



UCA

PONTIFICIA
UNIVERSIDAD CATÓLICA
ARGENTINA
Santa María de los Buenos Aires

FACULTAD DE INGENIERÍA Y CIENCIAS AGRARIAS

1- PROYECTO

1.1 Título:

“Gestión del Protocolo NETCONF Utilizando Técnicas de Aprendizaje Profundo”

(“NETCONF Protocol Management Using Deep Learning Techniques”)

1.2 Área Temática

Disciplina: Ingeniería en Informática

Especialidad: Protocolos para Comunicaciones – Aprendizaje Profundo

1.3 Área Prioritaria: Configuración y Nivel de Servicio en Internet

1.4 Tipo de Proyecto: Investigación Teórica y Desarrollo de Técnicas de Aprendizaje Profundo

1.5 Lugar de Trabajo: Facultad de Ingeniería y Ciencias Agrarias – UCA y Laboratorios Externos

2- RESPONSABLES

2.1 Director e Investigador 1:

Nombre y Apellido: Ing. Javier A. Ouret

Cargo Docente: Profesor Titular Ordinario

Dedicación: 10 hs semanales

Títulos académicos obtenidos: Ingeniero Civil (UCA) Especializado en Cálculo e Hidráulica Computacional. Especialista en Computación Gráfica y Redes de Alto Rendimiento. Especialización de Posgrado en Lógica Difusa y Bases de datos

relacionales (Stanford University – USA). Gerenciamiento técnico de sistemas para redes de datos metropolitanas y nacionales (Transition Networks – USA / Net2Edge UK).

3. PLAN DE INVESTIGACIÓN

3.1 Resumen

El protocolo NETCONF (Network Configuration Protocol) [11] define los mecanismos para instalar, manipular y eliminar configuraciones de dispositivos de las redes de comunicaciones y datacenters, utilizando llamadas a procedimientos remotos RPC (Remote Procedure Calls), XML (Extensible Markup Language) para la codificación de los datos y los mensajes del protocolo [11] y YANG como lenguaje de modelado de datos (relativos a la configuración, los datos de estado, las llamadas a procedimientos remotos y las notificaciones). Es considerado una mejora sustancial al Simple Network Management Protocol (SNMP) [12] que tiene serias limitaciones para atender los requerimientos actuales debido que fue desarrollado en 1987 para gestionar redes pero no para configurarlas. El protocolo NETCONF está definido en el RFC 4741 y fue publicado como una propuesta de estándar dentro del Internet Task Force (IETF) por el grupo de trabajo de NETCONF. El lenguaje YANG también fue desarrollado por el IETF dentro del Data Modeling Language Working Group (NETMOD) y está especificado en el RFC 7950.

Las redes de los proveedores de servicios múltiples de internet para datos, telefonía y video están compuestas por un sinnúmero de equipos de comunicaciones de distintas marcas y tecnologías. En esas redes convergen diferentes protocolos a nivel físico, de enlace, de red y de transporte. Esos equipos utilizan sistemas operativos embebidos que se configuran por medio de líneas de comando o lenguajes de configuración pseudo-estándar basados en IOS CLI (Command Line Interface), pero también existen otros mucho más evolucionados como Junos basados en XML/Json. Estos lenguajes fueron diseñados para la configuración individual de cada equipo, pero en la actualidad cada ISP, MSO, Carrier o Datacenter posee miles de dispositivos en su red, por lo que la configuración inicial, las actualizaciones o el monitoreo de cada equipo es difícil o imposible de ejecutarse en forma manual e individualmente para cada caso. Además es hoy muy usual la utilización de técnicas proactivas en vez de reactivas para detectar fallas o para introducir parámetros que mejoren la calidad del servicio [14]. En todos estos procesos hay involucrados varios dialectos de CLI, múltiples APIs (Application Programming Interface) y mantenimiento de rutinas (scripts). El protocolo NETCONF permite la automatización de las secuencias de configuración a través múltiples equipos heterogéneos tanto fijos, como móviles, y también en dispositivos de despliegue masivo como IOT, tanto del lado del operador como del lado del usuario. De acuerdo a lo que se observa en la industria, NETCONF será viable de ser utilizado en cada dispositivo disponible en el mercado, desde un ruteador hasta en equipos hogareños de uso diario.

Se busca por consiguiente investigar y proponer prácticas de inteligencia artificial para automatizar total o parcialmente los mecanismos de configuración y monitoreo utilizando NETCONF, YANG y RESTCONF, por medio de un método de aprendizaje profundo para la lectura de procesos existentes en diferentes dialectos de CLI. A los efectos de evaluar los resultados se utilizarán equipos avanzados ya disponibles en el mercado como demarcadores de servicio y ruteadores que soporten NETCONF/YANG, para redes basadas en fibras ópticas y 4G/LTE. Como plataforma de desarrollo e implementación a nivel de hardware se trabajará sobre una placa RBM33G de Mikrotik, módulos LTE Quectel EC25, microcomputadoras Raspberry PI, demarcadores de servicio basados en Microsemi VSC5619EV, SFPs con OTDR, ruteadores y conmutadores (switches) de uso masivo. A nivel de software se utilizarán sistemas operativos abiertos basados en OpenWRT, Python y

herramientas para monitoreo y captura de datos como RFC2544, Wireshark, Ostinato. El proceso de automatización estará basado en técnicas de aprendizaje profundo (deep learning) y lógica difusa (fuzzy logic) para el análisis de comandos y su comportamiento, generando configuraciones de uso. El aprendizaje profundo consiste en una serie de técnicas que permiten a los sistemas informáticos predecir, clasificar, ordenar, tomar decisiones y, en general, extraer conocimientos de los datos sin necesidad de definir explícitamente las reglas para realizar esas tareas. Con el crecimiento de la complejidad de las redes y de la cantidad de datos a analizar necesitamos cambiar de una inteligencia basada en reglas a una basada en datos. Se evaluará la viabilidad de utilizar TensorFlow [19] como herramienta IA de aprendizaje profundo de las configuraciones de CLI para ser automatizadas por NETCONF.

En base a la información obtenida, el sistema planteado podrá tomar decisiones que impacten en el proceso de automatización basado en NETCONF, optimizando el uso de la red, mejorando la calidad de servicio, así como la captura de datos de tráfico y parámetros de operación de los equipos.

3.2 Palabras claves

Machine Learning, Artificial Intelligence, AI, TensorFlow, NETCONF, YANG, SNMP, Fuzzy Logic, Carrier Ethernet, SLA, Service Level Agreement, OAM, Bandwidth Allocation, Junos, CLI,

3.3 Estado actual del conocimiento sobre el tema

NETCONF está estandarizado por medio de los RFC 4741, RFC 4742 (SSH transport v1), RFC 5277 (notifications), RFC 5717 (partial locking), RFC 6022 (monitoring), RFC 6241 (base:1.1), RFC 6242 (SSH transport v2), RFC 6243 (defaults filtering), RFC 6470 (notification events), RFC 6536 (access control), RFC 7230 (HTTP/1.1 Message Syntax and Routing), RFC 7231 (HTTP/1.1 Semantics and Content), RFC 7232 (HTTP/1.1 Conditional Requests), RFC 8040 (RESTCONF) y RFC 8071 (Call Home) [14]. Desde el punto de vista de la programación el RFC 8040 resuelve la necesidad de definir un mecanismo normalizado para que las aplicaciones web puedan acceder a datos de configuración, de estado, operaciones RPC y notificaciones de eventos de un dispositivo de red en forma modular y escalable. Esto se logra por medio del protocolo RESTCONF basado en HTTP para configurar los datos definidos en el modelo YANG usando los conceptos de almacén de datos (datastore) definidos por NETCONF [16]. Existen varias implementaciones de código abierto para desarrollar procesos de automatización basados en NETCONF como Netopeer2 que es un conjunto de aplicaciones que implementan herramientas de configuración de red basadas en el protocolo. Netopeer2 se basa en la nueva generación de las bibliotecas NETCONF y YANG: libyang y libnetconf2. El servidor Netopeer usa sysrepo como una implementación del almacén de datos NETCONF [17]. Otras herramientas disponibles son OpenDaylight, pyang YANG validator, yanglint, symd.py, YANG Explorer, YDK, YDK.py, etc.

Sin embargo debido a que la utilización del protocolo NETCONF es relativamente reciente la aplicación de técnicas de inteligencia artificial en el proceso de automatización todavía es incipiente, lo cual refuerza la importancia de realizar trabajos en esta línea de investigación. Un análisis preliminar nos indica que todavía hay pocos trabajos orientados en esta línea, que relacionen al aprendizaje profundo con las configuraciones basadas en el protocolo NETCONF y el lenguaje YANG. De hecho una búsqueda simple de la frase “*netconf machine learning*” da como resultado en primera página el trabajo del autor desarrollado en la UCA y publicado por el IEEE.

["Quality of Service Assesment Using Machine Learning ...](#)

by JA Ouret - 2018 - [Related articles](#)

Study of an unsupervised **machine learning** approach for the testing results defined by the RFC2544 ... Y1564 standard methodologies and the use of **NETCONF**.

En resumen, la complejidad de las redes actuales a nivel local o regional, con cientos o miles de equipos y velocidades de más 100 Gbps requieren el uso de aprendizaje profundo y clasificación inteligente de configuraciones y parámetros, usando diversos modelos y algoritmos de predicción, regresión lineal, clasificación o clustering, para la captura de configuraciones existentes y el muestreo adecuado que lleve a decisiones correctas. Además el crecimiento en el uso de internet por cientos de millones de usuarios en forma simultánea plantea grandes complejidades en la asociación de flujos de tráfico con las aplicaciones que los generan. La complejidad también se extiende hacia los mecanismos para realizar la configuración dinámica de los parámetros de ajuste, lo que nos lleva a requerir el uso de nuevas herramientas.

3.4 Objetivos e hipótesis de la investigación

- Investigar las mejores prácticas de inteligencia artificial para automatizar total o parcialmente los procesos de configuración y monitoreo utilizando NETCONF y YANG, por medio de técnicas de aprendizaje profundo.
- Desarrollar un prototipo funcional basado en OpenWRT, Netopeer2, NETCONF, YANG y Python, TensorFlow que implemente la utilización de Aprendizaje Profundo y NETCONF para la recolección de comandos de configuración en CLI, Junos o NETCONF mismo y que, en base a la información obtenida, tome decisiones que impacten en los dispositivos de red utilizando el protocolo NETCONF para la re-configuración automática o el monitoreo de los mismos.
- Validar las estructuras de datos sobre YANG.
- En base a lo anterior realizar pruebas de concepto (POC) para la recolección de datos de operación de dispositivos a nivel de usuarios (CPEs) sobre enlaces físicos con fibra óptica y con respaldo sobre 4G/LTE, en el laboratorio descrito anteriormente.
- Utilizar Yang/Netconf para configurar de forma dinámica [2] (Giardina, y otros, 2017) los parámetros de calidad de servicio en los equipamientos de telecomunicaciones, basado en los perfiles generados por el algoritmo genético.
- Escribir y presentar papers y artículos sobre lo anteriormente expuesto.

3.5 Metodología

- El método a seguir toma como base la línea de investigación aplicada a la extracción automática de datos estructurados de la web, web mining y XML. Las herramientas de extracción de información basadas en la inteligencia artificial y las herramientas de desarrollo de envolturas (wrapper) controladas por el usuario se complementan entre si. En teoría las

herramientas automatizadas de aprendizaje profundo generan una gran fracción de reglas de extracción que de otro modo se escribirían manualmente, dejando al usuario con el poder de definir algunas reglas de extracción que no se pueden aprender o que necesitan un ajuste manual [7].

- Utilizando Python, desarrollar un prototipo de extracción de datos por medio de aprendizaje profundo que tome como entradas las secuencias de comandos CLI de los dispositivos, los catalogue y devuelva una secuencia de tareas YANG/NETCONF para configurar los equipos de comunicaciones.
- Utilizar YANG y NETCONF para crear los comandos prototipo para la ejecución de la aplicación Ejecución de pruebas del prototipo. Mediciones y reportes.
- Elaboración de un documento de aplicación. Selección comparativa de las herramientas a utilizar. Comparación de reportes. Conclusiones. Planteo de desarrollos futuros. Presentación de papers.
- El proyecto estará bajo la dirección del Ing. Javier Ouret, con una dedicación de 10 horas semanales (40 hs/mes). La duración es de 2 años, con posibilidad de continuar.
- Participarán en este Proyecto de Investigación, además del Ing. Javier Ouret, 1 alumno de Ingeniería en Informática que presentará su Trabajo Final basándose en la temática de este trabajo de investigación.



3.6 Desarrollo del Trabajo

En las siguientes etapas:

Etapas 1 (Marzo 2020 - Agosto 2020)

- Revisión de antecedentes y trabajos relacionados.
- Revisión e investigación de las normas vigentes aplicables a este proyecto.
- Selección y ensayo de herramientas de software PyPi, PYANG, Netopeer2, etc.
- Preparación de entornos para hacer minería de datos de las configuraciones en CLI.
- Preparación de los entornos virtuales de redes utilizando Onos y Mininet.
- Preparación del laboratorio de equipos de demarcación para NETCONF.
- Preparación del RBM33G de Mikrotik con el sistema operativo OpenWRT incorporando NETCONF. Una vez depurado el software en esta placa de desarrollo se procederá a cargarlo en las microcomputadoras Raspberry PI para contar con múltiples dispositivos exactamente iguales para las pruebas.

Etapas 2 (Septiembre 2020 - Febrero 2021)

- Desarrollo de los pasos a implementar para procesos de automatización sobre NETCONF sin utilizar herramientas de inteligencia artificial.
- Selección y ensayo de las herramientas de inteligencia artificial para aprendizaje profundo.
- Evaluación de TensorFlow como posible herramienta para el aprendizaje profundo de datos de configuración CLI distribuidos.
- Escribir los modelos básicos en YANG con datos representados en XML (Tipos de datos, listas, contenedores. RPCs, dependencias, etc.)

- Validación de las estructuras de datos escritas utilizando pyang validator.
- Configuración de Netoppper2 y servidor/cliente NETCONF
- Ensayo sobre las configuraciones cargadas en las Raspberry Pi.
- Escribir y presentar un paper sobre lo anteriormente expuesto.
- Preparación de un seminario con los resultados parciales obtenidos.
- Implementación final del proceso de automatización basado en técnicas de aprendizaje profundo (deep learning) para el análisis de comandos y su comportamiento, generando patrones de configuración. Para los casos en que se observe discrepancias leves en cuanto al comando más adecuado para el resultado deseado, se realizará la selección por medio de métodos simples que incorporen lógica difusa, ya que trabajamos con información discreta no lineal ni continua.

Etapas 3 (Marzo 2021 – Agosto 2021)

- Preparación de los modelos de datos YANG de la información aprendida de los distintos equipos.
- Análisis de la estructura de la red por medio de NetworkX o herramienta similar para extraer contenedores de datos XML creados con el modelo YANG y validarlos.
- Evaluar los resultados obtenidos de la información de configuración. Se trabajará sobre parámetros de VLAN de servicio y cliente, así como el control de ancho de banda para 3 servicios diferenciados de datos.
- Aplicación de YANG/NETCONF para automatizar la configuración de los equipos involucrados en la red de prueba.
- Lectura de las configuraciones luego de impactarlas en los dispositivos utilizando el protocolo NETCONF.

Etapas 4 (Septiembre 2021 – Febrero 2022)

- Verificar los resultados obtenidos luego de la aplicación del sistema planteado por medio de herramientas de generación y captura de datos (RFC2544, Wireshark).
- Elaboración de conclusiones.
- Redacción del artículo completo sobre el trabajo.
- Escribir y presentar un paper sobre lo anteriormente expuesto.
- Preparación de un seminario con los resultados obtenidos.

3.7 Bibliografía

- [1] Changzhi, D., Zhang, H., Arens, E., & Zhiwei, L. (Marzo de 2017). Machine learning approaches to predict thermal demands using skin temperatures: Steady-state conditions. *Building and Environment*, Volume 114, págs. 1-10.
- [2] Giardina, P., Sambo, N., Dallaglio, M., Bernini, G., Carrozzo, G., Cugini, F., & Castoldi, P. (Marzo de 2017). Configuring monitoring entities through NETCONF and YANG in control and hierarchical management planes. *Optical Fiber Communication Conference 2017*. Los Ángeles, California, Estados Unidos: OSA
- [3] IEEE. (2013). Are we ready for SDN? Implementation challenges for software-defined networks. *IEEE Communications Magazine*.
- [4] Shafiq, M., Yu, X., & Wang, D. (2017). Network Traffic Classification Using Machine Learning Algorithms. *Advances in Intelligent Systems and Interactive Applications* , págs. 621-627.
- [5] Syahputra, R. (2017). Distribution Network Optimization Based on Genetic Algorithm. *Journal of Electrical Technology UMY*.
- [6] Zander, S., & Schmoll, C. (05 de Junio de 2018). Network Measurement and Accounting Meter (NETMATE). Obtenido de <https://github.com/DanielArndt/netmate-flowcalc>
- [7] Jussi Myllymaki, Jared Jackson. Robust Web Data Extraction with XML Path Expressions. IBM Research Division. 2002.
- [8] Jamuna .A, Vinodh Edwards S.E.(2013). Efficient Flow based Network Traffic Classification using Machine Learning. *International Journal of Engineering Research and Applications (IJERA)*
- [9] T.Kalaiselvi1, P.Shanmugaraja.(2016). Internet Traffic Classification Using supervised Learning Algorithms – A Survey. *International Research Journal of Engineering and Technology*.
- [10] RFC 6241. Network Configuration Protocol (NETCONF). (2011). Internet Engineering Task Force (IETF)
- [11] RFC 1157. A Simple Network Management Protocol (SNMP). (1990). Network Working Group.
- [12] Javier Ouret. (2018). Modelización Normalizada de la Calidad de Servicio en Redes de Acceso ME y GPON.
- [13] Javier Ouret. (2019) . Reconfiguración Automática de Parámetros de Calidad de Servicio Dispositivos por medio del Protocolo NETCONF. 48 JAIIO. Salta. Argentina.
- [14] RFC for NETCONF.
Disponible: <https://www.yumaworks.com/tools/yang-compiler/netconf-details>
- [15] RFC 8040. RESTCONF Protocol. Internet Engineering Task Force (IETF). 2017.
- [16] Netopeer2 – The NETCONF Toolset.
Disponible: <https://github.com/CESNET/Netopeer2>
- [17] NetworkX
- [18] Disponible: <https://networkx.github.io/documentation/stable/tutorial.html>
- [19] TensorFlow – Open-source software library for dataflow and differentiable programming across a range of tasks. Disponible: <https://www.tensorflow.org/tutorials/>

4. DESARROLLO DEL PROYECTO

1.1 Cronograma de Actividades

Etapas	Meses – Año 2020											
	1	2	3	4	5	6	7	8	9	10	11	12
1			X	X	X	X	X	X				
2									X	X	X	X
	Meses – Año 2021											
	1	2	3	4	5	6	7	8	9	10	11	12
2	X	X										
3			X	X	X	X	X	X				
4									X	X	X	X
	Meses – Año 2022											
	1	2	3	4	5	6	7	8	9	10	11	12
4	X	X										

- Se presentará un informe de avance bimensual con lo realizado y los resultados obtenidos.
- Se publicarán artículos o papers con los resultados obtenidos en Congresos con referato.
- Se publicará artículo o paper con los resultados obtenidos en una Revista con referato.

4.2 Actividades de Transferencia

Los resultados de este proyecto serán comunicados y transferidos mediante publicaciones, ponencias en congresos, eventos de la especialidad y en actividades de extensión en ámbitos académicos y de la industria.

4.3 Vinculación del proyecto con la actividad docente desarrollada en UCA

Los resultados del proyecto tendrán vinculación directa con materias de la carrera de grado de Ingeniería Informática, tales como Redes de Comunicaciones, Modelos y Simulación, Teoría de Lenguajes, Arquitectura de Computadoras y Sistemas Operativos.

4.4 Vinculación del proyecto con problemas de la Comunidad

El proyecto propone nuevas metodologías que contribuyan a la configuración y monitoreo de servicios en las redes de comunicaciones y centros de datos (Datacenters).

5. PERSONAL ASIGNADO AL PROYECTO

5.1 Completar la tabla de datos para cada uno de los integrantes en el siguiente orden: Director, Codirector, Investigadores e Investigadores en formación.

5.1.1. Por la UCA

Función:	Investigador		
Apellido y Nombre:	Javier A. Ouret		
Tipo y No. Documento:	DNI 13736141		
No. de Legajo en UCA:	85474-4		
Lugar y Fecha de Nacimiento:	Buenos Aires, 19/02/1960		
Nacionalidad:	Argentina		
Domicilio:	Vilela 2046 CABA		
TE Particular/celular:	4702 1435		
E -mail:	ingenieria.informaticaUCA@gmail.com		
Título de Grado:	Ingeniero Civil (1984)		
Máximo Título Obtenido:	Especialista		
Cargo Docente:	Profesor Titular Ordinario		
Si reviste como investigador en otra Institución (Ej.: CONICET, etc.), consignar: NO	Institución	Cargo	Dedicación

6. ALUMNOS COLABORADORES

6.1 Por la UCA

Función:			
Apellido y Nombre:			
Tipo y No. Documento:			
No. de Legajo en UCA:			
Lugar y Fecha de Nacimiento:			
Nacionalidad:			
Domicilio:			
TE Particular/celular:			
E -mail:			
Título de Grado:			
Máximo Título Obtenido:			
Cargo Docente:			
Si reviste como investigador en otra Institución (Ej.: CONICET, etc.), consignar: NO	Institución	Cargo	Dedicación

