

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

SCUOLA DI SCIENZE
Corso di Laurea in Informatica

DIFFERENTI APPROCCI
ALL'EVOLUZIONE
DI INTERNET

Tesi di Laurea in Informatica

Relatore:
Chiar.mo Prof.
VITTORIO GHINI

Presentata da:
MASSIMO BOLLIANDI

Sessione I
Anno Accademico 2013-2014

a Romina ed ai miei genitori Liliana e Giovanni

Indice

INTRODUZIONE	7
1 COGNITIVE NETWORK	11
1.1 Cos'è una Cognitive Network?	11
1.2 Il Knowledge Plane	12
2 QUALE FUTURO?	19
2.1 Cosa porterà l'internet del futuro all'uomo?	19
2.2 Perché l'internet del futuro sarà datacentrico?	21
2.3 Principi base dell'internet del futuro	23
2.4 Requisiti per la progettazione datacentrica	24
2.5 Architettura datacentrica logica	26
2.6 Architettura datacentrica autonoma	31
3 DIFFERENTI APPROCCI	35
3.1 NDN ed il suo protocollo CCN	35
3.2 PSIRP come base per PURSUIT	38
3.3 NetInf: da 4WARD a SAIL	43
3.4 DONA	47

4	CARATTERISTICHE	53
4.1	Naming e Sicurezza	53
4.2	Risoluzione dei nomi e Routing	54
4.3	Archiviazione in rete	56
4.4	API	57
5	PROPRIETÀ	61
5.1	Scalabilità	61
5.2	Sicurezza	63
5.3	Privacy	64
5.4	Mobilità	65
6	INCENTIVI	67
6.1	Distribuzione	67
6.2	Gli attori	68
6.3	Partecipanti ed investitori	70
7	SFIDE RIMANENTI	75
7.1	Naming	75
7.2	Risoluzione dei nomi	76
7.3	Routing, modifiche alla Topologia e Mobilità	77
7.4	Eterogeneità della rete	78
7.5	L'oggetto informazione	79
8	CONCLUSIONI	81
8.1	Obiettivo del documento	81
8.2	Considerazioni	82
	BIBLIOGRAFIA	86

INTRODUZIONE

Attuale architettura e limiti

Oggi, la maggioranza dell'utilizzo di internet riguarda contenuti e servizi di ricerca, di recupero e distribuzione dei contenuti di streaming ed accesso ai servizi web, in cui l'utente si preoccupa solo dei contenuti e servizi stessi dimenticandosi della loro locazione.

L'utente quindi sa che vuole, ad esempio, nuovi video o informazioni nel formato adatto e di qualità adatte, ma non è interessato su quale macchina risiedono i dati e servizi desiderati finché sa che sono garantite sicurezza e privacy.

Con le funzionalità appena descritte si realizza l'architettura internet attuale: content server o cache (sia professionali sia con contenuti creati da altri utenti) servizi centralizzati, decentralizzati o clustered server, compresi i motori di ricerca.

Ma anche server di supporto, ad esempio server DNS (Domain Name System), server AAA (Authentication, Authorization and Accounting), server DRM (Digital Rights Management), core router, edge router e gateway residenziali, utenti connessi attraverso terminali fissi, mobili o wireless.

Il primo step è la scoperta del contenuto da parte dei motori di ricerca: i motori di ricerca eseguono la scansione in internet per trovare, classificare ed identificare contenuti e servizi o in alternativa gli utenti possono pubblicare i contenuti ed informare i motori di ricerca.

Il secondo step è la scoperta dei contenuti da parte dell'utente: se l'utente non sa dove risiede il contenuto, interroga un motore di ricerca ed ottiene come risultato una lista di URL (Uniform Resource Locator) in cui il contenuto è memorizzato.

L'ultimo step è la distribuzione dei contenuti (streaming): l'utente seleziona un URL ed il contenuto viene consegnato in streaming e, nel caso di servizi live, gli utenti comunicano utilizzando i loro indirizzi IP (Internet Protocol) come riferimento. [1]

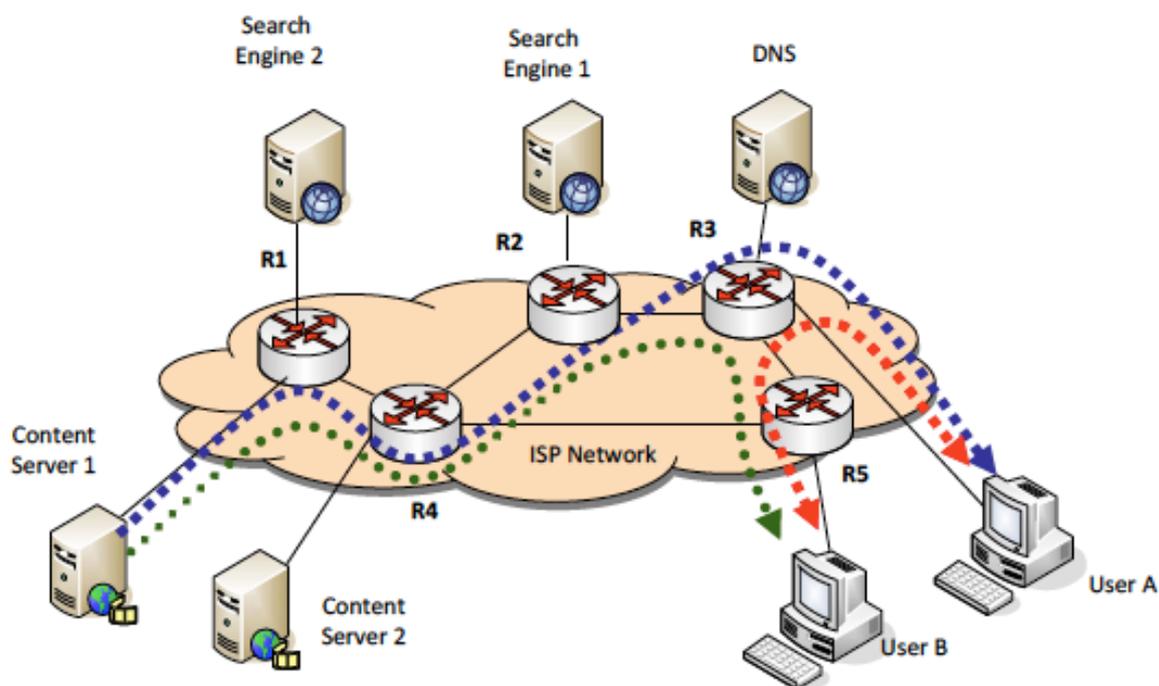


Figura 1: architettura di rete attuale [1]

Questo schema funziona a condizione che vi siano risorse sufficienti nel sistema: con il termine “risorse” possiamo identificare la capacità di banda in un determinato link o la capacità di instradare un pacchetto di dati con un ritardo sufficientemente basso.

Ma cosa accade quando:

- sono connessi miliardi di dispositivi in grado di essere interrogati e fornire informazioni che possono essere aggregate in servizi?
- gli utenti richiedono larghezze di banda maggiori di quelle che possono essere tipicamente supportate?
- sempre più utenti rendono critico il ritardo nelle comunicazioni video ed audio in tempo reale attraverso internet?

Tali cambiamenti saranno supportati dall'attuale internet solo attraverso grandi investimenti sapendo che l'architettura potrà presentare comunque caratteristiche di instabilità.

Un'intelligente evoluzione di internet porterà ad un utilizzo molto più efficiente delle risorse disponibili e fornirà un ambiente economico che incoraggerà gli investimenti.

Perché internet si deve evolvere?

Strade meno congestionate, forniture di energia elettrica migliori e più efficienti o trattamenti medici all'avanguardia gestiti da casa.

Sono questi alcuni esempi di cosa sarà possibile realizzare grazie all'internet del futuro che permetterà la connessione di oggetti fisici ed apparecchiature mobili per consentire di utilizzare in tempo reale un'enorme quantità di dati migliorandone funzionamento ed efficacia.

Poiché internet interconnette miliardi di terminali on-line per fornire informazioni sempre più sofisticate, le infrastrutture dovranno essere sufficientemente "intelligenti" per gestire in tempo reale un enorme flusso di dati.

L'obiettivo è di migliorare le infrastrutture fondamentali rendendole in grado di trattare enormi quantità di dati.

Da anni i fornitori di tecnologia cercano di combinare le reti dati e voce con l'obiettivo di fornire più servizi di comunicazione su un'unica rete ed a costi più bassi; l'infrastruttura di rete e del dominio di servizio si trova ad affrontare molti cambiamenti:

- la crescita esplosiva della velocità di trasmissione dati e delle capacità
- l'emergere di dati in massa e la distribuzione di contenuti e consumo
- l'evoluzione verso un'architettura di convergenza che si ha aumentando le combinazioni di servizi e gli utilizzi tra utenti, dispositivi, media e persino tra mondi reali e virtuali
- il coinvolgimento degli utenti che diventano anche produttori

Alcune modifiche allo schema tradizionale potrebbero portare ad un miglior utilizzo delle risorse disponibili e miglioramenti significativi possono essere pensati per gestire al meglio i mezzi di trasmissione del futuro.

Cosa accadrebbe se:

- il contenuto potesse essere memorizzato più vicino agli utenti finali in modo trasparente nella rete?
 - La distribuzione di contenuti sarebbe molto più efficiente!

- i router potessero identificare ed analizzare quali sono i contenuti che fluiscono attraverso di loro e fossero in grado di replicarli in modo efficace?
 - I motori di ricerca acquisirebbero la conoscenza della localizzazione dei contenuti!
- la rete potesse identificare dinamicamente qual è il percorso migliore il percorso migliore per l'utente?
 - Avrebbe fornito un modo migliore per consegnare i dati!
- il contenuto potesse essere adattato in modo interattivo e non soltanto in modo statico basato sulla rete e capacità dei terminali ma anche basato sulla selezione di contenuti interattivi da parte dell'utente?
 - L'esperienza utente sarebbe migliore!
- il contenuto potesse essere selezionato ed adattato al contesto?
 - L'utente avrebbe una vita molto più semplice utilizzando un'internet data-centrica!
- il contenuto potesse essere dinamico anziché statico?
 - Sarebbe molto più facile essere in grado di trovare i migliori contenuti!

Esistono due possibili approcci per sviluppare la nuova architettura, conosciuti come clean-state (PURSUIT e NetInf) ed overlay (DONA e NDN): con il primo metodo è necessario riprogettare l'intero stack protocollare mentre il secondo consente di creare un'infrastruttura avvalendosi delle funzionalità di una rete IP.

Questi approcci sono divenuti di recente centro di grande attenzione dato che sono in grado di risolvere molti dei problemi odierni che affliggono internet: incremento esponenziale del traffico, congestione di rete, non adeguato supporto alla mobilità, sicurezza e riduzione degli IP disponibili.

Nell'internet del futuro le reti datacentriche giocheranno un ruolo di primo piano poiché soddisfano al meglio le esigenze degli utenti, sempre più interessati al tipo di contenuto da acquisire attraverso la rete piuttosto che ai nodi remoti con cui interagire.

Capitolo 1

COGNITIVE NETWORK

1.1 Cos'è una Cognitive Network?

Una rete è cognitiva se attraverso un particolare processo è in grado di percepire le condizioni della rete, pianificare, decidere, agire su queste condizioni ed “imparare” dalle conseguenze delle proprie azioni.

Questo ciclo cognitivo percepisce l'ambiente di utilizzo, pianifica azioni in accordo con gli input e le policy della rete, decide quale scenario si adatta meglio allo scopo utilizzando un “ragionamento” ed infine agisce sullo scenario scelto.

Il sistema “impara” dal passato (situazioni, pianificazioni, decisioni, azioni) ed utilizza la propria “conoscenza” per migliorare le decisioni nel futuro. [2]

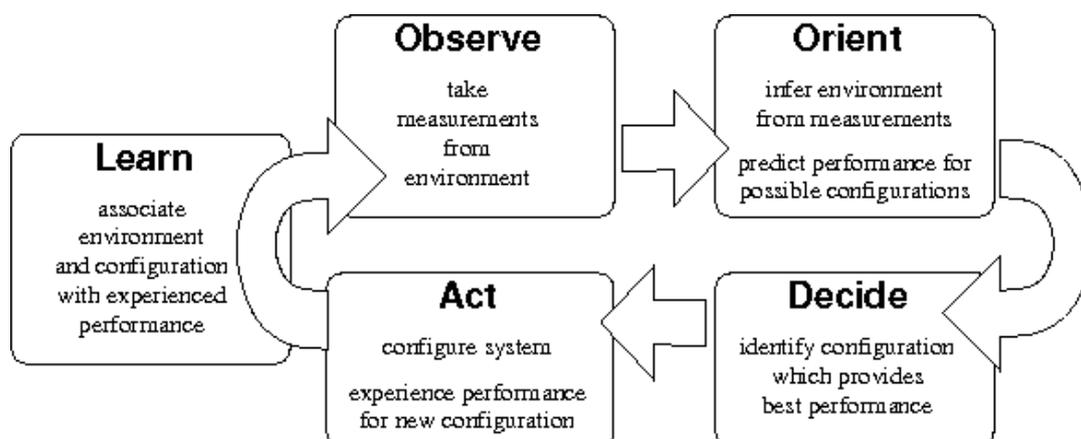


Figura 1.1: Cognitive Network [3]

L'architettura cross-layer di una cognitive network è anche chiamata EWI (Embedded Wireless Interconnection) e, a differenza del modello OSI, è basata su una nuova definizione di link.

I nuovi link astratti vengono ridefiniti come reciproci cooperatori in un insieme di nodi wireless vicini, a differenza della tradizionale rete wireless che si basa su link punto-punto virtualmente collegati tra loro.

1.2 Il Knowledge Plane

Bisogna creare una tecnica che sposti gli obiettivi dell'internet di oggi verso un nuovo orizzonte.

L'abilità della rete di sapere che cosa le si chiede di fare in modo che possa sempre "prenderci cura" di sé stessa senza dipendere più dagli utenti: i tradizionali approcci algoritmici non sono pronti a fornire gli strumenti richiesti!

Gli approcci dovranno offrire una capacità di astrazione in cui sia possibile isolare obiettivi di alto livello per azioni di basso livello, integrare ed agire su imperfezioni e contraddizioni ed imparare dalle azioni passate per migliorare le performance future: queste proprietà fanno parte di un sistema cognitivo chiamato knowledge plane.

Il knowledge plane è un sistema pervasivo all'interno della rete che costruisce e mantiene ad alto livello i modelli di ciò che la rete dovrebbe fare, al fine di fornire servizi e consulenze ad altri elementi della rete: il knowledge plane fa affidamento sugli strumenti di AI (Artificial Intelligence) e di sistemi cognitivi.

In un'applicazione di rete specifica, un approccio potrebbe essere quello di utilizzare ed incorporare tali knowledge specifici al centro del progetto della rete stessa, com'è già stato fatto nelle reti telefoniche.

Il knowledge plane include: [4]

- edge involvement
 - il principio end-to-end suggerisce che molte informazioni importanti sulle prestazioni della rete non nascono in rete, ma nei dispositivi e nelle applicazioni che le utilizzano
 - gran parte dei knowledge nel knowledge plane possono essere prodotti, gestiti e consumati al di là dei tradizionali edge della rete

- prospettive globali
 - ogni operatore gestisce la parte che possiede, ma è utile l'identificazione del problema che dipende dalla correlazione di osservazioni da parti differenti della rete
 - il knowledge plane dovrebbe essere in grado di estendere la propria prospettiva a tutta la rete globale
- strutture composizionali
 - se la portata dei knowledge plane è globale, allo stesso tempo essa deve essere progettata per tener conto della composizione in quanto i knowledge plane di due reti non collegate dovrebbero essere in grado di fondere le loro attività se le reti diventano connesse
 - una conseguenza del problema di composizione è il bisogno di operare in presenza di informazioni imperfette e contraddittorie
- approccio unificato
 - si potrebbe ipotizzare che i vari problemi possano essere facilmente risolti da meccanismi distinti, ma il knowledge plane è un singolo sistema unificato con standard e framework comuni per knowledge
 - necessario perché il mondo reale non è partizionato per task ma strutturato sulla base dei knowledge
- cognitive framework
 - il knowledge plane necessita di:
 - * formulare giudizi in presenza di informazioni parziali o in conflitto
 - * riconoscere ed inoltrare i conflitti nelle policy
 - * rispondere a problemi ed attacchi su frame in modo più efficiente
 - * eseguire le ottimizzazioni in ambienti ad alti livelli dimensionali troppo complicati per essere affrontati dall'uomo o da soluzioni analitiche
 - * automatizzare le funzioni che oggi devono essere effettuate da tecnici qualificati

La maggior parte delle discussioni sulle architetture di rete riconoscono due divisioni architettoniche: un data plane su cui il contenuto è trasmesso, ed un management plane (o control plane) che è usato direttamente e proporzionalmente sul data plane.

Quest'ultimo utilizza la stratificazione per nascondere la complessità e consente l'estensibilità, l'interoperabilità e la scalabilità.

I sistemi di controllo e di gestione sono progettati per tagliare attraverso la stratificazione, dare visibilità ed accesso a tutti gli aspetti della rete che deve essere monitorata e gestita, il management plane invece tende ad una scarsa scalabilità ed è difficile da cambiare.

Il knowledge plane chiaramente risiede in un posto diverso e, a differenza del management plane, tende a fornire una visione unificata piuttosto che partizionare: per questo è necessario che le informazioni di rete e quelle riguardanti i problemi osservati dall'utente possano essere nello stesso framework.

Gli obiettivi per il knowledge plane necessitano di soddisfare una serie di richieste: [4]

- devono funzionare in modo utile in presenza di informazioni incomplete, inconsistenti ed anche con quelle dannose o ingannevoli
- devono lavorare in modo adeguato in presenza di conflitti o di un obiettivo di alto livello inconsistente tra le diverse parti interessate di internet
- devono operare sempre in modo efficiente in presenza di nuove tecnologie ed applicazioni non concepite al momento della progettazione e, a fronte di un ambiente altamente dinamico, includere sia i cambiamenti a breve che a lungo termine nella struttura e complessità della rete sottostante

Queste richieste non possono essere soddisfatte dalle soluzioni analitiche in quanto richiedono di solito informazioni complete, formulazioni di problemi ben precisi ed un ambiente operativo realmente statico.

Quindi saranno necessarie tecniche cognitive: il vantaggio principale di queste tecniche è il loro potenziale di assolvere efficacemente, valutare e migliorare le proprie prestazioni in presenza di ambienti complessi inconsistenti, dinamici ed in continua evoluzione.

Nel contesto knowledge plane l'architettura cognitiva sarebbe distribuita e decentrata, invece il partizionamento sarebbe stato effettuato per sostenere interessi divergenti dei soggetti interessati alla rete.

Il knowledge plane deve essere in grado di argomentare e memorizzare in quanto l'apprendimento della conoscenza può avvenire attraverso molti mezzi.

Può avvenire ad esempio per tentativi, per errori, per istruzione, per generalizzazione, per analogia o attraverso la risoluzione di problemi.

In un ambiente con problemi statici è abbastanza semplice ammettere che la soluzione analitica di apprendimento sia irrilevante, ma gli IP di rete di progettazione sono in costante evoluzione in molte dimensioni ed in certe configurazioni sono infiniti!

Una funzione del knowledge plane sarà di sostenere la mediazione tra gli utenti e gli operatori i cui obiettivi possono essere in conflitto: l'inevitabilità di tali conflitti suggeriscono che si debbano sviluppare nuove tecniche di rappresentazione e "ragionamento" su vincoli e policy.

Bisogna unificare gli obiettivi per il knowledge plane: costruire una nuova generazione di rete consentendo di avere una visione di ciò che dovrebbe essere e che cosa si è supposto di fare: si può pensare ad un knowledge plane che possa risolvere un problema e che ne spieghi il motivo.

Il knowledge plane può comunicare con molte entità di un problema: la domanda di una correzione utente potrebbe innescare un messaggio di ritorno per l'utente stesso quando il programma è bloccato, ma potrebbe anche innescare un messaggio per l'ISP nel quale si comunica l'insoddisfazione dell'utente.

Una futura estensione è che il knowledge plane possa fornire un assistente per l'utente, cioè un agente che controlli cosa fa l'utente ed "impari" con il tempo ciò che è normale.

Altri programmi possono rilevare e segnalare che qualcosa non va, il knowledge plane dovrà decidere quanto dare credito a questi segnali in base a dove provengono.

Le diagnostiche possibili sono di controllare le funzioni a tutti i livelli di inoltro dei pacchetti alla funzione di applicazione: una volta che un nodo terminale ha eseguito ciò che la diagnosi può eseguire, la fase successiva è di aggiungere asserzioni di condivisione sul knowledge plane su cosa è stato scoperto.

Questo contributo al knowledge plane permette a tutti gli utenti della rete di costruire una visione globale della rete stessa e dello stato di servizio: i dati possono essere combinati con informazioni derivanti da attività di misurazione attualmente in corso su internet che tentano di costruire un modello di stato della rete.

Il configuration manager del knowledge plane deve avere sufficiente conoscenza della struttura di basso livello per capire se la rete è configurata correttamente, per rilevare sia se è disponibile una migliore configurazione alternativa sia se il sistema sembra essere danneggiato.

Il manager deve essere in grado di ricavare i setting di basso livello impostando obiettivi, priorità vincoli di alto livello e deve saper leggere l'impostazione attuale a basso livello e descriverne il conseguente comportamento in termini di alto livello.

Un problema interessante è quando il sistema rileva affermazioni contrastanti da due parti differenti: la configurazione delle operazioni non dovrebbe succedere una sola volta sulla rete ma dovrebbe accadere costantemente verificando il cambiamento delle condizioni della rete, le esigenze delle applicazioni e la modifica dei vincoli.

Questo compito può anche essere eseguito in modo ricorsivo: ogni regione configurerà in primo luogo sé stessa con gli obiettivi indicati a livello locale e dai vincoli, ma quando due regioni si connettono ci possono essere ulteriori vincoli che si impongono agli altri.

Se il knowledge plane dispone di informazioni sufficienti per configurare la rete, allora l'informazione stessa può anche essere utile alle altre applicazioni che si stanno configurando.

Ogni overlay di rete usa meccanismi basati su edge per valutare i risultati dei diversi percorsi possibili attraverso internet e cerca di costruire un insieme di percorsi di trasporto.

I percorsi instradano pacchetti applicativi attraverso quella che sembra essere la parte di internet più adatta alle esigenze applicative: le applicazioni di rete devono esplorare internet perché non esiste un meccanismo per conoscere le capacità del centro della rete.

Il knowledge plane può riuscire ad aggregare applicazioni, offrendo loro una migliore informazione sulla rete, ed accedere ai punti di controllo in cui il comportamento potrebbe essere modificato per aiutare a soddisfare le loro esigenze.

Al massimo livello, l'architettura del knowledge plane sarà definita da due caratteristiche fondamentali proprie: la struttura compositiva distribuita e la prospettiva di conoscenza potenzialmente globale.

Il knowledge plane deve rispettare e sfruttare il fatto che le reti hanno una struttura interna e sono dinamiche.

Grandi reti sono composte da interconnessioni più piccole, i partecipanti vengono e vanno, il rapporto tra proprietari, operatori ed utenti di reti differenti possono variare anche quando la struttura fisica non agisce.

Ciò implica che il servizio di una rete knowledge plane è una struttura automaticamente creata ricorsivamente, dinamicamente e continuamente composta e decomposta da piccole sottoparti.

Questi requisiti implicano che il knowledge plane: [4]

- sia distribuito
 - funzionalità knowledge plane per differenti regioni della rete sono fisicamente e logicamente decentralizzate
- sia bottom-up
 - entità semplici si possono comporre in grandi entità più complesse secondo la necessità e si decompongono dal sistema a seconda dei casi
 - un processo ricorsivo che può procedere su più livelli
- sia constraint driven
 - il principio fondamentale è che il sistema possa adottare un comportamento non specificamente limitato
- si sposti dal semplice al complesso
 - l'atto di comporre un insieme di reti per formare un più grande impiego con più requisiti e vincoli sul comportamento di ciascuna rete

Il primo obiettivo per l'architettura di sistema knowledge plane è supportare la prospettiva distribuita e compositiva con la necessità di permettere astrazione e capacità.

Il knowledge plane può essere in grado di raccogliere, filtrare, ridurre, osservare, concludere ed affermare i percorsi per le differenti parti della rete in punti utili, può inoltre accedere e gestire ciò che cognitive community chiamano sensori (entità che producono osservazioni) ed attuatori (entità che cambiano il comportamento).

L'aspetto centrale del knowledge plane è che supporta il calcolo cognitivo: questo è un problema impegnativo perché l'ambiente del knowledge plane dinamico e distribuito non è ben compensato né a livello dell'algoritmo del knowledge classico, né nell'architettura agente su cui si forma gran parte della corrente AI.

Gli algoritmi AI non sono progettati per lavorare in un contesto altamente distribuito, ciò che serve sono algoritmi robusti, trattabili ed on-line per ambienti dinamici, in parte osservabili, stocastici e soggetti ad errori.

Uno degli obiettivi del knowledge plane è fornire alle applicazioni la capacità di "imparare" e "ragionare" sul proprio ambiente: molti di questi problemi risiedono ai confini dell'AI.

Il knowledge plane è costruito a blocchi per avere una rete più affidabile e più robusta: una corretta implementazione dovrebbe continuare a migliorare la rete e quando la rete diventa più grande anche il knowledge plane stesso deve diventare più grande.

Nodi guasti possono iniettare malfornate osservazioni: alcuni nodi possono mentire sul loro comportamento ed alcuni esecutori possono tentare di interrompere o confondere il knowledge plane, o per attaccare la rete nel suo insieme, o per acquisire un certo vantaggio rispetto ad altri.

Il sistema knowledge plane dovrà dipendere su strategie quali la certificazione di consenso e controllo incrociato per rilevare comportamenti malfornati e dannosi: costruire un modello di fiducia richiede un'identità robusta e persistente.

Capitolo 2

QUALE FUTURO?

2.1 Cosa porterà l'internet del futuro all'uomo?

Se vogliamo rispondere alla domanda, dobbiamo considerare l'internet del futuro sia dal punto di vista socio-economico che dal punto di vista tecnologico.

Ci aspettiamo che l'internet del futuro ci porti nuove modalità di consumo e di creazione dei contenuti multimediali, al fine di coprire le diverse esigenze dell'uomo e preservare la generazione di reddito delle varie parti interessate.

Nel corso degli ultimi anni abbiamo visto un grande aumento di popolarità dei contenuti professionali forniti mediante la rete internet, comunicazioni multimediali a banda larga e contenuti generati dagli utenti possono aumentare rapidamente il traffico di rete.

I fornitori di servizi trarranno inoltre vantaggio dalle nuove funzionalità ed opportunità offerte dall'internet del futuro al fine di fornire nuovi servizi: la ricerca ed il recupero sono i primi passi per utilizzare il contenuto in quanto l'informazione è il risultato dell'estrazione, della combinazione e dell'aggregazione dei contenuti.

I servizi sono il risultato di manipolazione e di utilizzo dei contenuti o delle informazioni, la sicurezza e la privacy sono applicate sul contenuto: oggi utilizziamo internet o per il suo contenuto o per le informazioni che esso genera.

Quindi per dare una prima risposta a questa domanda, possiamo affermare che l'internet del futuro porterà nuove esperienze e nuove funzionalità di comunicazione.

L'internet del futuro potrebbe ospitare servizi, metodi, procedure e tecniche che permetteranno la creazione ed il consumo di nuove esperienze sia professionali che generate dagli utenti.

Questa risposta comprende nuovi servizi, nuovi mezzi di comunicazione, nuove applicazioni e nuove capacità di streaming: nell'internet del futuro, al fine di realizzare nuove esperienze, avremo bisogno di nuovi media e contenuti che possano catturare più sensi (vista, udito, tatto, ed anche l'olfatto).

Ci si aspetta che le nuove tecnologie possano rendere possibili contenuti emergenti: [1]

- tecniche di analisi automatica audio-visiva possono segmentare i contenuti
 - la tecnologia è attualmente utilizzata per ottimizzare i mezzi di codifica o per consentire la ricerca di multimedia, necessari nel breve e nel medio termine per decomporre i componenti che costituiscono il contenuto audio-visivo
- nel lungo periodo, dipenderemo dalla tecnologia di cattura component-preserving
 - sistemi avanzati di cattura con telecamere stereoscopiche, fotocamere array e time-of-flight o scanner in 3D preservano tre informazioni dimensionali, le loro componenti e relazioni spazio-temporali
- la modellazione 3D è usata per creare contenuti audio-visivi
 - tutte le informazioni sui componenti che costituiscono il processo di composizione saranno prontamente disponibili

La disponibilità delle componenti che costituiscono le relazioni spaziali e temporali aprono nuove opportunità per la creazione ed il consumo dei contenuti, inoltre: [1]

- il riutilizzo dei componenti da contenuti esistenti per la creazione di nuovi contenuti audio-visivi diventa molto meno ingombrante, permettendo un facile e rapido mash-up del media
- la collaborazione on-line per la compilazione dei contenuti audio-visivi prepara la strada per un vero Wikipedia multimediale
- la personalizzazione si evolve da una selezione di contenuti preparati ad una composizione just-in-time
- l'inserimento di contenuti audio-visivi in tempo reale sono facilitati

- la combinazione di cattura dei contenuti audio-visivi con contenuto sintetico 3D crea delle emozionanti esperienze di realtà mista

L'internet datacentrico del futuro offrirà all'utente finale una maggiore immersione in esperienze real-time che potrà indirizzare sia gli scenari di consumo che gli scenari di collaborazione e comunicazione di contenuti.

2.2 Perché l'internet del futuro sarà datacentrico?

Per rispondere alla seconda domanda invece dobbiamo sapere che l'internet del futuro non porterà solamente nuovi contenuti multimediali.

L'attenzione dovrebbe essere dedicata anche alla presentazione, alla comunicazione, ai servizi, alle infrastrutture ed alla privacy che sono componenti essenziali dell'internet del futuro.

Eppure il contenuto è ciò che gli utenti riceveranno e gradiranno: partendo dal presupposto che i contenuti e le rappresentazioni dei contenuti sono la base dell'internet del futuro e, seguendo un approccio bottom-up, possiamo supporre che se l'internet del futuro fosse datacentrico potrebbe maneggiarne facilmente i contenuti.

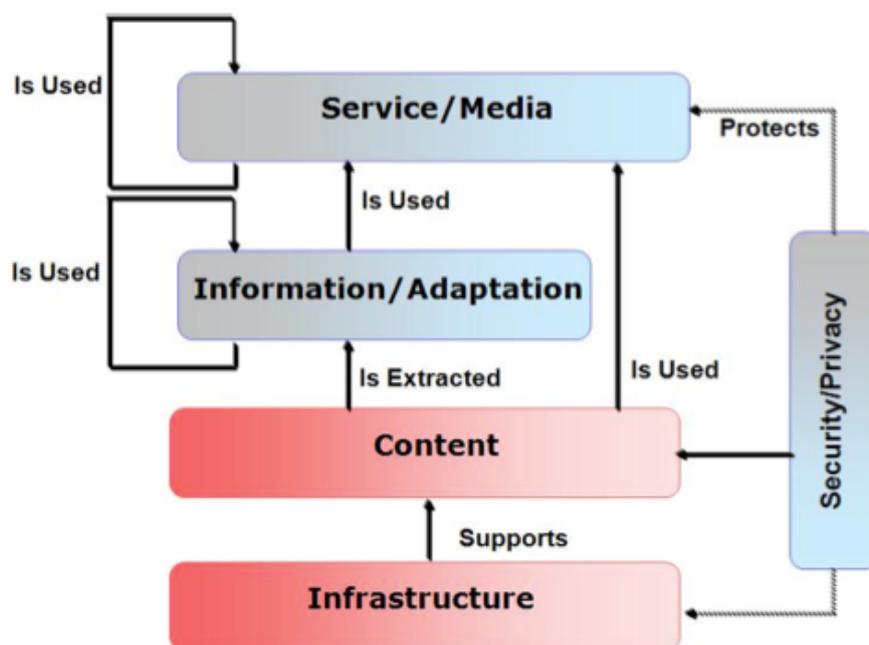


Figura 2.1: architettura datacentrica [1]

Definiamo in dettaglio la figura: [1]

- content
 - è un qualsiasi tipo e volume di media, può essere pre-registrato, cache o dal vivo, statico o dinamico, monolitico o modulare
 - può essere combinato, mixato o aggregato per generare nuovi contenuti e media
- information
 - è il prodotto di un certo numero di funzioni applicate al content o all'information in modo ricorsivo
 - combinando, estraendo, aggregando content e pezzi di information può essere estratta o generata la nuova informazione
- service
 - è il risultato di un insieme di funzioni applicate su content, su pezzi di information o ricorsivamente su service
 - manipolando, gestendo, combinando, personalizzando, adattando content, information o service possono essere composti o generati i nuovi servizi
- infrastructure
 - sarà costituito da funzioni di trasporto, di immagazzinamento e di processamento in modo distribuito
 - offre l'opportunità di "trattare" con l'attivazione di oggetti content, piuttosto che con sequenze di bit non strutturati nella rete
 - un'infrastruttura datacentrica intelligente può supportare in modo efficiente i contenuti ed i suoi derivati (informazioni e servizi)
- security and privacy
 - sarà una proprietà di content, information, service e infrastructure che consentirà un controllo molto più efficiente degli oggetti content

Anche se queste considerazioni possono sembrare semplificate, dobbiamo ammettere che la migliore manipolazione del contenuto che ci possa essere nell'internet del futuro è proprio datacentrica.

Questo consentirà alla prossima generazione di personalizzare i servizi, di salvare le registrazioni, contenuti live ed informazioni.

Ma sarà anche aperto un nuovo mondo di comunicazioni audio-visive di alta qualità, che richiedono contenuti da acquisire in streaming presentati in modo coerente con un ritardo minimo.

2.3 Principi base dell'internet del futuro

L'internet del futuro potrebbe essere realizzato dal punto di vista tecnologico in un gran numero di modalità alternative.

Ma sulla visione datacentrica emergono quattro principi fondamentali: [1]

- supportare modelli di business multipli e nuovi
 - l'internet del futuro dovrebbe essere in grado di supportare modelli di business flessibili
 - più parti interessate dovrebbero essere in grado di partecipare ad un ambiente aperto che sostenga ed incoraggi l'innovazione e la partecipazione di tutti senza barriere
 - tale principio include la possibilità di costruire un internet del futuro che possa ospitare in sé diverse architetture, una di quelle è datacentrica
- semplicità
 - nell'internet di oggi la complessità appartiene agli edge ed il layer IP di internet rimane il più semplice possibile
 - sistemi complessi sono generalmente meno affidabili e flessibili
 - problemi complessi a volte richiedono soluzioni complesse
 - l'internet del futuro fornirà funzionalità non banali in molti aspetti

- sostenibilità, scalabilità e robustezza
 - l’internet del futuro dovrebbe:
 - * essere concepito come una rete sostenibile offrendo un supporto incorporato per soluzioni efficienti di energia
 - * abbastanza flessibile da evolvere continuamente e da svilupparsi ed estendersi in risposta ai cambi di esigenze
 - * essere in grado di servire un numero molto elevato di entità (scalabilità)
 - * mantenere il proprio rapporto di funzionamento (disponibilità)
 - * recuperare facilmente nel caso di guasti (affidabilità)
 - * contenere e probabilmente espandere i principali sistemi di gestione come gli errori, le configurazioni, gli account, le prestazioni e la sicurezza
- loose coupling
 - quando le cose si fanno grandi molto spesso presentano un aumento dell’interdipendenza tra componenti
 - l’internet del futuro dovrebbe essere costruito tenendo conto di un “accoppiamento libero” dato che i sistemi liberamente accoppiati hanno una maggiore flessibilità di vincoli temporali, ambientali e di ordinamento rispetto a quelli accoppiati rigidamente

2.4 Requisiti per la progettazione datacentrica

Questi requisiti sono tratti da un centro di supporto analisi e di conseguenza spostano l’attenzione dei contenuti sulla rintracciabilità, sulla fiducia e sulla flessibilità: [1]

- ingegneria datacentrica
 - assumendo che l’architettura dell’internet del futuro sia datacentrica, allora l’efficiente contenuto ingegneristico dovrebbe essere al centro dei requisiti di progettazione
 - la progettazione dovrebbe comprendere nuove modalità di memorizzazione cifrando ed arricchendo la ricerca ed il rendimento dei contenuti, in cui il contenuto viene trattato in modo semantico piuttosto che essere solo un insieme di pixel incapsulati in pacchetti per la trasmissione in rete

- risoluzione dei nomi e rintracciabilità
 - gli utenti sono in generale interessati ad accedere ad una particolare porzione di contenuto piuttosto che sapere dove e quando il contenuto sia stato memorizzato, conoscerne il nome, URL o l'indirizzo IP del server ospitante
 - * ci si dovrebbe concentrare quindi sul contenuto e sulle informazioni che possono essere estratte dalla lavorazione dei contenuti pre-processati
 - la rintracciabilità invece contiene due operazioni primitive (individuare e registrare)
 - * gli utenti dovrebbero essere in grado di richiedere i dati di nome della rete utilizzando il ritrovamento primitivo
 - * i fornitori di contenuti dovrebbero essere in grado di pubblicare oggetti dati che saranno serviti agli utenti utilizzando il registro primitivo
- routing datacentrico
 - l'internet datacentrico del futuro dovrebbe essere in grado di sostenere il routing datacentrico, il quale permette agli utenti di scaricare particolari pezzi di informazioni dalla posizione più comoda o da posizioni basate sui requisiti dell'utente
 - un simile approccio porterebbe ad una diffusione di informazioni multiparty-to-multiparty piuttosto che a tradizionali conversazioni point-to-point e gli utenti dovrebbero essere in grado di trattare le informazioni senza occuparsi della posizione di tale informazione
- modelli flessibili di business dei contenuti
 - l'internet datacentrico del futuro dovrebbe essere in grado di sostenere un modello di business flessibile in cui più parti interessate possono partecipare in un ambiente aperto
 - sostiene ed incoraggia innovazioni e partecipazioni da parte di utenti, ISP, provider di contenuti e servizi, organizzazioni pubbliche e private e regolatori
 - questi requisiti di progettazione sono i possibili principali punti di differenziazione tra le NGN (Next Generation Networking) e l'internet datacentrico del futuro che dovrà essere progettato in "bagarre", cioè con lo scontro di interessi tra le parti interessate di internet

- affidabilità dei contenuti e dei servizi media
 - una delle sfide dell'internet datacentrico del futuro è:
 - * fornire servizi e loro contenuto in modo sicuro tra tutti gli utenti
 - * garantire la privacy di ciascun partecipante
 - * proteggere le reti contro i guasti e gli attacchi dannosi
 - ulteriori criteri di fiducia degli utenti devono essere soddisfatti:
 - * avere a disposizione servizi ed applicazioni verificati
 - * pieno controllo e consapevolezza di quali dati possono essere a disposizione
 - * metodi di gestione per l'accesso a contenuti non adatti a tutti
- scelta della fonte e della presentazione
 - l'internet datacentrico del futuro dovrebbe fornire la scelta di come e da chi un utente scarica un particolare pezzo di contenuti o informazioni
 - l'utente dovrebbe poter scegliere da dove prelevare contenuti o informazioni se memorizzati presso più provider
 - i contenuti dovrebbero essere disponibili in vari formati di presentazione ed il formato appropriato dovrebbe essere adattato alle preferenze ed alla situazione dell'utente
 - l'internet datacentrico del futuro dovrebbe inoltre sostenere l'allocazione automatica delle risorse appropriate per garantire la consegna e la presentazione di un particolare contenuto
- auto-organizzazioni decentralizzate
 - l'architettura di rete datacentrica proposta non dovrebbe richiedere una gestione o una configurazione manuale a larga scala, o coordinamento a livello globale, e l'obiettivo principale dovrebbe essere quello di prevedere una auto-organizzazione decentralizzata e sicura dei contenuti

2.5 Architettura datacentrica logica

Nell'internet del futuro l'intelligenza verrà spostata nella rete e trasformata in un'architettura datacentrica.

L'architettura può essere costituita da diverse gerarchie di nodi virtuali con diverse funzionalità, può essere facilmente scalata a più livelli di gerarchia, potrà essere basata su livelli di informazioni e sui requisiti di servizio di consegna e vincoli.

Al livello più basso c'è l'infrastruttura del provider di rete e servizi: i legami tra i nodi di rete attuali saranno la maggioranza dei nodi di rete per un certo numero di anni.

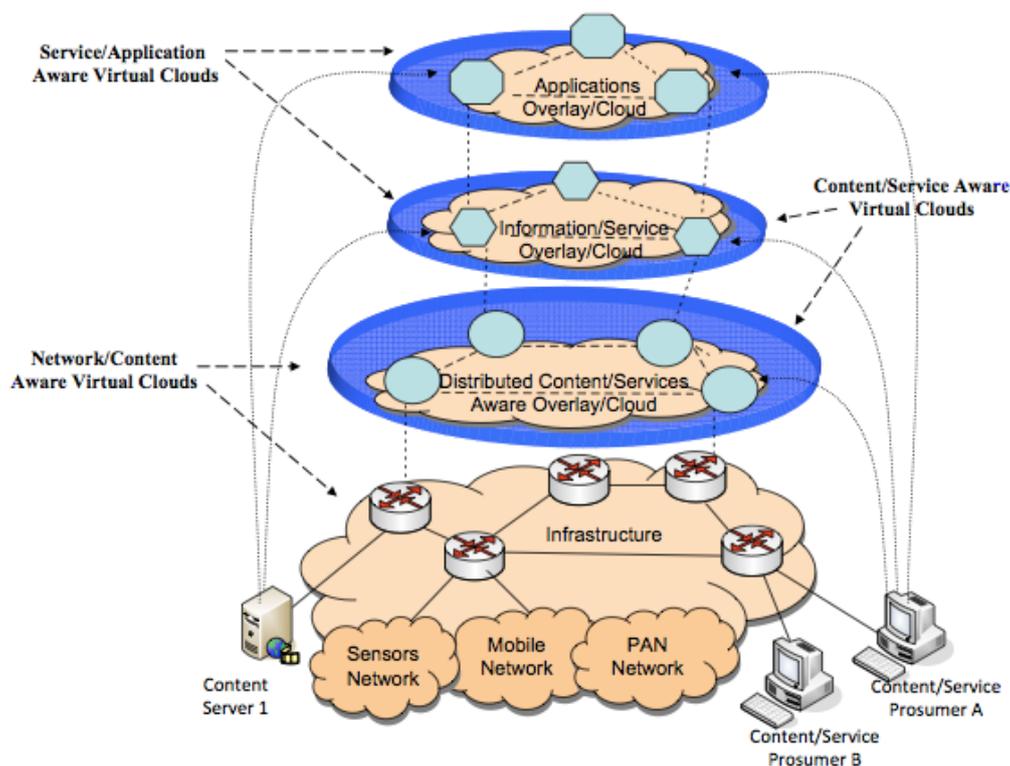


Figura 2.2: architettura datacentrica logica [1]

Questa infrastruttura di rete è il servizio che sarà composto da nodi con funzionalità ed intelligenza limitate in cui i contenuti saranno instradati assumendo i requisiti di qualità di base.

L'infrastruttura deve nascondere tutte le inutili complessità e fornire tutte le informazioni necessarie in modo che i nodi più "intelligenti" possano prendere tutte le decisioni necessarie per sostenere la funzionalità richiesta.

Così l'infrastruttura di rete: [1]

- avrà a che fare con:
 - oggetti dal contenuto attivo
 - sequenze di bit non strutturate

- fornirà:
 - il trasporto richiesto
 - il controllo di congestione
 - le policy
 - i protocolli di segnalazione e di elaborazione in modo distribuito
 - tutte le funzionalità di guasti, configurazioni, registrazioni, sicurezza e performance

- sosterrà:
 - la mobilità e la portabilità dei contenuti, delle informazioni, dei servizi, dei terminali e degli utenti
 - la virtualizzazione
 - l'auto-organizzazione e l'auto-gestione
 - la robustezza, la stabilità e la sopravvivenza

- offrirà:
 - di nascondere la complessità non necessaria
 - monitoraggio e controllo delle risorse di rete
 - hooks per sostenere il QoS (Quality of Service) e gli SLA (Service Level Agreements)
 - connettività a più amministratori
 - un dominio cross-layer
 - sicurezza e privacy dell'infrastruttura
 - capacità di interoperabilità e di interazione

Il livello di mezzo è l'overlay di conoscenza dei servizi e contenuti distribuiti: nodi di rete a conoscenza dei contenuti saranno situati in questo overlay che avranno l'intelligenza di filtrare i contenuti ed i servizi web che scorrono attraverso di essi o identificare le sessioni di streaming e traffico.

In alternativa il contenuto può essere considerato per formulare gli oggetti informazione come primo ordine di elementi nella rete: in questo modo è direttamente identificabile con i nodi della rete.

In entrambi i casi i nodi di questo gruppo riconosceranno e qualificheranno i contenuti: parte di queste informazioni possono essere memorizzate localmente e riferite al livello più alto della gerarchia.

L'overlay di conoscenza dei servizi e contenuti può essere frazionato in più cloud che può essere dinamicamente costruito a strati tra il contenuto e le informazioni di overlay.

Si possono considerare overlay per nascondere e classificare i contenuti, per monitorare la rete, per adattare i contenuti: rispetto alla fornitura dei contenuti, i nodi a questo layer possono operare come un client-server ibrido, peer-to peer o reti di cloud, secondo i requisiti di consegna.

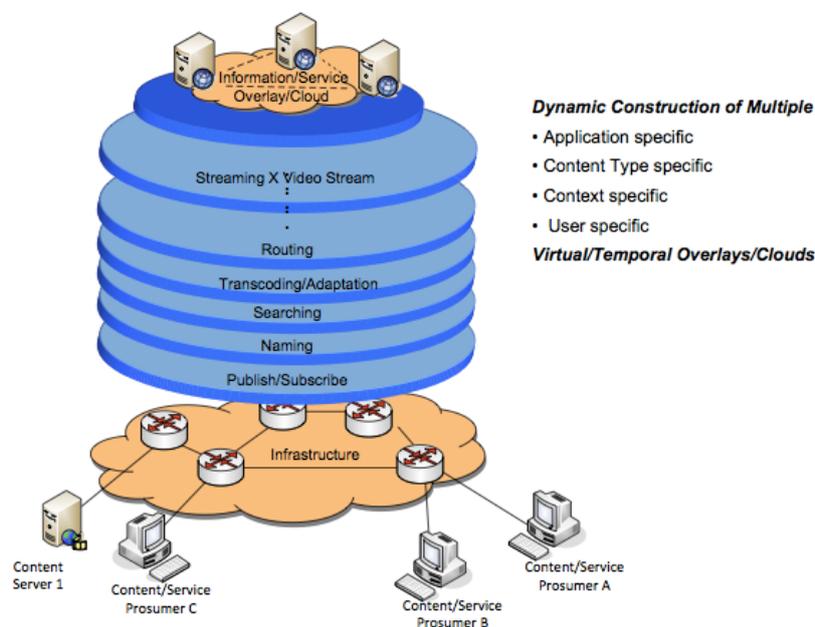


Figura 2.3: contrazione dinamica degli overlay [1]

In sintesi, l'overlay di conoscenza dei servizi e contenuti:

- fornirà:
 - la consapevolezza (ispezione, accettazione, categorizzazione, indicizzazione) dei contenuti
 - il caching dei contenuti e la rintracciabilità
 - l'adattamento e la personalizzazione dei contenuti

- offrirà meccanismi per:
 - l'estrazione, la combinazione e la creazione dei contenuti
 - la risoluzione dei nomi, percorso per nome e per tipo di capacità
 - la responsabilità e la fatturazione dei contenuti
 - la protezione dei contenuti

Ad un livello superiore c'è l'overlay di informazioni che si compone di nodi intelligenti o server che hanno una conoscenza distribuita sia dei contenuti e locazione dei servizi web e servizi nascosti, sia delle condizioni ed istanze di rete.

Basato sulla distribuzione ed istanziazione della rete attuale, su uno scenario o esigenze di servizio e sugli accordi di qualità dei servizi, i nodi possono variare da peer inaffidabili in una topologia next-P2P per garantire router aziendali o anche centri di dati nelle reti di cloud carrier-grade.

Il contenuto può essere memorizzato presso gli overlay information o ad un layer di gerarchia inferiore, anche se gli overlay information saranno sempre a conoscenza della locazione o caching dei contenuti e delle informazioni di rete.

Sulla base di queste informazioni si può decidere che il contenuto sia recuperato e consegnato in maniera ottimale per gli abbonati o utenti consumatori o servizi.

L'overlay information offrirà meccanismi per: [1]

- il controllo:
 - del contenuto
 - del servizio di rintracciabilità
 - della costruzione
 - della rappresentazione
 - della distribuzione
- la supervisione:
 - della locazione
 - caching e distribuzione dei contenuti
 - per le informazioni delle risorse di rete

- i contenuti:
 - pubblicazione
 - abbonamenti
 - responsabilità
 - fatturazione
- supporto:
 - alla rete tussle
 - ai nuovi modelli di business

Infine, al più alto livello è localizzato il layer applicazione: le applicazioni saranno usate efficientemente dai servizi, dalle informazioni e dai fornitori di media e contenuti per l'architettura datacentrica ed offrirà nuove esperienze media per gli utenti.

2.6 Architettura datacentrica autonoma

Dopo aver introdotto i componenti content, che possono essere generati e ricostruiti al volo e semanticamente arricchiti da mondi virtuali 3D al fine di creare una coinvolgente esperienza multimediale, ne espandiamo ulteriormente il concetto.

Un oggetto content (o semplicemente oggetto) è un contenitore che può essere costituito da mezzi di comunicazione, da norme, da comportamenti, da relazioni e da caratteristiche o da qualsivoglia combinazione di queste: [1]

- un media può essere una qualsiasi cosa che un essere umano può percepire e praticare con i suoi sensi
- le norme possono fare riferimento al modo in cui un oggetto viene trattato o manipolato da altri oggetti o dall'ambiente
- l'oggetto conosce il proprio scopo nell'esperienza multimediale integrata e quindi la sua priorità per il trasferimento
 - potrebbero essere incluse anche le opzioni per la manipolazione da parte dell'utente finale al momento della presentazione

- il comportamento può fare riferimento al modo in cui l'oggetto incide su altri oggetti o ambienti
- le relazioni tra due oggetti possono fare riferimento al tempo, allo spazio o problemi di sincronizzazione
- le caratteristiche descrivono il significato dell'oggetto e permettono di recuperare oggetti correlati

Gli oggetti possono essere organizzati gerarchicamente, come gli strumenti che costituiscono una band, e possono innescare la generazione di nuovi oggetti.

Un oggetto può essere diviso in oggetti nuovi o più oggetti possono essere combinati e concentrati per creare nuovi oggetti e gestire i servizi: tutte queste operazioni possono accadere durante il viaggio attraverso la rete.

Un oggetto può essere clonato, il clone mantiene le caratteristiche dell'oggetto principale ma sa che è un clone: questo può essere associato a problemi come l'incasso e la gestione dei DRM.

Gli oggetti autonomi viaggeranno attraverso la rete, si divideranno e si combineranno per creare nuovi servizi ed oggetti del mondo virtuale: attualmente è molto difficile immaginare come sarà un'architettura di rete che supporti oggetti.

Ecco un tentativo di mappare le caratteristiche di un approccio stratificato nel nuovo concetto layer-less di un oggetto, in cui uno o più strati vengono mappati da una o più caratteristiche dell'oggetto.

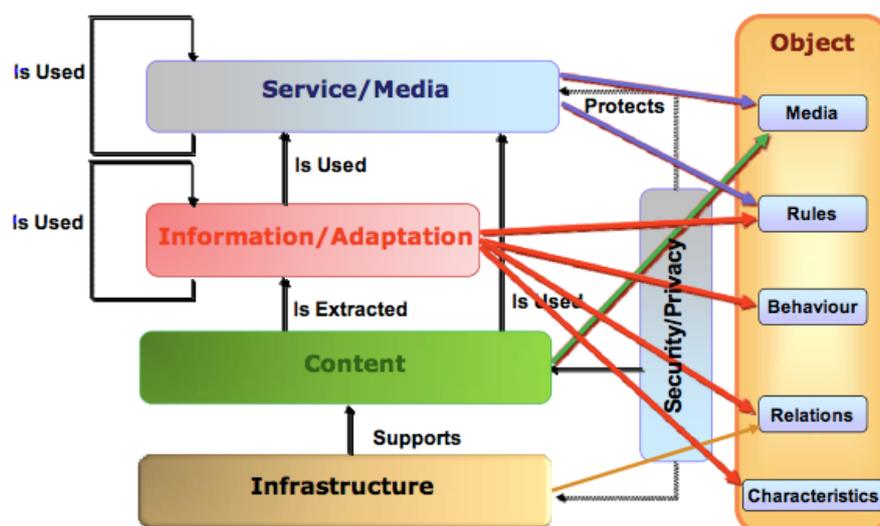


Figura 2.4: mappatura dell'architettura datacentrica [1]

Più nello specifico, il trasferimento è l'integrazione dell'oggetto ai fini della creazione di un'orchestrata esperienza che chiaramente richiede un'intelligenza tale da combinare applicazioni e contenuti informativi.

L'intelligenza potrebbe essere inserita nell'oggetto stesso, recuperare informazioni nella rete fornendo istruzioni per il routing e per la trasformazione o l'intelligenza potrebbe essere ospitata nei nodi della rete che cercano di soddisfare le richieste degli oggetti come sono descritti nelle norme, comportamenti e relazioni.

Infine le caratteristiche descrivono in modo significativo l'oggetto da prendere, soprattutto l'input del layer "information / adaptation".

Va notato che sia l'architettura datacentrica logica che autonoma forniscono una prospettiva olistica della visione di un internet datacentrico del futuro.

Oltre a servire le applicazioni esigenti del futuro in tempo reale, offrono anche la possibilità di ridurre in modo significativo le barriere e le limitazioni degli approcci esistenti nella gestione della rete.

Capitolo 3

DIFFERENTI APPROCCI

3.1 NDN ed il suo protocollo CCN

NDN (Named Data Networking) è un'architettura completamente nuova in cui i principi di design sono derivati dai successi dell'internet di oggi che riflettono la conoscenza dei punti di forza e di debolezza dell'attuale architettura.

Vengono applicati sei principi guida per disegnare l'architettura NDN: i primi tre derivano dai successi dell'attuale internet, mentre gli altri tre da dall'esperienza fatta negli anni.

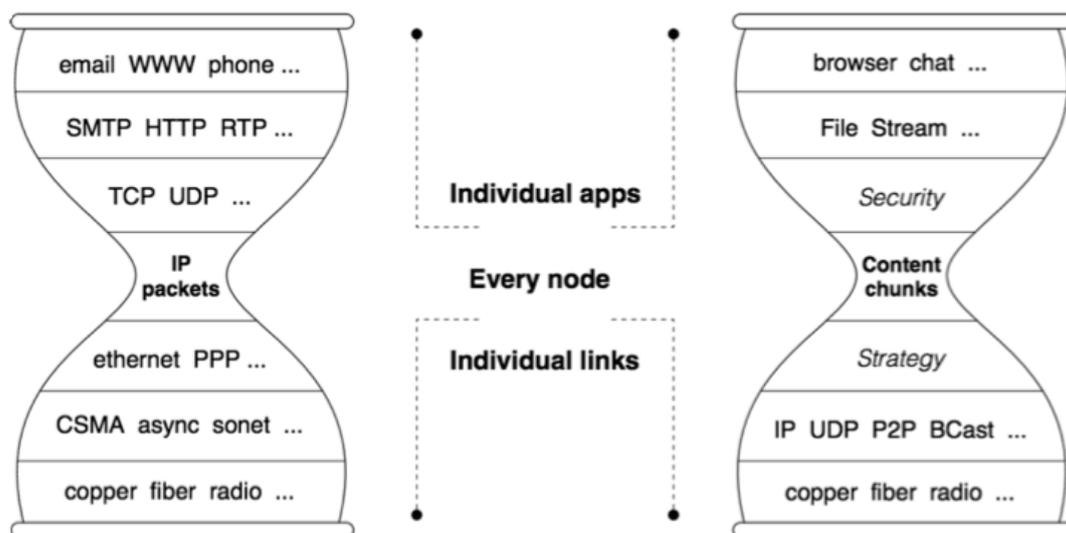


Figura 3.1: architettura internet e clessidra NDN [5]

L'architettura a clessidra è proprio ciò che fa diventare il design originale di internet elegante e potente in quanto è tutto centrato attorno un livello di rete universale che implementa le funzionalità minime necessarie ad un'interconnettività globale.

Questa “vita sottile” è stata una chiave che ha attivato una crescita di internet così esplosiva in quanto ha permesso ai layer inferiori e superiori di portare innovazione ed il principio end-to-end permette lo sviluppo di applicazioni robuste per contrastare guasti di rete.

La separazione di routing e forwarding è risultata necessaria nello sviluppo di internet in quanto permette al forward di funzionare finché il sistema di routing continua ad evolvere nel tempo. [5]

La sicurezza deve essere costruita nell'architettura: la sicurezza nell'attuale architettura di internet è in ripensamento in quanto non accoglie la domanda di un ambiente sempre più ostile; per questo NDN fornisce una sicurezza di base direttamente nella “vita sottile”.

Il traffico di rete si deve auto-regolare perché una consegna equilibrata del flusso di dati è essenziale per avere operazioni di rete stabili e, dato che IP esegue la consegna dei dati in un loop aperto, i protocolli di trasporto sono stati migliorati per fornire un traffico unicast bilanciato: per questo NDN disegna il bilanciamento di flusso nella “vita sottile”.

L'architettura dovrebbe facilitare le scelte degli utenti e la competizione ove possibile e, anche se non è stata un fattore rilevante nel disegno originale di internet, gli sviluppatori ci dicono oggi che l'architettura non è neutrale: ecco perché NDN fa uno sforzo per rafforzare gli utenti finali ed attivare la competizione.

Analogamente all'attuale architettura IP, la “vita sottile” è il centro dell'architettura NDN, la “vita sottile” del NDN utilizza data name anziché indirizzi IP di consegna per offrire un nuovo set di funzionalità di base: questo, che sembra un semplice cambiamento, conduce a significative differenze tra IP e NDN nelle loro operazioni di consegna dati.

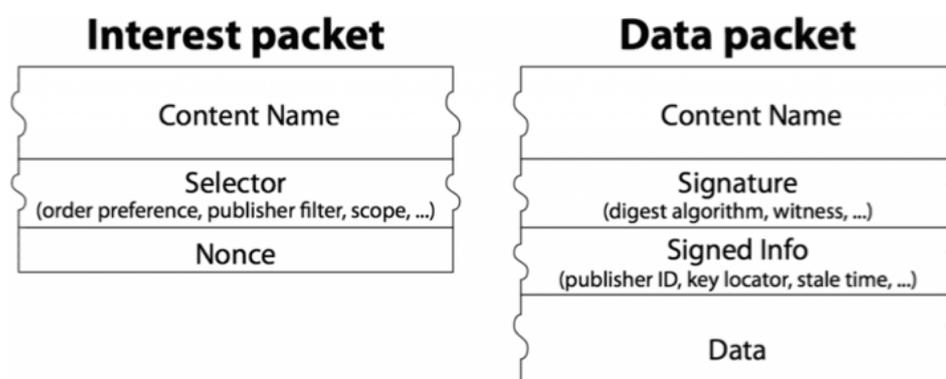


Figura 3.2: pacchetti nell'architettura NDN [5]

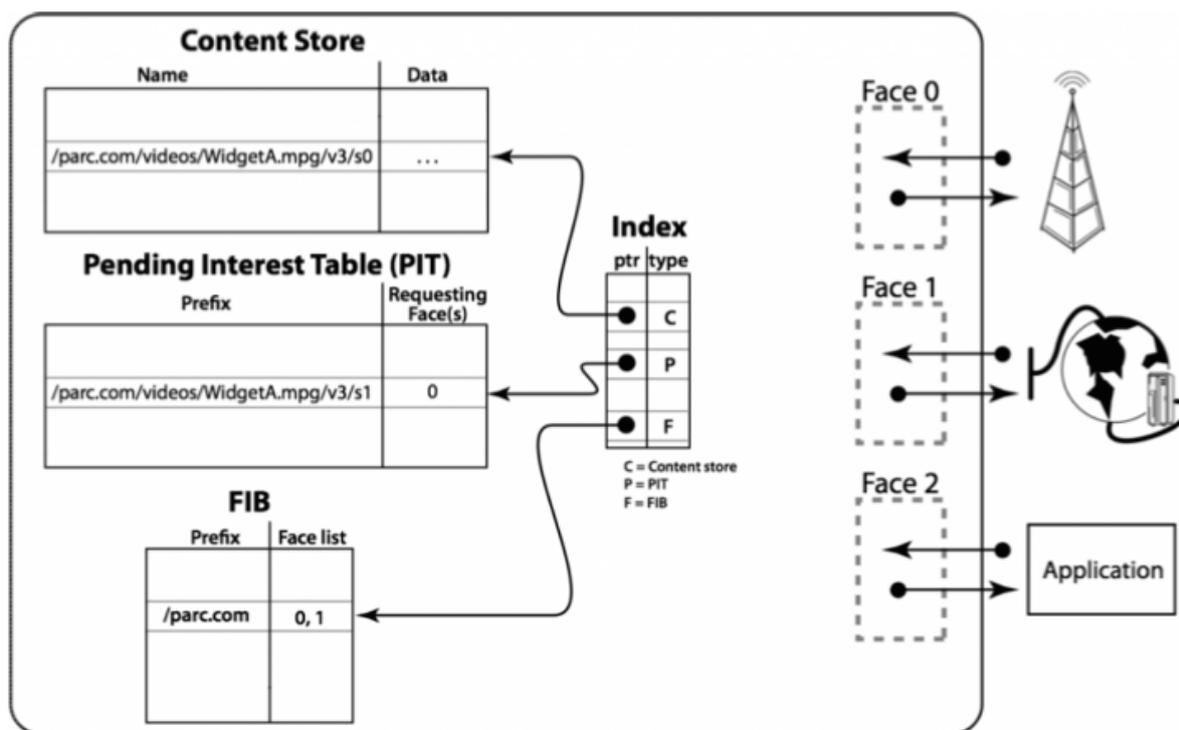


Figura 3.3: nodo NDN [5]

La comunicazione in NDN è guidata dal ricevente, infatti per ricevere il dato l'utente invia un pacchetto interest, il quale trasporta un nome che identifica il dato desiderato: il protocollo di trasporto costruito su NDN viene chiamato CCN (Content-Centric Networking).

CCN non ha nozione di host nel suo più basso livello, un pacchetto "indirizzo" indica il contenuto e non la posizione: il protocollo CCN consegna in modo efficiente il contenuto indicato anziché connettere un host ad altri host.

Tutti i pacchetti dati possono essere messi in cache in qualunque router CCN: ciò, combinato con un intrinseco supporto alla consegna multicast o broadcast, porta ad un utilizzo molto efficiente della rete nel momento in cui molte persone sono interessate al medesimo contenuto.

Il servizio include multipli inoltri per consegne end-to-end, controllo di flusso, consegne multicast trasparenti ed automatiche utilizzando registrazioni in buffer disponibili nella rete, inoltri a più percorsi esenti da cicli, verifica dell'integrità del contenuto con noncuranza del percorso di consegna ed una moltitudine di applicazioni arbitrarie.

Le applicazioni eseguono il protocollo CCN sopra a qualche servizio di comunicazione ad un layer più basso capace di trasmettere pacchetti, non ci sono restrizioni nella natura del servizio: può essere un trasporto fisico o un'altro protocollo di rete o trasporto.

Ad esempio, le applicazioni tipicamente eseguiranno il protocollo CCN sopra ad UDP per trarre vantaggio della connettività IP esistente: dato che il contenuto è indicato indipendentemente dalla posizione, potrebbe anche essere conservato indefinitamente nella rete fornendo efficacemente un servizio di filesystem distribuito.

Il protocollo CCN è generico e supporta un ampio spettro di applicazioni di rete: potrebbe essere naturale pensare ad applicazioni con contenuto memorizzato come distribuzioni di file video o documenti.

Ma il modello CCN supporta anche comunicazioni in tempo reale, protocolli di ricerca ed è abbastanza generico da trasportare conversazioni tra host come connessioni TCP (Transmission Control Protocol).

Il protocollo supporta un gran numero di applicazioni lasciando la scelta delle convenzioni sui nomi alle applicazioni, le complete specifiche sull'uso del protocollo CCN per una particolare applicazione richiederanno specifiche aggiuntive sulle regole dei nomi, formato dati e semantica dei messaggi.

Queste specifiche sui protocolli applicazione sopra al protocollo CCN (unite probabilmente alle specifiche API) sono chiamate profili.

Il protocollo CCN è disegnato per comunicazioni end-to-end tra applicazioni e perciò è inteso per essere integrato nell'elaborazione dell'applicazione anziché essere implementato in un layer separato.

3.2 PSIRP come base per PURSUIT

Esperti da tutto il Mondo stanno cominciando ad essere d'accordo sul fatto che una riforma dei paradigmi di internet e delle tecnologie di base sia necessaria per affrontare le sfide del nuovo millennio: PSIRP (Publish-Subscribe Internet Routing Paradigm) è stato proposto come un rimedio ai problemi che affliggono l'internet attuale.

In una rete publish-subscribe chi invia “pubblica” ciò che vuole inviare e chi riceve “sottoscrive” le pubblicazioni che vuole ricevere: all'inizio nessuno riceve del materiale per cui non ha espressamente richiesto un interesse tramite abbonamento con il risultato di avere un'infrastruttura potente e flessibile ed un alto grado di resilienza. [6]

Si può facilmente osservare che una larga parte dell'utilizzo di internet è già essenzialmente di natura publish-subscribe come la diffusione di aggiornamenti software, consegna di annunci e notizie, radiodiffusione e videodiffusione e la messaggistica periodica ed aperiodica.

Sembra promettente poter arrivare ad un'architettura di un nuovo internet basata sul paradigma publish-subscribe che conduca ad un nuovo disegno di tutti i layer di comunicazione internet.

In questo nuovo internet il multicast ed il caching saranno la normalità mentre sicurezza e mobilità saranno disegnate nell'architettura anziché aggiunte successivamente.

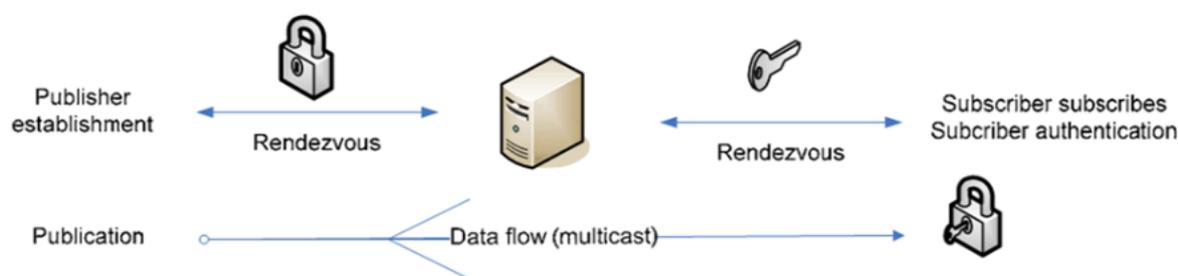


Figura 3.4: interazione publish-subscribe [7]

Il progetto PSIRP ridisegna l'architettura internet dal punto di vista publish-subscribe non dando nulla (nemmeno IP) per scontato e si focalizza sull'intersezione di sicurezza, routing, accessi senza fili, design dell'architettura ed economia di rete al fine di disegnare e sviluppare soluzioni efficienti ed efficaci.

La nuova architettura di rete publish-subscribe ripristinerà l'equilibrio degli incentivi economici di rete tra chi invia e chi riceve ed è ben adatta a per accogliere le sfide delle future applicazioni.

Gli obiettivi principali del progetto PSIRP sono: [8]

- disegnare una nuova architettura di rete basata sul paradigma publish-subscribe
- fare dell'informazione il centro di attenzione e rimuovere la spaccatura luogo-identità che affligge le reti attuali
- implementare innovative caratteristiche multicasting e di caching per ottimizzare prestazioni ed efficienza

- implementare la funzionalità di sicurezza come un componente di base dell'architettura
- implementare e validare la nuova architettura in scenari operazionali realistici
- collaborare con esperti chiave dell'industria e promuovere le loro idee alla comunità attraverso:
 - pubblicazioni
 - presentazioni
 - workshop
 - partecipazioni con altri progetti a livello Europeo ed internazionale
- diffondere i risultati

PSIRP è un progetto di disegno architeturale e perciò stabilisce un processo molto strutturato che permetterà di raggiungere traguardi ambiziosi: il design inizia con un'indagine SoA (State-of-the-Art) in cui vengono studiate le tecnologie e soluzioni attualmente avanzate e future.

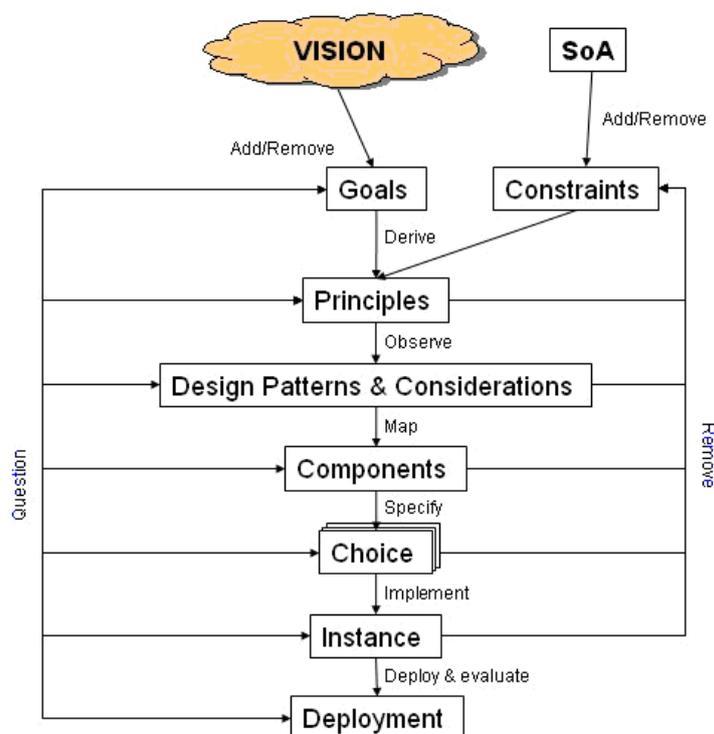


Figura 3.5: design architeturale dei processi [8]

Questa fase iniziale produce la terminologia utilizzata nel progetto, una tassonomia di soluzioni proposte e un'analisi di scritti scientifici chiave: serviranno tutti come pilastri di base per sostenere le future prospettive del progetto.

L'intenzione di seguire un approccio di design clean-state può essere riconosciuto da una formulazione di una chiara "vision" e dalla definizione dei "goal" del design.

Basato su questa metodologia di design, il progetto sviluppa incrementalmente una toolbox di protocolli completa e due implementazioni dell'architettura proposta: [8]

- un approccio evolutivo
 - in cui il publish-subscribe viene eseguito in sovrapposizione sopra all'IP
- un approccio rivoluzionario clean-state
 - in cui i layer più bassi di internet vengono completamente ridisegnati

L'architettura e la sua implementazione saranno validate da un "Red Team" che ha il compito primario di testare l'implementazione con le applicazioni reali in scenari operativi del mondo reale così come di tentare di "romperla" assoggettandola a DoS (Denial of Service) ed altri tipi di attacchi.

Le esperienze guadagnate dal processo di validazione saranno utilizzate per migliorare il disegno e l'implementazione in modo iterativo, conducendo ad una soluzione finale che comprenda la migliore combinazione di perfezionamento, affidabilità, sicurezza e prestazione.

Tra i punti chiave che PSIRP dovrà indirizzare troviamo: [8]

- scalabilità
 - metodi quantitativi valuteranno com'è scalabile la soluzione proposta
- sicurezza
 - non può essere trattata come un'entità separata ma come parte integrante del disegno e dell'implementazione
 - fra le sfide di sicurezza più difficili ci sono la protezione contro traffico non voluto ed il DoS
 - è valutata dal "Red Team" che cercherà di "romperne" il design

- efficienza
 - un'efficiente distribuzione di un grande ammontare di informazioni, includendo i video, si affida all'uso di multicasting e caching
 - nella nuova architettura non è l'eccezione ma la norma
- impatto socio-economico
 - disegnare un'architettura significa anche permettere prospettive di business
 - capire l'impatto sui business e la società del futuro come un tutt'uno
 - cruciale per valutare l'attuabilità della soluzione proposta
- sfruttamento
 - l'utilizzabilità di una innovazione tecnica è strettamente limitata dalla propria facilità d'uso
 - internet è uno strumento unico perché ha un'enorme base di utenti
 - le soluzioni PSIRP sono disegnate per fornire potenti vantaggi a tutti gli utenti

Ci si aspetta che il progetto abbia questi impatti: [8]

- incrementare la conoscenza Europea dell'architettura publish-subscribe e delle possibilità dell'architettura stessa
- influenzare la standardizzazione dell'internet del futuro e dare all'Europa la possibilità di influenzarne la direzione
- Portare a migliori servizi ICT (Information and Communication Technology) per:
 - la formazione
 - il business
 - la cura del malato e dell'anziano
 - lo svago del popolo Europeo
- dare ai gestori di telecomunicazioni Europei un vantaggio iniziale nel fornire prodotti e servizi per l'internet del futuro
- concedere alle SME (Small and Medium sized Enterprise) la possibilità di entrare nel mercato dell'internet del futuro come conseguenza dei risultati di PSIRP ed implementazioni open-source

Alla luce della sempre più crescente importanza del ICT in tutte le aree della società e del ruolo di internet come componente centrale del ICT, è vitale per l'Europa essere coinvolta attivamente nella creazione dell'internet del futuro.

La visione PSIRP è confluita in PURSUIT, continuazione del progetto che darà all'Europa l'opportunità di influenzare il disegno che interesserà la vita di tutte le persone nel Mondo e darà anche alle aziende Europee un buon punto di partenza per offrire prodotti e servizi del futuro. [9]

3.3 NetInf: da 4WARD a SAIL

L'assenza di strutture adeguate per la progettazione, ottimizzazione, ed interoperabilità di nuove reti attuali obbliga una convergenza verso un'architettura che sia ottimale per molte applicazioni che però non può sostenere le innovazioni di internet.

4WARD combina un insieme di approcci radicali architetturali sulla base di un forte background mobile e wireless per la progettazione di famiglie di architetture di rete complementari che possano interoperare.

4WARD ha l'obiettivo di: [10]

- migliorare la nostra capacità di progettare famiglie di architetture
- permettere la coesistenza di più reti su piattaforme comuni attraverso la virtualizzazione carrier-grade delle risorse di rete
- migliorare l'utilità delle reti rendendole auto-gestite
- aumentare la loro robustezza e la loro efficienza sfruttando le diversità
- migliorare il supporto di applicazione da un nuovo paradigma information-centric al posto del vecchio host-centric

Questo progetto mira ad accrescere la competitività dell'industria europea e di messa in rete per migliorare la qualità di vita dei cittadini europei mediante la creazione di una famiglia di reti affidabili ed interoperanti, fornendo un accesso diretto ed onnipresente alle informazioni.

Anche le reti wireless e wireline del futuro saranno concepite per essere facilmente adattabili alle esigenze presenti e future ad un costo accettabile.

L'obiettivo è rendere lo sviluppo delle reti e delle applicazioni di rete più semplice e veloce, guidando sia i servizi di comunicazione più avanzati sia quelli più accessibili.

In questo approccio si combinano da un lato le innovazioni necessarie per migliorare il funzionamento in ogni singola architettura di rete e dall'altro più architetture di rete diverse e specializzate che sono fatte per lavorare insieme in un quadro complessivo.

Si lavora:

- su innovazioni per superare le carenze delle attuali reti di comunicazione
- in un contesto che permetta la coesistenza, interoperabilità e la complementarità delle diverse architetture di rete
- in modo integrato, evitando le patch che oggi insidiano internet

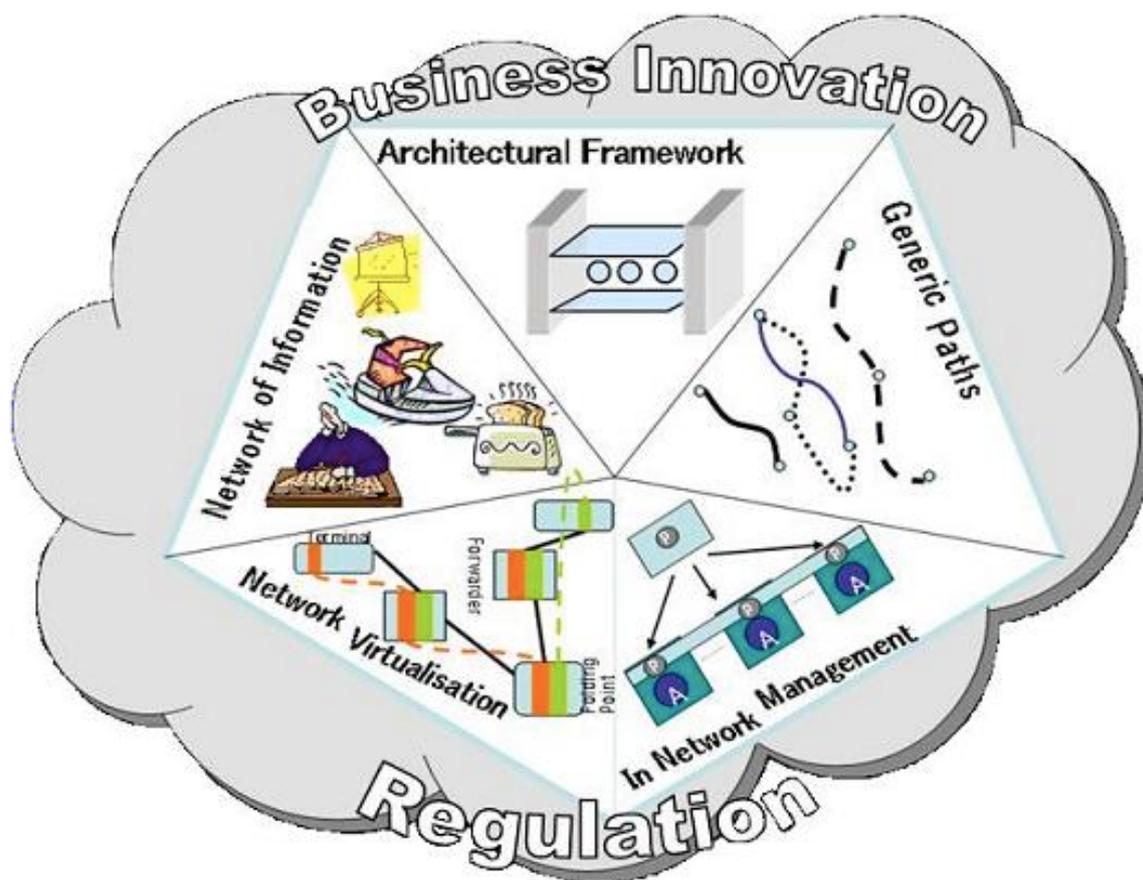


Figura 3.6: organizzazione del progetto [10]

Questo lavoro è strutturato in sei pacchetti di lavoro: tre prendono in considerazione le innovazioni per un'unica architettura (Generic Path, In-Network Management, Network of Information).

Un pacchetto utilizza la virtualizzazione per consentire a molteplici architetture di rete esistenti sulla stessa infrastruttura di operare, un'altro pacchetto esamina la progettazione e lo sviluppo di architetture interoperabili ed infine un pacchetto assicura che tutti gli sviluppi previsti non tengano conto dei problemi tecnici non essenziali.

Un consorzio di 25 partner tra cui i più grandi operatori di telecomunicazioni, rivenditori ed istituti di ricerca hanno fatto partire il 1 agosto 2010 il progetto SAIL (Scalable & Adaptive Internet soLutions) con l'obiettivo di disegnare l'architettura dell'internet del futuro partendo dagli studi di 4WARD chiuso ufficialmente il 30 giugno 2010. [11]

Il progetto SAIL si focalizza su tre principali obiettivi tecnici e fornirà soluzioni complete e funzionanti per rispondere ai cambiamenti integrando un Network of Information e Cloud Networking con Open Connectivity Services.

I modelli di interazione delle applicazioni emergenti non comporteranno più il semplice scambio di dati end-to-end, questi nuovi modelli sono centrati su pezzi di informazioni a cui si accede in vari modi.

Invece di accedere e manipolare le informazioni solo tramite un riferimento indiretto dei server che li ospitano, mettere gli oggetti informazione al centro della rete è attraente ed ha portato alla nozione di NetInf (Network of Information).

Il NetInf semplifica la comunicazione rispetto ad una rete end-to-end o store-and-forward e lo rende più adatto per applicazioni complesse a larga scala e per una migliore programmazione ed efficienza di rete.

Con il caching di più copie uguali dello stesso oggetto informazione, una NetInf supporta nativamente la distribuzione di contenuti su larga scala: queste proprietà la rendono adatta per una rete tollerante ai ritardi.

L'architettura della NetInf supporta nativamente la mobilità degli oggetti informazione, nodi, utenti ed applicazioni ma non è ancora chiaro come assicurarsi che un tale modello possa essere supportato in modo efficiente in rete in fase di esecuzione. [12]

Una tale integrazione offrirà notevoli benefici, ad esempio riducendo il ritardo di accesso alle informazioni e la larghezza di banda necessaria, migliorando l'affidabilità e semplificando il funzionamento della rete.

Applicazioni di rete possono variare rapidamente in popolarità ed in termini di quantità di interazione con l'utente: ciò rende il provisioning un problema difficile, sia lato server che stoccaggio, nonché sulla rete.

Il cloud computing ha affrontato con successo molte di queste sfide utilizzando la virtualizzazione come tecnica di base, tuttavia non è ancora chiaro come fornire il supporto di rete adatto per applicazioni altamente variabili quando vengono eseguite dentro le più complesse e diverse reti degli operatori di telecomunicazioni.

Anche se è possibile fornire le risorse computazionali, non è evidente come fornire dinamicamente il necessario supporto di rete e capacità o la complessa topologia di rete necessaria, ad esempio, ad una applicazione web su più livelli.

Un esempio di scenario complesso sarebbe una NetInf che non solo serve informazioni statiche ma anche dinamiche che richiedono risorse di calcolo allocate dinamicamente in una fetta di rete quando le entità NetInf stanno raggiungendo i limiti di capacità.

In linea di principio le tecnologie di base per trasportare efficientemente anche grandi quantità di dati sono già in atto, ma non si mescolano bene con gli attuali concetti di controllo internet.

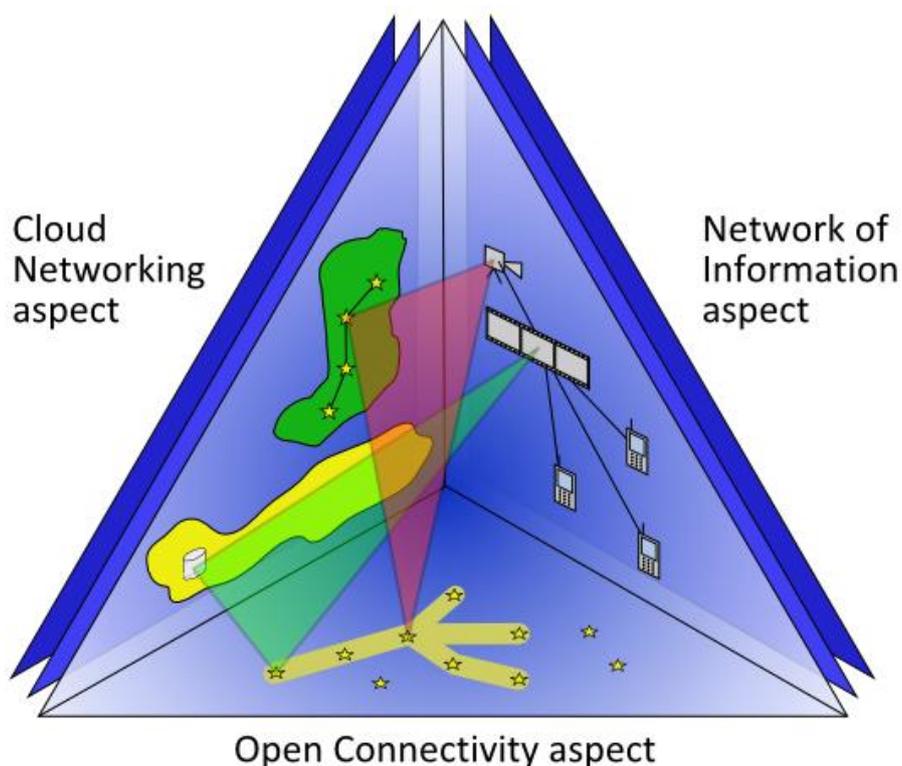


Figura 3.7: i tre obiettivi principali di SAIL [13]

Infatti le soluzioni attuali cozzano con l'eterogeneità delle reti distribuite e non possono integrare perfettamente il controllo di queste risorse: è ad esempio impossibile sfruttare le diversità esistenti su diverse tecnologie di comunicazione tra due punti finali.

Il Cloud Networking, e soprattutto il NetInf, cambierà in modo sostanziale i modelli di traffico ed i flussi di dati in internet e si raggiungeranno requisiti per un servizio di connettività ancora migliore.

SAIL si occupa di aspetti particolari di una vera e propria rete complessa: la figura mostra come le entità si riflettano in modo diverso in Network of Information, Cloud Networking e Open Connectivity Services, ma rimangano in contatto e relazione tra loro. [12]

La figura evidenzia inoltre che ci si aspetta che il Network of Information ed il Cloud Networking appaiano più volte, mentre il Open Connectivity Services costituisca un'unica base per gli altri due.

Inoltre ci sarà da lavorare anche su temi trasversali che influiscono su ciascuna delle aree tecniche e ci sarà anche da garantire la standardizzazione, migrazione e gli aspetti socio-economici, commerciali, normativi e dovranno essere adeguatamente considerati le attività di valorizzazione e diffusione.

3.4 DONA

Internet si è molto evoluto dalla sua originale incarnazione, ad esempio la gran maggioranza dell'uso che se ne fa oggi è per recuperare dati ed accedere a servizi quando invece l'architettura era stata disegnata attorno ad applicazioni host-to host come telnet e FTP.

Inoltre l'internet originario era un puro trasportatore di pacchetti in modo trasparente, adesso invece i consegnatari di rete utilizzano middlebox per aumentare la sicurezza ed accelerare le applicazioni.

Per adattare internet a questi cambiamenti, la proposta DONA (Data-Oriented Network Architecture) consiste in una riprogettazione completa del naming e risoluzione dei nomi di internet.

Il sistema di risoluzione dei nomi DNS (Domain Name System) è una parte fondamentale dell'internet di oggi essendo alla base di quasi tutto l'utilizzo, anche se è stato sviluppato abbastanza recentemente nell'evoluzione di internet dopo che molti pezzi di base dell'architettura erano già dati per assodati.

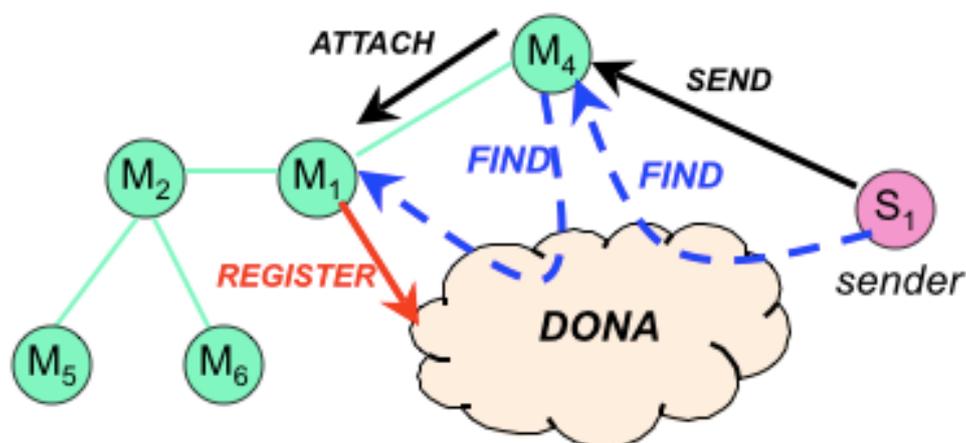


Figura 3.8: DONA [14]

Ad esempio, le sessioni TCP erano già limitate agli indirizzi IP e per questo il socket API (Application Programming Interface) fa riferimento agli indirizzi piuttosto che ai nomi: decisioni di design congelate, come queste, hanno limitato il modo in cui i nomi DNS (o qualsiasi altro sistema) possano permeare l'architettura.

Come risultato il ruolo corrente del naming nell'architettura è più un incidente di percorso che il risultato di un principio di design architetturale, ecco perché la proposta DONA guarda ad una completa riprogettazione di naming e risoluzione dei nomi.

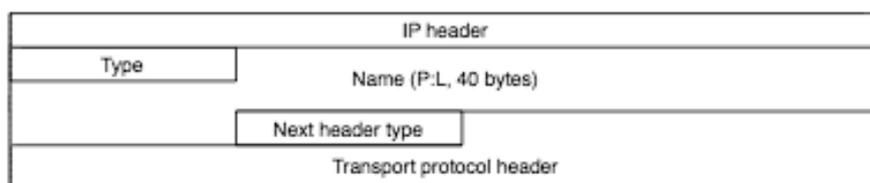


Figura 3.9: intestazione del protocollo FIND [15]

Lo scopo di ogni architettura è capire come soddisfare un largo spettro di potenziali utenti, sia quelli che incontriamo nel presente sia quelli che incontreremo nel futuro: il primo tema è spostare l'utilizzo delle applicazioni da host-centric a data-centric.

Le prime applicazioni internet, come il trasferimento di file e l'accesso remoto, erano strettamente basate su comunicazioni host-to-host: l'utente dirigeva esplicitamente la sorgente a comunicare con un altro host e l'unico ruolo della rete era trasportare i pacchetti a destinazione all'indirizzo scritto nell'intestazione.

L'architettura di internet era costruita attorno a questo modello host-to-host e, come risultato, si adatta bene per comunicazioni tra host fissi: oggi tuttavia l'utente è interessato al contenuto e non a dov'è localizzato.

Consideriamo questi tre temi rilevanti lato utente: [15]

- persistenza
 - una volta dato un nome per dei dati o servizi, l'utente vorrebbe che questo nome restasse valido finché il dato o servizio è disponibile
 - non ci dovrebbero essere l'equivalente odierno dei broken links quando i dati sono stati spostati da un'altra parte
 - oggi la ridirezione HTTP ed il DNS dinamico vengono utilizzati per minimizzare il problema ma non sono risposte sufficienti dato che nessuno dei due funziona se i dati cambiano di dominio a meno che l'operatore del dominio precedente fornisca perpetuo supporto
- disponibilità
 - dati e servizi dovrebbero avere alta disponibilità sia in termini di affidabilità che di bassa latenza
 - la disponibilità è fornita di solito replicando i punti di approvvigionamento ed il ruolo della rete è permettere alla richieste dell'utente di trovare copie vicine
 - la prima soluzione di larga scala a ciò fu data da Akamai utilizzando server DNS "intelligenti" ed una riscrittura delle URL
 - più recentemente meccanismi peer-to-peer come BitTorrent sono diventati prevalenti
 - non è del tutto chiaro come la disponibilità debba basarsi su un set ad hoc di meccanismi specifici per applicazioni
- autenticità
 - gli utenti vorrebbero sapere se i dati provengono da sorgenti appropriate
 - oggi ciò richiede un PKI (Public Key Infrastructure) che fornisca gli utenti di chiavi pubbliche del provider
 - oggi l'autenticità è tipicamente raggiunta rendendo sicuro il canale del sorgente piuttosto che autenticando esplicitamente i dati

Quindi molte delle funzionalità più naturali che l'utente vorrebbe per i servizi di accesso e recupero dati (persistenza, disponibilità ed autenticità) sono state create in modo troppo difficile per l'attuale modello host-to-host di internet, richiedendo molto spesso attività alternative maldestre e care.

Stabilita la discordanza tra il design storico host-centric e l'attuale utilizzo data-centric, sorge una domanda: come sarebbe l'architettura se fosse costruita attorno agli accessi a dati e servizi?

In maniera abbastanza sorprendente molti dei cambiamenti necessari risiedono nel come i nomi internet sono strutturati, ecco perché viene proposto di sostituire i nomi DNS con nomi lineari auto-certificati e sostituire la risoluzione dei nomi DNS con un anycast primitivo basato sui nomi che sta sopra il layer IP.

DONA migliora il recupero dati ed i servizi di accesso fornendo un supporto più forte ed architettualmente coerente per persistenza, disponibilità ed autenticità che può essere esteso per fornire supporto al caching ed aggiornamenti tipo RSS.

Tuttavia l'impatto di DONA non è limitato solo agli accessi a dati e servizi, saranno utilizzate queste applicazioni da esempio motivante in quanto forzeranno a pensare in modo differente ad altre tematiche fondamentali: come risultato il design globale di DONA ha implicazioni architetturali che vanno molto oltre gli accessi a dati e servizi.

L'anycast primitivo basato sui nomi di DONA è utile per trovare risorse, ad esempio può fornire le basi che riguardano SIP, supporto alla mobilità del host ed alla multidirezionalità, e stabilire stati di inoltro per multicast tra domini.

Posizionando anycast al layer naming, anziché al layer IP, permette di scrivere un'architettura votata alla funzionalità non limitata da preoccupazioni sulla scalabilità dato che il meccanismo non ha bisogno di operare alla velocità del collegamento.

C'è un altro tema in cui il design attuale è in disaccordo con l'utilizzo attuale: l'originale architettura di internet, seguendo il principio end-to-end, intendeva che la rete fosse un puro trasportatore trasparente di pacchetti.

Oggi invece i consegnatori di rete (come le aziende) utilizzano middlebox per aumentare la sicurezza (firewall, proxy) ed accelerare le applicazioni (cache).

Dato che il processo di risoluzione dei nomi anycast di DONA segue essenzialmente lo stesso percorso amministrativo come il seguire i pacchetti dati, DONA può comportarsi con i consegnatori durante il percorso come un rilevante attore internet: ciò permette a DONA di fornire un chiaro supporto per i middlebox.

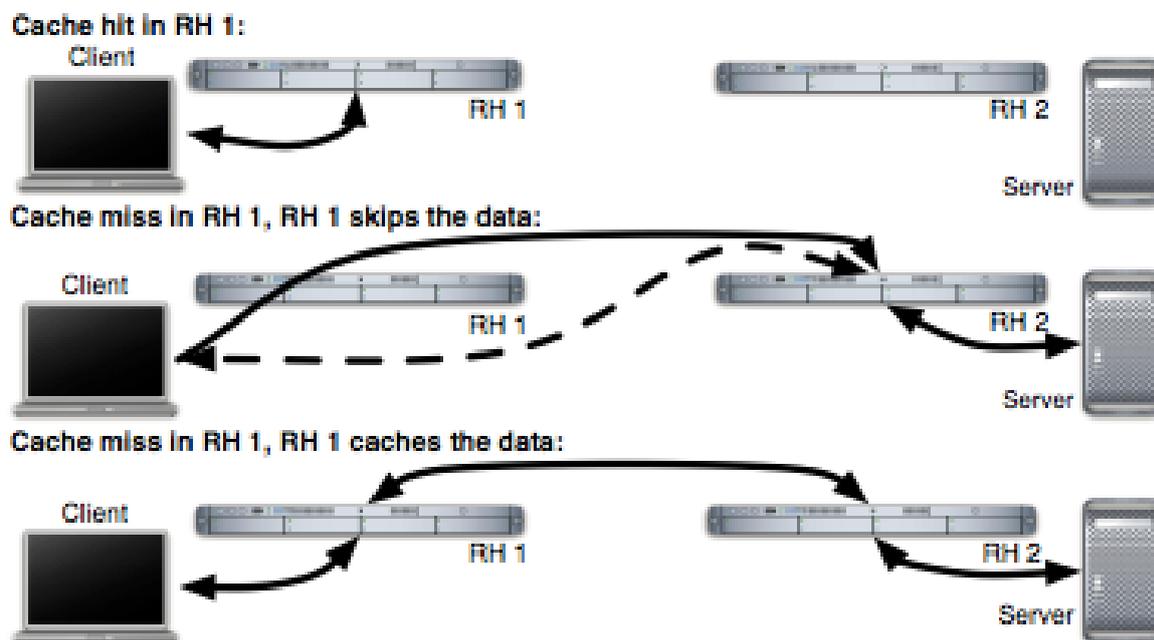


Figura 3.10: elaborazione del pacchetto FIND con supporto cache [15]

Questo non vuol essere un ripudio del principio end-to-end dato che la funzionalità è comunque fornita: è soltanto il riconoscimento che gli operatori dovrebbero avere a loro disposizione meccanismi architetturalmente coerenti per controllare che traffico attraversa le loro reti.

Più recentemente ci sono state molte mani alzate riguardo la scalabilità del routing nell'attuale paradigma di indirizzi.

L'anycast primitivo di DONA fornisce un meccanismo di ricerca che sta al di sopra del layer IP e ciò permette l'utilizzo di contrassegni che descrivano la natura delle associazioni tra oggetti anziché indirizzi globali, un approccio che darà come risultato piccole tabelle di routing tra domini.

Ad un livello più speculativo DONA rappresenta un parziale cambio di direzione da un approccio basato su chi manda ad uno basato maggiormente su chi riceve, uno dei compiti sarà capire quanto si potrà andare in questa direzione e cosa ciò comporterà per l'internet del futuro.

Queste implicazioni architetturali ci dicono che DONA non sarà solamente ristretto agli accessi di dati e servizi (che sono comunque davvero significativi per il loro dominante utilizzo nell'attuale internet), ma faciliterà anche miglioramenti in tutte le dimensioni.

Tuttavia ci sono una serie di tematiche non ancora discusse che meritano attenzione: internet necessita ancora di una miglior sicurezza (particolarmente contro DoS e router maliziosi o mal configurati), miglior maneggevolezza, miglior usabilità e molte altre proprietà.

La speranza è di riuscire a lavorare su questi problemi con un framework più largo che includa anche DONA.

Capitolo 4

CARATTERISTICHE

4.1 Naming e Sicurezza

Uno schema di denominazione che consenta l'identificazione degli oggetti informazione indipendentemente dalla posizione è probabilmente la parte più importante di una progettazione datacentrica.

Gli obiettivi del progetto per il naming di NetInf includono la persistenza, l'autocertificazione, l'autenticazione del proprietario e l'identificazione del proprietario.

La persistenza può essere mantenuta anche se cambiano la posizione di archiviazione, il contenuto, il proprietario del contenuto o la struttura organizzativa.

La separazione tra auto-certificazione, l'autenticazione del proprietario e l'identificazione del proprietario consente a quest'ultimo di mantenere l'anonimato e permette al publisher di essere diverso dal proprietario stesso.

Entrambe queste proprietà sono importanti come basi per la veridicità, la privacy e la gestione dei contenuti.

Così come i principali obiettivi di progettazione sono relativi alla sicurezza, i nomi NetInf sono essenzialmente relativi ad uno spazio piatto: non c'è, ad esempio, una struttura gerarchica che possa essere facilmente utilizzata per il routing.

I campi di tipo, autenticazione e le etichette possono comunque fornire una qualche struttura, specialmente il campo autenticatore può essere utile per rendere più efficiente la risoluzione dei nomi ed il routing.

Lo schema di denominazione CCN è gerarchico al fine di conseguire la migliore scalabilità di routing: i nomi sono radicati in un prefisso unico per ogni editore, il prefisso editore rende possibile per i client creare nomi validi per dati che ancora non esistono e gli editori sono in grado di rispondere con dati generati in modo dinamico. [16]

I nomi CCN sono utilizzati sia per la denominazione di informazioni che per il trasporto, la granularità dei nomi è molto fine perchè vengono nominati i singoli pacchetti.

PSIRP fa uso di due spazi principali, identificatori d'incontro e di spedizione, i primi (insieme con gli identificatori di scopo) nominano i punti d'incontro che vengono utilizzati per stabilire un contatto tra editori ed abbonati: non vi è uno spazio per un oggetto informazione di per sé, ma può essere utilizzato il nome stesso. [16]

Sia CCN che NetInf fanno il contrario: i nomi degli oggetti informazione possono anche essere utilizzati per nominare un punto di incontro: in linea di principio non c'è una differenza fondamentale, è più una questione di punti di vista. [16]

Il secondo spazio, gli identificatori di spedizione, viene utilizzato dal tessuto di inoltro di PSIRP per il trasporto dei dati, dopo che si è stabilito il contatto in un punto di incontro.

Entrambi i namespace sono piatti e possono essere basati sull'hash del contenuto o sull'hash di una chiave pubblica di un editore.

Gli URI DONA sono in formato P:L, in cui P è il "principale" che identifica l'editore dell'oggetto informazione, e L è l'etichetta (label) [16]: gli oggetti identificatore sono piatti e univoci in ogni namespace ed i "principali" sono globalmente univoci, i nomi sono intrinsecamente sicuri in quanto contengono l'hash della chiave pubblica dell'editore.

4.2 Risoluzione dei nomi e Routing

Risoluzione dei nomi di solito significa che un servizio di risoluzione venga interrogato ed uno o più locatori vengano restituiti, che poi può essere utilizzato per recuperare l'oggetto: questa fase viene chiamata risolvere / recuperare in due tempi.

Un alternativa è ritornare direttamente l'oggetto senza prima ritornare i locatori: questa fase viene chiamata risolvere / recuperare in un tempo.

Anche se quest'ultimo viene spesso definito come routing basato sul nome, i servizi di risoluzione dei nomi spesso utilizzano il routing basato sul nome per trovare la risposta ad una query.

Il routing basato sul nome può essere descritto come un metodo per comunicare attraverso il layer naming piuttosto che attraverso il layer di inoltro: ecco perché la risoluzione dei nomi ed il routing basato sul nome vengono considerati più o meno equivalenti.

La differenziazione importante è invece se i locatori sono o meno visibili al client: la scelta progettuale fatta per l'API NetInf è stata di nascondere i locatori dall'applicazione client e ciò significa che sia la fase a due tempi che quella ad un tempo sono supportate dal principio.

In NetInf sono stati sviluppati due meccanismi di resolution, uno che può fare la fase risolvere / recuperare in un tempo chiamato Multi-level DHT che utilizza una gerarchia di DHT per eseguire la resolution ed eventualmente anche il trasferimento dei dati.

C'è anche un approccio a due tempi LLC (Late Locator Construction) che si concentra sulla gestione di topologie di rete altamente dinamiche, comprese le reti mobili di grandi dimensioni.

Per rendere la risoluzione dei nomi "robusto" contro cambiamenti di topologia grandi e frequenti, i locatori sono costruiti al momento della richiesta di resolution: il vantaggio della scelta in due tempi è che possono essere utilizzati molti meccanismi di trasporto diversi, compresi quelli esistenti, facilitando la distribuzione nelle reti esistenti.

Con questo modello NetInf ha due livelli di routing, uno per la resolution dei name ed una per il trasporto: la fase ad un tempo è invece potenzialmente un approccio più semplice e più efficiente senza dipendenze da infrastrutture esistenti. [16]

CCN ha un meccanismo risolvere / recuperare in un tempo: i client chiedono un oggetto informazione inviando pacchetti di interesse che vengono instradati verso l'editore del prefisso nome utilizzando la corrispondenza più lunga. [16]

Quando si incontra una copia dell'oggetto informazione sul percorso, un pacchetto dati contenente l'oggetto richiesto viene inviato sul percorso inverso al client.

PSIRP utilizza un modello a due tempi in cui il decisore è chiamato punto d'incontro (rendezvous) e la trasmissione dei dati può, come in NetInf, prendere potenzialmente una strada di ritorno diversa per il client dalla resolution dei nomi o punto d'incontro che non deve essere né il percorso per l'editore né tenere una copia dei dati. [16]

I dati vengono trasmessi tramite routing con un filtro Bloom che descrive il percorso costruito dal punto d'incontro ed utilizzato dal richiedente per raggiungere la destinazione: il filtro è attaccato al pacchetto stesso e contiene tutti i nomi dei collegamenti che devono essere seguiti.

In DONA i nodi che sono autorizzati a servire i dati vengono registrati per l'infrastruttura di resolution e, solo una volta che un determinato contenuto è stato registrato, le richieste possono essere indirizzate ad esso: il comando register ha un TTL (Time To Live) che deve essere rinnovato dopo la sua scadenza. [16]

Gli RH (Resolution Handlers) hanno una struttura gerarchica: le richieste vengono instradate in base al nome in modo gerarchico ed ogni richiesta che non può essere soddisfatta viene inoltrata al RH padre.

L'infrastruttura di resolution richiede gli itinerari per nome e cerca di trovare una copia od il contenuto più vicino al client: l'infrastruttura di risoluzione è pensata per lavorare su IP.

Il client si connette alla fonte dei contenuti una volta che un nome viene risolto e la NAT (Network Address Translation) viene eseguita dai router di caching per memorizzare in modo trasparente il contenuto.

4.3 Archiviazione in rete

L'archiviazione in rete per il cache di oggetti informazione è una componente fondamentale di NetInf e di altri approcci datacentrici al networking in quanto è una delle caratteristiche distintive rispetto all'overlay ed alla tecnologia peer-to-peer.

Senza archiviazione in rete, NetInf non sarebbe una tecnologia a livello di rete ed il caching è strettamente su base opportunistica: una cache può eliminare un oggetto memorizzato in qualsiasi momento.

In NetInf ci sono due modi per utilizzare una copia cache di un oggetto: nel primo la copia può essere trovata direttamente interrogando il sistema di risoluzione dei nomi se la copia è esplicitamente registrata o trovata da un sistema di risoluzione dei nomi basato sulla ricerca locale se la copia è in fase di query.

Il secondo modo prevede che la copia possa essere trovata da un protocollo "consapevole" di trasporto NetInf in un percorso noto per contenere una copia, ad esempio una posizione recuperata dal sistema di risoluzione dei nomi. [16]

CCN indirizza una richiesta di dati verso l'editore e fa uso di copie cache lungo il percorso: le copie possono anche essere trovate mediante ricerca locale.

Questi casi sono funzionalmente abbastanza equivalenti in NetInf: poiché il trasporto CCN ed il routing basato sui nomi sono integrati non c'è alcun caso separato per il trasporto CCN. [16]

Si deve notare che in CCN gli oggetti atomici sono singoli pacchetti per cui è possibile che solo una parte di un oggetto più grande sia effettivamente in cache.

In PSIRP il caching è limitato al campo di applicazione del punto d'incontro per l'identificatore associato ad un oggetto, all'interno di tale ambito un oggetto può essere memorizzato in cache multiple. [16]

In DONA il caching è insito nell'architettura: ogni cache può rispondere ad una richiesta di find e servire il relativo oggetto informazione. [16]

4.4 API

Le API utilizzate negli attuali approcci hanno di solito solo alcune funzioni per lo più legate alla risoluzione e registrazione dei nomi e richieste di inoltra oggetti.

Ecco le funzioni chiave di ogni approccio: [16]

CCN

API comuni

ccn.connect		connette al servizio CCN locale
ccn.disconnect		disconnette dal servizio CCN locale

API trasmettitore

ccn.put	(PUT)	invia un oggetto dati
ccn.set_interest_filter		crea un filtro per ricevere uno specifico interesse

API ricevitore

ccn.express_interest	(GET)	esprime un interesse
ccn.get	(GET)	riceve un oggetto dati
ccn.verify_content		verifica un oggetto

PSIRP

API comuni

attachPolicy		assegna una policy al dato ambito o ID del punto d'incontro senza necessità di riabbonarsi
associate		associa una dato ID del punto d'incontro ad un dato ambito

API trasmettitore

forward		invia i dati ad un ricevitore utilizzato quando un editore è pronto a cominciare la trasmissione agli abbonati e richiede le informazioni di spedizione di agire in tal modo
register_provision		
withdraw_provision		annullare una disposizione in corso
publish	(PUT)	pubblica un oggetto
advertise	(PUT)	annunciare un dato oggetto

API ricevitore

listen		riceve i dati dallo stream
mute		smette di ricevere i dati dallo stream
register_interest	(GET)	utilizzato da un abbonato quando è pronto a ricevere i dati
withdraw_interest		annullare un interesse in corso
subscribe	(GET)	abbona al dato ID punto d'incontro
unsubscribe		annulla l'abbonamento attivo

NetInf

API trasmettitore

publish	(PUT)	pubblica un oggetto
revoke		revoca una precedente pubblicazione
send		stream dati ad una connessione tipo simil-canale

API ricevitore

resolve	(GET)	risolve un nome in una lista di descrittori
join		connette ad un servizio
receive	(GET)	riceve un oggetto od uno stream da una connessione

DONA

API trasmettitore

REGISTER	(PUT)	registra un oggetto al sistema di risoluzione dei nomi
UNREGISTER		annulla la registrazione di un oggetto al sistema di risoluzione dei nomi

API ricevitore

FIND	(GET)	ottiene l'oggetto corrispondente al nome dato
------	-------	---

Capitolo 5

PROPRIETÀ

5.1 Scalabilità

La tecnologia per le reti di comunicazione in tutto il Mondo deve essere progettata e valutata rispetto alla scalabilità e deve essere in grado di soddisfare i requisiti previsti attuali e futuri sia per le dimensioni della rete che delle prestazioni.

Scalabilità e prestazioni spesso vanno di pari passo dato che le prestazioni di solito diminuiscono all'aumentare delle dimensioni della rete.

Uno dei principali problemi della scalabilità delle reti IP di oggi è il numero di prefissi di indirizzi IP nelle tabelle di routing: per le reti datacentriche il problema di scalabilità principale è invece il numero di oggetti informazione o, più precisamente, la quantità di contabilità necessaria per tenere traccia di tutti gli oggetti informativi.

L'evidenza empirica mostra che è possibile operare una DHT (Distributed Hash Table) 24×7 che in un dato momento consista in più di due milioni di nodi: supportare quindi due milioni di nodi per l'infrastruttura di risoluzione non è quindi irragionevole. [17]

Stabilendo a 10^{15} il numero di oggetti che necessitano di essere supportati, tre ordini di grandezza superiori al numero di URL univoci attuali (al 2013), senza repliche e 100 B per la risoluzione del record ogni nodo ha bisogno di 50 GB di memoria, con 10 repliche e 1 kB per ogni record c'è invece bisogno di 5 TB di memoria.

50 GB di memoria possono essere implementati con tecnologia DRAM (Dynamic Random Access Memory) ma anche facilmente supportati con tecnologia SSD (Solid-State Drive), 5 TB sono possibili con tecnologia SSD allo stato dell'arte, oggi certamente costosa.

Il fatto che NetInf abbia un namespace piatto degli oggetti informazione significa che il sistema di risoluzione dei nomi di NetInf deve sostenere una voce per oggetto informazione, chiaramente un numero enorme. [16]

Se il routing può essere aggregato all'etichetta autenticatrice nei nomi NetInf, forse la contabilità viene ridotta al livello del numero di nomi di dominio di II livello.

I nomi degli oggetti in CCN sono di progettazione gerarchica, il sistema globale di routing CCN deve almeno essere in grado di gestire i prefissi a livello di editori.

Un altro problema di scala in CCN è che lo stato di inoltra per pacchetto è necessario lungo tutto il percorso end-to-end per guidare i pacchetti dati ai client: come plus, l'interesse per pacchetto permette il multicast senza infrastrutture aggiuntive. [16]

In PSIRP l'utilizzo di punto d'incontro e di identificatori di ambito rende possibile una struttura a due strati in modo tale che spazi di nomi piatti di punti d'incontro possono essere gestiti in ambiti separati; gli ambiti non hanno alcuna relazioni o strutture fissate.

La registrazione di punti d'incontro ed ambiti di identificazione è quindi dimensionata sulla base della conoscenza del locatore del luogo d'incontro all'interno di un ambito, la locazione degli ambiti può essere paragonata alle voci DNS tale che il numero di ambiti può almeno essere quanto il DNS serve al momento.

PSIRP ha una gerarchia del nome a due livelli con gli identificatori ambito e punto d'incontro: gli identificatori punto d'incontro sono registrati in un determinato ambito, gli ambiti non hanno alcuna relazione o struttura fissa, il che significa che è necessario prima trovare il campo di applicazione corretta al fine di trovare un certo punto d'incontro. [16]

Gli ambiti possono essere confrontati con i prefissi editore del CCN e l'etichetta autenticatore di NetInf: prima sono stati esclusivamente confrontati con il numero di domini di II livello in DNS.

Gli ambiti però possono anche essere confrontati con i numeri di AS (Autonomous System) del BGP (Border Gateway Protocol) che vengono utilizzati per attuare il routing globale in internet: attualmente ci sono poco più di 60.000 numeri AS assegnati di cui circa 34.000 sono annunciati in BGP.

I filtri Stateless Bloom sono utilizzati per l'inoltra dei pacchetti multicast in PSIRP ma ciò funziona bene per un piccolo gruppo di destinatari non adattandosi a gruppi con un numero elevato di destinatari: il passaggio tra il punto d'incontro ed il pacchetto d'inoltra del filtro Bloom è fatto con una funzione TF (Topology Formation).

In DONA il livello di risoluzione si basa su NRS (Name Resolution Service) perciò la scalabilità globale è ottenuta nello stesso modo dei DNS. [16]

5.2 Sicurezza

NetInf fornisce meccanismi per l'autenticazione autore e la verifica dell'origine nonché meccanismi per il controllo di integrità del contenuto.

I meccanismi sono integrati nel servizio di rete: queste funzioni sono inesistenti nella rete di oggi in cui la sicurezza si basa soprattutto sulla fiducia del server fornisce le informazioni.

Lo schema di denominazione NetInf prevede l'autocertificazione dell'integrità dei dati oggetto informazione [16]: autocertificazione significa che nessun sistema di terze parti o PKI debba essere consultato e quindi che l'integrità dei dati può essere consultata off-line.

Questa proprietà è raggiunta fornendo una firma hash del contenuto come metadati per l'oggetto, compreso l'hash di crittografia della chiave pubblica corrispondente alla chiave di firma nel campo autenticatore dell'identificatore stesso: in questo modo l'identificatore di oggetto è legato saldamente ai dati dell'oggetto stesso.

L'autenticazione e l'identificazione del proprietario si ottiene anche con le firme, ma per l'identificazione la fiducia nella chiave deve essere stabilita con mezzi supplementari come, ad esempio, una PKI basata sul certificato.

A causa degli obiettivi di progettazione di sicurezza dello schema dei nomi di NetInf, più precisamente l'autocertificazione, gli identificatori oggetto non sono molto amichevoli per l'uomo dato che sono costituiti più o meno da stringhe di bit casuali.

Questo compromesso significa che sono necessari ulteriori mezzi per legare saldamente i nomi ad un livello di applicazione più a portata d'uomo per gli identificatori NetInf.

Un approccio potrebbe essere fornire nomi firmati interpretabili dall'uomo come metadati per gli oggetti, ciò fornirebbe almeno un legame tra l'oggetto ed il nome interpretabile: i nomi CCN sono gerarchici a partire dal prefisso dell'editore fino al percorso in modo simile agli attuali URL.

Mentre i nomi CCN possono essere leggibili dall'uomo, la fiducia nella chiave firma deve essere sempre stabilita con mezzi esterni, anche per il controllo dell'integrità dei dati, poiché non vi è alcun legame diretto tra il contenuto ed il nome. [16]

PSIRP ha proprietà di sicurezza dirette con punti d'incontro e identificatori di ambito costruiti da hash dei contenuti e proprietà di sicurezza indirette attraverso l'hash della chiave dell'editore più un contrassegno.

PSIRP ha dunque gli stessi problemi coi nomi leggibili dall'uomo di NetInf ed in più viene aggiunto ad ogni pacchetto un autenticatore per combattere gli attacchi DoS. [16]

In DONA i nomi includono l'hash della chiave pubblica dell'editore (il "principale") [16] e ciò significa che se lo stesso contenuto sarà pubblicato da editori diversi avrà più nomi.

Non è però un grosso problema in quanto è possibile eseguire query con wildcard ed è possibile per gli editori delegare ad altre chiavi l'autorizzazione per pubblicare con lo stesso principio.

DONA ha quindi gli stessi problemi con la leggibilità da parte dell'uomo che affliggono NetInf e PSIRP.

5.3 Privacy

Rendendo le informazioni visibili ed identificabili a livello di rete, è facile affermare che i design datacentrici potrebbero minacciare la libertà di parola, o almeno fornire uno strumento più efficace per gli operatori di filtrare, dal loro punto di vista, contenuti o servizi indesiderati.

"Indesiderato" potrebbe significare illegale a causa di problemi di copyright, che l'operatore desideri differenziare il proprio servizio o che un governo voglia sopprimere la libertà di parola.

Poiché il naming NetInf supporta chiavi di firma univoci per oggetto, con conseguente unico autenticatore per i campi identificatore, l'editore può impedire un facile filtraggio di tutti gli oggetti da sé stesso.

La domanda è se gli operatori permetteranno chiavi univoche: forse operatori verranno applicati per ogni chiave utilizzata per pubblicare contenuti, soprattutto se il sistema di routing dovrà essere in grado di aggregare sulla base del campo autenticatore.

Ma, alla fine, c'è una differenza principale rispetto ad oggi? Molto materiale generato dagli utenti viene pubblicato su siti controllati da grandi attori (ad esempio YouTube), questa situazione è destinata a rimanere.

L'identificazione dell'oggetto informazione non aggiungeranno molto alla capacità di filtraggio, dal momento che l'identificatore oggetto molto probabilmente verrà creato utilizzando una chiave del grande attore e non del proprietario.

La conclusione è che il filtraggio sarà circa altrettanto facile o difficile come il filtraggio su indirizzi IP utilizzando l'ispezione dei pacchetti: l'unico modo per controllare chi sta scaricando un qualsiasi contenuto specifico CCN è di controllare tutti i router last-hop, il che può significare ogni singolo router della rete.

5.4 Mobilità

In NetInf la mobilità è ottenuta in modo simile all'IP mobile dato che è presente un punto d'incontro centrale per ogni apparecchiatura utente finale: è possibile una consegna scorrevole del client, anche la mobilità del fornitore di contenuti è possibile, ma non in modo scorrevole. [16]

In PSIRP, il client può semplicemente annullare l'iscrizione, cambiare rete e iscriversi nuovamente: un nuovo percorso o sottostruttura sarà calcolato dal livello di routing, buffering e numerazioni in sequenza permettono consegne scorrevoli. [16]

La mobilità del fornitore di contenuti è più complessa e coinvolge l'aggiornamento dello stato di routing nei nodi punto d'incontro: non si sa ancora esattamente come sarà fatto, ma questo processo avrà probabilmente una lenta convergenza.

In CCN la mobilità è inerente: un client può semplicemente passare ad un'altra rete e continuare a rilasciare i pacchetti d'interesse [16] mentre il layer strategia potrebbe notare lo switch e rilasciare nuovamente tutte le edizioni in sospenso senza attendere il loro time-out.

Un fornitore di contenuti dovrebbe aggiornare le tabelle di routing di tutti i nodi vicini, operazione lenta, ed inoltre molti fornitori di contenuti inquinerebbero le tabelle di routing con prefissi specifici, contrastando i vantaggi dell'aggregazione stessa.

In DONA la mobilità si ottiene in modo simile a PSIRP dato che i client possono annullare la registrazione dalla sede locale e registrarsi nuovamente presso la nuova sede: annullare la registrazione non è obbligatorio in quanto gli RH possono cancellare le voci riguardanti il contenuto che non riescono a trovare. [16]

Capitolo 6

INCENTIVI

6.1 Distribuzione

I differenti approcci datacentrici presentati tendono ad una riprogettazione dell'internet del futuro: bisogna considerare un possibile percorso di migrazione dalle reti di oggi alla nuova tecnologia ICN.

Si considerino i cinque attori principali: [16]

- gli utenti finali, siano essi privati o un'organizzazione che utilizza un servizio di rete
- gli operatori di accesso alla rete che forniscono il servizio agli utenti finali
- gli operatori di connettività di rete che forniscono il servizio ad altri operatori
- i fornitori di servizi o contenuti che hanno oggetti a disposizione per gli utenti (eventualmente anche a pagamento)
- gli sviluppatori di applicazioni che dovrebbero beneficiare di un accesso diretto agli oggetti informazione senza doversi preoccupare dei dettagli di rete

Il problema principale della distribuzione è che uno o più attori devono essere incentivati ad iniziare a distribuire in formato ICN.

Questa funzionalità si riduce a due componenti di rete principali: archiviazione in rete per gestire il caching degli oggetti informazione ed il servizio di risoluzione dei nomi o routing.

6.2 Gli attori

L'attuale tecnologia peer-to-peer implementa un servizio simile a ciò che la tecnologia ICN può fare: esistono reti di distribuzione basate sul peer-to-peer in cui gli utenti forniscono spazio e banda a beneficio degli altri utenti.

È quindi molto probabile che gli utenti finali siano incentivati ad implementare la tecnologia ICN: i personal computer comunemente utilizzati hanno già le capacità e le risorse necessarie per implementare la tecnologia ICN senza bisogno di alcun costoso investimento.

La distribuzione finale però non è sufficiente: se sono solo gli utenti finali ad implementarne la tecnologia, non c'è alcun miglioramento rispetto a quanto le altre modalità peer-to-peer possono fornire.

Per ottenere la piena potenzialità devono essere coinvolti anche i fornitori di rete: unica eccezione è la collaborazione locale in cui è possibile ottenere i principali vantaggi di ICN anche con una distribuzione tra utenti finali.

Gli operatori di accesso alla rete ottengono le loro entrate da parte degli utenti finali che servono e devono pagare gli operatori di rete per la connettività e per il traffico che i loro utenti generano in tutto il Mondo: il costo dipende dalla capacità del collegamento o dal volume di traffico in entrambe le direzioni.

Con questi presupposti l'operatore di accesso alla rete si sente incentivato a ridurre il volume di traffico per l'utente finale verso altri operatori dato che ciò riduce sia gli investimenti necessari per utente che il costo del traffico per utente.

Esiste però un costo dovuto a questo tipo di distribuzione dato che la memoria disponibile per il caching ICN non fa parte del dispositivo di rete che gli operatori di accesso attualmente hanno.

La questione è se questo costo d'investimento sia abbastanza basso per motivare la distribuzione: gli operatori di accesso sono gli attori critici e se non distribuiranno con tecnologia ICN non ci saranno miglioramenti rispetto a quanto faccia la tecnologia peer-to-peer.

Un confronto con la rete di distribuzione Akamai: quest'ultima sta facendo accordi con gli operatori di accesso alla rete per distribuire la loro archiviazione CDN (Content Delivery Network) all'interno della rete dell'operatore di accesso.

Akamai e l'operatore di rete non si pagano l'un l'altro in quanto trovano benefici dal loro accordo: la differenza rispetto al caso ICN è che l'operatore non deve sostenere costi per investimenti in quanto li sostiene direttamente Akamai. [16]

Dato che Akamai chiede un onere ai fornitori di contenuti per il servizio di caching, il quesito interessante riguardo ICN è se l'operatore di accesso debba anche pagare in qualche modo per la cache ICN al fine di motivare il costo di implementazione.

Un altro confronto è con la discussione presente nel gruppo di lavoro IETF (Internet Engineering Task Force) ALTO (Application-Layer Traffic Optimization).

L'obiettivo è di ridurre il traffico interoperatore indotto dagli operatori peer-to-peer fornendo un meccanismo di selezione tra pari che favorisca i peer presenti nella stessa rete dell'operatore.

Il caching ICN ha anche la capacità di localizzare il traffico all'interno della rete dell'operatore e, dato che è provato l'incentivo per la distribuzione ALTO, si deduce che sia incentivante distribuire ICN. [16]

Un discorso analogo si può fare anche per la memorizzazione necessaria per le funzionalità che il gruppo di lavoro IETF DECADE (DECoupled Application Data Enroute) sviluppa: ma questo argomento è più forte dal momento che DECADE ha bisogno di memorizzazione e quindi di investimenti più elevati rispetto ad ALTO. [16]

Gli operatori di connettività di rete ottengono le loro entrate dal traffico generato: normalmente fanno pagare in base alla quantità di traffico ricevuto dagli altri operatori di connettività di rete o in base al traffico totale scambiato con gli operatori di accesso alla rete.

In entrambi i casi gli operatori di connettività di rete sono incentivati ad aumentare la quantità di traffico e quindi se un operatore distribuisce ICN, l'entrata per l'operatore probabilmente diminuirà: con questo ragionamento gli incentivi funzionano contro la distribuzione ICN.

Questo esempio ci dice che ci potrebbe essere la necessità di introdurre nuovi modelli di business in concomitanza dell'introduzione di ICN: con ICN il traffico non viene spinto attraverso la rete ma viene generato in risposta alle richieste.

Sembrerebbe quindi logico ricompensare, e non penalizzare, i fornitori in grado di rispondere a queste richieste.

I fornitori di servizi o contenuti vogliono che i loro prodotti debbano essere sempre disponibili per i propri utenti: a volte pagano gli utenti per questo, a volte ottengono le loro entrate dalla pubblicità.

I grandi fornitori di contenuti oggi devono pagare Akamai, o similari reti di distribuzione, per rendere disponibile in banda larga i loro contenuti perché le risorse interne semplicemente non sono sufficienti.

I fornitori di contenuti stanno quindi indirettamente pagando per gli investimenti in cache CDN all'interno delle reti dell'operatore di accesso.

Ai fornitori certamente non importerebbe se la funzionalità CDN fosse stata costruita nella rete, come lo è per ICN: non avranno bisogno di utilizzare una rete CDN e sarà quindi possibile risparmiarne il costo.

È però chiaramente impraticabile per un fornitore di contenuti avere accordi con tutti gli operatori di accesso per utilizzare le loro cache per la distribuzione.

ICN aumenta quindi potenzialmente il braccio di ferro in corso tra operatori di rete e fornitori di contenuti e servizi laddove i primi vogliono una quota delle entrate degli ultimi al fine di finanziare l'investimento in apparecchiature di rete.

Un modello di business diverso rispetto ad oggi potrebbe essere necessario per risolvere questo problema: ICN potrebbe essere l'occasione per introdurre nuovi modelli di business che risolvano la bagarre tra operatori di rete e fornitori di servizi.

Inoltre per far sì che i fornitori di contenuti e servizi siano soddisfatti di una soluzione caching ICN è importante che ICN sia in grado di fornire un adeguato feedback su quanto frequentemente e da chi è accessibile il contenuto nella cache.

Per gli sviluppatori di applicazioni una API ICN che fornisca accesso diretto agli oggetti informazione necessari per l'applicazione deve essere di grande interesse soprattutto se si offre loro la stessa o una migliore efficienza rispetto alle soluzioni CDN e peer-to-peer di oggi.

Inoltre un'applicazione indipendente API ICN dovrebbe facilitare il riutilizzo su nuove applicazioni di elementi informativi creati da altre applicazioni: ciò dovrebbe portare ad un ecosistema applicazione più vitale.

6.3 Partecipanti ed investitori

NDN è stato fondato da NSF (National Science Foundation) con alla guida Lixia Zhang che dà vita ad un progetto Statunitense molto attivo, recentemente (settembre 2013) entrato nella "next phase" con partecipanti esclusivamente accademici: [18]

- Colorado State University (USA)
- Cornell Tech (USA)
- Northeastern University (USA)
- University of Arizona (USA)
- University of California, Irvine (USA)
- University of California, Los Angeles (USA)
- University of California, San Diego (USA)
- University of Illinois, Urbana-Champaign (USA)
- University of Maryland (USA)
- University of Memphis (USA)
- University of Michigan (USA)
- Vienna University of Technology (USA)
- Washington University (USA)

PSIRP nasce avendo per membri una delle più grandi compagnie di telecomunicazioni Europee (BT) [19] come progetto totalmente Europeo e che dalla sua evoluzione in PURSUIT (ottobre 2010 [9]) comprende 8 partner sia accademici che industriali: [20]

- partner accademici
 - Aalto University (Finlandia)
 - Athens University of Economics and Business (Grecia)
 - Cambridge University (Regno Unito)
 - Centre for Research and Technology Hellas (Grecia)
 - Essex University (Regno Unito)
 - RWTH Aachen (Germania)
- partner industriali
 - CTVC (Regno Unito)
 - Ericsson Research (Finlandia)

NetInf nella sua prima incarnazione 4WARD nasce da una moltitudine di partner accademici ed industriali prevalentemente Europei guidati da Ericsson AB [21], alcuni dei quali formeranno il consorzio che da agosto 2010 è a capo di SAIL: [22]

- partner accademici
 - Aalto University (Finlandia)
 - INRIA (Francia)
 - Institut Télécom (Francia)
 - Israel Institute of Technology (Israele)
 - National ICT Australia (Australia)
 - Royal Institute of Technology (Svezia)
 - Swedish Institute of Computer Science (Svezia)
 - Technical University of Lisboa (Portogallo)
 - Trinity College Dublin (Irlanda)
 - University of Bremen (Germania)
 - University of Cantabria (Spagna)
 - University of Paderborn (Germania)
- partner industriali
 - Alcatel-Lucent (Germania)
 - DOCOMO Communications Laboratories Europe (Germania)
 - Ericsson AB (Svezia)
 - France Télécom (Francia)
 - Fraunhofer AISEC (Germania)
 - Hewlett-Packard (Regno Unito)
 - Lyatiss (Francia)
 - NEC Europe (Regno Unito)
 - Nokia Siemens Networks (Finlandia)
 - Portugal Telecom Inovação (Portogallo)
 - Tecnalía (Spagna)
 - Telecom Italia (Italia)
 - Telefónica Investigación y Desarrollo (Spagna)

DONA infine è uno dei progetti datacentrici Statunitensi più vecchi (2006) ma, anche se è stato nel tempo preso in esame da un colosso del calibro di Cisco Systems [23], ad oggi risulta supportato (poco attivamente) solamente dalle due istituzioni accademiche che lo proposero inizialmente:

- International Computer Science Institute (USA)
- University of California, Berkeley (USA)

Capitolo 7

SFIDE RIMANENTI

7.1 Naming

Gli approcci specifici adottati per le denominazioni possono avere implicazioni: ad esempio mentre NetInf fornisce un sistema di denominazione sicuro, flessibile ed estensibile, i nomi non sono leggibili allo stesso modo degli URL di oggi. [16]

Ci sono un sacco di altre alternative per richiedere gli oggetti in modo fruibile per l'uomo come la selezione di alternative dai menu, icone, risultati di servizi di ricerca: per comunicare direttamente i riferimenti agli oggetti tra gli esseri umani, i nomi leggibili e parlabili sono ancora utili.

Per l'utente ciò significa che un servizio di directory e ricerca o qualsiasi altro tipo di supporto per le applicazioni è necessario per mappare concetti comprensibili dall'uomo agli ID NetInf: la questione critica è come questa mappatura possa essere protetta in modo da evitare gli attacchi.

In generale, una chiara comprensione dello scopo di un sistema di denominazione è importante: se lo scopo primario è un'unica e persistente identificazione oggetto, potrebbe non essere necessario (né auspicabile) utilizzare gli stessi nomi per la costruzione di reti ipertestuali.

D'altra parte aggiungere un ulteriore layer di riferimento indiretto può creare inefficienze di ricerca ed anche ambiguità riguardo l'utilizzo di questi layer per il collegamento tra oggetti.

Nell'internet di oggi è possibile creare un collegamento ad un oggetto con entrambi i nomi DNS o indirizzi IP nel URI e se il nome DNS si risolve allo stesso indirizzo IP, entrambi gli URI potrebbero collegarsi allo stesso oggetto: proprietà simili sarebbero auspicabili anche per un internet NetInf.

Il routing nel CCN viene eseguito sui nomi: i nomi CCN sono gerarchici e ciascun componente è una sequenza arbitraria di byte, così i nomi CCN possono potenzialmente (ma non necessariamente) essere leggibili. [16]

Il più grande svantaggio del routing sui nomi è la frammentazione dei percorsi: se ogni nome può essere indipendentemente instradabile, la tabella di instradamento nei router può diventare molto grande.

Una possibile soluzione a questo problema è di utilizzare i nomi CCN per il routing aggiungendo sopra un layer di denominazione leggibile.

Questo è l'approccio per NDN ma richiede considerazioni analoghe a quelle sopra descritte per i nomi piatti: ciò significa fissare il mapping tra nomi leggibili dall'uomo e nomi CCN.

La denominazione NetInf e CCN comprendono il concetto di versioning: il modo migliore di utilizzare questo concetto ha bisogno ancora di studio.

La denominazione PSIRP è organizzata su più layer, il più alto dei quali è il layer "identificatori dell'applicazione", in cui le applicazioni possono decidere i loro nomi: questo sistema soffre a causa delle stesse preoccupazioni di sicurezza della soluzione NetInf. [16]

I nomi DONA sono piatti, avendo così anch'essi le stesse problematiche dei nomi NetInf ma il routing in DONA può essere completamente basato sui nomi. [16]

7.2 Risoluzione dei nomi

L'architettura NetInf è aperta e così possono essere utilizzati NRS multipli; con gli oggetti informazione di autocertificazione è stato reso possibile l'utilizzo di NetInf senza dover avere instaurato un rapporto di fiducia con i NRS. [16]

Dato che possiamo avere più copie di oggetti informazione e che non ce ne possono essere versioni differenti, non ci può essere un certezza dello stato di un oggetto informazione soprattutto se si considera l'utilizzo di NRS e la possibilità di una rete partizionata.

È stata quindi discussa la possibilità di introdurre un'autorevole casa NRS per ogni oggetto informazione: se è fondamentale ottenere l'ultima versione di un oggetto, non solo la "migliore" copia disponibile, gli NRS dovrebbero essere utilizzati per risolvere l'ID NetInf.

La casa NRS potrebbe essere indicata nei metadati, oppure possono essere determinate ad esempio guardando una parte dell'ID a seconda della struttura dei nomi scelta.

Un altro aspetto che potrebbe richiedere un NRS attendibile o server attendibile è la revoca della chiave a seconda di come si gestiscono le revoche delle chiavi.

I nomi in PSIRP sono registrati al punto d'incontro casa, nel caso di DONA i nomi sono registrati nel NRS DONA: gli RH conoscono l'ubicazione degli oggetti informazione e possono restituire un locatore (non visibile all'applicazione) o l'oggetto stesso. [16]

Altri aspetti che si applicano a NetInf, DONA e PSIRP comprendono la sicurezza e la robustezza di NRS:

- come sono suscettibili agli attacchi DoS?
- che gestione degli NRS è necessaria?
 - manutenzione di associazioni!
 - meccanismi di aggiornamento!
 - raccolta dei rifiuti!

In CCN i nomi non vengono risolti da un servizio di risoluzione di rete, quindi le problematiche appena sopra descritte non gli si applicano. [16]

7.3 Routing, modifiche alla Topologia e Mobilità

Anche se l'idea generale di un internet datacentrico è di andare oltre l'indirizzamento basato sul locatore, ci deve essere una qualche relazione o mappatura della topologia basata sul nome alla topologia di rete organizzativa o fisica: alla fine il contenuto deve essere situato su qualche server di origine fisica o cache.

In una rete con topologia statica, un approccio con schema di denominazione gerarchico come CCN sembra attraente: i nomi possono essere strutturati secondo la topologia organizzativa come `/net/isp/com/example/media/video1.mp4`.

Le fonti possono registrare nomi con la rete, e la rete può aggregare tali nomi per consentire un routing efficiente: per i clienti di telefonia mobile c'è anche un problema in quanto si possono collegare a a diverse reti di accesso e re-inviare pacchetti di interesse come richiesto.

Il problema più grande è la mobilità delle fonti: se si dovesse muovere il nodo responsabile per `media/video1.mp4`, può ri-registrarsi utilizzando il nome completo della risorsa.

Il che potrebbe comunque pregiudicare l'aggregazione e la scalabilità del sistema di routing: l'altra alternativa è la ridenominazione che farebbe per l'appunto perdere la persistenza del nome.

L'aggiunta di un ulteriore layer di indirizzione, cioè un layer di denominazione coi nomi degli oggetti persistenti che vengono mappati a nomi più topologicamente rilevanti, può essere una soluzione: in realtà questa è una delle direzioni proposte per il progetto NDN.

In un approccio basato sulla risoluzione dei nomi, agilità rispetto alla mobilità e modifiche alla topologia dipendono dalla progettazione del sistema di risoluzione dei nomi: sembra difficile però da risolvere la "raggiungibilità senza soluzione di continuità" della risoluzione dei nomi di server o cache che si muovono velocemente.

7.4 Eterogeneità della rete

L'interruzione della rete è una particolare forma di cambiamento di topologia che comporta ulteriori sfide.

Le interruzioni della rete possono rendere la risoluzione dei nomi sia mancanti (ad esempio quando gli elementi di un servizio distribuito di risoluzione sono influenzati dal patizionamento della rete), sia inaffidabili (ad esempio quando i nomi vengono risolti a locatori che non sono più utilizzabili).

La ricerca sui DTN (Delay-Tolerant Networking) ha affrontato problemi simili: un approccio generalmente utile sembra essere una forma "tardiva", in cui i nomi vengono risolti il più tardi possibile. [16]

Relativamente a disagi e ritardi, sembra generalmente sfidante progettare un sistema ICN generale che funzioni bene e che performi bene attraverso diverse tipologie di rete, ad esempio infrastrutture ben collegate e reti di tipo DTN.

Alcuni problemi sono simili all'internet di oggi: non tutti i protocolli di trasporto funzionano bene end-to-end in presenza di alcuni collegamenti di rete invece altre peculiarità, come il problema della risoluzione dei nomi, sono invece specifiche di ICN.

Un possibile approccio è progettare un supporto per l'eterogeneità in un'architettura generale di ICN, cioè il supporto per la risoluzione dei nomi in un dominio specifico, le strategie di inoltro, routing e protocolli di trasporto.

A questo proposito, alcuni elementi dell'architettura DTN (oltre ai concetti di associazione tardiva) potrebbero essere utili, come ad esempio il concetto di layer di convergenza che consentano l'impiego di trasporti specifici in un layer più basso come richiesto.

7.5 L'oggetto informazione

È chiaro che non tutti gli oggetti debbano essere disponibili a tutti: con meccanismi di crittografia e controllo è possibile limitare l'accesso al contenuto di oggetti a chi dovrebbe giustamente accedervi.

Ma a volte la semplice esistenza di un oggetto potrebbe rivelare più informazioni rispetto a ciò che il proprietario dell'oggetto stesso vuole rivelare al mondo circostante: per evitare tutto questo, un NRS per un approccio di tipo NetInf dovrebbe avere un modo per decidere come rispondere ad una richiesta di risoluzione di un certo ID NetInf.

Un modo per raggiungere questo obiettivo è di introdurre un modo per l'editore di un oggetto informazione di aggiungere un ambito per la registrazione dell'oggetto informazione nel NRS.

Il campo di applicazione sarebbe quindi limitare quali utenti / applicazioni / oggetti NetInf otterranno una risposta alla loro richiesta di risoluzione di un determinato oggetto, gli altri semplicemente otterranno la risposta che l'oggetto non esiste.

Un'altra ragione per cui il campo di applicazione degli ID NetInf potrebbe essere una buona idea è la questione della scalabilità del NRS NetInf: sarebbe possibile limitare in che misura un certo ID NetInf si propagerà in tutta la NRS.

Il campo di applicazione in DONA e PSIRP è integrato e quindi non è necessario alcun lavoro supplementare. [16]

Capitolo 8

CONCLUSIONI

8.1 Obiettivo del documento

In questo lavoro è stata presentata una panoramica del campo di ricerca di un internet datacentrico discutendo quali problematiche sta affrontando ICN comprese:

- le questioni di distribuzione dei contenuti
- la scalabilità ed efficienza economica
- la necessità di una denominazione degli oggetti informazione unica e persistente
- come gli oggetti possano essere autenticati senza la necessità di “fidarsi” dell’infrastruttura che li consegna.

Inizialmente sono state analizzate le limitazioni e le barriere dell’internet di oggi e le ragioni sul perché l’approccio datacentrico sarà uno dei principali pilastri del futuro.

Poi sono stati descritti i requisiti di progettazione datacentrici più importanti ed i principi dell’architettura datacentrica del futuro concentrandosi principalmente sul migliorare l’esperienza multimediale degli utenti futuri.

Si ritiene che l’architettura datacentrica del futuro richieda un approccio di design olistico che spazia dalle infrastrutture di rete fino a contenuti e servizi, i quali a loro volta possono influenzare la progettazione delle infrastrutture di rete.

È stata presa la questione della definizione come punto di partenza e sono stati delineati un insieme di blocchi fondanti datacentrici che possano essere utilizzati per progettare un'architettura che soddisfi tutti i requisiti che possano derivare dai problemi presentati.

L'elemento più importante è ovviamente l'oggetto informazione stesso: un sistema di denominazione sicuro per l'oggetto informazione è fondamentale, così come un'applicazione di pubblicazione e sottoscrizione API per accedervi semplice da utilizzare.

Anche la memorizzazione ed il caching dell'oggetto informazione in una rete ICN sono componenti chiave come pure i protocolli necessari per il routing ed il trasporto.

Sono stati descritti i quattro approcci ICN più importanti e sono stati analizzati rispetto a come attuano i blocchi fondanti: sono state anche discusse le sfide generali di scalabilità, sicurezza, privacy e mobilità che tutti le reti ICN affrontano.

Dato che risolvere tutti i problemi tecnici non è sufficiente a garantire il successo di una nuova tecnologia, sono stati discussi anche i problemi di distribuzione: quali sono gli incentivi per distribuire ICN per i diversi attori?

Sono state delineate infine le sfide rimanenti e le sfide di ricerca in sospeso.

8.2 Considerazioni

Per far evolvere l'attuale internet a ciò che è stato definito internet del futuro potrebbero essere necessarie diverse architetture: l'internet del futuro dovrebbe essere composto da più di un'architettura, e quelle descritte ne potrebbero proprio essere la base.

Ognuna delle architetture descritte servirà per migliorare un aspetto o risolvere un problema che nell'internet attuale non è stato ancora risolto:

- l'architettura datacentrica è la più propensa a migliorare gli aspetti negativi che oggi affliggono la gestione ed il consumo dei contenuti audio-visivi
- per ora due reti di tipo diverso non possono comunicare tra loro ma grazie all'architettura framework sarà possibile interoperare attraverso meccanismi che nell'internet di oggi non sono disponibili
- le reti cognitive, a differenza di altri validi concetti, sono indispensabili per la creazione di un internet nuovo ed innovativo

Al fine di realizzare la visione di un internet del futuro pienamente adatto alle esigenze degli utenti futuri devono essere considerati diversi aspetti:

- complessità della struttura di rete contro semplicità dell'ingegneria di progettazione
- scalabilità contro garanzia di qualità e tempi di risposta
- efficienza contro facilità d'uso
- servizi e posizione dei contenuti
- utenti e mobilità della rete
- fiducia e sicurezza sociale

La decisione di seguire un approccio rivoluzionario o di tipo clean-state è ancora fortemente in discussione e molti concetti sono ancora per certi versi astratti, però probabilmente tra pochi anni se progettati correttamente potrebbero essere proprio l'innovazione di un internet che pensandolo ora sarebbe perfetto.

Bibliografia

- [1] FIA e Future Content Networks Group. *Why do we need a Content-Centric Future Internet?* European Commission Information Society and Media, 2009.
- [2] R. W. Thomas, L. A. DaSilva e A. B. MacKenzie. *New Frontiers in Dynamic Spectrum Access Networks*. The Institute of Electrical and Electronics Engineers, 2005.
- [3] <http://networks.cttc.es/mobile-networks/research/cognitive-network-optimization-techniques>. CTTC, 2009-2012.
- [4] D. D. Clark, C. Partridge, J. C. Ramming e J. T. Wroclawski. *A Knowledge Plane for the Internet*. ACM, 2003.
- [5] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs e R. L. Braynard. *Networking Named Content*. ACM, 2009.
- [6] N. Fotiou, G. C. Polyzos e D. Trossen. *Illustrating a Publish-Subscribe Internet Architecture*. European Commission Information Society and Media, 2009.
- [7] <http://www.psirp.org/overview/solution.html>. PSIRP, 2008-2010.
- [8] M. Ain, S. Tarkoma, D. Trossen, P. Nikander, T. Burbidge, A. Zahemszky, J. Rajahalme, D. Lagutin, M. Särelä, J. Riihijärvi e T. Rinta-aho. *Conceptual Architecture of PSIRP Including Subcomponent Descriptions*. European Commission Information Society and Media, 2008.
- [9] <http://www.psirp.org/News.html>. PSIRP, 2010.
- [10] H. Abramowicz, P. A. Aranda, J. Carapinha, C. Foley, M. Johnsson, H. Karl, G. Schultz e M. Soellner. *Introduction and Overview of the 4WARD Technical Results*. European Commission Information Society and Media, 2010.
- [11] <http://www.sail-project.eu/about-sail/>. SAIL, 2010.

- [12] T. Levä, J. Gonçalves e R. J. Ferreira. *Description of project wide scenarios and use cases*. European Commission Information Society and Media, 2011.
- [13] <http://www.sail-project.eu/about-sail/description>. SAIL, 2011.
- [14] Andrey Ermolinskiy e Mohit Chawla. <http://www.crimsonvoyage.com/posters/dona-poster.ppt>. DONA, 2011.
- [15] T. Koponen, A. Ermolinskiy, M. Chawla, K. H. Kim, I. Stoica, B. Chun e S. Shenker. *A Data-Oriented (and Beyond) Network Architecture*. ACM, 2007.
- [16] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher e B. Ohlman. *A Survey of Information-Centric Networking (Draft)*. Schloß Dagstuhl - Leibniz-Zentrum für Informatik, 2011.
- [17] M. Steiner, T. En-Najjary e E. W. Biersack. *Internet Measurement Conference*. ACM, 2007.
- [18] <http://named-data.net/project/participants>. NDN, 2013.
- [19] H. Flinck, D. Trossen e A. Karila. *Implementation Plan based on the Concept Architecture*. European Commission Information Society and Media, 2008.
- [20] A. Karila. *Project Presentation*. PURSUIT, 2010.
- [21] <http://www.4ward-project.eu/index.php?s=partners>. 4WARD, 2008.
- [22] <http://www.sail-project.eu/partners>. SAIL, 2010.
- [23] S. Shenker, M. Chawla, T. Koponen, K. Lakshminarayanan, A. Ramachanran, A. Tavakoli e I. Stoica. *Data-Oriented Network Architecture (DONA)*. Cisco Systems, 2007.