TCG enhancements on PowerPC

Nikunj A. Dadhania
nikunj@linux.vnet.ibm.com

Linux Technology Center, India, IBM

KVM Forum

25th August 2016

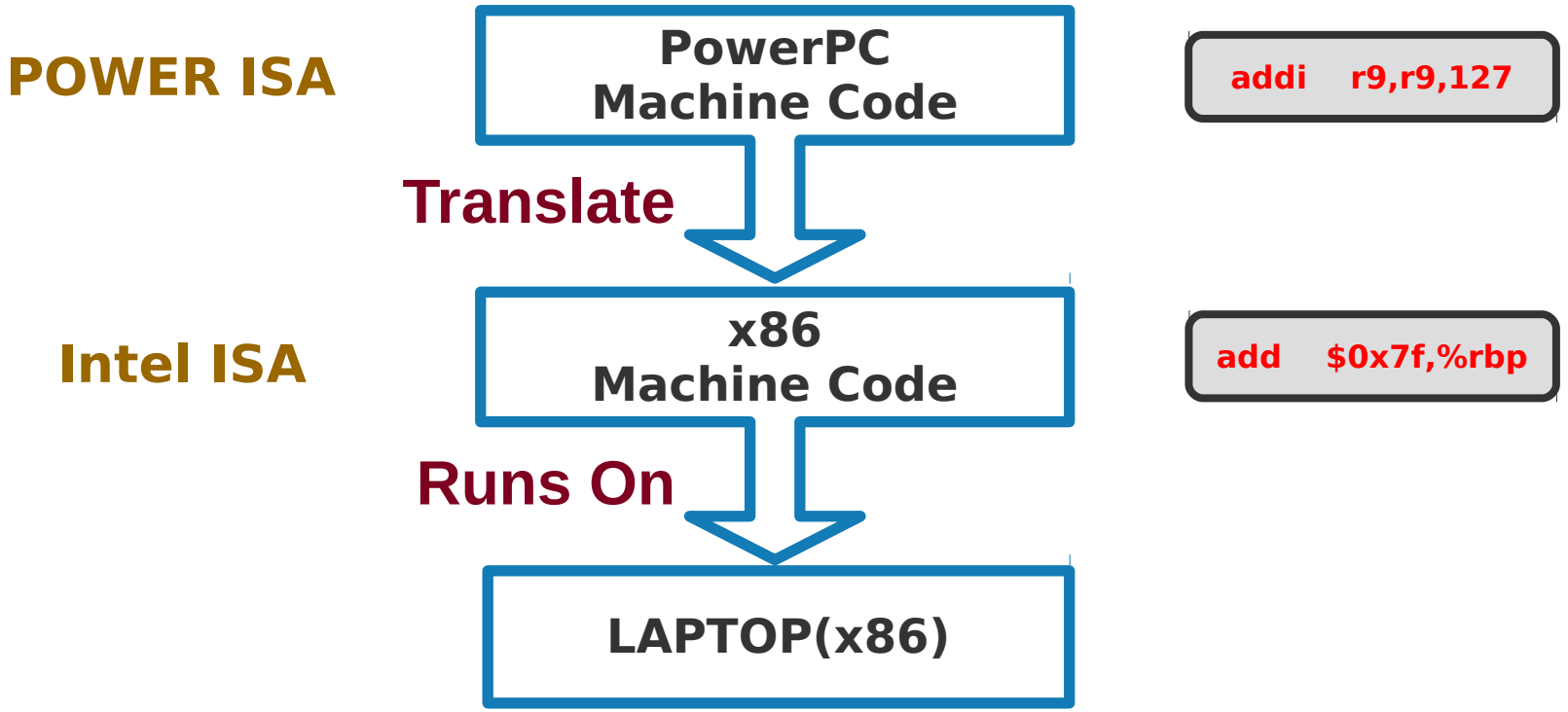
About me

- Guest firmware(SLOF) developer
- QEMU user/developer

Agenda

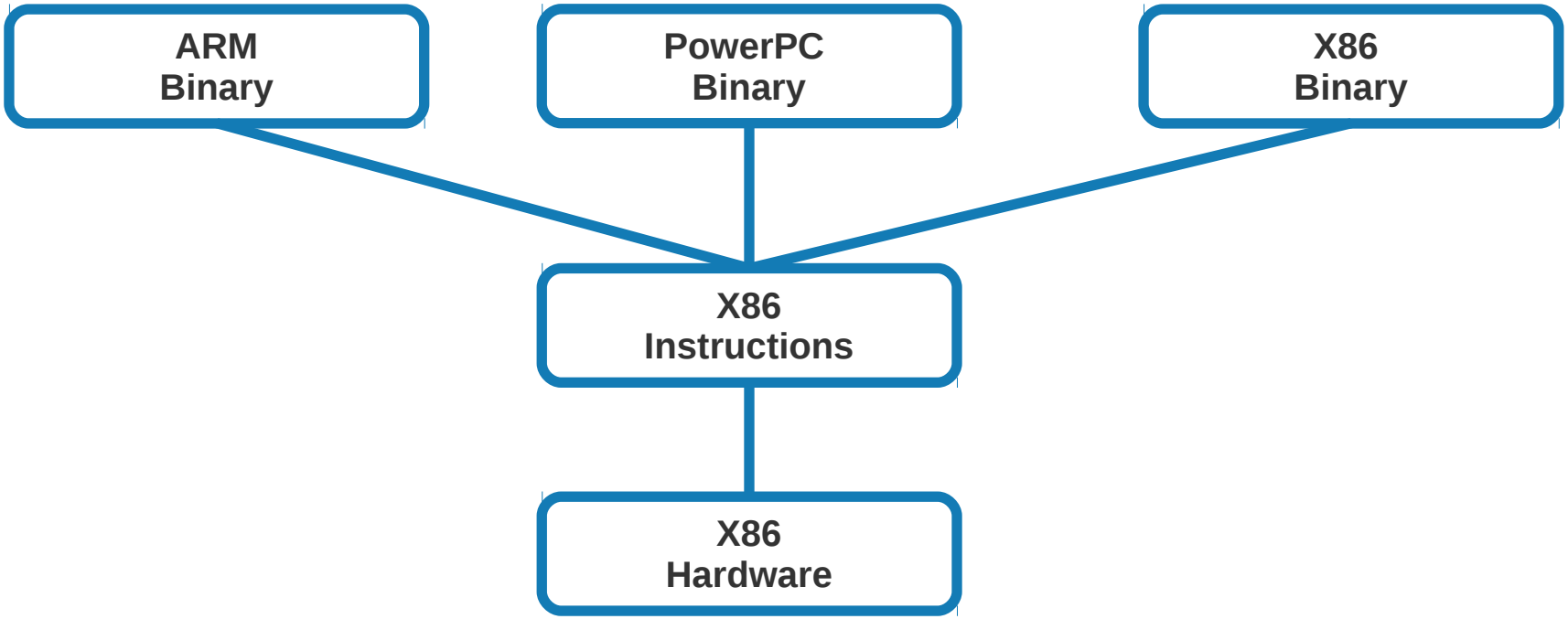
- QEMU TCG – Quick look
- Power ISA 3.0 Support
- PowerNV Platform
- PowerPC support for Multi-threaded TCG
- Other Optimizations
- Future work

How is emulation done ?



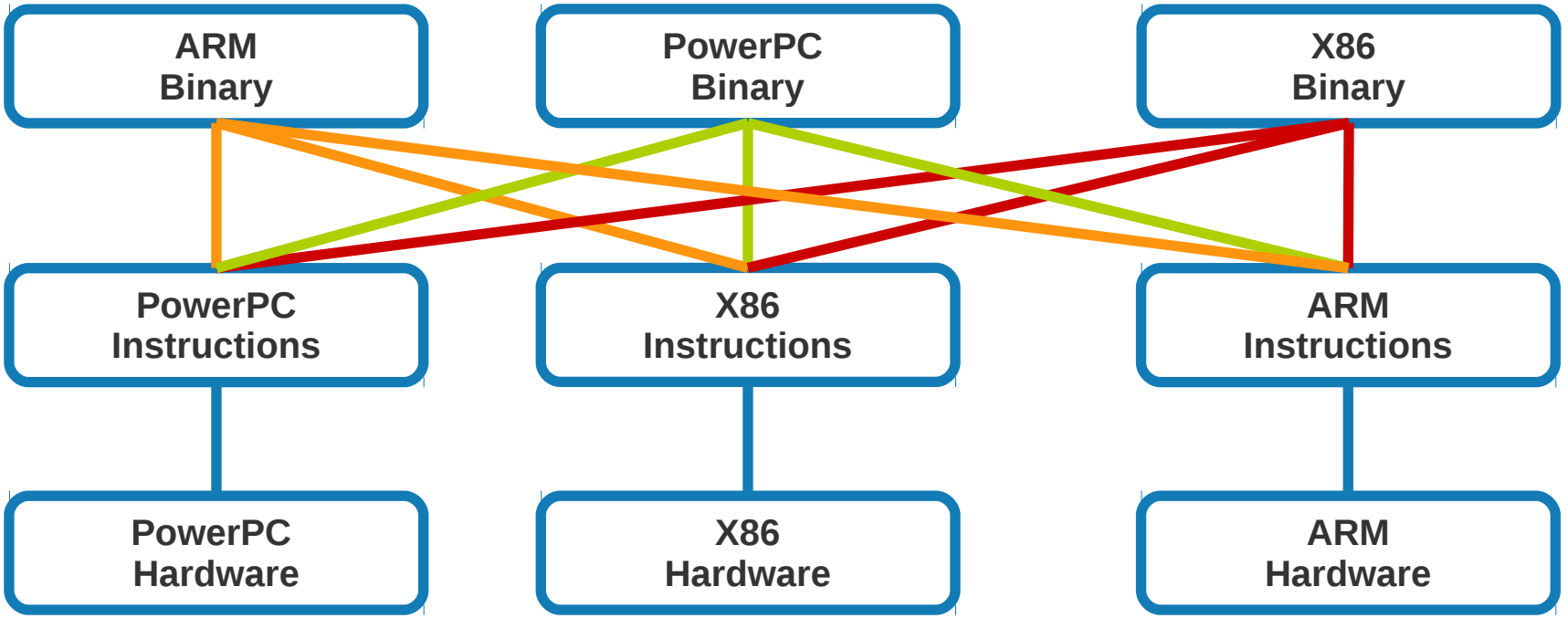
Credits: Alexander Graf's
QEMU's Recompilation Engine

More architectures !



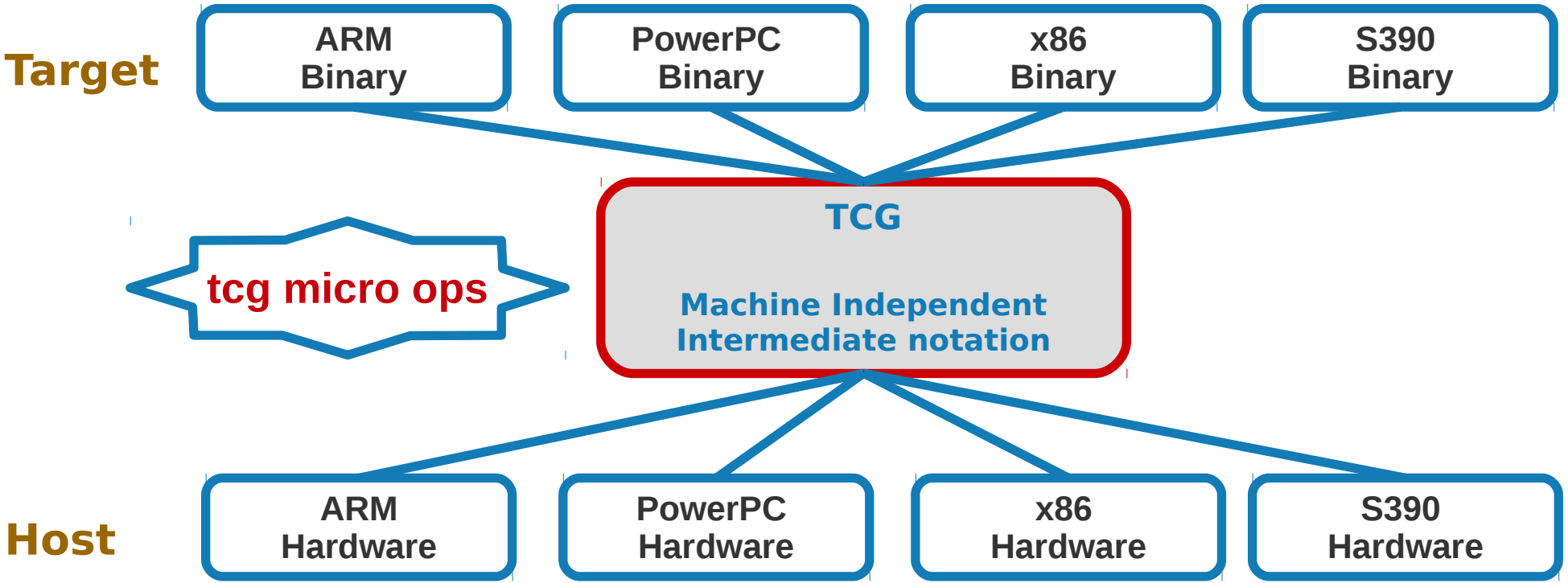
Credits: Alexander Graf's
QEMU's Recompilation Engine

N x N Support: Very complex



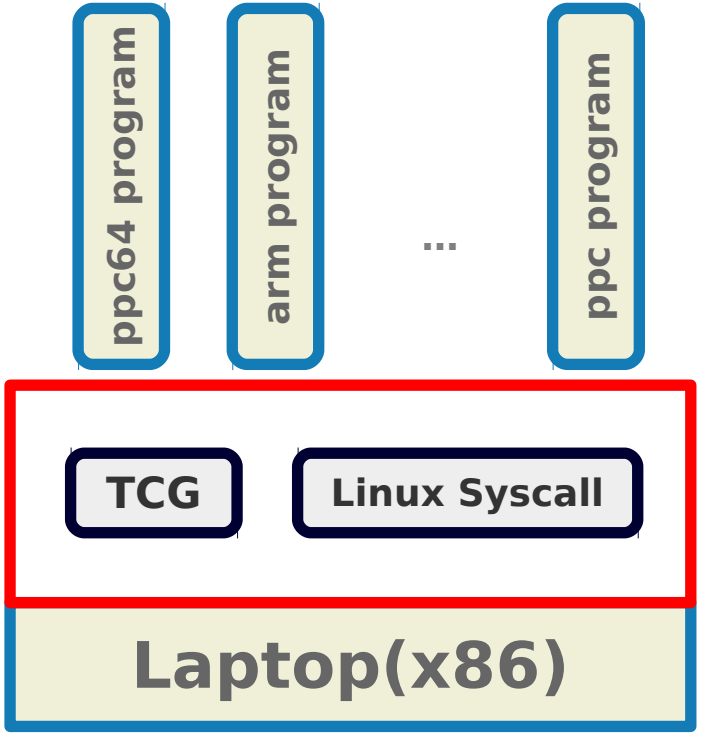
Credits: Alexander Graf's
QEMU's Recompilation Engine

QEMU TCG – Tiny Code Generator



Credits: Alexander Graf's QEMU's Recompilation Engine

QEMU Avatar – linux-user



- Input: Target binary and libraries
- Provides Linux system call emulation
- Emulates target ISA
- Can be used to debug user programs

POWER ISA 3.0

POWER ISA 3.0

- POWER (Performance Optimization With Enhanced RISC)
- Adds ~180 new instructions
- Various instructions added in different classes
  - Atomic memory operations
  - Hashing support operations
  - String operations (character testing, string processing)
  - Arithmetic operations(multiply-add, modulo)
  - ……
- **http://ibm.biz/power-isa3** (needs registration)

Status - POWER ISA 3.0

- 24 instructions queued in ppc-for-2.8
  - Modulo, Special compare
  - Vector absolute, compare, shift
- 24 instructions posted under review
  - Load/Store vector/scalar
  - Vector insert, extract, count trailing zeros
- 25 instructions under test

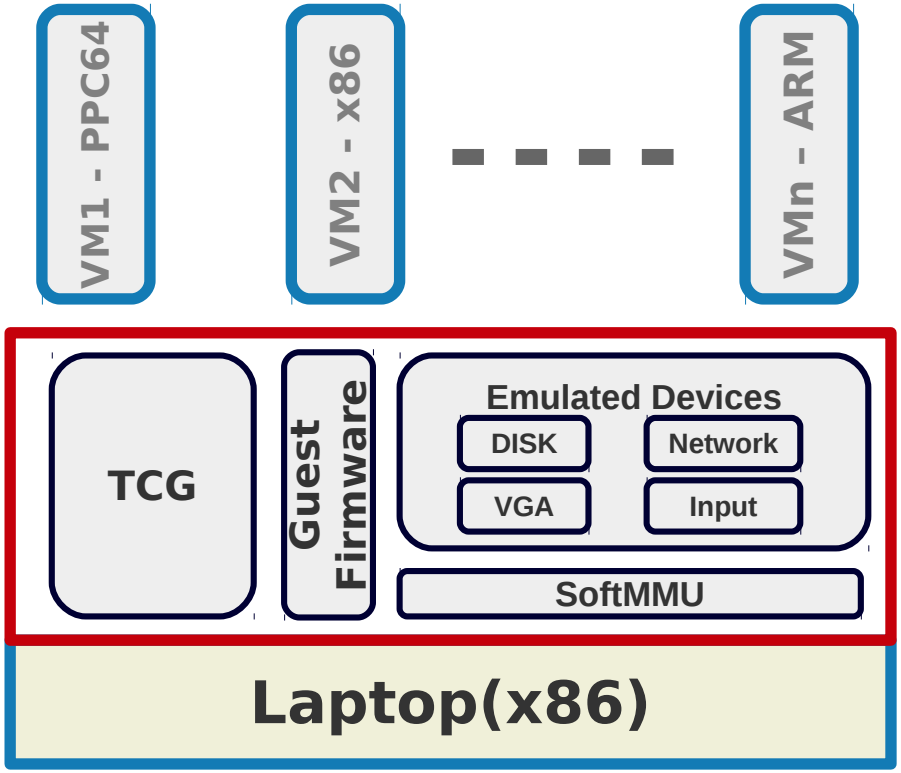
- `https://github.com/nikunjad/qemu/commits/p9-tcg`

Challenges: POWER ISA 3.0

- Testing and verifying the instructions
  - Correctness
  - Repeatability
  - Negative test cases
- Can use:
  - kvm-unit-test
  - QEMU qtest
- Anton Blanchard's instruction fuzzer
  - Compares physical CPU to QEMU emulation

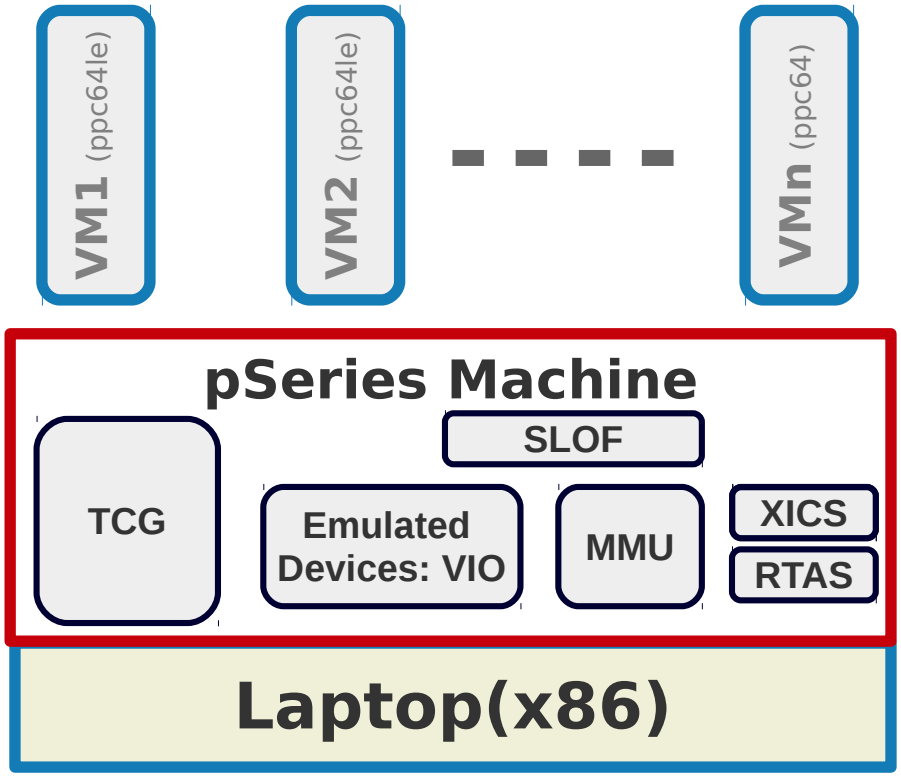
PowerNV Platform

QEMU Avatar – System Emulation



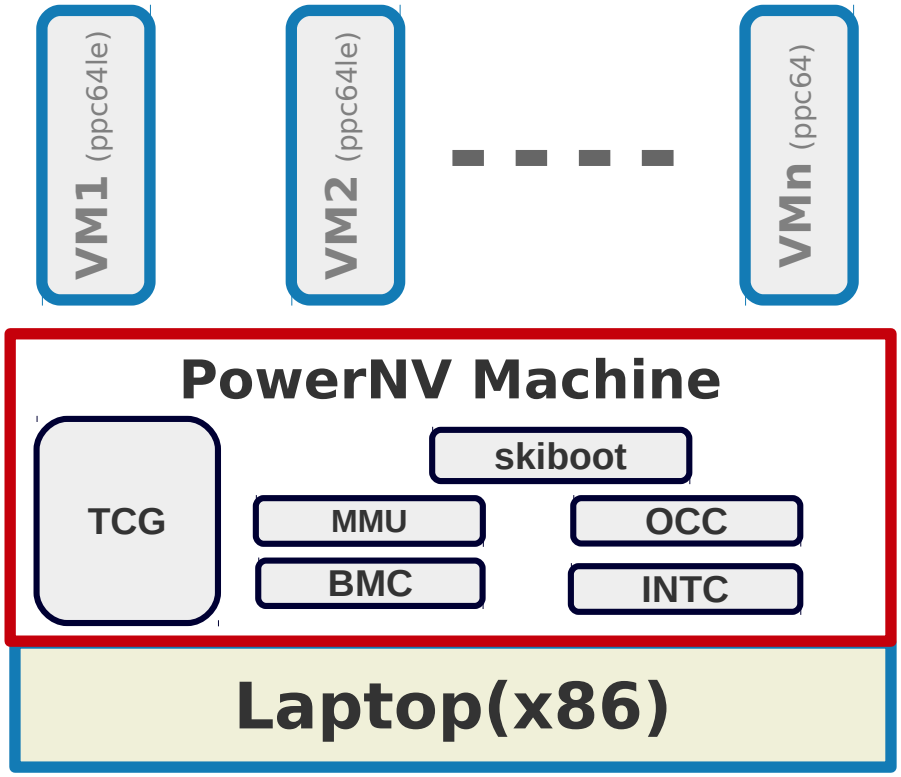
- Invoked as machines (-machine pseries)
- Runs isolated in its own memory space
- Can be used to debug firmware, kernel, etc.

pSeries Machine Emulation



- Based on sPAPR standard
- Guest Emulation
- Hyper-Call based
- Para-virtualized guest
- Has been supported since a while

PowerNV Machine Emulation



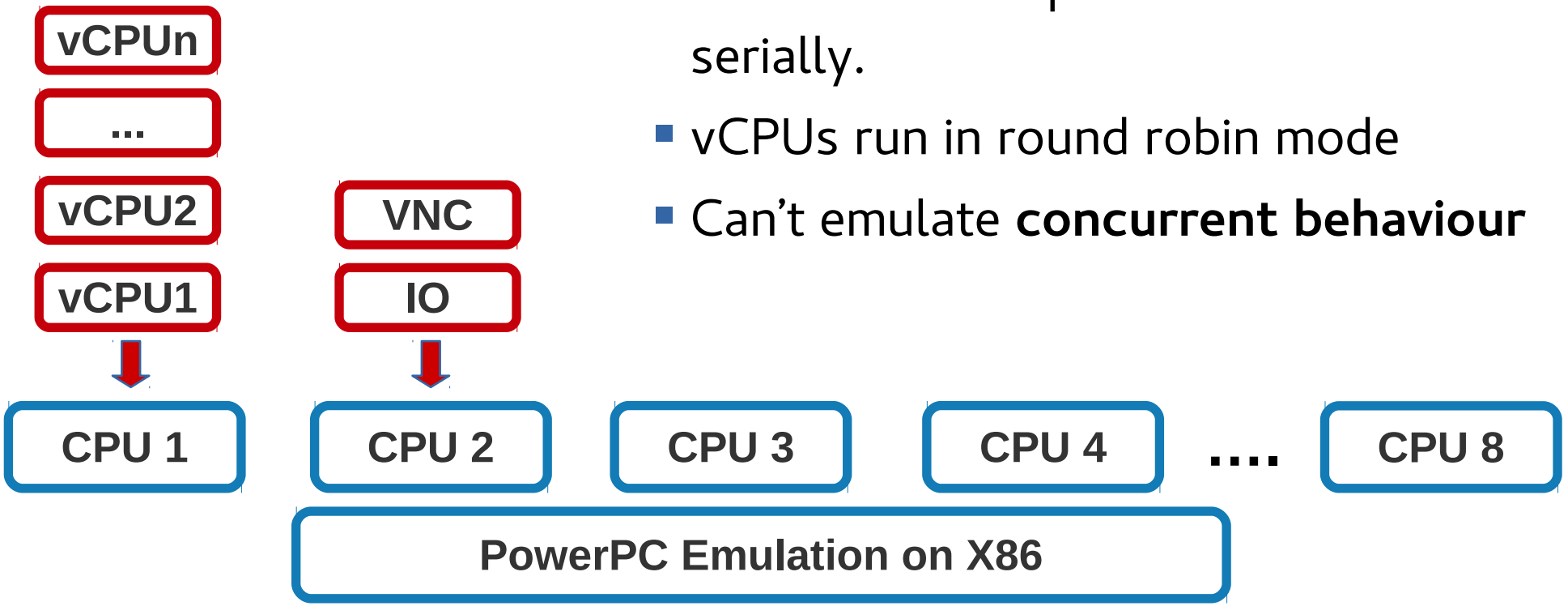
- Emulate Bare Metal POWER platform
- Model Board Management controller (BMC)
- Supports Hypervisor mode
- Can run nested guest
- Assists in early bringup
- Support IPMI

Status: PowerNV

- Initiated by Benjamin Herrenschmidt
- Cédric Le Goater developing and pushing patches upstream
- PowerNV ~50 preparatory patches upstream
  - POWER8 Hypervisor SPRs
  - Split Instruction and Data caches
  - Batching TLB flushes
  - XICS rework to support new native model

PowerPC support for Multi-threaded TCG

System emulation - runs vCPUs serially

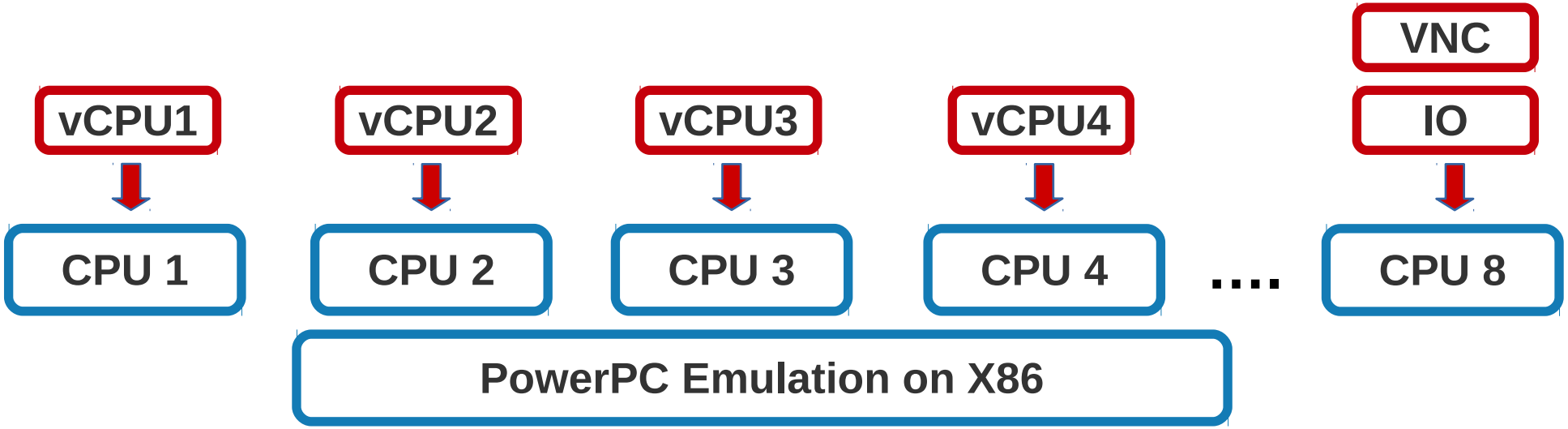


- Emulates multi-processor VM: but serially.
- vCPUs run in round robin mode
- Can't emulate **concurrent behaviour**

Credits: Alex Bennée
Towards Multithreaded TCG

QEMU Multi-threaded TCG

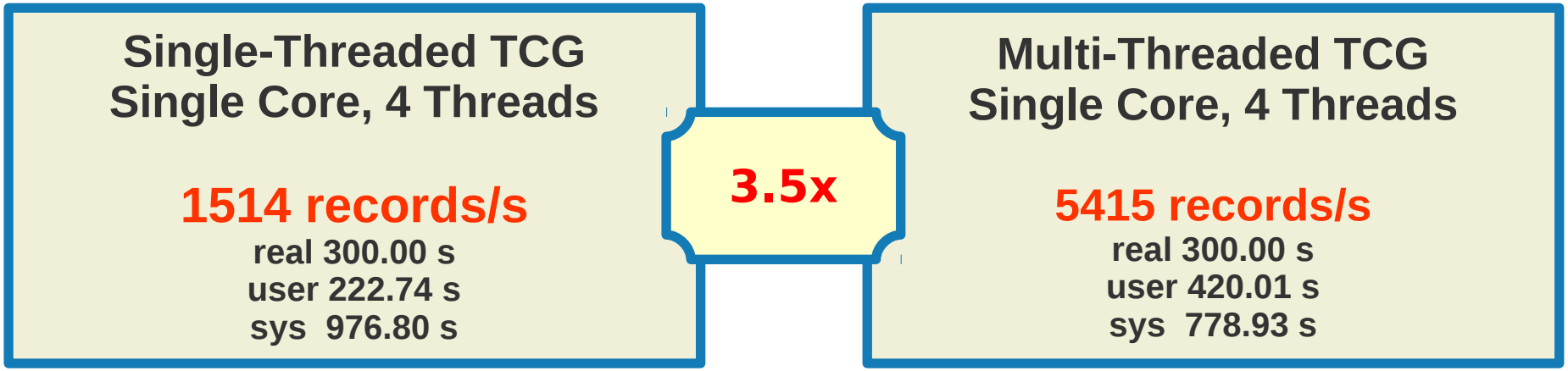
- QEMU for multi-core system bringup
- Community effort in progress
- Challenges: Atomics, Memory Barriers, TLB Flush, etc.



Credits: Alex Bennée
Towards Multithreaded TCG

Status – PPC support for MTTCG

- Based on MTTCG base patches and atomic cmpxchg
- Take iothread locks during hcalls
- Load with reservation(lwarx and family)
- Store conditional(stwcx. and family) with atomic cmpxchg micro-ops
- Booted VM with 4 vCPUs
- Ebizzy performance (ebizzy -S 300 -t 16)



Challenges: PPC support for MTTCG

- Still unstable
  - https://github.com/nikunjad/qemu/commits/pseries_mttcg_wip
- pSeries uses hcall for page table update/invalidate.
- Memory barriers
- Supporting PowerNV platform

Misc TCG Improvements

- Load/Store improvements – Benjamin Herrenschmidt
- Exception handling improvements – Benjamin Herrenschmidt
- Load/Store consolidation – Nikunj

Future

- Complete POWER ISA 3.0 support
- Upstreaming PowerNV in QEMU
- Future - POWER9 PowerNV support
- Stabilize MTTCG on POWER
- 128bit Load/Store support in TCG
- Testing mechanism for instructions

Credits

- Benjamin Herrenschmidt
- Cédric Le Goater
- Alexander Graf – QEMU's Recompilation Engine

  https://dl.dropboxusercontent.com/u/8976842/TCG.pdf
- Alex Bennée – Towards Multithreaded TCG

  http://www.linux-kvm.org/images/c/cf/02x02-Alex_Benee-Towards_Multithreaded_TCG.pdf

Legal statement

- This work represents the views of the author, and does not necessarily represent the view of IBM
- IBM and IBM (logo) are trademark of International Business Machines in the United States and/or other.
- Linux is a registered trademark of Linus Torvalds
- Other company, product, logos and service names may be trademarks or service marks of others
- This document is provided “AS IS”, with no express or implied warranties. Use the information in this document at your own risk
- Results mentioned in the presentation is for reference purposes only, and are not to be relied on in any manner.

धन्यवाद

Thank you