27 e 28 Giugno 2018

Conferenza dei Rettori delle Università Italiane - CRUI

I MAGNIFICI INCONTRI CRUI 2018

PIANO NAZIONALE UNIVERSITÀ DIGITALE

Infrastrutture tecnologiche e cybersecurity

Antonio Cisternino, Università di Pisa – Marco Baldi, Università Politecnica delle Marche – Sauro Longhi, Università Politecnica delle Marche – Marco Paganoni, Università di Milano Bicocca – Federico Ruggieri, GARR

Giugno 2018

Udine - Palazzo Garzolini di Toppo Wasserman, via Gemona 92

Tavolo 2B

Infrastrutture tecnologiche e cybersecurity

Antonio Cisternino, Università di Pisa – Marco Baldi, Università Politecnica delle Marche – Sauro Longhi, Università Politecnica delle Marche – Marco Paganoni, Università di Milano Bicocca – Federico Ruggieri, GARR

1. Introduzione

Le infrastrutture tecnologiche, con particolare riguardo a quelle informatiche e digitali sono di grande rilievo per la ricerca e l'alta formazione Universitaria. L'uso di tali infrastrutture è ormai prassi consolidata per tutte le attività: dalla creazione di dati scientifici e trasmissione ed elaborazione di tali dati, alle attività di didattica a distanza, di e-Learning e di utilizzo di strumenti collaborativi.

L'uso di reti trasmissive ad altissima velocità e sistemi di calcolo ad alte prestazioni ha reso possibile approcciare nuove metodologie di lavoro basate su condivisioni di grandi quantità di dati, di cicli rapidi di simulazioni ed elaborazioni, nonché sempre crescenti livelli di automazione e supporto alle decisioni.

Una nazione come l'Italia, per il rilievo scientifico ed economico che possiede, non può prescindere dal confrontarsi alla pari con le maggiori nazioni del mondo sul tema delle infrastrutture tecnologiche e del loro uso, in particolare, nel campo della ricerca ed istruzione che sono fra i pilastri fondanti di una nuova economia basata sulla conoscenza. Gli investimenti fatti negli anni passati devono trovare continuità e nuova linfa nei piani dei prossimi anni per garantire alla nostra nazione un ruolo leader e di primato nel campo scientifico, culturale e, conseguentemente, economico.

In un tale contesto, per sua natura altamente dinamico ed evolutivo, si sono sviluppate di pari passo necessità nuove di sicurezza che attengono non soltanto alla violazione di sistemi per fini illegali, ma anche al furto e manomissione dei dati ed alla salvaguardia delle informazioni personali e sensibili. Secondo i report periodici di numerosi ed autorevoli osservatori, l'incidenza di attacchi cyber è infatti in continuo aumento, ed il sistema universitario è tra i principali obiettivi di tali attacchi, sia relativamente alla continuità dei servizi che alla protezione della proprietà intellettuale. A ciò si aggiunge l'entrata in vigore della parte sanzionatoria, a partire dal 25 Maggio 2018, della nuova normativa europea sulla protezione dei dati (GDPR) [1], attiva già dal 24 Maggio 2016, che richiede aumentati livelli di attenzione e competenza da parte di tutti gli enti e le aziende coinvolti nel trattamento di dati personali dei cittadini europei.

La necessità di manipolazione di dati personali e sensibili è imprescindibile in molti contesti, si pensi ad esempio quello medico in cui ancora si stenta ad avere una cartella clinica elettronica uniforme a livello nazionale e, magari, europeo. Vale però la pena citare come, nella medesima direzione, si muova anche il settore dell'alta formazione universitaria con un panorama ancora non uniforme e non totalmente standardizzato dei curricula e delle certificazioni.

Negli ultimi due anni sono state pubblicate diverse norme e linee guida. Un ruolo rilevante in tale contesto è svolto dal Laboratorio Nazionale di Cybersecurity del CINI, che raccoglie 45 università italiane nella principale rete di eccellenza accademica nazionale sul tema della cybersecurity. In attuazione del "Quadro strategico nazionale per la sicurezza dello spazio cibernetico" [2], nel Febbraio 2016 il laboratorio nazionale ha rilasciato la prima versione del "Framework nazionale per la cybersecurity" [3], che raccoglie linee di indirizzo e buone pratiche per la cybersecurity di enti ed aziende. Il framework nazionale è stato recepito da diverse iniziative successive, come il documento dell'Agenzia per l'Italia Digitale (AgID) riportante "Misure Minime per la

sicurezza ICT della Pubblica Amministrazione"[1] [4] ed i controlli essenziali di cybersecurity per enti ed aziende definiti dal CIS Sapienza e dal Laboratorio Nazionale di Cybersecurity del CINI [5]. Rispetto a queste problematiche le Università (e anche gli enti pubblici di ricerca) hanno la necessità di adottare misure adeguate e condivise per garantire l'uniformità di regole e trattamento dei dati in loro possesso, nonché le metodologie di sicurezza e di intervento in caso di incidenti informatici. Risulta inoltre di primaria importanza la preparazione di risorse umane competenti e l'arricchimento dell'offerta formativa universitaria in tema di cybersecurity, in quanto il mercato del lavoro nazionale sta vivendo una fase in cui la richiesta supera l'offerta di personale specializzato in tale area. Queste ed altre linee di indirizzo sono raccolte nel libro bianco intitolato "Il Futuro della Cybersecurity in Italia: Ambiti Progettuali Strategici", pubblicato a Febbraio 2018 dal Laboratorio Nazionale di Cybersecurity del CINI [6].

2. Infrastrutture fisiche

Le infrastrutture fisiche sono quelle tangibili che possiamo suddividere tradizionalmente fra infrastrutture di rete e infrastrutture di calcolo e dati. La prima tipologia è gestita ed operata a livello italiano dal GARR, così come riconosciuto dal D. Lgs 218 del 2016. Le infrastrutture di calcolo e dati sono molteplici a livello nazionale e fanno capo a diverse tipologie architetturali e applicative come, ad esempio, High Performance Computing (HPC), High Throughput Computing, Grid e Cloud.

2.1 La Rete

Le Università e gli enti di ricerca italiani sono stati fra gli iniziatori delle attività di rete in Italia ed in Europa. La rete GARR da circa 30 anni ha raccolto il testimone e, sulla spinta della comunità stessa, ha realizzato e opera una delle reti accademiche e di ricerca fra le più avanzate nel mondo con oltre 15.000 km di fibra ottica di proprietà ed utilizzata per fornire accessi fino a 100 Gb e collegando più di 1000 sedi fra Università, Enti e centri di ricerca, scuole, musei ed istituzioni culturali.

Dall'ultimo Annual Report del GARR, in corso di pubblicazione, si possono ricavare le caratteristiche principali di questa infrastruttura fisica.

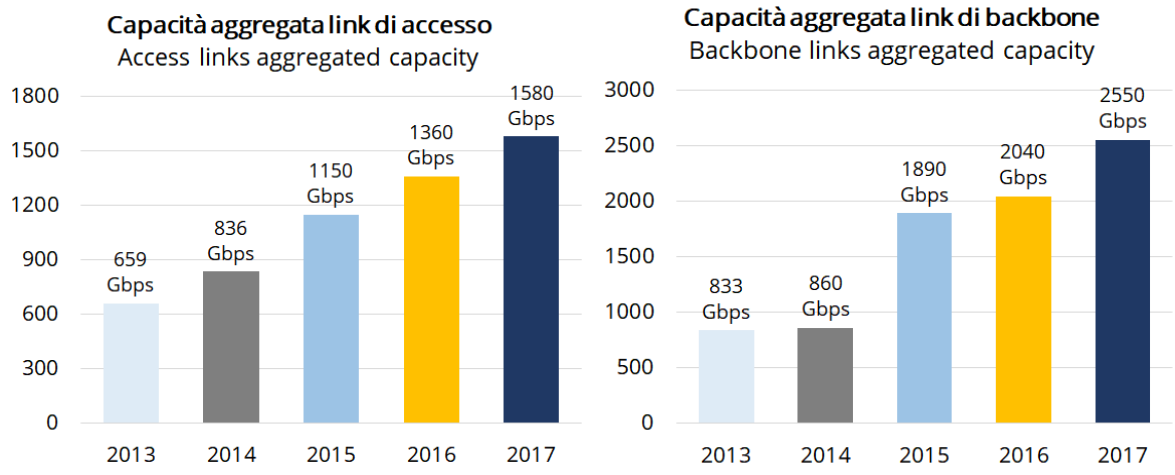

Figura 1. Capacità aggregata link di accesso e di backbone della rete GARR.

Dai dati in Figura 1 si nota come il traffico totale sull'infrastruttura di rete sia cresciuto considerevolmente a partire del 2016 quando è diventata operativa l'infrastruttura in fibra del Progetto GARR-X Progress (2013-2016) nelle regioni di convergenza (Campania, Calabria, Puglia e Sicilia) realizzata attraverso un

[1] http://www.agid.gov.it/sites/default/files/documentazione/misure_minime_di_sicurezza_v.1.0.pdf

finanziamento di 46,5 M€ del MIUR su fondi del Piano di Azione e Coesione (PAC). L'incremento del traffico, superiore a qualunque previsione, testimonia l'importanza e l'impatto degli investimenti in infrastrutture nelle aree a digital divide e come la realizzazione di reti con tecnologie e prestazioni all'avanguardia possano produrre effetti considerevoli sull'uso della rete da parte degli utenti di tali aree.

L'altra osservazione evidente è che la capacità totale della dorsale di rete (backbone) è superiore alla somma degli accessi. Tale caratteristica di infrastruttura "over-provisioned" è uno degli elementi che differenzia le reti della ricerca da quelle commerciali dove l'infrastruttura è invece di tipo "overbooked" offrendo prestazioni solo su base statistica.

La disponibilità di infrastruttura in fibra e di tecnologie avanzate di rete sopra di essa, con bande passanti elevate (10-100 Gb) e bidirezionali, consente alle Università ed Enti collegati di essere non solo fruitori di dati, ma anche di essere produttori e di offrire un accesso veloce ai dati stessi.

Come risulta dai dati riportati in Figura 2, il traffico verso altre reti della ricerca è prevalente e costituisce una dimostrazione evidente della necessità di avere delle Reti Nazionali Accademiche e di Ricerca che comunicano fra loro ad altissima velocità e senza intermediari.

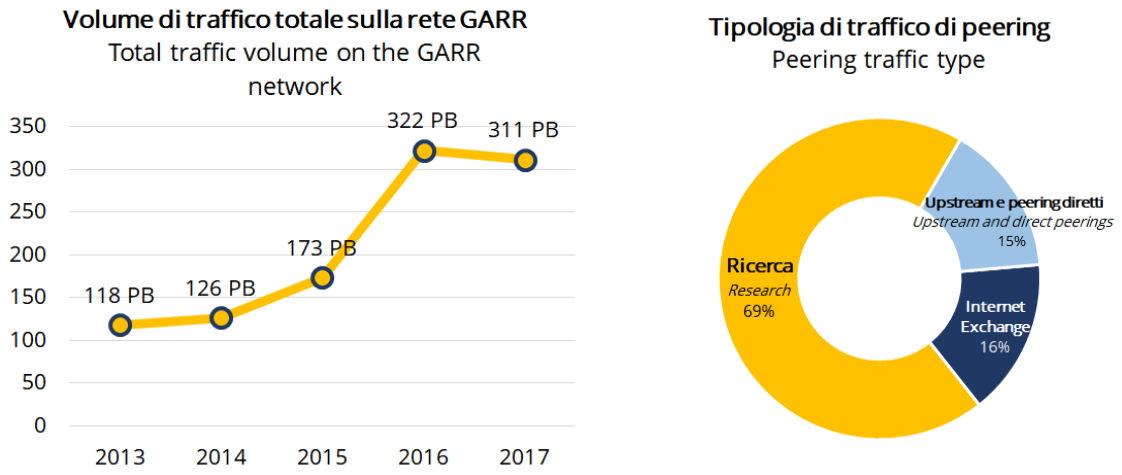

Figura 2. Volume e tipologia di traffico sulla rete GARR.

La grande banda passante a disposizione permette l'accesso ed il trasferimento di grandi quantità di dati fra centri di ricerca e data centre o centri di supercalcolo come il CINECA di Bologna. L'uso della rete è ormai indispensabile per accedere ad archivi digitali ed alle risorse di calcolo necessarie per processare enormi quantità di dati. La dipendenza di quasi tutte le attività delle Università e della ricerca dall'accesso alla rete è ulteriormente accentuata dalla necessità di creare sistemi affidabili attraverso le repliche dei data centre e la realizzazione di sistemi di disaster recovery. Recentemente GARR ha preso in considerazione l'uso di apparati trasmissivi per offrire connessioni punto-punto (p2p) tra sedi di data centre. La prima realizzazione pratica è stata quella fra il CINECA ed il Centro CNAF dell'INFN a Bologna dove è attualmente utilizzata una capacità trasmissiva iniziale di 400 Gb che può essere espansa fino a 1,2 Tb. A causa delle distanze limitate fra i due siti (circa 20 km reali di fibra ottica) i tempi di latenza sono tali da rendere equivalenti le risorse locali e quelle sull'altro sito, come se fossero nella stessa sala macchine. Naturalmente un allargamento di queste tecnologie a distanze geografiche maggiori è possibile ma con limitazioni dovute alla latenza (circa 5ms/1.000 km - velocità della luce nel materiale) derivante da maggiori distanze. Soluzioni basate su IP routing hanno prestazioni nettamente peggiori sia per la latenza che per il Jitter dovuti ai protocolli ed all'uso di buffer negli apparati di routing.

L'uso della fibra ottica è oramai una realtà dentro la comunità delle Università e della Ricerca e, anzi, molte Università sono state pioniere e protagoniste della realizzazione di reti metropolitane (es. Pisa o Trieste) o regionali (es. Cassino) in fibra. La fibra ha non solo la capacità di scalare a velocità di molti Tbps, ma anche di poter dare nuovi servizi alla comunità: interconnessione di Data Centres, segnali di tempo e fase (INRIM), nonché uso di trasferimenti real time per applicazioni come LOLA (Low Latency) nel campo audio-video.

Gli sviluppi necessari sono dunque quelli verso una maggiore resilienza attraverso l'accesso ridondato delle Università e l'introduzione di nuove infrastrutture trasmissive ottiche per aumentare la capacità futura della rete e l'utilizzo di nuovi servizi.

2.2 Il Computing e lo Storage dei Dati

Le infrastrutture di calcolo e dati a livello nazionale sono diverse, con caratteristiche differenti e gestite ed operate da più soggetti. Nel campo dello HPC sicuramente l'infrastruttura di riferimento è quella del Consorzio InterUniversitario CINECA, che opera supercalcolatori di rilievo europeo e mondiale. L'attuale macchina di punta del consorzio - Marconi - è presente al 14mo posto nell'ultima versione della lista dei Top500 di Novembre 2017. Il CINECA partecipa alla infrastruttura di supercalcolo europea PRACE ed alla infrastruttura europea di supporto alla gestione dai EUDAT.

Figura 3. Rete GARR e centri di HPC.

La Figura 3 mostra un quadro dei maggiori centri nazionali di Calcolo. L'infrastruttura di tipo HTC (High Throughput Computing) più grande a livello nazionale è quella dell'INFN che, con i suoi centri Tier1 (CNAF) e Tier2 (Bari, Catania, Frascati, Legnaro, Milano, Napoli, Pisa, Roma, Torino) partecipa alla infrastruttura mondiale di calcolo LHC Computing Grid (WLCG).

Infrastrutture nazionali sono state create anche da ENEA e INAF ed a queste si aggiungono anche centri di calcolo misti INFN/Università finanziati su fondi PON nel sud dell'Italia (es. Progetto RECAS - Bari, Catania, Cosenza e Napoli), il CRS4 in Sardegna e l'infrastruttura realizzata da GARR nell'ambito del già citato Progetto GARR-X Progress (Bari, Catania, Cosenza, Napoli e Palermo).

In questo ambito si inseriscono i Data Centre universitari che spesso soddisfano esigenze diverse: dall'amministrazione alla gestione degli studenti e alla ricerca. La varietà di tali infrastrutture è funzione delle dimensioni dell'ateneo, delle funzioni da svolgere e della quota parte di servizi esternalizzati.

Nella lunga lista citata di infrastrutture e data centre, solo alcuni possiedono caratteristiche di controllo degli accessi, ridondanza dei sistemi di alimentazione e condizionamento e relative certificazioni.

Nel 2017 si è costituito un gruppo di lavoro per la discussione e l'armonizzazione delle infrastrutture di calcolo e dati in Italia ICDI (Italian Computing and Data Infrastructure - www.icdi.it) con l'obiettivo di coordinare dal basso la partecipazione italiana alla European Open Science Cloud (EOSC) che è un'iniziativa della Commissione Europea (DG RTD e DG CNECT) per superare la frammentazione delle risorse a livello europeo nel campo del calcolo e dell'accesso ai dati. Nell'ambito di ICDI sono state lanciate alcune azioni:

- Un documento di intenti;
- Compilazione di un elenco delle risorse attualmente installate a livello nazionale dai vari enti;
- Creazione di una Assemblea di ICDI aperta a Infrastrutture di Ricerca e di calcolo e rete per discutere le problematiche e un possibile percorso di coordinamento;
- Redazione di un Memorandum of Understanding da sottoscrivere da parte degli enti che finanziano e operano le infrastrutture in Italia, con una possibile evoluzione verso una Joint Research Unit.

Le problematiche da affrontare sono molteplici: oltre alla frammentazione delle infrastrutture di calcolo e dati, manca ancora un'azione concreta per affrontare le odierne e future necessità in termini di accessibilità dei dati e della loro gestione e preservazione a medio e lungo termine.

3. Infrastrutture virtuali

"Software defined everything" è divenuto il mantra dell'industria IT, condizionando e ridefinendo il modo con cui si realizzano infrastrutture digitali. L'approccio, divenuto popolare nell'ambito della realizzazione di infrastrutture cloud, consiste nel disaggregare l'infrastruttura fisica dalla sua astrazione logica in modo da poter allocare dinamicamente le risorse di calcolo, memorizzazione e comunicazione in modo più efficace ed efficiente. Se la rivoluzione è cominciata con la virtualizzazione dei sistemi operativi, la virtualizzazione della rete (Software Defined Network, Network Function Virtualization, e Open Networking) e più recentemente quella dei sistemi di storage hanno completato la transizione.

Il modello di riferimento delle architetture cloud illustrato in Figura 4 è ormai alla base dell'organizzazione delle infrastrutture software-defined e spiega la necessità di introdurre uno strato di virtualizzazione nella gestione delle risorse di calcolo: il livello di servizio (service layer) che offre servizi in modalità self-service agli utenti richiede un livello di automazione che i livelli di orchestrazione e controllo possono raggiungere solo usando largamente la capacità di virtualizzare le risorse ed allocarle in modo da garantirne l'uso efficiente e robusto, capace di erogare servizio anche durante la manutenzione e/o il fallimento delle risorse al livello fisico.

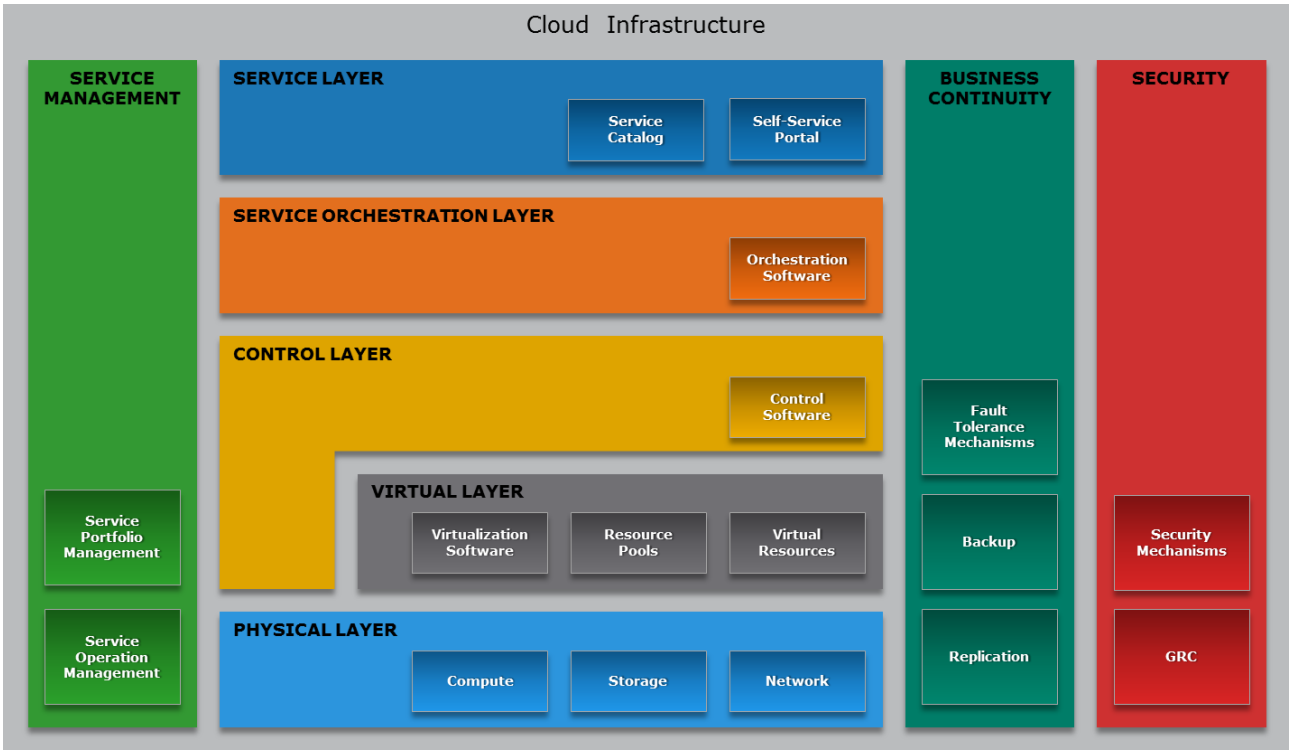

Figura 4. Modello di riferimento delle architetture cloud.

Negli ultimi anni la tecnologia dei container si è affiancata a quella della virtualizzazione per la condivisione delle risorse di calcolo. Questa tecnologia, resa popolare dallo startup *docker*, consente di partizionare le risorse di un sistema operativo ingannando i suoi processi e nascondendo le risorse incluso il file system. Concettualmente il ruolo svolto da questa tecnologia è lo stesso della virtualizzazione, ma avendo un overhead minore si presta ad essere impiegata anche in situazioni in cui la potenza di calcolo è un aspetto importante, come ad esempio il calcolo scientifico. Il progetto open *Singularity* ha come obiettivo quello di usare la tecnologia dei container per il calcolo ad alte prestazioni (HPC), in modo da poter rendere accessibili gli acceleratori grafici (GPU) oggi utilizzati per accelerare calcoli che fanno uso di matrici come ad esempio alcuni algoritmi di intelligenza artificiale.

Un'altra incarnazione di questo trend di virtualizzazione delle risorse di calcolo è quella della cosiddetta *iperconvergenza*, un approccio diverso nell'acquisizione e nella gestione delle risorse computazionali. Tradizionalmente, infatti, i servizi informatici si sono organizzati in tre silos che seguivano i tre aspetti fondamentali di un sistema di calcolo: l'elemento computazionale, la rete, e lo storage. Questa organizzazione, largamente indotta da caratteristiche tecnologiche, è stata messa in discussione dall'evoluzione che i dischi hanno subito negli ultimi anni, sia in termini di capacità che di velocità, rendendo possibile far crescere un sistema di calcolo aggiungendo server che allo stesso tempo aggiungono un po' dei tre elementi. In questo modo è possibile far crescere un sistema al crescere delle necessità, aggregando le risorse di più apparati grazie a software capaci di aggregare le risorse fisiche in un pool unico espandibile dinamicamente aggiungendo nuovi server.

3.1. La virtualizzazione: la didattica e la ricerca

Didattica e ricerca hanno necessità di evolvere costantemente per seguire i più recenti sviluppi, pertanto queste tecnologie che consentono di riallocare le risorse in modo dinamico offrono la capacità di adattare i sistemi a sempre nuove necessità pur mantenendo l'allocazione cosiddetta on-premise, ovvero vicino geograficamente ai ricercatori.

Il cloud è infatti uno dei massimi contribuenti allo sviluppo delle tecnologie della virtualizzazione, e molte di queste tecnologie sono state poi ingegnerizzare per consentire ad organizzazioni di realizzare cloud cosiddetti privati (ovverosia sistemi di calcolo che si comportano come un cloud pubblico ma sono eseguiti all'interno dei sistemi di calcolo di un'organizzazione). Se per alcune attività di didattica e ricerca il cloud è uno strumento importante, la crescente necessità di banda di rete per poter movimentare i dati raccolti e che richiedono l'analisi ai fini di ricerca è incompatibile con l'invio ad un cloud per poterli successivamente elaborare.

Nella tradizione del calcolo scientifico l'aspetto prevalente è quello della capacità di elaborazione, ma l'esplosione dei dati utilizzati da discipline che hanno solo recentemente cominciato ad avvalersi del calcolo scientifico per analizzare dati sta contribuendo alla ridefinizione dei requisiti che sottendono alla creazione di un'infrastruttura di calcolo scientifico al servizio della ricerca e della didattica.

Inoltre, l'esplosione del nuovo campo dell'Internet of Things (IoT) sta moltiplicando in maniera esponenziale il numero di dispositivi connessi ad Internet, e di conseguenza la necessità di trasferimento dati. La connessione di grandi numeri di dispositivi elementari alla rete, come ad esempio sensori di temperatura, richiede una capacità trasmissiva che rende necessario riavvicinare le risorse di calcolo al luogo di generazione dei dati. Cisco research ha dato un nome alla necessità di avvicinare il cloud a chi genera i dati: Fog computing. Il consorzio internazionale OpenFog (http://www.openfogconsortium.org) composto da industria e istituzioni accademiche, sta cercando di definire in modo più accurato l'architettura del Fog.

Allo stesso tempo le public cloud provider hanno cominciato ad indicare nel cloud ibrido la strada da seguire, ovverosia in una gerarchia di cloud gestiti da più organizzazioni in cui le risorse vengono allocate e spostate da un cloud all'altro. Ecco quindi che il cloud privato, locale ad un Ateneo, può essere esteso con un'infrastruttura cloud nazionale, e poi, magari da uno o più cloud pubblici.

La necessità di tornare ad un modello distribuito di uso delle risorse deriva quindi dal bisogno di avvicinare i processori ai dati in modo da poter beneficiare dall'enorme capacità di trasferimento dati che i dischi mettono oggi a disposizione. Non è un caso che *NVMe over fabric* [12] è stato uno dei temi caldi all'ultima edizione di SuperComputing 2018, storica conferenza che si occupa di HPC e calcolo scientifico.

Come ulteriore esempio basti pensare che l'ultima generazione di processori AMD è introdotto 128 linee PCI Express 3.0 (16 linee sono in grado di trasferire circa 15GB/s) direttamente all'interno della CPU per assicurare che i vari core possono ricevere dati il più efficientemente possibile. Questo significa che una singola CPU è virtualmente capace di ricevere in input quasi 1Tbps, cioè 10 volte la banda massima offerta da una connessione Ethernet.

3.2. Le infrastrutture nazionali e le Università

GARR ha cominciato ad offrire un'infrastruttura cloud per supportare i ricercatori nell'allocazione delle risorse di calcolo. L'infrastruttura è federata, nel senso che ciascun consorziato può entrare nella federazione mettendo a disposizione risorse computazionali proprie a ricercatori di altri enti.

L'idea di federare le risorse offre sicuramente un'opportunità di avvicinare computazioni a particolari set di dati. Se ad esempio un particolare Ateneo ha sviluppato un insieme di dati medici, questi possono essere messi a disposizione di altri ricercatori offrendo la possibilità di eseguire un calcolo presso l'Ateneo dove risiedono i dati invece di trasferire il dato per poterlo elaborare.

Sebbene questa scelta può ridurre la quantità di dati spostati lo spostamento di macchine virtuali e/o container può comunque richiedere la movimentazione di quantità significative di dati che sicuramente beneficiano da un'infrastruttura di rete sempre più performante come la rete GARR.

È importante sottolineare che un sistema di virtualizzazione interopera facilmente con altri, la federazione di cloud è quindi uno strumento che può facilitare lo scambio di immagini di macchine virtuali e di computazioni ma non è l'unica possibilità. Inoltre esistono già numerosi standard usati dalle differenti comunità scientifiche a livello mondiale, ed è plausibile ritenere che un Ateneo debba supportare più silos per offrire un adeguato supporto ai propri ricercatori.

In ogni caso, nella progettazione e nello sviluppo di un'infrastruttura nazionale, è necessario tenere conto della necessità di sviluppare non solo l'infrastruttura fisica, ma anche l'overlay virtuale necessario ad offrire la necessaria granularità nell'accesso alle risorse. Visto il probabile numero di istituzioni coinvolti è necessario disegnare forme di accoppiamento debole di sistemi in modo da consentire un'integrazione ed una interazione tra sistemi nel rispetto dell'autonomia e delle necessità dei singoli enti. Questo schema è di fatto stato realizzato dalla rete GARR utilizzando i protocolli di routing lasciando la necessaria autonomia a ciascun ente nella definizione della propria architettura di rete.

Sicuramente la pressione che AgID sta facendo sugli Atenei e sugli enti di Ricerca pubblici per fare uso delle convenzioni CONSIP come una qualsiasi Pubblica Amministrazione, o i vincoli che si intendono porre alla realizzazione ed allo sviluppo di Data Center può rappresentare un ostacolo significativo alla realizzazione di progetti di Ricerca, anche semplicemente a causa dell'aggravio burocratico imposto per giustificare il ricorso a strutture speciali. La necessità di razionalizzare la spesa della Pubblica Amministrazione non può danneggiare gli investimenti necessari allo sviluppo di infrastrutture nazionali adeguate alla Ricerca; l'aumento dell'attività di supporto all'approvvigionamento di servizi e sistemi ICT da parte di CRUI è apprezzabile e sarebbe auspicabile che crescesse in modo da supportare le necessità particolari di Didattica e Ricerca riducendo gli oneri amministrativi spesso necessari per poter ottenere gli strumenti necessari a poter svolgere queste missioni.

3.3. Cybersecurity

Come indicato nel modello di riferimento del cloud, la sicurezza è sicuramente un aspetto trasversale da tenere presente nella realizzazione di questo livello di virtualizzazione. Si tratta di un aspetto che pervade tutti i livelli dell'architettura e che oggi pone non solo sfide tecniche, ma anche normative, come l'applicazione delle misure minime di sicurezza AgID e del regolamento GDPR.

Nella visione tradizionale la sicurezza informatica viene costruita su tre livelli: sicurezza fisica, sicurezza logica e sicurezza procedurale. La capacità di definire perimetri in cui era ben definito un dentro ed un fuori rispetto alle risorse di un'organizzazione ha caratterizzato l'organizzazione della sicurezza negli ultimi 20 anni. Un firewall perimetrale verifica le connessioni provenienti dall'esterno nella ragionevole sicurezza che all'interno della rete elementi di sicurezza fisica logica e procedurale possano garantire l'opportuno livello di sicurezza.

L'IoT mette in discussione questa architettura poiché distribuendo geograficamente gli elementi di calcolo diventa quasi impossibile definire un perimetro che definisca un dentro e un fuori. In questo contesto è quindi necessario introdurre livelli di controllo sulle comunicazioni di rete che siano più pervasivi.

L'architettura che sembra più promettente in questo contesto è quella di microsegmentazione delle reti virtuali prevedendo l'introduzione di firewall fisici o virtuali per controllare piccoli segmenti di rete. Esistono soluzioni commerciali che cercano di seguire questo approccio, sia fornite dai produttori di apparati di rete che dai vendor di soluzioni di virtualizzazione.

In ogni caso le misure di cybersecurity a livello della virtualizzazione delle risorse seguono, da una parte la loro allocazione a tenant (gruppi di utenti/organizzazioni) differenti, dall'altra la necessità di assicurare canali di comunicazione sicuri che consentano lo scambio di dati in accordo alle misure di sicurezza imposte dalle varie normative.

4. Information governance e cybersecurity

La gestione dei dati e delle infrastrutture tecnologiche, così come la loro sicurezza, sono alla base dell'amministrazione dell'informazione, o information governance, che nello scenario attuale si pone come elemento fondamentale di competitività e sviluppo non solo degli atenei, ma dell'intera società. Si prevede infatti che nel 2020 la mole di dati a disposizione nell'universo digitale raggiungerà i 40 Zettabyte (ZB), ovvero un numero di byte pari a 57 volte la quantità di granelli di sabbia presenti su tutte le spiagge del pianeta. Già alla fine del 2018, questi dati saranno scambiati tra oltre 25 miliardi di dispositivi connessi in rete. È pertanto evidente come, tanto negli atenei quanto nel resto della società, tali moli di dati rischino di diventare incontrollabili in assenza di una adeguata information governance.

Due sono gli aspetti alla base di una information governance capace di proteggere adeguatamente gli asset digitali ed il loro valore: la preparazione del personale e l'uso di regolamentazioni e raccomandazioni adeguate.

Il primo di tali aspetti trova fondamento nel fatto che il fattore umano è ormai noto rappresentare l'anello debole della catena della cybersecurity, e si colloca al primo posto delle cause di vulnerabilità. Ciò è dovuto non soltanto al cattivo utilizzo degli strumenti tecnologici da parte di personale non adeguatamente addestrato, ma anche al proliferare di tecniche di attacco specificamente indirizzate alla componente umana di enti ed aziende, come le tecniche basate su social engineering e profilazione degli utenti [7]. All'esigenza interna di preparazione del personale di enti ed aziende, si aggiunge la crescente esigenza di risorse umane con profili tecnologici (e non solo) da assorbire presso aziende che forniscono servizi nell'ambito della cybersecurity. Negli ultimi anni la domanda di tali servizi è continuamente aumentata e le aziende che li offrono hanno difficoltà a reperire risorse umane necessarie a rispondere a tale aumentata domanda. Sulla base di tali premesse, riveste un ruolo fondamentale l'investimento in percorsi di formazione nell'area della cybersecurity, sia volti a creare consapevolezza del personale che gestisce o semplicemente usa le risorse digitali sia volti a preparare nuove figure altamente specializzate che trovano rapida collocazione nel mercato del lavoro.

In ambito accademico, il laboratorio nazionale di cybersecurity del CINI costituisce la più grande rete di eccellenza sul tema della cybersecurity, raccogliendo 250 tra professori e ricercatori appartenenti a 45 università Italiane. Scopi del laboratorio nazionale sono aiutare il sistema paese ad essere più resiliente alla minaccia cibernetica, migliorare la continuità di servizio dei sistemi critici, aumentare la consapevolezza nella società, migliorare le misure di protezione da attacchi cibernetici della pubblica amministrazione e delle imprese e supportare processi di definizione di standard e framework metodologici a livello nazionale. Per ciò che riguarda la formazione accademica nell'ambito della cybersecurity, il laboratorio nazionale promuove e coordina la creazione di specifici percorsi formativi presso gli atenei nazionali [8].

Per ciò che riguarda regolamentazioni e raccomandazioni nell'ambito della cybersecurity, va osservato che l'Italia si è dotata già dal 2013 di una strategia nazionale per la cybersecurity. Essa è stata introdotta col Decreto del Presidente del Consiglio dei Ministri 24 Gennaio 2013 ed aggiornata con Decreto del Presidente del Consiglio dei Ministri 17 Febbraio 2017, contenente la "Direttiva recante indirizzi per la protezione cibernetica e la sicurezza informatica nazionali". L'introduzione della strategia nazionale ha portato alla definizione nel Dicembre 2013 del "Quadro strategico nazionale per la sicurezza dello spazio cibernetico" da parte della Presidenza del Consiglio dei Ministri. Tra i sei indirizzi strategici identificati nel quadro nazionale figura la "promozione e diffusione della cultura della sicurezza cibernetica sia tra i cittadini che all'interno delle istituzioni, anche attraverso un sempre maggiore coinvolgimento del mondo della ricerca e

dell'università, al fine di accrescere il livello di consapevolezza e di conoscenza della minaccia e dei relativi rischi".

Nel Febbraio 2016 il laboratorio nazionale di cybersecurity del CINI ha introdotto il "Framework nazionale di cybersecurity" [3], che rappresenta un'importante raccolta di linee guida e buone pratiche per la cybersecurity, oltre che uno strumento di autovalutazione del rischio cyber a disposizione di enti ed aziende. Il framework nazionale trae ispirazione dal framework NIST (National Institute of Standards and Technology) per la protezione delle infrastrutture critiche, emanato nel 2014 e recentemente aggiornato alla versione 1.1 [9], da cui mutua la suddivisione delle linee guida in cinque categorie, legate alle fasi di una minaccia cibernetica, ovvero: identify, protect, detect, respond e recover. Il framework nazionale fornisce 98 sotto-categorie di raccomandazioni raccolte in tali cinque categorie, e per ciascuna di esse aggiunge la possibilità di specificare livelli di priorità e di maturità. Tale estensione rispetto al framework NIST consente di introdurre la nozione di contestualizzazione del framework alle diverse tipologie di organizzazioni. In termini pratici, una contestualizzazione del framework consiste nella selezione delle sotto-categorie del framework che si applicano alla tipologia di organizzazioni in esame, e la definizione dei relativi livelli di priorità e di maturità. Ciò permette l'applicazione del framework a varie tipologie di organizzazioni, sulla base delle loro caratteristiche e vulnerabilità tipiche.

Il framework nazionale di cybersecurity, oltre a fornire un riferimento normativo ed uno strumento di autovalutazione del rischio cyber, ha fornito la base di partenza per ulteriori iniziative volte a guidare l'innalzamento dei livelli di cybersecurity di enti ed aziende. Prima fra queste la pubblicazione di quindici controlli essenziali di cybersecurity da parte del laboratorio nazionale di cybersecurity del CINI [5], volti a fornire uno strumento più snello da affiancare al framework per l'esecuzione di un primo livello di screening dei livelli di cybersecurity di un ente oppure di un'azienda. Tali controlli, riportati in Tabella 1, sintetizzano in poche regole generali gli aspetti fondamentali per un'adeguata gestione della cybersecurity di un'organizzazione, ovvero: inventario di sistemi e dispositivi, governance, protezione da malware, gestione di password ed account, formazione e consapevolezza del personale, protezione dei dati, protezione delle reti, prevenzione e mitigazione delle minacce cyber. Il loro soddisfacimento pone quindi l'organizzazione in una condizione di adeguatezza rispetto ad un livello minimo di consapevolezza e gestione del rischio cyber. Ciò ha lo scopo di incentivare, a livello nazionale, l'avvio di iniziative di adeguamento dei sistemi e delle infrastrutture digitali di enti ed aziende al fine del raggiungimento di un livello diffuso, seppur minimo, di resilienza nei confronti di minacce cyber.

Inventario	1	Esiste ed è mantenuto aggiornato un inventario dei sistemi, dispositivi, software, servizi e applicazioni informatiche in uso all'interno del perimetro aziendale.
	2	I servizi web (social network, cloud computing, posta elettronica, spazio web, ecc. . .) offerti da terze parti a cui si è registrati sono quelli strettamente necessari.
	3	Sono individuate le informazioni, i dati e i sistemi critici per l'azienda affinché siano adeguatamente protetti.
	4	È stato nominato un referente che sia responsabile per il coordinamento delle attività di gestione e di protezione delle informazioni e dei sistemi informatici.
Gov.	5	Sono identificate e rispettate le leggi e/o i regolamenti con rilevanza in tema di cybersecurity che risultino applicabili per l'azienda.
Malw.	6	Tutti i dispositivi che lo consentono sono dotati di software di protezione (antivirus, antimalware, ecc...) regolarmente aggiornato.
Password e account	7	Le password sono diverse per ogni account, della complessità adeguata e viene valutato l'utilizzo dei sistemi di autenticazione più sicuri offerti dal provider del servizio (es. autenticazione a due fattori).
	8	Il personale autorizzato all'accesso, remoto o locale, ai servizi informatici dispone di utenze personali non condivise con altri; l'accesso è opportunamente protetto; i vecchi account non più utilizzati sono disattivati.
	9	Ogni utente può accedere solo alle informazioni e ai sistemi di cui necessita e/o di sua competenza.
Form. e cons.	10	Il personale è adeguatamente sensibilizzato e formato sui rischi di cybersecurity e sulle pratiche da adottare per l'impiego sicuro degli strumenti aziendali (es. riconoscere allegati e-mail, utilizzare solo software autorizzato, . . .). I vertici aziendali hanno cura di predisporre per tutto il personale aziendale la formazione necessaria a fornire almeno le nozioni basilari di sicurezza.
Prot. dati	11	La configurazione iniziale di tutti i sistemi e dispositivi è svolta da personale esperto, responsabile per la configurazione sicura degli stessi. Le credenziali di accesso di default sono sempre sostituite.
	12	Sono eseguiti periodicamente backup delle informazioni e dei dati critici per l'azienda (identificati al controllo 3). I backup sono conservati in modo sicuro e verificati periodicamente.
Reti	13	Le reti e i sistemi sono protetti da accessi non autorizzati attraverso strumenti specifici (es: Firewall e altri dispositivi/software anti-intrusione).
Prev. e mit.	14	In caso di incidente (es. venga rilevato un attacco o un malware) vengono informati i responsabili della sicurezza e i sistemi vengono messi in sicurezza da personale esperto.
	15	Tutti i software in uso (inclusi i firmware) sono aggiornati all'ultima versione consigliata dal produttore. I dispositivi o i software obsoleti e non più aggiornabili sono dismessi.

Tabella 1. Controlli essenziali di cybersecurity del CINI.

Il gruppo di cybersecurity dell'Università Politecnica delle Marche, in collaborazione con il CIS Sapienza, ha elaborato un questionario finalizzato a verificare il soddisfacimento di tali controlli essenziali da parte di enti ed aziende [10]. Tale questionario prevede, per ciascun controllo essenziale, la possibilità di indicare uno di tre livelli di soddisfacimento: totale, parziale o nullo. Il questionario è stato diffuso presso enti ed aziende, ed al momento della redazione di questo documento sono state raccolte circa 60 risposte. Per ciascuno dei quindici controlli essenziali, i dati raccolti in termini di percentuali di ciascuno dei tre esiti possibili sono mostrati in Figura 5.

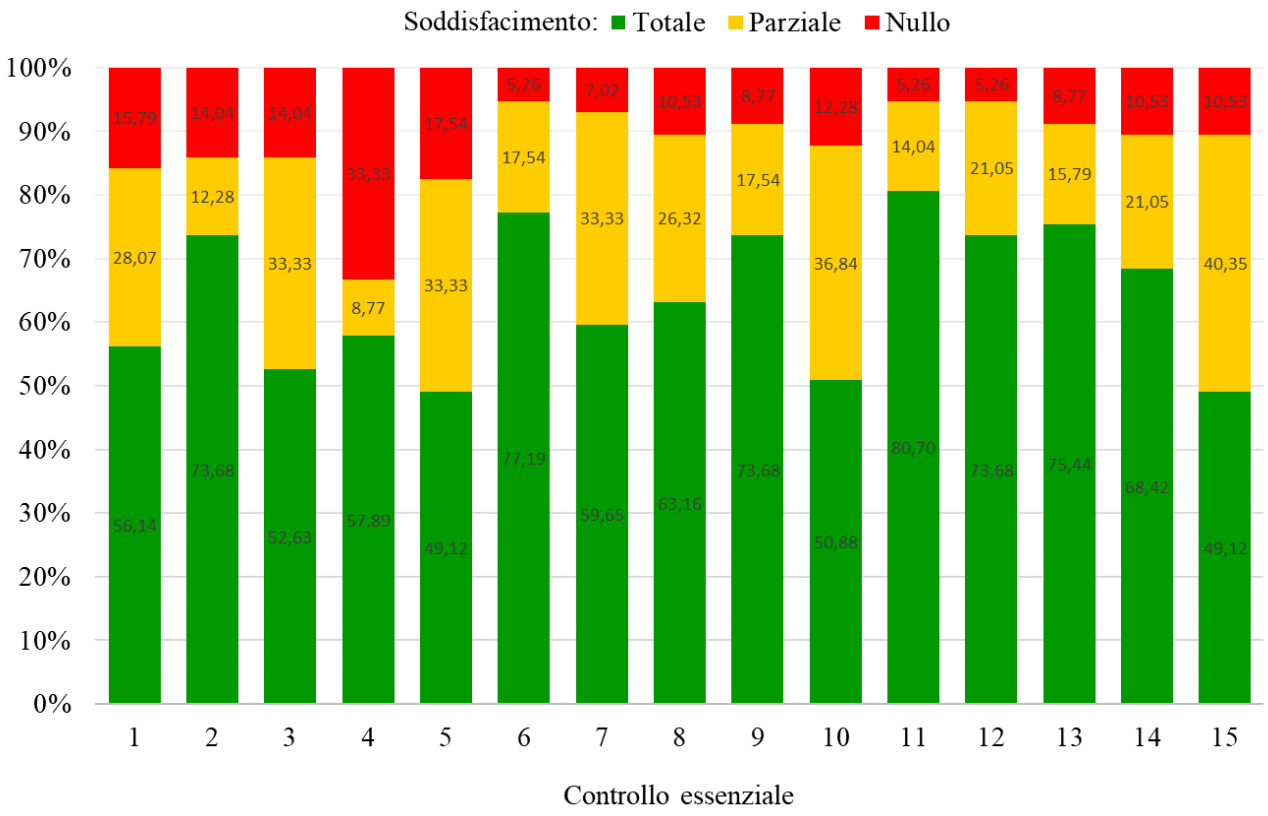

Figura 5. Risultati questionario sul soddisfacimento dei controlli essenziali.

Dalla figura si evince che, praticamente per tutti i controlli essenziali, risulta un soddisfacimento totale che si attesta almeno intorno al 50% del campione studiato. Cionondimeno, le percentuali di soddisfacimento parziale o addirittura nullo di molti controlli essenziali raggiungono livelli allarmanti. Ad esempio, per quanto riguarda il controllo numero 4, relativo alla nomina di un referente che sia responsabile per il coordinamento delle attività di gestione e protezione delle informazioni e dei sistemi informatici, la percentuale di soddisfacimento nullo è superiore al 30%.

Tale risultato appare particolarmente allarmante se visto alla luce della necessità di applicare, a partire dal 25 Maggio 2018, la nuova normativa europea sulla protezione dei dati (GDPR, Regolamento Europeo 2016/679), entrata in vigore già dal 24 Maggio 2016. Essa infatti impone un cambiamento di prospettiva rispetto alle precedenti regolamentazioni in materia di privacy, passando da un approccio basato sulla sola protezione dei dati ad uno più strutturato basato sulla governance dei dati. Ciò passa per la designazione di precisi soggetti e relativi ruoli nell'ambito della gestione dei dati dei cittadini europei, ovvero il titolare del trattamento ed il responsabile del trattamento, ai quali è affiancata la nuova figura del responsabile della protezione dei dati (DPO - Data Protection Officer). Pertanto, la governance dei dati e l'individuazione di tali figure riveste un ruolo di massimo rilievo non solo ai fini del raggiungimento di livelli minimi di cybersecurity, ma anche ai fini del rispetto della nuova normativa europea sulla protezione dei dati.

Per quanto riguarda la Pubblica Amministrazione, il framework nazionale di cybersecurity ha fornito ispirazione alla definizione, da parte dell'agenzia per l'Italia Digitale (AgID), di una serie di misure minime per la sicurezza ICT, che fungono da riferimento normativo e prassi metodologiche per il raggiungimento di livelli di sicurezza informatica adeguati per tutte le pubbliche amministrazioni [4]. I criteri che sono alla base di tali misure minime sono fondati sul framework nazionale di cybersecurity, ed il termine per la loro applicazione è stato fissato al 31 Dicembre 2017. Una delle prescrizioni previste dalle misure minime AgID consiste nella compilazione del "Modulo di implementazione delle misure minime di sicurezza per le pubbliche amministrazioni", che deve essere conservato e regolarmente aggiornato da ciascuna amministrazione

pubblica soggetta a tale regolamentazione. Le istituzioni universitarie pubbliche, in quanto pubblica amministrazione, rientrano nel campo di applicabilità di tali misure minime, ed alcuni atenei si sono già attivati ai fini della loro implementazione. Certamente il contesto accademico presenta delle peculiarità e criticità rispetto ad altre tipologie di pubblica amministrazione, non fosse altro per le caratteristiche peculiari del mondo della ricerca e dei relativi flussi informativi. Pertanto un'attività indispensabile all'interno degli atenei sarà quella di individuare aree omogenee di applicazione delle misure in questione, così come definire opportune procedure interne che coinvolgano le varie figure che condividono le responsabilità tecniche ed organizzative all'interno dell'ateneo (governance, referenti informatici, responsabili di laboratorio, etc.).

La situazione degli atenei nei confronti di questi processi e la loro velocità di adeguamento alle mutate condizioni operative e regolamentari meritano sicuramente attenzione. Dall'indagine svolta dalla CRUI sulla transizione al digitale [13] e l'applicazione di norme e regolamentazioni, nell'ambito della quale sono stati interpellati circa 60 atenei, emerge che nella maggior parte dei casi non è stato ancora nominato un responsabile della transizione al digitale e che poco più della metà degli atenei interpellati ha implementato le misure minime di sicurezza AgID. Parimenti, nella maggior parte degli atenei intervistati il DPO non è stato ancora nominato o è in fase di individuazione. Circa pari al 50% è inoltre la percentuale di atenei in cui si sono completati i regolamenti per l'attuazione della GDPR e si è predisposto il software per la gestione dei registri relativi al trattamento dei dati personali.

Per quanto concerne le recenti evoluzioni del quadro regolatorio europeo, va osservato che il 16 Maggio 2018 il Consiglio dei Ministri ha approvato il Decreto Legislativo per il recepimento e l'attuazione in Italia della direttiva europea NIS (Network & Information Systems) [11]. Sebbene sia prevista la possibile estensione dell'ambito di applicazione della direttiva a settori diversi da quelli elencati nella direttiva stessa, il Governo ha scelto di limitarne l'applicazione solo agli ultimi, ovvero: energia, trasporti, banche, mercati finanziari, sanità, fornitura e distribuzione di acqua potabile e infrastrutture digitali, motori di ricerca, servizi cloud e piattaforme di commercio elettronico. Il sistema universitario è quindi tendenzialmente escluso da tale obbligo, mentre lo stesso non può dirsi per altri enti pubblici che rientrino nei settori sopra elencati. Esistono tuttavia settori, come la sanità in quanto servizio essenziale, che rientrano nell'ambito di applicazione della direttiva stessa e che pure includono parte del contesto accademico. In linea generale, oltre ad obblighi specifici, tale direttiva prevede la definizione e l'implementazione di una strategia e di una visione a livello nazionale relativamente alla gestione della sicurezza di reti e sistemi digitali.

In conclusione, la information governance riveste un ruolo fondamentale e strategico nel presente e soprattutto nel futuro del tessuto accademico nazionale, così come della società tutta, sia per via di spinte interne dovute alla crescente dimensione ed al crescente valore degli asset digitali, sia per via di spinte esterne dovute al crescente numero di regolamentazioni e direttive nazionali ed europee. L'investimento in risorse umane, prima ancora che tecnologiche, rappresenta sicuramente il primo elemento di criticità con cui confrontarsi, essendo il fattore umano il primo elemento di vulnerabilità di reti e sistemi digitali ed essendo la richiesta di personale altamente qualificato in continua crescita. Per ciò che riguarda l'accademia, il laboratorio nazionale di cybersecurity del CINI, nel libro bianco 2018 [6], individua la necessità della formazione di una workforce (tecnici, ingegneri, esperti, ricercatori) importante e distribuita sul territorio nazionale ed auspica la definizione di un piano straordinario per l'assunzione di ricercatori e professori universitari del settore.

4. Conclusioni

La realizzazione di un'infrastruttura nazionale che supporti l'elaborazione digitale ha contribuito in modo determinante allo sviluppo del sistema della Ricerca nazionale, offrendo una piattaforma abilitante ai ricercatori di Atenei ed enti di Ricerca. Lo sviluppo e la manutenzione di questa infrastruttura è essenziale

per mantenere competitivo il sistema della Ricerca nazionale, soprattutto in un panorama in cui il calcolo scientifico permea ormai tutte le discipline.

L'infrastruttura fisica va sviluppata ed adeguata in termini di capacità e di capillarità, in modo da poter fronteggiare le sempre maggiori quantità di dati elaborati e trasferiti, e consentire di supportare la realizzazione di sistemi basati su tecnologia IoT. Unitamente all'infrastruttura fisica è necessario sviluppare quella virtualizzata, investigando come la capacità di renderla più versatile mediante approcci software-defined possa contribuire ad aprire nuovi scenari ed accelerare le ricerche.

Tali infrastrutture dovranno trovare nuovi equilibri necessari a garantire da una parte la conformità normativa relativamente alle tematiche digitali, e dall'altra assicurare quell'apertura e quella flessibilità che sono alla base della Ricerca.

Bibliografia

[1] Parlamento e Consiglio Europeo, Regolamento (UE) n. 2016/679 del 27 Aprile 2016 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati (regolamento generale sulla protezione dei dati).

[2] Decreto del Presidente del Consiglio dei Ministri 24 Gennaio 2013, "Direttiva recante indirizzi per la protezione cibernetica e la sicurezza informatica nazionale".

[3] Italian Cyber Security Report 2015, "Un Framework nazionale per la Cyber Security", a cura di R. Baldoni e L. Montanari, http://www.cybersecurityframework.it/sites/default/files/CSR2015_web.pdf

[4] Agenzia per l'Italia Digitale, circolare del 18/04/2017 n. 2/2017, recante "Misure minime di sicurezza ICT per le pubbliche amministrazioni".

[5] Italian Cybersecurity Report 2016 – Controlli Essenziali di Cybersecurity, a cura di R. Baldoni, L. Montanari e L. Querzoni, http://www.cybersecurityframework.it/sites/default/files/csr2016web.pdf

[6] Laboratorio nazionale di cybersecurity del CINI, libro bianco "Il Futuro della Cybersecurity in Italia: Ambiti Progettuali Strategici – Progetti e Azioni per difendere al meglio il Paese dagli attacchi informatici", a cura di R. Baldoni, R. De Nicola, P. Prinetto, Febbraio 2018.

[7] Rapporto Clusit 2018 sulla sicurezza ICT in Italia, disponibile su richiesta all'indirizzo https://clusit.it/rapporto-clusit/

[8] https://www.consorzio-cini.it/index.php/en/labcs-home/education-in-cyber-security-in-italy

[9] National Institute of Standards and Technology, Framework for Improving Critical Infrastructure Cybersecurity, http://www.nist.gov/cyberframework/

[10] Gruppo cybersecurity dell'Università Politecnica delle Marche, questionario sui controlli essenziali di cybersecurity del CINI, http://cybsec.univpm.it/controlli-essenziali-di-cyber-security

[11] Parlamento e Consiglio Europeo, Direttiva (UE) 2016/1148 del 6 Luglio 2016, recante "misure per un livello comune elevato di sicurezza delle reti e dei sistemi informativi nell'Unione".

[12] NVMe fabric overview http://www.nvmexpress.org/wp-content/uploads/NVMe_Over_Fabrics.pdf

[13] Stato di attuazione del GDPR e di altre normative ICT negli Atenei e negli enti di Ricerca italiani al 5/5/2018. http://bit.ly/gdpr-crui