ONNX parser is registered for optype [%s].

Possible Cause

No parser is registered for the optype in ONNX model conversion.

Solution

Submit an issue to request for support at <https://gitee.com/ascend>, or remove this type of operators from your model.

2.7 Vector Operator Errors

2.7.1 E80001 Invalid Config Value

Symptom

The %s information is missing in the %s configuration file of OPP package

Possible Cause

N/A

Solution

Try to reinstall the OPP package.

2.8 TBE Pass Compilation Tool Errors (Back-End)

2.8.1 EB0000 Invalid IR

Symptom

Failed to compile the operator. Reason:%s, %s

Possible Cause

The IR is invalid.

Solution

Check that the Compute and Schedule APIs are used correctly. For details, see the TBE Custom Operator Developer Guide. If the error persists, contact Huawei technical support.

2.8.2 EB1000 Primitives Error

Symptom

Primitives check failed. Reason: %s

Possible Cause

- 1.The Primitives are not properly used or not aligned.
2. Incorrect parameter types are used.
3. The correct primitive is not punched on the corresponding axis.

Solution

Check the schedule and compute statements to ensure that the schedule is correct.

2.8.3 EB3000 Memory Error

Symptom

Memory allocation check failed. Reason: %s

Possible Cause

The allocated memory is out of the allowed range.

Solution

Check whether the size of the specified memory is incorrectly calculated.

2.8.4 EB4000 Emit Insn Error

Symptom

Insn check failed. Reason: %s

Possible Cause

The command parameter is incorrect.

Solution

Check whether the offset address of the instruction map is aligned and whether the correct instruction is selected.

2.8.5 EB9000 Common Compilation Optimization Error

Symptom

Compiler Error check failed. Reason: %s

Possible Cause

The IR is invalid.

Solution

N/A

2.9 Auto Tune Errors

2.9.1 EC0000 Remote Device Connection Failure

Symptom

Failed to connect the remote device [IP: %s, Port: %s].

Possible Cause

N/A

Solution

1.Check the IP and port values in the config.json file in the \${install_path}/python/site-packages/auto_tune/estimate/python/ directory.

2.Check whether the remote device is normal.

2.9.2 EC0001 Remote Device Login Failure

Symptom

User [%s] failed to log in to the remote device.

Possible Cause

N/A

Solution

Check the username or password in the config.json file in the \${install_path}/python/site-packages/auto_tune/estimate/python/ directory.

2.9.3 EC0002 Invalid NFS Mount Path

Symptom

Remote mount path [%s] does not exist.

Possible Cause

N/A

Solution

Try again with a valid NFS mount path.

2.9.4 EC0003 File Not Found

Symptom

File [tss/bin/kernel_perf_comm] does not exist in remote mount path [%s].

Possible Cause

N/A

Solution

N/A

2.9.5 EC0004 Permission Denied

Symptom

The permissions for tss/bin or tss/bin/kernel_perf_comm in the remote mount path[%s] are insufficient.

Possible Cause

N/A

Solution

1. Check that the current user has the "**rw**x" permissions for tss/bin and tss/bin/kernel_perf_comm.
2. Check that the current user is the root user.

2.9.6 EC0005 Invalid Argument

Symptom

Value [%s] for parameter [--soc_version] is invalid or the config.json is not set correctly.

Possible Cause

The SoC version doesn't match the device or the config.json is not set correctly.

Solution

N/A

2.9.7 EC0006 Directory Operation Failure

Symptom

Directory operation failed. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.9.8 EC0007 Invalid Environment Variable

Symptom

The environment variable argument is invalid. Value [%s] for environment variable [%s] must be less than [%s].

Possible Cause

N/A

Solution

N/A

2.9.9 EC0008 Invalid Environment Variable

Symptom

Environment variable [%s] is not set. Use bash command [env] to check whether the environment variable is missing, and then use the [export] command to set it correctly.

Possible Cause

N/A

Solution

N/A

2.9.10 EC0009 Resources Insufficient

Symptom

Computing resources are insufficient. Reason: %s

Possible Cause

1. Memory resources are insufficient.
2. The number of file handles exceeds the limit.

Solution

1. Check whether the memory usage reaches the upper limit.
2. Check whether the file handle usage reaches the system limit.

2.9.11 EC0010 Failure to Import Python Module

Symptom

Failed to import Python module [%s].

Possible Cause

N/A

Solution

Check that all required components are properly installed and the specified Python path matches the Python installation directory. (If the path does not match the directory, run set_env.sh in the installation package.)

2.10 RTS Errors

2.10.1 EE1001 Invalid Argument

Symptom

The argument is invalid.Reason: %s

Possible Cause

N/A

Solution

- 1.Check the input parameter range of the function.
- 2.Check the function invocation relationship.

2.10.2 EE4001 Model Binding Errors

Symptom

Failed to bind the stream to the model. %s

Possible Cause

The stream has been bound to another model.

Solution

Remove the repeated binding operation on the stream from the code.

2.10.3 EE4002 Model Unbinding Errors

Symptom

Failed to unbind the stream to the model. %s

Possible Cause

- 1.The stream to be unbound is not bound to the model.
- 2.The model is running.

Solution

1. Check the code to ensure that the stream to be unbound is bound to the model.
2. Ensure that the model is not running.

2.10.4 EE4004 Profiling Enable Errors

Symptom

Failed to enable profiling. %s

Possible Cause

N/A

Solution

Do not enable profiling repeatedly.

2.11 LxFusion & AutoDeploy Errors

2.11.1 EF0000 Invalid Directory

Symptom

Failed to open the directory where repository file [%s] is located with read, write and execute permissions.

Possible Cause

N/A

Solution

Check the existence and permissions of the directory. If the directory does not exist, create one with read, write and execute permissions.

2.11.2 EF0001 Invalid SoC Version

Symptom

soc_version [%s] is invalid. AOE SGAT is not supported on current platform.

Possible Cause

N/A

Solution

N/A

2.11.3 EF0003 Invalid Environment Variable

Symptom

The environment variable argument is invalid. Value [%s] for environment variable [%s] must be [%s]

Possible Cause

N/A

Solution

N/A

2.11.4 WF0000 File Read Failure

Symptom

Failed to open config file [%s]

Possible Cause

N/A

Solution

Check whether the config file exists according to the file path, whether the file is in Json format, and whether the user has the read permission.

2.11.5 WF0001 Invalid Config Value

Symptom

Value [%s] for key [%s] is invalid. The value must be in the range of [%s]

Possible Cause

N/A

Solution

N/A

2.11.6 WF0002 Model Performance Not Optimal

Symptom

Model performance may not be optimal due to changes in UB fusion.

Possible Cause

N/A

Solution

Re-tune the subgraphs with AOE SGAT.

2.11.7 EF1001 Invalid Hierarchical Machine Configuration

Symptom

Configuration format %s is incorrect.

Possible Cause

N/A

Solution

Set the environment variable or resource file correctly.

2.11.8 EF1002 Repository File Not Exist

Symptom

Repository file %s does not exist.

Possible Cause

N/A

Solution

N/A

2.11.9 EF1003 Repository File Format Incorrect

Symptom

The repository file format is incorrect.

Possible Cause

The original repository file has been replaced by a file with a different format.

Solution

Make sure the format of the repository file is the same as that of the original repository file in the installed CANN compiler package.

2.11.10 EF1004 Fused Optimizer Operator Partition Failure

Symptom

Failed when processing node %s (%s), which should be connected with Variable or VariableV2.

Possible Cause

The input graph is incorrect.

Solution

Pass a correct graph.

2.11.11 EF1005 pulp Module Import Failure

Symptom

Failed to import the pulp module.

Possible Cause

N/A

Solution

Install the pulp module before running AutoDeploy.

2.11.12 EF1006 coin-or-cbc Not Installed or Correctly Configured

Symptom

coin-or-cbc Not Installed or Correctly Configured

Possible Cause

N/A

Solution

Perform operations by referring to the AutoDeploy user guide.

2.11.13 EF1007 Python Environment Failure

Symptom

Failed to call Python.

Possible Cause

Python is incorrectly installed or its path is incorrectly configure.

Solution

Recompile and install Python.

2.11.14 EF1008 Dynamic Shape Not Supported

Symptom

Dynamic shape is not allowed in computation graphs.

Possible Cause

N/A

Solution

N/A

2.11.15 EF1009 Invalid Slice Number for a Node

Symptom

The slice number %s that you set for node %s (%s) is invalid.

Possible Cause

An invalid character string is set as the slice number for the node.

Solution

Set the slice number to a positive integer, or do not set the slice number at all, in which case the default slice number will be used.

2.11.16 EF1010 Multiple Slice Numbers Set for a Node Scope

Symptom

Inconsistent custom slice numbers have been found in scope %s, for example %s and %s.

Possible Cause

Multiple slice numbers are set for a node scope to be sliced.

Solution

Set the expected slice number for only one node in the scope, or do not set the slice number at all, in which case the default slice number will be used.

2.12 AOE Errors

2.12.1 EG0000 Invalid Command Line Argument

Symptom

The command line argument is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.12.2 EG0001 Invalid Configuration Argument

Symptom

The configuration argument is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.12.3 EG0002 Invalid Environment Variable Argument

Symptom

The environment variable argument is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.12.4 EG0003 Invalid Option Argument

Symptom

The option argument is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.12.5 EG0010 Directory Operation Failure

Symptom

Directory operation failed. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.12.6 EG0011 File Operation Failure

Symptom

File operation failed. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.12.7 EG0012 Invalid Component

Symptom

Component loading failed. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.12.8 EG0013 Device Operation Failure

Symptom

Device operation failed. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.12.9 EG10000 Network Communication Failure

Symptom

Network communication failed. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.12.10 EG10001 Private Key/Certificate Verification Failure

Symptom

Verification of the private key or certificates failed. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.12.11 WG0000 Memory Insufficient

Symptom

Memory check failed. Reason: %s

Possible Cause

N/A

Solution

N/A

2.13 ACL Errors

2.13.1 EH0001 Invalid Argument

Symptom

Value [%s] for [%s] is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.13.2 EH0002 Null Pointer

Symptom

Argument [%s] must not be null.

Possible Cause

N/A

Solution

Try again with a correct pointer argument.

2.13.3 EH0003 Invalid Path

Symptom

Path [%s] is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.13.4 EH0004 Invalid File

Symptom

File [%s] is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.13.5 EH0005 Invalid AIPP Argument

Symptom

AIPP argument [%s] is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.13.6 EH0006 Unsupported Feature

Symptom

%s is not supported. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.14 HCCL Errors

2.14.1 EI0001 Invalid Environment Variable Configuration

Symptom

Environment variable [%s] is invalid. Reason: %s.

Possible Cause

The environment variable configuration is invalid.

Solution

Try again with valid environment variable configuration.

2.14.2 EI0002 Communication Operation Timeout

Symptom

The wait execution of the Notify register times out. Reason: The Notify register has not received the Notify record from remote rank [%]. base information: %s task information: %s

Possible Cause

1. An exception occurs during the execution on some NPUs in the cluster. As a result, collective communication operation failed.

2. The execution speed on some NPU in the cluster is too slow to complete a communication operation within the timeout interval. (default 1800s, You can set the interval by using HCCL_EXEC_TIMEOUT.)
3. The number of training samples of each NPU is inconsistent.
4. Packet loss or other connectivity problems occur on the communication link.

Solution

1. If this error is reported on part of these ranks, check other ranks to see whether other errors have been reported earlier.
2. If this error is reported for all ranks, check whether the error reporting time is consistent (the maximum difference must not exceed 1800s). If not, locate the cause or adjust the locate the cause or set the HCCL_EXEC_TIMEOUT environment variable to a larger value.
3. Check whether the completion queue element (CQE) of the error exists in the plog(grep -rn 'error cqe'). If so, check the network connection status. (For details, see the TLS command and HCCN connectivity check examples.)
4. Ensure that the number of training samples of each NPU is consistent.

2.14.3 EI0003 Invalid Collective Communication Op Argument

Symptom

In [%s], value [%s] for parameter [%s] is invalid. Reason: The collective communication operator has an invalid argument. Reason[%s]

Possible Cause

N/A

Solution

Try again with a valid argument.

2.14.4 EI0004 Invalid Ranktable Configuration

Symptom

The ranktable or rank is invalid,Reason:[%s]. Please check the configured ranktable. [%s]

Possible Cause

N/A

Solution

Try again with a valid cluster configuration in the ranktable file. Ensure that the configuration matches the operating environment.

2.14.5 EI0005 Inconsistent Collective Communication Arguments Between Ranks

Symptom

The arguments for collective communication are inconsistent between ranks: tag [%s], parameter [%s], local [%s], remote [%s]

Possible Cause

N/A

Solution

Check whether the training script and ranktable of each NPU are consistent.

2.14.6 EI0006 Get Socket Timeout

Symptom

Getting socket times out. Reason: %s

Possible Cause

N/A

Solution

1. Check the rank service processes with other errors or no errors in the cluster.
2. If this error is reported for all NPUs, check whether the time difference between the earliest and latest errors is greater than the connect timeout interval (120s by default). If so, adjust the timeout interval by using the HCCL_CONNECT_TIMEOUT environment variable.
3. Check the connectivity of the communication link between nodes. (For details, see the TLS command and HCCN connectivity check examples.)

2.14.7 EI0007 Allocation Failure

Symptom

Failed to allocate resource[%s] with info [%s]. Reason: Memory resources are exhausted.

Possible Cause

Failed to allocate memory or the Notify register due to resource insufficiency.

Solution

N/A

2.14.8 EI0008 Inconsistent CANN Versions

Symptom

The CANN versions are inconsistent: tag [%s], local_version [%s], remote_version [%s]

Possible Cause

N/A

Solution

Install the same CANN version.

2.15 HCCP Errors

2.15.1 EJ0001 HCCP Process Initialization Failure

Symptom

Failed to initialize the HCCP process. Reason: Maybe the last training process is running.

Possible Cause

N/A

Solution

Wait for 10s after killing the last training process and try again.

2.15.2 EJ0002 Environment Error

Symptom

The network port is down.Suggest: %s

Possible Cause

N/A

Solution

N/A

2.15.3 EJ0003 Bind Failed

Symptom

Failed to bind the IP port. Reason: %s

Possible Cause

N/A

Solution

N/A

2.15.4 EJ0004 Check Ip Failed

Symptom

Check ip fail. Reason: %s

Possible Cause

N/A

Solution

N/A

2.16 Profiling Errors

2.16.1 EK0001 Invalid Argument

Symptom

Value [%s] for [%s] is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.16.2 EK0002 Interface Call Sequence Error

Symptom

Failed to call %s before calling %s.

Possible Cause

N/A

Solution

N/A

2.16.3 EK0003 Profiling Config Setting Error

Symptom

Failed to set the %s to [%s]. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.16.4 EK0004 API Support Error

Symptom

[%s] is not supported in %s.

Possible Cause

N/A

Solution

N/A

2.16.5 EK0201 Memory Allocation Error

Symptom

Failed to allocate host memory for Profiling: %s.

Possible Cause

Available memory is insufficient.

Solution

Close unused applications.

2.16.6 EK9999 System Terminated

Symptom

An unknown error occurred. Please check the log.

Possible Cause

N/A

Solution

N/A

2.17 Driver Errors

2.17.1 EL0001 Device Absent/Abnormal

Symptom

Device not found.

Possible Cause

N/A

Solution

N/A

2.17.2 EL0002 Invalid Device ID

Symptom

The device ID is invalid.

Possible Cause

N/A

Solution

N/A

2.17.3 EL0003 Invalid Argument

Symptom

The argument is invalid.

Possible Cause

N/A

Solution

Try again with a valid argument.

2.17.4 EL0004 Memory Allocation Failure

Symptom

Failed to allocate memory.

Possible Cause

Available memory is insufficient.

Solution

Close applications not in use.

2.17.5 EL0005 Resource Busy

Symptom

The resources are busy.

Possible Cause

1. The resources have been occupied.
2. The device is being reset.
3. Software is not ready.

Solution

1. Close applications not in use.
2. Wait for a while and try again.

2.17.6 EL0006 Insufficient Resources

Symptom

The resources are insufficient.

Possible Cause

N/A

Solution

Close applications not in use.

2.17.7 EL0007 No Permission

Symptom

You have no permission.

Possible Cause

1. The resources have been occupied.
2. You have no permission on this API.

Solution

1. Close applications not in use.
2. Check that you have the required permission.

2.17.8 EL0008 Insufficient Event Resources

Symptom

The event resources are insufficient.

Possible Cause

N/A

Solution

Close applications not in use.

2.17.9 EL0009 Insufficient Stream Resources

Symptom

The stream resources are insufficient.

Possible Cause

N/A

Solution

Close applications not in use.

2.17.10 EL0010 Insufficient Notify Resources

Symptom

The Notify resources are insufficient.

Possible Cause

N/A

Solution

Close applications not in use.

2.17.11 EL0011 Insufficient Model Resources

Symptom

The model resources are insufficient.

Possible Cause

N/A

Solution

Close applications not in use.

2.17.12 EL0012 Full Service Queue

Symptom

The service queue is full.

Possible Cause

Available resources are insufficient due to busy services.

Solution

Reduce service loads.

2.17.13 EL0013 Insufficient CDQM Resources

Symptom

The CDQM resources are insufficient.

Possible Cause

Available resources are insufficient due to busy services.

Solution

Reduce service loads.

2.17.14 EL0014 Operation Unsupported

Symptom

The operation is unsupported.

Possible Cause

N/A

Solution

N/A

2.17.15 EL0015 Invalid Device Access

Symptom

The device access is invalid.

Possible Cause

1. The device has been bound to another container.
2. You have no permission on this device.

Solution

1. Check your configurations.
2. Check whether there are residual container processes.

2.18 Queue Schedule Errors

2.18.1 EM9001 Group Attaching Failure

Symptom

Failed to attach the queue_schedule process to the group.

Possible Cause

N/A

Solution

Submit an issue to request for support at <https://gitee.com/ascend> or contact Huawei technical support.

2.18.2 EM9002 Queue_Schedule Initialization Failure

Symptom

Failed to initialize the queue_schedule process.

Possible Cause

N/A

Solution

Submit an issue to request for support at <https://gitee.com/ascend> or contact Huawei technical support.

2.19 DVPP Errors

2.19.1 EN0001 Invalid Argument

Symptom

Value [%s] for [%s] is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.19.2 EN0002 Some Argument Invalid

Symptom

Some argument is invalid. See the %s API description in AscendCL API Reference.

Possible Cause

N/A

Solution

N/A

2.19.3 EN0003 Null Pointer

Symptom

Argument [%s] is null.

Possible Cause

N/A

Solution

Try again with a correct pointer argument.

2.19.4 EN0004 Unsupported Feature

Symptom

The operation is not supported. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.19.5 EN0005 Channel Resource Error

Symptom

An error occurred on channel resource %s. Reason: %s.

Possible Cause

N/A

Solution

Check the input channel ID or the number of created channels.

2.19.6 EN0006 Memory Address Error

Symptom

The passed memory address is incorrect.

Possible Cause

N/A

Solution

Use the `hi_mpi_dvpp_malloc` API to allocate memory.

2.19.7 EN0007 Operation Timeout

Symptom

%s times out. Reason: %s.

Possible Cause

1. The input timeout interval is improper;
2. The buffer is full; 3. An internal error occurred.

Solution

Set the timeout interval to a proper value.

2.19.8 EN0008 Unsupported JPEG File

Symptom

The input JPEG image is not supported. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.19.9 EN0009 Unsupported PNG File

Symptom

The input PNG image is not supported. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.20 AMCT Errors

2.20.1 EO0001 Invalid compression_optimize_conf Value

Symptom

The compression configuration value is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.20.2 EO0002 Invalid compression_optimize_conf Path

Symptom

The compression configuration path is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.20.3 EO0003 Invalid Model

Symptom

The model is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.20.4 EO0004 Command Error

Symptom

The command is invalid. Reason: %s.

Possible Cause

N/A

Solution

N/A

2.20.5 EO9999 System Terminated

Symptom

An unknown error occurred. Please check the log.

Possible Cause

N/A

Solution

N/A

2.21 Public Operator Errors

2.21.1 EZ0501 Unsupported Operator

Symptom

IR for Op [%s, optype [%s]], is not registered.

Possible Cause

IR for the operator type is not registered.

Solution

Submit an issue to request for support at <https://gitee.com/ascend>, or remove this type of operators from your model.

2.21.2 EZ3002 Unsupported Operator

Symptom

Optype [%s] of Ops kernel [%s] is unsupported. Reason: %s.

Possible Cause

The operator type is unsupported in the operator information library due to specification mismatch.

Solution

Submit an issue to request for support at <https://gitee.com/ascend>, or remove this type of operators from your model.

2.21.3 EZ3003 Unsupported Operator

Symptom

No supported Ops kernel and engine are found for [%s], optype [%s].

Possible Cause

The operator is not supported by the system. Therefore, no hit is found in any operator information library.

Solution

1. Check that the OPP component is installed properly.
2. Submit an issue to request for the support of this operator type.

2.21.4 EZ9010 Unsupported Operator

Symptom

No parser is registered for Op [%s, optype [%s]].

Possible Cause

No parser is registered for the operator type.

Solution

Submit an issue to request for support at <https://gitee.com/ascend>.

2.22 E*9999 Inner Error

系统内部错误码，请根据报错信息进行问题排查。如果问题仍然存在，请联系华为支持工程师：

登录<https://gitee.com/ascend>网站，提issue求助，华为支持工程师会联系您处理。

3 故障案例集

- [3.1 AscendCL常见故障](#)
- [3.2 GE常见故障](#)
- [3.3 Driver常见故障](#)
- [3.4 Runtime常见故障](#)
- [3.5 HCCL&HCCP常见故障](#)
- [3.6 getNext故障处理](#)

3.1 AscendCL 常见故障

3.1.1 背景知识

AscendCL (Ascend Computing Language) 提供Device管理、Context管理、Stream管理、内存管理、模型加载与执行、算子加载与执行、媒体数据处理等C API库供用户开发深度神经网络应用，用于实现目标识别、图像分类等功能。用户可以通过第三方框架调用AscendCL接口，以便使用昇腾AI处理器的计算能力；用户还可以使用AscendCL封装实现第三方lib库，以便提供昇腾AI处理器的运行管理、资源管理能力。

3.1.2 常见故障分析与处理

3.1.2.1 内存未释放

适用场景

- 业务场景：DVPP、模型推理、应用/算子开发
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

测试用例长稳运行时，出现内存泄漏的现象，内存占用持续上升。

可能原因

分析用户代码，可能存在以下故障原因：

系统存在只申请内存不释放内存的问题，正常情况下，内存申请与释放必须成对出现。

处理步骤

针对分析的故障可能原因，可以参考下面步骤处理：

排查所有内存申请和释放的地方，保证申请与释放一一对应。例如[aclrtMalloc](#)与[aclrtFree](#)，[aclrtMallocHost](#)与[aclrtFreeHost](#)、[aclrtCreateStream](#)与[aclrtDestroyStream](#)等。

3.1.2.2 APP 使用 dvpp 接口编译失败

适用场景

- 业务场景：DVPP
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

编译提示DVPP的相关接口未定义，编译报错，日志关键字包括：undefined reference to ***

可能原因

分析上述日志信息，可能存在以下故障原因：

DVPP与AscendCL已经分别打包到libacl_dvpp.so与libascendcl.so，测试用例使用了DVPP的相关接口，但没有链接libacl_dvpp.so。

处理步骤

针对分析的可能原因，可以参考下面步骤处理：

排查测试用例是否使用了预处理的接口，但未链接libacl_dvpp.so。如果未链接，需要在编译文件中链接libacl_dvpp.so。

需要排查CmakeLists中的target_link_libraries()选项是否连接了acl_dvpp这个target。

示例：

```
add_executable(main
  utils.cpp
  main.cpp)
target_link_libraries(main
  ascendcl acl_dvpp stdc++)
```

3.1.2.3 使用 dump 功能未获取 dump 结果

适用场景

- 业务场景：模型推理、应用/算子开发
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

日志显示正确执行了Dump功能，但在Dump结果路径下没有Dump的结果。日志信息包含了以下关键字：

```
[INFO] ASCENDCL ***** "HandleDumpConfig end in HandleDumpConfig."  
[INFO] ASCENDCL ***** "set HandleDumpConfig success in aclInit"
```

可能原因

分析上述日志信息，可能存在以下故障原因：Dump配置的模型名与实际的模型名不匹配。

处理步骤

针对分析的故障可能原因，可以参考下面步骤处理：

检查Dump配置文件acl.json，确保Dump配置文件合法，例如model_name是否配置正确。示例如下：

```
{  
  "dump":{  
    "dump_list":[  
      {  
        "model_name":"ResNet-50",  
        "layer":[  
          "convlconvl_relu"  
        ]  
      },  
      {  
        "model_name":"mxnet-model"  
      }  
    ],  
    "dump_mode":"output",  
    "dump_path":"/home/test/output/dump"  
  }  
}
```



```
}
```

通过ATC命令生成模型的json文件，在json文件中查找“name”字段对应值，查找模型名称和算子名称，模型名称在“graph”字段外、算子名称在“graph”字段内。

3.1.2.4 Event 数量超过上限导致 aclrtRecordEvent 接口返回失败

适用场景

- 业务场景：DVPP，模型推理，应用/算子开发
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

调用aclrtRecordEvent接口在Stream中记录一个Event时，日志中存在报错信息，关键日志信息如下，提示Event ID申请失败：

```
[ERROR] DRV *****[devdrv][drvEventIdAlloc 647] error.  
[ERROR] RUNTIME ***** EventIdAlloc: drvEventIdAlloc:errorCode = 7  
[INFO] RUNTIME EventIdAlloc:id = -1
```

可能原因

分析上述日志信息，可能存在以下故障原因：Event ID的数量超过上限。

处理步骤

多Stream之间同步等待的场景下，Event ID的资源是可以复用的，复用Event ID的流程是：在调用aclrtRecordEvent接口+aclrtStreamWaitEvent接口后，若指定的Event已完成，则需要及时调用aclrtResetEvent接口释放Event资源。

需要用户按照复用Event ID的流程优化代码逻辑。

3.1.2.5 执行单算子产生 coredump 的定位处理

适用场景

- 业务场景：应用/算子开发
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

单算子执行结束，出现重复释放内存，导致coredump，屏幕显示关键日志信息：

```
double free or corruption (! prev)
```

可能原因

分析屏显日志信息，可能存在以下故障原因：代码中出现重复释放内存的操作。

处理步骤

通过gdb挂载可执行文件，通过查看栈信息做排查：

- 重复释放内存代码是否是用户自身代码bug，如果是则需修复代码bug。
- 提供栈信息，通过<https://gitee.com/ascend>网站提交issue获取帮助。

具体步骤如下：

步骤1 gdb挂载可执行文件。

步骤2 执行gdb调试。

步骤3 查看调用栈。

如果该问题非用户代码问题，需要联系华为算子开发工程师定位排查。您可以通过<https://gitee.com/ascend>网站提交issue获取帮助。

----结束

3.1.2.6 进程异常退出后重新执行任务失败

适用场景

- 业务场景：DVPP、模型推理、应用/算子开发
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

进程异常退出时，包括强行终止任务（如ctrl + c或者kill命令终止进程）的场景，然后重新启动任务失败。

可能原因

进程异常退出时，只能依赖系统检测到程序退出后才进行资源释放，释放资源最长需要一分钟的执行时间。如果在未执行完资源释放前执行新的任务，可能导致新执行的任务失败。

处理步骤

进程异常退出后需要等待一分钟，才能保证下一次重新执行任务成功。

3.1.2.7 进程异常时资源清理的处理建议

适用场景

- 业务场景：DVPP、模型推理、应用/算子开发
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

用户捕获异常退出信号，并在信号处理函数中释放已申请资源，下一次执行时会报执行失败。此时查看日志，会发现unbind model stream failed报错。

可能原因

进程异常时，Host侧内核态驱动会自动检测并发起对应进程Device侧资源释放的流程，不需要用户捕获进程异常的信号并主动完成清理。若用户主动释放，会影响到系统的资源释放流程。

处理步骤

用户无需关注进程异常退出信号，不要对异常退出信号做处理。

3.1.2.8 单算子匹配失败

适用场景

- 业务场景：应用/算子开发
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

单算子执行过程中，出现匹配失败，日志显示如下类似信息。

离线加载执行场景：

EH9999 [Match][OpModel]failed to match model, opName = xxx Has not been compiled or loaded, Please make sure the op executed and the op compiled is matched, you can check the op type, op inputs number, outputs number, input format, origin format, datatype, memtype, attr, dim range, and so on.

在线加载执行场景：

EH9999 [Match][OpModel]MatchOpModel fail from static map or dynamic map. Please make sure the op executed and the op compiled is matched, you can check the op type, op inputs number, outputs number, input format, origin format, datatype, memtype, attr, dim range, and so on.

可能原因

- 离线场景：算子om文件未出现在aclSetModelDir中指定的路径下。
- 在线场景：一些特殊的匹配规则未适配，导致算子匹配失败。

处理步骤

- 离线场景：重新编译缺少的算子，并复制到aclSetModelDir中指定的路径下，余下步骤按离线单算子推理步骤进行。
- 在线场景：根据报错信息检查前后两次编译的算子的opType、dataType、format等信息是否一致。

3.1.2.9 ACL 执行无输出无报错

适用场景

- 业务场景：模型推理、应用/算子开发
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

ACL执行过程中，无接口报错，但是没得到预期结果。

可能原因

- ACL执行过程中链接到了stub中的so。
- 异步场景下，在拷贝输出数据时没有做流同步。

处理步骤

针对第一种情况：使用ldd命令查看执行文件链接的so是否正确，保证链接了有效的so。

针对第二种情况：请参考《[应用软件开发指南（C&C++）](#)》手册中“AscendCL API参考”章节的“aclmdlExecuteAsync”接口介绍，排查代码。

3.1.2.10 动态 shape 推理申请内存失败

适用场景

- 业务场景：DVPP、模型推理、应用/算子开发
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

模型推理过程中，申请了大小为0的内存，日志报错信息中包含以下关键信息：

```
[INFO] ASCENDCL ***** start to execute aclrtMalloc, size = 0  
[ERROR] ASCENDCL ***** malloc size must be greater than zero
```

可能原因

模型为动态shape模型，模型的输出shape中含有-1，所以直接调用aclmdlGetOutputSizeByIndex接口取到的size为0。

然后申请了大小为0的内存，导致失败。

处理步骤

请参考《[应用软件开发指南（C&C++）](#)》的“模型动态Shape输入推理”章节内容处理。

在aclmdlGetOutputSizeByIndex取到size为0时，用户需要预估一块较大的内存。

3.1.2.11 异步拷贝调用查询接口报错

适用场景

- 业务场景：推理
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

通过event实现H2D或D2H异步拷贝任务的同步等待时，在调用aclrtQueryEventStatus确认任务完成后，先调用aclrtFreeHost释放Host内存再调用aclrtDestroyEvent接口，可能会有如下报错信息打印：

```
[INFO] DRV(78295,python):2023-01-10-11:21:48.787.850 [ascend][coudid: 78295, 78295][drv][common][share_log_read 544][ascend] [ERR006] [devmm] <python:3940,3940> Set free error. (ref_lock=1; ref_free=0; ref_count=3)
[ascend] [ERR006] [devmm] <python:3940,3940> Oper address failed. (va=0x120043200000; ref_flag=0x108; ref_lock=0; ref_free=0; ref_count=3; context=1; says=0)
[ascend] [ERR006] [devmm] <python:3940,3940> Vaddress can not oper. (cmd=0x42204004; cmd_id=0x4; ret=-22)
[ERR006] DRV(78295,python):2023-01-10-11:21:48.787.846 [ascend][coudid: 78295, 78295][drv][devmm][devmm_ioctl_free_pages 135][errno:16, 17] Ioctl device error. (ret=-17)
[ERR006] DRV(78295,python):2023-01-10-11:21:48.787.851 [ascend][coudid: 78295, 78295][drv][devmm][devmm_virt_heap_free_pages 294][errno:16, 17] Devmm_ioctl_free failed. (ptr=0x120043200000; heap_type=4026417729)
[ERR006] DRV(78295,python):2023-01-10-11:21:48.787.856 [ascend][coudid: 78295, 78295][drv][devmm][devmm_free_physmem_heap_oper 323][errno:16, 17] Heap ops failed. (ret=-17; va=0x120043200000)
[ERR006] DRV(78295,python):2023-01-10-11:21:48.787.862 [ascend][coudid: 78295, 78295][drv][devmm][devmm_free_devicemem_heap_process 993][errno:16, 17] Free error. (va=0x120043200000; size=123731968; total=123731968; ret=-17)
[ERR006] DRV(78295,python):2023-01-10-11:21:48.787.870 [ascend][coudid: 78295, 78295][drv][devmm][devmm_free_to_normal_heap 704][errno:16, 17] Virt_heap_free_mem failed. (ret=-17; va=0x120043200000)
[ERR006] RUNTIME(78295,python):2023-01-10-11:21:48.788.013 [cpu_driver.cc:1531]78295 HostMemFree:[FINAL][FINAL]report error module_type=1, module_name=CL2959
[ERR006] RUNTIME(78295,python):2023-01-10-11:21:48.788.021 [cpu_driver.cc:1531]78295 HostMemFree:[FINAL][FINAL](drv api) halMemFree failed: drvRetCode=17
[ERR006] RUNTIME(78295,python):2023-01-10-11:21:48.788.075 [logger.cc:858]78295 HostFree:[FINAL][FINAL]free host memory failed.
[ERR006] RUNTIME(78295,python):2023-01-10-11:21:48.788.087 [api_2-cc:107]78295 aclrtFreeHost:[FINAL][FINAL]retCode=0x7859, desc=[Driver error:internal error], InnerCode=0x7020022
[ERR006] RUNTIME(78295,python):2023-01-10-11:21:48.788.102 [error_message_manage.cc:49]78295 FunErrorReason:[FINAL][FINAL]report error module_type=3, module_name=EE8888
[ERR006] RUNTIME(78295,python):2023-01-10-11:21:48.788.110 [error_message_manage.cc:49]78295 FunErrorReason:[FINAL][FINAL]aclrtFreeHost execute Failed, reason=[Driver error:internal error]
[ERR006] ASCENDCL(78295,python):2023-01-10-11:21:48.788.133 [memory.cpp:242]78295 aclrtFreeHost: [FINAL][FINAL]free host memory failed, runtime result = 507899
```

可能原因

报错是因为使用了异步拷贝任务之后下发了一个event record任务，期望使用aclrtQueryEventStatus查询到event record任务是否完成，从而判断异步拷贝任务是否完成，而后释放内存调用aclrtFreeHost。

实际上aclrtQueryEventStatus查询到的是Device执行完任务，并未透传到Host侧，所以此时释放内存，未先销毁Event会有时序问题导致报错。

处理步骤

处理该问题可以参考以下方案：

方案一：使用aclrtSynchronizeStream接口判断任务是否执行完成。

方案二：使用aclrtQueryEventStatus接口时，先调用aclrtDestroyEvent接口，再调用aclrtFreeHost接口，保证无时序问题。

3.1.3 数据预处理常见问题（昇腾 310P）

3.1.3.1 视频解码丢帧/丢包

适用场景

- 业务场景：DVPP
- 适用处理器：昇腾310P AI处理器
- 处理器形态：EP、RC

现象描述

视频解码丢帧，出现重影或不连续现象。查看Device侧日志，发现日志中存在以下几个报错的内容信息中的一个或多个。

- 日志报错信息（1）
[HiDvpp][A618] [Vfmw]:ppspps_check [Line]:6803 pps with this pic_parameter_set_id = %d haven't decode
[HiDvpp][A618] [Vfmw]:process_slice_header_first_part [Line]:7401 PPS or SPS of this slice not valid
[HiDvpp][A618] [Vfmw]:h264_dec_slice [Line]:7915 sliceheader dec err
- 日志报错信息（2）
[HiDvpp][A618] [Vfmw]:h264_dec_slice [Line]:7983 init pic err, find next recover point or next valid sps, pps, or exit
[HiDvpp][A618] [Vfmw]:h264_dec_slice [Line]:3716 dec list error, ret=-1
[HiDvpp][A618] [Vfmw]:receive_packet [Line]:10676 nal_release_err
- 日志报错信息（3）
[HiDvpp][A618] [Vfmw]:init_list_x [Line]:4829 for P slice size of list equal 0.ctx->dpb.ref_frames_in_buffer:0.
[HiDvpp][A618] [Vfmw]:dec_list [Line]:5068 init list error.
[HiDvpp][A618] [Vfmw]:h264_dec_list [Line]:4829 dec_list error, ret=-1
[HiDvpp][A618] [Vfmw]:h264_dec_one_nal [Line]:10298 slice_check failed, clear current slice.
- 日志报错信息（4）
[HiDvpp][A618] [Vfmw]:init_list_x [Line]:4865 for B slice size of two list all equal 0.
[HiDvpp][A618] [Vfmw]:dec_list [Line]:5068 init list error.
[HiDvpp][A618] [Vfmw]:h264_dec_list [Line]:4829 dec_list error, ret=-1
- 日志报错信息（5）
[HiDvpp][A618] [Vfmw]:hevc_vps_sps_pps_check [Line]:7300 pps with this pic_parameter_set_id = 0 haven't be decoded
[HiDvpp][A618] [Vfmw]:hevc_dec_slice_segment_header [Line]:3857 hevc_vps_sps_pps_check != HEVC_DEC_NORMAL
[HiDvpp][A618] [Vfmw]:hevc_inquire_slice_property [Line]:9004 hevc_dec_slice_segment_header dec err
[HiDvpp][A618] [Vfmw]:hevc_dec_decode_packet[Line]:9004 hevc_inquire_slice_property error.
- 日志报错现象（6）
[HiDvpp][A618] [Vfmw]:hevc_ref_pic_process [Line]:3474 ref frame(poc 15) lost.
[HiDvpp][A618] [Vfmw]:hevc_create_lost_picture [Line]:5839 DPB no suited fs for lost pic.
[HiDvpp][A618] [Vfmw]:hevc_create_lost_picture [Line]:5847 take poc(17) to create lost poc(15).

可能原因

分析上述日志报错信息现象，分别可能存在以下可能原因：

- 日志报错信息（1）可能原因：H264码流缺少IDR帧
- 日志报错信息（2）可能原因：H264码流缺少I帧
- 日志报错信息（3）可能原因：H264码流缺少P帧
- 日志报错信息（4）可能原因：H264码流缺少B帧
- 日志报错信息（5）可能原因：H265码流缺少IDR帧
- 日志报错信息（6）可能原因：H265码流缺少I帧或者P帧

处理步骤

针对可能原因分析，参考以下步骤处理：

步骤1 检查输入的源码流是否有问题。

使用第三方工具（如：eseye u）对输入码流进行检查，查看码流是否异常。

步骤2 若查看的源码流结果为正常，则可能码流在传输给设备侧VDEC的过程中遭到破坏，需要在调用发送码流接口之前，通过fwrite函数将输送给VDEC的码流保存下来。

- 使用第三方工具对保存的码流进行检查，如果码流异常，用户需自行排查将码流送进去之前是否有送流问题。

- 通过对应版本的sample，解码这段保留下来的码流，验证码流是否正常或VDEC是否支持该格式。

如果sample 解码正常，那就是开发代码有问题，可以参考VDEC示例代码，找到对应的视频解码的代码参考优化。

----结束

3.1.3.2 视频解码花屏

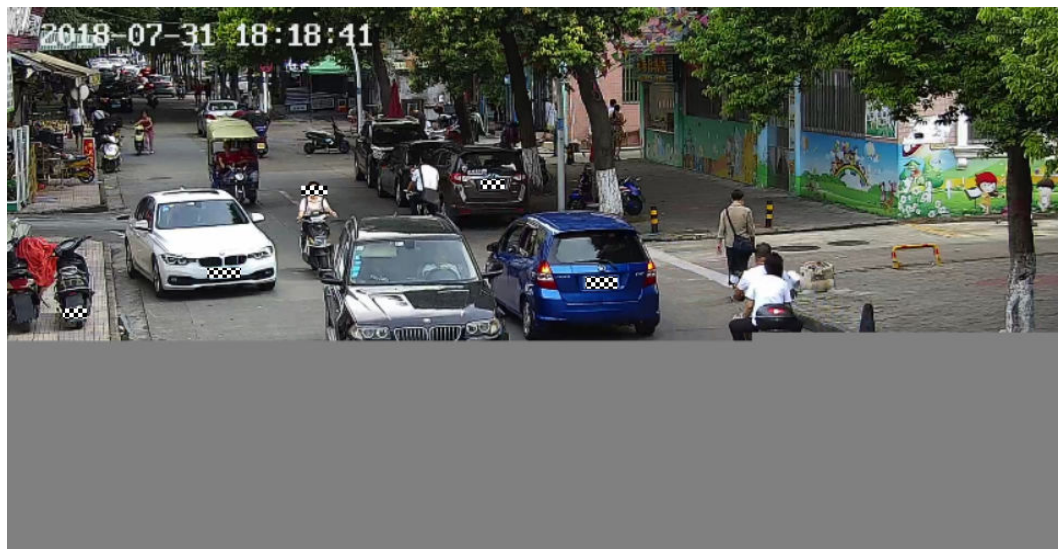
适用场景

- 业务场景：DVPP
- 适用处理器：昇腾310P AI处理器
- 处理器形态：EP、RC

现象描述

输入码流给VDEC进行解码，得到的解码数据不正确，产生花屏现象，如图3-1所示。并且日志中存在类似“decode error”、“input stream error, can't decode, report to user”、“num_ref_idx_l0_active(30) out of range(0,16)”信息。

图 3-1 视频花屏



可能原因

输入的码流中某些帧数据不完整、存在坏帧，导致硬件解码产生花屏。

处理步骤

针对可能原因分析，参考以下步骤处理：

步骤1 检查输入的源码流是否有问题。

使用第三方工具（如：eseye u等）对输入码流进行解码播放，查看是否存在花屏，若不花屏则进行**步骤2**；若花屏则替换输入码流。

步骤2 若查看的源码流结果为正常，则可能码流在传输给设备侧VDEC的过程中遭到破坏，需要在调用发送码流接口之前，通过fwrite函数将输送给VDEC的码流保存下来。

- 使用第三方工具对保存的码流进行检查，如果码流异常，用户需自行排查将码流送进去之前是否有送流问题。
- 通过对应版本的sample，解码这段保留下来的码流，验证码流是否正常或VDEC是否支持该格式。

如果sample 解码正常，那就是开发代码有问题，可以参考《[应用软件开发指南（C&C++）](#)》中的VDEC示例代码，找到对应的视频解码的代码参考优化。

----结束

3.1.3.3 视频解码性能问题

适用场景

- 业务场景：DVPP
- 适用处理器：昇腾310P AI处理器
- 处理器形态：EP、RC

现象描述

VDEC解码性能下降，小于VDEC对外发布的性能规格，导致发生卡顿等现象。

可能原因

分析VDEC性能下降导致的卡顿故障，可能原因有：

- 视频解码回调函数中耗时过长，影响解码性能。
- 输入码流中I帧比例过大，解码I帧耗时比解码P帧耗时长，影响解码性能。
- 输入码流中存在异常帧，影响解码性能。

处理步骤

针对上述可能的故障原因，可以参考以下方式进行处理：

- 步骤1** 在回调函数中打点测试耗时，查看耗时是否过长(回调函数允许的最大耗时和帧率相关，计算公式为：最大耗时=1/帧率，例如帧率=30fps，则最大耗时=1/（30fps）=0.033s）。
- 步骤2** 使用第三方工具打开输入码流，查看I帧比例是否过大。一般GOP值为30(即I帧间隔为30)，如果I帧比例过大，则需要替换成正常码流进行性能测试。
- 步骤3** 使用第三方工具打开输入码流，查看是否存在异常帧（如用第三方工具打开显示花屏或解码报错），如果存在异常帧，会造成不满足规格的现象。

----结束

3.1.3.4 视频解码失败

适用场景

- 业务场景：DVPP

- 适用处理器：昇腾310P AI处理器
- 处理器形态：EP、RC

现象描述

视频解码失败，日志中打印如下所示的异常信息：

- 日志信息：Invalid width or height
[HiDvpp][A618] [Vfmw] vdec_drv_check_video_wh [Line]:127 device 0 chn 0 type(255) pic size(32x32) is out of range[128,128]x[4096,4096].
- 日志信息：bit_depth_luma(*) not equal *
[HiDvpp][A618] [Vfmw] process_sps [Line]:8994 bit_depth_luma(%d) not equal 8.
[HiDvpp][A618] [Vfmw] hevc_process_sps [Line]:1462 chn 0, bit_depth_luma(9) is not supported, hevc only support 8 or 10 bit depth.

可能原因

针对上述日志分析，可能存在输入码流规格不符合以下要求：

- VDEC支持的码流宽高分别是width：128~4096，height：128~4096。
- VDEC支持8bit的h264码流，8bit或10bit的h265码流。

定位思路

- 1、排查输入的码流宽高是否满足VDEC要求。
- 2、排查输入的码流类型是否满足VDEC要求。

处理步骤

检查输入给VDEC的码流是否满足要求：

使用第三方工具（如：eseye u等）对输入码流进行检查，查看码流类型、码流宽高等信息是否符合要求。如果码流不满足要求，请替换符合要求的码流。

3.1.3.5 JPEG 图片解码失败

适用场景

- 业务场景：DVPP
- 适用处理器：昇腾310P AI处理器
- 处理器形态：EP、RC

现象描述

JPEGD模块解码失败，查看日志有类似如下报错信息：

日志信息（1）：

```
just support jpeg with YUV 444 422 420 400
do not support progressive mode
do not support arithmetic code, support huffman code only
```

日志信息（2）：

```
EOI segment of the stream is invalid
```

可能原因

分析上面日志信息，可能存在以下可能原因：

1. 数据格式不支持
JPEGD只支持huffman编码(colorspace: yuv, subsample: 444/422/420/400)，不支持算术编码，不支持渐进编码，不支持jpeg2000格式。
2. 图像数据不完整

处理步骤

针对上述可能原因，请按以下方式处理：

步骤1 针对目前不支持的超规格图像格式，建议用户自行使用第三方软件解码。

步骤2 针对图像数据不完整，根据报错提示，通过第三方软件查看原图像二进制进行确认。

例如“EOI segment of the stream is invalid”或“EOI segment of the stream is invalid, it should be FFD9. Try software decoding.”报错，表示图像缺失最后的EOI结束符，对应图像二进制类似下图所示。正常JPEG图片最后应该由标记码FF D9结束，该数据最后缺失FF D9标记码。

如果确认原图数据不完整，报错属于正常现象，需更换数据。

```
00005e90h: AB 79 F4 E3 4B CE 38 9C B8 B7 D2 B2 1C 59 4E A5 ; 坳齏K?湮分?YN?  
00005ea0h: 5F 73 B6 36 D6 FF 00 E5 7F FA BF F7 71 A2 47 FC ; _s??.? 齏 ?  
00005eb0h: 91 FC D3 FC 30 4C 6A 84 EB 8E AF 5B 8B 2A 56 BE ; 廖狱0Lj熬喝[?V?  
00005ec0h: 49 FA 60 D5 6D F7 67 3F 11 F9 4A D6 E2 CA 35 2A ; I鸱譚齏?.鸱肘?*<br>00005ed0h: C0 13 FA B6 FD B0 0A 9D C7 FD BF F8 60 CD 43 E4 ; ? .齏 馱蚌?  
00005ee0h: 83 F7 47 FF 00 B6 DE 31 80 2C BE E0 60 01 A7 93 ; 淦G .撥1€,距`.  
00005ef0h: FB 23 DF ED 8C FE 21 DF F4 7F D4 1F D5 8D 0D 7F ; ?唔岨!嘩?魯.0  
00005f00h: 33 F9 9F E3 8C F1 8C 7F 00 00 ; 3弱銓馱..
```

步骤3 如果原图像数据完整，可能数据在传输过程中存在损坏，需要在调用acldvppJpegDecodeAsync之前，通过fwrite函数将传输给JPEGD的码流保存下来，与原图进行二进制比较。如果不一致，传输过程出现数据缺失，需自行进一步定位传输过程数据缺失问题。

----结束

3.1.3.6 VPC 参数校验失败

适用场景

- 业务场景：DVPP图像处理
- 适用处理器：昇腾310P AI处理器
- 处理器形态：EP、RC

现象描述

调用VPC功能接口返回0xA0078003，即HI_ERR_VPC_ILLEGAL_PARAM，参数超出合法范围。查看日志有类似如下报错信息：

日志信息（1）：

```
resize width (224) is greater than output width (120)
```

日志信息 (2) :

```
crop width 120 cannot be greater than input width 100
```

日志信息 (3) :

```
resize interpolation [8] is not supported
```

日志信息 (4) :

```
buffer size(50176) is smaller than need buffer size(95264) when format is 1
```

日志信息 (5) :

```
on device 0, num 0 input addr (start 0x100020003000 end 0x100020004000) is illegal
```

可能原因

针对上面日志信息分析，可能存在以下对应原因：

- 日志信息 (1) : VPC缩放宽大于输出宽，
- 日志信息 (2) : VPC抠图宽大于输入宽
- 日志信息 (3) : 输入的缩放算法不正确
- 日志信息 (4) : 图片内存大小不够
- 日志信息 (5) : 图片地址校验失败

定位思路

1. 根据日志描述的错误信息，找到VPC对应的配置参数，根据提示进行修改。
2. 根据日志描述的错误信息，参考《[应用软件开发指南（C&C++）](#)》中DVPP接口章节VPC参数的约束修改。

处理步骤

根据提示的错误信息进行修改：

步骤1 如果为日志信息 (1)，修改resize宽，使其小于等于输出宽。

步骤2 如果为日志信息 (2)，修改抠图宽，使其小于输入宽。

步骤3 如果为日志信息 (3)，修改缩放算法为昇腾310P AI处理器支持的类型。

步骤4 如果为日志信息 (4)，需要按照格式申请足够的内存，并正确配置buffer_size

步骤5 如果为日志信息 (5)，使用hi_mpi_dvpp_malloc或acldvppMalloc 申请图片地址

----结束

3.1.3.7 VENC 创建通道失败

适用场景

- 业务场景：DVPP
- 适用处理器：昇腾310P AI处理器
- 处理器形态：EP、RC

现象描述

调用VENC创建通道的接口hi_mpi_venc_create_chn返回值非0，通道创建失败。

可能原因

导致创建通道失败可能原因有以下：

- 用户传入的通道ID超出了规定的合法范围，VENC规定通道ID在[0,255]总共256个通道。
- 用户传入的通道属性参数不在规定的合理范围内或设置了暂不支持的参数。
- 用户试图创建已经存在的通道，比如开始创建了通道号为0的通道，在这个通道还没销毁的情况下，又去创建通道号为0的通道。

处理步骤

针对分析的可能原因，请参考以下方法处理：

- 首先确定hi_mpi_venc_create_chn接口失败时返回的错误码是多少。
- 如果是0xa0088002，则说明用户传入的通道ID超出了规定的合法范围，需要用户修改通道ID在[0,255]内。
- 如果是0xa0088003或0xa0088008，则说明用户传入的通道属性参数不在规定的合理范围内或设置了暂不支持的参数，
具体是哪个参数传入有问题可以进一步查看内核日志打印，如下图所示，是传入的分辨率不正确。

```
[Venc]:venc_drv_check_resolution [Line]:342 max picture width (0) err! should in [128,4096]!
```

常见原因：1.入参的结构体没有进行memset初始化，导致有些参数如果没有主动设置就会是一些随机值；2.头文件不匹配，导致枚举类型传入和预期不符合；3.参数支持范围不了解，各个参数的支持范围可以详细查看DVPP对外接口文档。

- 如果是0xa0088004，则说明用户试图创建已经存在的通道，比如开始创建了通道号为0的通道，在这个通道还没销毁的情况下，又去创建通道号为0的通道

```
[Venc]:venc_create_chn [Line]:2449 device:0 chnl:0 had been created!
```

这种情况建议用户排查下通道号管理是否存在问题。

3.1.3.8 VENC 送入待编码图像帧失败

适用场景

- 业务场景：DVPP
- 适用处理器：昇腾310P AI处理器
- 处理器形态：EP、RC

现象描述

调用VENC发送帧的接口hi_mpi_venc_send_frame返回值非0，发送失败。

可能原因

导致发送帧失败可能原因有以下：

- 用户传入的图像帧参数不在规定的合理范围内或设置了暂不支持的参数。
- 用户送帧的频率太快，大于了性能的规格。

处理步骤

针对分析的可能原因，请参考以下方法处理：

- 如果是0xa0088003或0xa0088008，则说明用户传入的图像帧参数不在规定的合理范围内或设置了暂不支持的参数，
具体是哪个参数传入有问题可以进一步查看内核日志打印，如下图所示，是传入的YUV格式不正确。

```
[Venc]:hevc_check_pixel_format [Line]:1110 H.265 don't support format 5,should be NV12(1) or NV21(2)
```


常见原因：1.入参的结构体没有进行memset初始化，导致有些参数如果没有主动设置就会是一些随机值；2.头文件不匹配，导致枚举类型传入和预期不符合；3.参数支持范围不了解，各个参数的支持范围可以详细查看DVPP对外接口文档。
- 如果是0xa008800d，则说明VENC的输入空闲队列已满，此时无法再继续往内部送入数据帧，这种问题一般是由于用户送帧的频率太快，大于了芯片处理的速度，导致输入队列堆积，VENC的输入空闲队列长度为6帧，只要堆积到了6帧再继续往里送就会出现这个报错。这种情况建议用户控制调用hi_mpi_venc_send_frame的时间间隔，比如编码帧率30fps，调用的间隔可以控制在33ms一帧。

3.1.3.9 VENC 编码无输出

适用场景

- 业务场景：DVPP
适用处理器：昇腾310P AI处理器
- 处理器形态：EP、RC

现象描述

VENC通道创建，发送帧都成功，但调用VENC获取码流的接口hi_mpi_venc_get_stream返回0xa008800e，获取不到编码后码流数据。

可能原因

导致发送帧失败可能原因有以下：

- 用户传入的图像帧内存不是使用dvpp内存申请接口hi_mpi_dvpp_malloc申请的。
- 用户传入的图像帧内存和分辨率不匹配。
- OS内存管理出现问题。

处理步骤

针对分析的可能原因，请参考以下方法处理：

- 查看日志是否有出现看门狗相关的打印，如下图所示，如果出现看门狗，基本可以确认是内存使用存在问题

```
[Chnl]:chnl_watch_dog_timer_isr [Line]:1141 find VEDU_0 down,now reset it  
[Chnl]:chnl_watch_dog_blackbox [Line]:1097 vpu_id is 0, venc watchdog fail enter blackbx
```

- 排查用户代码中输入图像帧的内存申请方式，如果不是使用dvpp内存申请接口hi_mpi_dvpp_malloc申请的，VENC芯片将无法正常工作访问该内存，导致编码无输出，需要修改为使用hi_mpi_dvpp_malloc申请输入内存。
- 排查输入图像帧的内存和分辨率是否不匹配，比如NV12或NV21格式的YUV，一帧图像大小为宽x高x1.5，如果实际送给VENC的内存大小比设置的分辨率宽x高x1.5小的话，会产生访问越界等不可预期的行为，也会导致编码无输出。需要保证申请的内存大小和实际设置的分辨率参数匹配。
- 如果以上排查都不存在问题，可能是OS的内存管理出现问题，您可以通过<https://gitee.com/ascend>网站提交issue获取帮助。

3.1.3.10 模型推理精度下降

适用场景

- 业务场景：DVPP，精度调优
- 适用处理器：昇腾310P AI处理器
- 处理器形态：EP、RC

现象描述

JPEG图片集经过DVPP的JPEG解码模块，解码成yuv420格式，再经过VPC进行缩放处理，再经过AIPP进行色域变换处理成rgb格式，最后经过AI Core进行网络推理（推理网络为Resnet50，yolo3），与下面对比实验进行比对，图像集的整体推理精度存在下降。

对比实验：JPEG图片集经过opencv解码，解码成rgb格式，再经过opencv缩放，最后经过AI Core进行网络推理。

可能原因

可能原因分析：

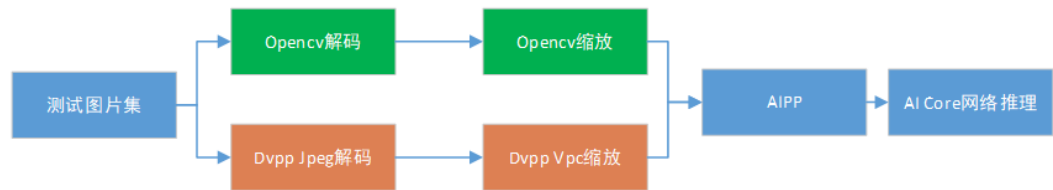
- DVPP的JPEG模块解码算法使用较新的libjpeg-turbo算法（Opencv3.4.4及以上），而已经训练好的模型，一般是使用libjpeg算法（Opencv2.4.9中的解码算法）训练得到的，造成这个模型不是最适配DVPP解码的模型。
- DVPP的VPC缩放算法采用的是自研算法，而已经训练好的模型，一般是使用双线性插值算法（Opencv默认算法）训练得到的，造成这个模型不是最适配DVPP缩放的模型。

定位思路

确定精度下降的原因是否为DVPP解码和缩放导致，具体过程如下：

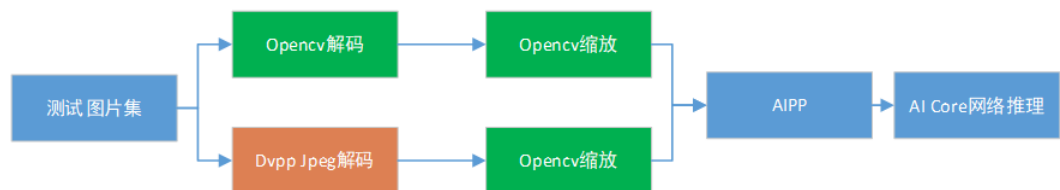
- 步骤1** 进行对比实验，实验过程如图3-2所示，对照组是用Opencv进行解码和缩放，如果与对照组精度相比下降，则认为这是由于DVPP解码和缩放导致。

图 3-2 DVPP 解码与缩放



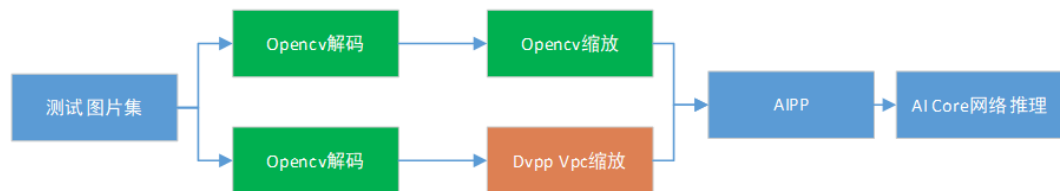
步骤2 进行对比实验，实验过程如图3-3所示，对照组仍然使用Opencv进行解码和缩放，如果与对照组精度相比下降，则认为DVPP解码对精度有影响。

图 3-3 DVPP 解码



步骤3 进行对比实验，实验过程如图3-4所示，对照组仍然使用Opencv进行解码和缩放，如果与对照组精度相比下降，则认为DVPP缩放对精度有影响。

图 3-4 DVPP 缩放



----结束

处理步骤

建议使用“DVPP JPEG解码+VPC缩放+AIPP色域转换”获取RGB图片，再对已有模型进行增量训练或者重新全量训练，得到新的模型用于推理。

您可以通过<https://gitee.com/ascend>网站提交issue获取帮助。

3.1.4 数据预处理常见问题（昇腾 310&昇腾 910）

3.1.4.1 视频解码丢帧/丢包

现象描述

视频解码丢帧，出现重影或不连续现象。查看Device侧日志，发现日志中存在以下几个报错的内容信息中的一个或多个。

- 日志报错信息（1）
[ERROR] DVPP:2020-12-31-23:51:51.339.518 [VDEC][PPSSPSCheckTmpId:7065][T3]
PPSSPSCheckTmpId: pps with this pic_parameter_set_id = 0 haven't decode
[ERROR] DVPP:2020-12-31-23:51:51.339.616 [VDEC][ProcessSliceHeaderFirstPart:7627][T3] PPS or SPS of this slice not valid

[ERROR] DVPP:2020-12-31-23:51:51.339.678 [VDEC][InquiresSceProperty:10582][T3] sliceheader dec err

- 日志报错信息 (2)

[ERROR] DVPP:2020-12-31-20:51:51.318.218 [VDEC][InitPic:6039][T3] line 6039: frame gap(=48) > dpb size(=2)
[ERROR] DVPP:2020-12-31-20:51:51.318.266 [VDEC][H264_DecSlice:8238][T3] init pic err, find next recover point or next valid sps, pps, or exit
[ERROR] DVPP:2020-12-31-20:51:51.318.336 [VDEC][H264_DecOneNal:10077][T3] DecList error, ret=-1
[ERROR] DVPP:2020-12-31-20:51:51.318.392 [VDEC][ReceivePacket:10400][T3] nal_release-err

- 日志报错信息 (3)

[ERROR] DVPP:2020-12-31-20:30:22.188.008 [VDEC][InitListX:4513][T3] for P slice size of list equal 0.
[ERROR] DVPP:2020-12-31-20:30:22.188.056 [VDEC][DecList:4832][T3] line: 4832 InitListX failed
[ERROR] DVPP:2020-12-31-20:30:22.188.128 [VDEC][H264_DecSlice:8260][T3] DecList error, ret=-1
[ERROR] DVPP:2020-12-31-20:30:22.188.199 [VDEC][H264_DecOneNal:10077][T3] Decoder Slice failed

- 日志报错信息 (4)

[ERROR] DVPP:2020-12-31-10:20:28.528.090 [VDEC][InitListX:4653][T3] for B slice size of two list all equal 0.
[ERROR] DVPP:2020-12-31-10:20:28.528.168 [VDEC][DecList:4830][T3] line: 4832 InitListX failed
[ERROR] DVPP:2020-12-31-10:20:28.528.266 [VDEC][H264_DecSlice:8257][T3] DecList error, ret=-1

- 日志报错信息 (5)

[ERROR] DVPP:2020-12-31-10:30:22.130.500 [VDEC][HEVC_VpsSpsPpsCheck:8084][T10] pps with this pic_parameter_set_id = 0 haven't be decoded
[ERROR] DVPP:2020-12-31-10:30:22.130.598 [VDEC][HEVC_DecSliceSegmentHeader:2793][T10] HEVC_VpsSpsPpsCheck != HEVC_DEC_NORMAL
[ERROR] DVPP:2020-12-31-10:30:22.130.686 [VDEC][HEVC_InquireSliceProperty:10169][T10] HEVC_DecSliceSegmentHeader dec err
[ERROR] DVPP:2020-12-31-10:30:22.130.789 [VDEC][HEVCDEC_DecodePacket:753][T10] HEVC_InquireSliceProperty error.

- 日志报错现象 (6)

[ERROR] DVPP:2020-12-31-11:22:28.800.158 [VDEC][HEVC_RefPicProcess:2480][T10] Ref frame(poc 15) lost.
[ERROR] DVPP:2020-12-31-11:22:28.800.236 [VDEC][HEVC_CreateLostPicture:6392][T10] DPB no suited fs for lost pic.
[ERROR] DVPP:2020-12-31-11:22:28.800.352 [VDEC][HEVC_RefPicProcess:2480][T10] Ref frame(poc 18) lost.
[ERROR] DVPP:2020-12-31-11:22:28.800.426 [VDEC][HEVC_CreateLostPicture:6392] [T10] DPB no suited fs for lost pic.
[ERROR] DVPP:2020-12-31-11:22:28.800.522 [VDEC][HEVC_RefPicProcess:2480] [T10] Ref frame(poc 18) lost.

- 日志报错信息 (7)

[ERROR] DVPP:2020-12-31-11:56:35.038.109 [VDEC][HEVC_RefPicProcess:2480] [T56] Ref frame(poc 15) lost.
[ERROR] DVPP:2020-12-31-11:56:35.038.283 [VDEC][HEVC_CreateLostPicture:6392] [T56] Take poc(17) to create lost poc(15).
[ERROR] DVPP:2020-12-31-11:56:35.038.502 [VDEC][FSP_SetRef:934] [T56] check condition: pstLogicFs->IsDummyFs == 0 fail
[ERROR] DVPP:2020-12-31-11:56:35.038.801 [VDEC][FSP_SetRef:934] [T56] check condition: pstLogicFs->IsDummyFs == 0 fail
[ERROR] DVPP:2020-12-31-11:56:35.039.128 [VDEC][FSP_SetRef:934] [T56] check condition: pstLogicFs->IsDummyFs == 0 fail

可能原因

分析上述日志报错信息现象，分别可能存在以下可能原因：

- 日志报错信息 (1) 可能原因：H264码流缺少IDR帧
- 日志报错信息 (2) 可能原因：H264码流缺少I帧
- 日志报错信息 (3) 可能原因：H264码流缺少P帧
- 日志报错信息 (4) 可能原因：H264码流缺少B帧

- 日志报错信息（5）可能原因：H265码流缺少IDR帧
- 日志报错信息（6）可能原因：H265码流缺少P帧
- 日志报错信息（7）可能原因：H265码流缺少I帧或者B帧

处理步骤

针对可能原因分析，参考以下步骤处理：

步骤1 检查输入的源码流是否有问题。

使用第三方工具（如：eseye u）对输入码流进行检查，查看码流是否异常。

步骤2 若查看的源码流结果为正常，则可能码流在传输给设备侧VDEC的过程中遭到破坏，需要在调用aclvdecSendFrame之前，通过fwrite函数将输送给VDEC的码流保存下来。

- 使用第三方工具对保存的码流进行检查，如果码流异常，用户需自行排查将码流从Host侧传输到Device侧的过程代码是否有导致码流异常的情况。
- 通过对应版本的sample，使用命令行方式编译和部署，解码这段保留下来的码流，验证码流是否正常或VDEC是否支持该格式。

如果sample解码正常，那就是开发代码有问题，可以参考《[应用软件开发指南（C&C++）](#)》中的VDEC示例代码，找到对应的视频解码的代码参考优化。

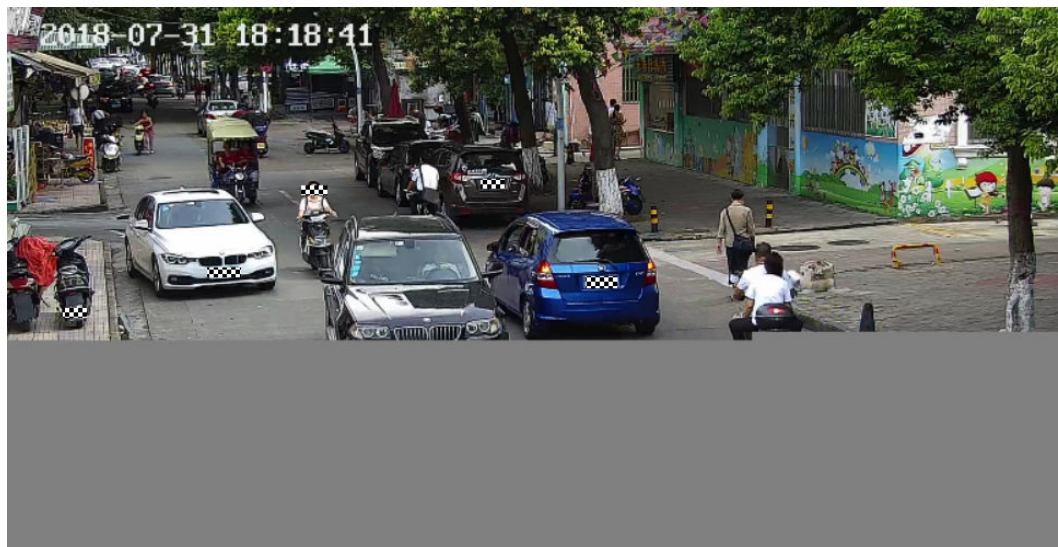
----结束

3.1.4.2 视频解码花屏

现象描述

输入码流给VDEC进行解码，得到的解码数据不正确，产生花屏现象，如[图3-5](#)所示。并且日志中存在类似“dvpp_vdec_vdm_process failed”、“Chan 0 ErrRatio = 44”信息。

图 3-5 视频花屏



可能原因

输入的码流中某些帧数据不完整、存在坏帧，导致硬件解码产生花屏。

处理步骤

针对可能原因分析，参考以下步骤处理：

步骤1 检查输入的源码流是否有问题。

使用第三方工具（如：eseye u等）对输入码流进行解码播放，查看是否存在花屏，若不花屏则进行[步骤2](#)；若花屏则替换输入码流。

步骤2 若查看的源码流结果为正常，则可能码流在传输给设备侧VDEC的过程中遭到破坏，需要在调用acldvdecSendFrame之前，通过fwrite函数将输送给VDEC的码流保存下来。

- 使用第三方工具对保存的码流进行解码播放，查看是否存在花屏。如果存在花屏，用户需自行排查将码流从Host侧传输到Device侧的过程代码是否有导致码流异常的情况。
- 通过对应版本的sample，使用命令行方式编译和部署，解码这段保留下来的码流，查看是否存在花屏。

如果sample解码正常，那就是开发代码有问题，可以参考《[应用软件开发指南（C&C++）](#)》中的VDEC示例代码，找到对应的视频解码的代码参考优化。

----结束

3.1.4.3 视频解码性能问题

现象描述

VDEC解码性能下降，小于VDEC对外发布的性能规格，导致发生卡顿等现象。

可能原因

分析VDEC性能下降导致的卡顿故障，可能原因有：

- 视频解码回调函数中耗时过长，影响解码性能。
- 输入码流中I帧比例过大，解码I帧耗时比解码P帧耗时长，影响解码性能。
- 输入码流中存在异常帧，影响解码性能。

处理步骤

针对上述可能的故障原因，可以参考以下方式进行处理：

步骤1 在回调函数中打点测试耗时，查看耗时是否过长(回调函数允许的最大耗时和帧率相关，计算公式为：最大耗时=1/帧率，例如帧率=30fps，则最大耗时=1/（30fps）=0.033s)。

步骤2 使用第三方工具打开输入码流，查看I帧比例是否过大。一般GOP值为30(即I帧间隔为30)，如果I帧比例过大，则需要替换成正常码流进行性能测试。

步骤3 使用第三方工具打开输入码流，查看是否存在异常帧（如用第三方工具打开显示花屏或解码报错），如果存在异常帧，会造成不满足规格的现象。

----结束

3.1.4.4 视频解码失败

现象描述

视频解码失败，日志中打印如下所示的异常信息。

- 日志信息：Invalid width or height
[ERROR] DVPP:2020-12-31-20:50:51.318.218 [VDEC][IsValidResolution:639][T56] Invalid width or height, valid range (w:128~4096) (h:128~4096), current width = 32, height = 32, realWidth = 18, realHeight = 18
[ERROR] DVPP:2020-12-31-20:50:51.318.365 [VDEC][upgrade_picture_info_in_detail:679][T56] check condition: ret == OMX_ErrorNone fail
[ERROR] DVPP:2020-12-31-20:50:51.318.429 [VDEC][EventHandLer:386][T56] Dynamic Resouces Unavailable now
[ERROR] DVPP:2020-12-31-20:50:51.368.106 [VDEC][hanlde_release_instance:1352][T56] wait Componment Exit Message Thread
- 日志信息：bit_depth_luma(*) not epual *
[ERROR] DVPP(13757,dvpp_performance):2020-12-31-21:18:26.998.657 [VDEC][ProcessSPS:9055][T26] bit_depth_luma(10) not epual 8.
[ERROR] DVPP(13757,dvpp_performance):2020-12-31-21:18:26.998.763 [VDEC][ProcessSPS:9070][T26] bit_depth_chroma(10) not epual 8.

可能原因

针对上述日志分析，可能存在输入码流规格不符合以下要求：

- VDEC支持的码流宽高分别是width：128~4096，height：128~4096。
- VDEC支持8bit的h264码流，8bit或10bit的h265码流。

定位思路

- 1、排查输入的码流宽高是否满足VDEC要求。
- 2、排查输入的码流类型是否满足VDEC要求。

处理步骤

检查输入给VDEC的码流是否满足要求：

使用第三方工具（如：eseye u等）对输入码流进行检查，查看码流类型、码流宽高等信息是否符合要求。如果码流不满足要求，请替换符合要求的码流。

3.1.4.5 视频帧解码失败不触发回调函数

现象描述

用户输入码流给VDEC解码，某些帧或所有帧都没有触发回调函数，用户收不到解码结果。

可能原因

码流中某些帧是坏帧，导致语法解析不出这些帧的含义，或者解码这些帧失败，从而不调用回调函数。

处理步骤

针对分析的可能原因，请参考以下步骤进行处理：

步骤1 查看日志中是否有**视频解码丢帧/丢包**中的日志报错信息，若有，则是因为异常帧解码失败导致没有回调。

步骤2 若没有**视频解码丢帧/丢包**中的日志报错信息，则调整DVPP模块或全局日志级别为info，查看下述三条日志打印的总次数是否和输入的帧数相等。

- (1) "The queue is empty, so call the non-intelligent pointer callback interface."
- (2) "The queue is not empty, so call the smart pointer callback interface."
- (3) "The queue is not empty, but hiai_data_sp is nullptr."

生成上述3种日志信息的场景如下：

- 未使用hiai_data_sp，成功解码返回，打印（ 1 ）日志，然后调用用户注册的回调函数。
- 每一帧对应设置一个hiai_data_sp，成功解码返回，打印（ 2 ）日志，然后调用用户注册的回调函数。
- N帧对应设置一个hiai_data_sp，第1帧成功解码返回，打印（ 2 ）日志，然后调用用户注册的回调函数；其他N-1帧，成功解码后打印（ 3 ）日志信息，并调用用户注册的回调函数。

不论上述场景中哪一种，均会调用用户注册的回调函数，即只要调用一次用户回调函数就说明解码返回一帧。所以上述三条日志出现的次数总和与用户输入总帧数相等，则说明无解码丢帧，用户需排查自身接收解码结果的统计是否有误。

----结束

3.1.4.6 JPEG 图片解码失败

现象描述

JPEGD模块解码失败，查看日志有类似如下报错信息：

日志信息（ 1 ）：

```
just support jpeg with YUV 444 422 420 400
do not support progressive mode
do not support arithmetic code, support huffman code only
```

日志信息（ 2 ）：

```
EOI segment of the stream is invalid
```

可能原因

分析上面日志信息，可能存在以下可能原因：

1. 数据格式不支持
JPEGD只支持huffman编码(colorsapce: yuv, subsample: 444/422/420/400)，不支持算术编码，不支持渐进编码，不支持jpeg2000格式。
2. 图像数据不完整

处理步骤

针对上述可能原因，请按以下方式处理：

- 步骤1** 针对目前不支持的超规格图像格式，建议用户自行使用第三方软件解码。
- 步骤2** 针对图像数据不完整，根据报错提示，通过第三方软件查看原图像二进制进行确认。

例如“EOI segment of the stream is invalid”或“EOI segment of the stream is invalid, it should be FFD9. Try software decoding.”报错，表示图像缺失最后的EOI结束符，对应图像二进制类似下图所示。正常JPEG图片最后应该由标记码FF D9结束，该数据最后缺失FF D9标记码。

如果确认原图数据不完整，报错属于正常现象，需更换数据。

```
00005e90h: AB 79 F4 E3 4B CE 38 9C B8 B7 D2 B2 1C 59 4E A5 ; 垓器K?涅分?YN?  
00005ea0h: 5F 73 B6 36 D6 FF 00 E5 7F FA BF F7 71 A2 47 FC ; s???.? 麟 ?  
00005eb0h: 91 FC D3 FC 30 4C 6A 84 EB 8E AF 5B 8B 2A 56 BE ; 廖狱0Lj熬喝[?V?  
00005ec0h: 49 FA 60 D5 6D F7 67 3F 11 F9 4A D6 E2 CA 35 2A ; I鸚譚鎗?.鴉肘?*  
00005ed0h: C0 13 FA B6 FD B0 0A 9D C7 FD BF F8 60 CD 43 E4 ; ? .瀝 駛卦?  
00005ee0h: 83 F7 47 FF 00 B6 DE 31 80 2C BE E0 60 01 A7 93 ; 淦G .撥1€,距`.  
00005ef0h: FB 23 DF ED 8C FE 21 DF F4 7F D4 1F D5 8D 0D 7F ; ?唔岨!嘩?魯.0  
00005f00h: 33 F9 9F E3 8C F1 8C 7F 00 00 ; 3弱趁駛..
```

- 步骤3** 如果原图像数据完整，可能数据在传输过程中存在损坏，需要在调用acldvppJpegDecodeAsync之前，通过fwrite函数将传输给JPEGD的码流保存下来，与原图进行二进制比较。如果不一致，传输过程出现数据缺失，需自行进一步定位传输过程数据缺失问题。

----结束

3.1.4.7 VPC 调用失败

现象描述

VPC模块调用失败，查看日志有类似如下报错信息：

日志信息（1）：

```
RoiNum(0), inputArea rightOffset is 1918, it should be odd!
```

日志信息（2）：

```
Output bufferSize(65536) should not be smaller than widthStride(256) * heightStride(256) * 3 / 2 = 98304
```

日志信息（3）：

```
Input widthStride(300) is not right, it should be 16 aligned!  
Input widthStride(16) is not right, it should not be smaller than 32!
```

日志信息（4）：

```
bareDataAddr(0xaaaaecccdd0), bareDataBufferSize(3133440) should be allocated by acldvppMalloc!
```

日志信息（5）：

```
Both RoiNum(1) outputAddr(0xaaaaecccdd0) and first roi outputAddr(0xffff00002000) should be  
allocated by acldvppMalloc!
```

日志信息（6）：

```
RoiNum(0): inputConfigure cropArea, leftOffset(26) should be smaller than rightOffset(25)!  
RoiNum(0): inputConfigure cropArea, upOffset(80) should be smaller than downOffset(79)!  
RoiNum(0): inputConfigure cropArea, cropWidth(1931) should not be bigger than width(1920)!  
RoiNum(0): inputConfigure cropArea, cropHeight(1270) should not be bigger than height(1088)!  
RoiNum(0): inputConfigure cropArea, cropWidth(9) should be between [10, 8192]!  
RoiNum(0): inputConfigure cropArea, cropHeight(4) should be between [6, 8192]!
```

```
RoiNum(0): inputConfigure cropArea, rightOffset(1921) should be smaller than width(1920)!  
RoiNum(0): inputConfigure cropArea, downOffset(1089) should be smaller than height(1088)!
```

日志信息（7）：

```
RoiNum(0): scale must be in [1/32, 16], cropWidth(1920), pasteWidth(10)!  
RoiNum(0): scale must be in [1/32, 16], cropHeight(6), pasteHeight(100)!
```

可能原因

针对上面日志信息分析，可能存在以下对应原因：

- 日志信息（1）：VPC抠图区域右偏移坐标需是奇数，日志信息里1918是偶数，不符合要求。
- 日志信息（2）：当输出内存大小应该大于等于宽stride*高stride*3/2，日志信息显示不满足这个条件。
- 日志信息（3）：输入图片的宽stride（即每行图像占用的内存大小）必须是16倍数、且最小值32。日志信息里宽stride是300，不满足16倍数的要求，需要将图像做对齐后，并将宽stride设置成对齐后的值。
- 日志信息（4）：VPC的输入内存需要使用acldvppMalloc接口申请。
- 日志信息（5）：VPC的输出内存需要使用acldvppMalloc接口申请。
- 日志信息（6）：VPC的抠图区域不符合约束要求，因此报错。
- 日志信息（7）：VPC的缩放范围为[1/32, 16]。日志信息提示了缩放范围，并且显示了抠图宽为1920，输出区域的宽为10，通过计算可以得到： $10/1920 < 1/32$ ，因此报错。

定位思路

1. 根据日志描述的错误信息，找到VPC对应的配置参数，根据提示进行修改。
2. 根据日志描述的错误信息，参考《[应用软件开发指南（C&C++）](#)》中VPC参数的约束修改。

处理步骤

根据提示的错误信息进行修改：

- 步骤1** 如果为日志信息（1），说明输入图片抠图区域的右偏移错误，应该设置为奇数。
- 步骤2** 如果为日志信息（2），说明输入内存的大小不正确，应该检查申请输入内存的代码，申请内存大小应该为 $1920*1088*3/2$ ，并且bareDataBufferSize这个值也要填写为 $1920*1088*3/2$ 。
- 步骤3** 如果为日志信息（3），说明输入图片的stride值不符合要求，需设置为16的倍数。
- 步骤4** 如果为日志信息（4）和日志信息（5），代码中申请内存时，需要使用acldvppMalloc接口申请。
- 步骤5** 如果为日志信息（6），需要修改抠图宽度。
- 步骤6** 如果为日志信息（7），需要修改缩放范围。

----结束

3.1.4.8 retCode 返回值设置错误，导致 VDEC 视频解码异常

现象描述

调用acldvdecSendFrame接口发送一帧码流后，继续复用output输出图片描述信息，进行后续帧码流的解码操作，结果反复出现解码不成功、解码异常的情况。

日志片段举例如下：

```
Channel[0]: success to acldvdecSendFrame, loop=1, count=7
get frame success, totalCount=7
packet.size is 26084.
Channel[0]: begin to send frame, loop=1, count=8
acldvppGetPicDescRetCode, retCode: 2.
Vdec ERROR!!!!!!!!!!!!!!
errCount is 3. total count is 3.
!!!!!!!!!!!!!!acldvppGetPicDescRetCode, retCode: 2.right_count:0,fail_count:3,total_count:3
Channel[0]: success to acldvdecSendFrame, loop=1, count=8
get frame success, totalCount=8
packet.size is 27927.
Channel[0]: begin to send frame, loop=1, count=9
acldvppGetPicDescRetCode, retCode: 2.
Vdec ERROR!!!!!!!!!!!!!!
errCount is 4. total count is 4.
!!!!!!!!!!!!!!acldvppGetPicDescRetCode, retCode: 2.right_count:0,fail_count:4,total_count:4
```

可能原因

根据日志中的提示，通过acldvppGetPicDescRetCode接口获取到的retCode为2，retCode为非0值时，表示解码异常。

再查看代码逻辑时，发现由于前一帧码流解码失败，retCode被置为2，在复用output输出图片描述信息时，retCode也继承了前一帧解码失败的状态值2，导致AscendCL在解码后续帧时，获取到retCode值为2，就一直判断解码是失败。

处理步骤

如果存在复用output输出图片描述信息的场景，需先调用acldvppSetPicDescRetCode设置为0，防止前一帧解码异常的状态影响后续解码。

3.1.4.9 模型推理精度下降

现象描述

JPEG图片集经过DVPP的JPEG解码模块，解码成yuv420格式，再经过VPC进行缩放处理，再经过AIPP进行色域变换处理成rgb格式，最后经过AI Core进行网络推理（推理网络为Resnet50，yolo3），与下面对比实验进行比对，图像集的整体推理精度存在下降。

对比实验：JPEG图片集经过opencv解码，解码成rgb格式，再经过opencv缩放，最后经过AI Core进行网络推理。

可能原因

可能原因分析：

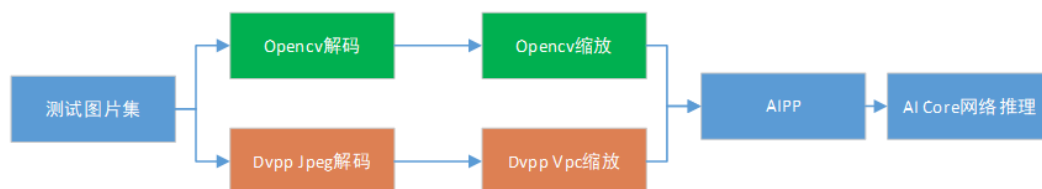
- DVPP的JPEG模块解码算法使用较新的libjpeg-turbo算法（Opencv3.4.4及以上），而已经训练好的模型，一般是使用libjpeg算法（Opencv2.4.9中的解码算法）训练得到的，造成这个模型不是最适配DVPP解码的模型。
- DVPP的VPC缩放算法采用的是自研算法，而已经训练好的模型，一般是使用双线性插值算法（Opencv默认算法）训练得到的，造成这个模型不是最适配DVPP缩放的模型。

定位思路

确定精度下降的原因是否为DVPP解码和缩放导致，具体过程如下：

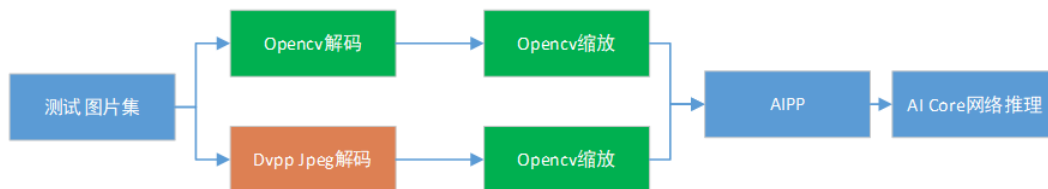
步骤1 进行对比实验，实验过程如图3-6所示，对照组是用Opencv进行解码和缩放，如果与对照组精度相比下降，则认为这是由于DVPP解码和缩放导致。

图 3-6 DVPP 解码与缩放



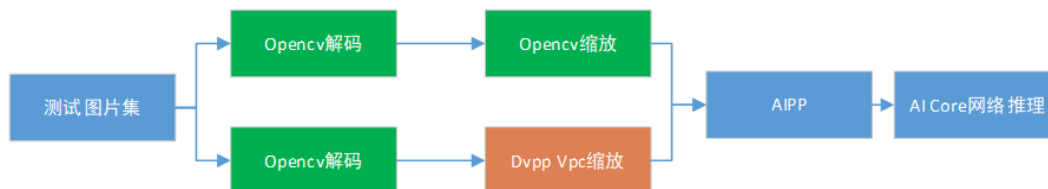
步骤2 进行对比实验，实验过程如图3-7所示，对照组仍然使用Opencv进行解码和缩放，如果与对照组精度相比下降，则认为DVPP解码对精度有影响。

图 3-7 DVPP 解码



步骤3 进行对比实验，实验过程如图3-8所示，对照组仍然使用Opencv进行解码和缩放，如果与对照组精度相比下降，则认为DVPP缩放对精度有影响。

图 3-8 DVPP 缩放



----结束

处理步骤

建议使用“DVPP JPEG解码+VPC缩放+AIPP色域转换”获取RGB图片，再对已有模型进行增量训练或者重新全量训练，得到新的模型用于推理。具体方法请咨询华为工程师。

3.2 GE 常见故障

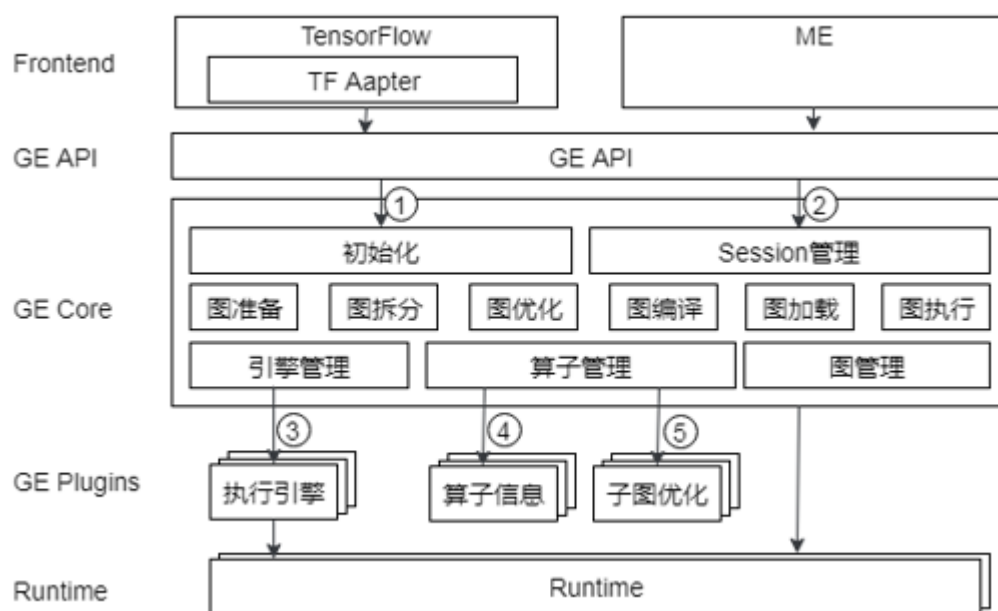
3.2.1 背景知识

GE模块从逻辑结构上划分，包括下面3个部分：

- GE API对外呈现GE Core中初始化、Session管理模块的接口，支持运行环境初始化，Session创建、销毁，图添加执行。
- GE Core从上往下，共分为三层，分别是执行控制层、业务功能层、数据管理层。
- GE Plugin：统一定义各插件模块需要提供的接口，具体的插件由不同的功能模块根据自己的能力提供，GE会根据不同的插件提供的能力，实现最优的执行选择。

GE逻辑框架如图3-9所示。

图 3-9 GE 逻辑框架



3.2.2 常见故障分析与处理

3.2.2.1 算子插件未注册报错

适用场景

- 业务场景：昇腾模型算法，模型转化、应用/算子开发、TensorFlow训练
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

查看日志，存在报错某个算子类型不支持：

```
Check op[%s]'s type[%s] failed, it is not supported.
```

或者

进行模型转换的时候，某个算子类型转换不符合预期，被转换成了frameworkop类型。

可能原因

根据日志分析，可能存在以下可能原因：

- 算子插件so未加载成功。
- 算子未注册映射关系，或者未编译到算子的插件so中。

解决措施

针对分析可能的故障原因，可以参考下面步骤处理：

步骤1 确认算子插件so是否加载成功。

1. 算子原型so加载成功打印类似信息：

```
plugin load /usr/local/Ascend/opp/built-in/framework/onnx/libops_all_onnx_plugin.so success.
```

2. 加载失败的告警关键信息：

```
dlopen failed, plugin name:%s. Message(%s).
```

步骤2 如果算子插件so加载成功，则需要继续确认算子注册的映射关系是否编译进加载的插件so中了。

使用nm命令查看so符号表，如果没有注册，则需要注册该算子插件，可以参考《[TBE&AI CPU算子开发指南](#)》的“算子适配”章节内容实现。

说明

nm -D命令可查看so文件符号表。

步骤3 如果算子插件so未加载成功，参考失败告警中Message提示内容处理。

----结束

3.2.2.2 算子原型未注册报错

适用场景

- 业务场景：昇腾模型算法，模型转化、应用/算子开发、TensorFlow训练
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

查看日志，存在报错某个算子没有原型定义：

```
op[%s] type[%s] have no ir factory.
```

或者

```
IR for op[%s] otype[%s] is not registered.
```

说明

op[%s] type[%s]中的%s分别表示具体的算子名称和算子类型。

可能原因

根据日志分析，可能存在以下可能原因：

- 算子原型so未加载成功。
- 算子未定义注册该类型算子，并编译到算子的原型so中。

解决措施

针对分析可能的故障原因，可以参考下面步骤处理：

步骤1 确认算子原型so是否加载成功。

1. 算子原型so加载成功打印类似信息：

```
OpsProtoManager plugin load /usr/local/Ascend/opp/built-in/op_proto/
libopsproto.so success.
```

2. 加载失败的告警关键信息：

```
OpsProtoManager dlopen failed, plugin name:%s. Message(%s).
```

步骤2 如果算子原型so加载成功，需要确认算子原型定义是否编译进加载的so中了。

使用nm查看so符号表，如果没有注册，则需要注册该算子原型，可以参考《[TBE&AI CPU算子开发指南](#)》的“算子原型定义”章节内容实现。

说明

nm -D命令可查看so文件符号表。

步骤3 如果算子原型so未加载成功，参考失败告警中Message提示内容处理。

----结束

3.2.2.3 动态 shape 模型用户输入和模型推导结果不匹配

适用场景

- 业务场景：昇腾模型算法，模型转化、应用/算子开发、TensorFlow训练
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

场景1：模型执行报错，日志信息中包含以下关键信息：“[Check][Size] %s(%s) index[%d] mem size out of range! Expected size: %ld, but given input size: %ld.”

```
[INFO] GE(2284607,ir_build):2021-11-08-19:31:33.920.888 [task_context.cc:257]2284607 AllocateOutput:To allocate output for node: add2, index = 0, tensor desc = [TensorDesc] DataType = 6, Format = 2, Shape = [50, ]  
[ERROR] GE(2284607,ir_build):2021-11-08-19:31:33.920.900 [task_context.cc:277]2284607 AllocateOutput: ErrorNo: 50331649() [LOAD][LOAD][Check]  
[Size] add2(Add) index[0] mem size out of range! Expected size: 128, but given input size: 2.  
[ERROR] GE(2284607,ir_build):2021-11-08-19:31:33.920.951 [aicpu_node_executor.cc:200]2284607 UpdateArgs: ErrorNo: 4294967295(failed) [LOAD][LOAD][Update][IoAddr] failed for Node[add2].  
[ERROR] GE(2284607,ir_build):2021-11-08-19:31:33.920.960 [execution_engine.cc:386]2284607 DoExecuteAsync: ErrorNo: 4294967295(failed) [LOAD][LOAD][Prepare][Task] For [add2(Add)] failed.
```

场景2：单算子场景执行报错，日志信息中包含以下关键信息：“[Check] [Param:outputs]Output size mismatch. index = %zu, model expect %zu, but given %zu(after align)”

```
[INFO] GE(130191,python3):2021-09-27-17:23:17.922.047 [single_op.cc:161]131511 ValidateArgsInput [1], aligned_size:96, inputs.length:40, input_size:96  
[INFO] GE(130191,python3):2021-09-27-17:23:17.922.054 [single_op.cc:186]131511 ValidateArgsOutput [0], aligned_size:3145760, outputs.length:3145760, output_size:18446744073709551615  
[ERROR] GE(130191,python3):2021-09-27-17:23:17.922.095 [jangle.cc:190]131511 ValidateArgs: ErrorNo: 145000[Parameter invalid] [FINAL][FINAL][Check][Param:outputs]Output size mismatch. index = 0, model expect 18446744073709551615, but  
[ERROR] ASCENDCL(130191,python3):2021-09-27-17:23:17.922.213 [jop_executor.cpp:58]131511 DoExecuteAsync: [FINAL][FINAL][Exec][Op]execute op failed, ge result = 145000  
[INFO] ASCENDCL(130191,python3):2021-09-27-17:23:17.922.280 [acl.cpp:242]131511 aclGetRecenterMsg: start to execute aclGetRecenterMsg.  
[INFO] ASCENDCL(130191,python3):2021-09-27-17:23:17.922.308 [acl.cpp:249]131511 aclGetRecenterMsg: execute aclGetRecenterMsg successfully.
```

可能原因

场景1：

此处为模型执行时，校验用户分配的output_buffer大小和模型经过一层层inershape后得到的模型输出的output_buffer大小不匹配。

场景2：

此处为单算子场景执行时，校验用户分配的output_buffer大小和单算子编译时inershape推出的output_size大小不匹配。

解决措施

针对分析的故障可能原因，可以参考下面步骤处理：

针对场景1：

若发生该错误，用户需要检查分配的output_buffer大小是否正确，应该与模型推导的大小保持一致。建议用户分配的output_buffer按照上报的ERROR中提示的大小进行分配。

针对场景2：

若发生该错误，用户需检查分配的output_buffer大小和算子inershape提示的size大小是否一致，建议按照报错提示的大小进行分配。

3.2.2.4 动态 shape 模型输入大小校验失败

适用场景

- 业务场景：昇腾模型算法，模型转化、应用/算子开发、TensorFlow训练
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

动态shape模型输入大小校验失败，日志信息中包含以下关键信息：tensor size mismatches. expected: ..., but given ...

```
plog-187626_20211011143937175.log:17480 [ERROR] GE(187626,python3):2021-10-11-14:39:37.270.942 [execution_engine.cc:450]191154 ValidateInputTensors: ErrorNo: 1343225860(Internal errors) [EXEC][EXEC][Check][Size]  
for [PartitionedCall_14[PartitionedCall]] Input[40] [tensor size mismatches. expected: 768032, but given 32]  
plog-187626_20211011143937175.log:17481 [ERROR] GE(187626,python3):2021-10-11-14:39:37.270.987 [execution_engine.cc:390]191154 DoExecuteAsync: ErrorNo: 4294967295(failed) [EXEC][EXEC][Validate][InputTensors] f  
or PartitionedCall_14[PartitionedCall] failed.  
plog-187626_20211011143937175.log:17482 [ERROR] GE(187626,python3):2021-10-11-14:39:37.270.992 [subgraph_executor.cc:470]191154 LaunchTasks: ErrorNo: 1343225860(Internal errors) [EXEC][EXEC][Invoke][ExecuteAs  
ync] failed for [PartitionedCall_14[PartitionedCall]]  
plog-187626_20211011143937175.log:17483 [ERROR] GE(187626,python3):2021-10-11-14:39:37.270.996 [subgraph_context.cc:166]191154 OnError: ErrorNo: 1343225860(Internal errors) [EXEC][EXEC][Check][Param:error][Ro  
o-Graph] Error:1343225860 occurred while executing graph.
```

可能原因

动态shape场景每个node都会在执行时对shape进行推导，该故障现象可能为算子的shape校验不合法。

解决措施

针对分析的故障可能原因，应根据图上的连接关系，对shape的来源进行排查，找出不合理的shape推导：

可搜索执行plog日志查看关键词：

"before_infershape when running"：显示算子shape推导前的输入、输出shape等信息。

"after_infershape when running"：显示算子shape推导后的输入、输出shape等信息。

从报错节点的输入、输出shape开始进行排查，检查当前节点shape推导结果是否正确（即判定根据输入shape推出的输出shape是否符合预期），如果是输入shape存在问题，则按照相同方法继续排查输入节点的shape推导。

```
DEBUG] GE(2284607,if_build):2021-11-08 19:31:33.920.683 [shape_refiner.cc:556]2284751 PrintInOutTensorShape:Shape dump [after_infershape when running], Node name: [add2], {input_0 tensor: [shape:[50]],(format:ND),(dtype:DT_INT16),(origin_shape:[50]),(origin_format:ND),(origin_dtype:DT_INT16),(shape_range:[(1,50)],(1,50)])} {input_1 tensor: [(shape:[50]),(format:ND),(dtype:DT_INT16),(origin_shape:[50]),(origin_format:ND),(origin_dtype:DT_INT16),(shape_range:[(1,50)],(1,50)])} {output_0 tensor: [(shape:[50]),(format:ND),(dtype:DT_INT16),(origin_shape:[50]),(origin_format:ND),(origin_dtype:DT_INT16),(shape_range:[(1,50)],(1,50)])}]
```

3.3 Driver 常见故障

3.3.1 背景知识

驱动作为软件与硬件的桥梁接口，主要实现如下功能：

- 提供接口，供Runtime调用。
- Device设备的发现、初始化、相关资源的管理。
- 向Task Scheduler发送CMD。
- 从Task Scheduler接收Task Report。
- 向其他模块提供Device设备的状态信息。
- 向其他模块提供接口对Device设备进行控制。

3.3.2 常见故障分析与处理

3.3.2.1 内核版本差异导致驱动子包安装失败

适用场景

- 业务场景：环境安装部署
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

安装run包时，如果当前OS操作系统的内核版本与run包中驱动要求的内核版本不一致时，可能会出现驱动安装失败，如图3-10所示。

图 3-10 日志信息：Drv_dkms_env_check failed, details in : ****

```
root@localhost:~# ./Ascend310-driver-1.75.t25.0.b250-centos7.6.x86_64.run --full
Verifying archive integrity... 100% All good.
Uncompressing ASCEND310 DRIVER RUN PACKAGE 100%
Start time: 2020-09-27 16:26:56
LogFile: /var/log/ascend_seclog/ascend_install.log
OperationLogFile: /var/log/ascend_seclog/operation.log
set username and usergroup, HwHiAiUser:HwHiAiUser
Drv dkms env check failed, details in : /var/log/ascend_seclog/ascend_install.log
Install driver failed, please retry after uninstall and reboot!
End time: 2020-09-27 16:26:58
```

可能原因

- 当前OS环境与run包要求环境不符。
- 源码编译驱动所需DKMS等工具未安装。

处理步骤

步骤1 使用**uname -rm**命令查看当前环境内核版本和架构，检查是否与安装run包资料里的版本要求一致。

如果内核版本不匹配，请准备正确的环境后再安装。或参考安装run包资料，进行驱动的源码编译后安装。

步骤2 或者，进一步检查DKMS等工具是否安装。

- Ubuntu操作系统，使用下面命令检查：
dpkg-query -s dkms
dpkg-query -s gcc
dpkg-query -s linux-headers-\$(uname -r)
- CentOS/EulerOS操作系统，使用下面命令检查：
rpm -qa | grep dkms
rpm -qa | grep gcc
rpm -qa | grep kernel-headers-\$(uname -r)
rpm -qa | grep kernel-devel-\$(uname -r)

如果有工具未安装，请先完成安装后再安装run包。

- Ubuntu操作系统使用命令：apt-get install 工具名
- Centos/EulerOS操作系统使用命令：yum install 工具名

----结束

3.3.2.2 PCI 总线上没检测到昇腾 AI 处理器

适用场景

- 业务场景：环境安装部署

- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

安装run包时，提示固件升级失败，并且日志中打印信息显示未找到昇腾AI处理器，如下所示。

```
Verifying archive integrity... 100% SHA256 checksums are OK. All good.
Uncompressing ASCEND910 FIRMWARE RUN PACKAGE 100%
[Firmware] [2021-06-23 19:27:30] [INFO]Start time: 2021-06-23 19:27:30
[Firmware] [2021-06-23 19:27:30] [INFO]LogFile: /var/log/ascend_seclog/ascend_install.log
[Firmware] [2021-06-23 19:27:30] [INFO]OperationLogFile: /var/log/ascend_seclog/operation.log
[Firmware] [2021-06-23 19:27:31] [INFO]base version is 1.78.T23.0.B230.
[Firmware] [2021-06-23 19:27:31] [WARNING]Do not power off or restart the system during the
installation/upgrade
[Firmware] [2021-06-23 19:27:31] [ERROR]Pcie device is missing, cold reset may fix it.
[Firmware] [2021-06-23 19:27:31] [INFO]End time: 2021-06-23 19:27:31
```

可能原因

服务器上昇腾处理器的标卡未插紧、接触不良或通风散热不好。

定位思路

如图3-11所示，通过指令`lspci | grep d100`查看当前环境PCI总线上是否检测到昇腾处理器，如果没有检测到昇腾处理器，则确认昇腾处理器的标卡状态异常，请按[处理步骤](#)检查。

图 3-11 检查 PCIE 设备

```
root@huawei-HP-ProDesk-600-G3-PCI-MT:~/xz/vmp# lspci | grep d100
root@huawei-HP-ProDesk-600-G3-PCI-MT:~/xz/vmp#
```

处理步骤

针对分析的可能原因，可以参考以下方式处理：

- 插紧PCIE标卡或更换插槽，解决未插紧或接触不良的情况。
- PCIE标卡非自主散热，如果温度过高，可增加外部散热。

3.3.2.3 调用 Device 失败

适用场景

- 业务场景：DVPP，昇腾模型算法，模型转化，精度调优、应用/算子开发，性能调优、TensorFlow训练
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

调用Device失败（容器内），日志中报以下类似信息：

日志信息（1）：日志中有打印ERROR信息“drvDeviceOpen”失败字样，如下所示。


```
[ascend][ERROR][2021-06-26-17:53:15:689064][curpid: 32707, 32707][devdrv][drvDeviceOpen 79]get phys failed, devid(0), phy_devid(0)
```

日志信息（2）：查看内核日志(dmesg)中有打印“device is in used”字样。如下所示。

```
[52783.010878] [ascend] [ERROR] [devdrv] [devdrv_manager_container_table_devlist_add_ns 832] <drv_hlt_dsmt_t:5615> device is in used
[52783.010881] [ascend] [ERROR] [tsdrv] [devdrv_open 194] <drv_hlt_dsmt_t:5615> add to list failed. dev_id(0)
[52783.013546] [ascend] [devdrv] [devdrv_manager_container_get_devnum 1306] <drv_hlt_dsmt_t:5615> weird device number, dev_num = 0
[52783.013547] [ascend] [ERROR] [devdrv] [devdrv_manager_container_get_devlist_ns 1210] <drv_hlt_dsmt_t:5615> some devices are used by other docker, mnt_ns = 0xffff8d33a5963720
[52783.013548] [ascend] [ERROR] [devdrv] [devdrv_manager_container_get_davinci_devlist 1248] <drv_hlt_dsmt_t:5615> get davinci devlist failed, ret(-22).
```

可能原因

针对调用Device失败，可能原因如下：

- 日志信息（1）：Docker容器启动时未映射设备。
- 日志信息（2）：Device被其他容器占用。

处理步骤

- 针对日志信息（1）可能原因，可以参考以下方法处理：
在容器内使用 **ls /dev** 指令查看是否映射Device。

```
[root@Euler ~]# ls /dev
console core fd full mqueue null ptmx pts random shm stderr stdin stdout tty urandom zero
```

如果查看到Docker容器内未映射设备，请参考《[CANN 软件安装指南](#)》资料启动容器，起Docker的命令中必须加挂载设备的参数，例如：

```
--device=/dev/davinciX --device=/dev/davinci_manager --device=/dev/devmm_svm --device=/dev/hisi_hdc
```

- 针对日志信息（2）可能原因，需要等使用该Device的容器被释放后再使用。
可以在容器内执行**ls /dev**命令查看当前容器内分配到的DavinciX，使用同样的方法查看其他容器分配到的DavinciX，然后通过手工停止Docker容器（例如：**docker stop 容器id**）的方式提前释放Device。

3.3.2.4 ssh 登录 Device 侧失败

适用场景

- 业务场景：环境检查
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

Host侧通过ssh登录Device侧失败，如[图3-12](#)和[图3-13](#)所示

图 3-12 Connection refused

```
root@huawie:~# ssh HwHiAiUser@192.168.1.199
ssh: connect to host 192.168.1.199 port 22: Connection refused
```


图 3-13 Connection timed out

```
root@ubuntu:~# ssh HwHiAiUser@192.168.1.199
ssh: connect to host 192.168.1.199 port 22: Connection timed out
root@ubuntu:~#
```

可能原因

根据ssh登录界面提示的错误信息，可能原因如下：

- 日志 “Connection refused”：Device侧ssh开关未打开。
- 日志 “Connection timed out”：Device侧未配置IP地址。

处理步骤

针对上述可能原因，可以参考以下方式处理

步骤1 参考《[DCMI API参考](#)》中的“dcmi_get_user_config”和“dcmi_set_user_config”接口原型查询配置项“ssh_status”，可读取和设置ssh开关状态。

“ssh_status”取值：

- 1：打开ssh服务
- 0：关闭ssh服务

步骤2 通过 ifconfig -a 指令查看Host侧 PCIe 虚拟网口是否配置IP，如[图3-14](#)所示表示未配置IP。

图 3-14 未配置 IP

```
root@ubuntu:~# ifconfig -a
endvnic  Link encap:Ethernet  HWaddr 10:1b:54:6d:89:d3
          BROADCAST MULTICAST  MTU:1500  Metric:1
          RX packets:0 errors:0 dropped:0 overruns:0 frame:0
          TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:0 (0.0 B)  TX bytes:0 (0.0 B)
```

如果未配置IP，可以使用root用户执行以下命令设置临时IP地址：

ifconfig endvnic 192.168.1.100

使用指令ssh HwHiAiUser@192.168.1.199登录Device侧，如果报“Host key verification failed”错误信息，可根据提示执行ssh-keygen -f "/root/.ssh/known_hosts" -R 192.168.1.199后再次登录。

说明

192.168.1.100、192.168.1.199为举例IP，请根据实际规划填写。

关于Device侧IP信息介绍、登录密码、以及如何设置静态IP地址等信息，请参考《[CANN 软件安装指南](#)》内容。

----结束

3.3.2.5 查询到的 Device 数量与实际设备数量不匹配

适用场景

- 业务场景：环境检查
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

环境上插了一张PCIE标卡（4个Device），使用npu-smi工具或upgrade-tool工具查询，只查到了1个Device。如图3-15所示。

图 3-15 设备查询

```
root@ubuntu:~# npu-smi info
+-----+
| npu-smi 20.0.0 | Version: 1.73.T5.0.B050 |
+-----+
| NPU   Name | Health | Power(W) | Temp(C) |
| Chip  Device | Bus-Id | AICore(%) | Memory-Usage(MB) |
+-----+
| 0     310   | OK     | 12.8     | 49      |
| 0     0     | 0000:01:00.0 | 0       | 2539 / 8192 |
+-----+

root@ubuntu:~# /usr/local/Ascend/driver/tools/upgrade-tool --device_index -1 --component -1 --version
{
  Get component version(1.73.5.0.B050) succeed for deviceId(0), componentType(0).
  {"device_id":0, "component":"nve", "version":"1.73.5.0.B050"}
  Get component version(1.73.5.0.B050) succeed for deviceId(0), componentType(1).
  {"device_id":0, "component":"xloader", "version":"1.73.5.0.B050"}
  Get component version(1.73.5.0.B050) succeed for deviceId(0), componentType(2).
  {"device_id":0, "component":"m3fw", "version":"1.73.5.0.B050"}
  Get component version(1.73.5.0.B050) succeed for deviceId(0), componentType(3).
  {"device_id":0, "component":"uefi", "version":"1.73.5.0.B050"}
  Get component version(1.73.5.0.B050) succeed for deviceId(0), componentType(4).
  {"device_id":0, "component":"tee", "version":"1.73.5.0.B050"}
}
```

可能原因

根据查询信息分析，可能存在以下原因：

- 主机散热不好，导致PCIE标卡温度过高，设备进入过温保护状态。
- Device的主机侧中断数量不足，无法自动加载驱动。
- Device通信线路问题：
 - Device硬件通信线路不通
 - Device通信线路断链

处理步骤

针对上述可能原因，可以参考以下方法处理：

步骤1 针对主机散热不好造成的Device异常，可以主机下电后再重新启动。

使用msnpureport工具导出并查看黑匣子日志，结合《[黑匣子日志参考](#)》中LPM3页签的内容，确认是否存在温度过高造成的异常码。

msnpureport工具使用方法请参见《[日志参考](#)》中“msnpureport工具使用”章节。

步骤2 使用dmesg命令查看打印信息，查看是否存在如下所示的日志信息。

如果存在类似日志信息，返回错误码-28（表示内核资源不足），可能为硬件环境MSI-X中断不足导致：

- 若安装昇腾AI处理器的主机为物理机，请查看硬件环境支持MSI-X中断的数量和BIOS配置，确保单芯片可申请116个中断。
- 若安装昇腾AI处理器的主机为虚拟机，请增加虚拟机的CPU配置。

```
[ 7.448019] devdrv_device_driver 0000:09:00.0: irq 503 for MSI/MSI-X
[ 7.448024] devdrv_device_driver 0000:09:00.0: irq 504 for MSI/MSI-X
[ 7.448030] devdrv_device_driver 0000:09:00.0: irq 505 for MSI/MSI-X
[ 7.448036] devdrv_device_driver 0000:09:00.0: irq 506 for MSI/MSI-X
[ 7.448042] devdrv_device_driver 0000:09:00.0: irq 507 for MSI/MSI-X
[ 7.448124] [drv_pcie] [devdrv_init_interrupt_normal 377] <systemd-udevd:605:605> vector_num -28
[ 7.448140] [ERROR] [drv_pcie] [devdrv_init_interrupt_normal 382] <systemd-udevd:605:605>
devdrv_device_driver: vector_num -28 error
[ 7.448143] [ERROR] [drv_pcie] [devdrv_probe 768] <systemd-udevd:605:605> devdrv_device_driver,
init interrupt failed. ret -1
[ 7.448374] devdrv_device_driver: probe of 0000:09:00.0 failed with error -1
[ 7.448387] [drv_pcie] [devdrv_probe 703] <systemd-udevd:605:605> probe driver IN. bdf:0a:00.0
[ 7.448511] [drv_pcie] [devdrv_set_startup_status 1404] <systemd-udevd:605:605> dev id -1 startup
status init jiffies 4294674711
[ 7.448641] [drv_pcie] [devdrv_register_pci_devctrl 1263] <systemd-udevd:605:605> devdrv_device_driver,
dev_id:2, bus:ffff8c1b39997c00
[ 7.448643] [drv_pcie] [devdrv_dev_startup_record 248] <systemd-udevd:605:605> probe new dev 2, add
to report,dev_num:3.
[ 7.448644] [drv_pcie] [devdrv_dev_startup_report 289] <systemd-udevd:605:605> dev startup no report
id:2
```

步骤3 使用lspci | grep d100 命令查询链接状态。

如图3-16，显示Device数量比实际数量少，原因为Device硬件通信线路不通。

图 3-16 通信线路不通

```
root@ubuntu:~# watch npu-smi info
root@ubuntu:~# lspci | grep d100
05:00.0 Processing accelerators: Device 19e5:d100 (rev 20)
root@ubuntu:~#
```

如图3-17，显示Device有“ff”状态，原因为Device通信线路断链。

图 3-17 通信线路断链

```
root@ubuntu:~# lspci |grep d100
03:00.0 Processing accelerators: Device 19e5:d100 (rev 20)
04:00.0 Processing accelerators: Device 19e5:d100 (rev ff)
05:00.0 Processing accelerators: Device 19e5:d100 (rev ff)
06:00.0 Processing accelerators: Device 19e5:d100 (rev ff)
```

可能是PCIE标卡与主机接触不良，则下电后重新插拔板卡，再上电启动。

如果以上操作无法解决该异常现象，请联系华为工程师处理。

----结束

3.3.2.6 I2C 通信失败

适用场景

- 业务场景：DVPP，昇腾模型算法，模型转化，精度调优、应用/算子开发，性能调优、TensorFlow训练
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

PCIE卡上昇腾AI处理器和MCU之间心跳连接失败，日志打印如图3-18类似信息。

图 3-18 心跳连接失败

```
[ 510.874320] i2c_designware 1309c0000.i2c: controller timed out
[ 512.874322] i2c_designware 1309c0000.i2c: controller timed out
[ 514.874321] i2c_designware 1309c0000.i2c: controller timed out
[ 516.874321] i2c_designware 1309c0000.i2c: controller timed out
[ 518.874322] i2c_designware 1309c0000.i2c: controller timed out
[ 520.874321] i2c_designware 1309c0000.i2c: controller timed out
[ 522.874320] i2c_designware 1309c0000.i2c: controller timed out
[ 524.874323] i2c_designware 1309c0000.i2c: controller timed out
[ 526.874323] i2c_designware 1309c0000.i2c: controller timed out
[ 528.874324] i2c_designware 1309c0000.i2c: controller timed out
[ 530.874321] i2c_designware 1309c0000.i2c: controller timed out
[ 532.874322] i2c_designware 1309c0000.i2c: controller timed out
[ 534.874327] i2c_designware 1309c0000.i2c: controller timed out
[ 536.874315] i2c_designware 1309c0000.i2c: controller timed out
[ 538.874321] i2c_designware 1309c0000.i2c: controller timed out
```

可能原因

针对上述现象，可能存在以下原因：

- PCIE标卡上昇腾AI处理器版本和MCU版本不匹配，导致不支持心跳功能。
- 昇腾AI处理器和MCU之间的通路有问题，导致心跳发送不出去。

处理步骤

针对上述可能原因，可以参考以下方法处理：

步骤1 查看PCIE标卡昇腾AI处理器的版本是否与MCU版本配套。

该场景为华为IT产品线定义的产品应用场景，请查看IT产品线提供的版本匹配兼容性列表。如果确认MCU版本不符合，请升级MCU版本，具体升级方法可以查看配套的产品手册。版本查询命令：

- 查询昇腾AI处理器版本：cat \${INSTALL_DIR}/xx模块/version.info
- 查询MCU版本：./npu-smi upgrade -b mcu -i <npu-id>

步骤2 如果昇腾AI处理器的版本与MCU版本配套正常，则可能是通信线路异常。

可以通过示波器检查线路，确认是否线路异常。如果是两者链路有问题，请联系华为售后工程师处理。

----结束

3.3.2.7 带外通信查询异常

适用场景

- 业务场景：DVPP，昇腾模型算法，模型转化，精度调优、应用/算子开发，性能调优、TensorFlow训练
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

通过BMC（Baseboard Management Controller）查询昇腾AI处理器相关状态，如温度、功耗等，查询失败，Device侧日志提示如图3-19所示。

图 3-19 查询昇腾 310 处理器状态失败

```
[ERROR] DRV:2019-09-12-10:17:01.717.120 [device monitor] [dsmi_check_out_valid 113] rcv msg data error code 3.  
[ERROR] DRV:2019-09-12-10:17:01.717.168 [device monitor] [dsmi_send_msg_rec_res 323] call dsmi_check_out_valid error:3.  
[ERROR] DRV:2019-09-12-10:17:01.717.225 [device monitor] [dsmi_cmd_passthru_mcu 929] dsmi_send_msg_rec_res failed
```

可能原因

- 昇腾AI处理器内部查询失败，导致无法给MCU返回正确结果。
- BMC的I2C通信线路异常，此时所有向昇腾AI处理器查询的命令都返回失败。

处理步骤

针对上述可能原因，可以参考以下方法处理：

步骤1 查看返回信息，根据返回的错误码确认原因并处理，返回值列表表3-1所示。

表 3-1 异常返回码

错误码	错误码含义	原因
1	获取数据超时。	设备侧出现异常，无法正常响应数据。
2	没有找到对应CMD的响应数据。	该查询CMD不支持。
3	该CMD的数据不可用。	该查询CMD数据不可用。
4	请求数据和相应数据长度不一致。	请求数据长度不对。
5	其他错误。	请求数据异常或者Device异常。

步骤2 如果所有查询命令都返回失败，则判断为通信线路异常，需联系华为售后工程师处理。

----结束

3.3.2.8 HDC 建链异常

适用场景

- 业务场景：DVPP，昇腾模型算法，模型转化，精度调优、应用/算子开发，性能调优、TensorFlow训练
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

查看dmesg信息，hdc驱动会话连接接口返回-512，有如下打印信息：

```
[ERROR][hdcdrv][hdcdrv_connect_wait_reply 2730]<hlt_hdc_test:41039>dev 3 connect wait error -512, local session_fd 153, remote session_fd ffffffff, pid -1  
[ERROR][hdcdrv][hdcdrv_connect 2857]<hlt_hdc_test:41039>device 3 service service_usr4 connect wait error -512, local session_fd 153, remote session_fd ffffffff, pid -1
```

可能原因

根据日志反馈码信息分析，可能是以下原因导致：

- 发起hdc会话连接的客户端进程，会等待服务端回复连接消息，如果在服务端回复连接消息前，对客户端进程执行ctrl+c或者kill -9操作，客户端进程退出，出现如上打印信息，属于正常现象，不需要处理。
- 使用gdb7.6跟踪调试程序，因gdb7.6通过tkill引入SIGSTOP信号，即使该操作是对于单个线程的，也会导致整个线程组中断崩溃，可能会出现如上打印，不符合预期执行结果。此时，需要客户新安装更高版本的gdb进行调试。

3.3.2.9 内存申请失败，出现 OOM 情况

适用场景

- 业务场景：DVPP，昇腾模型算法，模型转化，精度调优、应用/算子开发，性能调优、TensorFlow训练
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

内存申请失败，Host侧日志提示EL9999返回码，有如下打印信息：

```
[ERROR] DRV(2936187,python3):2022-04-21-14:19:39.429.481 [ascend][curpid: 2936187, 2969960][drv]  
[devmm][devmm_alloc_managed 182]<errno:12, 6> Heap_alloc_managed out of memory. (temp_ptr=0x1;  
bytesize=8592031776)  
[ERROR] RUNTIME(2936187,python3):2022-04-21-14:19:39.429.491 [npu_driver.cc:780]2969960  
DevMemAllocHugePageManaged:report error module_type=1, module_name=EL9999  
[ERROR] RUNTIME(2936187,python3):2022-04-21-14:19:39.429.495 [npu_driver.cc:780]2969960  
DevMemAllocHugePageManaged:[driver interface] halMemAlloc failed: device_id=1, size=8592031776,  
type=0, env_type=3, drvRetCode=6!
```


可能原因

根据日志信息分析, 判断为内存申请失败。可能原因:

1. 网络并行运行, 导致内存不足。
2. 网络运行需要内存过大, 导致内存申请失败。

处理步骤

步骤1 查看运行网络时是否存在并行情况。

步骤2 查询网络运行需要内存大小或者减少batchsize, 查看网络是否可以正常运行。

----结束

3.3.2.10 虚拟地址抢占导致 mmap 失败

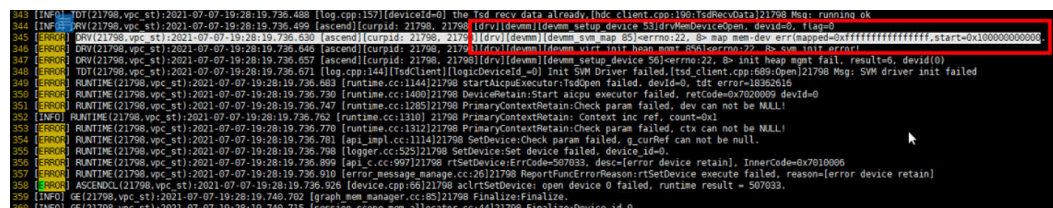
适用场景

- 业务场景: DVPP, 昇腾模型算法, 模型转化, 精度调优、应用/算子开发, 性能调优、TensorFlow训练
- 适用处理器: 昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态: EP、RC

现象描述

mmap申请失败, Host侧日志显示drvMemDeviceOpen失败, 出现如图3-20日志信息。

图 3-20 drvMemDeviceOpen 失败

A screenshot of a terminal log showing various error messages. The log includes timestamps and IP addresses. Key error messages include: 'drvMemDeviceOpen failed', 'mmap failed', 'SVM driver init failed', and 'PrimaryContextRetain: Check param failed'. A red box highlights a specific error message: 'mmap failed: No space left on device'. The log also shows 'INFO' and 'WARN' messages related to the device and context management.

可能原因

根据日志信息, 判断为mmap地址失败。可能原因:

- (1) 用户程序编译选项中启动了地址消毒 (-lasan): asan地址消毒会预留所有的虚拟地址空间, 而svm在mmap时也需要预留8T的地址空间, 二者冲突, 所以mmap时会失败。
- (2) 用户预留地址与SVM模块预留地址相同。

处理步骤

步骤1 关闭编译选项即可。

步骤2 查看用户预留地址是否与SVM模块预留地址重叠, 需要修改用户预留地址空间。

----结束

3.3.2.11 容器内 HDC 字符设备打开失败问题

适用场景

- 业务场景：容器场景
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP

现象描述

训练任务概率性失败，根据日志打印发现打开设备失败。

```
[drvHdcSetSessionReference 15333] <> Set reference open pcie device failed. (strerror=operation not permitted)
```

可能原因

根据日志信息，判断为hdc打开设备失败。可能原因：OS存在问题，概率性出现打开字符设备时设备权限和属组信息不正确。

处理步骤

确认OS是否支持连续打开字符设备，可以编写1个简单程序，连续打开系统字符设备dev/random，如出现问题请找OS定位。

3.3.2.12 DVPP 驱动引擎异常返回码

适用场景

- 业务场景：环境安装部署、DVPP，昇腾模型算法，模型转化，精度调优、应用/算子开发，性能调优、TensorFlow训练
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

视频解码失败，Device侧日志提示-512返回码：[dvpp_ioctl1_vpc 845] call proc failed:-512, engine_id:1，如图3-21所示。

图 3-21 DVPP 引擎返回-512 异常码

```
[ERROR] KERNEL(1200,sklogd):2019-11-29-21:42:07.421.714 [8078.067669] [dvpp] [dvpp_print_vpc_job 156] outctl out2_select:0
[ERROR] KERNEL(1200,sklogd):2019-11-29-21:42:07.421.740 [8078.067672] [dvpp] [dvpp_print_vpc_job 184] OUT
[ERROR] KERNEL(1200,sklogd):2019-11-29-21:42:07.421.764 [8078.067673] [dvpp] [dvpp_print_vpc_job 185] out use_flag:1
[ERROR] KERNEL(1200,sklogd):2019-11-29-21:42:07.421.789 [8078.067675] [dvpp] [dvpp_print_vpc_job 186] wr0 addr_frame_start:ffff35a00000
[ERROR] KERNEL(1200,sklogd):2019-11-29-21:42:07.421.823 [8078.067677] [dvpp] [dvpp_print_vpc_job 187] wr1 addr_frame_start:ffff35aca800
[ERROR] KERNEL(1200,sklogd):2019-11-29-21:42:07.421.853 [8078.067679] [dvpp] [dvpp_print_vpc_job 188] wr2 addr_frame_start:0
[ERROR] KERNEL(1200,sklogd):2019-11-29-21:42:07.421.880 [8078.067680] [dvpp] [dvpp_print_vpc_job 189] wr3 addr_frame_start:0
[ERROR] KERNEL(1200,sklogd):2019-11-29-21:42:07.421.905 [8078.068693] [dvpp] [dvpp_ioctl1_vpc 845] call proc failed:-512, engine_id:1
```

可能原因

根据日志信息，判断为VPC引擎返回的异常码，再根据-512返回码可以判断为等待视频解码异常中断。可能原因：用户态进程退出，导致等待解码时被异常中断。

处理步骤

查看Host侧的日志文件，检查用户态进程异常退出原因。

其他异常码参考

除样例VPC引擎返回的-512异常码外，DVPP驱动引擎还有其他引擎及各自的返回码，如表3-2所示。

您可以参考上述样例的方法，根据自身业务类型及Device侧日志文件中记录的异常返回码信息，先判断返回异常码的引擎类型，再根据返回码数值获取对应的返回码说明，初步判断故障原因。

表 3-2 DVPP 驱动引擎异常返回码

引擎类型	驱动IOCTL 返回码	返回码说明	常见错误日志参考
JPEGE	-22	参数错误或者不支持	常见错误打印： 1.返回值-35：日志 “JPEGE time wait, engine_id:xx” 编码超时 (执行器输入编码参数不 对等导致) 2.返回值-512：日志 “Signal interrupted, engine_id:xx” 编码错误 (编码被异常中断)
	-11	引擎处理异常， 需要重试	
	-14	拷贝用户态数据 时地址异常或者 释放引擎失败	
	-16	申请引擎失败	
	-35	编码超时	
	-5	编码失败且复位 成功	
	5	编码失败且复位 失败	
	-512	等待编码时被异 常中断	
PNGD	-22	参数错误或者不支持	常见错误打印： 1.返回值5/-5：日志 1 “Time out: wait failed xx” 解码超时(图片格式 不正确等码流异常等导 致) 1.返回值5/-5：日志 2 “Signal interrupted:xx, engine_id:xx” 解码时被 异常中断
	-14	拷贝用户态数据 时地址异常或者 释放引擎失败	
	-16	申请引擎失败	
	-5	解码失败且复位 成功	
	5	解码失败且复位 失败	

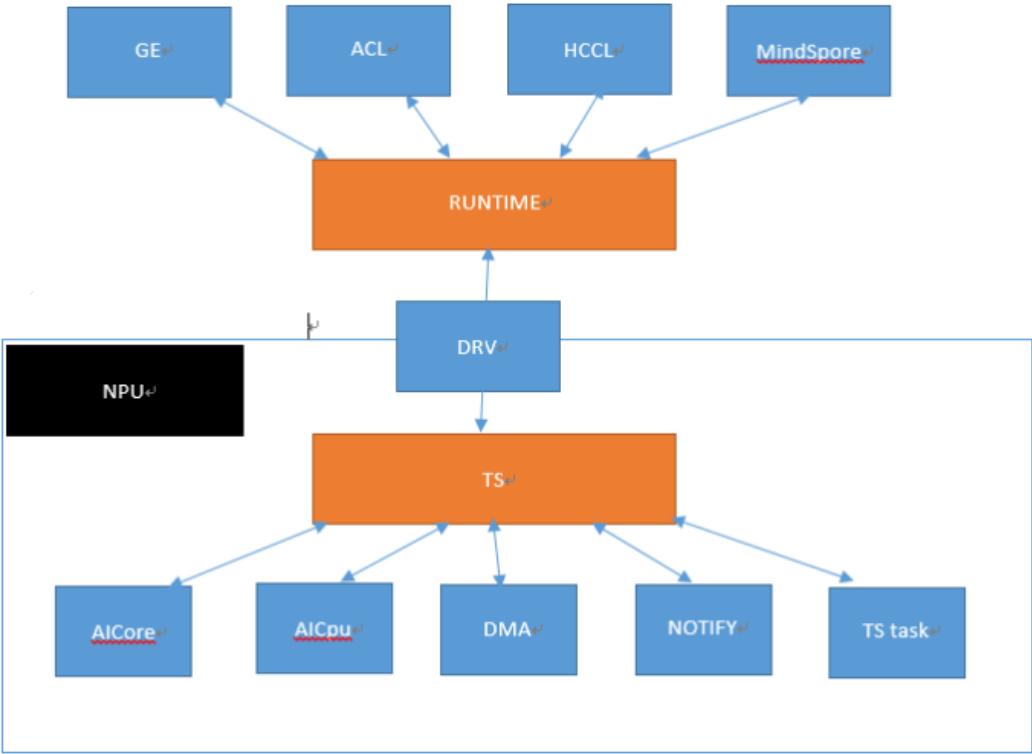
引擎类型	驱动IOCTL 返回码	返回码说明	常见错误日志参考
JPEGD	-22	参数错误或者不支持或者解码成功但是重置寄存器失败	常见错误打印： 1.返回值-1/-23：日志 “Decode abnormal, abnormal status:xx decode error! engine id:xx” 硬件上报异常中断 (图片格式不正确等码流异常导致) 2.返回值-2/-24：日志 “Decode abnormal, abnormal status:xx decode over time!, engine id:xx” 解码超时 (总线繁忙，系统缺页等导致)
	-14	拷贝用户态数据时地址异常或者释放引擎失败	
	-16	申请引擎失败	
	-1	解码失败且复位成功	
	-2/-3	解码超时且复位成功	
	-23	解码失败并且复位失败	
	-24/-25	解码超时并且复位失败	
VPC	-22	参数错误或者不支持	常见错误打印： 1. 返回值-35：日志 “dvpp_vpc_engine_p roc pipe wait time out: engine id:xx” 解码超时(总线繁忙、异常缺页等导致)。 2. 返回值-32：日志 “Hardware process failed, engine id:xx” 解码错误(码流异常等导致)。
	-11	引擎处理异常，需要重试	
	-14	拷贝用户态数据时地址异常或者释放引擎失败	
	-16	申请引擎失败	
	-35	解码超时	
	-512	等待解码时被异常中断	
	-32	硬件解码失败驱动捕捉到异常	
VPC_cmdlist	-22	参数错误或者不支持	常见错误打印： 1. 返回值-35：日志 “dvpp_vpc_engine_p roc pipe wait time out: engine id:xx” 解码超时(总线繁忙、异常缺页等导致)。 2. 返回值-32：日志 “Hardware process failed, engine id:xx”
	-11	引擎处理异常，需要重试	
	-14	拷贝用户态数据时地址异常或者释放引擎失败	
	-16	申请引擎失败	

引擎类型	驱动IOCTL 返回码	返回码说明	常见错误日志参考
	-35	解码超时	解码错误(码流异常等导致)。
	-512	等待解码时被异常中断	
	-32	硬件解码失败	
VENC	-22	参数错误或者不支持	常见错误打印： 1. 返回值-35：日志 “wait timeout, Ret value is xx” 未等到完成中断，编码失败(执行器配置参数错误或者总线繁忙等导致)。 2. 返回值-61：日志 “venc err interrupt state = xx, wait_time_remain = xx” 编码失败(执行器配置参数错误)。
	-11	引擎处理异常，需要重试	
	-14	拷贝用户态数据时地址异常或者释放引擎失败	
	-16	申请引擎失败	
	-35	未等到完成中断，编码失败	
	-61	上报编码超时中断	
VDEC	-22	参数错误或者不支持	常见错误打印： 1. 返回值-35：日志 “VDM wait time out,wait_time_remain = xx ,engine_id:xx” 解码超时(码流异常、总线繁忙等导致)。 2. 返回值=-61：日志 “wait_time_remain = xx,have vdm_error,p_vdm_backup_state->VdmState=xx ,engine_id:xx” 上报解码错误中断(码流异常等导致)。
	-11	引擎处理异常，需要重试	
	-14	拷贝用户态数据时地址异常或者释放引擎失败	
	-16	申请引擎失败	
	-35	解码超时	
	-61	上报解码错误中断	
	-4	等待解码被异常中断	

3.4 Runtime 常见故障

3.4.1 背景知识

Runtime和周边的关系：



上层模块

分类	序号	模块	功能
上层模块	1	GE	图引擎：负责，图编译，图下发，图运行的整体控制，包括图运行需要的内存，数据，算子的参数等信息的初始化。
	2	ACL	升腾对外编程接口，封装Runtime，提供对外基本功能，以及DVPP的功能封装。
	3	HCCL	训练中，负责集合通信（多P进程同步），以及训练间数据同步。
	4	Mind Spore	华为训练框架，负责训练的整体控制，调用GE，ACL HCCL 以及RUNTIME 实现。
	5	TF	第三方训练框架，负责训练的整体控制，调用GE，ACL HCCL 以及RUNTIME 实现。
	6	pytorch	第三方训练框架，负责训练的整体控制，调用GE，ACL HCCL 以及RUNTIME 实现。

底层模块

分类	序号	模块	功能
底层模块	1	DRV	1、负责Host和Device通信以及数据传输。 2、负责Device资源管理，（内存，stream，event，notify，model等）。 3、负责HDC，负责AI CPU的TDT数据通信。
	2	TS	整个NPU的调度执行中心，负责任务顺序调度执行。
	3	AI Core	执行AI Core算子。
	4	AI CPU	执行AI CPU算子和任务。
	5	DMA	负责数据搬移。
	6	notify	HCCL集合通信的底层功能实现。
	7	TS task	ts控制任务，由TS执行。

3.4.2 问题定位流程&方法

Runtime故障定位，一般采用至上而下的日志分析方法，根据业务流程逐步缩小到底层故障现象。

1. 查看plog日志，确认host的报错信息，plog日志路径为：/root/ascend/log/debug/plog/plog-pid_yyyymmddttttttt.log。
系统默认为ERROR级别的日志，如果需要更详细的日志，可以使用 **export ASCEND_GLOBAL_LOG_LEVEL=0** 或者**export ASCEND_GLOBAL_LOG_LEVEL=1**开启DEBUG或者INFO级别的日志。
注意：通过上面命令行方式，环境变量仅生效于执行界面，并且在需要重新执行任务才生效。
2. 查看device日志，确认device的报错信息。
使用/usr/local/Ascend/driver/tools/msnpureport命令将device侧收集日志到host侧，然后进行查看。

3.4.3 常见故障分析与处理

3.4.3.1 用户进程异常退出后重启进程失败

适用场景

- 业务场景：DVPP，昇腾模型算法，模型转化，精度调优、应用/算子开发，性能调优、TensorFlow训练、集群管理
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

用户进程卡住或者用户强制退出进程后，再次重启，重启后发现进程无法正常启动。类似的日志信息如下：

AscendCL日志信息：aclrtProcessReport failed

```
aclrtProcessReport failed, ret = 107012  
aclrtProcessReport failed, ret = 107012
```

在plog日志中Runtime打印了类似halResourceIdAlloc xxx failed的关键信息。

plog日志路径为：{install_path}/ascend/log/debug/plog/plog-pid_yyyymmddttttttt.log。

```
[ERROR] DRV(2086,rtstest_host):2021-06-09-02:14:46.034.368 [ascend][curpid: 2086, 2086][drv][tsdrv]  
[halResourceIdAlloc 477]id is exhausted, type(0 stream), range[0, 1024), dev_id(0), tsid(0).  
[ERROR] RUNTIME(2086,rtstest_host):2021-06-09-02:14:46.034.380 [npu_driver.cc:285]2086 StreamIdAlloc:  
[driver interface] halResourceIdAlloc streamid failed: device_id=0, tsid=0, drvRetCode=48!  
[ERROR] RUNTIME(2086,rtstest_host):2021-06-09-02:14:46.034.401 [stream.cc:448]2086 Setup:Failed to  
alloc stream id, retCode=0x702001a.  
[ERROR] RUNTIME(2086,rtstest_host):2021-06-09-02:14:46.034.416 [context.cc:1251]2086  
StreamCreate:Setup stream failed, retCode=0x702001a.  
[ERROR] RUNTIME(2086,rtstest_host):2021-06-09-02:14:46.034.440 [logger.cc:211]2086  
StreamCreate:Create stream failed, priority=7, flags=0.  
[ERROR] RUNTIME(2086,rtstest_host):2021-06-09-02:14:46.034.458 [api_c.cc:461]2086  
rtStreamCreateWithFlags:ErrCode=207008, desc=[driver error:no stream resource], InnerCode=0x702001a  
[ERROR] RUNTIME(2086,rtstest_host):2021-06-09-02:14:46.034.469 [error_message_manage.cc:26]2086  
ReportFuncErrorReason:rtStreamCreateWithFlags execute failed, reason=[driver error:no stream resource]
```

可能原因

通过日志分析无法正常重启的原因可能是taskid、stream id、eventid等资源申请不到引起的：

- 资源已经被其他进程占用完。
- 上一个进程退出时还未完全释放完资源。

处理步骤

针对上述可能原因，可以按以下方式处理：

- 等待一分钟后再重新启动进程，保证上一个进程资源释放完成。
- 停止其他进程或者等其他进程执行完成后再启动进程。
- 如果通过上述方式处理后仍然申请失败，建议检查是否超过了可用的资源上限，如果未超上限，则需要重启环境强行释放资源、恢复环境。

3.4.3.2 注册算子数超过最大规格

适用场景

- 业务场景：DVPP，昇腾模型算法，模型转化，精度调优、应用/算子开发，性能调优、TensorFlow训练、集群管理
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

推理过程中，用户load model出现报错。在plog日志中Runtime打印了类似ProgramRegister:Program register failed, program out of xxx和Register binary failed的关键信息。

plog日志在{install_path}/ascend/log/debug/plog路径下，日志格式为plog-pid_yyymmddhhmmss.log。

```
[ERROR] RUNTIME(3093,rtstest_host):2021-06-09-02:30:34.400.124 [runtime.cc:967]3093
ProgramRegister:Program register failed, program out of 40000000
[ERROR] RUNTIME(3093,rtstest_host):2021-06-09-02:30:34.400.155 [logger.cc:23]3093
DevBinaryRegister:Register binary failed.
[ERROR] RUNTIME(3093,rtstest_host):2021-06-09-02:30:34.400.182 [api_c.cc:127]3093
rtDevBinaryRegister:ErrCode=507032, desc=[program register num out of use], InnerCode=0x7090007
[ERROR] RUNTIME(3093,rtstest_host):2021-06-09-02:30:34.400.185 [error_message_manage.cc:26]3093
ReportFuncErrorReason:rtDevBinaryRegister execute failed, reason=[program register num out of use]
```

可能原因

通过日志分析报错的原因可能是一个进程内算子等资源注册超过最大规格：

- 模型太大，一个进程内的注册总算子数超过最大规格：Online模式规格为4000万，Offline模式规格为200万。

处理步骤

针对上述可能原因，可以按以下方式处理：

- 分析model，简化模型或者降低动态batch档次。
- 算子数是进程资源，model太大的情况下建议一个进程open一个device。
- 避免同一算子在不同模型中反复注册。
- 注册算子数不超过最大规格。

3.4.3.3 AI Core 算子执行报错

适用场景

- 业务场景：DVPP，昇腾模型算法，模型转化，精度调优、应用/算子开发，性能调优、TensorFlow训练、集群管理
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

Runtime执行报错，在plog日志中Runtime打印了类似fault kernel_name和func_name的关键信息。

plog日志在{install_path}/ascend/log/debug/plog路径下，日志格式为plog-pid_yyymmddhhmmss.log。

```
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.403.262 [engine.cc:1103]4150867
ReportExceptProc:[EXEC][DEFAULT]Task exception! device_id=0, stream_id=20, task_id=1, type=13,
retCode=0x91, [the model stream execute failed].
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.423 [device_error_proc.cc:495]4150867
PrintCoreErrorInfo:[EXEC][DEFAULT]The error from device(0), serial number is 193, there is an aicore error,
```

```
core id is 8, error code = 0x800000, dump info: pc start: 0x800120080047000, current: 0x1200800471cc, vec
error info: 0x7cafc4e, mte error info: 0x3000052, ifu error info: 0xc33f87bd7a80, ccu error info:
0xffd2bbd5005fe9d7, cube error info: 0x84, biu error info: 0, aic error mask: 0x65000200d000288, para base:
0x120080016300, errorStr: The DDR address of the MTE instruction is out of range.
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.443 [device_error_proc.cc:526]4150867
PrintCoreErrorInfo:[EXEC][DEFAULT]report error module_type=5, module_name=EZ9999
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.449 [device_error_proc.cc:526]4150867
PrintCoreErrorInfo:[EXEC][DEFAULT]The extend info from device(0), serial number is 193, there is aicore
error, core id is 8, aicore int: 0x10, aicore error2: 0, axi clamp ctrl: 0, axi clamp state: 0x1717, biu status0:
0x101d140000000000, biu status1: 0x80000201020000, clk gate mask: 0, dbg address: 0, ecc en: 0, mte ccu
ecc 1bit error: 0x2e80000000000000, vector cube ecc 1bit error: 0, run stall: 0x1, dbg data0: 0, dbg data1: 0,
dbg data2: 0, dbg data3: 0, dfx data: 0x8b
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.607 [task.cc:1021]4150867 PrintErrorInfo:
[EXEC][DEFAULT]Aicore kernel execute failed, device_id=0, stream_id=23, report_stream_id=20, task_id=24,
flip_num=0, fault kernel_name=16805736118314619649-1_0_1_Add_35,
func_name=te_add_729e2a87c649f49de98ac1a6fd491b3262ee7db9c1c2d6f4add7d7439aa3d22e_1__kernel0,
program id=22, hash=3338199064661472585.
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.618 [task.cc:3275]4150867 ReportErrorInfo:
[EXEC][DEFAULT]model execute error, retCode=0x91, [the model stream execute failed].
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.624 [task.cc:3247]4150867 PrintErrorInfo:
[EXEC][DEFAULT]model execute task failed, device_id=0, model stream_id=20, model task_id=1, flip_num=0,
model_id=3, first_task_id=65535
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.714 [stream.cc:929]4150867 GetError:[EXEC]
[DEFAULT]Stream Synchronize failed, stream_id=20, retCode=0x91, [the model stream execute failed].
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.742 [model.cc:581]4150867
SynchronizeExecute:[EXEC][DEFAULT]Fail to synchronize forbidden stream_id=20, retCode=0x7150050!
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.748 [model.cc:605]4150867
GetStreamToSyncExecute:[EXEC][DEFAULT]report error module_type=0, module_name=EE9999
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.753 [model.cc:605]4150867
GetStreamToSyncExecute:[EXEC][DEFAULT]Model synchronize execute failed, model_id=3!
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.774 [logger.cc:856]4150867 ModelExecute:
[EXEC][DEFAULT]Execute model failed.
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.787 [api_c.cc:2063]4150867 rtModelExecute:
[EXEC][DEFAULT]ErrCode=507011, desc=[the model stream execute failed], InnerCode=0x7150050
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.793 [error_message_manage.cc:49]4150867
FuncErrorReason:[EXEC][DEFAULT]report error module_type=3, module_name=EE8888
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.801 [error_message_manage.cc:49]4150867
FuncErrorReason:[EXEC][DEFAULT]rtModelExecute execute failed, reason=[the model stream execute failed]
```

可能原因

AI Core 算子执行失败，可能算子本身代码问题：数据输入不匹配、访问越界、计算溢出等异常。

查阅plog日志，根据fault kernel_name和func_name可获取报错算子名称和报错函数名称。

plog日志在/root/ascend/log/debug/plog路径下，日志格式为plog-pid_yyymmddhhmmss.log。

```
[ERROR] RUNTIME(4150867,msame):2022-09-22-09:27:46.404.607 [task.cc:1021]4150867 PrintErrorInfo:
[EXEC][DEFAULT]Aicore kernel execute failed, device_id=0, stream_id=23, report_stream_id=20, task_id=24,
flip_num=0, fault kernel_name=16805736118314619649-1_0_1_Add_35,
func_name=te_add_729e2a87c649f49de98ac1a6fd491b3262ee7db9c1c2d6f4add7d7439aa3d22e_1__kernel0,
program id=22, hash=3338199064661472585.
```

处理步骤

该类型错误，需要联系华为算子开发工程师定位排查。您可以通过<https://gitee.com/ascend>网站提交issue获取帮助。

可能导致的故障

模型下沉场景下，该问题可能导致ACL报错Execute model failed，并打印在plog日志中。


```
[ERROR] ASCENDCL(4150867,msame):2022-09-22-09:27:46.404.834 [model.cpp:699]4150867
ModelExecute: [EXEC][DEFAULT][Exec][Model]Execute model failed, ge result[507011], modelId[1]
[ERROR] ASCENDCL(4150867,msame):2022-09-22-09:27:46.404.857 [model.cpp:1547]4150867
aclmdlExecute: [EXEC][DEFAULT][Exec][Model]modelId[1] execute failed, result[507011]
```

非模型下沉场景下，该问题可能导致算子执行失败ACL报错get op desc failed，Runtime报错Aicore kernel execute failed，并打印在plog日志中。

```
[ERROR] RUNTIME(2856615,xacfk):2022-09-15-11:36:47.817.465 [task.cc:1058]2856939 PreCheckTaskErr:
[EXEC][DEFAULT]Kernel task happen error, retCode=0x26, [aicore exception].
[ERROR] RUNTIME(2856615,xacfk):2022-09-15-11:36:47.817.538 [task.cc:1029]2856939 PrintErrorInfo:
[EXEC][DEFAULT]Aicore kernel execute failed, device_id=0, stream_id=0, report_stream_id=0, task_id=615,
flip_num=0, fault kernel_name=12646006_1663210912148832_-1_0_while/transformer_0/decoder/
layer_0/rnn/rnn/while/Select,
func_name=te_select_7b314df6791292127cb82df985d04ddaf6d069cb31aaccec00e0b8ee2e997f20_1__kernel
0, program id=131, hash=14736095126365135477.
[ERROR] GE(2856615,xacfk):2022-09-15-11:36:47.818.283 [graph_execute.cc:557]2856939 GetOpDescInfo:
ErrorNo: 4294967295(failed) [EXEC][DEFAULT][Get][OpDescInfo] failed, device_id:0, stream_id:0, task_id:
615.
[ERROR] GE(2856615,xacfk):2022-09-15-11:36:47.818.308 [ge_executor.cc:1332]2856939 GetOpDescInfo:
ErrorNo: 4294967295(failed) [EXEC][DEFAULT][Get][OpDescInfo] failed, device_id:0, stream_id:0, task_id:
615.
[ERROR] ASCENDCL(2856615,xacfk):2022-09-15-11:36:47.818.315 [model.cpp:2216]2856939
aclmdlCreateAndGetOpDesc: [EXEC][DEFAULT][Get][OpDescInfo]get op desc failed, ge result[-1],
deviceId[0], streamId[0], taskId[615]
```

3.4.3.4 AI CPU 算子执行报错

适用场景

- 业务场景：DVPP，昇腾模型算法，模型转化，精度调优、应用/算子开发，性能调优、TensorFlow训练、集群管理
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器
- 处理器形态：EP、RC

现象描述

Runtime执行报错，在plog日志中Runtime打印了PrintAicpuErrorInfo的错误信息。

plog日志在{install_path}/ascend/log/debug/plog路径下，日志格式为plog-pid_yyymmddhhmmss.log。

```
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.791.865 [engine.cc:1103]16282
ReportExceptProc:Task exception! device_id=0, stream_id=7, task_id=2, type=1, retCode=0x2a, [aicpu
exception].
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.793.489 [device_error_proc.cc:669]16282
ProcessAicpuErrorInfo:report error module_type=0, module_name=E39999
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.793.498 [device_error_proc.cc:669]16282
ProcessAicpuErrorInfo:An exception occurred during AICPU execution, stream_id:7, task_id:2, errcode:5,
msg:aicpu execute failed.
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.793.932 [task.cc:1050]16282
PreCheckTaskErr:report error module_type=5, module_name=EZ9999
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.793.941 [task.cc:1050]16282
PreCheckTaskErr:Kernel task happen error, retCode=0x2a, [aicpu exception].
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.793.981 [task.cc:759]16282
PrintAicpuErrorInfo:report error module_type=0, module_name=E39999
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.793.990 [task.cc:759]16282
PrintAicpuErrorInfo:Aicpu kernel execute failed, device_id=0, stream_id=7, task_id=2.
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.794.116 [task.cc:777]16282
PrintAicpuErrorInfo:Aicpu kernel execute failed, device_id=0, stream_id=7, task_id=2, flip_num=0, fault
so_name=, fault kernel_name=, fault op_name=Unique, extend_info=(info_type:4, info_len:6,
msg_info:Unique).
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.794.384 [stream.cc:929]16243 GetError:[EXEC]
[DEFAULT]Stream Synchronize failed, stream_id=7, retCode=0x2a, [aicpu exception].
```

```
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.794.407 [stream.cc:932]16243 GetError:[EXEC]
[DEFAULT]report error module_type=0, module_name=E39999
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.794.419 [stream.cc:932]16243 GetError:[EXEC]
[DEFAULT]Aicpu kernel execute failed, device_id=0, stream_id=7, task_id=2, flip_num=0, fault_so_name=,
fault_kernel_name=, fault_op_name=Unique, extend_info=(info_type:4, info_len:6, msg_info:Unique)
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.794.482 [logger.cc:305]16243 StreamSynchronize:
[EXEC][DEFAULT]Stream synchronize failed, stream = 0x5643fe3e28d0
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.794.510 [api_c.cc:661]16243 rtStreamSynchronize:
[EXEC][DEFAULT]ErrCode=507018, desc=[aicpu exception], InnerCode=0x715002a
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.794.519 [error_message_manage.cc:49]16243
FuncErrorReason:[EXEC][DEFAULT]report error module_type=3, module_name=EE8888
[ERROR] RUNTIME(16243,msame):2022-09-22-11:27:01.794.532 [error_message_manage.cc:49]16243
FuncErrorReason:[EXEC][DEFAULT]rtStreamSynchronize execute failed, reason=[aicpu exception]
```

可能原因

AI CPU算子执行失败，可能算子本身代码问题：数据输入不匹配、访问越界、AI CPU线程挂死等问题。

比如通过查阅AI CPU的device日志，是数据输入不匹配导致。

device日志在/root/ascend/log/debug/device-x/路径下，日志格式为device-pid_yyyymmddhhmmss.log。

```
[ERROR] CCECPU(2309,aicpu_scheduler):2022-09-22-11:27:00.733.218 [aicpu_tf_kernel.cc:348][tid:2317]
[TFAdapter] AICPUKernelAndDevice::Run failure, kernel_id=0, op_name=Unique, op_type=UniqueExt,
error=Invalid argument: unique expects a 1D vector.
[ERROR] CCECPU(2309,aicpu_scheduler):2022-09-22-11:27:00.733.242 [tf_adpt_session_mgr.cc:74][tid:2317]
[TFAdapter] [sessionID:0] Failed to Run kernel, kernel_id=0.
[ERROR] CCECPU(2309,aicpu_scheduler):2022-09-22-11:27:00.733.261 [tf_adpt_session_mgr.cc:434][tid:
2317][TFAdapter] [sessionID:0] Run kernel on session failed.
[ERROR] CCECPU(2309,aicpu_scheduler):2022-09-22-11:27:00.733.277 [tf_adpt_api.cc:85][tid:2317]
[TFAdapter] [sessionID:0] Invoke TFOperateAPI failed.
[ERROR] CCECPU(2309,aicpu_scheduler):2022-09-22-11:27:00.733.296 [ae_kernel_lib_fw.cc:229]
[TransformKernelErrorCode][tid:2317][AICPU_PROCESSER] Call tf api return failed:5, input param to tf api:
0x124040017004
[ERROR] CCECPU(2309,aicpu_scheduler):2022-09-22-11:27:00.733.366 [aicpusd_event_process.cpp:1325]
[ExecuteTsKernelTask][tid:2317] Aicpu engine process failed, result[5].
```

处理步骤

该类型错误，需要联系华为算子开发工程师定位排查。您可以通过<https://gitee.com/ascend>网站提交issue获取帮助。

可能导致的故障

该问题会导致ACL报错Execute model failed，并打印在plog日志中。

```
[ERROR] ASCENDCL(16243,msame):2022-09-22-11:27:01.794.603 [model.cpp:599]16243
RuntimeV2ModelExecute: [EXEC][DEFAULT][Exec][Model]Execute model failed, ge result[4294967295],
modelId[2147483648]
[ERROR] ASCENDCL(16243,msame):2022-09-22-11:27:01.794.634 [model.cpp:1547]16243 aclmdlExecute:
[EXEC][DEFAULT][Exec][Model]modelId[2147483648] execute failed, result[500002]
```

3.4.3.5 Memcpy 异步拷贝算子执行报错

适用场景

- 业务场景：DVPP，昇腾模型算法，模型转化，精度调优、应用/算子开发，性能调优、TensorFlow训练、集群管理
- 适用处理器：昇腾310 AI处理器、昇腾310P AI处理器、昇腾910 AI处理器

- 处理器形态：EP、RC

现象描述

Runtime执行报错，在plog日志中Runtime打印了Memory async copy failed的错误信息。

plog日志在{install_path}/ascend/log/debug/plog路径下，日志格式为plog-pid_yyymmddhhmmss.log。

```
[ERROR] RUNTIME(420723,msame):2022-08-17-04:49:06.338.095 [engine.cc:1103]420723
ReportExceptProc:Task exception! device_id=0, stream_id=0, task_id=1, type=13, retCode=0x91, [the model
stream execute failed].
[ERROR] RUNTIME(420723,msame):2022-08-17-04:49:06.338.955 [device_error_proc.cc:648]420723
ProcessSdmaErrorInfo:report error module_type=3, module_name=EE8888
[ERROR] RUNTIME(420723,msame):2022-08-17-04:49:06.339.075 [task.cc:1582]420723
PrintErrorInfo:Memory async copy failed, device_id=0, stream_id=3, task_id=2, flip_num=0, copy_type=10,
memcpy_type=0, copy_data_type=0, length=4096.
[ERROR] RUNTIME(420723,msame):2022-08-17-04:49:06.339.079 [task.cc:1589]420723
PrintErrorInfo:Memory async copy failed, device_id=0, stream_id=3, task_id=2, flip_num=0, copy_type=10,
memcpy_type=0, copy_data_type=0, length=4096, src_addr=0x124080015000, dst_addr=0x124080016000.
[ERROR] RUNTIME(420723,msame):2022-08-17-04:49:06.339.082 [task.cc:3276]420723
ReportErrorInfo:model execute error, retCode=0x91, [the model stream execute failed].
```

可能原因

计算溢出、拷贝地址错误、多P训练时进程退出等。可以通过在plog日志中查找CQE status值确定具体原因（不同处理器平台，打印的日志形式不一样）：

```
[ERROR] RUNTIME(109239,ascend-dmi):2023-06-12-10:25:40.168.867 [device_error_proc.cc:1123]109239
ProcessSdmaErrorInfo:report error module_type=3, module_name=EE8888
[ERROR] RUNTIME(109239,ascend-dmi):2023-06-12-10:25:40.168.877 [device_error_proc.cc:1123]109239
ProcessSdmaErrorInfo:The error from device(chipId:0, dielId:0), serial number is 1633. there is a sdma
error, sdma channel is 11, sdmaBlkFsmState=0x7, dfxSdmaBlkFsmOstCnt=0x0, sdmaChFree=0x0,
irqStatus=0x220000, cqeStatus=0x3
[ERROR] RUNTIME(109239,ascend-dmi):2023-06-12-10:25:40.168.934 [task.cc:2251]109239
DoCompleteSuccess:mem async copy error, retCode=0x20b, [memcpy exception].
```

或

```
[ERROR] RUNTIME(92299,python):2023-06-08-12:25:55.259.270 [device_error_proc.cc:652]97319
ProcessSdmaErrorInfo:The error from device(0), serial number is 1. there is a sdma error, sdma channel is 0,
the channel exist the following problems: The SMMU returns a Terminate error during page table
translation.. the value of CQE status is 2. the description of CQE status: When the SQE translates a page
table, the SMMU returns a Terminate error.it's config include: setting1=0xc0000008c090000,
setting2=0xff009000ff004c, setting3=0x1f, sq base addr=0x6243d000
```

表 3-3 CQE 状态值描述

Value	Description
0000h	Successful completion: The command completed successfully.
0001h	表示可能出现如下错误中的一种： <ul style="list-style-type: none">● Submission Descriptor read response error。● SQE中出现非法OPCODE。
0002h	SQE进行页表翻译时，SMMU返回了Terminate错误。
0003h	Reserved。

Value	Description
0004h	SDMAA上报的错误，表示使用了非安全属性访问安全DDR空间。
0005h	SDMAA上报的错误，表示SDMAM下发的搬运地址在DAW没有映射。
0006h	SDMAA上报的错误，表示出现操作类型错误。
0007h	SDMAA上报的错误，表示SDMAA搬运过程中出现DDRC错误。
0008h	SDMAA上报的错误，表示SDMAA搬运过程中出现ECC错误。
0009h	SDMAA上报的错误，表示SDMAA搬运过程中出现COMPERR。
000Ah	SDMAA上报的错误，表示SDMAA搬运过程中出现COMPDATAERR。
000Bh	SDMAA上报的错误，表示reduce操作出现上溢错误。
000Ch	SDMAA上报的错误，表示reduce操作出现下溢错误。
000Dh	SDMAA上报的错误，表示reduce源数据格式不符合要求。
000Eh	SDMAA上报的错误，表示reduce目的数据格式不符合要求。
000Fh	SDMAA上报的错误，表示reduce源和目的数据格式不符合要求。
else	Reserved

处理步骤

根据错误码提示，进行后续处理。

例如，错误码0x2代表缺少页表。

- 地址错误，使用了未申请地址或者使用时地址已经释放。
- 拷贝长度错误，拷贝越界到未申请的地址上。
- 进程结束，地址被系统释放，导致错误。

这三种错误，由调用者分析具体错误情况。

例如，错误码：0x9、0xA、0xB、0xC、0xD、0xE、0xF。

- 拷贝数据错误，仅HCCL使用，一般是算子溢出，或者即将溢出出现了拷贝报错。通过开启溢出dump功能进行检测定位，发现溢出问题按照算子溢出进行深入定位。
- 如果通过算子溢出未定位到故障原因，则使用普通dump功能进行精度比对，进行网络精度分析。
- 如果仍然存在问题，需要联系华为HCCL开发工程师定位排查。您可以通过<https://gitee.com/ascend>网站提交issue获取帮助。

可能导致的故障

该问题会导致ACL报错Execute model failed，并打印在plog日志中。

```
[ERROR] ASCENDCL(16243,msame):2022-08-17-04:49:06.339.187 [model.cpp:599]21674
RuntimeV2ModelExecute: [EXEC][DEFAULT][Exec][Model]Execute model failed, ge result[4294967295],
modelld[0]
[ERROR] ASCENDCL(16243,msame):2022-08-17-04:49:06.339.193 [model.cpp:1547]21674 aclmdlExecute:
[EXEC][DEFAULT][Exec][Model]modelld[0] execute failed, result[500002]
...
[ERROR] ASCENDCL(16243,msame):2022-08-17-04:49:06.342.397 [model.cpp:599]21674
RuntimeV2ModelExecute: [EXEC][DEFAULT][Exec][Model]Execute model failed, ge result[4294967295],
modelld[1]
[ERROR] ASCENDCL(16243,msame):2022-08-17-04:49:06.342.399 [model.cpp:1547]21674 aclmdlExecute:
[EXEC][DEFAULT][Exec][Model]modelld[1] execute failed, result[500002]
```

3.4.3.6 Runtime 报 Notify wait 错误

适用场景

- 业务场景：DVPP，昇腾模型算法，模型转化，精度调优、应用/算子开发，性能调优、TensorFlow训练、集群管理
- 适用处理器：昇腾910 AI处理器
- 处理器形态：EP

现象描述

Runtime执行报错，在plog日志中Runtime打印了PrintErrorInfo:Notify wait execute failed的错误信息和PrintStreamTimeoutSnapshotInfo的Snapshot快照打印信息。

plog日志在{install_path}/ascend/log/debug/plog路径下，日志格式为plog-pid_yymmddhhmmss.log。

```
[ERROR] RUNTIME(935255,rtstest_host):2023-05-11-20:08:58.101.548 [engine.cc:1361]935255
ReportExceptProc:Real task exception! device_id=0, stream_id=1113, task_id=2, task_type=14 (NOTIFY_WAIT)
[ERROR] RUNTIME(935255,rtstest_host):2023-05-11-20:08:58.101.557 [engine.cc:1366]935255
ReportExceptProc:Task exception! device_id=0, stream_id=1867, task_id=1, type=13(MODEL_EXECUTE),
failuremode =0, retCode=0x91, [the model stream execute failed]
[ERROR] RUNTIME(935255,rtstest_host):2023-05-11-20:08:58.101.593 [device_error_proc.cc:968]935255
PrintStreamTimeoutSnapshotInfo:stream_id=1867, task_id=1, taskType=13 (MODEL_EXECUTE), .
[ERROR] RUNTIME(935255,rtstest_host):2023-05-11-20:08:58.102.330 [task.cc:3786]935255
ReportErrorInfo:model execute error, retCode=0x91, [the model stream execute failed].
[ERROR] RUNTIME(935255,rtstest_host):2023-05-11-20:08:58.102.336 [task.cc:3766]935255
PrintErrorInfo:model execute task failed, device_id=0, model stream_id=1867, model task_id=1, flip_num=0,
model_id=773, first_task_id=65535.
[ERROR] RUNTIME(935255,rtstest_host):2023-05-11-20:08:58.102.344 [task.cc:4106]935255
PrintErrorInfo:Notify wait execute failed, device_id=0, stream_id=1113, task_id=2, flip_num=0,
notify_id=535
[ERROR] RUNTIME(935255,rtstest_host):2023-05-11-20:08:58.102.350 [callback.cc:91]935255 Notify:notify
[_DEFAULT_MODEL_NAME_] task fail start.notify taskid:2 streamid:1113 retcode:507011
[ERROR] RUNTIME(935255,rtstest_host):2023-05-11-20:08:58.102.363 [stream.cc:1139]935255
ReportError:Stream Synchronize failed, stream_id=1867, retCode=0x91, [the model stream execute failed].
[ERROR] RUNTIME(935255,rtstest_host):2023-05-11-20:08:58.102.366 [stream.cc:1142]935255
GetError:report error module_type=2, module_name=EI9999
[ERROR] RUNTIME(935255,rtstest_host):2023-05-11-20:08:58.102.369 [stream.cc:1142]935255
GetError:Notify wait execute failed, device_id=0, stream_id=1113, task_id=2, flip_num=0, notify_id=535
[ERROR] RUNTIME(935255,rtstest_host):2023-05-11-20:08:58.110.670 [logger.cc:364]935255
StreamSynchronize:Stream synchronize failed
[ERROR] RUNTIME(935255,rtstest_host):2023-05-11-20:08:58.110.688 [api_c.cc:735]935255
rtStreamSynchronize:ErrCode=507011, desc=[the model stream execute failed], InnerCode=0x7150050
[ERROR] RUNTIME(935255,rtstest_host):2023-05-11-20:08:58.110.691 [error_message_manage.cc:
49]935255 FuncErrorReason:report error module_type=3, module_name=EE8888
[ERROR] RUNTIME(935255,rtstest_host):2023-05-11-20:08:58.110.695 [error_message_manage.cc:
49]935255 FuncErrorReason:rtStreamSynchronize execute failed, reason=[the model stream execute failed]
```

可能原因

可能是多P训练中，其中1P异常，其他P等待超时；也有可能是多P训练中其中1P计算不同步，导致其他P等待超时；还有可能是通信失败，导致通信丢失报错。

处理步骤

1. 在plog日志中，执行命令**grep -rns ERROR**查看所有训练设备的ERROR日志，排查Notify wait超时前的所有故障并完成处理。
2. 集合通信场景下已经定位到根节点且该类型报错为首报错，单机场景下定位到该节点且该错误类型为首报错，需要联系华为算子开发工程师定位排查。

执行asys.py工具复跑业务，收集维测日志信息，再通过<https://gitee.com/ascend>网站提交issue获取帮助。

可能导致的故障

该问题会导致HCCL报错Task run failed，任务类型为Notify Wait，并打印在plog日志中。

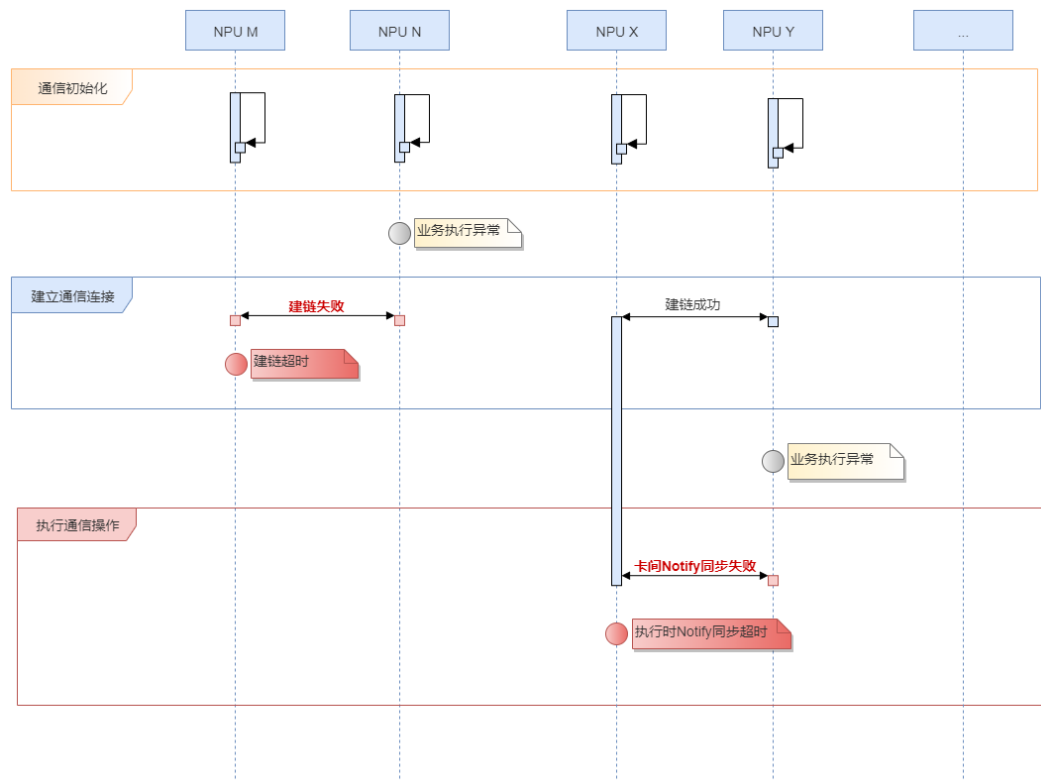
```
[ERROR]HCCL(22728,python3.7):2023-05-11-11:15:22.967.300 [task_exception_handler.cc:223][22728]
[64817][EXEC][EXEC][TaskExceptionHandler][Callback]Task run failed, base information is streamID:[67],
taskID[55], taskType[Notify Wait], tag[HcomAllGather_6629421139219749105_5].
[ERROR]HCCL(22728,python3.7):2023-05-11-11:15:22.967.335 [task_exception_handler.cc:225][22728]
[64817][EXEC][EXEC][TaskExceptionHandler][Callback]Task run failed, para information is notify id:
[0x000000003000000a0], stage:[ffffff], remote rank:[11].
```

3.5 HCCL&HCCP 常见故障

3.5.1 背景知识

HCCL提供了深度学习训练场景中服务器间高性能集合通信功能，其执行流程如下：

图 3-22 分布式训练通信阶段



通信过程分为三个阶段：

阶段一、通信初始化：获取必要的集合通信参数配置并初始化网络设备；

阶段二、建立通信连接：建立socket连接并交换通信两端的通信参数和内存信息；

阶段三、执行通信操作：通过Notify同步设备执行状态，传递内存数据。

通信初始化阶段不涉及昇腾设备之间的交互。

建立通信连接阶段，HCCL根据用户提供的集群信息结合网络拓扑与其他卡进行建链并交换用于通信的参数信息。如果在建链超时时间阈值（默认120s，可通过环境变量HCCL_CONNECT_TIMEOUT配置）内未得到其他卡的及时响应，会上报建链超时错误并退出训练。

通信操作执行阶段HCCL会根据通信算法编排NOTIFY/SDMA等task并通过runtime下发给昇腾设备task调度器，设备根据编排信息调度并执行task。其中Notify类task用于卡间同步，Notify wait会阻塞task流执行直到对应的Notify record到达，以确保后续的通信操作执行时彼此的内存处于ready状态。

单卡问题、通信链路问题均可能会导致集群出现大量的超时错误，在定位集群问题时
需要锁定问题出现位置，因此需要收集整个集群的日志信息。

3.5.2 常见故障分析与处理

3.5.2.1 通信操作超时 (EI0002)

适用场景

- 业务场景：集群训练
- 适用处理器：昇腾910 AI处理器
- 处理器形态：EP

现象描述

常见于算子执行阶段，屏显日志报错信息：The wait execution of the Notify register times out.

```
Error Message is :
EI0002: The wait execution of the Notify register times out. Reason: The Notify register has not received the Notify record from remote rank [4].base information: [streamID:[14], taskID:[6], taskType[Notify
ait], tag[HCCL@hdcube_6629421139219749105_0],] task information: [notify id:[0x0000000700000050], stage:[ffffff], remote rank:[4].
there are[15] 1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15.
serverid[199.188.106.111], deviceId[4], heartbeat Lost Occurred, Possible Reason: 1. Process has existed, 2. Network Disconnected

Possible Causes: 1. An exception occurs during the execution on some NPUs in the cluster. As a result, collective communication operation failed. 2. The execution speed on some NPU in the cluster is
o slow to complete a communication operation within the timeout interval. (default 600s, You can set the interval by using HCCL_EXEC_TIMEOUT).3. The number of training samples of each NPU is inconsistent.4
Packet loss or other connectivity problems occur on the communication link.
Solution: 1. If this error is reported on part of these ranks, check other ranks to see whether other errors have been reported earlier.2. If this error is reported for all ranks, check whether the
ror reporting time is consistent (the maximum difference must not exceed 600s). If not, locate the cause or adjust the locate the cause or set the HCCL_EXEC_TIMEOUT environment variable to a larger value.
Check whether the completion queue element (Cqe) of the error exists in the plog(grep -rn "error cqe"). If so, check the network connection status. (For details, see the TLS command and HCCL connectivity
eck examples.14. Ensure that the number of training samples of each NPU is consistent.
[drv api] halReportExecv failed: deviceId=7, taskId=4, type=2, cqeId=5, driverCode=18[FUNC:Loggic@hpcv2][FILE:cpu_driver.cc:LINE:803]
Notify wait execute failed, deviceId=7, streamId=14, taskId=6, flip_num=0, notify_id=11[FUNC:GetError][FILE:stream.cc:LINE:836]
rtStreamSynchronize execute failed, reason[the model stream execute failed][FUNC:FuncErrorReason][FILE:error_message_manage.cc:LINE:45]
```

plog日志中查询Notify，有如下类似信息：

```
[root@localhost ~]# grep -rn "Notify" plog
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 17:08:59.580.252 [task.cc:2994][218886] PrintErrorInfo:[EXEC][EXEC]Notify wait execute failed, device_id=0, stream_id=0, task_id=4, flip_num=0, notifi
y_id=7.
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 17:08:59.580.472 [task_exception_handler.cc:194][218874][218886][EXEC][EXEC][TaskExceptionHandler][Callback]Task faild base information is streamID:[1
0], taskID:[4], taskType[Notify Wait], tag[HCCL@hdcube_6629421139219749105_12].
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 17:08:59.625.785 [task.cc:2994][218886] PrintErrorInfo:[EXEC][EXEC]Notify wait execute failed, device_id=3, stream_id=12, task_id=13, flip_num=0, notifi
y_id=6.
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 17:08:59.626.924 [task_exception_handler.cc:194][218892][218886][EXEC][EXEC][TaskExceptionHandler][Callback]Task faild base information is streamID:[1
0], taskID:[13], taskType[Notify Wait], tag[HCCL@hdcube_6629421139219749105_12].
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 17:08:59.626.992 [task.cc:2994][218886] PrintErrorInfo:[EXEC][EXEC]Notify wait execute failed, device_id=4, stream_id=12, task_id=29, flip_num=0, notifi
y_id=6.
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 17:08:59.626.188 [task_exception_handler.cc:194][218855][218885][EXEC][EXEC][TaskExceptionHandler][Callback]Task faild base information is streamID:[1
12], taskID:[29], taskType[Notify Wait], tag[HCCL@hdcube_6629421139219749105_12].
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 17:08:59.626.806 [task.cc:2994][218884] PrintErrorInfo:[EXEC][EXEC]Notify wait execute failed, device_id=5, stream_id=16, task_id=11, flip_num=0, notifi
y_id=6.
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 17:08:59.627.607 [task.cc:2994][218883] PrintErrorInfo:[EXEC][EXEC]Notify wait execute failed, device_id=1, stream_id=16, task_id=4, flip_num=0, notifi
y_id=2.
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 17:08:59.628.743 [task_exception_handler.cc:194][218868][218883][EXEC][EXEC][TaskExceptionHandler][Callback]Task faild base information is streamID:[1
0], taskID:[1], taskType[Notify Wait], tag[HCCL@hdcube_6629421139219749105_12].
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 17:08:59.628.308 [task.cc:2994][218881] PrintErrorInfo:[EXEC][EXEC]Notify wait execute failed, device_id=6, stream_id=16, task_id=6, flip_num=0, notifi
y_id=1.
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 17:08:59.625.537 [task_exception_handler.cc:194][218893][218881][EXEC][EXEC][TaskExceptionHandler][Callback]Task faild base information is streamID:[1
0], taskID:[6], taskType[Notify Wait], tag[HCCL@hdcube_6629421139219749105_12].
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 17:08:59.628.814 [task.cc:2994][218887] PrintErrorInfo:[EXEC][EXEC]Notify wait execute failed, device_id=7, stream_id=16, task_id=19, flip_num=0, notifi
y_id=1.
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 17:08:59.629.950 [task_exception_handler.cc:194][218893][218887][EXEC][EXEC][TaskExceptionHandler][Callback]Task faild base information is streamID:[1
0], taskID:[19], taskType[Notify Wait], tag[HCCL@hdcube_6629421139219749105_12].
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 16:59:14.658.409 [engine.cc:967][218882] ReportExecProc:[EXEC][EXEC]Task exception! device_id=2, stream_id=1, task_id=3, retcode=0x01.
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 16:59:14.660.332 [task.cc:2720][218882] PrintErrorInfo:[EXEC][EXEC]Task kernel execute failed, device_id=2, stream_id=1, task_id=3, flip_num=0,
fault so name: fault kernel_name: fault op name:IteratorV2IteratorGetNext, extend_info:[info type=4, info len=25, msg info:IteratorV2IteratorGetNext].
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 16:59:14.661.631 [log.cc:281][218882] StreamSynchronize:[EXEC][EXEC]Stream synchronize failed, stream = 0x0000002790.
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 16:59:14.661.338 [task.cc:2720][218882] PrintErrorInfo:[EXEC][EXEC]model execute task failed, device_id=2, model stream_id=1, model task_id=63, flip_n
um=0, model_id=64, first task_id=65.
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 16:59:14.661.500 [stream.cc:730][218882] GetError:[EXEC][EXEC]Stream Synchronize failed, stream_id=1 retcode=0x01, [the model stream execute failed].
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 16:59:14.661.631 [log.cc:281][218882] StreamSynchronize:[EXEC][EXEC]Stream synchronize failed, stream = 0x0000002790.
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 16:59:14.661.635 [api.cc:610][218882] rtsyncStreamSynchronize:[EXEC][EXEC]Stream synchronize failed, InnerCqeId=0x7150650.
log-218874_2022021515550029.log:4492:[ERROR] HCCL[218874.pytho3]:2022-02-15 16:59:14.661.660 [error_message_manage.cc:40][218882] FuncErrorReason:[EXEC][EXEC]rtStreamSynchronize execute failed, reason[the model stream execute
```

可能原因

HCCL算子的task会在指定集群的每个Device上执行，并通过notify进行状态同步，若任何一个rank或者通信链路在执行前/中发生异常，则会导致集群同步失败，剩余卡会出现notify wait timeout。常见的原因主要有：

1. 部分卡被某些耗时较长的任务阻塞，在超过600秒（可通过HCCL_EXEC_TIMEOUT配置）后才执行到对应阶段。
2. 部分rank未执行到notify同步阶段。
3. 网络模型等原因导致某些rank间的task执行序列不一致。
4. 节点间通信链路不稳定。
5. Pytorch场景，用户设置的HCCL_EXEC_TIMEOUT小于HCCL_CONNECT_TIMEOUT时，建链失败，先报notify wait超时。

处理步骤

收集所有卡的plog日志后，按以下步骤排查：

- 步骤1** 检查所有卡的报错情况，若有卡未报notify超时错误，请通过训练或者plog日志检查此卡是否存在业务进程报错退出、卡死或core宕机的情况。

步骤2 若所有卡均上报notify超时错误, 则检查错误日志中最早报错和最晚报错时间差是否超过算子执行超时阈值, 若超过阈值请定位报错时间最晚的rank执行阻塞原因(如save checkpoint)或通过export HCCL_EXEC_TIMEOUT=3600(示例, 请根据需要配置)调整超时阈值。

步骤3 检查集群中是否存在Device网口通信链路不稳定的情况, 搜索所有卡的Device侧日志, 若存在error cq的打印, 时间位于业务区间内, 这种情况可能是发生网络拥塞, 需要排查交换机配置是否合理、训练过程中是否有网口link down等情况。

搜索命令样例: `grep -rn "err cq" | grep HCCL`

```
19907:[ERROR] HCCL[85111.python3]:2023-02-18 09:44:55.431.692 [heartbeat.cc:547][85111][94635][Heartbeat]cq err status[12], time:[2023-02-18 09:44:55.369944], ip:[192.0.1.7]
```

----结束

3.5.2.2 socket 建链超时 (EI0006)

适用场景

- 业务场景: 集群训练
- 适用处理器: 昇腾910 AI处理器
- 处理器形态: EP

现象描述

常见于算子加载阶段, 有以下2种情况:

1、Server间的建链超时现象, 日志报错信息如下:

```
[ERROR] HCCL[147373.python3]:2022-09-29 15:31:05.319.815 [comm_base.cc:934][147373][1482272][LOAD][LOAD] LINK_ERROR_INFO
[ERROR] HCCL[147373.python3]:2022-09-29 15:31:05.319.823 [comm_base.cc:935][147373][1482272][LOAD][LOAD] comm error, device[6] num[1]
[ERROR] HCCL[147373.python3]:2022-09-29 15:31:05.319.827 [comm_base.cc:936][147373][1482272][LOAD][LOAD] dest_ip(rank_id) src_ip(rank_id) Role Status
[ERROR] HCCL[147373.python3]:2022-09-29 15:31:05.319.831 [comm_base.cc:937][147373][1482272][LOAD][LOAD] .....
[ERROR] HCCL[147373.python3]:2022-09-29 15:31:05.319.848 [comm_base.cc:950][147373][1482272][LOAD][LOAD] 192.2.1.20(14) 192.2.1.16(6) server no connect
[ERROR] HCCL[147373.python3]:2022-09-29 15:31:05.319.853 [comm_base.cc:951][147373][1482272][LOAD][LOAD] the connection failure between this device and target device may be due to the following reasons:
[ERROR] HCCL[147373.python3]:2022-09-29 15:31:05.319.861 [comm_base.cc:952][147373][1482272][LOAD][LOAD] the connection between this device and the target device is abnormal.
```

2、Server内的建链超时现象, 日志报错信息如下:

```
[ERROR] HCCL[1869739.python3]:2022-09-27 15:57:37.497.793 [exchanger_network.cc:695][1869739][1880385][LOAD][LOAD] device[6] userrank[0] exchanger Status: run_step[1]
[ERROR] HCCL[1869739.python3]:2022-09-27 15:57:37.497.796 [exchanger_network.cc:696][1869739][1880385][LOAD][LOAD] dest_dev userrank Role connStatus
[ERROR] HCCL[1869739.python3]:2022-09-27 15:57:37.497.801 [exchanger_network.cc:697][1869739][1880385][LOAD][LOAD] .....
[ERROR] HCCL[1869739.python3]:2022-09-27 15:57:37.497.806 [exchanger_network.cc:716][1869739][1880385][LOAD][LOAD] 0 0 NA NA
[ERROR] HCCL[1869739.python3]:2022-09-27 15:57:37.497.810 [exchanger_network.cc:716][1869739][1880385][LOAD][LOAD] 1 1 server NO
[ERROR] HCCL[1869739.python3]:2022-09-27 15:57:37.497.814 [exchanger_network.cc:716][1869739][1880385][LOAD][LOAD] 2 2 client NO
[ERROR] HCCL[1869739.python3]:2022-09-27 15:57:37.497.817 [exchanger_network.cc:716][1869739][1880385][LOAD][LOAD] 3 3 server NO
[ERROR] HCCL[1869739.python3]:2022-09-27 15:57:37.497.821 [exchanger_network.cc:716][1869739][1880385][LOAD][LOAD] 4 4 client NO
[ERROR] HCCL[1869739.python3]:2022-09-27 15:57:37.497.824 [exchanger_network.cc:716][1869739][1880385][LOAD][LOAD] 5 5 server NO
[ERROR] HCCL[1869739.python3]:2022-09-27 15:57:37.497.827 [exchanger_network.cc:716][1869739][1880385][LOAD][LOAD] 6 6 client NO
[ERROR] HCCL[1869739.python3]:2022-09-27 15:57:37.497.830 [exchanger_network.cc:716][1869739][1880385][LOAD][LOAD] 7 7 server NO
[ERROR] HCCL[1869739.python3]:2022-09-27 15:57:37.497.834 [exchanger_network.cc:721][1869739][1880385][LOAD][LOAD] the connection failure between this device and the target device may be due to the following reasons:
[ERROR] HCCL[1869739.python3]:2022-09-27 15:57:37.497.841 [exchanger_network.cc:723][1869739][1880385][LOAD][LOAD] the connection between this device and the target device is abnormal.
```

可能原因

HCCL会在指定集群的每个Device上运行, 并在集群间建立socket链接, 若任一个rank或者通信链路在建链前/中发生异常, 则会导致集群建链失败。常见的原因包括:

1. 部分rank未执行到正确的建链阶段。
2. 部分卡被某些耗时较长的任务阻塞, 在超过120秒(可通过HCCL_CONNECT_TIMEOUT配置)后才执行到对应阶段。
3. 脚本等原因导致某些rank间的通信算子数量或者排序不一致。
4. 节点间通信链路不通或者不稳定。

处理步骤

收集所有卡的plog日志后, 按以下步骤排查:

- 步骤1** 检查所有卡的报错情况，若有卡未报建链超时错误，请通过训练或者plog日志检查此卡是否存在业务卡死情况，若存在请定位根因。
- 步骤2** 若所有卡均上报建链超时错误，则检查错误日志中最早和最晚时间差异是否超过超时阈值，若超过阈值请定位最晚报错的卡执行较慢的原因或通过增大建链超时时间尝试解决（例如通过配置export HCCL_CONNECT_TIMEOUT=1800将超时时间增大为1800s）。
- 步骤3** 若时间差异未超过阈值，请检查各rank执行的集合通信算子数量一致，算子加载的顺序一致。HCCL要求一个通信域内各个rank上算子加载顺序保持一致。

如需查询tag和计算图中node name的对应关系，需开启INFO日志，在host日志中搜索：

关键字1：GenerateOpTag:graph

关键字2：GenerateTask:graph

关键字1可以看到node的hash值，关键字2可以看到nodHash对应的NodeName，协助计算图排查。

- 步骤4** 检查集群中是否存在Device网口通信链路不通的情况，比较常见的原因：

1. IP不在同一网段或子网掩码配置存在问题。
2. IP冲突，集群中存在IP相同的两个rank。
3. 交换机配置不在同一个vlan。
4. 链路不通。
5. 各rank的TLS(安全增强)设置不一致时也会导致建链失败。

可通过hccn tool命令来确认TLS是否一致，如不一致请参考随软件包发布的《HCCN Tool 接口参考》进行TLS配置。

查询TLS状态命令：

```
hccn_tool -i 0 -tls -g
hccn_tool -i 1 -tls -g
hccn_tool -i 2 -tls -g
hccn_tool -i 3 -tls -g
hccn_tool -i 4 -tls -g
hccn_tool -i 5 -tls -g
hccn_tool -i 6 -tls -g
hccn_tool -i 7 -tls -g
```

TLS switch值为0表示关闭，1表示开启。如果提示no certificate found，也表示TLS功能关闭。如果各个rank的TLS情况不一致，可根据hccn tool使用指南配置TLS，或者关闭所有rank的TLS功能。

关闭TLS功能命令：

```
hccn_tool -i 0 -tls -s enable 0
hccn_tool -i 1 -tls -s enable 0
hccn_tool -i 2 -tls -s enable 0
hccn_tool -i 3 -tls -s enable 0
hccn_tool -i 4 -tls -s enable 0
hccn_tool -i 5 -tls -s enable 0
hccn_tool -i 6 -tls -s enable 0
hccn_tool -i 7 -tls -s enable 0
```

----结束

3.5.2.3 Get P2P status 超时

适用场景

- 业务场景：集群训练
- 适用处理器：昇腾910 AI处理器
- 处理器形态：EP

现象描述

常见于算子编译阶段，日志报错信息：Wait timeout for get P2P status

```
root@centos host-0]# grep ERROR host-0-20200916094117004.log
ERROR] HCCL(40469,python3.7):2020-09-16-09:43:19.508.988 [hccl/hccl/hccl_comm/impl/hccl_impl_base.cc:344][hccl-40469-0-1600220478-hccl_world_group][5]errNo[0x000000000500010] wait timeout for get P2P status, LocalId[5], PhysId[0] timeout[120]
ERROR] HCCL(40469,python3.7):2020-09-16-09:43:19.509.052 [hccl/hccl/hccl_comm/impl/hccl_impl_base.cc:210][hccl-40469-0-1600220478-hccl_world_group][5]index[0], check P2P vnic status failed
ERROR] HCCL(40469,python3.7):2020-09-16-09:43:19.509.087 [hccl/hccl/hccl_comm/impl/hccl_impl.cc:612][hccl-40469-0-1600220478-hccl_world_group][5]errNo[0x000000000500001] pre resource initialization failed
ERROR] HCCL(40469,python3.7):2020-09-16-09:43:19.509.110 [hccl/hccl/hccl_comm/hccl_comm.cc:92][hccl-40469-0-1600220478-hccl_world_group][5]errNo[0x000000000500010] hccl_initialize failed
ERROR] HCCL(40469,python3.7):2020-09-16-09:43:19.509.137 [hccl/hccl/hcom/src/hcom.cc:79][hccl-40469-0-1600220478-hccl_world_group][5]errNo[0x0000000005010010] hcclComm init error
ERROR] HCCL(40469,python3.7):2020-09-16-09:43:19.509.158 [hccl/hccl/hcom/src/hcom.cc:98][hccl-40469-0-1600220478-hccl_world_group][5]hcom init failed, return[83951632]
ERROR] HCCL(40469,python3.7):2020-09-16-09:43:19.509.492 [hccl/hcom_graph_adaptor/ge_plugin/plugin_manager.cc:226][hccl-40469-0-1600220478-hccl_world_group][5]errNo[0x0000000005010010] Initialize: hcom init failed
ERROR] HCCL(40469,python3.7):2020-09-16-09:43:19.509.515 [hccl/hcom_graph_adaptor/ge_plugin/plugin_manager.cc:183][hccl-40469-0-1600220478-hccl_world_group][5]Initialize: Initializehcom failed
ERROR] GE(40469,python3.7):2020-09-16-09:43:19.509.547 [framework/domi/opskernel_manager/ops_kernel_manager.cc:86]40469 Initialize: ErrorNo: 1343250441(There is no valid so about OpsKernelInfoStore or GraphOptimizer.)
ERROR] GE(40469,python3.7):2020-09-16-09:43:19.509.658 [framework/domi/init/gelib.cc:120]40469 InnerInitialize: ErrorNo: 1343250441(There is no valid so about OpsKernelInfoStore or GraphOptimizer.)
ERROR] GE(40469,python3.7):2020-09-16-09:43:19.509.731 [framework/domi/init/gelib.cc:80]40469 Initialize: ErrorNo: 1343250441(There is no valid so about OpsKernelInfoStore or GraphOptimizer.) Gelib initial failed.
ERROR] GE(40469,python3.7):2020-09-16-09:43:19.518.244 [framework/domi/client/ge_api.cc:116]40469 GEInitialize: ErrorNo: 1343229953(GEInitialize Failed.) geInitialize failed
error code = 1343229953
root@centos host-0]# cd ../
root@centos slog]# ll
total 0
```

可能原因

HCCL初始化会在AI Server内的各rank(Device)间建立P2P使能链接，若任一个rank在建链前/中发生异常，未拉起业务进程，则会导致本AI server初始化失败。

处理步骤

场景1 部分rank在P2P使能前异常时，会导致和其他rank间的P2P使能超时，处理步骤为：

1. 确认各rank上HCCL初始化是否成功（plog日志关键字为hcom init by file success，级别为TRACE），若未成功初始化则可能原因为：
 - （1）进入HCCL初始化之前有其他组件初始化失败
 - （2）上次训练过程存在进程残留
2. P2P使能前其他组件异常，导致没有走到HCCL初始化(典型问题：TSD没有拉起失败)。
3. 训练脚本有问题，训练拉起进程/docker数是否和训练卡数不一致，导致实际拉起的进程和集群描述不符。
4. TDT拉起失败/磁盘空间不足等，导致进程异常退出。

场景2 各rank上HCCL的调用时差超过120s，引起P2P使能超时：

确认各rank的通信初始化打印日志（plog日志关键字为Entry-HcomInitByFile，级别为EVENT）的时间，是否相差超过120s，若超过则为集合通信拉起时间相差过大。

常见场景：训练启动脚本有问题，多卡训练时应同时运行多个进程（每张卡对应一个进程），而不是一个进程结束后再启动下一个进程。

#典型错误

```
for i in ${DEVICE_LIST}
do
```

```
export DEVICE_ID=$i
export RANK_ID=$i
echo "start train ing device $i"
python3.7.5 /home/HwHiAiUser/train.py      # 阻塞执行
done
```

```
# -----
```

解决方法

```
for i in ${DEVICE_LIST}
do
    export DEVICE_ID=$i
    export RANK_ID=$i
    echo "start train ing device $i"
    python3.7.5 /home/HwHiAiUser/train.py &      # 后台执行
done
```

场景3 配置的rank_table和实际使用的Device不一致。

在日志（EVENT级别）中 grep 关键字 paraInfo(hcom init)(下沉模式) 或者 paraInfo(HcclCommInitClusterInfo)(单算子模式)，打开日志中打印的rank_table文件，确认该server上启动的rank数是否和rank_table文件上配置的一致。

例如，实际使用2P，但是rank_table是按照8P进行配置的，HCCL在初始化时会出现该现象。

3.5.2.4 HCCP ra 初始化失败，返回-17

适用场景

- 业务场景：集群训练
- 适用处理器：昇腾910 AI处理器
- 处理器形态：EP

现象描述

HCCL初始化网卡失败，HCCP报-17错误码：ra_rdev_init failed, ret [-17]

```
pool[2] device[0] rank[2] hccp[2] hccp[2] [2020-11-04-16:43:46.350.621] [drivers/network/hccp/rdma_agent/hdc/ra_hdc.c:223] ra_hdc_process_msg(223) : message head
check fail, ret[-17]
[ERROR] HCCP(83258.python):2020-11-04-16:43:46.350.633 [drivers/network/hccp/rdma_agent/hdc/ra_hdc.c:1166] ra_hdc_rdev_init(1166) : ra_hdc message
process failed, ret[-17]
[ERROR] HCCP(83258.python):2020-11-04-16:43:46.350.642 [drivers/network/hccp/rdma_agent/client/ra_host.c:316] ra_rdev_init(316) : ra_rdev init fail
ed, ret[-17]
[ERROR] HCCL(83258.python):2020-11-04-16:43:46.350.652 [hccl/hccl/hccl/hccl_comm/misc/network_manager/network_manager.cc:115] [hccl-83258-0-1604479
426-hccl_world_group][0]errNo[0x0000000050000002] ptr [rdmaHandle] is NULL, return HCCL_E_PTR
[ERROR] HCCL(83258.python):2020-11-04-16:43:46.350.660 [hccl/hccl/hccl/hccl_comm/impl/hccl_impl_base.cc:250] [hccl-83258-0-1604479426-hccl_world_gr
oup][0]start nic ipaddr[655cae80] failed
[ERROR] HCCL(83258.python):2020-11-04-16:43:46.350.669 [hccl/hccl/hccl/hccl_comm/impl/hccl_impl.cc:583] [hccl-83258-0-1604479426-hccl_world_group][
0]errNo[0x0000000050000002] ra resource initialization failed
[ERROR] HCCL(83258.python):2020-11-04-16:43:46.350.677 [hccl/hccl/hccl/hccl_comm/hccl_comm.cc:92] [hccl-83258-0-1604479426-hccl_world_group][0]errN
o[0x0000000050000002] hccl initialize failed
[ERROR] HCCL(83258.python):2020-11-04-16:43:46.350.685 [hccl/hccl/hcom/src/hcom.cc:81] [hccl-83258-0-1604479426-hccl_world_group][0]errNo[0x0000000
050100002] hcclComm init error
[ERROR] HCCL(83258.python):2020-11-04-16:43:46.350.693 [hccl/hccl/hcom/src/hcom.cc:100] [hccl-83258-0-1604479426-hccl_world_group][0]hcom init fail
```

可能原因

HCCL在初始化时会根据rank_table中的Device IP初始化Device网卡。如果初始化使用的Device IP和实际网卡的IP不一致，HCCP会初始化网卡失败并返回错误码 -17。

处理步骤

- 步骤1** 确认该Device的rank id，并在ranktable中找到对应的device_ip配置，rankid获取方式：

在用户态Host日志（需打开EVENT日志）中，grep关键字Entry-HcomInit，其identify中内容即为rankid。

步骤2 确认该server的Device IP是否配置正确，若出现ranktable中device_ip配置和查询结果不一致的情况，请以查询结果为准，并修改对应rank的ranktable的"device_ip"字段。

使用 hccn_tool 可查看Device 网卡信息。

```
hccn_tool -i 0 -ip -g
hccn_tool -i 1 -ip -g
hccn_tool -i 2 -ip -g
hccn_tool -i 3 -ip -g
hccn_tool -i 4 -ip -g
hccn_tool -i 5 -ip -g
hccn_tool -i 6 -ip -g
hccn_tool -i 7 -ip -g
或
for i in {0..7}; do hccn_tool -i $i -ip -g ; done
```

----结束

3.5.2.5 CANN 版本不一致

适用场景

- 业务场景：集群训练
- 适用处理器：昇腾910 AI处理器
- 处理器形态：EP

现象描述

常见于通信建链阶段，屏显日志报EI0008错误码：The CANN versions are inconsistent.

```
ERROR : GeOp5_0GEOP::DoRunAsync Failed
Error Message is :
EI0008: The CANN versions are inconsistent: tag [HcomAllReduce_6629421139219749105_0], local_version [1.83.T10.0.B206], remote_version [1.83.T10.0.B207]
Solution: Install the same CANN version.
Traceback (most recent call last):
Call ops_kernel_info_store LoadTask fail[FUNC:Distribute][FILE:hccn_task_info.cc][LINE:213]
```

可能原因

集合通信建链时会校验本端与对端Rank的CANN版本一致性，如果CANN版本不一致，HCCL会返回错误并打印错误码EI0008。

处理步骤

确认不同服务器安装的CANN runtime包版本是否一致，若不一致则需要安装一致的版本。

步骤1 在plog日志（INFO级别）中grep关键字CannVersion，确认各节点的CANN runtime包版本。

步骤2 重新安装一致的CANN runtime包版本。

----结束

3.6 getnext 故障处理

3.6.1 getnext 算子超时

适用场景

- 业务场景：TensorFlow/Mindspore训练
- 适用处理器：昇腾910 AI处理器、昇腾910B AI处理器
- 处理器形态：EP

现象描述

执行训练脚本时出现getnext算子超时。

```
Error Message is :
E30008: AI CPU operator execution time out.
Possible Cause: 1. For the GetNext operator, its preprocessing duration may be too long. 2. For a custom operator, its logic may be improper.
Solution: 1. For the GetNext operator, check its preprocessing or set OpExecuteTimeOut to a larger value. 2. For a custom operator, make sure its logic is proper.
TraceBack (most recent call last):
Aicpu kernel execute failed, device_id=0, stream_id=2, task_id=2, fault
op_name=aicpu_getnext_iteratorGetNext[FUNC:GetError][FILE:stream.cc][LINE:1133]
rtStreamSynchronizeWithTimeout execute failed, reason=[aicpu timeout][FUNC:FuncErrorReason]
[FILE:error_message_manage.cc][LINE:49]

[[{{node GeOp2_0}}]]
```

可能原因

可通过查询Device日志定界，日志中能记录getnext超时后Device侧驱动队列中的出队/入队相关信息：

```
debug/device-0/device-1927_20230406202333033.log:147:[ERROR] AICPU(3480,aicpu_scheduler):
2023-04-06-20:24:34.559.758 [kernel_util.cc:101][AICPU][operator():101][tid:3543]:device_id:0, queue_name:
12658048656348665736, queue_id:1, size:0, depth:2, status:1, workMode:1, type:2, enqueueCnt:0, dequeCnt=0,
enqueueFailCnt=0, dequeFailCnt=1, enqueueEventOk=0, enqueueEventFail=0, FullToNotFullEventOkCnt=0,
FullToNotFullEventFailCnt = 0, lastEnqueueTime.tv_sec:0, lastEnqueueTime.tv_usec:0, lastDequeueTime.tv_sec:0,
lastDequeueTime.tv_usec:0
```

可通过入队/出队信息确认数据集是否正常发送到了Device侧，如果入队数量很少，则可能是数据集生成不稳定或者数据集传输网络不稳定或者预处理阶段耗时较大。

处理步骤

针对以上可能原因，可参考以下步骤处理：

- 步骤1** 检查训练模型的输入数据集是否正常生成以及传输是否稳定。
- 步骤2** 检查host侧预处理过程处理逻辑是否存在耗时较大情况（数据集正常情况下，getnext超时后aicpu记录的出队/入队统计中enqueue数量较少或者lastEnqueueTime时间比较大说明预处理阶段耗时大），如果确认预处理阶段耗时较长，可通过OpExecuteTimeOut接口修改算子超时时间更长（默认超时时间28s）。

----结束

3.6.2 mbuf 入队场景出现驱动报错

适用场景

- 业务场景：TensorFlow/Mindspore训练
- 适用处理器：昇腾910 AI处理器、昇腾910B AI处理器
- 处理器形态：EP

现象描述

执行训练模型脚本过程中入队阶段出现如下报错：

- **报错现象一：**
Error Message is
EL9999: Inner Error!
EL9999 [drv api] halQueueEnQueueBuff failed: deviceId=0, qid=1, timeout=3000, drvRetCode=58.
[FUNC:MemQueueEnQueueBuff][FILE:npu_driver.cc][LINE:3205]
TraceBack (most recent call last):
rtMemQueueEnQueueBuff execute failed, reason=[invalid value][FUNC:FuncErrorReason]
[FILE:error_message_manage.cc][LINE:49]
Fail to execute acltdtSendTensor, device is 0, name is 17125841410799885412[FUNC:ReportInnerError]
[FILE:log_inner.cpp][LINE:145]
- **报错现象二：**
Error Message is
EL9999: Inner Error!
EL9999 [drv api] halQueueEnQueueBuff failed: deviceId=0, qid=1, timeout=3000, drvRetCode=27.
[FUNC:MemQueueEnQueueBuff][FILE:npu_driver.cc][LINE:3205]
TraceBack (most recent call last):
rtMemQueueEnQueueBuff execute failed, reason=[hdc send msg fail][FUNC:FuncErrorReason]
[FILE:error_message_manage.cc][LINE:49]
Fail to execute acltdtSendTensor, device is 0, name is 4174979144421111244[FUNC:ReportInnerError]
[FILE:log_inner.cpp][LINE:145]
- **报错现象三：**
Error Message is
EL9999: Inner Error!
EL9999 [drv api] halQueueEnQueueBuff failed: deviceId=0, qid=1, timeout=3000, drvRetCode=74.
[FUNC:MemQueueEnQueueBuff][FILE:npu_driver.cc][LINE:3190]
TraceBack (most recent call last):
rtMemQueueEnQueueBuff execute failed, reason=[driver error:copy data fail]
[FUNC:FuncErrorReason][FILE:error_message_manage.cc][LINE:49]
Fail to execute acltdtSendTensor, device is 0, name is 12934931840515960683[FUNC:ReportInnerError]
[FILE:log_inner.cpp][LINE:145]
- **报错现象四：**
Error Message is
EL9999: Inner Error!
EL9999 [drv api] halQueueEnQueueBuff failed: deviceId=0, qid=1, timeout=3000, drvRetCode=16.
[FUNC:MemQueueEnQueueBuff][FILE:npu_driver.cc][LINE:3205]
TraceBack (most recent call last):
rtMemQueueEnQueueBuff execute failed, reason=[report timeout][FUNC:FuncErrorReason]
[FILE:error_message_manage.cc][LINE:49]
Fail to execute acltdtSendTensor, device is 0, name is 16561572617881319536[FUNC:ReportInnerError]
[FILE:log_inner.cpp][LINE:145]

可能原因

以上报错场景，可能的原因如下：

- **报错现象一：可能为驱动入队地址非法。对应plog如下：**
[ERROR] DRV(3078,python3):2023-04-27-20:04:09.700.666 [ascend][curpid: 3078, 3420][drv]
[queuemng][QueueIoctl 133]Ioctl failed. (cmd=40585102; error=0; ret=58)
[ERROR] DRV(3078,python3):2023-04-27-20:04:09.700.674 [ascend][curpid: 3078, 3420][drv]

```
[queuemng][QueueSubmitEventSync 724]enqueue ioctl failed. (ret=58; event_id=4; gid=20; tid=0;
timeout=5000ms; subevent_id=1).
[ERROR] DRV(3078,python3):2023-04-27-20:04:09.700.677 [ascend][curpid: 3078, 3420][drv]
[queuemng][QueueSendQueueEventSyncTimeout 871]Submit event failed. (ret=58; devId=0; qid=1)
[ERROR] RUNTIME(3078,python3):2023-04-27-20:04:09.700.686 [npu_driver.cc:3205]3420
MemQueueEnQueueBuff:report error module_type=1, module_name=EL9999
[ERROR] RUNTIME(3078,python3):2023-04-27-20:04:09.700.688 [npu_driver.cc:3205]3420
MemQueueEnQueueBuff:[drv api] halQueueEnQueueBuff failed: deviceId=0, qid=1, timeout=3000,
drvRetCode=58.
[ERROR] RUNTIME(3078,python3):2023-04-27-20:04:09.700.721 [api_c.cc:3738]3420
rtMemQueueEnQueueBuff:ErrCode=107000, desc=[invalid value], InnerCode=0x7020025
```

- 报错现象二：可能为HDC通道断链导致Host数据发送异常。对应plog记录如下：
[ERROR] DRV(253915,python3):2023-05-16-22:35:00.917.147 [ascend][curpid: 253915, 254303][drv]
[queuemng][QueueIoctl 133]ioctl failed. (cmd=40585102; error=0; ret=27)
[ERROR] DRV(253915,python3):2023-05-16-22:35:00.917.161 [ascend][curpid: 253915, 254303][drv]
[queuemng][QueueSubmitEventSync 724]enqueue ioctl failed. (ret=27; event_id=4; gid=20; tid=0;
timeout=5000ms; subevent_id=90).
[ERROR] DRV(253915,python3):2023-05-16-22:35:00.917.174 [ascend][curpid: 253915, 254303][drv]
[queuemng][QueueSendQueueEventSyncTimeout 871]Submit event failed. (ret=27; devId=0; qid=1)
[ERROR] RUNTIME(253915,python3):2023-05-16-22:35:00.917.197 [npu_driver.cc:3205]254303
MemQueueEnQueueBuff:report error module_type=1, module_name=EL9999
[ERROR] RUNTIME(253915,python3):2023-05-16-22:35:00.917.209 [npu_driver.cc:3205]254303
MemQueueEnQueueBuff:[drv api] halQueueEnQueueBuff failed: deviceId=0, qid=1, timeout=3000,
drvRetCode=27.
[ERROR] RUNTIME(253915,python3):2023-05-16-22:35:00.917.347 [api_c.cc:3783]254303
rtMemQueueEnQueueBuff:ErrCode=507051, desc=[hdc send msg fail], InnerCode=0x7110013
- 报错现象三：可能为PCIE链路异常导致发起DMA拷贝失败。对应plog记录如下：
[ERROR] RUNTIME(179263,python3):2023-05-17-09:35:01.035.824 [npu_driver.cc:3205]179650
MemQueueEnQueueBuff:report error module_type=1, module_name=EL9999
[ERROR] RUNTIME(179263,python3):2023-05-17-09:35:01.035.946 [npu_driver.cc:3205]179650
MemQueueEnQueueBuff:[drv api] halQueueEnQueueBuff failed: deviceId=0, qid=1, timeout=3000,
drvRetCode=74.
[ERROR] RUNTIME(179263,python3):2023-05-17-09:35:01.035.988 [api_c.cc:3783]179650
rtMemQueueEnQueueBuff:ErrCode=507052, desc=[driver error:copy data fail], InnerCode=0x7020024
[ERROR] RUNTIME(179263,python3):2023-05-17-09:35:01.035.993 [error_message_manage.cc:
49]179650 FuncErrorReason:report error module_type=3, module_name=EE8888
[ERROR] RUNTIME(179263,python3):2023-05-17-09:35:01.035.998 [error_message_manage.cc:
49]179650 FuncErrorReason:rtMemQueueEnQueueBuff execute failed, reason=[driver error:copy data
fail]
- 报错现象四：可能为DEVICE侧CP进程异常导致H2D拷贝失败。对应plog如下：
[ERROR] DRV(2622,python3):2023-05-17-10:05:11.057.003 [ascend][curpid: 2622, 3009][drv]
[queuemng][QueueIoctl 133]ioctl failed. (cmd=40585102; error=0; ret=16)
[ERROR] DRV(2622,python3):2023-05-17-10:05:11.057.125 [ascend][curpid: 2622, 3009][drv]
[queuemng][QueueSubmitEventSync 724]enqueue ioctl failed. (ret=16; event_id=4; gid=20; tid=0;
timeout=5000ms; subevent_id=1).
[ERROR] DRV(2622,python3):2023-05-17-10:05:11.057.132 [ascend][curpid: 2622, 3009][drv]
[queuemng][QueueSendQueueEventSyncTimeout 871]Submit event failed. (ret=16; devId=0; qid=1)
[ERROR] RUNTIME(2622,python3):2023-05-17-10:05:11.057.143 [npu_driver.cc:3205]3009
MemQueueEnQueueBuff:report error module_type=1, module_name=EL9999
[ERROR] RUNTIME(2622,python3):2023-05-17-10:05:11.057.147 [npu_driver.cc:3205]3009
MemQueueEnQueueBuff:[drv api] halQueueEnQueueBuff failed: deviceId=0, qid=1, timeout=3000,
drvRetCode=16.
[ERROR] RUNTIME(2622,python3):2023-05-17-10:05:11.057.187 [api_c.cc:3783]3009
rtMemQueueEnQueueBuff:ErrCode=507012, desc=[report timeout], InnerCode=0x711000c

处理步骤

该类型错误需要联系华为Driver开发工程师定位排查。您可以通过<https://gitee.com/ascend>网站提交issue获取帮助。

3.6.3 mbuf 入队场景出现 OOM 报错

适用场景

- 业务场景：TensorFlow/Mindspore训练
- 适用处理器：昇腾910 AI处理器、昇腾910B AI处理器
- 处理器形态：EP

现象描述

执行训练模型脚本过程中出现如下报错：

```
Error Message is
EL0004: Failed to allocate memory.
Possible Cause: Available memory is insufficient.
Solution: Close applications not in use.
TraceBack (most recent call last):
rtMemQueueEnQueueBuff execute failed, reason=[driver error:out of memory][FUNC:FuncErrorReason]
[FILE:error_message_manage.cc][LINE:49]
Fail to execute acltdtSendTensor, device is 0, name is 9238275185956894463[FUNC:ReportInnerError]
[FILE:log_inner.cpp][LINE:145]
```

可能原因

其他业务同步执行导致可用内存不足。

处理步骤

针对内存不足情况，可以参考以下步骤处理：

- 步骤1** 检查环境上业务同步执行情况，关闭其他暂无关业务。
- 步骤2** 检查脚本中内存申请相关设置，确认是否batch_size等参数导致需要申请的内存过大导致申请失败，可适当调整参数。

----结束

4 故障定位工具

[4.1 故障诊断工具（一键收集）](#)

[4.2 集群故障日志分析](#)

[4.3 集群故障日志分析（云环境）](#)

4.1 故障诊断工具（一键收集）

4.1.1 工具介绍

为提高系统故障维测效率，提供故障诊断工具一键式收集故障信息：

- 信息收集：一次性收集进程故障现场信息，为故障定位效率提升提供有效输入。
- 业务复跑：一次性复跑业务并收集故障信息，提高故障问题复现以及收集信息效率。

工具可以收集以下信息：

- 软件信息：涉及软件包版本，环境变量，软件依赖，系统信息。
- Host侧CANN软件栈日志。
- Host侧message日志。
- Device侧固件日志：device-*日志（需root权限）。
- Device侧系统日志：message日志，device-os日志（需root权限）。
- 黑匣子、Stackcore文件（需root权限）。
- 任务打屏日志。
- run包安装日志（需run包安装帐号与用例执行帐号一致才可收集）。
- GE dump图。
- TF Adapter dump图。
- 发生Aicore err时生成的dump文件。
- 算子编译.o, .json文件。
- 用户用例执行的命令信息。

4.1.2 工具使用方法

使用约束

1. 不支持原有执行脚本内部直接后台执行的方式。
如：原有用例通过命令行执行 `sh cmd.sh`来拉起用例，而`cmd.sh`的实现里执行 `python3 test.py &`，用后台的方式执行，此种用例由于无法感知结束点，暂不支持使用。
2. 相同用户、相同时间段内，同机器同时作业时，收集到的数据会有交叉。
3. 非root用户，获取到的数据范围会受限，具体限制参考[工具介绍](#)权限要求。
4. 工具支持形态：EP场景(不包括昇腾310 AI处理器)。
5. 集群、容器、虚拟机、云场景不支持一键式工具收集故障信息。

工具使用

脚本工具存放在Toolkit包安装路径`{install_path}/ascend_system_advisor`下，针对不同的信息收集场景，分别提供命令行指令，命令行格式如下。

- 信息收集：

```
python3 {install_path}/tools/ascend_system_advisor/asys/asys.py collect  
[--task_dir="path1"] [--output="path2"]
```

📖 说明

也可以通过`source ${install_path}/latest/bin/setenv.bash`命令使环境变量生效，然后直接输入`asys`命令符，不用输入脚本全路径。

`task_dir`参数是可选参数，为任务执行产生的算子编译.o、.json文件和dump图的搜索根目录，可指定为训练或推理任务脚本执行的目录，也可以指定为用户自定义的输出目录。若不指定`task_dir`参数，则不开启算子编译文件、GE dump图和TF Adapter dump图收集功能。

`output`参数是可选参数，其值作为`asys`工具的结果输出目录的路径前缀，即最终输出目录为`{output}/asys_output_timestamp`。命令行中不带`output`参数时，输出结果存放在命令行执行目录下；若`output`指定值为空、无效字符串、或指定路径目录无写权限、或创建目录失败，则`asys`工具退出执行并报错。

- 业务复跑：

```
python3 {install_path}/tools/ascend_system_advisor/asys/asys.py launch  
--task="sh ../app_run.sh" [--output="path"]
```

📖 说明

命令行中，`task`参数是必须参数，为复跑业务的执行命令，需要填写完整的命令，例如`"sh ../app_run.sh"`，`sh`为执行任务的工具，`../app_run.sh`为要执行的任务文件。

本工具会通过环境变量`NPU_COLLECT_PATH`指定的路径来收集dump图及算子编译.o文件；如果该环节变量未配置或不存在，则以`task`任务执行路径收集。

业务复跑默认开启算子编译文件、GE dump图和TF Adapter dump图收集功能。

执行完命令后，生成本次收集的信息文件，存放路径如下：

— asys_output_timestamp	
— software_info.txt	//保存安装包版本、环境变量、依赖软件、系统信息
— dfx	
— bbox	//保存Device侧的黑匣子信息
— data-dump	//发生Aicore err时，生成的dump文件
— graph	//保存用例执行时，生成的dump图信息，包含GE与TF Adapter的dump图

ops	//保存用例执行时, 生成的算子编译.o, .json文件信息
stackcore	//保存报错触发coredump时的core文件信息
log	
device	
firmware	//保存固件生成的日志
system	//保存常驻进程生成的日志
host	
message	//保存message/syslog日志
install	//保存包历史安装情况的日志
cann	//保存Host侧应用类日志。
screen.txt	//保存打屏日志。(如果内容为空, 则可能训练任务中设置了重定向。)
user_cmd	//保存用户用例执行的命令

进阶功能

1. 自定义配置软件信息software_info.txt中的依赖软件收集项。

{install_path}/tools/ascend_system_advisor/conf/dependent_package.csv

```
make,make --version
cmake,cmake --version
unzip,unzip -v
zlib1g,dpkg -l zlib1g| grep zlib1g| grep ii
zlib1g-dev,dpkg -l zlib1g-dev| grep zlib1g-dev| grep ii
libsqlite3-dev,dpkg -l libsqlite3-dev| grep libsqlite3-dev| grep ii
openssl,dpkg -l openssl| grep openssl| grep ii
libssl-dev,dpkg -l libssl-dev| grep libssl-dev| grep ii
libffi-dev,dpkg -l libffi-dev| grep libffi-dev| grep ii
```

说明

每行对应一个配置项, 逗号分割依赖项名字和查询指令, 逗号后无空格。

2. 配置业务复跑任务的算子编译文件和dump图收集功能开关, 以及配置默认环境变量值。

通过修改{install_path}/tools/ascend_system_advisor/conf/asys.ini文件中参数, 从而打开或关闭开启算子编译文件和dump图收集等功能。

```
[launch]
dump_ge_graph = 2          //控制dump图的内容多少, 对应环境变量DUMP_GE_GRAPH
dump_graph_level = 3       //控制dump图的个数, 对应环境变量DUMP_GRAPH_LEVEL
log_level = INFO           //应用类日志的全局日志级别及各模块日志级别, 对应环境变量
ASCEND_GLOBAL_LOG_LEVEL
log_event_enable = TRUE    //应用类日志是否开启Event日志, 对应环境变量
ASCEND_GLOBAL_EVENT_ENABLE
log_print_to_stdout = FALSE //是否开启日志打屏, 对应环境变量ASCEND_SLOG_PRINT_TO_STDOUT
```

说明

asys工具启动时, 环境变量默认使用该文件里的配置值, 但如果复跑任务的脚本中设置了这些环境变量为其他值, 则会发生环境变量覆盖, 以脚本中设置的值为准。因此, 可能造成收集到的维测信息不满足定位需求。

4.1.3 工具使用 FAQ

故障现象

先使用ctrl+z中止业务复跑task, 接着再次拉起业务复跑task, 屏显日志显示业务复跑task错误。如图4-1所示。

图 4-1 业务复跑报 task occurred error

```
2023-07-03 16:51:22,641 [ASYS] [INFO]: asys start.
2023-07-03 16:51:22,641 [ASYS] [INFO]: asys output directory: /npu/xxxxxxxx/infer/auto_infer/testcase/1971_aoe_ID2664_yolov5_bs64/Asys_
path/asys_output_20230703165122641
2023-07-03 16:51:22,641 [ASYS] [INFO]: launch task start, running:
/^Z
[1]+  Stopped                  python3 /usr/local/Ascend/latest/toolkit/tools/ascend_system_advisor/asys/asys.py launch --task="sh ./run.
sh" --output="./Asys_path"
root@huawei:/npu/xxxxxxxx/infer/auto_infer/testcase/1971_aoe_ID2664_yolov5_bs64# head 0703test.log
root@huawei:/npu/xxxxxxxx/infer/auto_infer/testcase/1971_aoe_ID2664_yolov5_bs64# asys launch --task="sh ./run.sh" --output="./Asys_path"

2023-07-03 16:51:48,600 [ASYS] [INFO]: asys start.
2023-07-03 16:51:48,600 [ASYS] [INFO]: asys output directory: /npu/xxxxxxxx/infer/auto_infer/testcase/1971_aoe_ID2664_yolov5_bs64/Asys_
path/asys_output_20230703165148600
2023-07-03 16:51:48,600 [ASYS] [INFO]: launch task start, running:
2023-07-03 16:54:58,889 [ASYS] [WARNING]: task occurred error, output:
Segmentation fault (core dumped)
Process ForkServerPoolWorker-7:
Traceback (most recent call last):
  File "/usr/local/python3.7.5/lib/python3.7/multiprocessing/pool.py", line 127, in worker
    put((job, i, result))
  File "/usr/local/python3.7.5/lib/python3.7/multiprocessing/queues.py", line 364, in put
    self._writer.send_bytes(obj)
  File "/usr/local/python3.7.5/lib/python3.7/multiprocessing/connection.py", line 200, in send_bytes
    self._send_bytes(m[offset:offset + size])
  File "/usr/local/python3.7.5/lib/python3.7/multiprocessing/connection.py", line 403, in _send_bytes
    self._send(header + buf)
  File "/usr/local/python3.7.5/lib/python3.7/multiprocessing/connection.py", line 367, in _send
    n = write(self._handle, buf)
BrokenPipeError: [Errno 32] Broken pipe
```

故障原因

执行ctrl+z操作等导致任务异常终止，但还存在任务进程残留（且还进行重定向写文件操作等操作），与后面新启动的asys复跑任务相互冲突，导致复跑异常。

处理方法

在asys复跑前查询是否存在运行中的推理/训练进程id，需要手工kill相关进程，然后再重新asys复跑。

4.2 集群故障日志分析

4.2.1 工具介绍

适用场景

集群训练任务失败时，面对大量的日志，用户较难快速锁定问题rank，影响集群问题定位效率。本工具会通过自动解析（HCCL组件的）日志，快速给出集群中未报错的/最早报错的rankid，并辅助定界/定位一些集群常见问题，通过打屏呈现给用户。适用于：

1. 判断是否有卡训练未拉起或异常退出导致集群训练失败，并给出问题的rankid。
2. 判断是否存在卡间异步时间过长引起集群互相等待超时导致集群训练失败，并给出最早/最晚拉起集合通信任务的rankid。
3. 判断是否存在常见的HCCL内task执行失败（SDMA计算溢出、memcpy访问越界）导致的集群训练失败。

环境准备

使用plog解析日志工具前，请确保环境已经安装以下模块：

已安装python软件，且使用的python版本为3.7.5及以上版本，ubuntu系统下可以使用apt-get install python3命令安装。

工具获取

plog解析工具存放在toolkit包中，安装后，plog日志解析工具会自动安装到\${HOME_ToolKit}/tools目录下，文件名称：parse_plog_hccl.py。

约束与限制

1. 使用前请确保所有卡日志齐全，并保留原始文件名称，不支持日志老化场景。
2. 所有卡日志需要放入同一个文件夹下，并保证不同的server的日志放在不同的子目录下。
3. 运行日志时使用的python版本必须要为Python 3.7.5及以上版本。

4.2.2 工具使用方法

命令行使用说明

命令行格式: python3 parse_plog_hccl.py *path*

Path参数说明:

1. 需要解析的plog日志所在的路径，并包含所有卡的日志文件。
2. path可以是绝对路径也可以是相对路径。

命令执行示例:

```
Python3 parse_plog_hccl.py /home/HwHiAiUser/modelarts-job-ec7b5289-e9fc-4fa6-9467-cbd3dad4eb9a
```

屏显结果分析

脚本执行过程中，会有对应的日志输出，主要输出的内容说明如下：

1.对于HCCL错误输出格式如下：

- 第一行：主要是错误的结果日志，表明错误是不是属于hccl组件，如果是，则会给出错误的原因。
- 第二行：打印最先报错的rankid及报错时间。
- 第三行：打印最晚报错的rankid及报错时间。
- 第四行：建链/notify超时时间阈值。
- 第五行：集群rank数量。
- 第六行及以后：打印最早报错的rank里面的详细的ERROR日志。

使用示例：

图 4-2 日志分析示例

```
root@huawei:/home/hua/# python3 parse_plog_hccl.py /home/hua/1/modelarts-job-34c75f48-58bf-4d94-b8ce-336b74d39d41/ascend/process_log/
2022-03-21 18:22:17,356 - INFO: The cause of this error is 'notify timeout', please turn to the relevant engineer to solve this problem.
2022-03-21 18:22:17,356 - INFO: The first rank to report the error is rank443 and time is 2022-03-14-17:25:46.996120
2022-03-21 18:22:17,356 - INFO: The last rank to report the error is rank728 and time is 2022-03-14-17:27:22.147111
2022-03-21 18:22:17,356 - INFO: The default timeout period is used.
2022-03-21 18:22:17,356 - INFO: Total of rankid is 768
2022-03-21 18:22:17,356 - INFO: Error-log of the first rank that reported the error:
=====detail log=====
[ERROR] RUNTIME(45282,python):2022-03-14-17:25:46.996.120 [engine.cc:964]45806 ReportExceptProcTask exception! device_id=3, stream_id=30, task_id=1, type=13, retCode=0x91.
[ERROR] RUNTIME(45282,python):2022-03-14-17:25:46.997.352 [task.cc:2843]45806 PrintErrorInfo:Notify wait execute failed, device_id=3, stream_id=777, task_id=3, notify_id=62.
[ERROR] RUNTIME(45282,python):2022-03-14-17:25:46.997.369 [task.cc:2607]45806 ReportErrorInfo:model execute error, retCode=0x91, [the model stream execute failed].
[ERROR] RUNTIME(45282,python):2022-03-14-17:25:46.997.376 [task.cc:2580]45806 PrintErrorInfo:model execute task failed, device_id=3, model stream_id=30, model task_id=1, model_id=0, first task_id=6535
[ERROR] HCCL(45282,python):2022-03-14-17:25:46.998.809 [task_exception_handler.cc:190][45282][45806][TaskExceptionHandler][Callback]Task form HCCL run failed.
[ERROR] HCCL(45282,python):2022-03-14-17:25:46.998.836 [task_exception_handler.cc:193][45282][45806][TaskExceptionHandler][Callback]Task failed base information is streamId: [777], taskID: [3], taskType: [Notify Wait], tag: [16-66211557098207583]backward_0_10_5].
[ERROR] HCCL(45282,python):2022-03-14-17:25:46.998.871 [task_exception_handler.cc:155][45282][45806][TaskExceptionHandler][Callback]Task failed para information is notify id: [0x000000009000001f0], stage: [ffffff], remote rank: [491].
[ERROR] RUNTIME(45282,python):2022-03-14-17:25:46.998.969 [stream.cc:712]45806 GetError:Stream Synchronize failed, stream_id=30 retCode=0x91, [the model stream execute failed].
```

2.对于非HCCL错误（包括日志老化，环境问题等），对日志输出无固定格式。

使用示例：

图 4-3 日志分析示例

```
[root@ms-aiserver-ei-56-42 tools]# python3 parse_plog_hccl.py /npu/debug00854014/plog
['rank8', 'rank9', 'rank10', 'rank11', 'rank12', 'rank13', 'rank14', 'rank15', 'rank16', 'rank17', 'rank18', 'rank19', 'rank20', 'rank21', 'rank22'] no errors in the log. The possible cause is that the process is hung.
[root@ms-aiserver-ei-5
```

4.3 集群故障日志分析（云环境）

4.3.1 工具介绍

适用场景

云上集群训练任务失败时，面对大量的日志，用户较难快速锁定问题rank，影响集群问题定位效率。本工具会通过自动解析（HCCL组件的）日志，快速给出集群中未报错的/最早报错的rankid，并辅助定界/定位一些集群常见问题，通过打屏呈现给用户。适用于：

1. 有心跳检测版本。
2. 判断是否有卡训练未拉起或异常退出导致集群训练失败，并给出问题的rankid。
3. 判断是否存在卡间异步时间过长引起集群互相等待超时导致集群训练失败，并给出最早/最晚拉起集合通信任务的rankid。
4. 判断是否存在常见的HCCL内task执行失败（SDMA计算溢出、memcpy访问越界）导致的集群训练失败。

环境准备

使用plog解析日志工具前，请确保环境已经安装以下模块：

已安装python软件，且使用的python版本为3.7.5及以上版本，ubuntu系统下可以使用apt-get install python3命令安装。

工具获取

plog解析工具存放在toolkit包中，安装后，plog日志解析工具会自动安装到\${HOME_ToolKit}/tools目录下，文件名称：parse_plog_hccl.py。

约束与限制

1. 使用前请确保所有卡日志齐全，并保留原始文件名称，不支持日志老化场景。
2. 所有卡日志需要放入同一个文件夹下，并保证不同的server的日志放在不同的子目录下。
3. 运行日志时使用的python版本必须要为Python 3.7.5及以上版本。

4.3.2 工具使用方法

命令行使用说明

命令行格式: `python3 parse_plog_hccl.py path --Force_kill=1`

参数说明:

1. path可以是绝对路径也可以是相对路径。需要指定为待解析的plog日志所在的路径，该路径下包含所有卡的日志文件。
2. --Force_kill=1 表示是云上版本。

命令执行示例:

```
Python3 parse_plog_hccl.py /home/HwHiAiUser/modelarts-job-ec7b5289-  
e9fc-4fa6-9467-cbd3dad4eb9a --Force_kill=1
```

屏显结果分析

脚本执行过程中，会有对应的日志输出，主要输出的内容说明如下。

对于HCCL错误输出格式如下：

- 第一行：无心跳的rankid。
- 第二行：主要是错误的结果日志，表明错误是不是属于HCCL组件，如果是，则会给出错误的原因。
- 第三行：打印最先报错的rankid及报错时间。
- 第四行：打印最晚报错的rankid及报错时间。
- 第四行：建链/notify超时时间阈值。
- 第五行：集群rank数量。
- 第六行及以后：打印最早报错的rank里面的详细的ERROR日志。