Panoptes: cluster HA di monitoraggio della rete

Giuseppe Sava, INFN Sez. di Catania, Netgroup
Andrea Francesco Fornaia, INFN Sez. di Catania
Gianni Mario Ricciardi, Korea Institute of Science and Technology Information

Di cosa parliamo

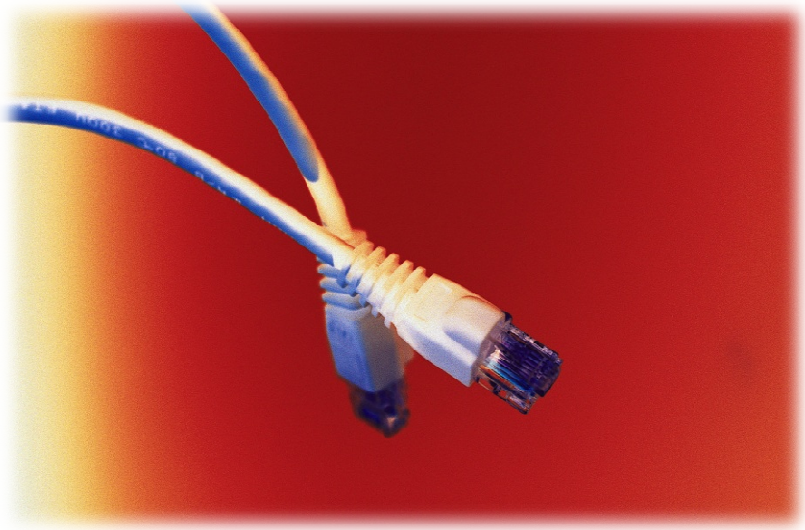
- Il network monitoring è il processo di controllo di apparati, host e servizi contenuti in una rete
- Un amministratore che ha il compito di monitorare una rete distribuita con differenti entità necessita di facility centralizzate per il monitoraggio
- Questo controllo permette agli amministratori di rete di mantenere una rete affidabile e ad alta qualità del servizio

Osserva, raccogli, agisci, notifica

- Un sistema di monitoraggio deve osservare costantemente lo stato della rete
- Raccogliere informazioni su eventi e cambiamenti di stato dei vari elementi
- Agire in maniera autonoma, se necessario, per ripristinare lo stato ottimale del sistema
- Notificare gli amministratori in caso di anomalie

La rete da monitorare

- LAN dipartimentale:
  - Circa 30 apparati di rete, Data e VoIP
  - Circa 20 servizi
  - 3 armadi di piano
  - Sala macchine



Non software ma architettura

- Non si parla semplicemente di "software" di monitoraggio ma di "architettura"
- Potrebbero essere necessari più applicativi per soddisfare a pieno le nostre esigenze
- Potrebbero essere necessari più servizi dislocati in vari punti della rete
  - Probe Acquiring

Qualè il problema

- Per poter effettuare un monitoraggio completo servono differenti programmi
- Ognuno di essi ha una propria GUI...
- Interfacce complete, ma non sempre intuitive...
- Molte informazioni...

... serve davvero tutto?

- The Swiss Knife
  - Molte funzionalità allettanti
  - Superflue nella maggior parte dei casi
  - Ma un giorno potrebbero servirmi...

Selezione & Aggregazione

- Si potrebbe invece continuare ad utilizzare i sistemi già installati, configurati e testati
- Selezionare le pagine di utilizzo frequente
- Aggregare le informazioni in un'unica interfaccia

Personalizzare & Estendere

- Creare un'interfaccia consona alle proprie esigenze e al proprio "modus operandi"
- Fare in modo che, all'occorrenza, il sistema possa essere esteso con nuove funzionalità
  - Aggiunta nuovi applicativi
  - Selezione di differenti informazioni
  - Riorganizzazione delle informazioni precedenti

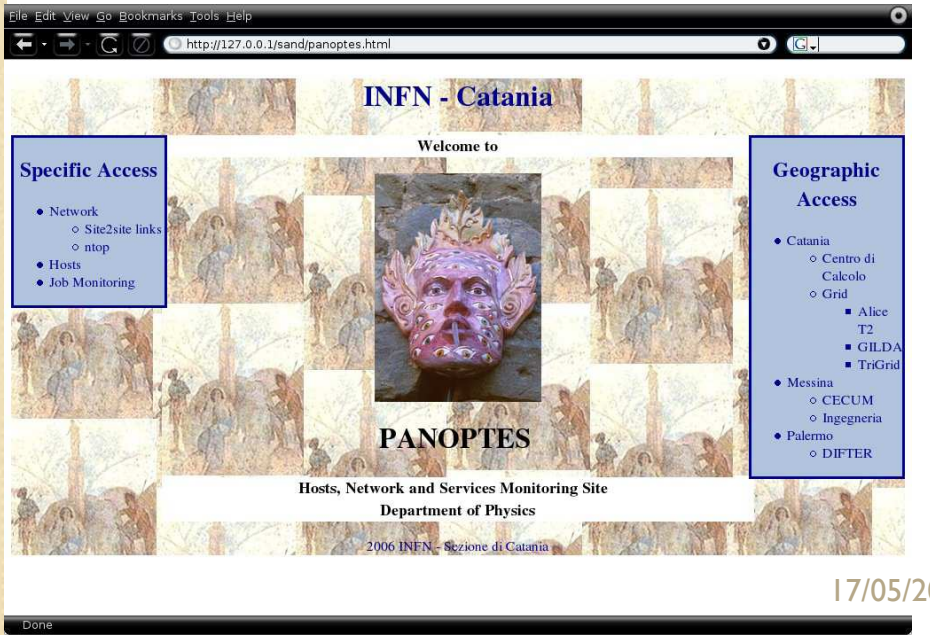
Panoptes

- Architettura basata su sistemi open source per il monitoraggio della rete e notifica di failure
- consente agli amministratori di avere in tempo reale tutte le informazioni utili, mediante un'interfaccia web all-in-one-page
- Panoptes è stato inoltre progettato per inoltrare notifiche di failure anche in assenza di connettività di rete mediante un sms-gateway
- La costante disponibilità del servizio di monitoraggio e notifica è garantita dall'architettura cluster in HA

Home Page Panoptes v.2

La storia del sistema

Giano
Basato su NetSaint
2001



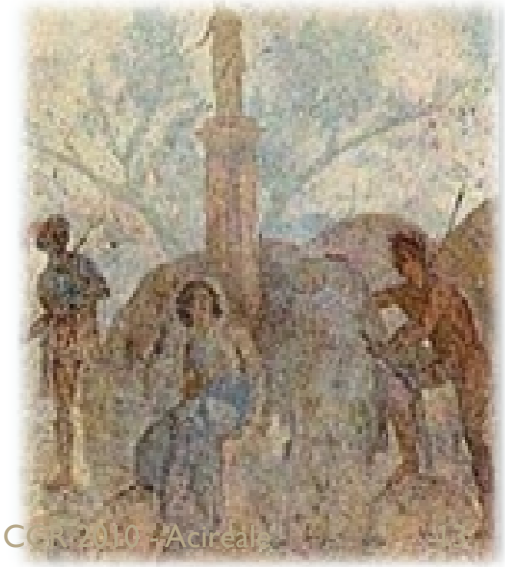
Panoptes v. 1
Versione precedente
del sistema attuale
2007

Argus Panoptes

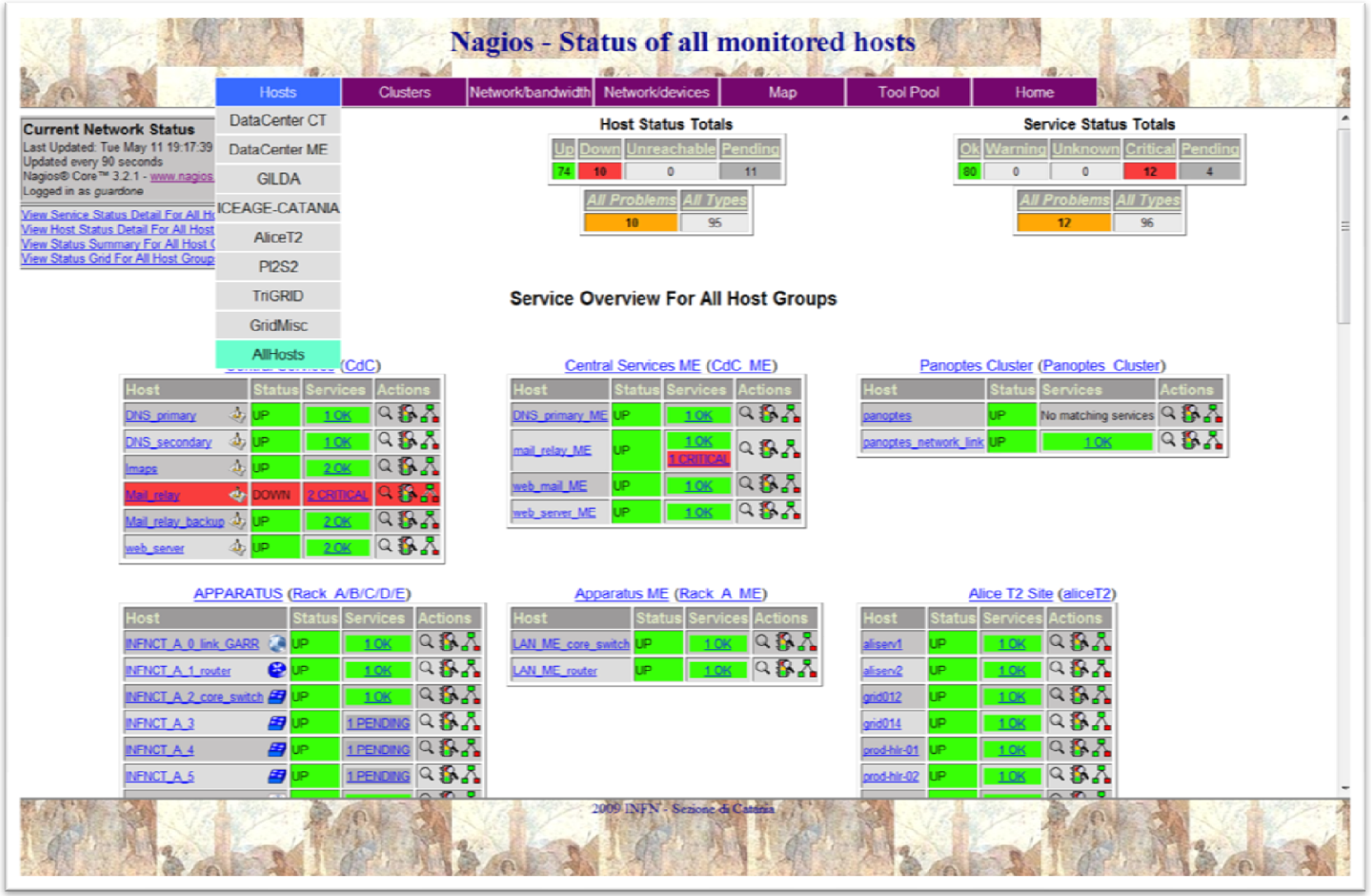
Nella mitologia greca, Argus Panoptes
era un gigante con 100 occhi
che dormiva tenendone chiusi solo alcuni (grande esempio di capacità di monitoraggio !).

Secondo il mito, Argo era stato inviato a sorvegliare Io.
Zeus, per liberarla, mandò Ermes.

In uno splendido affresco del Palatino, a Roma, sono raffigurati Ermes, Io e Argo.



II Layout All-in-one-page



Make it easy!

- Panoptes non si propone come un nuovo programma di monitoraggio, ma come un sistema di aggregazione e selezione di informazioni ottenute da differenti prodotti open source
- L'interfaccia web all-in-one-page permette di raggiungere le funzionalità più importanti, rendendo più semplice l'utilizzo quotidiano

Make it fast!

- Attraverso un drop-down menù multilivello, sempre visibile, è possibile raggiungere le varie sezioni del portale
- Il contenuto delle pagine viene visualizzato al di sotto del menù
- Solo scroll delle pagine interne

Make it personal!

- Si possono aggiungere e rimuovere facilmente nuove pagine al portale
- Il sistema è pensato per essere dinamico, in continua evoluzione

Tipi di strumenti

- Overview
  - Attraverso grafici prestazionali e di stato della rete si può avere rapidamente informazioni sullo lo stato della rete
- Notifica
  - In caso di anomalie o di eventi rilevanti, gli amministratri vengono notificati via e-mail e sms
- Indagine
  - Attraverso pagine contenenti informazioni più dettagliate o utilizzando direttamente le GUI dei sistemi utilizzati (Tool Pool)

Nagios: il cuore di Panoptes

- Software Open Source, rilasciato sotto licenza GPL, per il monitoraggio di computer e risorse di rete
- Versione corrente : 3.2.1. (9 Marzo 2010)
- Sito ufficiale: www.nagios.org

Di che si tratta

- Consiste in un daemon che a intervalli prestabiliti esegue controlli su host e sui servizi specificati usando plugins esterni
- Se vengono riscontrate anomalie il daemon può mandare notifiche in diverse forme (email, sms, etc.) ai contatti amministrativi o mandare in esecuzione routine di gestione
- La struttura a plugin rende il sistema altamente personalizzabile
- Lo storico delle notifiche e degli eventi può essere consultato via interfaccia web

Gli oggetti principali: Host

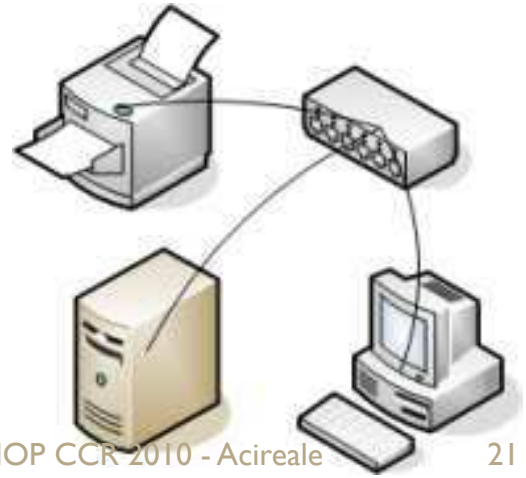
Attraverso i file di configurazione l'amministratore fornisce al sistema le informazioni necessarie, definendo i seguenti oggetti principali:

- Host
  - Rappresentano tipicamente host fisici o virtuali, stampanti di rete,

    apparati..
  - Hanno un indirizzo associato
  - Hanno uno o più servizi

    associati
- Host Group
  - Aggregazione di più host
  - Semplificano la visualizzazione

    nell'interfaccia web

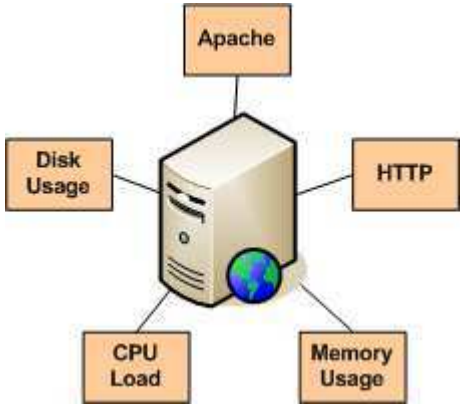


Esempio definizione di un host

```
define hostgroup{
    hostgroup_name              Panoptes_Cluster
    alias                       Panoptes Cluster
    members                     panoptes, ha-324, ha-325
}
define host{
        use                     cdc-host-CRIT
        host_name               panoptes
        alias                   panoptes cluster
       contact_groups  cdc_admins_sms, cdc_admins_email
        address                 192.167.0.229
        parents                 INFNCT_A_1_router
        }
```

Gli oggetti principali: Service

- Service
  - Costituiscono la parte centrale del sistema di monitoraggio
  - Possono riferirsi ad attributi interni dell'host (CPU load, disk usage...)
  - Possono riferirsi a servizi forniti (HTTP, SSH...)
  - Tutto ciò che si vuole!
  - Ad ogni servizio corrisponde un comando di check, per determinarne lo stato
- Service Group
  - Aggregazione di più servizi



Esempio definizione di un service

```
define service{
        use                             cdc-service-CRIT
        hostgroup_name                  Panoptes_Cluster
        service_description             PING
        normal_check_interval           3
        retry_check_interval            1
        max_check_attempts              4
        check_command
          check_ping!100.0,20%!500.0,60%
        }
```

Gli oggetti principali: Contact

- Contact
  - Identificano un contatto da notificare in caso di determinati eventi (alto livello di personalizzazione) e quando è possibile inoltrare notifiche
  - Si specifica il metodo di notifica (e-mail, SMS... Indicando il relativo comando)
  - É possibile associare un'utenza sul sistema ad un contatto
- Contact Group
  - Aggregazione di contatti
  - Gli host e servizi indicano come destinatario delle notifiche il gruppo e non il singolo contatto

Esempio definizione di un contact

```
define contact{
        contact_name                    afornaia_email
        use                             generic_contact
        alias                           Andrea Francesco Fornaia
        email                           fornaia.andrea@libero.it
        service_notification_period     24x7
        host_notification_period        24x7
        service_notification_options    w,u,c,r,f,s
        host_notification_options       d,u,r,f,s
        service_notification_commands   notify-service-by-email
        host_notification_commands      notify-host-by-email
        }
```

Esempio definizione di un contactgroup

```
define contactgroup{
        contactgroup_name       cdc_admins_email
        alias                   CdC Administrators (EMAIL)
        members                 afornaia_email, gsava_email
        }
```

Gli oggetti principali: Command

- Command
  - È il metodo con cui Nagios effettua azioni sugli elementi del sistema
  - Ad ogni comando è associato un programma, script... tipicamente chiamato Plugin
    - sono facili da realizzare, basta attenersi a delle interfacce di output per poter permettere a Nagios di interpretarne il risultato

Esempio definizione di un command

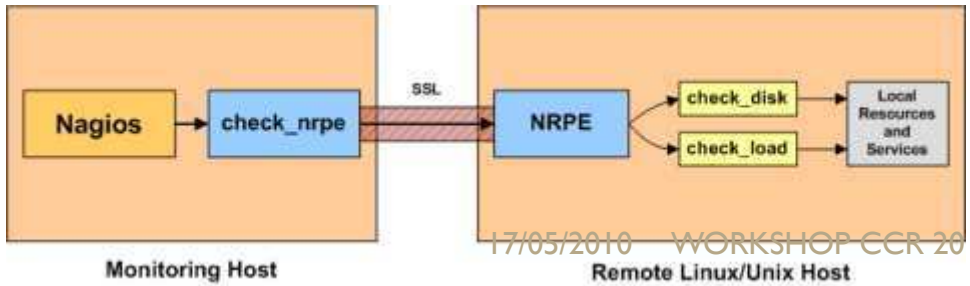
```
define command{
        command_name    notify-service-by-sms
        command_line    /usr/local/bin/sendsinglesms
   $CONTACTPAGER$ "PANOPTES:
   $NOTIFICATIONTYPE$ with $SERVICEDESC$ on
   $HOSTALIAS$:$HOSTADDRESS$: state
   $SERVICESTATE$ in $LONGDATETIME$">/dev/null
      }
```

Check attivo e Check passivo

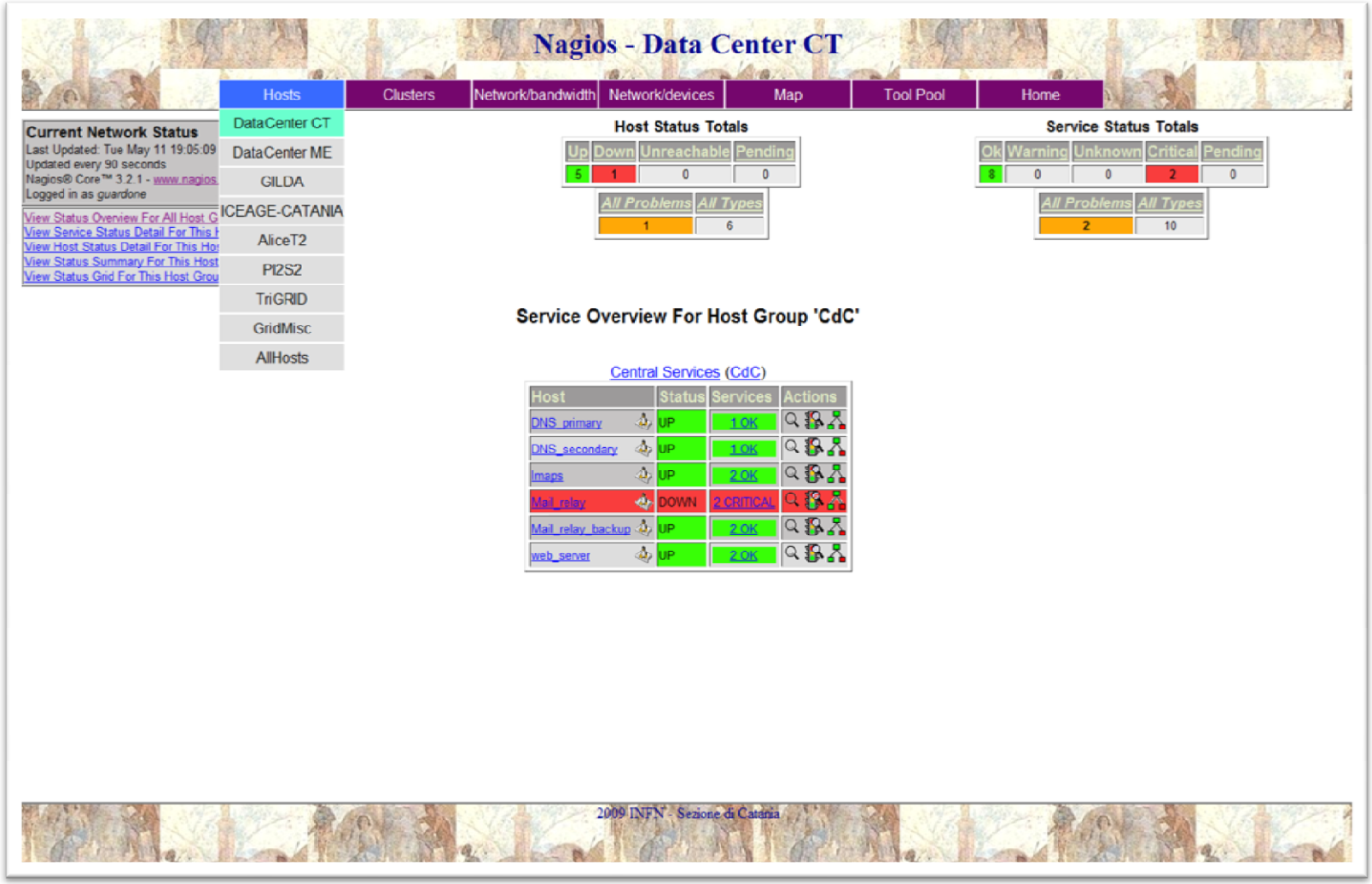
- Nel caso di check attivo, servizi e host vengono monitorati dal server Nagios mandando in esecuzione, a intervalli regolari i comandi di check indicati
- Nel caso di check passivo, saranno altre applicazioni a scrivere, in maniera asincrona, il risultato di un check su di un file di spool, che Nagios controllerà a intervalli regolari

NRPE: Nagios Remote Plugin Executor

- Tipicamente i check vengono dal server, avendo quindi solo informazioni esterne sullo stato di host e servizi
- Attraverso NRPE, un Addon per Nagios, è possibile inserire delle probe negli host per eseguire i check localmente, rendendo disponibili informazioni interne all'host (disk usage, process number...)



Nagios in Panoptes



Dettaglio sulla nomenclatura

APPARATUS (Rack_A/B/C/D/E)

Host	Status	Services	Actions
INFNCT_A_0_link_GARR	UP	1 OK	
INFNCT_A_1_router	UP	1 OK	
INFNCT_A_2_core_switch	UP	1 OK	
INFNCT_A_3	UP	1 OK	
INFNCT_A_4	UP	1 OK	
INFNCT_A_5	UP	1 OK	
INFNCT_A_VOIP_1	UP	1 OK	
INFNCT_A_VOIP_2	UP	1 OK	
INFNCT_A_VOIP_6509	UP	1 OK	
INFNCT_A_VOIP_CCM	UP	1 OK	
INFNCT_A_VOIP_INFN-CUCM4-3	UP	1 OK	
INFNCT_A_VOIP_Orion	UP	1 OK	
INFNCT_A_VOIP_Unity	UP	1 OK	
INFNCT_A_VOIP_VG248	UP	1 OK	
INFNCT_A_VOIP_router_3640	UP	1 OK	
INFNCT_B_1_stack	UP	1 OK	

- Tutti gli apparati in un unico hostgroup per facilitarne la visualizzazione
- Nomenclatura per la rapida collocazione dell'apparato
  - Armadio
  - Altezza
  - Tipo
  - Nome significativo

Le sezioni di Nagios in Panoptes

- Hosts: è possibile visionare lo stato degli host, e servizi associati, suddivisi per hostgroup, o tutti nella sezione All
- Network/devices: è possibile visionare lo stato degli apparati di rete, Data e VoIP
- Map: rappresentazione grafica dello stato della rete, rispettando le notazioni gerarchiche degli apparati e degl host

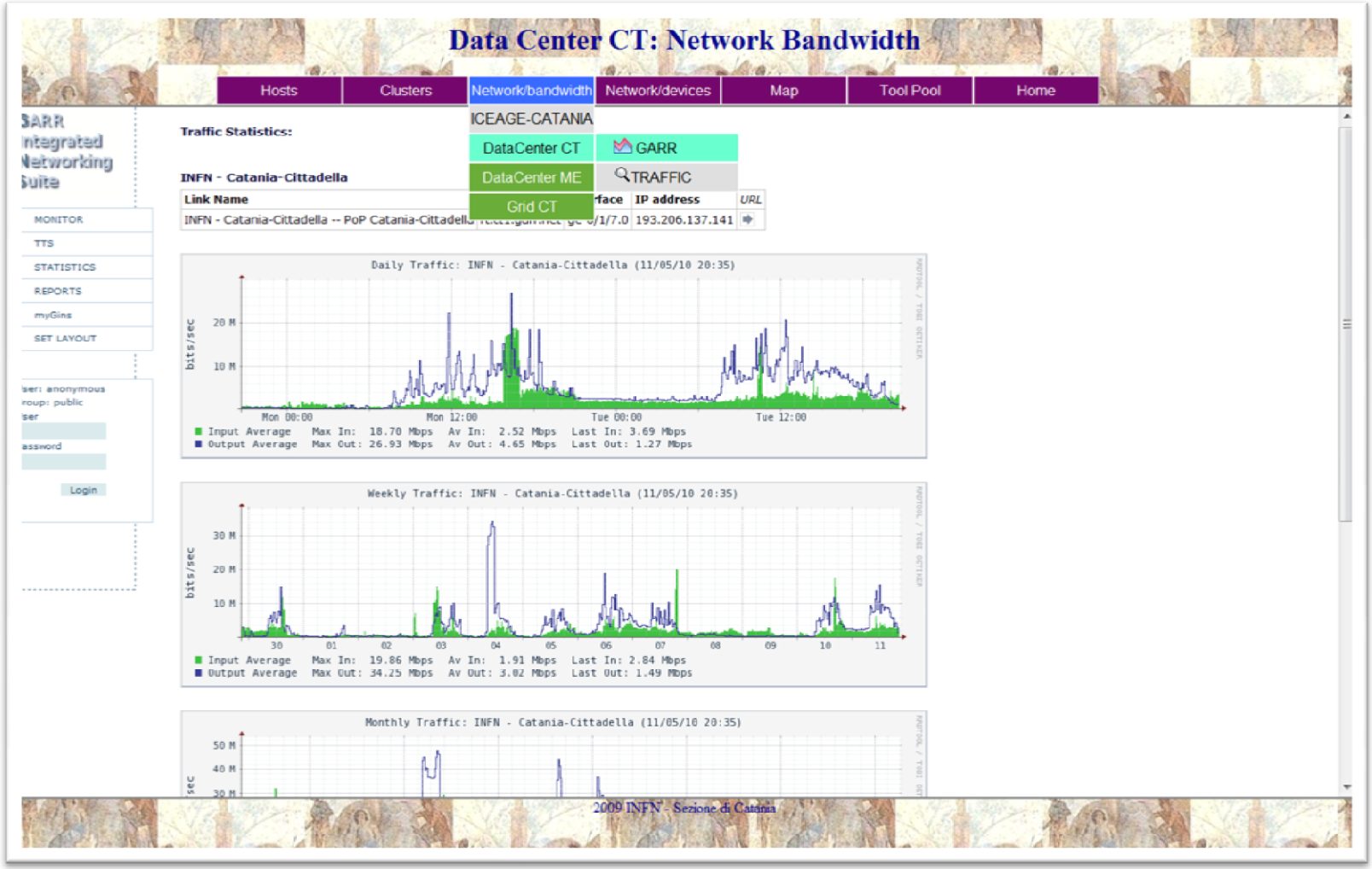
Non solo Nagios: GARR GINS

- GARR Inegrated Network Suite: statistiche sull'utilizzo di banda nel tempo fornite dal Provider del Dipartimento
- https://www.noc.garr.it/GINS/
- Utilizzato come strumento di monitoraggio e di indagine

GARR Integrated Networking Suite

Owned by: sw.dev@garr.it

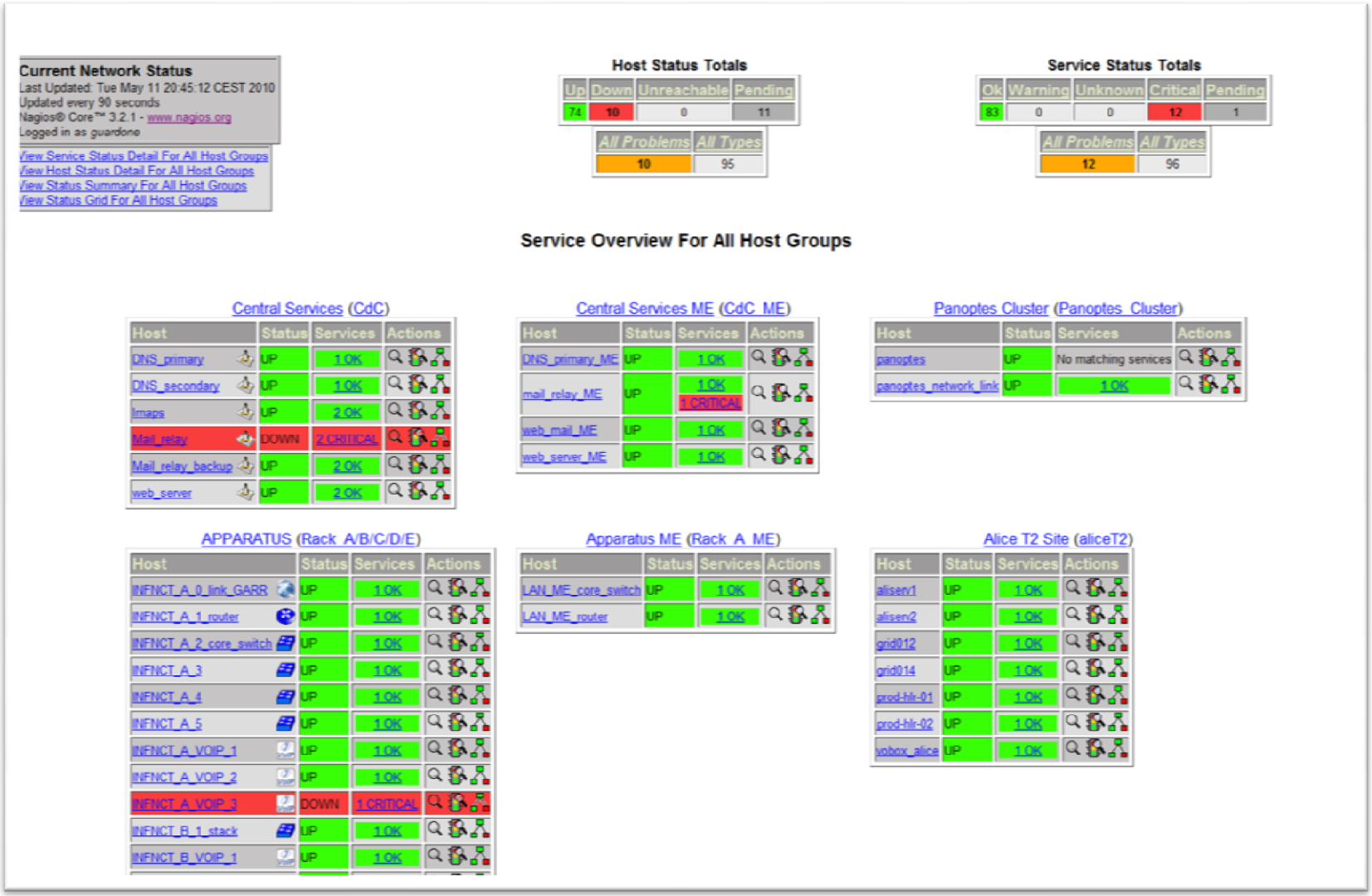
Screenshot di GINS in Panoptes



Non solo Nagios: Eagleye

- È stata realizzato un semplice applicativo web in grado di prendere in input l'output di Nagios relativo alla visualizzazione dello stato di tutti gli host e presentarlo a schermo effettuando l'auto-scroll della pagina se troppo lunga, in maniera continua
- È possibile settare velocità, quantità di incremento in pixel, e la sorgente
- Utilizzato attualmente al Centro di Calcolo, l'output viene mostrato su di uno schermo di monitoraggio sempre visibile agli amministratori

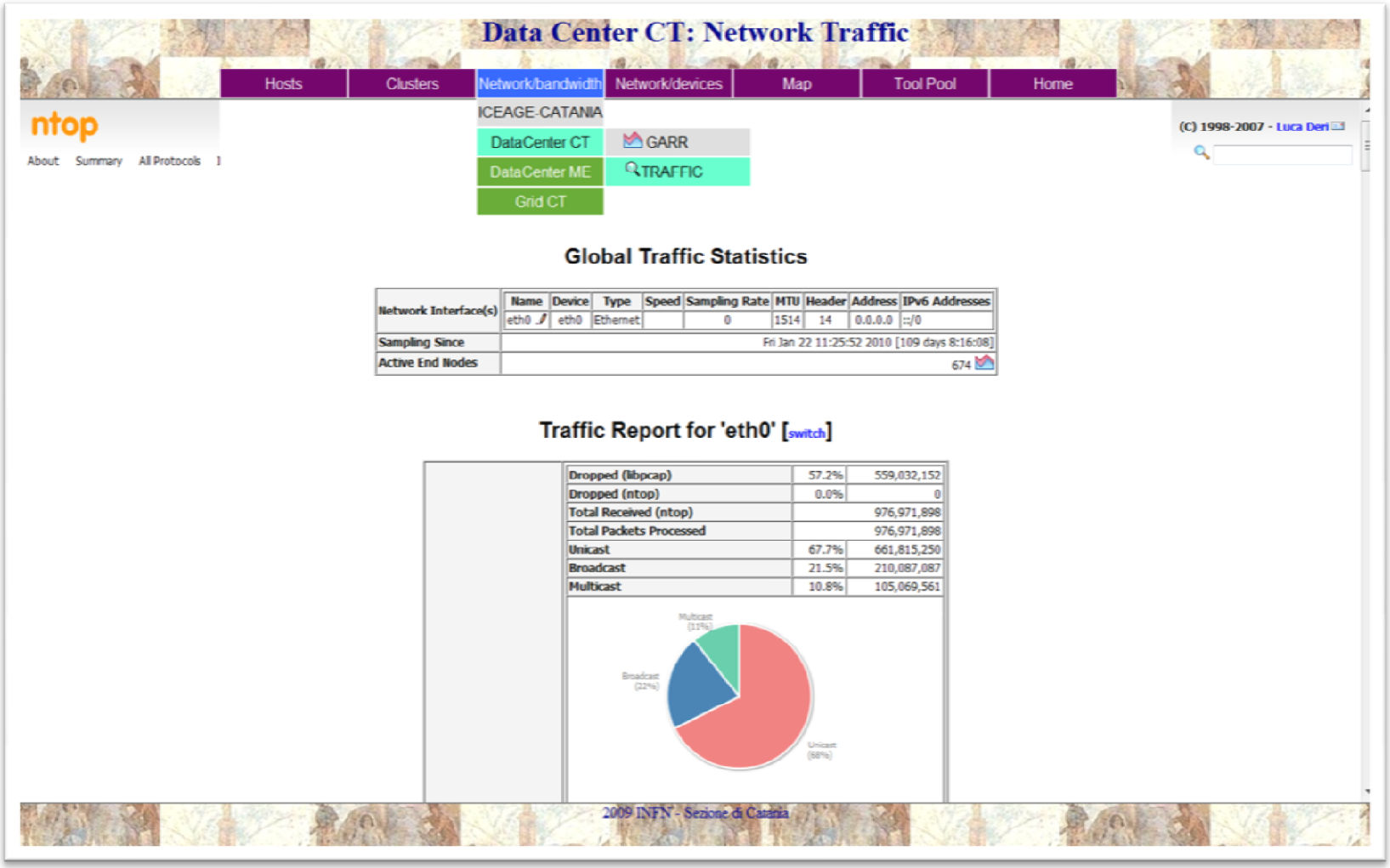
Screenshot di Eagleye



Non solo Nagios: Ntop

- Network traffic probe per il monitoraggio del traffico di rete
- Fornisce una web-interface
- Informazioni sul traffico suddiviso per protocolli
- Storico e Statistiche
- Generazione di grafici

Screenshot di ntop in Panoptes



Non solo Nagios: LogZilla

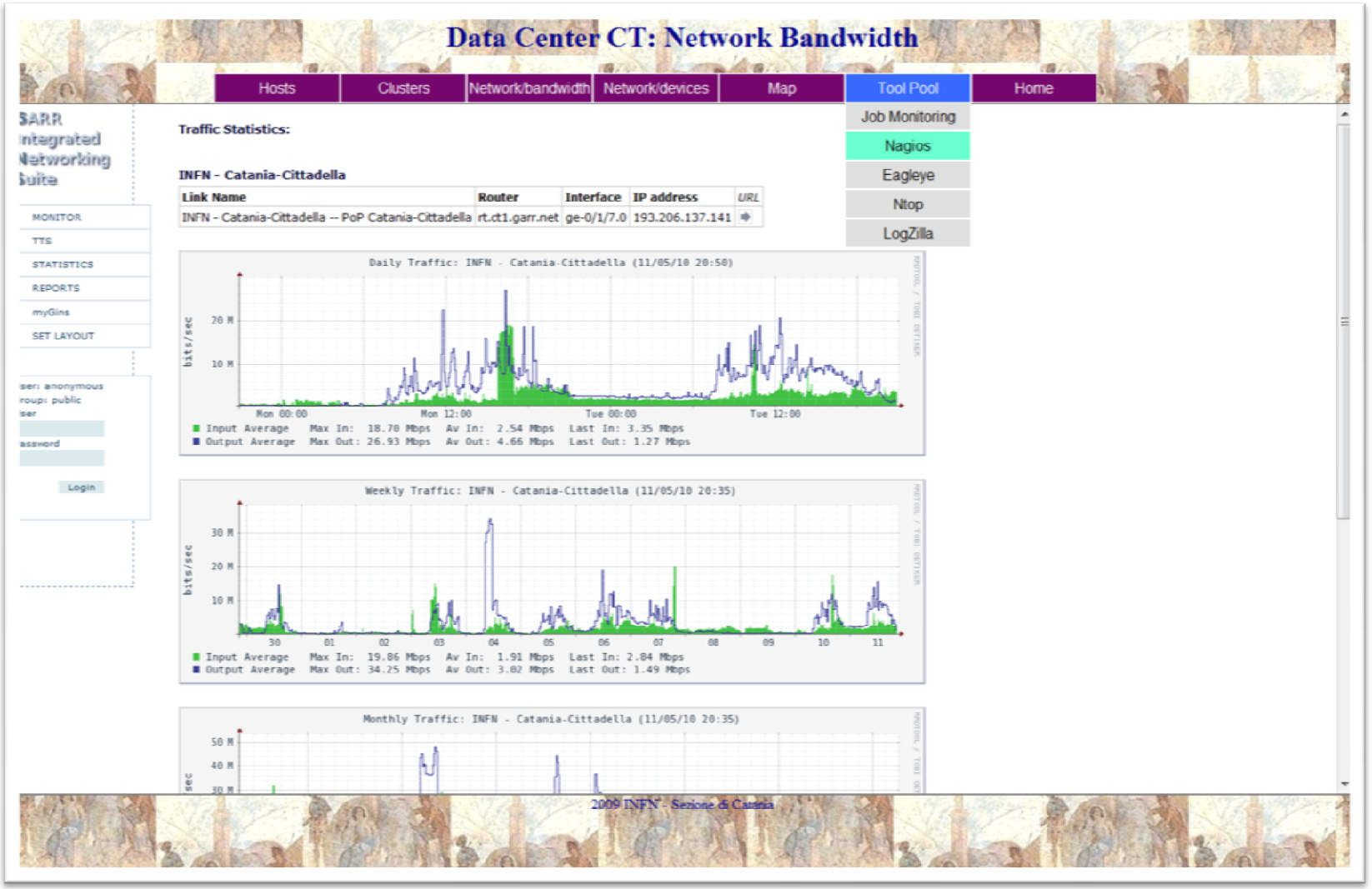
- logserver
- Dispone di un web front-end per effettuare query all'interno di un data-base MySQL che funge da storage per i messaggi del syslog-ng server
- Log provenienti da apparati di rete e servizi
- Maschere di ricerca dettagliate
- Elaborazione di grafica dei dati

Screenshot di LogZilla

Tool Pool

- Dalla sezione Tool Pool è possibile raggiungere le home degli applicativi di cui è composto Panoptes
- Panoptes è utile per avere una visione rapida dello stato della rete ma in fase di indagine sull'origine di un problema può essere utile utilizzare specifiche funzionalità dei software usati oppure di usare software dedicati (auditing, traffic monitoring...)
- Mantenere collegamenti ad utility

Screenshot del Tool Pool



Gestione delle Notifiche

- Non tutti gli eventi hanno la stessa rilevanza
- Al momento sono previsti due livelli di criticità
- Gli eventi "NORM" vengono notificati solo tramite e-mail (es. down su di un servizio di backup)
- Gli eventi "CRIT" vengono notificati sia tramite e-mail che sms (es. Down sulla connessione con la rete esterna)

Perché notificare con SMS

- Raggiungono direttamente l'amministratore
- Se ci si appoggia ad un gateway esterno non vi sono altri vantaggi
  - in caso di failure sulla connettività della INTRANET con INTERNET sarà impossibile inoltrare notifiche sia via e-mail che sms

Perché costruire un Gateway SMS interno

- Permette di appoggiarsi direttamente alla rete wireless GSM, indipendente dalla connessione di rete wired
- Unico modo per notificare problemi di connettività con l'esterno
- Più economico, se non si ha la necessità di Bulk sms
- Effettivo HA del servizio di notifica

Come realizzare un Gateway SMS

- Occorrente
  - Modem GSM
  - Gateway SMS Server
  - SMS Client
- Come macchina server è stato scelto di utilizzare la stessa macchina che si occupa del monitoraggio della rete

Modem GSM

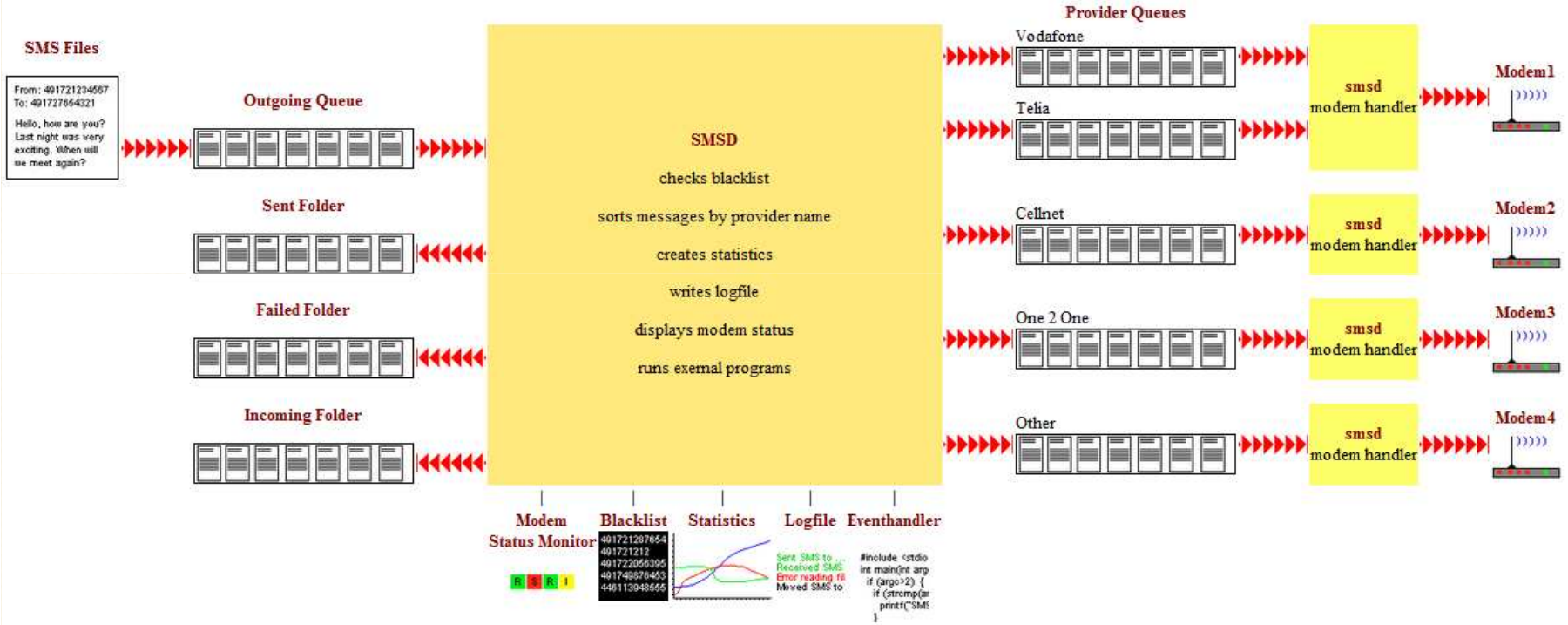
- Ve ne sono molti in commercio, più o meno professionali, anche per soluzioni industriali
- In realtà e sufficiente un cellulare GSM
- Oppure una "Internet key"
  - Se connessi ad un PC vengono visti come sistemi di archiviazione di massa
  - I produttori forniscono il software necessario per effettuare lo switching da storage device a modem device
  - Sottoutilizzati, non sono semplici modem GSM
  - Utilizzabili con i comuni comandi AT per la gestione dei Modem GSM

Gateway SMS Server: SMS Server Tools 3

- È stato usato SMS Server Tools 3
  - Software Open Source rilasciato sotto licenza GPL
  - La versione attuale è la 3.1.8. (5 Maggio 2010)
  - Disponibile per Windows, GNU/Linux, Solaris, MacOS, FreeBSD
  - Sito di riferimento: http://smstools3.kekekasvi.com/

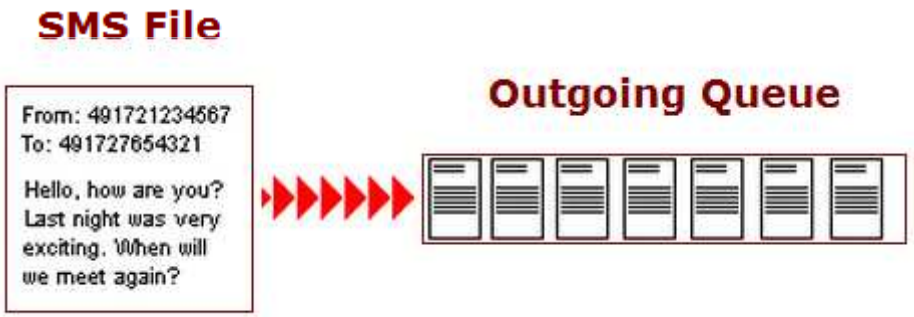
SMS Server Tools 3

Cenni sul funzionamento



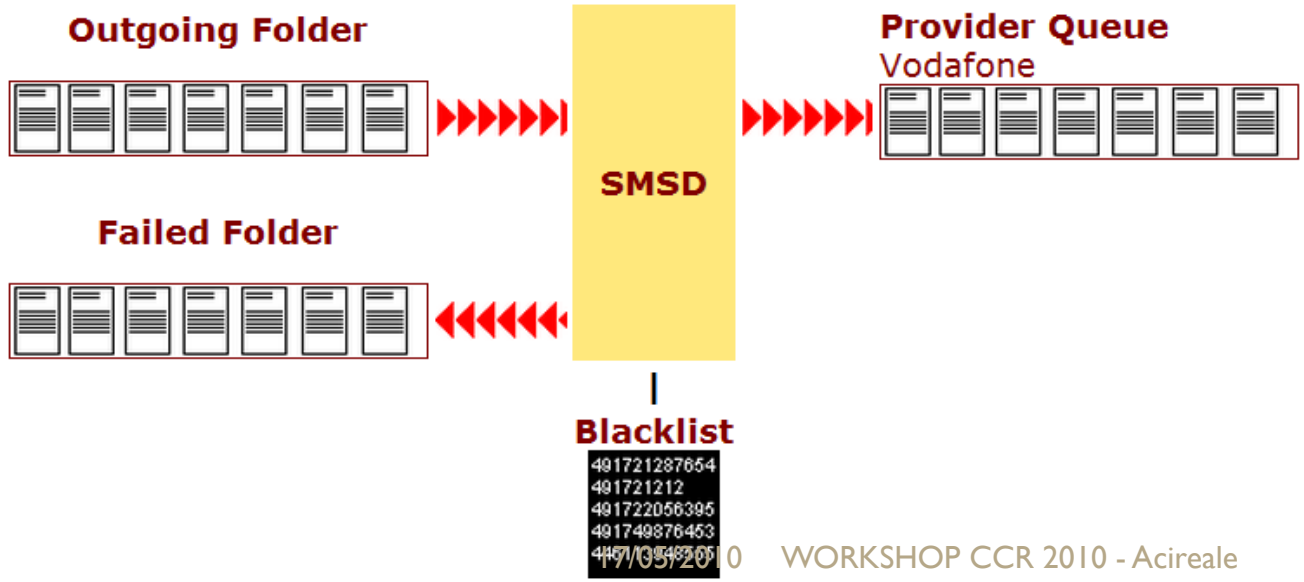
Cenni sul funzionamento

- Il client deve avere i permessi di scrittura nella cartella /var/spool/sms/outgoing



Cenni sul funzionamento

- il server di spool controlla periodicamente la coda di outgoing.
- Se trova file da inviare effettua la convalida dell'invio (es. black-list check)
- Se può essere inviato, lo sposta nella coda checked, altrimenti in failed
- Se è attivo il sistema di gestione delle code per provider (risparmio!) lo sposta nella coda appropriata



Configurazione dell'/etc/smsd.conf

```
devices = GSM1
sent = /var/spool/sms/sent
outgoing = /var/spool/sms/outgoing
checked = /var/spool/sms/checked
failed = /var/spool/sms/failed
logfile = /var/log/smsd.log
loglevel = 5

[GSM1]
device = /dev/ttyUSB1
#hardware handshake disabled to solve problem: "modem is not clear to send"
rtscts = no
#reduce baudrate to limit error probability (original br: 115200)
baudrate = 9600
incoming = no
#pin = 1111
```

SMS sender

```
#!/bin/sh
DEST=$1
TEXT=$2

if [ -z "$DEST" ]; then
  printf "Destination: "
  read DEST
fi

if [ -z "$TEXT" ]; then
  printf "Text: "
  read TEXT
fi

FILE=`mktemp /var/spool/sms/outgoing/sms_XXXXXX`
echo "To: $DEST" >> $FILE
echo "" >> $FILE
echo -n "$TEXT" >> $FILE
```

Notifiche SMS con Nagios

- Possibile soluzione
- Si crea per ogni contatto effettivo due contatti:
  - Specificando come metodo di notifica l'invio della e-mail (es. afornaia_email)
  - Specificando come metodo di notifica l'invio di un sms (es. afornaia_sms)
- Si creano due gruppi di notifica distinti
  - Contatti notificabili via e-mail (es. cdc_admins_email)
  - Contatti notificabili via SMS (es. cdc_admins_sms)
  - Fare uso dei template per semplificare la creazione dei contatti
- Gli eventi a bassa criticità notificheranno solo il gruppo cdc_admins_email
- Gli eventi ad alta criticità notificheranno sia il gruppo cdc_admins_email che cdc_admins_sms

Notifiche SMS con Nagios

```
define contactgroup{
        contactgroup_name       cdc_admins_sms
        alias                   CdC Administrators (SMS)
        members                 afornaia_sms, gsava_sms
        }

define contactgroup{
        contactgroup_name       cdc_admins_email
        alias                   CdC Administrators (EMAIL)
        members                 afornaia_email, gsava_email
        }
```

Notifiche SMS con Nagios

```
### Andrea Fornaia

define contact{
        contact_name                    afornaia_sms
        use                             notified_by_sms
        alias                           Andrea Francesco Fornaia
        pager                           +393331234567
        }

define contact{
        contact_name                    afornaia_email
        use                             notified_by_email
        alias                           Andrea Francesco Fornaia
        email                           fornaia.andrea@libero.it
        }
```

Notifiche SMS con Nagios

```
define contact{
        name                            notified_by_email
        use                             generic-contact
        service_notification_period     24x7
        host_notification_period        24x7
        service_notification_options    w,u,c,r,f,s
        host_notification_options       d,u,r,f,s
        service_notification_commands   notify-service-by-email
        host_notification_commands      notify-host-by-email
        register                        0
        }
```

Notifiche SMS con Nagios

```
define contact{
        name                            notified_by_sms
        use                             generic-contact
        service_notification_period     awaketime
        host_notification_period        awaketime
        service_notification_options    w,u,c,r,f,s
        host_notification_options       d,u,r,f,s
        service_notification_commands   notify-service-by-sms
        host_notification_commands      notify-host-by-sms
        register                        0
        }
```

Notifiche SMS con Nagios

```
define command{
        command_name    notify-service-by-email
        command_line    /usr/bin/printf "%b" "***** Nagios *****\n\nNotification Type:
    $NOTIFICATIONTYPE$\n\nService: $SERVICEDESC$\nHost: $HOSTALIAS$\nAddress:
    $HOSTADDRESS$\nState: $SERVICESTATE$\n\nDate/Time:
    $LONGDATETIME$\n\nAdditional Info:\n\n$SERVICEOUTPUT$" | /usr/bin/mail -s "**
    $NOTIFICATIONTYPE$ Service Alert: $HOSTALIAS$/$SERVICEDESC$ is
    $SERVICESTATE$ **" $CONTACTEMAIL$
        }

#notify-service-by-sms' command definition
define command{
        command_name    notify-service-by-sms
        command_line    /usr/local/bin/sendsinglesms $CONTACTPAGER$ "PANOPTES:
    $NOTIFICATIONTYPE$ with $SERVICEDESC$ on $HOSTALIAS$:$HOSTADDRESS$:
    state $SERVICESTATE$ in $LONGDATETIME$">/dev/null
        }
```

Notifiche SMS con Nagios

```
define host{
        name                    cdc-host-NORM
        use                     host-NORM
        contact_groups          cdc_admins_email
        register                0
}
define host{
        name                    cdc-host-CRIT
        use                     host-CRIT
        contact_groups          cdc_admins_sms, cdc_admins_email
        register                0
}
```

Notifiche SMS con Nagios

```
define host{
        use                     cdc-host-CRIT
        host_name               Mail_relay
        alias                   Mail Server Cdc
        address                 192.84.150.109
        parents                 INFNCT_A_1_router
        }
```

Gateway SMS per HA sul servizio di notifica… ma è sufficiente?

- Cosa succede se l' SMS Gateway "va in Crash"?
- Se succedesse al server di monitoraggio?
- Bisogna irrobustire l'architettura implementando un sistema di High Availability del servizio di monitoraggio e di notifica

High Availability

- Alta Disponibilità
- Realizzare un'infrastruttura in HA vuol dire adoperare degli accorgimenti mediante i quali si garantisce continuità del servizio in relazione a determinati eventi di failure previsti
- In base alla criticità del servizio tali accorgimenti saranno più dispendiosi, ridondando un maggior numero di risorse
  - Connessioni di rete
  - Apparati
  - Dischi
  - Server
  - [...]
- Si implementa tipicamente realizzando un High-availability Cluster, visto all'esterno come un'unica entità
- Per eliminare Single Points of Failure

High-availability Cluster

- Detto anche Failover Cluster
- Nel nostro caso, parliamo di server ridondati (nodi) che possono esporre tutti lo stesso servizio
- Solo uno per volta (master) è in grado di modificare le risorse ed offrire il servizio
- Gli altri nodi (slave) controllano la disponibilità del servizio fornito dal master e in caso di failure, uno degli slave viene eletto master per garantire continuità del servizio

Reliability del singolo nodo

- Si cerca in questo caso di irrobustire il singolo nodo, duplicando le risorse:
  - Disk mirroring
  - Redundant Network Connections
  - Redundant Elettrical Power
- Se si realizza un High-availability Cluster, a seconda della criticità del servizio, tali accorgimenti devono essere presi in considerazione

Come si realizza

- Node Replication
- Data Mirroring / Sharing
- Heartbeat software
- Resource Manager

Node Replication

- Stesse applicazioni installate
- Stessa configurazione software
- Non necessariamente stessa configurazione HW, ma preferibile

Data Mirroring / Sharing

- Replicazione o condivisione delle risorse necessarie a mantenere lo stato del sistema aggiornato
- Solo il master deve avere accesso in scrittura, per garantire la consistenza

Heartbeat software

- Software per il controllo della reperibilità del Master, eleggendo uno Slave a Master in caso di Failure

Resource Manager

- Software che si occupa della gestione dell'acquisizione del controllo sulle risorse dell'entità virtuale
  - Mastering sulle risorse di storage
  - Mastering sulla configurazione di rete del server virtuale (Cluster Logical Host)
  - Avvio dei servizi
- Si occupa anche del rilascio

Soluzione adottata

- Il cluster metterà a disposizione in HA il servizio di
  - Accesso al sistema di monitoraggio (Panoptes)
  - Nagios
  - Notifiche via SMS
- Useremo la combinazione di due software Open Source: Heartbeat + DRBD

Nagios®

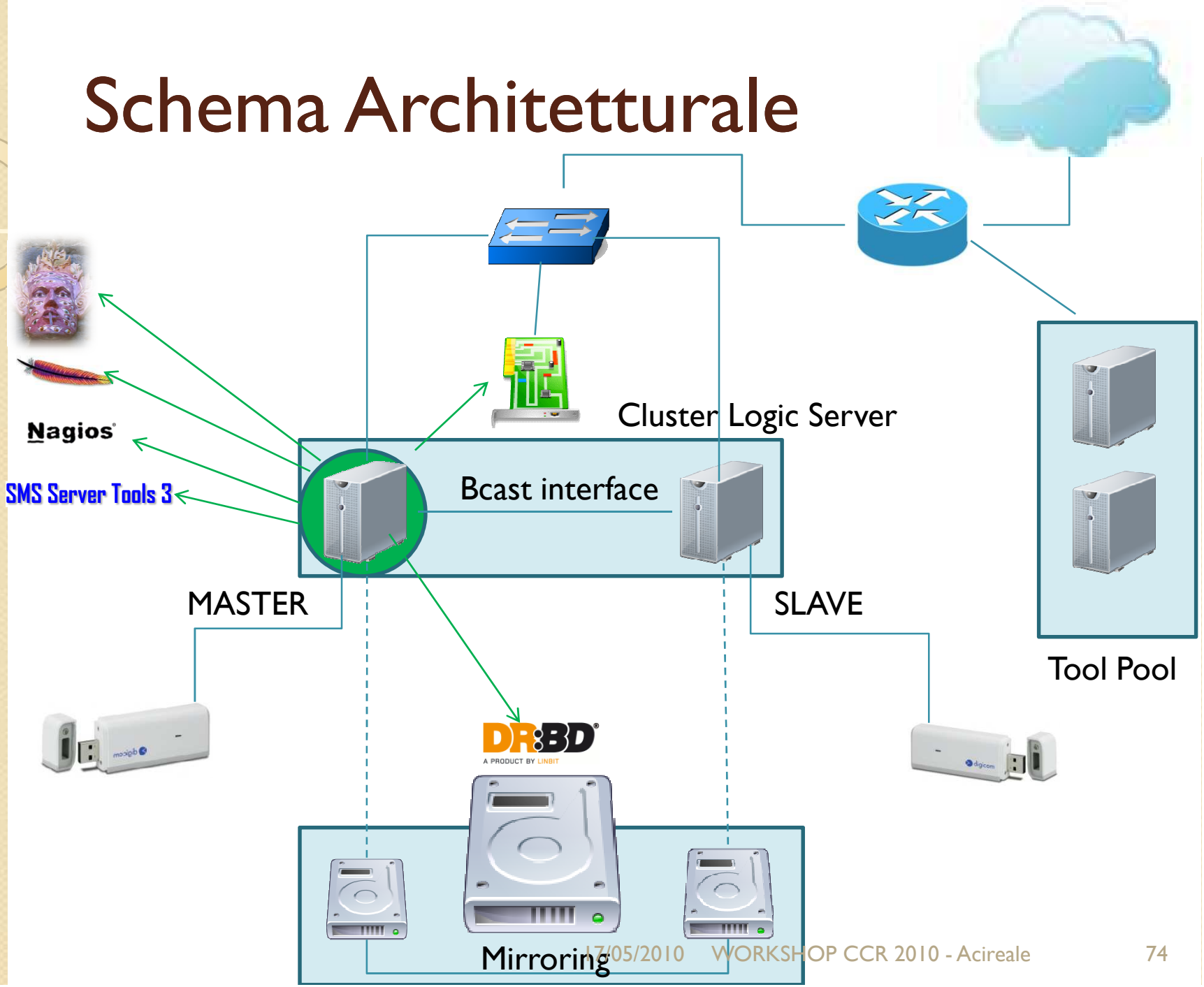
SMS Server Tools 3

HighAvailability

DRBD®

A PRODUCT BY LINBIT

Schema Architetturale



Heartbeat

- Software Open Source per la realizzazione di cluster HA
- Fino alla versione 2.99 integra sia il layer di Heartbeat che il Cluster Resource Manager, ora inserito in un progetto a parte , Pacemaker
- Usato tipicamente per 2-nodes clusters, ma non solo
- Sito ufficiale: http://www.linux-ha.org/wiki/Main_Page

Funzionamento di Heartbeat

- Servizio installato e attivo su tutti i nodi
- Si configura la "Risorsa" da condividere
  - Interfaccia virtuale del Cluster Logical Host
  - Accesso in scrittura al disco mirrorato con DRBD
  - Avvio dei servizi necessari (Apache, Nagios, smsd)
- Solo un nodo, il master, detiene la risorsa
- Su rete broadcast master e slave comunicano attraverso scambio di "battiti"
- Se il master non risponde, lo slave si dichiara master e detiene le risorse
- Se il master torna up, lo slave rilascia la risorse

DRBD

- Software per la realizzazione di un servizio di mirroring di dischi su rete (network based RAID-1)
- Sito ufficiale: http://www.drbd.org/home/what-is-drbd/

Funzionamento di DRBD

- Due nodi con partizioni delle stesse dimensioni
- Viene creato un device aggiuntivo DRBD su entrambi i nodi, agganciato alla coppia di device
- Solo uno dei nodi ha accesso in scrittura sul device virtuale e può montarlo per effettuare modifiche (Primary /Secondary)
- Le modifiche fatte dal Primary vengono riportate su entrambi i dischi (mirroring)

Evitare lo Split Brain

- Situazione in cui lo slave non riesce a contattare il master e si elegge tale prendendo il controllo delle risorse
- Il master, anche se non sente lo slave non se ne preoccupa, e non rilascia le risorse
- Capita nel caso di failure sulla rete di broadcast
- Entrambi detengono le risorse
- Inconsistenza del sistema soprattutto sui dati mirrorati
- Una soluzione è ridondare l'interfaccia di broadcast. Heartbeat può usare anche porte seriali in caso di scarsità di interfacce ethernet, utilizzando cavi null-modem

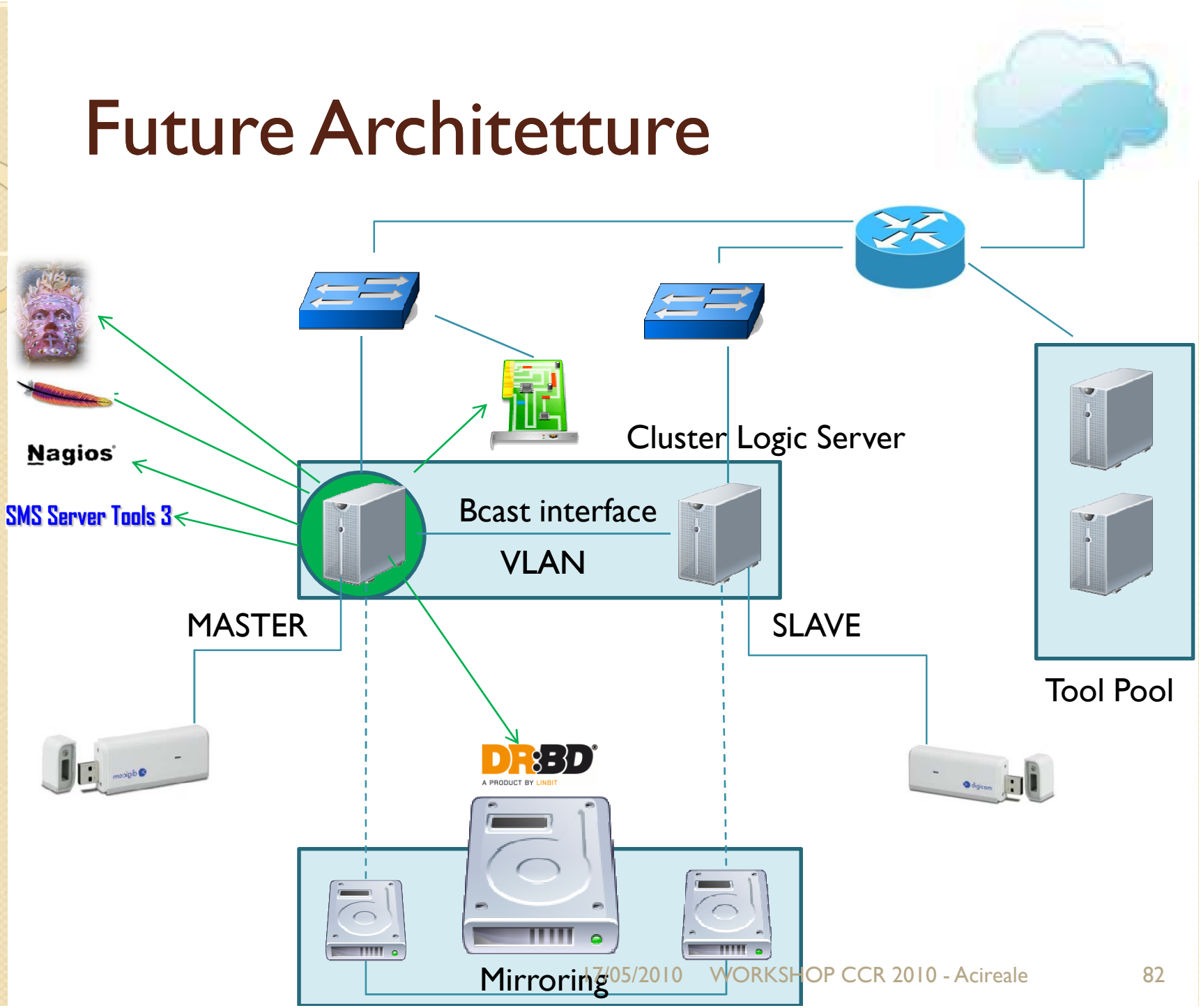
Nagios come sensore di problemi sul Master

- Non sempre l'host deve andare DOWN per poter dichiarare che il servizio non è più disponibile
  - Impossibilità di raggiungere il gateway (es. failure sull'interfaccia)
  - Crash del servizio (es. per eccessivo carico)
- Nagios in esecuzione sul master può monitorare con i plugin lo stato di tali servizi, e in generale lo stato interno del master.
- In caso di failure, attraverso l'event handling può forzare l'arresto di heartbeat sul nodo, facendo in modo che sia lo slave a dichiararsi master
- È solo una soluzione, vi sono molti altri modi

Configurazione Nagios

```
define host{
        use                     cdc-host-CRIT
        host_name               panoptes_network_link
        alias                   panoptes network link
        notes                   panoptes link with router
        address                 192.167.0.1
        parents                 INFNCT_A_1_router
        event_handler           swap_heartbeat_master
        max_check_attempts      3
        }
```

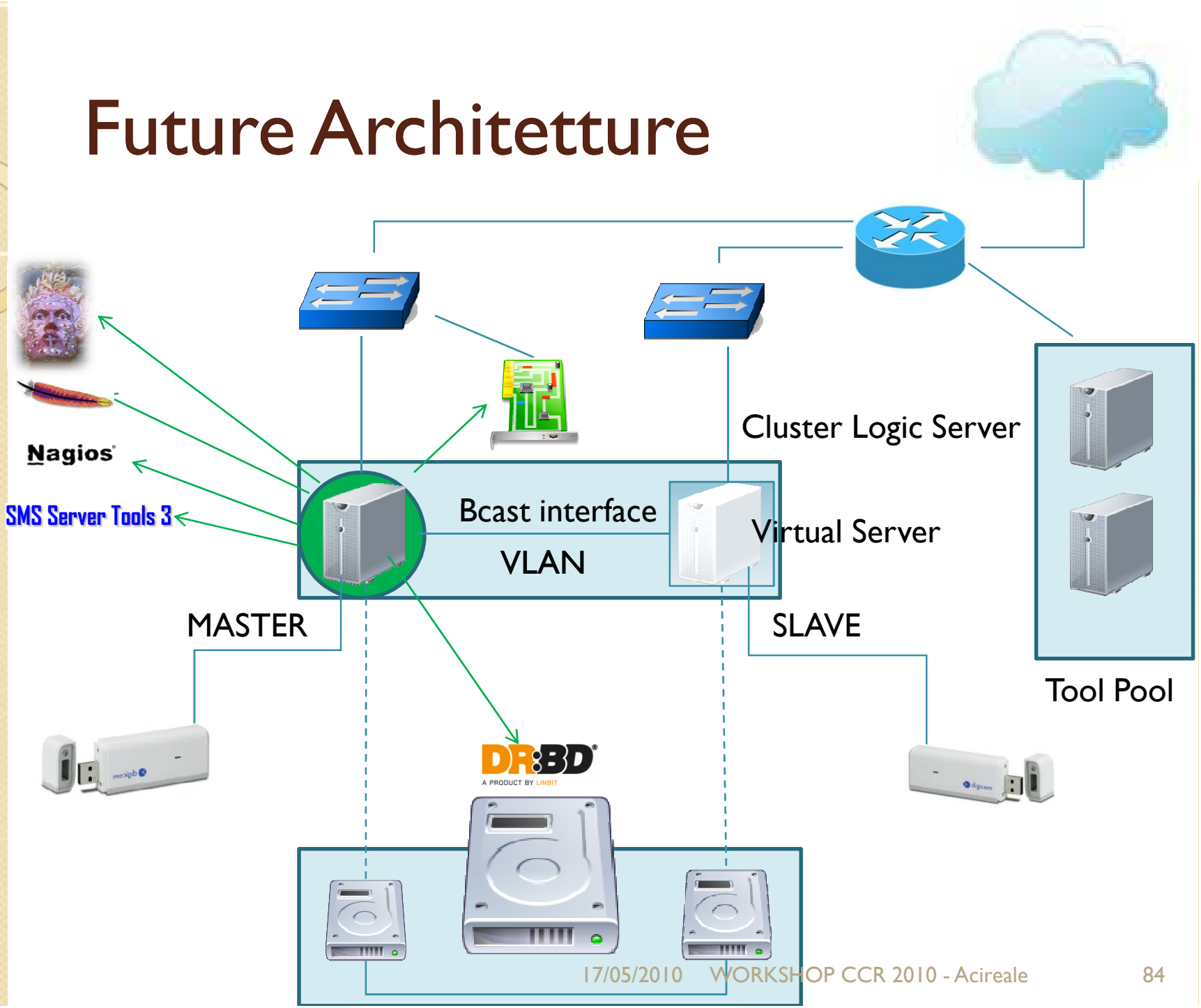
Future Architetture



Future Architetture

- Dislocare le macchine del cluster in posizioni differenti
- Percorsi verso l'esterno differenti (limitare i points of failure)
- Bcast interface utilizzando VLAN dedicata

Future Architetture



Future Architetture

- Per limitare il dispendio energetico, lo slave potrebbe essere un Virtual Server, realizzato con sistemi di virtualizzazione (es. XenServer)

Conclusione

La progettazione e l'estensione di un sistema di monitoraggio della rete, LAN e WAN, costituisce anche un'ottima opportunità di formazione per un futuro sistemista

Suggerimenti?