TEST Engineering & Management

January - February 2020
ISSN: 0193 - 4120 Page No. 6046 - 6050

Traits of Outsourcing Bigdata in V's Term with Safe Secure Cryptosystem

[2]Dr. A. Neela Madheswari,, Supervisor Anna University,Chennai
[2]Dr. R. S. D Wahida Banu ,Supervisor, Anna University,Chennai
[1]R.Menaka ,Research Scholar, Anna University,Chennai,menaka.murugesan@gmail.com

Article Info
Volume 82
Page Number: 6046 - 6050
Publication Issue:
January-February 2020

Article History
Article Received*: 18 May 2019*
Revised*: 14 July 2019*
Accepted*: 22 December 2019*
Publication*: 29 January 2020*

Abstract:
Throughout the advent of the Information Era, dependencies of computing paradigm model must rapport and support the concept called Big data. Big data is collection of dataset of complex structured and unstructured data that cannot be processed or stored by traditional method. Such an accumulation of data that to be processed, accessed, stored and management of data streams arrival from different sector has increased enormously. From the internet, social media and IoT existence, unstructured formatted data is collected and enables the data captivity from any source. For discovering newly emerging product or service through text messages and tweet, many corporations started analysing the big data to pluck the optimised effects. Some organisation work with cloud or utilise the web farm to personalize the corporate product plan for legal ramifications. Web farm in the sense indirectly refer the big data concept. The point in the graph is represented by two dimensions x and y axis and the point in the plane is represented by three dimensions like x,y and z axis . Like this, big data is characterised by many numbers of chained beads as traits. Along with the discussion on 17 specific important beads in big data aspect, its dimension enlargement of data increases the need of security level to protect the data in transit or in session also becomes must

Keywords*: corporate product, two dimensions, computing paradigm model*

I. **Introduction**:

Our new era of the internet community walk around with massive on-the-fly resource usage application. This high tech community [1] prevails with the fast access to large quantities of data and commensurately fast computational resource. Each and every organization have its own local authorities to administrate the resources which may be in the form of software or hardware. It is impractical to execute this immense computational or storage entailing application using resources exist in individual organisation. The internet technolgy grants the facilities to share the data from any location . Now the arriving computing technology grasp the space or power of the interconnected resource as a single entity there by gaining more than terabyte space requirements or the processing power of a supercomputer for a fraction of the cost respectively . The user running an application [2] which is the collection of subnet's resources belong to different domains requires assurance of the machine retaining its integrity, to ensure that proprietary application remains safe.

Any processes in this web are represented in the form of digital. This digitization rise up or enlargement of the web platform with varying social media, blogs, deployment of sensor variant, existence of palmtop devices and wearable devices support out blast of internet usage and data generation by everyone ,everything and everywhere in continuous basis. None can negate the drastic involvement of internet in success of business, government, education and lifestyle of netizens. The internet adoption by kids to senior

January - February 2020
ISSN: 0193 - 4120 Page No. 6046 - 6050

citizens increases the rate of data generation along with more storage space requirement. Assuming the rise might top up to 35 trillion gigabytes at the coming year of 2020. This high rate of data generation with velocity, variety and volume by the support of internet coined the new concept named as Bigdata. Big data almost covers the processing, storing and analysing the prevailing data for predicting the future happening or suggestion for promoting scheme such that to achieve precision score, maintain time sharp and perfection. General intention of Bigdata revolution predict the behaviour of business entity and its personalized scope for marketing goal reachable, comprehending the customer well wishes, preferring low cost treatment to patient, defending the angle deviation spot at military offenses and e-governance possession[3].

II. Data evolution and its scope:

Data is collection of symbols or character in the character set. Information is the collection of data that interpret some meaning. Bigdata is bulky volume of data and information that may not be possible to store and processed by outmoded method within the given time slot. So , data stored beyond the capacity and processed beyond is called as Bigdata.Data architecture and technology starts from OLTP, Online Transaction Processing that uses the Database management system. Then it is improved to use Data Warehouse with OLAP, Online Analytical Processing. Now, Big data directed toward the RTAP, Real-time Analytics.

Our daily life activities like reading e-book, listening to music, surfing, online purchasing and transaction, downloading and uploading movies ,photo, video generates data. It is very hard to find the activities that never generate data. The large voluminous of data that is left as digital trace by electronic media and digital equipment accompanied for social , business, offense ,defence, education and governance of multiple millionaires is called as Big data . This increased volume of data alone cannot suggest this buzzword. Beyond these volume characteristics, the big data have many beads to fulfil its scope.

Beads of 17Vs from Bigdata:

The beads that frames the distributed data dimension space is given below:

- Volume: Elasticity and Scalability of spawning and gathering of data. The size that is greater than multiple form of GB/TB/PB/EB or > large size.
- Velocity: The speed at which data is generated. The rate changes for every earth revolution like 3 millions of email for every second, 20 hours of video uploaded per minutes and 50 millions of tweets every day
- Variety: Different forms of data. Data can be represented by semi-structured, structured unstructured and binary. Interaction nodes vary by person to machine, machine to machine and person to person [4].
- Veracity: Accuracy and Authentication. The literal meaning for veracity is “exactness or keep an aware of dirty data”. The information is collected from massive resources that are declared as public by citizens or corporate is dumped in web by the support of internet. The gathered big data should be cleansed before taken for analysis because it might be relevant or uncertain. Additionally, the fact like confidential and quality of data is need to be assured for further segregation or scrutiny.
- Value: importance of data. The credible data pave a way for cost cutting measure, cross-selling chance, new product idea, cure the problematic disease.
- Validity: Data quality with its life scope. Its strength is how data is essence for intended use. Assurance of till date validation of data, valid for specific sessional transaction or corporate and valid format in accordance with technologies or protocol.

January - February 2020
ISSN: 0193 - 4120 Page No. 6046 - 6050

- Variability: evolution of data technology
- Venue: Data exist in remote, local, heterogeneously or homogenously distributed.
- Vocabulary: Semantics that describe the abstract data type and terminology.
- Vagueness: uncertainty of data needs and its tools.
- Volatility: Duration of usefulness and duration of data stored (like booting and cache content)
- Visualization: picture representation of statistical analysis.
- Virality: broadcasting speed to 'n' user and again to n* n users
- Verbosity (density of noise or redundancy data): Data from the different source might be unfinished, flaws and expired data. Such a collection might cause additional processing time and storage space. Procedure to be initiated to check for trusted and pertinent data at initial stage to avoid time and space consumed in sake of garbage data.
- Voluntariness (make avail willingly): It is dedicated data set that assist or predict the impact factor for enterprises efficiency. Using this trader gets the customer preferences, visualization of ecosystem modelling and traffic patterns, entrepreneurs predict the artefact issues to retain their productivity demand, Health department to get alert of infection and refine the patient health, revolutionize the research science and to ascertain fraud.
- Versatility (used differently for different context):Big data is flexible and adaptable nature like hour glass model that drops the needed resources alone in needed milestone to satisfy the needs of multiple organisations, researchers and Government.

The chain of beads by 17Vs[5] cannot be processed, stored or analysed by traditional technology and mechanism. It needs new tools and techniques. As the dimension of data increases, security enhancement to protect that data in transit or in session becomes must .

The development of information technology shift the users to access different services and perform various activities through web services. The web services provide easy adaptation and accessibility to the users. As the web services are loosely coupled they have more chance of being caught by middlemen and subject to various threats. To improve the security of outsourcing data and services , there are many approaches being discussed, but lags with security performance. To improve the security an session based random elliptic curve are generation is highlighted with cryptographic techniques proposed with Session based ECC Encryption and Decryption Algorithmn(SeBliwECC)is discussed.. The proposed method provide significant computational infeasibility in breaking the keys in session..

Supporting ECC traits:

Elliptic curve cryptography is based on mathematical function applied to coordinates of elliptic curve and binded to elliptic curve discrete logarithmic problem(ECDLP). Security of any systems depends upon the quantities randomness and time complexity. It assures high security due to the lack of known full or sub exponential time algorithm to solve that problem. Implementing ECC with 160 bit offers the equivalent security level as RSA of 1026 bits . So it is best suited for application that enforces security in Integrated Chips of embedded device. Cubic Equation of Elliptic curve use in ECC are is $y^2 \bmod q = x^3 + ax + b$ where a,b,x and y are integers .

January - February 2020
ISSN: 0193 - 4120 Page No. 6046 - 6050

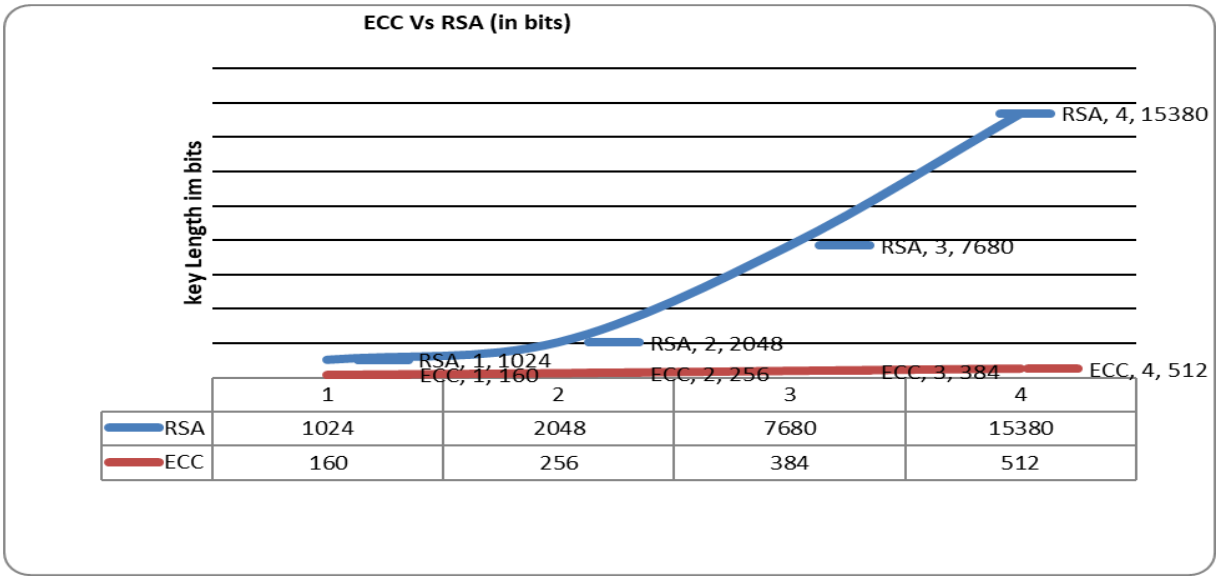

Figure –1. Computation comparison with RSA and ECC.

Due the key size variant Key length of ECC is closely very low while compared with RSA

Elliptic Curve Discrete Logarithm Problem (ECDLP)

Let Ep(a,b) is the elliptic curve over some finite field,p is a prime number,P and Q be a point of order on Ep(a,b) and d be integer with the range 1,n-1. Based on variation in a and b different elliptical curve arises. The strength of ECC depends on ECDLP that gives complex to find the integer d from the given Q=d*P and P,where d*P will be calculated by adding of points P by d times (P + P+ P+…+P). Here the different way of procedure is mentioned to generate the elliptical curve randomly.

Procedure: Elliptic curve Fp generated randomly
Input: Fp, the finite field of prime element p

Let Lp = floor($\log_2(p)$)
Ls= floor (Lp-1/160)
Lh=Lp-160*Ls

Step 1: Choose the random string of 160 bit as SEED(S), l- length of SEED in bits .

Step 2: Compute H(S) and pick the rightmost Lh bits (R)
Step 3: Initialise leftmost bit of R to zero and assign to LWo(Lh bits)
Step 4: loop i=1to Ls
Calculate LW_i=H(S+i)mod $2^l$and then concatenate the LW result after each iteration
Step 5: Find the binary expansion of LW and name as randno
Step 6: Choose the value for a and b such that r*b^2 congruent to a^3 mod p
Step 7: if the equation $4a^3$+ 27 b^2=0 mod p then start again from step 1
else accept the curve with a and b .

Session based light weight Elliptic curve cryptography (SeBliwECC) is proposed to improve the performance of big data service realm by initializing the points generated by the ECC with random key as values at each session and these values are sent to the user at the start of the session. Whenever the user accesses the service, the method selects the point randomly from the given key pool and uses the key as the encryption key to produce the SOAP request. The generated XML based SOAP file is sent to the server side. The receiver system performs decryption operation using keys found in key pools and extract the original content.

January - February 2020
ISSN: 0193 - 4120 Page No. 6046 - 6050

Encryption and Decryption methods

Procedure: Session based ECC Encryption and Decryption Algorithmn(SeBliwECC):

Step 1: Both the communicating parties have the agreement to choose the elliptic curve of finite field p - Ep(a,b) and the random point C from it.
Step 2: Person A choose the random number $\alpha<p$, the point A on the curve and compute $A1=\alpha(C+A), A2=\alpha A$
Step 3: Person B choose the random number $\beta<p$, the point B on the curve and compute $B_1=\beta(C+B), B_2=\beta B$
Step 4: Both of them keep α,β,A and B point as private respectively and remaining as GLOBAL parameter..
Step 5: Compute $pub_a=\alpha B_2$ and $pub_b=\beta A_2$
Step 6: Encryption by B:
Choose session key $\gamma<p$ and RM is the Request message
Calculate $E_1=\gamma C$ and $E_2=RM+(\beta+\gamma)A_1-\gamma A_2+pub_a$
Step 7: Decryption by A:
Receiving E_1 and $E_{2,}$, compute $RM=E_2-T$; where the $T=\alpha E_1+\alpha B_1+pub_b$

III. Conclusions:

Any agency that plays the role with defence, offenses, real-time prediction, real-time scheduling, completeness, perfection, expecting the surrounding happenings highly depends upon online massive collection of data. If their exist the digital system that automatically gather that data, track the system status and gives the alarm for replacing, predict the repairing moment and recalling the tragedy snap without any help from human being dynamically is most saluted. Sensor started replacing human through which data is continually captured and received by the targeted computers that are ready for further analysis or processing. Since sensors can capture data at speeds and in quantities that no human could ever match, they have led to the phenomenon called Big Data – or the acquisition and analysis of extremely large data sets. There is fabulous opening for IoT in manufacturing and data utilising corporate. So analysing, Big data in all dimension is well and good. Additionally outsourcing data confidentiality is proposed with SeBliwECC crypto process

IV. References:

1. Menaka R, Wahida Banu R. S. D, Ashadevi B, Survey on Signatured Xml Encryption for Multi-Tier Web Services Security , Indian Journal of Science and Technology, 2016,April , 9(16): 1-10.
2. Menaka R, Wahidha Banu R.S.D, Security Analysis in Hierarchical Resource Arrangements in Grid Computing, International Journal of Computer Applications,2011, 24(6):34-36.
3. Arvind Murali,Role-of-data-in-digital-transformation, https://blogs.perficient.com/2017/01/06, 2017,January.
4. Tom Shafer, The 42 V's of Big Data and Data Science ,https://www.kdnuggets.com/2017/04/42-vs-big-data-data-science.html, April, 2017.
5. Hiba Jasim Hadi, Ammar Hameed Shnain, Sarah Hadishaheed, Haji Ahmad, Big Data And Five V's Characteristics, International Journal of Advances in Electronics and Computer Science, 2015,January, 2(1):16-23.