

Výkonnostní měření dávkového systému Torque

Šimon Tóth

January 18, 2010

Abstract

Tento dokument mapuje výkonnostní limity dávkového systému Torque v jednotlivých dostupných verzích a vliv implementovaných experimentálních vlastností na tyto limity a celkové chování systému.

Pro měření byl sestaven testovací nástroj postavený nad lehkými linuxovými kontejnery, které umožnili spouštění všech experimentů na jednom fyzickém stroji. Tento systém poslouží v budoucnu k verifikaci nových vlastností.

První sekce obsahuje krátký popis sestaveného nástroje. Následuje popis metodiky a významu jednotlivých měření následovaný popisem jednotlivých provedených experimentů a jejich výsledků. V závěru se pak nachází přehledové grafy a tabulky pro všechny zaznamenané experimenty.

1 Systém pro výkonnostní testování Torque

Pro otestování výkonu Torque, stability systému a vlivu jednotlivých implementovaných úprav byla vytvořena sada testovacích nástrojů. Tato sada byla navržena s ohledem na minimální ovlivňování výsledků experimentů.

Celá sada je postavena na technologii Linuxových kontejnerů¹, které jsou použity místo virtuálních strojů jako jejich velmi odlehčená alternativa. Díky velmi nízké režii jednotlivých spuštěných kontejnerů bylo možné spouštět poměrně rozsáhlé testy na jednom fyzickém stroji (Intel(R) Core(TM)2 Duo CPU E8400 @ 3.00GHz s 4GB RAM).

S podobnou filozofií byl navržen i systém pro spouštění samotným experimentů. V průběhu experimentu se provádí pouze dohled nad experimentem a kolekce dat. Veškeré výpočty se provádějí až po skončení experimentu. Zároveň veškerá průběžná kolekce probíhá pouze formou postupného čtení logů, které Torque v průběhu běhu generuje, aby nedošlo k zatížení komunikačního cyklu serveru dotazy na aktuální stav.

1.1 Části systému

Celý systém se skládá ze čtyř hlavních částí:

- Hlavní dohlížecí program
 - ten dle zvoleného nastavení buďto spustí experiment, nebo provede zpracování již dříve sesbíraných logů. Zároveň je zodpovědný za dohled na běžícím experimentem a detekci konce experimentu (ať už úspěšného nebo neúspěšného).
- Sada stage skriptů
 - ty se starají o samotné spouštění jednotlivých částí experimentů (install, prepare, submit, terminate, collect, cleanup). Jednotlivé skripty jsou vytvořeny poměrně obecně, za účelem maximalizace znovupoužití.

¹<http://lxc.sourceforge.net/>

- Obalovací vrstva nad správcí kontejnerů
 - lehké kontejnery jsou poměrně čerstvou novinkou Linuxového jádra a nástroje pro jejich správu se nechovají velmi spolehlivě (typicky označují virtuální stroj za vypnutý nebo zapnutý i když stroj v tomto stavu zatím není)
- Sada konverzních zpracovávajících skriptů a jiných drobných programů
 - které se starají o zpracovávání samotných výstupů měření a jejich převod do vizuální podoby

1.2 Tvorba nových experimentů

Pro vytvoření nového experimentu obvykle stačí vytvořit nový konfigurační soubor popisující daný experiment (viz. Obrázek 1) a v některých případech vytvořit mírně upravenou verzi stage skriptů.

```
# stage skripty
install_cmd="/light-vm/install-torque-world-2.4"
prepare_cmd="/light-vm/prepare-200n"
submit_cmd="/light-vm/submit-5000j-concurrent"
terminate_cmd="/light-vm/stop-all"
collect_cmd="/light-vm/collect-1x1"
cleanup_cmd="/light-vm/delete-vm.sh"

# popis experimentu
description="World base of 2.4 branch, 200 nodes 5000 jobs, concurent submit"

# adresare obsahujici logy, ktere se maji sledovat
account_log_dir="/light-vm/vm001/var/spool/torque/server_priv/accounting/"
server_log_dir="/light-vm/vm001/var/spool/torque/server_logs/"

# ocekavany pocet uloh a pocet vypocetnich uzlu
job_count=5000
node_count=200

# generovane soubory
job_graph="/light-vm/job-graph-data"
timestamps="/light-vm/timestamps.data"
events="/light-vm/events.log"
```

Figure 1: Ukázka konfiguračního souboru pro experiment

Nejdůležitějším parametrem jsou cesty k stage skriptům. Stage skripty spouštějí jednotlivé fáze experimentu.

- install
 - v této fázi dojde k instalaci potřebného SW na bazový stroj (ze kterého se pak klonují ostatní stroje)
- prepare
 - cílem této fáze je připravit potřebný počet kontejnerů do stavu ve kterém je možné spustit experiment (běžící servery a plánovače, potřebný počet výpočetních uzlů)

- submit
 - tato fáze má za úkol do systému zadat potřebný počet testovacích úloh
- stop
 - tato fáze je vyvolána po detekci ukončení experimentu (všechny sledované úlohy doběhly, nebo 5 minut nebyla zaznamenána žádná událost) a má za cíl vypnout běžící kontejnery a připravit je na sběr dat (logy, accounting. . .)
- collect
 - tato fáze posbírání potřebné data od jednotlivých kontejnerů (převážně logy a nastavení)
- cleanup
 - závěrečný úklid, který má za cíl uvést hostitelský systém do stavu ve kterém bude možné spustit další experiment

Pro korektní detekci konce potřebuje systém znát celkový počet úloh. Experiment je ukončen po skončení všech zaregistrovaných úloh nebo pokud uběhnou 5 minutový timeout od poslední zaznamenané události.

Tento timeout byl přidán z důvodu aktuálního nedostatku v Torque kdy dochází k občasné ztrátě úlohy mezi výpočetním uzlem a serverem. V průběhu spouštění experimentů došlo k této situaci cca. v jednom z 10 až 20 experimentů (odpovídá jedné z 50tis. až 100tis. úloh).

2 Metodika

Veškeré experimenty byly provedeny na jednom fyzickém stroji (Intel(R) Core(TM)2 Duo CPU E8400 @ 3.00GHz s 4GB RAM) s využitím již zmiňované technologie lehkých kontejnerů. Jako základní nastavení byla zvolena konfigurace 1-5 serverů 1-5 plánovačů a 200 výpočetních uzlů, které mohou být asymetricky přiděleny k jednotlivým serverům. Do systému se vložilo pomocí standardních nástrojů 5000 úloh s prakticky nulovou délkou běhu (`/bin/true`), zapnutým uchováváním výstupů na straně výpočetních uzlů (v základním nastavení odesílají výpočetní uzly výstupy serveru pomocí `scp`). Všechny úlohy byly spuštěny pod jedním uživatelem (z důvodu prakticky nulové délky úloh nemá smysl uvažovat fairshare).

2.1 Záznamy experimentů

Záznamy všech experimentů jsou uloženy pro pozdější zpracování (logy + nastavení). Systém dokáže pracovat jak s běžícím experimentem tak offline s již sesbíranými logy. Následující výstupy jsou produktem, který vzniká zpracováváním logů.

Výstupem každého experimentu jsou naměřené výsledky: délka trvání zadávání úloh (doba od zadání první úlohy do zadání poslední úlohy), délka čistého běhu úloh (doba od prvního startu úlohy do ukončení poslední) a celková doba (doba od zadání první úlohy do ukončení poslední úlohy).

Dále se uchovává informace o zaregistrovaných událostech (zadání úlohy, starty úloh, konce úloh), logy serverů a plánovačů a nastavení se kterým byl experiment spuštěn a sesbírané data pro průběhový graf experimentu.

Ze zaznamenaných událostí dále systém generuje průběhový graf experimentu (viz. obrázek 2). Tento sestává ze čtyř částí. První je záznam průběhu počtu známých úloh (úlohy, které vstoupily do clusteru), počtu úloh aktuálně v systému (běžící nebo ve frontě), počet úloh ve frontách a počet ukončených úloh. Zbylé části pak zaznamenávají aktuální počet obsazených uzlů, aktuální počet spuštění úloh v dané vteřině a aktuální počet přesunů úloh v dané vteřině.

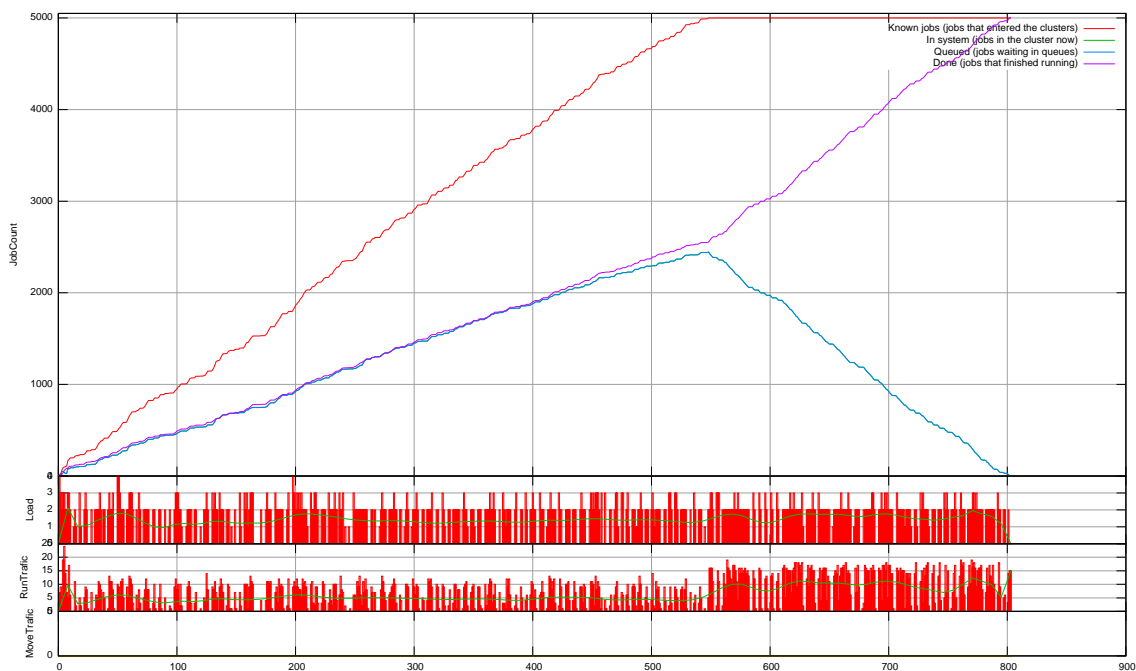


Figure 2: Průběhový graf experimentu

Pomocí zpracovávajících skriptů se dále generují statistické grafy (viz. obrázek 3) pro 7 nejlepších z 10 celkových naměřených experimentů (dle celkové délky experimentu, pro eliminaci anomálních měření).

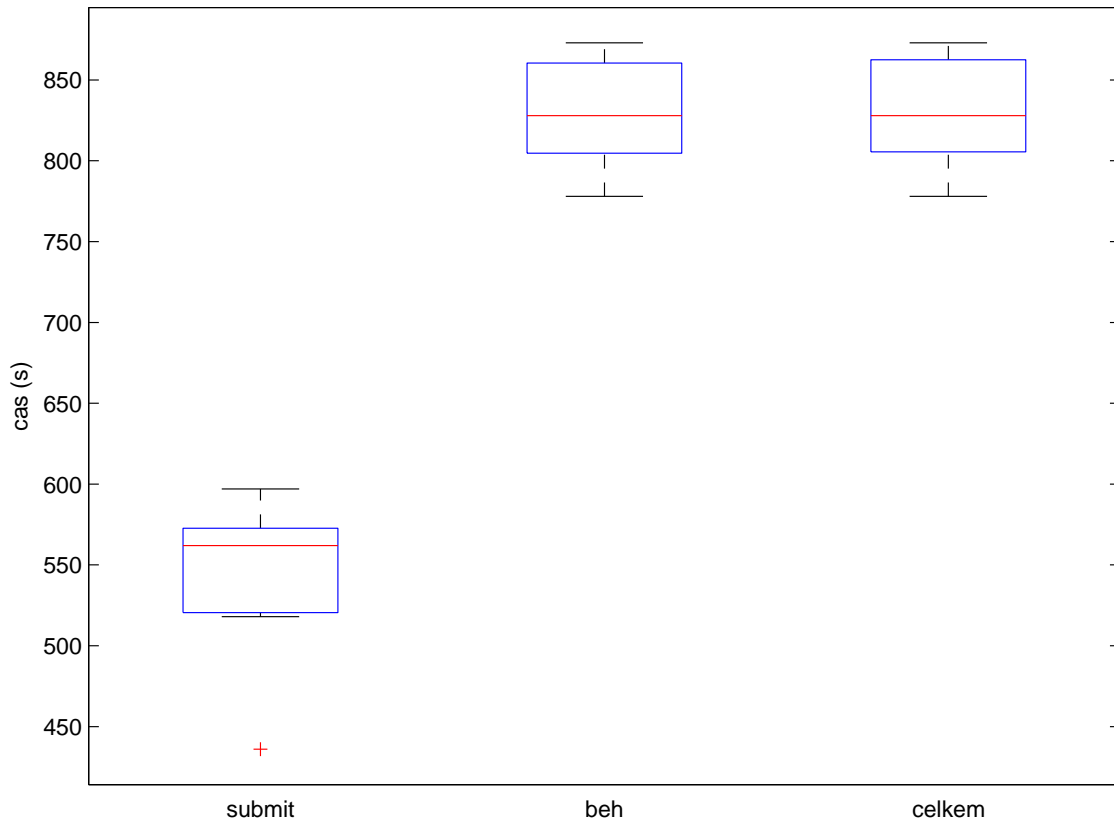


Figure 3: Statistický graf (7 nejlepších výsledků z 10 celkových měření)

Statistický graf zaznamenává pro každou měřenou hodnotu její maximum, minimum, medián a 25% a 75% kvantily (viz. obrázek 4).

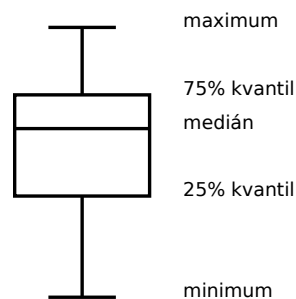


Figure 4: Význam statistického grafu

Posledním výstupem jsou předpřipravené buňky pro přehledové latexové tabulky (viz. obrázek 5).

souběžný submit				3 plánovače (zamykání)			3 plánovače (bez zamykání)		
	min	med.	max	min	med.	max	min	med.	max
submit	436	528	562	427	465	521	578	526	573
beh	778	822	847	783	861	943	836	904	953
celkem	778	822	852	783	866	944	840	904	954

souběžný submit				3 plánovače (zamykání)			3 plánovače (bez zamykání)		
	min	med.	max	min	med.	max	min	med.	max
submit	100.0%	121.1%	128.9%	97.9%	106.7%	119.5%	132.6%	120.6%	131.4%
beh	100.0%	105.7%	108.9%	100.6%	110.7%	121.2%	107.5%	116.2%	122.5%
celkem	100.0%	105.7%	109.5%	100.6%	111.3%	121.3%	108.0%	116.2%	122.6%

Figure 5: Ukázka přehledových tabulek

3 Přehled provedených experimentů

3.1 Experimenty pro porovnání jednotlivých verzí Torque a obecného vlivu implementovaných funkcí na výkon

Tento experiment měl změřit rozdíl mezi aktuální vývojovou verzí Torque a její stabilní verzí (2.3) a také vliv implementovaných vlastností na rychlost systému (v tabulkách označováno jako World patche). V průběhu testování vyšla nová stabilní verze (2.4) a proto byla do testu zahrnuta i ona.

Systém se chová velice rozdílně pokud je v souběhu se zadáváním úloh povoleno spouštění úloh. Proto byla ve všech případech otestovaná varianta jak se souběžným spouštěním úloh tak varianta, kdy je spouštění úloh povoleno až po ukončení zadávání úloh. U každého experimentu byl také otestován vliv uchovávání informací o ukončených úloh v systému (varianty bez uchovávání v tabulkách označeny jako nokeep).

Experiment	Průměrná doba běhu
Torque 2.3 Stable, souběžný submit	808 sec
Torque 2.3 Stable, dedikovaný submit	648 sec
Torque 2.3 Stable, souběžný submit, nokeep	716 sec
Torque 2.3 Stable, dedikovaný submit, nokeep	580 sec
Trunk snapshot, souběžný submit	756 sec
Trunk snapshot, dedikovaný submit	638 sec
Trunk snapshot, souběžný submit, nokeep	727 sec
Trunk snapshot, dedikovaný submit, nokeep	582 sec
Trunk snapshot + World patche, souběžný submit	798 sec
Trunk snapshot + World patche, dedikovaný submit	711 sec
Trunk snapshot + World patche, souběžný submit, nokeep	740 sec
Trunk snapshot + World patche, dedikovaný submit, nokeep	656 sec
Torque 2.4 Stable, souběžný submit	774 sec
Torque 2.4 Stable, dedikovaný submit	663 sec
Torque 2.4 Stable, souběžný submit, nokeep	722 sec
Torque 2.4 Stable, dedikovaný submit, nokeep	582 sec
Torque 2.4 Stable + World patche, souběžný submit	831 sec
Torque 2.4 Stable + World patche, dedikovaný submit	762 sec
Torque 2.4 Stable + World patche, souběžný submit, nokeep	755 sec
Torque 2.4 Stable + World patche, dedikovaný submit, nokeep	687 sec

Experimenty prokázaly měřitelné zlepšení výkonu mezi verzemi 2.3 a 2.4 (viz. statistické grafy 6 7 8 9). Nejlepší výkonu poskytuje snapshot vývojové verze.

Dedikovaný submit snižuje celkový běh experimentů, varianta nokeep opět snižuje celkový běh a snižuje rozptyl naměřených výsledků (viz. statistický graf 10).

Díky oddělenému submitu generují varianty s dedikovaným submitem výrazně odlišný průběh než varianta s souběžným submitem (viz. průběhové grafy 18 a 19).

3.2 Experimenty pro otestování vlivu souběžných qstat dotazů a komunikace plánovačů s MOM

V průběhu testování starší verze systému jsme naráželi na výkonnostní problémy. Z těchto problémů byl podezříván qstat, který byl použit pro sledování stavu experimentu. Následující experimenty měli za cíl zjistit jak moc velký tento vliv ve skutečnosti je.

K těmto testům byl přidán i test odstavení komunikace plánovačů s výpočetními uzly, kterou plánovač používá k získání podrobnějších informací o stavu uzlů.

Experiment	Průměrná doba běhu
Torque 2.4 Stable + World patche, qstat 0.5 sec perioda	836 sec
Torque 2.4 Stable + World patche, qstat 1 sec perioda	852 sec
Torque 2.4 Stable + World patche, qstat 2 sec perioda	814 sec
Torque 2.4 Stable + World patche, bez komunikace s MOM	838 sec

Experimenty prokázali, že při uvedených frekvencích (0.5-2Hz) dotazů na server nedochází k měřitelnému zpomalení běhu experimentu (viz. statistický graf 11). S podobným výsledkem skončil i test odstavení komunikace plánovačů s výpočetními uzly.

3.3 Experimenty pro otestování zátěže vznikající paralelním zapojením (1 server, víc plánovačů)

Víc plánovačů zapojených na jeden server představuje větší zátěž, protože dochází k znásobení načítání stavových informací. Zároveň dochází ke kolizím při spouštění úloh (plánovače se budou souběžně snažit spouštět stejné úlohy, uspět může ale pouze jeden z nich).

Zároveň s tímto testem byl otestován patch pro zamykání serveru (který by měl eliminovat souboj plánovačů).

Experiment	Průměrná doba běhu
Torque 2.4 Stable + World patche, 3 plánovače, bez zamykání	918 sec
Torque 2.4 Stable + World patche, 3 plánovače, se zamykáním	889 sec
Torque 2.4 Stable + World patche, 5 plánovačů, bez zamykání	934 sec
Torque 2.4 Stable + World patche, 5 plánovačů, se zamykáním	955 sec

S počtem plánovačů se postupně zpomaluje běh experimentu, nicméně tento vliv je poměrně nízký. Varianty bez zamykání vykazují mírně horší výsledky, než varianty se zamykáním (viz. statistický graf 12).

3.4 Experimenty pro otestování vlivu periody pro vynucené kontaktování plánovačů

Plánovač dostává od serveru informace o vzniklých událostech, ty jsou ale odeslány pouze pokud se serveru povede plánovač ve velmi krátké době kontaktovat. Pokud plánovač "prospí" veškeré události, je další možností probuzení plánovačů časová událost, která se odesílá vždy po uplynutí timeoutu.

Tento experiment ověřoval vliv délky uvedeného timeoutu.

Experiment	Průměrná doba běhu
Torque 2.4 Stable + World patche, 5 plánovačů, 5 sec, se zamykáním	1029 sec
Torque 2.4 Stable + World patche, 5 plánovačů, 5 sec, bez zamykání	974 sec
Torque 2.4 Stable + World patche, 5 plánovačů, 10 sec, se zamykáním	1034 sec
Torque 2.4 Stable + World patche, 5 plánovačů, 10 sec, bez zamykání	1018 sec
Torque 2.4 Stable + World patche, 5 plánovačů, 30 sec, se zamykáním	1268 sec
Torque 2.4 Stable + World patche, 5 plánovačů, 30 sec, bez zamykání	1488 sec

Vliv se postupně zvyšuje s délkou timeoutu (viz. statistický graf 14), problém který vzniká jsou období "klidu", v průběhu kterých server i plánovač čekají (viz. průběhový graf 22).

3.5 Experimenty pro otestování vlivu zapojení více serverů na jeden plánovač

S rostoucím počtem serverů zapojených na jeden plánovač roste zátěž plánovačů a vznikají potencionální problémy soubroje jednotlivých serverů.

Tento experiment měl za cíl ověřit vliv zapojení více serverů na jeden plánovač.

Experiment	Průměrná doba běhu
Torque 2.4 Stable + World patche, 2 servery, 2x poloviční submit	856 sec
Torque 2.4 Stable + World patche, 2 servery, 2x kompletní submit	1568 sec
Torque 2.4 Stable + World patche, 5 serverů, 5x pětinový submit	811 sec

Při zapojení více serverů na jeden plánovač a rovnoměrném rozložení úloh odpovídá běh experimentu přibližně době potřebné ke zpracování stejného počtu úloh na jednom serveru (viz. statistický graf 13).

3.6 Víc serverů a asymetrické zapojení

Tento experiment měl za cíl prověřit výkon v případě M:N zapojení a vliv jednotlivých vlastností implementovaných do plánovačů (konkrétně zamykání plánovače a přesuny jobů mezi servery). U podpory pro přesun se očekávalo výrazné zpomalení, protože v současné implementaci není optimalizována na práci s velkým počtem úloh s nízkými požadavky.

Experiment	Průměrná doba běhu
Torque 2.4 Stable + World patche, 2 servery, 2 plánovače, 2x poloviční submit	792 sec
Torque 2.4 Stable + World patche, 2 servery, 2 plánovače, 2x poloviční submit, zamykání plánovače	760 sec
Torque 2.4 Stable + World patche, 2 servery, 2 plánovače, 2x poloviční submit, přesuny	948 sec
Torque 2.4 Stable + World patche, 2 servery, 2 plánovače, 2x poloviční submit, přesuny, zamykání plánovače	994 sec

Při zapojení stejného počtu plánovačů a serverů dochází ke zkrácení doby běhu experimentu (pro stejný počet úloh). Zamykání plánovače opět snižuje dobu běhu experimentu (viz. obrázek 15).

Po zapnutí podpory pro přesun úloh mezi servery dochází k výraznému zpomalení. Zamykání plánovače je v této variantě neúčinný, naopak výsledek ještě mírně zhoršil. Díky rozsáhlému množství volných uzlů si servery neustále přehazovaly úlohy mezi sebou (viz. obrázek 23).

3.7 Úlohy s nenulovou délkou

Pro porovnání byl proveden experiment s úlohami délky 1 sec (`sleep 1`).

Experiment	Průměrná doba běhu
Torque 2.4 Stable + World patche, 2 servery, 2 plánovače, 1 sec úlohy	805 sec
Torque 2.4 Stable + World patche, 2 servery, 2 plánovače, 1 sec úlohy, zamykání plánovače	835 sec
Torque 2.4 Stable + World patche, 2 servery, 2 plánovače, 1 sec úlohy, přesuny	791 sec
Torque 2.4 Stable + World patche, 2 servery, 2 plánovače, 1 sec úlohy, přesuny, zamykání plánovače	816 sec

I při experimentu s úlohami délky 1 sekund zůstává cca. 90% uzlů neobsazených (viz. obrázek 24). Na době běhu experimentu se toto nastavení v podstatě neprojevovalo.

3.8 Porovnání s PBS pro

PBS Pro v testovacím prostředí vykazovalo nečekaně nízký výkon, pokud bylo k serveru připojeno víc uzlů. Tento problém byl způsoben pravděpodobně samotnou virtualizací, nebo privátními adresami pod kterými server a uzly běžely (konkrétní zdroj problému byl ponechán k

pozdějšímu prozkoumání). Z tohoto důvodu byl otestováno pouze zapojení: jeden server, jeden scheduler, jeden výpočetní uzel.

Experiment	Průměrná doba běhu
PBS Pro 1 CPU	1704 sec
PBS Pro 4 CPU	939 sec
PBS Pro 8 CPU	705 sec
Torque 2.4 Stable 1 CPU	2315 sec
Torque 2.4 Stable 4 CPU	807 sec
Torque 2.4 Stable 8 CPU	834 sec

Naměřené hodnoty PBS Pro odpovídají očekávaným hodnotám (dle zkušeností z provozu), Torque dosahuje velice podobný výkon (viz. obrázek 17). Průběh experimentu je převážně způsoben mnohem vyšší náhodností chování Torque vůči PBS Pro (viz. obrázek 25, 27 a 29 vs. 26, 28 a 30).

4 Závěr

Cílem tohoto dokumentu bylo zmapovat výkon jednotlivých verzí batch systému Torque, vliv jednotlivých nastavení serveru a scheduleru na tento výkon a v neposlední řadě také vliv implementovaných rozšíření a jejich celkovou stabilitu.

V rámci tohoto procesu byl sestaven testovací systém který umožnil provést všechny testy na jednom fyzickém stroji. Tento systém bude díky své modularitě v budoucnu požit k ověření stability a výkonu nových vlastností přidávaných do Torque.

Z jednotlivých experimentů byli získané následující znalosti:

- stabilní verze vykazují slabší výkon než verze vývojové
- mladší stabilní verze vykazují vyšší výkon
- implementované patche zpomalují běh, nicméně tento vliv není drastický
- očekávaný vliv periodické načítání stavu serveru nebyl na měřených periodách potvrzen
- zapojení více schedulerů na jeden server snižuje výkon serveru, nicméně pouze mírně
- při zapojení více serverů na jeden scheduler a rovnoměrnému rozdělení úloh dochází k zanedbatelnému zrychlení
- naopak znásobení počtu úloh (při rovnoměrném rozdělení) způsobí prakticky identické (odpovídající násobku) zpomalení
- zvyšování timeoutu pro kontaktování scheduleru způsobuje náhodně rozložené prodlevy
- zapojení M:N pro dva servery a dva schedulery dosahuje lepších výsledků než zapojení 1:1 pro stejný počet úloh
- aktuální implementace přesunů úloh mezi servery není vhodná pro velký počet úloh s nízkými požadavky
- systém se chová prakticky identicky při úlohách s nulovou a jednovteřinovou délkou
- PBS Pro je výkonově stabilnější a také výkonnější než Torque, nicméně rozdíl ve výkonu není nikterak propastný

Budoucí práce se bude soustředit na lokalizaci a odstranění původců naměřených výkonnostních nedostatků.

A Shrnující tabulky

	Stabilní verze 2.3.7			Stabilní verze 2.4.2			Trunk snapshot			Trunk snapshot + Patch			Stabilní 2.4.2 + Patch		
	min	med.	max	min	med.	max	min	med.	max	min	med.	max	min	med.	max
souběžný submit	submit	404	441	475	463	533	480	427	435	453	525	459	436	528	562
	beh	740	804	821	740	751	791	718	744	752	761	790	778	822	847
	celkem	741	804	821	741	752	791	718	745	753	761	790	778	822	852
dedikovaný submit	submit	108	138	135	125	151	160	111	113	155	115	135	123	159	121
	beh	525	496	506	507	518	514	489	521	497	546	557	595	597	656
	celkem	634	635	642	633	670	674	601	634	653	675	692	720	757	778
souběžný submit, nokeep	submit	402	446	520	410	412	488	487	516	546	414	455	508	452	475
	beh	663	699	736	676	703	747	676	722	735	686	733	729	751	762
	celkem	663	699	737	677	703	747	677	723	735	686	734	729	751	762
dedikovaný submit, nokeep	submit	110	145	152	126	158	164	146	109	133	105	126	148	156	147
	beh	423	433	445	414	420	432	388	443	446	495	518	510	535	550
	celkem	533	579	598	542	579	597	534	552	580	613	657	636	692	698

	Stabilní verze 2.3.7			Stabilní verze 2.4.2			Trunk snapshot			Trunk snapshot + Patch			Stabilní 2.4.2 + Patch		
	min	med.	max	min	med.	max	min	med.	max	min	med.	max	min	med.	max
souběžný submit	submit	100.0%	109.2%	117.6%	114.6%	131.9%	118.8%	107.7%	107.7%	112.1%	130.0%	113.6%	123.5%	130.7%	139.1%
	beh	100.0%	108.6%	110.9%	100.0%	101.5%	106.9%	97.0%	100.5%	101.6%	102.8%	106.8%	110.3%	111.1%	114.5%
	celkem	100.0%	108.5%	110.8%	100.0%	101.5%	106.7%	96.9%	100.5%	101.6%	102.7%	106.6%	110.1%	110.9%	115.0%
dedikovaný submit	submit	100.0%	127.8%	125.0%	115.7%	139.8%	148.1%	102.8%	104.6%	143.5%	106.5%	125.0%	117.6%	147.2%	112.0%
	beh	100.0%	94.5%	96.4%	96.6%	98.7%	97.9%	93.1%	99.2%	94.7%	104.0%	106.1%	111.4%	113.7%	125.0%
	celkem	100.0%	100.2%	101.3%	99.8%	105.7%	106.3%	94.8%	100.0%	103.0%	106.5%	109.1%	114.4%	119.4%	122.7%
souběžný submit, nokeep	submit	100.0%	110.9%	129.4%	102.0%	102.5%	121.4%	121.1%	128.4%	135.8%	103.0%	113.2%	129.1%	112.4%	118.2%
	beh	100.0%	105.4%	111.0%	102.0%	106.0%	112.7%	102.0%	108.9%	110.9%	103.5%	110.6%	113.7%	113.3%	114.9%
	celkem	100.0%	105.4%	111.2%	102.1%	106.0%	112.7%	102.1%	109.0%	110.9%	103.5%	110.7%	113.3%	113.3%	114.9%
dedikovaný submit, nokeep	submit	100.0%	131.8%	138.2%	114.5%	143.6%	149.1%	132.7%	99.1%	120.9%	95.5%	114.5%	134.5%	141.8%	133.6%
	beh	100.0%	102.4%	105.2%	97.9%	99.3%	102.1%	91.7%	104.7%	105.4%	117.0%	122.5%	119.1%	126.5%	130.0%
	celkem	100.0%	108.6%	112.2%	101.7%	108.6%	112.0%	100.2%	103.6%	108.8%	115.0%	123.3%	124.6%	129.8%	131.0%

souběžný submit		dedikovaný submit		souběžný submit, nokeep		dedikovaný submit, nokeep		bez komunikace s MOM								
min	med.	max	min	med.	max	min	med.	max	min	med.	max					
submit	436	528	562	123	159	121	508	452	475	125	156	147	472	586	570	
beh	778	822	847	595	597	656	729	751	762	510	535	550	792	844	855	
celkem	778	822	852	720	757	778	729	751	762	636	692	698	793	844	855	
souběžný submit		souběžný qstat (0.5sec)		souběžný qstat (1 sec)		souběžný qstat (2 sec)										
min	med.	max	min	med.	max	min	med.	max	min	med.	max					
submit	436	528	562	442	541	489	468	486	514	441	504	453				
beh	778	822	847	794	813	864	810	856	872	779	816	822				
celkem	778	822	852	794	813	866	810	856	872	779	816	823				
souběžný submit		3 plánovače (zamykání)		3 plánovače (bez zamykání)		5 plánovačů (zamykání)		5 plánovačů (bez zamykání)								
min	med.	max	min	med.	max	min	med.	max	min	med.	max	min	med.	max		
submit	436	528	562	427	465	521	578	526	573	497	539	496	572	585	534	
beh	778	822	847	783	861	943	836	904	953	884	921	962	850	914	954	
celkem	778	822	852	783	866	944	840	904	954	884	922	963	851	914	955	
souběžný submit		5 plánovačů, 5sec (zamykání)		5 plánovačů, 5sec (bez zamykání)		5 plánovačů, 10sec (zamykání)		5 plánovačů, 10sec (bez zamykání)								
min	med.	max	min	med.	max	min	med.	max	min	med.	max	min	med.	max		
submit	436	528	562	389	593	472	567	627	499	631	627	558	564	564	1709	
beh	778	822	847	929	1004	1041	938	947	973	1139	1358	1709	1713	1713		
celkem	778	822	852	929	1004	1043	940	948	974	1186	1231	1276	1140	1359	1713	
souběžný submit		5 plánovačů, 10sec (bez zamykání)		5 plánovačů, 30sec (zamykání)		5 plánovačů, 30sec (bez zamykání)										
min	med.	max	min	med.	max	min	med.	max	min	med.	max					
submit	436	528	562	519	586	553	457	632	631	627	558	564				
beh	778	822	847	976	1007	1025	1186	1230	1275	1139	1358	1709				
celkem	778	822	852	977	1008	1025	1186	1231	1276	1140	1359	1713				
souběžný submit		2 servery, 2x poloviční submit		2 servery, 2x plný submit		5 serverů, 5x pětinnový submit										
min	med.	max	min	med.	max	min	med.	max	min	med.	max					
submit	436	528	562	176	172	211	597	573	685	280	306	258				
beh	778	822	847	832	854	864	1187	1546	1580	744	764	783				
celkem	778	822	852	833	854	865	1473	1547	1580	798	808	813				
souběžný submit		2 servery, 1 plánovač		2 servery, 2plánovače		2 servery, 2plánovače, zamykání										
min	med.	max	min	med.	max	min	med.	max	min	med.	max					
submit	436	528	562	176	172	211	274	280	374	275	272	314				
beh	778	822	847	832	854	864	746	716	816	661	717	719				
celkem	778	822	852	833	854	865	747	761	817	730	757	766				
souběžný submit		2 servery, 2plánovače, moving		2 servery, 2plánovače, moving, zamykání												
min	med.	max	min	med.	max	min	med.	max	min	med.	max					
submit	436	528	562	579	629	653	624	635	665	624	635	665				
beh	778	822	847	879	864	903	928	992	1009	928	992	1009				
celkem	778	822	852	879	881	995	944	993	1009	944	993	1009				

souběžný submit		2 servery, 2plánovače, 1sec			2 servery, 2 plánovače, 1sec, zamykání										
	min	med.	max	min	med.	max									
submit	436	528	562	351	260	273									
beh	778	822	847	755	771	825									
celkem	778	822	852	768	838	846									
		2 servery, 2plánovače, 1sec, moving			2 servery, 2 plánovače, 1sec, moving, zamykání										
	min	med.	max	min	med.	max									
	615	644	649	598	687	700									
	892	890	966	845	910	984									
	931	937	991	915	974	992									
souběžný submit		dedikovaný submit			souběžný submit, nokeep			dedikovaný submit, nokeep			bez komunikace s MOM				
	min	med.	max	min	med.	max	min	med.	max	min	med.	max	min	med.	max
submit	100.0%	121.1%	128.9%	28.2%	36.5%	27.8%	116.5%	103.7%	108.9%	28.7%	35.8%	33.7%	108.3%	134.4%	130.7%
beh	100.0%	105.7%	108.9%	76.5%	76.7%	84.3%	93.7%	96.5%	97.9%	65.6%	68.8%	70.7%	101.8%	108.5%	109.9%
celkem	100.0%	105.7%	109.5%	92.5%	97.3%	100.0%	93.7%	96.5%	97.9%	81.7%	88.9%	89.7%	101.9%	108.5%	109.9%
souběžný submit		souběžný qstat (0.5sec)			souběžný qstat (1 sec)			souběžný qstat (2 sec)							
	min	med.	max	min	med.	max	min	med.	max	min	med.	max	min	med.	max
submit	100.0%	121.1%	128.9%	101.4%	124.1%	112.2%	107.3%	111.5%	117.9%	101.1%	115.6%	103.9%	101.1%	115.6%	103.9%
beh	100.0%	105.7%	108.9%	102.1%	104.5%	111.1%	104.1%	110.0%	112.1%	100.1%	104.9%	105.7%	100.1%	104.9%	105.7%
celkem	100.0%	105.7%	109.5%	102.1%	104.5%	111.3%	104.1%	110.0%	112.1%	100.1%	104.9%	105.8%	100.1%	104.9%	105.8%
souběžný submit		3 plánovače (zamykání)			3 plánovače (bez zamykání)			5 plánovačů (zamykání)			5 plánovačů (bez zamykání)				
	min	med.	max	min	med.	max	min	med.	max	min	med.	max	min	med.	max
submit	100.0%	121.1%	128.9%	97.9%	106.7%	119.5%	132.6%	120.6%	131.4%	114.0%	123.6%	113.8%	131.2%	134.2%	122.5%
beh	100.0%	105.7%	108.9%	100.6%	110.7%	121.2%	107.5%	116.2%	122.5%	113.6%	118.4%	123.7%	109.3%	117.5%	122.6%
celkem	100.0%	105.7%	109.5%	100.6%	111.3%	121.3%	108.0%	116.2%	122.6%	113.6%	118.5%	123.8%	109.4%	117.5%	122.8%
souběžný submit		5 plánovačů, 5sec (zamykání)			5 plánovačů, 5sec (bez zamykání)			5 plánovačů, 10sec (zamykání)			5 plánovačů, 10sec (bez zamykání)				
	min	med.	max	min	med.	max	min	med.	max	min	med.	max	min	med.	max
submit	100.0%	121.1%	128.9%	89.2%	136.0%	108.3%	130.0%	143.8%	114.4%	112.2%	121.8%	122.7%	131.2%	134.2%	122.5%
beh	100.0%	105.7%	108.9%	119.4%	129.0%	133.8%	120.6%	121.7%	125.1%	122.5%	128.1%	137.8%	109.3%	117.5%	122.6%
celkem	100.0%	105.7%	109.5%	119.4%	129.0%	134.1%	120.8%	121.9%	125.2%	122.5%	128.5%	137.9%	109.4%	117.5%	122.8%
		5 plánovačů, 10sec (bez zamykání)			5 plánovačů, 30sec (zamykání)			5 plánovačů, 30sec (bez zamykání)							
	min	med.	max	min	med.	max	min	med.	max	min	med.	max	min	med.	max
submit	119.0%	134.4%	126.8%	104.8%	145.0%	144.7%	143.8%	145.0%	144.7%	143.8%	128.0%	129.4%	143.8%	128.0%	129.4%
beh	125.4%	129.4%	131.7%	152.4%	158.1%	163.9%	152.4%	158.1%	163.9%	146.4%	174.6%	219.7%	146.4%	174.6%	219.7%
celkem	125.6%	129.6%	131.7%	152.4%	158.2%	164.0%	152.4%	158.2%	164.0%	146.5%	174.7%	220.2%	146.5%	174.7%	220.2%
souběžný submit		2 servery, 2x poloviční submit			2 servery, 2x plný submit			5 serverů, 5x pětinnový submit							
	min	med.	max	min	med.	max	min	med.	max	min	med.	max	min	med.	max
submit	100.0%	121.1%	128.9%	40.4%	39.4%	48.4%	136.9%	131.4%	157.1%	64.2%	70.2%	59.2%	64.2%	70.2%	59.2%
beh	100.0%	105.7%	108.9%	106.9%	109.8%	111.1%	152.6%	198.7%	203.1%	95.6%	98.2%	100.6%	95.6%	98.2%	100.6%
celkem	100.0%	105.7%	109.5%	107.1%	109.8%	111.2%	189.3%	198.8%	203.1%	102.6%	103.9%	104.5%	102.6%	103.9%	104.5%

souběžný submit		2 servery, 1 plánovač			2 servery, 2plánovače			2 servery, 2 plánovače, zamykání				
	min	med.	max	min	med.	max	min	med.	max	min	med.	max
submit	100.0%	121.1%	128.9%	40.4%	39.4%	48.4%	62.8%	64.2%	85.8%	63.1%	62.4%	72.0%
beh	100.0%	105.7%	108.9%	106.9%	109.8%	111.1%	95.9%	92.0%	104.9%	85.0%	92.2%	92.4%
celkem	100.0%	105.7%	109.5%	107.1%	109.8%	111.2%	96.0%	97.8%	105.0%	93.8%	97.3%	98.5%
		2 servery, 2plánovače, moving			2 servery, 2plánovače, moving			2 servery, 2plánovače, moving, zamykání				
	min	med.	max	min	med.	max	min	med.	max	min	med.	max
	132.8%	144.3%	149.8%	113.0%	111.1%	116.1%	113.0%	113.2%	127.9%	143.1%	145.6%	152.5%
	113.0%	111.1%	116.1%	113.0%	113.2%	127.9%	121.3%	127.9%	129.7%	121.3%	127.6%	129.7%
souběžný submit		2 servery, 2plánovače, 1sec			2 servery, 2plánovače, 1sec, zamykání							
	min	med.	max	min	med.	max	min	med.	max	min	med.	max
submit	100.0%	121.1%	128.9%	80.5%	59.6%	62.6%	67.4%	82.6%	90.4%	67.4%	82.6%	90.4%
beh	100.0%	105.7%	108.9%	96.8%	98.6%	97.6%	97.0%	99.1%	106.0%	99.1%	106.0%	106.0%
celkem	100.0%	105.7%	109.5%	100.3%	101.4%	104.1%	98.7%	107.7%	108.7%	98.7%	107.7%	108.7%
		2 servery, 2plánovače, 1sec, moving			2 servery, 2plánovače, 1sec, moving, zamykání							
	min	med.	max	min	med.	max	min	med.	max	min	med.	max
	141.1%	147.7%	148.9%	114.1%	114.4%	124.2%	137.2%	157.6%	160.6%	137.2%	157.6%	160.6%
	114.7%	114.4%	124.2%	119.7%	120.4%	127.4%	108.6%	117.0%	126.5%	108.6%	117.0%	126.5%
	119.7%	120.4%	127.4%	117.6%	125.2%	127.5%	117.6%	125.2%	127.5%	117.6%	125.2%	127.5%

B Přehledové statistické grafy

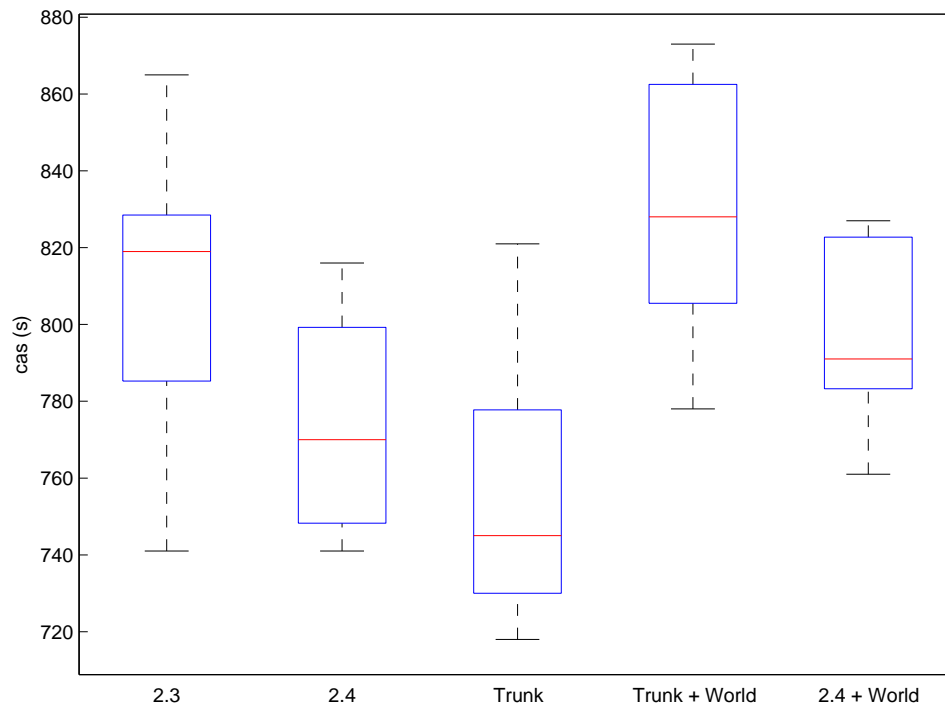


Figure 6: Porovnání verzí - souběžný submit

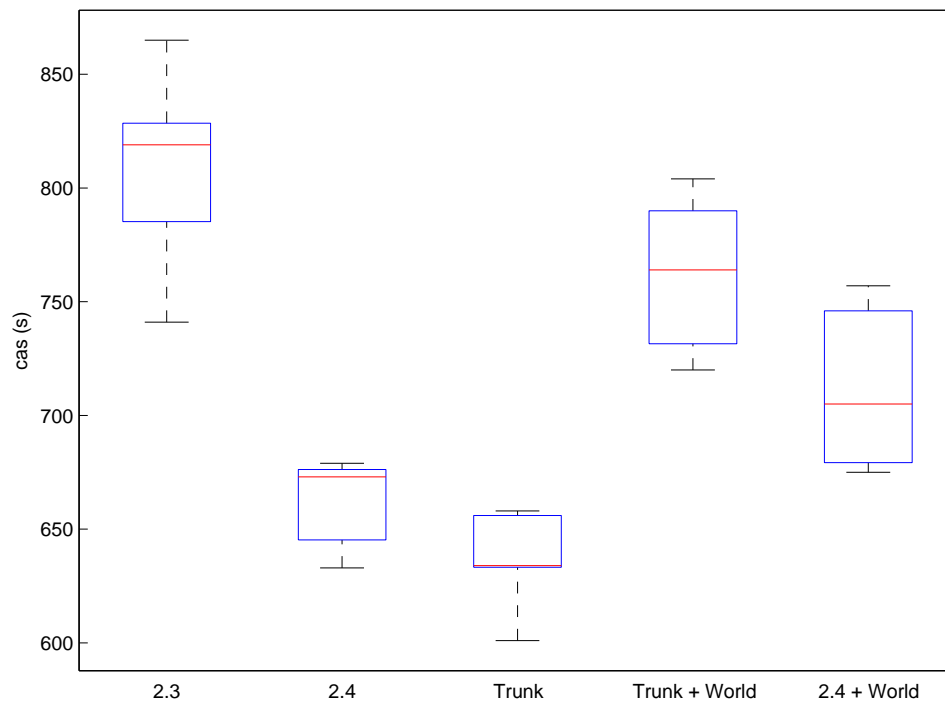


Figure 7: Porovnání verzí - dedikovaný submit

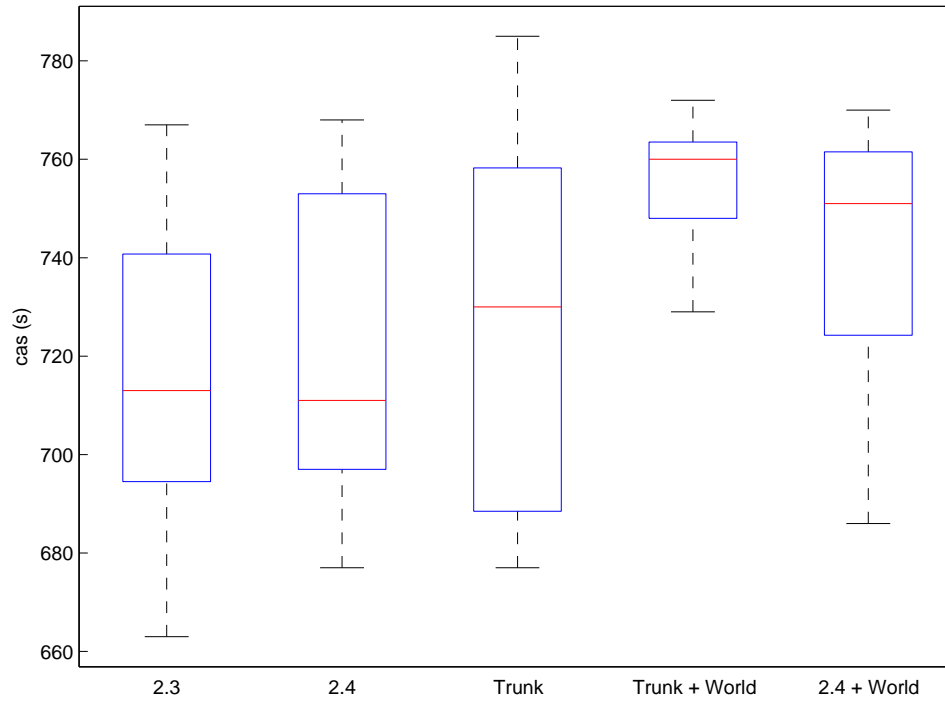


Figure 8: Porovnání verzí - souběžný submit, nokeep

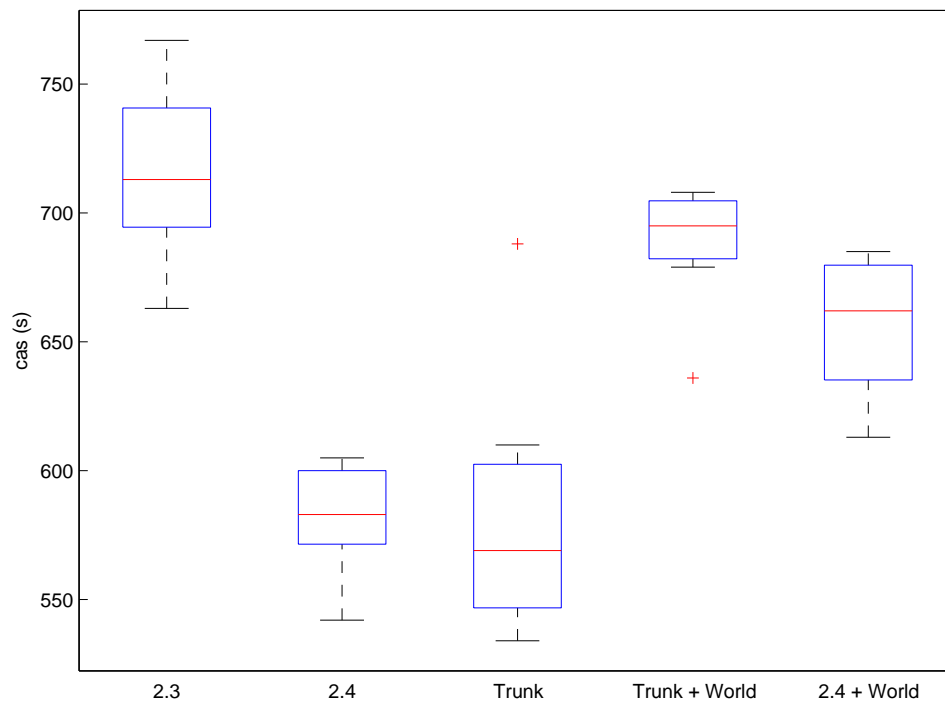


Figure 9: Porovnání verzí - dedikovaný submit, nokeep

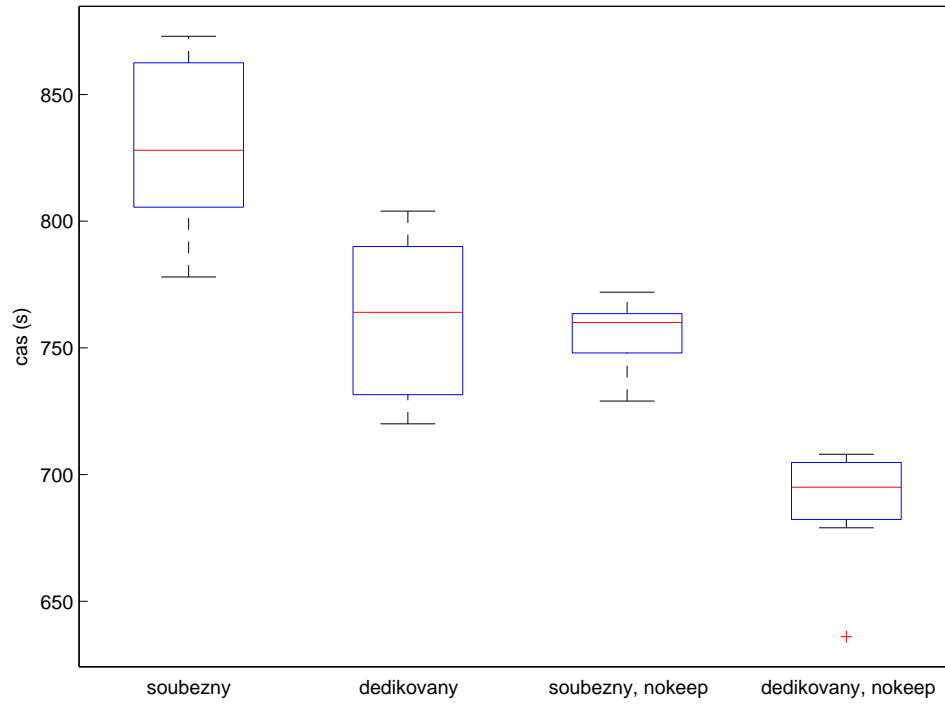


Figure 10: World patche 2.4, různé konfigurace

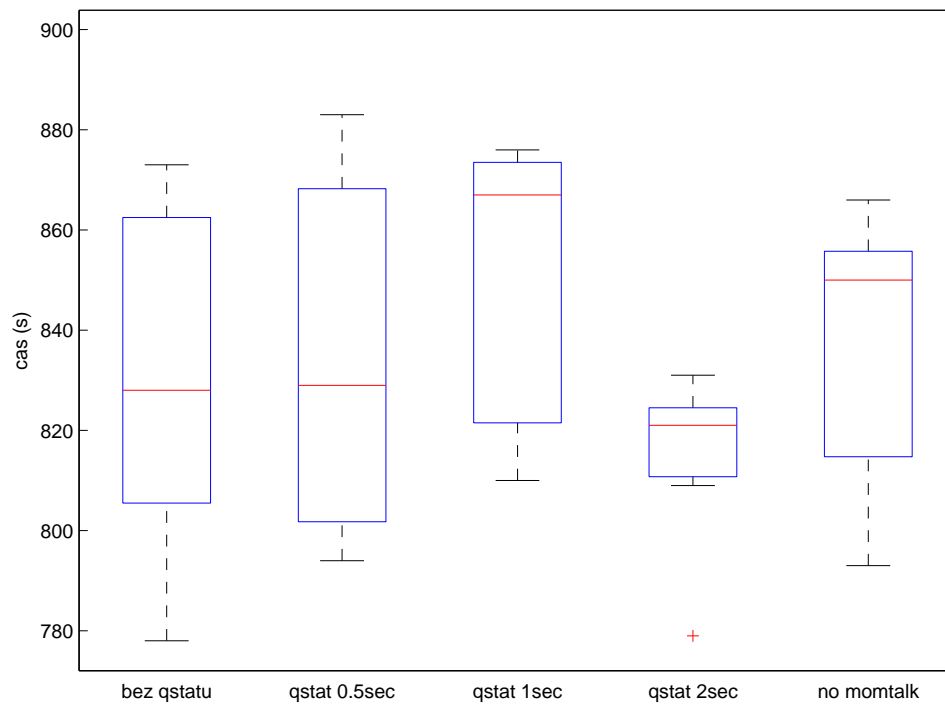


Figure 11: Vliv qstatu

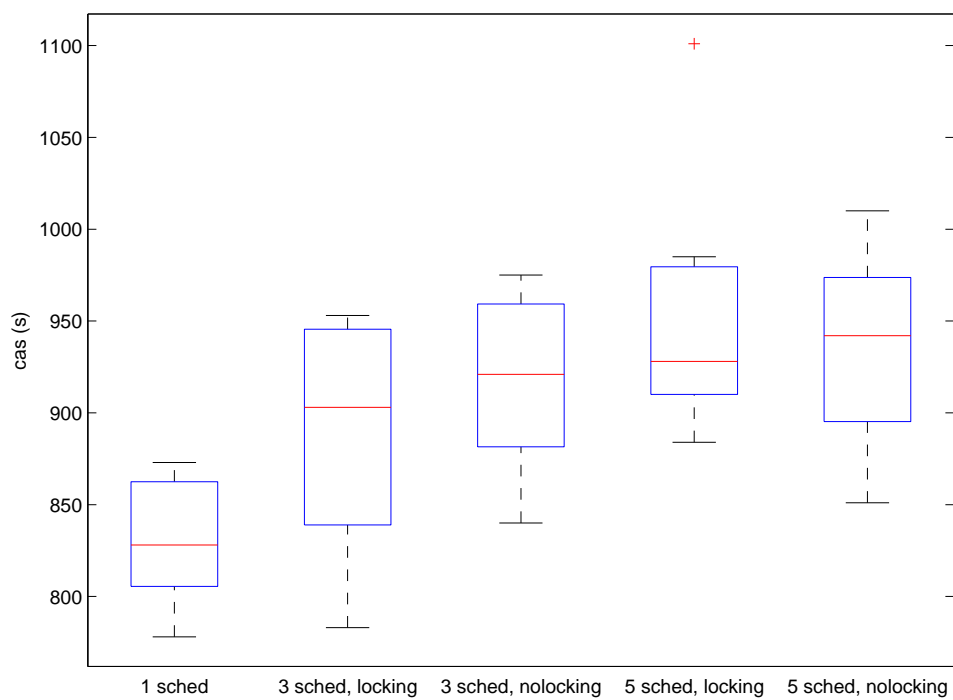


Figure 12: Zapojení více plánovačů

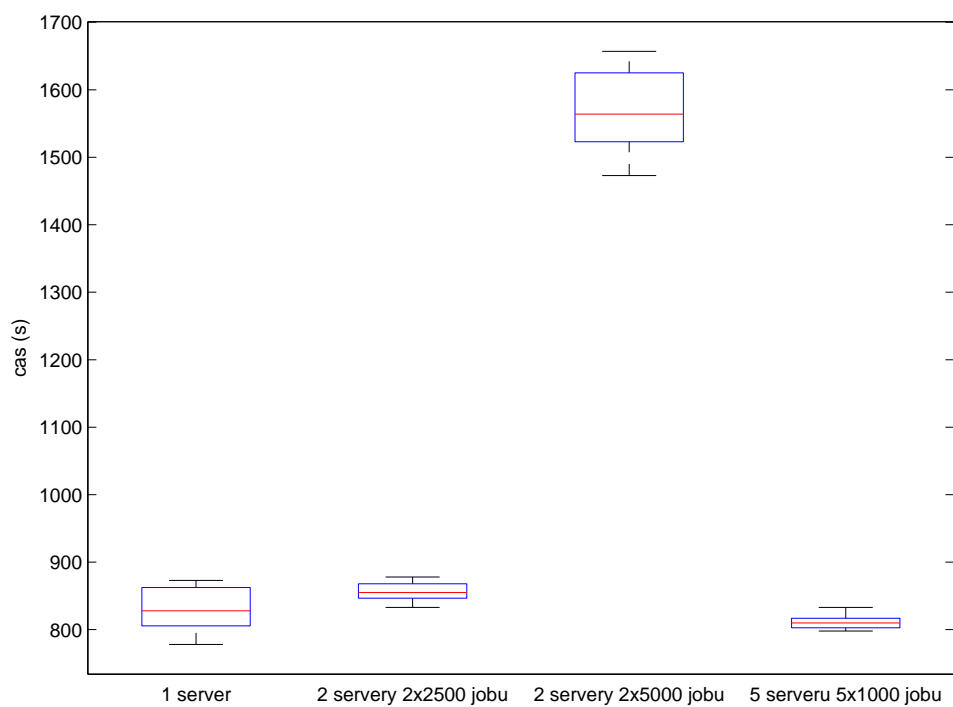


Figure 13: Zapojení více serverů

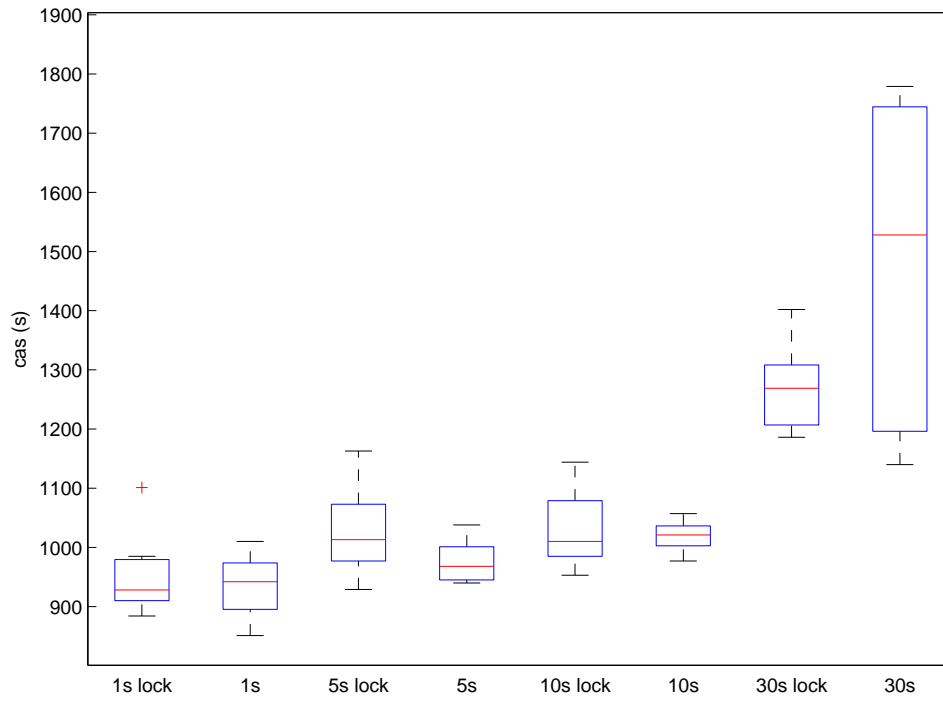


Figure 14: Vliv timeoutu pro kontaktování plánovačů (5 plánovačů)

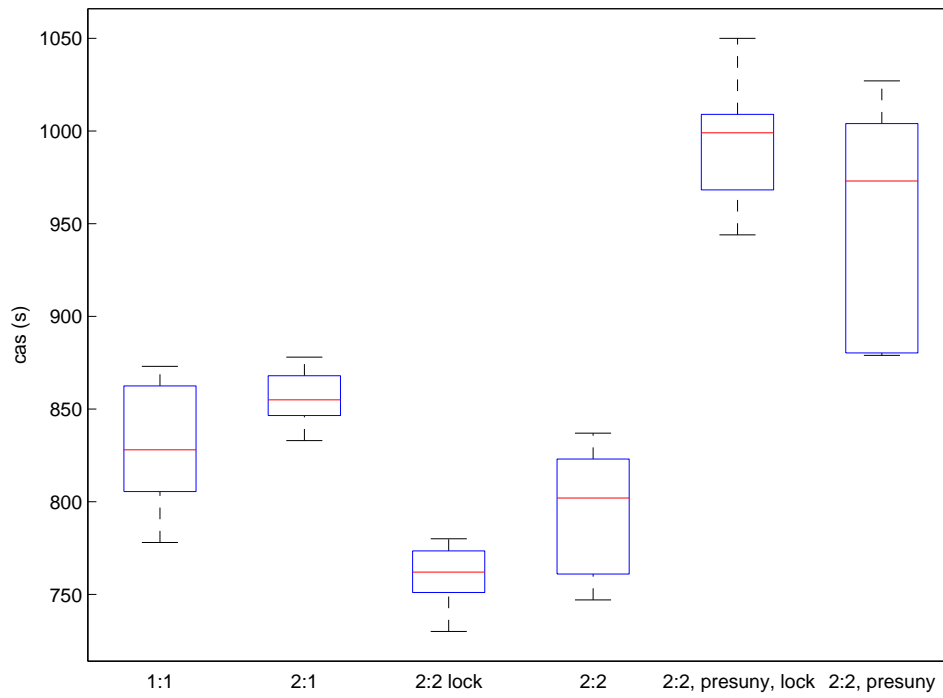


Figure 15: Zapojení M:N (počet serverů:počet plánovačů)

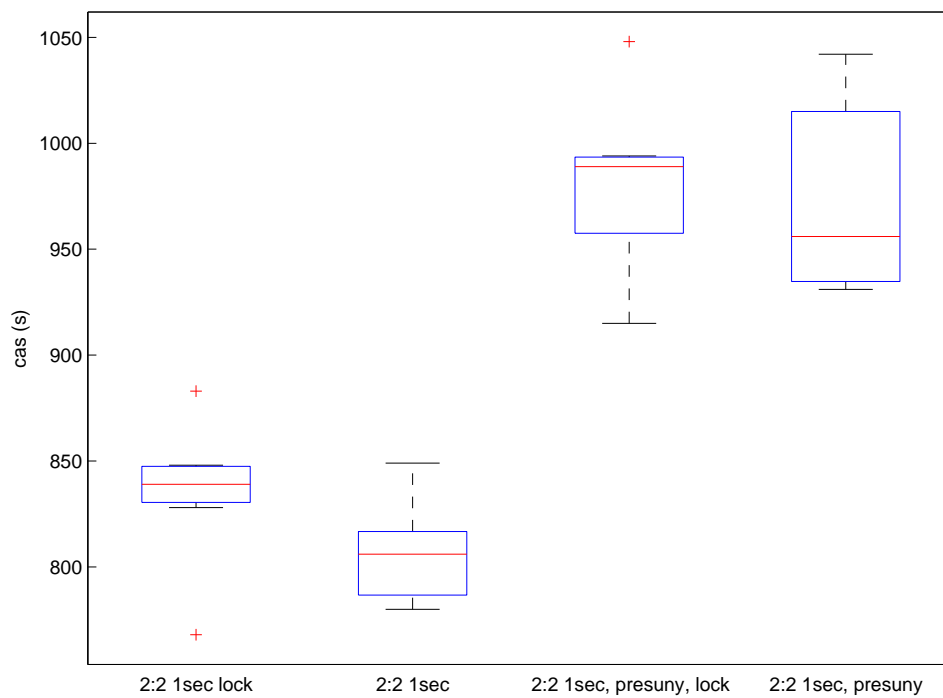


Figure 16: Nenulová délka úloh

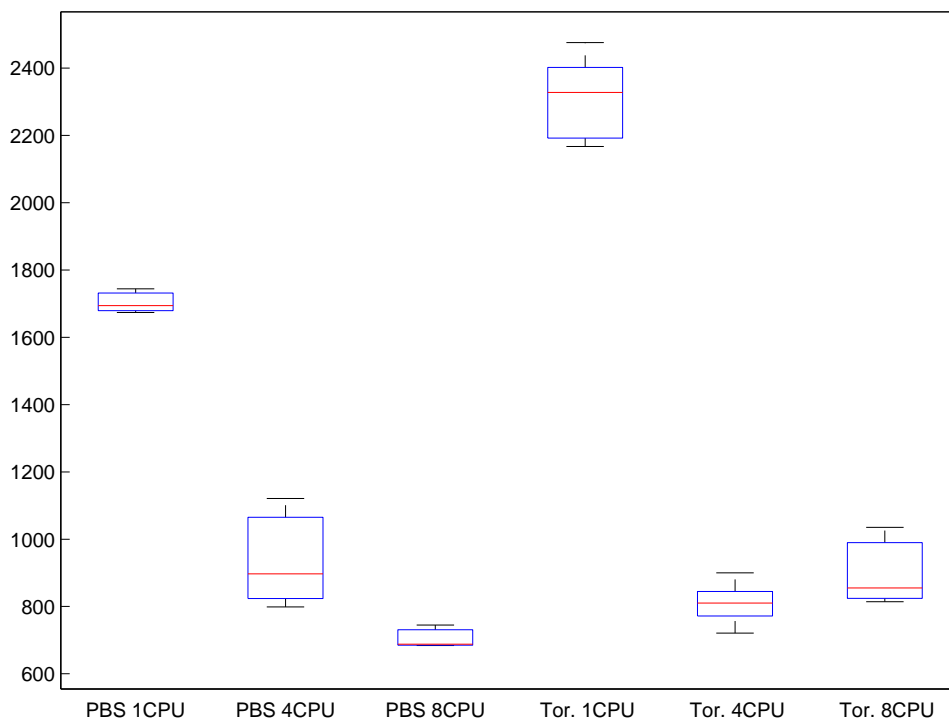


Figure 17: Porovnání Torque s PBS Pro

C Průběhové grafy

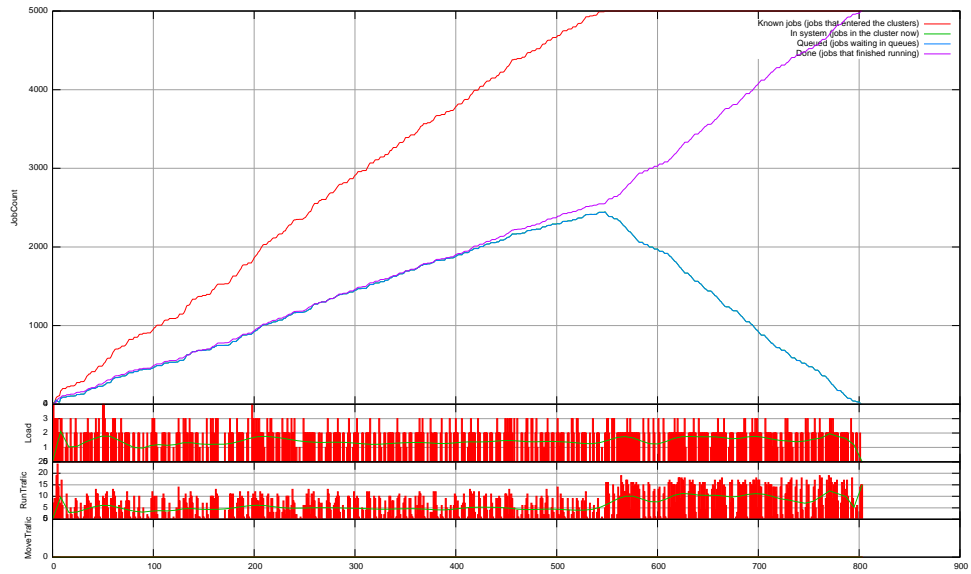


Figure 18: Průběhový graf - 2.4 - souběžný submit

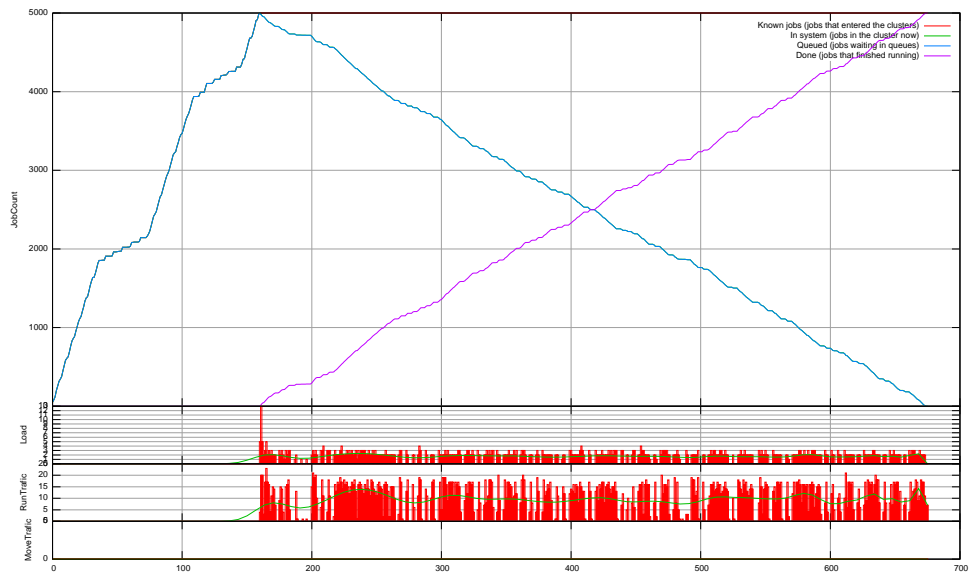


Figure 19: Průběhový graf - 2.4 - dedikovaný submit

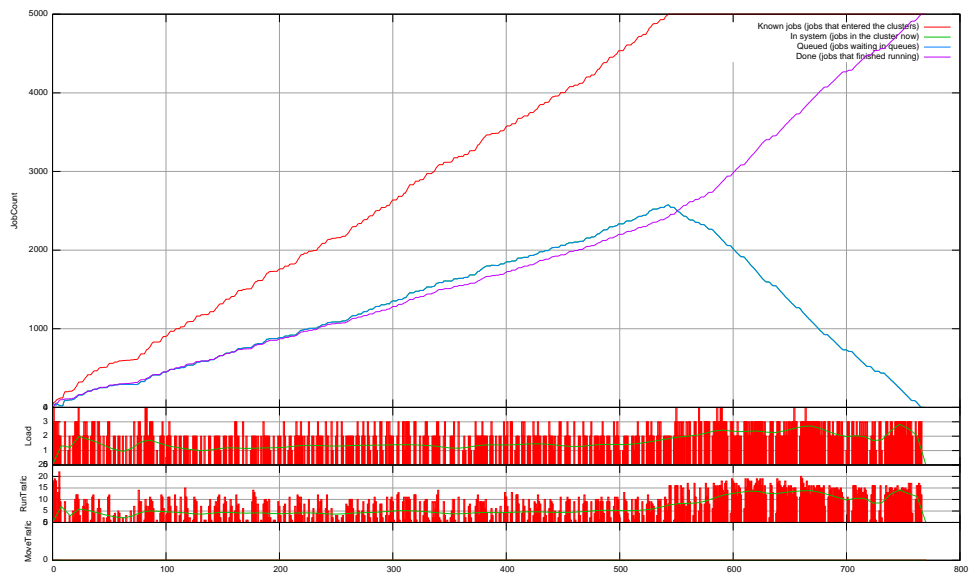


Figure 20: Průběhový graf - 2.4 - souběžný submit, nokeep

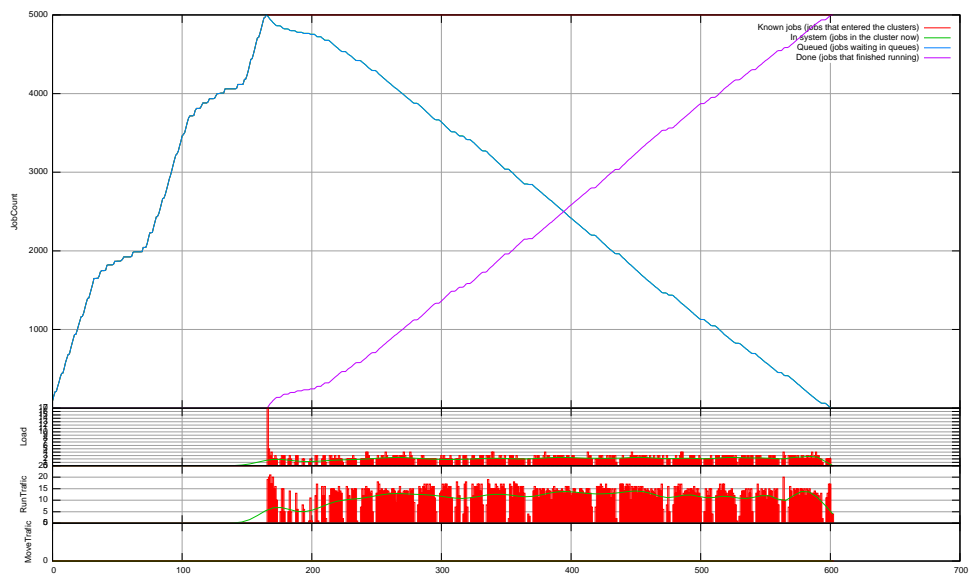


Figure 21: Průběhový graf - 2.4 - dedikovaný submit, nokeep

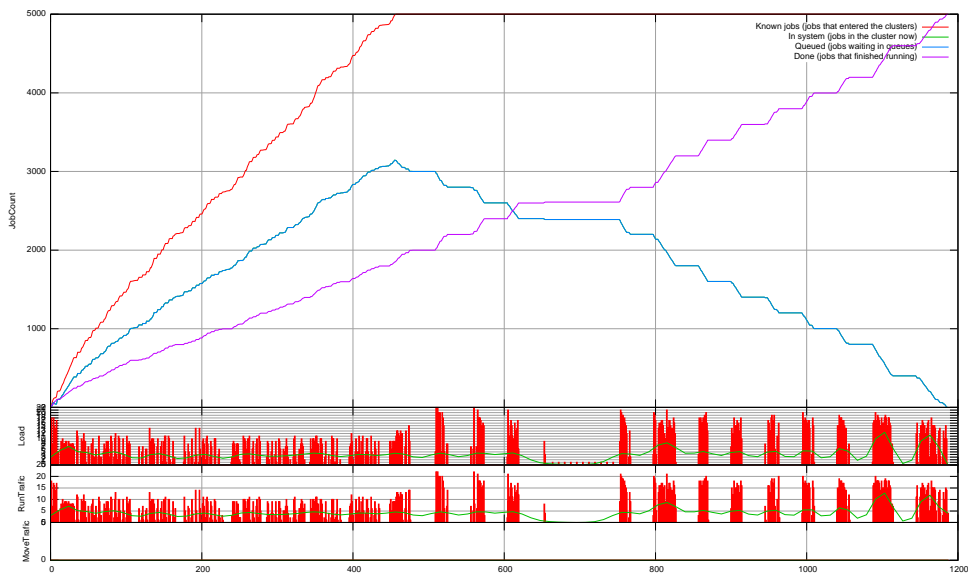


Figure 22: Průběhový graf - World patche - 2.4 - 30 sec wakeup 5 plánovačů, zamykání

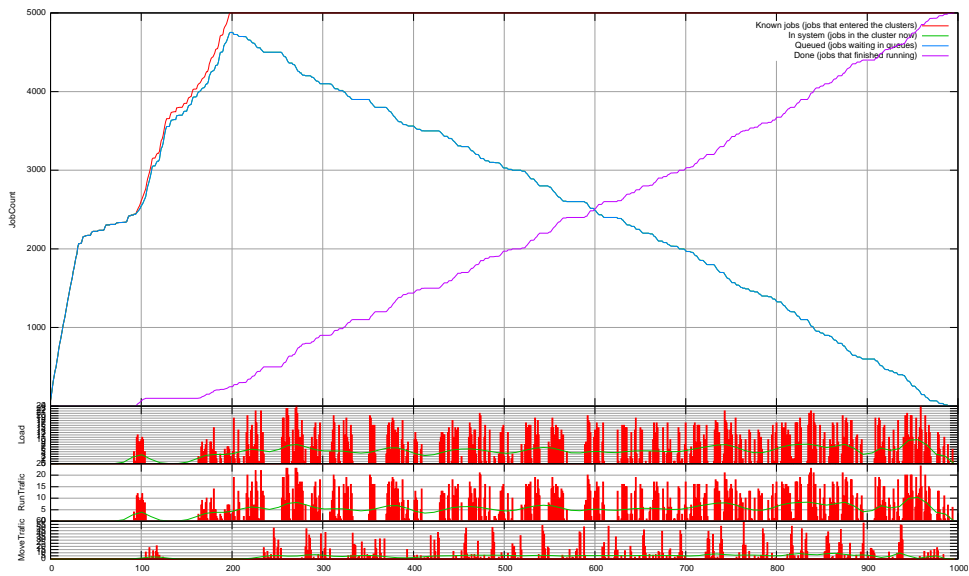


Figure 23: Průběhový graf - World patche - 2.4 - 2 servery, 2 plánovače, přesuny úloh

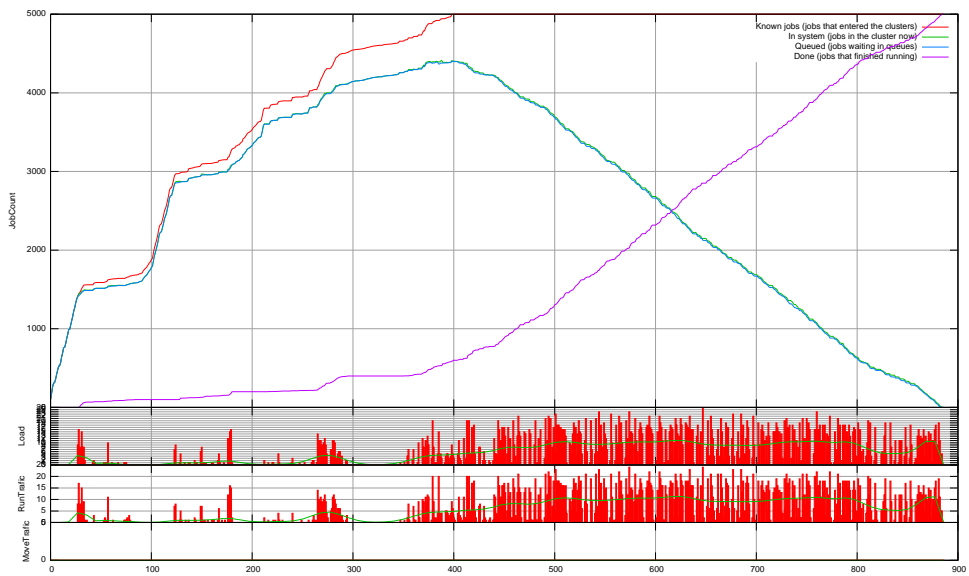


Figure 24: Průběhový graf - World patche - 2.4 - 2 server, 2 plánovače, 1sec úlohy

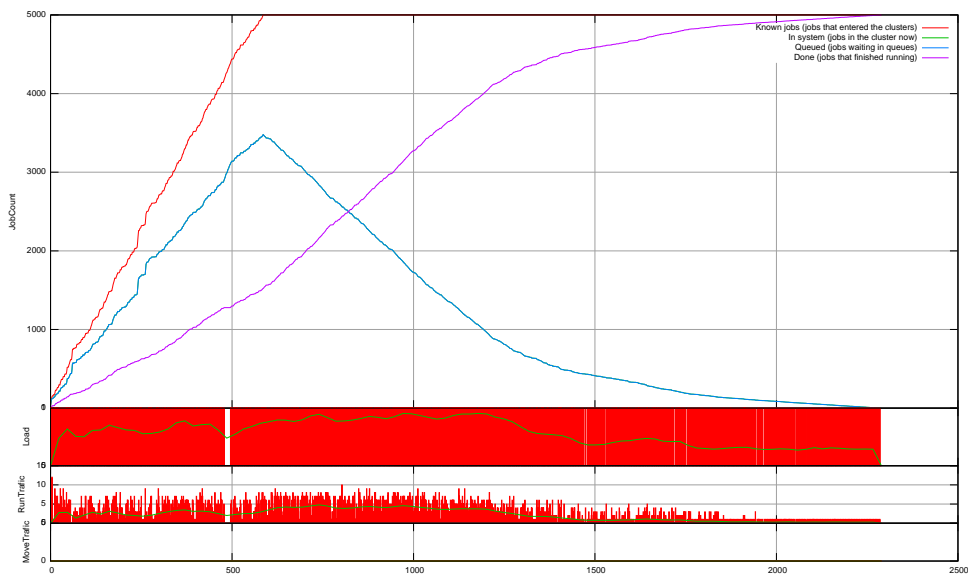


Figure 25: Průběhový graf - Torque 2.4 Stable - 1 server, 1 scheduler, 1 MOM, 14 CPU

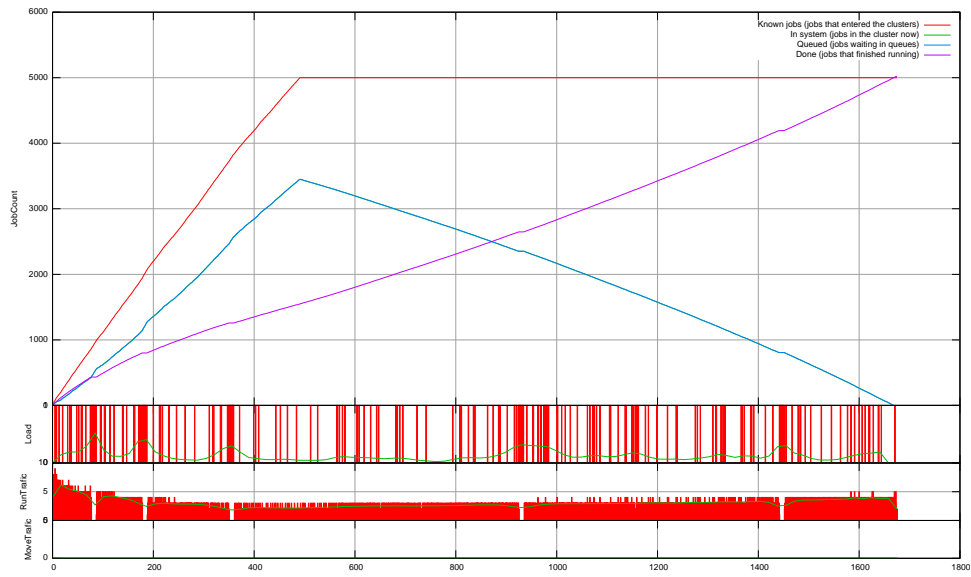


Figure 26: Průběhový graf - PBS Pro - 1 server, 1 scheduler, 1 MOM, 1 CPU

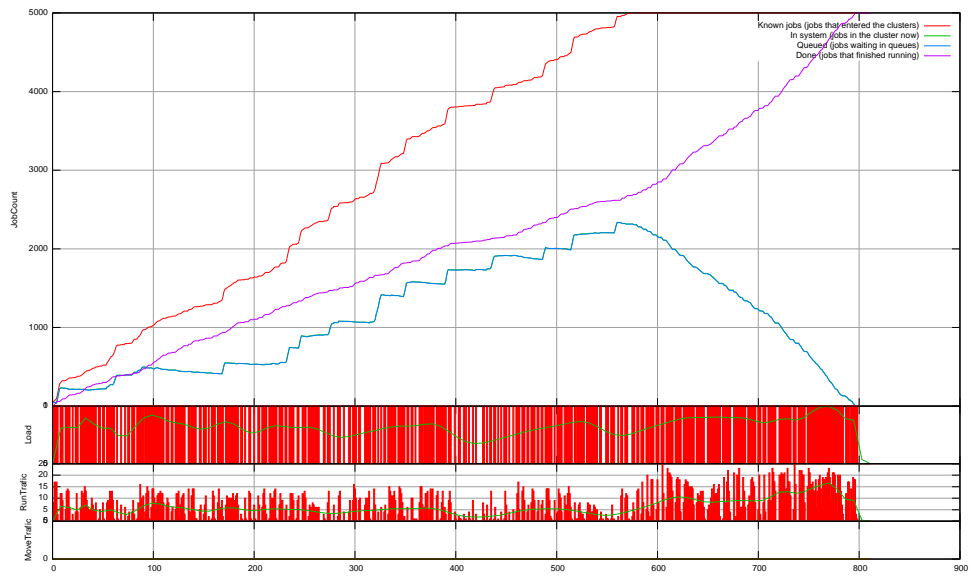


Figure 27: Průběhový graf - Torque 2.4 Stable - 1 server, 1 scheduler, 1 MOM, 4 CPU

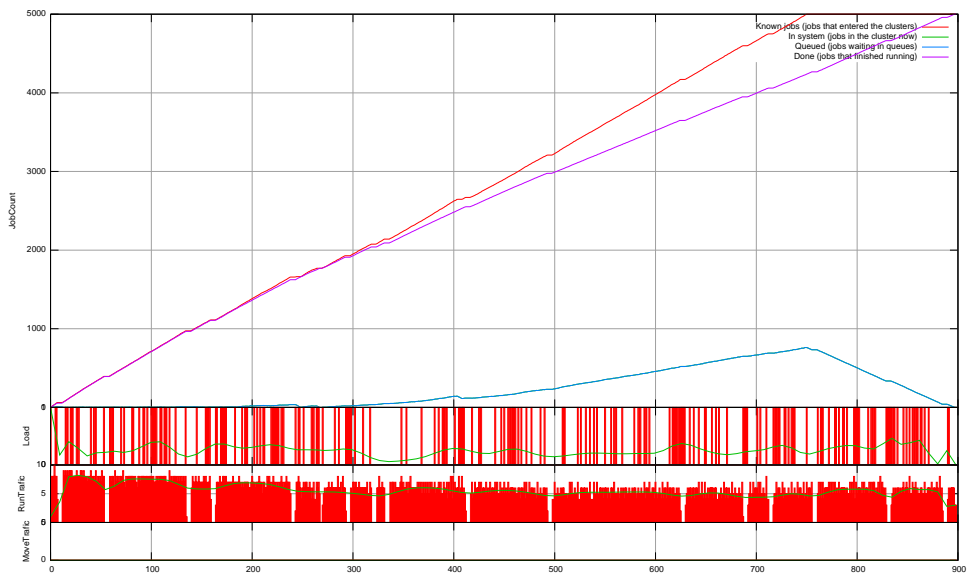


Figure 28: Průběhový graf - PBS Pro - 1 server, 1 scheduler, 1 MOM, 4 CPU

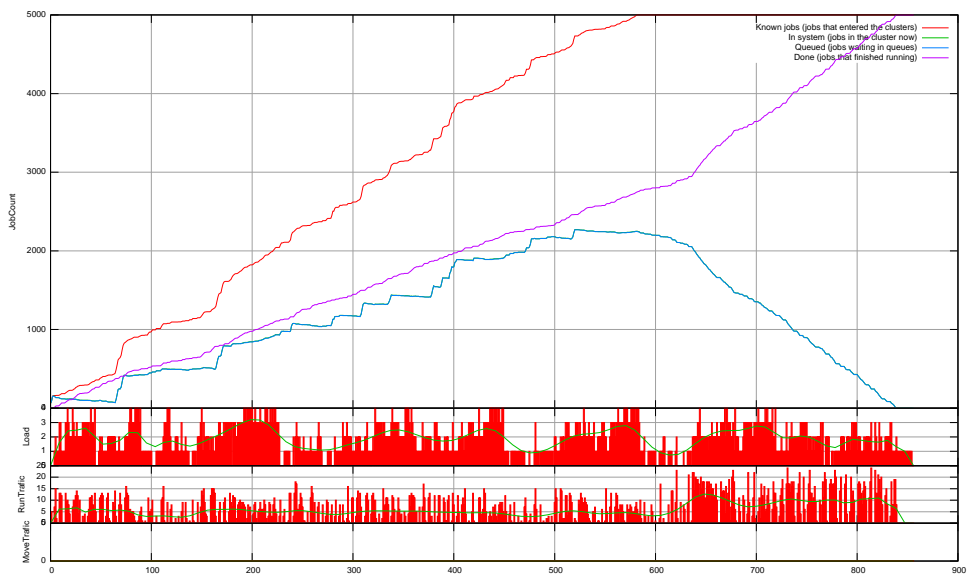


Figure 29: Průběhový graf - Torque 2.4 Stable - 1 server, 1 scheduler, 1 MOM, 8 CPU

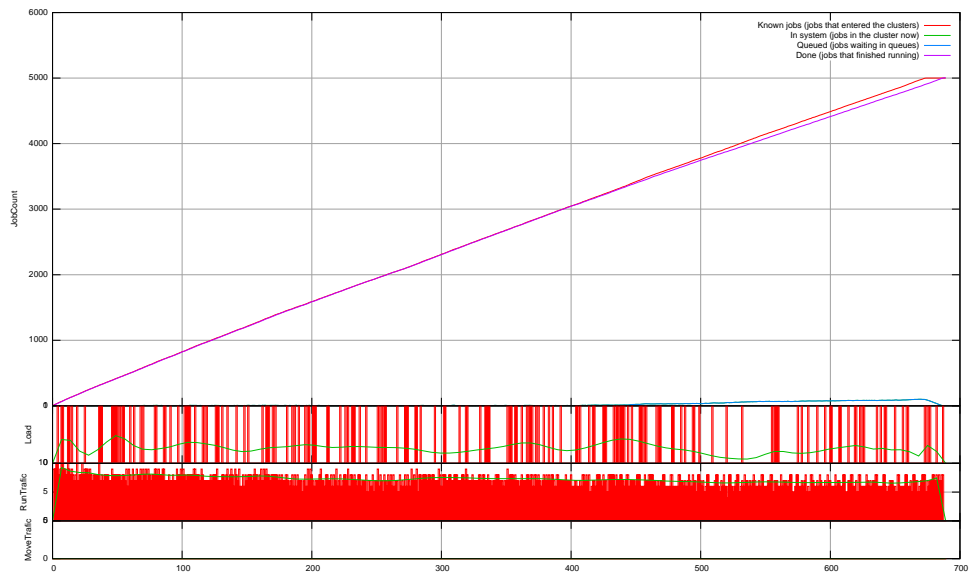


Figure 30: Průběhový graf - PBS Pro - 1 server, 1 scheduler, 1 MOM, 8 CPU

D Statistické grafy zaznamenaných experimentů

D.1 Stabilní verze 2.3.7

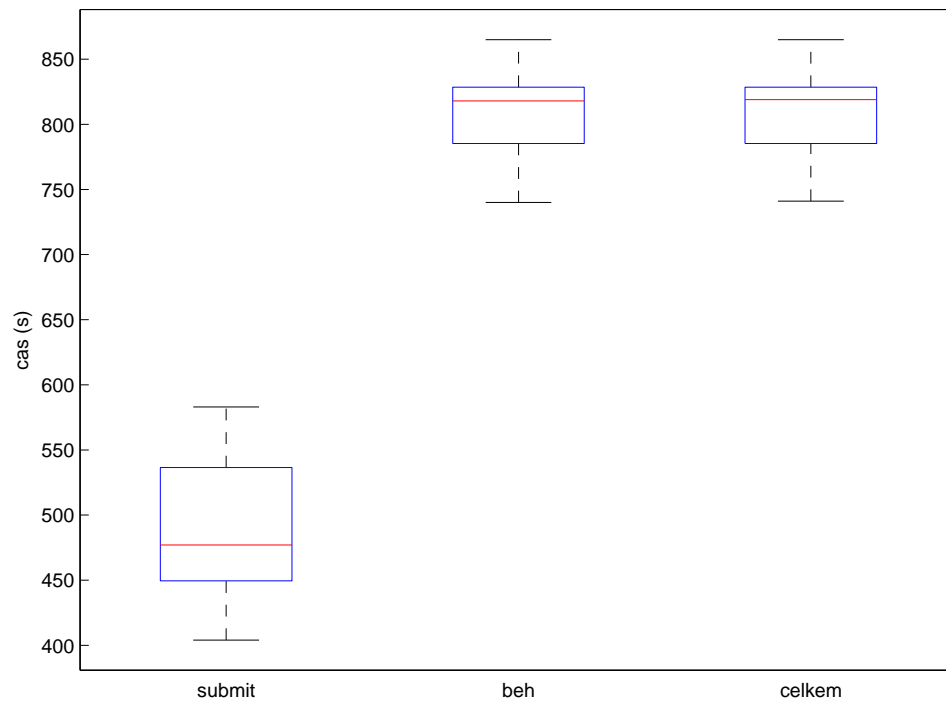


Figure 31: 2.3.7 - souběžný submit

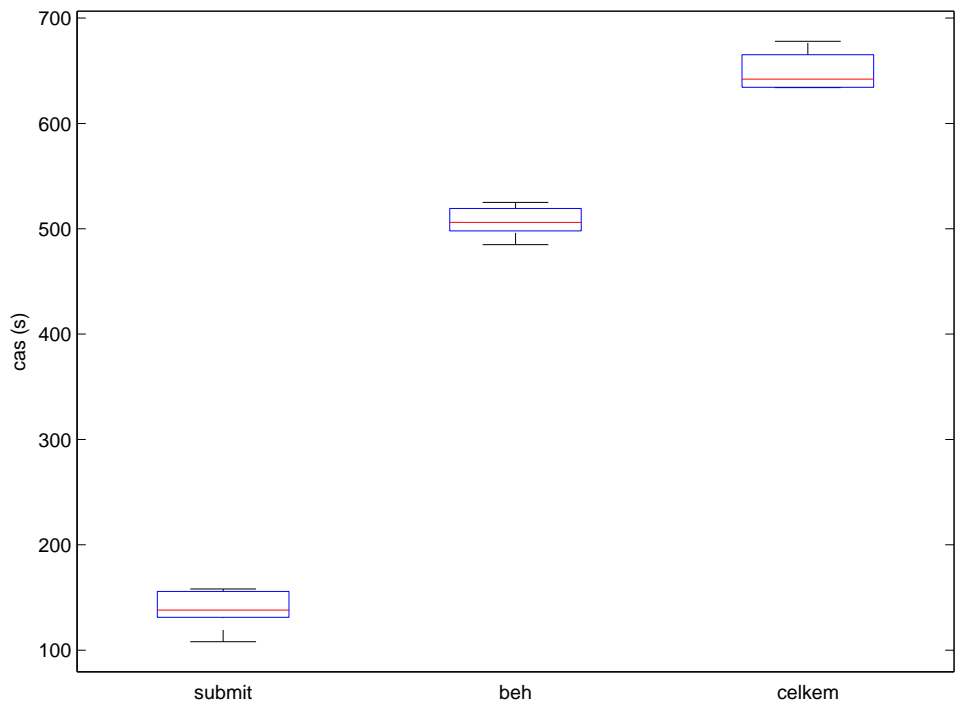


Figure 32: 2.3.7 - dedikovaný submit

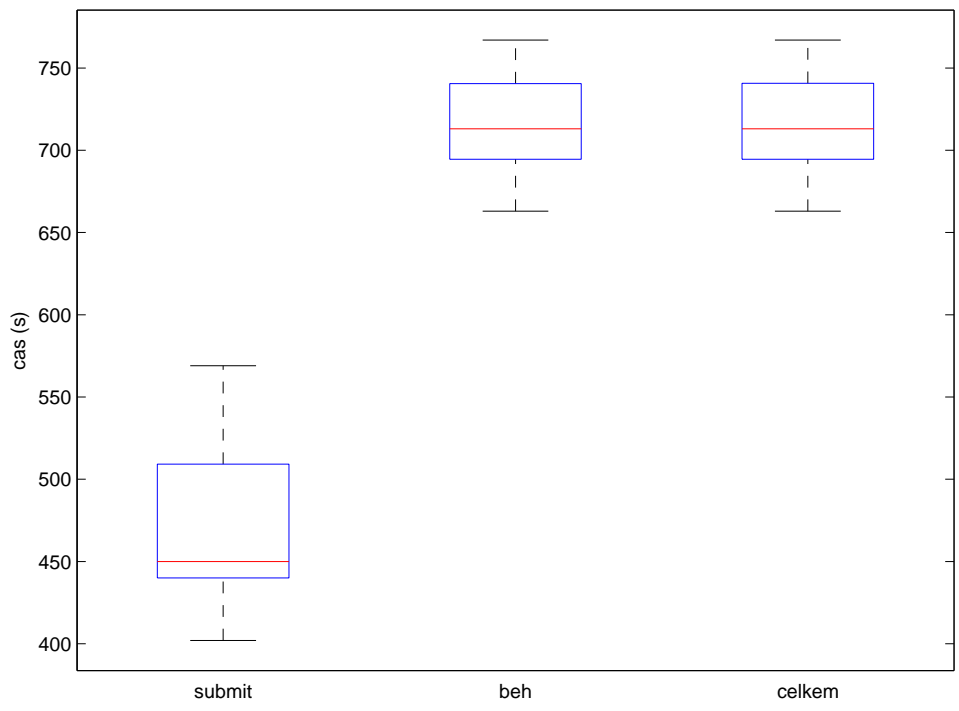


Figure 33: 2.3.7 - souběžný submit, nokeep

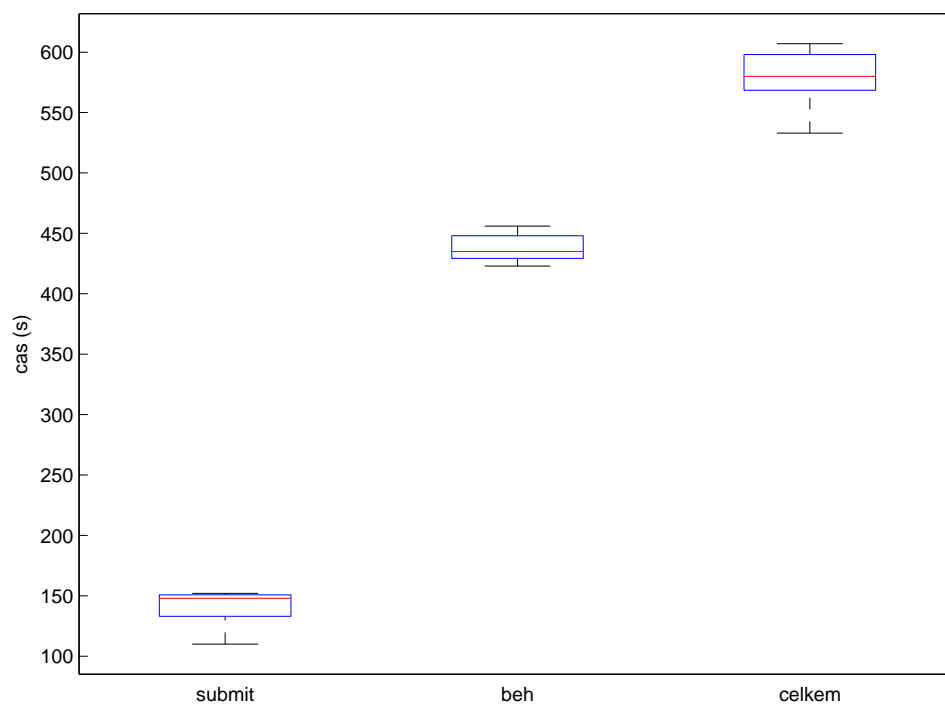


Figure 34: 2.3.7 - dedikovaný submit, nokeep

D.2 Stabilní verze 2.4.2

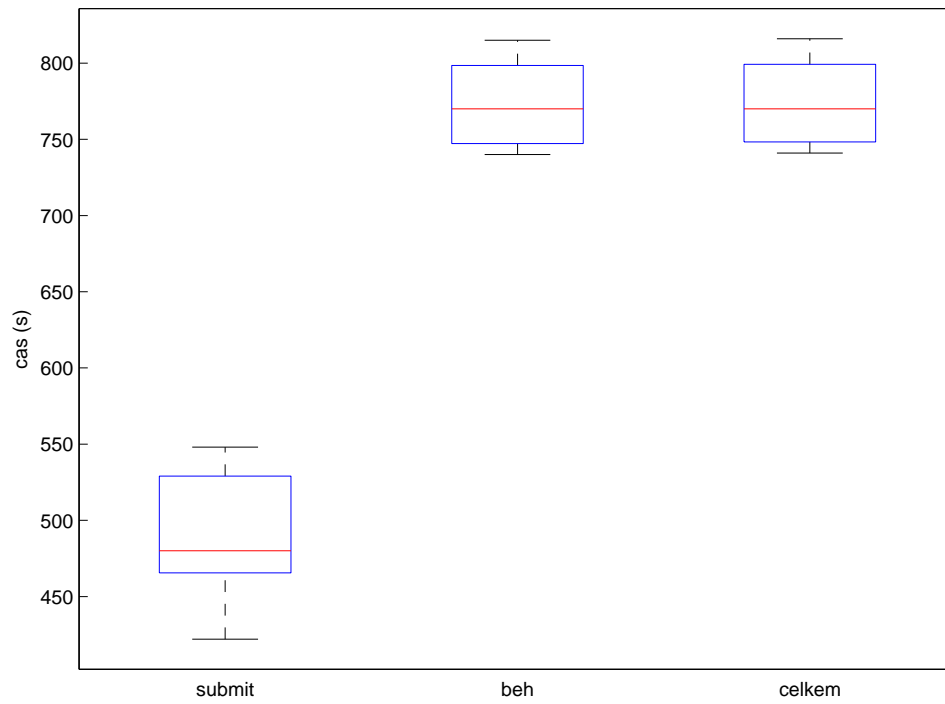


Figure 35: 2.4.2 - souběžný submit

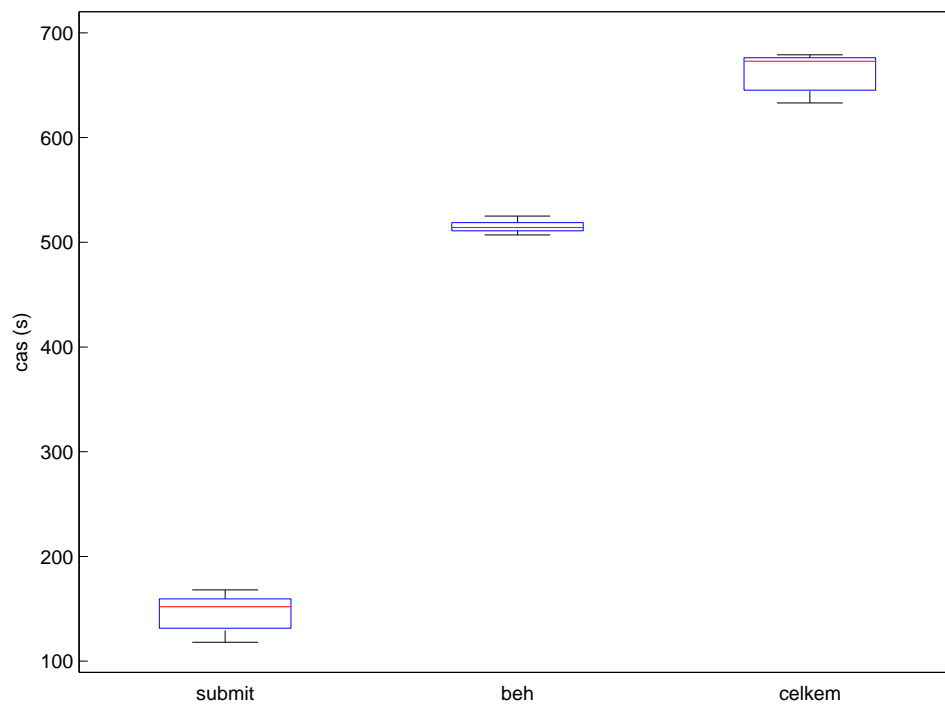


Figure 36: 2.4.2 - dedikovaný submit

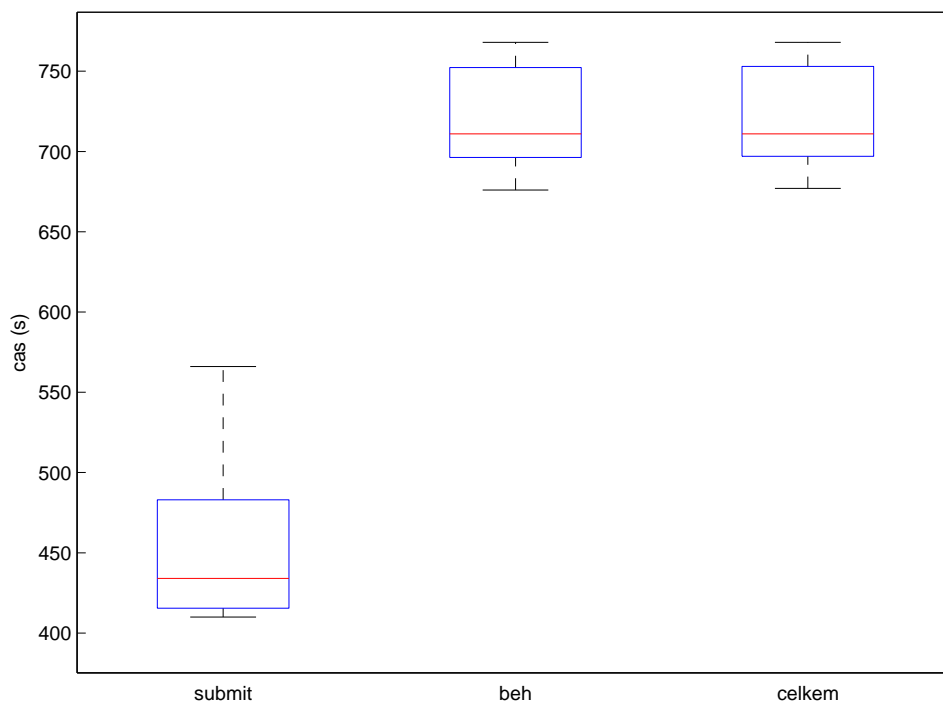


Figure 37: 2.4.2 - souběžný submit, nokeep

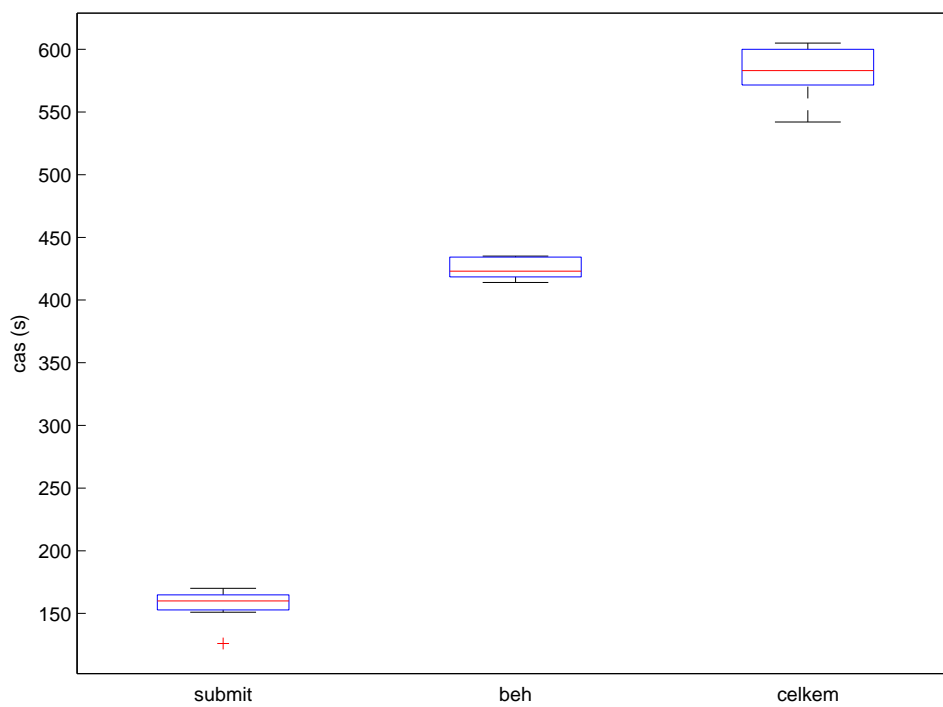


Figure 38: 2.4.2 - dedikovaný submit, nokeep

D.3 Trunk snapshot 19.10. 2009

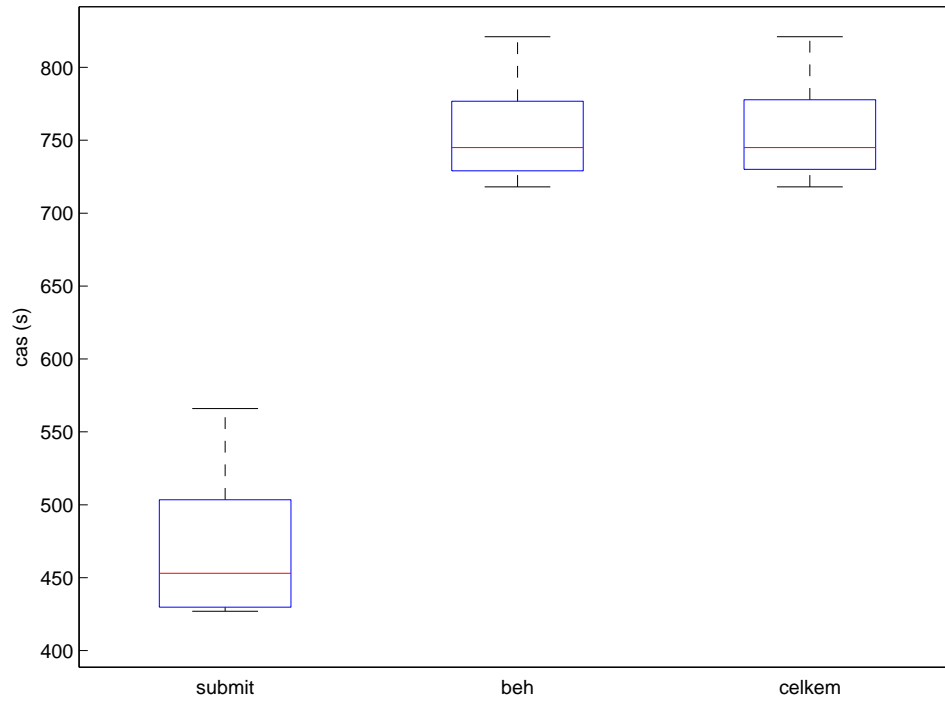


Figure 39: Trunk - souběžný submit

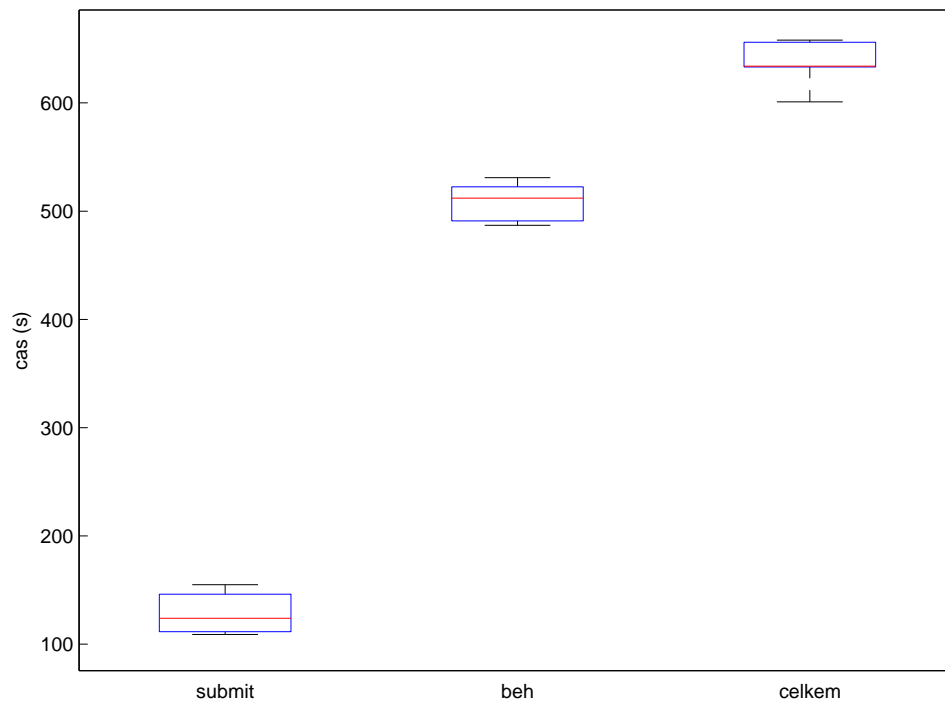


Figure 40: Trunk - dedikovaný submit

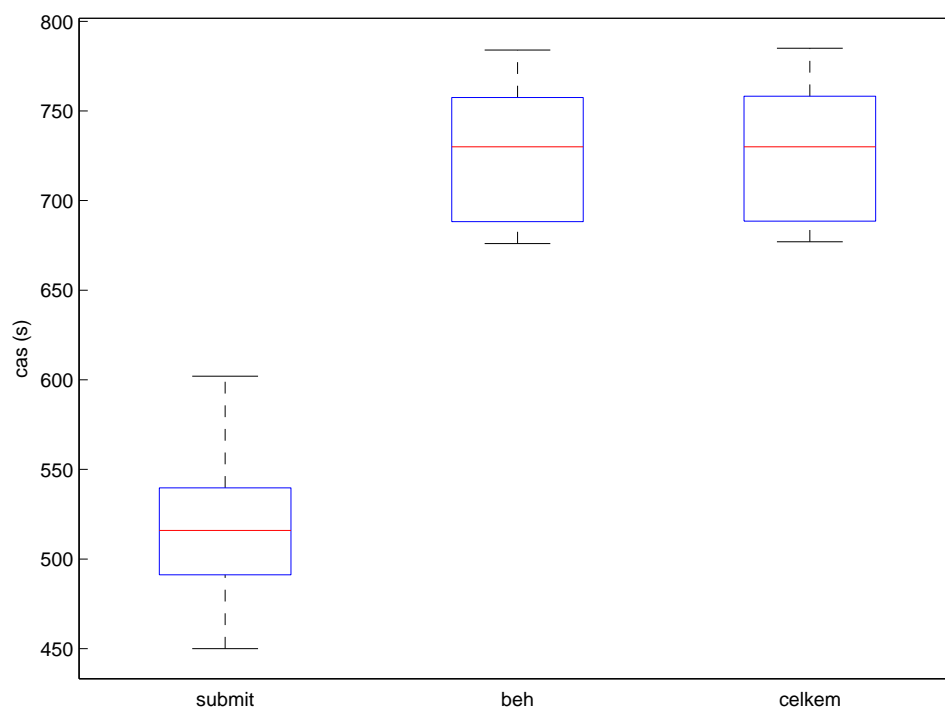


Figure 41: Trunk - souběžný submit, nokeep

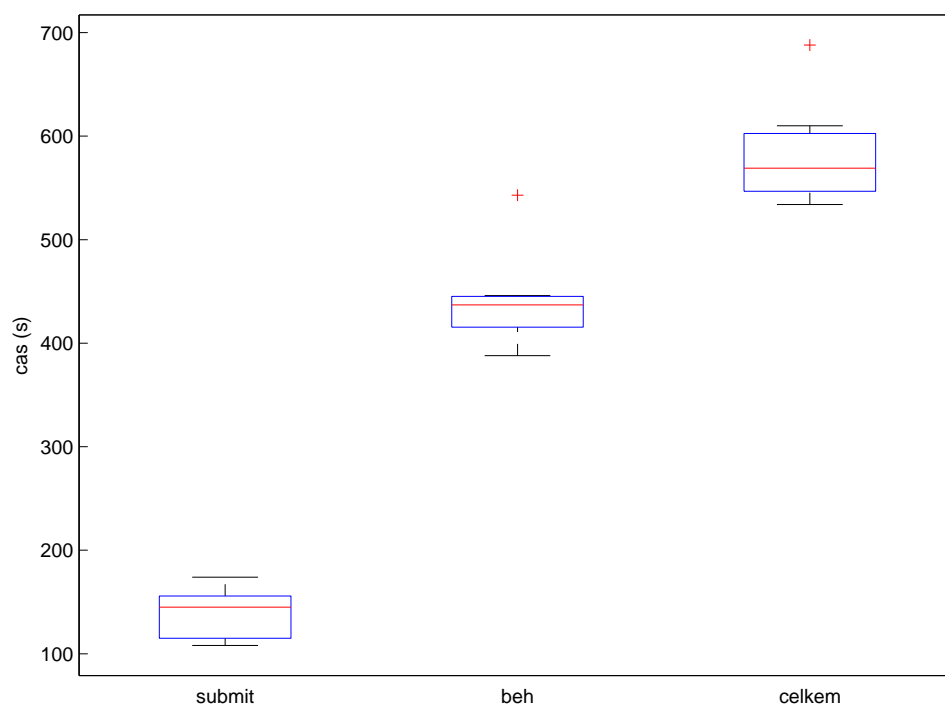


Figure 42: Trunk - dedikovaný submit, nokeep

D.4 Trunk snapshot 19.10. 2009 + patch

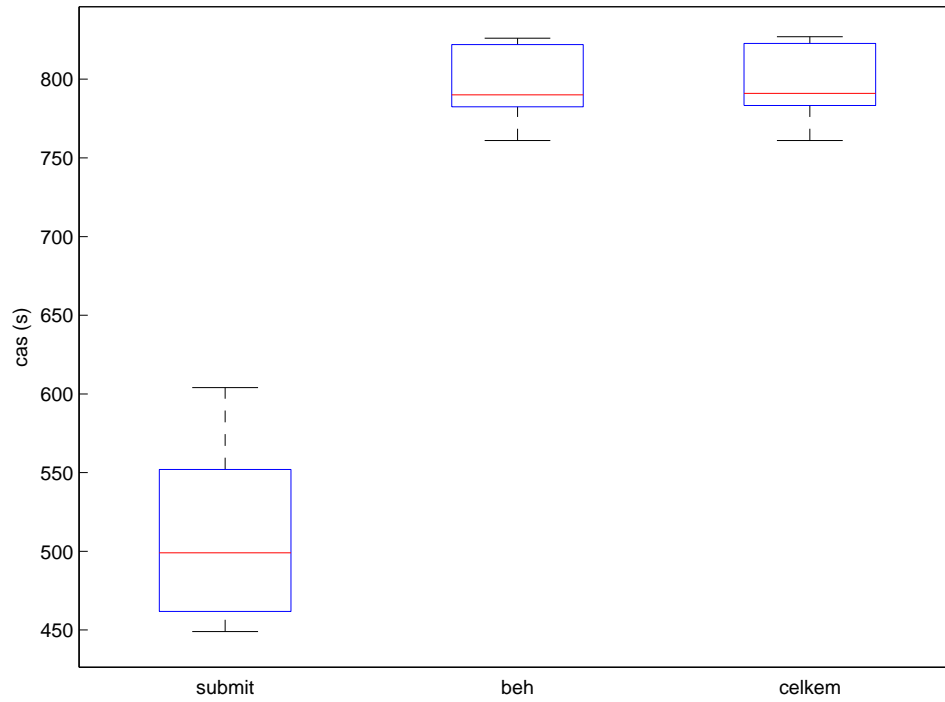


Figure 43: World nad Trunk - souběžný submit

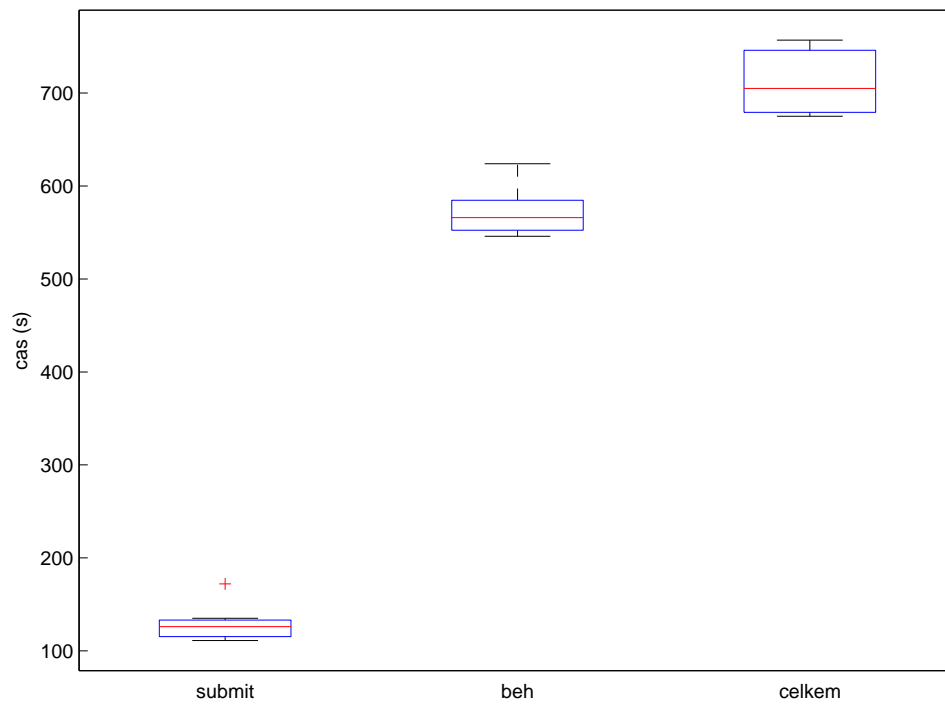


Figure 44: World nad Trunk - dedikovaný submit

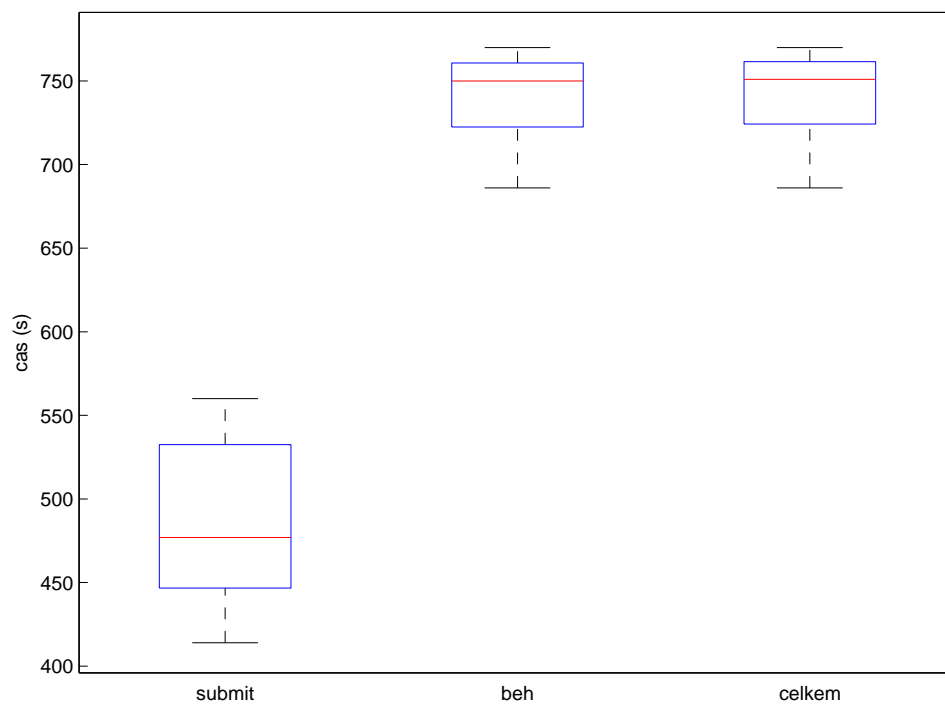


Figure 45: World nad Trunk - souběžný submit, nokeep

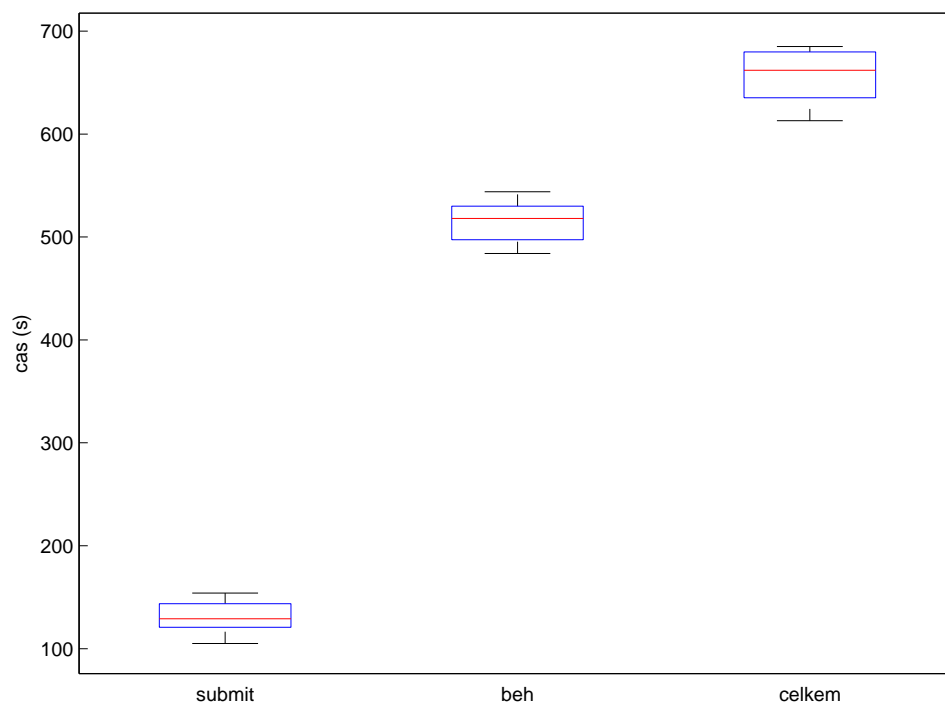


Figure 46: World nad Trunk - dedikovaný submit, nokeep

D.5 Stabilní verze 2.4.2 + patch

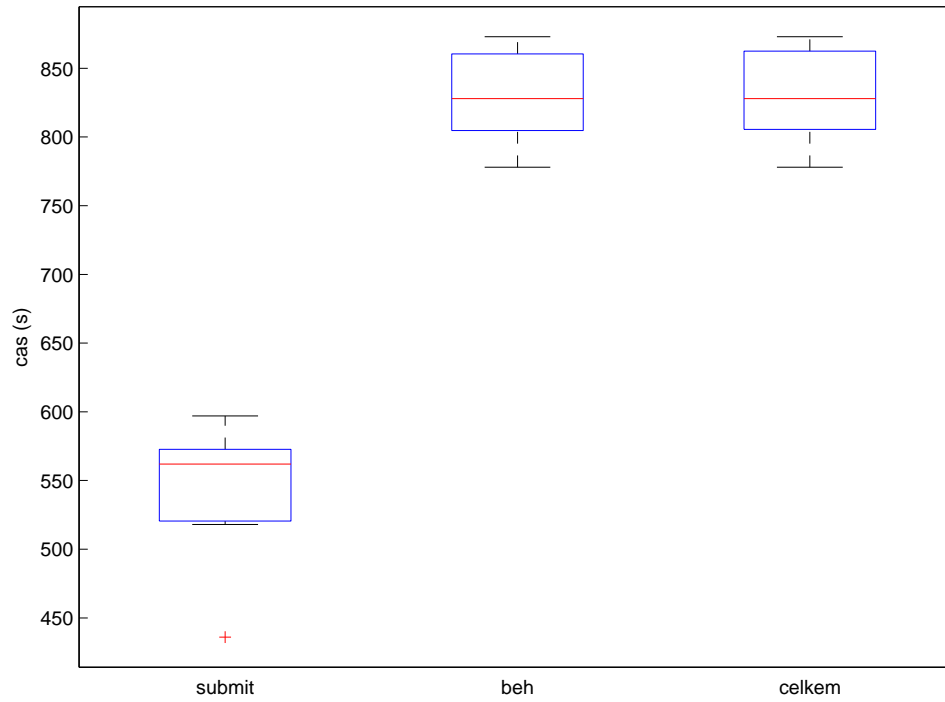


Figure 47: World nad 2.4.2 - souběžný submit

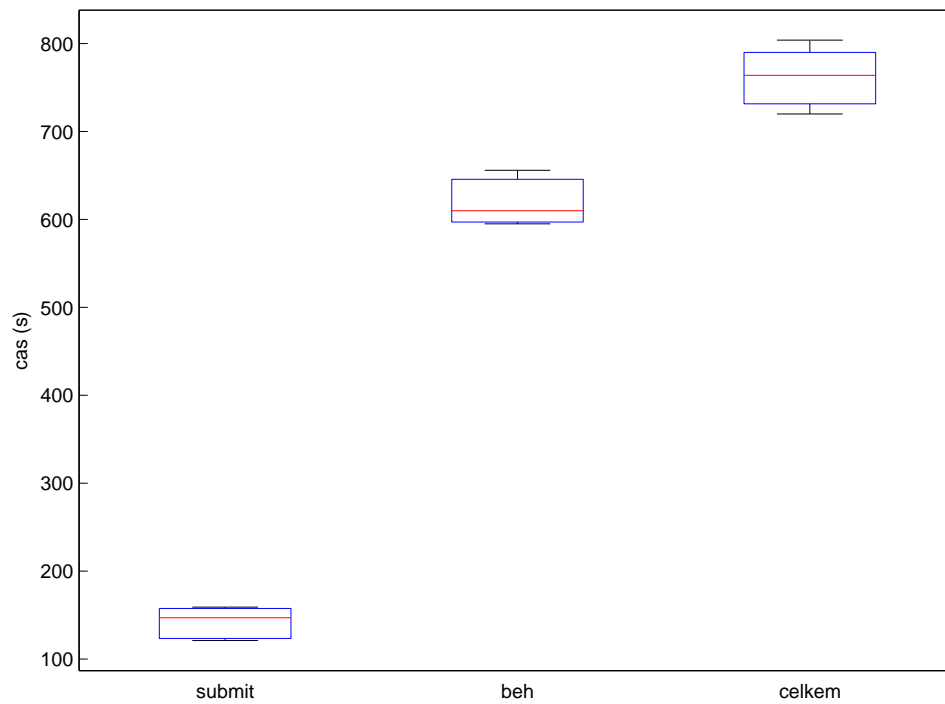


Figure 48: World nad 2.4.2 - dedikovaný submit

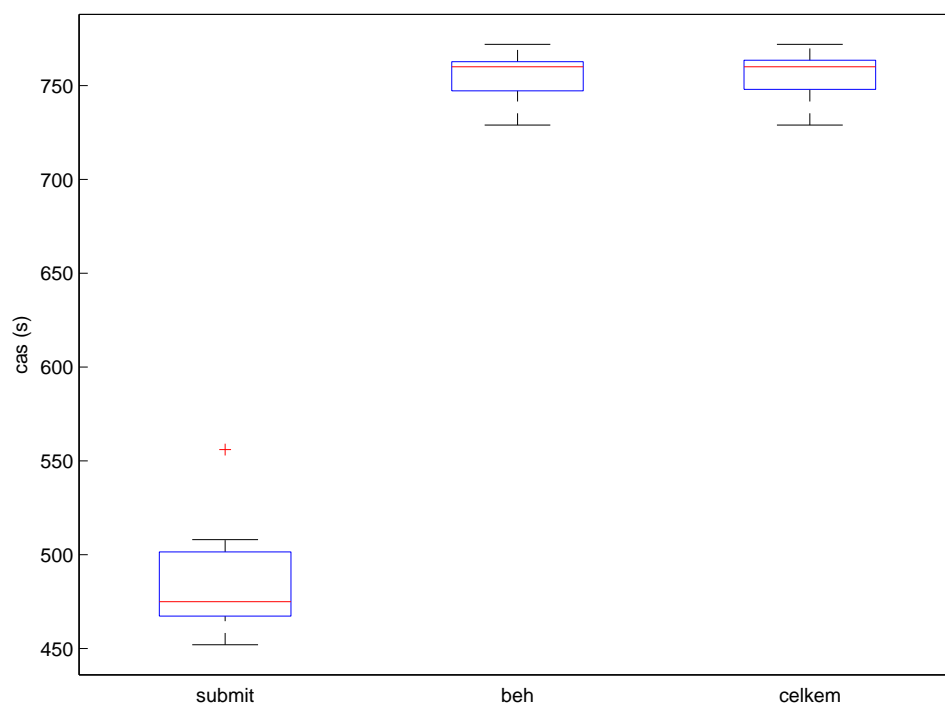


Figure 49: World nad 2.4.2 - souběžný submit, nokeep

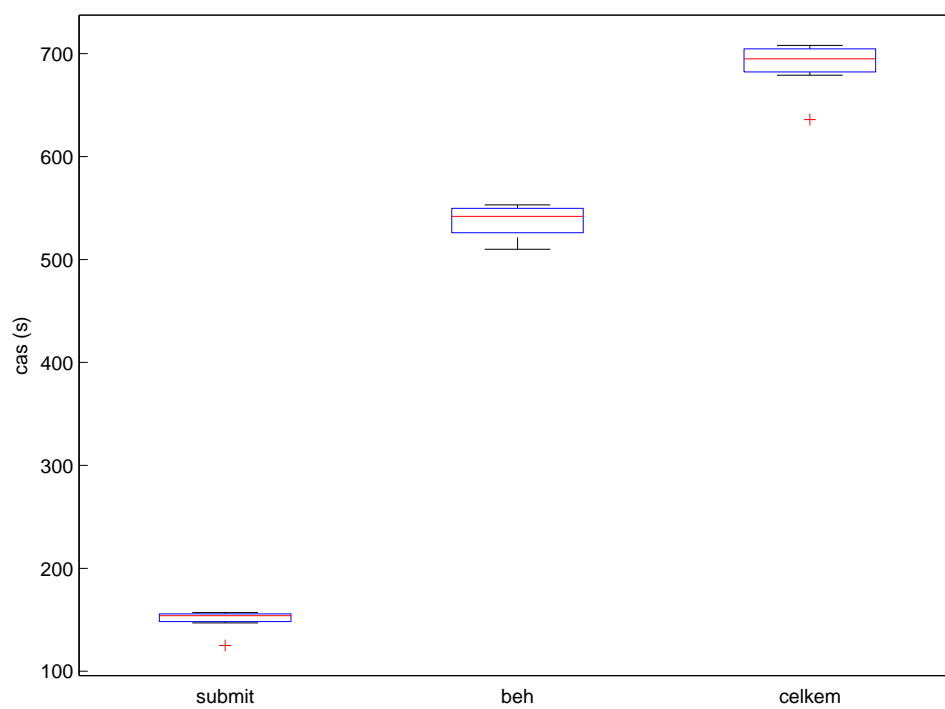


Figure 50: World nad 2.4.2 - dedikovaný submit, nokeep

D.6 Testy vlivu qstatu

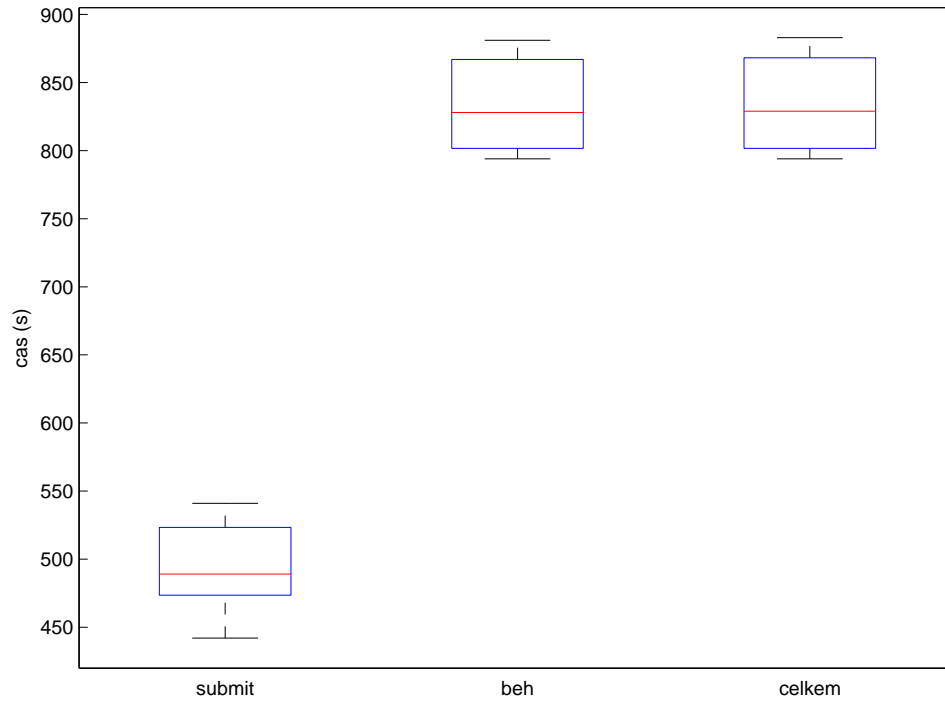


Figure 51: World nad 2.4.2 - qstat 0.5s

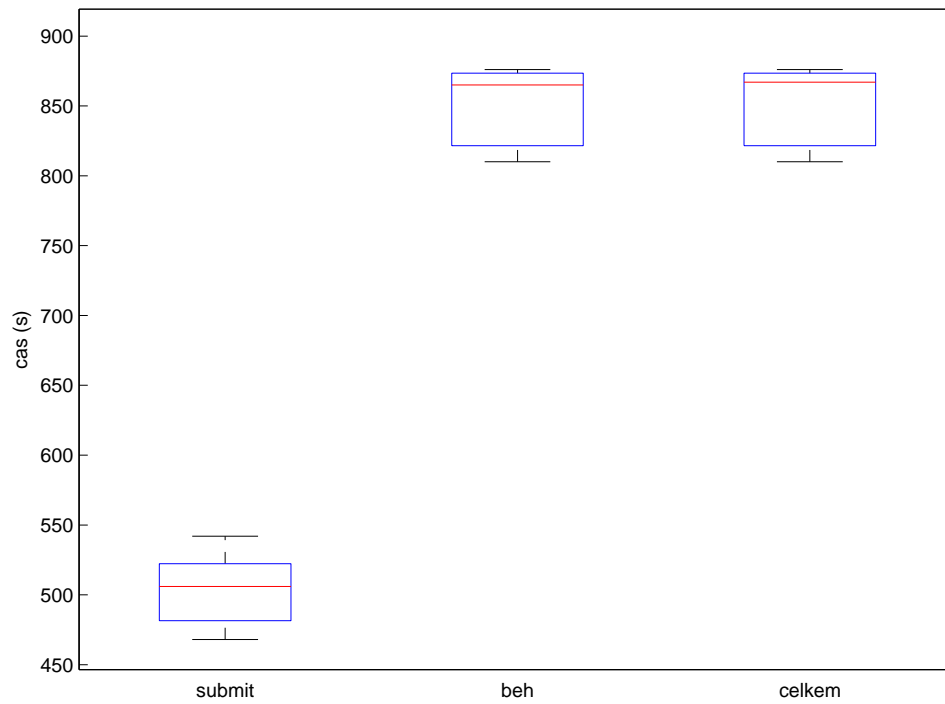


Figure 52: World nad 2.4.2 - qstat 1s

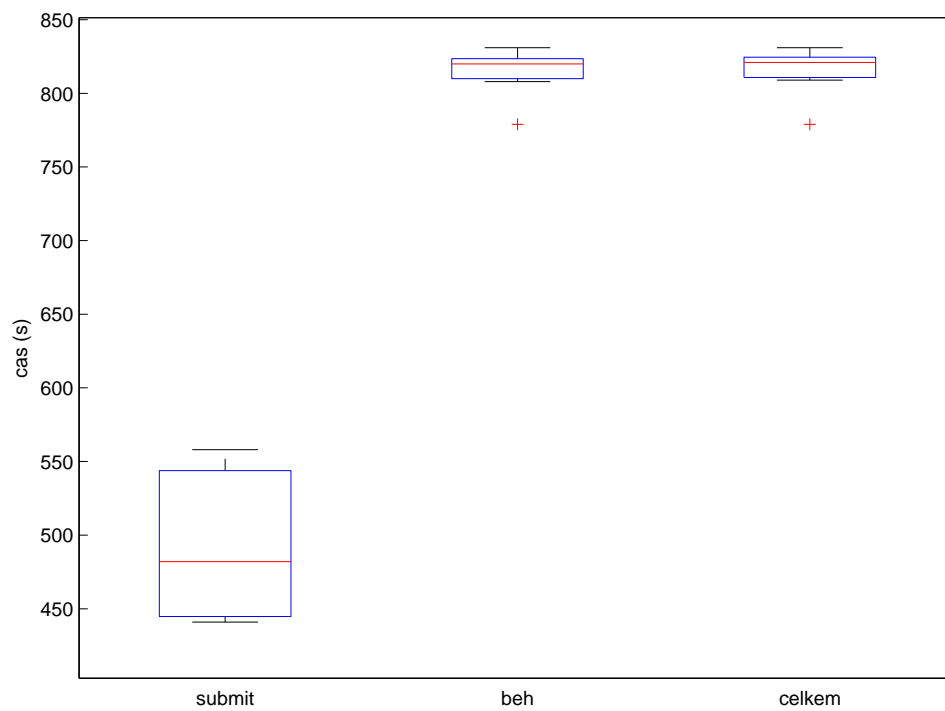


Figure 53: World nad 2.4.2 - qstat 2s

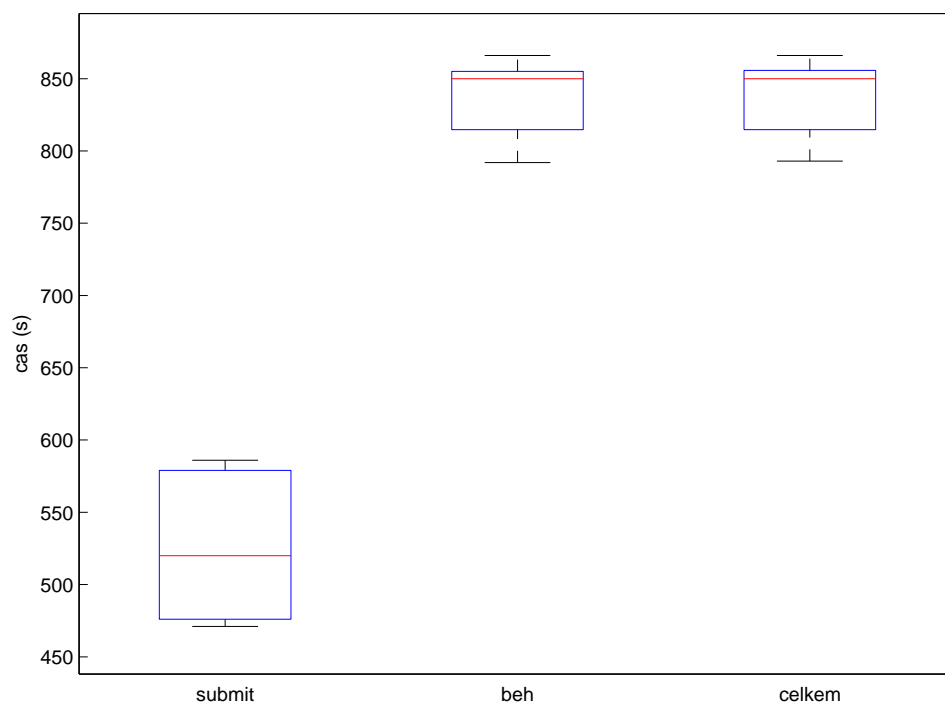


Figure 54: World nad 2.4.2 - bez komunikace plánovač-MOM

D.7 Testy zapojení více plánovačů

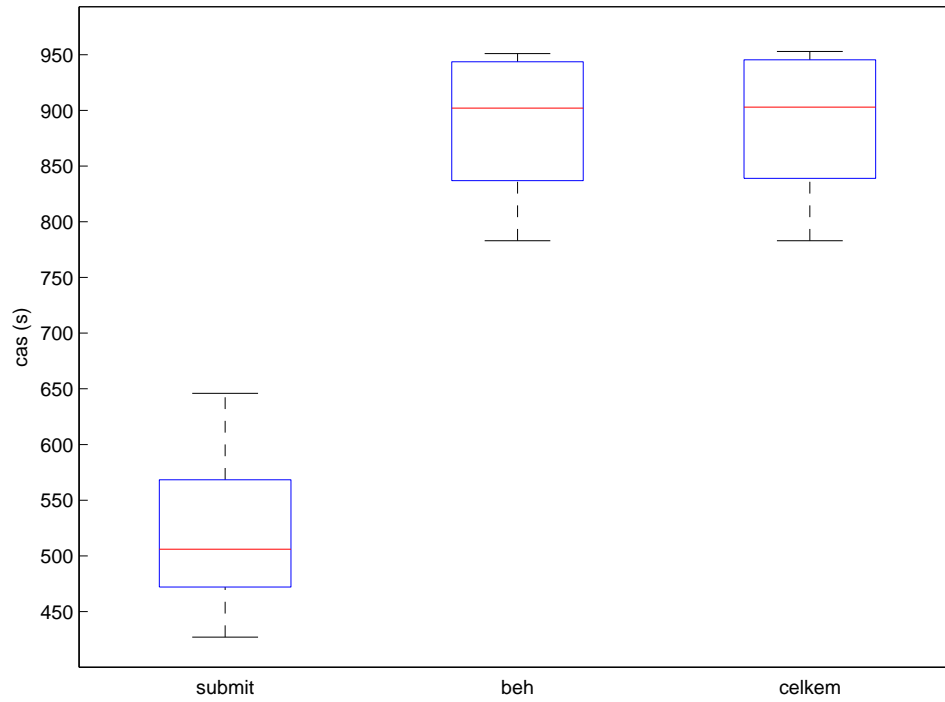


Figure 55: World nad 2.4.2 - 3 plánovače, se zamykáním

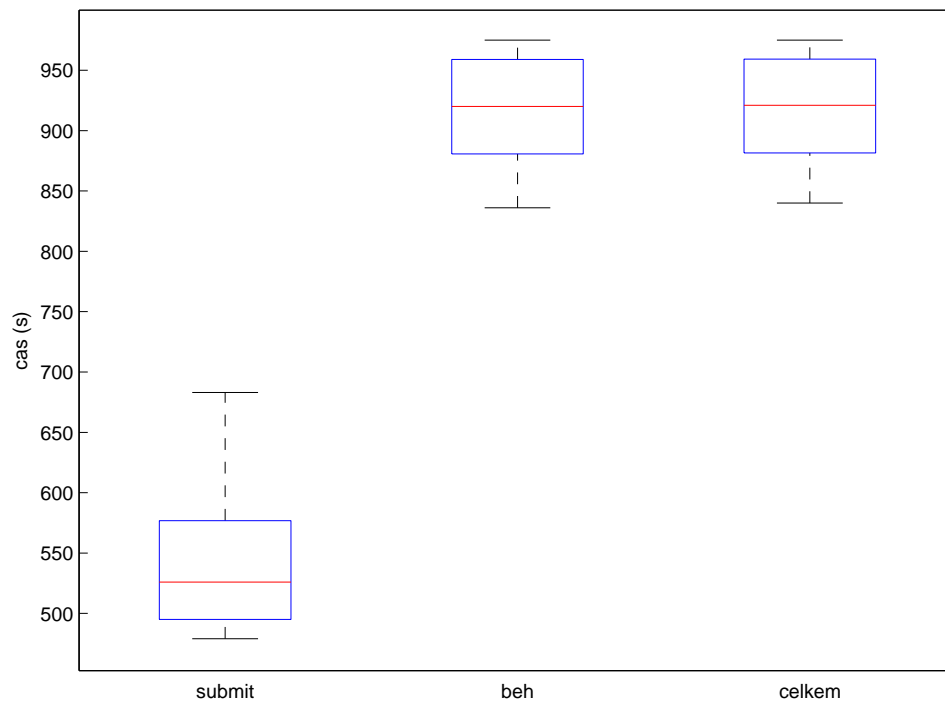


Figure 56: World nad 2.4.2 - 3 plánovače bez zamykání

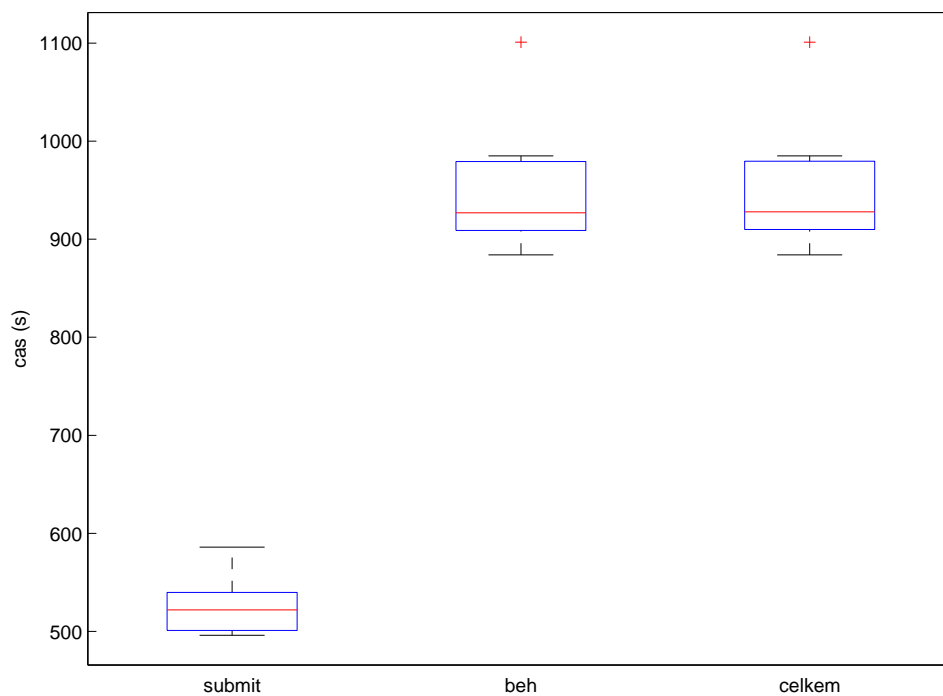


Figure 57: World nad 2.4.2 - 5 plánovačů se zamykáním

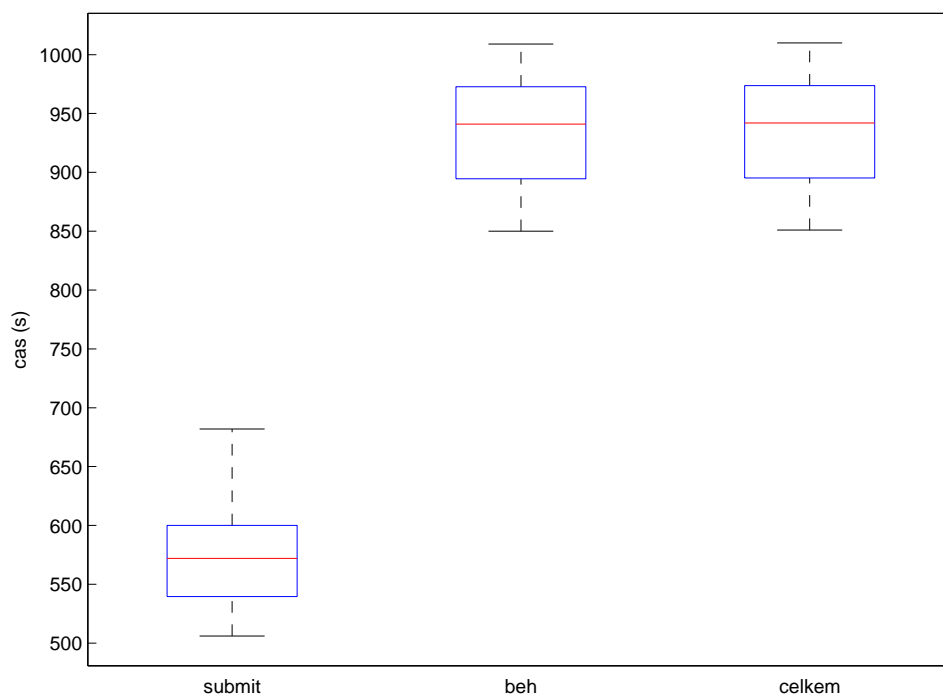


Figure 58: World nad 2.4.2 - 5 plánovačů bez zamykání

D.8 Testy vlivu wakeup timeoutu

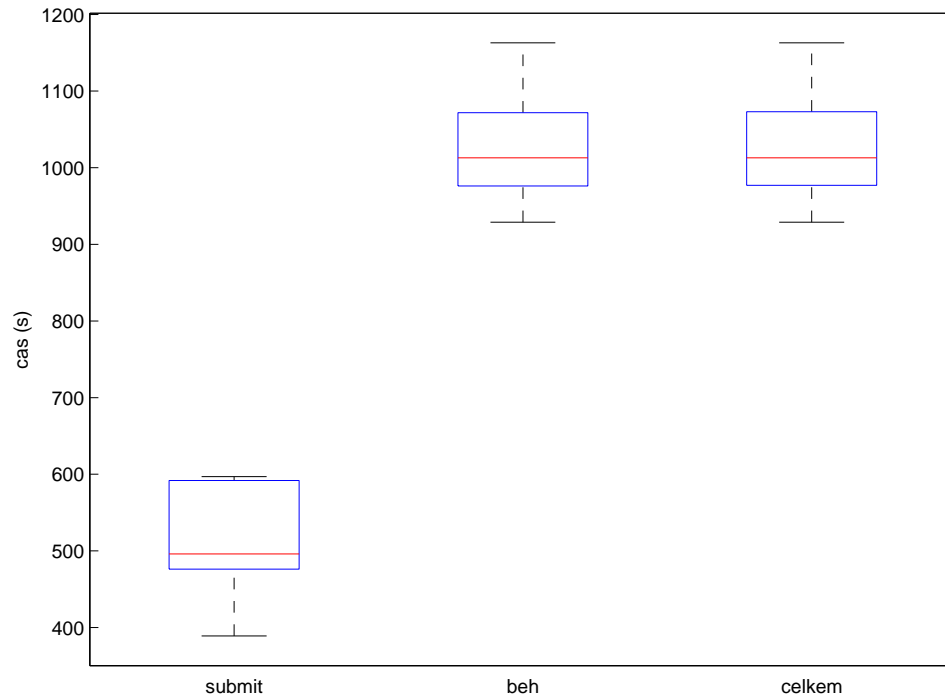


Figure 59: World nad 2.4.2 - 5 plánovačů, 5 sec. timeout plánovače, se zamykáním

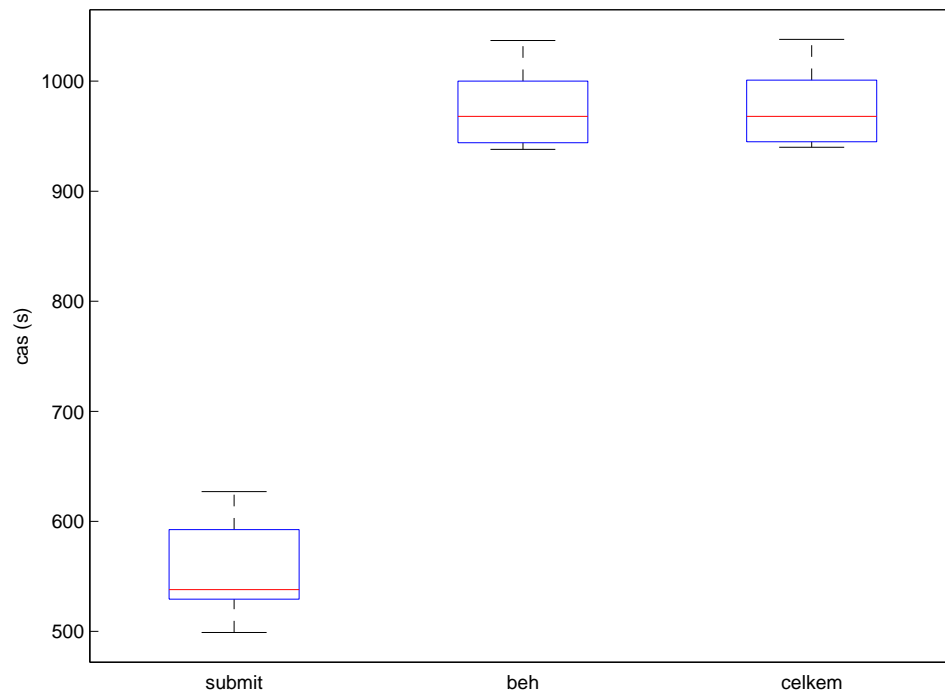


Figure 60: World nad 2.4.2 - 5 plánovačů, 5 sec. timeout plánovače, bez zamykání

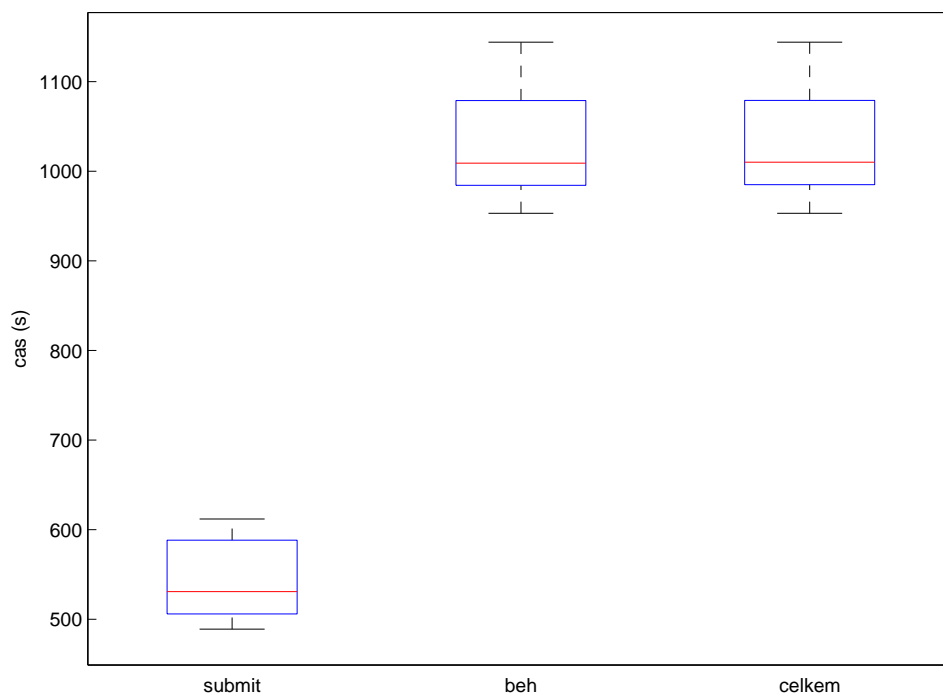


Figure 61: World nad 2.4.2 - 5 plánovačů, 10 sec. timeout plánovače, se zamykáním

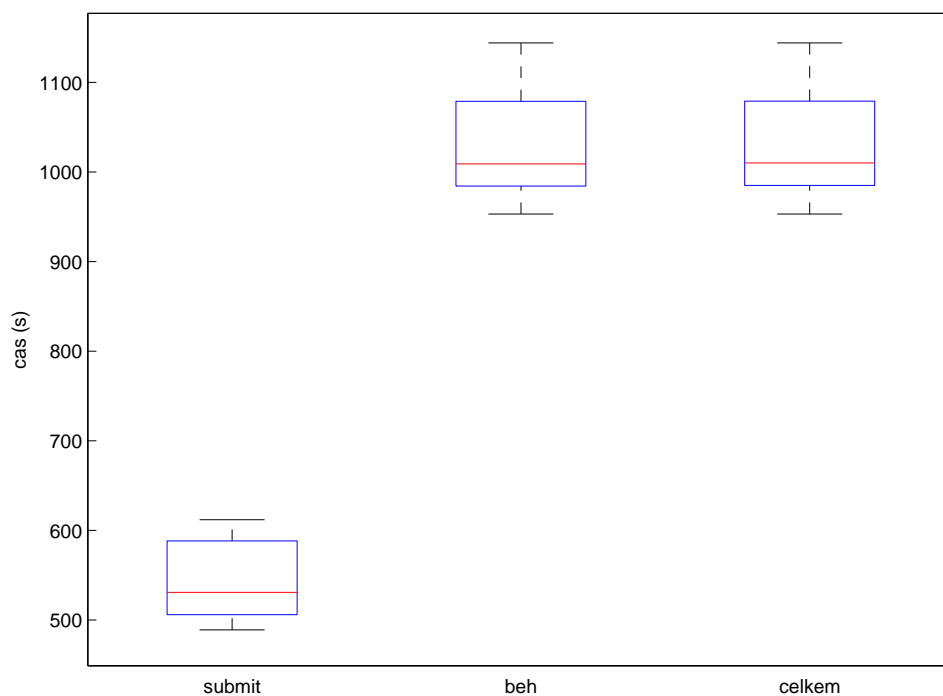


Figure 62: World nad 2.4.2 - 5 plánovačů, 10 sec. timeout plánovače, bez zamykání

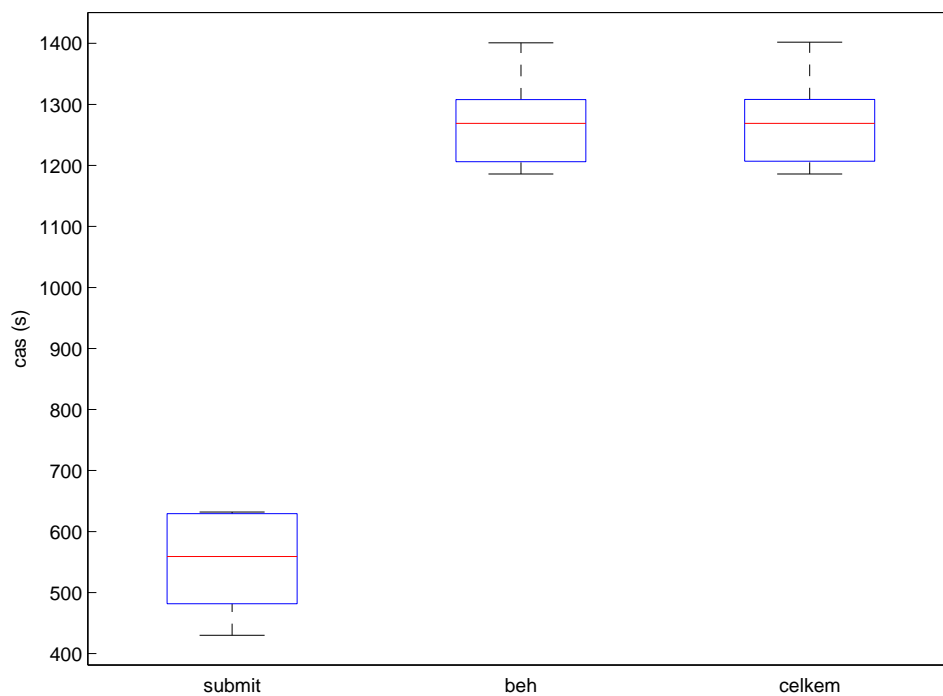


Figure 63: World nad 2.4.2 - 5 plánovačů, 30 sec. timeout plánovače, se zamykáním

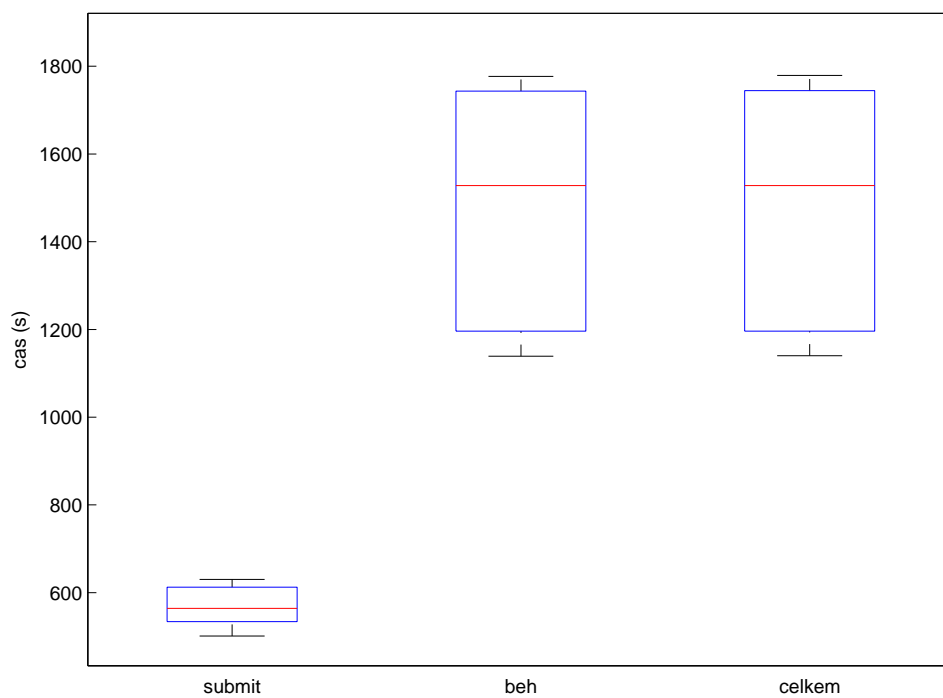


Figure 64: World nad 2.4.2 - 5 plánovačů, 30 sec. timeout plánovače, bez zamykání

D.9 Test zapojení více serverů na jeden plánovač

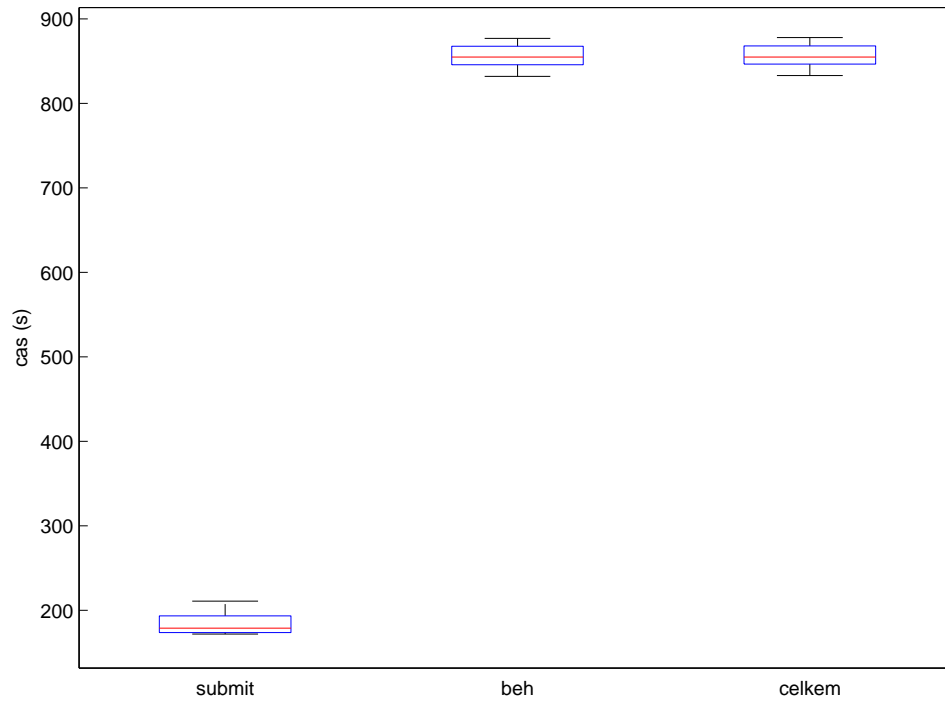


Figure 65: World nad 2.4.2 - 2 servery, 2x poloviční submit

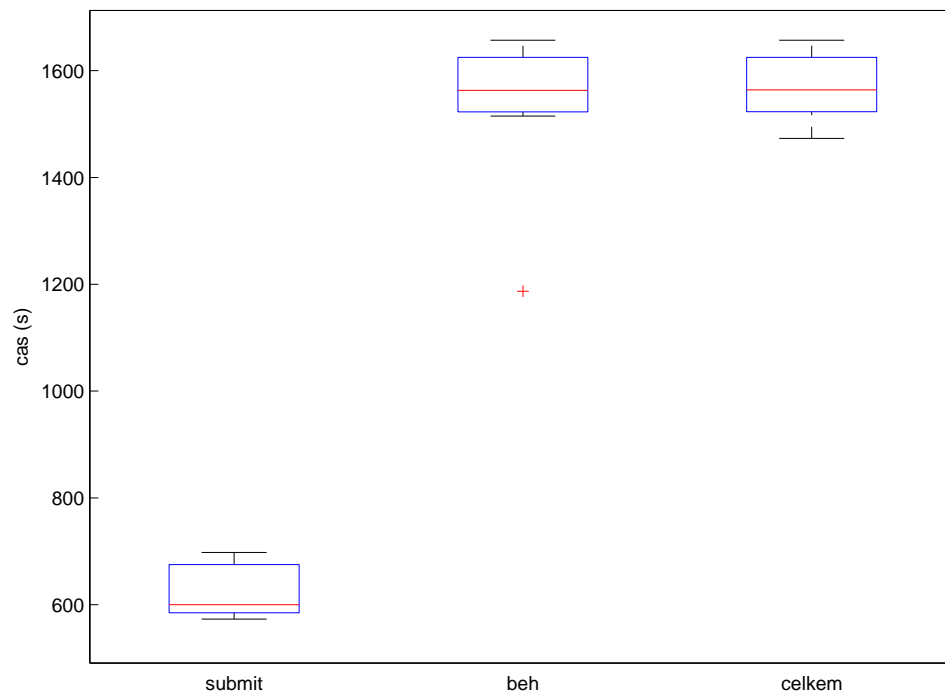


Figure 66: World nad 2.4.2 - 2 servery, 2x plný submit

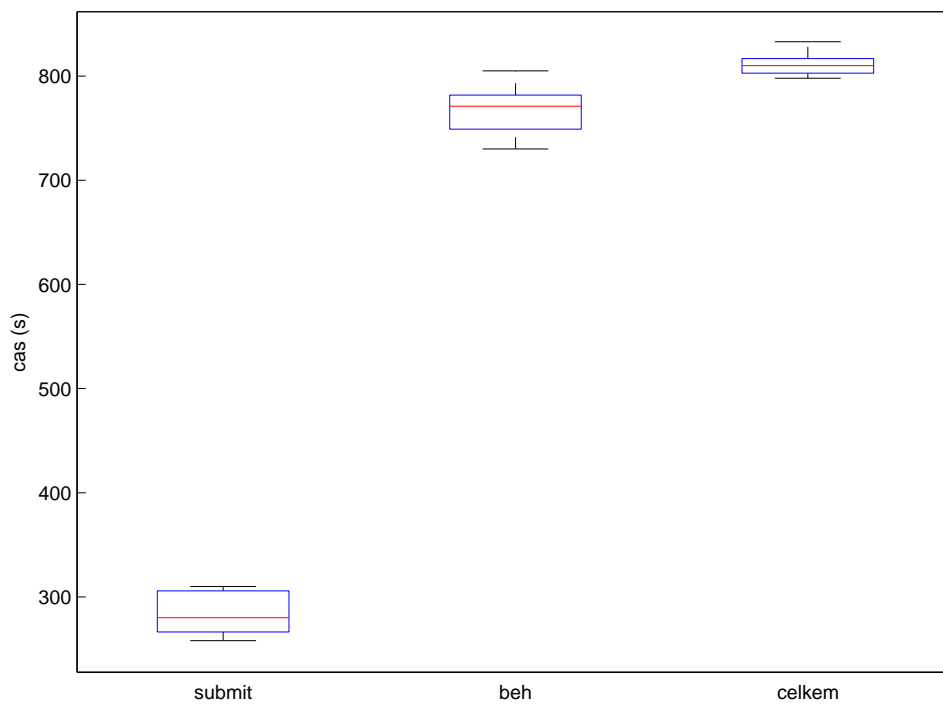


Figure 67: World nad 2.4.2 - 5 servery, 5x pětínový submit

D.10 Test zapojení M:N

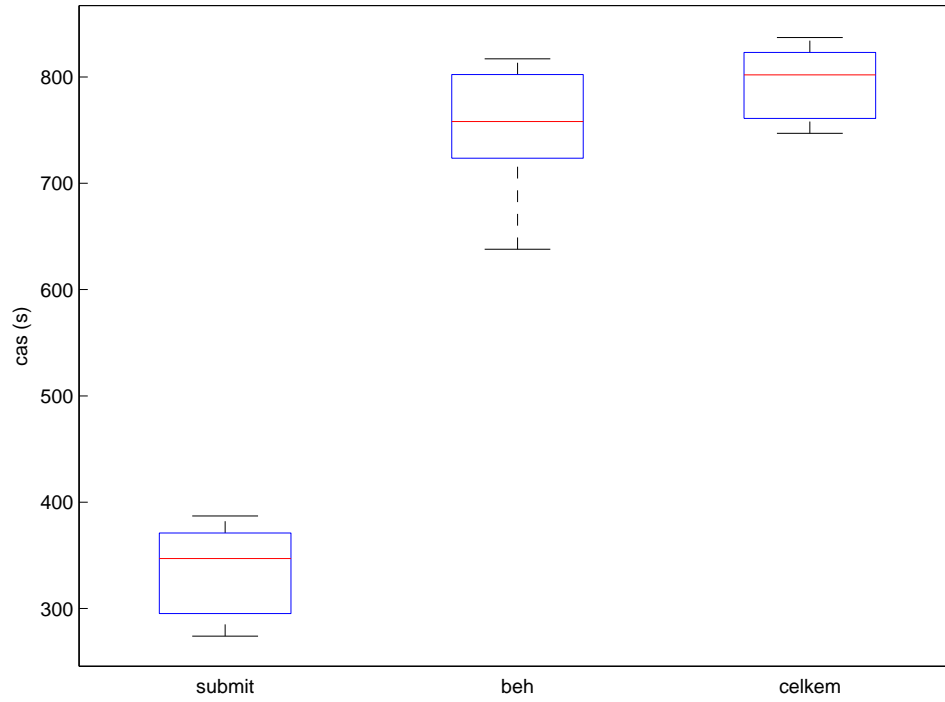


Figure 68: World nad 2.4.2 - 2 servery, 2 plánovače

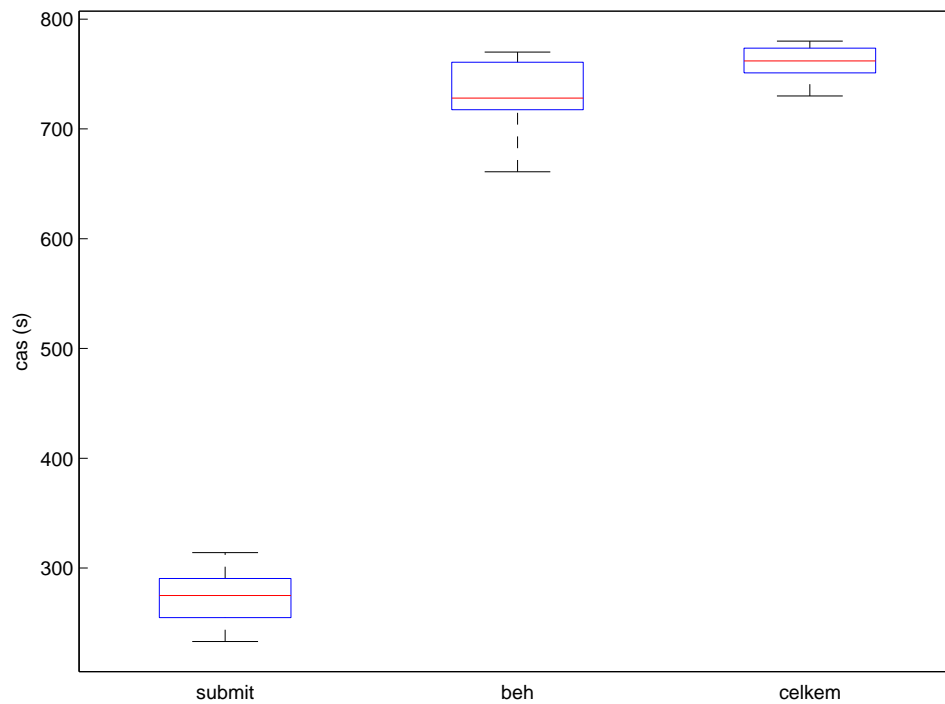


Figure 69: World nad 2.4.2 - 2 servery, 2 plánovače, zamykání

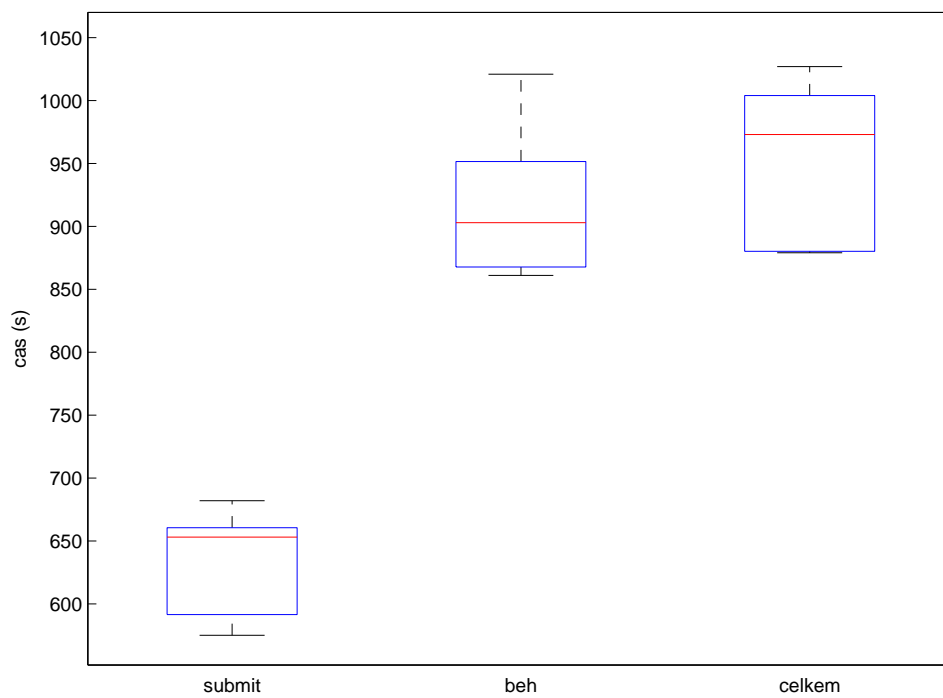


Figure 70: World nad 2.4.2 - 2 servery, 2 plánovače, přesuny

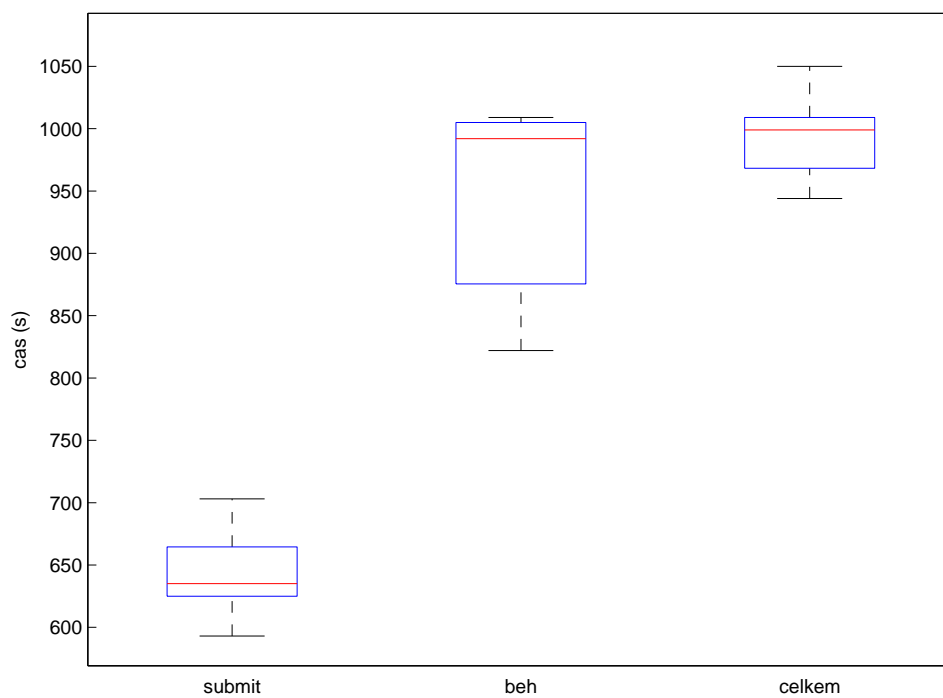


Figure 71: World nad 2.4.2 - 2 servery, 2 plánovače, přesuny, zamykání

D.11 Test vlivu nenulové délky úlohy

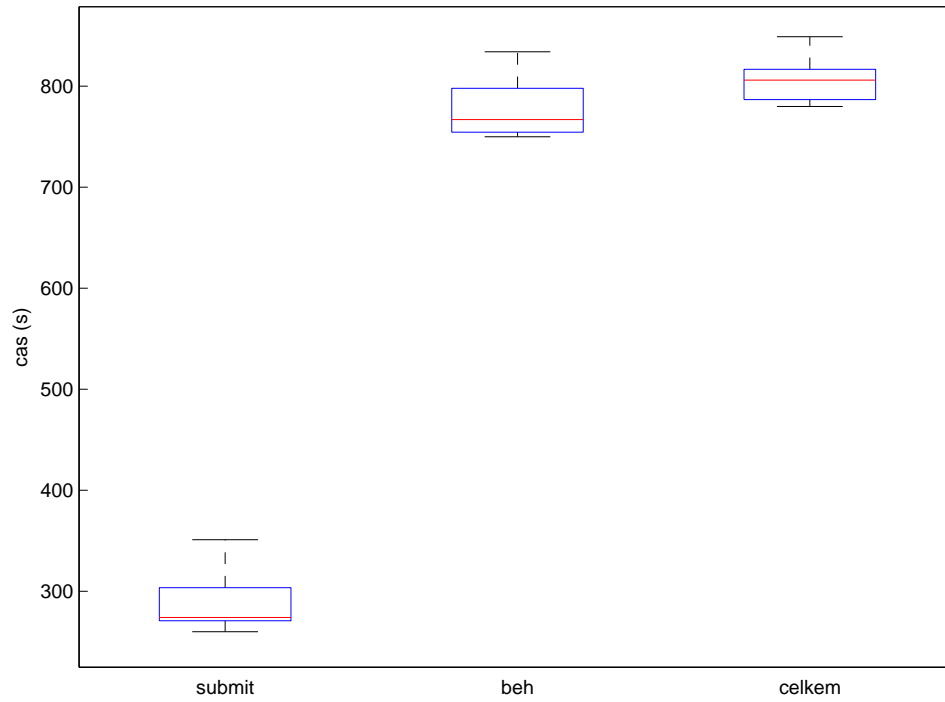


Figure 72: World nad 2.4.2 - 2 servery, 2 plánovače, 1sec úlohy

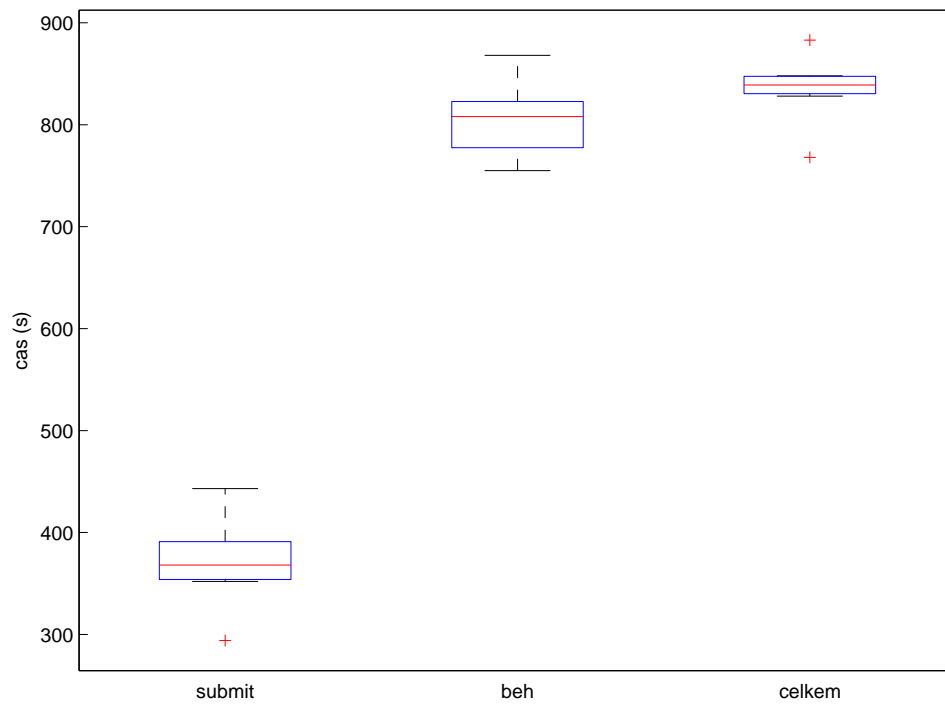


Figure 73: World nad 2.4.2 - 2 servery, 2 plánovače, 1sec úlohy, zamykání

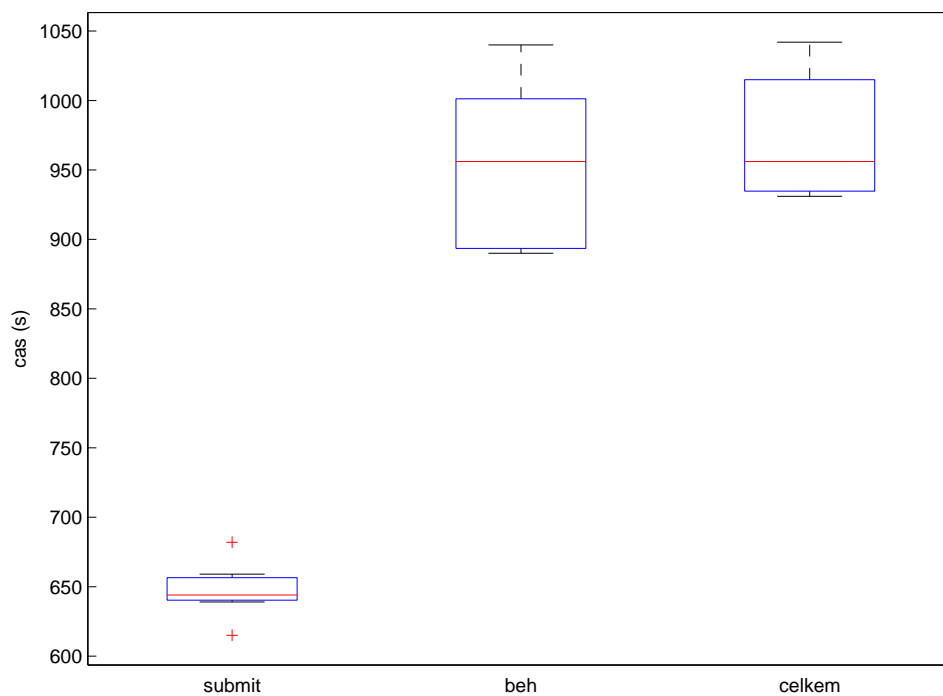


Figure 74: World nad 2.4.2 - 2 servery, 2 plánovače, 1sec úlohy, přesuny

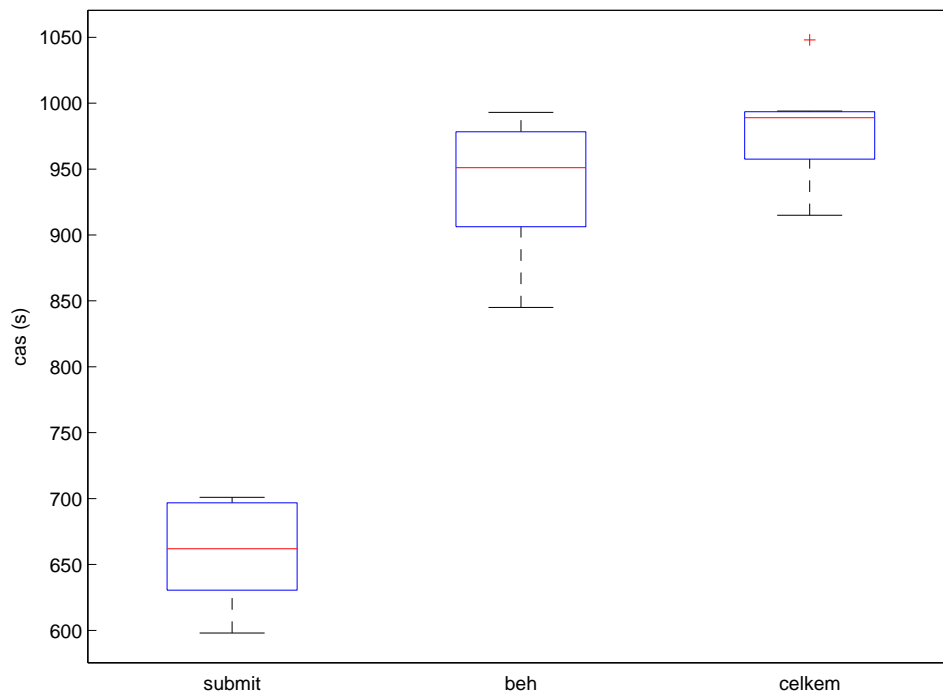


Figure 75: World nad 2.4.2 - 2 servery, 2 plánovače, 1sec úlohy, přesuny, zamykání