

Ψηφιακές υποδομές για την κοινότητα ELIXIR-GR



Θανάσης Βεργούλης
ΙΠΣΥ, Ε.Κ. «ΑΘΗΝΑ»

Ταυτότητα της κοινότητας ELIXIR-GR

- **ELIXIR: διακυβερνητικός οργανισμός για την οργάνωση, ολοκλήρωση και συντήρηση των πόρων βιοπληροφορικής στην Ευρώπη**
 - **Ίδρυση: 2014** (εν μέρει χρηματοδότηση από ΕΕ, RI – Horizon2020)
 - **23 μέλη-κόμβοι**
 - Κεντρικός κόμβος
 - 22 χώρες-μέλη (Ελλάδα επίσημα από 9/2018)
 - 1 χώρα-παρατηρητής (Κύπρος)
 - **180+ ερευνητικοί οργανισμοί**
 - Συντονιστής: EBI του EMBL
 - **Τύποι πόρων:** σύνολα δεδομένων, εργαλεία, υπολογιστικές υπηρεσίες, πρότυπα διαλειτουργικότητας, εκπαιδευτικές δραστηριότητες & υλικό



Το έργο ELIXIR-GR

- **Θέμα:** Αναβάθμιση & ολοκλήρωση της ελληνικής υποδομής για τη διαχείριση & ανάλυση δεδομένων στις Βιοεπιστήμες.
- **Έναρξη:** 1/11/2017
- **Διάρκεια:** 36 μήνες
- **Χρηματοδότηση:** Επιχειρησιακό Πρόγραμμα «Ανταγωνιστικότητα Επιχειρηματικότητα και Καινοτομία 2014-2020» (ΕΥΔ-ΕΠΑΝΕΚ)
- Συμμετέχουν **15 οργανισμοί**
 - Ανάμεσά τους το ΕΚ Αθηνά & ο ΕΔΕΤ
 - Συντονισμός ΕΚΕΒΕ «Αλ. Φλέμινγκ»

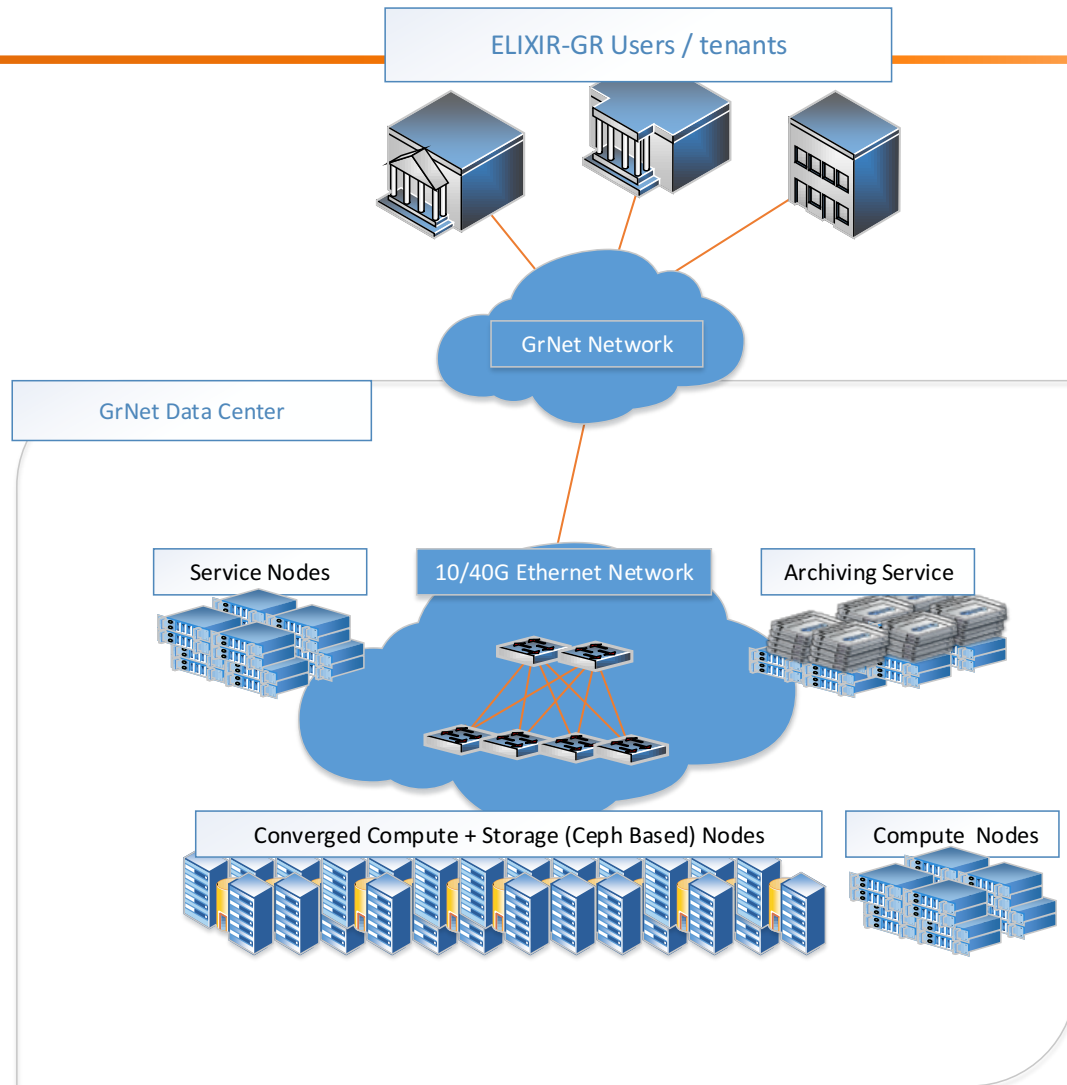
ELIXIR-GR: ο ρόλος του ΕΔΕΤ & του ΕΚ «Αθηνά»

- Το ΕΔΕΤ:
 - είναι ο **πάροχος (βασισμένων στο Νέφος) IaaS** της κοινότητας ELIXIR-GR
 - Χρήση openstack πλατφόρμας διαθέσιμης μέσω horizon UI
 - παρέχει **ειδικού τύπου ασφαλείς & γρήγορες γραμμές για τη διασύνδεση όλων των οργανισμών-εταίρων**
 - Χρήση ασφαλών, αφιερωμένων κυκλωμάτων
 - 10 Gbps δίκτυο
- Το ΕΚ «Αθηνά» έχει αναλάβει τη φροντίδα σχετικά με τις **αναλύσεις έντασης δεδομένων** (μεταξύ άλλων)
- Μαζί υποστηρίζουν τον πλήρη κύκλο ζωής της υποδομής:
 - Καταγραφή & ανάλυση απαιτήσεων
 - Ορισμός σεναρίων χρήσης & profiling
 - Σχεδιασμός υποδομής (με βάση τις απαιτήσεις)
 - Ολοκλήρωση προμήθειας του απαιτούμενου εξοπλισμού
 - Προετοιμασία υποδομής για παροχή υπηρεσιών
 - Παραγωγική φάση λειτουργίας υποδομής

Καταγραφή & ανάλυση απαιτήσεων

- **Ερωτηματολόγια** καταγραφής απαιτούμενων πόρων
 - ανά οργανισμό
 - ανά εργαλείο / υπηρεσία
- **Αποτελέσματα:**
 - Κάλυψη αναγκών με υπηρεσίες Νέφους διαφόρων τύπων
 - Thin nodes: Γενικού σκοπού servers (καλύπτουν πλειοψηφία εργαλείων)
 - Fat nodes: Δυνατοί servers για απαιτητικές αναλύσεις (μνήμη, cores κτλ)
 - GPU nodes: Κόμβοι εξοπλισμένοι με accelerators.
 - I/O nodes: Servers με υψηλό I/O throughput στο δίσκο.
 - Χρήση προηγμένων δικτυακών τεχνικών για ταχεία & ασφαλή δικτύωση
 - Ανάγκες για αποθήκευση σε δίσκο από όλα τα εργαλεία.
 - Εξυπηρέτηση μελών της ELIXIR-GR κοινότητας & την ευρύτερη κοινότητα βιοπληροφορικής

Αρχιτεκτονικός σχεδιασμός



- Η υποδομή θα είναι διασυνδεδεμένη με τα data centers IP Fabric του ΕΔΕΤ.
- Η προτεινόμενη λύση θα βασίζεται σε τεχνολογίες Openstack και Kubernetes.
- Θα υποστηρίζεται η δυνατότητα για self-provisioning
- Θα παρέχεται δικτυακή πύλη διαχείρισης για τους διαχειριστές της κοινότητας.

Προδιαγραφές υλικού

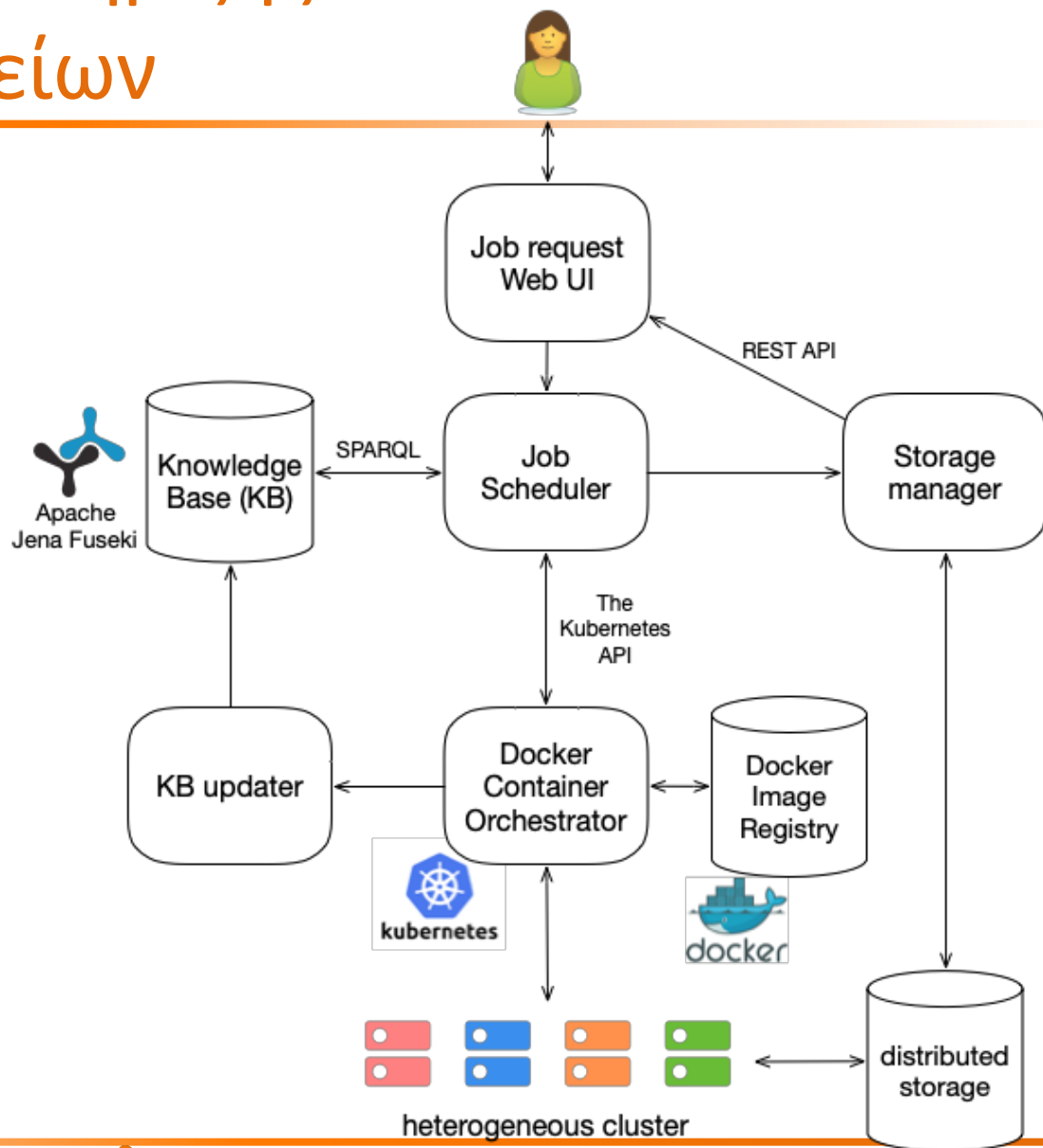
A/A	Brief Specs	cores / threads	mem	gpu	disk	form factor	pcs	disk (tb)
1data Switch	• 48x 10G SFP+ ports • 4x 40G QSFP ports • 2x 1G RJ45 ports • Switching Capacity:>1000 Gbps • Wire-speed Layer 3 Routing • 1:1 Non-blocking connectivity • Link Aggregation and Jumbo Frame Support					1 U	12	
IPMI switch / Management Fabric	• 48x 1G RJ45 ports - 48 RJ45 Copper ports and - 4 RJ45 or SFP ports • Switching Capacity:>150 Gbps • Wire-speed Layer 3 Routing • 1:1 Non-blocking connectivity • Link Aggregation and Jumbo Frame Support					1 U	6	
infrastructure nodes	Intel® Xeon® Gold family 2.6G,14C/28T,Intel UPI 10.4GT/s, 19M Cache ή μεγαλύτερη,Turbo,HT (140W) DDR4-2666, nbd support - see details tab	28	128		2x2TB SATA	<=1 U	12	
4Fat Client	Intel® Xeon® Platinum 8160M 2.1G, 24C/48T, Intel UPI 10.4GT/s, 33M Cache ή μεγαλύτερη, Turbo, HT (150W) 1.5TB DDR4-266, nbd support - see details tab	48	1024			1 U	2	
5GPU Client	Intel® Xeon® Gold 6132 2.6G,14C/28T, Intel UPI 10.4GT/s, 19M Cache ή μεγαλύτερη,Turbo,HT (140W) DDR4-2666, nbd support - see details tab	28	512		2	1 U	3	
6I/O Node	Intel® Xeon® Gold 6132 2.6G,14C/28T,Intel UPI 10.4GT/s, 19M Cache ή μεγαλύτερη,Turbo,HT (140W) DDR4-2666, nbd support - see details tab	28	512		2x2TB NVMe	1 U	8	16
7CCS	Intel® Xeon® Gold 6132 2.6G,14C/28T,Intel UPI 10.4GT/s, 19M Cache ή μεγαλύτερη,Turbo,HT (140W) DDR4-2666, nbd support - see details tab	28	256		12x10TB SATA	2 U	32	1280

Ανάπτυξη σε 3 racks



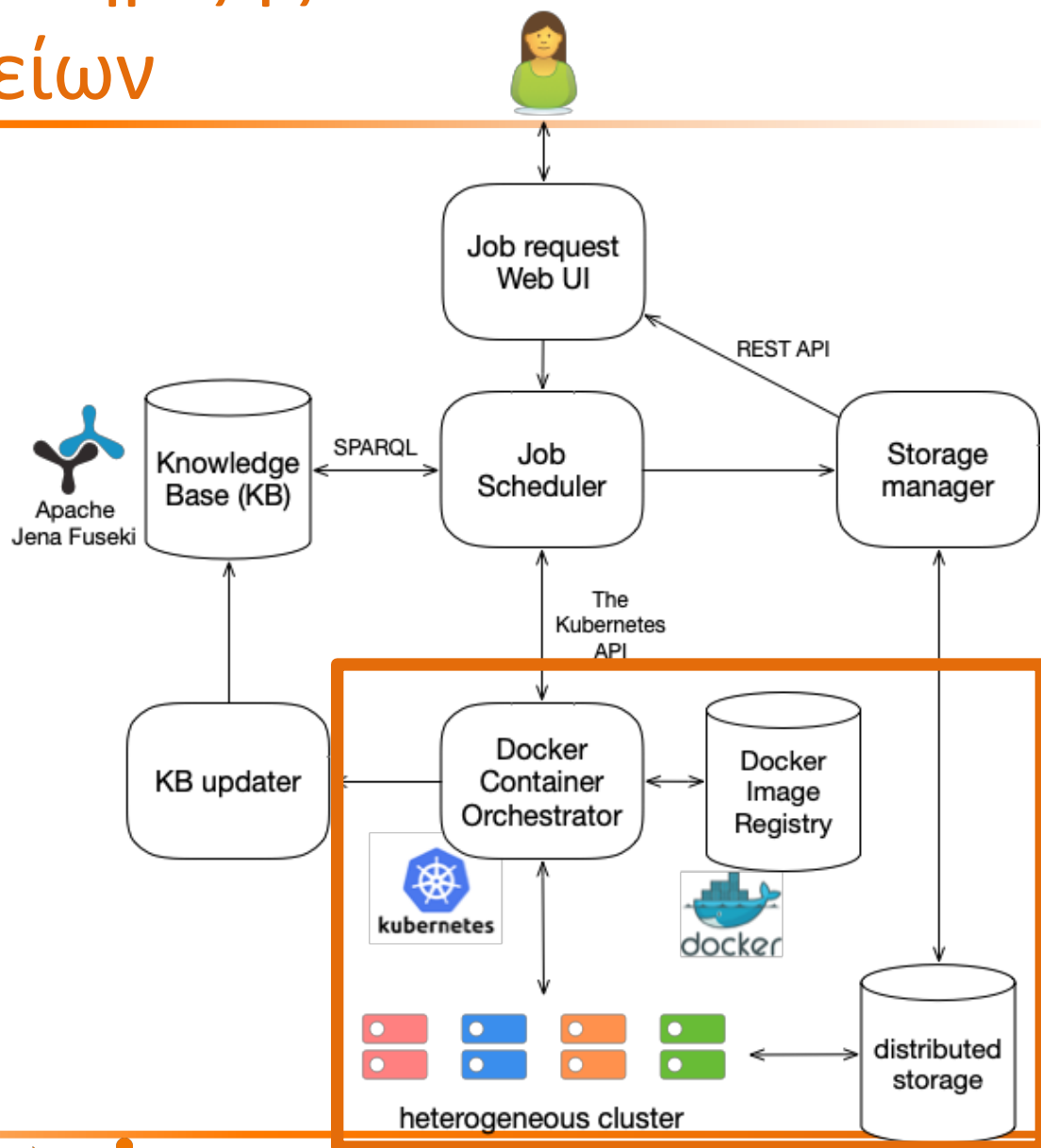
Υποσύστημα υποστήριξης dockerized εργαλείων

- Πρόσθετα από τις υπηρεσίες self-provisioning θα παρέχεται **υποσύστημα υποστήριξης βιοπληροφορικών εργαλείων σε μορφή docker.**
- Θα εγγυάται **αυτοματοποιημένη κατανομή υπολογιστικών φορτίων & αποθήκευσης δεδομένων.**
- Ο βιοπληροφορικός δεν προβληματίζεται με θέματα όπως στήσιμο VM, load balancing κτλ.



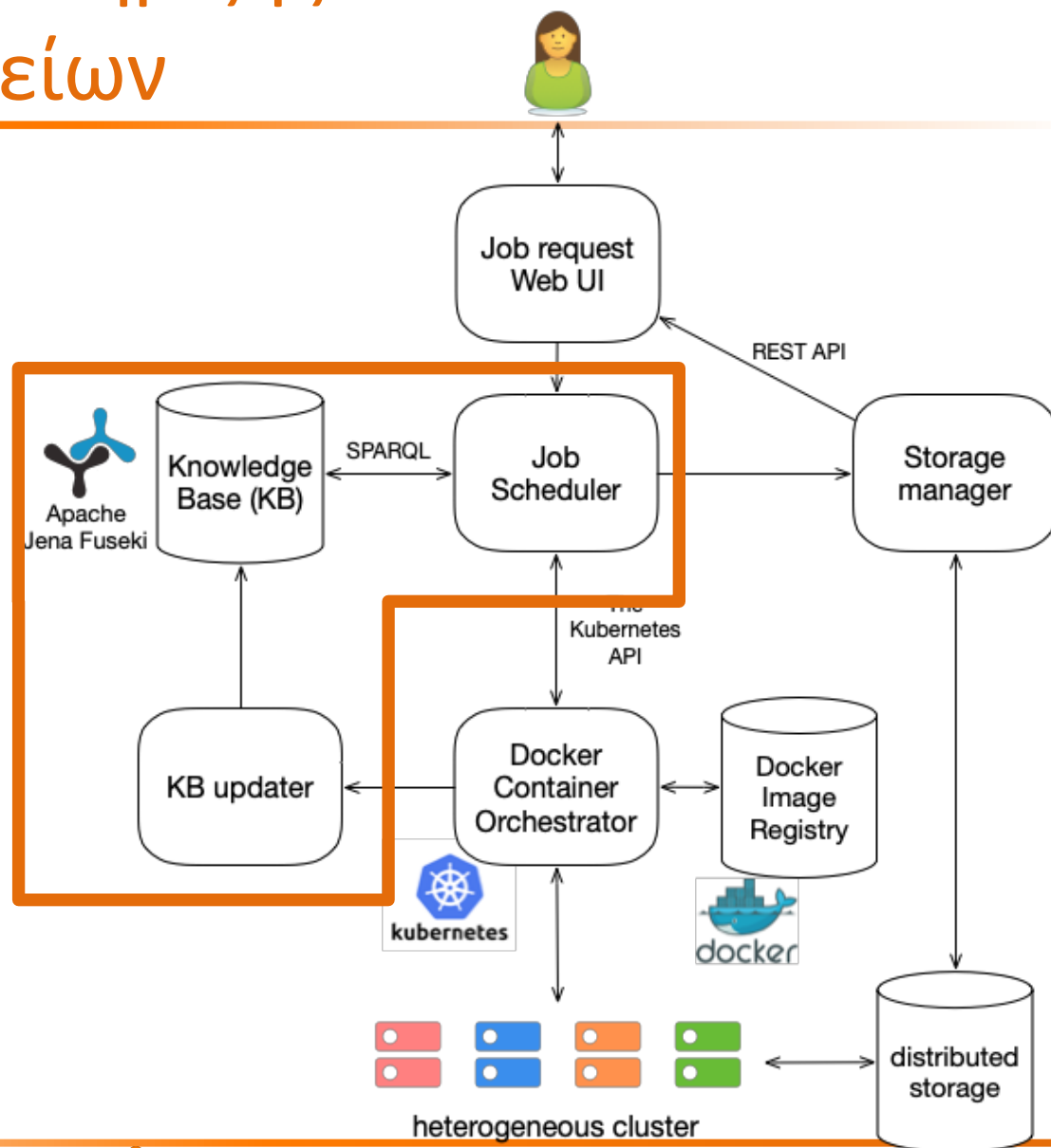
Υποσύστημα υποστήριξης dockerized εργαλείων

- **Docker Image Registry:** σύνολο dockerized εργαλείων κοινότητας
- Εκτέλεση σε ετερογενές cluster με κόμβους διαφορετικών δυνατοτήτων.
- Διαχείριση & παρακολούθηση εκτέλεσης με βάση το **Kubernetes**.
 - Υποβοήθηση από custom ενδιάμεσο λογισμικό
- Είσοδος/έξοδος δεδομένων από/σε **κατανεμημένο μέσο αποθήκευσης**



Υποσύστημα υποστήριξης dockerized εργαλείων

- **Βάση γνώσης (Knowl. Base - KB):**
 - Προεπιλεγμένα χαρακτηριστικά διαθέσιμων εργαλείων.
 - Τύποι υπολογισμού που εξυπηρετούνται καλύτερα κάθε είδος κόμβου.
- **Συμπερασμός** για τοποθέτηση εισερχόμενων job στο κατάλληλο είδος κόμβου
- **KB updater:** ανατροφοδοτεί τη βάση γνώσης για διόρθωση & προσθήκη νέας γνώσης με βάση το ιστορικό των εκτελέσεων (αναφορές εκτέλεσης kubernetes)



Ευχαριστώ

- Θανάσης Βεργούλης
vergoulis@imis.athena-innovation.gr