How SAP is using Python to test its database SAP HANA

Christoph Heer
EuroPython 2017 - July 11
SAP HANA

- In-Memory database
- Column & Row store
- OLAP + OLTP
- Single host or distributed
- SQL
- C++ & Python
Python + SAP HANA

- SQLAlchemy
- Django
- app.py
- SAP HANA Python Client
  - DBAPI
  - SQLDBC
- Database
Testing a database

End-to-end tests
Python

Integration tests
Python

Unit/Component tests
C++
Gerrit

Pending changes

Central Repository

Pre-merge Quality Assurance

Build
Test
Static code analysis
Style checks
C++ sanitizers

Review
Implementation - 2010

Gerrit

Jenkins

Configuration

Gerrit Trigger

Job Queue

Node 1
Node 2
Node 3
... Node n

Pull Build Setup Test

Web UI

QA Database
600 Developers

700 Commits per day

13 mil. lines Python tests

36,000h Test runtime per day

1,300 Jenkins nodes

408 TB Memory
Scaling domains

- Test Runtime
- Test Scheduling
- Artifacts
- Healthy Test Environment
Test Runtime: Divide and conquer test

Pull | Build | Setup | Test

Pull | Build | Store | Setup | Test

Pull | Build | Store | Setup | Test

Pull | Build | Store | Setup | Test

Setup | Test

Setup | Test

Setup | Test

Ready for review

1  2  3  4  5  6  7
Test Scheduling: Handle test failures

<table>
<thead>
<tr>
<th>Command</th>
<th>Test Case</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>git 003adae</td>
<td>TestBaseFunctionality</td>
<td>FAILED</td>
</tr>
<tr>
<td>git 003adae</td>
<td>TestBaseFunctionality_rerun</td>
<td>OK</td>
</tr>
</tbody>
</table>

- Regression, sporadic failure or infrastructure problem
- Rerun test to
  - confirm stable failure => regression
  - exclude infrastructure problem
- Who restarts failed tests?
Test Scheduling

Configuration:

• What tests should run?
• Branch specific configuration
• Layered testing
• Exclude current broken tests

Observe and react:

• Re-schedule failed tests
• Automated review
• Notify about completion
Test Scheduling: Queuing

- Nightly test runs should be complete next morning
- Test bug fixes before new features
- Finish testing of 95% tested commit

Jenkins: First in – first out
Test Scheduling: Queuing

Prioritized Test Queue

<table>
<thead>
<tr>
<th>Priority</th>
<th>Description</th>
<th>Git SHA</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>Feature git c91abba</td>
<td>412fdae</td>
<td>Test 1</td>
</tr>
<tr>
<td>30</td>
<td>Feature git c91abba</td>
<td>412fdae</td>
<td>Test 2</td>
</tr>
<tr>
<td>30</td>
<td>Feature git c91abba</td>
<td>412fdae</td>
<td>Test 3</td>
</tr>
<tr>
<td>50</td>
<td>Bugfix git afa25ce</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>Restart git 003adae</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>nightly git 412fdae</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>nightly git 412fdae</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>nightly git 412fdae</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Waiter

Queue processor

Jenkins 1

Jenkins 2

Jenkins n
Artifacts

Installer binaries (15 GB)

NFS share

Test data (2 MB – 800 GB)

NFS share
NFS share
NFS share

Fetch  Setup  Test

Local cache  Local cache

9PB data transfer per week hitting bandwidth limits

66% less traffic to 3PB per week
Healthy Test Environment

- External dependencies
  - will fail
- Parallel testing on host
  - Noisy neighbors
- Health check before and while testing
  - Failure marks test as invalid
  - Reschedule
  - Invalid tests are invisible for developers
Implementation – 2017

- Gerrit
  - master

- Build Infrastructure

- QA Database

- Jenkins
  - Job Queue
  - Node 1
  - Node 2
  - Node 3
  - Node n

- Web UI
- QA Database

- Check
- Fetch
- Setup
- Test
Learning curve
Non-developers can write tests

Community
virtualenv, pip, Fabric, Sentry etc.

Python

Development velocity
idea to production <1 day

Platform independent
CPU architecture + OS
Outlook

- Scale for ~3000 nodes
  - Resource based scheduling with Apache Mesos

- Isolation with Linux Containers
  - Guaranteed resources: Memory, CPU, IO

- Python 3
  - asyncio, type annotations
  - even better standard library
Thank you.

We are hiring!

Open Python positions
(Junior) Python Developer: https://jobs.sap.com/s/kEui5u

(Junior) Quality Engineer: https://jobs.sap.com/s/lKieYY

(Senior) Quality Engineer: https://jobs.sap.com/s/2Cz9B8

Linux System Administrator: https://jobs.sap.com/s/5uvE3u
This presentation is licensed to the public under Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0).

The information contained herein may be changed without prior notice. Some software products marketed by SAP SE and its distributors contain proprietary software components of other software vendors. National product specifications may vary.

These materials are provided by SAP SE or an SAP affiliate company for informational purposes only, without representation or warranty of any kind, and SAP or its affiliated companies shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP or SAP affiliate company products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

In particular, SAP SE or its affiliated companies have no obligation to pursue any course of business outlined in this document or any related presentation, or to develop or release any functionality mentioned therein. This document, or any related presentation, and SAP SE’s or its affiliated companies’ strategy and possible future developments, products, and/or platform directions and functionality are all subject to change and may be changed by SAP SE or its affiliated companies at any time for any reason without notice. The information in this document is not a commitment, promise, or legal obligation to deliver any material, code, or functionality. All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements, and they should not be relied upon in making purchasing decisions.

SAP and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP SE (or an SAP affiliate company) in Germany and other countries. All other product and service names mentioned are the trademarks of their respective companies. See http://global.sap.com/corporate-en/legal/copyright/index.epx for additional trademark information and notices.