Fish4Knowledge

WP 4

High Performance Storage and Execution Architecture

NCHC, NARL, TW

Team members & Stakholders

Team based on NSC funded Project for F4K

NCHC, NARL (Nat'l Center for HPC)

- Fang-Pang Lin
- Yun-Te Lin (System)
- Chao-Wen Huang (system)
- Hsiu-Mei Chou (DB)
- Shi-Wei Lo (Simulation)
- Yi-Haur Shiao (Viz)
- Yi-Hsuan Chen (SensorNet)
- Tom Cheng (system)

TORI, NALR (Taiwan Ocean Research Institute)

- Chien-Shin Chen (Coral Reef Biology)
- Shi-Hu Ho (Marine SensorNet)
- Chen-Ping Chen (Fish Taxonomy)

NMMBA (Nat'l Museum of Marine Biology and Aquarium)

- Tung-Yun Fan (Coral Reef biology)

Academia Sinica

- Kwang-Tsao Shao (Marine Fish)

Stakholders

- NCHC
- TORI
- NMMBA
- Academia Sinica
- 3rd Nuclear Power Plant, Tai-Power
- Kenting National Park Bureau

Objectives

- O4.1
  - Achieve scalable long term real time capturing and buffering for **multiple undersea video stream**.
- O4.2
  - **Build a Tera-scale data service platform** consisting of repositories for the video data, for the metadata, for the processed data and for the live stream data, and a computational cluster to support analysis.
- O4.3
  - Achieve **high performance** data store and computation access for the data service platform.

Description of Work

- T4.1
  - Enhance current video capturing and storage
- T4.2
  - Build data Storage facility.
- T4.3
  - Develop data process execution interfaces
- T4.4
  - Develop distributed data and computational methods.
- T4.5
  - Support parallelisation.

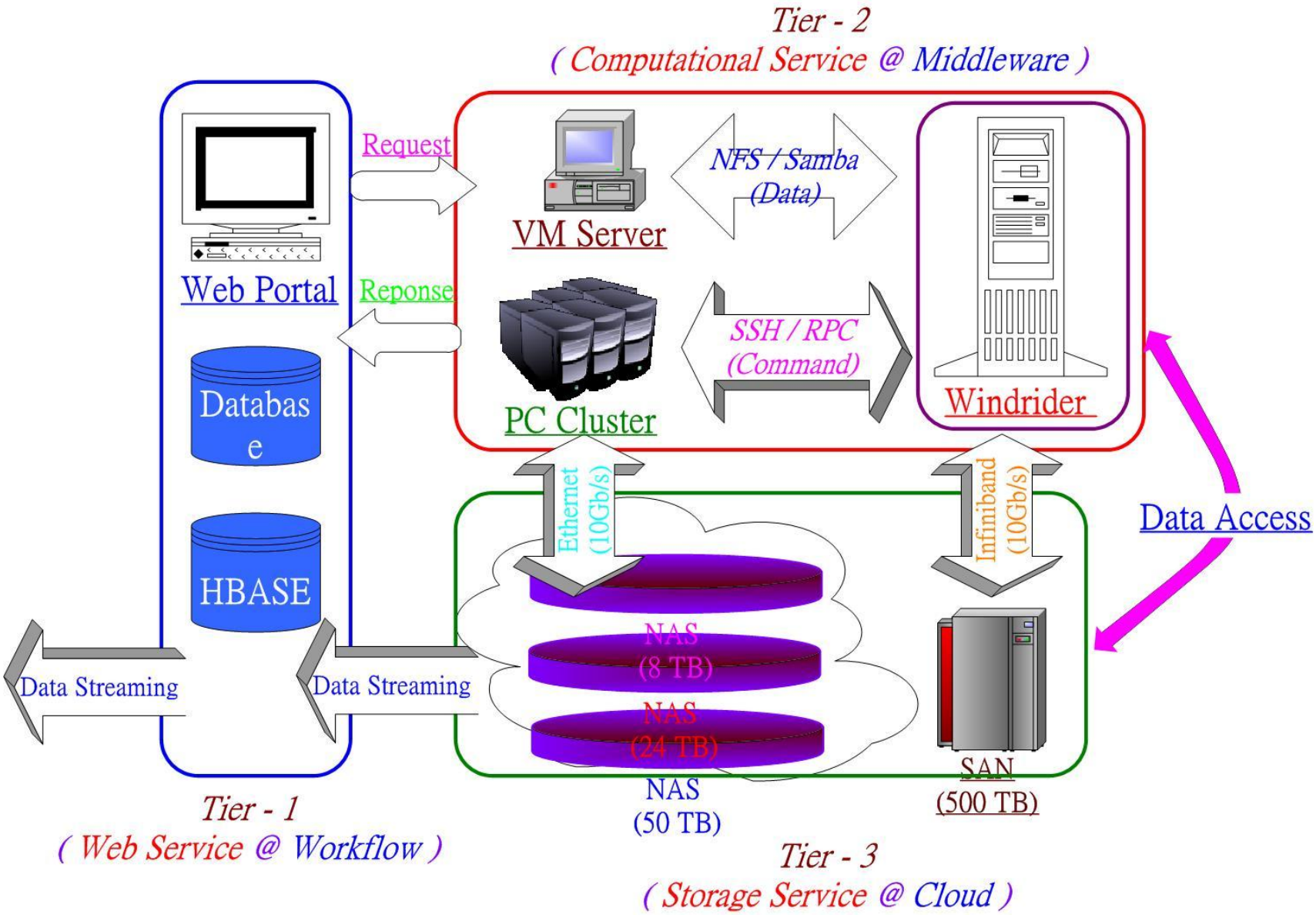
Deliverables

- D4.1: (month 12)
  - *Video and RDF **store**, plus **access***
- D4.2: (month 12)
  - *Workflow **computational platform***
- D4.3: (month 24)
  - Process execution

Summary of Current System development

- 6 N7700 NAS, provide about 48 TB storage size
- 2 WindStar NAS, currently installed totally 16 TB storage and will be extended to 48TB
- Underwater camera, using HD camera and CCTV to provide eco-video as data source to the F4K system.
- 48 cores (4 dies x12 core) servers, provides virtual machine running platform. 96 CPUs from Wind Rider supercomputer (WR) are used for compute service.
- Video Query Portal for accessing the recorded ecology video.
- Customized VM Portal to access virtual machine in the system.

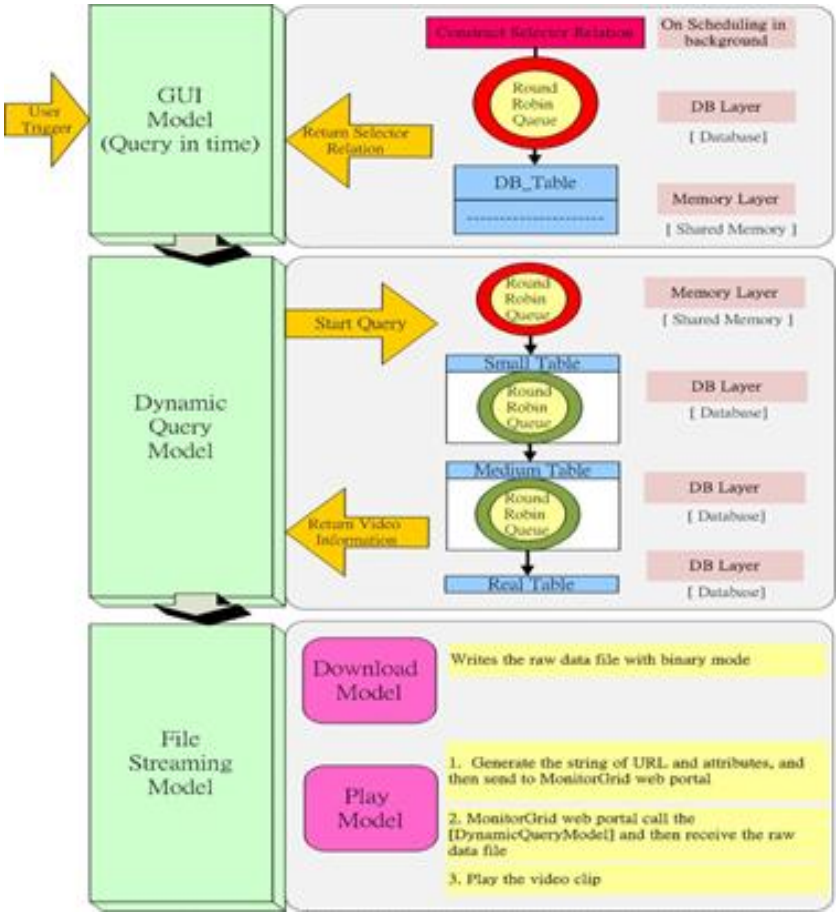
3 Tiers Architecture in NCHC (v1)

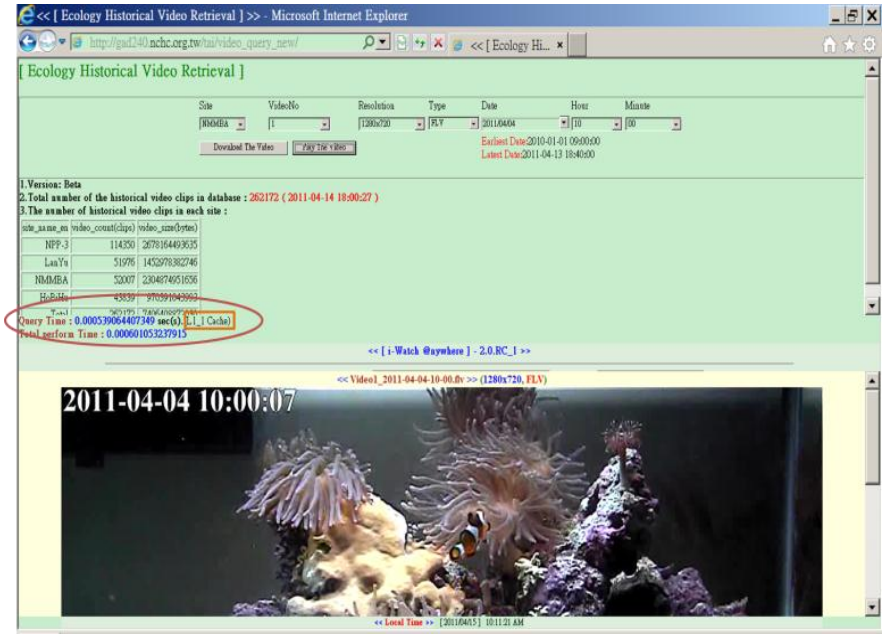


Description of Work

- **T4.1**
  - **Enhance current video capturing and storage**
- T4.2
  - Build data Storage facility.
- T4.3
  - Develop data process execution interfaces
- T4.4
  - Develop distributed data and computational methods.
- T4.5
  - Support parallelisation.

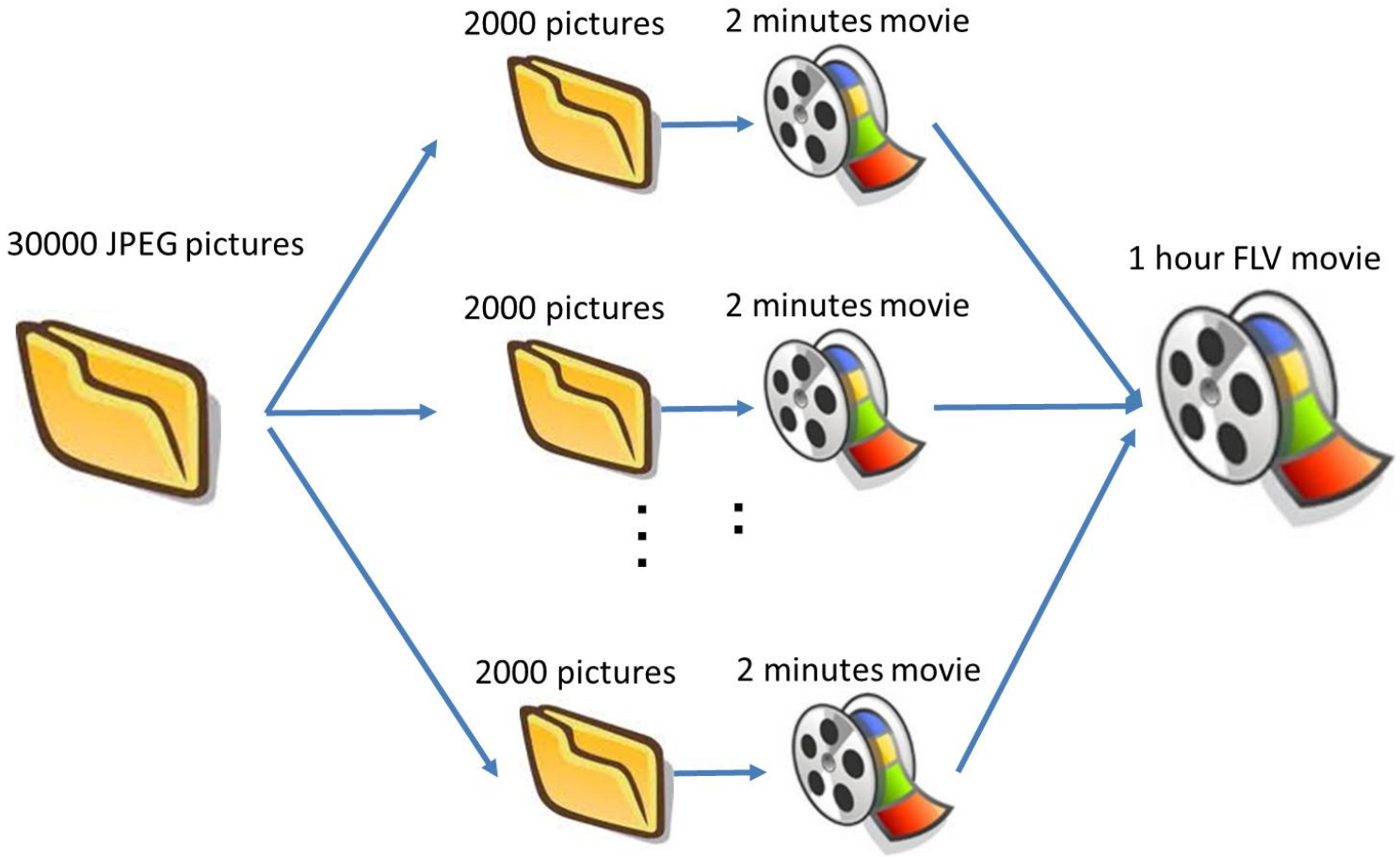
Video capturing and storage



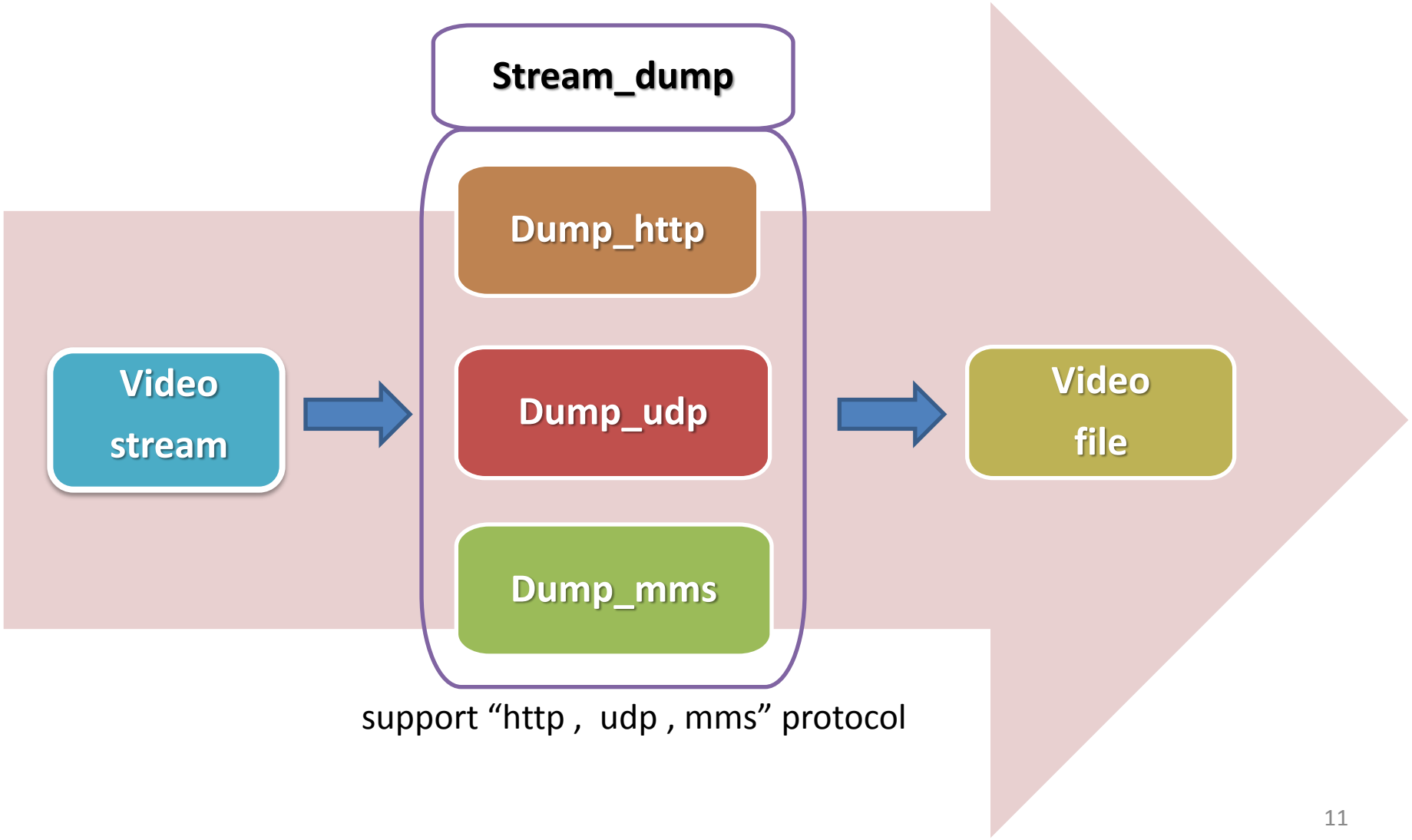


Parallel processing for large data

- modify ffmpeg to support "controlable amount of the image count of each sub-directory"



High Performance stream extracting/encoding



support “http , udp , mms” protocol

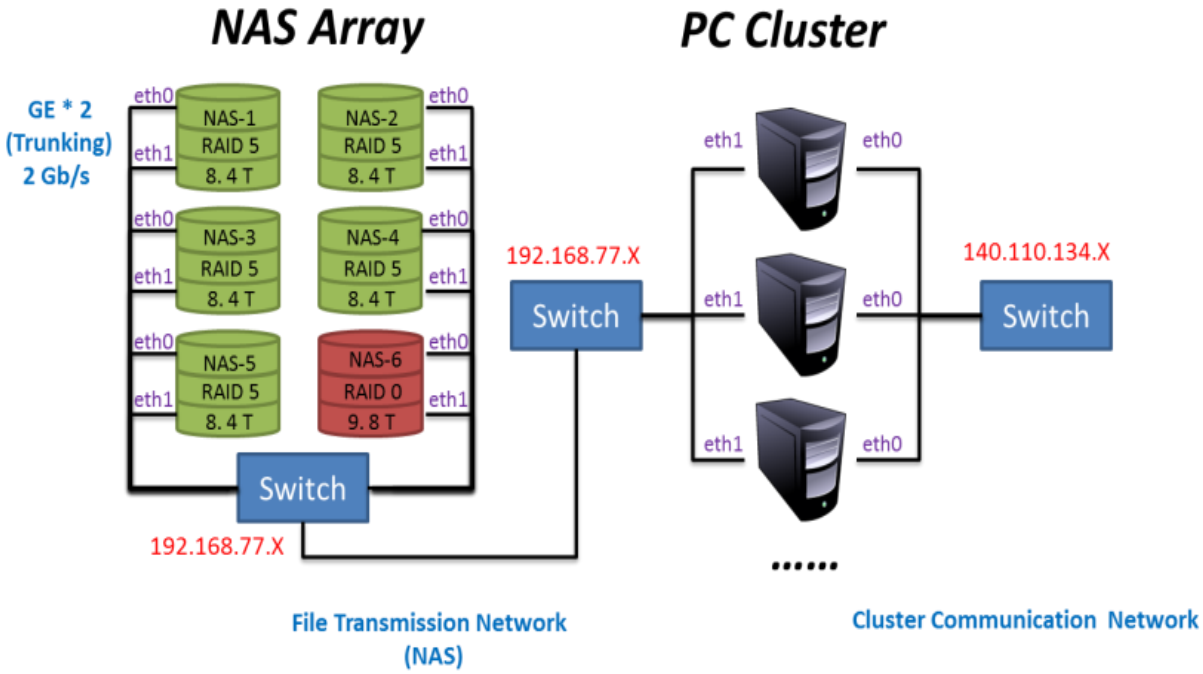
Provide the High Quality Video Clips

- Adapt the current historical video clips production code into the NPP-3 local server.
- Apply the data policy for the high quality video transmission between NPP-3 and NCHC. (used for 4M/1M ADSL network bandwidth limitation)
- Provide the high quality video pre-processing unit in NPP-3 local server to support some kind of the intelligent pre-detection extracting.

Description of Work

- T4.1
  - Enhance current video capturing and storage
- **T4.2**
  - **Build data Storage facility.**
- T4.3
  - Develop data process execution interfaces
- T4.4
  - Develop distributed data and computational methods.
- T4.5
  - Support parallelisation.

Data storage facility



NARL F4K NAS storage Array available for 1GE and 10 GE network interface.

NAS Overview

- Manufacturer – Thecus
- Module – N7700Plus & N7700Pro
- SATA Controller – 7 SATA Disks for Internal; 1 eSATA for External
- RAID Support – JBOD, 0, 1, 5, 6, 10
- Capability – 14TB (XFS, ZFS) / 8TB (ext3) per volume

NAS Implementation

- 6 NASs (N7700Plus) in NCHC Taichung site

	NAS 1	NAS 2	NAS 3	NAS 4	NAS 5	NAS 6
File System	XFS	XFS	EXT4	XFS	XFS	XFS
RAID System	5	5	5	5	5	0
Total Space	8.2 TB	8.2 TB	7.7 TB	8.2 TB	8.2 TB	9.6 TB
Space Used	4.2 TB	1.7 TB	61 GB	620 MB	7.8 TB	161 GB
Space Available	4.0 TB	6.6 TB	7.69 TB	8.13 TB	417 GB	9.4 TB

NAS Implementation

- 1 NAS (N7700Pro) in Kenting site (Taipower southern exhibition)

	NAS in Taipower southern exhibition
File System	EXT4
RAID System	5
Total Space	7.7 TB
Space Used	697 GB
Space Available	7.0 TB

Video Format Overview

- Old Historical Video Format

	CCTV	HD
Format	FLV	FLV
FPS	8	4
Capturing Method	Distributed Extracting/Encoding	Distributed Extracting/Encoding
File Size (10 mins)	20MB ~ 35MB	10MB ~ 20MB
Capturing Period	6:00 AM ~ 18:00 PM (13 hours)	9:00AM ~ 18:00 PM (10 hours)
Data Size Generated Per Day	35MB*6*13 = 2730 MB ≈ 2.73 GB	20MB*6*10 = 1200 MB ≈ 1.2 GB
Data Size Generated Per Month	2.73 GB * 9 (cameras) * 30 (days)= 737 GB	1.2GB * 3 (cameras) * 30 (days) = 108 GB
Data Size Generated Per Year	737GB * 12 (monthes) = 8844 GB	108GB * 12 (monthes) = 1296 GB

Video Format Overview

- New Historical Video Format

	CCTV	HD
Format	FLV	FLV
FPS	24	30
Capturing Method	Stream dump	Stream dump
File size (10 mins)	20MB ~ 35MB	60MB ~ 140 MB
Capturing Period	6:00 AM ~ 18: 00PM (13 hours)	9:00 AM ~ 18:00 PM (10 hours)
Data Size Generated Per Day	35MB*6*13 = 2730MB ≈ 2.73GB	140MB*6*10 = 8400MB ≈ 8.4 GB
Data Size Generated Per Month	2.72GB * 9 (cameras) * 30 (days) = 737 GB	8.4GB * 3 (cameras) * 30 (days) = 756 GB
Total File Size Generated Per Year	737GB * 12 (monthes) = 8844 GB	756GB * 12 (monthes) = 9072 GB

Historical Video Clips – Video Count

- Clip Recorded (count : numbers)

Type / Site	Old Format Resolution: 320x240 Fps: 8	New Format Resolution: 640x480 Fps: 24	Total Count
NPP-3	142,402	57,476	199,878
LanYu	62,410	35,942	98,352
NMMBA	61,940 (1280x720; 4 fps)	25,331 (1280x720; 8 fps)	87,271
HoBiHu	48,683	6,580	55,263
Total	315435	125329	440764

- Total number of video clips : 440764

Historical Video Clips - Video Size

- Data Generated (size : MB)

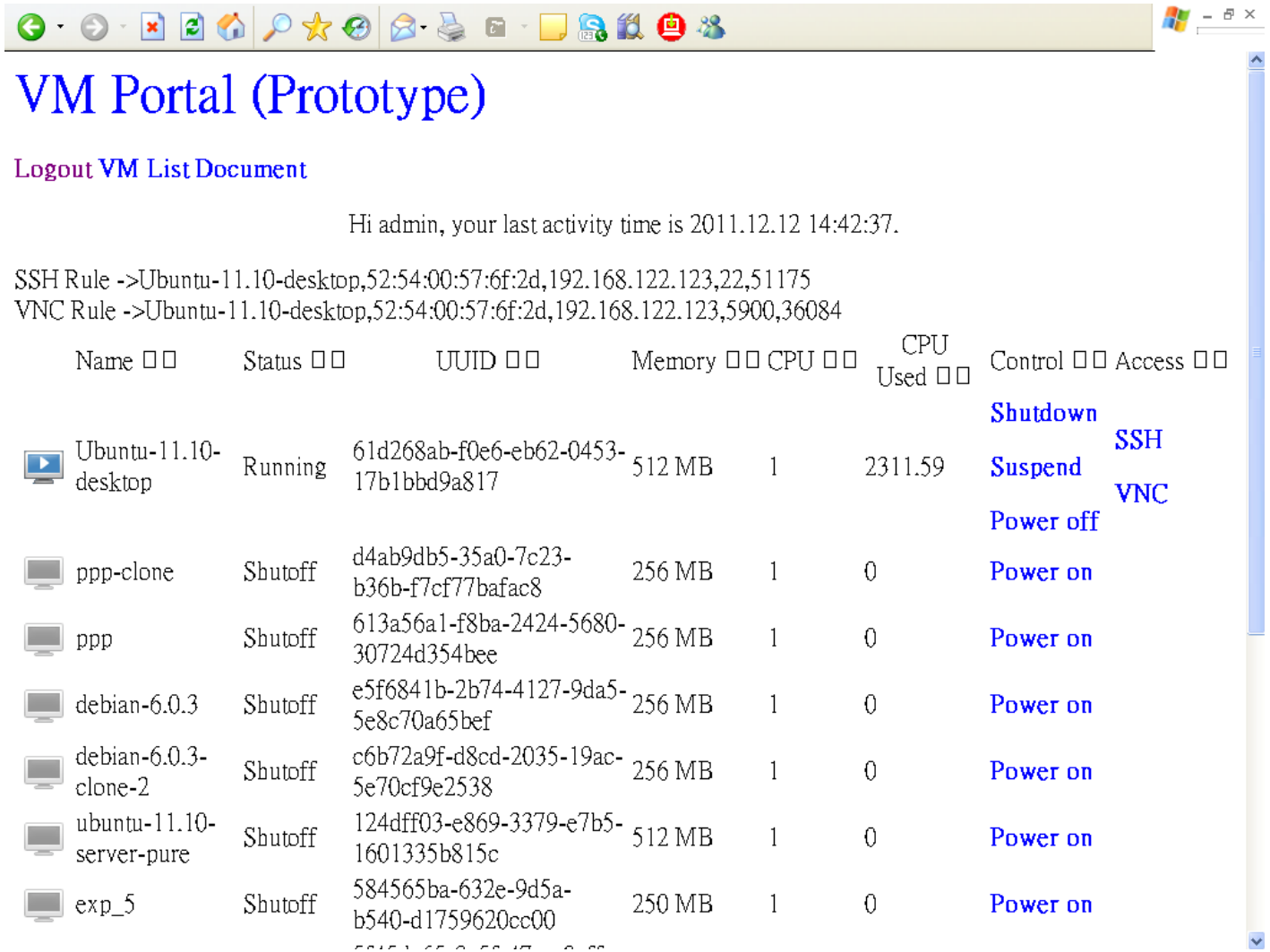
Type / Site	Old Format Resolution: 320x240 Fps: 8	New Format Resolution: 640x480 Fps: 24	Total Data Size per sites
NPP-3	2,965,218 MB	1,337,821 MB	4,303,039 MB
LanYu	1,394,357 MB	906,978 MB	2,301,335 MB
NMMBA	2,252,216 MB	1,483,641 MB	3,735,857 MB
HoBiHu	1,085,636 MB	437,513 MB	1,523,149 MB
Total	7,697,427 MB	4,165,953 MB	11,863,380 MB

- Total size of data : 11,863,380 MB ≈ 11,585 GB ≈ 11.31 TB

Description of Work

- T4.1
  - Enhance current video capturing and storage
- T4.2
  - Build data Storage facility.
- **T4.3**
  - **Develop data process execution interfaces**
- T4.4
  - Develop distributed data and computational methods.
- T4.5
  - Support parallelisation.

Process execution interfaces: VM Portal



VM Portal (Prototype)

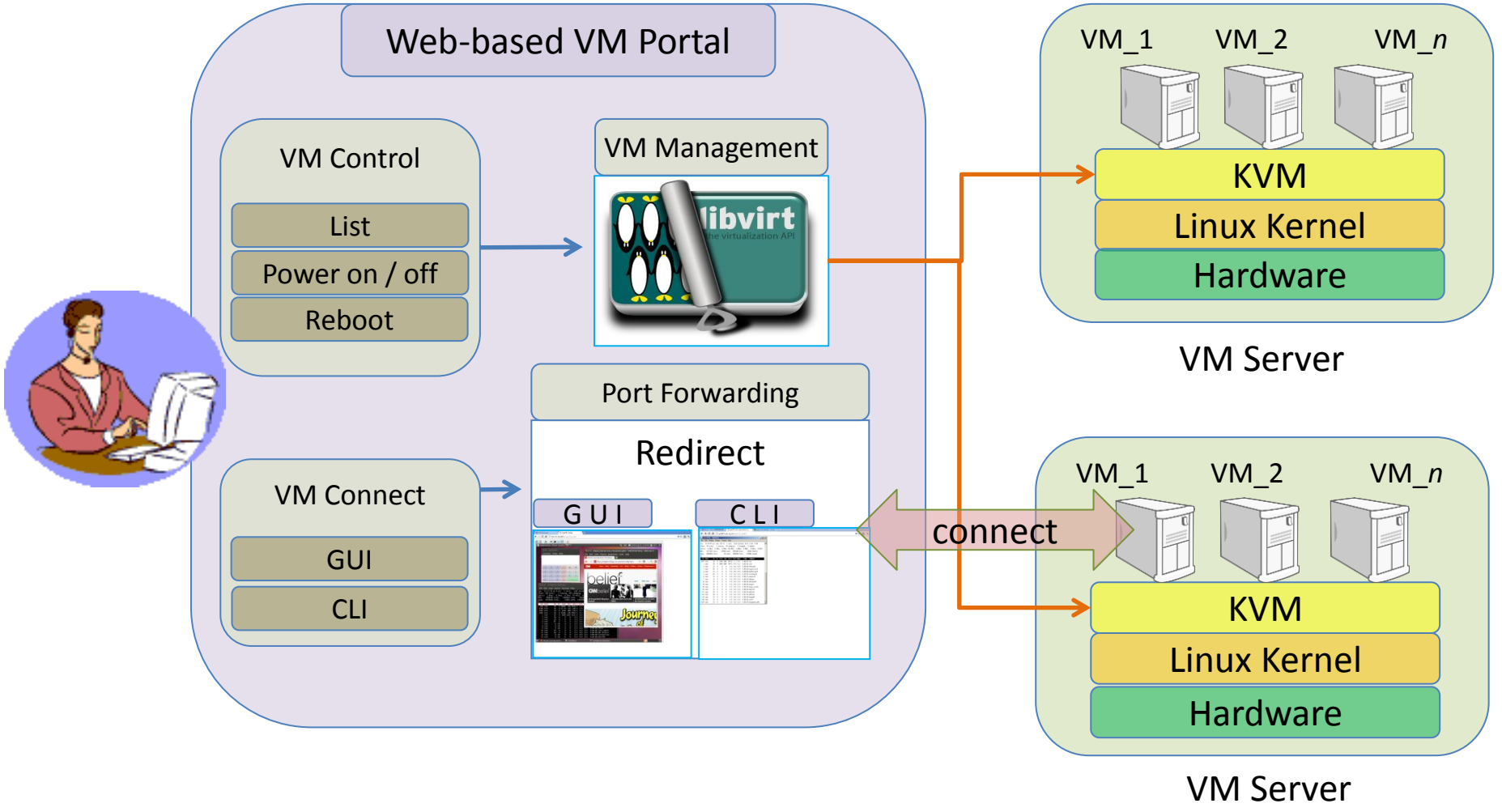
Logout VM List Document

Hi admin, your last activity time is 2011.12.12 14:42:37.

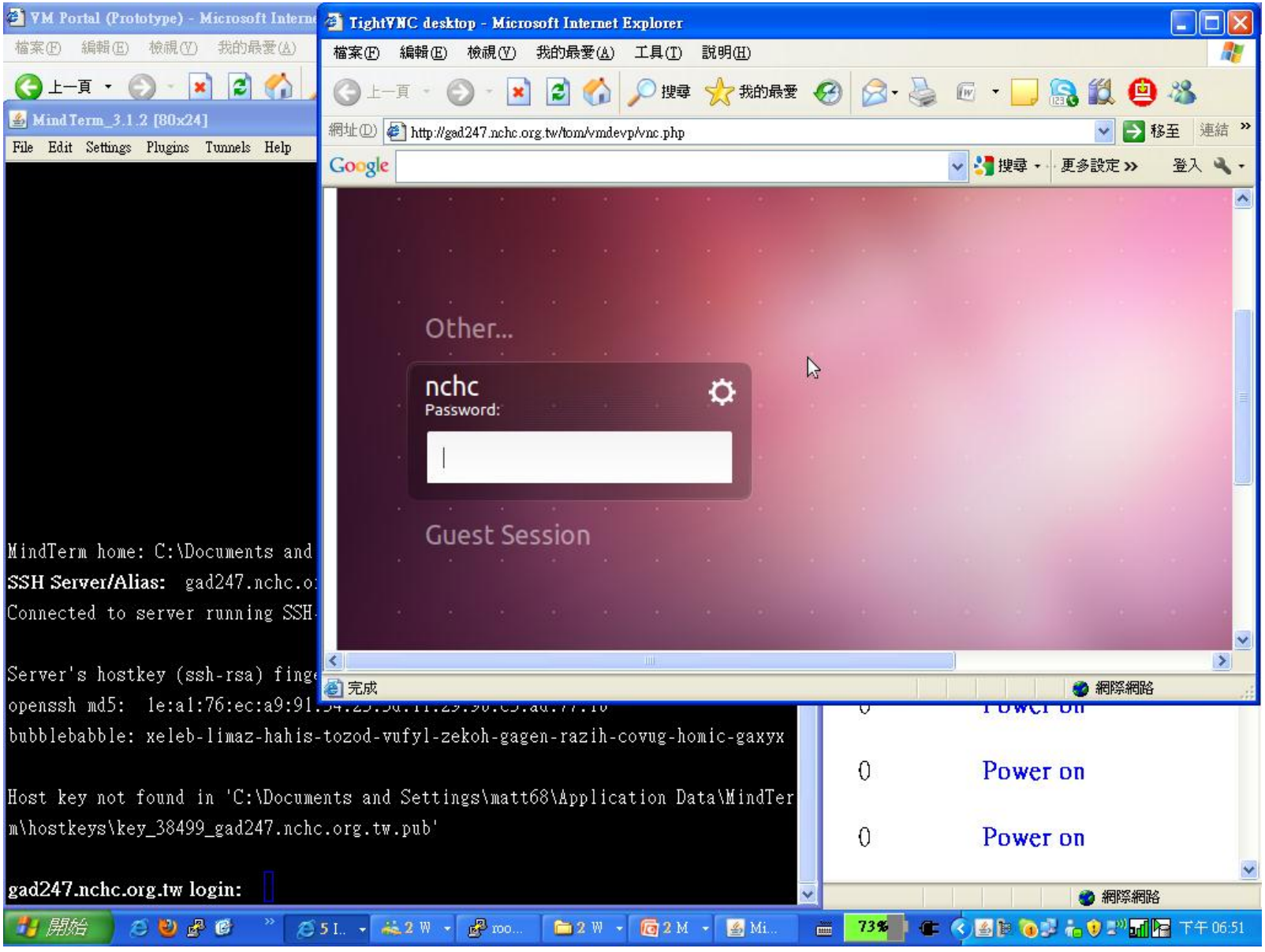
SSH Rule ->Ubuntu-11.10-desktop,52:54:00:57:6f:2d,192.168.122.123,22,51175
VNC Rule ->Ubuntu-11.10-desktop,52:54:00:57:6f:2d,192.168.122.123,5900,36084

Name □□	Status □□	UUID □□	Memory □□	CPU □□	CPU Used □□	Control □□	Access □□
Ubuntu-11.10-desktop	Running	61d268ab-f0e6-eb62-0453-17b1bbd9a817	512 MB	1	2311.59	Shutdown Suspend Power off	SSH VNC
ppp-clone	Shutoff	d4ab9db5-35a0-7c23-b36b-f7cf77bafac8	256 MB	1	0	Power on	
ppp	Shutoff	613a56a1-f8ba-2424-5680-30724d354bee	256 MB	1	0	Power on	
debian-6.0.3	Shutoff	e5f6841b-2b74-4127-9da5-5e8c70a65bef	256 MB	1	0	Power on	
debian-6.0.3-clone-2	Shutoff	c6b72a9f-d8cd-2035-19ac-5e70cf9e2538	256 MB	1	0	Power on	
ubuntu-11.10-server-pure	Shutoff	124dff03-e869-3379-e7b5-1601335b815c	512 MB	1	0	Power on	
exp_5	Shutoff	584565ba-632e-9d5a-b540-d1759620cc00	250 MB	1	0	Power on	

Virtual Machine Portal



VM WebPortal



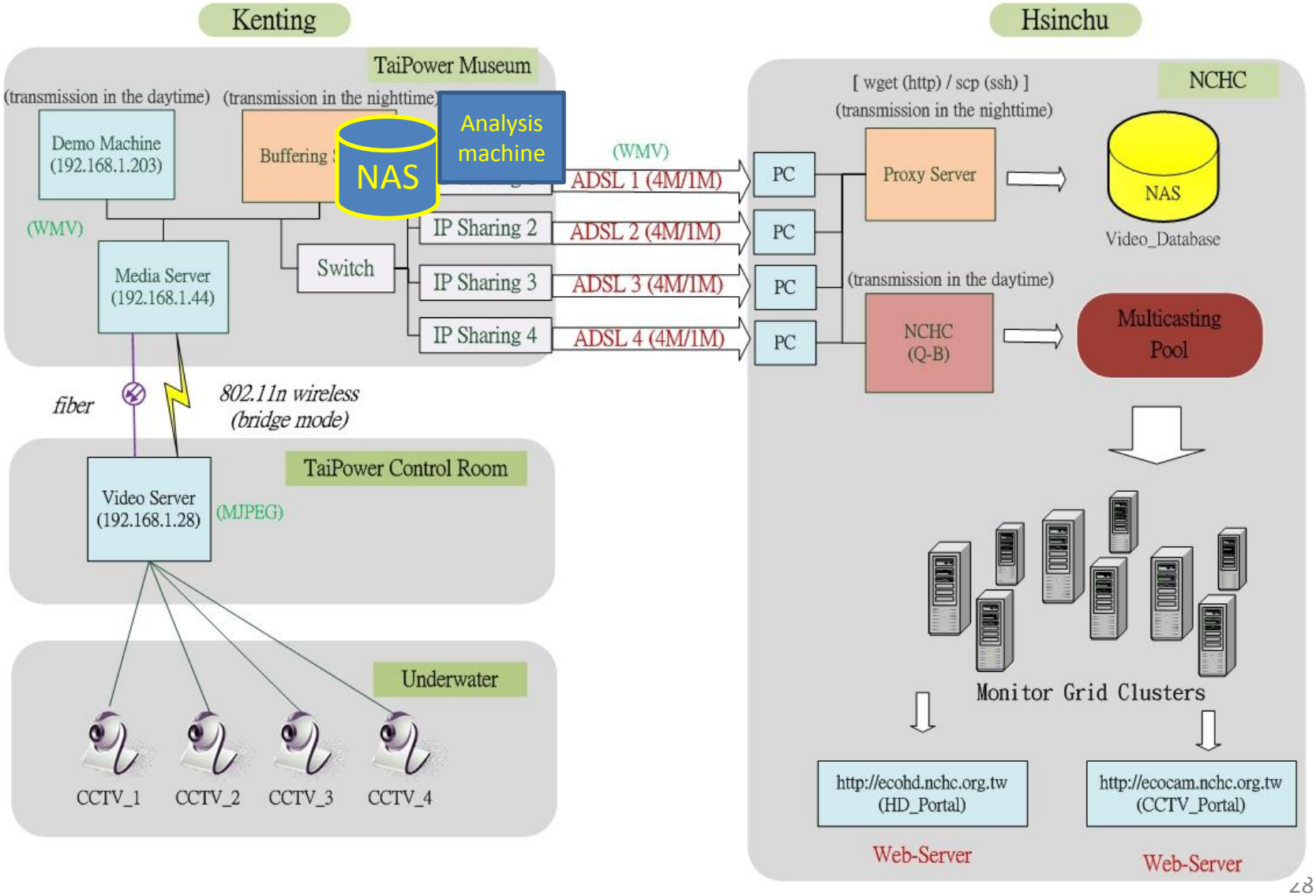
Description of Work

- T4.1
  - Enhance current video capturing and storage
- T4.2
  - Build data Storage facility.
- T4.3
  - Develop data process execution interfaces
- **T4.4**
  - **Develop distributed data and computational methods.**
- T4.5
  - Support parallelisation.

Methodologies investigated (on-going)

- Concurrency (multi-threading)
- Locality (in-situ computing)
  - New machine (8 core/16GRAM/8TB Storage) moved to Kenting and software installed.
- Multi-core computing (Share memory)
- GPU Computing (Accelerators)
  - Tested on Monte Carlo Sampling in random numbers & Large scale Algebra Calculations.

Ex. NPP-3 site architecture (explore locality)



Summary of Main Achievement in Year 1

- Main project machine up and running
- Virtual machine execution environment
- > 6000 hours video captured
- Exploring different SQL options

Next Step for year 2

- **D4.3: (month 24) - Process execution**

- **Maintain quality video Data Provisioning**
  - Prepare upgrade to 1K CCTVs in NPP3 site and support more sophisticated installation.
  - Recover broken underwater cables in Lanyu and Hobihu.
- **Enhance Compute Services**
  - Working with all partners on porting workflow and analysis modules and on all necessary tests.
  - Tuning performance for the modules w.r.t. handling large video dataset.
  - Provide user manual and spec documents
  - Provide training workshop in March.
- **Expand Data facility**
  - Dedicated storage totally ~60TB, expected to extend to ~100TB. We are now applying 500TB for 2013.
- **Developing distributed data and computational methods**
  - Following the 4 basic methodologies to increase the entire performance of the system.

http://ecocam.nchc.org.tw/

http://ecohd.nchc.org.tw/

http://ecocam.nchc.org.tw/tdw/

http://ecosite.nchc.org.tw/

http://gad240.nchc.org.tw/tai/video_query_new/

Thank You !!

Appendix: Compute server configuration, status, access, packages

WindRider (apls) - 48 cores Machine x 2

Host & Compute Node	
Machine Name	ALPS — Acer AR585 F1 Cluster
Machine Structure	SMP Cluster (2 nodes, 96 cores)
Processors	AMD Opteron 6174, 12 cores, 2.2GHz (compute nodes)
Main Memory (per node)	128 GB (compute nodes) 4 HT Links with “25.6 GB/s” bandwidth (the speed of each HT Links is “6.4GT/s”)
Operating System	Novell SuSE Linux Enterprise 11 SP1
Job Scheduler & Queuing System	Platform LSF (Load Sharing Facility) 7.06

Software Stack	
Installing Software	ABINIT, Amber, CASINO, CHARMM, Chemsoft, CPMD, DL_POLY, GAMESS, Gaussian, GROMACS, Molpro, NAMD, NWChem, octopus, OpenMX, Quantum ESPRESSO, siesta, VASP, WIEN2k
Development Tools	Intel Cluster Toolkit, PGI CDK/Server, x86 Open64, GCC
MPI/OpenMP	Platform MPI (formerly HP-MPI), Intel MPI, PGI MPI/OpenMP, MVAPICH/MVAPICH2
Math	Intel MKL (Math Kernel Library), AMD ACML (AMD Coe Math Library)

VM Server (gad246) - 48 cores Machine

Host & Compute Node	
Machine Name	**Gad246 — Tyan 8812F48W8HR**
Machine Structure	**SMP System**
Processors	**AMD Opteron 6176 SE, 12 cores, 2.3GHz**
Main Memory (per node)	**128 GB (compute nodes)** **4 HT Links with "25.6 GB/s" bandwidth (the speed of each HT Links is "6.4GT/s")**
Operating System	**Ubuntu 10.04.1 LTS**

PC Cluster (gad247) – 8 + 1 Cluster

Host & Compute Node	
Machine Name	Gad246 — ASUS MD-710
Machine Structure	SMP System
Processors	Intel core i7 2600, 4 cores, 3.4G Hz
Main Memory (per node)	4 GB (compute nodes)
Operating System	Ubuntu 10.04.1 LTS