

CNAF – Reloaded

WP 2.1 Datacenter – Cloud SGSI

Barbara Martelli

10/12/2020

Outline



1- status quo (il punto di partenza, con punti di forza e debolezza)



2- evoluzione dei requirement



3- proposte tecnologiche in prospettiva 2021-2025



4- stima di costi, personale e rischi



0- Cos'è un SGSI



Cos'è un SGSI

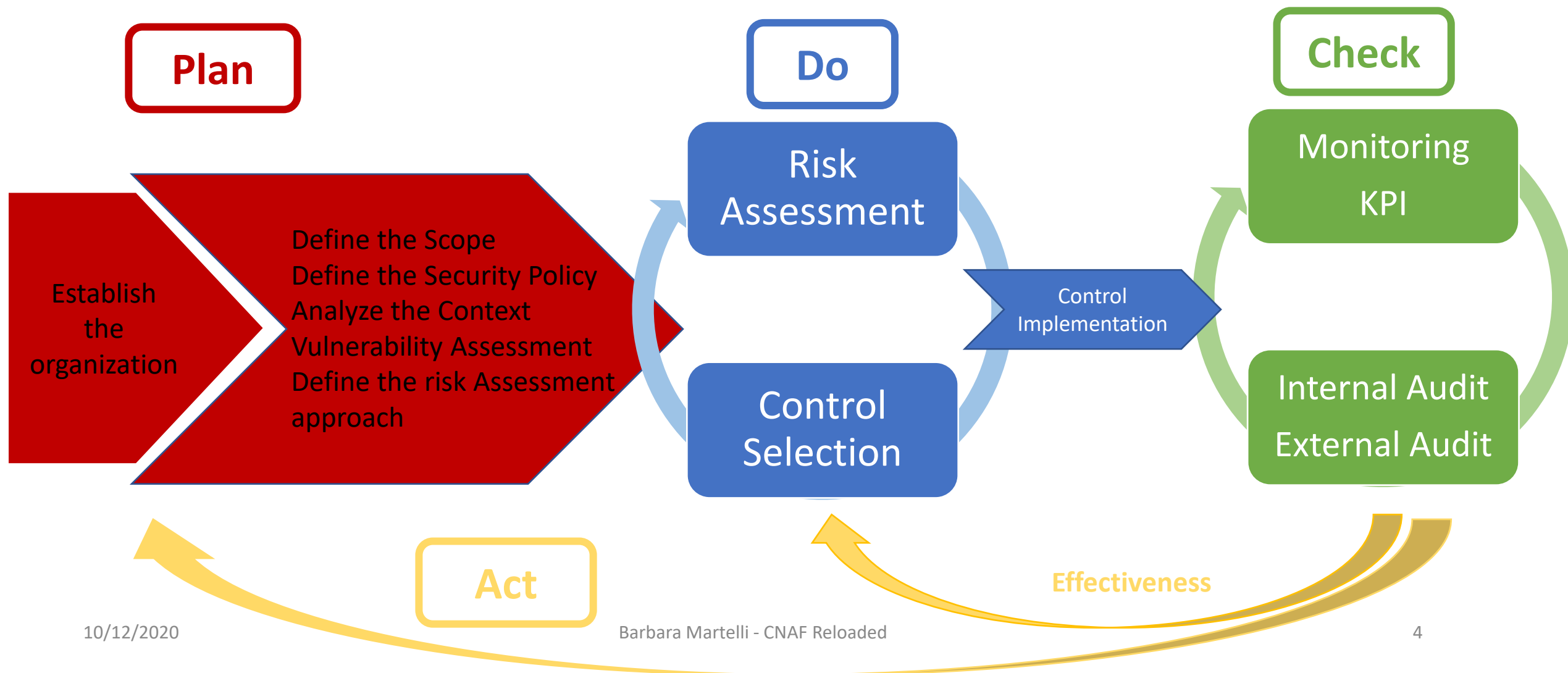
ISO/IEC 27001 Annex A

- Sistema di Gestione della Sicurezza delle Informazioni (ISMS in inglese)
https://en.wikipedia.org/wiki/Information_security_management
- è un **framework organizzativo** che collega tutti gli elementi che impattano la sicurezza delle informazioni, in modo da assicurare che le policy, i processi, le procedure e gli obiettivi di sicurezza siano implementati, comunicati, valutati e **migliorati** nel tempo -> **ciclo di Deming**
- è centrato sul **risk assessment** -> tutte le decisioni vengono prese in seguito ad una **valutazione dei rischi**
- è finalizzato al raggiungimento degli **obiettivi dell'organizzazione** in termini di requisiti di **sicurezza** delle informazioni, in modo **sostenibile**
 - Sicurezza delle informazioni = RID = Riservatezza + Integrità + Disponibilità
 - Sostenibile = al minor costo possibile + capacity management





Ciclo di Deming e Risk Assessment





L'SGSI del CNAF - Genesi



- Nato nel 2017 dall'esigenza del Progetto IMI2 Harmony - Healthcare Alliance for Resourceful Medicine Offensive against Neoplasms in Hematology www.harmony-alliance.eu
 - Ospitare la big data platform Harmony per l'analisi di dati genomici
 - Requisito: gestire la piattaforma adottando un Sistema di Gestione della Sicurezza delle Informazioni certificato secondo lo standard internazionale ISO/IEC 27001
- Nel 2020 è stato aggiunto lo use case di Alleanza Contro il Cancro
 - Contratto di 24 mesi per la fornitura di un servizio IaaS + consulenza su security, sync&share e AuthN/AuthZ per la gestione di dati genomici



Certificazione ISO/IEC 27001



- Un SGSI può esistere ed essere efficace anche *senza essere certificato*
- Il CNAF ha certificato il suo Sistema perchè questo era un requisito preciso del Progetto Harmony
- Scope del certificato attualmente in vigore: ***“Hosting di sistemi fisici e virtuali per la conservazione e l’accesso a dati biomedici e gestione applicativi di analisi dati finalizzati alla ricerca in campo biomedico/genomico”***
 - [Link al certificato CNAF](#)
- A cosa serve il certificato: dimostra fattivamente agli stakeholder **trasparenza** e **affidabilità** nel garantire i requisiti di sicurezza e la **conformità** alle leggi, ai regolamenti e ai contratti/accordi stipulati
 - Ci si sottopone ad audit annuali svolti da **organismi terzi e indipendenti**
 - Ogni tre anni si rinnova il certificato, affrontando una revisione completa svolta da organismi terzi e indipendenti



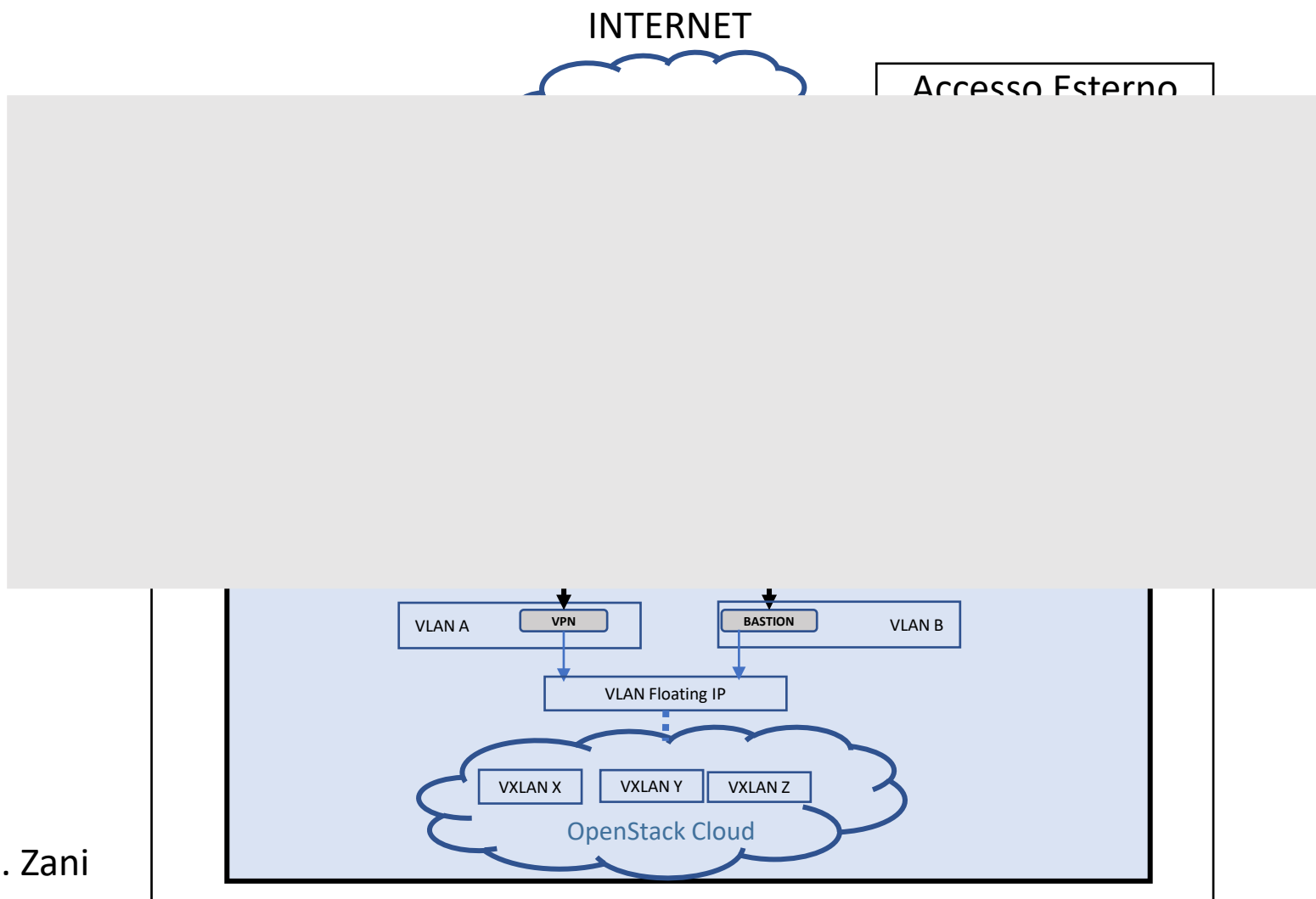
Perimetro fisico



- Due rack al Tier-1 Sala 1
- Un server per il backup secondario
- Terzo backup presso altra sede INFN



Schema logico di rete oggi



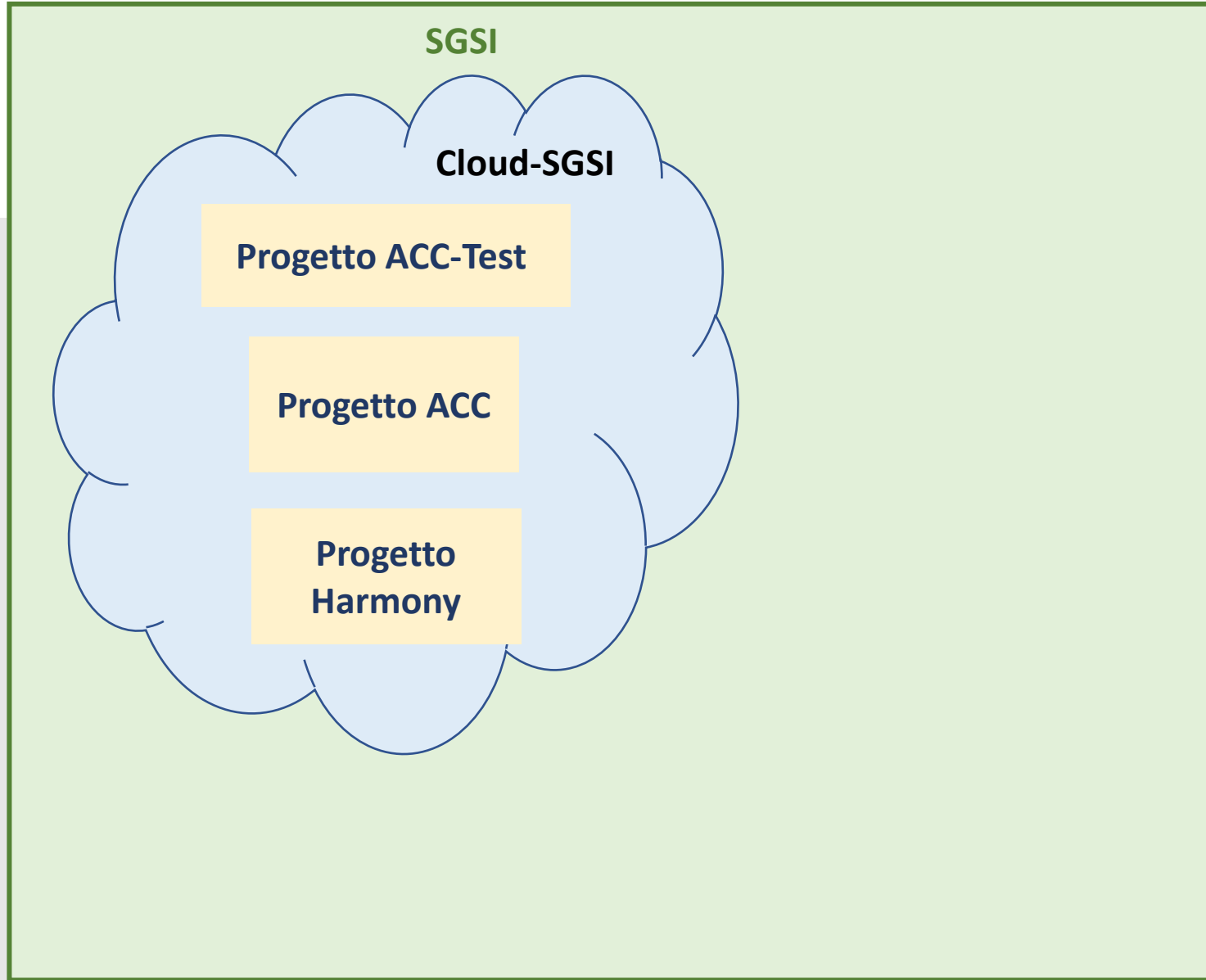
S. Zani

Servizi utilizzati da SGSI e dipendenze esterne



-  SGSI VM/Service (oVirt)
-  SGSI VM/Service
-  VM/Service Extra SGSI

Extra CNAF





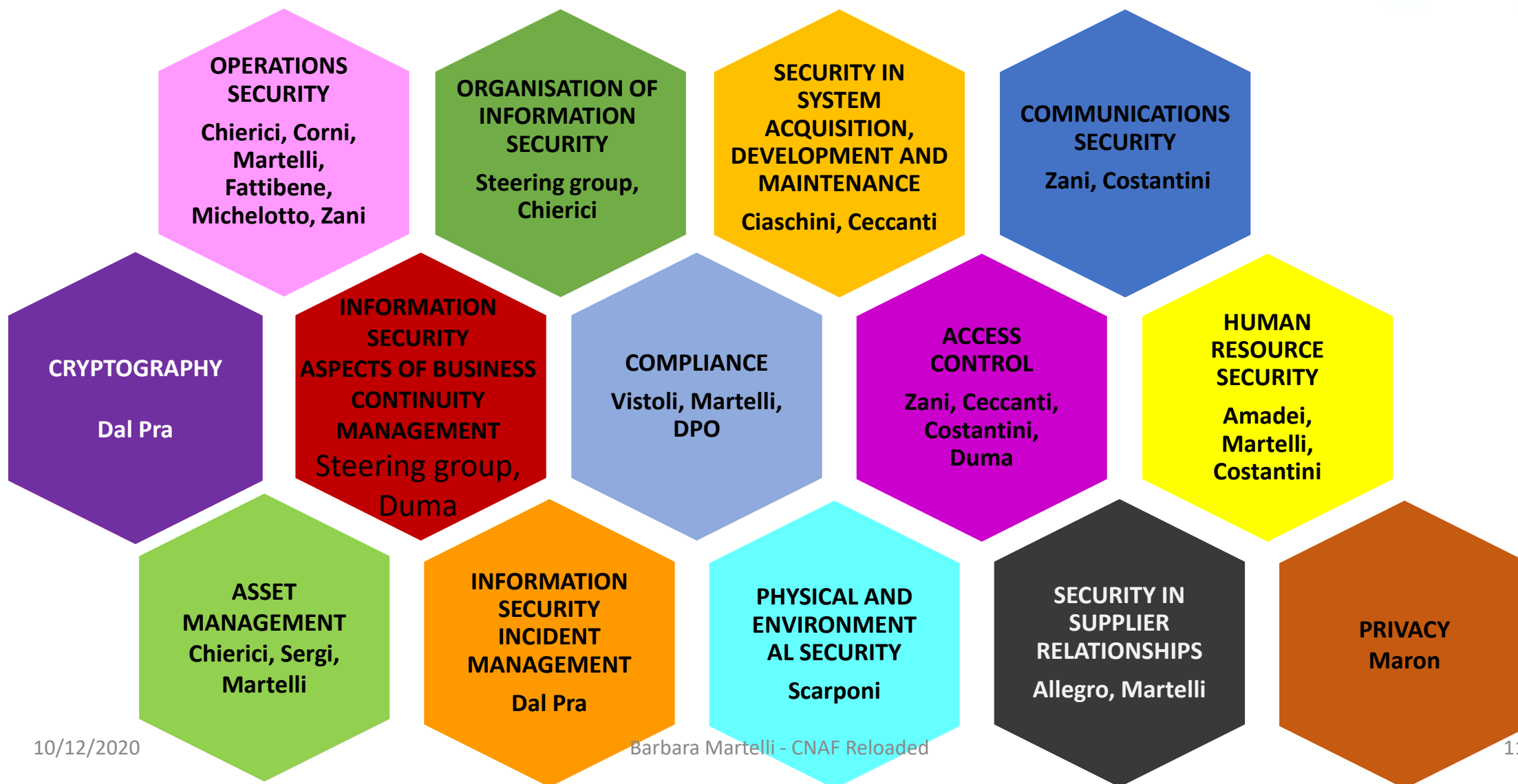
Organizzazione



- **Steering Group (Maron, Salomoni, Vistoli, dell'Agnello, Cesini, Martelli):**
 - Ha l'autorità di allocare le risorse materiali e umane necessarie per raggiungere gli obiettivi di sicurezza
 - Ha la responsabilità di garantire il mantenimento della sicurezza delle informazioni in modo continuo, anche in situazioni avverse (incidenti, disastri, crisi)
- **Responsabile SGSI (Martelli):**
 - Ha la responsabilità di stabilire, implementare, mantenere e migliorare il SGSI e assicurare il rispetto dei Requisiti di Sicurezza.
 - Si confronta mensilmente con lo Steering Group per assicurarsi che vengano allocate le risorse e vengano attribuite le priorità necessarie a raggiungere gli obiettivi di sicurezza
 - Riferisce mensilmente sia alla Direzione (Steering Group) che internamente sulle performance del SGSI
 - Coordina e gestisce i processi di risk assessment, audit interni, security incident
- **Security Group (Cesini, Chierici, Ciaschini, Duma, Zani, Fattibene, Martelli, Scarponi):**
 - Ha la responsabilità di assicurare che l'SGSI sia conforme ai requisiti della ISO 27001 e alle norme e leggi applicabili
 - Approva tutti i documenti che definiscono o modificano Requisiti di Sicurezza
- **Security Coordinator (Ciaschini):**
 - Coordina il Security Group
 - Riferisce al responsabile dell'SGSI e allo Steering Group sullo stato della sicurezza
- **Responsabili Obiettivi di Controllo:** hanno la responsabilità di definire e implementare i controlli dell'Annex A dello standard. **Matrice RACI** (Responsible, Accountable, Consulted, Informed) per ogni area di responsabilità



Responsabili obiettivi di sicurezza





Attività in corso



Terminato il primo ciclo triennale



Avviato a giugno 2020 un progetto interno per

- **aggiungere i controlli** previsti per gli ambienti cloud che trattano dati personali (ISO/IEC 27017 e ISO/IEC 27018);
- **ampliare lo scope** del certificato per includere **software e applicazioni** non strettamente legate all'analisi dei dati genetici
 - Dati clinici, dati personali
 - Dati confidenziali (per esempio collegati ad attività di collaborazione con imprese sotto NDA)
- **ampliare lo scope** per includere la possibilità di aggiungere **altri data center**.



Scope del nuovo certificato “Servizi IaaS, PaaS e SaaS per la gestione di dati biomedici e genomici. I servizi in scope saranno specificati alla pagina: <https://www.infn.it/compliance/services-in-scope>. I servizi in scope possono essere erogati presso i centri di calcolo specificati alla pagina: <https://www.infn.it/compliance/locations-in-scope>”

Ad oggi i servizi in scope sono:

1. Servizio IaaS per la gestione di dati biomedici e genomici, sede: CNAF – Cloud SGSI
2. Servizio IaaS, sistema di autenticazione e autorizzazione, sistema “sync&share”, sede: CNAF – Cloud SGSI



Audit esterno schedato nei giorni 21,22,25,26,27,28 gennaio 2021



Punti di forza e debolezza



- ⊕ Grandissimo **sforzo organizzativo** verso:
 - ⊕ Chiarezza su responsabilità e azioni da fare
 - ⊕ **accountability**
 - ⊕ Procedure definite e documentate su change management, user access, project onboarding, incident management, security in supplier relationships, ecc...
 - ⊕ **Robustezza** e **sicurezza** del deployment
 - ⊕ Introduzione della metodologia di risk assessment
 - ⊕ definizione più **razionale** delle priorità
- ⊕ Progetto **orizzontale** che coinvolge tutte le UF del CNAF
 - ⊕ Occasione di **collaborazione** e scambio di conoscenze
- ⊕ Maggiore **garanzia dei requisiti** di Riservatezza, Integrità e Disponibilità
- ⊖ Necessaria **maggiore integrazione** nelle attività del CNAF
 - ancora casi di duplicazione di lavoro
 - le persone collaborano, ma senza una percentuale di tempo chiaramente definita
- ⊖ Progetto **orizzontale** che coinvolge tutte le UF del CNAF
 - **Difficoltà nel coordinamento** delle attività



2- evoluzione dei requirement



Richieste da parte di possibili use case con forti requisiti RID



- Harmony Plus (evoluzione di Harmony)
- Health Big Data Project (evoluzione di ACC) obiettivo specifico di questo progetto è la creazione di una *piattaforma tecnologica* che consenta la raccolta, condivisione ed analisi di dati clinici e scientifici dei pazienti di ciascun IRCCS
- Spallanzani (Covid-19)
- PLANET (CSN5 2021)
- San Raffaele
- Regina Elena
- SUPER
- Excalate4Covid
- Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori Meldola
- Molte richieste di collaborazione per **co-creazione di piattaforme di analisi di dati medici** con toolset ad hoc
- Molti bandi europei aperti ai quali noi per ora abbiamo sempre **reagito** -> in futuro potremmo affrontarli in maniera più **proattiva**



Dettaglio richieste per PLANET/Spallanzani/CovidStat (prime stime)



	PLANET	Spallanzani	CovidStat
Storage	1TB / 2 TB totali		
CPU-Servizi	16 cores totali (Va incluso IAM)		
RAM-Servizi	64 GB Totali ((va incluso IAM)		
CPU-analisis	100 Cores	100 Cores (Stima)	128
RAM-analysis	250 GB	500GB	non ancora specificato
GPU	NON in prima priorità	NON in prima priorità	si

Servizi

MinIO

JupyterHUB

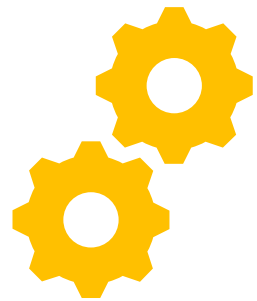
IAM

K8S (da valutare, dipende dai nodi a disposizione)

IP pubblici 6

Backup via TSM (capire se necessarie licenze ulteriori o se utilizzabile attuale client condiviso con altri progetti)

VPN oppure possibilità di esporre i servizi -> da valutare



3- proposte tecnologiche in prospettiva 2021-2025



Linee guida generali



Linee guida generali:

- **Uniformità**

- riutilizzare il più possibile gli **strumenti e le tecnologie scelte dal CNAF** -> no creazione di ulteriore silos
- Policy e procedure come “specificazione” delle policy e procedure CNAF -> vedere esempio slide 25

- **+ efficacia – formalità**

- Limitare al massimo la burocrazia: procedure il più possibile **automatizzate** attraverso i tool già in utilizzo, audit e review il più possibile automatiche (cron jobs, task jira ricorrenti, ecc), registri aggiornati in modo automatico (log, db, inventory, ecc), **documentation as a code**, ...

- **Integrazione**

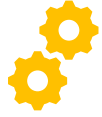
- ove possibile, gestire gli “esperimenti” SGSI come tutti gli altri use case del CNAF



Possibili evoluzioni (1/4)

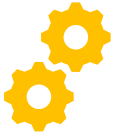


- **Uniformare** il servizio di supporto utenti includendo Cloud SGSI tra gli “esperimenti” gestiti da **User Support**
 - Attualmente si utilizza **RedMine**, utile convergere su uno stesso tool CNAF
- Gestione dello **Storage unificata** sotto il gruppo Storage del Tier-1?
 - Attualmente **Ceph**
- Inclusione della Cloud SGSI in un **unico sistema di monitoring CNAF**?
 - Attualmente in uso monitoring basato su **Sensu/InfluxDB/Grafana**
- **Estensione** delle **metodologie organizzative** SGSI al resto del CNAF?
 - Collaborazione con **WP1**?
- **Federazione** della Cloud SGSI con Cloud@CNAF o con INFN Cloud? **T2.2**
 - L'AuthN/AuthZ implementata è già ispirata al modello WLCG (**IAM/keycloak/FreeIPA**)



Possibili evoluzioni (2/4)

- Sfruttare **Legnaro** come ulteriore sito di Business Continuity e/o Disaster Recovery?
- Adottare un tool di **Asset Inventory**
- Valutare se e quanto sia possibile **astrarre il perimetro** dell'SGSI per **ottimizzare** l'uso delle risorse
 - VM e/o container orchestrati su hypervisor diversi?
 - Rack chiusi singolarmente? Aree segregate? TVCC? Quanto la segregazione fisica è efficace? Collaborazione con **WP3**
 - **Segregazione** di tenant e reti **a che livello?**
 - Valutare un'architettura **Zero Trust** (network/security)?



Possibili evoluzioni (3/4)

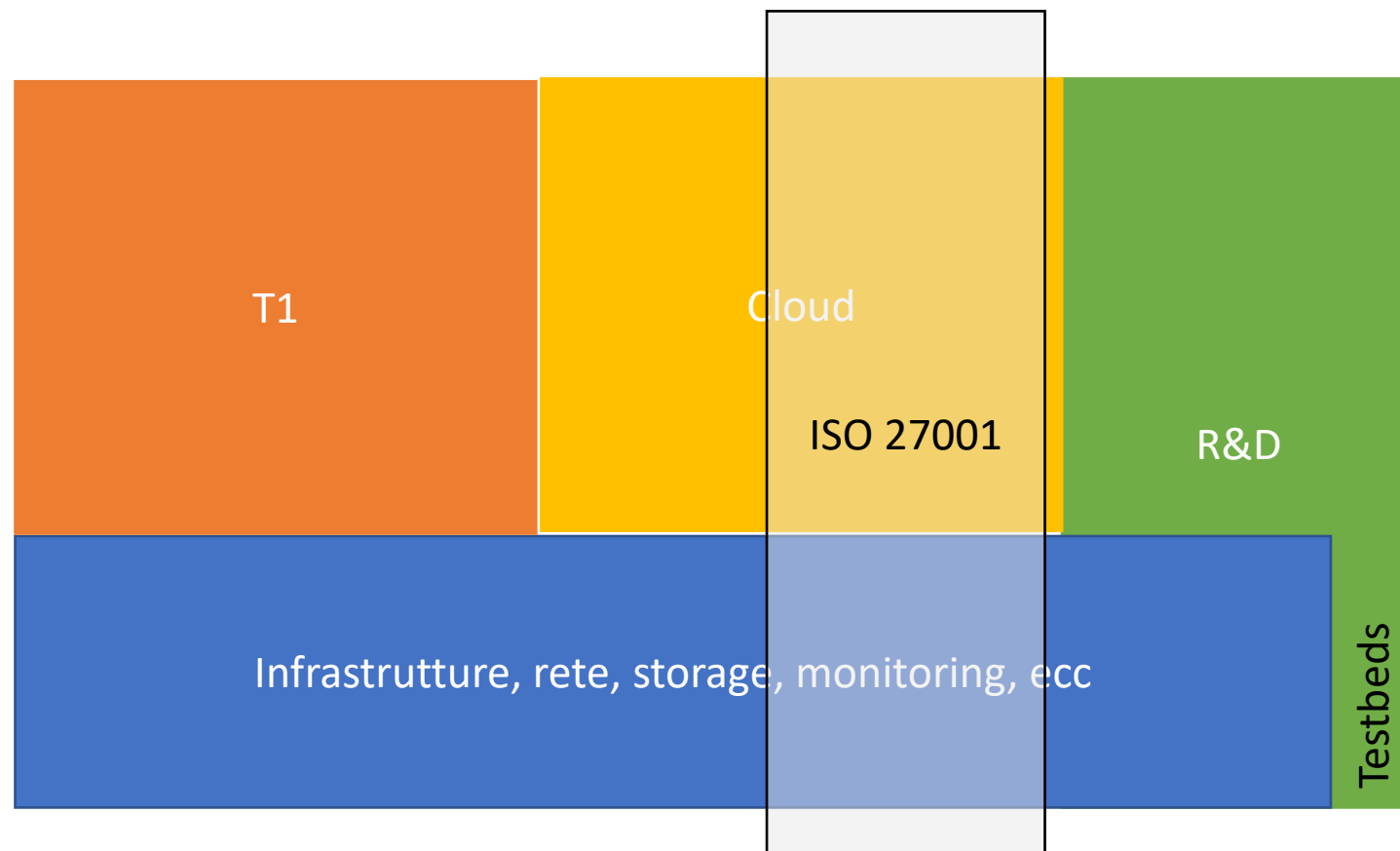
- Utilizzare un SIEM (Security Information and Event Management) per analizzare velocemente gli **alert** di sicurezza, gestire gli **incidenti** e gli **eventi** di sicurezza -> **T2.3** vedere presentazione Ciaschini
- (ISO/IEC 27017) Verificare contratti stipulati per i tool di terze parti utilizzati -> insieme a **WP1?**
 - verificare la presenza e coerenza con SGSI delle **clausole standard** per il cloud computing
 - mitigare i rischi contrattuali (vendor lock-in in primis)
 - Limits to provider power of modification, acceptable use of the service-breaches of contract, security, privacy, data erasure, **databack delivery**, **access data after contract termination**, customer liability-idemnification, provider liability-idemnification, **data portability**, SLA
- Verificare contratti di assistenza per l'hardware (ISO/IEC 27001) -> insieme a **WP1?**

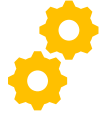


Possibili evoluzioni (4/4)

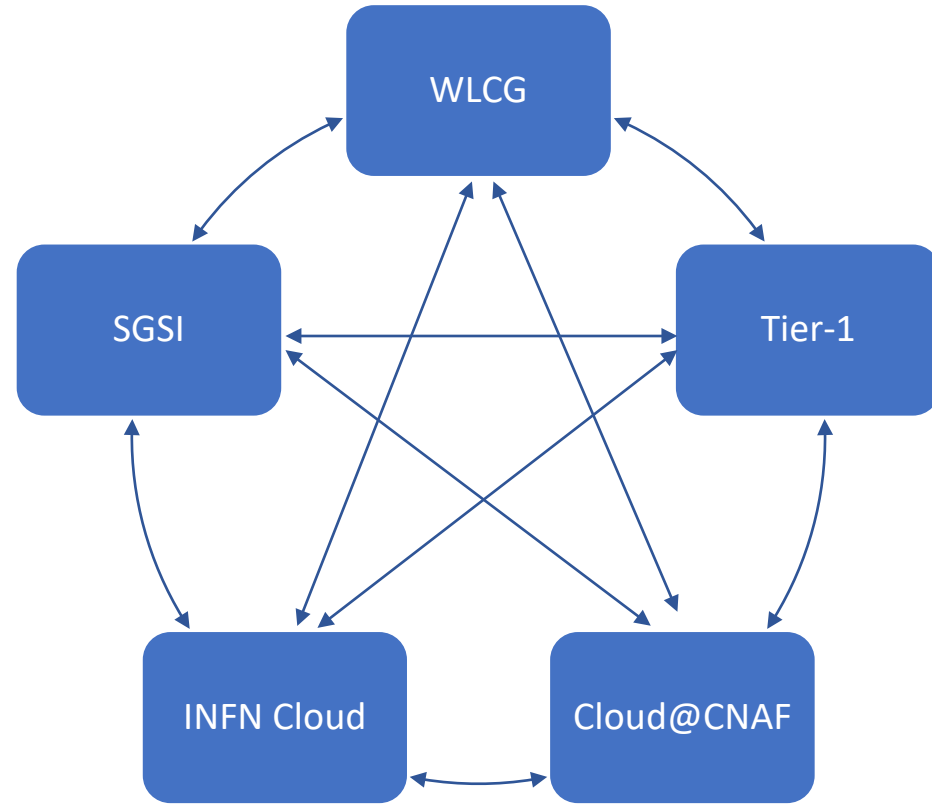
Modello organizzativo integrato

- Procedure il più possibile uniformi
- Procedure ISO ereditate da quelle CNAF con modifica dei soli aspetti necessari a garantire gli ulteriori requisiti RID





Proposta operativa



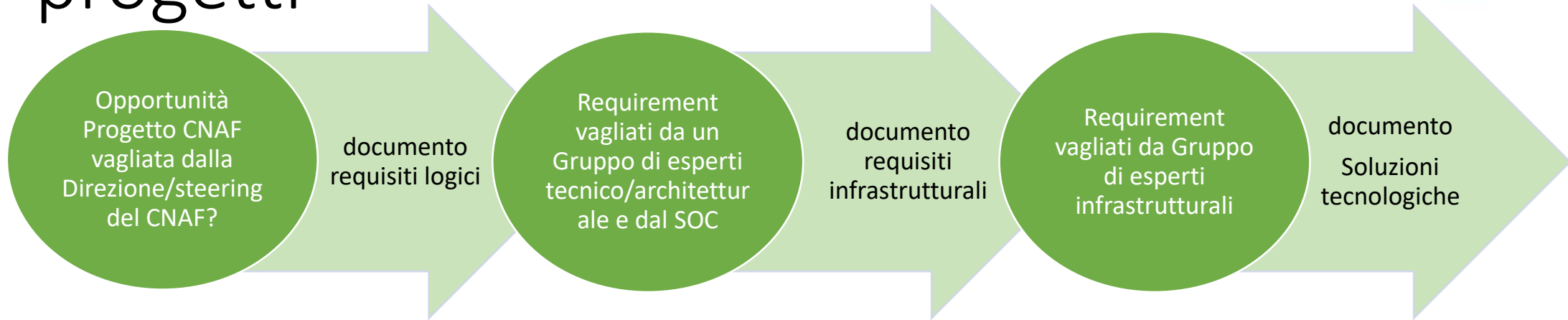
- Dove ci siano policy o procedure INFN, CNAF, WLCG, INFN Cloud già approvate
 - Partire da esse
 - Creare policy e procedure specifiche solo se necessario a garantire gli ulteriori requisiti di sicurezza
- Dove non ci siano policy o procedure INFN, CNAF, WLCG, INFN Cloud già approvate
 - Crearle per SGSI
 - Verificare con il resto del CNAF come generalizzarle all'intero centro
- Ove possibile fare il **merge** e **razionalizzare**



Esempio di processo di acquisizione progetti

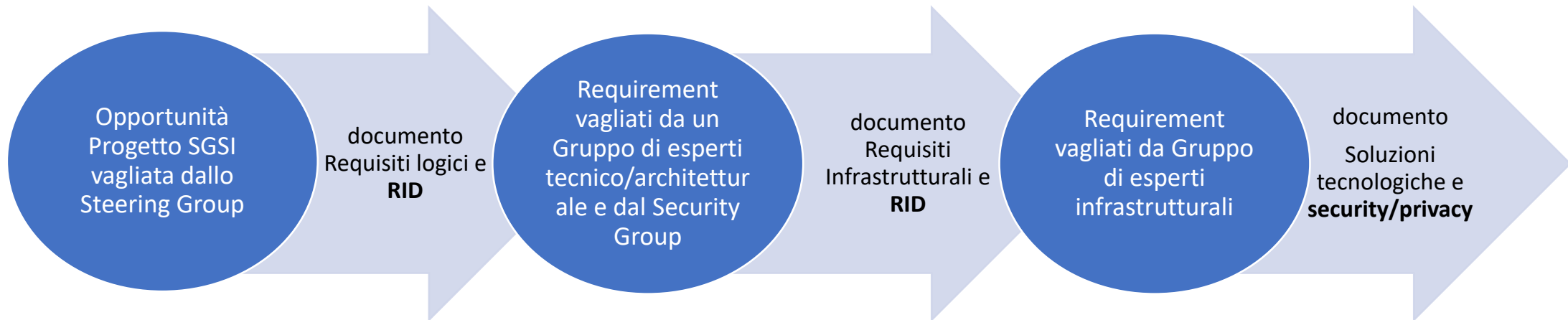


Ambito Etra SGSI



Unica differenza: in ambito SGSI i requisiti RID devono essere considerati in ogni fase

Ambito SGSI





4- Stima dei costi, personale, rischi



4- Stima FTE

Il setup dell'organizzazione richiesta per gestire un SGSI certificato è molto oneroso

PERÒ

1. Gran parte del lavoro da svolgere per SGSI è **sovrapponibile** a quello necessario a costruire processi e procedure per il tecnopolo
2. Gran parte del personale CNAF è **già coinvolto** in SGSI

Necessario distinguere tra “fare” e “sapere”

- Formazione continua su
 - modalità di lavoro secondo processi e procedure ISO 10 giorni/anno per ogni afferente
 - Tecniche/tecnologie in ambito security/privacy 15 giorni/anno per ogni afferente -> sinergia con **T2.3**

Approccio il più possibile **personalizzato** e *learning-by-doing*

- Attività prettamente SGSI difficile da quantificare ora -> nell'ipotesi di implementazione di un modello organizzativo integrato l'effort specifico per SGSI sarà più basso



Rischi (1/3)

Dipendenza da servizi esterni all'SGSI: la robustezza della catena è pari a quella dell'anello più debole, **non è possibile replicare tutti i servizi** all'interno dell'isola SGSI

- Possibile mitigazione 1: durante il vulnerability assessment valutare approfonditamente tutte le **dipendenze da servizi esterni** ed **adeguarli ad un livello di sicurezza accettabile rispetto** ai requisiti di sicurezza delle applicazioni **genomiche**
- Possibile mitigazione 2: riguardo ai tool condivisi, assicurarsi di **includere i requisiti SGSI durante la scelta dei tool**
- Possibile mitigazione 3: implementazione del modello organizzativo integrato





Rischi (2/3)

Incompatibilità della timeline di certificazione con il resto delle deadline che coinvolgono il CNAF

- Possibile mitigazione 1: **creazione di un Gantt CNAF** che includa Gantt SGSI e Gantt degli altri progetti in corso
- Possibile mitigazione 2: implementazione del modello organizzativo integrato





Rischi (3/3)

Creazione di un **silos SGSI** all'interno del nuovo datacenter, in cui le stesse persone devono adottare procedure, strumenti e modalità organizzative diverse, a causa dell'organizzazione verticale a Unità Funzionali

- Possibile mitigazione 1: implementazione del modello organizzativo integrato
- Possibile mitigazione 2: creare **canali di Comunicazione chiari e stabili** tra SGSI e Unità Funzionali/Servizi/Reparti
 - Documentazione aperta a tutto il CNAF
 - Riunioni (ma senza esagerare)
 - Attività trasversali

Questo sforzo è già in atto -> da potenziare



Domande?

