

ComputationWorld 2021

Panel Discussion

Data Processing on Clouds: Complexity and Security Challenges

Moderator Herwig Mannaert, University of Antwerp, Belgium

Panellists Hans-Werner Sehring, Tallence AG, Germany Robert Duncan, University of Aberdeen, Scotland, UK Minoru Sasaki, Ibaraki University, Japan Claudio Zandron, Università Milano Bicocca, Italy Christoph Reich, Furtwangen University of Applied Science, Germany

Panel

Data Processing on Clouds: Complexity and Security Challenges (complexity, clustering, patterns, parallelism, privacy, trust, etc.)

ComputationWorld 2021

Cloud computing from Wikipedia

IARIA





ComputationWorld 2021

Challenges and questions for discussion

- Trust and privacy issues
 - Ever increasing data collection and customer consumption profiles
 - Trust and privacy not just for data, but for processing algorithms as well
 - Not just confidentiality, but also data integrity and non-repudiation
- Governance issues
 - Policy decisions should be made by democratic governments instead of big tech
 - Do governments have the ability to impose and enforce cloud data policies related to trust services ?
- Technical challenges
 - Can we leverage the possibilities offered by massive parallel data processing in the cloud ?
 - Distributed systems are known to pose major integration issues, cloud processing is (distributed computing)²
- Semantic challenges
 - Cloud computing integrations will require common dictionaries and semantics
- Any further challenges ...



ComputationWorld 2021

Topics from the panellists

Hans-Werner Sehring, Tallence AG, Germany

- The degree of trust required depends on the data and on the processing to be applied, and on the privacy they require.
- Some architectural considerations on the coordination of Cloud and on-premise computing.

Robert Duncan, University of Aberdeen, Scotland, UK

• Consider the forthcoming massive increase in internet speed through fibre and enhanced 5G on cybersecurity?

Minoru Sasaki, Ibaraki University, Japan

The reliability of word meanings in online dictionaries and how word meanings change over time.

Claudio Zandron, Università Milano Bicocca, Italy

How to implement parallel data processing in an evolving network ?

Christoph Reich, Furtwangen University of Applied Science, Germany

The security for AI as a Service.

Herwig Mannaert, University of Antwerp, Belgium

The need for cloud computing integration strategies and patterns.



ComputationWorld 2021

Panellist Position

Data Processing on clouds: integration issues

Herwig Mannaert, University of Antwerp, EU-BE <u>herwig.mannaert@uantwerp.be</u>

- Many generations of distributed computing: DCE, CORBA/DCOM, Web/REST Services
- Application Integration has always been tedious and cumbersome:
 - Messaging protocols, semantics, cross-cutting concerns, vendor lock-in, ...
- Cloud computing will require lots of integration, with extra challenges:
 - Highly distributed: micro-services, IoT integration
 - Lots of cross-cutting concerns: trust, privacy, authorization, accountability, ...
 - Fast evolving proprietary cloud platforms with limited customer participation

 \rightarrow Cloud computing integration will become complex and costly

 \rightarrow Integration strategies and patterns are needed





ComputationWorld 2021

Panellist Position

The Ongoing Cybersecurity Challenges

Bob Duncan, University of Aberdeen, UK bobduncan@abdn.ac.uk

- Ever increasing complexity of new cloud services can lead to security vulnerabilities
- The common failure to properly protect system logs and audit trails
- The ease with which attackers manipulate forensic records to cover their tracks
- The resulting non-compliance for corporates
- Leading to ever increasing regulatory fines

 \rightarrow This is not a new issue

ightarrow It has been a challenge for decades before cloud arrived

 \rightarrow It is still an unresolved challenge

→ Without a proper resolution, new super-cloud solutions will only make the outcome worse





ComputationWorld 2021

Panellist Position

The Security for AI as a Service

Christoph Reich, Furtwangen University of Applied Science, Germany rch@hs-furtwangen.de

- Professor of Computer Science in distributed system and IT security
 - Current lectures: IT security, network security, IoT, mobile computing, distributed systems, and machine learning
- Head of Institute of Data Science, Cloud Computing, and IT Security
- Research projects in cloud computing, security, machine learning mainly in the area of industry 4.0
- For SMEs often machine learning (ML) pipelines are hosted in cloud infrastructures. The data quality is critical of the inference results of ML models. Therefor the security of the infrastructure concerning the data manipulation is of great interest and mitigation solutions are needed, for each layer.
- Security measurements are needed for the data collection layer, preprocessing layer, ML training layer, ML inference layer, ML presentation layer





ComputationWorld 2021

Panelist Position

Implementing Parallel Data Processing in an Evolving Network

Claudio Zandron, DISCo, Universita' Milano Bicocca, Italy claudio.zandron@unimib.it

- How to Implement Efficient Parallel Data Processing?
- Structural (entities) and functional (functions) evolution
- One possible proposal to be investigated: application of Bio-inspired models
- Advantages: flexibility, efficiency
- Disadvantages: description of complex data structures





Panellist Position

Architectural Considerations of Trusted Information Systems in the Cloud

Hans-Werner Sehring, Tallence AG, Germany hans-werner.sehring@tallence.com

- Secured data centers vs. trusted Cloud services vs. communication in the public Internet
- Protected data vs. content publication; public services vs. competitive advantage through algorithms
- Custom solutions vs. customizable service offerings
- Data shipping vs. function shipping
- Architectural considerations on the integration of Cloud and on-premise computing.
 - \rightarrow Application-specific requirements
 - ightarrow Classes of domain-specific solutions
 - \rightarrow Architecture best practices and patterns.



ComputationWorld

2021



Panellist Position

Understanding the Meaning of Words Using Dictionary

Minoru Sasaki, Ibaraki University, Japan minoru.sasaki.01@vc.ibaraki.ac.jp

- Find out the meaning of the words
- Example Sentence Extensions for Words
- Reliability of Sense Definition of the Dictionary
- Possibility to use multiple versions of the WordNet dictionary
 - \rightarrow Dictionaries will be more useful.
 - → Humans and computers will be able to capture the characteristics of word meanings by using multiple version of the dictionary.

 \rightarrow We may be able to capture the changes in word senses over time and commonly used senses.



ComputationWorld 2021

Computation World 2021

Data Processing on Clouds: Complexity and Security Challenges (complexity, clustering, patterns, parallelism, privacy, trust, etc.)

Cloud Security Challenges by Dr Bob Duncan





Dr Bob Duncan

30 years in industry as a corporate accountant

6 years as a lecturer at University of Aberdeen

Research area cloud security and corporate compliance





Presentation Outline

- What are the cloud security challenges?
- What is the problem?
- What can we do about it?
- Why was this not done years ago?
- What else could this be used for?
- Summary



What are the cloud security challenges?

- The most serious challenge is the difficulty in securing the forensic trail of all major transactions and events
- All too often attackers are able to modify or even delete the system logging and audit trail records to cover their tracks
- This is not a new challenge
- This was a major problem a couple of decades before cloud arrived on the scene
- Many solutions were proposed, yet very few attackers have been caught



What is the problem?

- If there were solutions available before cloud, why do they not work for cloud?
- The solutions proposed before cloud are not foolproof
- The major challenge surrounds the fact that system logs and audit trails have always been captured in traditional database management systems
- And there is nothing to prevent attackers from doing what they like with these systems
- What does not work for traditional systems will equally not work for cloud



What can we do about it?

- We need to accept the reality that we have to do something more constructive to resolve these issues
- First, we need to accept that there is a need for using immutable databases, where only new records can be added, with no modifications or deletions allowed
- Second, the old approach of capturing multiple copies is good practice
- Third, we need to ensure that these are safely distributed in immutable databases, with off-cloud copies safely stored



Why was this not done years ago?

- We have had immutable databases for decades, so why was this never done before?
- Early immutable databases were clunky to operate and could not support key fields, meaning that as the volume of data increased, so the performance of the database slowed more and more, rendering them useless over time
- Thankfully, we have seen recent advancements in this area
- Highly functional immutable databases are now available
- And for a belt and braces approach, we have the option of Blockchain and Smart Contracts available to use for this now



Will it work?

- Providing the right steps are put in place, it is possible to install a serious security system that can safeguard these system logs and audit trail records, thus ensuring the forensic records can be preserved
- This will prove beneficial for corporates, because the prospect of reducing regulatory fines will be welcome
- The prospect of being able to pass on evidence to the authorities will mean there will be more chance of security conviction against attackers



What else could this be used for?

- The obvious way forward is to leverage the existence of complete forensic records by using data analytics, machine learning and artificial intelligence for actively monitoring systems to both learn about normal patterns of behaviour and to develop tailored solutions to detect anomalous behaviour
- This could lead to very capable intrusion detection systems evolving to be able to stop possible attacks before the attacker can complete their dastardly deeds



Summary

Thus, it may be possible to considerably improve the security of cloud systems, while adding improved security at the same time, by introducing a relatively simple approach to make it all happen

For those of you who are interested, the same approach could be used to resolve this issue for conventional systems, where the same problem still persists to this day







The Security for AI as a Service

Christoph Reich Christoph.Reich@hs-furtwangen.de

Institut für Data Science, Cloud Computing und IT Security

Prof. Dr. Ch. Reich Hochschule Furtwangen rch@hs-furtwangen.de idacus.hs-furtwangen.de





Security Attacks in CPS and Cloud



Prof. Dr. Christoph Reich (rch@hs-furtwangen.de)





Security Attacks ML-Specific



Evasion attack



"panda" 57.7% confidence **"gibbon"** 99.3% confidence

Prof. Dr. Christoph Reich (rch@hs-furtwangen.de)





Towards Robust, Secure, Private, and Accountable Machine Learning

- End-to-end security (CPS, cloud, and ML)
 - Trustchains (Blockchain could be the solution)
- Monitoring security and privacy
 - Anomaly detection at all process steps
 - Boundary definition
- Countermeasurments
 - Distributed CNNs
 - Introducing randomness in CNNs
 - Homomorphic encryption





Institute for Data Science, Cloud Computing and IT Security rch@hs-furtwangen.de idacus.hs-furtwangen.de



Prof. Dr. Christoph Reich (rch@hs-furtwangen.de)



Literatur

- Apruzzese, G., Colajanni, M., Ferretti, L., and Marchetti, M. (2019). "Addressing adversarial attacks against security systems based on machine learning," in 2019 11th International conference on cyber conflict (CyCon), Tallinn, Estonia, May 28–31, 2019 (IEEE), 900, 1–18
- Chen, S., Xue, M., Fan, L., Hao, S., Xu, L., Zhu, H., et al. (2018). Automated poisoning attacks and defenses in malware detection systems: an adversarial machine learning approach. Comput. Secur. 73, 326–344. doi:10.1016/j.cose.2017.11.007
- Gu, T., Liu, K., Dolan-Gavitt, B., and Garg, S. (2019). BadNets: evaluating backdooring attacks on deep neural networks. IEEE Access 7, 47230–47244. doi:10.1109/access.2019.2909068
- Jan Stodt, Daniel Schönle, Christoph Reich, Fatemeh Ghovanlooy Ghajar, Dominik Welte, and Axel Sikora; Security Audit of a Blockchain-Based IndustrialApplication Platform; MDPI
- Qayyum Adnan, Ijaz Aneeqa, Usama Muhammad, Iqbal Waleed, Qadir Junaid, Elkhatib Yehia, Al-Fuqaha Ala; Securing Machine Learning in the Cloud: A Systematic Review of Cloud Machine Learning Security; Frontiers in Big Data; 3, 2020

Bio-inspired Models to implement Parallel Data Processing in Evolving Networks

Claudio Zandron claudio.zandron@unimib.it

DISCo - Universita' di Milano-Bicocca, Italy

Future Computing Conference 2021



DIPARTIMENTO DI INFORMATICA, SISTEMISTICA E COMUNICAZIONE



Claudio Zandron



Claudio Zandron got the PhD in Computer Science from the University of Milan in 2002.

Since 2006 he is Associate Professor at the Department of Informatics, Systems and Communication of the University of Milano-Bicocca, Italy.

His research interests concern the areas of formal languages, molecular computing models, DNA computing, Membrane Computing and Computational Complexity.

- How to Implement Efficient Parallel Data Processing?
- Structural (entities) and functional (functions) evolution
- One possible proposal to be investigated: application of Bio-inspired models
- Advantages: flexibility, efficiency
- Disadvantages: description of complex data structures, programming

Implementing Parallel Data Processing

- Multicore vs Distributed Computing
- Shared Memory vs Distributed Memory
- Centralized vs Decentralized Management
- Synchronous vs Asynchronous
- Different types of evolving networks: Structural (entities) and Functional (functions)

Membrane systems

- Each membrane defines a REGION (compartment) in the membrane structure
- The most external membrane separates the system and the environment
- Chemicals (lons, molecules, proteins, etc) are distributed among regions
- Reactions modify chemicals and communicate results through membranes
- All regions evolve in parallel, all reactions are applied following a maximally parallel semantic

Bio-inspired models: Membrane systems

- Idea to be investigated
 - Regions Entities
 - Chemicals Data
 - Reactions Functions
- Simulate Membrane Systems in silico?

Apply Membrane Systems to Parallel Data Processing in an Evolving Network?

- Advantages
 - Flexible structure: membranes can be added or removed easily, without impacting the whole system
 - Functions within a region can be easily modified/deleted
 - Efficiency: all regions evolve in parallel, all reactions within a region are applied in a maximally parallel manner
 - Multi-level architecture: a node can represent both a single membrane as well as a (sub)group of membranes
- Disadvantages
 - How to represent complex data structures (linked lists, binary trees, stacks, queues)?
 - How to "program" reactions together?
 - How to organize parallel computations to be as much efficient as possible?

Discussion

- Automatic Parallelization
- Fault recovery
- Data Distribution among Entities
TA//ENCE

SOME ARCHITECTURAL CONSIDERATIONS ON THE COORDINATION OF CLOUD AND ON-PREMISE COMPUTING

Panel position, Panel 2: Data Processing on Clouds: Complexity and Security Challenges The Thirteenth International Conference on Creative Content Technologies, CONTENT 2021, April 18, 2021 to April 22, 2021 - Porto, Portugal

Hans-Werner Sehring

// Some discussion points

- 1 Secured data centers vs. trusted Cloud services vs. communication in the public Internet
- 2 Protected data vs. content publication; public services vs. competitive advantage through algorithms
- 3 Custom solutions vs. customizable service offerings
- 4 Data shipping vs. function shipping
- 5 Architectural considerations on the integration of Cloud and on-premise computing



// Integrating Cloud solutions with on-premise systems

Secured data centers vs. trusted Cloud services vs. communication in the public Internet



Dr. Hans-Werner Sehring / Content Management in the Cloud, Computation World 2013, Valencia, Spain

May 31st, 3 **//**

// Content distribution and service provision in the Cloud



Regional playout components for ...

Dr. Hans-Werner Sehring / Content Management in the Cloud, Computation World 2013, Valencia, Spain

May 31st, 2013 4

// Protection and visibility of data and services

Geben Sie inten Ontentiter ein

Protected data vs. content publication

- > Data is a valuable resource that needs to be protected.
- > For some data, legal and compliance regulations for privacy.
- > Content, e.g., marketing information, is to be distributed.

Competitive advantage through algorithms vs. public services

- > In many domains, algorithms are part of the business value. Example: risk assessment in insurance and banking domains.
- > Many software solutions provide services over public APIs that might be used in unintended ways.



// Custom solutions vs. customizable service offerings

Geben Sie Ihren Untertitel ein



- > Custom solutions
 - > Deploy your services in the Cloud
 - > Integrate available Cloud services in your solution
- > Service offerings
 - > Platforms providing services or applications.
 Typical examples: "Marketing Cloud", "Sales Cloud", "Commerce Cloud"
 - > Customizable, sometimes in the form of a framework that allows to register own code that is called back

// Data shipping vs. function shipping



Titel / Autor / Copyright

// Secure communication between Cloud solutions and on-premise systems

Architectural considerations on the integration of Cloud and on-premise computing



TA//ENCE

THANKS.

Hamburg

Tallence AG Neue Gröningerstraße 13 20457 Hamburg

T +49 40 36 09 35 100 F +49 40 36 09 35 200 info@tallence.com

Frankfurt/Main

Tallence AG Walther-von-Cronberg-Platz 1 60594 Frankfurt am Main

T +49 69 401 50 76 60 F +49 40 36 09 35 200 frankfurt@tallence.com

Karlsruhe

Tallence AG Erbprinzenstraße 23 76133 Karlsruhe

T +49 721 38 13 49 10 F +49 40 36 09 35 200 karlsruhe@tallence.com

Görlitz

Tallence AG Neißstraße 27 02826 Görlitz

T +49 40 36 09 35 100 F +49 40 36 09 35 200 goerlitz@tallence.com

Marburg

Tallence AG Am Grün 40 35037 Marburg

T +49 40 36 09 35 100 F +49 40 36 09 35 200 info@tallence.com



Panel Data Processing on Clouds: Complexity and Security Challenges (complexity, clustering, patterns, parallelism, privacy, trust, etc.)

Panellist Position

Understanding the Meaning of Words Using Dictionary

Minoru Sasaki, Ibaraki University, Japan minoru.sasaki.01@vc.ibaraki.ac.jp

- Find out the meaning of the words
- Example Sentence Extensions for Words
- Reliability of Sense Definition of the Dictionary
- Possibility to use multiple versions of the WordNet dictionary
 - \rightarrow Dictionaries will be more useful.
 - → Humans and computers will be able to capture the characteristics of word meanings by using multiple version of the dictionary.

 \rightarrow We may be able to capture the changes in word senses over time and commonly used senses.



ComputationWorld 2021

Find out the meaning of the words

- How do we behave in the digital society in order to know the meanings of words?
 - Wikipedia
 - English Dictionary, Translation Dictionary (e.g. English-Japanese)
 - WordNet
- Recently, Neural word sense disambiguation systems make use of external resources (e.g. dictionary, thesaurus)
- What is a useful dictionary for both humans and computers to find the meaning?
 - We need to further enrich the content of the existing dictionaries.

Example Sentence Extensions for Words

- Word Sense Disambiguation (WSD)
 - Identify which sense of a target polysemous word is used in an example.
 - Example sentences assigned with word senses can be added to the dictionary.
- WSD is used in many NLP tasks.
 - machine translation, question answering, information extraction, etc.



Reliability of Sense Definition

- Existing researches show sense definition or glosses is a useful resource for improving WSD.
 - GlossBERT(Huang, 2019), EWISE(Kumar, 2019), EWISER(Bevilacqua, 2020) etc.
- In WordNet, the sense definition is different depending on its version.
 - It is more reliable to use the latest version of the dictionary.
 - However, is the new version of the dictionary still applicable for old documents?

Sense Definition of the word "cool" (adjective) Red (added), Green (Modified)

• WordNet 1.7

- 02451284 neither warm or very cold; giving relief from heat; "a cool autumn day"; "a cool room"; "cool summer dresses"; "cool drinks"; "a cool breeze"
- 2. 00504017 marked by calm self-control (especially in trying circumstances); unemotional; "play it cool"; "keep cool"; "stayed coolheaded in the crisis"; "the most nerveless winner in the history of the tournament"
- 3. 02453801 inducing the impression of coolness; used especially of greens and blues and violets; "cool greens and blues and violets"
- 02452905 psychologically cool and unenthusiastic; unfriendly or unresponsive or showing dislike; "relations were cool and polite"; "a cool reception"; "cool to the idea of higher taxes"
- 5. 01862201 (informal; of a number or sum) without exaggeration or qualification; "a cool million bucks"
- 00933202 (informal) fashionable and attractive at the time; often skilled or socially adept; "he's a cool dude"; "that's cool"; "Mary's dress is really cool"; "it's not cool to arrive at a party too early"

• WordNet 3.1

- 02540350 (neither warm nor very cold; giving relief from heat) "a cool autumn day"; "a cool room"; "cool summer dresses"; "cool drinks"; "a cool breeze"
- 2. 00532977 (marked by calm self-control (especially in trying circumstances); unemotional) "play it cool"; "keep cool"; "stayed coolheaded in the crisis"; "the most nerveless winner in the history of the tournament"
- 02542621 (inducing the impression of coolness; used especially of greens and blues and violets when referring to color) "cool greens and blues and violets"; "the cool sound of rushing water"
- 4. 02541827 (psychologically cool and unenthusiastic; unfriendly or unresponsive or showing dislike) "relations were cool and polite"; "a cool reception"; "cool to the idea of higher taxes"
- 5. 02088253 (being satisfactory or in satisfactory condition) "an all-right movie"; "the passengers were shaken up but are all right"; "is everything all right?"; "everything's fine"; "things are okay"; "dinner and the movies had been fine"; "another minute I'd have been fine"
- 6. 01920631 (used of a quantity or amount (especially of money) for emphasis) "a cool million bucks"
- 7. 00974839 (fashionable and attractive at the time; often skilled or socially adept) "he's a cool dude"; "that's cool"; "Mary's dress is really cool"; "it's not cool to arrive at a party too early"

Possibility to use multiple versions of the WordNet dictionary

- Use multiple versions of the dictionary
 - It is possible to capture the changes in word senses over time and how word senses change.
 - The word senses added in the new version could be helpful in analyzing recently created documents and conversations among young people.
 - There are some word senses that do not change in multiple versions of the dictionary.
 - It is possible to capture commonly used meanings and general concepts for words.

Conclusion

- Dictionaries will be more useful "for people to understand the meaning of words" and "for WSD systems to improve the performance".
- By using multiple versions of the dictionary,
 - It may be possible to capture the changes in word senses over time and how word senses expand.
 - It may be possible to capture commonly used senses and general concepts for words.







COMPUTATIONWORLD 2021

Panel 2 Data Processing on Clouds *DETAILED POSITION* The Need for Integration Strategies

HERWIG MANNAERT



APRIL, 2021

Universiteit Antwerpen

The Quest for Distributed Plug & Play



- Monolithic applications dominated 1960's and 1970's
- Distributed architectures & standards emerged from 1980's:
 - DCE/RPC in 1980's 1990's
 - CORBA in 1990's 2000's
 - XML/RPC Web Services in 2000's 2010's
 - JSON/RPC REST Services in 2010's 2020's
 - Service Mesh Sidecar Proxy in 2020's ...
- Architectures are now being implemented on clouds
- Business objects and capabilities
 - Business capabilities keep being (re-)implemented
 - Integration issues remain relevant and challenging



HTTP

Integration Challenges – Size of Aggregation



- The customer requires more end-to-end business capabilities
 - across multiple organizations
 - across different types of systems
- The individual micro-services become more fine-grained
- Internet of Things technologies require integration of
 - increasing amount of devices
 - increasing number of protocols

Size and heterogeneity of service aggregation are increasing rapidly

Integration Challenges – Cross-cutting Concerns



- Cross-cutting concerns are often intertwined with business logic:
 - Authorization
 - Access control
 - Authentication
 - Load balancing
 - Logging, archiving
 - ...
- Standardized decoupled solutions are not available

Cross-cutting concerns lead to duplications and compatibility issues

Integration Challenges – Vendor Lock-in



- The cloud platforms provide growing list of features:
 - infrastructure services
 - computing services
 - data storage services
 - trust and security services
 - ...
- The cloud platforms are proprietary and evolving separately

Cloud platform coupling and vendor lock-in are increasing rapidly

Conclusion – Need for Integration Strategies



- The integration challenges are significant:
 - Increasing levels of service aggregation
 - Increasing number of cross-cutting concerns
 - Increasing amounts of proprietary coupling

→ We require integration strategies and patterns for cloud computing

QUESTIONS ?

herwig.mannaert@uantwerp.be