

CANN
5.0.4.6

故障处理 (开放态, Ascend 310)

文档版本 01
发布日期 2023-04-12



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。

前言

概述

本文档以故障案例形式介绍昇腾AI处理器常见异常现象及处理方法，同时提供了常见错误码及故障定位工具相关介绍。






读者对象

本文档主要适用于开发人员。开发人员必须具备以下经验和技能：

- 熟悉昇腾AI处理器内部模块功能、业务流程。
- 有故障维护经验，熟悉查阅日志等。

符号约定

在本文中可能出现下列标志，它们所代表的含义如下。

| 符号 | 说明 |
|---|---|
|  危险 | 表示如不避免则将会导致死亡或严重伤害的具有高等级风险的危害。 |
|  警告 | 表示如不避免则可能导致死亡或严重伤害的具有中等级风险的危害。 |
|  注意 | 表示如不避免则可能导致轻微或中度伤害的具有低等级风险的危害。 |
|  须知 | 用于传递设备或环境安全警示信息。如不避免则可能会导致设备损坏、数据丢失、设备性能降低或其它不可预知的结果。 “须知”不涉及人身伤害。 |
|  说明 | 对正文中重点信息的补充说明。 “说明”不是安全警示信息，不涉及人身、设备及环境伤害信息。 |

目 录

前言..... ii

1 概述..... 1

1.1 内容简介..... 1

1.2 处理流程简介..... 1

2 Error Code 故障..... 4

2.1 使用说明..... 4

2.2 Command Line Errors..... 5

2.2.1 E10001 Invalid Argument..... 5

2.2.2 E10002 Invalid --input_shape Argument..... 5

2.2.3 E10003 Invalid Argument..... 5

2.2.4 E10004 Invalid Argument..... 6

2.2.5 E10005 Invalid Argument..... 6

2.2.6 E10006 Invalid Argument..... 6

2.2.7 E10007 Invalid Required Argument..... 7

2.2.8 E10008 Invalid Argument..... 7

2.2.9 E10009 Invalid Dynamic Shape Argument..... 7

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..... 8

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 Argument, Node Not Found..... 9

2.2.17 E10017 Invalid Argument, Node Not Found..... 10

2.2.18 E10018 Invalid Dynamic Shape Argument..... 10

2.2.19 E10019 Invalid --input_shape Argument..... 10

2.2.20 E10020 Invalid --dynamic_image_size Argument..... 11

2.2.21 E10021 Invalid Argument..... 11

2.2.22 E10022 Invalid Argument..... 11

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..... | 12 |
| 2.2.27 E10027 Invalid --singleop Argument..... | 13 |
| 2.2.28 E10029 Invalid --singleop Argument..... | 13 |
| 2.2.29 E10030 Invalid --singleop Argument..... | 13 |
| 2.2.30 E10031 Invalid --input_shape Argument..... | 14 |
| 2.2.31 E10034 Invalid --input_fp16_nodes Argument..... | 14 |
| 2.2.32 E10035 Invalid Dynamic Shape Argument..... | 14 |
| 2.2.33 E10036 Invalid Dynamic Shape Argument..... | 15 |
| 2.2.34 E10037 Invalid Dynamic Shape Argument..... | 15 |
| 2.2.35 E10038 Invalid Dynamic Shape Argument..... | 15 |
| 2.2.36 E10039 Invalid Dynamic Shape Argument..... | 16 |
| 2.2.37 E10040 Invalid --input_shape Argument..... | 16 |
| 2.2.38 E10041 Invalid --framework or --model Argument..... | 16 |
| 2.2.39 E10045 Invalid Dynamic Shape Argument..... | 17 |
| 2.2.40 E10046 Invalid Dynamic Shape Argument..... | 17 |
| 2.2.41 E10047 Invalid Argument..... | 17 |
| 2.2.42 E10048 Invalid --input_shape_range Argument..... | 18 |
| 2.2.43 E10049 Invalid --input_shape_range Argument..... | 18 |
| 2.2.44 E10050 Invalid --input_shape_range Argument..... | 18 |
| 2.2.45 E10052 Invalid AIPP Configuration..... | 18 |
| 2.2.46 E10410 Invalid Argument..... | 19 |
| 2.2.47 E11001 Caffe Model Data Error..... | 19 |
| 2.2.48 E11002 Caffe Model Data Error..... | 19 |
| 2.2.49 E11003 Caffe Model Data Error..... | 20 |
| 2.2.50 E11004 Caffe Model Data Error..... | 20 |
| 2.2.51 E11005 Invalid --input_shape Argument..... | 20 |
| 2.2.52 E11008 Caffe Model Data Error..... | 21 |
| 2.2.53 E11009 Unsupported Caffe Operator..... | 21 |
| 2.2.54 E11012 Caffe Model Data Error..... | 21 |
| 2.2.55 E11014 Caffe Model Data Error..... | 21 |
| 2.2.56 E11015 Caffe Model Data Error..... | 22 |
| 2.2.57 E11016 Invalid --out_node Argument..... | 22 |
| 2.2.58 E11017 Invalid --out_node Argument..... | 22 |
| 2.2.59 E11018 Caffe Model Data Error..... | 23 |
| 2.2.60 E11021 Caffe Model Data Error..... | 23 |
| 2.2.61 E11022 Caffe Model Data Error..... | 23 |
| 2.2.62 E11023 Caffe Model Data Error..... | 23 |
| 2.2.63 E11024 Caffe Model Data Error..... | 24 |
| 2.2.64 E11027 Caffe Model Data Error..... | 24 |
| 2.2.65 E11029 Caffe Model Data Error..... | 24 |
| 2.2.66 E11032 Caffe File Error..... | 25 |
| 2.2.67 E11033 The Caffe weight file is invalid..... | 25 |

| | |
|---|----|
| 2.2.68 E11035 Caffe Model Data Error..... | 25 |
| 2.2.69 E11036 Caffe Model Data Error..... | 25 |
| 2.2.70 E11037 Caffe Model Data Error..... | 26 |
| 2.2.71 E12009 TensorFlow Model Data Error..... | 26 |
| 2.2.72 E12010 TensorFlow Model Data Error..... | 26 |
| 2.2.73 E12013 TensorFlow Model Data Error..... | 27 |
| 2.2.74 E12029 TensorFlow Model Data Error..... | 27 |
| 2.2.75 E16001 ONNX Model Data Error..... | 27 |
| 2.2.76 E16002 ONNX Model Data Error..... | 28 |
| 2.2.77 E16004 ONNX Model Data Error..... | 28 |
| 2.2.78 E16005 ONNX Model Data Error..... | 28 |
| 2.2.79 E20100 Invalid GE Initialization Argument..... | 28 |
| 2.2.80 E20101 Invalid GE Initialization Argument..... | 29 |
| 2.2.81 E20102 Invalid GE Initialization Argument..... | 29 |
| 2.2.82 E20103 Invalid Platform Initialization Argument..... | 29 |
| 2.2.83 W21000 Empty Path..... | 30 |
| 2.2.84 E21001 Failure to Open File..... | 30 |
| 2.2.85 E21002 Failure to Read File..... | 30 |
| 2.2.86 E21003 Empty Path..... | 31 |
| 2.2.87 W40001 Invalid Path..... | 31 |
| 2.2.88 W40002 Failure to Create Directory..... | 31 |
| 2.2.89 E40000 Failure to Import te.platform.log_util..... | 31 |
| 2.2.90 E40001 Failure to Import Python Module..... | 32 |
| 2.2.91 E40002 Failure to Call Python Function..... | 32 |
| 2.2.92 E40003 Failure to Import Auto Tiling Manager of Python Module..... | 32 |
| 2.2.93 E40004 Failure to Import RL Tiling Manager of Python Module..... | 33 |
| 2.2.94 E40008 Failure to Import Python Module..... | 33 |
| 2.2.95 E40010 Invalid Auto Tune Mode..... | 33 |
| 2.2.96 E40011 Invalid Argument..... | 34 |
| 2.2.97 EE1000 Invalid SoC Version..... | 34 |
| 2.2.98 E76002 ONNX Model Data Error..... | 34 |
| 2.3 Resource Errors..... | 35 |
| 2.3.1 E10044 Insufficient Memory..... | 35 |
| 2.3.2 E19022 Insufficient Memory..... | 35 |
| 2.3.3 E19023 Too Large OM Model..... | 35 |
| 2.4 Invalid Argument..... | 36 |
| 2.4.1 E10051 Invalid --job_id Argument..... | 36 |
| 2.4.2 E10401 Invalid Operator Input Count..... | 36 |
| 2.4.3 E10402 Invalid Input Buffer Allocation for Operator Execution..... | 36 |
| 2.4.4 E10403 Invalid Operator Output Count..... | 37 |
| 2.4.5 E10404 Invalid Output Buffer Allocation for Operator Execution..... | 37 |
| 2.4.6 E10405 Inconsistent Input Buffer Count and Input Tensor Count for Operator Execution..... | 37 |

| | |
|---|----|
| 2.4.7 E10406 Inconsistent Output Buffer Count and Output Tensor Count for Operator Execution..... | 38 |
| 2.4.8 E14001 Invalid Argument for Operator Compilation..... | 38 |
| 2.4.9 E19009 Operator Name Conflict..... | 38 |
| 2.4.10 E19014 Operator Data Verification Failure..... | 38 |
| 2.4.11 E19018 Failure to Parse File..... | 39 |
| 2.4.12 E19025 Input Tensor Error..... | 39 |
| 2.5 System Support Errors..... | 39 |
| 2.5.1 E10501 Unsupported Operator..... | 39 |
| 2.5.2 E13002 Unsupported Operator..... | 40 |
| 2.5.3 E13003 Unsupported Operator..... | 40 |
| 2.5.4 E19010 Unsupported Operator..... | 40 |
| 2.5.5 EZ0501 Unsupported Operator..... | 41 |
| 2.5.6 EZ3002 Unsupported Operator..... | 41 |
| 2.5.7 EZ3003 Unsupported Operator..... | 41 |
| 2.5.8 EZ9010 Unsupported Operator..... | 41 |
| 2.6 Configuration Errors..... | 42 |
| 2.6.1 E12004 Operator Prototype Registration Error..... | 42 |
| 2.6.2 E19024 Invalid Environment Variable..... | 42 |
| 2.7 File Errors..... | 42 |
| 2.7.1 E19000 Invalid Directory..... | 42 |
| 2.7.2 E19001 Failure to Open File..... | 43 |
| 2.7.3 E19002 Too Long File Directory..... | 43 |
| 2.7.4 E19003 Failure to Read File..... | 43 |
| 2.7.5 E19004 Failure to Write File..... | 44 |
| 2.7.6 E19005 Failure to Parse File..... | 44 |
| 2.7.7 E19015 File size invalid..... | 44 |
| 2.8 Invalid Path Name..... | 44 |
| 2.8.1 E19026 Input Path Name Error..... | 44 |
| 2.9 Minor Error..... | 45 |
| 2.9.1 W11001 Operator Missing High-Priority Performance..... | 45 |
| 2.10 ACL API Errors..... | 45 |
| 2.10.1 EH0001 Invalid Argument..... | 45 |
| 2.10.2 EH0002 Null Pointer..... | 45 |
| 2.10.3 EH0003 Invalid Path..... | 46 |
| 2.10.4 EH0004 Invalid File..... | 46 |
| 2.10.5 EH0005 Invalid AIPP Argument..... | 46 |
| 2.10.6 EH0006 Feature Unsupported..... | 47 |
| 2.11 Environment Errors..... | 47 |
| 2.11.1 E30000 Invalid Configuration File..... | 47 |
| 2.11.2 E30001 Invalid Configuration File..... | 47 |
| 2.11.3 E30002 Invalid Configuration File..... | 48 |
| 2.11.4 EE2000 Device Error..... | 48 |

| | |
|--|-----------|
| 2.11.5 EE2001 Failure to Allocate Memory..... | 48 |
| 2.11.6 EE3001 Driver Failure..... | 48 |
| 2.11.7 EI0001 Invalid Environment Variable Configuration..... | 49 |
| 2.11.8 EI0002 SO File Not Found..... | 49 |
| 2.11.9 EI0004 Invalid Ranktable Configuration..... | 49 |
| 2.11.10 EI0007 Allocation Failure..... | 50 |
| 2.12 Operator Compilation Errors..... | 50 |
| 2.12.1 E20001 Failure to Precompile Op..... | 50 |
| 2.12.2 E20003 Failure to Compile Op..... | 50 |
| 2.12.3 E20004 Failure to Compile Op..... | 51 |
| 2.12.4 E40009 Failure to Compile Op..... | 51 |
| 2.13 Invalid Operator Arguments..... | 51 |
| 2.13.1 E20002 Failure to Precompile Op..... | 51 |
| 2.13.2 E20006 Failure to Calculate Data Edge Size..... | 52 |
| 2.13.3 E40005 Failure to Execute Function check_supported..... | 52 |
| 2.13.4 E40006 Failure to Obtain Op Format..... | 52 |
| 2.13.5 E40007 Failure to Precompile Op..... | 53 |
| 2.13.6 EI0003 Invalid Collective Communication Op Argument..... | 53 |
| 2.13.7 EI0005 Inconsistent Collective Communication Arguments Between Ranks..... | 53 |
| 2.13.8 EI0006 Invalid Collective Communication Allocation..... | 53 |
| 2.14 Fusion Pass Errors..... | 54 |
| 2.14.1 E20007 Failure to Execute Fusion Pass..... | 54 |
| 2.14.2 E20008 Failure to Execute Fusion Pass..... | 54 |
| 2.14.3 E20009 Failure to Execute Fusion Pass..... | 54 |
| 2.15 Task Generation Errors..... | 55 |
| 2.15.1 E20005 Failure to Generate Task..... | 55 |
| 2.16 TBE Compiler Errors..... | 55 |
| 2.16.1 EB0000 Invalid IR..... | 55 |
| 2.17 Task Schedule Execution Errors..... | 55 |
| 2.17.1 EE4001 Model Execution Error..... | 56 |
| 2.17.2 EE4002 Model Execution Error..... | 56 |
| 2.17.3 EE4003 Profiling Execution Error..... | 56 |
| 2.17.4 EE4004 Profiling Execution Error..... | 56 |
| 2.17.5 EE5001 Task Schedule Resource Error..... | 57 |
| 2.17.6 EE5002 Task Schedule Resource Error..... | 57 |
| 2.18 Default Error..... | 57 |
| 2.18.1 E19999 System Terminated..... | 57 |
| 3 AscendCL 常见故障..... | 58 |
| 3.1 背景知识..... | 58 |
| 3.2 常见故障..... | 58 |
| 3.3 常见故障分析与处理..... | 58 |
| 3.3.1 内存未释放..... | 59 |

| | |
|--|-----------|
| 3.3.2 APP 使用 dvpp 接口编译失败..... | 59 |
| 3.3.3 使用 dump 功能未获取 dump 结果..... | 60 |
| 3.3.4 Event 数量超过上限导致 aclrtRecordEvent 接口返回失败..... | 61 |
| 3.3.5 执行单算子产生 coredump 的定位处理..... | 62 |
| 3.3.6 进程异常退出后重新执行任务失败..... | 64 |
| 3.3.7 进程异常时资源清理的处理建议..... | 64 |
| 4 ATC 常见故障..... | 65 |
| 4.1 背景知识..... | 65 |
| 4.2 常见故障..... | 66 |
| 4.3 常见故障分析与处理..... | 66 |
| 4.3.1 目标模型中包含算子原型库未注册的算子..... | 66 |
| 4.3.2 目标模型中因为未找到算子归属引擎导致转换失败..... | 67 |
| 4.3.3 缺省 input_shape 参数导致转换失败..... | 67 |
| 4.3.4 out_nodes 参数指定的输出节点不存在..... | 68 |
| 4.3.5 out_nodes 参数输入校验失败导致转换失败..... | 69 |
| 4.3.6 input_fp16_nodes 参数指定的节点名字不存在..... | 69 |
| 4.3.7 算子 infershape 失败导致模型转换失败..... | 70 |
| 4.3.8 算子找不到算子信息库导致模型转换失败..... | 70 |
| 4.3.9 insert_op_conf 参数文件路径不存在导致转换失败..... | 71 |
| 4.3.10 insert_op_conf 参数指定的 aipp 配置文件校验失败导致转换失败..... | 71 |
| 4.3.11 dynamic_image_size 或 dynamic_batch_size 设置分辨率数值过大或档位过多导致运行环境异常缓慢..... | 72 |
| 4.3.12 算子信息库中缺少算子..... | 72 |
| 5 GE 常见故障..... | 74 |
| 5.1 输入数据的大小远小于模型输入的大小导致推理失败..... | 74 |
| 5.2 输入数据的个数大于模型输入的个数导致推理失败..... | 75 |
| 6 DVPP 常见故障..... | 76 |
| 6.1 背景知识..... | 76 |
| 6.2 常见故障..... | 77 |
| 6.3 常见故障分析与处理..... | 78 |
| 6.3.1 视频解码丢帧/丢包..... | 78 |
| 6.3.2 视频解码花屏..... | 79 |
| 6.3.3 视频解码性能问题..... | 80 |
| 6.3.4 视频解码失败..... | 81 |
| 6.3.5 视频解码格式不支持..... | 82 |
| 6.3.6 视频帧解码失败不触发回调函数..... | 83 |
| 6.3.7 设置帧序号但回调接收的 hiai_data 对象却为空..... | 84 |
| 6.3.8 JPEG 图片解码失败..... | 84 |
| 6.3.9 VPC 调用失败..... | 85 |
| 6.3.10 多次调用 VPC 触发硬件保护机制, 导致 VPC 所有引擎不可用..... | 87 |
| 6.3.11 模型推理精度下降..... | 88 |
| 6.3.12 VDEC 或 VPC 超时说明..... | 89 |

| | |
|--------------------------------|-----------|
| 6.3.13 DVPP 输入和输出内存申请方式说明..... | 89 |
| 6.3.14 HFBC 格式介绍..... | 91 |
| 7 Runtime 常见故障..... | 93 |
| 7.1 用户进程异常退出后重启进程失败..... | 93 |
| 7.2 注册算子数超过最大规格..... | 94 |

1 概述

- 1.1 内容简介
- 1.2 处理流程简介

1.1 内容简介

本文主要以昇腾310 AI处理器场景案例的方式描述开发者在开发过程中、调试程序时可能遇到的各类异常故障现象，并通过分析可能原因及提供通用处理方法，方便开发者快速定位并解决故障。

为方便快速问题处理，内容按模块故障场景分类，直接指定到故障发生点，提供点到点的处理方案，内容更加实用。

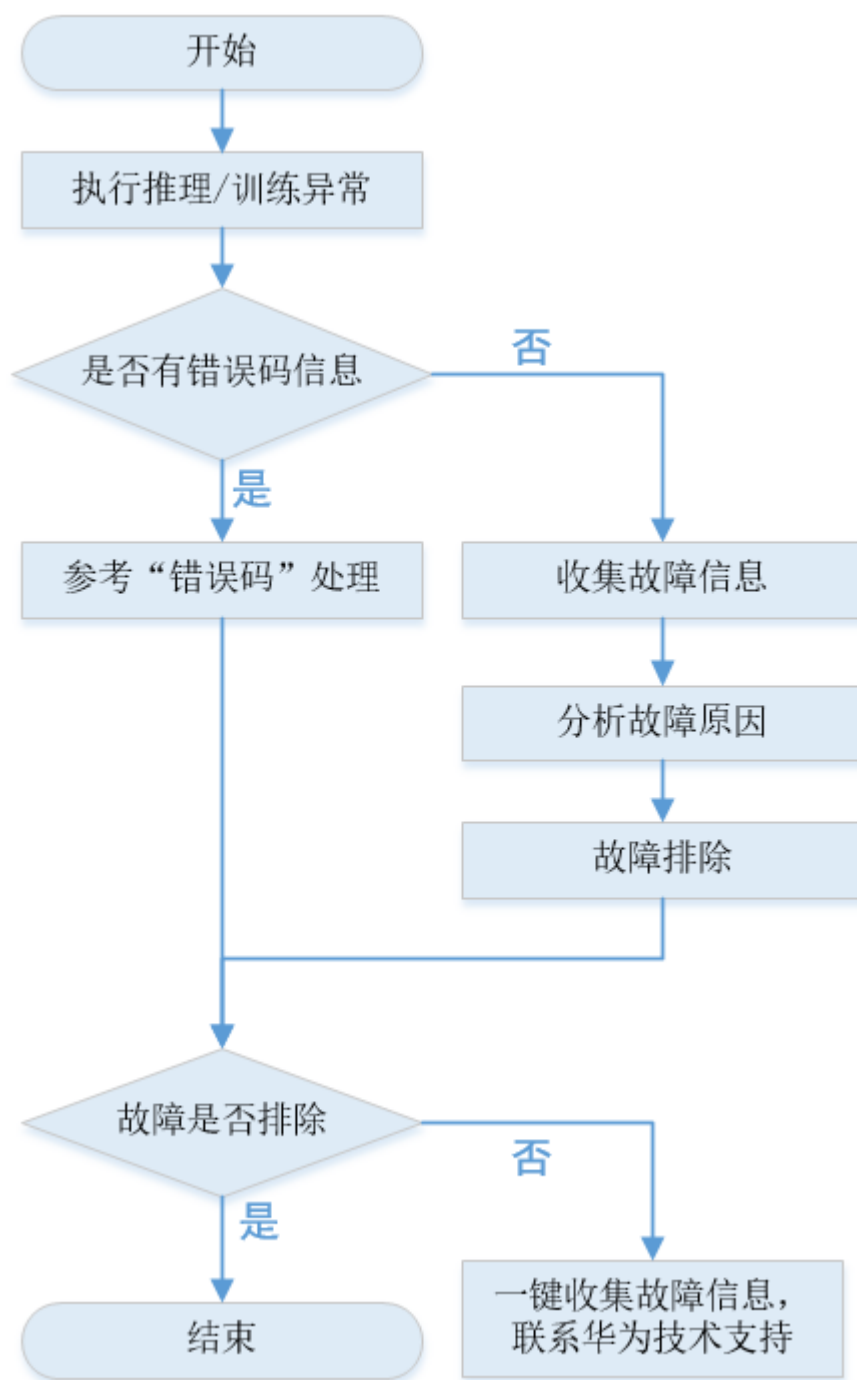
说明

本文档**Error Code故障**内容可以适配昇腾310 AI处理器、昇腾710 AI处理器、昇腾910 AI处理器使用；但其他章节仅适配昇腾310 AI处理器使用。

1.2 处理流程简介

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 Command Line Errors
- 2.3 Resource Errors
- 2.4 Invalid Argument
- 2.5 System Support Errors
- 2.6 Configuration Errors
- 2.7 File Errors
- 2.8 Invalid Path Name
- 2.9 Minor Error
- 2.10 ACL API Errors
- 2.11 Environment Errors
- 2.12 Operator Compilation Errors
- 2.13 Invalid Operator Arguments
- 2.14 Fusion Pass Errors
- 2.15 Task Generation Errors
- 2.16 TBE Compiler Errors
- 2.17 Task Schedule Execution Errors
- 2.18 Default 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E19000错误码在手册中的表示形式为:

E19000: Path[%s] is empty. Reason: %s.

用户界面实际报错示例:

E19000: Path[./test_argmax_case_v1.onn] is empty. Reason: No such file or directory.

2.2 Command Line Errors

2.2.1 E10001 Invalid Argument

Symptom

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

Possible Cause

Invalid Argument.

Solution

Try again with a valid argument.

2.2.2 E10002 Invalid --input_shape Argument

Symptom

Value [%s] for [--input_shape] is invalid. Reason: %s. Must be formatted as [%s].

Possible Cause

The [--input_shape] argument in the atc command line is invalid.

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 the [--%s] argument is invalid. Reason: %s.

Possible Cause

Invalid Argument.

Solution

Try again with a valid argument.

2.2.4 E10004 Invalid Argument

Symptom

Value for [--%s] is empty.

Possible Cause

The argument must not be empty.

Solution

Try again with a valid argument.

2.2.5 E10005 Invalid Argument

Symptom

Value [%s] for [--%s] is invalid. Must be either [true] or [false].

Possible Cause

The argument is invalid. Must be either [true] or [false].

Solution

Try again with a valid argument.

2.2.6 E10006 Invalid Argument

Symptom

Value [%s] for [--%s] is invalid. Must be either 1 or 0.

Possible Cause

The argument is invalid. Must be either 1 or 0.

Solution

Try again with a valid argument.

2.2.7 E10007 Invalid Required Argument

Symptom

[--%s] is required. Must be [%s].

Possible Cause

The argument is invalid.

Solution

Try again with a valid argument.

2.2.8 E10008 Invalid Argument

Symptom

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

Possible Cause

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

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], and [--dynamic_dims] are mutually exclusive.

Possible Cause

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

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 [--log] is invalid. Must be selected from [debug], [info], [warning], [error], and [null].

Possible Cause

The [--log] argument is invalid.

Solution

Valid values include [debug], [info], [warning], [error], and [null].

2.2.11 E10011 Invalid --input_shape Argument

Symptom

Argument [%s] for [--input_shape] is invalid. Shape values must be positive integers. The error value is: %s.

Possible Cause

In static shape scenarios, set the shape values of each input node in the [--input_shape] to positive integers.

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

As [--dynamic_batch_size] is included, the dimension count of the dynamic-shape input configured in [--input_shape] must not be 0.

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.

Possible Cause

The argument is out of range.

Solution

Try again with a valid argument.

2.2.14 E10014 Invalid Argument

Symptom

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

Possible Cause

The argument is invalid.

Solution

Try again with a valid argument. 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

Try again with a valid argument. For details, see [ATC Instructions].

2.2.16 E10016 Invalid Argument, Node Not Found

Symptom

Input Op [%s] specified in [--%s] is not found in the model.

Possible Cause

The node specified in the argument does not exist in the model.

Solution

Try again with a valid node name.

2.2.17 E10017 Invalid Argument, Node Not Found

Symptom

Input Op [%s] specified in [--%s] is invalid.

Possible Cause

The input node specified in the argument is invalid.

Solution

Try again with a valid input node name.

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 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 [--dynamic_image_size] is invalid. Must be formatted as [imagesize1_height,imagesize1_width;imagesize2_height,imagesize2_width].

Possible Cause

In the [--dynamic_image_size] argument, each profile must have two dimensions.

Solution

Try again with a valid [--dynamic_image_size] argument. Make sure that each profile has two dimensions.

2.2.21 E10021 Invalid Argument

Symptom

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

Possible Cause

The specified directory is too long.

Solution

Try again with a valid directory.

2.2.22 E10022 Invalid Argument

Symptom

Directory [%s] for [--%s] does not include the file name

Possible Cause

The specified directory does not include the file name.

Solution

Add the file name to the directory.

2.2.23 E10023 Invalid --singleop Argument

Symptom

Value [%s] for [--singleop] is invalid. Must be a proper directory.

Possible Cause

The [--singleop] argument is not a proper directory.

Solution

Try again with a valid argument.

2.2.24 E10024 Invalid --singleop Argument

Symptom

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

Possible Cause

Failed to open the file specified by the [--singleop] argument.

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

Failed to parse the file specified by the [--singleop] argument in JSON format.

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 the [--singleop] argument.

Possible Cause

The file specified by the [--singleop] argument contains an unnamed Op.

Solution

Check that no Op name is empty in the file.

2.2.27 E10027 Invalid --singleop Argument

Symptom

[%s] and [%s] for tensor indexed [%s] of Op [%s] are invalid in the file specified by the [--singleop] argument.

Possible Cause

The file specified by the [--singleop] argument contains invalid tensor attributes.

Solution

Try again with 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 the [--singleop] argument.

Possible Cause

The file specified by the [--singleop] argument contains an unnamed attribute.

Solution

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

2.2.29 E10030 Invalid --singleop Argument

Symptom

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

Possible Cause

The file specified by the [--singleop] argument contains an empty attribute.

Solution

Check that no Op attribute value is empty 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 E10034 Invalid --input_fp16_nodes Argument

Symptom

Nodes (for example, [%s]) connected to aipp must not be of type fp16. Check the [--insert_op_conf] and [--input_fp16_nodes] options in your atc command line.

Possible Cause

Do not set dtype to fp16 for a node connected to aipp.

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.

2.2.32 E10035 Invalid Dynamic Shape Argument

Symptom

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

Possible Cause

The number of configured profiles is less than the minimum.

Solution

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

2.2.33 E10036 Invalid Dynamic Shape Argument

Symptom

[--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims] has [%s] profiles, which exceeds the maximum [%s].

Possible Cause

The number of configured profiles exceeds the maximum.

Solution

Ensure that the number of profiles configured in the [--dynamic_batch_size], [--dynamic_image_size], or [--dynamic_dims] argument does not exceed the maximum.

2.2.34 E10037 Invalid Dynamic Shape Argument

Symptom

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

Possible Cause

The configured profiles have inconsistent dimension counts.

Solution

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

2.2.35 E10038 Invalid Dynamic Shape Argument

Symptom

Dimension size [%s] is invalid. Must be greater than 0.

Possible Cause

The shape values of each profile must be positive.

Solution

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

2.2.36 E10039 Invalid Dynamic Shape Argument

Symptom

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

Possible Cause

The configured profiles must be unique.

Solution

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

2.2.37 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.

Possible Cause

In dynamic shape scenarios, the corresponding shape value in the input node must be set to -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.

2.2.38 E10041 Invalid --framework or --model Argument

Symptom

Failed to load the model from [%s]. Check the model file or the [--framework] argument.

Possible Cause

The model file is invalid.

Solution

1. Check that the model file is valid.
2. Check that the [--framework] argument matches the actual framework of the model file.

2.2.39 E10045 Invalid Dynamic Shape Argument

Symptom

The number of -1 axes overflow for data [%s] with shape [%s].

Possible Cause

[--dynamic_dims] is included for more than one input, but the number of -1 axes in the [--input_shape] argument does not match the dimension count per profile in [--dynamic_dims].

Solution

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

2.2.40 E10046 Invalid Dynamic Shape Argument

Symptom

The number of -1 axes of data [%s] with shape [%s] is greater than the dimension count per profile.

Possible Cause

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

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].

2.2.41 E10047 Invalid Argument

Symptom

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

Possible Cause

The options are mutually exclusive.

Solution

Remove either of them and try again.

2.2.42 E10048 Invalid --input_shape_range Argument

Symptom

Value [%s] for [--input_shape_range] is invalid. Reason: %s. Must be formatted as [%s].

Possible Cause

Invalid --input_shape_range Argument

Solution

Try again with a valid argument.

2.2.43 E10049 Invalid --input_shape_range Argument

Symptom

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

Possible Cause

The dimension count in the [--input_shape_range] argument does not match the dimension count of the node.

Solution

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

2.2.44 E10050 Invalid --input_shape_range Argument

Symptom

Current dimension size [%s] is out of range [%s-%s] specified by [--input_shape_range].

Possible Cause

The dimension size is out of the range specified by [--input_shape_range].

Solution

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

2.2.45 E10052 Invalid AIPP Configuration

Symptom

AIPP configuration is invalid. Reason: %s.

Possible Cause

The AIPP configuration is invalid.

Solution

Try again with valid AIPP configuration.

2.2.46 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.

2.2.47 E11001 Caffe Model Data Error

Symptom

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

Possible Cause

[--input_dim] and [--input_shape] are mutually exclusive in NetParameter of the source Caffe model.

Solution

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

2.2.48 E11002 Caffe Model Data Error

Symptom

Caffe model has no input.

Possible Cause

The source Caffe model has no input.

Solution

Modify your Caffe model and try again.

2.2.49 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

The number of [input_dim] fields in the source Caffe model is not 4x the input count.

Solution

Modify your Caffe model and try again.

2.2.50 E11004 Caffe Model Data Error

Symptom

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

Possible Cause

The number of input shapes of the source Caffe model does not match the number of inputs.

Solution

Modify your Caffe model and try again.

2.2.51 E11005 Invalid --input_shape Argument

Symptom

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

Possible Cause

The shape of the input in the source Caffe model is not defined.

Solution

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

2.2.52 E11008 Caffe Model Data Error

Symptom

Optype DetectionOutput is unsupported. Modify the model file and use an explicit type, such as FSRDetectionOutput or SSDDetectionOutput.

Possible Cause

The source Caffe model contains an unsupported optype, DetectionOutput.

Solution

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

2.2.53 E11009 Unsupported Caffe Operator

Symptom

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

Possible Cause

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

Solution

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

2.2.54 E11012 Caffe Model Data Error

Symptom

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

Possible Cause

The bottom blob in the source Caffe model has no connected top blob.

Solution

Modify your Caffe model and try again.

2.2.55 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.

2.2.56 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.

2.2.57 E11016 Invalid --out_node 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

The node index specified by the [--out_node] argument is out of the range of outputs of the corresponding node in the source Caffe model.

Solution

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

2.2.58 E11017 Invalid --out_node Argument

Symptom

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

Possible Cause

The node specified by the [--out_node] argument does not exist in the source Caffe model.

Solution

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

2.2.59 E11018 Caffe Model Data Error

Symptom

Op name string [%s] is invalid. Allowed characters include: letters, digits, hyphens (-), periods (.), underscores (_), and slashes (/).

Possible Cause

The Op name string in the source Caffe model is invalid. Allowed characters include: letters, digits, hyphens (-), periods (.), underscores (_), and slashes (/).

Solution

Modify the Op name string and try again.

2.2.60 E11021 Caffe Model Data Error

Symptom

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

Possible Cause

The Caffe model contains [layers] structures, which have been deprecated in Caffe and unsupported by ATC. Replace [layers] with [layer].

Solution

Try again with a valid Caffe model.

2.2.61 E11022 Caffe Model Data Error

Symptom

Invalid prototxt file. No [layer] structures are found in the model.

Possible Cause

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

Solution

Try again with a valid Caffe model.

2.2.62 E11023 Caffe Model Data Error

Symptom

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

Possible Cause

The Caffe weight file contains [layers] structures, which have been deprecated in Caffe and unsupported by ATC. Replace [layers] with [layer].

Solution

Try again with a valid Caffe weight file.

2.2.63 E11024 Caffe Model Data Error

Symptom

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

Possible Cause

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

Solution

Try again with a valid Caffe weight file.

2.2.64 E11027 Caffe Model Data Error

Symptom

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

Possible Cause

The source Caffe model has an input node with shape size 0.

Solution

Try again with a valid Caffe model.

2.2.65 E11029 Caffe Model Data Error

Symptom

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

Possible Cause

The node exists in the Caffe model file but is not found in weight file.

Solution

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

2.2.66 E11032 Caffe File Error

Symptom

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

Possible Cause

The Caffe .proto file is invalid.

Solution

Try again with a valid .proto file.

2.2.67 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.

2.2.68 E11035 Caffe Model Data Error

Symptom

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

Possible Cause

The top size of the data node in the source Caffe model is not 1.

Solution

Try again with a valid Caffe model.

2.2.69 E11036 Caffe Model Data Error

Symptom

Data nodes have duplicate top blobs [%s].

Possible Cause

The data nodes in the source Caffe model have duplicate top blobs.

Solution

Try again with a valid Caffe model.

2.2.70 E11037 Caffe Model Data Error

Symptom

Op [%s] has zero outputs.

Possible Cause

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

Solution

Try again with a valid Caffe model.

2.2.71 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.

2.2.72 E12010 TensorFlow Model Data Error

Symptom

Model has no Placeholder or _Arg node.

Possible Cause

No node of type Placeholder or _Arg is found in the model.

Solution

Try again with a valid TensorFlow model.

2.2.73 E12013 TensorFlow Model Data Error

Symptom

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

Possible Cause

No subgraph .proto description is found based on the subgraph name.

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].

2.2.74 E12029 TensorFlow Model Data Error

Symptom

Failed to find the subgraph library. As the model contains function operators, run the [func2graph.py] script in the ATC installation directory to save the subgraphs to [graph_def_library.pbtxt].

Possible Cause

If the model to convert contains function subgraphs, the [graph_def_library.pbtxt] file is required.

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].

2.2.75 E16001 ONNX Model Data Error

Symptom

Model has no [%s] node.

Possible Cause

The node is not found in the ONNX model.

Solution

Try again with a valid ONNX model.

2.2.76 E16002 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.

2.2.77 E16004 ONNX Model Data Error

Symptom

ONNX model has no graph.

Possible Cause

The ONNX model has no graph.

Solution

Try again with a valid ONNX model.

2.2.78 E16005 ONNX Model Data Error

Symptom

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

Possible Cause

The source ONNX model has more than one [--domain_version] argument.

Solution

Try again with a valid ONNX model.

2.2.79 E20100 Invalid GE Initialization Argument

Symptom

Value [%s] is invalid for option [ge.opSelectImplmode]. Must be either [high_precision] or [high_performance].

Possible Cause

The operator implementation mode specified for GE initialization is invalid. Must be either [high_precision] or [high_performance].

Solution

Try again with a valid [opSelectImplmode] argument, which must match the [--op_select_implmode] argument in the ATC command line. Alternatively, the error is due to a GE translation failure.

2.2.80 E20101 Invalid GE Initialization Argument

Symptom

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

Possible Cause

See the error message for details.

Solution

Try again with a valid argument.

2.2.81 E20102 Invalid GE Initialization Argument

Symptom

Value [%s] is invalid for option [ge.engineType]. Must be either [AiCore] or [VectorCore].

Possible Cause

The engine type argument for GE initialization is invalid. Only [AiCore] or [VectorCore] is supported.

Solution

Contact Huawei technical support to check the engine type specified for GE initialization.

2.2.82 E20103 Invalid Platform Initialization Argument

Symptom

Value [%s] for option [ge.aicoreNum] is out of range (0, %s].

Possible Cause

The number of AI Cores configured for platform initialization is out of the valid range. This is a Huawei internal parameter and is user-irrelevant.

Solution

Contact Huawei technical support to check the value of [aicoreNum] in the [platform.ini] file.

2.2.83 W21000 Empty Path

Symptom

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

Possible Cause

This error does not terminate the process. The path is not specified or the specified path is empty.

Solution

Try again with a valid path.

2.2.84 E21001 Failure to Open File

Symptom

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

Possible Cause

The file does not exist or cannot be opened.

Solution

Check the error message and error log.

2.2.85 E21002 Failure to Read File

Symptom

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

Possible Cause

The file path is empty or the file does not exist.

Solution

Check the error message and error log.

2.2.86 E21003 Empty Path

Symptom

The file path of node [%s] is empty. Reason: %s.

Possible Cause

The compilation result of this node is empty, possibly due to a compilation failure.

Solution

1. Check the error message to see if a compilation failure has occurred.
2. If no compilation failure has occurred, contact Huawei technical support for fault locating.

2.2.87 W40001 Invalid Path

Symptom

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

Possible Cause

The path is empty or does not exist.

Solution

Try again with a valid path.

2.2.88 W40002 Failure to Create Directory

Symptom

Failed to create directory [%s]. Reason: %s.

Possible Cause

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

Solution

Modify the permission or directory name and try again.

2.2.89 E40000 Failure to Import te.platform.log_util

Symptom

Failed to import te.platform.log_util.

Possible Cause

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

Solution

1. Check that all required components are properly installed, Python is properly installed, and the specified Python path matches the Python installation directory.
2. Check the configured environment variables such as [PYTHONPATH] and [ASCEND_OPP_PATH].

2.2.90 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.

2.2.91 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.

2.2.92 E40003 Failure to Import Auto Tiling Manager of Python Module

Symptom

Failed to import Auto Tiling Manager of Python module: [%s]

Possible Cause

A component installation is missing or its installed version does not match.
Alternatively, the specified [PYTHONPATH] is invalid.

Solution

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

2.2.93 E40004 Failure to Import RL Tiling Manager of Python Module

Symptom

Failed to import RL Tiling Manager of Python module: [%s]

Possible Cause

A component installation is missing or its installed version does not match.
Alternatively, the specified [PYTHONPATH] is invalid.

Solution

Check that the ATC and FwkACCLib components are properly installed and the specified [PYTHONPATH] matches the Python installation directory.

2.2.94 E40008 Failure to Import Python Module

Symptom

Failed to import the Python module from [%s].

Possible Cause

A component installation is missing or its installed version does not match.
Alternatively, the specified [PYTHONPATH] is invalid. Check that Python is properly installed, and the specified Python path matches the Python installation directory.

Solution

Check that the ATC and FwkACCLib components are properly installed and the specified [PYTHONPATH] matches the Python installation directory.

2.2.95 E40010 Invalid Auto Tune Mode

Symptom

Auto Tune mode [%s] is invalid.

Possible Cause

The specified Auto Tune mode is invalid.

Solution

Try again with a valid argument.

2.2.96 E40011 Invalid Argument

Symptom

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

Possible Cause

An invalid argument is passed.

Solution

Try again with a valid argument. If the error persists, check the argument verification logic in the code.

2.2.97 EE1000 Invalid SoC Version

Symptom

Value %s is invalid for the SoC version. Valid values are: %s

Possible Cause

The SoC version argument is invalid.

Solution

Try again with a valid argument.

2.2.98 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.

2.3 Resource Errors

2.3.1 E10044 Insufficient Memory

Symptom

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

Possible Cause

Auto Tune requires at least 2 GB available memory.

Solution

Stop unnecessary processes and ensure that at least 2 GB system memory is available. Alternatively, change a machine with 2 GB or more of available memory.

2.3.2 E19022 Insufficient Memory

Symptom

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

Possible Cause

The memory size required by the model exceeds the system limit.

Solution

1. Retrain the model with a reduced batch size.
2. Reduce the model size.

2.3.3 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

Try again with reduced model size.

2.4 Invalid Argument

2.4.1 E10051 Invalid --job_id Argument

Symptom

Value [%s] for [--job_id] is too long, exceeding the allowed maximum [%s].

Possible Cause

The length of the [--job_id] argument is too long.

Solution

Try again with a valid argument.

2.4.2 E10401 Invalid Operator Input Count

Symptom

The operator takes [%s] input(s) according to specification, but [%s] input(s) are configured.

Possible Cause

The number of inputs configured for operator execution does not match the operator specification.

Solution

Try again with a valid number of inputs.

2.4.3 E10402 Invalid Input Buffer Allocation for Operator Execution

Symptom

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

Possible Cause

The input buffer allocation for operator execution does not match that required.

Solution

Try again with a valid buffer allocation.

2.4.4 E10403 Invalid Operator Output Count

Symptom

The operator returns [%s] output(s) according to specification, but [%s] output(s) are configured.

Possible Cause

The number of outputs configured for operator execution does not match the operator specification.

Solution

Try again with a valid number of outputs.

2.4.5 E10404 Invalid Output Buffer Allocation for Operator Execution

Symptom

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

Possible Cause

The output buffer allocation for operator execution does not match that required.

Solution

Try again with a valid buffer allocation.

2.4.6 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

The number of input buffers for operator execution does not match the number of input tensors.

Solution

Ensure that the number of input buffers for operator execution matches the number of input tensors.

2.4.7 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

The number of output buffers for operator execution does not match the number of output tensors.

Solution

Ensure that the number of output buffers for operator execution matches the number of output tensors.

2.4.8 E14001 Invalid Argument for Operator Compilation

Symptom

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

Possible Cause

The argument for operator compilation is invalid.

Solution

Try again with a valid argument.

2.4.9 E19009 Operator Name Conflict

Symptom

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

Possible Cause

The graph contains operators with duplicate names.

Solution

Ensure that the operators in the graph have unique names.

2.4.10 E19014 Operator Data Verification Failure

Symptom

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

Possible Cause

Verification of operator data fails.

Solution

Fix the error according to the error message.

2.4.11 E19018 Failure to Parse File

Symptom

Failed to parse file [%s] through [google::protobuf::TextFormat::Parse]. Try again with a valid protobuf file.

Possible Cause

Failed to parse the file using the protobuf library function. The file may not be in valid protobuf format.

Solution

Try again with a valid file.

2.4.12 E19025 Input Tensor Error

Symptom

Input tensor is invalid. Reason: %s.

Possible Cause

The input tensor does not match the graph is invalid, resulting in an execution failure of the training or inference job.

Solution

Fix the error according to the error message.

2.5 System Support Errors

2.5.1 E10501 Unsupported Operator

Symptom

IR for Op [%s, otype [%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.

2.5.2 E13002 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.

2.5.3 E13003 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.

2.5.4 E19010 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>.

2.5.5 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.

2.5.6 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.

2.5.7 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.

2.5.8 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>.

2.6 Configuration Errors

2.6.1 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] (the abs value) must be less than the input count [%s]

Possible Cause

To optimize an input to an attribute during operator registration, if the input index is less than 0, the absolute value of the index must be less than the actual number of inputs of the operator.

Solution

Try again with a valid operator prototype.

2.6.2 E19024 Invalid Environment Variable

Symptom

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

Possible Cause

The configured environment variable is invalid.

Solution

Fix the error according to the error message.

2.7 File Errors

2.7.1 E19000 Invalid Directory

Symptom

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

Possible Cause

The file does not exist.

Solution

Try again with a valid directory.

2.7.2 E19001 Failure to Open File

Symptom

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

Possible Cause

Failed to open the file.

Solution

Fix the error according to the error message.

2.7.3 E19002 Too Long File Directory

Symptom

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

Possible Cause

The file directory is too long.

Solution

Try again with a valid file directory.

2.7.4 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.

2.7.5 E19004 Failure to Write File

Symptom

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

Possible Cause

Failed to write the file.

Solution

Fix the error according to the error message.

2.7.6 E19005 Failure to Parse File

Symptom

Failed to parse file [%s]. Try again with a valid protobuf file or check the version of your protobuf installation.

Possible Cause

Failed to parse the file.

Solution

Check that a matched protobuf version is installed and try again with a valid file.

2.7.7 E19015 File size invalid

Symptom

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

Possible Cause

File size is not valid.

Solution

Try again with a valid file.

2.8 Invalid Path Name

2.8.1 E19026 Input Path Name Error

Symptom

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

Possible Cause

The input path does not exist or is invalid.

Solution

Try again with a valid path.

2.9 Minor Error

2.9.1 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>.

2.10 ACL API Errors

2.10.1 EH0001 Invalid Argument

Symptom

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

Possible Cause

The argument is invalid.

Solution

Try again with a valid argument.

2.10.2 EH0002 Null Pointer

Symptom

Argument [%s] must not be null.

Possible Cause

A null pointer is passed to the parameter.

Solution

Check that memory is properly allocated to the pointer and try again with a valid pointer.

2.10.3 EH0003 Invalid Path

Symptom

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

Possible Cause

The input path does not exist or is invalid. Alternatively, you do not have enough permission to access the path.

Solution

Try again with a valid path.

2.10.4 EH0004 Invalid File

Symptom

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

Possible Cause

The file is invalid.

Solution

Try again with a valid file.

2.10.5 EH0005 Invalid AIPP Argument

Symptom

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

Possible Cause

The AIPP argument is invalid.

Solution

Try again with a valid AIPP argument.

2.10.6 EH0006 Feature Unsupported

Symptom

%s is not supported. Reason: %s.

Possible Cause

The feature is not supported.

Solution

Check if parameters of unsupported features are used according to your environment setup.

2.11 Environment Errors

2.11.1 E30000 Invalid Configuration File

Symptom

System error occurred. Failed to load AI CPU Op library information from file [%s].
Reason: %s.

Possible Cause

The configuration file does not exist or the file is invalid.

Solution

Reinstall the FwkACCLib or ACLlib component or edit the file.

2.11.2 E30001 Invalid Configuration File

Symptom

System error occurred. Failed to load AI CPU kernel info from JSON file [%s].
Reason: %s.

Possible Cause

The configuration file does not exist or the file is invalid.

Solution

Reinstall the OPP component or edit the file.

2.11.3 E30002 Invalid Configuration File

Symptom

System error occurred. Failed to load IR configuration file [%s]. Reason: %s.

Possible Cause

The configuration file does not exist or the file is invalid.

Solution

Reinstall the FwkACCLlib or ACLlib component or edit the file.

2.11.4 EE2000 Device Error

Symptom

Device [%s] is not running properly.

Possible Cause

Failed to connect to the device, possibly due to a bus error or device error.

Solution

Send the log to Huawei technical support for fault locating.

2.11.5 EE2001 Failure to Allocate Memory

Symptom

Failed to allocate memory for [%s] due to insufficient memory. The requested allocation is %s.

Possible Cause

System memory is insufficient for the network.

Solution

1. Free the system memory and try again.
2. If no more system memory can be freed, reduce the network size and try again.

2.11.6 EE3001 Driver Failure

Symptom

The process has lost the connection between the host and device.

Possible Cause

The execution of a particular operator times out or the host and device connection is unstable.

Solution

Fix the timeout operator or reconnect the host and device, and then restart the app.

2.11.7 EI0001 Invalid Environment Variable Configuration

Symptom

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

Possible Cause

The environment variable configuration is invalid.

Solution

Try again with valid environment variable configuration.

2.11.8 EI0002 SO File Not Found

Symptom

SO file [%s] is not found.

Possible Cause

The file does not exist or is damaged.

Solution

Try again with a valid FwkACCLib installation.

2.11.9 EI0004 Invalid Ranktable Configuration

Symptom

Invalid ranktable, with rankID [%s] and local devID [%s]. Check that ranktable [%s] is valid and the environment setup matches the ranktable.

Possible Cause

The cluster configuration in the ranktable file does not match that of the actual operating environment, or the configuration format is invalid

Solution

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

2.11.10 EI0007 Allocation Failure

Symptom

Failed to allocate resource [%s], with info [%s].

Possible Cause

Failed to allocate memory or notify due to resource insufficiency.

Solution

Clean up unnecessary allocations.

2.12 Operator Compilation Errors

2.12.1 E20001 Failure to Precompile Op

Symptom

Failed to precompile Op [%s, optype [%s]] of graph_id [%s].

Possible Cause

The operator has an invalid argument.

Solution

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

2.12.2 E20003 Failure to Compile Op

Symptom

Failed to compile Op [%s, optype [%s]] of graph_id [%s].

Possible Cause

The operator has an invalid argument.

Solution

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

2.12.3 E20004 Failure to Compile Op

Symptom

Failed to compile Op [%s, optype [%s]] of graph_id [%s] in thread [%s] of task [%s].

Possible Cause

The operator has an invalid argument.

Solution

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

2.12.4 E40009 Failure to Compile Op

Symptom

Failed to compile Op [%s, oppath [%s], optype [%s]] of taskID [%s].

Possible Cause

See the error message for details.

Solution

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

2.13 Invalid Operator Arguments

2.13.1 E20002 Failure to Precompile Op

Symptom

Failed to precompile Op [%s, optype [%s]] of graph_id [%s] in thread [%s] of task [%s].

Possible Cause

The operator has an invalid argument.

Solution

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

2.13.2 E20006 Failure to Calculate Data Edge Size

Symptom

Failed to calculate the tensor size of Op [%s, optype %s].

Possible Cause

Failed to calculate the data edge size, possibly due to an invalid shape or format.

Solution

Inspect the error log, perform a partial dump by using **npucollect.sh** or the `[export DUMP_GE_GRAPH=2]` command, and then send the partial dump to Huawei technical support for fault locating.

2.13.3 E40005 Failure to Execute Function `check_supported`

Symptom

Failed to execute the `[check_supported]` function of Op [%s, oppath [%s], optype [%s]].

Possible Cause

An invalid argument is passed or the Python function `[check_supported]` has a bug. This error is most likely caused by an internal problem.

Solution

This error is caused by an internal problem. Contact Huawei technical support to check the Python function call in the code and the implementation of the `[check_supported]` function.

2.13.4 E40006 Failure to Obtain Op Format

Symptom

Failed to obtain Op [%s]'s format. Reason: %s.

Possible Cause

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

Solution

This error is caused by an internal problem. Contact Huawei technical support to check the Python function call in the code and the implementation of the `[op_select_format]` function.

2.13.5 E40007 Failure to Precompile Op

Symptom

Failed to precompile Op [%s, oppath [%s], optype [%s]].

Possible Cause

See the error message for details.

Solution

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

2.13.6 EI0003 Invalid Collective Communication Op Argument

Symptom

In [%s], value [%s] of parameter [%s] is invalid. Tips: %s.

Possible Cause

The collective communication operator has an invalid argument.

Solution

Try again with a valid argument.

2.13.7 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

The arguments for collective communication are inconsistent between ranks.

Solution

Try again with valid arguments. Ensure that the arguments are consistent between ranks.

2.13.8 EI0006 Invalid Collective Communication Allocation

Symptom

[%s] failed with info [%s]. Please check the passed [%s] resource.

Possible Cause

The memory or notify allocation passed to collective communication is invalid.

Solution

Try again with a valid memory or notify allocation.

2.14 Fusion Pass Errors

2.14.1 E20007 Failure to Execute Fusion Pass

Symptom

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

Possible Cause

The fusion pass code has a bug or the source graph does not meet the fusion pass requirements.

Solution

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

2.14.2 E20008 Failure to Execute Fusion Pass

Symptom

Failed to run pass [%s]. Reason: %s.

Possible Cause

The fusion pass code has a bug or the source graph does not meet the fusion pass requirements.

Solution

Check the pass code.

2.14.3 E20009 Failure to Execute Fusion Pass

Symptom

Failed to do topological sorting after graph fusion, graph is cyclic, graph name: %s, pass name: %s.

Possible Cause

The fusion pass code has a bug.

Solution

Check the pass code.

2.15 Task Generation Errors

2.15.1 E20005 Failure to Generate Task

Symptom

Failed to generate a task for Op [%s, optype [%s]] of graph_id [%s].

Possible Cause

A compilation argument is invalid, the operator is not compiled, or the GE memory allocation is invalid.

Solution

Check the error message and error log.

2.16 TBE Compiler Errors

2.16.1 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.

2.17 Task Schedule Execution Errors

2.17.1 EE4001 Model Execution Error

Symptom

This stream has been bound to a model. Repeated binding is not allowed. %s

Possible Cause

This stream has been bound to a model. Repeated binding is not allowed.

Solution

Remove redundant bindings from your code.

2.17.2 EE4002 Model Execution Error

Symptom

Failed to bind the stream to the model. %s

Possible Cause

The model ID does not match the input arguments or the specific stream is busy.

Solution

See the error message for details and modify the rtModelBindStream arguments.

2.17.3 EE4003 Profiling Execution Error

Symptom

Failed to profile the HWTS log, as online profiling is in process. Please stop online profiling first. %s

Possible Cause

Failed to profile the HWTS log, as online profiling is in process.

Solution

Stop online profiling and try again.

2.17.4 EE4004 Profiling Execution Error

Symptom

Failed to enable profiling. %s

Possible Cause

Profiling is enabled. See the error message for details.

Solution

Disable profiling and try again.

2.17.5 EE5001 Task Schedule Resource Error

Symptom

Insufficient HWTS submit queue(SQ) resource. %s

Possible Cause

The HWTS SQ resource are not sufficient to launch any new tasks.

Solution

Wait until sufficient HWTS SQ resource is available.

2.17.6 EE5002 Task Schedule Resource Error

Symptom

Insufficient task schedule memory. %s

Possible Cause

The task schedule memory are not sufficient to allocate new memory.

Solution

Reduce the number of tasks of your model and try again.

2.18 Default Error

2.18.1 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 **npucollect.sh** to collect the logs generated when the fault occurs and locate the fault based on the logs.

3 AscendCL 常见故障

3.1 背景知识

3.2 常见故障

3.3 常见故障分析与处理

3.1 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。

3.2 常见故障

根据故障场景发生的不同，AscendCL可能存在以下情况的故障分类：

- 内存未释放
- APP使用DVPP接口编译失败
- 使用Dump功能未获取Dump结果
- Device侧调用aclrtMallocHost/aclrtFreeHost接口失败
- Event数量超过上限导致aclrtRecordEvent接口返回失败
- 执行单算子产生coredump的定位处理
- 强行终止任务后重新执行任务失败

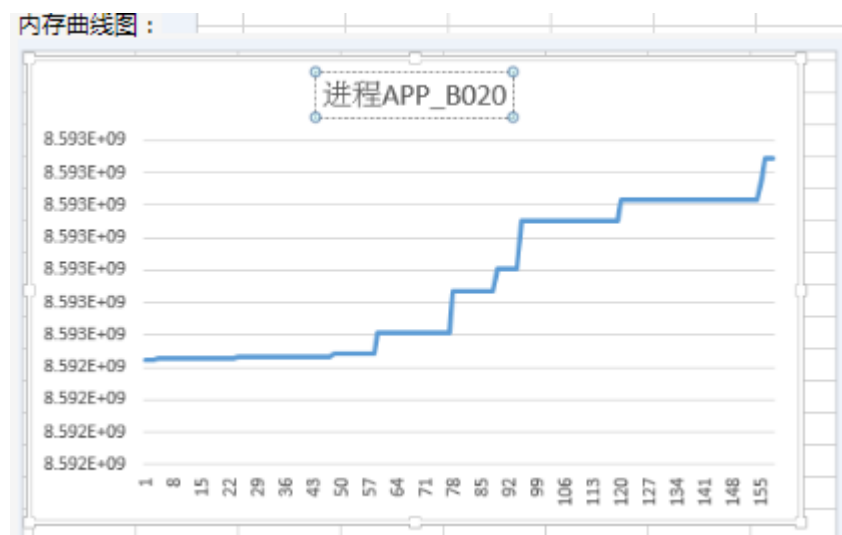
3.3 常见故障分析与处理

3.3.1 内存未释放

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。如图3-1所示。

图 3-1 内存占用持续上升



可能原因

分析上述日志信息，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：

排查所有的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。

3.3.2 APP 使用 dvpp 接口编译失败

现象描述

编译提示DVPP的相关接口未定义，如图3-2所示。

图 3-2 日志信息: undefined reference to ***

```
sample_process.cpp:(.text+0x82e7): undefined reference to `acldvppDestroyPicDesc'
CMakeFiles/main.dir/dvpp_process.cpp.o: In function `DvppProcess::ProcessResize()':
dvpp_process.cpp:(.text+0x86e): undefined reference to `acldvppVpcResizeAsync'
CMakeFiles/main.dir/dvpp_process.cpp.o: In function `DvppProcess::DestroyResizeResource()':
dvpp_process.cpp:(.text+0x907): undefined reference to `acldvppFree'
dvpp_process.cpp:(.text+0x930): undefined reference to `acldvppDestroyPicDesc'
dvpp_process.cpp:(.text+0x959): undefined reference to `acldvppDestroyPicDesc'
CMakeFiles/main.dir/dvpp_process.cpp.o: In function `DvppProcess::DestroyResource()':
dvpp_process.cpp:(.text+0x991): undefined reference to `acldvppDestroyResizeConfig'
dvpp_process.cpp:(.text+0x9b8): undefined reference to `acldvppDestroyChannel'
dvpp_process.cpp:(.text+0x9ec): undefined reference to `acldvppDestroyChannelDesc'
CMakeFiles/main.dir/sample_process.cpp.o: In function `SampleProcess::Process()':
sample_process.cpp:(.text+0x608): undefined reference to `acldvppFree'
sample_process.cpp:(.text+0x668): undefined reference to `acldvppFree'
sample_process.cpp:(.text+0x68c): undefined reference to `acldvppFree'
sample_process.cpp:(.text+0x720): undefined reference to `acldvppFree'
sample_process.cpp:(.text+0x775): undefined reference to `acldvppFree'
```

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

针对分析的故障可能原因,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。

图 3-3 排查结果示例

```
add_executable(main
    utils.cpp
    dvpp_process.cpp
    model_process.cpp
    singleOp_process.cpp
    sample_process.cpp
    main.cpp)

target_link_libraries(main
    ascendcl acl_dvpp stdc++)
```

3.3.3 使用 dump 功能未获取 dump 结果

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。如图3-4所示。

图 3-4 Dump 日志信息

```
[INFO] ASCENDCL(18454,acltest_host):2020-03-29-08:26:14.071.950 18454 HandleDumpConfig:acl/toolchain/dump.cpp:103: "HandleDumpConfig end in HandleDumpConfig"
[INFO] ASCENDCL(18454,acltest_host):2020-03-29-08:26:14.071.953 18454 aclInit:acl/common/acl.cpp:45: "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，如图3-5所示。

图 3-5 检查 model 名称是否正确



```
{
  "dump": {
    "dump_list": [
      {
        "model_name": "ResNet-50",
        "layer": [
          "res5c",
          "res4d_branch2c",
          "res2a_branch2c",
          "res3b_branch2a",
          "res3d_branch2a",
          "res4c_branch2b",
          "res3b_branch2b",
          "res4b_branch2a",
          "res4e_branch2c",
          "res3a_branch2c",
          "res2b_branch2c"
        ]
      },
      {
        "model_name": "ResNet-18",
        "layer": [
          "res5c",
          "res4d_branch2c",
          "res2a_branch2c",
          "res3b_branch2a",
          "res3d_branch2a",
          "res4c_branch2b",
          "res3b_branch2b",
          "res4b_branch2a",
          "res4e_branch2c",
          "res3a_branch2c",
          "res2b_branch2c"
        ]
      }
    ],
    "dump_path": "/myproject"
  }
}
```

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。

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

现象描述

调用aclrtRecordEvent接口在Stream中记录一个Event时，日志中的报错如下，红框中是关键日志信息，提示Event ID申请失败：

```
[ERROR] RUNTIME(18401,main):2020-05-29-17:53:11.569.381 18408 StreamSynchronize:stream Synchronize failed
[ERROR] RUNTIME(18401,main):2020-05-29-17:53:11.569.404 18408 StreamSynchronize:stream Synchronize failed, error = 14
[ERROR] ASCENDCL(18401,main):2020-05-29-17:53:11.569.427 18408 aclDvppDestroyChannel:acl/single_op/dvpp/channel.cpp:193: "synchronize stream failed, result =
[INFO] RUNTIME(18401,main):2020-05-29-17:53:11.569.453 18407 ReceivingRun:report[0].task_id=32772
[INFO] RUNTIME(18401,main):2020-05-29-17:53:11.569.473 18407 TryDeRecordedTask:del public task from stream, stream_id=835, tailTaskId=32772, delTaskId=32772,
[ERROR] RUNTIME(18401,main):2020-05-29-17:53:11.569.491 18407 TaskFinished:device_id=0, stream_id=835, sq_id=835, task_id=32772, task_type=7,task_finish_num=4
[ERROR] RUNTIME(18401,main):2020-05-29-17:53:11.569.517 18408 StreamDestroy:stream is not in current ctx
[ERROR] RUNTIME(18401,main):2020-05-29-17:53:11.569.536 18408 StreamDestroy:destroy stream failed, error = 6
[ERROR] ASCENDCL(18401,main):2020-05-29-17:53:11.569.553 18408 DestroyMotiifyAndStream:acl/single_op/dvpp/channel.cpp:41: "fail to destroy stream, ret = 6"
[INFO] ASCENDCL(18401,main):2020-05-29-17:53:11.569.586 18408 aclDvppDestroyChannelDesc:acl/types/dvpp.cpp:401: "destroy DvppChannelDesc info: channelIndex =
mdListLen = 2097152"
[ERROR] DRV(18401,main):2020-05-29-17:53:11.569.787 [devdrv] [drvEventIdAlloc 647] error.
[ERROR] RUNTIME(18401,main):2020-05-29-17:53:11.569.804 18409 EventIdAlloc:drvEventIdAlloc:errorcode = 7
[INFO] RUNTIME(18401,main):2020-05-29-17:53:11.569.819 18409 EventIdAlloc:id = 1
[ERROR] RUNTIME(18401,main):2020-05-29-17:53:11.569.834 18409 Record:event id alloc error, error = 14
[WARNING] RUNTIME(18401,main):2020-05-29-17:53:11.569.849 18409 Record:fail to init record task
[ERROR] RUNTIME(18401,main):2020-05-29-17:53:11.569.865 18409 Synchronize:fail to record
[ERROR] RUNTIME(18401,main):2020-05-29-17:53:11.569.880 18409 StreamSynchronize:stream Synchronize failed
[ERROR] RUNTIME(18401,main):2020-05-29-17:53:11.569.895 18409 StreamSynchronize:stream Synchronize failed, error = 14
[ERROR] ASCENDCL(18401,main):2020-05-29-17:53:11.569.911 18409 aclDvppDestroyChannel:acl/single_op/dvpp/channel.cpp:193: "synchronize stream failed, result =
[ERROR] RUNTIME(18401,main):2020-05-29-17:53:11.569.927 18409 StreamDestroy:stream is not in current ctx
[ERROR] RUNTIME(18401,main):2020-05-29-17:53:11.569.942 18409 StreamDestroy:destroy stream failed, error = 6
[ERROR] ASCENDCL(18401,main):2020-05-29-17:53:11.569.958 18409 DestroyMotiifyAndStream:acl/single_op/dvpp/channel.cpp:41: "fail to destroy stream, ret = 6"
[INFO] ASCENDCL(18401,main):2020-05-29-17:53:11.569.982 18409 aclDvppDestroyChannelDesc:acl/types/dvpp.cpp:401: "destroy DvppChannelDesc info: channelIndex =
mdListLen = 2097152."
```

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的资源时可以复用的，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。

3.3.5 执行单算子产生 coredump 的定位处理

现象描述

单算子执行结束，出现重复释放内存，导致coredump，屏幕显示如[图3-6](#)所示报错。

图 3-6 屏显信息：double free or corruption (!prev)

```
3330 1798 800 4263 2030 2444 1234 132 0 2130 143
[INFO] Write output[0] success. output file = result_files/output_0.bin
[INFO] Run op success
double free or corruption (!prev)
Aborted (core dumped)
[root@localhost out]# 2020-08-18 03:06:28,204 4371 PCOMPILER Master process dead. worker process quitting..
2020-08-18 03:06:28,331 4372 PCOMPILER Master process dead. worker process quitting..
2020-08-18 03:06:28,576 4374 PCOMPILER Master process dead. worker process quitting..
2020-08-18 03:06:28,577 4373 PCOMPILER Master process dead. worker process quitting..
2020-08-18 03:06:28,578 4375 PCOMPILER Master process dead. worker process quitting..
2020-08-18 03:06:28,579 4377 PCOMPILER Master process dead. worker process quitting..
```

可能原因

分析上述日志信息，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，联系华为工程师分析定位。

具体步骤如下：

步骤1 gdb挂载可执行文件。


```
[root@localhost out]# gdb execute_mul_op
GNU gdb (GDB) EulerOS 8.2-3.h4.eulerosv2r8
Copyright (C) 2018 Free Software Foundation, Inc.
License GPLv3+: GNU GPL version 3 or later <http://gnu.org/licenses/gpl.html>
This is free software: you are free to change and redistribute it.
There is NO WARRANTY, to the extent permitted by law.
Type "show copying" and "show warranty" for details.
This GDB was configured as "aarch64-Huawei-linux-gnu".
Type "show configuration" for configuration details.
For bug reporting instructions, please see:
<http://www.gnu.org/software/gdb/bugs/>.
Find the GDB manual and other documentation resources online at:
<http://www.gnu.org/software/gdb/documentation/>.

For help, type "help".
Type "apropos word" to search for commands related to "word"...
```

步骤2 执行gdb调试。

```
(gdb) run
Starting program: /home/xzz/acl_execute_mul_fwkacllib_op_compile_execute/run/out/execute_mul_op
Missing separate debuginfos, use: dnf debuginfo-install glibc-2.28-9.h12.eulerosv2r8.aarch64
[Thread debugging using libthread_db enabled]
Using host libthread_db library "/lib64/libthread_db.so.1".
[New Thread 0xffffb644de0 (LWP 26566)]

Thread 1 "execute_mul_op" received signal SIG61, Real-time event 61.
0x0000ffffbf135408 in kill () from /lib64/libc.so.6
Missing separate debuginfos, use: dnf debuginfo-install libgcc-7.3.0-20190515.h13.eulerosv2r8.aarch64 libstdc++
aarch64
(gdb) c
Continuing.
[New Thread 0xffff9baacde0 (LWP 26647)]
[New Thread 0xffff9ba8bde0 (LWP 26648)]
```

```
[INFO] Run op success
double free or corruption (!prev)
[Thread 0xffff3eff6de0 (LWP 28408) exited]
[Thread 0xffff477f7de0 (LWP 28400) exited]
[Thread 0xffff3e7f5de0 (LWP 28409) exited]
[Thread 0xffff3bfff0de0 (LWP 28414) exited]
[Thread 0xffff47ff8de0 (LWP 28399) exited]
[Thread 0xffff457f3de0 (LWP 28404) exited]
[Thread 0xffff467f5de0 (LWP 28402) exited]
[Thread 0xffff46ff6de0 (LWP 28401) exited]
[Thread 0xffff45ff4de0 (LWP 28403) exited]
[Thread 0xffff44ff2de0 (LWP 28405) exited]
[Thread 0xffff3ff7f7de0 (LWP 28407) exited]
[Thread 0xffff3c7f1de0 (LWP 28413) exited]
[Thread 0xffff3dff4de0 (LWP 28410) exited]
[Thread 0xffff3d7f3de0 (LWP 28411) exited]
[Thread 0xffff3fff8de0 (LWP 28406) exited]
[Thread 0xffff3cff2de0 (LWP 28412) exited]
[Thread 0xffff9baacde0 (LWP 28223) exited]
[Thread 0xffff9ba8bde0 (LWP 28224) exited]
[Thread 0xffff6b53bde0 (LWP 28383) exited]
[Thread 0xffff6ad3ade0 (LWP 28384) exited]

Thread 1 "execute_mul_op" received signal SIGABRT, Aborted.
```

步骤3 查看调用栈。

```
(gdb) bt
#0 0x0000ffffbf135100 in raise () from /lib64/libc.so.6
#1 0x0000ffffbf1364ac in abort () from /lib64/libc.so.6
#2 0x0000ffffbf1707b0 in __libc_message () from /lib64/libc.so.6
#3 0x0000ffffbf170844 in malloc_printerr () from /lib64/libc.so.6
#4 0x0000ffffbf178448 in __int_free () from /lib64/libc.so.6
#5 0x0000ffff9b0f3400 in std::string::Rep::M_dispose (__p=..., this=<optimized out>) at /usr/include/c++/7.3.0/bits/stl_tree.h:1861
#6 std::basic_string<char, std::char_traits<char>, std::allocator<char>>::_basic_string (this=0x12aa9e0, __in_chrg=<optimized out>) at /usr/include/c++/7.3.0/bits/basic_string.h:3621
#7 te::fusion::TbAttrParserParam::TbAttrParserParam (this=0x12aa9e0, __in_chrg=<optimized out>) at tensor_engine/te_fusion/fusion_op.h:78
#8 std::pair<te::ATTR_DTYPE const, te::fusion::TbAttrParserParam>::pair (this=<optimized out>, __in_chrg=<optimized out>) at /usr/include/c++/7.3.0/bits/stl_pair.h:198
#9 __gnu_cxx::new_allocator<std::pair<te::ATTR_DTYPE const, te::fusion::TbAttrParserParam>>::_M_destroy_node (this=<optimized out>, __p=<optimized out>) at /usr/include/c++/7.3.0/ext/new_allocator.h:140
#10 std::allocator_traits<std::allocator<std::pair<te::ATTR_DTYPE const, te::fusion::TbAttrParserParam>>>::_M_destroy_node (this=<optimized out>, __p=<optimized out>) at /usr/include/c++/7.3.0/bits/alloc_traits.h:487
#11 std::pair<te::ATTR_DTYPE const, te::fusion::TbAttrParserParam>::~SelectList<std::pair<te::ATTR_DTYPE const, te::fusion::TbAttrParserParam>>, std::allocator_traits<std::allocator<std::pair<te::ATTR_DTYPE const, te::fusion::TbAttrParserParam>>>::_M_destroy_node (this=<optimized out>, __p=<optimized out>) at /usr/include/c++/7.3.0/bits/stl_tree.h:650
#12 std::pair<te::ATTR_DTYPE const, te::fusion::TbAttrParserParam>::~SelectList<std::pair<te::ATTR_DTYPE const, te::fusion::TbAttrParserParam>>, std::allocator_traits<std::allocator<std::pair<te::ATTR_DTYPE const, te::fusion::TbAttrParserParam>>>::_M_drop_node (this=<optimized out>, __p=<optimized out>) at /usr/include/c++/7.3.0/bits/stl_tree.h:650
```

如果该问题非用户代码问题，联系华为工程师解决。

----结束

3.3.6 进程异常退出后重新执行任务失败

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。

3.3.7 进程异常时资源清理的处理建议

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，会发现如图3-7所示报错。

图 3-7 unbind model stream failed

```
device-0-2020090421510557.log:407494:[ERROR] TSCHE-1.null:2020-09-04 21:51:13.452.762 63696 (cpuid:0) task_scheduler_engine.c:707 proc_model_stream_unbind: unbind model stream failed, stream is running, stream->model_id=512, model_id=512, task_sq_id=514, task_task_id=5
```

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。

4 ATC 常见故障

4.1 背景知识

4.2 常见故障

4.3 常见故障分析与处理

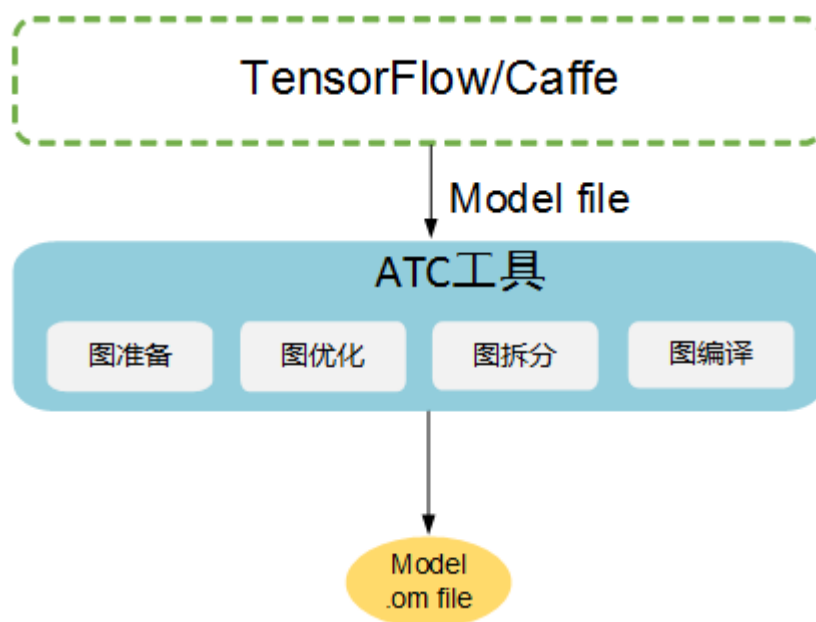
4.1 背景知识

ATC是一个模型转换工具，将开源框架的网络模型（如Caffe、TensorFlow等）以及单算子Json文件转换成昇腾AI处理器支持的离线模型，模型转换过程中可以实现算子调度的优化、权值数据重排、内存使用优化等，可以脱离设备完成模型的预处理。

工具功能架构

ATC工具功能架构如图4-1所示。

图 4-1 ATC 工具功能架构



4.2 常见故障

根据故障场景发生的不同, ATC可能存在以下情况的故障分类:

- 目标模型中包含算子原型库未注册的算子。
- 目标模型中因为未找到算子归属引擎导致转换失败。
- 缺省input_shape参数导致转换失败。
- out_nodes参数指定的输出节点不存在。
- out_nodes参数输入校验失败导致转换失败。
- input_fp16_nodes参数指定的节点名字不存在。
- 算子infershape失败导致模型转换失败。
- 算子找不到算子信息库导致模型转换失败。
- insert_op_conf参数文件路径不存在导致转换失败。
- insert_op_conf参数指定的aipp配置文件校验失败导致转换失败。
- dynamic_image_size或dynamic_batch_size设置分辨率数值过大或档位过多导致运行环境异常缓慢
- 算子信息库中缺少算子

4.3 常见故障分析与处理

4.3.1 目标模型中包含算子原型库未注册的算子

现象描述

模型转换失败, 日志中有类似no OpProto of [XXX] registered的日志, 如图4-2所示。

图 4-2 日志信息: no OpProto of [XXX] registered

```
INFO GE:2020-03-17-01:59:30.359.576 DFSTopologicalSorting:common/graph/.compute_graph.cpp:413:"node_vec.push_back fc1000"
INFO GE:2020-03-17-01:59:30.359.610 DFSTopologicalSorting:common/graph/.compute_graph.cpp:413:"node_vec.push_back prob"
INFO GE:2020-03-17-01:59:30.359.643 DFSTopologicalSorting:common/graph/.compute_graph.cpp:413:"node_vec.push_back Node_Output"
INFO GE:2020-03-17-01:59:30.359.737 DumpGraph:common/graph/.utils/graph_utils.cpp:410:"Start to dump on txt: 4"
WARNING GE:2020-03-17-01:59:30.359.873 SerializeEdge:common/graph/.model_serialize.cpp:68:"Input Anchor PeerOutAnchor Is Empty nodeName data nodeType Data"
WARNING GE:2020-03-17-01:59:31.166.120 EncodeNodeLink:common/graph/.utils/ge_ir_utils.cpp:508:"Input Data Anchor PeerOutAnchor Is Empty nodeName data nodeType Data"
INFO GE:2020-03-17-01:59:31.209.384 DumpGraphToOnnx:common/graph/.utils/graph_utils.cpp:497:"Start to dump ge onnx file: 4"
INFO GE:2020-03-17-01:59:32.030.831 SetDefaultParams:framework/domi/graph/preprocess/insert_op/ge_aipp_op.cc:141:"parse aipp params:input format:1, csc_switch:1."
INFO GE:2020-03-17-01:59:32.030.857 SetDefaultParams:framework/domi/graph/preprocess/insert_op/ge_aipp_op.cc:144:"parse aipp params:mean_chn_0:104, mean_chn_1:117, mean_chn_2:123."
INFO GE:2020-03-17-01:59:32.030.874 SetDefaultParams:framework/domi/graph/preprocess/insert_op/ge_aipp_op.cc:147:"parse aipp params:min_chn_0:0.000000, min_chn_1:0.000000, min_chn_2:0.000000."
WARNING GE:2020-03-17-01:59:32.031.446 CreateOperator:common/graph/.operator_factory_impl.cpp:33:"no OpProto of [aipp] registered"
ERROR GE:2020-03-17-01:59:32.031.472 InsertAippToGraph:framework/domi/graph/preprocess/insert_op/base_insert_op.cc:60:"aipp is not registered Error Code:0x50100001()"
ERROR GE:2020-03-17-01:59:32.031.542 InsertAippOps:framework/domi/graph/preprocess/insert_op/util_insert_aipp_op.cc:83:"insert op to graph failed"
ERROR GE:2020-03-17-01:59:32.031.563 TryDoAipp:framework/domi/graph/preprocess/graph_preprocess.cpp:521:"ErrorNo: 1343242252(Graph which insert dynamic op failed.) "TryDoAipp: insert aipp op ret failed, ret:1343242252"
ERROR GE:2020-03-17-01:59:32.031.580 Prepare:framework/domi/graph/preprocess/graph_preprocess.cpp:650:"ErrorNo: 1343242252(Graph which insert dynamic op failed.) "Run graph prepare fail, ret:1343242252"
ERROR GE:2020-03-17-01:59:32.031.600 PreRun:framework/domi/graph/manager/graph_manager.cpp:259:"ErrorNo: 1343242252(Graph which insert dynamic op failed.) "OMS RunGraph input compute graph is NULL"
ERROR GE:2020-03-17-01:59:32.031.619 StartForRunGraph:framework/domi/graph/manager/graph_manager.cpp:384:"ErrorNo: 1343242252(Graph which insert dynamic op failed.) "PreRun failed."
ERROR GE:2020-03-17-01:59:32.031.639 BuildGraph:framework/domi/graph/manager/graph_manager.cpp:577:"ErrorNo: 1343242268(PreRun failed.) "[BuildGraph] StartForRunGraph failed"
```

可能原因

分析上述日志信息, 可能存在以下故障原因:

aipp算子没有原型库注册。

解决措施

步骤1 确认报错算子，如aipp是否是新增算子，是否有算子原型库，并编译进libopsproto.so。

步骤2 确认是否是版本时间差异。

----结束

4.3.2 目标模型中因为未找到算子归属引擎导致转换失败

现象描述

模型转换失败，报错的日志中有类似Can not find engine of op type Data信息，如图4-3所示。

图 4-3 日志信息：Can not find engine of op type Data

```
is TFKernel
[ERROR] GE:2020-03-26-07:51:40.581.887 GetOpsKernelInfo: ErrorNo: -1(failed) Failed to get opsKernelInfo object by type: Data.
[ERROR] GE:2020-03-26-07:51:40.581.902 Run: ErrorNo: 1343229963(GE is not yet initialized or is finalized.) Can not find engine of op type Data
[ERROR] GE:2020-03-26-07:51:40.581.914 Initialize: ErrorNo: -1(failed) Engine placer run failed.
[ERROR] GE:2020-03-26-07:51:40.581.926 Partition: ErrorNo: 1343242240(Failed to initialize graph.) [GraphPartitioner]: initialize failed
[ERROR] GE:2020-03-26-07:51:40.581.937 PreRun: ErrorNo: -1(failed) Graph partition failed
[ERROR] GE:2020-03-26-07:51:40.581.951 StartForRunGraph: ErrorNo: -1(failed) PreRun failed.
[ERROR] GE:2020-03-26-07:51:40.581.962 BuildGraph: ErrorNo: 1343242268(PreRun failed.) [BuildGraph] StartForRunGraph failed!
[ERROR] GE:2020-03-26-07:51:40.581.974 BuildModel: ErrorNo: 1343266819(Graph manager build graph failed.) graphManager BuildGraph failed, id: 0
[ERROR] GE:2020-03-26-07:51:40.581.986 GenerateModel: ErrorNo: 1343266819(Graph manager build graph failed.) Build model failed
[ERROR] GE:2020-03-26-07:51:40.582.022 GenerateModel: "GE GenerateOfflineModel execute failed"
[ERROR] GE:2020-03-26-07:51:40.582.033 GenerateModel: "ATC Generate execute failed"
[DEBUG] GE:2020-03-26-07:51:40.582.050 Finalize: "Finalizing the FEOpsKernelStore..."
[DEBUG] GE:2020-03-26-07:51:40.582.064 FinalizeSubStore: "Finalizing the subOpsKernelInfoStore begin"
```

可能原因

分析上述日志信息，可能存在以下故障原因：

这个算子的引擎so没有被加载：需要确认这个算子是否有归属引擎，如果有归属引擎，则确认引擎是否没有被正确加载。

解决措施

步骤1 确认报错算子的归属引擎。例如报错日志中的Data是归属于libge_local_engine.so。

步骤2 搜索info日志关键字OPTION_EXEC_EXTERN_PLUGIN_PATH，后面的内容就是加载引擎的路径，查看环境上是否有这个路径；或者搜WARNING日志是否出现Failed to get realpath of日志。

步骤3 如果路径加载不正确，确认是否配置了ASCEND_ENGINE_PATH环境变量，若配置了则代码中会取该环境变量中的路径下去加载，需确保环境变量配置的路径正确。

----结束

4.3.3 缺省 input_shape 参数导致转换失败

现象描述

模型转换失败，日志中有“input x shape is empty”的内容信息。如图4-4所示。

图 4-4 日志信息：“input x shape is empty”

```
are::InferOriginFormat1 is (9047) micro second.
[ERROR] GE(9478,atc):2020-07-02-00:10:08.261.524 [ops/built-in/op_proto/nn_calculation_ops.cpp:2612][OP_PROTO] Conv2DVerify:2612 OpName:[InceptionV3/InceptionV3/Mixed_Sc/Branch_0/Conv2D_0a_1x1/Conv2D] "input x shape is empty."
[ERROR] GE(9478,atc):2020-07-02-00:10:08.261.559 [common/graph/.shape_refiner.cc:338]9478 InferShapeAndType: ErrorNo: -1(failed) Verifying InceptionV3/InceptionV3/Mixed_Sc/Branch_0/Conv2D_0a_1x1/Conv2D failed.
[ERROR] GE(9478,atc):2020-07-02-00:10:08.261.577 [framework/domi/graph/passes/infer_shape_pass.cc:27]9478 Run: ErrorNo: 1343242270(Prepare Graph infer_shape failed) infer_shape failed. node: InceptionV3/InceptionV3/Mixed_Sc/Branch_0/Conv2D_0a_1x1/Conv2D
[ERROR] GE(9478,atc):2020-07-02-00:10:08.261.591 [framework/domi/graph/passes/base_pass.cc:88]9478 RunPasses: ErrorNo: 1343225860(Internal errors) Failed to process pass InferShapePass on node InceptionV3/InceptionV3/Mixed_Sc/Branch_0/Conv2D_0a_1x1/Conv2D, result 1343242270, the passes will be terminated immediately.
[ERROR] GE(9478,atc):2020-07-02-00:10:08.261.608 [framework/domi/graph/passes/base_pass.cc:216]9478 RunPassesOneGraph: ErrorNo: 1343242270(Prepare Graph infer_shape failed) Ex
```


可能原因

分析上述日志信息，可能存在以下故障原因：

原始的TensorFlow模型输入是动态shape，如“ $? \times 299 \times 299 \times 3$ ”，转换模型的时候缺省了input_shape选项导致报错。

解决措施

步骤1 使用Netron工具查看模型，确认模型输入的shape。

步骤2 对于动态shape的模型，转换时必须加上input_shape参数。命令行举例：

```
atc --model=./resnetv2.pb --framework=3 --input_shape="input_tensor:8,299,299,3" --input_format="NHWC" --output=./resnetv2 --soc_version=Ascend310
```

----结束

4.3.4 out_nodes 参数指定的输出节点不存在

现象描述

模型转换失败，报错日志中有“can not find node: *”或者“node * not found”的内容信息。如图4-5所示。

图 4-5 日志信息：can not find node

```
[DEBUG] GE:2020-04-03-03:21:34.848.601 AddEdges:Start add edge: From data1:0 to conv1:0.
[ERROR] GE:2020-04-03-03:21:34.848.619 AddEdgeForUserOutNodes: ErrorNo: 50331649() Can not find out node:conv2, you should check --out_nodes
[ERROR] GE:2020-04-03-03:21:34.848.632 Parse:"Caffe parser add edges for user out nodes failed."
[ERROR] GE:2020-04-03-03:21:34.848.669 ParseGraph:"ATC model parse ret fail. Error Code:0x3000001(Parameter's invalid!)"
[ERROR] GE:2020-04-03-03:21:34.848.700 GenerateModel:"ATC Parse graph dom1:FAILED"
[ERROR] GE:2020-04-03-03:21:34.848.712 GenerateModel:"ATC Generate execute failed"
```

可能原因

分析上述日志信息，可能存在以下故障原因：

- 转模型使用了out_nodes参数指定输出节点（算子名称），但是指定的输出节点在Graph中不存在。
- 指定的输出节点在Graph中存在，但是依然报找不到，可能是这个节点在转换过程中被融合了。

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：

步骤1 检查Graph中是否存在out_nodes参数指定的输出节点。

如果指定的节点在Graph中不存在，需要重新指定正确的节点。

步骤2 如果Graph中有指定的输出节点，依然报找不到错误，可以通过以下方法确认该节点在转换过程中是否被融合了：

尝试先不带out_nodes参数进行转换，然后最后通过查看最终build的图是否有这个节点，若没有则说明已经被融合掉，这种情况需要重新指定输出节点。

----结束

4.3.5 out_nodes 参数输入校验失败导致转换失败

现象描述

模型转换失败，报错日志中有“The input format of --out_nodes is invalid”，如图4-6所示。

图 4-6 日志信息：The input format of --out_nodes is invalid

```
ATC is still working now, please wait for a moment.
[ERROR] GE:2020-03-26-08:13:34.634.276 ParseOutNodes: ErrorNo: 50331649() The input format of out_nodes is invalid, the correct format is {opname:index}, while the actual input is conv1.
[ERROR] GE:2020-03-26-08:13:34.634.303 ParseGraph: "ATC Generate parse out nodes fail"
[ERROR] GE:2020-03-26-08:13:34.634.320 GenerateModel: "ATC Parse graph dom1::FAILED"
[ERROR] GE:2020-03-26-08:13:34.634.333 GenerateModel: "ATC Generate execute failed"
ATC run failed, Please check the detail log
```

可能原因

分析上述日志信息，可能存在以下故障原因：

转模型使用了out_nodes参数指定输出节点（算子名称），但未指定第几个输出，输入格式校验错误。

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：

确认out_nodes入参的格式，需要指定第几个输出，例如--out_nodes="node_name1:0;node_name1:1;node_name2:0"

4.3.6 input_fp16_nodes 参数指定的节点名字不存在

现象描述

模型转换失败，报错日志中有“Input parameter[--input_fp16_nodes]'s opname[***] is not exist in model”，如图4-7所示。

图 4-7 日志信息：“Input parameter[--input_fp16_nodes]'s opname[***] is not exist in model”

```
[INFO] GE(4564,atc):2020-09-02-10:53:30.534.185 [framework/dm1/offline/./session/omg.cc:146]4564 CheckInputFp16Nodes:The input fp16 nodes is set dat
[ERROR] GE(4564,atc):2020-09-02-10:53:30.534.283 [framework/dm1/offline/./session/omg.cc:154]4564 CheckInputFp16Nodes: ErrorNo: 1343225857(Parameter's invalid!) [input param
er[--input_fp16_nodes]'s opname[dat] is not exist in model]
[ERROR] GE(4564,atc):2020-09-02-10:53:30.534.604 [framework/dm1/offline/main.cc:946]4564 GenerateModel: ErrorNo: -1(failed) ATC Parse graph dom1::FAILED
[ERROR] GE(4564,atc):2020-09-02-10:53:30.534.622 [framework/dm1/offline/main.cc:947]4564 GenerateModel: ErrorNo: -1(failed) ATC Generate execute failed
[INFO] GE(4564,atc):2020-09-02-10:53:30.534.639 [framework/dm1/graph/execute/graph_execute.cc:118]4564 FreeInOutBuffer:[GraphManager] not malloc buffer.
```

可能原因

分析上述日志信息，可能存在以下故障原因：

转模型使用了input_fp16_nodes参数指定支持输入数据类型为FP16，该参数后面带的是输入节点的名字，该处报错是指这个节点名字没有找到。

📖 说明

如果日志报类似这样的信息：input_fp16_nodes: res5c_relu is not a input node name，表示input_fp16_nodes参数指定的节点不是输入节点。

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：

打开原始model确认输入节点的名字与input_fp16_nodes后面的是否一致。

4.3.7 算子 inershape 失败导致模型转换失败

现象描述

目标模型转换失败，报错日志中有以下类似信息：Run ge_passes inershape for preprocess failed，如图4-8所示。

图 4-8 日志信息：Run ge_passes inershape for preprocess failed

```
root@ubuntu:/var/dlog# grep ERROR host-0/host-0_20200229170651057.log
[ERROR] GE(1537.mg):2020-02-29-17:07:00.868.643 RandomShapeWithDataType:Layer1_Dropout/random_uniform/RandomUniform: get dtype attr failed.
[ERROR] GE(1537.mg):2020-02-29-17:07:00.868.649 callInferFunc:Layer1_Dropout/random_uniform/RandomUniform call infer func. ret: 4284907295
[ERROR] GE(1537.mg):2020-02-29-17:07:00.868.653 InferShapeWithType:Layer1_Dropout/random_uniform/RandomUniform call infer function failed.
[ERROR] GE(1537.mg):2020-02-29-17:07:00.868.657 Run: ErrorNo: 1343242270(Prepare Graph inershape failed) inershape failed, node: Layer1_Dropout/random_uniform/RandomUniform
[ERROR] GE(1537.mg):2020-02-29-17:07:00.868.661 RunPasses: ErrorNo: 1343225866(Internal errors) failed to process pass InferShapePass on node Layer1_Dropout/random_uniform/RandomUniform, result 1343242270, the passes will be terminated immediately.
[ERROR] GE(1537.mg):2020-02-29-17:07:00.868.665 Run: ErrorNo: 1343225866(Internal errors) failed to process passes on node Layer1_Dropout/random_uniform/RandomUniform type RandomUniform, error code: 1343242270
[ERROR] GE(1537.mg):2020-02-29-17:07:00.868.682 InferShapeForPreprocess: ErrorNo: 1343225866(Internal errors) Run ge_passes inershape for preprocess failed, ret:1343225866.
[ERROR] GE(1537.mg):2020-02-29-17:07:02.643.879 FormatAndShapeProcess: ErrorNo: 1343242270(Prepare Graph inershape failed) Prepare Graph inershape failed
[ERROR] GE(1537.mg):2020-02-29-17:07:02.643.902 Preprocess: ErrorNo: 1343242270(Prepare Graph inershape failed) FormatAndShape process failed
[ERROR] GE(1537.mg):2020-02-29-17:07:02.643.988 Prepare: ErrorNo: 1343242270(Prepare Graph inershape failed) Run graph prepare fail, ret:1343242270
[ERROR] GE(1537.mg):2020-02-29-17:07:02.643.911 PreRun: ErrorNo: 1343242270(Prepare Graph inershape failed) ATC RunGraph input compute graph is NULL
[ERROR] GE(1537.mg):2020-02-29-17:07:02.643.924 StartForRunGraph: ErrorNo: 1343242270(Prepare Graph inershape failed) PreRun failed.
[ERROR] GE(1537.mg):2020-02-29-17:07:02.643.917 BuildGraph: ErrorNo: 1343242260(PreRun failed) (BuildGraph) StartForRunGraph failed!
[ERROR] GE(1537.mg):2020-02-29-17:07:02.643.921 BuildModel: ErrorNo: 1343266819(Graph manager build graph failed.) graphManager BuildGraph failed, id: 0
[ERROR] GE(1537.mg):2020-02-29-17:07:02.643.924 GenerateOfflineModel: ErrorNo: 1343266819(Graph manager build graph failed.) Build model failed
[ERROR] GE(1537.mg):2020-02-29-17:07:02.644.023 GenerateModel:framework/dm/offline/main.cpp:917: "GE GenerateOfflineModel execute failed"
[ERROR] GE(1537.mg):2020-02-29-17:07:02.644.033 GenerateModel:framework/dm/offline/main.cpp:916: "ATC Generate execute failed"
root@ubuntu:/var/dlog# ll
total 24
```

可能原因

分析上述日志信息，可能存在以下故障原因：报错算子的输入有问题。

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：

步骤1 定位报错算子实现的inershape失败的原因，确认异常输入的上个算子inershape实现是否正确。

步骤2 不断向上回溯到inershape实现异常的问题算子。

----结束

4.3.8 算子找不到算子信息库导致模型转换失败

现象描述

模型转换失败，报错日志中有“Can't find any supported ops kernel and engine of %s, type is %”，如图4-9所示。

图 4-9 日志信息：Can't find any supported ops kernel and engine of %s, type is %

```
t failed, kernel_name is AicoreEngine, op type is RoiPooling, op name is roi_pooling
[ERROR] GE(26576.atc):2020-03-16-20:40:32.744.432 [framework/dm/engine_manager/dmengine_manager.cc:223]GetDNNEngineName: ErrorNo: 1343242282(assign engine failed) GetDNNEngineName:Op type RoiPooling of ops kernel AicoreEngine is unsupported, reason:Op roi_pooling not supported reason: The dtype, format or shape of input in op desc is not supported in op store, check the dtype, format or shape of input between the op store and the graph. Op store name is the-builtin.
The preCompileFunc is nullptr, check whether this type of op does not exist in the-builtin and then go to the the-plugin, Op store name is the-plugin. The type of this op is not found in op store, check whether the op store has this type of op. Op store name is cce-builtin.
[ERROR] GE(26576.atc):2020-03-16-20:40:32.744.458 [framework/dm/engine_manager/dmengine_manager.cc:226]GetDNNEngineName: ErrorNo: 1343242282(assign engine failed) Can't find any supported ops kernel and engine of roi_pooling, type is RoiPooling
[ERROR] GE(26576.atc):2020-03-16-20:40:32.744.483 [framework/dm/graph/partition/engine_place.cc:54]Run: ErrorNo: 1343229963(GE is not yet initialized or is finalized.) Can not find engine of op type RoiPooling
[ERROR] GE(26576.atc):2020-03-16-20:40:32.744.507 [framework/dm/graph/partition/graph_partition.cc:523]Initialize: ErrorNo: -1(failed) Engine placer
```

可能原因

分析上述日志信息，可能存在以下故障原因：

该算子在算子信息库中没找到。

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：

步骤1 确认这个算子是否在算子信息库中支持。

步骤2 如果这个算子的dtype、format、shape不匹配算子信息库中的信息，则需要具体查看该算子推导出来的过程，排查是否是由于输入的节点异常导致的。

----结束

4.3.9 insert_op_conf 参数文件路径不存在导致转换失败

现象描述

模型转换失败，报错日志中有“insert op config file not found”，如图4-10所示。

图 4-10 日志信息：insert op config file not found

```
[ERROR] GE(9851,atc):2020-07-02-00:41:08.831.825 [framework/domi/offline/./ir_build/atc_ir_common.cc:326]9851 CheckInsertOpConfParamValid: ErrorNo: 1343225857(Parameter's in valid) insert op config file not found: /home/test/aipp.cfg
[ERROR] GE(9851,atc):2020-07-02-00:41:08.831.858 [framework/domi/offline/main.cc:370]9851 CheckFlags: ErrorNo: -1(failed) check insert op conf failed!
```

可能原因

分析上述日志信息，可能存在以下故障原因：

insert_op_conf入参是在aipp场景下才需要的，后面是aipp配置文件的路径，路径访问不到。

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：

步骤1 确认是否是aipp场景，若不是，则不需要此入参。

步骤2 如果是aipp场景，确定路径是否正确，若路径正确则查看属组权限是否正常。

----结束

4.3.10 insert_op_conf 参数指定的 aipp 配置文件校验失败导致转换失败

现象描述

模型转换失败，报错日志中有“Read AIPP conf file error”，如图4-11所示。

图 4-11 日志信息：Read AIPP conf file error

```
[libprotobuf ERROR google/protobuf/text_format.cc:317] Error parsing text-format domi.InsertNewOps: 3:5: Unknown enumeration value of "crop" for field "aipp_mode".
[ERROR] GE(9913,atc):2020-07-02-00:48:05.525.876 [framework/domi/common/util.cc:288]9913 ReadProtoFromText: ErrorNo: 0(success) Parse file[./bak.cfg] through [google::protobuf::TextFormat::Parse] failed. Please check whether the file is a valid protobuf format file.
[ERROR] GE(9913,atc):2020-07-02-00:48:05.525.952 [framework/domi/graph/preprocess/insert_op/util_insert_aipp_op.cc:72]9913 Parse: ErrorNo: -1(failed) Read AIPP conf file error: ./bak.cfg Error Code:0xFFFFFFFF(failed)
[ERROR] GE(9913,atc):2020-07-02-00:48:05.525.976 [framework/domi/graph/preprocess/graph_preprocess.cc:1786]9913 TryDoAipp: ErrorNo: 1343242285(OMG parse dynamic node config file failed.) TryDoAipp: parse config file ./bak.cfg failed
[EVENT] GE(9913,atc):2020-07-02-00:48:05.526.000 [framework/domi/graph/preprocess/graph_preprocess.cc:2087]9913 PrepareDynShape: [GEPERFTRACE] The time cost of Prepare::TryDoAipp is [273] micro second.
[ERROR] GE(9913,atc):2020-07-02-00:48:05.526.024 [framework/domi/graph/preprocess/graph_preprocess.cc:2087]9913 PrepareDynShape: ErrorNo: 1343242285(OMG parse dynamic node config file failed.) failed to process Prepare_TryDoAipp
[EVENT] GE(9913,atc):2020-07-02-00:48:05.526.045 [framework/domi/graph/manager/graph_manager.cc:371]9913 PreRun: [GEPERFTRACE] The time cost of GraphManager::graph_preparer_.PrepareDynShape is [7534] micro second.
[ERROR] GE(9913,atc):2020-07-02-00:48:05.526.103 [framework/domi/graph/manager/graph_manager.cc:371]9913 PreRun: ErrorNo: 1343242285(OMG parse dynamic node config file failed.) Failed to process GraphManager_graph_preparer_.PrepareDynShape
```

可能原因

分析上述日志信息，可能存在以下故障原因：
insert_op_conf入参指定的aipp文件解析失败。

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：
确认aipp配置文件是否正确，是否是protobuf文件格式。

4.3.11 dynamic_image_size 或 dynamic_batch_size 设置分辨率数值过大或档位过多导致运行环境异常缓慢

现象描述

执行动态shape模型转换时，Host侧运行环境缓慢，无法执行其他操作，需重新启动昇腾处理器才能恢复正常，查看日志有如下类似信息，如图4-12所示。

图 4-12 一直打印等待日志

```
[INFO] TEFUSION:2020-04-17-11:52:30.311.097 WaitAllFinished Get finished compilation task list size[0]. graphId[140048978401024]
[INFO] TEFUSION:2020-04-17-11:52:30.577.297 WaitAllFinished Get finished compilation task list size[0]. graphId[140048961615616]
[INFO] TEFUSION:2020-04-17-11:52:30.833.047 WaitAllFinished Get finished compilation task list size[0]. graphId[140048953222912]
[INFO] TEFUSION:2020-04-17-11:52:31.106.662 WaitAllFinished Get finished compilation task list size[0]. graphId[140048970008320]
[INFO] TEFUSION:2020-04-17-11:52:31.315.255 WaitAllFinished Get finished compilation task list size[0]. graphId[140048928044800]
[INFO] TEFUSION:2020-04-17-11:52:31.415.745 WaitAllFinished Get finished compilation task list size[0]. graphId[140048886081280]
[INFO] TEFUSION:2020-04-17-11:52:31.520.568 WaitAllFinished Get finished compilation task list size[0]. graphId[140048877688576]
[INFO] TEFUSION:2020-04-17-11:57:41.538.210 WaitAllFinished Get finished compilation task list size[0]. graphId[140048953222912]
```

可能原因

动态shape模型转换时，算子编译会占用较大内存，如果Host主机内存配置不够，会出现Host侧内存占用光，且将swap交换区间作为内存继续处理，造成IO写满，导致Host运行环境变慢。

解决措施

针对执行动态shape模型转换场景下，尤其是dynamic_image_size或dynamic_batch_size设置分辨率数值过大或档位过多，建议执行**swapoff -a**命令关闭swap交换区间作为内存的功能，防止出现由于内存不足，将swap交换空间作为内存继续调用，导致运行环境异常缓慢的情况。

4.3.12 算子信息库中缺少算子

现象描述

执行模型转换时，报错日志信息：“The type of this op is not found in op store”，如图4-13所示。

图 4-13 日志信息：The type of this op is not found in op store

```
[ERROR] GE(18669.python3):2020-07-03-03:52:51.235.767 [framework/domi/engine_manager/dnnengine_manager.cc:241]18817 GetDNNEngineName: ErrorNo: 1343242282(assign engine failed) GetDNNEngineName:Op type Range of ops kernel AicoreEngine is unsupported, reason:Op range not supported reason: The type of this op is not found in op store, check whether the op store has this type of op. Op store name is the builtin. The precompilerunc is nullptr, check whether this type of op does not exist in the builtin and then go to the the-plugin. Op store name is the-plugin.
[ERROR] GE(18669.python3):2020-07-03-03:52:51.235.788 [framework/domi/engine_manager/dnnengine_manager.cc:246]18817 GetDNNEngineName: ErrorNo: 1343242282(assign engine failed) Can't find any supported ops kernel and engine of range, type is Range
[ERROR] GE(18669.python3):2020-07-03-03:52:51.235.800 [framework/domi/graph/partition/engine_place.cc:55]18817 Run: ErrorNo: 1343229963(GE is not yet initialized or is finalized.) Can not find engine of op type Range
[ERROR] GE(18669.python3):2020-07-03-03:52:51.235.812 [framework/domi/graph/partition/graph_partition.cc:504]18817 Initialize: ErrorNo: -1(failed) Engine placer run failed.
[ERROR] GE(18669.python3):2020-07-03-03:52:51.235.847 [framework/domi/graph/partition/graph_partition.cc:855]18817 PartitionSubGraph: ErrorNo: 1343242248(Failed to initialize graph.) [GraphPartitioner]: initialize failed
[ERROR] GE(18669.python3):2020-07-03-03:52:51.235.870 [framework/domi/graph/partition/graph_partition.cc:826]18817 Partition: ErrorNo: -1(failed) Sub graph partition failed.
[ERROR] GE(18669.python3):2020-07-03-03:52:51.235.880 [framework/domi/graph/manager/graph_manager.cc:2326]18817 OptimizeSubgraph: ErrorNo: -1(failed) Graph partition failed.
[ERROR] GE(18669.python3):2020-07-03-03:52:51.235.900 [framework/domi/graph/manager/graph_manager.cc:392]18817 PreRun: ErrorNo: -1(failed) Failed to process GraphManager OptimizeSubgraph
```

可能原因

分析上述日志信息，可能原因：

aicore或aicpu的算子信息库中没有该算子，当前产品不支持该算子。

处理步骤

针对分析的故障可能原因，当前版本并不支持该算子，可以通过以下方式确认：

检查\${INSTALL_DIR}/opp/op_impl/built-in/aicpu/tf_kernel/config/tf_kernel.json或\${INSTALL_DIR}/opp/op_impl/built-in/ai_core/tbe/config/<soc_version>文件，确认执行的算子是否在aicpu或aicore算子信息库中，如果不存在，说明当前不支持该算子，不可以使用。

说明

\${INSTALL_DIR}表示安装路径，请根据实际环境替换。

5 GE 常见故障

5.1 输入数据的大小远小于模型输入的大小导致推理失败

5.2 输入数据的个数大于模型输入的个数导致推理失败

5.1 输入数据的大小远小于模型输入的大小导致推理失败

现象描述

用户通过acl调用GE的executor接口进行推理时，报错日志有“Input size can not be smaller than op size”，如图5-1所示。

图 5-1 日志信息：Input size can not be smaller than op size

```
The time cost of GraphLoader::InitBeHandle is [548] micro second, call num is 171
[EVENT] GE(7475,main):2020-07-02-16:04:04.340.080 [framework/domi/executor/./graph/load/new_model_manager/davinci_model.cc:581]7475 Init:[GEPERFTRACE] The
time cost of GraphLoader::DoTaskSink is [98085] micro second.
[ERROR] GE(7475,main):2020-07-02-16:04:04.340.963 [framework/domi/executor/./graph/load/new_model_manager/davinci_model.cc:2830]7475 CheckInputAndModelSize
: ErrorNo: -1(failed) Input size [63488] can not be smaller than op size [602144] after 64-byte alignment
[ERROR] GE(7475,main):2020-07-02-16:04:04.340.981 [framework/domi/executor/./graph/load/new_model_manager/davinci_model.cc:2902]7475 UpdateIoTaskArgs: Error
No: -1(failed) Check input size and model size failed
[ERROR] GE(7475,main):2020-07-02-16:04:04.340.992 [framework/domi/executor/./graph/load/new_model_manager/davinci_model.cc:2846]7475 CopyModelData: ErrorNo
: 1343225857(Parameter's invalid!) [ZCPV] Update input data to model failed.
[ERROR] GE(7475,main):2020-07-02-16:04:04.341.002 [framework/domi/executor/./graph/load/new_model_manager/davinci_model.cc:3260]7475 NnExecute: ErrorNo: -1
(failed) Copy input data to model failed.
[ERROR] GE(7475,main):2020-07-02-16:04:04.341.011 [framework/domi/executor/./graph/load/graph_loader.cc:325]7475 ExecuteModel: ErrorNo: 1343225860(Internal
errors) Execute model failed, model_id:1.
[ERROR] ASCENDCL(7475,main):2020-07-02-16:04:04.341.022 [acl/model/model.cpp:575]7475 ModelExecute:acl/model/model.cpp:575: "Execute model failed, ge result
[1343225860], modelId[1]"
[ERROR] ASCENDCL(7475,main):2020-07-02-16:04:04.341.034 [acl/model/model.cpp:593]7475 aclmdlExecute:acl/model/model.cpp:593: "aclmdlExecute failed, result[5
80802], modelId[1]"
[EVENT] TDT(7475,main):2020-07-02-16:04:04.359.840 [tdt/host/./common/src/log.cpp:149][TsdClient] Close [phyDeviceId=0][logicDeviceId=0] [sessionId=1] hcc
p and computer enter,[tdt/host/src/tsd/tsd_client.cpp:921:Close]7475
```

可能原因

分析上述日志信息，可能存在以下故障原因：

用户申请的input、output内存的大小与模型输入和输出size的大小做校验不一致，该日志信息表明用户申请的数据小于离线模型中节点的size。

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：

步骤1 通过报错size大小，可以确认是用户申请的输入或输出size远小于模型输入或输出的size。

步骤2 确认后查看调用acl处，数据文件和模型是否有误，或者数据类型有误。

----结束

5.2 输入数据的个数大于模型输入的个数导致推理失败

现象描述

用户通过acl调用GE的executor接口进行推理时，报错日志有“Verify input data num failed”，如图5-2所示。

图 5-2 日志信息：Verify input data num failed

```
[EVENT] RUNTIME(15633,demo_acl):2020-07-03-00:30:48.962.569 [runtime/feature/src/logger.cc:1096]15636 TaskLaunchedEx:device_id=0, stream_id=512, sq_id=512, task_id=2, task_type=0, task_launched_num=17
[ERROR] GE(15633,demo_acl):2020-07-03-00:30:48.963.547 [framework/domi/executor/./graph/load/new_model_manager/davinci_model.cc:2867]15634 UpdateIoTaskArgs: ErrorNo: -1(failed) Verify input data num failed: model requires 2, but user actually feeds 3
[ERROR] GE(15633,demo_acl):2020-07-03-00:30:48.963.566 [framework/domi/executor/./graph/load/new_model_manager/davinci_model.cc:2833]15634 CopyModelData: ErrorNo: 1343225857(Parameter's invalid!) [ZCPY] Update input data to model failed.
[ERROR] GE(15633,demo_acl):2020-07-03-00:30:48.963.580 [framework/domi/executor/./graph/load/new_model_manager/davinci_model.cc:3238]15634 NnExecute: ErrorNo: -1(failed) Copy input data to model failed.
[ERROR] GE(15633,demo_acl):2020-07-03-00:30:48.963.605 [framework/domi/executor/./graph/load/graph_loader.cc:325]15634 ExecuteModel: ErrorNo: 1343225860(Internal errors) Execute model failed, model id:1
[ERROR] ASCENDCL(15633,demo_acl):2020-07-03-00:30:48.963.619 [acl/model/model.cpp:565]15634 ModelExecute:acl/model/model.cpp:565: "Execute model failed[1343225860], modelId[1]"
[ERROR] ASCENDCL(15633,demo_acl):2020-07-03-00:30:48.963.634 [acl/model/model.cpp:587]15634 aclmdlExecute:acl/model/model.cpp:587: "aclmdlExecute failed[500002], modelId[1]"
```

可能原因

分析上述日志信息，可能存在以下故障原因：

离线模型输入和推理时所用的数据文件个数不一致，导致校验报错。

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：

步骤1 确认加载的离线模型是否正确。

步骤2 如果离线模型确认正确，则确认推理时输入的数据文件个数是否正确。

----结束

6 DVPP 常见故障

6.1 背景知识

6.2 常见故障

6.3 常见故障分析与处理

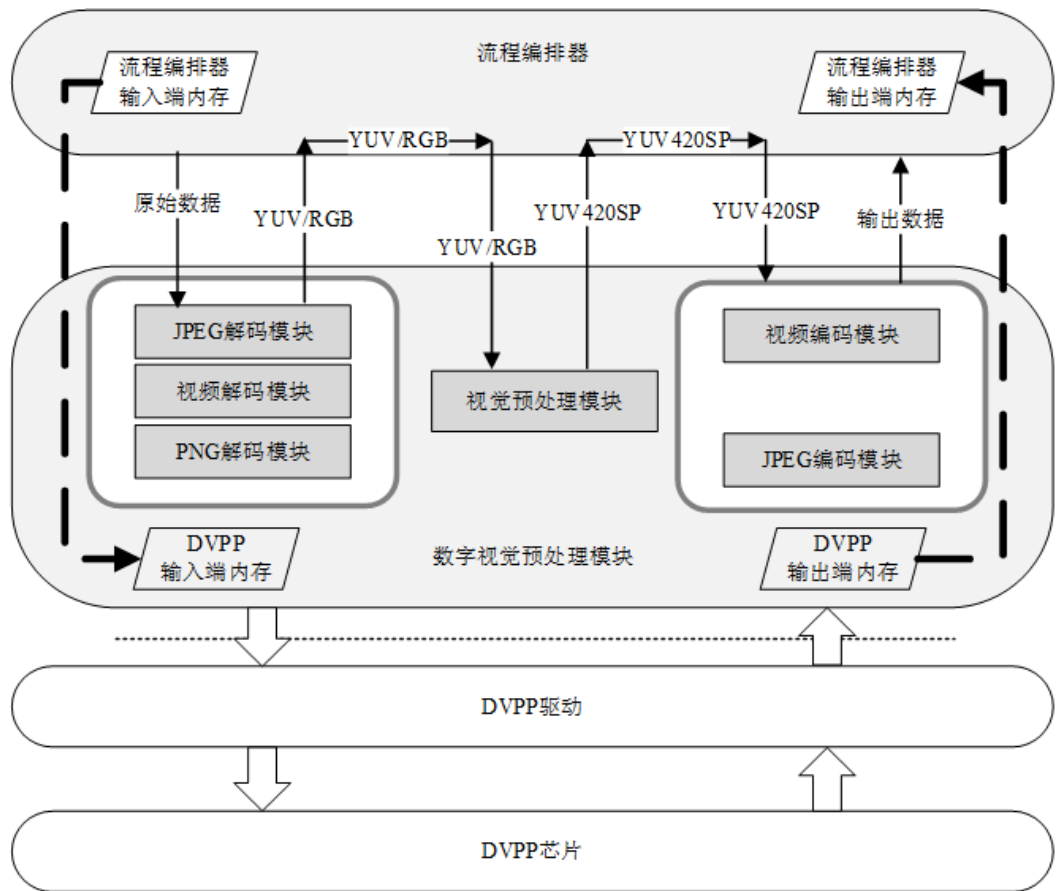
6.1 背景知识

数字视觉预处理模块（DVPP）作为整个软件流执行过程中的编解码和图像转换模块，为神经网络发挥着预处理辅助功能。当来自系统内存和网络的视频或图像数据进入昇腾AI芯片的计算资源中运算之前，由于达芬奇架构对输入数据有固定的格式要求，如果数据未满足架构规定的输入格式、分辨率等要求，就需要调用数字视觉处理模块进行格式的转换，才可以进行后续的神经网络计算步骤。

DVPP全称Digital Vision Pre-Processor，主要实现视频解码（VDEC）、视频编码（VENC）、JPEG编解码（JPEGD/JPEGE）、PNG解码（PNGD）、图像预处理（VPC）。

DVPP的执行流程如图6-1所示，需要由流程编排器、数字视觉处理模块、DVPP驱动和DVPP硬件模块共同协作完成。位于框架最上层是流程编排器，负责调度DVPP中的功能模块进行相应处理以及管理数据流。数字视觉预处理模块位于功能架构的中上层，为流程编排器提供调用视频图形处理模块的编程接口，通过这些接口可以配置编解码或视觉预处理模块的相关参数。DVPP驱动位于功能架构的中下层，最贴近于DVPP的硬件模块，主要负责设备管理、引擎管理和引擎模组的驱动，DVPP驱动会根据数字视觉预处理模块下发的任务，分配对应的DVPP硬件引擎，同时还对硬件模块中的寄存器进行读写，完成其它一些硬件初始化工作。最底层的是真实的硬件计算资源DVPP硬件模块，是一个独立于昇腾AI处理器中其它模块的一个单独的专用加速器，专门负责执行与图像和视频相对应的编解码和预处理任务。

图 6-1 DVPP 功能架构



6.2 常见故障

DVPP故障场景有解码失败、异常码流、参数错误、使用限制等，如：

- 视频解码丢帧/丢包
- 视频解码花屏
- 视频解码性能问题
- 视频解码失败
- 视频帧解码失败不触发回调函数
- 设置帧序号但回调接收的hiai_data对象却为空
- JPEG图片解码失败
- VPC调用失败
- 多次调用VPC触发硬件保护机制，导致VPC所有引擎不可用
- 模型推理精度下降
- VDEC或VPC超时说明
- DVPP输入和输出内存申请方式说明
- HFBC格式介绍

6.3 常见故障分析与处理

6.3.1 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 havn'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][InquiresSlceProperty: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 havn'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).
```

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VdecCtl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《[DVPP API参考（开放态, Ascend310）](#)》中的VDEC示例代码，找到对应的视频解码的代码参考优化。

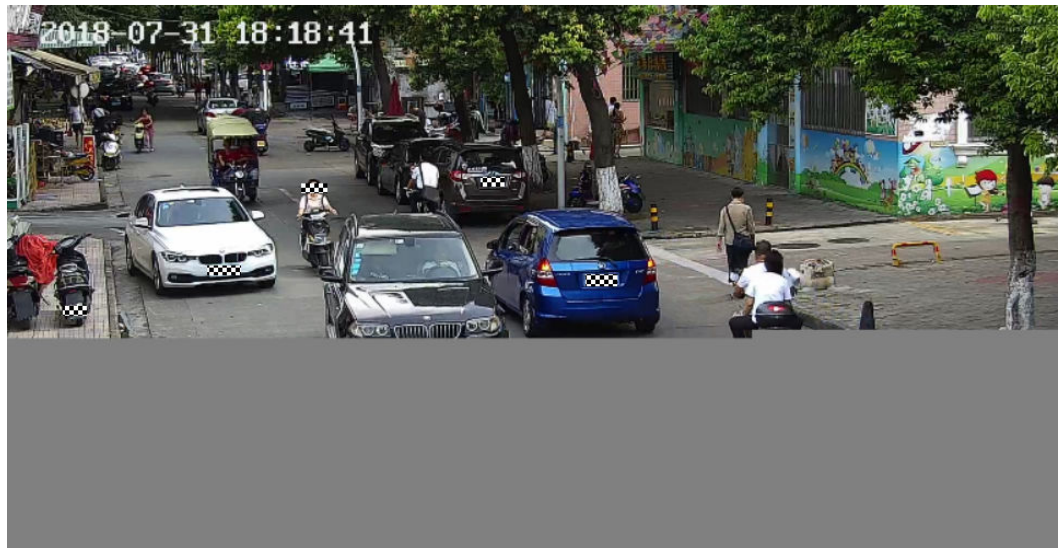
----结束

6.3.2 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6-2](#)所示。并且日志中存在类似“dvpp_vdec_vdm_process failed”、“Chan 0 ErrRatio = 44”、“Vdm ioctl fail, errno != EAGAIN. No retry.”信息。

图 6-2 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VdecCtl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《[DVPP API参考（开放态, Ascend310）](#)》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

6.3.3 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

6.3.4 视频解码失败

现象描述

视频失败，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 equal *
[ERROR] DVPP(13757,dvpp_performance):2020-12-31-21:18:26.998.657 [VDEC][ProcessSPS:9055][T26] bit_depth_luma(10) not equal 8.
[ERROR] DVPP(13757,dvpp_performance):2020-12-31-21:18:26.998.763 [VDEC][ProcessSPS:9070][T26] bit_depth_chroma(10) not equal 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。

6.3.5 视频解码格式不支持

现象描述

原始视频是mp4格式，每一帧都能正常播放，但是送入DVPP解码的每一帧均上报格式不支持错误，导致解码异常。日志信息如下：

```
[ERROR] DVPP(19789,aicpu_scheduler):2021-11-23-19:06:47.901.870 [omx_vdec.cpp:1101][VDEC]
[MessageProcessBranchEx:1101] [T208] ERROR: VDEC_EVT_REPORT_SYNTAX_ONEFAME_FAIL, chanId =
0, frame index = 187
[ERROR] DVPP(19789,aicpu_scheduler):2021-11-23-19:06:47.901.933 [omx_vdec.cpp:2158][VDEC]
[EventProcessBranchEx:2158] [T208] OMX_ErrorSyntaxOneFrameFail ERROR, chanId = 0, frame index = 187
[ERROR] DVPP(19789,aicpu_scheduler):2021-11-23-19:06:47.904.481 [frame_tracer.c:103][VDEC]
[ReportFrameFailure:103] [T208] FRAME_TRACE [0] ID:[188] [2 0 0 0 3] [2 254] [189 0 0] [188(0 187 0 0)]
[ERROR] DVPP(19789,aicpu_scheduler):2021-11-23-19:06:47.904.532 [omx_vdec.cpp:1101][VDEC]
[MessageProcessBranchEx:1101] [T208] ERROR: VDEC_EVT_REPORT_SYNTAX_ONEFAME_FAIL, chanId =
0, frame index = 188
[ERROR] DVPP(19789,aicpu_scheduler):2021-11-23-19:06:47.904.599 [omx_vdec.cpp:2158][VDEC]
[EventProcessBranchEx:2158] [T208] OMX_ErrorSyntaxOneFrameFail ERROR, chanId = 0, frame index =
188[ERROR] DVPP(19789,aicpu_scheduler):2021-11-23-19:06:47.906.386 [frame_tracer.c:103][VDEC]
[ReportFrameFailure:103] [T208] FRAME_TRACE [0] ID:[189] [2 0 0 0 3] [2 254] [190 0 0] [189(0 188 0 0)]
[ERROR] DVPP(19789,aicpu_scheduler):2021-11-23-19:06:47.906.470 [omx_vdec.cpp:1101][VDEC]
[MessageProcessBranchEx:1101] [T208] ERROR: VDEC_EVT_REPORT_SYNTAX_ONEFAME_FAIL, chanId =
0, frame index = 189
```

可能原因

导致该错误的可能原因如下：

- 送入的裸码流为芯片不支持的H261等格式；或者解码器配置为H264协议，但送入码流却为H265格式，反之亦然。
- 裸码流是H264或H265，但目前只支持annex-B格式的裸码流，不支持AVCC格式的裸码流。

annex-B格式裸码流，码流样式如下，固定以0x00000001开始(H264为例)：

```
<pre><code>0x0000 | 00 00 00 01 67 64 00 0A AC 72 84 44 26 84 00 00
```

AVCC格式裸码流，每一个NALU包都加上了一个指定其长度(NALU包大小)的前缀，样式如下（H264为例）：

```
0x0000 | 01 64 00 0A FF E1 00 19 67 64 00 0A AC 72 84 44
0x0010 | 26 84 00 00 03 00 04 00 00 03 00 CA 3C 48 96 11
```

处理步骤

针对上述2种可能的错误原因，请分别参考下面方法处理：

步骤1 使用ffmpeg命令，检查封装的码流是否为DVPP支持格式，目前仅支持H264/H265。

命令格式：ffmpeg -i [input file name]

如输入文件FHD_1frame_1920x1080.mp4:

ffmpeg -i FHD_1frame_1920x1080.mp4

```
ffmpeg version 4.3.1 Copyright (c) 2000-2020 the FFmpeg developers
built with gcc 7 (Ubuntu 7.5.0-3ubuntu1-18.04)
configuration: --enable-shared
libavutil      56. 51.100 / 56. 51.100
libavcodec     58. 91.100 / 58. 91.100
libavformat    58. 45.100 / 58. 45.100
libavdevice    58. 10.100 / 58. 10.100
libavfilter     7. 85.100 / 7. 85.100
libswscale     5.  7.100 / 5.  7.100
libswresample  3.  7.100 / 3.  7.100
Input #0: mov,mp4,m4a,3gp,3g2,mj2, from 'FHD_1frame_1920x1080.mp4':
Metadata:
  major_brand      : isom
  minor_version    : 512
  compatible_brands: isomiso2avc1mp41
  encoder         : Lavf58.10.100
  Duration: 00:00:00.04, start: 0.000000, bitrate: 12755 kb/s
  Stream #0:0(und): Video: h264 (High) (avc1 / 0x31637661), yuv420p, 1920x1080, 12595 kb/s, 25 fps, 25 tbr, 12800 tbn, 50 tbc (default)
  Metadata:
    handler_name    : VideoHandler
```

该命令会显示相应的输入码流样式，如果不是当前支持码流，需要更换码流；如果支持，用户排查代码配置协议参数video_format和实际码流格式是否相符。

步骤2 针对mp4等解封装后的H264/H265裸码流，如果芯片不支持，建议调用ffmpeg接口进行裸码流转码，接口调用样式如下（代码不能直接使用，需要根据实际情况调整，注意防止内存泄漏）：

```
AVBitStreamFilterContext* pBSFContext_ = av_bitstream_filter_init("h264_mp4toannexb"); // H265为
hevc_mp4toannexb
tmpPkt = packet;
ret = av_bitstream_filter_filter(pBSFContext_, pCodecCtx_, nullptr, &tmpPkt.data, &tmpPkt.size, packet.data,
packet.size, packet.flags & AV_PKT_FLAG_KEY);
if (ret > 0)
{
    av_free_packet(&packet);
    tmpPkt.buf = av_buffer_create(tmpPkt.data, tmpPkt.size, av_buffer_default_free, nullptr, 0);
}
packet = tmpPkt; // 新的数据包packet
av_bitstream_filter_close(h264bsfc);
```

----结束

6.3.6 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debug，查看下述四条日志打印的总次数是否和输入的帧数相等。

- (1) "The non-intelligent pointer callback interface start, channelId = 0"

- (2) "The non-intelligent pointer callback interface end, channelId = 0"
- (3) "The smart pointer callback interface start, channelId = 0"
- (4) "The smart pointer callback interface end, channelId = 0"

生成上述4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, 回调函数执行成功后打印 (2) 日志。
- 每一帧对应设置一个hiai_data_sp, 成功解码返回, 打印 (3) 日志, 然后调用用户注册的回调函数, 回调函数执行成功后打印 (4) 日志。
- N帧对应设置一个hiai_data_sp, 第1帧成功解码返回, 打印 (3) 日志, 然后调用用户注册的回调函数, 回调函数执行成功后打印 (4) 日志。

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

6.3.7 设置帧序号但回调接收的 hiai_data 对象却为空

现象描述

回调函数传给用户的形参hiai_data为NULL。

回调函数格式: void(*call_back)(FRAME* frame,void * hiai_data)

可能原因

形参hiai_data为NULL的可能原因有以下:

- VDEC本身支持多帧码流一次性调用VdecCtl送入, 但用户若设置了hiai_data_sp对象, 就只能每次调用VdecCtl送入一帧。而此问题可能是用户设置了一个hiai_data_sp对象, 但此对象对应多帧码流一次性调用VdecCtl送入导致。
- 码流参考帧间隔超过30帧, 参考帧周期过长导致帧序号小的hiai_data_sp对象从队列中被丢弃。

处理步骤

针对分析的可能原因, 请参考以下方法处理:

检查日志中是否出现 “not find node, currFid = ***, reqFid = ***, chanid = ***” 内容信息, 若有则说明是第1条可能原因所致, 请排查传入数据正确性。

6.3.8 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) :

```
Unsupported subsample format, just support jpeg with YUV 444 440 422 420 400
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
JPEG D只支持huffman编码(colorspace: yuv, subsample: 444/440/422/420/400), 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

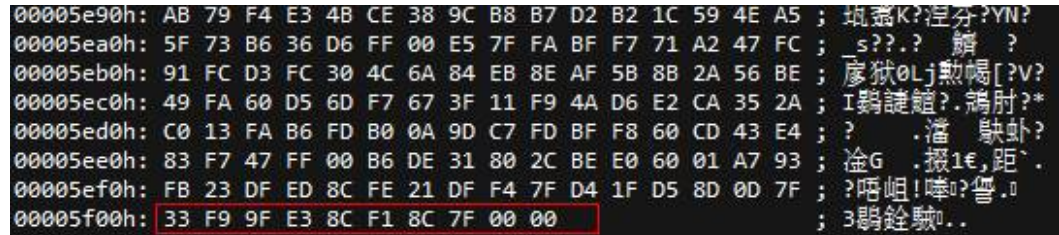
处理步骤

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。



- 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DvppCtl之前, 通过fwrite函数将传输给JPEG D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

6.3.9 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(0xaaaadeccdc0), bareDataBufferSize(3133440) should be allocated by acldvppMalloc!
```

日志信息 (5) :

```
Both RoiNum(1) outputAddr(0xaaaadeccdc0) 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《[DVPP API参考 \(开放态, Ascend310\)](#)》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

6.3.10 多次调用 VPC 触发硬件保护机制，导致 VPC 所有引擎不可用

现象描述

代码存在bug的情况下，在调试过程中连续调用VPC模块处理失败，查看日志有如下类似错误信息：

```
p_engine_para->fail_count< (15)
```

修复代码bug后，VPC处理仍然失败，需重启昇腾处理器才能恢复。

除了VPC模块，DVPP其他模块也存在这个问题，即出现"`p_engine_para->fail_count<(15)`"的错误打印，都需要重启昇腾处理器才能恢复。

可能原因

对于VPC模块，可能由于内存地址异常，多次调用VPC，触发硬件保护机制，导致VPC所有引擎不可用。

DVPP其他模块可能是由于码流或者参数配置有问题。

定位思路

对于VPC模块，检查VPC输入输出内存存在调用VPC处理之前是否已经释放，是否变成野指针在使用。

对于DVPP其他模块，检查码流和参数配置是否正确。

处理步骤

- 步骤1** 检查输出地址是否存在野指针情况，以及检查码流和参数配置是否正确。

例如：将JPEGD的输出内存作为VPC的输入内存时，这个内存释放应该在VPC之后，不能在调用JPEGD之后就进行释放，导致VPC输入内存地址成为野指针。

- 步骤2** 重启系统。

----结束

6.3.11 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AIIPP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：

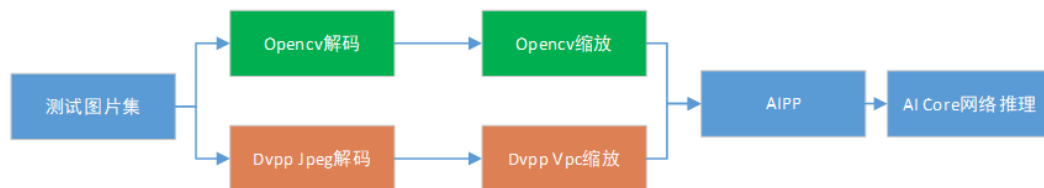
- DVPP的JPEG模块解码算法使用较新的libjpeg-turbo算法（Opencv3.4.4及以上），而已经训练好的模型，一般是使用libjpeg算法（Opencv2.4.9中的解码算法）训练得到的，造成这个模型不是最适配DVPP解码的模型。
- DVPP的VPC缩放算法采用的是自研算法，而已经训练好的模型，一般是使用双线性插值算法（Opencv默认算法）训练得到的，造成这个模型不是最适配DVPP缩放的模型。

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：

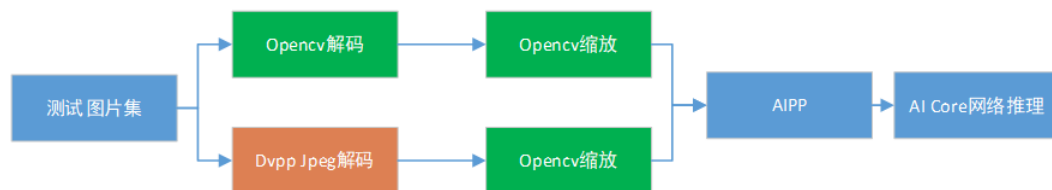
步骤1 进行对比实验，实验过程如图6-3所示，对照组是用Opencv进行解码和缩放，如果与对照组精度相比下降，则认为这是由于DVPP解码和缩放导致。

图 6-3 DVPP 解码与缩放



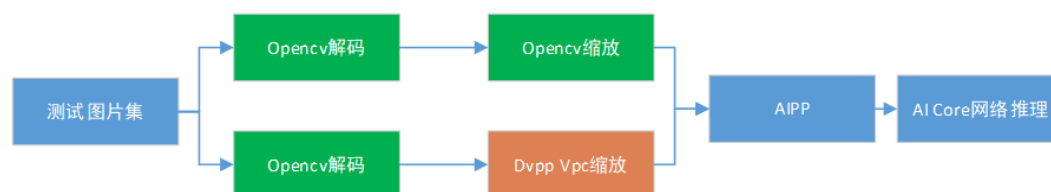
步骤2 进行对比实验，实验过程如图6-4所示，对照组仍然使用Opencv进行解码和缩放，如果与对照组精度相比下降，则认为DVPP解码对精度有影响。

图 6-4 DVPP 解码



步骤3 进行对比实验，实验过程如图6-5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 6-5 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。

6.3.12 VDEC 或 VPC 超时说明

- VPC业务在正常处理流程中，低概率偶现性能抖动，约300ms。VDEC业务实现视频的解码，内部数据经过VPC处理后，输出YUV420SP格式（包括NV12和NV21）的图片，因此，VDEC也存在低概率偶现抖动现象。

抖动发生时，内核日志（ERROR级别）中有如下关键词：

日志信息：

```
dvpp_vpc_cmdlist_engine_proc wait time out,engine id:0  
PIPE/CVDR OK,engine id:0  
CMDLIST not last frame ,engine id:0
```

- a. 该日志为VDEC/VPC业务抖动发生时的一个提示，不影响业务正常运行。
- b. 若用户业务代码对于DVPP专用内存的使用方式是临时申请与释放，且因为上述抖动造成队列堆积，从而消耗较多内存导致内存不够时。建议业务中建立反压机制，如内存池等。

- 用户做性能统计与分析时，若需要对性能抖动有具体了解，建议设置日志级别为warning级别。

日志信息：

```
dvpp_vpc_cmdlist_engine_proc wait time out,engine id:0  
PIPE/CVDR OK,engine id:0  
CMDLIST last frame ,engine id:0
```

当上述日志出现，业务性能也会有300ms左右抖动。

6.3.13 DVPP 输入和输出内存申请方式说明

DVPP输入输出内存申请要求说明：

表 6-1 内存要求说明

| 功能模块 | 输入内存 | 输出内存 |
|-------------|---|---|
| VPC | <p>请使用AscendCL提供的接口申请/释放内存：</p> <ul style="list-style-type: none"> 使用acldvppMalloc接口申请内存，内存满足DVPP的要求（内存起始地址16对齐）； 使用acldvppFree接口释放内存。 | <p>请使用AscendCL提供的接口申请/释放内存：</p> <ul style="list-style-type: none"> 使用acldvppMalloc接口申请内存，内存满足DVPP的要求（内存起始地址16对齐）； 使用acldvppFree接口释放内存。 |
| JPEGE和JPEGD | <p>请使用AscendCL提供的接口申请/释放内存：</p> <ul style="list-style-type: none"> 使用acldvppMalloc接口申请内存，内存满足DVPP的要求（内存起始地址128对齐）； 使用acldvppFree接口释放内存。 | <ul style="list-style-type: none"> 由用户指定输出内存时，由用户自行释放内存。请使用AscendCL提供的接口申请/释放内存： <ul style="list-style-type: none"> 使用acldvppMalloc接口申请内存，内存满足DVPP的要求（内存起始地址128对齐）。在申请内存前，可以调用DvppGetOutParameter接口获取输出内存大小。 使用acldvppFree接口释放内存。 不由用户指定输出内存时，DVPP内部申请内存，需由用户调用JPEGE/JPEGD出参结构体中的cbFree()回调函数释放内存，并将内存地址指针置为空。 |

| 功能模块 | 输入内存 | 输出内存 |
|-----------|---|--|
| PNGD | <p>请使用AscendCL提供的接口申请/释放内存：</p> <ul style="list-style-type: none"> 使用acldvppMalloc接口申请内存，内存满足DVPP的要求（内存起始地址128对齐）； 使用acldvppFree接口释放内存。 | <ul style="list-style-type: none"> 由用户指定输出内存时，由用户自行释放内存。请使用AscendCL提供的接口申请/释放内存： <ul style="list-style-type: none"> 使用acldvppMalloc接口申请内存，内存满足DVPP的要求（内存起始地址128对齐）。在申请内存前，可以调用DvppGetOutParameter接口获取输出内存大小。 使用acldvppFree接口释放内存。 如果不由用户指定输出内存时，DVPP内部申请内存，需由用户调用PNGD出参结构体中的FreeOutputMemory()函数释放内存。 |
| VDEC和VENC | <p>对内存无要求，支持调用malloc/free、new/delete等原生接口申请/释放内存，也支持调用AscendCL提供的acldvppMalloc/acldvppFree接口申请/释放内存。</p> | <p>VDEC输出的HFBC格式数据直接作为VPC的输入。</p> <p>VENC输出内存是DVPP内部管理，用户在使用时可以拷贝输出内存中的数据。</p> |

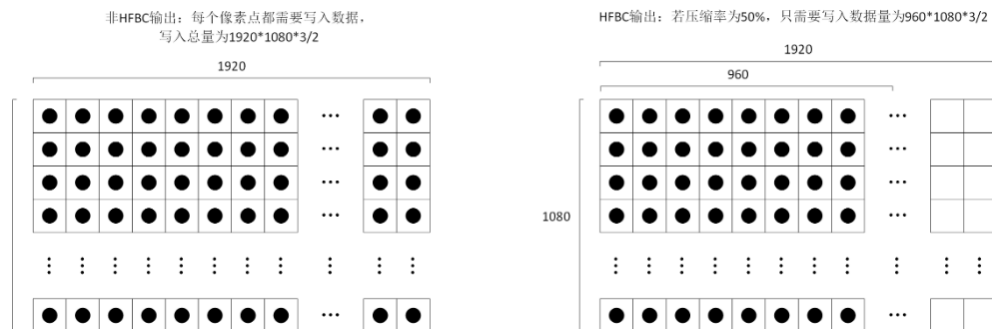
6.3.14 HFBC 格式介绍

HFBC格式是一种数据压缩格式，是VDEC解码后输出的每一帧数据的格式，而对此格式进行解压缩由VPC完成，因VDEC和VPC是两个独立模块，所以用户在使用过程中可以看到HFBC数据。

- 为什么要有HFBC格式？

因为VDEC解码每一帧后，将此帧图像进行HFBC压缩之后再写入内存，既可以降低写入内存的数据量，又可以减少输出此帧的耗时，从而获取更高解码性能。

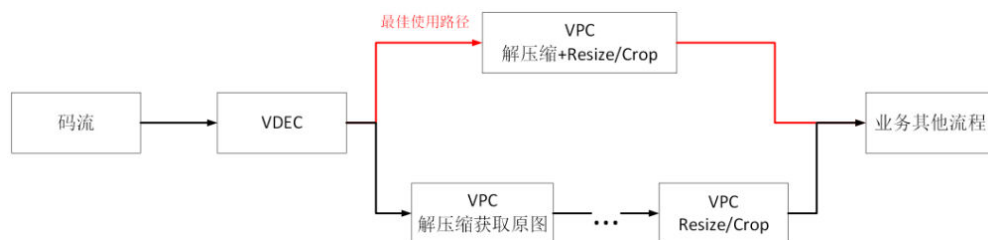
举例：如下图，比如此帧图像分辨率为1920*1080，正常一帧图像大小为 $1920 \times 1080 \times 3/2 = 3110400\text{B}$ ，所以需要写入内存的数据量为3110400B，而经过HFBC压缩后（一般压缩率为50%），写入内存数据量仅有1555200B，所以既减少了写入内存的数量，也减少了写入内存的耗时，从而性能更高。



- 怎么使用HFBC格式更优？

因用户最终需要的是YUV数据，而VDEC输出的是HFBC数据，所以用户需要调用VPC模块对HFBC数据进行解压缩得到YUV数据。而一般场景下，用户最终想要的是送入网络大小的图（比如 224×224 ），并不是原始图像（比如 1920×1080 ），所以当VDEC解码后输出每帧HFBC数据后，用户可以将HFBC数据送入VPC并同时配置好resize/crop参数，驱使VPC直接输出需要的图像数据，即一次性完成【HFBC解压缩+resize/crop】操作，这样是最佳使用方式。

若用户需要原始图像，则VDEC输出HFBC格式相比VDEC直接输出YUV格式并没有优势（因为VDEC输出HFBC后，还需要VPC读入HFBC数据解压缩后再输出YUV数据），但当前芯片VDEC不能直接输出YUV数据，所以只能使用HFBC格式数据给VPC进行处理。



7 Runtime 常见故障

7.1 用户进程异常退出后重启进程失败

7.2 注册算子数超过最大规格

7.1 用户进程异常退出后重启进程失败

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
```

Runtime日志信息：halResourceIdAlloc xxx failed

```
[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 [npd_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public 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。

7.2 注册算子数超过最大规格

现象描述

推理过程中，用户load model出现报错。出现如下关键日志信息：

ProgramRegister:Program register failed, program out of xxx和Register binary failed，例如：

```
[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：

- 反复load、unload模型。
- 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。
- 避免反复load、unload操作，做到一个model load一次，执行多次。