الثورة الصناعية الرابعة و أدواتها
(HADOOP-SAPHAHN-NOSQL)
أنموذجاً

The fourth industrial revolution and its tools:
(HADOOP-SAPHAHN-NOSQL)
as a model

د.نهال أحمد الشاذلي
مدرس المكتبات و المعلومات-كلية الآداب-جامعة المنوفية

ملخص
هذه الدراسة تتلخص في تحديد العديد من أدوات تحليل البيانات الضخمة و انتقاء أشهر و أهم ثلاث أدوات و هم هادوب-ساب هنا- نو إس كيو إل- لإجراء المقارنة واستنتاج الأعلي جودة و الأفضل وإظهار النسبة المئوية فيما بينهم و قد أظهرت النتائج تقارب شديد بين الأدوات الثلاثة محل الدراسة بالرغم من تقدم ساب هنا علي الأداتين الأخرتين و لكن تظل الأدوات الثلاثة مشتركون في العديد من المميزات و أشارت الدراسة إلي بعض المصطلحات التي تستخدم في البحث والتي يستمع إليها القاريء مثل الحوسبة السحابية علي سبيل المثال فقد يختلط علي البعض وجود تقارب بينها و بين أدوات تحليل البيانات إلا أن مع استعراض المصطلح وجد أن لا يوجد مجال للمقارنة بينهم لعدم اشتراكهم في أي من المؤشرات.

أهمية الدراسة :-

إن أهمية أدوات و تقنيات تحليل البيانات الضخمة تعد أهم من وجود البيانات الضخمة متراكمة لا فائدة سوي بالقليل منها و بذلك فإن تجاهل البيانات أكثر تكلفة من عملية الحفاظ عليها بل إن تسليم المعلومات في الوقت الحقيقي أو شبه الحقيقي هي واحدة من الخصائص المميزة لتحليل البيانات الضخمة، و بذلك فإن من الضروري البحث عن أدوات تحليل البيانات الضخمة و نظرا لأن هذه الدراسة تقوم علي الادوات الأكثر شهرة و من ثم تستمد الدراسة أهميتها من أهمية الأدوات موضوع الدراسة و التي يسهل الاستفادة من البيانات الضخمة و التي بدورها تساعد المؤسسات التعليمية و غيرها في :

- اكتشاف الرؤى المهمة القابلة للتنفيذ .
- تحديد البيانات الأكثر أهمية للفئات التي تستهدفها من الطلاب و الهيئة التدريسية و غيرهم من ذوي المصالح المشتركة و توجيه القرارات المستقبلية .

- تفهم احتياجات المستفيدين بشكل أكثر و التعرف علي كيفية زيادة المعلومات الاستخبارية و زيادة الكفاءة و تحسين العمليات و خدمات العملاء و تحسس مواطن الخطر و تحسين إدارة المخاطر .

ظاهرة الدراسة :-

تواجه المؤسسات التي تمتلك بيانات ضخمة تحديا كبيرا في كيفية الاستفادة من الكم الهائل من البيانات التي تمتلكها و ذلك لان حجم الاستفادة التي تحصل عليه تلك المؤسسات يعطي ميزة تنافسية لانها تعطي فهما أعمق عن الأشياء و الأشخاص و تساعد علي اتخاذ القرارات الصائبة و يري بعضهم أن البيانات الضخمة تتألف من بيانات منظمة و تشكل جزءاً ضئيلاً يصل إلي 10% و البيانات الغير منظمة تمثل بقية المحتوي.

و ستتناول الباحثة في هذه الدراسة أدوات و تقنيات تحليل البيانات الضخمة مع توضيح الفرق بينهم و كيفية الاستفادة من البيانات الضخمة في ضوء استخدامهم.

أهداف الدراسة:

تهدف الدراسة إلي مجموعة من العناصر التالية:

1. تعريف التحديات التي تواجه المؤسسات التي لديها بيانات ضخمة .
2. انتقاء أدوات تحليل البيانات الضخمة. (البنية و الانتشار و الأهمية)
3. تحديد الفروق بين أدوات تحليل البيانات الضخمة موضوع الدراسة.

تساؤلات الدراسة :

1. ما البيانات الضخمة مفهومها و أنواعها ومدي انتشارها طرق الاستفادة منها؟
2. ما أدوات تحليل البيانات الضخمة (البنية-الانتشار –الاهمية)
3. ما الفروق بين أدوات تحليل البيانات الضخمة موضوع الدراسة ؟

منهج الدراسة و أدوات جمع البيانات :-

تعتمد الدراسة علي المنهج الوصفي التحليلي في الادوات المختلفة المنوطه بالبحث و المنهج المقارن للتعرف علي الفروق بين أدوات تحليل البيانات موضوع الدراسة.

بداية نبدأ بالتعرف علي مصطلحات الدراسة و التي منها:-

البيانات الضخمة:-

تعرف البيانات الضخمة بأنها مجموعة من البيانات سواء غير مهيكلة و التي تتمثل في الرسائل النصية و البريد و الفيديوهات و الشبه مهيكلة و التي تتمثل في ملفات الـ xml و لها معمارية معينة و البيانات المهيكلة و التي تتمثل في قواعد البيانات الهادوب HADOOP

هو أحد برامج ال Open Source مصمم للتعامل مع البيانات الضخمه (Big data) يستخدم لغة الجافا و عن طريقه تستطيع أن تحول البيانات الضخمة إلى إحصائيات لتبني عليها قرارات

ساب هنا SAPHANA

أداة لتحليل البيانات الضخمة يساعد مستخدميه في القيام بعمليات التحليل المتزامن للبيانات الضخمة و التي ترد إلي منصة النظام مما يساعد علي سرعة اتخاذ القرار

نو إس كيو إل NOSQL

قواعد بيانات تتيح تخزين و استرجاع البيانات بعد نمذجتها بتقنيات تختلف عن تلك المستخدمة في نمذجة البيانات في قواعد البيانات العلائقيه و تستخدم هذه التكنولوجيا لما تقدمه من بساطه في التصميم و قابلة للتوسع و لذلك استخدمت في تطبيقات البيانات الضخمة و تطبيقات الزمن الحقيقي للمساعدة في التخلص من التكلفة و البطء الشديد لقواعد البيانات العلائقية

الحوسبة السحابية CLOUD COMPUTING

هي تكنولوجيا متطورة تعتمد علي نقل المعالجة و مساحة التخزين الخاصة بالحاسوب إلي ما يسمي السحابية و هي جهاز خادم يتم الوصول إليه عن طريق الانترنت لتحول برامج تكنولوجيا المعلومات من منتجات إلي خدمات .

تمهيد :-

إن الثورة الصناعية الرابعة هي أحد أهم متغيرات القرن الحالي بما فيه من تطورات في عالم التكنولوجيا و انترنت الأشياء و غيرها من محدثات العصر و لذلك أصبح من

الضروري عقد مؤتمرات و ندوات و كافة حلقات العلم و النقاش حول هذا المفهوم الجديد حتي يصدر العديد من المؤلفات و التي منها يستفيد طائفة كبيرة من المتخصصين و من هنا وجدت الباحثة أن أدوات هذه الثورة هو أحد العوامل الهامة و التي تؤثر علي أداء المؤسسات و الأفراد في ظل البيانات الضخمة التي تحتاج إلي تحليل و الوصول من خلالها لنتائج هامة لأصحابها و ستعتمد الباحثة علي أهم و أشهر ثلاث أدوات و هم الهادوب و ساب هنا و نو اس كيو إل و ذلك من خلال البحث في الإنتاج الفكري العربي و الأجنبي.

أولاً : البيانات الضخمة

1/1 المفهوم :-

هي مجموعة من البيانات سواء غير مهيكلة و التي تتمثل في الرسائل النصية و البريد و الفيديوهات و الشبه مهيكلة و التي تتمثل في ملفات الـ xml و لها معمارية معينة و البيانات المهيكلة و التي تتمثل في قواعد البيانات و بالرغم من ذلك نجد العديد من البرامج التي تعالج جميع أنواع البيانات الضخمة[1] و التي سنتحدث عنها في السطور القادمة.

قبل الحديث عن البرمجيات المختلفة المخصصة لتحليل البيانات الضخمة يجب الاشارة إلي

1/2خصائص البيانات الضخمة :-

هي عبارة عن 3V و التي يوضحها شكل[2] (1) :

[1] Altahat,Khalid .suliman, taha.2018.moving towards the cloud.
المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاحmandumah.com/870100

[2] الأكلبي،على بن ديب.2015.أهمية تحليل البيانات الضخمة في اتخاذ القرار في جامعة الملك سعود.. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870185

شكل (1) خصائص البيانات الضخمة

بالنظر إلي الشكل رقم (1) السابق نجد أن :-

بناء علي ما يعرضه الشكل السابق و هو معروض في شكل فن ثلاث كرات متقاطعة نجد أن تتوافر عوامل رئيسية كي تتكون عناصر البيانات الضخمة و هي:-

1. الحجم: وهو عدد التيراباتيت من البيانات التي نطلقها يومياً من المحتوى.
2. التنوع: وهو تنوع هذه البيانات ما بين مهيكلة وغير مهيكلة ونصف مهيكلة
3. السرعة: مدى سرعة تواتر حدوث البيانات، مثلاً تختلف سرعة نشر التغريدات عن سرعة مسح أجهزة الاستشعار عن بعد لتغييرات المناخ.

و في السطور القادمة سيتم توضح مدي استفادة المؤسسات من البيانات الضخمة من خلال أدوات و تقنيات تحليل البيانات ومنها Hadoop, saphana, nosql .

ثانياً: الهادوب HADOOP

هو أحد برامج ال Open Source مصمم للتعامل مع البيانات الضخمه (Big data) يستخدم لغة الجافا و عن طريقه تستطيع أن تحول البيانات الضخمة إلى إحصائيات لتبني عليها قرارات.

كما أن آليه العمل في برنامج Hadoop مبنية على توزيع البيانات الضخمة والعمليات المختلفة على مجموعة من الأجهزة المرتبطة ببعضها بإحكام شديد . والبرنامج مصمم بحيث اذا حدث أي مشكلة في أحد أجزاء ال Hardware لا يمنع البرنامج من العمل ولكن سيتم تجاوز المشكلة يستمر البرنامج في العمل[3] .

2/1 نشأة الهادوب HADOOP:-

و تعود نشأة هادوب إلي بداية جوجل لاختراع برنامج لحل مشكلة البيانات الضخمة ثم استفاد ياهو بما نشره جوجل عام 2004 تحت عنوان the google file system لإمكانية قراءة الملفات الموزعة بكفاءة عالية ثم نشر في نفس العام ورقة أخري لحل الجزء الثاني من المشكلة و هي معالجة البيانات الضخمة في بيئة معالجة متوازية و علي بيانات موزعة تحت عنوان :

Mapreduce:simplified data processing on large clusters

و بناء علي هذين الورقتين استفاد دوق كتنDoug cutting و من ثم بني نظاماً ناجحاً إلي أن تم تعيينه في ياهو و قام بتطوير النظام بالاعتماد علي فريق العمل و سمي هذا النظام هادوب و قررت ياهو أن تجعله مفتوح المصدر و كانت النسخة الأولي عام 2011 Hadoop 1.0.0 و كان مكون من جزئين هما[4]:

نظام الملفات الموزع ، و إطار عمل برمجة موزعة اسمه mapreduce

إلي أن ظهرت النسخة الثانية عام 2012 hadoop2.0.0 و تم إضافة جزء ثالث و هو YARN ثم صدرت النسخة الكاملة عام 2017 hadoop3.0.0 إلي ان توالت النسخ hadoop 3.2.0 لعام 2019 .

2/2مكونات الهادوب:

و يتكون برنامج هادوب من أربع أجزاء رئيسية [5]وهي :

* Hadoop Common
* Hadoop Distributed File System (HDFS)
* MapReduce

[3] Lakhe, bhushon.practical hadoop security:friends of apress.

[4]//: itwadi.com/Apache_Hadoop

[5] Leetaru,kalew h.2018.the user of future bigdata and remaginig how we think about information. : المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870072

* Yet Another Resource Negotiator (YARN)

و نظراً لأهمية هذه الأجزاء فسوف نتناول وظيفة كل جزء كالآتي:-

1 - Hadoop Common : ويعد مسئول عن الخدمات اللى بتدعم باقي مكونات البرنامج .

2 - Hadoop Distributed File System : يضع كل الداتا علي سيرفرات ويقسم الداتا علي بلوكات و تتوزع علي النودز و اللي بتنقسم إلي Name Node, data node و تعد الـــname node هي أساس العمل داخل الـــcluster لانها تعلم وجود جميع النودز أو الـــblocks .

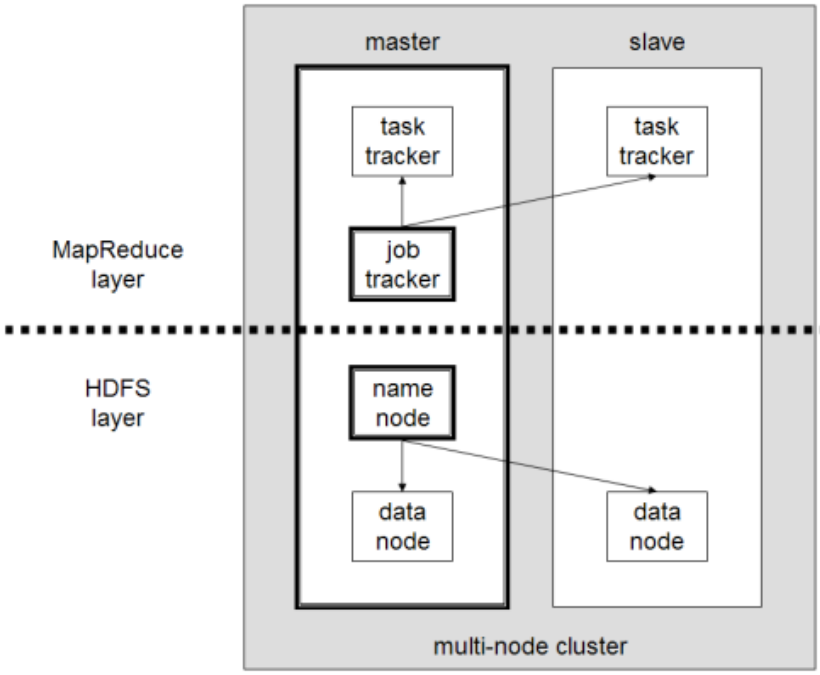
3 - MapReduce : يجعل كل النودز تكلم بعضها و يوزع المهام فمثلا عند معالجة نود يسأل الـــname node ثم يوزع الـــ job علي كل النودز ثم تظهر النتيجة في حوالي دقيقتان و نصف و بذلك يعمل MapReduce علي تقليل الوقت و توزيع المهام و سؤال الـــname node

الإطار الذي يتم فيه كتابة البرامج التي تقدر أن تتعامل مع البيانات الضخمة سواء كانت منظمة او لا والتي يتم توزيعها على الأجهزة المختلفة بالتوازي .كما هو موضح بالجدول[6] رقم (1) التالي:-

Job Name	اسم المهمة
Executable (Jar) Class	الملف التنفيذي الرئيسي – هنا الفئة WordCount
Mapper Class	الفئة التي تحتوي الدالة map وهي TokenizerMapper
Combiner	الفئة التي تحتوي على الدالة reduce وهي IntSumReducer
Reducer	الفئة التي تحتوي على الدالة reduce وهي IntSumReducer
Output Key	نوع المخرج (المفتاح) وهو هنا Text
Output Value	نوع المخرج (القيمة) وهي هنا IntWritable
File Input Path	تحديد مكان المدخل هنا (Path(args[0] (سيتم تمرير المسار كمعطى للدالة main)
File Output Path	تحديد مكان المخرج هنا (Path(args[1] (سيتم تمرير المسار كمعطى للدالة main)

4 - YARN : هي ببساطة Technology لإدارة الأجهزة اللى بيتم توزيع البيانات خلالها وتعتبر الجيل القادم من ال Mapreduce

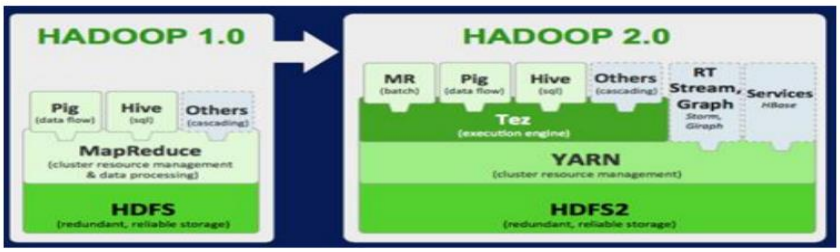
[6] Wadkar, Sameer,…etal.pro apache hadoop.friend of apress.



شكل(3) [7]كيفية عمل الـهادوب

و تتعدد الاصدارات لبرنامج هادوب كما هو موضح بالملحق رقم (1)

و الشكل (4) التالي يوضح الفرق بين أول نسختين للنظام:-



بالتعليق علي الشكل 4 السابق يتضح العديد من التعديلات و التطورات علي نسخة هادوب منها إضافة YARN و RT Stream Graph و Services و هو توافر

[7] https:// https://www.facebook.com/OpenitEG/photos/a.218684984974491.1073741828.202314549944868/463163927193261/?type=1&theater

رسومات و خدمات أخري بالإضافة إلي وجود جميع المكونات الأخري مثل المابريديوس و نظام إدارة الملفات.

2/3 مميزات الهادوب:

يتميز الهادوب بالخصائص التالية :-

1 - Scalable : تتم آلية العمل من خلال مجموعة أجهزة مرتبطة ببعضها . بحيث تمكن المستفيد من القدرة علي تزويد عدد الأجهزة دي أو الموارد الموجودة بدون الحاجة إلي تغيير النظام .

2 - Flexible :له القدرة علي التعامل مع أي نوع بيانات سواء كانت منظمة أو غير منظمة ويستطيع استقبال البيانات من أي مصدر ويستطيع تحليلها بشكل أفضل .

3 - Cost effective : و ذلك بالاستغلال الأمثل لمكونات الجهاز بحيث إن العمليات داخل البرنامج بتتم على التوازي لتحقيق أقصى استفادة ممكنه من المكونات وتقليل التكلفة .

4 - Fault tolerant : في حالة حدوث أي مشكلة لا يمنع البرنامج من اتمام العمل ولكنه بينتقل لمكان تاني لمعالجة البيانات الموجوده فيه بمعني أخر اكثر وضوحاً و هو يستغل البديل الاحتياطي من النودز للحماية من الأعطال و للعلم أن كل نود يكرر 3 مرات كبديل احتياطي و في حالة حدوث عطل أو خطأ أقدر أعمل نود جديد أو أوزع علي البلوكات كل النودز بحيث يوجد 3 نسخ في الـ cluster .

5- single write, multiple read : معلومات ثابتة موجودة بحجم ضخم مثل ما يوجد من بيانات في وكالة الفضاء ناسا و تقوم بمعالجة البيانات الثابتة و المستقرة .

6 – PiG, Hive : بالنسبة للـــ PiG هي خاصة بالسيكوال سيرفر بمعني أنها خاصة ببيانات قاعدة البيانات و تكتب بأى لغة برمجة حتي و إن اختلفت مع لغة الـMapreduce و لكن يفضل لغة الـــ Java كأفضل لغه مع الـــ MapReduce .

و بالنسبة للـــ Hive عبارة عن Project جاهزة بتستخدم علشان تنفذ مهام معينة و تساعد علي تحليل البيانات.[8]

2/4إمكانات الهادوب:-

تتمثل إمكانات الهادوب في السبعة عناصر التالية[9] :-

[8] //: itwadi.com/Apache_Hadoop

[9] https://www.sap.com/dam/application/shared/images/charts-diagrams/sap-hana-product-structure.png.adapt.800_-1.false.false.true.false.png

1. نمذجة المخاطر Risk Modeling
2. Customer churn Analysis
3. Recommendation engine
4. Ad targeting
5. Transactional Analysis
6. Threat analysis
7. Search Quality

ثانيا:ساب هانا sap Hana (in-memory)

3/1 المفهوم:

أداة لتحليل البيانات الضخمة يساعد مستخدميه في القيام بعمليات التحليل المتزامن للبيانات الضخمة و التي ترد إلي منصة النظام مما يساعد علي سرعة اتخاذ القرار .و بذلك فإنها تضع تحت السيطرة قوة البيانات و يسرع في ابتكار النتائج الموثوق بها بواسطة تطوير الذكاء و الحلول المباشرة في الوقت المناسب، و بما أن التقارير و أسلوب العمل مبني علي نسخة من البيانات الفردية فإن البرنامج سيدعم تعاملات الجيل القادم من المعالجة و التحليل و العمل علي إضافة مجموعة من التحليلات المتقدمة من خلال السحابية في البيانات المتعددة و المزدوجة[10]. و سيتم توضيح بعض المصطلحات الهامة كالآتي:-

CLOUD : السحابية، الاستدلالية ونشر البيانات المختلطة.

SECURE : الأمن = المستقبل و هو بيانات مجهزة داخل الذاكرة .

Live : الذكاء المباشر مع سرعة معالجة البيانات .

[10] Mira, Hasan .Sap hana (2019).available at :

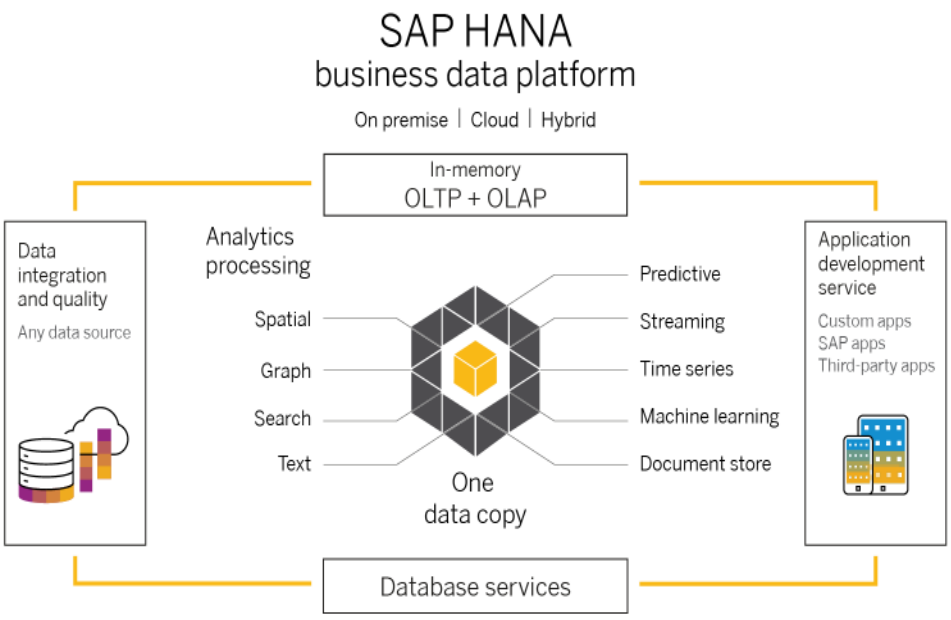
https://www.youtube.com/watch?v=OFEwPwdIKW0

Broadest advanced analytics : تحليل متقدم كامل مع توحيد معالجة الـOLAP(إجراء نظام المعالجة علي الخط المباشر) OLTP, (نظام معالجة التحليلات علي الخط المباشر) من خلال نسخ البيانات الفردية .

2/3 أجزاء الـــSAP-HANA[11] :-

- SAPHANA DB = تشير إلي قواعد البيانات
- SAPHANA APPLIANCE= تشير إلي الجمع بين قواعد البيانات و الاستديو و أي أدوات أخري ترسل كـ جهاز.
- SAPHANA STUDIO= تشير إلي نمذجة الأدوات .
- SAPHANA APPLICATION CLOUD= التطبيق السحابي : يشير إلي البنية التحتية للقواعد السحابية للتوصيل من الجهاز .

و الشكل (5)[12] التالي يوضح مكونات و طريقة عمل البرنامج :-



شكل (5) كيفية العمل برمجيات الـ sap hana

[11] نفس المرجع السابق.

[12] Sap hana (2019). available at:

https://www.youtube.com/watch?v=7RSSy7Azvr0

3/3 مبادئ عمل الـــSAPHANA [13]:-

1. In-memory computing : لأن لديها معالج و عدد كبير من الـــRAM بدل الهارد ديسك.
2. Column Based Storing : بدل الأعمدة تستبدل بالملفات لعملية الاسترجاع .
3. Parallel Processing: يدعم المعالجة المتوازية بحيث تقطع الكود إلي أجزاء و توزعها علي الـCPU المتعددة و بذلك تسرع عملية المعالجة .

4/3 مميزات استخدام الـــ SAPHANA :

1. تقليل التعقيدات.
2. العمل في أي مكان .
3. نتائج حقيقية .

4/3/1: تقليل التعقيدات :-

ببساطة أنه يعمل من خلال منصة تطبيقات نقل التحليلات باستخدام الـــ SAPHANA لتحليل البيانات المباشرة لدعم العمليات التجارية في الوقت المناسب بالإضافة إلي تقليل فائض البيانات، التغطية الكاملة، الهاردوير،و العمليات المستخدمة.

4/3/2: العمل في أي مكان :-

يفعل الوظائف سواء بشكل عام أو باستخدام تقنيات السحابية بشكل خاص (مركز البيانات البسيط أوترتب الأجزاء أو العناصر لأكثر من 1000 شهادة من 13 موزع .

4/3/3: نتائج حقيقية :-

الوصول إلي أفضل نتائج تجارية و زيادة نسبة الابتكار و تقليل تكلفة إدارة البيانات [14].

[13] Saphana(2019). Available at :http://hana.sap.com/abouthana.html

[14] https://www.sap.com/dam/application/shared/images/charts-diagrams/sap-hana-product-structure.png.adapt.800_-1.false.false.true.false.png

5/3 الامكانيات Capabilities :

- Transform database management : تغير إدارة قواعد البيانات وذلك ببساطة يتم باستخدام العمليات من خلال نسخ فردية من مشروعات البيانات و منصة حديثة لأمن البيانات .
- Leverage advanced analytics :توفير السيولة المالية باستخدام التحليل المتقدم : توفير السيولة المالية بأداة لمعالجة البيانات المتقدمة للبيانات النصية، الصور الجرافيكية، و البيانات الهامة.
- Process data in-memory : معالجة البيانات داخل الذاكرة : معالجة التعاملات و تحليل البيانات بالإضافة إلي وجود البيانات المستخدمة المتكاملة المكررة داخل الذاكرة .
- Develop next generation applications: تطوير الجيل القادم : الاجراءات المزدوجة لتوفير السيولة المالية و القدرة علي معالجة التحليلات لبناء تطبيقات الجيل القادم[15] .

6/3 الاصدارات:-

يوضح الجدول رقم (2)[16] الاصدارات المختلفة لبرمجيات الـsap :-

وجه المقارنة	Saphana enterprise edition	Saphana standard edition	Saphana express edition	Saphana runtime edition
1-database service				
-Columnar oltp&olap	√[1]	√	√	√
-columnar RDBMS	√	√	√	√
-online analytical processing (olap)	√	√	√	√
-online transactional processing (oltp)	√	√	√	√
-multicore and parallelization	√	√	√	√
-advanced	√	√	√	√

[15] نفس المرجع السابق

[16] https://www.sap.com/mena/products/hana.html

Saphana runtime edition	Saphana express edition	Saphana standard edition	Saphana enterprise edition	وجه المقارنة
				compression
√	√	√	√	
√	√	√	√	-multitenancy
[3]√√	-[2]	√	√	-multitier storage
√√	√	√	√	-data modeling
√√	√	√	√	-openness (standard interface)
√	√	√	√	-administration and security
√	-	√	√	-high availability and disaster recovery
				2- application services
√√	√	√	√	-webserver
√√	√	√	√	-Java script
√√	√	√	√	-sap fioriux
√	√	√	√	-graphic modeler
√√	√	√	√	-application lifecycle management
				3-processing services
√√	√	[4]0	√	-spatial
√√	√	0	√	-graph
√√	√	0	√	-predictive
√√	√	0	√	-search
√√	√	0	√	-text analytics
-	√	0	√	-streaming analytics
√√	√	√	√	-series data
√√	√	√	√	-business functions
				4- integration and quality services
√√	√	√	√	-data virtualization
√√	-	0	√	-etland replication
√√	-	0	√	-data quality

وجه المقارنة	Saphana enterprise edition	Saphana standard edition	Saphana express edition	Saphana runtime edition
-hadoop and spart integration	√	0	√	√√
-remote data synch	√	0	√	-

نستنتج من الجدول رقم (2) السابق ما يلي:-
وجود أربع قطاعات تندرج منهم العديد من المحاور و التي تميز بين الاصدارات المختلفة و هما:-

- خدمات قاعدة البيانات.
- خدمات التطبيقات.
- خدمات المعالجة .
- التكامل و جودة الخدمة.

نجد أن الاصدار الأول من الـsap يشتمل علي جميع المحاور بنسبة 100%مما يعطيه ميزة عن باقي الاصدارات أما بالنسبة للإصدارة الثانية تتوافر فيها الاصدارات بنسبة 100% أيضاً و لكن مع وجود 10 محاور بشكل اختياري و نجد أن الاصدار الثالث هو الأقل في عدد المحاور لتصل النسبة إلي 87,096 % و الاصدار الرابع يشكل نسبة 93,54% مع وجود بعض المحاور صُممت خصيصاً في برمجيات الـsap فقط.

رابعاً: تكنولوجيا (NOSQL (NOT ONLY SQL :

1/4 المفهوم:

ظهر هذا التطبيق كبديل للنموذج العلائقي لتكون لها القدرة علي التعامل مع البيانات الضخمة و طبيعتها من حيث التنوع و الاختلاف و يمكن تعريفها بأنها : قواعد بيانات تتيح تخزين و استرجاع البيانات بعد نمذجتها بتقنيات تختلف عن تلك المستخدمة في نمذجة البيانات في قواعد البيانات العلائقيه و تستخدم هذه التكنولوجيا لما تقدمه من بساطه في التصميم و قابلة للتوسع و لذلك استخدمت في تطبيقات البيانات الضخمة و تطبيقات الزمن الحقيقي للمساعدة في التخلص من التكلفة و البطء الشديد لقواعد البيانات العلائقية[17].

[17] A Comparison Of NoSQL Database Management Systems And Models(2019) available

2/4 الفرق بين SQL, NOSQL:-

فيما يلي جدول رقم (3)[18] لتوضيح الفرق بين الـــ SQL و NOSQL

أوجه المقارنة	SQL	NOSQL
هيكلـــة و نـــوع البيانـات المحـتفظ بها	تتطلـب هيكلـة ذات خصـائص محـددة للحفـاظ علي البيانات	تسمح بعمليات انسياب حر free- flaw) (operation
الاستعلام	تسـتخدم كـل قواعـد البيانـات العلائقيـة معيـار SQL إلي حد ما، و لهذا يمكن الاستعلام فيها بلغـة SQL Structured Query Language	لا تسـتخدم طريقـة محـددة للعمـل علـي البيانـات التـي تديرها
التحجيم	عمودي (أي بزيادة موارد النظام)	أفقـي (عـن طريـق انشـاء شبكة عنقودية cluster من أجهزة متعددة)
المتانـــــــــــة Reliability	تتمتع بمتانة أفضل من أي قاعدة بيانات	تتمتع بالمتانة

نستنتج من الجدول رقم (3) السابق ما يلي :-

أن برمجيات nosql هـي الأكثـر مرونـة و تطورهـا مـن لغـة sql فهـي متفوقـة فـي المساحة و الاستعلام و ايضاً في هيكلتها و نوع البيانات المستخدمة بها بحيث تتعامل بشكل اكبر مع جميع أنواع البيانات و امكانيات تمديد المساحة لانها أفقية و ليست عمودية و لذلك فإن استخدام nosql يعتبر الأفضل من حيث الامكانيات التي تواكب التطورات التكنولوجية .

3/4 أنظمة قواعد البيانات NoSQL:

1. Key-Value stores

تدل علي وجود علاقة بين المفتاح و القيمة كما هو الحال مع وجود هاش (Hash Map) و قد تتمثل القيمة في سلسلة عناصر Serialized Object مثلاً و غياب الـــData typing في هذا الـــ Model لـه تـأثير مهم جداً علـي الـــ Querying لأن التواصل مع قاعدة البيانات سيقتصر علي ثلاث عمليات فقط : Delete, put, Get

at:https://www.digitalocean.com/community/tutorials/a-comparison-of-nosql-database-management-systems-and-models

[18] نفس المرجع السابق.

2. Column Oriented

يتشابه بشكل كبير مع RDBMS إلا أن عدد أعمدته ديناميكي علي عكس الجداول التقليدية حيث يمكن وجود عدة records بأعمدة مختلفه مما يسمح لك بتفادي وضع Null في الخانات الزائدة.

3. Document-based

يعتمد هذا الـModel علي نموذج Key-Values القيمة في هذه الحالة هي ملف JSON أو XML الهدف من تمثيل كهذا هو إمكانية الحصول علي مجموعة بيانات مرتبه بشكل شجري من خلال مفتاح واحد فقط، نفس العملية تقابلها مجموعة من الـJOINS في الـــRDBMS التقليدية.

4. Graph Data Model

يعتمد هذا الـــModel علي الـــ Graph Theory حيث تحتوي كل عقده (علي الاقل) علي مؤشر يحمل عنوان العقده التالية و بالتالي لا حاجة في استخدام الـــindexes يستخدم هذا النموذج عادة في الشبكات الاجتماعية حيث يسمح بالحصول علي معلومات الأعضاء بسرعة، كما يمكن أيضاً من إيجاد العلاقات التي تربط الأعضاء فيما بينهم بشكل سريع [19].

4/4 مميزات الـNOSQL

1. تقوم بتخزين البيانات في ملفات Documents بدلأمن تخزينها في جداول Tables.
2. لا تتبع تصميم ثابت Schema
3. تدعم التوسع و إتاحة البيانات في جميع الأوقات بشكل ممتاز performance and availability [20].

خامساً:أنظمة أخري لإدارة البيانات الضخمة[21] :

1. Mongo DB:

[19] Corbellini, alejandro2017.persisting big-data:the nosql landscape.date18/12/2018 availanle at : www.elsevier.com/locate/Infosys

[20] Biceuska, zane.oditis,Ivo.2017.towards nosql-based data warehouse solutions date27/07/2018.available at : http://creativecommons.org/licenses/by-nc-nd/4-01

[21] Chandra, dekaganesh.2015.base analysis of nosql database.date18/12/2018available at : http://www.elsevier.com/locate/f9cs

منصة حرة و مفتوحة المصدر تصنف ضمن فئة قواعد البيانات Nosql و تقوم بتخزين و معالجة البيانات .

2. Cassandra :

هي عبارة عن نظام لإدارة قواعد البيانات الموزعة كما أنها مفتوحة المصدر و قد تم إنشاؤها من قبل شركة فيسبوك في أول الأمر.و هي قادرة علي معالجة كمية كبيرة من البيانات الموزعة علي عدة خوادم.

3. Spark :

هي واحدة من أكثر مشاريع Apache نشاطاً و هي منصة حوسبة عنقودية مفتوحة المصدر.

4. Google Fusion Tables :

منصة ذات إمكانيات كبيرة في تحليل البيانات و تجسيدها علي شكل رسوم بيانية و خرائط و لاشك أن خرائط جوجل لها فضل كبير في جعل هذه الأداة واحدة من أهم أدوات تحليل البيانات.

5. Node XL :

هي منصة لتحليل و تجسيد العلاقات و الشبكات و هي حرة و مفتوحة المصدر و تمتلك القدرة علي تجميع البيانات من منصات التواصل الاجتماعي مما يجعلها واحدة من أهم أدوات الإحصاء و تحليل البيانات الكبيرة .

6. Couch DB :

قاعدة بيانات مفتوحة المصدر من فئةNosql قادرة علي تخزين و معالجة البيانات الضخمة في صيغة Json .

7. Apache Hive :

واحدة من أفضل منصات تحليل البيانات الضخمة.

8. Tableau :

هي أداة مجانية تمكن من التأكد من صحة الفرضيات و استكشاف البيانات بشكل سريع .

9. Rapid Miner :

هي منصة تحليل البيانات مفتوحة المصدر تساعد علي تمثيل البيانات و معالجتها إضافة إلي النشر و التقييم و التحليلات التنبؤية. كما توفر بيئة متكاملة للتحليلات الخاصة بالأعمال business analytics و التحليلات التنبؤية و التعليم الآلي Machine learning .

10. Open Refine :

تعرف أيضا باسم Google Refine و هي أداة فعالة للعمل علي البيانات غير المنظمة فهي قادرة علي تنظيف و ترتيب البيانات و تحويلها إلي من صيغة إلي أخري كما أنها قادرة علي تجميع البيانات من مواقع الويب المختلفة .

11. Knime :

تعد واحدة من أهم منصات تحليل البيانات الضخمة و التي تساعد علي اكتشاف الإمكانيات الخفية في البيانات و التنبؤ بخصائص جديدة .

12. Wolfram Alpha :

هو محرك بحث متطور قادر علي تحليل البيانات المتعلقة بموضوع البحث و من ثم تقديم جداول و رسوم بيانية حول الأسعار، أحدث المعاملات، مقارنة الأداء .

13. Solver :

هي أداة مضافة علي برنامج Microsoft office excel إنها أداة معالجة متقدمة تساعد علي تحليل البيانات بشكل سريع .و يوضح جدول رقم (4) [22]التالي مقارنة بين بعض أدوات تحليل البيانات الضخمة :-

الأداة/المنصة	الامكانيات	مواطن الضعف
Tableau Public	- لا يحتاج إلي مهارات برمجية. - يسمح بمشاركة الرسومات البيانية و	- لايقبل بيانات سوي في صورة اكسيل أو TXT .

[22] Hamad , f(…etal).2018.bigdataand analytics strategies in academic libraries:opportunities and challenges,the case of the university of Jordon library. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870281

الأداة/المنصة	الامكانيات	مواطن الضعف
	تحليلاتها علي مواقع الويب و شبكات التواصل الاجتماعي.	حجم البيانات المعالجة محدود.
Open Refine	- تحويل البيانات من صيغة إلي أخري. - تجميع البيانات من مواقع الويب. - تنظيم و ترتيب البيانات غير المنظمة أو الناقصة.	- لايستخدم مع مجموعة بيانات كبيرة
KNME	- لا تحتاج إلي مهارات برمجية و لا إلي كتابة أكواد بالإضافة إلي أنه يدعم مختلف لغات البرمجة.	- امكانيات بسيطة علي تصوير و تجسيد البيانات في شكل رسوم بيانية.
Rapid Miner	- يلائم الأنشطة التجارية و إدارة الأعمال . - توفير بيئة متكاملة لتحليل البيانات و التنبؤ بها .	- يحتاج في تنصيبه إلي أجهزة ذات قدرات كبيرة . - قيود علي حجم البيانات التي يتم معالجتها .
Google Fusion Tables	- القدرة علي معالجة البيانات المختلفة في صيغ مختلفة في نفس الوقت.	- حجم البيانات المعالجة محدود .

يتبين من الجدول (4) السابق أن أغلب أدوات تحليل البيانات محدودة في عملية المعالجة من حيث الكم و لذلك تري الباحثة أنها غير صالحة لمعالجة البيانات الضخمة فيماعدا Hadoop, NoSQL, SapHana و هم من أشهر الادوات و سيتم القاء الضوء علي الفروق فيما بينهم في السطور القليلة القادمة كما ستطرق الباحثة علي التعرف علي علاقة الحوسبة السحابية بهم .

سادساً الحوسبة السحابية :

1/6تعريف الحوسبة :

هي تكنولوجيا متطورة تعتمد علي نقل المعالجة و مساحة التخزين الخاصة بالحاسوب إلي ما يسمي السحابية و هي جهاز خادم يتم الوصول إليه عن طريق الانترنت لتحول برامج تكنولوجيا المعلومات من منتجات إلي خدمات[23] .

كما يوجد تعريف أخر و هو الخدمات التي تتم عبر أجهزة و برامج متصلة بشبكة خوادم تحمل بياناتها في سحابة افتراضية تضمن اتصالها بشكل دائم دون انقطاع مع أجهزة مختلفة كمبيوتر، هواتف ذكية، و غيرها بعد وضع كود خاص لفتح قفل الشبكة و بالتالي يتم الدخول إليها من أي مكان و في أي زمان .

[23] Pentecost,Michael j .2014.bigdata.

و تعرف أيضاً بـ cloud computing بالإنجليزية و يعود ظهور هذا المصطلح إلي الستينيات و ذلك عندما بدأت بعض الشركات بالتفكير بمدي الفائدة التي ستحقق جراء جعل الحوسبة خدمة و أداة مساعدة تقدم إلي الأفراد عبر الانترنت، و يشير هذا المصطلح إلي إمكانية حصول الشركات و الأفراد علي بعض الخدمات كالتخزين و قواعد البيانات و البرمجيات و غيرها من الخدمات المختلفة من خلال شبكة الانترنت و يوجد في العالم العديد من الجهات الرائده في مجال تقديم الخدمات السحابية كـ Google Cloud PlatForm و يوجد خدمة الحوسبة السحابية العديد من أنواع الخدمات التي تقدمها حيث تقدم بعض الشركات خدمة البنية الأساسية أو ما يعرف بـــInfrastructure As aSerivice و اختصارها هو IAAS و تقدم خدمات التخزين عبر الانترنت من خلال هذا النوع من الخدمات السحابية و هو Software as a Service و الذس يشار إليه بـSAAS و يعني هذا النوع من الخدمات السحابية بتقديم خدمات البرمجيات المختلفة عبر الانترنت و يعتبر برنامج Miocrosoft office 365 هو أحد البرمجيات الشائعة التي يمكن الحصول عليها من خلال SAAS إذا ما هو الـ SAAS,IAAS,PAAS؟

يتشارك البرمجيات الثلاث في أنهم نماذج للحوسبة علي أساس الخدمة و فيما يلي شرح كل علي حده:-

أولاً : البنية التحتية كخدمة infrastructure as a service (IAAS)

تقوم المؤسسات بتوفير الأجهزة فقط للمستخدمين، بينما يقوم المستفيدون بتحميل ما يحتاجونه من نظام تشغيل و البرامج المختلفة بالإضافة إلي خدمة S3 من أمازون، وبذلك فتعتمد هذه الطبقة علي توفير العتاد التي تمثل البنية التحتية فقط دون الخوادم و غيرها.

ثانياً : المنصة كخدمة Platform as a service (PAAS)

تقدم المنصة البنية التحتية اللازمة لتطوير التطبيقات و تقاسمها بمعني أن المستفيد يستطيع اختيار نظام التشغيل ثم عمل بعض الإضافات عليخ ليكون جاهزاً لهذا النوع من الحوسبة مثل google App Engine, Windows Azure [24].

ثالثاً :البرمجيات كخدمة Software as a service (SAAS)

[24] البسيوني، بدوية محمد.2015.برمجيات الحوسبة السحابية و دورها في تطوير خدمات المعلومات:دراسة تطبيقية علي مواقع التخزين السحابي.القاهرة:الجمعية المصرية للمكتبات و المعلومات و الأرشيف .مج2،ع3. تاريخ الاتاحة2019/9/20.متاح علي علي http://search.mandumah.com/677920

هي مجموعة من التطبيقات المطوره بواسطة مزود الخدمة لاستخدام البنية التحتية و منصة السحابة، و هنا يتم تسويق برمجيات جاهزة يستطيع المستفيد الوصول إليها غالباً بواسطة "المتصفح" مثل "البريد الالكتروني" أو خدمة محرر النصوص التي تقدمها مايكروسوفت فهي تشمل التطبيقات التي يحتاجها المستخدم مثل تطبيقات مايكروسوفت و تطبيقات جوجل.[25]

و تتعد تطبيقات الحوسبة السحابية و التي ستذكرها الباحثة في السطور القليلة القادمة علي سبيل المثال لا الحصر:-

Sky Drive
Windows live
Microsoft office
Google chrome os
Windows Azure
Pixlr editor
Google docs
Jaycut
Aviary music creator

1. Windows live :خدمة تقدمها شركة مايكروسوفت لجميع المشتركين بها و هي خدمة مجانية و تتوفر مساحة تخزين مجانية تصل إلي 25GB بحيث لأيزيد حجم الملف الواحد عن 50MB كما يمكن رقع 5 ملفات من جهاز المستخدم في نفس الوقت و يمكن من مشاركة الملفات و تخزين مجموعات من الصور في مجلدات يقوم بإنشائها المستخدم و مشاركتها و يزود المستخدم أيضاً بإمكانية الإنشاء و التعديل و التخزين العديد من المستندات الخاصة ببرنامج Microsoft Office مثل مستندات الوورد و الاكسيل و الباوربوينت و السماح للأصدقاء بمشاهدتها و تعديلها، كما يوفر الاشتراك في خدمة RSS الخاصة بالملفات العامة الخاصة بالملفات العامة Public Files التي يقوم مستخدمين أخرين برفعها و ذلك من خلال

[25] Altahat,Khalid .suliman, taha.2018.moving towards the cloud.

المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل

المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة

2019/10/18.متاح علي http://search.mandumah.com/870100

مشـاهدة أيقونـات أو صـور تحـدد اسـم الملـف و نوعـه وروابـط التحميـل الخاصـة بالملفات.[26]

2. Google Chrome OS :

هـو عبـارة عـن نافـذه متصـفح يتـيح للمسـتخدم اسـتخدام كـل التطبيقـات المتاحـة علـي الانترنـت و الموجـودة علـي جوجـل ويـب سـتور Google Web Store الـذي اثبـت نجاحـا كبيـراً فـي تقـديم تطبيقـات و بـرامج تغطـي كـل الوظـائف التـي تقـوم بهـا أنظمـة التشغيل العادية بداية من كتابة رسائل التذكير و حتي إنتاج الأفلام.

3. Windows Azure :

هو نظام يستخدم اساليب في تطوير التطبيقات الخاصة به و الذي يقوم المطور بانشاء برنامج أو تطبيق ليعمل علي Windows Azure فكل ما يحتاجه هو استخدام نفس التقنيات التي استخدمها المطوروين في إنشاء برامجهم مثل C++, JAVA, PHP, Microsoft SQL و العديد من التقنيات المألوفهلدي معظم المطورين بالإضافة إلي أنهيعتمد علي تقليل الوقت المستخدم.

4. Pixlr editor :

هو محرر صور يعمل باحترافية عالية لتصميم و تحرير الصور علي الانترنت دون تنصيب أو تحميل أو شراء أو اشتراك لأنه مجاني و متاح لكل المستخدمين علي جميع المتصفحات و أنظمة التشغيل و لايحتاج لأى مشغل فلاش Flash Player لكي يعمل علي المتصفح و يمكن للمستخدم أن يقوم بتسجيل حساب و يقوم بتخزين صوره علي حساب يمكنه من تحريرها في أي وقت، تتساوي إمكانيات البرنامج مع Adobe Photo Shop و لكنه مجاني و في أي مكان .[27]

---

[26] خدمات الحوسبة (2019). متاح علي : https://almaarik.wordpress.com/%D8%A7%D9%84%D8%AD%D9%88%D8%B3%D8%A8%D8%A9-%D8%A7%D9%84%D8%B3%D8%AD%D8%A7%D8%A8%D9%8A%D8%A9-%D9%81%D9%8A-%D8%A7%D9%84%D8%AA%D8%B9%D9%84%D9%8A%D9%85-cloud-computing-in-education/

[27] البيشي، ساره.2018.معالجة البيانات الضخمة بالتخزين السحابي في المكتبات و مؤسسات المعلومات الجامعية. المكتبة المركزية بجامعة الأميرة نورهبنت عبد الرحمن نموذجاً. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870344

5. Google Docs :

هو عبارة عن مجموعة مكتبية متاحة مجاناً من جوجل لجميع مستخدميها و هي تتضمن الخدمات داخل Gmail حيث يمكن لكل المستخدمين من مشاهدة مرفقاتهم علي الانترنت دون الحاجة إلي تحميلها و هذه المجموعة تتضمن برامج المكتب المهمة مثل:

للكتابة، عمل عروض تقديمية، عمل جداول حسابات المخططات، الاستبانات، و كل ذلك بشكل متزامن مع حساب جوجل و بإتاحة صلاحيات مختلفة لكل مستند و مشاركته مع الأخرين.[28]

6. Jaycut:

هو برنامج يمكن المستخدم من إنشاء أفلام فيديو دون تحميل و عناء فهو يعمل علي الانترنت و لا يحتاج إلي مواصفات خاصة ليعمل و لأيحتاج لمتصفح و لا مشغل فلاش و هو محمل بالعديد من الوظائف في تحرير و إنتاج الأفلام منها التسجيل من كاميرا الويب و تصدير الأفلام مباشرة إلي يوتيوب و هو متاح بالعديد من اللغات لسهولة الاستخدام [29].

7. Aviary Music Creator :

هو برنامج يمكنك من إنتاج مقاطع الموسيقي من البداية حتي تصنع مقطوعات كاملة فمن خلاله يمكن تقطيع الأصوات و إضافة مؤثرات و إضافة مؤثرات و إضافة آلات موسيقية و تحديد فترات لعمل كل آلة موسيقية علي حده و يعمل باحترافية خاصة فب عالم برامج إنتاج الموسيقي و هو نموذج ناجح لبرامج الوسائط المتعددة و التي تعمل علي السحاب علي شبكة الانترنت .[30]

---

[28] عبد الحسيب،جمال رجب.موسي،أحمد محمد.2017.تصور مقترح للاستفادة من خدمات الحوسبة السحابية بالجامعات المصرية في ضوء التوجه نحو مجتمع المعرفة الرقمي.بنها:جامعة بنها-كلية التربية.مج28،ع111.تاريخ الاتاحة 2019/05/14.متاح علي:
http://search.mandumah.com/861653

[29] عمر،معاوية مصطفي محمد.2014.دور الحوسبة السحابية في تعزيز احتياجات مجتمع المعرفة الرقمي العربي من المعلومات .السودان:جامعة الخرطوم.ع33.تاريخ الاتاحة 2018/06/06.متاح علي: http://search.mandumah.com/642594

[30] الزهري، سعد بن سعيد.2018.الحوسبة السحابية و استثمارها المستقبلي في المكتبات العامة السعودية:رؤية استشرافية. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870093

2/6مميزات الحوسبة:-

1. الدخول إلي البيانات من أي مكان تتوافر فيه خدمة الانترنت .
2. قلة التكلفة في التجهيزات المادية .
3. ستغني المستخدم سواء فرد أو شركة من شراء تراخيص البرمجيات.
4. لا تحتاج المؤسسات أو الشركات إلي تخصيص مكان للأجهزة التي تدير العمل .
5. عدم الحاجة إلي الدعم الفني داخل المنشأة.
6. يمكن الاعتماد عليها في الابحاث الكبيرة و المعقدة مما يوفر الوقت.
7. الحفظ و النسخ الاحتياطي يخدم المستخدم و خاصة المبرمجين .
8. خاصية المشاركة للملفات و التي تخفض المساحة التخزينية .
9. التحجيم أو ربط التكلفة باستخدام Scalability.[31]

3/6 تحديات الحوسبة السحابية :-

1. الحاجة إلي مساحة تخزينية كبيرة.
2. الحفاظ علي أمن و سرية البيانات الخاصة بالعملاء .
3. توافر التطبيقات المناسبة لجميع المؤسسات و جدواها الاقتصادية .
4. توافر التغطية الشاملة للوصول السريع للانترنت .
5. اقناع المسؤلين و أصحاب القرار بجدوي الانتقال للحوسبة السحابية .[32]

من خلال العرض السابق للأدوات المختلفة ومدي جدواها و مكوناتها و امكانياتها ننتقل للجزء الأخير من البحث و هو توضيح الفروق بين الأدوات السابق عرضها من خلال الجدول رقم (5) التالي :-

وجه المقارنة	Hadoop	Saphana	Nosql
لا تتبع تصميم ثابت	-	-	√
المرونة	√	√	√

[31] مميزات الحوسبة (2019) . متاح علي :
https://mawdoo3.com/%D9%85%D8%A7_%D9%87%D9%8A_%D8%A7%D9%84%D8%AD%D9%88%D8%B3%D8%A8%D8%A9_%D8%A7%D9%84%D8%B3%D8%AD%D8%A7%D8%A8%D9%8A%D8%A9

[32] محمود، ممدوح علي.2016.التخزين السحابي للبيانات و أمن المعلومات: دراسة تقييمية.القاهرة:الجمعية المصرية للمكتبات و المعلومات و الأرشيف.مج3.ع4.تاريخ الاتاحة 2019/02/12.متاح علي:http://search.mandumah.com/806203

وجه المقارنة	Hadoop	Saphana	Nosql
التوسع	√	√	√
تكلفة اقل	√	√	√
تصحيح المشاكل تلقائياً	√	√	√
معالجة البيانات الثابتة	√	√	√
سرعة معالجة البيانات	√	√	√
Olap	-	√	-
Oltp	-	√	-
المعالجة داخل الذاكرة	-	√	-
الاسترجاع بالملفات	-	√	√
المعالجة المتوازية	√	√	√
مفتوح المصدر	√	√	√
التعامل مع البيانات الضخمة	√	√	√
المجموع	10	13	11
النسبة %	%71.42	%92.85	%78.57

نستنتج من الجدول رقم(5) السابق ما يلي:-

تفوق برمجيات الـsap hana علي منافسيها بمعدل ممتاز و ذلك لتوافر جميع أنواع المعالجة و التحليلات علي الخط المباشر مع السرعة في معالجة البيانات الضخمة بالرغم من حداثته مقارنة بالهادوب بل إن Nosql لدية ميزة إضافية عن الهادوب و هي الاسترجاع بالملفات، و لكن بصورة عامة يعد الهادوب هو الكثر شعبية و الأسهل استخداماً و مألوفاً لدي المستخدم مما أعطاه شعبية كبيرة و أخيراً فتعد البرمجيات الثلاث لديهم من الجودة ما يتوافق مع احتياجات المؤسسات بشتي أنواعها.

بعد استعراض و الاجابة علي تساؤلات الدراسة يتبقي لدينا التساؤل الأخير و هو الفرق بين أدوات تحليل البيانات الضخمة و الحوسبة السحابية فنجد إختلاف كبير بينها و بين أدوات تحليل البيانات لأنها لا تحتاج إلي خبرة في التعامل و لأنها لا تستخدم في تحليل البيانات و ليس منها إصدارات.

من هنا نستنتج أن أدوات تحليل البيانات الضخمة يمكن أن تحل محل الحوسبة السحابية و لكن لا يمكن للحوسبة السحابية أن تحل محل أدوات تحليل البيانات الضخمة .

نتائج الدراسة :-

تتمثل نتائج الدراسة في النقاط التالية :-

- أن المؤسسات التر تمتلك بيانات ضخمة تواجه تحدياً كبير لتحليل بياناتها بأشكالها المختلفة سواء المهيكلة أو غير المهيكلة.
- أن أفضل أدوات تحليل البيانات الضخمة هم Hadoop-Nosql-saphana .
- تتعدد إصدارات نسخ الهادوب لتصل إلي 61 إصدارة لـ 3 نسخ مما يدل علي التحديث المستمر لإمكانيات النظام .
- تتعدد إصدارات الـSaphana مع اختلاف الامكانيات المتاحة فيما بينهم فالإصدارة الأولي Saphana enterprise edition تشتمل علي جميع الامكانيات المتاحة بنسبة 100% بينما تتراوح باقي الاصدارات ما بين 87% إلي 93% و هي نسبة جيدة جدا أيضاً.
- تتعدد الفروق بين تكنولوجيا الـNosql, sql فنجد تفوق الثانية في المرونة و المساحة و الاستعلام و الهيكلة و نوع البيانات المستخدمة
- توجد فروق طفيفة بين الأنظمة الغير شهيرة في معالجة البيانات و لكن علي الرغم من أنها مصممة لمعالجة البيانات الضخمة إلا أن أغلبها محدود في معالجة البيانات من ناحية الكم.
- تتعدد الفروق بين خدمة الـ SAAS, IAAS,PAAS المستخدمة في الحوسبة السحابية فالأولي (SAAS) هي الأكثر ميزة لتوفير البرمجيات كخدمة للمستفيد.
- تفوق برمجيات الـ saphana علي باقي الأدوات محل الدراسة و ذلك لتوافر جميع أنواع المعالجة و التحليلات علي الخط المباشر مع السرعة في معالجة البيانات الضخمة بالرغم من حداثته بالمقارنة بالـ hadoop .
- تتميز تكنولوجيا الـNosql عن الهادوب في عملية الاسترجاع بالملفات علي الرغم من شعبية و انتشار الهادوب.
- يعتبر الهادوب هو الأسهل في الاستخدام و الأكثر أُلفه لدي المستخدم,
- يتوافر بالأدوات الثلاث الجودة بما يجعلها تتوافق مع احتياجات المؤسسات بشتي أنواعها.
- تعد تطبيقات جوجل السحابية أحد أبرز الطرق لتكوين حوسبة سحابية نشطة و مفتوحة للجميع و يمكن الوصول إليها من أي مكان و في أي وقت علي عكس أدوات تحليل البيانات الضخمة فهي غير مجانية.
- أن كلاً من شركة جوجل و ميكروسوفت تقدم العديد من الخدمات و البرامج في بيئة الحوسبة السحابية مثل البريد الالكتروني المجاني و التراسل وغيرها .
- يعد تصحيح المشاكل بشكل تلقائي من أهم مميزات أدوات تحليل البيانات الضخمة و المتوافرة في أدوات الدراسة.
- أن أدوات تحليل البيانات الضخمة يمكن أن تحل محل الحوسبة السحابية و ليس العكس.

التوصيات:-

توصي الباحثة بما يلي :-

- توفير أدوات لتحليل البيانات الضخمة مجانية و مفتوحة المصدر.
- اهتمام المؤسسات المختلفة بتحليل بياناتها باستخدام أدوات تحليل البيانات المناسبة لها .
- ضرورة توافر خاصية التحليل المتزامن للبيانات الضخمة في أدوات تحليل البيانات الضخمة خاصة في الـ hadoop, nosql .
- ضرورة إنشاء أدوات تحليل بيانات ضخمة عربية تمتاز بالبساطة و الدقة و التوسع و مفتوحة المصدر و متاحة للمستخدم العربي.
- إدخال تطبيقات الحوسبة السحابية في دعم مؤسسات المعلومات و أنظمة المكتبات خاصة المكتبات الصغيرة ضعيفة الميزانية
- ضرورة توافر المعالجة و التحليلات علي الخط المباشر في جميع أدوات تحليل البيانات الضخمة و خاصة hadoop, nosql .

قائمة المراجع و المصادر:-

1. Mira, Hasan .Sap hana (2019).available at : https://www.youtube.com/watch?v=OFEwPwdIKW0
2. A Comparison Of NoSQL Database Management Systems And Models(2019) available at:https://www.digitalocean.com/community/tutorials/a-comparison-of-nosql-database-management-systems-and-models
3. Altahat,Khalid .suliman, taha.2018.moving towards the cloud.

المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870100

4. Biceuska, zane.oditis,Ivo.2017.towards nosql-based data warehouse solutions date27/07/2018.available at : http://creativecommons.org/licenses/by-nc-nd/4-01
5. Chandra, dekaganesh.2015.base analysis of nosql database.date18/12/2018available at : http://www.elsevier.com/locate/f9cs
6. Corbellini, alejandro2017.persisting big-data:the nosql landscape.date18/12/2018 availanle at : www.elsevier.com/locate/Infosys
7. Hamad , f(…etal).2018.bigdataand analytics strategies in academic libraries:opportunities and challenges,the case of the university of Jordon library. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870281
8. Lakhe, bhushon.practical hadoop security:friends of apress.
9. Leetaru,kalew h.2018.the user of future bigdata and remaginig how we think about information. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870072
10. Pentecost,Michael j .2014.bigdata.
11. Sap hana (2019). available at: https://www.youtube.com/watch?v=7RSSy7Azvr0
12. Saphana(2019). Available at :http://hana.sap.com/abouthana.html
13. Wadkar, Sameer,…etal.pro apache hadoop.friend of apress.
14. الأكلبي،على بن ديب.2015.أهمية تحليل البيانات الضخمة في اتخاذ القرار في جامعة الملك سعود.. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870185
15. البسيوني، بدوية محمد.2015.برمجيات الحوسبة السحابية و دورها في تطوير خدمات المعلومات:دراسة تطبيقية علي مواقع التخزين السحابي.القاهرة:الجمعية المصرية للمكتبات و المعلومات و الأرشيف .مج2،ع3. تاريخ الاتاحة2019/9/20.متاح علي http://search.mandumah.com/677920
16. بوعناقة،سعاد.2018.البيانات الضخمة في قطاع المكتبات.نقاط القوة و الضعف،الفرص و التهديدات. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870247
17. فوزي،محمد.2019.البيانات الضخمة .متاح علي : http://www.youtube.com/watch?v=18n8speyyk4

18. سرحان،عماد عمر .(2019). البيانات الضخمة و انترنت الأشياء.متاح علي: https://www.slideshare.net/EmadOmarSarhan/ss-88037740
19. البيشي، ساره.2018.معالجة البيانات الضخمة بالتخزين السحابي في المكتبات و مؤسسات المعلومات الجامعية. المكتبة المركزية بجامعة الأميرة نورهبنت عبد الرحمن نموذجاً. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870344
20. تعريف الحوسبة (2019).متاح علي :https://www.tech-/wd.com/wd/2013/01/21/cloud-computing-ksa
21. الحسيني، فواز ماطر ...وأخ.2018.متطلبات صناعة المعلومات في ظل عصر البيانات الضخمة:دراسة تحليلية لعدد من المختصين و الدارسين في تخصص دراسات المعلومات في دولة الكويت. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870247
22. خدمات الحوسبة (2019). متاح علي :

https://almaarik.wordpress.com/%D8%A7%D9%84%D8%AD%D9%88%D8%B3%D8%A8%D8%A9-%D8%A7%D9%84%D8%B3%D8%AD%D8%A7%D8%A8%D9%8A%D8%A9-%D9%81%D9%8A-%D8%A7%D9%84%D8%AA%D8%B9%D9%84%D9%8A%D9%85-cloud-computing-in-education/

23. خليفة،زينب محمد حسن.2015.الحوسبة السحابية خدماتها و دورها في العملية التعليمية .القاهرة:جامعة عين شمس-كلية التربية-مركز التعليم الجامعيزع31.تاريخ الاتاحة2018/12/12.متاح علي http://search.mandumah.com/741589
24. الزهري، سعد بن سعيد.2018.الحوسبة السحابية و استثمارها المستقبلي في المكتبات العامة السعودية:رؤية استشرافية. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870093
25. السالمي، جمال بن مطر بن يوسف.بني عرابة، سعيد.2018.البيانات الضخمة و دورها في دعم اتخاذ القرار و التخطيط الاستراتيجي دراسة وصفية. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870077
26. السالمي،خلود بنت خالد بن سليمان.الهنوية،هاجر سليمان ناصر.2018.واقع استخدام البيانات الضخمة في هيئة الوثائق و المحفوظات الوطنية :دراسة لستطلاعية. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870085

27. سوهام،بادي.بوخالفه،خديجة.2018.مواكبة برامج أقسام المكتبات و المعلومات في الوطن العربي لمفهوم البيانات الضخمة:دراسة تحليلية. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870112

28. الشوابكه، يونس أحمد إسماعيل.2018.الوعي بمفهوم البيانات الضخمة(bigdata) لدي العاملين في المكتبات الأكاديمية:دراسة حالة لمكتبة الجامعة الأدردنية. . المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870244

29. الشيخ، فطيمه. بوخالفة، خديجة.2018.مواكبة برامج أقسام المكتبات و المعلومات في الوطن العربي لمفهوم البيانات الضخمة:دراسة تحليلية. . المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870112

30. عبد الحسيب،جمال رجب.موسي،أحمد محمد.2017.تصور مقترح للاستفادة من خدمات الحوسبة السحابية بالجامعات المصرية في ضوء التوجه نحو مجتمع المعرفة الرقمي.بنها:جامعة بنها-كلية التربية.مج28،ع111.تاريخ الاتاحة 2019/05/14.متاح علي: http://search.mandumah.com/861653

31. عبد القادر،أمل حسين.2018.إسهام البيانات و المعلومات في التنمية المستدامة:رؤية مستقبلية. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870079

32. العتيبي،آمال بنت سعود.2018.البيانات الضخمة و صناعة المعلومات في المكتبات و مراكز المعلومات : مكتبة الملك فهد الوطنية نموذجاً. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870339

33. عمر،معاوية مصطفي محمد.2014.دور الحوسبة السحابية في تعزيز احتياجات مجتمع المعرفة الرقمي العربي من المعلومات .السودان:جامعة الخرطوم.ع33.تاريخ الاتاحة 2018/06/06.متاح علي: http://search.mandumah.com/642594

34. الفاعوري، أريج.الفاعوري، شذى.2017.تطبيقات الحوسبة السحابية في المكتبات.المؤتمر الدولي الثالث في النشر الالكتروني لمكتبة الجامعة الأردنية:نحو مكتبات حديثة-الجودة و الاعتمادية.تاريخ الاتاحة2019/09/03.متاح علي : http://search.mandumah.com/837307

35. محمود، ممدوح علي.2016.التخزين السحابي للبيانات و أمن المعلومات: دراسة تقييمية.القاهرة:الجمعية المصرية للمكتبات و المعلومات و الأرشيف.مج3.ع4.تاريخ الاتاحة 2019/02/12.متاح علي:http://search.mandumah.com/806203

36. مرسال، محي الدين.2018.دور و أهمية المكتبات الجامعية في إدارة البيانات الضخمة:دراسة حالة المكتبة المركزية لجامعة الخرطوم. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870188

37. مقداني،جدينة.شبيلة، مقدم.2018.دور البيانات الضخمة في دعم التنمية المستدامة بالدول العربية. المؤتمر الرابع و العشرون :البيانات الضخمة و آفاق استثمارها : الطريق نحو التكامل المعرفي.مسقط:جمعية المكتبات المتخصصة فرع الخليج العربي. تاريخ الاتاحة 2019/10/18.متاح علي http://search.mandumah.com/870145

38. مميزات الحوسبة (2019) . متاح علي :
https://mawdoo3.com/%D9%85%D8%A7_%D9%87%D9%8A_%D8%A7%D9%84%D8%AD%D9%88%D8%B3%D8%A8%D8%A9_%D8%A7%D9%84%D8%B3%D8%AD%D8%A7%D8%A8%D9%8A%D8%A9

39. https://
https://www.facebook.com/OpenitEG/photos/a.218684984974491.1073741828.202314549944868/463163927193261/?type=1&theater

40. https://ar.karizma.io/%D8%A3%D9%87%D9%85%D9%8A%D8%A9-%D8%AA%D8%AD%D9%84%D9%8A%D9%84-%D8%A7%D9%84%D8%A8%D9%8A%D8%A7%D9%86%D8%A7%D8%AA-%D8%A7%D9%84%D8%B6%D8%AE%D9%85%D8%A9-%D9%88%D8%A3%D8%A8%D8%B1%D8%B2-%D9%85%D9%86%D8%B5/

41. https://www.sap.com/dam/application/shared/images/charts-diagrams/sap-hana-product-structure.png.adapt.800_-1.false.false.true.false.png

42. //:itwadi.com/Apache_Hadoop

ملحق 1 اصدارات برنامج هادوب

تم إطلاق 3 نسخ و 61 إصدارة من الهادوب موضحة بالجدول التالي:-

الاصدار	تاريخ الاطلاق
3.2.0	2019
3.1.1	8 أغسطس 2018
3.0.1	31 مايو 2018
2.9.1	3 مايو 2018
2.8.4	15 مايو 2018

الاصدار	تاريخ الاطلاق
2.7.7	20يونيو 2018
2.6.5	8 اكتوبر 2016
1.2.1	1أغسطس 2013
0.23.9	8 يونيو 2013
2.0.5	6 يوليو 2013
0.23.8	5 يوليو 2013
1.2.0	13 مايو 2013
2.0.4	25 أبريل 2013
0.23.7	18 أبريل 2013
1.1.2	15 فبراير2013
2.0.3	14فبراير 2013
0.23.6	7 فبراير 2013
1.1.1	1 ديسمبر 2012
0.23.5	28 نوفمبر 2012
0.23.4	15 اكتوبر 2012
1.1.0	13 أكتوبر 2012
1.0.4	12 أكتوبر 2012
2.0.2	9 أكتوبر 2012
0.23.3	17 سبتمبر 2012
2.0.1	26 يوليو 2012
2.0.0	23 مايو 2012
1.0.3	16 مايو 2012
1.0.2	3 أبريل 2012
1.0.1	10 مارس 2012
0.23.1	27 فبراير 2012
1.0.0	27 ديسمبر 2011
0.22.0	10ديسمبر 2011
0.23.0	11 نوفمبر 2011
0.20.205.0	17 اكتوبر 2011

الاصدار	تاريخ الاطلاق
0.20.204.0	5سبتمبر 2011
0.20.203.0	11 مايو 2011
0.21.0	23 أغسطس 2010
0.20.2	26 فبراير 2010
0.20.1	14 سبتمبر 2009
0.19.2	23 يونيو 2009
0.20.0	22 أبريل 2009
0.19.1	24 فبراير 2009
0.18.3	29 يناير 2009
0.18.2	3 نوفمبر 2008
0.18.1	17 سبتمبر 2008
0.18.0	12 أغسطس 2008
0.17.1	23 يوليو 2008
0.17.0	20 مايو 2008
0.16.3	16 أبريل 2008
0.16.2	12 أبريل 2008
0.16.1	13 مارس 2008
0.16.0	7 فبراير 2008
0.15.3	18 يناير 2008
0.15.2	2 يناير 2008
0.15.1	27 نوفمبر 2007
0.14.4	26نوفمبر 2007
0.15.0	29 اكتوبر 2007
0.14.3	19 اكتوبر 2007
0.14.1	4 سبتمبر 2007

abstract

This study boils down to Identify several big data analysis tools And the selection of the most famous and most important of the three tools they are HADOP-SAB HANA-NoSQL To make a comparison and deduce the highest quality and the best And show the percentage in between The results showed a strong convergence between the three tools Despite sabhana's progress over the other two tools However, the three tools still share many advantages And the study referred to some terms that are used in the research Which the reader listens to like cloud computing Some people may be confused by the existence of convergence between them and data analysis tools However, with the review of the term, it was found that there is no comparison between them because they do not participate in any of the indicators.

Inclouded [1]
[2] Not inclouded
[3] Only for sap applications
[4] optional