

INFN WAN Condor Pool: Status Report

P.Mazzanti, F.Semeria



Workshop sulle problematiche di calcolo e reti nell'INFN. 6-9 Maggio 2002
La Biodola, Isola d'Elba

Sommario

- › Condor e Progetto Condor INFN
 - Configurazione e stato del Pool
- › Utilizzazione del pool e reazioni dagli utenti
- › Possibili sviluppi



Workshop sulle problematiche di calcolo e reti nell'INFN. 6-9 Maggio 2002
La Biodola, Isola d'Elba

Cosa e` Condor

Condor converte un insieme di
workstation scorrelate in un sistema
di calcolo

high-throughput

E` un sistema non invasivo



Cosa e` Condor (cont.)

- › Condor cerca di runnare un job anche se alcune macchine:
 - Vanno in crash
 - Vengono disconnesse
 - Sono tolte o aggiunte al pool

Utilita` di Condor

- › Condor utilizza risorse che altrimenti andrebbero perdute:
1000 cpu non usate per 15 ore al giorno = **1.7 anni di cpu al giorno sprecati**

Policy d'uso generale

Ogni gruppo deve **mantenere il massimo controllo e la massima priorita`** sulle proprie macchine.



The ‘Condor on WAN’ INFN Project

- › Approved by the Computing Committee on February 1998.
- › Goal: install Condor on the INFN WAN and evaluate its effectiveness for the our computational needs.
- › ~30 people involved.
- › Collaboration INFN-CS Madison-Wisconsin



Organization of the Project

- › Coordinator: Paolo Mazzanti.
- › Central Management: (condor-admin@infn.it)
- › Local Management: ~25 people.
(condor@infn.it):
- › Steering Committee
- › Software support and maintenance contract with the Madison Condor Team

Central management

- > The Admin Group has to provide:
 - Configuration, tuning and overall maintenance of the INFN Condor Wan pool.
 - Management tools.
 - Activity reports.
 - Condor resource usage statistics (CPU, Network, Ckpt-server).
 - Help desk for users and local administrators.
 - Interface to Condor support in Madison.

Local management

› Local management has to provide:

release installation in collaboration with
the central management.

local Condor usage policies.

Steering Committee

- › The Steering Committee has to:
 - Interact with the Condor Team and suggest possible modifications of the system.
 - Define the general policy of the Condor pool.
 - Organize meeting for Condor administrators and users.

INFN-WAN Pool

ALPHA/OSF1	107
INTEL/LINUX	122
SUN/SOLARIS	6
INTEL/WNT	1

Total	236

INFN-WAN Pool (cont.)

Il primo esempio in Europa di un sistema di calcolo distribuito a livello nazionale.



Il problema del checkpoint

- › Il checkpoint di un job puo' essere molto grande (~100MB): possibili problemi di rete se fatto su WAN.

Checkpoint Server Domains

Per ridurre il traffico di rete si divide il pool in “domini di checkpoint”, ognuno con un checkpoint server.

Un job deve rimanere il piu’ possibile nello stesso dominio.

Checkpoint Server Domains(cont.)

Numero di macchine associate ai ckpt domain

BO: 42

CNAF: 71

LNGS: 12

MI: 19

NA: 11

PD: 25

PV: 17

TO: 32

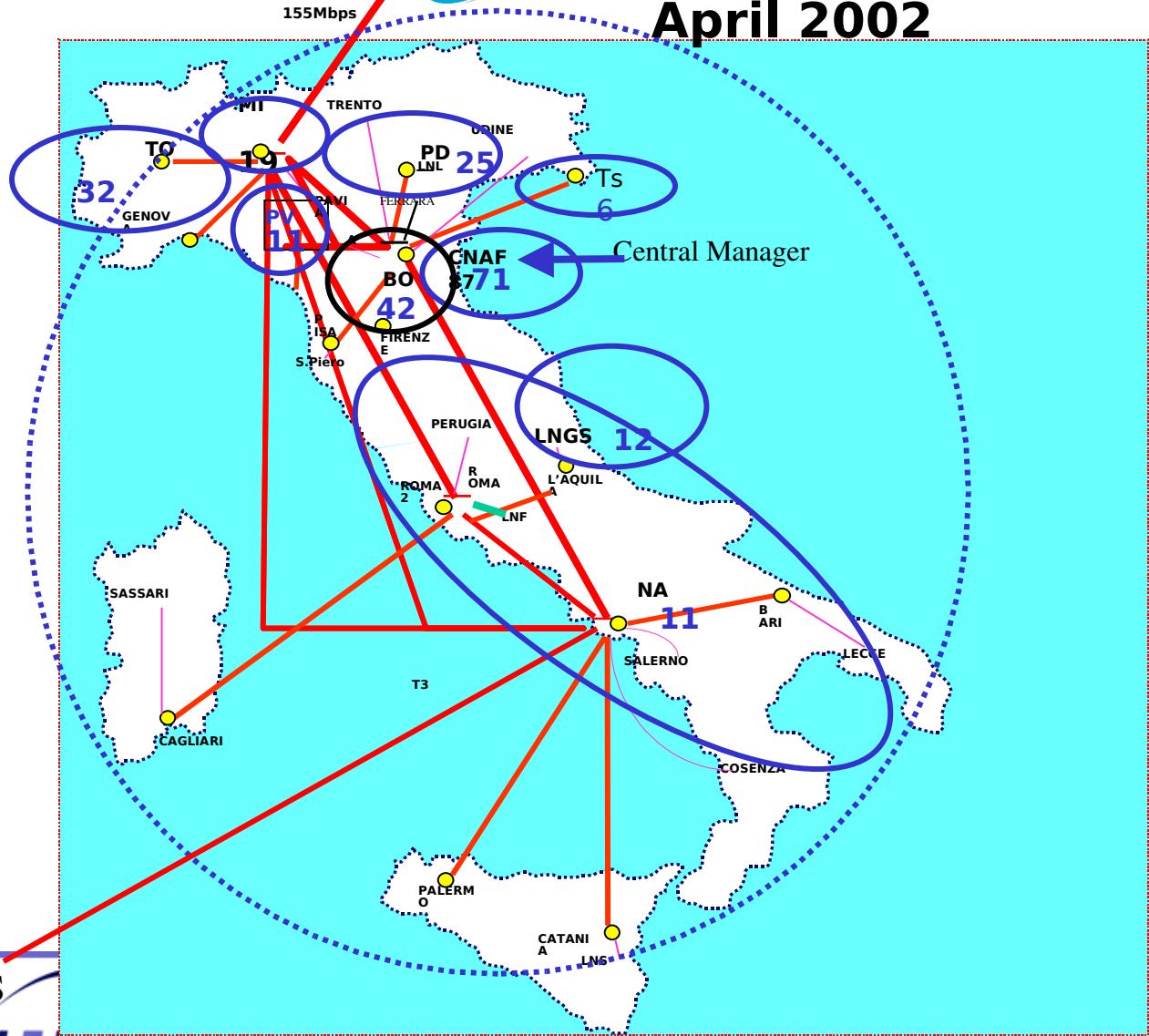
TS: 6



Europe/US



INFN Condor Pool on WAN: checkpoint domains April 2002



GARR-B Topology

155 Mbps ATM based Network

access points (PoP)

main transport nodes

radio wave bridge 34M
user access E1-E3

CKPT domain
hosts

Default ckpt
domain in cnaf

Maggio 2002

Software installed Condor releases

Official release: 6.2.0/1

However a large number of machines, mainly Linux are running old versions
6.1.x (x=8,10,12,16) → problems



APPLICATIONS

- › Simulation of Cherenkov light in the atmosphere (CLUE).
- › MC integration in perturbative QCD.
- › Dynamic chaotic systems.
- › Extra-solar planets orbits.
- › Stochastic differential equations.
- › Maxwell equations.
- › MC simulation, liquid crystals

Il numero di utenti che hanno usato il pool per piu' di 100 ore: 45.

Un utente ha utilizzato circa 100.000 ore di allocation time.

Possibilita' di runnare per tempi lunghissimi.



Totale ore utilizzate Periodo Gennaio-Dicembre 2001

~ 722000 > 82 anni!

Outside INFN ...

- **CERN: Chorus exp.**
- **NASA : ~ 300 cpus**
- **Japan : ~ 300 cpus**
- **Oracle: ~ 1500 cpus**
- **Madison: ~ 1800 cpus in 10 campus pools**

**More than 350 condor pool with more than
9000 cpus world wide.**

Grado di soddisfazione:

Utenti che lo usano in modo sistematico: ottimo

Ma:

- Alcuni programmi non danno buone prestazioni... I/O ? (Ma: caching..)**
- Problemi associati al CKPTng (non omogeneita' delle release ?)**
- Il pool DEVE essere mantenuto aggiornato.**

Alcuni utenti trovano difficile usare Condor !!!

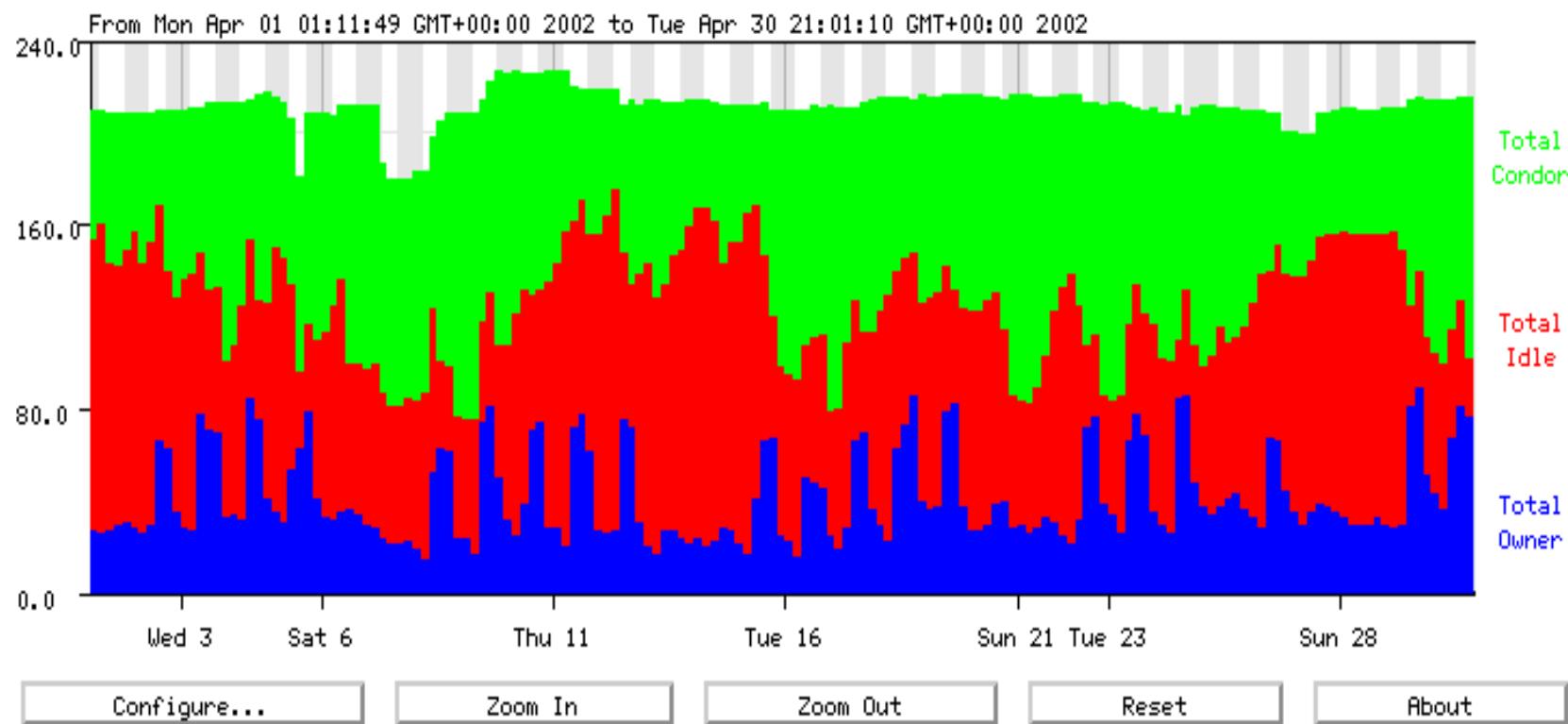
Caching

Condor keeps a buffer of recently-used data for each file a job accesses.

This buffer is used both to cache commonly-used data and to consolidate small reads and writes into larger operations that get better throughput.

Caching only apply to standard-universe jobs.

INFN Condor Pool Machine Statistics for Apr



[Graph Hints: The Y-axis is number of machines, the X-axis is time. When graph finishes updating, press "Configure..." to view different Architecture or State data. Also, you can use the mouse to draw a rectangle on the graph and then press "Zoom In". Press "Reset" to center/resize the data after Configure or when done zooming. Nighttime shows up on graph background as grey.]

Arch	Owner Average	Condor Average	Idle Average	Owner Peak	Condor Peak
Total	41.6 (19.6%)	86.1 (40.7%)	83.8 (39.6%)	95 (44%)	138 (65%)

Sviluppi

- Migliorare l'interazione con gli utenti.
- Migliorare la gestione del pool.
- Installare la prossima release stabile: indispensabile maggiore omogeneita`.
- Incrementare il numero delle cpu nel pool: **includere farm (flocking?)**.

Uso di Condor come sistema batch.

Sviluppi (flocking)

Flocking allows jobs from one Condor pool to run in another Condor pool if there are insufficient resources in the local pool.

Sviluppi (batch)

A batch system may be set up by using the following features:

- more than one job may be executed on the same cpu at the same time
- a job time limit may be defined

As a consequence QUEUES may be easily (!!) defined and configured.

New Release 6.4.0

Arch. and OS.:

- **Complete support for RedHat 7.x,
HP, Sun, SGI, Alpha.**
- **W2K and MAC OS X 'clipped supported'
(no ckpt,no parallel functionalities,
no remote system calls).**

New Release 6.4.0 (cont)

Security:

Secure network connection (Kerberos, X509, NT LanMan)
Encryption: all ‘condor’ network traffic can be encrypted
Authorization based on host or USER.

Firewall support

E.g.: **LOWPORT = X**

HIGHPORT = Y

All dynamic ports will be between x and y



New Release 6.4.0 (cont.)

New Job Universe:

- MPI universe
- Globus Universe (GRID)
- Java Universe.

Possibility to run condor jobs on different architectures.

New Release 6.4.0 (cont.)

File Transfer:

**possibility to transfer files from the submitting machine
to the executing one.**

Example:

Transfer_input_files = x. y. z.

Transfer_output_files = a. b. c.

Transfer_files = [ALWAYS | ONEXIT]

default: send back new/changed files

Conclusioni

**L'INFN WAN condor pool si e' dimostrato un efficiente sistema di calcolo distribuito.
Limitati i costi di gestione e mantenimento.**

Soddisfa ad un requisito fondamentale per quanto riguarda l'utilizzo delle risorse: permette l'uso di risorse di calcolo che altrimenti andrebbero perdute.

L'estrema flessibilita' del software rende possibile moltissime diverse configurazioni.

Si continua?