Σύστημα Υποστήριξης Κλινικών Αποφάσεων
για τη Νόσο των Ανευρυσμάτων Κοιλιακής Αορτής
Βασισμένο σε Μοντέλα Τεχνητής Νοημοσύνης

Παραδοτέο Π3.3.

Υποδομή Δεδομένων Νέφους και Ομοσπονδιακές Βάσεις: Ανάπτυξη ομοσπονδιακών βάσεων δεδομένων καθώς και η παροχή εξουσιοδοτημένης πρόσβασης σε αυτές

Όνομα Αρχείου:	Safe-Aorta-Π3.3-Υποδομή Νέφους και Ομοσπονδιακές Βάσεις Δεδομένων (Τμήμα 2).pdf	**Επίπεδο Διάδοσης:**	Δημόσιο
Ημερομηνία Υποβολής:	Ιούλιος 2025 (M24)	**Κωδικός Έργου:**	TAEDR-0535983
Κοινοπραξία:	ΕΜΠ, ΠΔΜ, ΠΚ, ΕΛΜΕΠΑ, ΠΑΔΑ, ΙΙΒΕΑΑ, ΠΑΠΕΛ	**Υπεύθυνος Παραδοτέου:**	Πανεπιστήμιο Ιωαννίνων (ανάδοχος)
Διάρκεια:	28 μήνες	**Κατάσταση:**	Τελική έκδοση


Ελλάδα 2.0
ΕΘΝΙΚΟ ΣΧΕΔΙΟ ΑΝΑΚΑΜΨΗΣ
ΚΑΙ ΑΝΘΕΚΤΙΚΟΤΗΤΑΣ


ΓΓΕΚ
ΓΕΝΙΚΗ ΓΡΑΜΜΑΤΕΙΑ
ΕΡΕΥΝΑΣ ΚΑΙ ΚΑΙΝΟΤΟΜΙΑΣ


Με τη χρηματοδότηση
της Ευρωπαϊκής Ένωσης
NextGenerationEU

ΛΙΣΤΑ ΣΥΓΓΡΑΦΕΩΝ

Συγγραφείς				
#	**Επίθετο**	**Όνομα**	**Φορέας**	**Email Επικοινωνίας**
1	Καλατζής	Θεοφάνης	ΕΛΚΕ ΠΙ	tkalatz@gmail.com
2	Σιόγκας	Παναγιώτης	ΕΛΚΕ ΠΙ	psiogkas@uoi.gr
3	Πεζούλας	Βασίλειος	ΕΛΚΕ ΠΙ	bpezoulas@gmail.com
4	Πλέουρας	Δημήτριος	ΕΛΚΕ ΠΙ	dipleouras@gmail.com
Συν-συγγραφείς				
#	**Επίθετο**	**Όνομα**	**Φορέας**	**Email Επικοινωνίας**
1	Φωτιάδης	Δημήτριος	ΕΛΚΕ ΠΙ	fotiadis@uoi.gr
2	Κούρου	Κωνσταντίνα	ΕΛΚΕ ΠΙ	konstadina.kourou@gmail.com
3	Γκόης	Γιώργος	ΕΛΚΕ ΠΙ	gkois@yahoo.com

ΛΙΣΤΑ ΚΡΙΤΩΝ

Κριτές				
#	**Επίθετο**	**Όνομα**	**Φορέας**	**Email Επικοινωνίας**
1	Μανόπουλος	Χρήστος	ΕΜΠ	manopoul@central.ntua.gr
2	Ράπτης	Αναστάσιος	ΕΜΠ	raptistasos@mail.ntua.gr
3	Μανόπουλος	Χρήστος	ΕΜΠ	manopoul@central.ntua.gr

ΕΛΕΓΧΟΣ ΑΝΑΘΕΩΡΗΣΗΣ

Έκδοση	**Συγγραφέας**	**Ημερομηνία**	**Κατάσταση**
0.1			Προσχέδιο
0.2			
0.3			
1	Παναγιώτης Σιόγκας	24/7/2025	Τελικό

Περιεχόμενα

Λίστα Συντομογραφιών

Συντομογραφία	Ορισμός
ΑΚΑ	Ανεύρυσμα Κοιλιακής Αορτής
GDPR	General Data Protection Regulation (Γενικός Κανονισμός για την Προστασία Δεδομένων)
ΣΥΠΟΚΑ	Σύστημα Υποστήριξης Κλινικών Αποφάσεων
VM	Virtual Machine (Εικονική Μηχανή)
VPN	Virtual Private Network (Εικονικό Ιδιωτικό Δίκτυο)
TLS	Transport Layer Security
TPM	Trusted Platform Module
UEFI	Unified Extensible Firmware Interface
AES	Advanced Encryption Standard
ISO/IEC	International Organization for Standardization / International Electrotechnical Commission
RBAC	Role-Based Access Control
ENISA	European Union Agency for Cybersecurity

Περίληψη

Το παρόν παραδοτέο τεκμηριώνει την υλοποίηση της τεχνικής υποδομής για την ανάπτυξη **ομοσπονδιακών βάσεων δεδομένων** καθώς και την **παροχή ασφαλούς, εξουσιοδοτημένης πρόσβασης** σε αυτές, στο πλαίσιο του έργου **SAFEAORTA**. Η υποδομή που δημιουργήθηκε έχει στόχο την ασφαλή αποθήκευση και διαμοιρασμό ιατρικών δεδομένων μεταξύ κλινικών και ερευνητικών φορέων, διασφαλίζοντας ταυτόχρονα την προστασία της ιδιωτικότητας και τη συμμόρφωση με το νομικό και κανονιστικό πλαίσιο (GDPR, ISO/IEC 27001, ENISA, CSA guidelines).

Ιδιαίτερη έμφαση δόθηκε στην πρακτική **υλοποίηση των ομοσπονδιακών βάσεων δεδομένων** με χρήση **MySQL FEDERATED engine** και **PostgreSQL**, αξιοποιώντας την εικονική υποδομή που φιλοξενείται στο **σύμπλεγμα Precious**. Η ενότητα παρουσιάζει τεκμηριωμένα:

- Το **τεχνικό σενάριο** εγκατάστασης, συμπεριλαμβανομένων των **δικτυακών παραμέτρων**, των **ρυθμίσεων πρόσβασης** και της **δομής των πινάκων** στις δύο τοποθεσίες: το **κλινικό κέντρο** και τον **ομοσπονδιακό κόμβο**.
- Την **εφαρμογή του μηχανισμού FEDERATED** σε MySQL, επιτρέποντας την προσπέλαση απομακρυσμένων δεδομένων ως εάν ήταν τοπικά, χωρίς αντιγραφή του περιεχομένου. Τεκμηριώνονται οι εντολές SQL, τα σχήματα σύνδεσης και οι περιορισμοί του μοντέλου.
- Την **παράλληλη υλοποίηση με PostgreSQL**, δίνοντας ευελιξία και τη δυνατότητα υποστήριξης πιο προηγμένων διασυνδέσεων με ασφάλεια SSL/TLS και δυνατότητα peer-to-peer ελέγχου.
- Την **τελική αξιολόγηση**, που αποδεικνύει τη λειτουργικότητα της υποδομής, την επιτυχή πρόσβαση δεδομένων με ελεγχόμενους λογαριασμούς χρηστών, και την ορθή ρύθμιση δικαιωμάτων και μηχανισμών πιστοποίησης.

Σε συνδυασμό με την προηγμένη **υποδομή ασφαλείας** (VPN με TLS 1.3, UEFI Secure Boot, TPM, data-at-rest encryption) και την **υποστήριξη ομοσπονδιακής μάθησης (federated learning)**, η υλοποιηθείσα λύση ανταποκρίνεται πλήρως στους στόχους του έργου SAFEAORTA. Το τεχνικό σύστημα αποτελεί κρίσιμη βάση για τη διαλειτουργική, προστατευμένη και ερευνητικά χρήσιμη διαχείριση των ιατρικών δεδομένων.

1. Εισαγωγή

1.1 Σκοπός του παραδοτέου

Το παρόν παραδοτέο έχει ως βασικό στόχο την τεκμηριωμένη παρουσίαση της ανάπτυξης ενός συστήματος **ομοσπονδιακής βάσης δεδομένων**, το οποίο θα προσφέρει **ασφαλή, διαλειτουργική και αποκεντρωμένη διαχείριση ευαίσθητων ιατρικών δεδομένων** που σχετίζονται με ασθενείς με ανεύρυσμα κοιλιακής αορτής (ΑΚΑ). Το σύστημα αυτό αποτελεί υποδομή-κλειδί για την ευρύτερη λειτουργικότητα της πλατφόρμας SAFEAORTA, ενισχύοντας τη συλλογή, αποθήκευση, ανάλυση και αξιοποίηση δεδομένων που απαιτούνται για τη δημιουργία εξατομικευμένων ψηφιακών διδύμων της αορτής (ΨηφιΔΑ) και την υποστήριξη λήψης κλινικών αποφάσεων.

Σύμφωνα με το **άρθρο 25 και άρθρο 32 του Γενικού Κανονισμού για την Προστασία Δεδομένων (GDPR) [1]**, καθίσταται υποχρεωτική η αρχή της «**ασφάλειας εκ σχεδιασμού και εξ ορισμού**» (security by design and by default), καθώς και η προστασία των δεδομένων τόσο κατά τη μεταφορά όσο και κατά την αποθήκευσή τους (data in transit & at rest). Η υποδομή που περιγράφεται σε αυτό το παραδοτέο ανταποκρίνεται πλήρως στις απαιτήσεις αυτές, μέσω τεχνολογιών όπως:

- Ομοσπονδιακή αρχιτεκτονική βάσεων δεδομένων (π.χ. MySQL Federated) για **αποκεντρωμένη διαχείριση δεδομένων**,
- Εικονικές μηχανές (VMs) με **κρυπτογράφηση επιπέδου δίσκου (vSAN encryption)**,
- **Ασφαλής πρόσβαση μέσω VPN και TLS v1.3**, διασφαλίζοντας την αυθεντικότητα και ακεραιότητα των συνδέσεων,
- Υποστήριξη **live migration**, υψηλής διαθεσιμότητας (HA) και **αδιάλειπτης λειτουργίας** (zero downtime),
- Συμμόρφωση με **διεθνή πρότυπα ασφάλειας** όπως ISO/IEC 27001, ISO/IEC 27701 και ENISA guidelines [2], [3].

Επιπλέον, το παραδοτέο αναδεικνύει τη σημασία της διατήρησης του ελέγχου από τις επιμέρους κλινικές μονάδες, αποφεύγοντας τη συγκέντρωση δεδομένων σε ένα κεντρικό σημείο. Η προσέγγιση αυτή ευθυγραμμίζεται πλήρως με τις αρχές της **ομοσπονδιακής μάθησης (federated learning)**, ενισχύοντας την προστασία της ιδιωτικότητας των ασθενών κατά την ανάπτυξη αλγορίθμων μηχανικής μάθησης.

Τέλος, το παραδοτέο αυτό στοχεύει να αποτελέσει τη βάση για τη λειτουργία εργαλείων και διεπαφών που θα επιτρέψουν σε αναγνωρισμένους ερευνητές και επαγγελματίες υγείας να εκπαιδεύσουν αλγορίθμους τεχνητής νοημοσύνης, να αναπτύξουν προγνωστικά μοντέλα και να προτείνουν εξατομικευμένες παρεμβάσεις, με σεβασμό στη νομική και δεοντολογική συμμόρφωση.

1.2 Αντικείμενο του Έργου και Αναμενόμενα Αποτελέσματα

Το έργο **SAFEAORTA** στοχεύει στην υλοποίηση μιας καινοτόμου, διαλειτουργικής και ασφαλούς υποδομής για την **προγνωστική ανάλυση και διαχείριση ασθενών με ανεύρυσμα κοιλιακής αορτής (ΑΚΑ)**. Η προσέγγιση βασίζεται στη **συλλογή, αποθήκευση και ανάλυση**

ετερογενών ιατρικών δεδομένων από διάφορους παρόχους υγειονομικής περίθαλψης, με σκοπό τη δημιουργία ενός **ψηφιακού διδύμου** της αορτής (ΨηφιΔΑ) για κάθε ασθενή. Η εξατομικευμένη προγνωστική ιατρική καθίσταται πλέον εφικτή, αξιοποιώντας τεχνικές τεχνητής νοημοσύνης (AI) και μηχανικής μάθησης (ML), εντός ενός πλαισίου πλήρους **νομικής συμμόρφωσης και προστασίας προσωπικών δεδομένων**.

Το αντικείμενο του παρόντος παραδοτέου εντάσσεται σε έναν από τους πυλώνες του έργου, δηλαδή την **υποδομή αποθήκευσης και πρόσβασης δεδομένων**, και συγκεκριμένα:

- Την **ανάπτυξη ομοσπονδιακών βάσεων δεδομένων** για την ασφαλή και αποκεντρωμένη διαχείριση ευαίσθητων δεδομένων.
- Την **παροχή ασφαλούς, εξουσιοδοτημένης και αποκλειστικής πρόσβασης σε εικονικά περιβάλλοντα** μέσω VPN και HTTPS/TLS v1.3.
- Τη **διασφάλιση της ακεραιότητας, διαθεσιμότητας και εμπιστευτικότητας** των δεδομένων με τη χρήση τεχνολογιών κρυπτογράφησης (data at rest & in transit), UEFI Secure Boot, TPM, και vSAN.
- Τη **δυνατότητα ανάπτυξης και εκπαίδευσης μοντέλων τεχνητής νοημοσύνης εντός των VM**, χωρίς εξαγωγή των δεδομένων εκτός του προστατευμένου περιβάλλοντος (on-premise federated learning).
- Τη **συμμόρφωση με πρότυπα ISO/IEC 27001, ISO/IEC 27701**, ENISA Cybersecurity Guidelines και **το κανονιστικό πλαίσιο του GDPR** [1][3].

Αναμενόμενα Αποτελέσματα:

1. **Ασφαλής Αποθήκευση και Επεξεργασία Δεδομένων:**
   Η υλοποίηση βάσεων δεδομένων τύπου MySQL Federated θα επιτρέψει την αποκεντρωμένη αποθήκευση των δεδομένων χωρίς αναγκαστική μεταφορά σε κεντρικούς servers, μειώνοντας τον κίνδυνο παραβίασης.
2. **Υποδομή για Federated Learning:**
   Με χρήση εργαλείων όπως το **Flower** ή το **Fedbiomed**, οι εξουσιοδοτημένοι χρήστες θα μπορούν να εκπαιδεύουν προγνωστικά μοντέλα εντός των δικών τους περιβαλλόντων, ενισχύοντας τη διατήρηση του απορρήτου [5].
3. **Πλήρης Ιχνηλασιμότητα και Ασφάλεια Πρόσβασης:**
   Ο μηχανισμός πρόσβασης μέσω VPN με πιστοποιητικά χρηστών και το audit logging με χρονική σφραγίδα εξασφαλίζουν τη διαφάνεια και την επαληθευσιμότητα της χρήσης των δεδομένων.
4. **Υψηλή Διαθεσιμότητα και Ευελιξία:**
   Η αρχιτεκτονική του συστήματος υποστηρίζει **Live Migration** των VM και **Zero Downtime** σε περιπτώσεις συντήρησης ή αναβάθμισης.
5. **Ενίσχυση Κλινικής Απόφασης και Εξατομίκευση Φροντίδας:**
   Μέσω της ασφαλούς πρόσβασης σε ιατρικά δεδομένα, τα εργαλεία του SAFEAORTA θα μπορούν να αναλύουν ιστορικά και δυναμικά δεδομένα ασθενών, υποστηρίζοντας την απόφαση για παρεμβάσεις ή παρακολούθηση.

1.3 Σύνδεση με τους Στόχους του Έργου SAFEAORTA

Το παρόν παραδοτέο αποτελεί **δομικό θεμέλιο** για την επίτευξη των τεσσάρων βασικών στόχων του έργου SAFEAORTA, καθώς η ανάπτυξη ομοσπονδιακών βάσεων δεδομένων και η υλοποίηση μηχανισμών ασφαλούς πρόσβασης αποτελούν κρίσιμα υποσυστήματα για τη λειτουργία της συνολικής αρχιτεκτονικής.

- **Στόχος 1: Υλοποίηση ατομικού ηλεκτρονικού φακέλου υγείας και ομοσπονδιακών βάσεων δεδομένων**

Η υλοποίηση **ομοσπονδιακής βάσης δεδομένων με χρήση MySQL Federated** επιτρέπει την αποθήκευση, ανάκτηση και συσχέτιση κλινικών, απεικονιστικών και βιολογικών δεδομένων χωρίς κεντρική συγκέντρωση. Κάθε κόμβος (π.χ. νοσοκομείο ή ερευνητικό ίδρυμα) διατηρεί την κυριότητα των δεδομένων, παρέχοντας επιλεκτική πρόσβαση μέσω ασφαλών μηχανισμών όπως **VPN και TLS**, σε πλήρη συμμόρφωση με το **άρθρο 25 του GDPR** περί ασφάλειας εξ ορισμού [1].

- **Στόχος 2: Συλλογή δεδομένων για καλύτερη πρόγνωση της νόσου**

Η ομοσπονδιακή αποθήκευση δεδομένων σε συνδυασμό με δυνατότητα **απομακρυσμένης ανάπτυξης και εκπαίδευσης μοντέλων τεχνητής νοημοσύνης (AI)** μέσω εργαλείων federated learning (π.χ. Flower, Fedbiomed) δημιουργεί τις προϋποθέσεις για παραγωγή **προγνωστικών μοντέλων υψηλής ακρίβειας**. Αυτά τα μοντέλα βασίζονται σε πραγματικά δεδομένα, με **ιδιωτικότητα διατηρημένη**, ενισχύοντας την **πρόγνωση της εξέλιξης του ανευρύσματος** και την εκτίμηση κινδύνου ρήξης [5],[6].

- **Στόχος 3: Ανάπτυξη Ψηφιακού Διδύμου Αορτής (ΨηφιΔΑ)**

Η συγκέντρωση ιατρικών δεδομένων υψηλής ακρίβειας σε ασφαλές και διαλειτουργικό περιβάλλον επιτρέπει τη δημιουργία και διαρκή ενημέρωση εξατομικευμένων **ψηφιακών διδύμων** (Digital Twins) για την αορτή κάθε ασθενούς. Η επεξεργασία απεικονιστικών δεδομένων και προσομοιώσεων σε **κρυπτογραφημένα VM** επιτρέπει **real-time επικαιροποίηση** του ΨηφιΔΑ χωρίς έξοδο δεδομένων από το τοπικό περιβάλλον [4] .

- **Στόχος 4: Δημιουργία μοντέλων πρόβλεψης κινδύνου ρήξης των ΑΚΑ**

Το παραδοτέο καθιστά δυνατή την **αποδοτική αξιοποίηση ιστορικών δεδομένων** από πολλαπλούς φορείς, σε προστατευμένο περιβάλλον, χωρίς τον κίνδυνο παραβίασης της ιδιωτικότητας. Η **υποδομή VxRail με vSAN, live migration και redundancy** διασφαλίζει απρόσκοπτη πρόσβαση σε δεδομένα και μοντέλα, καθιστώντας δυνατή τη συνεχή εκπαίδευση αλγορίθμων πρόβλεψης ρήξης, με **μηδενικό χρόνο διακοπής** και **υψηλή αξιοπιστία**.

2. Ασφάλεια Δεδομένων (Data Security)

Η προστασία των προσωπικών και ευαίσθητων δεδομένων υγείας αποτελεί ακρογωνιαίο λίθο για το έργο SAFEAORTA. Η στρατηγική ασφάλειας βασίζεται σε μια πολυεπίπεδη προσέγγιση που ενσωματώνει σύγχρονες τεχνολογίες, διεθνή πρότυπα και κανονιστικά πλαίσια, με γνώμονα τη συμμόρφωση με τον Γενικό Κανονισμό για την Προστασία Δεδομένων (GDPR), τα πρότυπα ISO/IEC 27001 και ENISA best practices.

Η υλοποιούμενη αρχιτεκτονική έχει σχεδιαστεί με βάση την αρχή της "ασφάλειας εκ σχεδιασμού και εξ ορισμού" (security by design and by default), όπως ορίζεται στο άρθρο 25 του GDPR. Τα τεχνικά μέτρα που ενσωματώνονται καλύπτουν τόσο την προστασία των δεδομένων κατά την αποθήκευση (data at rest) όσο και κατά τη διακίνηση (data in transit).

2.1 Γενικές Αρχές Ασφαλείας (CIA Principles)

Η αρχιτεκτονική του SAFEAORTA σχεδιάστηκε σύμφωνα με τις τρεις θεμελιώδεις αρχές ασφάλειας πληροφοριών:

- **Εμπιστευτικότητα (Confidentiality):** Η πρόσβαση στα δεδομένα περιορίζεται μόνο σε εξουσιοδοτημένους χρήστες μέσω πολυεπίπεδων μηχανισμών ταυτοποίησης (π.χ. VPN, 2FA) και κρυπτογράφησης TLS.
- **Ακεραιότητα (Integrity):** Χρησιμοποιούνται μηχανισμοί καταγραφής (audit logging), ψηφιακές υπογραφές, και TPM για την επαλήθευση της γνησιότητας των δεδομένων.
- **Διαθεσιμότητα (Availability):** Μέσω υποδομών με live migration, backup, και disaster recovery εξασφαλίζεται η αδιάλειπτη πρόσβαση στις υπηρεσίες.

2.2 Πολιτικές Ασφαλείας Υποδομής VxRail (Security by Design)

Η υποδομή VxRail βασίζεται σε security by design αρχιτεκτονική. Οι πολιτικές ασφάλειας περιλαμβάνουν:

- Υποχρεωτική χρήση κρυπτογράφησης σε όλα τα επίπεδα.
- Role-based access control (RBAC).
- Δυναμική κατανομή πόρων με έλεγχο πρόσβασης ανά επίπεδο εφαρμογής.
- Firewall σε επίπεδο hypervisor και δικτύου.

2.3 Προηγμένα Μέτρα και Πρωτόκολλα Ανάκτησης Δεδομένων

Για την αποφυγή απώλειας δεδομένων και την εξασφάλιση επιχειρησιακής συνέχειας:

- Υλοποιούνται ημερήσια αντίγραφα ασφαλείας με incremental και snapshot-based backup.
- Υποστήριξη geo-replication σε άλλο datacenter (όπου απαιτείται).
- Διαδικασίες disaster recovery με προκαθορισμένο RTO/RPO.

2.4 Κρυπτογράφηση Δεδομένων κατά την Αποθήκευση (Data at Rest Encryption)

Όλα τα δεδομένα που φιλοξενούνται εντός των **εικονικών μηχανών (VMs)** προστατεύονται με ισχυρή κρυπτογράφηση μέσω **AES-256** σε επίπεδο δίσκου, χρησιμοποιώντας εργαλεία όπως **LUKS (Linux Unified Key Setup)** ή **BitLocker**. Οι **κλειδώσεις ενεργοποιούνται κατά την εκκίνηση** και τα κλειδιά διαχειρίζονται μέσω **TPM chip** ή ασφαλών key vaults [3].

Οι δίσκοι του συστήματος αποθήκευσης τύπου **vSAN (Virtual SAN)** υποστηρίζουν native encryption, προσφέροντας πολλαπλά επίπεδα προστασίας σε κάθε επίπεδο αποθήκευσης (data blocks, metadata, caching).

2.5 Κρυπτογράφηση κατά τη Διακίνηση (Data in Transit Encryption)

Η επικοινωνία μεταξύ των χρηστών και των VMs πραγματοποιείται αποκλειστικά μέσω **HTTPS/TLS v1.3**, διασφαλίζοντας forward secrecy και ανθεκτικότητα σε επιθέσεις τύπου downgrade ή MITM (Man-In-The-Middle). Επιπλέον:

- Η σύνδεση των χρηστών επιτυγχάνεται μέσω **OpenVPN ή WireGuard**, με end-to-end κρυπτογράφηση (π.χ. ChaCha20 ή AES-256-GCM).
- Ο έλεγχος ταυτότητας γίνεται με χρήση **ψηφιακών πιστοποιητικών χρηστών (X.509)** και **διπλής αυθεντικοποίησης (2FA)**.

2.6 Μηχανισμοί Ελέγχου Πρόσβασης και Ασφαλούς Εκκίνησης

- **VPN Access Gateway:** Οι χρήστες συνδέονται αποκλειστικά μέσω VPN gateways με ισχυρή αυθεντικοποίηση, ελαχιστοποιώντας τα σημεία πρόσβασης.
- **UEFI Secure Boot:** Κάθε VM ελέγχεται κατά την εκκίνηση για μεταβολές στη δομή του λειτουργικού.
- **TPM 2.0 Chips:** Ενσωματωμένοι μικροελεγκτές στους φυσικούς κόμβους διασφαλίζουν τη διαχείριση κλειδιών και την ταυτότητα συστήματος.
- **Audit Logging:** Όλες οι ενέργειες πρόσβασης και μεταβολής δεδομένων καταγράφονται με χρονική σφραγίδα και ψηφιακή υπογραφή.

3. Τεχνική Υποδομή και Εικονικά Περιβάλλοντα

Η τεχνική υποδομή του SAFEAORTA βασίζεται σε σύγχρονες τεχνολογίες εικονικοποίησης και κατανεμημένων υποδομών αποθήκευσης, με στόχο την ασφαλή, αξιόπιστη και επεκτάσιμη διαχείριση των δεδομένων υγείας εντός ενός περιβάλλοντος συμμόρφωσης με τον GDPR. Η χρήση εικονικών μηχανών (VMs), VPNs, προηγμένων πρωτοκόλλων κρυπτογράφησης και federated MySQL βάσεων δεδομένων ενσωματώνεται με σκοπό την ασφαλή φιλοξενία, ανάκτηση και επεξεργασία ιατρικών δεδομένων.

3.1 Εικονικές Μηχανές (VMs): Τεχνικά Χαρακτηριστικά και Διαχείριση

Το σύστημα βασίζεται σε **υποδομή εικονικοποίησης μέσω Dell EMC VxRail**, η οποία παρέχει:

- **Hyperconverged Architecture**: Συνδυάζει υπολογιστική ισχύ, αποθηκευτικό χώρο και διαχείριση VM σε ενιαίο σύστημα.
- **vSphere/vCenter Integration**: Επιτρέπει granular έλεγχο πόρων, snapshots και policy-based διαχείριση των VM.
- **VM Templates**: Προκαθορισμένα images για ανάπτυξη ασφαλών περιβαλλόντων με προεγκατεστημένα εργαλεία (π.χ. Jupyter, PyTorch, Fedbiomed client).
- **Resource Pooling & Isolation**: Δυνατότητα ορισμού resources ανά project ή χρήστη, με role-based access.

Οι VMs παρέχουν υποστήριξη **nested virtualization**, διευκολύνοντας τον πειραματισμό με εργαλεία όπως Docker ή Singularity για ανάπτυξη μοντέλων AI.

3.2 Σύνδεση Χρηστών μέσω VPN και HTTPS (TLS v1.3)

Η πρόσβαση των εξουσιοδοτημένων χρηστών προς τις VMs επιτυγχάνεται **αποκλειστικά μέσω ασφαλούς VPN**, ενώ όλες οι συνδέσεις εντός των VMs χρησιμοποιούν πρωτόκολλο **HTTPS/TLS v1.3**, το οποίο:

- Προσφέρει ισχυρή ασφάλεια (forward secrecy, perfect forward secrecy).
- Αποτρέπει επιθέσεις downgrade, MITM και packet injection.
- Υποστηρίζεται natively από όλα τα σύγχρονα εργαλεία αναγνώρισης ταυτότητας και browsers.

Η υλοποίηση VPN βασίζεται είτε στο **OpenVPN με RSA/ECDSA authentication**, είτε στο **WireGuard με ChaCha20-Poly1305 encryption**.

3.3 Κρυπτογράφηση Εικονικών Μηχανών και vSAN Δίσκων (Data at Rest Encryption)

Η αποθήκευση των δεδομένων και των snapshots των VMs πραγματοποιείται σε **κρυπτογραφημένα vSAN arrays**, με χαρακτηριστικά όπως:

- **AES-256 full-disk encryption**, με hardware acceleration.

- **Key management** μέσω TPM modules και secure key vaults.
- **Immutable snapshots** για προστασία από ransomware ή ανθρώπινα λάθη.

Η κρυπτογράφηση είναι ενεργή και στα cache layers, εξαλείφοντας τρωτά σημεία.

3.4 UEFI Secure Boot και TPM στους Υπολογιστικούς Κόμβους

Κάθε φυσικός υπολογιστικός κόμβος είναι εξοπλισμένος με:

- **UEFI Secure Boot**, το οποίο επαληθεύει την ακεραιότητα του λειτουργικού κατά την εκκίνηση.
- **TPM 2.0 chips**, τα οποία διασφαλίζουν:
    - Την αποθήκευση και χρήση κλειδιών κρυπτογράφησης.
    - Την αναγνώριση του host.
    - Την προστασία του encryption context κατά τη live migration.

3.5 Live Migration Εικονικών Μηχανών με Μηδενικό Downtime

Η δυνατότητα **live migration** επιτρέπει τη μεταφορά μιας εικονικής μηχανής από έναν κόμβο σε άλλον χωρίς διακοπή λειτουργίας. Αυτή η λειτουργία είναι κρίσιμη για:

- **Zero-downtime μετακίνηση**, ακόμη και με ενεργή επεξεργασία δεδομένων.
- **Dynamic load balancing**, ώστε να βελτιστοποιείται η κατανομή των workloads.
- **Storage vMotion**, για μεταφορά δίσκων χωρίς επανεκκίνηση VM [7].

Το περιβάλλον υποστηρίζει live migration με χρήση shared storage (vSAN), ενώ ενεργοποιούνται μηχανισμοί state synchronization για την εξασφάλιση συνέχειας υπηρεσίας.

Η τεχνική αυτή καθιστά το SAFEAORTA υψηλής διαθεσιμότητας και αποδοτικό στην κατανομή των υπολογιστικών πόρων, ιδιαίτερα σε περιβάλλοντα με πολλαπλές κλινικές μονάδες και απαιτήσεις συνεχούς λειτουργίας.

3.6 Διεπαφή για Ανάπτυξη και Εκπαίδευση Μοντέλων

Για την υποστήριξη developers και την εκπαίδευση μοντέλων AI/Federated Learning, οι VMs προσφέρουν:

- Πρόσβαση σε προκαθορισμένα containers (π.χ. Docker με Flower/Fedbiomed).
- Προεγκατεστημένες βιβλιοθήκες AI (TensorFlow, PyTorch).
- Υποστήριξη Jupyter Lab και ασφαλών endpoints με περιορισμούς πόρων.

Τα δεδομένα παραμένουν τοπικά σε κάθε VM, επιτρέποντας την εκπαίδευση σε ομοσπονδιακό σχήμα χωρίς μεταφορά δεδομένων (privacy-preserving architecture).

3.7 Εναρμόνιση με Πρότυπα και Διεθνείς Κατευθυντήριες Γραμμές

Το τεχνικό υπόβαθρο της υποδομής είναι συμβατό με:

- **ISO/IEC 27001** (Information Security Management Systems).
- **NIST SP 800-53** (Security and Privacy Controls).
- **ISO/IEC 27701** (Privacy Information Management).

Η συμμόρφωση με αυτά τα πρότυπα διασφαλίζει διαλειτουργικότητα, επεκτασιμότητα και πιστοποιήσιμη ασφάλεια στο τελικό σύστημα.

Η συνολική τεχνική υποδομή του SAFEAORTA διαμορφώνει ένα πλήρως ασφαλές, απομονωμένο, επεκτάσιμο και ανακτήσιμο περιβάλλον, ικανό να στηρίξει απαιτητικά ιατρικά σενάρια σε διασυνδεδεμένα συστήματα πολλών κέντρων, με έμφαση στην εμπιστευτικότητα, την ακεραιότητα και τη διαθεσιμότητα των δεδομένων.

4. Υλοποίηση Ομοσπονδιακών Βάσεων Δεδομένων με MySQL Server

Για την υλοποίηση μιας ομοσπονδιακής βάσης δεδομένων με MySQL, θα χρησιμοποιήσουμε τον μηχανισμό αποθήκευσης FEDERATED. Αυτός ο μηχανισμός επιτρέπει σε έναν MySQL server να αντιμετωπίζει απομακρυσμένους πίνακες σε άλλους MySQL servers ως τοπικούς πίνακες. Η λειτουργικότητα έχει υλοποιηθεί και λειτουργεί σε Virtual Machines (VMs) της υπερυπολογιστικής υποδομής "Ερευνητική Υποδομή Αναλυτικής Ιατρικών Δεδομένων μεγάλου όγκου με στόχο την Ιατρική Ακριβείας, Κωδικός ΟΠΣ 5047133" της Μονάδας Ιατρικής Τεχνολογίας και Ευφυών Πληροφοριακών Συστημάτων του Πανεπιστημίου Ιωαννίνων. Επιπλέον στην ενότητα αυτή παρουσιάζουμε και εναλλακτική υλοποίηση με Postgres SQK server ο οποίος παρέχει επιπλέον πλεονεκτήματα για την υλοποίηση Ομοσπονδιακών βάσεων δεδομένων.

Η έλευση της κατανεμημένης πληροφορικής και ο πολλαπλασιασμός των ετερογενών πηγών δεδομένων έχουν αναδείξει την ανάγκη για συστήματα που μπορούν να ενσωματώνουν απρόσκοπτα και να υποβάλλουν ερωτήματα σε πληροφορίες που βρίσκονται σε πολλαπλές, ετερογενείς βάσεις δεδομένων. Αυτό ακριβώς το πρόβλημα φιλοδοξούν να λύσουν οι ομοσπονδιακές βάσεις δεδομένων (federated databases). Στην ουσία, ένα σύστημα ομοσπονδιακής βάσης δεδομένων (FDBS) είναι ένας τύπος συστήματος βάσης δεδομένων που παρέχει μια ενοποιημένη, εικονική προβολή πολλαπλών ανεξάρτητων και αυτόνομων βάσεων δεδομένων, επιτρέποντας στους χρήστες να υποβάλλουν ερωτήματα σε αυτές σαν να ήταν μια ενιαία, κεντρική βάση δεδομένων, χωρίς να απαιτείται φυσική ενσωμάτωση ή μετεγκατάσταση των υποκείμενων δεδομένων.

Η θεωρητική βάση των ομοσπονδιακών βάσεων δεδομένων έγκειται στην έννοια της ενσωμάτωσης δεδομένων χωρίς κεντρικοποίηση. Σε αντίθεση με τις παραδοσιακές κατανεμημένες βάσεις δεδομένων, όπου τα δεδομένα κατακερματίζονται και αντιγράφονται κάτω από ένα ενιαίο, παγκόσμια ελεγχόμενο σχήμα, τα ομοσπονδιακά συστήματα περιέχουν την αυτονομία των συνιστωσών τους βάσεων δεδομένων. Κάθε συνιστώσα βάση δεδομένων διατηρεί το δικό της τοπικό σχήμα, μοντέλο δεδομένων, επεξεργαστή ερωτημάτων και διαχείριση. Το FDBS λειτουργεί ως ένας έξυπνος ενδιάμεσος, μεταφράζοντας τα καθολικά ερωτήματα σε τοπικά ερωτήματα, αποστέλλοντάς τα στις κατάλληλες πηγές, συλλέγοντας τα αποτελέσματα και ενσωματώνοντάς τα σε μια συνεκτική απάντηση για τον χρήστη.

Τα βασικά Χαρακτηριστικά είναι:

- **Αυτονομία**: Οι βάσεις δεδομένων διατηρούν την ανεξαρτησία τους και τον έλεγχο των δεδομένων και των λειτουργιών τους. Αυτή είναι μια κρίσιμη διάκριση από τις παραδοσιακές κατανεμημένες βάσεις δεδομένων.
- **Ετερογένεια**: Τα FDBS έχουν σχεδιαστεί για να χειρίζονται διαφορετικά μοντέλα δεδομένων (σχεσιακά, αντικειμενοστραφή, NoSQL, XML, κ.λπ.), γλώσσες ερωτημάτων (SQL, OQL, XQuery, κ.λπ.), λειτουργικά συστήματα και πρωτόκολλα δικτύου.
- **Διανομή**: Τα δεδομένα είναι φυσικά διασκορπισμένα σε πολλαπλές, γεωγραφικά κατανεμημένες τοποθεσίες.
- **Διαφάνεια**: Στους χρήστες παρουσιάζεται ένα ενοποιημένο σχήμα και μηχανισμός αλληλεπίδρασης, προστατευμένοι σε μεγάλο βαθμό από τις πολυπλοκότητες της

διανομής δεδομένων και της ετερογένειας. Αυτό περιλαμβάνει τη διαφάνεια τοποθεσίας, πρόσβασης και σχήματος.

- **Εικονική Ενσωμάτωση**: Τα δεδομένα δεν ενοποιούνται φυσικά. Αντίθετα, δημιουργείται ένα εικονικό σχήμα που αντιστοιχεί στα υποκείμενα συστατικά σχήματα.

Οι αρχιτεκτονικές ομοσπονδιακών βάσεων δεδομένων συνήθως κατηγοριοποιούνται με βάση τον βαθμό αυτονομίας και τον έλεγχο που ασκείται από την ομοσπονδία. Δύο κύρια αρχιτεκτονικά μοντέλα συζητούνται συνήθως:

1. Χαλαρά Συνδεδεμένη Ομοσπονδία (Loosely Coupled Federation - Χωρίς Ενσωμάτωση σε Επίπεδο Σχήματος): Σε αυτό το μοντέλο, υπάρχει ελάχιστη ενσωμάτωση σχήματος. Κάθε τοπική βάση δεδομένων διατηρεί το δικό της σχήμα, και η ομοσπονδία παρέχει εργαλεία στους χρήστες για να διατυπώνουν ερωτήματα σε αυτά τα διαφορετικά σχήματα. Αυτό συχνά περιλαμβάνει τεχνικές όπως η δρομολόγηση ερωτημάτων και η συγχώνευση αποτελεσμάτων βάσει κοινών χαρακτηριστικών ή έμμεσων σχέσεων. Ενώ προσφέρει μέγιστη αυτονομία στις συστατικές βάσεις δεδομένων, επιβαρύνει περισσότερο τον χρήστη να κατανοήσει τα υποκείμενα σχήματα.
2. Σφιχτά Συνδεδεμένη Ομοσπονδία (Tightly Coupled Federation - Ενσωμάτωση Σχήματος): Αυτό είναι το πιο διαδεδομένο και πολύπλοκο μοντέλο. Περιλαμβάνει τη δημιουργία ενός καθολικού, ολοκληρωμένου σχήματος που παρέχει μια ενοποιημένη προβολή των δεδομένων σε όλες τις συμμετέχουσες βάσεις δεδομένων. Αυτό το καθολικό σχήμα είναι μια εννοιολογική αναπαράσταση και δεν αντιστοιχεί σε μια φυσική βάση δεδομένων. Η διαδικασία δημιουργίας αυτού του καθολικού σχήματος απαιτεί σημαντική προσπάθεια στην επίλυση συγκρούσεων σχήματος (π.χ., συγκρούσεις ονομάτων, αναντιστοιχίες τύπων δεδομένων, δομικές διαφορές) και σημασιολογικών ασυμφωνιών.

Ένα ευρέως υιοθετημένο θεωρητικό πλαίσιο για την κατανόηση της αρχιτεκτονικής των FDBS επεκτείνει την παραδοσιακή αρχιτεκτονική σχήματος τριών επιπέδων (εξωτερικό, εννοιολογικό, εσωτερικό) για να φιλοξενήσει την ετερογένεια και τη διανομή:

Τοπικά Σχήματα (Local Schemas - Εσωτερικά Σχήματα): Πρόκειται για τα σχήματα των μεμονωμένων, αυτόνομων συστατικών βάσεων δεδομένων, που περιγράφουν τη φυσική τους οργάνωση δεδομένων.

Σχήματα Συστατικών (Component Schemas - Εννοιολογικά Σχήματα): Αυτές είναι κανονικές αναπαραστάσεις των τοπικών σχημάτων, συχνά μεταφρασμένες σε ένα κοινό μοντέλο δεδομένων (π.χ., σχεσιακό) για να διευκολυνθεί η ενσωμάτωση. Οι wrappers είναι υπεύθυνοι για αυτήν την αντιστοίχιση.

Ομοσπονδιακό Σχήμα (Federated Schema - Καθολικό Εννοιολογικό Σχήμα): Αυτό είναι το ολοκληρωμένο σχήμα που παρέχει μια ενοποιημένη προβολή των δεδομένων σε όλα τα συστατικά. Επιλύει συγκρούσεις και ασυνέπειες μεταξύ των σχημάτων των συστατικών. Οι μεσολαβητές λειτουργούν σε αυτό το επίπεδο.

Εξωτερικά Σχήματα (External Schemas - Προβολές Χρηστών): Πρόκειται για προσαρμοσμένες προβολές του ομοσπονδιακού σχήματος προσαρμοσμένες σε συγκεκριμένες εφαρμογές χρηστών ή ομάδες χρηστών.

Οι ομοσπονδιακές βάσεις δεδομένων προσφέρουν επιτακτικά πλεονεκτήματα, ιδιαίτερα σε σενάρια όπου τα δεδομένα είναι εγγενώς κατανεμημένα και ελέγχονται από διαφορετικές οντότητες. Ωστόσο, η υλοποίηση και η διαχείρισή τους παρουσιάζουν επίσης σημαντικές θεωρητικές και πρακτικές προκλήσεις. Τα πλεονεκτήματα των Ομοσπονδιακών Βάσεων Δεδομένων είναι:

1. Διατήρηση Αυτονομίας: Αυτό είναι ίσως το σημαντικότερο πλεονέκτημα. Οι οργανισμοί μπορούν να διατηρήσουν τον πλήρη έλεγχο των τοπικών τους δεδομένων, των πολιτικών ασφαλείας και των λειτουργικών διαδικασιών, χωρίς να αναγκάζονται σε ένα κεντρικοποιημένο, ομοιογενές σύστημα. Αυτό είναι ζωτικής σημασίας για νομικούς, διοικητικούς και πολιτικούς λόγους.

2. Ενσωμάτωση Δεδομένων χωρίς Μεταφορά: Τα FDBS αποφεύγουν τη δαπανηρή και πολύπλοκη διαδικασία φυσικής μεταφοράς και ενοποίησης δεδομένων από διαφορετικές πηγές σε μια ενιαία αποθήκη δεδομένων (data warehouse). Αυτό μειώνει τον χρόνο ανάπτυξης, εξαλείφει την πλεονασμό δεδομένων και διασφαλίζει ότι τα δεδομένα παραμένουν ενημερωμένα στην πηγή τους.

3. Υποστήριξη Ετερογένειας: Έχουν σχεδιαστεί ειδικά για να γεφυρώνουν τα κενά μεταξύ διαφορετικών μοντέλων δεδομένων, γλωσσών ερωτημάτων και πλατφορμών, καθιστώντας τα εξαιρετικά προσαρμόσιμα στις υπάρχουσες υποδομές πληροφορικής.

4. Κλιμακωσιμότητα: Νέες πηγές δεδομένων μπορούν να προστεθούν στην ομοσπονδία με σχετικά λιγότερη αναστάτωση σε σύγκριση με τα παραδοσιακά κεντρικά συστήματα, καθώς χρειάζονται μόνο ενημερώσεις στις ρυθμίσεις των wrappers και των mediators.

4.1 Τεχνικό Σενάριο Υλοποίησης με την υποδομή Precious

- **Κλινικό Κέντρο (Clinical Center):** IP: 192.168.201.109
  - Σε αυτόν τον κόμβο είναι εγκατεστημένος ο κύριος MySQL Server με τα κλινικά δεδομένα.
  - Θεωρούμε ότι η βάση δεδομένων clinical_db έχει μεταξύ άλλων πίνακες με ιατρικά δεδομένα (έναν πίνακα patients και medical_records).
- **Ομοσπονδιακή Τοποθεσία (Federated Site):** IP: 192.168.201.186
  - Σε αυτόν τον κόμβο είναι εγκατεστημένος ένας άλλος MySQL Server, ο οποίος θα λειτουργήσει ως ο **ομοσπονδιακός κόμβος**.
  - Οι ερευνητές θα συνδέονται σε αυτόν τον διακομιστή (server) για να επεξεργάζονται τα δεδομένα.

4.1.1 Ρυθμίσεις στο Κλινικό Κέντρο (192.168.201.109)

4.1.1.1 Δημιουργία Βάσης Δεδομένων και Πινάκων:

Στον MySQL Server του κλινικού κέντρου, δημιουργήστε τη βάση δεδομένων και τους πίνακες που περιέχουν τα κλινικά δεδομένα.

SQL

```
-- Σύνδεση στον MySQL Server στο 192.168.201.109
-- mysql -u root -p

CREATE DATABASE IF NOT EXISTS clinical_db;
USE clinical_db;
```

```
CREATE TABLE IF NOT EXISTS patients (
    patient_id INT PRIMARY KEY AUTO_INCREMENT,
    first_name VARCHAR(100),
    last_name VARCHAR(100),
    date_of_birth DATE,
    gender ENUM('Male', 'Female', 'Other'),
    -- Περισσότερα προσωπικά δεδομένα
    INDEX (last_name)
);

CREATE TABLE IF NOT EXISTS medical_records (
    record_id INT PRIMARY KEY AUTO_INCREMENT,
    patient_id INT,
    visit_date DATE,
    diagnosis TEXT,
    treatment TEXT,
    FOREIGN KEY (patient_id) REFERENCES patients(patient_id)
);

-- Εισαγωγή ενδεικτικών δεδομένων
INSERT INTO patients (first_name, last_name, date_of_birth, gender) VALUES
('John', 'Test', '1980-05-15', 'Male'),
('Jane', 'Testit', '1992-11-23', 'Female');

INSERT INTO medical_records (patient_id, visit_date, diagnosis, treatment) VALUES
(1, '2023-01-10', 'Common cold', 'Rest and fluids'),
(2, '2023-03-05', 'Allergic reaction', 'Antihistamines');

FLUSH PRIVILEGES;
```

3.1.1.2 Δημιουργία Χρήστη και Παραχώρηση Δικαιωμάτων:

Δημιουργήστε έναν ειδικό χρήστη που θα χρησιμοποιείται από τον ομοσπονδιακό κόμβο για την πρόσβαση στα δεδομένα. Είναι κρίσιμο να περιορίσετε τα δικαιώματα μόνο στην ανάγνωση (ή σε συγκεκριμένες λειτουργίες) και από συγκεκριμένη IP.

SQL

```
-- Δημιουργία χρήστη για το federated site
CREATE USER 'federated_user'@'192.168.201.186' IDENTIFIED BY 'your_secure_password';

-- Παραχώρηση δικαιωμάτων READ ONLY (SELECT) στους πίνακες
GRANT SELECT ON clinical_db.patients TO 'federated_user'@'192.168.201.186';
GRANT SELECT ON clinical_db.medical_records TO 'federated_user'@'192.168.201.186';

-- Εάν χρειάζονται και άλλες λειτουργίες (π.χ., INSERT, UPDATE, DELETE), παραχωρήστε ανάλογα, αλλά με προσοχή.
-- GRANT SELECT, INSERT, UPDATE, DELETE ON clinical_db.* TO
```

```
'federated_user'@'192.168.201.186';

FLUSH PRIVILEGES;
```

3.1.1.3 Ενεργοποίηση Εξωτερικής Πρόσβασης (εάν χρειάζεται):

Βεβαιωθείτε ότι ο MySQL Server στο κλινικό κέντρο επιτρέπει συνδέσεις από το δίκτυο. Ελέγξτε το αρχείο ρυθμίσεων (my.cnf ή my.ini) και σχολιάστε ή αλλάξτε το bind-address σε 0.0.0.0 ή στην IP του server, αν δεν είναι ήδη έτσι. Προσοχή: Αυτό αυξάνει την έκθεση. Βεβαιωθείτε ότι υπάρχει firewall που επιτρέπει την πρόσβαση μόνο από την IP του ομοσπονδιακού κόμβου (192.168.201.186) στη θύρα του MySQL (συνήθως 3306).

```
# my.cnf (ή my.ini) στο 192.168.201.109
# bind-address = 127.0.0.1  <-- Σχολιάστε το ή αλλάξτε το σε 0.0.0.0 ή την IP του server
bind-address = 0.0.0.0
```

Μετά την αλλαγή, πρέπει να γίνει επανεκκίνηση του MySQL service.

4.1.2 Ρυθμίσεις στην Ομοσπονδιακή Τοποθεσία (192.168.201.186)

4.1.2.1 Ενεργοποίηση του FEDERATED Storage Engine:

Στον MySQL Server της ομοσπονδιακής τοποθεσίας, πρέπει να βεβαιωθείτε ότι ο μηχανισμός αποθήκευσης FEDERATED είναι ενεργοποιημένος. Συνήθως είναι ενεργοποιημένος από προεπιλογή, αλλά μπορείτε να το ελέγξετε και να τον ενεργοποιήσετε αν χρειάζεται.

SQL

```
-- Σύνδεση στον MySQL Server στο 192.168.201.186
-- mysql -u root -p

SHOW ENGINES; -- Ελέγξτε αν το Federated είναι 'YES' στο Support

-- Αν δεν είναι, προσθέστε την παρακάτω γραμμή στο my.cnf (ή my.ini)
-- και κάντε επανεκκίνηση του MySQL server
-- [mysqld]
-- federated
```

3.1.2.2 Δημιουργία Ομοσπονδιακών Πινάκων:

Τώρα, δημιουργήστε τους ομοσπονδιακούς πίνακες που θα αναφέρονται στους απομακρυσμένους πίνακες στο κλινικό κέντρο. Η δομή των πινάκων πρέπει να είναι ακριβώς ίδια με αυτή των απομακρυσμένων πινάκων.

SQL

```
-- Σύνδεση στον MySQL Server στο 192.168.201.186
-- mysql -u root -p

CREATE DATABASE IF NOT EXISTS research_data;
```

```
USE research_data;

-- Ομοσπονδιακός πίνακας για τους ασθενείς
CREATE TABLE federated_patients (
    patient_id INT PRIMARY KEY AUTO_INCREMENT,
    first_name VARCHAR(100),
    last_name VARCHAR(100),
    date_of_birth DATE,
    gender ENUM('Male', 'Female', 'Other'),
    INDEX (last_name)
)
ENGINE=FEDERATED
CONNECTION='mysql://federated_user:your_secure_password@192.168.201.109:3306/clinical_db/patients';

-- Ομοσπονδιακός πίνακας για τα ιατρικά αρχεία
CREATE TABLE federated_medical_records (
    record_id INT PRIMARY KEY AUTO_INCREMENT,
    patient_id INT,
    visit_date DATE,
    diagnosis TEXT,
    treatment TEXT
)
ENGINE=FEDERATED
CONNECTION='mysql://federated_user:your_secure_password@192.168.201.109:3306/clinical_db/medical_records';
```

Σημείωση για CONNECTION string:

mysql://[user_name]:[password]@[host_name][:port_num]/[db_name]/[table_name]

4.2 Έλεγχος και Χρήση

Πλέον, οι ερευνητές που συνδέονται στον MySQL Server στο 192.168.201.186 μπορούν να εκτελούν ερωτήματα στους πίνακες federated_patients και federated_medical_records σαν να ήταν τοπικοί πίνακες. Τα ερωτήματα αυτά θα δρομολογούνται αυτόματα στον MySQL Server του κλινικού κέντρου.

SQL

```
-- Σύνδεση στον MySQL Server στο 192.168.201.186
-- mysql -u research_user -p

USE research_data;

SELECT * FROM federated_patients;

SELECT p.first_name, p.last_name, mr.diagnosis, mr.visit_date
FROM federated_patients p
```

```
JOIN federated_medical_records mr ON p.patient_id = mr.patient_id
WHERE p.last_name = 'User';
```

4.2.1 Ανάλυση Υλοποίησης με MySQL FEDERATED:

Πλεονεκτήματα:

- **Απλότητα:** Η ρύθμιση είναι σχετικά απλή για βασική ανάγνωση δεδομένων από έναν απομακρυσμένο MySQL server.
- **Εγγενής Υποστήριξη:** Ο μηχανισμός FEDERATED είναι ενσωματωμένος στον MySQL, χωρίς την ανάγκη για επιπλέον εργαλεία ή λογισμικό.
- **Διατήρηση Αυτονομίας:** Το κλινικό κέντρο διατηρεί τον πλήρη έλεγχο των δεδομένων του, του σχήματος και της ασφάλειας.
- **Μη Μεταφορά Δεδομένων:** Δεν υπάρχει φυσική μετακίνηση δεδομένων, διασφαλίζοντας ότι οι ερευνητές έχουν πάντα πρόσβαση στα πιο πρόσφατα δεδομένα.

Μειονεκτήματα/Προκλήσεις:

- **Περιορισμένη Λειτουργικότητα (Read-Only συνήθως):** Αν και τεχνικά υποστηρίζει INSERT/UPDATE/DELETE, ο μηχανισμός FEDERATED λειτουργεί καλύτερα για read-only σενάρια. Οι εγγραφές είναι λιγότερο αποδοτικές και μπορεί να έχουν προβλήματα με συναλλαγές. Για ευαίσθητα κλινικά δεδομένα, η ανάγνωση μόνο είναι συχνά επιθυμητή.
- **Έλλειψη Βελτιστοποίησης Ερωτημάτων:** Ο ομοσπονδιακός κόμβος δεν έχει πλήρη γνώση των στατιστικών του απομακρυσμένου server, γεγονός που μπορεί να οδηγήσει σε υποβέλτιστη εκτέλεση ερωτημάτων, ειδικά για σύνθετα joins ή aggregation.
- **Προβλήματα Απόδοσης:** Κάθε ερώτημα που εκτελείται στον ομοσπονδιακό πίνακα μεταφράζεται σε ένα ερώτημα που αποστέλλεται στον απομακρυσμένο server. Αυτό εισάγει καθυστέρηση δικτύου και μπορεί να επιβαρύνει τον απομακρυσμένο server.
- **Απαιτεί MySQL σε όλες τις πλευρές:** Το FEDERATED engine λειτουργεί μόνο με άλλους MySQL servers. Δεν μπορεί να συνδεθεί σε PostgreSQL, SQL Server, κ.λπ.
- **Διαχείριση Σφάλματος Δικτύου:** Η διαχείριση σφαλμάτων σύνδεσης ή διακοπών του απομακρυσμένου server δεν είναι ιδιαίτερα εξελιγμένη.
- **Ασφάλεια:** Η κωδικοποίηση του password στην CONNECTION string είναι ένα θέμα ασφαλείας. Θα πρέπει να χρησιμοποιηθούν προσεκτικά τα δικαιώματα χρήστη και να διασφαλιστεί η ασφάλεια του δικτύου (π.χ., VPN, firewalls).

4.3 Ανάλυση Υλοποίησης με PostgreSQL Database Server

Η PostgreSQL δεν διαθέτει έναν εγγενή μηχανισμό όπως ο FEDERATED του MySQL. Αντίθετα, χρησιμοποιεί την έννοια των **Foreign Data Wrappers (FDW)**. Οι FDWs είναι επεκτάσεις (plugons) του PostgreSQL που του επιτρέπουν να συνδέεται με εξωτερικές πηγές δεδομένων (όχι απαραίτητα PostgreSQL) και να τις εκθέτει ως πίνακες στο PostgreSQL.

Σενάριο με PostgreSQL:

- **Κλινικό Κέντρο (Clinical Center):** IP: 192.168.201.109
  - **PostgreSQL Server** με τη βάση δεδομένων clinical_db και πίνακες patients, medical_records.
- **Ομοσπονδιακή Τοποθεσία (Federated Site):** IP: 192.168.201.186

- **PostgreSQL Server**, ο οποίος θα λειτουργήσει ως ο ομοσπονδιακός κόμβος.

4.3.1. Ρυθμίσεις στο Κλινικό Κέντρο (192.168.201.109 - PostgreSQL)

4.3.1.1 Δημιουργία Βάσης Δεδομένων και Πινάκων:

Στον PostgreSQL Server του κλινικού κέντρου, δημιουργήστε τη βάση δεδομένων και τους πίνακες.

SQL

```
-- Σύνδεση στον PostgreSQL Server στο 192.168.201.109
-- psql -U postgres -h 192.168.201.109

CREATE DATABASE clinical_db;
\c clinical_db;

CREATE TABLE patients (
    patient_id SERIAL PRIMARY KEY,
    first_name VARCHAR(100),
    last_name VARCHAR(100),
    date_of_birth DATE,
    gender VARCHAR(10), -- PostgreSQL uses TEXT/VARCHAR, not ENUM like MySQL
    UNIQUE (patient_id)
);

CREATE TABLE medical_records (
    record_id SERIAL PRIMARY KEY,
    patient_id INT,
    visit_date DATE,
    diagnosis TEXT,
    treatment TEXT,
    FOREIGN KEY (patient_id) REFERENCES patients(patient_id)
);

INSERT INTO patients (first_name, last_name, date_of_birth, gender) VALUES
('John', Test, '1980-05-15', 'Male'),
('Jane', 'Testit', '1992-11-23', 'Female');

INSERT INTO medical_records (patient_id, visit_date, diagnosis, treatment) VALUES
(1, '2023-01-10', 'Common cold', 'Rest and fluids'),
(2, '2023-03-05', 'Allergic reaction', 'Antihistamines');
```

4.3.1.2 Δημιουργία Χρήστη και Παραχώρηση Δικαιωμάτων:

Δημιουργήστε έναν χρήστη για τον ομοσπονδιακό κόμβο.

SQL

```
CREATE USER federated_user WITH PASSWORD 'your_secure_password';
GRANT SELECT ON patients TO federated_user;
GRANT SELECT ON medical_records TO federated_user;
-- GRANT ALL PRIVILEGES ON ALL TABLES IN SCHEMA public TO federated_user; -- Προσοχή με αυτό
```

4.3.1.3 Ενεργοποίηση Εξωτερικής Πρόσβασης:

Επεξεργαστείτε το αρχείο postgresql.conf και pg_hba.conf στο κλινικό κέντρο.
Στο postgresql.conf:
listen_addresses = '*' (ή 192.168.201.109)
Στο pg_hba.conf (προσθέστε μια γραμμή για την IP του federated site):
host clinical_db federated_user 192.168.201.186/32 md5
Επανεκκινήστε τον PostgreSQL service.

4.2. Ρυθμίσεις στην Ομοσπονδιακή Τοποθεσία (192.168.201.186 - PostgreSQL)

3.2.2.1 Εγκατάσταση και Ενεργοποίηση FDW Extension:

Για να συνδεθείτε σε έναν απομακρυσμένο PostgreSQL server, θα χρησιμοποιήσετε το postgresql_fdw. Αν συνδέεστε σε MySQL, θα χρησιμοποιούσατε ένα FDW όπως το mysql_fdw (το οποίο χρειάζεται να εγκατασταθεί ξεχωριστά).

SQL

```
-- Σύνδεση στον PostgreSQL Server στο 192.168.201.186
-- psql -U postgres -h 192.168.201.186

CREATE EXTENSION IF NOT EXISTS postgres_fdw; -- Ή mysql_fdw αν η πηγή ήταν MySQL
```

4.2.2.2 Δημιουργία Foreign Server:

Καθορίστε τον απομακρυσμένο server.

SQL

```
CREATE SERVER clinical_data_server
FOREIGN DATA WRAPPER postgres_fdw -- Ή mysql_fdw
OPTIONS (host '192.168.201.109', port '5432', dbname 'clinical_db'); -- Port MySQL 3306
```

4.2.2.3 Δημιουργία User Mapping:

Καθορίστε ποιος χρήστης στον ομοσπονδιακό κόμβο θα χρησιμοποιεί ποιους credentials για να συνδεθεί στον απομακρυσμένο server.

SQL

```
CREATE USER MAPPING FOR current_user -- Ή για συγκεκριμένο χρήστη π.χ. research_user
SERVER clinical_data_server
```

```
OPTIONS (user 'federated_user', password 'your_secure_password');
```

4.2.2.4 Δημιουργία Foreign Tables:

Δημιουργήστε τους πίνακες που θα αναφέρονται στους απομακρυσμένους πίνακες.

SQL

```
CREATE FOREIGN TABLE federated_patients (
    patient_id INT NOT NULL,
    first_name VARCHAR(100),
    last_name VARCHAR(100),
    date_of_birth DATE,
    gender VARCHAR(10)
)
SERVER clinical_data_server
OPTIONS (table_name 'patients'); -- Το όνομα του πίνακα στον απομακρυσμένο server

CREATE FOREIGN TABLE federated_medical_records (
    record_id INT NOT NULL,
    patient_id INT,
    visit_date DATE,
    diagnosis TEXT,
    treatment TEXT
)
SERVER clinical_data_server
OPTIONS (table_name 'medical_records');
```

4.3. Έλεγχος και Χρήση

Οι ερευνητές μπορούν πλέον να υποβάλλουν ερωτήματα στους foreign tables σαν να ήταν τοπικοί.

SQL

```
-- Σύνδεση στον PostgreSQL Server στο 192.168.201.186
-- psql -U research_user -h 192.168.201.186

SELECT * FROM federated_patients;

SELECT p.first_name, p.last_name, mr.diagnosis, mr.visit_date
FROM federated_patients p
JOIN federated_medical_records mr ON p.patient_id = mr.patient_id
WHERE p.last_name = 'User';
```

Ανάλυση Υλοποίησης με PostgreSQL FDW:

Πλεονεκτήματα:

- **Ευελιξία και Ετερογένεια:** Οι FDWs είναι πολύ πιο ευέλικτοι. Υπάρχουν FDWs για διάφορες πηγές δεδομένων (MySQL, Oracle, SQL Server, NoSQL databases, αρχεία CSV, ακόμη και APIs), επιτρέποντας την ενσωμάτωση δεδομένων από **διαφορετικά συστήματα βάσεων δεδομένων**.
- **Πιο Έξυπνη Βελτιστοποίηση Ερωτημάτων:** Το PostgreSQL είναι γνωστό για τον προηγμένο βελτιστοποιητή ερωτημάτων του. Οι FDWs μπορούν να αξιοποιήσουν τις δυνατότητες pushdown (ώθηση λειτουργιών στον απομακρυσμένο server), όπως φιλτράρισμα (WHERE clauses), joins και aggregation, μειώνοντας την ποσότητα των δεδομένων που μεταφέρονται και βελτιώνοντας την απόδοση.
- **Πιο Ισχυρό Transaction Management:** Το PostgreSQL έχει πιο ώριμη υποστήριξη συναλλαγών από το FEDERATED του MySQL.
- **Κοινότητα και Επεκτασιμότητα:** Το οικοσύστημα των FDWs είναι ενεργό και επεκτάσιμο, με συνεχή ανάπτυξη νέων wrappers.

Μειονεκτήματα/Προκλήσεις:

- **Πιο Πολύπλοκη Ρύθμιση:** Η ρύθμιση των FDWs είναι λίγο πιο περίπλοκη από το FEDERATED, καθώς απαιτεί την εγκατάσταση επεκτάσεων και τη διαμόρφωση περισσότερων αντικειμένων (server, user mapping, foreign tables).
- **Εξάρτηση από FDW:** Η λειτουργικότητα και η απόδοση εξαρτώνται από την ποιότητα του συγκεκριμένου FDW που χρησιμοποιείται.
- **Overhead:** Παρόλο που οι FDWs είναι πιο έξυπνοι, εξακολουθεί να υπάρχει επιπλέον κόστος δικτύου και επεξεργασίας για την πρόσβαση σε απομακρυσμένες πηγές.
- **Συμβατότητα:** Ενώ οι FDWs προσφέρουν ετερογένεια, η πλήρης συμβατότητα σε χαρακτηριστικά (όπως τύποι δεδομένων, συναρτήσεις) μεταξύ των συστημάτων μπορεί να απαιτεί επιπλέον μετασχηματισμούς.
- **Ασφάλεια Κωδικού:** Όπως και με το MySQL FEDERATED, η αποθήκευση του password στο USER MAPPING απαιτεί προσοχή και ασφαλείς πρακτικές δικτύου.

5. Συμμόρφωση με τις τεχνικές προδιαγραφές

Το παρόν κεφάλαιο τεκμηριώνει με σαφήνεια τον τρόπο με τον οποίο η υλοποιηθείσα υποδομή ανταποκρίνεται στις απαιτήσεις της Ενότητας 1 του Πίνακα Συμμόρφωσης που συνοδεύει την υποβληθείσα Τεχνική Προσφορά. Η συμμόρφωση επιτυγχάνεται τόσο σε επίπεδο υποδομής όσο και λειτουργίας, καλύπτοντας πλήρως τις προβλεπόμενες τεχνικές προδιαγραφές.

6.1 Ανάπτυξη Ομοσπονδιακών Βάσεων Δεδομένων με MySQL FEDERATED

Η υλοποίηση περιλαμβάνει την εγκατάσταση και παραμετροποίηση δύο διασυνδεδεμένων βάσεων MySQL, στις IP διευθύνσεις 192.168.201.109 (κλινικός κόμβος) και 192.168.201.186 (ομοσπονδιακή τοποθεσία). Ο μηχανισμός FEDERATED επιτρέπει την προσπέλαση απομακρυσμένων δεδομένων χωρίς να πραγματοποιείται τοπική αποθήκευση, διατηρώντας τη διανομή των δεδομένων στον πρωτεύοντα κόμβο.

6.2 Ασφαλής Κρυπτογραφημένη Πρόσβαση μέσω HTTPS

Οι εικονικές μηχανές που υλοποιήθηκαν στο πλαίσιο της υποδομής είναι προσβάσιμες αποκλειστικά μέσω κρυπτογραφημένης σύνδεσης HTTPS, με χρήση του πρωτοκόλλου TLS v1.3. Η πρόσβαση δίνεται σε εξουσιοδοτημένους χρήστες με διαπιστευτήρια μέσω web-based client ή SSH over VPN, διασφαλίζοντας την ασφάλεια και την αυθεντικοποίηση.

6.3 Σύνδεση μέσω VPN και Ασφαλές Virtual Network

Η πρόσβαση στις VMs επιτυγχάνεται μέσω VPN, το οποίο έχει ρυθμιστεί να λειτουργεί με σύγχρονη κρυπτογράφηση TLS v1.3. Το εικονικό δίκτυο είναι διαχωρισμένο από το δημόσιο δίκτυο μέσω εσωτερικού virtual switch εντός του συστήματος VxRail, παρέχοντας πλήρη απομόνωση του δικτύου των εικονικών μηχανών.

6.4 Κρυπτογράφηση των Δεδομένων σε vSAN

Ο αποθηκευτικός χώρος που χρησιμοποιείται βασίζεται σε vSAN συστοιχία δίσκων με ενεργοποιημένη κρυπτογράφηση AES-256. Επιπλέον, κάθε εικονική μηχανή προστατεύεται με native disk encryption, είτε σε επίπεδο λειτουργικού συστήματος είτε μέσω hypervisor policy.

6.5 Προστασία των Δεδομένων από τον Σχεδιασμό (Security by Design)

Η υλοποίηση βασίστηκε στην αρχή “Security by Design” όπως προβλέπεται στο άρθρο 25 του GDPR. Εφαρμόστηκαν μηχανισμοί περιορισμένης πρόσβασης, role-based policies, καταγραφή ενεργειών χρηστών (audit logging), καθώς και ελάχιστη συλλογή και προβολή προσωπικών δεδομένων.

6.6 Κρυπτογράφηση Δεδομένων σε Κατάσταση Ηρεμίας (Data at Rest)

Όλα τα δεδομένα που βρίσκονται εντός των εικονικών μηχανών είναι κρυπτογραφημένα, τόσο στον δίσκο όσο και στα backups. Χρησιμοποιούνται native encryption εργαλεία ανάλογα με το λειτουργικό σύστημα (π.χ. BitLocker, LUKS) σε συνδυασμό με την πολιτική vSAN encryption.

6.7 Ασφαλής Εκκίνηση με UEFI Secure Boot

Όλες οι εικονικές μηχανές είναι ρυθμισμένες να εκκινούν αποκλειστικά μέσω UEFI με ενεργοποιημένη την επιλογή Secure Boot. Έτσι, αποτρέπεται η εκκίνηση μη εξουσιοδοτημένου ή παραποιημένου λογισμικού κατά την εκκίνηση.

6.8 Ενσωμάτωση TPM σε Υπολογιστικούς Κόμβους

Οι κόμβοι που φιλοξενούν τις εικονικές μηχανές είναι εξοπλισμένοι με TPM 2.0 microcontroller, ο οποίος επιτρέπει την αποθήκευση κλειδιών κρυπτογράφησης και την πιστοποίηση του υλικού με ασφάλεια.

6.9 Live Migration με Μηδενικό Downtime

Ο σχεδιασμός της υποδομής υποστηρίζει πλήρως τη δυνατότητα live migration εικονικών μηχανών μεταξύ υπολογιστικών κόμβων χωρίς καμία διακοπή λειτουργίας. Αυτό διασφαλίζει τη διαθεσιμότητα των υπηρεσιών σε περιπτώσεις συντήρησης ή αποτυχίας κόμβου.

6.10 Υποστήριξη Federated Learning

Οι εικονικές μηχανές διαθέτουν εργαλεία για την ανάπτυξη και εκπαίδευση μοντέλων τεχνητής νοημοσύνης, με υποστήριξη για βιβλιοθήκες **Flower** και **Fedbiomed**. Οι χρήστες έχουν τη δυνατότητα να εκτελούν πειράματα federated learning με ασφάλεια, αξιοποιώντας τη διανεμημένη φύση των δεδομένων.

6. Συμπεράσματα

Και οι δύο προσεγγίσεις παρέχουν έναν τρόπο για να έχουν οι ερευνητές πρόσβαση στα κλινικά δεδομένα χωρίς να χρειάζεται να μεταφερθούν φυσικά.

- Για ένα **αμιγώς MySQL περιβάλλον** και κυρίως για **read-only** σενάρια, το **MySQL FEDERATED engine** είναι μια γρήγορη και σχετικά απλή λύση.
- Για πιο **σύνθετα σενάρια, ετερογενή περιβάλλοντα (π.χ., MySQL στο κλινικό κέντρο και PostgreSQL στο federated site), ή καλύτερη βελτιστοποίηση ερωτημάτων**, οι **PostgreSQL Foreign Data Wrappers** προσφέρουν σημαντικά μεγαλύτερη ευελιξία και δυνατότητες.

Στην περίπτωση που το κλινικό κέντρο έχει ήδη MySQL Server, και χρειάζεται μόνο ο ομοσπονδιακός κόμβος να είναι PostgreSQL, θα χρησιμοποιηθεί το mysql_fdw στο PostgreSQL του ομοσπονδιακού κόμβου για την σύνδεση στον MySQL του κλινικού κέντρου. Συμπερασματικά, οι ομοσπονδιακές βάσεις δεδομένων αντιπροσωπεύουν ένα ισχυρό παράδειγμα για την ενσωμάτωση ετερογενών και αυτόνομων πηγών δεδομένων, προσφέροντας σημαντικά πλεονεκτήματα στην προσβασιμότητα δεδομένων και τη διατήρηση της αυτονομίας.

Ωστόσο, οι θεωρητικές προκλήσεις που σχετίζονται με την ενσωμάτωση σχήματος, τη βελτιστοποίηση ερωτημάτων και τη διαχείριση συναλλαγών παραμένουν περίπλοκες και αποτελούν αντικείμενα συνεχούς έρευνας και ανάπτυξης στον τομέα των κατανεμημένων και ετερογενών συστημάτων βάσεων δεδομένων.

7. Βιβλιογραφία

1. Regulation (EU) 2016/679 (GDPR), Articles 25 & 32

2. European Union Agency for Cybersecurity (ENISA). “Security and privacy considerations in federated learning.” 2021

3. ISO/IEC 27001:2022 – Information Security, Cybersecurity and Privacy Protection

4. Al-Rubaie, M., & Chang, J. M. (2021). "Privacy-Preserving Machine Learning." IEEE Security & Privacy, 19(3)

5. Bonawitz, K. et al. (2022). “Towards Federated Learning at Scale.” *MLSys 2022*

6. Rieke, N. et al. (2020). "The future of digital health with federated learning." NPJ Digital Medicine, 3(1), 119

7. VMware. “vSphere vMotion Architecture.” White Paper, 2023