D0 MC simulace - manuál

20. září 2006

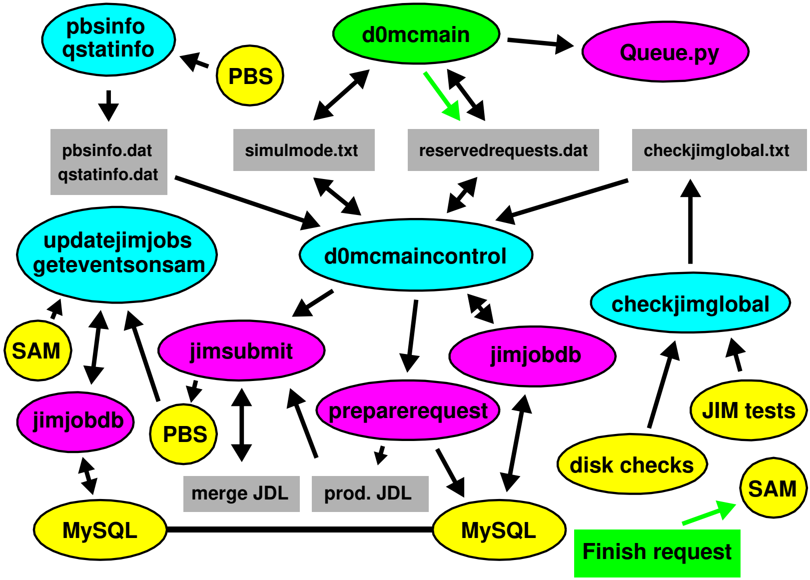
- Ovládací software
    - Stav a struktura
    - Hlavní komponenty
    - Co dělá uživatel
    - Co by neměl dělat uživatel
- Monitoring

Ovládací software - Stav

- Ovládání simulací je plně automatizováno
- Všechny joby, produkční i mergovací, se spouští automaticky
- Stav každého jobu je automaticky kontrolován a s aktuálním počtem eventů ukládán do databáze
- Automaticky testována dostupnost potřebných zdrojů JIM, PBS, d0prod fronta, diskový prostor, cardfiles, mc_runjob, d0 release
- Automatické vypnutí simulací v případě nedostatku diskového prostoru

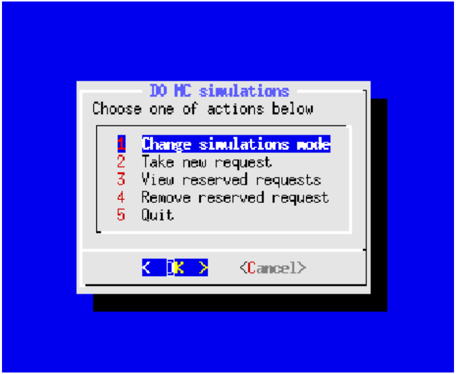
Ovládací software - Struktura

- Bash skripty spouštěné periodicky z crontabu
  - spustitelné pouze uživatelem 'soustruz'
  - uživatel nespouští a nepoužívá
- Bash skripty (+ 1 python skript) volané z 'cron' bash skriptů
  - spustitelné pouze uživatelem 'soustruz'
  - uživatel nespouští a nepoužívá
- **Ovládací bash skript**
  - spustitelný všemi členy skupiny 'd0mcsim' (V. Hynek, K. Soustružník, P. Vokáč)
  - jediný skript určený ke spouštění uživatelem
- Textové soubory
  - přenos informací a povelů mezi 'cron' skripty
- MySQL databáze
  - uchovávání informací o requestech a jobech
- SAM, JIM, PBS, lokální disky


pbsinfo
qstatinfo
PBS
d0mcmain
Queue.py
pbsinfo.dat
qstatinfo.dat
simulmode.txt
reservedrequests.dat
checkjimglobal.txt
updatejimjobs
geteventsonsam
d0mcmaincontrol
checkjimglobal
SAM
jimsubmit
jimjobdb
jimjobdb
PBS
preparerequest
JIM tests
disk checks
merge JDL
prod. JDL
SAM
MySQL
MySQL
Finish request

Hlavní komponenty - d0mcmain

- **d0mcmain** skript - ovládací skript
  - spouštěn uživatelem, členem d0mcsim skupiny, na **samu**
    sam : /home/soustruz/scripts/d0mcmain
  - zobrazuje aktuální status (mód) simulací
  - umožňuje měnit status simulací, rezervovat si nový request



Hlavní komponenty - d0mcmain # 2

Status (mód) simulací :

- run
  - spouští se jak produkční, tak mergovací joby
- pause
  - spouští se pouze joby mergovací
- stop
  - nespouští se žádné joby
- abort
  - nespouští se žádné joby, navíc všechny joby běžící a čekající ve frontě jsou pobity

Hlavní komponenty - d0mcmain # 3

Rezervace nového requestu :

- Pomocí *Queue.py* skriptu spouštěného z menu *d0mcmain* skriptu
- Odpovědět na otázky *Queue.py* skriptu
  1. *Does your farm have a maximum number of events/request ?*
     No.
  2. *Does your farm have a minimum number of events/request ?*
     Yes.
  3. *What is the minimum number of events per request ?*
     51000.
  4. Nabídne request (číslo a počet eventů)
     Do you want to run this request ?
     Yes.
  5. *What is your facility name ?*
     FZU_GRID.
- Zadat číslo rezervovaného requestu : *d0mcmain* skript

Hlavní komponenty - d0mcmaincontrol(-test)

- **d0mcmaincontrol-test** skript - hlavní kontrolní bash skript
  - spouštěn 1x za hodinu z crontabu uživatele *soustruz* na *samu*

  - zjišťuje status (mód) simulací (může ho i změnit - vždy jen zastavit simulace)
  - přesouvá zarezervované requesty ze souboru do databáze
  - kontroluje dostupnost zdrojů (diskový prostor, JIM, PBS)
  - kontroluje vytížení *d0prod* fronty
  - spouští produkční a mergovací joby (JIM joby) (podle módu simulací a vytížení *d0prod* fronty)
  - kontroluje stav JIM jobů a zajišťuje produkci požadovaného počtu eventů

Co dělá uživatel # 1

- Zajišťuje dostatek requestů ke zpracování
  - rezervuje nové requesty pomocí skriptu *d0mcmain*
- Označuje *completed* requesty jako *finished*
  - *sam modify request --requestId=requestID --status=finished*
- Monitoruje běh simulací a řeší případné problémy
  - hlavně hledá příčiny pádu jobů
    - log soubory jednotlivých lokálních jobů
      */raid3_d0/jim_sandbox_area/box***/std_out*†
      */raid3_d0/jim_sandbox_area/box***/std_err*†
      kde "***" je číslo sandboxu daného JIM jobu a "†" je pořadové číslo lokálního jobu pocházejícího z daného JIM jobu
    - číslo sandboxu daného JIM jobu lze zjistit z WWW stránky SAMGRIDu pro daný JIM job ("BS" link na seznamu JIM jobů spuštěných u nás) - viz. Monitoring
    - log soubory lokálních jobů jsou přístupné pouze uživateli *samgrid*
  - hlavní log file sady skriptů
    /raid4_d0/d0/soustruz/d0mcmonitor/d0mcmaincontrol.dat

Co dělá uživatel # 2

- Doplňuje chybějící cardfily
  - seznam cardfilů *sam : /sam/cache3/boo/cardFile*.tar.gz*
  - aktualizovat nejnovější filedataset
    sam get dataset --defname="dataset_name"
  - definice filedatasetů v *sam :*
    /home/soustruz/samgrid/mcp17/jim_prod/datasets
- Provádí aktualizace produkčního softwaru
  - aktuální verze simulačního softwaru na WWW stránkách D0 MC production skupiny
  - vytvořit nový filedataset obsahující tar soubory D0 release, magnetické pole, mc_runjob a cardfily
    sam define dataset
    --file_name=seznam-souborů-oddělených-čárkami
    --def_name="jméno-datasetu" --group=dzero
  - stáhnout nový filedataset ze SAMu
    sam get dataset --defname="jméno-datasetu"

Co by **neměl** dělat uživatel

- Brát request jinak než pomocí *Queue.py* skriptu
  - může způsobit chaos, neboť stejný request si může vzít více farem
- Spouštět produkční joby “ručně”
  - job nebude v databázi a tudíž ani započten do celkové sumy vyprodukovaných eventů, zbytečná nadprodukce
  - “ručně” spuštěný job se musí “ručně” i zmergovat
- “Ručně” spouštět mergovací joby
  - o “ručním” mergování nejsou údaje v databázi a skripty se tak job uvedený v databázi pokusí zmergovat i automaticky, hrozí vícenásobné zmergování týchž tmb souborů
- Předčasně označovat requesty za hotové
  - uživatelé si mohou vytvořit sam datasety ze souborů tohoto requestu a po jeho skutečném dokončení mohou být tyto datasety chybné, chyby se mohou dále propagovat do analýz
  - v případě potíží s předčasně označeným requestem nelze ze SAMu odstranit soubory použité v datasetech

Monitoring *hpv.farm.particle.cz:8080/∼d0mc/db/d0mcrequests.php*

D0 MC requests done in Prague

Request ID	Group	Request description	Events req.	Events done	Events on SAM	Start date	Finish date	Request status	Release	Cardfile	Done by
38152	dzero	CSMC Alpgen+Pythia w+2lp->lnu+2lp excl	200000	0	0	2006-09-19	2006-09-19	booked	p17.09.06	v00-09-44	JLM
38151	dzero	CSMC Alpgen+Pythia w+2lp->lnu+2lp excl	200000	0	0	2006-09-19	2006-09-19	booked	p17.09.06	v00-09-44	JLM
37790	dzero	p17.09.06 CSMC Pythia gam-z(+jets)->taut	200000	187000	0	2006-09-14	2006-09-14	running	p17.09.06	v00-09-07	JLM
37789	dzero	p17.09.06 CSMC Pythia gam-z(+jets)->taut	200000	3000	197500	2006-09-14	2006-09-14	running	p17.09.06	v00-09-07	JLM
37884	dzero	CSMC Alpgen+Pythia w+0lp->lnu+0lp excl	200000	25500	0	2006-09-11	2006-09-11	running	p17.09.06	v00-09-44	JLM
37638	dzero	CSMC Pythia W(+jets)->enu(+jets)	200000	51500	211500	2006-09-08	2006-09-08	completed	p17.09.06	v00-09-07	JLM
37637	dzero	CSMC Pythia W(+jets)->enu(+jets)	200000	112500	225000	2006-09-08	2006-09-08	completed	p17.09.06	v00-09-07	JLM
37564	dzero	Pythia gam-z(+jets)->mumu(+jets) m=5-15	200000	80500	210500	2006-09-06	2006-09-13	done	p17.09.06	v00-09-43	JLM
37563	dzero	Pythia gam-z(+jets)->ee(+jets) m=5-15 (r	200000	16500	216500	2006-09-06	2006-09-11	done	p17.09.06	v00-09-43	JLM
37562	dzero	Pythia gam-z(+jets)->ee(+jets) m=5-15 (r	200000	154500	214500	2006-09-06	2006-09-11	done	p17.09.06	v00-09-43	JLM
32684	dzero	CSMC Alpgen+Pythia w+0lp->lnu+0lp excl	100000	4000	99000	2006-07-27	2006-08-30	done	p17.09.01	v00-09-20	JLM

Monitoring *hpv.farm.particle.cz:8080/∼d0mc/db/d0mcrequests.php*

- Stránka D0 MC simulací v Praze (*www-hep2.fzu.cz/d0/d0mcsimulations.html*) : link "Production status" (zelené vlevo dole)
- Přehledný seznam requestů zpracovávaných na FZU farmě, chronologicky řazený (nejnovější nahoře)
- Barevně odlišeny různé stavy requestů
- Odkazy na podrobnější informace o jednotlivých requestech (číslo requestu v prvním sloupci zleva)
- pravidelně aktualizované počty vyprodukovaných a zmergovaných eventů

Monitoring *hpv.farm.particle.cz:8080/∼d0mc/db/runningjimjobs.php*

Simulations status

RUN

Running JIM jobs

PBS server **PBS scheduler**

JIM			
/raid4_d0	84%	/raid4_work	79%
/data	31%	sam /	86%
/sam/cache3	96%	/raid3_d0	88%

Request ID	JIM job ID	Job type	Running	Queued	Done %
37789	soustruz_sam.farm.particle.cz_132000_6383	mcprod	1	0	85
37790	soustruz_sam.farm.particle.cz_182115_22074	merge	5	10	0
37884	soustruz_sam.farm.particle.cz_142003_9181	mcprod	117	248	15
Total			123	258	

Monitoring *hpv.farm.particle.cz:8080/∼d0mc/db/runningjimjobs.php*

- Stránka D0 MC simulací v Praze (*www-hep2.fzu.cz/d0/d0mcsimulations.html*) : link “Running JIM jobs” (zelené vlevo dole)
- Barevně zvýrazněný status (mód) simulací
- Seznam běžících (již registrovaných v PBS) JIM jobů, produkčních i mergovacích
- Uvedeny počty běžících i čekajících lokálních jobů náležejících každému JIM jobu
- Ukazuje relativní počet dokončených eventů
- Informace o stavu PBSka (PBS server i PBS scheduler)
- Ukazuje aktuální obsazení disků důležitých pro simulace
- Ukazuje stav JIM serverů

Monitoring - další užitečné stránky

- *www.particle.cz/farm/index.aspx?page=monqueuesgolias*
  - WWW stránky FZU farmy s grafy vytíženosti a počty jobů v jednotlivých frontách
  - z našich stránek D0 simulací (*www-hep2.fzu.cz/d0/d0mcsimulations.html*) link "Our farm page" (vpravo, 2. link shora) → "Monitoring" → "Queues" → "Jobs on Golias"
- *samgrid.fnal.gov:8080/known_scheduler.php?schedulername =sam.farm.particle.cz*
  - WWW stránka SAMGRIDu se seznamem JIM jobů spuštěných u nás
  - linky na stránky s informacemi o jednotlivých lokálních jobech - ve sloupci "Global Job ID" pod "(BS)"
  - z našich stránek D0 simulací link "JIM job monitoring" (zelené vlevo dole) → "sam.farm.particle.cz"