



# Ceph @ MIMOS

Presented by Hong Ong

---

18th<sup>n</sup> November 2015





# Outline

- MIMOS: An Overview
- Ceph@MIMOS Use Cases:
  - Backup
  - Virtual Desktop Infrastructure (VDI)
- Moving Forward





# MIMOS: An Overview

Enhancing ICT industry growth through indigenous technologies

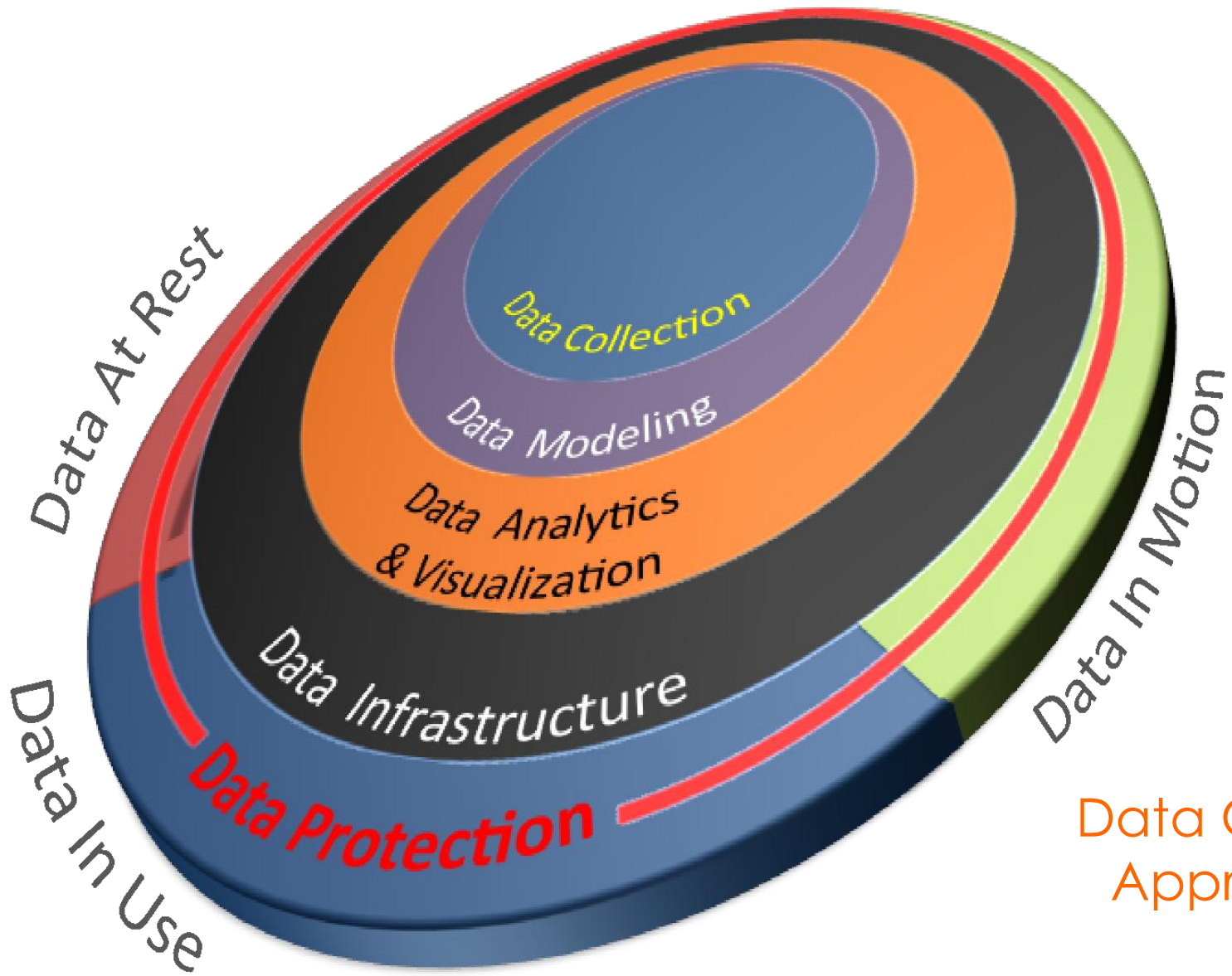
## R&D in ICT

<http://www.mimos.my>





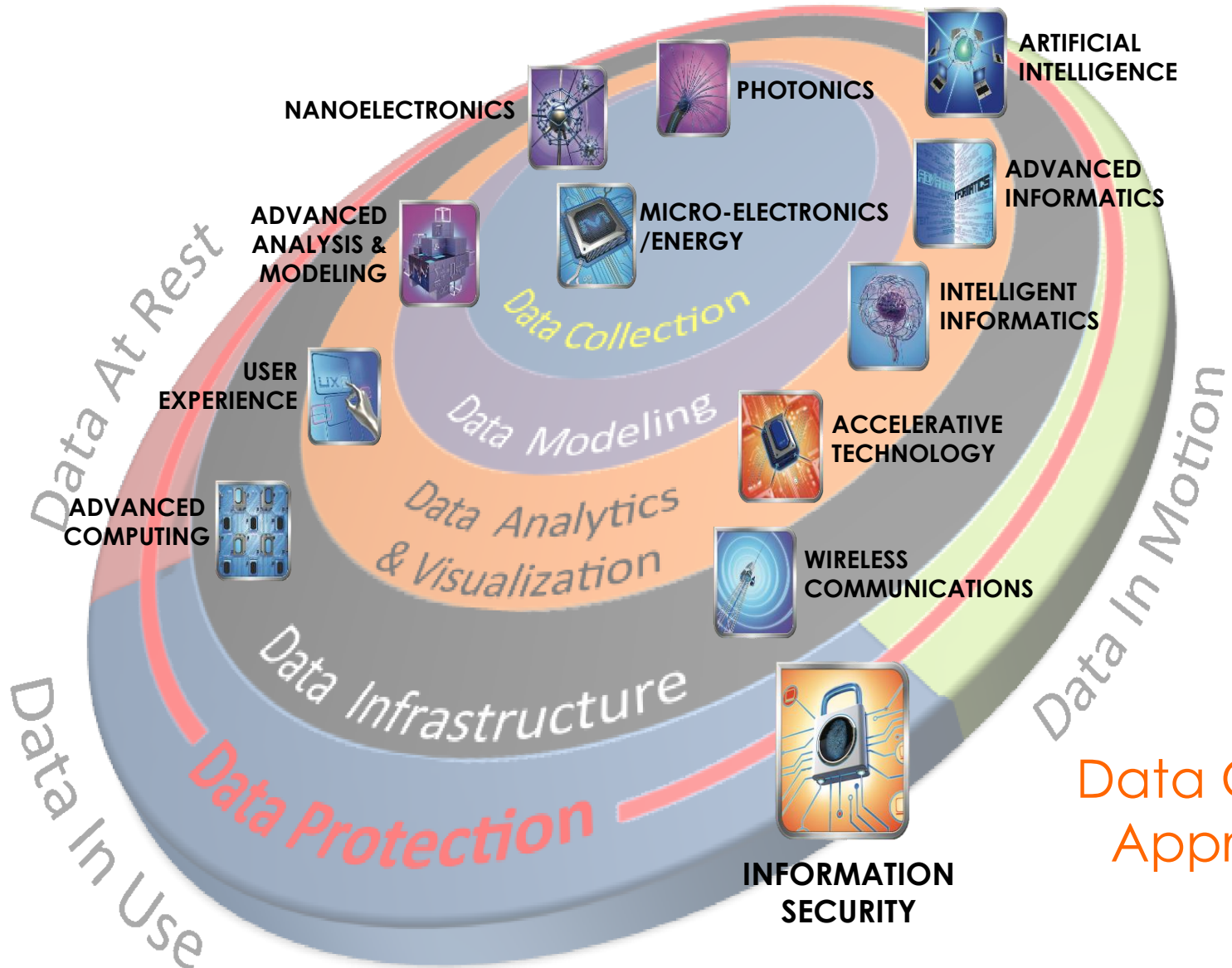
# MIMOS R&D Labs



Data Centric  
Approach



# MIMOS R&D Labs



Data Centric Approach





# Ceph@MIMOS Use Case #1: Backup

“Can we move backups to Ceph and distribute the objects across 2 datacenters, which is about 400km away?”





# Testbed Environment

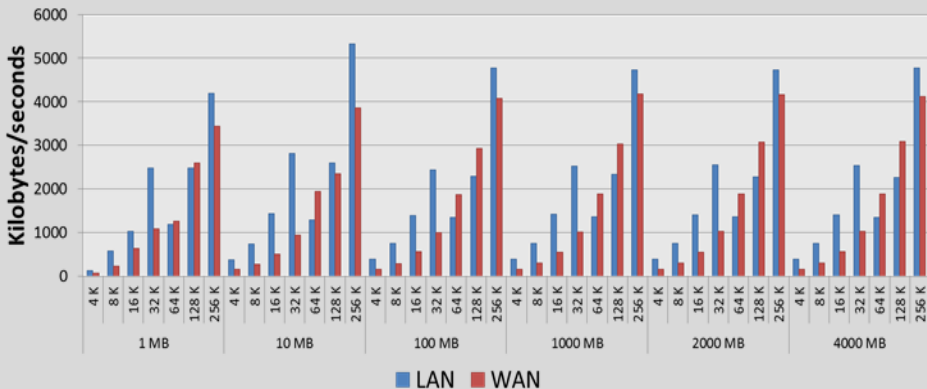
- Started in late 2013
  - 3 MONs (with 3 MDSs) and 2 OSD Servers
    - Each with 12 x 4TB SATA, bonded 2 x GigE for both public and cluster networks
    - Added 5 more OSD servers; each with 10 x 4TB SATA and 2 SSDs in 2015
  - Initially with Emperor and Ubuntu 12.04
    - Upgraded to Hammer and Ubuntu 14.04 in late May 2015
  - Backup software: BackupPC
  - Volume/Mount Point: CephFS (kernel module)



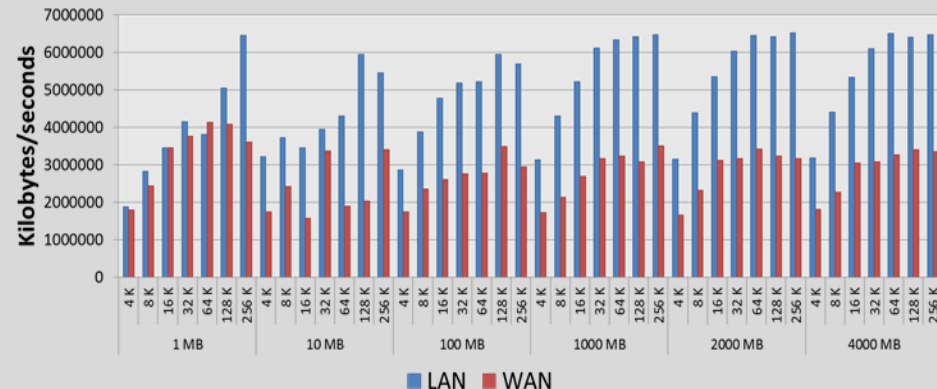
# Benchmark Results

Round-trip time in ms	Bandwidth (Mbps)	2 TCP Iperf	Min	Avg	Max	Mdev
<b>DC1 KL to DC1</b>	250	96%	13.149	13.491	16.167	0.684
<b>Kulim</b>						
<b>DC2 KL to DC1</b>	250	96%	13.176	14.004	17.665	1.079
<b>Kulim</b>						
<b>DC1 KL to DC2</b>	1000	86%	0.422	0.490	1.203	0.136
<b>KL</b>						

Ceph Sequential Write Profile



Ceph Sequential Read Profile



1. Irrespective of LAN or WAN, Ceph scaled well in all scenarios whether its Read or Write.
2. WAN based Ceph cluster is a doable scenario if we have sufficient bandwidth





# Issues and Challenges - Backup

- MDSs started to crash after 4 months+ of stable operation
  - Patched by *Yan, Zheng (special thanks)*; system back into normal operation
- One of the new server's OS randomly "froze"
  - Eventually downgraded kernel of the problematic server from 3.19.x to 3.16.x *after noticed other servers are using 3.16.x.*
  - Perhaps regression issue in the kernel (?)



# Ceph@MIMOS Use Case #2: VDI

- Due to project requirement, we needed a controlled environment where users remotely access a windows desktop/client for development
- Solution – Open Nebula + Ceph (Emperor)
  - 60+ Windows 7 VMs with RDP
  - 30+ development VMs
  - Additional attached storage





# Issues and Challenges - VDI

- Couple Issues:
  - Responsiveness of VMs - takes more than 3 minutes to get Eclipse to start
    - Workaround: use journal-less OSDs (LevelDB keyvalue store backend with Ceph Firefly)
  - Frequent crashes due to OSD servers ran out of RAM and Placement Groups (PGs) in inconsistent state
    - Workaround: added cron jobs to drop cache, release heap and fix PGs



# Moving Forward

- Extending and expanding to support more in-house initiatives (primarily to reduce storage cost)
  - MIMOS Cloud Platform (Mi-Cloud)
    - Open Nebula + Qemu/KVM + librbd
    - Current up to 300+ VMs
  - Common file sharing
    - Currently support NFS and SAMBA mount
    - iSCSI (next)
- Development
  - Developed an simple NAS “appliance” to simplify deployment and management
  - Tune and optimize for other mix workloads (databases, web apps, big data, HPC, etc.)



# Recent Publications and Patents

- Improving Performance of Database Appliances on Distributed Object Storage. ICCCRI 2015
- A Study of Distributed File System for Database Hosting in a Cloud Environment, International Conference on Cloud Computing Research and Innovation. ICCCRI 2014
- Simulate Distributed File System on Wide Area Network using Virtualized Environment, Asia Pacific Advanced Network. APAN 38 – NRW, 2014
- Challenges of deploying Wide-Area-Network Distributed Storage System under network and reliability constraints – A case study. PRAGMA 26, 2014
- Preliminary Study of Two Distributed File Systems for Cloud Infrastructure Storage: Ceph vs. Gluster (PRAGMA 25, 2013)
- *System and Method for Distributed, Secured Storage in Torus Network. PI 2014701657*
- Management of Block Device Image and Snapshot in Distributed Storage of Torus Network Topology. PI 2015700043
- Method to Fulfil Multi-Class Distributed Storage SLA and QoS Using Dynamic Network Load and Location Awareness. PI 2015702101



Hong Ong – [hongong@computer.org](mailto:hongong@computer.org)



# Thank You

---

© 2015 MIMOS Berhad. All rights reserved.