

VM and Cloud Computing (Research projects @ HKU.CS.SRG)



Cho-Li Wang The Systems Research Group (SRG) Department of Computer Science

Systems Research Group @HKU

Our Motto: solving <u>REAL</u> problems with the use of <u>REAL</u> computing resources (e.g., PC, PDA, mobile phones, *clusters*).



"Self-Made" Gideon 300 cluster in 2002



300 Pentium4 PCs @355 Gflops; Ranked #175 in TOP500 (11/2002)

HKU Computing Resources

- Total # of cores: 3004 CPU + 5376 GPU
- RAM Size: 8.34 T
- Disk storage: 130 T
- Peak computing power: 27.05 TFlops



Gideon-II & MDRP Cluster



GPU-Cluster (Nvidia M2050): **7.62 Tflops**



HKU Gideon-II Cluster (HKU/SRG)



China National Grid (CNGrid) Project



Ongoing Projects

- WAVNet:
 - Wide-Area Network Virtualization Technique for Private Cloud
- vBalance:
 - VM I/O Performance Isolation Techniques (Not Discussed)
- BetterLife 2.0:
 - Personalized recommendation service on Cloud
- eXCloud:
 - Multi-granularity Migration Support for Scaling Mobile Applications in Cloud : (1) <u>JESSICA2</u>: A distributed JVM with transparent thread migration. (2) <u>SODEE</u>: A "Stack-On-Demand" (SOD) execution model for elastic Mobile Cloud Computing.

Multicore and GPU Computing:



<u>PI/ GRF</u>: Japonica : Java with <u>Auto-Parallelization ON</u> Graph<u>Ics Co-processing Architecture (9/2011-8/2013)</u>

<u>PI/GRF</u>: Software Coherence for Future Tiled Architectures, (9/2012-8/2015)

WAVNet: Live VM Migration over WAN

- A P2P cloud with Live VM • **Migration over WAN**
- **NAT hole punching**
 - Establish direct host-to-host connection
 - Free from proxies

Virtual LAN 1

VM Layer

Desktob Layer

Rendezvous Laver

Virtual LAN

Able to traverse most NATs

Virtual LAN 2

Physical host

NAT/firewall

Switch

۲

1 0.8

0.5

CAN Overlay



WAVNet : Team Members









Bolla Xu (Graduated)



Weida Zhang

Luwei Cheng



2012/7/18

WAVNet: Live VM Migration over WAN

o Experiments at Pacific Rim Areas



(2) WAVNet: Live VM Migration over WAN

Ping RTT drops after migration Freeze Time: 0.6s ~ 2.1s



"WAVNet: Wide-Area Network Virtualization Technique for Virtual Private Cloud" (ICPP2011).

StoryTelling@Home on WAVNet



StoryTelling@HOME

Hello, Eric! Log out

Upload new story Add new friends View your profile

List of recent stories

Story Upload (Apache MyFaces) Story Search Streaming/Downloading

13/3									
	sid	Author Path		Publish time	Live Streaming Like/Unlike Comm		Comments	Direct Download	
	29	Eric	<u>192.168.123.179/01. Instrumental 1.mp3</u>	Wednesday Apr 13, 04:00:00		<u>like</u>	<u> View comments(0)</u>	<u>Download</u>	
	30	Eric	<u>192.168.123.179/05. </u>	Wednesday Apr 13, 04:00:00	0	<u>unlike</u>	<u> View comments(0)</u>	<u>Download</u>	
	31	Eric	<u>192.168.123.179/Que te Pasa.mp3</u>	Wednesday Apr 13, 04:00:00		<u>like</u>	<u> View comments(0)</u>	<u>Download</u>	

List of the stories from yourself

1.3/3									
	sid	Author	Path	Publish time	Live Streaming	Direct Download			
	29	Eric	192.168.123.179/01. Instrumental 1.mp3	Wednesday Apr 13, 04:00:00		<u>Download</u>			
	30	Eric	192.168.123.179/05.	Wednesday Apr 13, 04:00:00	0	<u>Download</u>			
	31	Eric	192.168.123.179/Que te Pasa.mp3	Wednesday Apr 13, 04:00:00		<u>Download</u>			

Prototyped by Eric Wu, Queena Fung (2011)

BetterLife 2.0: Context-aware Recommendation System

Case-based Reasoning (CBR): Solve new problems by finding previous similar experiences



13

BetterLife 2.0: Shopping Recommender

103 user created by Elgg social networking Locations of convenient stores in Hong Kong Based on social closeness of these 103 users.



Prototyped by Lo Fung, Kong Kwai Yee, Wong Kwok Kit

BetterLife 2.0: Query by Image Example

Similar Photo Search

- Off-line feature extraction using MapReduce (~1M images)
- Locality Sensitive Hashing (LSH) for high-dimensional indexing and searching
- Weight Learning by Evolution Algorithm
- Context info: GPS (used in the test)





BetterLife 2.0: Query by Image Example





Profile

Wikipedia Photo Gallery Google Map

Feature Type	Dimension	Expertise			
TAMURA	18	Texture			
GABOR	60	Lines detection			
SCALABLE COLOR	64	Color in HSV space			
EDGE HISTOGRAM	80	Five Types of Edges			
CEDD	144	Color and Texture			
FCTH	192	Color and Texture, enhanced CEDD			
ACC	256	Color Correlogram			
SURF	128	Scale Invariant Interest Point			

Total features points: 942

Prototyped by Jason Wan (2010) and Hao Liu (2011)



eXCloud: Multi-granularity Migration Support





History and Roadmap of JESSICA

JESSICA V1.0 (1996-1999)

- Execution mode: Interpreter Mode
- JVM kernel modification (Kaffe JVM)
- Global Heap: built on top of TreadMarks (Lazy Release Consistency + homeless)

JESSICA V2.0 (2000-2006)

- Execution mode: JIT-Compiler Mode (full speed)
- JVM kernel modification (Kaffe JVM)
- Lazy Release Consistency + migratinghome protocol

JESSICA V3.0 (2008~2010)

- Built above JVM (JVMTI)
- Support Large Object Space
- For any JVM. Run @ full speed of the underlying JVM.
- JESSICA v.4 (2010~)
 - Software transactional memory model
 - Multicore/GPU cluster (Japonica, 2012-2015)



Past Members



King Tin LAM,

Chenggang Zhang





Kinson Chan

Ricky Ma

Current Members

eXCloud = WAVNet+ J2+ SOD



Thanks!

For more information:

Cloud Computing Projects http://i.cs.hku.hk/~clwang/projects/cloud.html

C.L. Wang's webpage: http://www.cs.hku.hk/~clwang/

JESSICA Related Publications

- JESSICA2: A Distributed Java Virtual Machine with Transparent Thread Migration Support, (CLUSTER 2002).
- High Performance Computing on Clusters : The Distributed JVM Approach, High Performance Computing: Paradigm and Infrastructure, <u>John Wiley & Sons, Inc</u>. 2004.
- A Performance Study of Clustering Web Application Servers with Distributed JVM, The 14th IEEE International Conference on Parallel and Distributed Systems (ICPADS'08).
- "Path-Analytic Distributed Object Prefetching" (I-SPAN 2009).
- "Web Application Server Clustering: the Distributed JVM Approach", <u>Handbook of Research on Scalable Computing</u> <u>Technologies</u>, IGI Global, pp.658-681 (Chapter 28), July 2009
- "Adaptive Sampling-Based Profiling Techniques for Optimizing the Distributed JVM Runtime," (IPDPS 2010)

Other Publications

VM Migration:

 "WAVNet:Wide-Area Network Virtualization Technique for Virtual Private Cloud" (ICPP2011).

VM Performance Isolation

- Cooperative Scheduling in Defeating Network Jitter for Virtual Machines, submitted to 4th IEEE/ACM International Conference on Utility and Cloud Computing (Best Student Paper Award)
- Network Performance Isolation for Latency-sensitive Cloud Applications, to appear in *Future Generation Computer Systems*

Stack on Demand

- "A Stack-On-Demand Execution Model for Elastic Computing" (ICPP2010)
- "eXCloud: Transparent Runtime Support for Scaling Mobile Applications," 2011 IEEE International Conference on Cloud and Service Computing (<u>CSC2011</u>) (Best Paper Award)
- Lightweight Application-level Task Migration for Mobile Cloud Computing, The 26th IEEE International Conference on Advanced Information Networking and Applications (<u>AINA-2012</u>), Fukuoka, Japan, March 26-29, 2012

Other Publications

Resource Allocation

- Probabilistic Best-fit Multi-dimensional Range Query in Self-Organizing Cloud (<u>ICPP2011</u>)
- "Conflict-minimizing Dynamic Load Balancing for P2P Desktop Grid", The 11th IEEE/ACM International Conference on Grid Computing (Grid2010)
- Dual-phase Just-in-time Workflow Scheduling in P2P Grid Systems, The 39th International Conference on Parallel Processing (<u>ICPP2010</u>)
- Gossip-based Dynamic Load Balancing in an Autonomous Desktop Grid, The 10th International Conference on High-Performance Computing in Asia-Pacific Region, (<u>HPC Asia</u> <u>2009</u>)
- Decentralized Proactive Resource Allocation for Maximizing Throughput of P2P Grid, submitted to Journal of Parallel and Distributed Computing (JPDC)
- Dynamic Optimization of Multi-Attribute Resource Allocation in Self-Organizing Clouds, to appear in IEEE Transactions on Parallel and Distributed Systems (TPDS).

Comparison of Migration Latency

Sys	SODEE on Xen (Stack mig.)		JESSICA2 on Xen (Thread mig.)		G-JavaMPI on Xen (Process mig.)			JDK on Xen (VM live mig.)				
A n n	Exec. ti	me (sec)	МО	Exec. ti	me (sec)	МО	Exec. ti	me (sec)	МО	Exec. time (sec)		МО
App	w/ mig	w/o mig	(ms)	w/ mig	w/o mig	(ms)	w/ mig	w/o mig	(ms)	w/ mig	w/o mig	(ms)
Fib	12.77	12.69	83	47.31	47.21	96	16.45	12.68	3770	13.37	12.28	1090
NQ	7.72	7.67	49	37.49	37.30	193	7.93	7.63	299	8.36	7.15	1210
TSP	3.59	3.58	13	19.54	19.44	96	3.67	3.59	84	4.76	3.54	1220
FFT	10.79	10.60	194	253.63	250.19	3436	15.13	10.75	4379	12.94	10.15	2790

•SOD has the smallest migration overhead

Frame: Thread : Process : VM = 1 : 3 : 10 : 150