國立中興大學

選送碩、博士研究生/博後/教授赴紐約IBM短期研習進修(Intern)機會公告

 本研習機會由科技部駐波士頓科技組與大紐約區美洲中國工程師學會共同合作，提供共計25個自費至紐約IBM短期研習名額，**申請對象以碩、博生、博士後研究員及教授為主**，計畫項次分『**A類：Cloud Computing and Cloud System**』及『**B類：Power Converters**』，詳細研習內容與申請資格請參考附件研習計畫說明。執行期間為今年五月中至八月中，可做彈性調整。

 本校擬支援至多5個名額參與此次的研習計畫，有意願者請於3月13日(星期五)前將申請表格填妥並寫明有興趣參加之計畫項次寄至本校永續能源與奈米科技研究中心nanocenter@nchu.edu.tw。

**選送研究生/博後/教授赴IBM短期研習(Intern)專案補助辦法**

1. 專案提供補助，以**獎學金形式補助，20~30萬/人，限於2020年前歸國完成。**
2. 敬請欲報名之**碩、博士研究生、博士後研究員及教授填寫本赴IBM短期研習進修資料表，並於3/13日以前回傳至能源奈米中心辨公室(nanocenter@nchu.edu.tw)**。
3. 選送人員在歸國後，須繳交研習成果與心得書面報告。

**國立中興大學**

**高等教育深耕計畫─109年度發展國際優勢特色**

**選送碩、博士研究生/博士後研究員/教授赴IBM短期研習進修(Intern)** **進修資料表**

|  |
| --- |
| **2020 IBM Research Intern Project Application From** |
| **中文姓名Chinese Name** |  | **性別Gender** |  |
| **英文姓名English Name** |  | **國籍Nationality** |  |
| **學院College** |  | **身分別Degree** |  |
| **系所/年級****Department/** **Academic Year** |  | **指導教授姓名****Name of Advisor** |  |
| **擬赴IBM研習計畫編號及名稱****Name of Research Intern Project Proposals****(請參考附件研習計畫說明)** | **□ A類：Cloud Computing and Cloud System**Project Name: A ( ) . |
| **□ B類：Power Converters**Project Name: B ( ) . |
| **個人專長及研究主題****Specialty & Research Area****(請描述個人專長及研究)** |  |
| **英文能力證明Certification of English****(請檢附證明文件)** |  |

註：請檢附個人資料表於下頁，格式自訂。

個人資料表(格式自訂)

**2020 IBM Research Intern Project Proposals**

**計畫項次分為A類：cloud computing and cloud system；B類：power devices**

The project names with red-color font are with higher priority looking for people.

==============================================================================

**A類**

**22 internships collected from Dr. I-hsin Chung 鍾一新博士 (****ihchung@us.ibm.com** **)**

As per our discussion, please find the descriptions of intern projects at the bottom of this email.

The project names with red-color font are with higher priority looking for people.

The 22 intern projects listed are only from my department which focuses on the system and cloud computing. There are other IBM Research departments that I can help to bridge later.

The general requirements for the interns are

a. graduate students are preferred - this is due to the visa requirement. The visa for undergraduate students is processed via a different route and may take a longer period of time.

b. reasonable English communication skills - reading, writing, listening and speaking

c. good working/learning attitude

d. can live independently in the US environment

It will be great to help with non-IBM-paid intern positions by supporting their transportation and living expenses.

IBM Research will help with mentors, office and working environment.

In terms of the visiting time:

Mid-May to mid-Aug is the time where most interns are here so students can build the network with other interns.

Mid-Aug there is usually a poster session. Most interns use that as a target to present their intern work.

Other than the above time, any time period of the year is normally fine as long as the mentors are available. The collaboration can start once both parties are committed and can continue after the visit.

**Proposals**

**Hardware**

**A01.**
Project name
Resonant gate control chip for smaller, more efficient DC-DC converters in cloud systems
Project description
Design a resonant gate control chip, designed to replace the secondary gate driver chip in today's Power Block DC-DC converter. The IBM Yorktown power, packaging and cooling team has laboratory hardware providing 98% AC-DC conversion efficiency, a 54 V intermediate bus and 93% DC-DC conversion efficiency, for an overall AC to processor core efficiency of over 91%. We expect to easily improve this to 93% overall AC-to-core efficiency. This is far superior to the power delivery efficiency in any IBM system today.
The main limitation to applying this power delivery technology to future cloud hardware with compact water-cooled and mezzanine-accelerator packaging requirements is the form factor of today's IBM Power Block DC-DC converter. The Power Block has been designed for superior area efficiency when mounting the converter on motherboard in a 1U or taller drawer. In order to physically shrink the height of the Power Block to fit into denser formats, we will have to increase the switching efficiency and shrink the transformer. Maintaining efficiency while increasing frequency will require resonant gate control. Ultimately we will want a space-efficient, fully integrated Power Block control chip. However, the first step will be to create a discrete resonant gate control chip, which is a drop-in replacement for the secondary gate driver chip in today's Power Block.
Skill requirement
Analog chip design. DC-DC converter design.
Degree requirement.
M.S. in EE or a related field.
Project length (time)
3 to 6 months
Expected deliverable/consumer
If 3 months, architecture and schematic design for resonant gate control. If 6 months, tapeout of gate control chip.

**A02.**
project name: "Composable IO-Fabric" - attachment of accelerators and memory into the x86 ecosystem via CXL
project description: (Intern would have a role in the following larger project) The value proposition for composable systems is that compute resources (CPUs, memory, accelerators, I/O...) can be matched to a wide range of workload demands, which has the potential to improve utilization, TCO, efficiency, simplify packaging, as well as decouple upgrade cycles of various hardware components. In order to attach accelerators (or NVMe/DRAM etc.) into the x86 ecosystem, the CXL protocol suite appears to be very promising, having wide support from all major hardware suppliers (Intel esp.) as well as multiple hyperscalers. The goal of this subproject is to do a proof-of-concept implementation of a CXL-based IO-fabric on an FPGA demo platform. Specifically, (1) existing VHDL code (e.g. from Synopsis or IBM) of the CXL-stack has to be implemented on an FPGA. (2) a CXL-based link between host and device will be brought up, and in preparation for (3) benchmarking of selected workloads which is a longer range goal. The upfront plan is to have the summer intern work on task #1.
skill requirement: FPGA programming skills, VHDL or Verilog knowledge, Python
degree req: MS/BS in ECE or CS
project length:  tbd, this is longer than the summer, but intern piece could be shorter
expected deliverable/customer:  CXL based demo / Cloud

**A03.**
Project name:  Network Optimization for specific workloads (e.g. AI)
Project description:  This project is intended to help address the question of the benefit of specialized networks for specific workloads, taking AI/ML training as a first example.  AI/ML training is one of the more likely high volume workloads that could make it worthwhile to deploy with a specialized network or interconnect (E.g. Google TPUs operate in a torus cluster "appliance"), while general Cloud PoDs mostly use folded Clos networks and some leading HPC systems have used Dragonfly or Fat Tree networks.  The intern would assist in implementing the AI models into the Venus tool and evaluating different network solutions. As examples, we could model the Gen2/3/4 IBM Cloud networks, as well as AI-based clusters (e.g. working with the Sentient team and/or David Kung's ML team) or optimized networks for specialized target workloads.
Skill requirement:  Network modeling (e.g. OMNET++),  knowledge of ML and HW architectures for AI training,  strong programming skills
Degree requirement:  PhD candidate
Project length (time):  Would be part of a larger project, might be able to contain Intern work to just summer.  TBD
Expected deliverable/consumer:  Performance comparisons for different AI workloads on various networks, learning that could apply to other specialized workloads / "3-5 year" cloud strategy - Cloud team

**A04.**
Project name: Hybrid cloud simulation and modeling framework
Project description: develop a simulation and/or modeling framework for the hybrid cloud, allowing for architects and users to estimate the advantages of adding a new feature or increasing existing capabilities (e.g. increasing network bandwidth).
Skill requirement: computer and network architecture, proficiency on C/C++ programming. Familiarity with simulation (e.g. event driven simulators) and modeling.
Degree requirement: computer science or electrical engineering (undergrad student acceptable, grad student would be preferred)
Project length (time): 6 months ???
Expected deliverable/consumer: model proposed architecture for next generation cloud, identify pros and cons of the new features proposed, for instance what would be the impact of increasing the network bandwidth, or adding a local SSD to cache the file system. The consumer for this deliverable are the cloud architects. Depending on the skill sets from the intern, we would choose a subset of the modeling and simulation work that is appropriate.

**A05.**

Project name: Linux kernel driver for secure hypervisor channel

Project description:
Develop a Linux kernel driver for a new PCIe hardware device that provides secure communication between hypervisor and management nodes.

Skill requirement:
Computer and network architecture, proficiency on C/C++ programming. Kernel and PCIe experience preferred.
Degree requirement:
Computer science or electrical engineering (undergrad student acceptable, grad student would be preferred)
Project length (time): 3 months

Expected deliverable/consumer:
Develop driver, bring up driver and and test for different use scenarios. If time permits, attempt to test channel for vulnerabilities and assess performance for multiple use cases.

**AI Operations**

**A06.**
Project name
Using recurrent neural network to detect anomalies in the cloud environment
Project description
Anomaly detection is very important in the cloud AIOps for forecasting potential problems and delve into the root cause. The current DeCorus system uses statistical methods (univariate and multivariate analyses). The team would like to come back to evaluate deep learning approaches. This team have applied various auto encoders extensively, and notably in the MuMMI project to find outlier states in the Ras protein simulation on the cell membrane, and discovered several biologically important outliers. The approach used there so far does not take the time dimension into consideration. We expect the intern to utilize recurrent neural networks to capture the dynamic time sequence of signals. Another new aspect is the dealing with many signals.
Skill requirement
Mandatory: Pytorch or TF
Degree requirement
PhD Candidate or MS
Project length (time)
3 months
Expected deliverable/consumer
An implementation that may supplement current anomaly detection. AIOPs

**Software**

**A07.**
Project name
Performance evaluation and assessment of KubeVirt on OpenShift
Project description
The scope of the project is to run KubeVirt on Kubernetes/OpenShift and evaluate
performance mainly on two aspects: 1) Kubernetes' scheduler scalability and 2) VM provisioning time.
After a first assessment, further investigations would include performance analysis in VM-to-VM networking as well
as storage I/O, in order to identify potential bottlenecks beforehand.
Skill requirement
Mandatory: Kubernetes, VMs, Linux
Optional: Golang
Degree requirement
PhD Candidate
Project length (time)
3/4 months
Expected deliverable/consumer
A report on
performance of OpenShift and KubeVirt in VMs provisioning and scheduling
I/O performance evaluation for VMs deployed with KubeVirt
investigate state-of-the-art for VM fast/dynamic provisioning techniques

**A08.**
Project name: TrustPlatformModule/Keylime in NextGen cloud computing
Project description: George describes multiple projects in his challenge. An intern can help with setup/experimentation, device driver, analysis, etc.
Skill requirement: Embedded systems, Linux
Degree requirement: None
Project length (time): 3 months
Expected deliverable/consumer: Depends on the specific challenge workstream

**A09.**
Project name: NextGen cloud computing workloads
Project description: We spent time containerizing various worklaods to run on ICP (MuMMI components, an F1 workload, Spark/GATK4, CORAL benchmarks, etc). We should migrate some of these to NG and run them regularly to benchmark performance.
Skill requirement: Some knowledge of cloud and CI/CD environments
Degree requirement: None
Project length (time): 3 months
Expected deliverable/consumer: Workloads that run on NG. We will be the initial consumers.

**A10.**
Project name
Secure End-to-end Connection for Cloud (Broadcom Stingrays)
Project description
Enable end-to-end encryption using crypto engines on Broadcom Stingray smartNIC for IBM Cloud and performance analysis
Skill requirement
Experienced B.S. or M.S. (Working experience of network products would be a big plus)
Degree requirement
Ph.D. Candidate
Project length (time)
3 Months
Expected deliverable/consumer
Development report and performance analysis

**A11.**
Project name
Intrusion Detection and Deep packet Inspection
Project description
Enable Intrusion Detection and deep packet inspection for high speed network (100GB) in Cloud environment
Skill requirement
Network background, Cloud research, Intrusion detection background (e.g., Snort, Bro)
Degree requirement
Ph.D. Candidate
Project length (time)
3 Months
Expected deliverable/consumer
Development report and performance analysis

**A12.**
Project name
Reliable and secure End-to-end Connection for Cloud (using ARM smartNIC and TrustZone)
Project description
Enable secure connection between IBM Cloud hosts even if network boundary is broken
Skill requirement
Network background, SmartNIC experience, Cloud experience
Degree requirement
Experienced B.S. or M.S. (Working experience with network products would be a big plus)
Project length (time)
3 Months
Expected deliverable/consumer
Development report and performance analysis

**Linux**

**A13.**
Project name
Next Gen IaaS Cloud deployment on ContainerLinux/CoreOS (aka Demonstrate the deployment of IBM Next Gen IaaS Cloud framework on top of a minimalistic, compact Operating System)
Project description
This project targets to demonstrate the viability of the deployment of the Next Gen IaaS Cloud core framework (codenamed "GenCtl") on a minimalistic, small attack-surface Operating System. For this Proof of Concept, we selected Container Linux (formerly known as CoreOS):  a small, compact linux distribution specifically targeted for deployment of container based runtimes.  The project will investigate the migration of GenCtl from its current Ubuntu-based Operating System base towards ContainerLinux, identifying all the relevant required changes that need to be made in the CI/CD process. This will include the deployment of both the Master, Control and Compute Nodes atop of the new OS.
In addition to this main task, we will examine reduced linux kernel configurations, which - without compromising the minimal required functionality - can enhance security by reducing the attack surface further. This security evaluation shall be based on measurable quantities that are to be defined during the course of this project, also leveraging known scientific results (e.g. popular paths). The milestones to achieve these goals are
(i) identify the changes required to allow the deployment of base services on ContainerLinux (e.g., specific kernels, custom network drivers, IBM proprietary SDN).
(ii) identify the changes required on core components of GenCtl.
(iii) demonstrate a functional prototype of Next Gen IaaS cloud framework on Container Linux.
(iv) identify opportunities for minimization of attack surface on the kernel.
(v) identify opportunities for additional hardening of host os through configuration protection.
Skill requirement
Mandatory: Operating Systems, System Management
Advanced: Hypervisors, Virtualization Technology
Optional: Software Security
Degree requirement
Undergrad with mandatory skills
PhD with Mandatory and Advanced skills
Project length (time)
4 months
Expected deliverable/consumer
Demonstrate a prototype of IBM Next Gen IaaS cloud framework deployed on ContainerLinux/CoreOS.
This is something requested by IBM Cloud infrastructure partners.

**A14.**
Project name
Next Gen Cloud Hypervisor and Virtual Machine Manager based on RustVMM Technology
Project description
This project main target is the identification of an existing Rust-based Virtual Machine Monitor (VMM) that is functionally equivalent to a production-level QEMU system, but with improved security and performance features. Several Cloud companies are now developing or deploying Cloud compute instances that feature hypervisors based on Rust language technology. The scientific foundation of this is that Rust is a “safer” language than C/C++, which are instead the de-facto standard used in the vast majority of production level system software. We do not expect functional equivalence between QEMU and the Rust-based VMM to come out of the box: development work will be performed as part of this project to achieve it. One aim of this project is to draw a comparison of both security and functionality between QEMU and  the chosen Rust-based virtual machine monitor. Security evaluation will be based on measurable quantities that are to be defined during the course of this project, considering known scientific results (e.g. popular paths). The milestones to achieve these goals are: (i) identification and evaluation of Rust-based VMM, with comparison in functionality to thin hypervisor based on QEMU; (ii) extension of such Rust VMM to fill-the-gap of some functional differences (e.g. device support); (iii) identification of security metrics for VMM comparison; (iv) full functional and security comparative analysis of improved Rust VMM and QEMU.
Skill requirement
Mandatory: Operating Systems, Programming Languages,
Advanced: Hypervisors, Virtualization Technology
Optional: Software Security
Degree requirement
Undergrad with mandatory skills
PhD with Mandatory and Advanced skills
Project length (time)
4 months
Expected deliverable/consumer
Rust-based VMM software and associated scientific paper.
This is something requested by IBM Cloud infrastructure partners.

**A15.**

Project Name:
Automated Environment for Determining Correlations between Code Coverage and Security Exploits.

Project description:
The project is intended to provide and environment in which we can continuously evaluate Linux Kernel and subsystem for coverage (or lack there off) based on configurations.

Skill Requirements:
Linux Kernel skills, Security skills, cloud deployement (desirable)

Degree requirements:
Master or undergraduate with experience in listed skill set

Project length (time):
3-5 month Summer intern project. Can adjust scope to available time

Expected deliverable/consumer:
(1) Functional environment for automated configuration exploration for Linux Kernel and VMM subsystem.
(2) Integration of said system with IBM NextGen cloud use-cases.

**Experiment and Integration**

**A16.**
C. Evangelinos (Cambridge based)
Project name
On-demand scalable distributed and parallel filesystem investigation for Cloud VM and Container usage
Project description
Experiment with different options for on-demand scalable and distributed filesystems for the Cloud building on work done (at the bare metal, on-premise level) during 2019 IAP MIT externship. Idea is to use prototype local storage (NVMe or SSD) and RAMdisks to instantiate a common namespace with good performance for distributed applications and workflows with significant I/O needs (bandwidth and/or IOPS). Experimentation options include Lustre, GPFS, BeeGFS, OrangeFS and others. This work is to be done at the VM and at the Container level - ideally given the breadth we could have 1 intern for each type or prioritize work for either VM or Containers and pick the rest of the work up with a follow-on internship.
Skill requirement
Mandatory: Linux at the root/sysadmin level, C and shell script coding
Advanced: Experience with Parallel Filesystems, Experience with performance analysis tools
Optional: Parallel computing skills, Linux kernel module experience
Degree requirement
Undergrad with Mandatory skills
PhD student with Mandatory and Advanced skills
Project length (time)
3 months
Expected deliverable/consumer
A technical presentation and a more detailed report comparing and contrasting the options in terms of usability, stability and performance. An initial set of scripts allowing the automation of the process (not production level). A conference paper or better.

**A17.**
Project name
Tuned communication libraries for IBM Cloud (VMs and Containers)
Project description
With SR-IOV enabled Cloud software coming in 2H2020 we can proceed in evaluating the performance (and tuning the parameters of - in a virtuous cycle) of different communication libraries used in technical computing and ML: MPI implementations (MVAPICH, Intel MPI, Spectrum MPI (ppc64le and x86\_64), OpenMPI, MLNX MPI), various OpenSHMEM implementations, GASNet, NCCL etc. This way we can offer our customers an optimized recipe for utilizing our hardware and be able to compete against AWS and Azure that are already offering that. The impact is going to be felt for both distributed learning and more traditional technical computing and analytics fields that require efficient and scalable communication. Comparison to performance on AWS and Azure would be helpful in setting a baseline.
Skill requirement
Mandatory: Linux at the root/sysadmin level, C and shell script coding
Advanced: Experience with MPI (at least), Experience with performance analysis tools, C++
Optional: Linux kernel module experience
Degree requirement
Undergrad with Mandatory skills
PhD student with Mandatory and Advanced skills
Project length (time)
3 months
Expected deliverable/consumer
A technical presentation and a more detailed report comparing and contrasting evaluated software in terms of usability, stability and performance. Recipes for tuned building/setup of the libraries. Possible code changes documented and submitted upstream. A conference paper or better.

**A18.**
Project name
FastBoot for Remote Virtual Machine Images
Project description
Currently the image of virtual machines are stored in remote storage devices. Before the VM can boot up, the image must be completely loaded to the local memory or disk. This project investigates the strategies to diminish the delay of booting up the virtual machine.
Skill requirement
This project may need to modify QEMU and/or Linux Kernel Virtual Machine (KVM) to implement new VM image management strategies and conduct the performance evaluation. C/C++ Programming is a must. Better to have experiences on QEMU/KVM.
Degree requirement
M.S. or Ph.D. student (preferred)
Project length (time)
3 months or longer, depending on the whether IBM Cloud may want it or not.
Expected deliverable/consumer
A technical report and a prototype for the proposed strategy. The prototype may be integrated into NextGen. The outcome could also become a conference paper and/or be patented as an IBM property mechanism, or it may be submitted to the upstream.

**A19.**
Project name
Encrypted Real Partition Support for Virtual Machines
Project description
QEMU can support mounting a real partition to the VM. However, the acceleration of data encryption for this scenario was not investigated yet. (IF NOT TRUE, DISCARD THIS PROPOSAL.) This project investigates the possible way to enforce the fast encryption for the real partition on NVMe devices.
Skill requirement
This project may need to modify QEMU and/or Linux Kernel Virtual Machine (KVM) to implement the encryption mechanism and conduct the performance evaluation. C/C++ Programming is a must. Better to have experiences on QEMU/KVM. Experiences on HW acceleration of cryptographic algorithm is optional.
Degree requirement
M.S. or Ph.D. student (preferred)
Project length (time)
3 months or longer, depending on whether IBM Cloud may want it or not.
Expected deliverable/consumer
A technical report and a prototype for the proposed strategy. The prototype may be integrated into NextGen. The outcome could also become a conference paper and/or be patented as an IBM property mechanism.

**A20.**
Project name
Accelerating the Access of Remote Large Dataset for AI Training
Project description
Since the dataset for AI training could be very large and mostly stored in the cloud object store, the time cost to retrieve the data from the cloud object store may become the performance bottleneck. This project is to investigate the possible protocol design to diminish the time to retrieve the dataset from the remote data store.
Skill requirement
Understanding network architecture and protocol. Golang programming or C/C++ programming.
Degree requirement
M.S. or Ph.D. student (preferred)
Project length (time)
3 months or longer, depending on whether IBM Cloud/HPC team may want it or not.
Expected deliverable/consumer
A technical report and a prototype for the proposed strategy. The outcome could also be patented as an IBM property mechanism.

**A21.**
Project name: Analysis and optimization of etcd software performance bottlenecks
Project description: Analyze software bottlenecks in etcd and determine the root cause of the lack of performance improvement with Optane DC memory. Restructure etcd to remove these bottlenecks and improve the performance on Optane DC.
Skills requirements: Go language, database programming, software performance analysis debugging, persistent memory programming/PMDK
Degree requirement: 2nd year PhD or more
Project length (time): 3 months
Expected deliverable/consumer: Improve scaling of Kubernetes through etcd, release fixes back to etcd project maintainers

**A22.**
Project name: Jitter/variability/scalability study on Clouds
Project description: Research and study the Jitter/variability/scalability on both IBM and competitors' cloud offerings.
Skills requirements: CS/EE major (knowledge of OS, architecture, cloud computing), performance analysis
degree requirement: graduate student
project length: 3+ months
Expected deliverable/consumer: competitive analysis, feedback to the IBM cloud team

==============================================================================

**B類**

**3 internships collected from Dr. Ko-Tao Lee李克濤博士(** **leeko@us.ibm.com****)**

**B01.**

**Topic:** GaN solutions for 48 V system infrastructures

**Organization:** IBM T. J. Watson Research Center

**Contact:** Ko-Tao Lee

**Email:** leeko@us.ibm.com
**Details:** Phase 1 - International collaboration on GaN device modeling and fabrication

 Phase 2 - On-site characterization and system integration

**B02.**

**Topic:** Resonant gate driver for high power density power converters for cloud systems

**Organization:** IBM T. J. Watson Research Center

**Contact:** Xin Zhang/Ko-Tao Lee

**Email:** leeko@us.ibm.com
**Details:** Phase 1 – analysis and modeling of high power density power converters.

 Phase 2 – design and simulation of resonant gate driver for power converters with optimum efficiency

**B03.**

**Topic:** Reinforcement Learning with Active Learning for Efficient Electrical Power Converter Design

**Organization:** IBM T. J. Watson Research Center

**Contact:** Xin Zhang/Ko-Tao Lee

**Email:** leeko@us.ibm.com
**Details:** Phase 1 – formulate the design problem as Mixed Integer Optimization with surrogate models to be intelligently and automatically solved using an RL-based optimizer.

 Phase 2 - train RL-informed surrogate models with the infrequent physics-based models via intelligent strategies such as active learning

==============================================================================