

全国产计算数据库性能测试报告

一、 测试背景

由于众所周知的世界经贸形势，拥有全面自主可控的 IT 解决方案成为摆在国家面前的一项重要任务。但是，由于目前国产芯片的性能还相对较弱，大多只能应用于 OA 等边缘业务，要用到关键业务中时必须加大成本以扩大集群规模，这将进一步推高因为生态成熟度还不足而导致本身成本就不低的全国产技术方案的应用门槛，抑制了全国产技术的推广应用范围。

SPL 计算数据库是一项从理论模型到代码实现都拥有完全自主知识产权的新技术，它采用了创新代数体系，允许程序员实现更高性能的低复杂度算法，从而获得比传统关系数据库更好的运算性能。因此，可以通过软件弥补硬件，使较低性能的国产芯片不需要扩大集群规模就能够适应原来必须使用国外芯片的应用场景，并且还有一定程度的性能提升。

二、 测试方法

采用国际上公认的数据库性能测试工具 TPCH，通过与国外数据库产品的杰出代表 Oracle12.1.0 及另一国产数据库(下文中简称国产数据库)进行对比测试。

TPCH 是由 TPC(Transaction Processing Performance Council)事务处理性能委员会公布的一套针对数据库决策支持能力的测试基准，通过模拟数据库中业务相关的复杂查询考察数据库的综合处理能力，获取数据库操作的响应时间。

TPCH 基准模型中定义了一个数据库模型，容量可以在 1GB~10000GB 的 8 个级别中进行选择。数据库模型包括 CUSTOMER、LINEITEM、NATION、ORDERS、PART、PARTSUPP、REGION 和 SUPPLIER 8 张数据表，涉及 22 条复杂的 SQL 查询语句。

TPCH 测试具有广泛的权威性和被认可度，涉及内容丰富，可以避免片面性，并确保测试结论的说服力。

传统数据库都是用 SQL 语言进行查询，在 SPL 数据库中，需要将 TPCH 提供的 SQL 翻译成 SPL 脚本来运行。

三、 测试环境

本场测试分别在五台不同配置的独立服务器上作运行测试，分别命名为 intel2670、intel3014、龙芯、飞腾 FT1500、飞腾 FT2000，硬件配置如下表：

名称	CPU	内存	硬盘类型
Intel2670	2 个 intel2670 CPU，主频 2.6G，共计 16 核	128G	SSD 固态硬盘
Intel3014	2 个 intel3014 CPU，主频 1.7G，共计 12 核	64G	
龙芯	2 个 CPU，共计 8 核	64G	
飞腾 FT1500	飞腾 CPU，共计 16 核	32G	
飞腾 FT2000	飞腾 CPU，共计 64 核	256G	

国产数据库只在龙芯上测试，Oracle 只在 intel 上测试，SPL 在五台服务器上均做了测试。

四、 数据准备

测试数据均采用 TPCB 提供的数据生成程序产生，产生了 100G 和 200G 两种规模的测试数据，100G 数据用在 intel3014、龙芯、飞腾 FT1500、飞腾 FT2000 四台服务器上，200G 数据用在 intel2670、飞腾 FT2000 两台服务器上。

TPCH 产生原始文本文件数据以后，需要导入数据库表中供数据库软件查询使用。SPL 数据库使用的数据存储硬盘组表文件中，也需要由原始文本数据转换生成。下表是转换成所需要的测试数据所花的时间(共有 8 张表，只比较数据量较大的 3 张表)。

数据表名		LINEITEM	ORDERS	PARTSUPP
100G	Oracle(intel3014)	09:40:59	01:31:34	00:59:02
	SPL 组表(intel3014)	01:40:00	00:17:48	00:10:47
	国产数据库(龙芯)	01:55:37	00:17:41	00:10:18
200G	Oracle(intel2670)	37:47:28	07:43:54	03:12:30
	SPL 组表(intel2670)	01:52:57	00:25:20	00:17:23

可以看出，Oracle 数据导入的时间非常长。SPL 和国产数据库的导入时间相当，这方面均比 Oracle 有数倍的性能优势。

五、 测试结果

测试的三种软件均支持并行查询，即用多线程共同协调完成一次查询任务。然而并行的路数并不是越多越好，因此在每台测试服务器上需要测出查询速度最快的并行数。我们用 TPCB 的第一条查询任务作测试，在每台测试服务上的并行查询结果如下：

TPCH1 测试响应时间表(秒)

数据量	服务器	并行数	1	2	4	8	12	16
100G	Intel3014	Oracle	570	356	219	170	131	141
		SPL 组表	336	174	91	46	38	40
	龙芯	国产数据库	1620	961	553	507	579	694
		SPL 组表	1521		525	275	284	
	飞腾 FT1500	SPL 组表	711	367	198	95		62
	服务器	并行数	8	16	32	40	48	64
	飞腾 FT2000	SPL 组表		37	22	19	21	40
200G	Intel2670	Oracle	318	316	318			
		SPL 组表	63	40	41			
	飞腾 FT2000	SPL 组表				36		

上表中红色数字对应的并行数，即为各测试服务器上最佳的并行数，于是 TPCB 后面各查询测试中就以此并行数来进行测试。

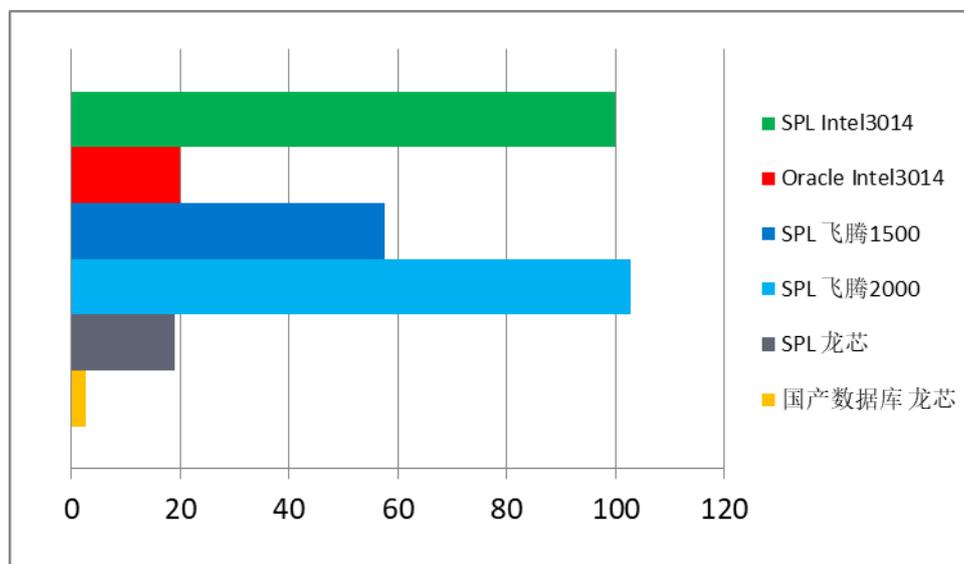
测试结果如下表：

TPCH 共 22 条查询测试响应时间表(秒)

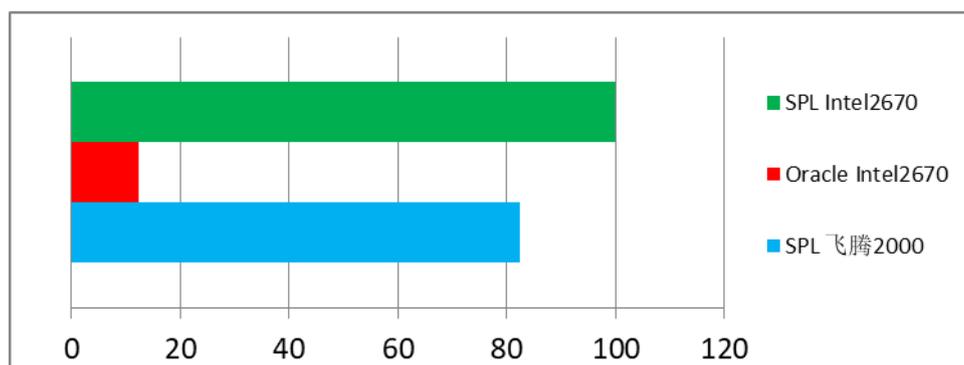
TPCH 编号	100G 数据						200G 数据		
	Intel3014 (12 并行)		FT1500 (16 并行)	FT2000 (40 并行)	龙芯 (8 并行)		Intel2670 (16 并行)		FT2000 (40 并行)
	SPL	Oracle	SPL	SPL	SPL	国产数据库	SPL	Oracle	SPL
1	38	131	62	19	275	507	40	325	36
2	4	27	8	6	18	247	8	73	13
3	16	222	33	22	97	4451	23	582	35
4	12	207	27	18	89	1790	21	454	43
5	20	225	36	24	72	1761	25	463	45
6	9	135	22	6	60	757	11	352	12
7	16	184	32	20	91	700	22	496	30
8	29	192	46	48	93	1611	29	485	80
9	68	234	125	65	517	1066	85	636	135
10	23	215	35	22	99	1634	34	493	42
11	5	33	12	6	29	165	9	63	11
12	25	184	72	38	173	647	52	464	55
13	57	37	114	85	335	2209	135	103	135
14	22	157	65	12	142	500	65	368	38
15	18	155	60	26	103	506	61	358	46
16	10	13	19	12	53	105	14	71	22
17	21	165	48	9	100	963	40	349	19
18	21	344	35	13	163	2382	25	966	26
19	23	154	65	12	137	518	60	345	29
20	18	175	57	11	110	594	55	442	20
21	233	326	222	190	901	3349	191	790	398
22	22	48	37	27	99	139	30	99	49
合计	710	3563	1232	691	3756	26601	1035	8777	1319

六、 结论

将测试结果制成统计图示：



TPCPH 100G



TPCCH 200G

可以看出如下一些结论：

1. 在同一台测试机上使用同一份 TPCCH 测试数据，比较 22 条查询所用的总时间，使用了新算法的 SPL 比 Oracle 要快 5-7 倍，创新理论和算法确实有效。
2. 在飞腾 FT1500 上比 intel3014 上的 Oracle 快了近 3 倍，在飞腾 FT2000 上比 intel2670 上的 Oracle 快了 5 倍还多。在飞腾芯片上搭载的 SPL 数据库，远远超过了 Intel 上 Oracle 的运算性能，应用于关键业务完全没有问题。
3. 龙芯相对略弱，但在 SPL 的支持下，也能和 intel3014 上的 oracle 速度基本相当（接近 95%），基本可应用于关键业务中了。
4. 龙芯上国产数据库性能较差，只能达到 Intel3014 上 Oracle 的七分之一左右，基本不具备应用于关键业务的能力。