

DSD

Elosztott
Rendszerek
Osztály

**A KOPI
Plágiumkereső
terhelésének
elosztása cloud
környezetben**

Micsik András,
Pataki Máté,
Garzó András

MTA SZTAKI

A KOPI Plágiumkereső

DSD

Elosztott
Rendszerek Osztály

- A kopi.sztaki.hu portál 2004-ben indult egynyelvű plágium keresési szolgáltatással
- 2011-ben a világon elsőként bemutattuk a fordítási plágiumkeresőt
 - Ez képes detektálni, ha valaki például az angol Wikipedia-ból lefordított bekezdéseket használ fel
 - Az új algoritmus számítási igénye nagyságrendekkel nagyobb, mint az egynyelvű plágiumkeresésé

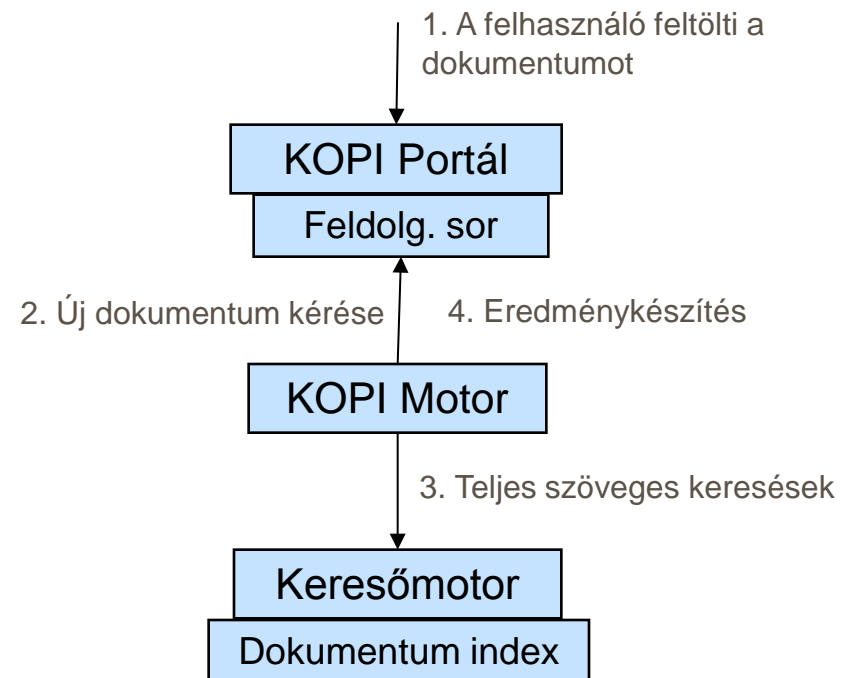
A plágiumkeresés folyamata

DSD

Elosztott
Rendszerek Osztály

■ A plágiumkeresés folyamata

- A felhasználó feltölti a dokumentumot, amely a KOPI Portál feldolgozási sorába kerül
- A KOPI Motor új feladatot kér a Portáltól
- A KOPI Motor feldolgozza a dokumentumot, és közben teljes szöveges kéréseket ad ki a Keresőmotornak
- A KOPI Motor összeállítja az eredményt és visszaküldi a Portálnak
- A felhasználó értesítést kap, hogy az eredmény elkészült.
- Az eredmény egy listát tartalmaz az esetlegesen másolt részekről és a plágium valószínűségéről

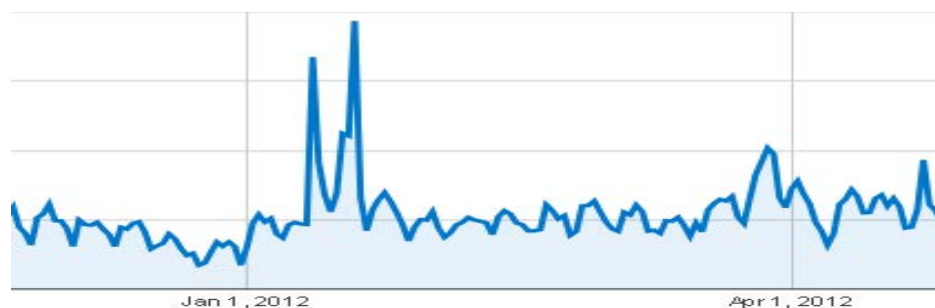


A kísérletről

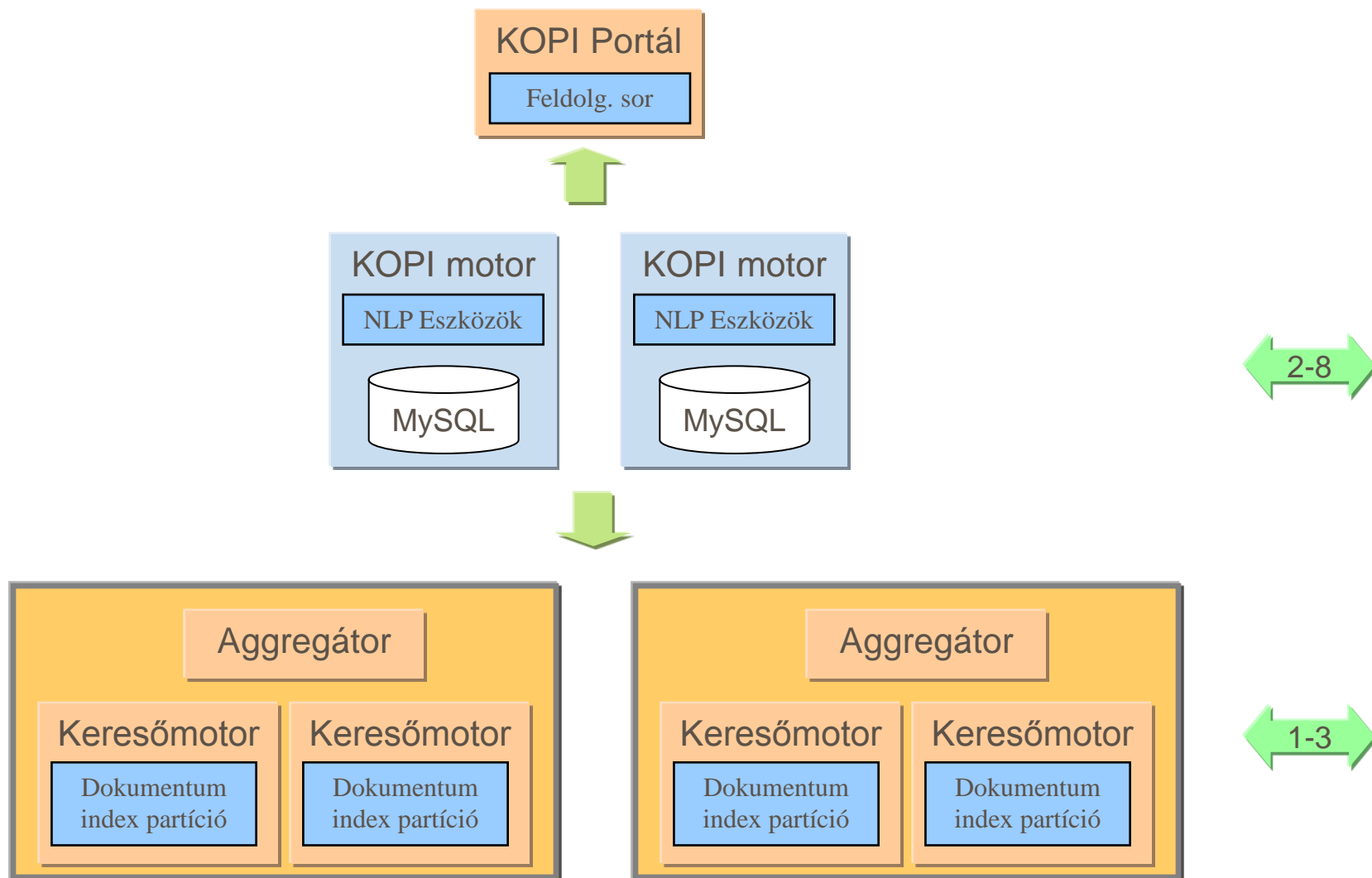
DSD

Elosztott
Rendszerek Osztály

- A cél: stabil szolgáltatásminőség fenntartása
 - Pl. az eredményre várakozás idejének felső korlátot szabni
 - Egy dokumentum ellenőrzése általában 30-50 perc
 - Amikor túl sok dokumentum érkezik be, ez akár 24 órára is nőhet
- A kísérlet során
 - Modellezzük a tipikus felhasználói tevékenységet
 - Különbféle skálázási módszereket mérünk heterogén felhőszövetségekben



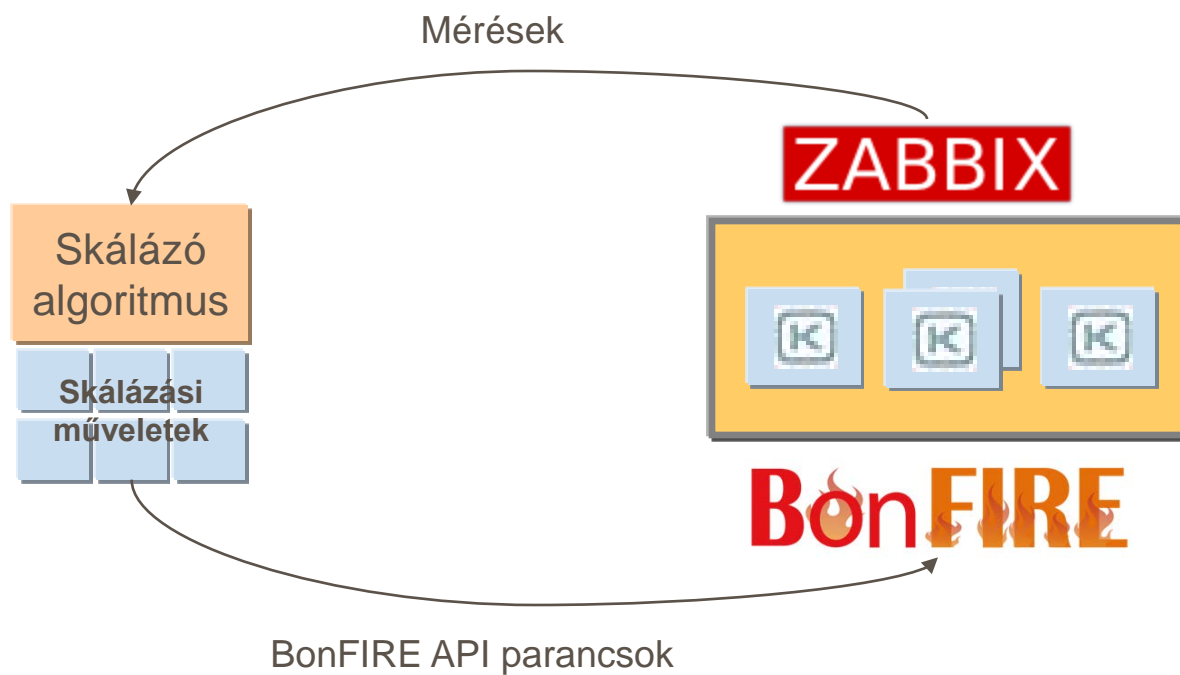
Skálázási lehetőségek



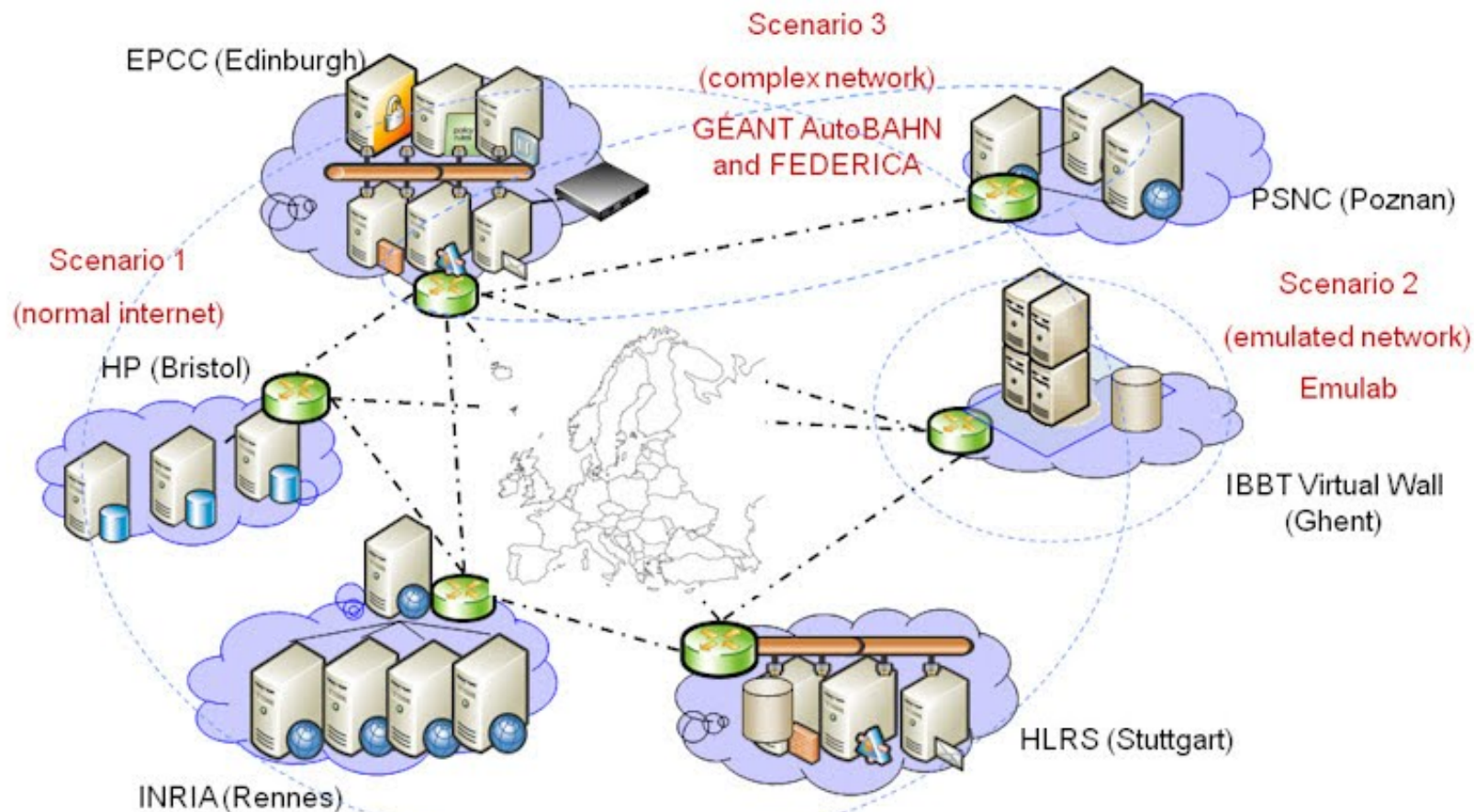
DSD

Elosztott
Rendszerek Osztály

A KOPFire kísérlet



BonFIRE: elosztott felhő tesztkörnyezet



Permanent (~350cores / 30TB) & On-Request (theoretically 3000+ cores) infrastructures

Note: network links indicative only

BonFIRE lehetőségek

- VM típusok (8 típus + egyedi)
- Adatblokkok
 - OS vagy DATA, perzisztens, shared, stb.
 - Több blokk is kapcsolható egy VM-hez
- Hálózat
 - Internet, AutoBahn, Virtual Wall, FEDERICA, háttérforgalom generálás, stb.
- Monitorozás
- Elasticity as a Service
- Értesítések (RabbitMQ)

DSD

Elosztott
Rendszerek Osztály

A BonFIRE felhasználói portál



BonFIRE



micsik

BonFIRE >> Experiments >> Experiment Details

Experiment Details - *KOPFire 1* (ready)

Experiment ID: /experiments/23617
User ID: micsik
Groups: kopfire
Creation Time: Tue, 02/12/13 14:07:12 UTC
Last Updated: Tue, 02/12/13 14:07:12 UTC
Expires: Wed, 02/12/14 14:07:12 UTC
[\(Show XML\)](#)

✖ Delete ■ Stop ▶ GO

Resources

Site Interconnection

Elasticity

Monitoring

Compute Resources

Filter by Site: <all>

Name	Id	Type	Cpus	Mem	VM Image	Wan IP	SSH	State	Del.
fulltext-engin	/locations/de-hlrs/computes/11426	small	1	1024MIB	BonFIRE Debian Squ	172.18.2.152	Not active - Host is	DONE	✖
fulltext-engin	/locations/de-hlrs/computes/11425	small	1	1024MIB	BonFIRE Debian Squ	172.18.2.151	Not active - Host is	DONE	✖
fulltext-engin	/locations/de-hlrs/computes/11424	small	1	1024MIB	BonFIRE Debian Squ	172.18.2.150	Not active - Host is	DONE	✖
fulltext-engin	/locations/de-hlrs/computes/11420	small	1	1024MIB	BonFIRE Debian Squ	172.18.2.145	Not active - Host is	DONE	✖
fulltext-engin	/locations/de-hlrs/computes/11419	small	1	1024MIB	BonFIRE Debian Squ	172.18.2.144	Not active - Host is	DONE	✖
frontend-hlrs	/locations/de-hlrs/computes/11098	small	1	1024MIB	BonFIRE Debian Squ	172.18.2.156	OK	RUNNING	✖
fulltext-engin	/locations/be-ibbt/computes/891	Large-EN	4	4048MIB	BonFIRE Debian Squ	<unassigned>	Not active - Host is	PENDING	✖
fulltext-index	/locations/be-ibbt/computes/890	Large-EN	4	4048MIB	BonFIRE Debian Squ	<unassigned>	Not active - Host is	PENDING	✖
fulltext-index	/locations/be-ibbt/computes/889	Large-EN	4	4048MIB	BonFIRE Debian Squ	<unassigned>	Not active - Host is	PENDING	✖
fulltext-index	/locations/be-ibbt/computes/888	Large-EN	4	4048MIB	BonFIRE Debian Squ	<unassigned>	Not active - Host is	PENDING	✖

Page 2 of 2

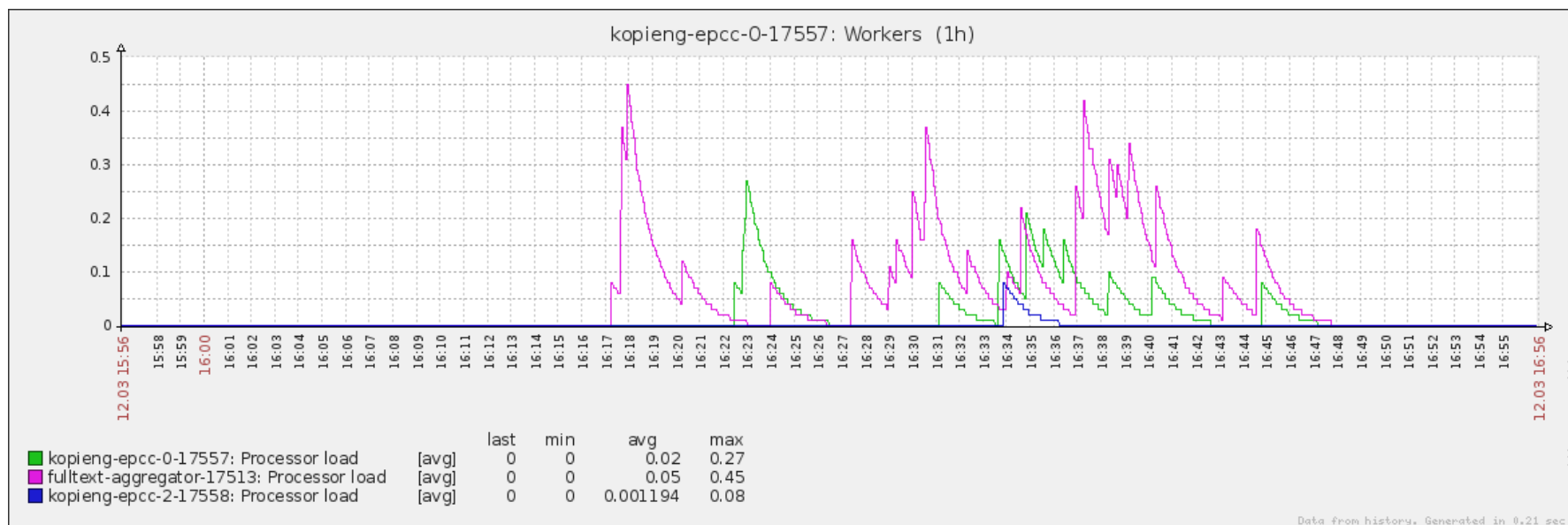
Add Compute at be-ibbt

Storage Resources

Filter by Site: <all>

Name	Id	Description	Type	Size	FS	State	Pub.	Per.	Del.
fulltext-index-small-2	/locations/be-ibbt/storages/44	2. part of small index	DATABLOCK	2500MB	ext3	<n/a>	false	true	✖
fulltext-index-small-0	/locations/uk-epcc/storages/2676	0. part of the smaller index	DATABLOCK	2500MB	ext3	USED	false	false	✖

- Egységes monitorozási lehetőség Zabbix-szal
 - Fizikai gépek
 - Virtuális gépek
 - Saját mérések



- REST API
 - Experiment descriptor: JSON vagy OCCI


```
<compute xmlns="http://api.bonfire-project.eu/doc/schemas/occi">
  <name>my-vm</name>
  <instance_type>lite</instance_type>
  <disk> <storage href="/locations/fr-inria/storages/165" /></disk> <nic>
  <network href="/locations/fr-inria/networks/47" /> </nic>
  <location href="/locations/fr-inria" />
</compute>
```
- Restfully (Ruby)


```
experiment.computes.submit(
  :name => "VM#{experiment['id']}",
  :instancetype => "small",
  :disk => [{
    :storage => inria.storages.find{|s|
      s['name'] == SERVER_IMAGE_NAME},
    :type => "OS"    }],
  :location => inria )
```
- CLI (parancssor)
 - `bfcompute create 'vm0' '/locations/de-hlrs/storages/2088' 23617`

- A KOPI terhelési problémáinak tesztelésére nagyon jó lehetőséget ad a BonFIRE
- A BonFIRE platform egy felhőszövetség (cloud federation)
 - Elérhető lesz legalább 2014 őszéig
 - Nyílt hozzáférés
 - Jelentkezni lehet kísérletek elvégzésére:
<http://portal.bonfire-project.eu/en/public/propose-experiment/>

Bon FIRE

Building service testbeds on FIRE

Köszönöm a figyelmet!