

การเปรียบเทียบประสิทธิภาพการจัดเก็บและสืบค้นข้อมูล Log ระบบห้องสมุดอัตโนมัติ (ALIST) ระหว่าง ORACLE และ SOLR

Performance Comparison of Storing and Retrieving for Automated Library System (ALIST) Data Logs between ORACLE and SOLR

ชำนาญ อินทสโร^{1*}

Chumnan Inthasaro^{1*}

บทคัดย่อ

จุดประสงค์ของงานวิจัยนี้คือเปรียบเทียบพื้นที่จัดเก็บข้อมูล และเวลาของการสืบค้นข้อมูล Log ระหว่างฐานข้อมูล ORACLE และฐานข้อมูล SOLR จากข้อมูล Log ของระบบสืบค้นและระบบการยืมคืนของห้องสมุด ของระบบห้องสมุดอัตโนมัติ ALIST ผลการวิจัยพบว่าพื้นที่จัดเก็บข้อมูลที่เก็บใน SOLR มีขนาดเล็กกว่าที่จัดเก็บใน ORACLE ประมาณ 42 เปอร์เซ็นต์ ระยะเวลาในการสืบค้นข้อมูลแบบตัวอักษร วันที่ รวมถึงการนับจำนวนแยกตามกลุ่มข้อมูลด้วย SOLR ใช้เวลาน้อยกว่า ORACLE ประมาณ 75 เปอร์เซ็นต์ ในระหว่างที่การสืบค้นข้อมูลแบบเงื่อนไขตัวเลข และสืบค้นด้วยตัวอักษรแบบเงื่อนไขเดียวการสืบค้นบน ORACLE ใช้เวลาน้อยกว่าบน SOLR ประมาณ 30 เปอร์เซ็นต์

คำสำคัญ: ระบบห้องสมุดอัตโนมัติ ALIST ORACLE SOLR

Abstract

The purpose of this research was to compare the data log size and retrieving time between ORACLE database and SOLR database of library searching log (Online Public Access Catalog) and library circulation system log of the automated library system (ALIST). The research result found that the data size in SOLR was smaller than in ORACLE about 42 percent. Retrieving time of alphabetic criteria, date criteria and data grouping and counting (facet) with SOLR was about 75 percent small than ORACLE while retrieving time of searching with numerical criteria and single alphabetic criteria on ORACLE was smaller than SOLR approximately 30 percent.

Keywords: Automated Library System, ALIST, ORACLE, SOLR

บทนำ

ระบบห้องสมุดอัตโนมัติเพื่อสถาบันอุดมศึกษาไทย ALIST [1-2] เป็นซอฟต์แวร์สำหรับบริหารจัดการระบบห้องสมุด ใช้ระบบฐานข้อมูล ORACLE สำหรับจัดเก็บข้อมูลในรูปแบบตารางข้อมูล (Tables) และจัดเก็บกิจกรรมการใช้งานระบบลงในตาราง Log (Log Tables) เมื่อปริมาณข้อมูลที่บันทึกใน Log Tables มีจำนวนมากขึ้นและขนาดของตารางข้อมูลที่โตขึ้นอย่างรวดเร็ว การทำงานเช่น การสืบค้นประวัติข้อมูล การจัดทำรายงานต่างๆ จาก Log Tables ใช้เวลานานมาก ดังนั้นงานวิจัยนี้จึงมีแนวคิดเปรียบเทียบฐานข้อมูลอื่นๆ กับฐานข้อมูล ORACLE เพื่อแก้ปัญหาดังกล่าว

¹ นักวิชาการคอมพิวเตอร์ ศูนย์คอมพิวเตอร์ มหาวิทยาลัยสงขลานครินทร์ สงขลา 90112

¹ Computer Technical Officer, Computer Center, Prince of Songkla University, Songkhla, 90112

* Corresponding author: Tel.: 074-289090. E-mail address: chumnan.i@psu.ac.th

ระบบฐานข้อมูลถือได้ว่าเป็นสิ่งจำเป็นพื้นฐานสำหรับการจัดเก็บและสืบค้นข้อมูล ระบบฐานข้อมูลในกลุ่ม SQL แบ่งออกเป็น 3 ประเภท คือ 1. ระบบฐานข้อมูลแบบ SQL 2. ระบบฐานข้อมูลแบบ NoSQL และ 3. ระบบฐานข้อมูลแบบ NewSQL [3] โดยระบบฐานข้อมูลในกลุ่ม SQL ได้แก่ ORACLE, SQL Server, MySQL เป็นต้น ระบบฐานข้อมูลแบบ NoSQL ได้แก่ SOLR, MongoDB เป็นต้น และระบบฐานข้อมูลแบบ NewSQL ได้แก่ SciDB, C-Store, H-Store เป็นต้น งานวิจัยนี้ได้ศึกษาและเปรียบเทียบระบบฐานข้อมูลในกลุ่ม SQL และ NoSQL เท่านั้น

ระบบฐานข้อมูลแบบ SQL และ NoSQL มีความแตกต่างกันคือ ฐานข้อมูลแบบ SQL มีลักษณะการจัดเก็บข้อมูลในรูปแบบตารางข้อมูล (Table) สามารถเข้าถึงข้อมูลได้อย่างรวดเร็วผ่านดัชนี (Index) สามารถสร้างความสัมพันธ์ระหว่างตารางข้อมูล (Table Relationship) ได้ เหมาะกับการทำงานบน Transaction Table ในขณะที่ฐานข้อมูลแบบ NoSQL เช่น SOLR มีลักษณะการจัดเก็บข้อมูลแบบไม่เป็นโครงสร้าง รองรับการทำงานกับข้อมูลปริมาณมากได้ดี สืบค้นข้อมูลได้อย่างรวดเร็ว [4-6] แต่ทั้งนี้ฐานข้อมูลทั้ง 2 สามารถทำงานร่วมกันได้ดี [7]

งานวิจัยนี้จะเปรียบเทียบพื้นที่จัดเก็บข้อมูล Log ของการสืบค้น (Searching Logs) และข้อมูล Log จากระบบการยืมคืน (Circulation Logs) บน ORACLE เปรียบเทียบกับพื้นที่จัดเก็บบน SOLR และเปรียบเทียบเวลาของการสืบค้นข้อมูล Log ของตารางข้อมูลทั้งสองฐานข้อมูล ซึ่งเป็นการสืบค้นจากเครื่องแม่ข่ายฐานข้อมูลหลัก ALIST Database Server ไปยัง ORACLE Log Server และ SOLR Log Server โดยมีจุดประสงค์คือ 1. เพื่อศึกษาว่าการจัดเก็บข้อมูล Logs บนระบบฐานข้อมูลแบบใดใช้พื้นที่น้อยกว่า และ 2. เพื่อศึกษาว่าการสืบค้นข้อมูล Logs จากระบบข้อมูล ORACLE และ SOLR ระบบใดใช้เวลาน้อยกว่ากัน เพื่อจะได้นำผลวิจัยมาแก้ปัญหาเรื่องพื้นที่จัดเก็บข้อมูลและความเร็วในการสืบค้นข้อมูลห้องสมุดต่อไป

การออกแบบและดำเนินการวิจัย

1. ศึกษาข้อมูล Log ที่เก็บในฐานข้อมูลระบบห้องสมุด

จากการศึกษาระบบห้องสมุดอัตโนมัติเพื่อสถาบันอุดมศึกษาไทย ALIST ซึ่งใช้ฐานข้อมูล ORACLE สำหรับบริหารจัดการข้อมูลและจัดเก็บข้อมูลของระบบ ข้อมูลทั้งหมดถูกจัดเก็บในรูปแบบตาราง ได้แก่ ข้อมูลบรรณานุกรม หนังสือ, ข้อมูลตัวเล่มทรัพยากร, ข้อมูลพื้นฐานต่างๆ ของระบบ รวมถึงจัดเก็บกิจกรรม Log ต่างๆ ได้แก่ การเพิ่มหรือปรับปรุงข้อมูลบรรณานุกรม การให้บริการยืมคืนของเจ้าหน้าที่ การสืบค้นข้อมูลของสมาชิกผ่านเครือข่ายอินเทอร์เน็ต การจองหรือยืมต่อของสมาชิกผ่านเว็บไซต์ห้องสมุด เป็นต้น

จากการศึกษาพบว่าจำนวนข้อมูลที่เก็บใน Log Tables ของระบบห้องสมุดอัตโนมัติในขั้นตอนการยืมคืนหนังสือได้บันทึกข้อมูลลงในหลายตาราง ได้แก่ 1. ตารางข้อมูล Log การยืมคืนหนังสือ 2. ตารางข้อมูลคิวการจอง 3. ตารางข้อความการเตือน 4. ตารางข้อมูลการจ่ายค่าปรับ เป็นต้น และในส่วนการสืบค้นข้อมูลห้องสมุด ข้อมูล Log ซึ่งบันทึกอยู่ในตาราง Log ของการสืบค้น ซึ่งข้อมูล Log การสืบค้นมีปริมาณข้อมูลเพิ่มขึ้นอย่างรวดเร็วเช่นเดียวกันเนื่องจากผู้ใช้บริการสามารถสืบค้นข้อมูลห้องสมุดได้สะดวกในหลายช่องทาง เช่นการสืบค้นผ่านเครื่องคอมพิวเตอร์ส่วนบุคคล (Personal Computers) หรือจากอุปกรณ์แบบพกพา (Mobile Devices)

จากปัญหาในเรื่องพื้นที่จัดเก็บข้อมูลและเวลาของการสืบค้นข้อมูลห้องสมุดของ ผู้วิจัยได้นำข้อมูล Log จากฐานข้อมูลห้องสมุดทั้ง 6 ห้องสมุดของมหาวิทยาลัยสงขลานครินทร์ ซึ่งได้ถูกจัดเก็บในฐานข้อมูล ORACLE มาแสดงโดยในตารางที่ 1 แสดงถึงจำนวนข้อมูลและขนาดของพื้นที่ Log การยืมคืนหนังสือ ส่วนตารางที่ 2 แสดงจำนวนข้อมูลและขนาดของพื้นที่ Log การสืบค้นหนังสือตามลำดับ

ตารางที่ 1 จำนวนข้อมูล Log ของการยืมคืนหนังสือและขนาดพื้นที่บนฐานข้อมูล ORACLE

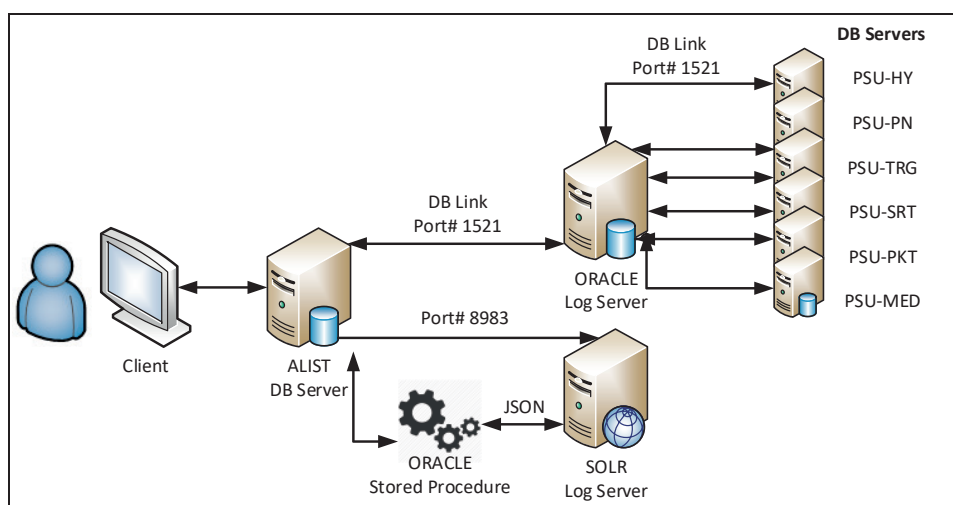
ห้องสมุดสาขา	จำนวนข้อมูล (Records)	ขนาดพื้นที่ (MB)
1. ห้องสมุดวิทยาเขตหาดใหญ่ (PSU-HY)	6,284,368	1,341
2. ห้องสมุดวิทยาเขตปัตตานี (PSU-PN)	3,067,507	688
3. ห้องสมุดวิทยาเขตตรัง (PSU-TRG)	420,413	96
4. ห้องสมุดวิทยาเขตสุราษฎร์ธานี (PSU-SRT)	386,618	80
5. ห้องสมุดวิทยาเขตภูเก็ต (PSU-PKT)	184,196	40
6. ห้องสมุดวิทยาศาสตร์สุขภาพ (PSU-MED)	1,137,156	248
รวม	11,480,258	2,493

ตารางที่ 2 จำนวนข้อมูล Log ของการสืบค้นและขนาดข้อมูลบนฐานข้อมูล ORACLE

ห้องสมุดสาขา	จำนวนข้อมูล (Records)	ขนาดพื้นที่ (MB)
1. ห้องสมุดวิทยาเขตหาดใหญ่ (PSU-HY)	4,127,610	448
2. ห้องสมุดวิทยาเขตปัตตานี (PSU-PN)	4,496,094	480
3. ห้องสมุดวิทยาเขตตรัง (PSU-TRG)	1,069,377	118
4. ห้องสมุดวิทยาเขตสุราษฎร์ธานี (PSU-SRT)	713,901	88
5. ห้องสมุดวิทยาเขตภูเก็ต (PSU-PKT)	1,153,307	120
6. ห้องสมุดวิทยาศาสตร์สุขภาพ (PSU-MED)	1,139,754	132
รวม	12,700,043	1,386

2. จัดเตรียม SOLR Server และ ORACLE Server

ผู้วิจัยได้ออกแบบสถาปัตยกรรมของระบบดังภาพที่ 1 ซึ่งเป็นการเชื่อมต่อกันของเครื่องแม่ข่าย ประกอบด้วย 1. เครื่องแม่ข่าย ORACLE สำหรับสืบค้นข้อมูล (ALIST DB Server) 2. เครื่องแม่ข่าย ORACLE สำหรับเก็บข้อมูล Log รวมจากฐานข้อมูลทั้ง 6 ห้องสมุด (ORACLE Log Server) 3. เครื่องแม่ข่าย SOLR สำหรับจัดเก็บข้อมูล Log รวมจากฐานข้อมูลทั้ง 6 ห้องสมุด (SOLR Log Server) 4. เครื่องแม่ข่ายฐานข้อมูล 6 ห้องสมุด (DB Servers) ซึ่งมีรายละเอียดดังนี้



ภาพที่ 1 รูปแบบการเชื่อมต่อเครื่องแม่ข่าย ALIST DB Server กับ ORACLE Log Server และ SOLR Log Server

1. ALIST Database Server ทำหน้าที่เป็นเครื่องแม่ข่ายหลักจัดตั้งฐานข้อมูล ORACLE ไว้ภายในสำหรับสืบค้นข้อมูลจาก ORACLE Log Server และ SOLR Log Server
2. ORACLE Log Server ทำหน้าที่เป็นเครื่องแม่ข่ายฐานข้อมูล ORACLE สำหรับรวบรวมข้อมูล Log จาก ORACLE Database Server ทั้ง 6 ห้องสมุด และส่งผลการสืบค้นข้อมูลไปยัง ALIST Database Server โดย ORACLE Log Server มีคุณลักษณะดังนี้ CPU Intel® Xeon® CPU E7-4870 @2.40 GHz RAM 8 GB HD 200 GB และติดตั้ง ORACLE XE ไว้ภายใน
3. SOLR Log Server ทำหน้าที่เป็นเครื่องแม่ข่ายฐานข้อมูล SOLR จัดเก็บข้อมูล Log ในรูปแบบ XML ที่แปลงมาจากข้อมูลใน ORACLE Log Server โดยส่งผลการสืบค้นข้อมูลในรูปแบบ JSON ไปยัง ALIST Database Server ซึ่งข้อมูลรูปแบบ JSON นี้เองจะถูกแปลงด้วย ORACLE Stored Procedure เป็นรูปแบบตารางก่อนนำไปแสดงผล โดย SOLR Log Server มีคุณลักษณะเดียวกันกับ ORACLE Log Server
4. DB Servers เครื่องแม่ข่ายฐานข้อมูลทั้ง 6 ห้องสมุด จัดเก็บข้อมูลระบบห้องสมุดบนฐานข้อมูล ORACLE กระจ่ายกันไปตามวิทยาเขตต่างๆ ของมหาวิทยาลัยสงขลานครินทร์

การเชื่อมต่อแลกเปลี่ยนข้อมูลต่อระหว่างเครื่องแม่ข่ายใช้หมายเลขพอร์ตมาตรฐานของแต่ละระบบคือ การเชื่อมต่อของฐานข้อมูล SOLR จะใช้หมายเลขพอร์ต 8983 ส่วนของฐานข้อมูล ORACLE จะใช้หมายเลขพอร์ต 1521

3. การจัดการข้อมูล Log ของระบบห้องสมุดอัตโนมัติ

ขั้นตอนการจัดการข้อมูล Log ของระบบห้องสมุดอัตโนมัติสำหรับการทดสอบประกอบด้วย 1. การจัดเตรียมและนำเข้าข้อมูลสำหรับ ORACLE Log Server และ 2. การจัดเตรียมและนำเข้าข้อมูลสำหรับ SOLR Log Server ซึ่งมีรายละเอียดดังนี้

3.1 การจัดเตรียมและนำเข้าข้อมูลสำหรับ ORACLE Log Server

เชื่อมต่อฐานข้อมูล ORACLE Log Server ที่ได้เตรียมไว้เพื่อดึงข้อมูลจากฐานข้อมูลห้องสมุดทั้ง 6 สาขา คือ PSU-HY, PSU-PN, PSU-TRG, PSU-SRT, PSU-PKT และ PSU-MED มาสร้างเป็นตารางข้อมูลแบบรวมฐาน ซึ่งประกอบด้วยตารางข้อมูล Log การยืมคืนหนังสือ Circulation Log Tables และตารางข้อมูล Log การสืบค้น Searching Log Tables

3.2 การจัดเตรียมและนำเข้าข้อมูลสำหรับ SOLR Log Server

นำข้อมูล Log การยืมคืนหนังสือและข้อมูล Log การสืบค้นหนังสือจาก ORACLE Log Server มาแปลงในรูปแบบไฟล์ข้อมูล XML และนำเข้าสู่ SOLR Log Server

การสืบค้นข้อมูลจาก SOLR Log Server จะได้ผลลัพธ์ในรูปแบบข้อมูล JSON ซึ่งเป็นโครงสร้างในรูปแบบ (Key : Value) ดังภาพที่ 2 ดังนั้นเพื่อให้ข้อมูลรูปแบบ JSON สามารถแสดงในรูปแบบ Table บน ORACLE ได้ จึงจำเป็นต้องจัดทำโปรแกรม ORACLE Stored Procedure บน ALIST DB Server เพื่อแปลงข้อมูลรูปแบบ JSON ให้มีรูปแบบ Table พร้อมสำหรับการแสดงผลบน ORACLE

<pre>{ "responseHeader": { "status":0, "QTime":0, "params":{ "q":"branch_code:PSUHY"}, "response":{"numFound":5351564, "start":0, "docs":[{ "id":"LC-eae88a07476241b8963c10cd55535d5d", "log_type_no":1, "operator_username":["nikon.h"], "ip_address":["192.168.26.52"], "patron_no":54873, "item_no":165001, "detail":0, "log_datetime":["2010-08-02T21:17:48Z"], "branch_code":["PSUHY"], "_version_":1594227364096639023, "operator_username_str":["nikon.h"], "ip_address_str":["192.168.26.52"], "branch_code_str":["PSUHY"]}], }</pre>	<pre>{ "responseHeader": { "status":0, "QTime":191, "params":{ "q":"branch_code:PSUHY", "facet.field":"searchtype", "rows":"10", "facet":"on"}, "response":{"numFound":4127610, "start":0, "docs":[{ "id":"pushy-5707483", "searchtext":"A 224969 224", "searchgroup":"B", "searchtype":"ti=", "branch_code":"PSUHY", "ipaddress":"172.28.15.56", "recordcount":519, "log_datetime":"2015-01-16T19:29:43Z", "_version_":1594838404288217155},]}</pre>
--	---

ภาพที่ 2 รูปแบบข้อมูลแบบ JSON จากการสืบค้นข้อมูลบน SOLR Log Server
ของ Circulation Logs และ Searching Log ตามลำดับ

ผลการวิจัย

การทดสอบการวิจัยได้ทดสอบโดยใน 2 ประเด็นคือขนาดพื้นที่จัดเก็บข้อมูลและความเร็วของการสืบค้นข้อมูล ผลการวิจัยพบว่า 1. ขนาดพื้นที่จัดเก็บข้อมูลบนฐานข้อมูล SOLR ใช้พื้นที่จัดเก็บข้อมูลขนาดเล็กกว่าฐานข้อมูล ORACLE และ 2. เวลาของการสืบค้นข้อมูลใน 4 กลุ่มเงื่อนไข คือ ตัวเลข ตัวอักษร วันที่ และนับจำนวนในกลุ่ม โดยทดสอบใน 10 รูปแบบ พบว่าการสืบค้นจากฐานข้อมูล ORACLE ใช้เวลาน้อยกว่าในเงื่อนไขตัวเลข และตัวอักษรแบบเงื่อนไขเดียว ส่วนเงื่อนไขอื่นๆ การสืบค้นจากฐานข้อมูล SOLR จะใช้เวลาน้อยกว่าโดยมีรายละเอียดผลการทดสอบดังนี้

1. การทดสอบพื้นที่จัดเก็บข้อมูล จากการวิจัยพบว่า เมื่อนำข้อมูล Log การยืมคืนหนังสือ (Circulation Logs) ที่จัดเก็บบน ORACLE Log Server มาจัดเก็บบน SOLR Log Server ขนาดพื้นที่ลดลงประมาณ 48 % ดังตารางที่ 3 และข้อมูล Log การสืบค้นหนังสือ (Searching Logs) ที่จัดเก็บใน SOLR Log Server มีขนาดของพื้นที่ที่จัดเก็บลดลงประมาณ 42 % ดังตารางที่ 4 ตามลำดับ

ตารางที่ 3 ขนาดพื้นที่การจัดเก็บข้อมูล Log การยืมคืนหนังสือและเปอร์เซ็นต์การลดลงของขนาดพื้นที่บน ORACLE Log Server เทียบกับบน SOLR Log Server

ห้องสมุดสาขา	จำนวนข้อมูล (Records)	ขนาดพื้นที่ (MB) บน ORACLE	ขนาดพื้นที่ (MB) บน SOLR	ขนาดพื้นที่ (MB) ที่ลดลง (%)
1. ห้องสมุดวิทยาเขตหาดใหญ่ (PSU-HY)	6,284,368	1,341	608	55
2. ห้องสมุดวิทยาเขตปัตตานี (PSU-PN)	3,067,507	688	420	39
3. ห้องสมุดวิทยาเขตตรัง (PSU-TRG)	420,413	96	56	42
4. ห้องสมุดวิทยาเขตสุราษฎร์ธานี (PSU-SRT)	386,618	80	32	60
5. ห้องสมุดวิทยาเขตภูเก็ต (PSU-PKT)	184,196	40	27	33
6. ห้องสมุดวิทยาศาสตร์สุขภาพ (PSU-MED)	1,137,156	248	153	38
รวม	11,480,258	2,493	1,296	48

ตารางที่ 4 ขนาดพื้นที่จัดเก็บข้อมูล Log การสืบค้นข้อมูลหนังสือ และเปอร์เซ็นต์การลดลงของขนาดพื้นที่บน ORACLE Log Server เทียบกับบน SOLR Log Server

ห้องสมุดสาขา	จำนวนข้อมูล (Records)	ขนาดพื้นที่	ขนาดพื้นที่	ขนาดพื้นที่
		(MB)	(MB)	(MB)
		บน ORACLE	บน SOLR	ที่ลดลง (%)
1. ห้องสมุดวิทยาเขตหาดใหญ่ (PSU-HY)	4,127,610	448	289	35
2. ห้องสมุดวิทยาเขตปัตตานี (PSU-PN)	4,496,094	480	292	39
3. ห้องสมุดวิทยาเขตตรัง (PSU-TRG)	1,069,377	118	59	50
4. ห้องสมุดวิทยาเขตสุราษฎร์ธานี (PSU-SRT)	713,901	88	38	57
5. ห้องสมุดวิทยาเขตภูเก็ต (PSU-PKT)	1,153,307	120	60	50
6. ห้องสมุดวิทยาศาสตร์สุขภาพ (PSU-MED)	1,139,754	132	63	52
รวม	12,700,043	1,386	801	42

จากการวิจัยพบว่าฟิลด์ข้อมูลที่จัดเก็บใน Log table มีหลายประเภทดังแสดงในภาพที่ 2 ได้แก่ ข้อมูลประเภทตัวเลข (patron_no, item_no), ประเภทตัวอักษร (id, operator_username, ip_address) และประเภทวันที่ (log_datetime) เป็นต้น ดังนั้นการสืบค้นข้อมูลจึงสามารถระบุเงื่อนไขได้ในทั้ง 3 ประเภทข้างต้น ดังตัวอย่างในภาพที่ 3 และยังสามารถสืบค้นแบบนับจำนวนแยกตามกลุ่มข้อมูลได้อีกด้วย โดยรายละเอียดดังนี้

```
{
  "responseHeader": {
    "status":0,
    "QTime":0,
    "params":{"q":"operator_username:praneet.w",
      "fl":["branch_code,operator_username,log_datetime"]},
    "response":{"numFound":80464, "start":0, "docs":[
      {
        "operator_username":["praneet.w"],
        "log_datetime":["2014-11-13T10:23:24Z"],
        "branch_code":["PSUHY"]},
      {
        "operator_username":["praneet.w"],
        "log_datetime":["2014-11-13T10:12:41Z"],
        "branch_code":["PSUHY"]},
      {
        "operator_username":["praneet.w"],
        "log_datetime":["2014-11-13T10:12:45Z"],
        "branch_code":["PSUHY"]},
      {
        "operator_username":["praneet.w"],
        "log_datetime":["2014-11-13T10:23:32Z"],
        "branch_code":["PSUHY"]}]}]}
```

ภาพที่ 3 ตัวอย่างการสืบค้นด้วยตัวอักษรแบบเงื่อนไขเดียวฟิลด์ประเภทตัวอักษร (operator_username) และการสืบค้นแบบหลายเงื่อนไขด้วยฟิลด์ตัวอักษรและวันที่ (operator_username AND log_datetime)

2. การทดสอบความเร็วของการสืบค้นข้อมูล ผู้วิจัยได้ทดสอบเปรียบเทียบความเร็วของการสืบค้นระหว่างฐานข้อมูล ORACLE และ SOLR ในเงื่อนไข 4 กลุ่ม คือ 1. กลุ่มเงื่อนไขตัวเลข (รูปแบบที่ 1 และ 2) 2. กลุ่มเงื่อนไขตัวอักษร (รูปแบบที่ 3 และ 4) 3. กลุ่มเงื่อนไขวันที่ (รูปแบบที่ 5 และ 6) และ 4. กลุ่มเงื่อนไขสำหรับการสืบค้นแบบนับจำนวนกลุ่ม (รูปแบบที่ 7 ถึง 10) พบความเร็วที่แตกต่างกันดังตารางที่ 5

ตารางที่ 5 ความเร็วการสืบค้นข้อมูลจาก ALIST Database Server ไปยัง ORACLE Log Server เปรียบเทียบกับ SOLR Log Server หน่วยเวลาความเร็วเป็นมิลลิวินาที (ทดสอบ 10 รอบในแต่ละรูปแบบ)

รูปแบบการสืบค้น/เรียงลำดับ	ความเร็วการสืบค้นข้อมูล (มิลลิวินาที)		เปรียบเทียบความเร็ว
	บน ORACLE	บน SOLR	
1. สืบค้นด้วยตัวเลขแบบเงื่อนไขเดียว	91	141	ORACLE เร็วกว่า 35 %
2. สืบค้นด้วยตัวเลขแบบหลายเงื่อนไข	110	161	ORACLE เร็วกว่า 32 %
3. สืบค้นด้วยตัวอักษรแบบเงื่อนไขเดียว	94	110	ORACLE เร็วกว่า 15 %
4. สืบค้นด้วยตัวอักษรแบบหลายเงื่อนไข	974	140	SOLR เร็วกว่า 86 %
5. สืบค้นด้วยวันที่แบบเงื่อนไขเดียว	9,766	125	SOLR เร็วกว่า 99 %
6. สืบค้นด้วยวันที่แบบหลายเงื่อนไข	15,061	125	SOLR เร็วกว่า 99 %
7. สืบค้นแบบนับจำนวนกลุ่มตัวเลข	5,340	875	SOLR เร็วกว่า 84 %
8. สืบค้นแบบนับจำนวนกลุ่มตัวอักษร	8,410	406	SOLR เร็วกว่า 95 %
9. สืบค้นแบบนับจำนวนกลุ่มวันที่	8,448	1,922	SOLR เร็วกว่า 77 %
10. สืบค้นแบบนับจำนวนหลายกลุ่มข้อมูล	20,052	3,180	SOLR เร็วกว่า 84 %

จากผลการทดสอบการสืบค้นข้อมูลตามตารางที่ 5 พบว่า การสืบค้นข้อมูลบนฐานข้อมูล ORACLE ทำงานได้เร็วในการสืบค้นข้อมูลด้วยเงื่อนไขตัวเลข และการสืบค้นข้อมูลด้วยเงื่อนไขตัวอักษรแบบเดียว (รูปแบบที่ 1-3) ส่วนการสืบค้นประเภทอื่นๆ รวมถึงแบบนับจำนวนแยกตามกลุ่ม (รูปแบบที่ 4-10) การทำงานบน SOLR จะทำงานได้เร็วกว่าบนฐานข้อมูล ORACLE มาก

การอภิปรายผลและสรุปผลการวิจัย

จากการวิจัยพบว่าเมื่อนำข้อมูล Log การยืมคืนหนังสือและข้อมูล Log การสืบค้นข้อมูล ที่จัดเก็บบนฐานข้อมูล ORACLE ไปจัดเก็บบนฐานข้อมูล SOLR พบว่าขนาดพื้นที่จัดเก็บบนฐานข้อมูล SOLR มีขนาดเล็กลงประมาณ 40 % เมื่อเปรียบเทียบกับบนฐานข้อมูล ORACLE ส่วนการสืบค้นข้อมูลบนฐานข้อมูล SOLR ใช้เวลาน้อยกว่าบนฐานข้อมูล ORACLE ประมาณ 75 % ในเกือบทุกรูปแบบเงื่อนไข ยกเว้น แบบเงื่อนไขตัวเลขและเงื่อนไขตัวอักษรแบบเงื่อนไขเดียว ซึ่งการค้นบนฐานข้อมูล ORACLE ใช้เวลาน้อยกว่าบนฐานข้อมูล SOLR ประมาณ 30 % ดังนั้นจึงสรุปได้ว่าเมื่อแยกข้อมูล Log ออกจากฐานข้อมูลหลัก ORACLE ไปจัดเก็บในฐานข้อมูล SOLR ขนาดพื้นที่จัดเก็บข้อมูล Log จะมีขนาดเล็กลงและเวลาของการสืบค้นข้อมูลน้อยลง ทำให้ในการแก้ปัญหาเรื่องพื้นที่จัดเก็บและความเร็วของการสืบค้น อาจจำเป็นต้องปรับโครงสร้างฐานข้อมูลของระบบห้องสมุดอัตโนมัติ ALIST ให้มีการทำงานประสานกันของฐานข้อมูลทั้ง 2 ระบบ แต่ทั้งนี้ยังอาจมีประเด็นที่จะต้องศึกษาเพิ่มเติมต่อไป ในเรื่องความคุ้มค่าของการจัดเตรียมเครื่องแม่ข่ายที่เพิ่มขึ้น เสถียรภาพของเครือข่าย และการแลกเปลี่ยนข้อมูลของฐานข้อมูลทั้ง 2 ระบบ ให้มีความถูกต้องและทันสมัย เพื่อให้การทำงานของระบบห้องสมุดอัตโนมัติทั้งระบบมีความถูกต้อง มีเสถียรภาพและมีความเชื่อถือ

เอกสารอ้างอิง

- [1] Automated Library System for Thai Higher Education Institutes. (2017). **General Information** (Online). Retrieved 28 September 2017, from <http://www.alist.psu.ac.th>.
- [2] Inthasaro, C. (2014). “Database Management System for Automated Library System ALIST”, **PULINET Journal**. 1(3), 85-90.

- [3] Kepner, J., Gadepally, V., Hutchison, D., Jananthan, H., Mattson, T., Samsi, S. and Reuther, A. (2016). “Associative Array Model of SQL, NoSQL, and NewSQL Databases”, **2016 IEEE High Performance Extreme Computing Conference (HPEC)**. Waltham, MA, 2016, 1-9.
- [4] Sirisup, K. and Utakrit, N. (2014). “On the Application of Solr Software for Helpdesk and Retrieval Issue TOT IPTV”, In **The 10th National Conference on Computing and Information Technology**. 884-889. 8-9 May 2014, Angsana Laguna Phuket. Bangkok: King Mongkut’s University of Technology North Bangkok.
- [5] Pengmusaw, T., Witosurapot, S. and Tangkuptanon, W. (2009). “On the Application of Solr Software for Thai-Script Searching in Courses”, In **The Prince of Songkla University Academic Conference, PSU Open Week**. 18 August 2009, The 60th Anniversary of His Majesty the King’s Accession to the Throne International Convention Center. Songkhla: Prince of Songkla University.
- [6] Chanta, S. and Porrawatpreyakorn, N. (2013). “A Web News Information Classification and Retrieval System using Multilayer Perceptron Neural Network”, **Information Technology Journal**. 9(2), 14-19.
- [7] Timasornwichakit, R., Paoon, W. and Phongpech, W. (2016). “Comparative Study between Big Data Technology and Relational Database Query System Case Study: Healthcare Dataset”, **Journal of the Thai Medical Informatics Association**. 2, 134-145.