SAP EWM 9.1 on SAP HANA DB: High-Performance Warehouse Management Based on HANA DB

Table of Contents

1 About This Document .................................................................................................................. 2
2 Disclaimers and Legal Information .......................................................................................... 3
3 Introduction and Summary of Test Results .............................................................................. 4
4 Technical Landscape and Process Setup .................................................................................. 6
4.1 Targeted Production Landscape .......................................................................................... 6
4.2 Test Environment .................................................................................................................. 7
4.2.1 System Under Test .......................................................................................................... 7
4.3 Outbound Process .................................................................................................................. 8
5 Performance Results .................................................................................................................. 9
5.1 Methodology .......................................................................................................................... 9
5.2 Outbound Scenario ................................................................................................................ 11
5.2.1 Dialog Performance ....................................................................................................... 11
5.2.2 Dialog Performance of Picking and Staging .................................................................. 12
5.2.3 Scalability and Dialog Performance of Pick Confirmation in High Throughput Scenario .... 16
5.2.4 Outlier Analysis of Dialog Response Time for Picks ..................................................... 18
5.2.5 Request Performance of SAP HANA DB when Moving Data Changes ....................... 19
5.2.6 Performance of Picking Wave Release as an Example for Mass Data Transactional Processing 20
5.2.7 Processing of Incoming Outbound Delivery Orders and Wave Creation ..................... 20
5.2.8 Performance of Telegram Processing for Material Flow System (MFS) in the EWM Case 21
6 Conclusions and Outlook ......................................................................................................... 24
1 About This Document

SAP has been using rigorous performance tests to both check the reliability and scalability of SAP Extended Warehouse Management (SAP EWM) and to provide essential and comprehensive up-to-date performance and configuration information for optimization of retailer industry investment, in particular.

The results of earlier such tests published in 2010 were based on EWM 7.00 using DB2 as the database. For more information, see Retail Warehouse Management with SAP Software, Benefit from Proof of Technology and Performance and SAP EWM 7.00: High-Performance Warehouse Management for Retailers on SAP Help Portal at http://help.sap.com/extendedwarehousemanagement -> SAP Extended Warehouse Management 9.1 -> Additional Information -> How-To Guides.

The SAP EWM 7.00: High-Performance Warehouse Management for Retailers document is intended to guide and facilitate SAP EWM implementation projects by providing detailed technical insights into the EWM process configuration setup and applied technical optimizations. Furthermore, it is designed to enable successor test projects to run analogous application test scenarios.

This document provides the latest test results of SAP EWM 9.1 on the SAP HANA database system. The most important performance KPIs, ‘dialog response time in RF picking’ and ‘wave release system runtime’ were measured in what is called the ‘Scale A’ outbound process of the test setup, which is described in the document covering the test of SAP EWM 7.00 for retailers.

Both KPIs offer insight into the behavior of EWM in combination with the SAP HANA database system (HANA DB) in two technically extreme domains:

- High frequency of very short transactions
- Large mass data transactions

Under these two domain conditions, it is possible to validate the performance of EWM on HANA DB within a wide range of technical parameter spaces.

This document shall not reiterate all details of scenario setup (that is, Customizing); instead it shall provide information on the differences in hardware setup in comparison to the SAP EWM 7.00 test for retailers. As such, the document will focus on the results of performance KPIs for EWM on HANA DB as opposed to describing a general process implementation for retailers.

This test was conducted by the following contributors, listed in alphabetical order by department:

- Igor Becker, Center of Excellence (CoE), SAP SE, Walldorf, Germany.
- Jan Vandieken, Center of Excellence (CoE), SAP SE, Walldorf, Germany.
- Tobias Adler, EWM Standard Development, SAP SE, Walldorf, Germany.
- Dr. Andreas Daum, EWM Standard Development, SAP SE, Walldorf, Germany.
- Shiny Annie Besant, Suite Test Center, SAP SE, Bangalore, India.

This document is published in the same location as the earlier test conducted in 2010 that is mentioned above. See SAP EWM 9.1 on HANA DB: High-Performance Warehouse Management Based on HANA DB on SAP Help Portal at http://help.sap.com/extendedwarehousemanagement -> SAP Extended Warehouse Management 9.1 -> Additional Information -> How-To Guides.
2 Disclaimers and Legal Information

General Disclaimer

SAP does not represent or endorse the accuracy or reliability of any of the information, content, or advertisements (collectively, the "Materials") contained on, distributed through, or linked, downloaded, or accessed from any of the services contained on this Web site (the "Service"), nor the quality of any products, information, or other materials displayed, purchased, or obtained by you as a result of an advertisement or any other information or offer in or in connection with the Service (the "Products"). You hereby acknowledge that any reliance upon any Materials shall be at your sole risk. SAP reserves the right, in its sole discretion and without any obligation, to make improvements to, or correct any error or omissions in any portion of the Service or the Materials.

THE SERVICE AND THE MATERIALS ARE PROVIDED BY SAP ON AN "AS IS" BASIS, AND SAP EXPRESSLY DISCLAIMS ANY AND ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WITH RESPECT TO THE SERVICE OR ANY MATERIALS AND PRODUCTS. IN NO EVENT SHALL SAP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES OF ANY KIND WHATSOEVER WITH RESPECT TO THE SERVICE, THE MATERIALS, AND THE PRODUCTS.

© Copyright 2015 SAP SE or an SAP affiliate company. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of SAP SE or an SAP affiliate company.

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. Please see www.sap.com/corporate-en/legal/copyright/index.epx#trademark for additional trademark information and notices.
3 Introduction and Summary of Test Results

The test was conducted with SAP EWM 9.1 SP01 as an Add-On to an SAP SCM system. We have used an SCM system for this EWM test because this system was available within the SAP internal test system landscape for performance validations of SCM/APO on HANA in parallel. Given the fact that EWM has a single code line for the different deployment options, the test results provided in this document are valid for all EWM deployment options.

Since the release of SAP EWM 9.0, it is possible to deploy a dedicated EWM system as an Add-On to Application Server ABAP NetWeaver without the need of having a full-fledged SCM system instance. This smaller system is the default deployment option for new SAP EWM installations as of release 9.0. For more information about best-practice EWM deployment options, see SAP Note 1606493.

The major results of this test are summarized in the three tables below. The overview of system dialog response time in the table shows only the major step which runs very complex or mass functions in the system. Besides these major dialog steps, there are also minor dialog steps in the process with shorter response times of less than 500 milliseconds (ms), thus providing real-time system interaction. The measured RF dialog response times exclude the runtime on the frontend device, for example UI rendering times, because this highly depends on the device hardware and browsers used. When using modern hardware and browsers this runtime is in the order of 0.2 to 0.3 seconds (sec).

The throughput numbers are normalized to a unit frequency of ‘per hour’; although in several cases the measurement time span was shorter than one hour. For raw measurement data, see section 5. The measured throughput numbers are based on the test procedure which was derived from the required KPIs. The listed throughput numbers do not approach the technical limitations of SAP EWM. The exploration of these technical limits was not a test goal, it was only experimentally confirmed that the required throughput KPIs were fulfilled. The test thus focused on the validation of a well-defined business scenario rather than investigating technical limits, which often heavily depend on many factors that are independent of SAP EWM application architecture and could lead to unreliable results.

<table>
<thead>
<tr>
<th>Table 1. Response time.</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Performance KPIs</strong></td>
<td><strong>Goal</strong></td>
<td><strong>Measure</strong></td>
</tr>
<tr>
<td>Picking dialog response (RF)</td>
<td>Each step &lt; 0.9 sec (avg)</td>
<td>Major step 0.7 sec (avg)</td>
</tr>
<tr>
<td>Staging dialog response (RF)</td>
<td>Each step &lt; 0.9 sec (avg)</td>
<td>Major step 0.3 sec (avg)</td>
</tr>
<tr>
<td>Guaranteed response time in RF picking (excluding frontend device runtimes )</td>
<td>100% &lt; 2.0 sec 99.9% &lt; 1.5 sec 99.5% &lt; 1.0 sec</td>
<td>100% &lt; 1.7 sec 99.9% &lt; 1.0 sec 98.5% &lt; 0.9 sec 92.4% &lt; 0.8 sec</td>
</tr>
<tr>
<td>Picking wave release and picking warehouse order (WO) creation (desktop transaction)</td>
<td>42,000 delivery/wave items in &lt; 20 min</td>
<td>Wave with 30,000 items: First/last pick WO after 6/7 min Wave with 60,000 items: First/last pick WO after 14/15 min</td>
</tr>
<tr>
<td>MFS case conveyor routing decision response time</td>
<td>&lt; 500 ms with less than 1% having &gt; 500 ms</td>
<td>99.4% &lt; 400 ms, 99.92% &lt; 500 ms</td>
</tr>
</tbody>
</table>
Table 2. System sizing and process implementation.

<table>
<thead>
<tr>
<th>Business KPIs</th>
<th>Goal</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute hardware sizing</td>
<td>Lower than or equal to retail industry typical range, that is, in the range of 100,000 SAPS</td>
<td>Processing of picking wave outbound process and wave release for outbound wave was possible with a system which has in total approx. 80,000 SAPS, while observing 40% CPU utilization for picking and staging, with a pick frequency of 134,000 picks per hour, and 12% CPU utilization when releasing a wave with 60,000 items within 15 min.</td>
</tr>
<tr>
<td>Linear scalability of the software when increasing throughput / number of users</td>
<td>Hardware resource consumption increases linear with the load of the system, system (dialog) response time remains stable</td>
<td>Linear increasing CPU utilization and constant response times in picking and staging up to a picking frequency of 134,000 picks per hour</td>
</tr>
</tbody>
</table>

Table 3. Throughput.

<table>
<thead>
<tr>
<th>KPIs</th>
<th>Goal</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Picking (RF) and staging (RF)</td>
<td>42,000 delivery items picked into 1,680 HUs and staged in &lt; 1 hour</td>
<td>Extrapolated rate (Scale A): 134,000 picks per hour 5,360 ship HU staging per hour</td>
</tr>
</tbody>
</table>
4 Technical Landscape and Process Setup

This section covers the technical setup of the test landscape and the basic scenario description. For more details about the EWM scenario setup, see section 5 of the EWM 7.00 test documentation indicated above.

4.1 Targeted Production Landscape

The principle layout of a typical EWM production landscape is shown in Figure 1. In the center is the EWM backend, which is technically an SAP SCM system. Although EWM can be deployed on top of an SAP ERP system, we do not recommend this option in a high-performance environment. We recommend a dedicated system instance in a high-performance EWM environment. This could be either a dedicated SAP SCM system or a dedicated NetWeaver system with EWM deployed as an Add-On. For a description of backend installation, see https://service.sap.com/instguides → SAP Business Suite Applications → SAP EWM → Using SAP EWM 9.1.

To download EWM product standard guides, such as the Master Guide for SAP EWM 9.1 and the Application Operations Guide for EWM 9.1, see SAP Service Marketplace at https://service.sap.com/instguides → SAP Business Suite Applications → SAP SCM → SAP EWM → Using SAP EWM 9.1. The operations guide is especially important with regard to periodic batch jobs that need to be scheduled in a productive environment.

The EWM backend exchanges business documents (such as delivery notifications) with an SAP ERP system. The documents are transferred via qRFCs. In addition, the EWM backend exchanges telegrams with the Programmable Logic Controller (PLC) for control of the automated high rack storage or automated conveyor systems. The PLC is connected to the backend via SAP Plant Connectivity (SAP PCo). The PCo must be run in a Microsoft Windows environment; see the SAP Product Availability Matrix. For guides related to the PCo release employed in this test, see Plant Connectivity 15.0 on SAP Service Marketplace at https://service.sap.com/instguides → SAP Business Suite Applications → SAP Manufacturing → SAP Manufacturing Integration and Intelligence → Plant Connectivity 15.0.

End users can be broadly grouped into two classes: those who work in the office with the SAP frontend (SAPGUI) running on desktop PCs and those who work at the handling units (such as pallets and material handling trucks) with mobile devices that are wirelessly connected to the backend. These devices are usually equipped with a scanner so that bar codes on the handling units can be scanned and the related data be sent to the backend. The devices are connected via SAP ITS Mobile technology over the http protocol.
4.2 Test Environment

This section covers the technical details of the test environment. The setup of the system for testing, including crucial hardware/software parameters, is described in subsequent sections.

4.2.1 System Under Test

The system under test consists of one server on which the EWM/SCM ABAP NetWeaver system and the HANA system are installed. The hardware of the server comprises four CPUs of model Intel(R) Xeon(R) CPU E7- 4870 @ 2.40GHz with 30 MB cache, each having 10 cores. Each of the processing cores has two threads and thus provides two logical CPUs to OS level. All together, the operating system has 80 CPUs. The HANA database and AS ABAP server both share the 4 * 10 core CPUs and are based on Linux OS. The server has 1.0 TB of memory assigned to HANA.

According to benchmark certification for SAP component Sales and Distribution, the system has approximately 80,000 SAPS. For more information, see SAP Standard Application Benchmarks, certification number 2011015 at http://download.sap.com/download.epd?context=40E2D9D5E00EEF7C2E1EC28E273D05AA5BABF57108CEC9C65D0BF4DBFB4628AB.

Successor hardware systems available since then offer higher scalability performance (that is, more CPU cores) and faster single computing unit speed, leading to, for example, faster system dialog response time in RF.
SAP EWM on HANA DB can be deployed on many different hardware systems. The hardware system that was used in this test serves as an example system.

The http load for RF picking (via ITS Mobile) is generated with SAP LoadRunner by HP, running on a MS Windows server.

The following AS ABAP configuration was used:

- 45/21/6/2/6 DIA/UPD/BTC/SPO/UPD2 work processes (in total 80 WPs)
- Extended memory 16 GB

### 4.3 Outbound Process

Automatic wave assignment of outbound delivery order items is used to organize the daily picking workload. Therefore, to set up the wave template determination, the condition technique was used. The objective is the grouping of the outbound deliveries according to their planned goods issue (PGI) date and time. In EWM, the out-of-yard planned date can be derived from this PGI date, which is retrieved from ERP. (This ERP PGI date and time can be calculated with Transport and Shipping Scheduling in SD, including the retail route schedule as an added option.) Wave template options are introduced on an hourly basis. Based on this, the delivery items are assigned to a wave according to their planned departure from yard and route.

The creation of the warehouse order in outbound is used for two purposes: to create warehouse orders with a maximum of 50 items all belonging to the same consolidation group (in this case, goods recipient); and second, to create pick HUs of two per warehouse order with a maximum of 25 items.

The appropriate definition rule for warehouse order creation is based on a new warehouse process type for picking. This warehouse process type is also linked to a process-oriented storage control process, which consists of the steps for picking, staging and loading of goods onto transportation units. The determination of the warehouse process type is based on a special indicator for warehouse process type, which is set in the product master.

In picking, the warehouse clerk picks not by individual piece but by cartons, which consist of six pieces. Cartons are therefore the operative unit of measure. For this purpose, packaging specifications are set up and the appropriate determination is maintained. Moreover, the alternative unit of measure for cartons is maintained in every product affected.

The determination of the search sequence for a picking storage type is based on the warehouse process type, which picks goods from the storage type for fix bin. To use fix bin, individual fix bins are assigned to products used and, according to a given maximum quantity. The automated high rack is set up to replenish fix bins before they can become empty. Replenishment of fix bins is done by using a special warehouse order creation rule and warehouse process type. The latter is connected to a search sequence of storage type to determine the automated high rack storage type.

For more information about outbound process setup in the EWM application, see sections 8.1.4 and 8.2.2 in the appendix of the test document, *SAP EWM 7.00: High-Performance Warehouse Management for Retail*.
5 Performance Results

5.1 Methodology

User interaction with the EWM system for SAP EWM mobile device transactions (RF) was simulated in the test by using the technical tool, SAP LoadRunner by HP. This tool generated the network load on the SAP system, which would come from the user’s mobile devices in productive system landscapes. For this network load of mobile device transactions, the http protocol was used. For SAP EWM desktop transactions, the SAPGUI load was created by manual user interactions for single user with a classic SAPGUI installed on the user’s PC.

Most of the response times indicated in this document are based on LoadRunner measurements on http protocol level. This means that these response times include network latency time and all backend system runtimes like SAP EWM application runtimes and SAP ITS for mobile, but exclude any runtime which would be needed for presentation frontend software. Among others, these frontend runtimes typically comprise rendering runtimes for transferring the network protocol data onto UI screen elements.

In the case of desktop transactions, these local frontend runtimes are less than 0.1 sec with modern PC hardware. In the case of mobile devices, the frontend runtime strongly depends on the operating system and hardware used. In the case of high-performing OS and hardware, frontend runtimes of less than 0.15 sec can be achieved, while in other worse cases about 1 sec frontend runtime is possible. Note that in the latter case, the frontend runtimes would be higher than the total response time of all other system components. To improve these frontend runtimes of mobile devices, we strongly recommend a hardware evaluation phase covering the mobile devices in a EWM implementation project.

Some response times indicated in this document are based on AS ABAP statistical performance records (as provided by transaction STAD), which measure the response times of AS ABAP and thus exclude network times. Given the fact that during the test a fast LAN was used, the response times based on STAD data are nearly similar to the response times as measured on the LoadRunner protocol levels (which means the response times including network).

There is another important parameter in the simulation of user interaction, the so-called ‘think time.’ The ‘think-time’ is defined as the time span between screen presentation by the system and dialog action initiated by the user (for example, when pressing a button). Within the LoadRunner emulation, this ‘think-time’ was always randomly calculated according to a uniform distribution with a well-defined average and maximum / minimum values. The random distribution is needed in order to prevent possible artificial pile-up effects during recurring cycles of the automated test procedure. The average think-time was chosen per business step in order to reflect real circumstances of high throughput warehouses. For example, in picking, a think-time of 5 or 10 seconds on average was chosen for the dialog step which presents the next bin for picking while the system is waiting for the user to scan the pick bin (thus simulating fast picking as with pick-by-light or voice), while for the other ‘multiple data input’ dialog picking steps there was no think-time used. This simulates the data input of multiple screen fields in sequence by means of a mobile device enabled for scanning.

The measurements and the results presented below follow a general scheme that was introduced in the EWM 7.00 retail test. Although this test based on HANA DB is limited to a subset of process steps, we adhere to this same scheme in order to allow a direct comparison between the performance measurements in this test on HANA DB and the EWM 7.00 retail test.

- Each scenario has an abbreviation; example: ‘OB’ for outbound scenario
- Each scenario is split into several independent process phases, technically implemented by means of LoadRunner test scripts. Each script consists of a well-defined business process step or multiple steps which belong to the same process phase. Typically, all of these steps are
performed by the same group of concurrent users, often within a single transaction code. For example, ‘OB3B’ signifies warehouse workers doing picking and staging.

- Each process phase/test script consists of multiple user actions, called dialog steps (one request in the UI to the server and the server’s reply). Many dialog steps have a response time much shorter than one second, while some are optional (for example, the user can inspect data by toggling between tab strips or scrolling within tables). In the test procedure/LoadRunner script, the optional interactions were omitted, and the steps with very fast response time were not treated as important measurement results (instead they were combined into a single technical measurement via ‘defined transactions’ within the HP LoadRunner). The test procedure focuses only on the resource-intensive dialog steps, which were handled by LoadRunner as defined transactions with an individually measured average response time. The corresponding screenshots of the steps are shown in the scenario specification in the appendix of the EWM 7.00 retail performance test. The tables in the following sections display an overview of the results from response times of the major functional steps, only.

- Throughput results are given in terms of extrapolated throughput. The extrapolated throughput is calculated by number of processed items (or HUs) in the test divided by the test’s runtime. In addition to the throughput numbers of a process phase, the system resource consumption of the EWM backend was also measured and is listed for system sizing estimation purposes.

- The response time results are provided for multi-user test runs (multiple concurrent users) with a different number of users and different frequency of transactions. The comparison of these could provide insights into the concurrency management capabilities of the technical EWM implementation in conjunction with using HANA DB as the primary database system for EWM. Ideal concurrency management retains system data consistency while having a zero impact on response times. In that sense, the response times as measured in the multi-user test are always at the upper limit of the response times obtained with a smaller number of concurrent users.

The following value set of performance relevant characteristics was used for the outbound scenario. (For more details, see the EWM 7.00 test documentation):

**Outbound Scale A:** Supply to a high number of medium-sized stores: 420 outbound trucks per day in a time frame of about 16 warehouse working hours considered. Each truck has a load for five medium-sized stores. Two-hundred products in eight material handling trucks are delivered to each store. This means that each day there is a distribution to 2,100 stores. The load of each truck is five deliveries with 1,000 delivery product items in total. The load consists of 40 material handling trucks, each containing 25 different products. Each product is packed into one or several small boxes, that is, three boxes each containing six pieces of a product. These boxes are picked by means of 14 picking waves throughout the day, each wave having a size of 30,000 items and planned for 30 truckloads. Three-hundred pickers and 30 loaders are working simultaneously.

The test focused only on picking wave release and picking (including staging). The other process steps were not measured. The number of concurrent pick users was limited to less than 300 because setup of AS ABAP memory allocation was limited to a maximum use of 16 GB, which was configured somewhat too small for supporting 300 pick users with their transactional memory and their connected background processes (such as update tasks). Note that AS ABAP is capable of supporting significantly more than 300 pick users with a different setup. For example, in the EWM 7.00 retailer test, more than 20 GB memory was assigned to AS ABAP with 400 concurrent pick users.

Although the business scenario foresees picking waves with 30,000 items, the performance of releasing a wave with 60,000 items was also tested.
5.2 Outbound Scenario

5.2.1 Dialog Performance

This section summarizes dialog performance of important steps in the outbound scenario. This test focuses on wave release and picking & staging. For more information about other dialog steps, see the documentation on the EWM 7.00 retailer test. For purposes of comparison, this document describes the same technical IDs of process steps as used in the EWM 7.00 retailer test. In Figure 2, related processes in outbound are shown. Generally, there are two groups of processes; those that can be seen as pre-steps to the processing of a wave (OB2A-C) and those that comprise the outbound wave processing (OB3A-F).

The process OB2A creates replenishment warehouse tasks (WTs), which are then processed by OB2B. In the process OB2C, the transport units (TUs) are created and the ODOs assigned. Upon releasing wave (OB3A), the processing of an outbound wave starts, followed by picking and staging (OB3B).

Wave release (OB3A) is done via the EWM monitor (transaction /SCWM/MON). This is not part of a multi-user test because there is only one user releasing the outbound wave. The degree of technical processing in parallel in the EWM system, however, can be controlled by EWM Customizing. The results listed in the next table were achieved with a 10-fold parallelization. After the release wave dialog returns with a success message, further processing is done in update tasks. Because the first pickers can already start working on the first warehouse tasks (WTs) after the success message is displayed, it is worth reporting the response times with and without update processing.

The picking process (OB3B) consists of three steps in one iteration: After the user logs in via ITS Mobile and navigates to the system-guided selection where he is presented with a list of open WTs, he presses F4 (OB3B_Src_HU_1). He then confirms picking of one item (OB3B_Src_HU_2), which is done in this process 25 times for the 25 items. This is repeated one more time, because there are two pick HUs per user. He then transports the two HUs to the staging area and drops them into the destination bin, for confirmation (OB3B_DestBin_Confirm). This step is called staging.

---

**Figure 2. Outbound processes.**
For processing a single wave, each user performs two iterations. Up to 273 users performing picking and staging were tested.

<table>
<thead>
<tr>
<th>Table 4. Outbound scenario response times of the most complex and resource-intensive dialog steps.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OB3A</strong></td>
</tr>
<tr>
<td><strong>TCODE:</strong></td>
</tr>
<tr>
<td><strong>/SCWM/MON</strong></td>
</tr>
<tr>
<td><strong>Release_Wave</strong></td>
</tr>
<tr>
<td>(All UPD WO)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Release_Wave</strong></td>
</tr>
<tr>
<td>(incl. first UPD WO)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

5.2.2 Dialogue Performance of Picking and Staging

The following graphs cover a test run for wave picking with 30,000 items and 250 concurrent pick users having 10 seconds think-time, leading to a pick frequency of approximately 73,000 picks per hour in a period where all 250 pickers are active. The graphs are derived from HP LoadRunner measurements. The response times are measured on the http protocol level and include network latency, EWM ABAP system backend runtime (including DB requests), and ITS runtimes. The graphs show:

- Number of concurrent users who process picking and staging
- Number of LoadRunner transactions (these are not technical transactions / LUWs in ABAP server but the number of loops in the LoadRunner script which equals the number of picks)
- Number of transferred bytes per second as measured at HTTP interface level of LoadRunner. This data might be of relevance for network bandwidth sizing.
- Total system load as measured with HANA studio, that is, the CPU utilization during the test period caused by HANA DB processes and AS ABAP processes
- Dialog response times as measured at the HTTP request level of HP LoadRunner Tool

In this test run with 73,000 picks per hour, the CPU utilization was still at a low level of 20 percent.
Figure 3. Number of concurrent users processing picking and staging via the standard EWM RF transaction.

Figure 4. Number of HP LoadRunner tool script loops per second, named as ‘transactions per second’ in the figure. The blue line (on top) indicates the number of picks per second performed by concurrent users processing picking and staging via standard RF transaction. The green and red lines below it show the staging confirmations of the first and second HU from the resource- to the staging bin. The bottom yellow line shows the starting phase of the test run, from login to the system-guided picking by several users.
Figure 5. Used network bandwidth by user interface connectivity as measured at the HTTP level of HP LoadRunner tool.

Figure 6. System load as observed with SAP HANA Studio. The red dotted line shows CPU utilization at a level of 20 percent. This utilization consists of load generated by both AS ABAP work processes and HANA database request processing. The system load as observed with SAP HANA studio was cross-checked against system load monitoring at OS level; both showed consistent data, (that is, similar CPU utilization).
Figure 7. Response times of the major dialog steps as measured at HTTP request level of HP LoadRunner tool in outbound picking and staging with 73,000 picks per hour in the time period of 10:20 to 10:55. The step with highest response time is pick confirmation with 0.6 to 0.7 sec. The green color depicts staging confirmation when the pick HU is moved to the staging area. Note that the asynchronous delivery update was not activated. Otherwise this would lead to faster dialog response times of approximately 200 ms for pick confirmation than within the domain of less than 0.5 sec. The higher response times at the beginning of the test run can be explained by caching mechanisms in several areas of the system, such as, caching of customizing tables in AS ABAP and caching of DB SQL execution plans in the HANA DB system. After the first few executions, all of these caches are in place as it would be in a productive system. Therefore, the starting phase of the test has artificially longer dialog response times.
5.2.3 Scalability and Dialog Performance of Pick Confirmation in High Throughput Scenario

The test run described above with 73,000 picks per hour already shows very good system dialog response times in a throughput range that is sufficient for nearly any warehouse. This means that it shows the capability of processing picking for more than 500,000 outbound delivery items per day in a single eight-hour shift.

Only a few warehouses operate with a throughput in the order of 1,000,000 outbound delivery items per day. In a domain of more than 125,000 picks per hour, the system was validated for this kind of high throughput scenario by shortening the think time of the HP LoadRunner script to five seconds between picks. The two performance KPIs are:

I. Average dialog response time of AS ABAP (including DB requests) independent of system load, such as, picking throughput in terms of picks per hour, and being below or equal to 0.7 sec, thus allowing for 0.3 sec rendering time for frontend device UIs and leading to an overall dialog response of less than 1.0 sec at device UI level

II. Dialog response time of AS ABAP (including DB requests) of short duration, that is, a narrow distribution graph of dialog response times when measured many times, especially for a guaranteed upper limit in long dialog response time of outliers.

Because this test focused on SAP EWM based on HANA DB, network effects between AS ABAP and the frontend client were factored out by excluding the use of dialog response times measured with HP LoadRunner at http request level. Instead, the AS ABAP Business Transaction Analysis (transaction code STAD) was used. This data on dialog response time excludes any network fluctuations, but still includes all system performance fluctuations of AS ABAP and HANA DB. The difference in measurements between average dialog response time at AS ABAP and at HP Loadrunner HTTP request levels was very small – in the order of less than 10 ms. This is because a fast LAN was used. To account for outlier statistics, this AS ABAP level measurement approach was taken; see section 5.2.4.

<table>
<thead>
<tr>
<th>No. of Picks per Hour</th>
<th>System CPU Utilization</th>
<th>Avg Dialog Response Time</th>
<th>Standard Deviation Avg Response Time</th>
<th>ABAP CPU Runtime of Dialog WP</th>
<th>ABAP CPU Runtime of Update WP</th>
<th>DB Request Time of Dialog WP</th>
<th>DB Request Time of Update WP</th>
</tr>
</thead>
<tbody>
<tr>
<td>28,000</td>
<td></td>
<td>650</td>
<td>52</td>
<td>307</td>
<td>91</td>
<td>132</td>
<td>133</td>
</tr>
<tr>
<td>53,000</td>
<td></td>
<td>641</td>
<td>46</td>
<td>309</td>
<td>88</td>
<td>130</td>
<td>130</td>
</tr>
<tr>
<td>73,000</td>
<td>20%</td>
<td>653</td>
<td>53</td>
<td>316</td>
<td>89</td>
<td>131</td>
<td>132</td>
</tr>
<tr>
<td>117,000</td>
<td>30%</td>
<td>702</td>
<td>56</td>
<td>344</td>
<td>94</td>
<td>139</td>
<td>140</td>
</tr>
<tr>
<td>134,000</td>
<td>35 - 40%</td>
<td>718</td>
<td>61</td>
<td>350</td>
<td>96</td>
<td>144</td>
<td>144</td>
</tr>
</tbody>
</table>

The following qualitative test results can be derived from the above measurements:

- The average dialog response time of the ABAP system with EWM-on-HANA DB is very stable and nearly independent of system load.
- The very small increase in dialog response time in the order of 50 ms is primarily caused by the increased AS ABAP runtimes.
The SAP HANA DB request times are increasing by eight percent, between 28,000 and 134,000 picks per hour. The AS ABAP work process runtimes increase by 12 percent; see Figure 7. SAP HANA DB thus shows a very good scalability behavior with increasing DB request throughput. This holds true for DB read accesses (primarily present in dialog work processes) and also for DB insert/update/delete accesses which are primarily present in the update work process of DB requests.

The system shows a linear scalability in terms of CPU utilization in accordance with system throughput; the tests validated this linear behavior in a domain of up to 134,000 picks per hour.

Figure 7. Dialog response time as measured at AS ABAP level and runtime of HANA DB request processing in a range between 28,000 and 134,000 picks per hour. With increasing system throughput, the DB response time remains very stable.

Figure 8 shows the dialog response times of the test run with highest throughput of 134,000 picks per hour as measured at the HTTP request level of HP LoadRunner tool, including network times. For this test run, a sample of 50,000 delivery items and warehouse tasks was used and picking processing was done by 250 concurrent RF users with an average think-time of five seconds.
Figure 8. Response times of the major dialog steps as measured at HTTP request level of HP LoadRunner tool in outbound picking and staging with 134,000 picks per hour in the time period 11:54 to 12:08, when all 250 concurrent pick users were active. The pick confirmation step has the highest response time of approximately 0.7 sec. The green line is the staging confirmation step when the pick HU is moved to the staging area. The asynchronous delivery update was not activated because this would otherwise lead to faster response times of approximately 200 ms for pick confirmation than within the domain of 0.5 sec.

5.2.4 Outlier Analysis of Dialog Response Time for Picks

Service level agreements commonly have to guarantee a certain response time. This means that the majority of users expect to experience good (that is, low) response time. The stipulation of a threshold for average response time may often result in misleading data for analysis because, for example, 90 percent of the users might have a very low response time while 10 percent of the users might have a very high response time that they find unacceptable; in this case, the average response time would seem to still be ‘good’.

In Table 5, the average dialog response time and statistical standard deviation for pick confirmation is shown. Because the statistical distribution curve of system runtimes in computer systems is typically a non-symmetrical curve (that is, not a Gauss curve), the standard deviation is not a good measure for indicating the probability of events with a much higher runtime than on average. For non-symmetric statistical curves, the median value of the statistical distribution is the better measure; and, for highly non-symmetric
distribution curves, the outlier analysis is the better approach and methodology. The table below shows the results from a test run based on 134,000 picks per hour in the time period between 11:54 to 12:08. In this very high throughput scenario, 0.1 percent of pick events have a dialog response time of more than 1.0 sec – the maximum response time being 1.7 sec. HANA DB and the complete EWM 9.1 system clearly fulfill the aforementioned required guaranteed response times.

<table>
<thead>
<tr>
<th>Dialog Response Time</th>
<th>Number of Events/Picks from 31,241 in Total</th>
<th>Percentage of All Test Run Picks</th>
<th>Probability According to Gauß Curve, 718/61 (mean/standard deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 0.8 sec</td>
<td>2,367</td>
<td>07.60%</td>
<td>09.00%</td>
</tr>
<tr>
<td>&gt; 0.9 sec</td>
<td>476</td>
<td>01.52%</td>
<td>0.14%</td>
</tr>
<tr>
<td>&gt; 1.0 sec</td>
<td>34</td>
<td>0.11%</td>
<td>0.00%</td>
</tr>
<tr>
<td>&gt; 1.1 sec</td>
<td>12</td>
<td>0.04%</td>
<td>-</td>
</tr>
<tr>
<td>&gt; 1.2 sec</td>
<td>09 (max. value 1.7 sec)</td>
<td>0.03%</td>
<td>-</td>
</tr>
</tbody>
</table>

5.2.5 Request Performance of SAP HANA DB when Moving Data Changes

In the SAP HANA database system, you use the MERGE DELTA statement to move data organized in column store at tablespace level from the delta to the main storage structure.

As a column store table is being read, optimized, and compressed, deltas are introduced to optimize insert or update operations. All data insertions are passed to the delta storage. At a certain point in time the delta changes to a table can be merged into the table main storage. For more information, see Managing Tables in section 2.6 of the SAP HANA Administration Guide on SAP Help Portal at https://help.sap.com/hana_platform -> System Administration and Maintenance Information -> SAP HANA Administration Guide.

The system used in this test was set up so that delta merge operations were automatically triggered by the HANA DB system per single table and completely independent of any EWM application logic. The HANA system writes a log for each delta merge operation. Logs of delta merge operations on DB tables fundamental to EWM and involving picking (for example, the warehouse task table and the stock table) were searched for a correlation between longer DB request runtimes and the point-in-time of delta merge operations. The results of this analysis for a correlation were based on the STAD data of AS ABAP and the log data available in the HANA table, M_DELTA_MERGE_STATISTICS, which also records the start time of a merge run on the precision level of a ms; see Figure 8.

It was found that there was no visible correlation between the small number of pick confirmations with longer DB request runtime and HANA DB delta merge operations. This study was done with data from a test run of 73,000 picks per hour.
5.2.6 Performance of Picking Wave Release as an Example for Mass Data Transactional Processing

EWM on HANA DB was tested for high-frequency pick execution in the domain of concurrent short transactions in large numbers and SAP LUWs. At the same time, the complementary technical domain of large mass transactions with mass reading and mass insert/update database accesses was investigated. The most important business example of this kind of transaction in SAP EWM is the picking wave release. The release of waves was tested with 30,000 items according to the example for Scale A process in outbound (section 5.1) and with wave sizes doubled at 60,000 items. Technically, a 10-fold parallelization was used by means of standard EWM Customizing as was used in the earlier EWM 7.0 retailer test published in 2010. This 10-fold parallelization creates only an additional CPU utilization of 12 percent in the test system; see Figure 10. This release step creates for each wave item warehouse pick tasks and warehouse orders, which each contain 50 tasks. At the point in time when the first warehouse order is saved in the database, the system-guided pick warehouse workers can start pick execution for the wave.

Table 7. System runtime with release of large pick waves.

<table>
<thead>
<tr>
<th>Wave Size</th>
<th>Until First Pick Warehouse Order Saved in the DB</th>
<th>Until Last Pick Warehouse Order Saved in DB</th>
</tr>
</thead>
<tbody>
<tr>
<td>30,000 items</td>
<td>06 min</td>
<td>07 min</td>
</tr>
<tr>
<td>60,000 items</td>
<td>14 min</td>
<td>15 min</td>
</tr>
</tbody>
</table>

5.2.7 Processing of Incoming Outbound Delivery Orders and Wave Creation

A beneficial outcome from the test procedures for picking and wave release was the validation of the good scalability of the outbound delivery order creation, including picking wave creation and assignment. The processing of outbound delivery orders (ODOs) received from the ERP system is simulated by an ABAP
program, which creates corresponding qRFC queues. The same test procedure was applied as in the 2010 EWM 7.00 retailer test. (See in that test document, section 6.3.1).

Figure 10 shows a test run of 30,000 outbound delivery items, which were created via the standard SAP ECC -> EWM qRFC interface within 3.5 min while generating a CPU utilization load to the system at around 40 percent. This corresponds to a throughput of 514,000 delivery items and assigned wave items per hour.

For this test run, 150 outbound delivery orders were created via the ECC -> EWM interface in the time period of 10:58 to 11:02, each having 200 items. This also includes picking wave creation by automatic wave assignment of deliveries through PPF actions. The release of the wave with 30,000 items was performed in the time period of 11:06 to 11:13. Finally, picking and staging was run in the time period of 11:23 to 11:48 with a maximum of 273 concurrent users and a peak frequency of 117,000 picks per hour in the time period of 11:34 to 11:37, when all users were active.

Figure 10. System load as observed in HANA Studio for a test run of 30,000 items with outbound delivery order creation, picking wave creation and assignment, wave release and picking and staging execution. Creation of outbound delivery orders with wave creation and assignment was observed with a throughput of 514,000 deliveries and wave items per hour at a system CPU utilization level of approximately 40 percent.

5.2.8 Performance of Telegram Processing for Material Flow System (MFS) in the EWM Case Conveyor Scenario

Performance testing of the EWM MFS case conveyor scenario was conducted. In this scenario, automatic case conveyors are controlled by routing decisions derived from EWM, which records in real time case stock location on a specific conveyor segment. The movement of the case on the conveyor is not halted to await the routing decision; instead, the routing decision is taken fast enough so that the physical case movement is not slowed down. An example of a case conveyor scenario in practice might be as follows:
1. The programmable logic controller (PLC) of the case conveyor system sends a ‘Routing Request’ telegram to EWM when a case passing in front of the conveyor system’s fork is scanned.

2. EWM immediately sends back in acknowledgement to the PLC a logical confirmation telegram in receipt of the ‘Routing Request’ telegram. It is necessary to notify the PLC of the successful receipt of telegram in order to prevent repeated transmission of the ‘Routing Request’ telegram to EWM, which otherwise supposedly failed to send.

3. EWM derives the routing decision for this specific HU/case based on the rules which are configured in EWM and sends a ‘Planned Routing’ telegram back to the PLC.

4. The PLC immediately sends back in acknowledgement to EWM a logical confirmation telegram in receipt of the ‘Planned Routing’ telegram.

5. The PLC steers the hardware fork of the conveyor system according to the information received by the ‘Planned Routing’ telegram and then sends to EWM an ‘Actual Routing’ telegram to notify EWM system about the actual physical routing execution.

6. EWM immediately sends back in acknowledgement to the PLC a logical confirmation telegram in receipt of the ‘Actual Routing’ telegram.

7. EWM performs asynchronous processing of a HU warehouse task movement to the conveyor segment where the HU was physically routed.

Per single case routing, this integration scenario consists of three telegrams received by EWM, three telegrams sent by EWM, and one HU warehouse task with stock update. The most important performance KPI is the time difference as measured at the PLC between sending the ‘Routing Request’ telegram and receiving the ‘Planned Routing’ telegram. This time period has to be small enough to guarantee that the HU/case has not already arrived at the conveyor fork before the steering of the fork can be triggered by the PLC based on the routing decision received from EWM.

In this test we used a simulation tool to simulate four conveyor system PLCs. The test simulated a throughput of 92,000 case routings in total per hour, leading to an overall number of 552,000 telegrams per hour. EWM received and sent 77 telegrams per second. In addition, 92,000 HU warehouse move tasks were processed per hour. Figure 11 shows a test excerpt on the telegram response time over a period of five minutes for both sending and receiving at the PLC level.
It is important to have guaranteed response times with only a very small number of outliers to ensure a smooth operation of the conveyor system without the need to move cases/HUs to clarification zones or buffer cycle loops because a routing decision had come too late for the steering of the hardware fork. In Table 8 below, a stable guaranteed response time with only six out of 1,000 telegrams having longer than 400 ms response time, and only eight out of 10,000 having longer than 500 ms response time is shown.

Table 8. Telegram response time for 92,000 case routing decisions per hour.

<table>
<thead>
<tr>
<th>Response Time</th>
<th>&lt; 200 ms</th>
<th>&lt; 300 ms</th>
<th>&lt; 400 ms</th>
<th>&lt; 500 ms</th>
<th>&lt; 750 ms</th>
<th>&gt; 750 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of Occurrence</td>
<td>78.7%</td>
<td>96.3%</td>
<td>99.4%</td>
<td>99.92%</td>
<td>100%</td>
<td>none observed</td>
</tr>
</tbody>
</table>

Figure 11. Telegram response time in millisecond as measured at PLC level between sending of ‘Routing Request’ and receiving of ‘Routing Planned’ telegram.
6 Conclusions and Outlook

The result of this latest test further proves the capability of SAP EWM-on-HANA DB in the domain of up to one million delivery items and manual RF picks in a single eight-hour shift and for MFS case conveyor scenarios with more than 700,000 case routings in an eight-hour shift for use in high throughput warehouses.

The HANA DB system fulfills quality requirements for fast and guaranteed dialog response times in processing time-critical information daily on the level of a single stock keeping unit (SKU), such as RF pick confirmations. These latest test results for RF picking in the domain of short and high frequency transactions are equal to the test results obtained with EWM 7.0 on DB2.

The system demonstrates very good performance when processing mass transactions, especially during picking wave release. The runtime of releasing large picking waves with EWM 9.1 on HANA DB is approximately only half the runtime observed in the EWM 7.00 on DB2 retailer test.

This test is a first step in a roadmap towards realization of SAP EWM-on-HANA DB by assuring the business that EWM transactional operations can continue to run even better without disruption when using the SAP HANA DB system. The next obvious step toward realization of SAP EWM-on-HANA DB is to couple the strengths of HANA technology together with the EWM application domain. This can be done, for example, by combining OLTP and OLAP in a single system in especially the area of operational analytics and reporting with such new products as Smart Business for EWM and HANA Live for EWM. For more information about these new products, see the SCN blog at https://scn.sap.com/docs/DOC-55154. Within the EWM application domain based on standard AS ABAP, you can also use HANA DB for embedded application functions, such as Labor Demand Planning.