8

# 配电网监测数据微批处理的血统链标记容错法

屈志坚,彭 翔,王群峰,王汉林 (华东交通大学 电气与自动化工程学院,江西 南昌 330013)

摘要:针对分布式配电自动化系统存在数据量井喷、海量监测数据缺乏高效的分布式故障容错机制的问题,提出一种血统链标记容错新方法。利用弹性分布式数据集、微批计算的记录级容错和血统链标记序列融合处理的设计技巧,实现了分布式数据容错中血统链的追溯和条件标记的自动选择。以铁路配电网监测采集的数据为算例,搭建了4机集群的调度监控平台进行容错测试。以发生频次最高的单数据节点故障为例,测试结果表明:对于包含3×10<sup>6</sup>条监测数据记录的弹性分布式数据集,血统链标记容错模型的集群 CPU 平均占用率波动小于1.5%,磁盘占用率下降4.2%;当迭代次数达到600、800次时,迭代运算耗时分别降低24.3%和42.9%;所提方法实现500 ms 流处理延时的同时,对集群资源的使用情况也具有较好的优化效果,验证了该方法对分布式集群容错的有效性。

关键词:配电自动化系统;配电网监测数据;分布式集群;微批计算;血统链标记;流计算;容错

中图分类号:TM 73

文献标识码:A

DOI:10.16081/j.issn.1006-6047.2019.04.002

# 0 引言

配电网处于电力系统的末端,并直接与用户相连,具有地域分布广、电网规模大、设备种类多、网络连接多样、运行方式多变等特征[1]。随着电力信息化的推进,配电网数据的规模和种类快速增长[2-5]。在配电网监测方面,为了捕获各种状态信息,满足应用系统全景状态监测的需求[6],设备的采样频率越来越高,监测数据量呈几何级数增长[7],其数据结构日趋复杂[8-9],数据量井喷易导致监测界面反应速度慢的问题,严重时易造成关键信息的迟报、漏报或误报[10],若调度系统的容错性能不足,就可能造成故障决策的误判,从而威胁电网的安全。

数据容错的实现方法主要有以下 3 类。第一类方法是数据检查点技术,当系统发生故障时,利用保存于可靠存储介质中的检查点文件,使系统恢复到发生故障前的某一状态[11]。然而,数据检查点的操作成本高,易耗尽网络宽带资源,并不适用于处理配电网海量实时监测大数据等庞大的数据集[12-13]。第二类方法引入了记录数据更新的实时容错模式[14-16],与需付出高昂代价的检查点机制不同,记录级容错可通过历史操作记录文件来重建所有丢失的

收稿日期:2018-07-04;修回日期:2019-01-09

基金项目:国家自然科学基金资助项目(51567008);江西省 杰出青年人才计划项目(20162BCB23045);江西省自然科学 基金资助项目(20171BAB206044);江西省科技厅应用培育 项目(20181BBE58010)

Project supported by the National Natural Science Foundation of China (51567008), the Foundation Plan for Distinguished Young Scholars in Jiangxi Province (20162BCB23045), the Natural Science Foundation of Jiangxi Province (20171BAB206044) and the Application Research Project of Jiangxi Provincial Science and Technology Department (20181BBE58010)

分区[17-18]。对于不同的计算框架,其实现记录级容 错的方式也不相同,例如流计算框架 Strom 采用上 游数据备份、消息确认的机制,保障消息在失败后能 被重新处理.尽管此类方法能保证不会出现数据丢 失的事件,但丢失的数据可能被重复多次执行,导致 集群资源浪费。第三类方法是一类可解决丢失数据 重复执行的记录级容错。Apache 基金会组织的顶 级项目 Spark 提出了弹性分布式数据集(RDD)的概 念,一个 RDD 包含了如何从其父 RDD 衍生的关联 信息,从而无需检查点操作就可重构丢失的数据分 区,相较于其他计算框架的记录级容错,由于 RDD 可在不同集群节点中分发使用,故其在数据并行类 应用的适应性更好[19-20]。但由于配电网调度监控系 统的实时监测数据呈海量、异构、更新快等特点,如 配电网调度自动化监控系统中要求小于 1.5 s 的响应 时间,配电数据的更新时间更是要求达到百毫秒级, 且各数据间具有强关联性,若记录的 RDD 中的衍生 关系过于冗长,就有可能导致数据恢复的时间随之变 长,从而影响调度监控系统的实时性、运行效率。

鉴于此,本文提出一种配电网监测数据故障后快速恢复的带条件标记血统链 CLM (Conditional Lineage Mark)容错新方法。该方法基于分布式微批框架编程模型,通过记录级容错和血统链标记序列融合处理的设计技巧,实现快速、可靠和高效的分布式数据容错。利用配电自动化工程中采集获取的配电网监测数据集,在多机分布式集群上搭建了基于Spark 开源计算框架的流计算平台,对配电网监测的数据集进行了集群容错测试和验证。

#### 1 配电网监测大数据流式处理

1.1 **分布式配电自动化系统大数据的处理** 随着智能配电网中负荷管理、网络分析、故障处

理、数据采集与监视控制(SCADA)系统和管理信息系统(MIS)等应用的内部数据量快速增长和分布式智能量测体系的搭建,监测集群数据节点的故障处理需求越来越紧迫<sup>[21]</sup>。智能配电网的典型分层结构如图 1 所示。

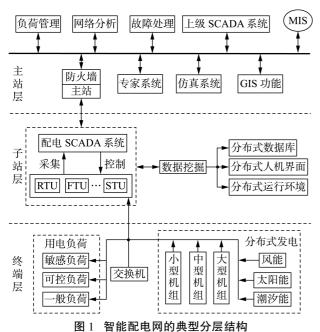


Fig.1 Typical hierarchical structure of smart distribution network

在海量监测数据分布式处理的背景下,建立容错能力强的新一代高可靠性配电自动化集群调度系统是目前实施配电网智能调度支撑平台的瓶颈。

# 1.2 分布式微批处理技术

微批处理以 RDD 为核心数据结构,通过如图 2 所示的分布式主从结构,实现调度监控工作站和数据节点的微批处理及流计算的融合。

集群中主要有主节点和工作节点这2类节点。 主节点上运行着微批处理界面端,负责接收调度监 控客户端发来的监测数据流,向工作节点发布任务, 并监听资源管理器的资源请求。资源管理器同样运 行在主节点上,主要负责将工作节点的资源申请推 送到任务调度分发器,再将其返回的资源合理地分 配至工作节点,并启动执行器对任务进行处理。每 个工作节点上都同时运行着若干个执行器进程,该 进程负责执行开关量变位判定、模拟量越限处理、设 定电量阈值等任务,并且将任务处理完成后得到的 监测数据存储在内存或者磁盘中。工作节点在聚合 各个执行器发来的处理信息后,将最终的执行结果 返回至监控主机所在的主节点,最终聚合计算结果, 完成对监控应用的分布式计算处理。配电网监测实 时信息流的传输过程遍布于配电网子站与主站之间 交互处理与调度监控的各个环节,为配电网调度提 供及时的实时信息,快速高效的信息流容错机制是 配电网数据节点故障处理的基础。

0

# 2 微批处理的容错机制

#### 2.1 RDD 的操作

作为微批框架中流计算的核心数据结构,RDD被设计为只读属性,其结构相对简单、统一,当对配电网监测数据流进行并行计算时,可不用考虑数据互斥问题,避免了监测系统在同时处理多种异构数据时可能发生的数据读写冲突。其次,RDD的高可靠性是通过记录足够的计算操作过程,在需要时从头或从某个镜像重新计算来恢复,并非通过检查点和数据复制实现,如当某节点发生故障,导致配电网监测数据流中某一环节的内容失效时,可从任一记录的计算过程中进行恢复,在保证监测系统有足够容错性能的前提下,节省大量的内存及网络带宽。

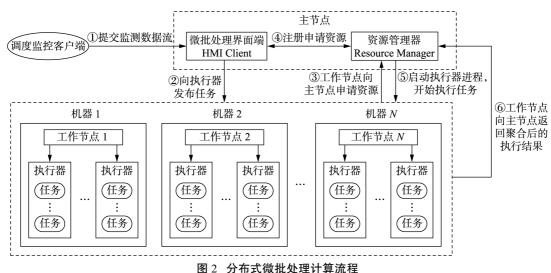


Fig.2 Computing process of distributed micro-batching

st<sub>4</sub>},每个站所都具有一定数量的量测点,现欲对站所{st<sub>1</sub>,st<sub>2</sub>,st<sub>3</sub>}供电臂低压侧 A 相出线电压值和电流值进行处理,转换过程如图 3 所示。RDD 一般采用流水线形式生成,这不适用于图 3 中量测线电压和线电流的求均值及按站所排序等应用处理,由于后续监测信息处理中仍有多个计算任务依赖于操作前的 RDD,因此需将之前的部分 RDD 缓存于内存中以避免重复计算。现场终端装置采集到的各类实时电量数据,经远程通信网络接入微批处理计算框架,通过过滤、求均值、合并、排序等操作后,再对处理后的数据进行整合与缓存。

在图 3 的量测线电压处理阶段中,首先由 filter 算子过滤{st<sub>1</sub>, st<sub>2</sub>, st<sub>3</sub>, st<sub>4</sub>} 采集的量测线电压,取前 3 个站所的线电压值,再通过 map 算子求均值,按站所排序后,得到(st<sub>1</sub>,231.6 kV)、(st<sub>2</sub>,233.0 kV)、(st<sub>3</sub>,231.7 kV)这 3 条电压记录。在量测线电流处理阶段中,电流记录(st<sub>1</sub>,96.9 A)所处数据分区在之前的计算中已被缓存,故可直接调用,按站所排序后,得到(st<sub>1</sub>,96.9 A)、(st<sub>2</sub>,97.1 A)、(st<sub>3</sub>,97.0 A)这 3 条电流记录。在分布式微批计算集群中,线电压、线电流记录。在分布式微批计算集群中,线电压、线电流的处理是并行执行的。获取得到{st<sub>1</sub>,st<sub>2</sub>,st<sub>3</sub>}这 3 个站所的平均线电压和平均线电流记录后,调用join 连接算子,得到(st<sub>1</sub>,231.6 kV,96.9 A)、(st<sub>2</sub>,233.0 kV,97.1 A)、(st<sub>3</sub>,231.7 kV,97.0 A)这 3 条电量记录,最后通过 cache 操作将该记录缓存至内存中,用于后续的计算。

#### 2.2 微批处理的监测数据流计算模型

微批处理的流计算模型具有高吞吐量、高容错性和高扩展性的特点。配电网采集监测数据是通过数据报文连续传输的,而对于微批处理的流计算过程,其数据处理是按批进行的。因此,在流计算模型中需预先设置批处理间隔 Δt,当输入监测数据流超过批处理间隔时,就会将其汇总成一批具有相同处理间隔的数据流并提交给流计算模型进行处理,该数据流即为原始数据流。窗口宽度由其窗口内部的批处理数据量决定,滑动间隔指的是窗口滑动一次并形成新的窗口所需的时间,窗口宽度和滑动间隔大小一定为批处理间隔的整数倍。

微批处理的流计算模型如图 4 所示,批处理间隔为 1 个时间单位,窗口宽度为 3 个时间单位,滑动间隔为 2 个时间单位。对于初始窗口 $\{t_1,t_2,t_3\}$ ,只有窗口宽度满足了 3 个时间单位才触发数据的处理。每经过 2 个时间单位,窗口滑动一次移去最早 2 个时间单位的数据,并与最近 2 个时间单位内流入的新数据流进行汇总形成新窗口 $\{t_3,t_4,t_5\}$ 。

根据已设置的批处理间隔  $\Delta t$ ,流计算模型将输入监测数据分成离散数据队列,然后每一段离散数据都被转换成微批处理中的 RDD。这样,流计算模型对监测数据流的处理就转变成了对 RDD 的处理。

### 2.3 监测数据流计算的血统链容错技巧

在微批计算中,每一个 RDD 都会记忆创建该数据集所需的操作,并跟踪记录其父级 RDD 与其子级

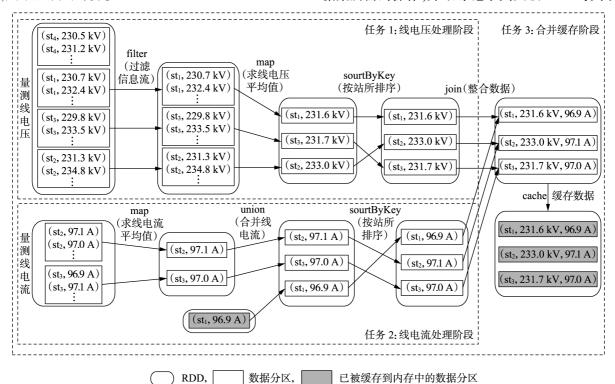


图 3 调度监控系统的 RDD 转换

Fig.3 Transformation of RDD in dispatch monitoring system

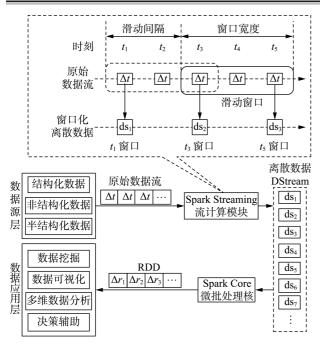


图 4 微批处理的流计算模型

Fig.4 Stream computing model of micro-batching

RDD 之间的继承关系,这种关系称为血统链(lineage),对于一个 RDD 而言,它的历史操作越多,血统链就越长。微批处理的血统链容错机制如图 5 所示,站所 st<sub>1</sub> 的母线电流记录(st<sub>1</sub>,96.8 A)存储在分区 0 中,站所 st<sub>2</sub> 的母线电流记录(st<sub>2</sub>,97.3 A)存储在分区 1 中,依此类推,当配电网调度监控系统突发故障,某工作节点宕机,导致 RDD<sub>1</sub> 中分区 1 的数据丢失时,可以通过血统链机制来获取足够的信息,通过其父级 RDD(RDD<sub>0</sub>)中的分区 1 来重新运算和恢复丢失的数据分区,由于 RDD 是不可改写的,可保证供配电监测数据恢复操作前、后数据的一致性。

假如工作节点 1 的 RDD<sub>1</sub> 中分区 1 的数据丢失,若需对 RDD<sub>3</sub> 进行计算,那么它所依赖的 RDD<sub>1</sub> 中分区 0、2 和 3 的缓存仍可用,无需再次计算,减轻了部分运算压力。但由于 RDD<sub>1</sub> 中分区 1 的缓存已丢失,需从 RDD<sub>1</sub> 的父级 RDD(即 RDD<sub>0</sub>)的分区 1 重新开始计算。因此,配电网调度监控系统信息处理的数据量越大,RDD 间的依赖关系越复杂,血统链就会越长,记录数据更新这种方法所产生的冗余计算开销也就越大。

#### 3 CLM 容错法

#### 3.1 流计算集群中的分布式 CLM 模型

为避免配电网调度监控中由于血统链过长导致的额外计算开销问题,在原生记录级容错模型的基础上对条件标记进行定义,当血统链延伸至某时刻满足设置条件时,立即对 RDD 的状态信息进行保存,系统发生故障后可由该条件标记进行恢复而无需重新执行血统链,可降低资源管理器向微批处理界面

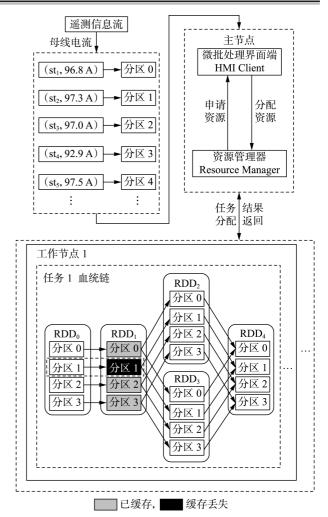


图 5 微批处理的血统链容错机制

Fig.5 Lineage chain fault-tolerant mechanism of micro-batching

端发起的资源注册申请频率,减少冗余计算开销。

将每一条血统链用一个 4 元组  $\tau_i(T_i, E_i, L_i, n_i)$  定义,其中  $\tau_i$  为标记的血统链, $T_i$  为血统链的周期,  $E_i$  为该血统链的执行时间, $L_i$  为血统链的时限, $n_i$  为血统链的分段数, $n_i$ —1 为每个周期设置的条件标记数量,即  $n_i$ —1 个条件标记将血统链  $\tau_i$  划分为  $n_i$  段。使用条件标记序列存在一定的处理开销,令  $C_i^A$  为执行一次接收测试所需时间, $C_i^S$  为保存血统链状态信息的时间开销,则建立一个条件标记的计算时间开销为  $C_i = C_i^A + C_i^S$ 。另外,将监控系统的数据节点发生故障后恢复到最近条件标记的时间开销记为  $C_i^R$ 。

# a. 单链血统链标记模型最坏执行时间。

对于只有 1 条血统链的血统链集合  $S\{\tau_i(T_i, E_i, L_i, n_i), i=1\}$ ,假设  $n_i$ -1 个条件标记将血统链划分为  $n_i$  段,在  $\tau_i$  的执行期间只发生一处节点故障,则单链血统链  $\tau_i$  的最坏执行时间  $E_i^0$  可表示为:

$$\begin{split} E_{i}^{\text{ft}} = & E_{i}' + (n_{i} - 1) C_{i} + C_{i}^{\text{R}} + E_{i}' / n_{i} + C_{i} = \\ & (1 + 1 / n_{i}) E_{i}' + n_{i} C_{i} + C_{i}^{\text{R}} \\ & n_{i} > 1, C_{i} > 0 \end{split} \tag{1}$$

其中, $E_i'$ 为单链血统链引入条件标记序列前的正常运行时间; $(n_i-1)C_i$ 为最坏情况下,即单链血统链执行到最后一个条件标记才发生故障时,建立所有条件标记的时间开销; $E_i'/n_i$ 为单链血统链恢复到最近一个条件标记,重新运行的附加时间开销。

# b. 原生记录级容错模型的最坏执行时间。

若使用原生记录级容错方法进行故障恢复,而不使用条件标记序列,则在式(1)中,令 $n_i$ =1, $C_i$ =0,可得其最坏执行时间 $\tilde{E}_i^n$ 为:

$$\widetilde{E}_{i}^{ft} = \widetilde{E}_{i} + 0 + \widetilde{C}_{i}^{R} + \widetilde{E}_{i} + 0 = 2 \, \widetilde{E}_{i} + \widetilde{C}_{i}^{R} \tag{2}$$

其中, $\tilde{E}_i$ 为原生记录级容错模型血统链 $\tilde{\tau}_i$ 的正常运行时间; $\tilde{C}_i^R$ 为原生记录级容错模型恢复血统链状态的开销。

# c. 血统链标记模型的最坏执行时间。

另外,考虑到在通常情况下,式(1)中的血统链并非被平均划分为若干等份,因此,对式(1)进行修正,血统链 $\hat{\tau}$ ,的最坏执行时间 $\hat{E}^n$ 可表示为:

$$\hat{E}_{i}^{ft} = \hat{E}_{i} + \hat{E}_{i} / \hat{n}_{i} + \hat{n}_{i} \hat{C}_{i} + \hat{C}_{i}^{R}$$
(3)

其中, $\hat{E}_i$  为  $\hat{\tau}_i$  引入条件标记序列前的正常运行时间; $\hat{E}_i$ / $\hat{n}_i$  为血统链恢复到最近一个条件标记,重新运行的附加时间开销; $\hat{n}_i\hat{C}_i$  为建立所有条件标记的时间开销; $\hat{C}_i^R$  为从条件标记中恢复血统链状态的开销。

修正式(3)中 $\hat{n}_i$ 与 $\hat{E}_i^n$ 呈正相关的关系,根据此血统链和条件标记序列的混合修正计算模型,本文提出一种适用于配电网监测大数据的分布式 CLM 容错新方法。

# 3.2 血统链的追溯和条件标记的选择

通过实时监测应用程序编程接口,生成大量不同的 RDD,运算过程中的每次迭代计算也生成了许多中间 RDD。未设置条件标记的原生记录级容错机制仅保存每次操作的结果,易忽略中间 RDD 的运算过程,但事实上,中间 RDD 详细地记录了每一次的运算过程,并用于下一次的迭代计算。RDD 的血统链追溯如图 6 所示。

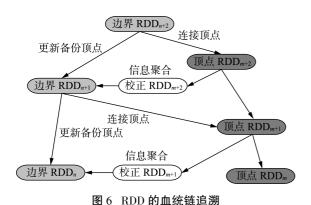


Fig.6 Retrospection method of lineage chain in RDD

图 6 中的箭头表示 2 个 RDD 之间的依赖关系,如边界 RDD<sub>n+2</sub> 依赖于边界 RDD<sub>n+1</sub> 和顶点 RDD<sub>m+2</sub>。 其中,顶点 RDD 是每次迭代计算所产生的结果,边界 RDD 作为中间 RDD 也被用于下一次的迭代运算。按照血统链进行追溯,能够找到并保存中间 RDD,通过重新计算中间 RDD,就可以再次获取某次操作中丢失的任何一个 RDD。

在未设置条件标记序列的情况下,每次迭代运算的分布式内存回收开销随着迭代次数的增加呈几何倍数增长。通过将虚拟机的旧数据空间使用率设置为条件标记的时间门槛,实现条件标记的自动选择,可使原本在第一次内存回收前急剧提升的系统内存占用率得到缓解。此外,由于冗长的血统链被条件标记所切断,每次迭代运算中用于保持血统链各个阶段之间链接的完整性所需要的额外计算开销被降低到第一次迭代运算的水平,保证分布式集群在血统链变长的情况下仍有良好的容错能力。

#### 3.3 CLM 容错处理流程

在初始化网络参数和指定血统链之后,首先判断该带条件标记序列的血统链是否为空链,在非空链的前提下随机地对该血统链中的某一  $RDD_n$  进行预标记。如果预标记的  $RDD_n$  存在父级  $RDD_m$  进行条件标记选择判定:将旧数据空间使用率设置为 K,当前空间使用率设置为 T,在 T>K 且父级  $RDD_m$  未设置条件标记的条件下,对父级  $RDD_m$  进行条件标记设置。若父级  $RDD_m$  还有多个平行父级  $RDD_m$  则继续对其进行条件标记选择判定。CLM 容错方法处理流程见图 7。

#### 4 典型工程配置下的算例测试

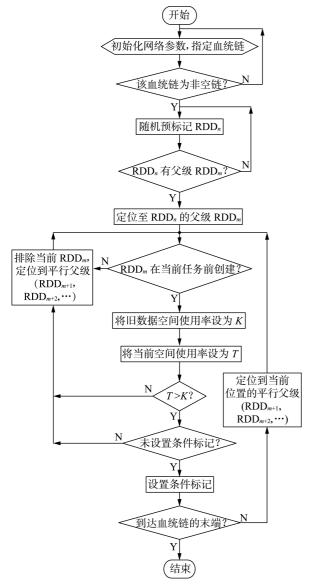
根据石家庄—德州段铁路配电网调度监控系统工程中的典型调度主站配置,设置 2 台数据服务器以及 2 台主、备调度工作站搭建 4 机集群,且以铁路 10 kV 配电网调度监控工程的数据为算例,部署 1 个主节点和 3 个工作节点的流计算集群,如图 8 所示。

取系统下属站所中"低压侧三相电压、三相电流"的电量信息作为输入,以连续的电压、电流队列形式交由流计算集群进行实时处理,在微批框架执行聚集计算过程中分别对流计算集群节点进行测试。

#### 4.1 节点故障测试

以 3×10<sup>6</sup> 条监测数据记录作为输入,分别对原生记录级容错模型和 CLM 容错模型进行集群节点宕机测试,模拟配电网监测过程中的数据节点故障情况,取宕机故障 30 min 内主节点的 CPU 占用率和磁盘占用率这 2 个指标进行测试,验证带条件标记序列的分布式容错模型的优化效果。





#### 图 7 CLM 容错方法的流程

Fig.7 Flowchart of CLM method

# a. 主节点 CPU 占用率方面的对比。

在正常运行与宕机节点数为 1—3 的条件下,分别对原生记录级容错模型、CLM 容错模型的 CPU 占用情况进行测试,测试结果如图 9 所示。

由图 9 可以得出,单节点故障对 CPU 占用率的 影响较小,当发生单节点故障时,原生记录级容错模 型和 CLM 容错模型的 CPU 平均占用率与正常运行 状态相比仅分别增加了 2.7%和 1.5%,表明微批框 架对单节点故障情况有较好的容错性;由于分布式 集群的核心数据结构——RDD 本身具有高可靠性 并可在多次计算间重复使用,在宕机节点数为 2、3 的情况下,CLM 容错模型分别将 CPU 平均占用率降低了 2.7%和 2.5%。上述结果表明,CLM 容错方法可以改善原生记录级容错机制的 CPU 占用情况。

#### b. 主节点磁盘占用率方面的对比。

在正常运行与宕机节点数为 1—3 的条件下,分别对原生记录级容错模型与 CLM 容错模型的磁盘占用情况进行测试,结果如图 10 所示。

由图 10 可知,磁盘占用率与宕机节点数呈正相关的关系,宕机节点数越多,则工作节点向主节点申请的磁盘资源越多。相比磁盘平均占用率,磁盘占用率的加权平均值能更加准确地反映磁盘的真实繁忙程度。在正常运行与宕机节点数为 1—3 的的情况下,采用 CLM 容错模型后,磁盘占用率的加权平均值分别下降了 4.2%、3.6%、3.4% 和 2.8%,表明CLM 容错方法能改善分布式流计算集群的磁盘占用情况,节省了部分额外计算开销。

#### 4.2 迭代性能测试

以 3×10<sup>6</sup> 和 6×10<sup>6</sup> 条监测数据记录作为输入, 在分布式集群中对所选算例的低压侧三相电压、三 相电流的电量信息进行过滤、求均值、合并、排序等 聚集计算,分别记录原生记录级容错模型和 CLM 容

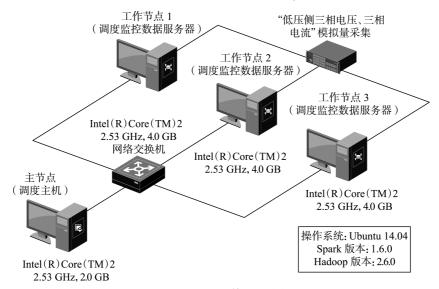
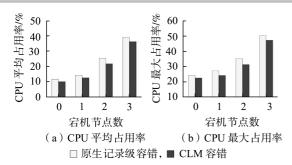


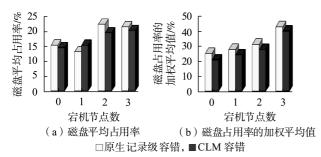
图 8 分布式流计算集群的部署

Fig.8 Deployment of distributed stream computing clusters



#### 图 9 2 种容错模型的 CPU 占用率对比

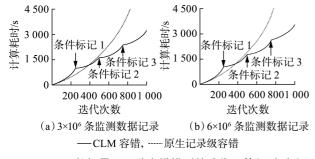
Fig.9 Comparison of CPU utilization rate between two fault-tolerant models



# 图 10 2 种容错模型的磁盘占用率对比

Fig. 10 Comparison of disk utilization rate between two fault-tolerant models

错模型的迭代运算耗时,测试结果如图 11 所示。



# 图 11 不同数据量下 2 种容错模型的迭代运算耗时对比

Fig.11 Comparison of iteration time between two fault-tolerant models under different data sizes

由图 11 可知,迭代次数与计算耗时呈正相关,符合式(3)中 $\hat{n}_i$ 与 $\hat{E}_i^a$ 的正相关关系。在前 400 次迭代运算中,因为条件标记时间开销  $C_i$ 、 $C_i^R$ 的存在,原生记录级容错模型的计算耗时稍低于 CLM 容错模型;当迭代大于 400 次后,迭代次数越大,CLM 容错模型对计算性能的提升越显著。微批计算框架的迭代次数能真实反映血统链的复杂程度,迭代次数越大,表示 RDD 间的依赖关系越复杂,血统链越长。

当输入  $3\times10^6$  条监测数据记录时, $E_i=3$  236.4 s,设置 3 个条件标记,血统链  $\tau_i$  被分成 4 段,即  $n_i=4$ ,令  $C_i=10$  s, $C_i^R=20$  s,按照式(3)可计算得到故障发生在血统链最后一个条件标记的最坏执行时间约为 4 105.5 s,即平均每次迭代的最坏执行时间约为 4.1 s;当迭代次数为  $600\ 800$  时,CLM 容错模型分别将计算耗时降低了 24.3%和 42.9%。综上表明,CLM 容

错方法利用血统链标记序列,将冗长的血统链切断后在集群节点中进行多机并行处理,降低了迭代运算的额外计算开销,可有效提升容错计算性能。

#### 4.3 延时性能测试

以 3×10<sup>6</sup> 条监测数据记录作为输入,分别对原生记录级容错模型和 CLM 容错模型进行集群延时性能测试,观察在不同处理延时指标下流计算集群的资源使用情况,取集群节点 CPU 平均占用率与磁盘占用率的加权平均值这 2 个指标进行测试,验证基于 CLM 的分布式容错模型的优化效果,测试结果如图 12 所示。

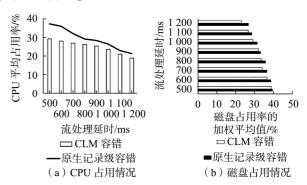


图 12 不同处理延时下 2 种容错模型的 CPU、 磁盘占用情况对比

Fig.12 Comparison of CPU and disk utilization between two fault-tolerant models under different processing time delays

由图 12 可知,流处理延时设置得越小,所消耗的集群资源越多,如对于原生记录级容错模型而言,相比于 800 ms 的流处理延时,流处理延时为 500 ms 时的集群 CPU 和磁盘占用率分别增加了 22.1%与9.4%;采用 CLM 容错模型进行调优后,流处理集群在不同延时指标下的集群资源使用情况都得到了一定程度的改善,如 500 ms 流处理延时下的 CPU 平均占用率相比原生记录级容错模型降低了 21.8%,表明 CLM 容错方法对不同流处理延时下集群资源的使用情况具有较好的优化效果。

#### 5 结论

- a. 利用分布式微批编程模型搭建流计算集群,融合血统链标记序列和记录级容错处理的设计技巧,提出一种适用于配电网海量监测数据的 CLM 容错新方法,通过比较 CLM 容错模型与原生记录级容错模型的分布式集群测试,验证了 CLM 容错方法的有效性。
- b. 以 10 kV 配电网调度监控数据集为例,进行了微批框架的流计算集群节点故障测试和迭代性能测试。结果表明,在应对突发数据节点故障时,CLM容错方法不仅可降低集群计算节点的 CPU 平均占用率和磁盘占用率,还可以减少迭代运算的计算耗时,在不同的流处理延时条件下均能改善集群资源



# 的使用情况。

# 参考文献:

- [1] 邓建玲. 能源互联网的概念及发展模式[J]. 电力自动化设备, 2016,36(3):1-5.
  - DENG Jianling. Concept of energy internet and its development modes[J]. Electric Power Automation Equipment, 2016, 36(3): 1-5.
- [2] 胡丽娟, 刁嬴龙, 刘科研,等. 基于大数据技术的配电网运行可靠性分析[J]. 电网技术, 2017, 41(1): 265-271.

  HU Lijuan, DIAO Yinglong, LIU Keyan, et al. Operational reliability analysis of distribution network based on big data technology [J]. Power System Technology, 2017, 41(1): 265-271.
- [3] 葛磊蛟,王守相,瞿海妮. 智能配用电大数据存储架构设计[J]. 电力自动化设备,2016,36(6):194-202. GE Leijiao,WANG Shouxiang,QU Haini. Design of storage framework for big data of SPDU[J]. Electric Power Automation Equipment,2016,36(6):194-202.
- [4] 曲广龙,杨洪耕,熊晓祎. 可在线编辑电网地理接线信息的电能质量监测系统设计与实现[J]. 电力自动化设备,2015,35 (11):165-174.

  QU Guanglong, YANG Honggeng, XIONG Xiaoyi. Design and realization of power quality monitoring system with online editable geographical wiring information of power grid[J]. Electric Power Auto-
- [5] 邵振国,吴瑾樱,苏文博. 面向海量历史监测数据的谐波污染用户统计建模方法[J]. 电力自动化设备,2016,36(8):110-114. SHAO Zhenguo, WU Jinying, SU Wenbo. Statistical modeling based on massive historical monitoring data for harmonic pollution customer[J]. Electric Power Automation Equipment,2016,36(8): 110-114.

mation Equipment, 2015, 35(11):165-174.

- [6] 李洋,蔡志远,刘海涛,等. 广域分布式电源运营管控系统[J]. 电力自动化设备,2018,38(3):102-108. LI Yang,CAI Zhiyuan,LIU Haitao, et al. Management and control system of wide-area distributed generation[J]. Electric Power Automation Equipment,2018,38(3):102-108.
- [7] 屈志坚,徐振清,周锐霖,等. 按列存储的配电网监测数据包区间编码正规化压缩处理[J]. 电力自动化设备,2018,38(3);56-62.
  - QU Zhijian, XU Zhenqing, ZHOU Ruilin, et al. Range encoding normalization compression processing for distribtion network monitoring data pack with column storage [ J ]. Electric Power Automation Equipment, 2018, 38(3):56-62.
- [8] 曲广龙,杨洪耕,李兰芳. 主动配电网电能质量实时监测系统设计与实现[J]. 电力系统自动化,2015,39(10):117-123. QU Guanglong,YANG Honggeng,LI Lanfang. Design and realization of power quality real-time monitoring system for active distribution networks[J]. Automation of Electric Power Systems,2015,39(10):
- [9] 宋墩文,温渤婴,杨学涛,等.广域量测信息大数据特征分析及应用策略[J]. 电网技术,2017,41(1):157-163.

  SONG Dunwen,WEN Boying, YANG Xuetao, et al. Big data feature analysis and application strategy of wide area measurement information[J]. Power System Technology,2017,41(1):157-163.

117-123.

- [10] 张宇帆,艾芊,肖斐,等. 数据驱动电能质量分析现状及其支撑技术与展望[J]. 电力自动化设备,2018,38(11):187-196.
  ZHANG Yufan,AI Qian,XIAO Fei,et al. Present situation, supporting technologies and prospect of data driven power quality analysis [J]. Electric Power Automation Equipment,2018,38(11):187-196.
- [11] BOUGUERRA M S, TRYSTRAM D, WAGNER F. Complexity ana-

- lysis of checkpoint scheduling with variable costs [J]. IEEE Transactions on Computers, 2013, 62(6):1269-1275.
- [12] WANG H, CHEN H P, DU Z W, et al. BeTL: MapReduce check-point tactics beneath the task level[J]. IEEE Transactions on Services Computing, 2017, 9(1):84-95.
- [ 13 ] HE Z Z, CHEN Y J, MEN C G, et al. Schedulability of fault tolerant real time system based on local optimum checkpoint under priority mixed strategy[ J]. Chinese Journal of Electronics, 2015, 24(2): 236-244.
- [14] 王守相,张卫涛,葛磊蛟. 智能配电网统一模型云与应用服务技术系统[J]. 电力自动化设备,2015,35(2):49-54.
  WANG Shouxiang,ZHANG Weitao,GE Leijiao. Unified model cloud and application service system for smart distribution network[J]. Electric Power Automation Equipment,2015,35(2):49-54.
- [15] 范贵生,虞慧群,陈丽琼,等. 基于效用的云计算容错策略和模型[J]. 中国科学(信息科学),2014,44(1):158-176.
  FAN Guisheng, YU Huiqun, CHEN Liqiong, et al. Utility-based fault tolerant strategy and model for cloud computing[J]. Scientia Sinica (Informationis),2014,44(1):158-176.
- [16] 张治斌,李燕歌. 云计算下 MapReduce 多组容错机制架构的分析与研究[J]. 微电子学与计算机,2014,31(1):52-55,59. ZHANG Zhibin,LI Yange. Analysis and study on under cloud computing multiple sets of fault tolerance strategy architecture of Map-Reduce[J]. Microelectronics & Computer, 2014,31(1):52-55,59.
- [ 17 ] WEI H W, CHEN T Y, CHEN S H, et al. A QoS-aware data reconstruction strategy for a data fault-tolerant storage system [ C ] // 2015 IEEE 7th International Conference on Cloud Computing Technology and Science (CloudCom). Vancouver, BC, Canada: IEEE, 2015:274-281
- [18] SINGH S, LIU Y. A cloud service architecture for analyzing big monitoring data [J]. Tsinghua Science and Technology, 2016, 21(1): 55-70.
- [19] 王保义,王冬阳,张少敏. 基于 Spark 和 IPPSO\_LSSVM 的短期 分布式电力负荷预测算法[J]. 电力自动化设备,2016,36(1): 117-122.
  - WANG Baoyi, WANG Dongyang, ZHANG Shaomin. Distributed short-term load forecasting algorithm based on Spark and IIPPSO\_LSSVM [J]. Electric Power Automation Equipment, 2016, 36(1):117-122
- [20] ALSHEIKH M A,NIYATO D,LIN S W,et al. Mobile big data analytics using deep learning and apache spark [J]. IEEE Network, 2016,30(3):22-29.
- [21] 刘健,赵树仁,张小庆. 中国配电自动化的进展及若干建议[J]. 电力系统自动化,2012,36(19):6-10,21.

  LIU Jian,ZHAO Shuren,ZHANG Xiaoqing. Development of distribution automation in China and some suggestions[J]. Automation of Electric Power Systems,2012,36(19):6-10,21.

#### 作者简介:



屈志坚(1978—),男,江西南昌人,教授,博士,主要研究方向为智能监控理论与信息处理技术(E-mail:08117324@bjtu.edu.cn);

彭 翔(1991—),男,江西萍乡人,硕士研究生,主要研究方向为智能配电网大数据容错:

王群峰(1992—),男,江西南昌人,硕士研究生,主要研究方向为智能配电网大数据技术;

王汉林(1994—),男,河北沧州人,硕士研究生,主要研究方向为电网大数据的处理。



# Lineage chain mark fault-tolerant method for micro-batching of distribution network monitoring data

QU Zhijian, PENG Xiang, WANG Qunfeng, WANG Hanlin

(School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 330013, China)

**Abstract**: Aiming at the existing problems of monitoring message blowout and lacking of efficient distributed fault-tolerant mechanism for huge amounts of monitoring data in the distributed distribution automation system, a novel lineage chain mark fault-tolerant method is proposed. The retrospection of lineage chain and the automatic selection of conditional marks in the distributed data fault-tolerant progress are realized by applying the design technique of fusing the resilient distributed data set, the record-level fault-tolerant of micro-batching computing and the lineage chain mark sequence. The monitoring data of railway distribution network are taken as the simulation example and the dispatch monitoring platform of a 4-generator cluster is built to carry out the fault-tolerant test. A single-node fault with the highest frequency is taken as an example, whose test results show that, for the resilient distributed data set including 3×10<sup>6</sup> monitoring data records, the average CPU utilization rate fluctuation is less than 1.5% and the disk utilization rate decreases by 4.2%; when the number of iteration increases to 600 and 800, the time-consuming decreases by 24.3% and 42.9% respectively; the proposed method not only realizes the stream processing delay of 500 ms, but also has a better optimization effect on the usage of cluster resources, verifying the effectiveness of the method for distributed cluster fault-tolerant.

**Key words**: distribution automation system; distribution network monitoring data; distributed cluster; micro-batching computing; lineage chain mark; stream computing; fault-tolerant

-----

(上接第7页 continued from page 7)

# Comprehensive optimization method of capacity configuration and ordered installation for distributed energy storage system accessing distribution network

JIA Yulong<sup>1</sup>, MI Zengqiang<sup>1</sup>, LIU Liqing<sup>2</sup>, YIN Qukai<sup>1</sup>

(1. State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Baoding 071003, China;

2. Electric Power Research Institute of State Grid Tianjin Electric Power Company, Tianjin 300092, China)

Abstract: With the continuous technical economy improvement of electric energy storage, it has become a development trend to assess a large number of DESSs (Distributed Energy Storage Systems) in the distribution network. Aiming at the planning and operation problems faced by ordered access of DESSs, a comprehensive optimization method for capacity configuration and ordered installation of DESS is proposed. The upper layer model is the capacity configuration model, which considers the investment income of DESS and slowing down the capacity increase of distribution network. The lower layer model is the installation spot optimization model, which considers the characteristics and capacity limit of DESS, the load regulation for peak-load shifting and power fluctuation smoothing. The validity of the proposed method is verified by the example of IEEE 33-bus system, based on which, the placement scheme of DESS accessing to distribution network is obtained. Simulative results show that, the proposed comprehensive optimization method is suitable for different application scenarios with the integration of DESS of different capacity configurations and different installation spots and can provide a technical support for the planning and managing of large-scale DESS accessing to the distribution network.

**Key words**: distributed energy storage system; distribution network; ordered installation; comprehensive optimization model; placement scheme; capacity configuration