

IT168首页 > 数据库 > 数据库评论 > 正文

创新是国产数据库的唯一出路 | DTCC2022

作者: 卿云 编辑: 任朝阳 2022-12-21 15:04 IT168网站原创 (7) 分享

2022年12月14日~16日,由IT168联合旗下ITPUB、ChinaUnix两大技术社区主办的第13 届中国数据库技术大会(DTCC2022)在线上降重召开。大会以"数据智能价值创新"为主 题,设置2大主会场,20+技术专场,超百位行业专家,为广大数据领域从业人士提供了一 场年度的饕餮盛宴。



正如ITPUB总经理、中国数据库技术大会创办人唐川所说,从2010年开始到目前为 止,DTCC中国数据库技术大会可谓是见证了数据库技术的发展与变革。ITPUB社区也走 过了二十二个年头,伴随着多少技术人一起走过青春年华。近12年来,国内数据库市场不 断有新的厂商涌现,而本次数据库技术大会更是汇集了华为、腾讯、阿里、达梦、亚马逊 等国内外主流的数据库厂商,可谓是百花齐放。

如今的DTCC已然成为国内数据库领域的技术风向标,今年云原生数据库、分布式数 据库是行业重点发展趋势,金融业的数据库探索实践更是引发广泛讨论和关注。本次大会 设置了金融行业数据库技术实践、云原生数据库开发与实践、分布式数据库应用最 佳实践、图数据库技术与应用创新等专场,分享数据库最前沿的技术动态和探索实践。

金融业数据库升级提速

在主会场,渤海银行生产运行中心副主任、首席数据库专家王飞鹏带来"商业银行核心数据库从集中式到分布式架构转型最佳实践"的主题演讲,介绍了商业银行数据库架构演进趋势,王飞鹏指出,金融科技的本质是利用信息科技为金融发展全面赋能。在金融科技潮流之下,数据库做为商业银行承载业务数据的核心,其在选型规划中务必坚持战略思维,要从根本性、全局性、长远性的问题进行科学谋划,具体来说,稳定优先,数据库架构要首先满足银行对系统可用性、数据一致性、系统安全性等核心要求;随着银行对业务快速创新的要求日益迫切,数据库系统应支撑业务快速上线投产;还要满足自主可控要求,数据库架构要从传统商业封闭式向开源、开放、自主可控的架构转型;随着银行交易量和数据量的爆发式增长,数据库架构从集中式向分布式转型已成为必然趋势;还有降本增效,银行经营环境的变化,要求运行数据库的基础设施从昂贵的IBM小型机向PC服务器、从EMC集中式存储向本地盘等低成本设施转型。

王飞鹏介绍,银行引入分布式数据库面临着选型、开发、迁移、运维四个方面的挑战。选型方面,目前分布式数据库在国内取得快速发展,初步具备对国外数据库的替换能力,但也面临各厂商能力参差不齐的问题,需要对分布式数据库进行选型,选择最合适的产品;开发方面,对行内的应用开发人员而言,大家习惯使用Oracle等数据库,应用都是基于Oracle进行开发,迁移到分布式数据库面临代码的应用和迁移问题;迁移方面,把一些系统从集中式迁移到分布式技术栈,存在比较多的工作和风险;运维方面,不同信创数据库的实际功能、高可用能力、可扩展能力和灾备能力存在差异,为在生产运维环境做好数据库运维提出了挑战。分布式数据库产品的成熟度、稳定性还需要经过真实生产环境的打磨,才能真正稳定。

武汉达梦数据库股份有限公司技术总监黄海明指出,当前外部形势非常动荡,金融数据库的安全风险和供应链风险值得关注。当前时间紧、任务重,金融数据库国产化转型的关键因素在于解决工程化难题,要以工程化方式快速、大批量替代海外数据库,需要考虑风险、效率和成本等多方面因素。达梦的解决方案可以在架构层面实现同等替代,可以最大程度上降低替代的风险,保护用户既有投资。

创新是国产数据库发展的唯一出路

达梦数据技术(江苏)有限公司总经理付新指出,当下,数据库甚至整个软件领域最流行的两个技术架构相关的关键词是分布式和云原生。此外,可以看到数据库技术发展的重要趋势是融合,从不同的角度都可以看到融合趋势。在软件和硬件一体化方面,当下国内数据库一体机的热度开始增加,特别是基于国产软硬件平台的数据库一体机非常契合当下党政、央企等用户的需要。此外还有多模数据库、HTAP、湖仓一体、流批一体等数据技术融合发展。在融合之外,数据库技术还向着专精特新的细分方向发展。

付新强调,数据库技术发展的最终目标是持续不断降低运维成本、降低运行成本,最 终降低用户的数据库总体拥有成本。对于未来国产数据库的发展,达梦认为创新才是国产 数据库发展的唯一出路,自主创新是达梦公司坚定不移的技术发展路线。创新离不开人 才、技术、资本和需求四大要素,而中国已经具备了所有的基础条件。

云数据库加速融合发展

腾讯云数据库总经理王义成带来了题为"腾讯云数据库赋能企业释放数据生产力"的主题演讲,王义成指出,云数据库的发展经历了三个发展阶段,云数据库1.0,是云数据库发展初期,特征是原生数据库的平台化和自动化。云数据库2.0,是云原生数据库+国产化数据库多样化发展。云数据库3.0,即下一代数据库的核心趋势是自研+融合一体化输出。腾讯云数据库坚持自研和融合两个战略。其中,融合分为产品能力融合和生态融合,产品能力融合包括HTAP、软件硬件等融合,产业融合则要做到产业上下游体系真正地融合。

华为云计算技术有限公司华为云数据库服务产品部总经理苏光牛指出,数据库的发展需要自主创新,就像攀登珠穆朗玛峰时,一种是从北坡攀登,一种是从南坡攀登,数据库技术需要迎难而上,克服最核心的技术来攀上珠峰。苏光牛指出,从数据库角度来说,分布式和云基础能力的结合将是数据库市场的未来。他表示,在技术实践上,在面向金融政企的国产分布式数据领域,GaussDB在高可靠、安全、性能方面进行多项新特性发布,全面超越国外主流数据库产品;在面向互联网的云原生领域,云原生数据库完成三大转变,实现Serverless、Regionless、Modeless的全新架构演进。

阿里巴巴集团副总裁、阿里云数据库产品事业部负责人李飞飞在"PolarDB向云原生一体化的演进和发展"的主题演讲中表示,PolarDB向云原生一体化方向快速演进,提供企业级云原生数据处理能力,例如serverless、共享内存、多租户多写、分布式等,针对海量数据提供实时融合分析与计算的云原生HTAP能力以及结合ADB提供一体化HTAP能力。

阿里云认为数据库将向云原生一站式数据管理与服务纵深发展,并提出云原生化、平台化、一体化、智能化"四化"趋势,其中云原生化中最关键的能力是serverless。未来,云原生要持续做强,基于资源池化、资源解耦、微服务化、容器化等构建下一代原生数据库。随着物理世界数字化、生物世界数字化,如何更好支持多模态数据也将面临挑战。接下来三到五年,数据库市场需要关注云原生、多模、智能化、集中分布式一体化,以及软硬结合创新,未来有竞争力的数据库系统一定是和云基础设施、云平台深度融合,做软硬协同平台化创新。未来数据库需要提升数据的可信与安全,最终实现搭积木式的高性能、低成本、多场景一站式能力。

分布式数据库大有可为

OceanBase首席架构师杨志丰在"OceanBase 4.0: 单机分布式一体化的技术演进"的主题演讲中表示,从2014年OceanBase 0.5版本开始,经过多年多版本的更新,OceanBase逐渐演进的越来越成熟,并在TPC-C国际权威的OLTP评测中表现突出,具有很好的扩展性。

OceanBase 4.0作为单机分布式一体化数据库,实现了单机部署并兼顾分布式架构的扩展性与集中式架构的性能优势,不仅突破了分布式数据库单机性能的瓶颈,还实现了单机性能 赶超集中式数据库的跨越,可以更好支持不同规模的企业应用。

金篆信科副总经理秦延涛介绍了中兴通讯专为拓展分布式数据库产业而成立的控股子公司——金篆信科,分享了GoldenDB的发展历程、v7.0技术优势等。GoldenDB在金融、运营商等行业积累了大量的数据库迁移实践经验,秦延涛介绍,在系统迁移过程中,除了需要做数据全量迁移、增量迁移比对以外,一般推荐客户做回流方案,即设计逃生通道,当老系统迁移到新系统以后,GoldenDB支持数据回流到老系统数据库,如果新系统里出现异常,可以返回到原来的老系统里,确保业务不受影响,为客户增强迁移的信心,此外,从集中式系统迁移到分布式架构,需要一个好的运维管控平台,需要一套完善的运维体系来做支撑,确保运维团队快速了解系统运行状态。

图数据库的未来在于生态发展

在图数据技术与应用创新专场,来自字节跳动、蚂蚁集团、等方面的专家重点介绍了 图数据库技术的前沿发展与创新应用。去年的图数据库专场,大家谈论比较多的技术细 节,比如万亿大图的点切以及边切等,今年谈论更多的是图数据库的应用以及生态建设 等。

字节跳动研发工程师陈超带来了"字节跳动图数据架构演进——索引和执行优化"的主题演讲,聚焦于图数据库中的索引加速和执行优化,分享 ByteGraph 在性能优化方面的一些进展。ByteGraph是字节跳动自研的分布式图数据库存储系统,采用计算存储分离架构,支持有效图模型,支持Gremlin图查询语言,读写吞吐可以扩展到千万QPS,目前ByteGraph已经部署了1000多个集群,遍布全球多个机房,支持头条、抖音、西瓜视频、电商风控图谱等多个业务。在社交网络中,图计算系统会存在超级节点问题,例如抖音中网红大V会有千万或者上亿粉丝,超级节点会影响单机的可用性,针对超级节点,ByteGraph会对每条查询分为轻查询和重查询两类,同时会把单机资源划分到新查询池和

重查询池,查询时将对应类别查询,通过这种方式,限制重查询使用的资源量,避免少量 重查询打满单机CPU,后续ByteGraph也会持续优化自适应限流机制。针对超级节点,除了 限流,还会做一些自适应局部索引,在系统内部会衡量识别查询代价和构建索引后的执行 代价。

蚂蚁集团图数据库负责人洪春涛带来了"蚂蚁高性能图数据库TuGraph-DB技术思考及实践"的主题演讲,关联数据爆炸性增长对图计算高效处理提出迫切需求,图计算利用图形结构描述事物联系,成为分析复杂关系数据的重要使能技术。蚂蚁集团联合清华大学研发了大规模图计算系统TuGraph。TuGraph图计算系统由图数据库TuGraph-DB,流式图计算系统TuGraph-DataFlow,离线图计算系统TuGraph-Compute等多个子系统构成。其中TuGraph-DB是一个高性能分布式图数据库,它历经蚂蚁万亿级业务的实际场景锤炼,已经成熟应用于金融风控、搜索推荐、设备管理等内外部场景。

洪春涛指出,图计算是一种很通用的、更高层次的抽象计算方法,上个世纪70年代末80年代初其实就已经有图数据库概念,后来关系型数据库得以快速发展,而图数据库没有那么快发展起来,主要原因在于彼时计算机还没有那么强大,关系数据库这种二维表的表达对计算机会更友好,而现在多跳查询等复杂关系的查询,SQL会有挑战,图计算在多跳查询等方面会有很大的优势。洪春涛认为目前图计算系统的发展还处于一个比较早期的阶段,未来走向普及还有很多工作要做,包括易用性、性能和成本以及标准化等问题。图计算系统的发展最重要的是生态,需要有更多的用户、更多的场景来应用,才能不断推动图计算系统的发展。

第十三届中国数据库技术大会(DTCC2022)获得了圆满成功,有太多的不易,有太多需要感谢,有太多值得被记录。一点一滴,创新永无止境,众志成城,未来必将行远。中国数据库技术发展的新历史正在被书写,创新是国产数据库的唯一出路,未来的辉煌需要每一个数据库从业者共同去创造。我们有幸是时代的见证者,我们有幸是时代的参与者,在这个数据库的好时代,我们满怀热情,脚踏实地。我们期待下一次相遇,因为每一次相遇都会有前进的欢喜。DTCC2023,第十四届中国数据库技术大会,不见不散!