
Il Progetto di Centro di Reprocessing di BaBar: Monitoring e Simulazione



Moreno Marzolla

Email marzolla@pd.infn.it

WWW: <http://www.pd.infn.it/~marzolla>

INFN Padova – BaBar Collaboration &
Dip. Informatica, Università di Venezia

Sommario

- Introduzione alla farm di Reprocessing dell'esperimento BaBar
- Requisiti per il monitoring
 - *ASC: Asynchronous SNMP Collector*
- Requisiti per la simulazione

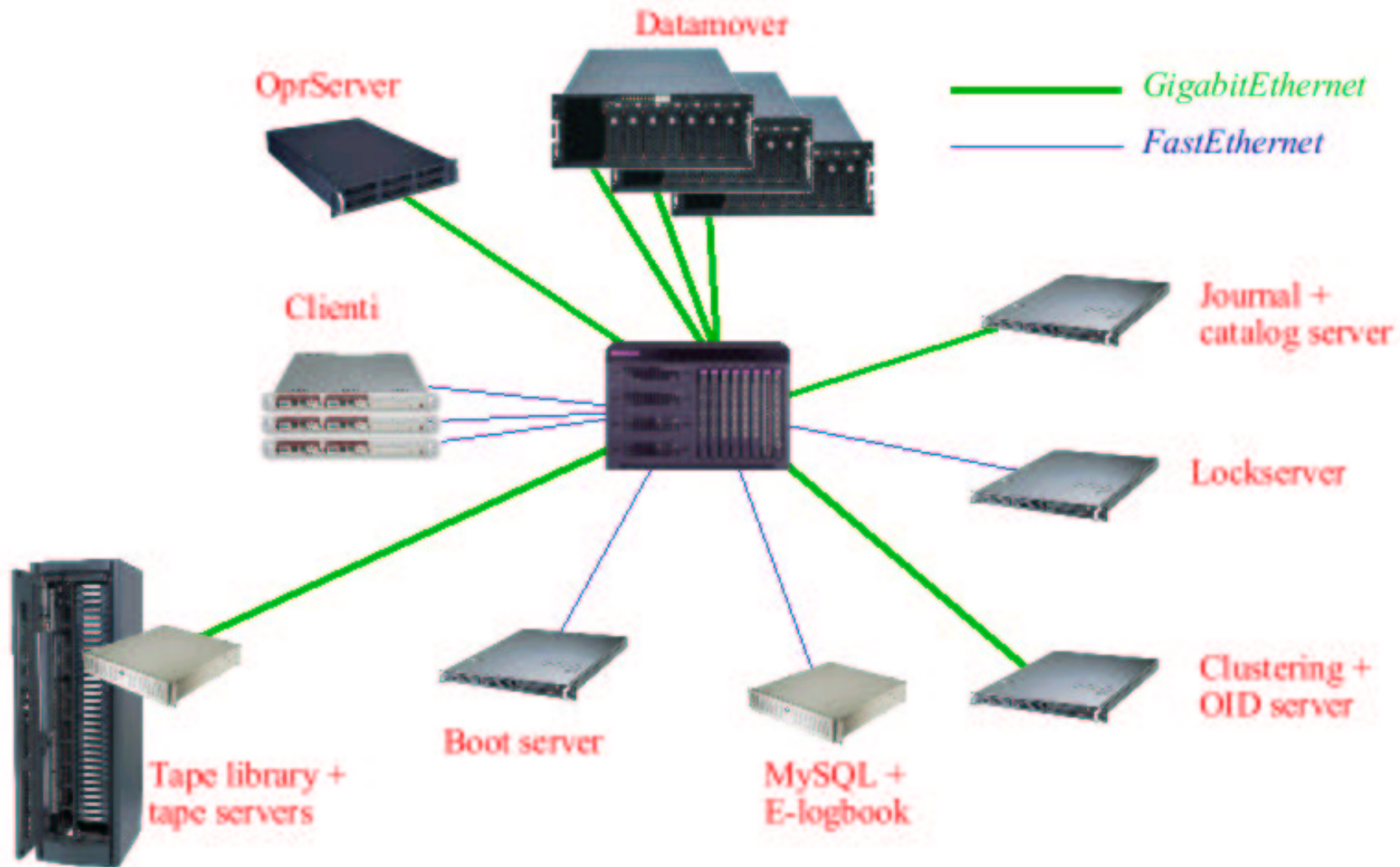
La farm di Reprocessing (1)

- Deve riprocessare ogni anno la statistica raccolta dall'esperimento BaBar negli anni precedenti
 - Tempo di riprocessamento \approx 6 mesi
- Sarà composta da circa 150 nodi client.
Ciascuno ha:
 - 2 CPU Pentium, almeno 1GHz
 - Dischi locali
 - 1 GB Ram
 - 1 Ethernet 100Mbit/s
- Il numero di nodi aumenterà ogni anno.

La farm di Reprocessing (2)

- In più vi sono un certo numero di server, con caratteristiche leggermente diverse
 - Più disco locale
 - Scheda di rete Gigabit Ethernet
- Altre componenti:
 - Libreria robotizzata di nastri
 - 20 TB di spazio disco
 - Switch di rete ad alte prestazioni
 - Gruppi di continuità
 - Sistema di condizionamento ambientale

Schema della Farm



Requisiti per il Monitoring (1)

- I parametri da monitorare appartengono alle categorie seguenti:
 - Stato dell'Hardware
 - Crash generico delle macchine
 - Crash dei dischi
 - Altri guasti hardware diagnosticabili
 - Stato della libreria dei nastri
 - Temperatura delle CPU
 - Spazio libero sui dischi

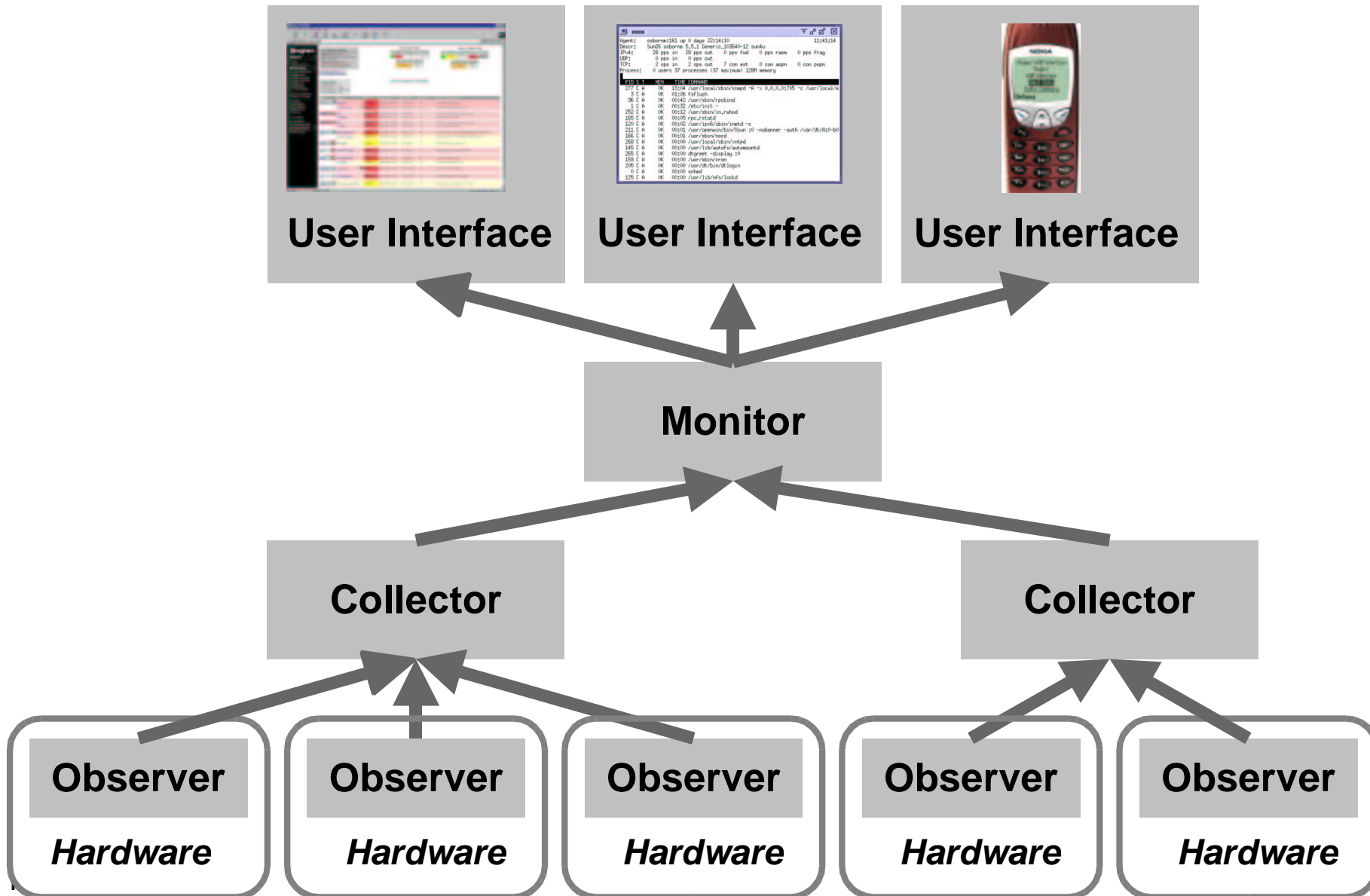
Requisiti per il Monitoring (2)

- Stato dei processi
 - Utilizzazione delle CPU nei client e nei server
 - Utilizzazione della rete
 - I/O su disco, specialmente nei server
 - Crash dei processi server o client
- Parametri ambientali
 - Temperatura nella stanza che ospita la Farm
 - Stati dei condizionatori
 - Stato dell'UPS

Altri requisiti

- Il sistema di monitoring dovrebbe essere:
 - *Scalabile*
 - Efficiente (richiedere la minor quantità di risorse possibili)
 - Flessibile e facilmente adattabile a nuove esigenze
 - *Facile da configurare*
 - Possibilmente gestibile anche da riga di comando
- È importante che possa correlare eventi diversi per evidenziare eventuali problemi

II monitoring in generale



Observer: SNMP

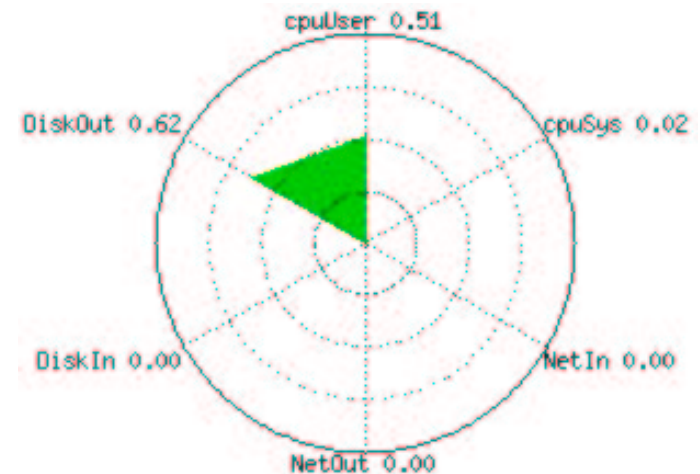
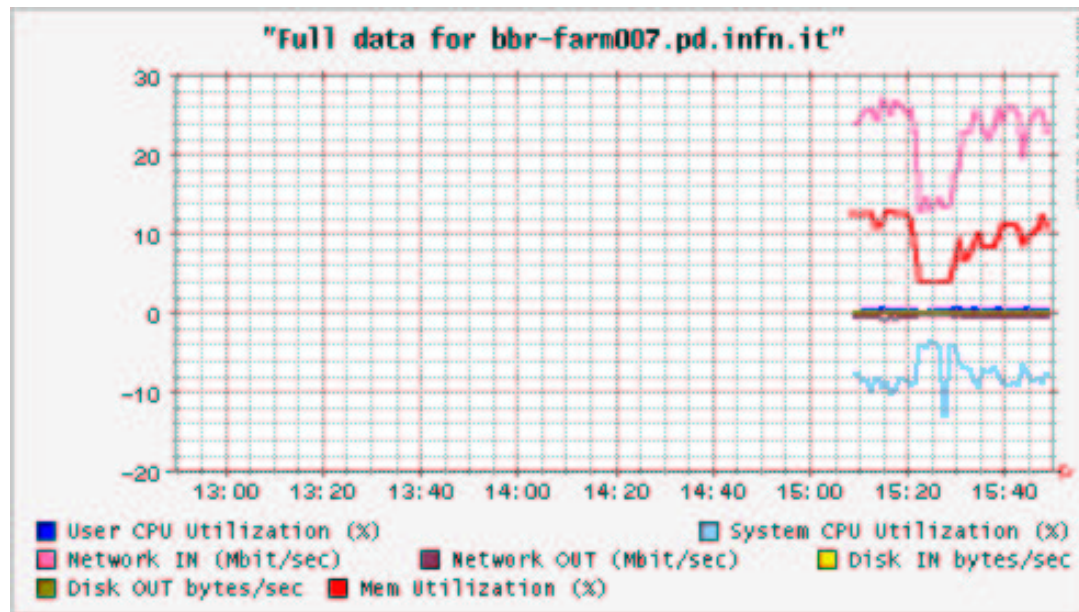
- In base ad alcuni esperimenti preliminari, sembra ragionevole basare il futuro sistema di monitoring su **SNMP** (*Simple Network Management Protocol*)
 - Si tratta di un protocollo ormai maturo
 - È supportato da una vasta gamma di dispositivi
 - È ragionevolmente flessibile
 - Esiste una ottima implementazione *open source* per Unix/Linux (<http://ucd-snmp.ucdavis.edu/>)

Collector: ASC

- È un prototipo in fase di sviluppo
- Features:
 - Polling SNMP asincrono;
 - File di configurazione in formato XML;
 - Usa Round Robin Database per memorizzare dati con granularità diversa;
- Si sta dimostrando molto efficiente
 - Legge \approx 20 MIBs SNMP su 20 macchine in circa 2 secondi;
 - Uso della CPU praticamente nullo;

Observer: httpd

- Per visualizzare i dati raccolti in forma grafica si fa uso di un piccolo server http
 - Riceve richieste di file grafici
 - Produce i grafici sul momento.

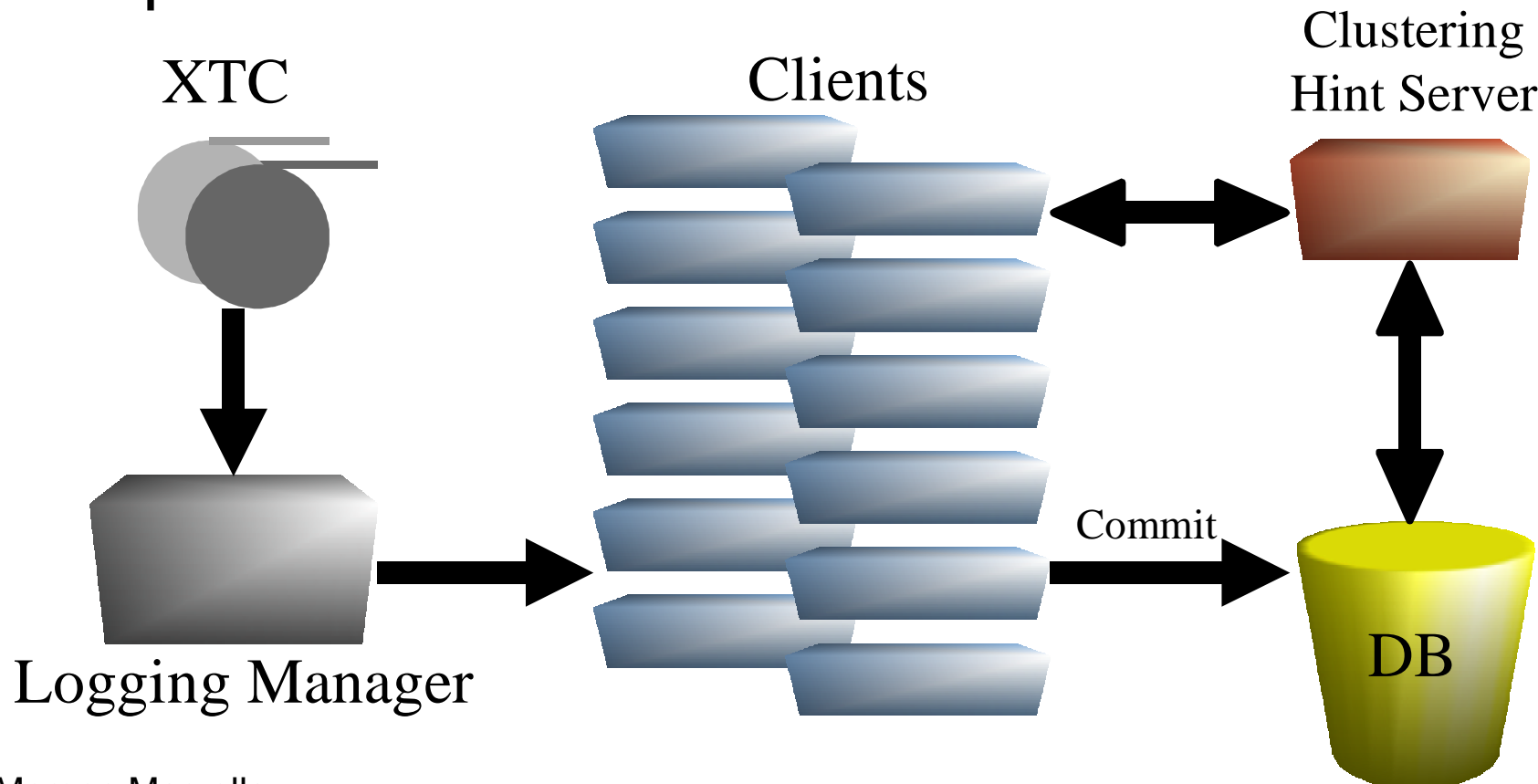


Attività proposte

- Estendere il tool per aggiungere nuove funzionalità
 - Configurabilità dei grafici da generare;
 - Generare direttamente le pagine WEB riassuntive complete di grafici;
- Studiare l'impatto del monitoring sulla rete
 - Ossia: quanta banda viene "consumata" dal monitor?
- Studiarne i limiti di scalabilità
 - Come cresce il tempo necessario per interrogare tutti i nodi, in funzione del numero di nodi?

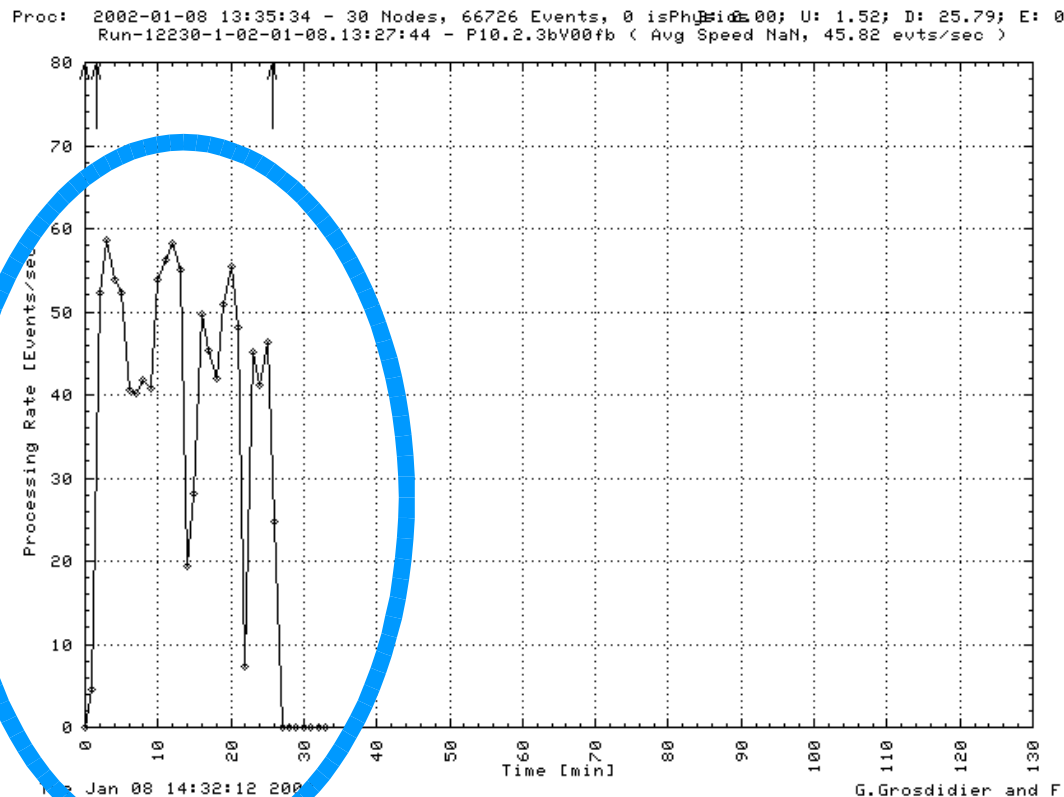
Simulazione

- La farm è un sistema altamente complesso, che deve garantire un certo livello di prestazioni.



Problemi

- Nella pratica si osserva che il tasso a cui le macchine client elaborano gli eventi è estremamente fluttuante. **Perché?**



Piano di lavoro

- Sviluppare un modello del funzionamento della farm;
- Implementare il modello con un linguaggio/libreria di simulazione (es. SIMULA)
- Misurare i parametri di funzionamento del sistema reale
- Cercare di riprodurre nel modello il comportamento osservato nel sistema reale

Una proposta di modello

Rete di code aperta, con blocco

