

Zero Architecture

Huawei's Next-Generation Virtualization 3.0 Integrates Hardware and Software

Jinsong Liu <liu.jinsong@Huawei.com>
Zhichao Huang <Huangzhichao@Huawei.com>

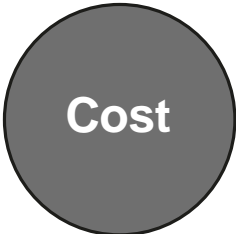


Gaps Between Current Virtualization and Cloud Service Requirements



Performance

3%–15% virtualization overhead
15% computing loss



Cost

10%–20% CPU reserved
10%–20% memory reserved



Security

VM escape
Data leak



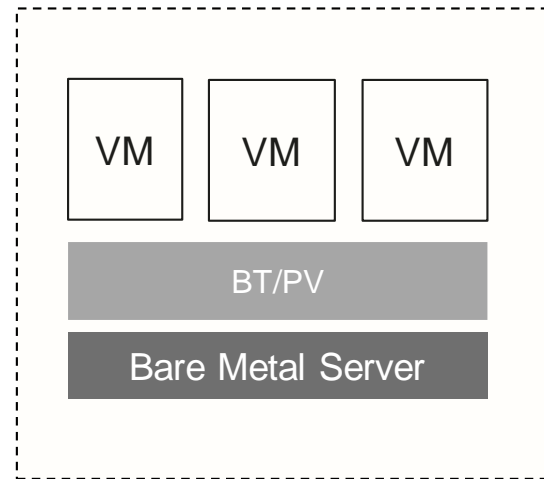
Stability

Neighbor noise
Unstable performance

Evolution of Virtualization Technology

Virtualization 1.0

Till 2003

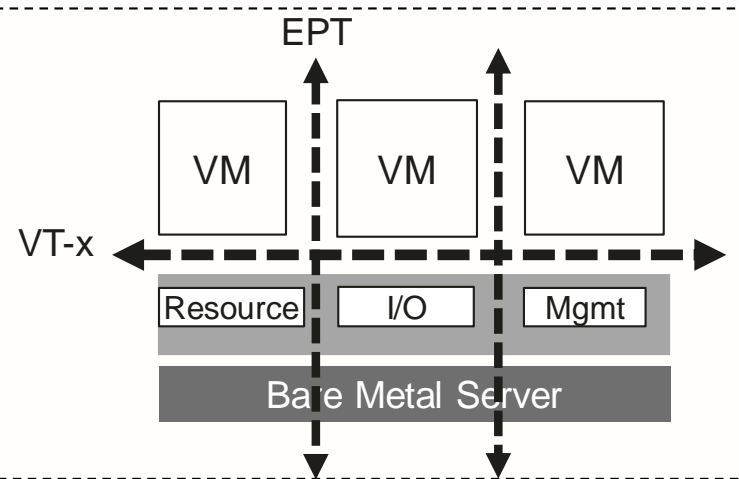


Software Virtualization

- Stanford: Binary Translation
- Cambridge: Para Virtualization
- Complicated and insecure

Virtualization 2.0

2004–2017

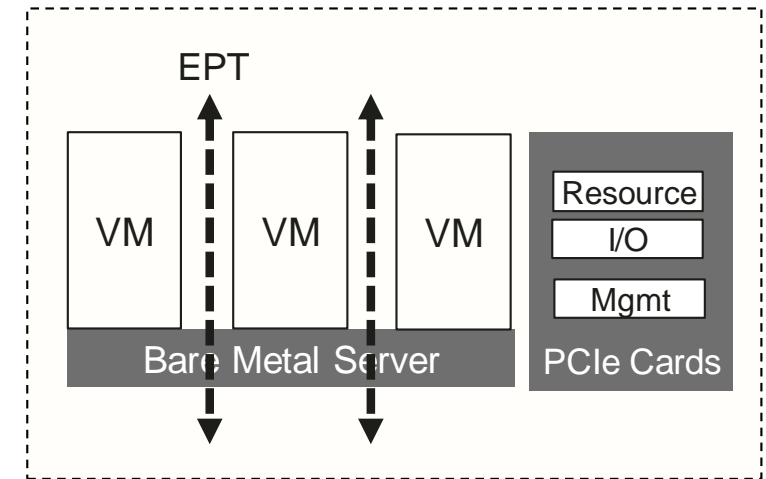


Intel & AMD: Hardware-assisted

- CPU: VT-x with high overhead
- Memory: EPT enables memory isolation with low overhead
- I/O: VT-d with ecosystem and scalability issues

Virtualization 3.0

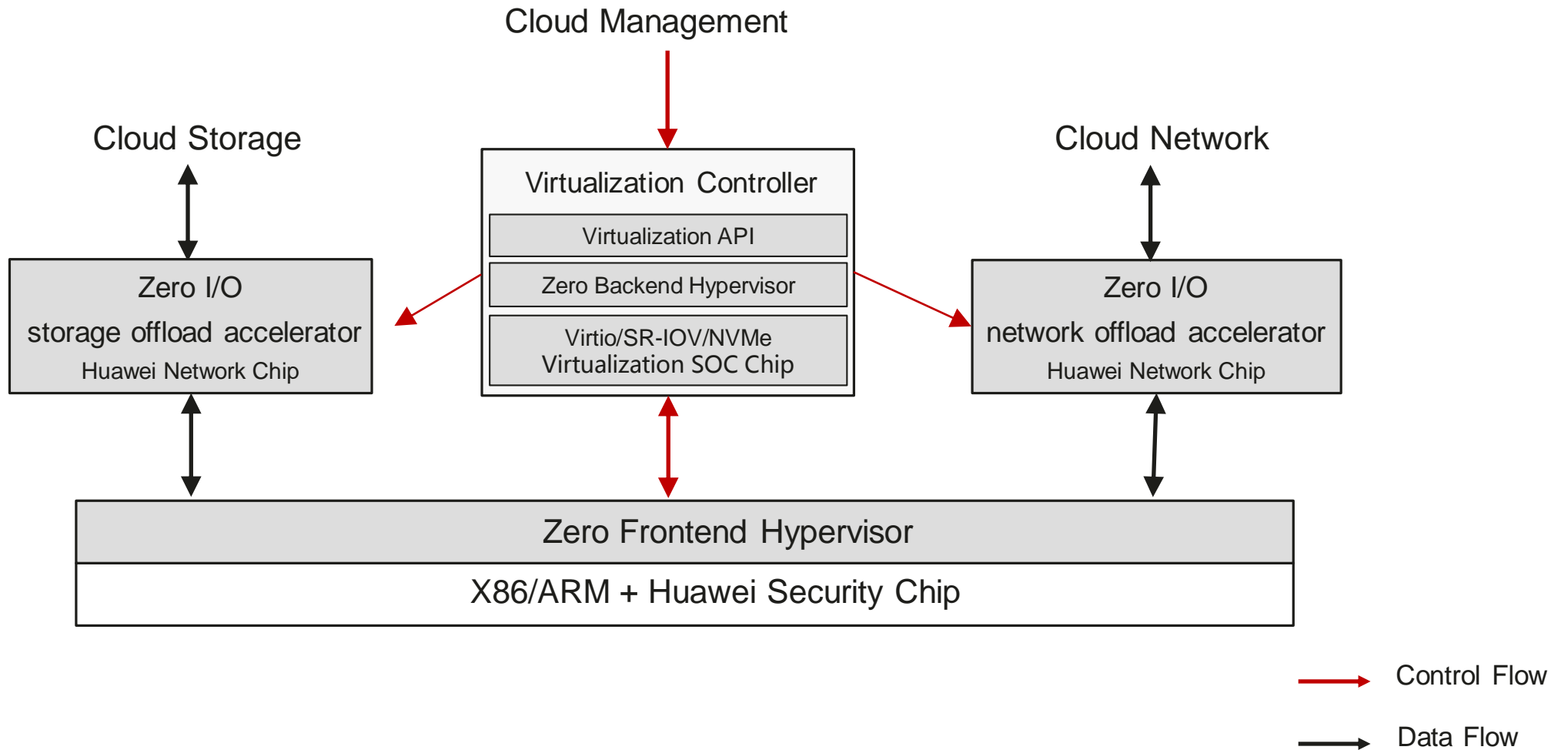
2018 to now



Cloud Vendors: Customized Hypervisors & Chips

- CPU: ultra-simplified customization with no virtualization overhead
- Memory: adopts mature isolation mechanisms
- I/O: Hardware offloading and acceleration with high and scalable I/O performance

Zero Architecture, Huawei's Next-Generation Virtualization 3.0 Integrates Hardware and Software



Zero Components

Zero = Zero Hypervisor + Zero Virtualization Controller + Zero I/O



Near-Bare Metal Computing

- Split-hypervisor: ~0 resource reserved
- Ultra-simplified customization: ~0 computing loss
- Performance close to BMS: ~0 jitter
- Enterprise-class virtualization: ~0 overhead

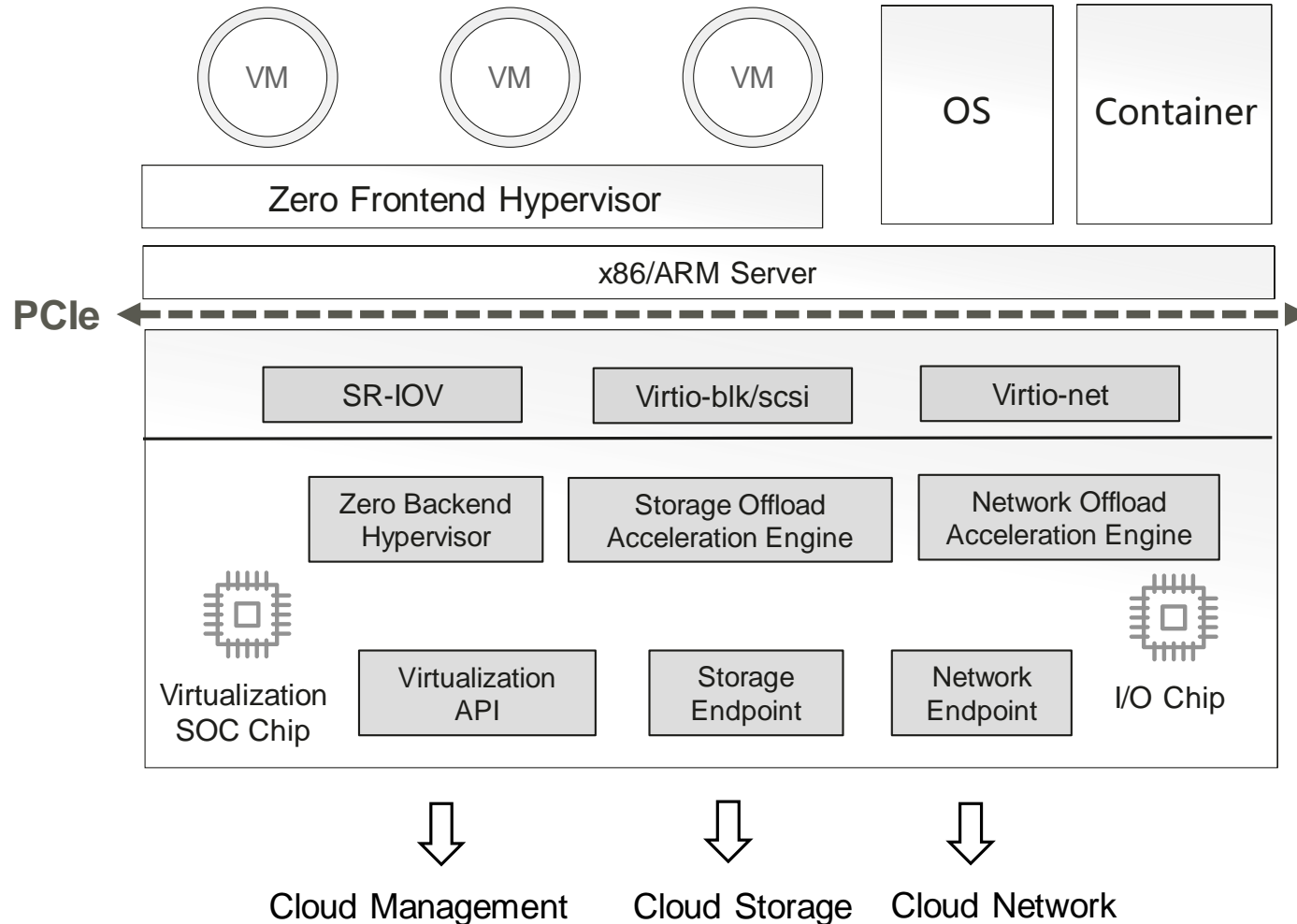
High Scalability and Security

- Virtualization SOC Chip
- Online configuration of Virtio, SR-IOV, and NVMe
- Unified management of VMs and BMSs
- Supporting pass-through live migration
- Supporting x86 and ARM
- Supporting Huawei security chip and providing enterprise-class Roots of Trust
- Hardware isolation and mini-TCB, providing enterprise-class trusted baseline

High-Performance I/O

- Huawei Network chip, I/O offload and acceleration
- High performance network: 18 Mpps @full stack + 45 Gbit/s
- High performance storage: 1 M IOPS + 100 μ s

Zero Supports BMSs, VMs, and Containers



Unified Data Plane

- Supports BMSs, VMs, and containers



Near-Bare Metal Computing

- Performance and stability close to BMSs



High-Performance I/O

- EVS: 1M iops, 100 μ s
- VPC: 18M pps, 45 Gbit/s



Pass-Through Live Migration

- Online upgrade

Zero Computing System

Single control card

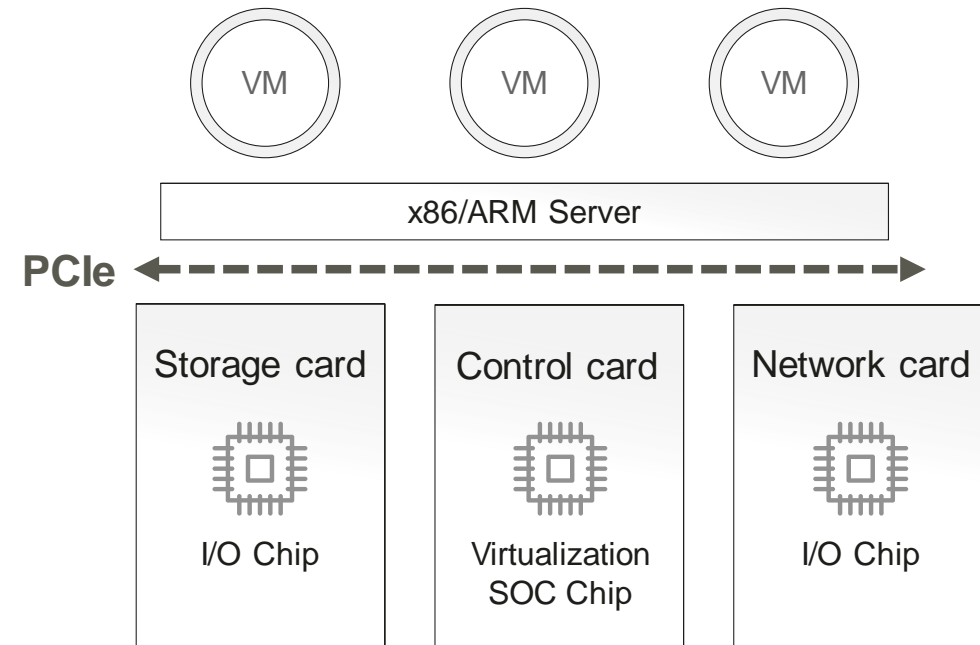
- Provides management API
- Coordinates all other cards and host

Multi I/O cards

- Optional and configurable
- Support native VPC interface
- Support native EBS interface

Host

- No local disk, no network
- Tiny Linux, least TCB
- ECS mode, BMS mode, Container mode, fast switch
- All CPU/MEM sold



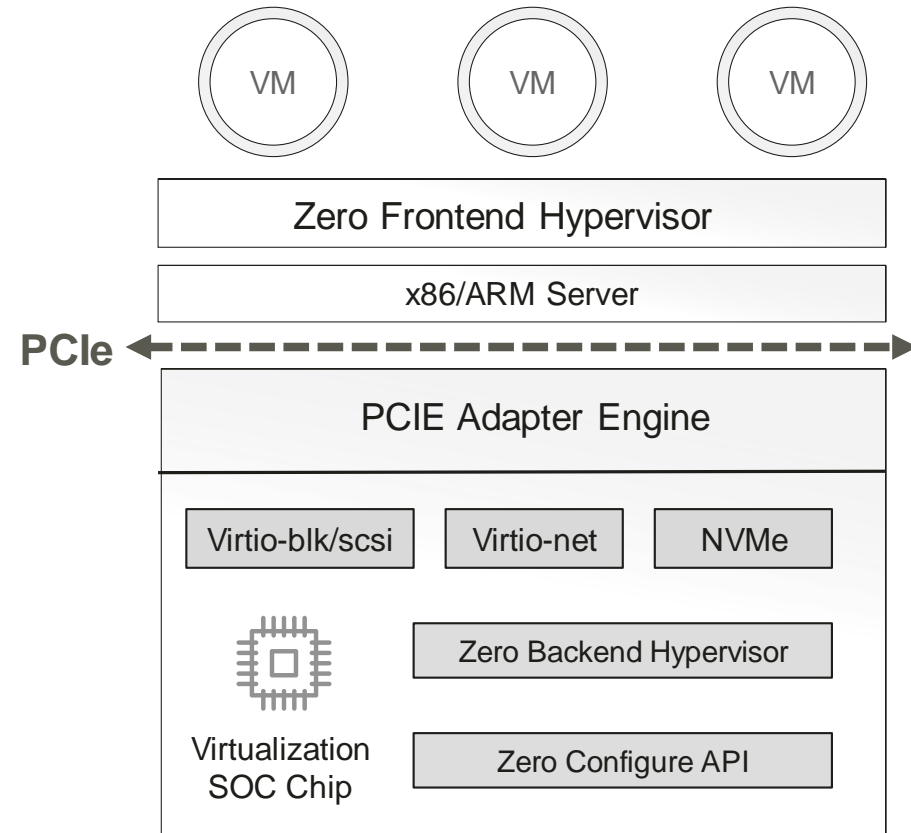
Zero I/O System

Software define chip function

- PCIe : Virtio-blk/Virtio-scsi/Virtio-net/NVMe/...
- Fast switch among above modes
- Un-modified guest OS

Bare Metal-like performance

- PCIe pass-through
- Descriptor prefetch
- Batch processing requests
- Reduce DMA frequency



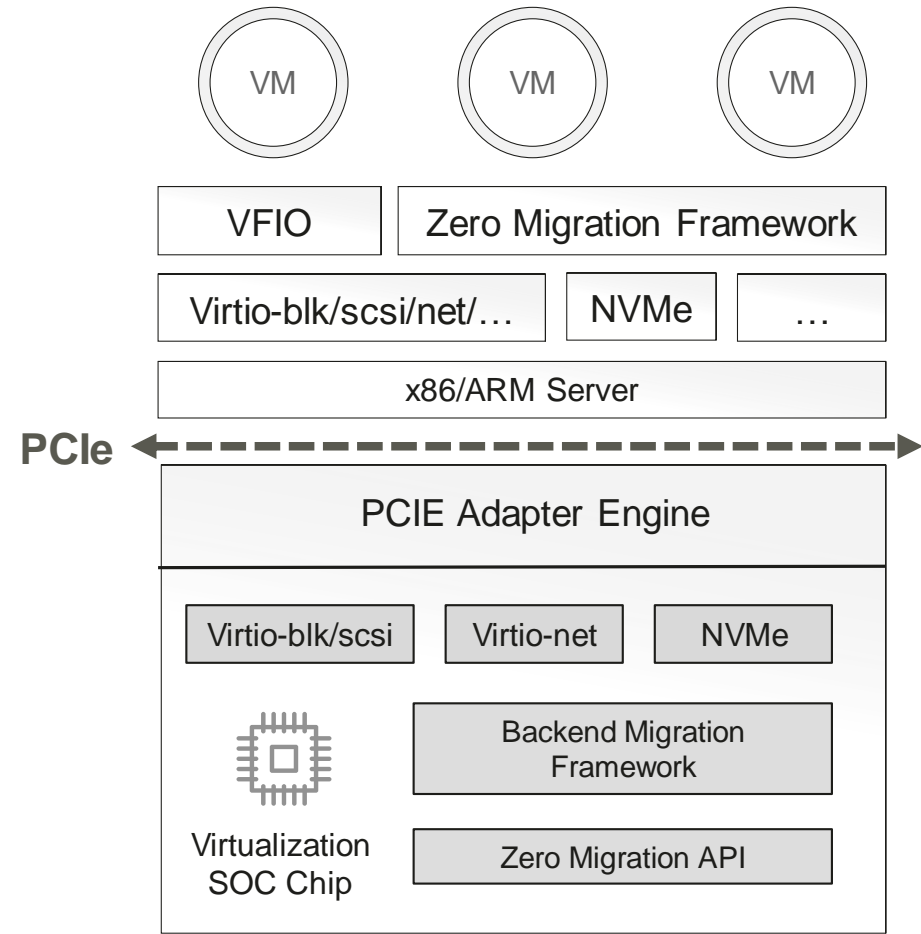
Zero O&M: Pass-Through Live Migration

Pass-through migration framework:

- Virtio-blk/Virtio-scsi/Virtio-net/NVMe/...
- Native guest OS, no modification
- Dirty page tracking
- Device state save and restore

Best performance:

- Concurrent log sync/ save/ restore
- Pause and resume asynchronously
- Downtime: 50ms



Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home and
organization for a fully connected,
intelligent world.

**Copyright©2019 Huawei Technologies Co., Ltd.
All Rights Reserved.**

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

