

Zhang Z, Jin PQ, Xie XK *et al.* Online nonstop task management for storm-based distributed stream processing engines. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 39(1): 116–138 Jan. 2024. DOI: 10.1007/s11390-021-1629-9

Online Nonstop Task Management for Storm-Based Distributed Stream Processing Engines

Zhou Zhang (张洲), Pei-Quan Jin (金培权), Xi-Ke Xie (谢希科),

Xiao-Liang Wang (王晓亮), Rui-Cheng Liu (刘睿诚), Shou-Hong Wan (万寿红)

Research Objectives

- Design an online nonstop task management scheme for Storm-based DSPEs, which consists of two modules:
 - An online task migration scheme that can migrate target tasks without affecting other deployed tasks.
 - An online task deployment scheme that can smoothly adjust the task deployment.

Research Method

- Propose N-Storm, which is a task-resource decoupling DSPE. It allows tasks allocated to resources to be changed at runtime, which is implemented by a thread-level scheme for task migrations.
- Propose OTD, which can gradually adjust the current task deployment to an optimized one based on the communication cost and the runtime states of resources.

Research Results

- N-Storm can avoid the system stop and save up to 87% of the performance degradation time, compared with Storm and other state-of-the-art approaches.
- OTD can increase the average CPU usage by 51% for computation-intensive applications and reduce network communication costs by 88% for communication-intensive applications.

Research Conclusions

- Design an online nonstop task management mechanism for DSPEs to adapt to the time-varying data flows.
 - Propose N-Storm, which is a task-resource decoupling DSPE that supports thread-level online task migrations.
 - Propose OTD to realize online task deployment.
- The experimental results confirm the effectiveness of N-Storm and OTD.