

# Author Index

- Akiva, Navot 257
- Barolli, Admir 145  
Barolli, Leonard 145  
Ben Hadji, Hend 473
- Chakraborty, Sandip 89  
Chakraborty, Suchetana 89  
Chen, Ying-ping 187  
Cheptsov, Alexey 167  
Chevallard, Sophie Vallet 473  
Ciancaglioni, V. 405  
Ciobanu, Radu-Ioan 29
- de Paula, Luciano Bernardes 1  
Dobre, Ciprian 29
- Emaldi, Mikel 443
- Gall, Franck Le 473  
Gazis, Vangelis 257  
Ghinea, Gheorghita 367  
Gluhak, Alex 473  
Goncalves, Allan 211  
Goto, Keisuke 113  
Grieco, L.A. 405  
Grønli, Tor-Morten 367
- Hansen, Jarle 367  
Hara, Takahiro 113
- Kanzaki, Akimitsu 113  
Karmakar, Sushanta 89  
Kasai, Hiroyuki 385  
Koller, Bastian 167
- Lázaro, Jon 443  
Leu, Fang-Yie 187  
Lin, Jia-Chun 187  
Liquori, L. 405  
López-de-Ipiña, Diego 443  
Loti, R. 405
- Magalhães, Maurício Ferreira 1  
Matsuo, Kazuya 113  
Moreno-Cano, María V. 341  
Morreale, Patricia 211
- Nandi, Sukumar 89  
Nishio, Shojiro 113
- Okada, Yoshihiro 231  
Omori, Yusuke 385
- Palmieri, Francesco 283  
Pasquini, Rafael 1  
Peña, Oscar 443  
Perner, Petra 57  
Piro, G. 405
- Santa, José 341  
Sawada, Yasuharu 385  
Scarfò, Antonio 283  
Shinkuma, Ryoichi 385  
Silva, Carlos 211  
Simms, David 319  
Skarmeta, Antonio F. 341  
Spaho, Evjola 145  
Strohbach, Martin 257
- Takahashi, Tatsuro 385

Villaça, Rodolfo da Silva 1

Walravens, Nils 473

Xhafa, Fatos 29, 145

Xueli, Zhang 473

Yamaguchi, Kazuhiro 385

Younas, Muhammad 367

Zamora-Izquierdo, Miguel A. 341

Ziekow, Holger 257

# Subject Index

3G 37, 38, 50  
6LowPAN 355

## A

academic 37, 43, 45, 46  
ACID 302  
ACK 117, 120  
Adult Data Set 3, 20  
adult vectors 3, 4  
ageing of relations 402  
agent data 119, 137  
Agriculture 491  
air quality control 358  
altruism 31, 47–49  
always connected 284  
anomaly detection 214  
Apache 304  
architectural perspective 352  
architecture definition 407  
automation 354

## B

bandwidth 37, 48, 50  
benchmark 480, 482, 495  
Benchmarking 473  
benchmarking framework 473, 476, 479,  
480, 482, 483, 484, 489, 494  
best practices 473, 482, 484  
Big Data 1–3, 6, 27, 283, 452  
big data analysis 211  
big data analytics 307  
Big Data Models 312

Bluetooth 31, 37  
bottleneck 32, 33, 36  
bucket 16, 17  
buffer 123  
Building 491  
business 473, 474, 475, 479, 480, 481, 483,  
484, 488, 489, 491, 493–496  
Business ecosystem 473, 474  
business model 481, 483, 484, 494, 495,  
496  
business roles 481

## C

cache 394  
caching algorithm 397  
centrality 40, 41, 43–45, 389  
child 120  
CIFS 295  
City Explorer 353  
City management 490, 491  
Cloud 287  
Clustergram 310  
clustering 389  
clustering coefficient 391  
Cold-standby redundancy 188  
collaborative 29, 47, 51  
columnar storage 302  
comfort 352  
common node 388  
communication range 116  
communication route 114  
community 34, 39–45, 47–50  
connectivity 389

computational intelligence 351  
 Consistency 145, 146, 148, 152  
 contact 31–33, 37, 40–47, 51  
 contents 3, 7, 10, 11, 19, 20  
 content delivery networks 396  
 Correlation 3, 17, 18, 22, 24, 311  
 cosine 2–5, 7, 17–19, 22, 24  
 cooperative filtering 393  
 crosswise distribution 127  
 crosswise route 125  
 crosswise tree 124  
 Crowd-sourced media contents 405  
 crowdsourcing 343

**D**

Dashboards 311  
 Data 473, 475, 482–488, 490–494, 496  
 data aggregation 114, 122  
 Data analysis 446, 449  
 Data as a service 287  
 Data availability 145, 149, 152, 163  
 Data Center 1, 3, 27  
 data cleaning 211  
 Data Collection 103  
 Data Compression 294  
 Data curation 458  
 Data Deduplication 293  
 Data discovery 446, 451, 454  
 Data Gathering 90  
 Data integration 448  
 Data life cycle 443, 445, 447, 448, 450, 451  
 Data management 444, 447, 452, 455  
 data mining 211  
 Data publication 462  
 Data quality 454  
 Data replication 145–148, 153, 156  
 data restoration 125  
 Data storage 460  
 data-centric applications on big-scale, 167  
 decoder 387  
 delivery 29, 30, 38, 39, 41–43, 45, 47–50  
 dense MWSNs 113  
 deployment 473–477, 479–486, 491–496  
 detection 33, 40–42, 45, 47–50, 52  
 devices 30, 32–41, 46, 48, 50, 474, 475, 483, 484, 491–493  
 DGUMA 114, 119  
 DGUMA/DA 114, 121  
 DGUMA/DAwoRC 139

Dijkstra's algorithm 396  
 dimension 3–6, 12, 15, 24–26  
 dimensionality 2, 6, 26  
 direct-attached storage 295  
 dissemination 30, 32, 37–40, 42, 43, 48–52  
 distribution path 394  
 Distributed Hash Table 3, 6  
 domotics 354  
 DTN 37  
 Dynamic Storage Tiering 297  
 Dynamic warm-up mechanism 187

**E**

Ecoinformatics 447  
 ecosystem 474, 475, 479, 480, 488, 491, 493, 494, 496  
 Education 491  
 electric lighting 358  
 encoder 387  
 encounter 31, 32, 36–44, 46–49  
 energy management 352  
 energy savings 358  
 Enron (Enron Corp.) 389  
 enroute caching 397  
 Environment 491  
 environmental 476, 484, 486, 488, 490  
 environmental sustainability 212  
 epidemic 30, 39, 40  
 ETL 299  
 Euclidean distance 2  
 evaluation 9, 13, 17, 19, 20, 22, 25, 27  
 evaluation criteria 478  
 evaluation framework 476  
 evolution 289

**F**

Failure detection 196  
 Failure detection algorithm 196  
 Financial model 482  
 fingers 10, 21  
 fixed-route release message 125  
 forwarding 31–33, 36, 41, 43, 44, 48, 49  
 forwarding area 116  
 forwarding route control 114, 124  
 forwarding tree 119  
 frequency distribution 3, 19, 20, 22  
 Fully parallel redundancy 192  
 Fuzzy logic 157

**G**

geographical granularity 114, 115  
 geo-routing 116  
 get 3, 7, 13, 20, 22  
 GHCN 219  
 Google Maps 215  
 government 290  
 Gray Code 10–12, 14, 17, 21  
 grid 115  
 gridID 115  
 growth of relations 402  
 GSOD 223

**H**

Hadoop 187, 304  
 HAM 354  
 Hamming DHT 1, 3, 6, 7, 9–11, 17, 19–22, 27  
 Hamming distance 2, 3, 5–7, 10, 13–15, 17, 19, 21, 24  
 Hamming similarity 1, 3, 5–7, 9, 10, 15, 17, 18, 22, 24, 27  
 HBase 306  
 HCube 1, 3, 6, 7, 9, 12–17, 22, 24, 25, 27  
 HDFS 304  
 Healthcare 491  
 Hilbert 6  
 history 38, 39, 43–45, 49  
 History flow 311  
 Hive 306  
 HMI 351  
 hop 32, 45, 47, 49  
 hops 7, 10, 11, 15, 16, 20–22, 24–26  
 Hot-standby redundancy 188  
 HSO 189  
 HSS 188  
 Human factors 485, 487  
 human-centric 348  
 HVAC 358  
 Hybrid Redundant System 187  
 hypercube 6

**I**

ICN 49, 50, 52  
 ICT 283  
 identifier 1–3, 5–7, 10–14, 16–22, 24, 26, 27

IEEE 802.11p 137  
 IMDG 299  
 In memory architectures 299  
 incentive 33, 47–49, 52  
 incentive mechanism 351  
 indexing 6  
 indicators 479, 480, 483  
 indoor localization 360  
 Information centric networking 406, 438  
 Information Life cycle Management 284  
 infrastructure 482, 483, 484, 487, 488, 490, 492, 493, 496  
 Infrastructure as a Service 287  
 infrastructure 32, 34, 36–39, 50  
 in-memory data grid 299  
 innovation 289  
 intellectual 290  
 intelligent system 359  
 Internet 30, 31, 36, 38, 49–52  
 Internet of Things 286, 341  
 Interoperability 475, 482, 487  
 investment 484, 494  
 IOP 193  
 IoT 49–52, 473  
 iot applications 344  
 IoT infrastructure 473, 482, 483, 487, 492  
 IoT Technologies 344

**J**

Java bindings for Message-Passing Interface 169  
 JobTracker 188

**K**

key performance indicators 311  
 Key-Value Store 302

**L**

Large Hadron Collider 285  
 Legal 485  
 lengthwise distribution 127  
 lengthwise route 125, 128  
 lengthwise tree 124  
 LHC 285  
 liability 290  
 link strength 389

## Linked Data

  Linked Open Data 443–445, 454

Linked Data 444, 450, 458, 464

load balance 399

Locality Sensitive Hashing 3, 5–7

lookup 20

**M**

machine learning 351

machine-to-machine 287

MANET 116

Mapper 189

MapReduce 187, 303

MapReduce cluster 188

massively parallel processing 300, 301

memory 39, 48

message 29–33, 39–41, 43–49

meta-business 312

Metadata 447

middleware 355

mobile 29–38, 40, 41, 45, 46, 48, 50, 52

mobile agent 114, 119

mobile telecommunication services 285

Mobility 31, 32, 40, 43, 44, 46, 47, 288

model 38, 42, 47, 48, 50

multi-hop wireless communication

113

MWSNs 113

MySQL 221, 299

**N**

Naïve Bayes Algorithm 225

name space design 406

NameNode 188

NDN-based service platform 405, 424

nearest neighbors 6

network bandwidth 114

network controller 394

network graph 386

network-attached storage 295

networking 29–36, 38–52

NFS 295

NL-SATA 295

NOAA 228

NoSQL 301

NR 189

NS2 199

**O**

OCP platform 355

on-line social network 405

Ontology 458

Open Data 443, 445, 457

opendata project 362

opportunistic 29–42, 44–51

Orange 228

Ordinary parallel 191

OSPF 393

overlay 34, 40, 41

**P**

P2P 1, 3, 6, 7

P2P applications 145, 147, 156,

P2P Systems 145, 146, 147, 148

packet 120, 122

packet collision 141

packet header 116

parent 120

participatory sensing 113

Payment Card Industry 292

Performance 475, 483

performance analysis and optimization tools  
169

performance evaluation 407

Periodical P-metadata backup/update 195

physical link 393

physical phenomenon 115

physical router 393

Platform as a Service 287

P-metadata 193

Policy makers 473, 478, 489, 496

popularity 41, 43–45, 47, 394

prediction 29, 34, 39, 43–47, 52

predictive modeling 214

Privacy 290, 385, 475

privacy-conscious delivery 402

profiles 2–9, 17–20, 22, 24–27

Provenance 450, 454, 457, 463

pub/sub-based controller 394

Public

  Public partnership model, 476–479, 483,  
484, 487–490, 494–496

Public safety 491

Public Sector 289

publish/subscribe 40, 41, 49, 50

put 3, 20, 22  
power law 391

## Q

query 2, 3, 7, 8, 12, 13, 17, 19, 20, 22, 24,  
25, 27

## R

Random Hyperplane Hashing 1–3, 5, 6, 13,  
17–19  
random waypoint mobility model 137  
RapidMiner 224  
RDF 445, 450, 455  
Real-time analytics 299  
real-time media services 405  
Recall 1, 7, 13, 15, 20–23, 25–27  
Reducer 189  
Regulation 485  
regulatory 473, 475  
ReHRS 187  
relational databases 301  
relational graph 385  
retrieval latency 385  
relational metric 385  
relational metric-based controller 394  
Reliability 100  
remote control 356  
Replica distribution 153,  
Replication factor 145, 146, 151, 157, 158,  
160  
Replication requirements 146, 155  
Replication techniques 145–148, 153  
retrieval 2, 6, 9, 11, 13, 20, 22, 25, 26  
RFID 286  
route fix message 125  
routing 11, 12, 16, 17, 27, 29–32, 35,  
38–41, 43–51  
RR-FOP 193  
RTP 196  
RU-FOP 193

## S

SAS 295  
SCADA 354  
Scale-out storage 294

security 32–34, 50, 51, 290, 352, 394, 475,  
490  
selfishness 33, 47–49, 52  
sensing area 115  
sensing cycle 115  
sensing point 115  
sensing time 114, 119  
sensor 29, 30, 34, 36, 37  
sensor data 120, 385  
Sensor Networks 89, 214  
sensor node 115  
sensor reading 120  
Sensors 451, 456, 466  
servers 2, 3, 7, 12–17, 22, 25, 27  
Service Level Agreements 296  
services and applications 474, 483, 493,  
496  
shared-nothing 301  
Similarity Search 1–3, 6, 7, 9, 10, 12–15,  
17, 19, 21, 22, 25–27  
sink 115  
Small world 391  
smart buildings 343  
smart cities 343  
Smart City 443, 444, 451, 452, 454,  
476–478, 485, 487–491, 496  
Smart Data 444  
smart environments 341  
smart management 346  
smartphone 30, 35, 37, 50  
Smartphones 444, 457  
social 32, 34, 37–42, 44–50, 52  
SocialCast 385  
social communication 285  
Social networks 285, 451, 456  
social perspective 350  
socio-economic 476  
Software as a Service 287  
Space Filling Curve 6, 7, 14  
SPARQL 445, 461, 466  
Spatial information flow 311  
spatial sensing range 402  
SSD 295  
stakeholders 473–476, 478, 479, 483–485,  
487–489, 491, 493, 494, 496  
Stakeholders 475, 489  
standardisation 480, 483  
State/metadata synchronization 193  
STEEPLE 480  
storage area networks 295

Storage tiering 296  
 store-carry-and-forward 31  
 STP 196  
 Symmetric Multiprocessor 301

**T**

tagging 310  
 Takeover delays 189  
 Takeover process 187  
 taxonomy 32, 39  
 technical 475, 476, 477, 479, 481  
 Technologies 473, 474  
 Technology Transfer Center 357  
 tele-assistance 356  
 temporal sensing range 402  
 Thin Provisioning 295  
 time series data analysis 220  
 timer 121  
 T-metadata 193  
 Tourism 491  
 transmission 30, 34, 36, 40, 47, 50  
 Transportation 491  
 tree information 125  
 Tree Maintenance 96  
 Tree Management 95  
 Trust 454  
 Trusted IoT 345

**U**

unified collaboration 285  
 University of Murcia 352  
 URI 445  
 user feedback 353  
 user involvement 346  
 user-centric 342

**V**

valid area 116  
 valid data 116  
 Value network 479–482, 494, 495  
 Value proposition 482  
 Vector Space Model 3  
 Visualisation 452, 467  
 Visualization Systems 310

**W**

Web 444  
   Semantic Web 444  
 Weka 212  
 WiFi 31, 35, 37, 38, 50  
 wireless 29, 30, 37, 38, 50  
 WSO 189  
 WSS 188  
 XOR 12, 13, 16, 26, 27



# Acronyms

3G/4G	3 <sup>rd</sup> /4 <sup>th</sup> Generation (of mobile telecommunications technology)
6LowPAN	IPv6 over Low Powered Wireless Personal Area Networks
ACID	Atomicity, Consistency, Isolation, Durability
ACK	ACKnowledgement
ADSL	Asymmetric Digital Subscriber Line
API	Application Program Interface
API	Application Programming Interface.
API	Application Programming Interface
BS	base station
C2DM	Cloud to Device Messaging
CAN	Controller Area Network
CAPIM	Context-Aware Platform using Integrated Mobile services
CC	Convergecast Controller
CCN	Content-Centric Networking
CDN	Content Delivery Network
CDS	Connected Dominating Set
CI	Computational Intelligence
CIFS	Common Internet File System
CoAP	Constrained Application Protocol
COMET	COntent Mediator architecture for content-aware nETworks
CPU	Central Processing Unit
CPU	Central Processing Unit
CRUD	Create, Read, Update, Delete
CS	Content Store
CSV	Comma Separated Value
CSV	Comma Separated Values.
DaaS	Data as a Service
DBMS	Distributed Database Management Systems
DFS	Depth First Search
DGUMA	Data Gathering method Using Mobile Agents
DGUMA/DA	DGUMA with Data Aggregation

DGUMA/DAwRC	DGUMA/DA without Route Control
DHT	Distributed Hash Table
DONA	Data-Oriented Network Architecture
DPG	Data Placement Graph
DRAM	Dynamic Random Access Memory
DSN	Deep Space Network
DSN	Detected Social Network
DTN	Delay-Tolerant Network
Eclat	Equivalence Class Clustering and Bottom-up Lattice Traversal
EDGE	Enhanced Data Rates for GSM Evolution
EIB	European Installation Bus
ESDA	Exploratory Spatial Data Analysis
ETSI	European Telecommunications Standards Institute
ETT	Electronic Triage Tag
EWIDS	Extreme Wireless Distributed Systems
FC	Fiber Channel
FIB	Forwarding Information Base
FIFO	First Input First Output
FL	Fuzzy Logic
FRB	Fuzzy Rule Base
FTP	File Transfer Protocol
GHCN	Global Historical Climatology Network
GIS	Geographic Information System
GPFS	General Parallel File System
GPS	Geographical Positioning System
GPS	Global Positioning System
GSG	Global Serialization Graph
GSOD	Global Summary of Day
HAM	Home Automation Module
HDD	Hard Disk Drive
HDFS	Hadoop Distributed File System
HMI	Human Machine Interface
HPC	High Performance Computing
HPC	High Performance Computing
HTML	HyperText Markup Language.
HTTP	HyperText Transfer Protocol.
HTTP	HyperText Transfer Protocol
HVAC	Heating, Ventilation and Air Conditioning
I/O	Input/Output
ICN	Information-Centric Network
ICN	Information Centric Network
ICNRG	ICN Research Group

ICT	Information & Communication Technology
ICT	Information and Communication Technologies
ID	Identity
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
ILM	Information Life-Cycle Management
IMDG	In Memory Data Grid
IO	Input-Output
IOPS	IO per Second
iOS	iPhone Operating System
IoT	Internet of Things
IoT	Internet of Things
IoT	Internet of Things
IP	Internet Protocol
IRTF	Internet Research Task Force
iSCSI	Internet SCSI
ISO	International Organization for Standardization
ISP	Internet Service Provider
IT	Information Technology
ITS	Intelligent Transport System
JNI	Java Native Interface
JPG	joint photographic expert group
JS	JavaScript.
JSON	JavaScript Object Notation.
KML	Keyhole Markup Language.
LAN	Local Area Network
LarKC	Large Knowledge Collider
LFU	Least Frequently Used
LHC	Large Hadron Collider
LLD	Linked Life Data
LOD	Linked Open Data.
LRU	Least Recently Used
LSG	Local Serialization Graph
LSH	Locality Sensitive Hashing
MAETT	Mobile Agent Electronic Triage Tag
MANET	Mobile Ad hoc NETWORK
MANET	Mobile Ad-Hoc Network
MFV	Most Frequently Visited
MLN	Most Likely Next
MMP	massively parallel-processing
MPI	Message-Passing Interface
MR	MapReduce
MWSNs	Mobile Wireless Sensor Networks
N3	Notation3.
NAS	Network Attached Storage
NDN	Named Data Networking
NDO	Named Data Object

NDP	Number of Documents per Peer
NER	Named Entity Recognition.
NFC	Near Field Communication
NFS	Network File System
NLP	Natural Language Processing.
NL-SATA	Near Line Serial ATA
OAuth	Open Authorization
OCF	Open Context Platform
OLTP	Online transaction processing
OMPI	Open MPI
OMPIJAVA	Java bindings for Open MPI
ON	Opportunistic Network
OPAL	Open Portable Access Layer
ORTE	Open Run-Time Environment
OWL	Web Ontology Language.
P2P	Peer-to-Peer
P2P	Peer-to-Peer
P2P	Peer-to-Peer
PaaS	Platform-as-a-Service
PC	Personal Computer
PIT	Pending Interest Table
PMPI	Profiling Interface for Open MPI
PNG	Portable Network Graphics
POI	Point Of Interest.
PSN	Pocket-Switched Network
PSNR	Peak Signal-to-Noise Ratio
PTP	Parcel Transfer Protocol
PURSUIT	Publish Subscribe Internet Technology
QoE	Quality of Experience
QoS	Quality of Service
RAM	Random Access Memory
RAM	Random Access Memory
RDF	Reach Data Framework
RDF	Resource Description Framework.
RDFa	Resource Description Framework in Attributes.
REST	REpresentational State Transfer.
REST	Representational State Transfer
RF	Replication Factor
RF	Radio Frequency
RFID	Radio Frequency Identification Device
RFID	Radio Frequency Identification
RHH	Random Hyperplane Hashing

RP	Replication Percentage
RPC	Remote Procedure Call
RSS	Really Simple Syndication.
RSS	Really Simple Syndication
SaaS	Software as a Service
SAIL	Scalable and Adaptive Internet Solutions
SAN	Storage Area Network
SAS	Serial Attached SCSI
SCADA	Supervisory Control and Data Acquisition
SCF	Store-Carry-and-Forward
SFC	Space Filling Curve
SLA	Service Level Agreement
SLC/MLC-SSD	Single-level cell/ Multi-level cell solid state drive
SN	Social Network
SNC	Saami Network Connectivity
SOA	Service Oriented Architectures
SOAP	Simple Object Access Protocol.
SP	Super-Peer
SPARQL	SPARQL Protocol and RDF Query Language.
SQL	Structured Query Language, computer language to manipulate relational database
SQL	Structured Query Language.
SRP	Scale of Replication per Peer
SRSN	Self-Reported Social Network
SSD	solid state drives/devices
SWIM	Shared Wireless Infostation Model
TECD	Time-Evolving Contact Duration
TMM	Tree Management Module
TSV	Tab Separated Values.
URI	Uniform Resource Identifier.
US	Uniform Social
VANET	Vehicular Ad-Hoc Network
VSM	Vector Space Model
W3C	World Wide Web Consortium.
WLAN	Wireless Local Area Network
WMS	Web Map Service.
WOM	Word-of-Mouth
XML	eXtensible Markup Language.
XML	Extensible Markup Language
XOR	eXclusive OR

# Glossary

$\alpha$ : The total number of consecutive heartbeats that a server has not received from another server.

$H_{r'}$ : A hash value stored in the hash pool.

$H_r$ : The hash value of log record  $L_r$  where  $r$  ranges from  $maxID+1$  to  $Z$ .

$L_v$ : A log record generated by the ReHRS master server whenever a write operation is initiated or completed, where  $v \geq 1$  is an incremental record ID.

$M_{bef}$ ,  $H_{bef}$ , and  $W_{bef}$ : They are nodes that act as the master server, HSS, and WSS before a state transition, respectively.

$M_{aft}$ ,  $H_{aft}$ , and  $W_{aft}$ : They are nodes that act as the master server, HSS, and WSS after a state transition, respectively.

$N_u^S$ : It represents node  $N_u^S$  state (also called role) is  $S$ , where  $u = 1, 2, 3$  and  $S \in M, H, W, n/a$  in which M, H, W, and n/a respectively stand for the master server, HSS, WSS, and unavailable

$T_{early}$ : The time period from the moment when the WSS is requested to warm itself up to the moment when it is requested to take over for the HSS.

$T_{remain}$ : The time required by the ReHRS's WSS to finish its warm-up process

$T_{total}$ : The total time required by the WSS to finish its warm-up process when it receives a warm-up request from the commander.

$T - metadata$ : The metadata that is frequently updated.

**K**: It refers to the largest record ID of the log record currently collected in the HSS's journal

**masID**: It is the maximum ID of the log records currently collected in the WSS's journal

**Analytics**: Using software-based algorithms and statistics to derive meaning from data

**Anomaly detection**: is the task of identifying events or measured characteristics which differ from other data or the expected measurements.

**Assurance**: the process of obtaining and using accurate and current information about the efficiency and effectiveness of policies and operations, and the status of compliance with the statutory obligations, in order for management to control an organization's activities.

**Audit comfort**: the support needed to draw conclusions provided by audit procedures.

**Audit procedures:** the various tests of details, analytical procedures, confirmations and other activities performed to gain audit evidence.

**Batch layer:** Refer to the corresponding processing layer of the Lambda architecture [38] that processes data in batches.

**Batch processing:** Refers to processing methods that collect data over a certain amount of time and then process several data sets as a whole.

**Big data:** the accumulation of datasets from different sources and of different types that can be exploited to yield insights.

**Big Data analytics framework:** An architectural reference model including architectural principles and a set of tools suitable for analytical challenges that require model learning and near real-time decision making based on large data sets.

**Big data:** refers to very large datasets such the data cannot be processed using techniques for smaller data collections; large data management and data processing tools have been developed for big data analysis.

**Big Data:** We use the term *Big Data* both to refer to "datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze" [37].

**Business Intelligence:** software system to reports, analyzes and presents data. These tools use data previously stored in data warehouse/Data Mart

**Cassandra:** an open source database management for huge amount of data on a distributed systems. Currently part of apache software foundation

**CEP:** *See* Complex Event Processing.

**Clickstream analytics:** The analysis of users' Web activity through the items they click on a page

**Cloud computing:** a computing paradigm in which highly scalable computing resources, often configured as a distributed system, are provided as a service through a network.

**Cold-standby:** redundancy When this scheme is employed, the task fails on one node will be executed by another node from the very beginning.

**Complex Event Processing:** Refers to processing methods that are tailored for application of logic over continuous data flows. Complex event processing derives complex events from more simple events.

**Dashboard:** A graphical reporting of static or real-time data on a desktop or mobile device. The data represented is typically high-level to give managers a quick report on status or performance

**Data center:** A physical facility that houses a large number of servers and data storage devices. Data centers might belong to a single organization or sell their services to many organizations

**Data cleaning:** refers to the process of detecting incorrect data in a dataset and removing or correcting the data in such a manner that it would be useable, such as correcting the data format.

**Data Mart:** subset of Data Warehouse ready to be used by Business Intelligence tools

**Data mining:** refers to retrieving data from a dataset, looking for patterns within this data, and displaying it in an intelligible way.

**Data mining:** The process of deriving patterns or knowledge from large data sets

**Data set:** A collection of data, typically in tabular form

**Data Warehouse:** a suite of applications and databases optimized for storing and processing large amount of structured data

**DB, DBMS:** Database management system. Software that collects and provides access to data in a structured format

**Design effectiveness:** that internal control procedures are appropriate in respect of the control objective they are intended to achieve and the risk they are intended to address.

**Device:** Technical physical component (hardware) with communication capabilities to other IT systems. A device can be either attached to or embedded inside a Physical Entity, or monitor a Physical Entity in its vicinity.

**E:** *E* could be either the HSS or the WSS. If *E* is the HSS, Whenever RTP times out, it checks to see whether it has received a heartbeat from the WSS during the RTP or not. But if *E* is the WSS, then it checks to see whether it has received a heartbeat from the HSS during the RTP or not.

**ETL process:** Extract, Transform and Load, tools to extract data form sources and transform them for operational needs and load in data store architectures (e.g. data warehouse)

**Fully parallel redundancy:** It refers to a set of schemes that consists of several servers. Each of them can offer a transparent takeover when another server fails.

**Hadoop:** An open-source MapReduce implementation developed by Apache.

**Hadoop:** Open Source framework to process large data set on distributed systems. Currently part of apache software foundation

**Hot-standby redundancy:** This scheme provides a backup node to maintain an up-to-date copy of the state of a master server. When the master server fails, the backup node can continue the operation of master server.

**HSO:** A hot-standby-only scheme.

**HSS:** A hot-standby server employed by the ReHRS.

**Hybrid Redundant System:** The full name of ReHRS.

**In-Stream processing:** Refers to methods that process data continuously (cf *batch processing*).

**Intelligent Transportation Systems:** A distributed architecture based on ICT components (e.g., sensors, actuators, servers, etc.) and M2M communications that is used to manage and control a transportation infrastructure.

**Internal controls:** the activities performed to ensure that policies and procedures are implemented and operated consistently and effectively, allowing errors and omissions to be prevented or detected.

**Internet of Things:** A collection of things having identities and virtual personalities operating in smart spaces using intelligent interfaces to connect and communicate within social, environmental, and user contexts.

**Internet of Things:** A global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on, existing and evolving, interoperable information and communication technologies

**Interoperability:** The ability to share information and services. The ability of two or more systems or components to exchange and use information. The ability of systems to provide and receive services from other systems and to use the services so interchanged to enable them to operate effectively together.



**IOP:** It is an initiated operation.

**IoT:** *See* Internet of Things.

**ITS:** *See* Intelligent Transportation Systems.

**JobTracker:** One of the master server of MapReduce. It coordinates all jobs running on the MapReduce clusters, performs task assignment for each job, and monitors the progresses of all map and reduce tasks

**Lambda architecture:** A high level architecture for Big Data systems defined by Marz and Warren [38].

**M2M:** *See* Machine-to-machine communication.

**Machine-to-machine communication:** A communication capacity whereby a set of distributed components (e.g., sensors, actuators, servers, etc.) communicate to support a computational process in realizing its established policy and where the process itself has no human participation (i.e., in terms of neither a provider of inputs to the process nor a consumer of outputs of the process).

**Map task:** A task of a MapReduce job used to execute the user-defined map function.

**Mapper:** A machine/worker used to execute a map task

**MapReduce:** A distributed programming model introduced by Google to process a vast amount of data in parallel on large-scale machines/clusters.

**MapReduce cluster:** A cluster consisting of a set of commodity machines/workers for conducting MapReduce jobs.

**MapReduce:** method for distributing tasks across nodes

**Mashup:** functionality that combines data presentation or presentation from two or more sources to provides new services

**Metadata:** information that describe content and context of other data like files

**Model learning:** Summarizes algorithms that characterize large data sets by considerably smaller data sets.

**NameNode:** Another MapReduce master server, which manages the distributed filesystem namespace and processes all read and write requests.

**noSQL:** Refers to data storage systems that use non-relational data models.

**NR:** A No-Redundant scheme, which is similar to the redundant scheme adopted by Hadoop for JobTracker.

**NS2:** A simulation tool supporting TCP, routing, and multicast protocols over wired and wireless networks.

**Operating effectiveness:** that internal control procedures are functioning as de-signed, regularly and consistently.

**Ordinary parallel:** The same to hot-standby redundancy

**P-metadata:** It is the execution result of an operation and is infrequently or never changed since it is generated.

**Predictive modeling:** involves the study of data and outcomes from earlier measurements or experiences, and developing a model to predict future data or experiences.

**Q:**  $Q$  could be either the HSS or the WSS. If  $Q$  is the HSS, it sends a heartbeat to the WSS whenever STP times out. But if  $Q$  is the WSS, it sends a heartbeat to the HSS whenever STP times out.

**Reduce task:** A task of a MapReduce job used to execute the user-defined reduce function.

**Reducer:** A machine/worker used to execute a reduce task

**Redundancy mechanisms:** A set of common methods to improve system reliability.

**ReHRS :** A reliable hybrid redundant system.

**Relational database:** data stored in relation-shaped tables

**Representational State Transfer:** An application model whereby behavior is structured on the basis of elementary operations (e.g., create, update, delete, etc.) upon distinct resources whose identity is represented in a URL notation.

**REST:** *See* Representational State Transfer.

**RR-FOP:** It is a response-required finished operation. It means that  $L_v$  records information concerning an operation requested by a client/worker, and this operation have been finished by the master server.

**RTP:** It is a predefined receiving time period.

**RU-FOP:** It is a response-unrequired finished operation. It means that  $L_v$  records information concerning an operation both initiated and finished by the master server

**Scalability:** The ability of a system or process to maintain acceptable performance levels as workload or scope increases

**Schema:** The structure that defines the organization of data in a database system

**SCL:** *See* Service Capability Layer.

**Security triad:** the three key information security objectives of Confidentiality, Integrity and Availability.

**Semi structured data:** data not resides in fixed structures but that contain markers to organize data elements

**Service Capability Layer:** A set of services offered to applications utilizing M2M capacities that is organized through a layer of software (e.g., on top of middleware platform).

**SIM:** *See* Subscriber Identity Module.

**SQL:** Standard Query Language. Refers to the standardized query language for relational database management systems.

**STP:** It is a predefined sending time period. Every STP, the HSS and WSS in the ReHRS mutually send a heartbeat to each other.

**Stream processing:** Refers to technologies that are tailored for applying processing logic over continuous streams of linearly ordered data with a given schema.

**Streaming layer:** Refers to the corresponding processing layer of the *Lambda architecture* [38] that uses *stream processing* for processing data continuously.

**Structured data:** data resident in statically dimensioned structures like matrixes or tables

**Subscriber Identity Module:** An embedded component that provides cryptographic mechanisms (e.g., authentication, authorization, etc.) of accessing the identity information of a subscriber to a service (e.g., a cellular mobile data service) and the profile information associated to that subscription.

**Takeover delays:** The time period from the moment when a server fails to the moment when another node acts as the server. It comprises the server failure detection time, IP address reconfiguration time, P-metadata retrieval time, and T-metadata retrieval time.

**Time series:** refers to a sequence of observations which are ordered in some way according to time or space. For example, temperature samples collected over time could

represent time series data, if the temperatures were gathered in sequence, over time.

**Unstructured data:** data sitting outside structured and organized data repositories such as relational databases.

**Unstructured data:** data that not resides in fixed structures like matrixes, tables etc.

**Visualization:** tools for providing a synoptic view of information

**Warm-standby redundancy:** When this scheme is employed, the states of the master server are periodically replicated to a warm-standby node. When the master server fails, the state replica can be used to restart the operation of the master server.

**WSO:** A warm-standby-only scheme.

**WSS:** A warm-standby server employed by the ReHRS

**x:** A pre-defined threshold to request the WSS to warm itself up.

**y:** A pre-defined threshold of taking over the failed server.

**Z:** The largest ID of the log records collected in the journal of the commanders

**ZFS:** open source distributed file system solution implemented by Sun Microsystems