

SPEC性能数据中台

郭健美

2022年2月

课程项目作业

- 动手做一个数据中台，实现四大核心能力
 - 汇聚整合（第5章）：数据获取和存储
 - 爬虫采集
 - 格式转换
 - 提纯加工（第6、7章）：数据处理和数据体系建设
 - 预处理、离线计算/实时分析引擎、可视化/Notebook建模
 - 贴源数据ODS、统一数仓DW、标签数据TDM、应用数据ADS
 - 服务可视化（第8、9章）：数据资产化和服务化
 - 数据资产门户
 - 查询服务、分析服务、推荐服务、圈人服务
 - 数据大屏、数据报表、智能应用
 - 价值变现（第10、11章）：中台运营
 - 可阅读、易理解、好使用、有价值
 - 价值挖掘和模式创新

数据来源 SPEC <https://spec.org/>

- Benchmark reports (三级类目)
 - CPU
 - SPEC CPU 2017 (25061 records)
 - Floating Point Rates
 - Floating Point Speed
 - Integer Rates
 - Integer Speed
 - SPEC CPU 2006 (48381 records)
 - CINT
 - CFP
 - CINT Rates
 - CFP Rates
 - Java
 - SPECjbb 2015 (582 records)
 - Composite
 - MultiJVM
 - Distributed
 - SPECjvm 2008 (12 records)
 - Power
 - SPECpower_ssj 2008 (778 records)

数据结构

- Record

CPU2017 Integer Rates

Hardware Vendor	System	Peak Result	Base Result	Energy Peak Result	Energy Base Result	# Cores	# Chips	Published	Disclosure
Huawei	Huawei TaiShan 200 Server (Model 2280) (2.6 GHz, Huawei Kunpeng 920 7260)	--	318	--	--	128	2	Jun-2020	HTML CSV PDF PS Text Config

- Report

- <https://www.spec.org/cpu2017/results/res2020q2/cpu2017-20200529-22566.html>

SPEC CPU®2017 Integer Rate Result			
Copyright 2017-2020 Standard Performance Evaluation Corporation			
Huawei (Test Sponsor: Peng Cheng Laboratory) Huawei TaiShan 200 Server (Model 2280) (2.6 GHz, Huawei Kunpeng 920 7260)		SPECrate®2017_int_base = 318 SPECrate®2017_int_peak = Not Run	
CPU2017 License: 5036 Test Sponsor: Peng Cheng Laboratory Tested by: Peng Cheng Laboratory		Test Date: May-2020 Hardware Availability: Jun-2019 Software Availability: Jul-2020	

Benchmark result graphs are available in the [PDF report](#).

Hardware								Software								
CPU Name: Huawei Kunpeng 920 7260								OS: kylin release 10 (Azalea)								
Max MHz: 2600								4.19.90-5.ky10.aarch64								
Nominal: 2600								Compiler: C/C++/Fortran: Version 9.1.0 of GCC, the GNU Compiler Collection								
Enabled: 128 cores, 2 chips								Parallel: No								
Orderable: 1,2 chips								Firmware: Huawei Corp. Version 1.16 released Feb-2020								
Cache L1: 64 KB I + 64 KB D on chip per core								File System: xfs								
L2: 512 KB I+D on chip per core								System State: Run level 5 (multi-user graphical)								
L3: 64 MB I+D on chip per chip								Base Pointers: 64-bit								
Other: None								Peak Pointers: Not Applicable								
Memory: 512 GB (16 x 32 GB 2Rx4 PC4-2933Y-R)								Other: jemalloc memory allocator V5.2.1								
Storage: 1 x 1.92 TB SAS SSD								Power Management: BIOS set to prefer performance at the cost of additional power usage								
Other: None																
Results Table																
Benchmark	Base								Peak							
	Copies	Seconds	Ratio	Seconds	Ratio	Seconds	Ratio	Copies	Seconds	Ratio	Seconds	Ratio	Seconds	Ratio		
500.perlbenc_r	128	545	374	543	375	563	362									
502.gcc_r	128	680	266	679	267	686	264									

SPEC性能数据中台

- Step 1 汇聚整合（第5章）：数据获取和存储
 - 实现爬虫程序，以及数据的动态获取和统一格式存储
- Step 2 提纯加工（第6、7章）：数据处理和数据体系建设
 - 实现数据预处理和清洗功能
 - 实现一个明确的分析引擎接口和四层数据体系接口：贴源数据ODS、统一数仓DW、标签数据TDM、应用数据ADS
 - (Optional) 离线 & 实时分析引擎
- Step 3 服务可视化（第8、9章）：数据资产化和服务化
 - 实现一个数据资产门户，包括查询服务、数据报表、数据大屏功能
 - 实现数据分析服务、推荐服务、圈人服务，以及可能的任一智能应用
- Step 4 价值变现（第10、11章）：中台运营
 - 依据可阅读、易理解、好使用、有价值，实现至少一种价值变现模式