

项目编号: 2022PI0AC03

之江实验室科研项目 任务书

项目名称: 高并发分布式图计算系统

项目负责人: 张宇

牵头中心: 图计算研究中心

申请时间: 2022 年 4 月

项目周期: 2022 年 5 月 至 2024 年 4 月

之江实验室

2021 年制

一、基本信息

项目名称		高并发分布式图计算系统				
牵头中心		图计算研究中心				
项目经费概算 (万元，人民币)		2600				
项目周期		2022 年 5 月~2024 年 4 月				
项目 负责 人	姓 名	张宇	学 位	博士	专 业	计算机系统 结构
	职 称	副教授	身份证/护照 号码	431122198707170515		
	所在单位	华中科技大学/之江实验室				
	电子邮箱	zhyu@hust.edu.cn		移动电话	13971637644	
	参研方式	特聘		每年项目 在岗时间	6 个月	
项目 核心 成员	姓 名	职 称	单 位	参研方式	身份证/护照号码	每年项目 在岗时间 (个月)
	廖小飞	教授	华中科技大学 /之江实验室	拟特聘	370502197803143239	3
	蒋文斌	教授	华中科技大学 /之江实验室	特聘	430402197501101018	6
	赵进	助理研 究员	华中科技大学 /之江实验室	拟双站博 后	51138119960228575X	12
	张吉	教授	南昆士兰大学 /之江实验室	特聘	320101197709041036	9
	刘海坤	教授	华中科技大学	项目聘	429001198105184432	6
	郑然	副教授	华中科技大学	项目聘	429001197709160023	6
	毛伏兵	助理教 授	华中科技大学	项目聘	429004198512031459	6
	胡侃	副教授	华中科技大学	项目聘	42010619621227533X	6
	王雄	副教授	华中科技大学	项目聘	42090219911226773X	6
	余辉	助理研 究员	华中科技大学	项目聘	350322199305110514	12
项目	序 号	单位名称		单位性质	组织机构代码	

合作 单位	1	华中科技大学	事业单位	12100000441626842D
	2			
	3			

备注:

1. 项目负责人因故无法执行项目的, 由牵头中心向之江实验室提出项目负责人变更申请。
2. 项目周期变更的, 项目负责人应提前 1 个月发起延期申请, 中心审批后提交办理。
3. 上述基本信息中其他内容的变更应当符合《之江实验室科研项目过程管理办法 (试行)》相关规定。

二、研究目标与内容

(一) 研究目标

大量应用需在极短时间处理海量图数据并支持运行时特征和需求复杂的图分析任务, 这些应用对算力和可扩展性等方面有极高需求。然而, 现有高并发分布式图计算技术面领着分布式资源利用率低、通信开销大、可扩展性欠佳、容错成本高等问题。

科学问题 1: 面向复杂图数据的高效分布式图数据访存技术

复杂的依赖关系导致分布式图数据访存效率极低; 同时, 复杂的依赖关系也使得难以通过有效的图划分消除各计算节点之间通信。如何设计高效的图数据划分机制和图数据访存机制, 以及利用新型存储硬件 (例如, NVM 和 Smart SSD) 来有效提高分布式环境下的图数据访存效率成为了亟需研究的问题。

科学问题 2: 异构体系结构下高性能高可靠的图计算运行时支撑技术

利用 GPU、FPGA 等新型硬件来提供高性能的图计算算力支撑已成为研究热点。然而, 由于图数据之间复杂冗长的依赖关系与大量底层并行计算资源的不匹配, 导致底层并行计算资源利用率低, 图计算任务计算效率差。同时, 现有分布式容错机制开销高, 分布式图计算可靠性差, 对复杂图计算支持不足。为此, 如何设计面向异构体系结构的高性能图计算引擎以及面向复杂图计算场景的容错机制已成为亟需解决的问题。

科学问题 3: 大规模分布式环境下图数据通信技术

数据之间的复杂依赖导致多计算节点之间通信效率低, 在多计算节点之间进行数据同步开销大。此外, 随着计算节点愈多, 计算效率越低, 使得分布式计算资源利用率随计算节点数量增加而降低, 分布式可扩展性差。为此, 如何设计高效的分布式通信优化机制以及利用智能网卡等新型硬件来降低分布式通信开销、降低通信时延已成为亟需解决的问题。

为此，本项目将解决高并发分布式图处理面临的分布式资源利用率低、通信开销大、可扩展性欠佳、容错成本高等问题，突破高效分布式图数据存储和访问、大规模图数据计算和高吞吐低延迟图数据通信等关键技术。项目将基于以上关键技术的研究，预期构建一套高效、高可扩展的高并发分布式图计算系统。

关键技术 1：分布式图数据存储和访问

该技术拟从分布式图数据存储和访问层面开展研究，探索高效的硬件加速的图数据存储和访问技术，高效支持大规模复杂图数据存储、划分和访问，提高数据局部性和访存效率。

关键技术 2：大规模图数据计算

该技术拟从大规模图数据计算层面开展研究，探索硬件加速的分布式图计算技术，充分利用底层大规模异构硬件资源来提高分布式图计算效率，并研究高效的分布式图计算容错机制来保证高可靠性。

关键技术 3：高吞吐、低延迟图数据通信

该技术拟从高吞吐、低延迟图数据通信层面开展研究，探索基于智能网卡的分布式图计算通信机制，降低分布式图计算通信开销和通信时延，保证分布式图计算的高可拓展性。

目标成果应用价值和影响：研制的高并发分布式图计算系统能够支持大豆育种、社会治理等 1-2 项数字反应堆应用，具备支持千亿图顶点规模的超大规模图计算的能力，性能达到业界领先水平。

(二) 研究内容

随着图数据规模的爆发式增长，图计算应用对实时性的要求越来越高，以及图计算应用的需求和特征日益复杂多样，现有方案在支持大规模分布式图计算场景时面临着计算效率低、响应速度慢等诸多挑战，难以有效满足现实应用对复杂图计算的实时响应和高可靠性等需求。为了实现高效、高可扩展的高并发分布式图计算系统，本项目计划从分布式图数据存储和访问、大规模图数据计算和高吞吐、低延迟图数据通信三个任务来展开研究：

序号	任务名称	研究内容	任务负责人	骨干成员
任务一	分布式图数据存储和访问	从分布式图数据存储和访问层面开展研究，探索高效的硬件加速的图数据存储和访问技术，高效支持大规模复杂图数据存储和访问。	蒋文斌	王雄 郑然 赵建军 赵金玮 徐嘉鸿
任务二	大规模图数据计算	从大规模图数据计算层面开展研究，探索硬件加速的分布式图计算技术，从而高效支持大规模图计算相关应用。	廖小飞	赵进 彭达 毛伏兵 余辉 齐豪 杨赞
任务三	高吞吐、低延迟图数据通信	从高吞吐、低延迟图数据通信层面开展研究，探索基于智能网卡的分布式图计算通信机制，降低分布式图计算通信开销，实现高可拓展性。	张吉	余婷 王彪 王振 许增辉 张阳 姜婷 Jaward Bah

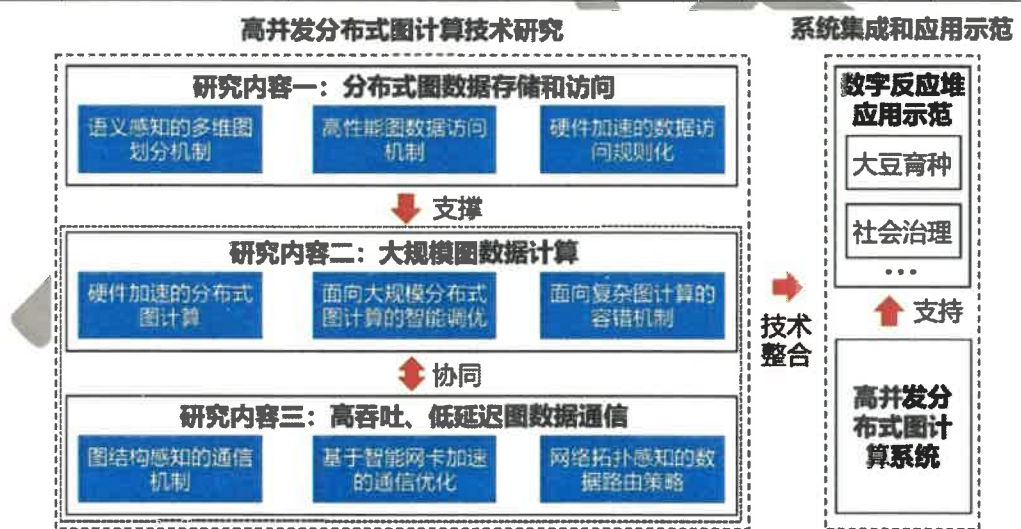


图 1 研究内容关系

本项目从分布式图数据存储、计算和通信三个层面开展研究，旨在实现一套高效、高可扩展的高并发分布式图计算系统，从而高效满足大量应用在极短时间处理海量数据和对算力高可扩展性的需求。项目研究内容关系如图 1 所示，研究内容一为研究内容二和研究内容三提供支撑，研究内容二和研究内容三相互协同形成高性能分布式图计算引擎，三个研究内容通过技术整合形成一整套高并发分布式图计算系统，并为相关应用提供支持。项目的三个任务采用模块化的研发方式，保证任务间的并行性；同

时模块之间设计统一化接口，保证项目一体化实施。

任务一：分布式图数据存储和访问

任务负责人：蒋文斌

参与人员：刘海坤、王雄、郑然、闫建荣、董雨康、何东澳、潘凡星、文苏洋、吴锐、徐嘉鸿、杨小康、姚佳波、于泽霖、张培煜、张宇健、赵建军、赵金玮

本任务重点研究语义感知的多维图划分机制、高性能图数据访问机制、硬件加速的数据访问规则化等关键技术，高效支持大规模复杂图数据存储、划分和访问，提高数据局部性和访存效率。本任务包括如下子任务：

（1）语义感知的多维图划分机制

支持多维混合的划分方式，通过对图数据特征和机器特征进行分析，对不同特征数据和机器特征自适应采用合适数据划分方式。具体而言，该任务首先对不同图划分模式特征进行分析，挖掘不同划分机制在不同数据集以及不同机器配置时的性能特征；其次，采用自适应的划分方式来根据挖掘的特征进行高效图划分，保证各计算节点的负载均衡和数据局部性。同时，在任务执行过程中，大量图顶点会不断收敛，导致各计算节点之间的负载不均衡。为此，拟进一步研究负载感知的动态划分机制来充分保证各计算节点的负载均衡。

（2）高性能图数据访问机制

通过挖掘图数据访问特征，进而从访问特征感知的数据缓存、访问特征感知的图数据预取和局部性友好的数据访存管理三个方面开展研究，设计特征感知的图数据缓存策略和访问机制。同时，利用非易失性内存（NVM）和智能 SSD（Smart SSD）等新型存储设备以及分布式异构内存（NVM+DRAM）技术来有效加快图数据的访存效率。

（3）硬件加速的数据访问规则化

通过对图分析任务依赖关系的高效动态挖掘和依赖关系感知的数据访问规则化等技术开展研究，设计基于硬件（例如 FPGA、Smart SSD 等）加速的数据访问规则化机制，从而利用硬件高效规则化图分析任务的数据访问行为，减少随机访问和数据访问冲突。

任务二：大规模图数据计算

任务负责人：廖小飞

参与人员：赵进、胡侃、毛伏兵、余辉、齐豪、何东皓、黄志颖、姜新宇、李其锟、李仕俊、廖宇健、宁鑫、彭达、沈千格、杨赞、王涛、吴凯博、夏武

本任务重点研究硬件加速的分布式图计算、面向大规模分布式图计算的智能调优、面向复杂图计算的容错机制等关键技术，充分利用底层大规模异构硬件资源来提高分布式图计算效率，并保证高可靠性。本任务包括如下子任务：

（1）硬件加速的分布式图计算

通过活跃图顶点驱动的数据提取、高性能动态负载均衡技术、关联性感知的数据重划分和差分的图数据缓存策略四个核心技术开展研究，设计基于 GPU、FPGA 等硬件单元的高性能大规模异构图计算技术，提高图计算性能。

（2）面向大规模分布式图计算的智能调优

通过图算法运行时特征分析、图算法优化空间探索、分布式图计算智能调优模型以及资源智能适配和映射四个核心技术开展研究，以此高效来挖掘图计算运行时特征，实现硬件资源的自适应配置、部署和计算高效映射，充分利用底层并行异构计算资源来加速图计算任务收敛速度。

（3）面向复杂图计算的容错机制

通过图计算任务故障和失效特征分析、主被动协同的容错机制和面向大规模复杂图数据处理的容错机制等技术开展研究，分析分布式图计算中故障和失效特征，探索高性能大规模图计算容错技术，保证大规模分布式图计算系统的高可靠性。

任务三：高吞吐、低延迟图数据通信

任务负责人：张吉

参与人员：余婷、王彪、王振、许增辉、张阳、姜婷、Mohamed Jaward Bah、侯琳琳

本任务重点研究图结构感知的通信机制、基于智能网卡加速的通信优化、网络拓扑感知的数据路由策略等关键技术，有效减少通信传输量和通信时延，保证分布式图计算的高可扩展性。本任务包括如下子任务：

（1）图结构感知的通信机制

通过研究图结构感知的通信压缩策略、非阻塞式图顶点状态同步机制和高效地址转化策略等关键技术，挖掘图计算通信特征，从而支持高性能图顶点状态同步，减少通信开销，并提高计算效率。

（2）基于智能网卡加速的通信优化

通过面向图计算任务的数据访问合并策略、面向图计算任务的规则化消息发送机制

等核心技术开展研究，以此来设计基于智能网卡（Smart NIC）加速的通信优化机制，使其能够智能适配复杂图计算通信需求，为图计算提供高吞吐、低延迟的通信能力。

（3）网络拓扑感知的数据路由策略

通过高性能网络拓扑抽取和拓扑感知的通信分发策略等技术开展研究，利用图分区结构不变性等特征来高效合并和压缩多机间的通信消息，并实现智能路由。

（三）研究思路与技术路线

本项目拟以“核心需求驱动、关键技术突破、典型应用验证”为指导，结合图应用的发展规律，对研究目标、研究内容及其层次关系进行细分，如图 2 所示。采用多层并行、多层互动的研究方法，从大规模图计算动态多变、复杂关联等特征出发，围绕高性能、高可靠性等现实需求，按照拟定的 3 个科学问题和 3 项主要研究内容，设置 3 个任务，开展研究工作。



图 2 项目研究思路

任务一：分布式图数据存储和访问

本任务拟从语义感知的多维图划分机制研究出发，随后研究高性能图数据访问机制和硬件加速的数据访问规则化技术，从而高效支持大规模复杂图数据存储、划分和访问，提高数据局部性和访存效率。

（1）语义感知的多维图划分机制

该研究内容首先计划对不同图划分模式特征进行分析，挖掘不同划分机制在不同数据集以及不同机器配置时的性能特征；其次，拟采用自适应的划分方式来根据挖掘的特征进行高效图划分，保证各计算节点的负载均衡和数据局部性。同时，拟进一步研究负载均衡的动态划分机制来充分保证各计算节点的负载均衡。

（2）高性能图数据访问机制

该研究内容通过挖掘图数据访问特征，拟研究访问特征感知的数据缓存机制和局部性友好的数据访存管理机制，高效利用分布式异构内存技术提高图数据访存效率；其次，拟研究访问特征感知的图数据预取机制，进一步降低数据访存时延。

（3）硬件加速的数据访问规则化

该研究内容首先对图分析任务依赖关系进行高效动态挖掘，随后研究依赖关系感知的数据访问规则化技术，提高数据访问局部性，减少随机访问和数据访问冲突；最后基于 FPGA、Smart SSD 等新型硬件来加速以上机制。

任务二：大规模图数据计算

本任务拟首先研究硬件加速的分布式图计算机制；最后随后研究面向大规模分布式图计算的智能调优，充分利用底层大规模异构硬件资源来提高分布式图计算效率；最后研究面向复杂图计算的容错机制，保证高性能图计算的高可靠性。

（1）硬件加速的分布式图计算

通过研究活跃图顶点驱动的数据提取机制，挖掘每轮迭代中需要处理的图数据，实现更细粒度的执行任务管理；随后研究高性能动态负载均衡技术，保证各计算核心、各异构计算资源之间的负载均衡；其次，研究关联性感知的数据重划分和差分的图数据缓存策略，进一步提高数据利用率，降低数据访存开销，提高计算效率。

（2）面向大规模分布式图计算的智能调优

该研究内容首先通过对图算法运行时特征进行分析以及对图算法的优化空间进行探索，提出分布式图计算智能调优模型；随后基于以上模型，设计资源智能适配和映射机制，实现硬件资源的自适应配置、部署和计算高效映射，充分利用底层并行异构计算资源来加速图计算任务收敛速度。

（3）面向复杂图计算的容错机制

该研究内容首先通过对图计算任务故障和失效特征进行分析，设计主被动协同的容

错机制；其次，针对现实应用的复杂多样特征，设计面向大规模复杂图数据处理的容错机制，保证大规模分布式图计算系统的高可靠性。

任务三：高吞吐、低延迟图数据通信

本任务拟首先研究图结构感知的通信机制，通过挖掘图结构特征来降低图数据通信开销；其次，研究基于智能网卡加速的通信优化，通过新型硬件来进一步加速分布式图数据通信，降低通信时延；最后，拟研究网络拓扑感知的数据路由策略，从而有效减少网络拓扑中的通信传输量和通信时延，保证分布式图计算的高可扩展性。

(1) 图结构感知的通信机制

该研究内容首先拟研究图结构感知的通信压缩策略，以此降低分布式节点之间的通信量，减少通信开销；其次，拟研究非阻塞式图顶点状态同步机制，通过高效重叠计算和通信，进一步隐藏通信时延；最后，研究高效地址转化策略，降低通信过程中全局 ID 到本地 ID 之间的转换开销，提高通信效率。

(2) 基于智能网卡加速的通信优化

该研究内容首先拟研究面向图计算任务的数据访问合并策略和面向图计算任务的规则化消息发送机制，充分合并图计算任务的通信消息，降低数据传输量以及数据传输次数，降低通信开销；随后，设计基于智能网卡（Smart NIC）加速的通信优化机制，使其能够智能适配复杂图计算通信需求，为图计算提供高吞吐、低延迟的通信能力。

(3) 网络拓扑感知的数据路由策略

该研究内容首先研究高性能网络拓扑抽取机制，在运行时高效挖掘网路图拓扑关系；随后，研究拓扑感知的通信分发策略，利用图分区结构不变性等特征来高效合并和压缩多机间的通信消息，实现智能路由。

三、项目实施计划与人才队伍建设计划

(一) 项目实施计划

从项目启动开始，按照自然年半年度一个节点开展，各节点计划如下：

时间节点	实施计划
2022 年 05 月 - 2022 年 10 月	1. 项目方案论证，技术路线制定 2. 构建通用图计算编程框架 3. 开展分布式图数据存储和访问系统研究
2022 年 11 月 - 2023 年 04 月 (中期节点)	1. Graph500 冲榜 2. 完成高并发分布式图计算系统原型设计 3. 构建硬件加速的高并发分布式图计算模块 4. 构建面向大规模分布式图计算的智能调优模块
2023 年 05 月 - 2023 年 10 月	1. Graph500 冲榜 2. 完成高并发分布式图计算容错机制 3. 构建高并发分布式图数据高吞吐通信模块
2023 年 11 月 - 2024 年 04 月 (结题节点)	1. 完成复杂图数据结构设计和存储 2. 完成整个高并发分布式图计算系统集成和优化 3. 完成高并发分布式图计算系统的测试与验证
任务一：分布式图数据存储和访问	
2022 年 05 月 - 2022 年 10 月	1. 对不同图划分模式特征分析，挖掘不同划分机制在不同数据集以及不同机器配置时的性能特征 2. 通过挖掘图数据访问特征，完成访问特征感知的数据缓存
2022 年 11 月 - 2023 年 04 月 (中期节点)	1. 完成自适应划分方式根据挖掘的特征进行高效图划分 2. 完成访问特征感知的图数据预取机制 3. 完成图分析任务依赖关系的高效动态挖掘
2023 年 05 月 - 2023 年 10 月	1. 完成负载感知的动态划分机制 2. 完成局部性友好的数据访存管理机制，利用智能 SSD 进行数据访问加速 3. 完成基于分布式异构内存(NVM+DRAM)技术的图数据访存机制 4. 完成依赖关系感知的数据访问规则化机制
2023 年 11 月 - 2024 年 04 月 (结题节点)	1. 完成复杂图数据结构设计和存储 2. 完成分布式图数据存储和访问系统整体集成、测试与验证，并集成到高并发分布式图计算系统
任务二：大规模图数据计算	
2022 年 05 月 - 2022 年 10 月	1. 完成活跃图顶点驱动的数据提取机制 2. 完成图算法运行时特征分析
2022 年 11 月 - 2023 年 04 月 (中期节点)	1. 完成高性能动态负载均衡和关联性感知的数据重划分机制 2. 完成基于 GPU 或 FPGA 等硬件加速的高性能大规模异构图计算模块 3. 完成图算法优化空间探索和分布式图计算智能调优模型设计 4. 完成图计算任务故障和失效特征分析
2023 年 05 月 - 2023 年 10 月	1. 完成差分的图数据缓存策略 2. 完成资源智能适配和映射机制 3. 完成主被动协同的容错机制

	4. 面向大规模复杂图数据处理的容错机制
2023 年 11 月 - 2024 年 04 月 (结题节点)	1. 完成硬件加速的分布式图计算引擎整体集成、测试与验证，并集成到高并发分布式图计算系统
任务三：高吞吐、低延迟图数据通信	
2022 年 05 月 - 2022 年 10 月	1. 完成图计算通信特征分析 2. 完成图结构感知的通信压缩策略
2022 年 11 月 - 2023 年 04 月 (中期节点)	1. 完成非阻塞式图顶点状态同步机制和高效地址转化策略 2. 完成面向图计算任务的数据访问合并策略和面向图计算任务的规则化消息发送机制 3. 完成高性能网络拓扑抽取机制
2023 年 05 月 - 2023 年 10 月	1. 完成高效地址转化策略 2. 完成基于智能网卡(Smart NIC)加速的通信优化机制
2023 年 11 月 - 2024 年 04 月 (结题节点)	1. 完成拓扑感知的通信分发策略 2. 完成基于智能网卡的分布式图计算通信系统整体集成、测试与验证，并集成到高并发分布式图计算系统

(填写说明：请明确每半年（每一节点）计划实施的研究内容、完成事项等。其中项目中期节点、结题节点应在下述项目实施计划中予以明确。)

(二) 里程碑节点 (阶段性考核目标)

里程碑节点	里程碑目标 (技术目标、论文、专利等)	考核方式
2023 年 04 月	1. 完成分布式异构环境下通用图计算编程框架，支持 10 个及以上典型图算法； 2. 完成高并发分布式图计算系统原型研制，在大规模分布式集群上测试(之江实验室 100 节点以上的 GPU 集群)，并支持千亿图顶点规模的超大规模图计算	1. 现场测试支持 10 个及以上典型图算法正常运行； 2. 在之江实验室 100 节点以上的 GPU 集群进行现场测试
2024 年 04 月	1. 完成高并发分布式图数据高吞吐通信系统研制； 2. 完成高并发分布式图计算系统整体集成和优化，完成 1-2 项数字反应堆中的典型图计算应用示范，并且，对于数字反应堆中相应的典型图计算场景，当运行在之江实验室整个 GPU 集群上时，相对于典型分布式图计算系统 GraphX，在支持图计算操作时也获得 10 倍以上性能提升，即图计算操作执行时	1. 现场测试高并发分布式图计算系统和 1-2 项数字反应堆中的典型图计算应用示范； 2. 在之江实验室 100 节点以上的 GPU 集群（或者太湖之光等其他超算平台之一）上进行测试； 3. 在之江实验室 100 节点以上的 GPU 集群上进行现场测试，与典型分布式图计算系统 GraphX 进行性能

	间缩短至 1/10 以内 3. 在之江实验室 100 节点以上的 GPU 集群（或者太湖之光等其他超算平台之一）上进行测试；系统具有达到 120,000GTEPS 以上的计算能力 4. 在之江实验室 100 节点以上的 GPU 集群上，高性能分布式图计算系统，相比于典型分布式图计算系统 GraphX 能够获得 10 倍以上性能提升，即图计算执行时间缩短至 1/10 以内； 5. 发表高水平学术论文 1-2 篇，申请发明专利 2-4 项	比较
任务一：分布式图数据存储和访问		
2023 年 04 月	1. 完成数据特征感知的高效图划分策略，支持 TB 级分布式图数据存储； 2. 完成分布式图数据存储和访问原型系统，相对于典型分布式图计算系统 GraphX，图数据访问时间减少 5%以上	1. 在之江实验室 100 节点以上的 GPU 集群上进行现场测试，与典型分布式图计算系统 GraphX 进行性能比较
2024 年 04 月	1. 完成基于分布式异构内存(NVM+DRAM)技术的图数据访存机制； 2. 完成分布式图数据存储和访问系统整体集成和优化，相对于典型分布式图计算系统 GraphX，图数据访存时间减少 10%以上	1. 在之江实验室 100 节点以上的 GPU 集群上进行现场测试，与典型分布式图计算系统 GraphX 进行性能比较
任务二：大规模图数据计算		
2023 年 04 月	1. 完成分布式图计算引擎的通用图计算编程框架，支持 10 个及以上典型图算法 2. 完成硬件加速的高并发分布式图计算模块开发； 3. 完成硬件加速的分布式图计算引擎原型研制，并与任务一和任务三的研究内容进行整合，实现高并发分布式图计算系统原型系统，在大规模分布式集群上测试(之江实验室 100 节点以上的 GPU 集群)，并支持千亿图顶点规模的超大规模图计算	1. 现场测试支持 10 个及以上典型图算法正常运行 2. 在之江实验室 100 节点以上的 GPU 集群进行现场测试

	4. 在之江实验室 100 节点以上的 GPU 集群上，在采用本项目设计的硬件加速的分布式图计算引擎后，相对于典型分布式图计算系统 GraphX，图计算效率提升 4 倍以上，即图计算执行时间缩短至 1/4 以内	
2024 年 04 月	1. 完成面向大规模分布式图计算的智能调优机制 2. 完成面向复杂图计算的容错机制 3. 完成硬件加速的分布式图计算引擎整体集成和优化，在之江实验室 100 节点以上的 GPU 集群上，在采用本项目设计的硬件加速的分布式图计算引擎后，相对于典型分布式图计算系统 GraphX，图计算效率提升 8 倍以上，即图计算执行时间缩短至 1/8 以内	1. 现场测试支持 10 个及以上典型图算法正常运行 2. 在之江实验室 100 节点以上的 GPU 集群进行现场测试
任务三：高吞吐、低延迟图数据通信		
2023 年 04 月	1. 完成图结构感知的通信机制机制 2. 完成分布式图计算通信系统原型研制，在采用本项目设计的分布式图计算通信系统后，相对于典型分布式图计算系统 GraphX，通信量减少 20%-50%，通信时间减少 5%以上	1. 在之江实验室 100 节点以上的 GPU 集群进行现场测试
2024 年 04 月	1. 完成基于智能网卡加速的通信优化机制 2. 完成网络拓扑感知的数据路由策略 3. 完成基于智能网卡的分布式图计算通信系统整体集成和优化，在之江实验室 100 节点以上的 GPU 集群上，在采用本项目设计的基于智能网卡的分布式图计算通信系统后，相对于典型分布式图计算系统 GraphX，通信时间减少 10%以上 4. 完成在公共卫生(如传染病分析与防治)、大豆育种和社会	1. 在之江实验室 100 节点以上的 GPU 集群进行现场测试

	治理领域等典型图计算场景上的 1 到 2 项应用示范，当运行在之江实验室 GPU 集群上时，相对于典型分布式图计算系统 GraphX，在支持图计算操作时获得 10 倍以上性能提升，即图计算操作执行时间缩短至 1/10 以内	
--	---	--

(三) 人才队伍建设计划

项目组按照节点提出各个阶段的用人需求，配合人力等部门完成人员招聘工作,完成团队建设，队伍建设节点和团队规模预期如下：

人才队伍 建设节点	全职人员 (K\E4 及以下)	全职人员 (K\E5 及以上)	全职双聘
2022 年 07 月	4	4	2
2022 年 12 月	6	4	4
2023 年 12 月	12	8	8
2024 年 06 月	18	12	10

备注：

- 项目验收时间节点，预计团队总人数不少于__40__人（包含全职人员、全职双聘、项目成员、研究生、博士生等）。
- 项目参与人员应当在参与时签订《项目参与人员承诺书》。

四、成果要求与考核指标

（一）成果要求

项目成果包括硬件装置、软件系统、应用系统、相关知识产权以及人才队伍建设成果，具体包括：

1. 新理论发展

项目将设计高效分布式图数据存储和访问原理和方法；建立大规模图数据计算软硬协同设计机理；提出高吞吐、低延迟图数据通信模型。详见考核指标。

本部分成果呈现形式为提供相应的技术总结报告。

2. 新技术开发

项目将突破分布式图数据存储和访问、大规模图数据计算和高吞吐低延迟图数据通信等关键技术，开发高效、高可扩展的高并发分布式图计算系统。详见考核指标。

本部分成果呈现形式为提供相应的技术总结报告和高并发分布式图计算系统。

3. 应用及拓展

研发的分布式图计算编程框架和支撑库能够支持大豆育种、社会治理等 1-2 项数字反应堆应用。详见考核指标。

4. 人才培养与队伍建设：

利用项目研究支撑，培养一支具备原创能力和国际竞争力的高水平人才队伍。预计在图计算相关领域，团队建设一支 40 人左右的团队。团队将全面掌握高并发分布式图计算的设计和开发技术，为我国计算机系统软件和体系结构领域发展提供人才保障。

（二）考核指标

1. 技术指标

本项目将研制一款高效、高可扩展的高并发分布式图计算系统。其中包括一套分

布式图数据存储和访问系统（任务一）、一套硬件加速的分布式图计算引擎（任务二）和一套基于智能网卡的分布式图计算通信系统（任务三）。

（1）高并发分布式图计算系统总体指标

指标一：高性能分布式图计算系统，支持之江实验室 100 节点以上的 GPU 集群，支持异构计算，支持千亿图顶点规模的超大规模图计算。在之江实验室 100 节点以上的 GPU 集群（或者太湖之光等其他超算平台之一）上进行测试，具有达到 120,000GTEPS 以上的计算能力。

指标二：高性能分布式图计算系统，在之江实验室 100 节点以上的 GPU 集群上，在开源数据集（例如 com-Friendster）上运行典型图算法（例如 PageRank, SSSP, CC, BFS），相比于典型分布式图计算系统（GraphX）能够获得 10 倍以上性能提升，即图算法执行时间缩短至 1/10 以内。

指标三：通用分布式图计算编程框架和支撑库，在数字反应堆中的大豆育种和社会治理等典型图计算场景开展 1-2 项应用示范，并且，对于数字反应堆中相应典型图计算场景，当运行在之江实验室整个 GPU 集群上时，相对于典型分布式图计算系统 GraphX，在支持图计算操作时获得 10 倍以上性能提升，即图计算操作执行时间缩短至 1/10 以内。

指标四：力争在 Graph 500 中性能排名第一。

（2）各任务指标

■ 任务一：分布式图数据存储和访问系统指标

完成数据特征感知的高效图划分策略，支持 TB 级分布式图数据存储；

在之江实验室 100 节点以上的 GPU 集群上，在采用本项目的分布式图数据存储和访问系统后，相对于典型分布式图计算系统 GraphX，图数据访存时间减少 10%以上；

■ 任务二：硬件加速的分布式图计算引擎指标

完成分布式图计算引擎的通用图计算编程框架，支持 10 个及以上典型图算法；

在之江实验室 100 节点以上的 GPU 集群上，在采用本项目设计的硬件加速的分布式图计算引擎后，相对于典型分布式图计算系统 GraphX，图计算效率提升 8 倍以上，即图计算执行时间缩短至 1/8 以内；

■ 任务三：基于智能网卡的分布式图计算通信系统指标

在之江实验室 100 节点以上的 GPU 集群上，在采用本项目设计的基于智能网卡的分布式图计算通信系统后，相对于典型分布式图计算系统 GraphX，通信量减少 20%以上。

在之江实验室 100 节点以上的 GPU 集群上，在采用本项目设计的基于智能网卡的分布式图计算通信系统后，相对于典型分布式图计算系统 GraphX，通信时间减少 10%以上。

基于项目搭建的通用分布式图计算编程框架和支撑库，在公共卫生(如传染病分析与防治)、大豆育种和社会治理领域等典型图计算场景开展 1-2 项应用示范，当运行在之江实验室 GPU 集群上时，相对于典型分布式图计算系统 GraphX，在支持图计算操作时获得 10 倍以上性能提升，即图计算操作执行时间缩短至 1/10 以内。



2. 国际对标

(1) 对标国际上最流行的图计算系统 GraphX 最新版本（Spark 3.1.3 中 GraphX）

(2) 对标 Graph 500 排名中目前国际最好图计算软硬协调系统性能

测试基准	Affiliation	性能 (GTEPS)
项目研发的高并发分布式图计算系统		120,000
BFS	RIKEN Center for Computational Science (Japan)	102,956
	National Supercomputing Center in Wuxi (China)	23,755.7
	Information Technology Center The University of Tokyo (Japan)	16,118
	EuroHPC/CS (Finland)	8,467.71
	Oak Ridge National Laboratory (USA)	7,665.7
SSSP	National Supercomputer Center in Tianjin (China)	2,054.35

	Leibniz Rechenzentrum (Germany)	1,053.93
	BDTS && SCTS && CGCL (China)	884.361
	NERSC/LBNL (USA)	558.833
	NERSC/LBNL (USA)	229.188

注：全球超级计算大会的 Graph500 和 Green Graph500 是国际上评价超级计算机图计算性能的最权威榜单，直接反映超级计算机的图数据处理能力，全球所有超级计算机都常年参与评测。

3. 评测及考核方式

(1) 性能评测方法：在大规模分布式集群上测试，并支持千亿图顶点规模的超大规模图计算；高并发分布式图计算系统具有达到 120,000GTEPS 以上的计算能力。GTEPS 计算方法参考 2010 年超级计算机大会上提出的 Graph 500 中的相关定义 (<https://graph500.org/>)，具体公式为 E/T，其中 E 代表图算法遍历边的总量，T 代表运行图算法的执行时间。

(2) 各项指标的具体评测方法

总体指标一：基于之江实验室 100 节点以上的 GPU 集群，进行功能测试；基于在之江实验室 100 节点以上的 GPU 集群（或者太湖之光等其他超算平台之一），测试高并发分布式图计算系统计算性能，进行 Graph 500 打榜。

总体指标二：基于之江实验室 100 节点以上的 GPU 集群，与典型分布式图计算系统 GraphX 进行性能比较。

总体指标三：基于之江实验室 GPU 硬件平台，测试 10 个及以上典型图算法和 1-2 项数字反应堆中的典型图计算应用示范。

总体指标四：测试结果在 Graph 500 排名榜单中排名第一

任务一指标：基于之江实验室 100 节点以上的 GPU 集群，与典型分布式图计算系统 GraphX 进行性能比较。

任务二指标：基于之江实验室 GPU 硬件平台，测试 10 个及以上典型图算法，与典型分布式图计算系统 GraphX 进行性能比较。

任务三指标：基于之江实验室 100 节点以上的 GPU 集群，测试 1-2 项数字反应堆中的典型图计算应用示范，以出具的应用证明为依据，并与典型分布式图计算系统 GraphX 进行性能比较。

(3) 申请专利以受理通知书或授权为准，论文以学术期刊或国际学术会议的论文录用通知或检索号为准。

4. 论文与知识产权

发表中国计算机学会高质量论文（CCF A/B 及 ACM/IEEE Trans.）2 篇，申请发明专利 4 项。

(三) 团队建设

在之江实验室建设一支 40 人左右的之江全职研究团队，其中全职人员 30 人左右，全职双聘 10 人。

五、知识产权归属、成果分配及其他

本项目产生的项目成果及其知识产权归属和收益分配按之江实验室与华中科技大学于 2021 年 12 月 31 日签订的《华中科技大学与之江实验室联合建立华中科技大学-之江实验室图计算联合研究中心合作协议书》约定执行。约定如下：

1. 联合研究中心成立前，甲乙双方已有的技术成果归原持有方所有。
2. 联合研究中心成立后，甲乙双方在联合研究中心合作形成的知识产权为双方共有，其中：(1) 共有知识产权的权利人排序依贡献大小为据：以共有知识产权申报科技奖励时，经由双方协商一致后确定报奖单位、人员排序。(2) 对共有知识产权双方共享收益，共有知识产权的使用、许可等由双方在行为实施前另行协商约定。

六、项目经费概预算

名 称	币种：人民币		单位：万元	
	总概算		第一年度预算	
	预算额度	总外拨	预算额度	外拨额度
项目总额	2600.00	660.00	1300.00	300.00
一、直接费用	2048.80	264.00	1040.00	120.00
（一）设备费	364.00	0.00	260.00	0.00
（二）业务费	441.80	56.00	158.50	16.00
（三）劳务费	1243.00	208.00	621.50	104.00
二、间接费用	551.20	396.00	260.00	180.00
（一）管理费	208.00	52.80	104.00	24.00
（二）激励费	343.20	343.20	156.00	156.00

备注 1：外拨经费额度，华中科技大学 660 万元。

备注 2：

1. 项目经费预算（包含外拨预算）为项目的最高拨付/外拨额度，实际项目经费按照项目开展情况及项目经费使用情况分期度拨付。
2. 设备费主要列支项目实施过程中购置或试制专用仪器设备，对现有仪器设备进行升级改造，以及租赁外单位仪器设备而发生的费用。
3. 业务费主要列支项目实施过程中消耗的各种材料、辅助材料等低值易耗品的采购、运输、装卸、整理等费用，发生的测试化验加工、燃料动力、出版/文献/信息传播/知识产权事务、会议/差旅/国际合作交流等费用，以及其他相关支出；
4. 劳务费主要列支项目实施过程中支出给参与项目的研究生、博士后、访问学者和项目聘用的研究人员、科研辅助人员等的劳务性费用；以及支出给临时聘请的咨询专家的费用等
5. 管理费指在项目（课题）实施过程中对使用现有仪器设备及房屋，日常水、电、气、暖消耗以及其他有关管理费用的补助支出。管理费按照项目总经费的 8%编制（含外拨经费）；
6. 激励费按照外拨经费的直接费用扣除设备购置费后的一定比例按需编制，原则上按照不超过实验室同类人员的月薪酬额发放。在实验室领取薪酬的人员，不得再发放激励费。
7. 设备费原则上不外拨。
8. 项目经费拨付按照之江实验室科研经费使用及财务支付管理的相关规定执行。

七、其他条款

1. 本任务书一式[陆]份，之江实验室科研发展部保管[贰]份，财务资产部保管[壹]份，研究中心或项目组保管[叁]份。
2. 本任务书系之江实验室内部文件，任务书履行中的任何问题均按照之江实验室内部规章制度和相关政策规定进行解释和执行。
3. 双方共同确认：本任务书与基于本任务书而签订的《之江实验室科研项目课题任务书》是一个不可分割的整体。如本任务书解除，《之江实验室科研项目课题任务书》一并解除。
4. 本任务书项目负责人，项目成员和课题有关人员明晰《之江实验室自设科研项目立项管理办法（试行）》《之江实验室科研项目过程管理办法（试行）》《之江实验室科研项目验收与评价管理办法（试行）》《之江实验室自设科研项目经费管理办法（修订）》、____/____和本任务书的相关内容，项目周期内将遵守相关规定和约定。如本任务书中涉及的相关管理办法发生变更的，以最新规定为准。
5. 其他：_____。

【以下无正文】

八、任务书签署

项目牵头中心（盖章）：

中心负责人（签字）：

项目负责人（签字）：

日期：2022年7月19日

项目委托单位（盖章）：

项目委托单位法定代表人（签章）：

日期：2022年7月19日



附件一：

项目负责人承诺书

本人在此郑重承诺：严格遵守《之江实验室自设科研项目立项管理办法（试行）》《之江实验室科研项目过程管理办法（试行）》《之江实验室科研项目验收与评价管理办法（试行）》《之江实验室自设科研项目经费管理办法（修订）》以及其他相关规定，将严格按下述方式履行本任务书：

（一）在本项目实施过程中，将严格遵守有关项目管理及经费管理规定，按照“政策相符性、目标相关性和经济合理性”的原则，科学、合理、真实地管理使用经费，经费不用于与本项目无关的支出。

（二）对于科研所需的材料将按照实验室的规定，在材料管理系统中认真做好材料的出入库管理。

（三）在项目实施过程中严格遵守有关保密规定，尊重他人知识产权，对于项目（课题）产出的科研成果，将及时做好专利申请、软件著作权登记等知识产权保护工作，保障实验室合法权益。

此外，在项目周期内，恪守职业规范和科学道德，遵守评审规则和工作纪律，杜绝以下行为：

1. 抄袭、剽窃他人科研成果或者伪造、篡改研究数据、研究结论；
2. 购买、代写、代投论文等；
3. 弄虚作假，违规使用经费，骗取、套取科研经费以及奖励、荣誉等；
4. 严重缺岗，在岗期间违犯之江实验室相关规章制度和管理要求；
5. 参与项目、提交的里程碑目标与项目成果违犯对境内外第三方承担的在先义务或侵犯境内外第三方的合法权益；
6. 在项目周期内，未经之江实验室同意，擅自离职或申请调出；
7. 违反保密义务，违约披露保密信息等。

项目负责人（签字）：

2022年07月19日